



# Generation Challenge Programme

CULTIVATING PLANT DIVERSITY FOR THE RESOURCE-POOR

## Medium-Term Plan

# 2005–2007

August 2004



Generation Challenge Program  
Cultivating Plant Diversity for the Resource Poor

**MEDIUM-TERM PLAN 2005 – 2007**

August 2004

# Contents

Acronyms and Abbreviations	iii
Introduction	1
Program Structure	1
Strategic Overview	3
Program Outputs, Milestone, and Activities	8
Collaboration	9
Highlights of Achievements of 2004	10
Subprogram 1: Genetic Diversity of Global Genetic Resources	10
Subprogram 2: Comparative Genomics for Gene Discovery	14
Subprogram 3: Trait Capture for Crop Improvement	17
Subprogram 4: Genetic Resources, Genomic, and Crop Information Systems and Bioinformatics	19
Subprogram 5: Capacity Building	22
Administration and Governance	24
Financial Highlights	27
 <b>ANNEXES</b>	
Annex 1: GCP Subprograms and Projects (Outputs)	30
Subprogram 1	30
Subprogram 2	36
Subprogram 3	42
Subprogram 4	49
Subprogram 5	56
Annex 2: Logframe	65
Annex 3: GCP Consortium Members	70

## Acronyms and Abbreviations

AGROPOLIS	International Complex for Research and Higher Education in Agriculture (France)
ARI	Advanced research institute
BAC	Bacteria artificial chromosome
BioMOBY	Open source biological web services project, “Model Organization Bring Your Own Database”
CAAS	Chinese Academy of Agricultural Sciences
CBR	C-repeat binding factor
cDNA	complementary DNA
CENARGEN	National Center for Genetic Resources and Biotechnology Research, Brazil
CGIAR	Consulative Group on International Agricultural Research
CIAT	International Center for Tropical Agriculture (Centro Internacional de Agricultura Tropical)
CIMMYT	International Maize and Wheat Improvement Center (Centro Internacional de Mejoramiento de Maíz y Trigo)
CIP	International Potato Center (Centro Internacional de la Papa)
COS	Conserved orthologous sets
DNA	Deoxyribonucleic acid
DREB	Dehydration responsive element binding
EMBRAPA	Brazilian Agricultural Research Corporation (Empresa Brasileira de Pesquisa Agropecuaria)
eQTL	Expressed quantitative trait loci
EST	Expressed sequence tag
FAO	Food and Agriculture Organization of the United Nations
Gramene	A comparative mapping resource for grains
ICARDA	International Center for Agricultural Research in the Dry Areas
ICIS	International Crop Information System
ICRISAT	International Crops Research Institute for the Semi-Arid Tropics
IITA	International Institute of Tropical Agriculture
INIBAP	International Network for Improvement of Banana and Plantain
IPGRI	International Plant Genetics Resources Institute
IRIS	International Rice Information System
IRRI	International Rice Research Institute
IWIS	International Wheat Information System
JIC	John Innes Centre, UK
LD	Linkage disequilibrium
LIMS	Laboratory information management systems
MAS	Marker-assisted selection
MTA	Material transfer agreement
NARS	National agricultural research system
NCGR	National Center for Genome Resources
NGO	Non-governmental organization
NIAS	National Institute of Agrobiological Sciences, Japan
PAC	Program Advisory Committee
PCR	Polymerase chain reaction
PRMT	Program Research Management Team
PSC	Program Steering Committee
QTL	Quantitative trait loci

RFLP	Restriction fragment length polymorphism
RIL	Recombinant inbred line
SGRP	System-wide Genetic Resources Programme
SINGER	System-wide Information Network for Genetic Resources
SME	Small- and medium-sized enterprises
SNP	Single nucleotide polymorphism
SSR	Simple sequence repeat
TRIPS	Trade-Related Aspects of Intellectual Property Rights
USDA	United States Department of Agriculture
WARDA	West African Rice Development Association
WTO	World Trade Organization

## **Introduction**

The inception year of the Generation Challenge Program (GCP) has been one of impressive scientific productivity, exceptional institutional growth, expansion of exciting scientific and institutional partnerships, and increasing international recognition beyond the CGIAR. The productive and promising arrival of this new model for the application of cutting-edge science to addressing problems of the world's poorest people is proving to be attractive to leading scientists from around the world. This appeal extends to traditional donors to the CGIAR as well and offers a new portal for non-traditional donors and the private sector to participate in the CGIAR.

This first medium-term plan (MTP) for GCP is based upon the program presented to the Executive Council of the CGIAR in 2003. The GCP was approved for a one year inception phase in July of that year that now extends to September 2004. This MTP translates the original conceptual framework of the GCP into an operational document with specific verifiable activities and outputs with associated time frames. It is the product of an extremely intensive year-long interaction among scientists and research managers from 15 participating institutions and their partners.

Mapping the broad approaches of our original document onto a well-anchored and concrete MTP began with the initial work planning meeting held at the end of August 2003, in Wageningen, Netherlands – less than one year ago. At that meeting, representatives from the institutions that form the core consortium of the GCP (see Annex 3) initiated the development of the first year work plan. This plan, subjected to external peer review, was approved by the two largest donors at that time, the World Bank and the European Commission, became operational in early 2004 (see <http://www.generationcp.org> for both the original Challenge Program proposal and the Year 1 Work Plan).

The first year work plan laid out three main domains of activity. The first domain was to initiate the gathering and application of essential genomic information of the target CGIAR crop species. The second domain was to begin the establishment of the technical and scientific foundation for the future activities of the GCP. Both of these domains involved a range of specific research and capacity building undertakings whose progress will be summarized under each Subprogram section. The third activity domain was to establish the institutional structure that would support such a complex international undertaking.

## **Program Structure**

The GCP activities for 2005 – 2007 are organized within five subprograms that fall into two broad objectives (described in Annex 1). Each contributes directly to the GCP purpose of creating a freely available public platform to access and utilize the vast genetic diversity held in germplasm collections of crops and their wild relatives. In addition to gene/trait discovery and application, the GCP subprograms also establish the mechanisms at a CGIAR level for capturing, storing, analyzing, accessing and interpreting the vast amount of biological data that the GCP and its partners will generate. Integrated into all the subprograms is a strong capacity building component that assures

that scientists from developing countries will be active partners in the Program and help ensure that the products of GCP research will ultimately reach the intended beneficiaries.

The application of new and existing tools of modern plant genomics and comparative biology to this germplasm will identify new genes for traits of importance that will be incorporated into crop improvement programs targeting the poorest of people in developing countries. Each of the subprograms has a set of clearly defined and measurable outputs that relate directly to its rationale and objectives. These outputs are presented in a time frame that will allow unambiguous assessment of progress and productivity. The ultimate beneficiaries will be resource poor farmers and consumers as varieties with improved tolerance to difficult environments contribute to improved productivity and use of natural and applied resources. Considering the global priority given to water use and managing its scarcity, and its relevance to all production systems, the GCP has selected drought tolerance as the over-arching trait around which to organize and focus its activities.

Thus, in light of the attention it will pay to policies surrounding intellectual property, genetic resources and the field application of molecular biology, the GCP is well positioned in each of the five CGIAR outputs.

The GCP operational objectives and Subprograms are:

### ***Objective 1***

Develop a platform for, and undertake analysis of, genetic diversity in international crop genetic resources and apply this to improve major crops for drought tolerance and other related traits of importance to resource-poor farmers.

- **Subprogram 1: Genetic Diversity of Global Genetic Resources**

*This subprogram aims to characterize the diversity of the crop germplasm collections held by the CGIAR and its partners. This characterization includes an assessment of the genetic structure of the collections as well as the phenotypic variation associated with that structure. As many of the policy questions that confront the GCP are associated with access to and application of genetic resources, much of the GCP's policy-related activities are incorporated in this subprogram.*

- **Subprogram 2: Comparative Genomics for Gene Discovery**

*This subprogram focuses on genomic tools, technologies, and approaches to achieve an understanding of gene systems across many species of importance to developing country agriculture. Comparative biology and genomics will be used to discover and validate the function of key genes central to the practical objectives of the GCP.*

- **Subprogram 3: Trait Capture for Crop Improvement**

*This subprogram focuses on the validation and refinement of molecular breeding systems and the resultant enhanced germplasm with the primary purpose of increasing the efficiency, speed, and scope of plant breeding gains. This includes a substantial commitment to create appropriate technologies for the application of marker-assisted selection in national breeding programs, to provide technical assistance for the rapid and effective uptake of molecular breeding in tropical staple*

*crops, and to foster the development of communities of practice supported by regional centers of excellence and state of the art technologies and approaches.*

- **Subprogram 4: Genetic Resource, Genomic, and Crop Information Systems**

*The value of the data generated in the first three subprograms will largely depend on the way they are stored, managed, analyzed, and made accessible. The way they can be analyzed will, in turn, depend on the way analysis tools and other information sources are made available. Subprogram 4 addresses the challenge of linking and integrating these information components and analysis tools into a coherent information gateway. A bioinformatics, biometric, and advanced data management system will be designed to support an integrated genetic resources, genomics, and crop improvement information network.*

### **Objective 2**

Strengthen the capacity of NARS and GCP consortium scientists to apply the tools of genomics, molecular biology and bioinformatics to the analysis of genetic diversity held in germplasm collections, and to use this knowledge to improve crop breeding programs and to develop new stress-tolerant varieties.

- **Subprogram 5: Capacity Building**

*Like Subprogram 4, this subprogram addresses a major cross-sectional theme of the GCP. It has two dimensions: one is to better enable GCP members to carry out this cutting edge research agenda. The second is to empower national program scientists to participate in GCP activities. In combination, these two activities create mechanisms by which GCP products can reach crop improvement programs and farmers.*

## **Strategic Overview**

Three simultaneous technological revolutions over the past decade have had dramatic impact on the CGIAR development-oriented research program. The revolutions in the fields of molecular biology and genetics, data storage and management, and communications introduce capabilities and opportunities that were undreamt of as the CGIAR was taking shape over thirty years ago. The Generation Challenge Program, more than any other, represents how the CGIAR is demonstrating the flexibility to respond positively to dramatic changes in its operating environment.

The spectacular advances in pharmacology and human genetics made possible from the Human Genome Project and model mammalian systems projects (e.g., the mouse) are harbingers of what is to come for plant systems. The application of cross species comparative genomics and association genetics has revealed surprisingly simple genetic relationships for physiologically complex syndromes in humans (e.g., the recent identification of gene(s) controlling “asthma”). It is reasonable to expect that complex traits, such as tolerance to drought and other abiotic stresses, may be deciphered in important food crops using approaches similar to those undertaken during the completion of the decoding of the rice genome and the dicot model species *Arabidopsis thaliana* and



*Medicago truncatula*. Indeed, there is a steady flow of reports in the scientific literature describing advances in our understanding of the relationship between gene sequence, function, plant physiology, and performance. That the private sector is now investing heavily in the development of drought tolerant cultivars through the application of plant genomics is a clear signal that practical results are possible for even the most difficult traits.

Nonetheless, the extraordinary discoveries of plant molecular biology, largely led by advanced research institutions in the “North” have yet to be used in ways that will benefit the world’s poor; likewise, the rich pools of genetic resources collections held by national agricultural research systems (NARS) and the CGIAR have yet to be tapped in a systematic way. The GCP creates a strong coalition of institutions dedicated to alleviating poverty by applying the recent advances of the biological sciences. This alliance aims to harness the powerful tools of the genomics revolution to unlock the genetic potential within crop germplasm to address the needs of the resource-poor. One of the principal products of the Program will be a unique *public platform* for accessing and developing new genetic resources using new molecular technologies and traditional means. This Program will make available as public goods an unprecedented array of genomic and genetic resources, ready for direct use in plant improvement, first to the national agricultural research systems (NARS) of developing countries that have plant improvement programs, and later to any other entities that have crop improvement goals, especially those dedicated to resource-poor farmers. These products will be in the form of enabling technologies and intermediate products for crop improvement programs in NARS and elsewhere.

### ***Why a Challenge Program?***

There is a clear and convincing case that the revolutions in biology, data management, and communications provide tremendous opportunities for solving some of the world’s most serious agricultural and food security issues. But, what advantage comes with creating a new program, rather than simply providing additional funding to the individual commodity programs? There are compelling scientific, operational, and financial arguments that lead to the conclusion that a unifying program like the GCP is required to take full advantage of the opportunities offered by science and to build upon the past investments of the CGIAR.

From a strictly scientific perspective, it is clear that the major advances in crop improvement will be derived via comparative biology and comparative genomics. The profound insights afforded by the discovery of broad synteny in genome organization among related species – first illustrated within the cereals – argue that it is far more informative and efficient to study genetic variation in sets of related species than to focus on one alone. The GCP offers a mechanism to envision and support research programs that cross institutional boundaries. Under a model where support for an individual crop species were simply increased, it would be extremely difficult to develop a program to assess, for example, differences in floral development under water deficit in cereals where four or more CGIAR centers should participate. Similarly, fundamental questions related to the genetics and physiology of adaptation to drought, root architecture, and/or low phosphorus in legumes must be addressed by teams that involve scientists from

several centers. By creating a framework in which cross species research can be conducted, the GCP captures the expertise and knowledge specific to individual crops in different centers and incorporates this into a common research program.

This framework for multi-institutional research extends to the involvement of ARIs and NARS. A mechanism that is available to a cross institutional program like the GCP, and for all practical purposes unavailable to individual centers, is a competitive grants program. The GCP has initiated such a program as a transparent, merit-based process to attract new and powerful partners to address its research agenda. The strong response from the research community indicates that this will be a successful model. For example, in the 80 pre-proposals submitted in response to the GCP's first call for proposals, over 130 non-consortium institutions were recruited as partners. This represents on one hand the ability of the GCP to tap into the enormous store of global scientific capacity, and on the other the attractiveness of the GCP's agenda and framework.

Although there was no required specification of in-kind contributions in competitive grant pre-proposal submissions, where there was in-kind specified, the amount averaged over 50% of the requested funding. Thus, a \$5 million competitive grants program will likely translate into an additional \$2.5 million. Since this is mostly accounted for in salaries, the partnerships created by the GCP are really means by which the CGIAR can procure specialized expertise without incurring long-term commitments. It is noteworthy that institutions typically did not include highly specialized and costly equipment, such as high throughput DNA sequencers, microarray spotters and readers, or mass spectroscopy facilities in their in-kind contributions. Therefore the value of their in-kind contributions is most likely greatly underestimated.

An important component of the GCP comparative genomics approach is that common procedures for data collection and measurement must be developed. As an independent program, the GCP provides the global forum appropriate to create these common operating procedures. In addition to standard and agreed upon means of data collection, a comparative biology approach demands that data from many experiments be accessible to the broader community. This means that there must be common standards for data capture and storage, as well as assured interoperability among databases and the analytical tools used to query them. The GCP offers the means to develop these standards and to create a global public platform for data access, analysis, and interpretation.

The creation of this platform requires a major investment in computer hardware, software, and personnel. Broadband internet access allows us to query remote data bases and apply tools hosted on distinct servers. Within the GCP framework much of the investments in hardware, software, and analytical capacity need be made only once, rather than replicated across centers. It is not necessary for every CGIAR center to make a massive investment in bioinformatics if the GCP provides the means to rationalize investments. For example, IRRI invests over \$400,000 per year in bioinformatics expertise. The GCP provides the mechanism to capture and leverage this investment such that it serves the entire community. Similarly, by negotiating favorable terms on cluster computing facilities and associated grid access, the GCP helps create a resource that empowers all members of the international agricultural research community. Rather than purchase its own computer cluster, any GCP member or its partner institutions, can

simply use its broadband internet access to avail itself of the massive computing power required to undertake comparative genomics analyses.

In order to fully participate in the genomics revolution, the international agricultural research community must have a voice in how international plant genomics data bases and web services are constructed, accessed, and managed. Otherwise the requirements imposed by the needs of our clients and beneficiaries may go unmet. It is not possible for each international center to be represented in the fora where such decisions are made. The specialized expertise would be difficult to justify for each center, and it is unlikely that the fora would welcome that many participants. However, the GCP is already a participant in the Plant Gene Ontology Consortium and contributes to BIO MOBY, the emerging standard for biological web services. By assembling the needs and perspectives of all the CGIAR centers with crop programs, the GCP ensures that research needs will be met.

### ***Program Strategic Focus***

The GCP's development goal is to increase food security and improve livelihoods in developing countries by unlocking the genetic potential and enhancing the use of public genetic resources in plant breeding programs through the concerted generation, management, dissemination, and application of comparative biological knowledge. In pursuit of this goal, the Challenge Program will create an integrated platform for dissecting genetic diversity in crop plant genetic resources, identifying important genes to reduce the impacts of environmental and biotic stresses on crop productivity, enhancing yield, and improving nutritional quality of crop products. Beyond this, the Challenge Program will identify, manipulate, and validate gene expression resulting in plants with potential value far beyond present-day crops. These plants, through seeds or vegetative propagules, will be transferred to breeding programs. An important GCP contribution will be to enhance the capacity of NARS scientists to participate in this program.

This Challenge Program will serve as a platform to assemble and use the intricacies of applied genomic sciences for the benefit of crop improvement efforts by NARS and others targeting the world's poorest regions. A major challenge, however, is how to satisfy the needs of very large set of stakeholders within and outside the CGIAR. The key feature of the GCP platform will be its applicability to any crop and any trait, thereby ensuring that all 22 CGIAR mandate crops will be supported by the platform. The platform will also be applicable to the Water for Food, Harvest Plus, and Sub-Saharan Africa Challenge Programs.

Despite the broad applicability of the GCP platform, there is still a need for focusing GCP activities. Even considering the power of comparative genomics and biology, resources must be allocated to only a limited set of crops for primary analysis. Likewise, traits and crops must be selected so as to benefit the greatest numbers of the resource poor as soon as possible; consequently, regional considerations will be important in setting research priorities.

The GCP is establishing three processes to establish and maintain its focus and relevance. First, with the support of GFAR we have created a diverse global Stakeholders' Committee comprised of representatives of NARS from the regional fora, NGOs, farmer

associations, and the private sector. Second, we will work with the new CGIAR program on Institutional Learning and Change (ILAC) to establish a learning-oriented monitoring and evaluation system. Third, with IFPRI we will develop a crop/region/trait prioritization model within the context of applied crop genomics and improvement. In addition to these GCP prioritization activities we expect to receive guidance from the Science Council as it completes its prioritization exercise.

As we refine our priorities over the next 12 months, we will work within the following guidelines that have established our program to date.

- **Poverty alleviation:** The world's greatest absolute numbers of the very poor are in South Asia (SA), the greatest proportion of the population that is poor is in Sub-Saharan Africa (SSA), and a very large zone of stagnant agricultural productivity associated with recalcitrant rural poverty is in Central and West Asia and North Africa (CWANA). Northeast Brazil and the Andean zone also contain pockets of serious poverty. Thus, our crop x trait focus must first and foremost address these areas of greatest need.
- **Crop targets:** Our comparative biology approach will in the first phase have greatest impact within three crop groups: cereals, legumes, and clonal crops. Within the cereals both the availability of scientific tools and poverty alleviation indicate that our initial concentration should be on rice (SA), maize (SSA), and wheat (CWANA). Research on these species will be complementary in that they will generate knowledge of broader applicability. For example, rice will focus on functional genomics, maize on the development of association genetics capacity, and wheat on gene identification taking advantage of global genetic stocks. The legumes are behind the cereals, yet investment in *Phaseolus* and cow pea genomics will have important impact for breeding programs targeting SSA. Modest and targeted investment in potato should yield insights into carbohydrate metabolism and starch accumulation that is relevant to cassava and *Musa*, as well. Since we expect the "orphan crops" to benefit substantially from our investments in the major crops, we will complete the initial characterization of their germplasm collections begun in the GCP inception phase and modestly assist in development of genetic stocks. This will permit these crops to more effectively use genomics tools and insights derived from other crops in the coming years. If we are successful in securing additional funding this will be expanded in SSA.
- **Trait targets:** Drought was chosen as the long-term case study because drought affects all of the CGIAR mandate crops, it is the main constraint in the three largest poverty stricken areas of the world, and it has resisted resolution using conventional approaches. This effort will be reinforced by the long history of drought research and by current global interest in water conservation. Furthermore, drought and associated water use efficiency emerged as the top priorities in the CGIAR System Priorities (preliminary results as presented by Alain de Janvry at the USAID SOP conference in Davis, California, June, 2004). The interaction of plants with water deficit ("drought") is a complex phenomenon, as are the genetic and physiological strategies by which plants manage water deficit. Thus we will examine a range of traits associated with this broader target. Non-drought related traits – especially those with a shorter time horizon for impact such as disease and pest resistance, food quality,

and plant architecture – will be addressed if they contribute substantially to tool and technique development. The GCP will also address the need for parallel analysis of stress-response. An important aspect of the comparative biology underpinnings of the GCP is to make use of well-characterized systems, not only for bringing near-term results, but for enabling identification and manipulation of “drought tolerance genes.” Understanding and predicting the interaction of stress-response traits, either synergistic or antagonistic, is critical to assembling useful gene combinations for pre-breeding products.

We do not underestimate the challenge of going from “gene” to “trait” to “breeding program” to “crop” for a trait such as drought. Therefore we will incorporate a significant modeling component to critically evaluate our assumptions and predict the consequences of various approaches *in silico*. This will present opportunities to understand the interactions of multiple pathways with bearings on whole-plant stress response.

## **Program Outputs, Milestones, and Activities**

The Challenge Program will generate new science-based enabling and intermediate technologies. While our direct products are intermediate in nature, we are establishing strong links to crop improvement and other programs that directly reach end users in the world’s most impoverished regions. For example, we have developed an agreement with NEPAD to cooperate in developing joint capacity building and crop improvement support programs for eastern Africa. The GCP Subprogram leader for SP 3 will be located at the East Africa Biosciences facility in the ILRI campus to directly manage this integration. Likewise, we agreed with the Harvest Plus Challenge Program to develop a common strategy to reach end users. Harvest Plus will take the lead in eastern Africa. We are working with a donor to secure funding to initiate a rainfed lowland rice improvement network for South Asia, where the world’s largest concentration of rural poor lives and over half of the rice grown is subject to drought and other abiotic stresses.

A technology transfer plan will be designed in coordination with Harvest Plus and our consortium members to ensure that the products of research undertaken by the Challenge Program will be delivered to and used by plant breeders and farmers. Such a plan will include consideration of *inter alia*: joint venture agreements; license agreements with humanitarian clauses and market segmentation provisions; and material transfer agreements. The management of GCP intellectual property as put forth in our Consortium Agreement is consistent with the Guiding Principles for the CGIAR on Intellectual Property Rights Related to Genetic Resources, the Convention on Biological Diversity, and the International Treaty on Plant Genetic Resources for Food and Agriculture. The principle aim of intellectual property management in the Challenge Program will be to ensure that research outcomes remain accessible in the public domain for the benefit of NARS and other plant breeders, and ultimately, the resource-poor farmers in developing countries.

The specific outputs, milestones, verifiable indicators, and time frames for each of the Subprograms are presented in Annex 1.

## **Collaboration**

The Challenge Program was originally presented by a founding group of three CGIAR Centers (CIMMYT, IPGRI, and IRRI), two NARS (CAAS, China and EMBRAPA, Brazil), and five advanced research institutes (ARIs) (Cornell University, USA; AGROPOLIS, France; John Innes Centre, UK; National Institute of Agrobiological Sciences, NIAS, Japan; and Wageningen University, the Netherlands). Subsequent to a Stakeholders Meeting in Alexandria, Egypt, the following institutions were added to the Consortium and are represented on the Program Steering Committee: GFAR, CIAT, CIP, ICARDA, ICRISAT, and IITA. It is expected that additional institutions, both private and public, will join the Challenge Program as mutual benefits of their membership are identified. ILRI, as host to the Biosciences for Eastern and Central Africa facility, is an important partner.

As indicated earlier, both the competitive and the commissioned grants programs are intended to broaden participation in the GCP beyond the Consortium members. Thus we expect significant collaboration with ARIs and NARS beyond those included in the Consortium. For competitive grants, it is required to include at least one partner from a developing country. Likewise, NARS capacity building is seen as an integral and central part of our collaborative strategy.

The Generation Challenge Program will champion a new architecture of innovation for agricultural development based on creating new collaborative projects linking ARIs, IARCs, NARS, NGOs, and the private sector to generate systemic teams stretching from innovation to product delivery and impact. This mission and the goals of the GCP are very much allied to that of NEPAD's Biosciences initiative, with its first hub (BECA) currently being established at ILRI in Nairobi, Kenya, aimed at serving the East and Central Africa region. Thus, GCP and BECA have sought to build an alliance to harness the many synergies offered by bringing together their congruent and complementary programs. This will be achieved through the establishment of a joint capacity building and product delivery program whereupon both initiatives can collectively establish strategic partnerships to ensure delivery and impact of their science innovation outputs in East Africa. Preliminary joint ventures have included the co-sponsorship and co-organization of an intensive molecular breeding training program and joint fundraising efforts for establishing BECA as a genotyping hub for the germplasm and breeding activities in the region. In particular, the GCP wishes to create regional molecular breeding communities of practice served by centralized facilities for shuttle genotyping and training.

We now envisage the creation of a joint competitive grants system for projects that will be predominantly implemented by NARS scientists carrying out key upstream activities at the Biosciences facility, while all other activities will be carried out in their host institution, which would also play a critical role in national infrastructural development and product delivery. The most compelling common research priority areas appear to be the development of host plant solutions to drought stress and poor soil fertility, although the inclusion of developing host plant resistance to insect pests and mycotoxins might also be considered.

We will seek to maximize the efficient use of our scientists' time by coordinating as much as possible our research and capacity building activities with the other Challenge Programs. Where we have common objectives, partners, and/or targets we will develop joint programs. The GCP and HarvestPlus have already agreed to develop a joint research and training group in East and Southern Africa. Research activities will likely focus on germplasm characterization of maize, beans, and cassava but potentially also include rice, wheat, and sweet potato. Expansion to include marker-assisted selection projects is also envisioned. The primary goals of the training program will be to effectively integrate MAS into national breeding programs through an intensive program of sensitization, training, and support. This collaborative program is at the same time building critical mass in seed systems development and product delivery in the region in order to synergize a product-driven approach.

Much of our collaboration with NARS crop improvement programs will follow a network model. We recognize that some networks have been far more successful than others. To learn from past experience, we are undertaking a meta analysis of selected networks from around the world with the McKnight Foundation. The results of this analysis should allow us to create a next generation of crop improvement networks that will increase the potential for impact of the Challenge Programs.

We anticipate fertile ground for collaboration with the SSA Challenge Program, should it be launched.

## **Highlights of Achievements of 2004**

Recognizing most activities are at the initial stage of implementation, we requested each research team funded for the implementation phase to briefly report what have been started, the status of the work, and any deviation from the original plan. Reports from the research teams describe a range of activities at different stages of progression. Since work on different species progresses at different rates, the balance of results shown below does not necessarily reflect the balance of investment. There has been quite satisfactory progress of the GCP research agenda as put forth in the Work Plan derived from the Wageningen meeting and subsequent consultations. The noteworthy achievements in the first six months include:

### ***Subprogram 1***

Genotyping activities in first six months of 2004 were new activities for all of the participants, which required extensive coordination and planning. This was accomplished by several workshops in the first half of the year. Twenty seven participants representing the eleven crops selected for genotyping in the first year met on January 9 at the Plant and Animal Genome meeting to determine marker selection, sampling strategy, laboratory protocols, data collection, and deadlines. A composite germplasm set was identified for genotyping for each crop, generally following the criteria set forth in the Year 1 Work Plan. The list of accessions to be genotyped with structural markers often depended on negotiations among the various partners to assure that their priority germplasm was included. The global organization of the work was reviewed at a "data analysis workshop" that had been planned at the PAG workshop and took place in

Zaragoza, Spain, in late June. In several instances the identification of the composite set was performed in two steps, a first list of accessions being agreed upon early in the year to allow genotyping to start. The choice of the complementary set was left to the second part of the year so that information generated on the first list of germplasm could be used as a guide. The projected completion of the genotyping in the first year is presented in the following table. In general the work is on track. The delays experienced were primarily a result of delays in equipment arrival and the care in negotiations to be sure that the proper germplasm was being analyzed.

Crop	Partners	Objective 2004 (acc x marker loci)	# acc chosen	Expected Completion date
Barley	CAAS, ICARDA	500 x 50	500	December 2004
Wheat	CIMMYT, AGROPOLIS, ICARDA, CAAS	3000 x 50	2600	May 2005
Maize	CIMMYT AGROPOLIS, CAAS	1700 x 50	1300	May 2005
Sorghum	ICRISAT, AGRO, CAAS	700 x 30	700	December 2004
Rice	IRRI, CIAT, AGROPOLIS, EMBRAPA, WARDA	3000 x 50	1536	January 2005
Potato	CIP	1079 x 50	1079	December 2004
Cassava	CIAT, EMBRAPA, IITA, IPGRI	3000 x 36, 500 x DArTs	3000	December 2004
Cowpea	IITA, CAAS	(100*+2000) x 50		May 2005 (40 gSSR)
Chickpea	ICRISAT, ICARDA	288 x 50	288	October 2004
Common bean	CIAT, EMBRAPA	2060	1400	March 2005
Musa	IPGRI, AGROPOLIS, IITA	960 x 50	192	February 2005

\* 100 accessions analysed for 10 plants individually

At the data analysis workshop in June the participants agreed to use a common file format for data exchange. The fields (mandatory ones in bold) are: **Laboratory/Institute**, **Species**, **Sample ID**, Germplasm ID, **Locus**, Name of internal standard (=name of the molecular weight standard for peak size estimation), Dye, **Allele (size in bp)**, Peak size, Quality (scale from 1 to 100), Peak height, Volume (area under the curve), **Allele amount (2n, 3n, 4n,..., bulk)**. The content of the file can be pasted in a web site and converted to various input file formats adapted to various software packages using a Web tool box. The suggested softwares are global packages such as SAS, Genstat, PowerMarker, and others, or specific softwares such as DarWin, Structure, Partition, Mstrat, etc. The input formats will include SAS files; Individual x allele matrices with various column types (1 column per ploidy level, alleles separated by a "/" or concatenated alleles); disjunctive tables (1 column per allele); fully disjunctive tables (2 columns per allele). The list is not restrictive. SP4 will provide the Web interfaces for these conversions as well as clear any intellectual property issues that could arise from using copyrighted software.

The composition of the initial composite set includes wild materials (W), landraces (L) and improved (I) materials. Although there is no single rule for the diverse crops, a



distribution of 5%W:75%L:20%I was agreed to be an acceptable target by the participants. The issue of heterogeneous accessions was considered: when possible and efficient (suitable multiplication rate), it was agreed to extract DNA from a single plant per accession and to self it (inbreeders) and use these seeds as the basis for analysis. For outbreeders not clonally propagated, such as maize, a range of methods is possible; CIMMYT and INRA have acquired considerable experience in the handling of bulks to evaluate allele frequencies. However, there is concern that there is not adequate statistical power to reveal functional associations with available within-accession diversity.

The application of phenotyping with the view to conducting association studies requires an extrapolation from the original genotyped collections for field/growth chamber/greenhouse experiments. The extrapolation from the composite set to yield reference sample that would be preferentially used for association studies must include:

- representatives of the main components of the diversity to cover the range of allelic diversity
- continuous coverage of the global range of genotypic diversity, ideal for species-wide association studies
- those sectors of the diversity that seem derived from recombination between two distinct components, ideal for Linkage Disequilibrium mapping
- some components with large continuous variation, ideal for subspecific association studies.

The value of using as many as possible common accessions across experiments for subsequent integration of information (and cross-comparisons) is recognized by GCP participants. Reference samples will be standard, but it was agreed that more specific materials such as preferred checks for each institution/environment will be included and that those materials that are not adapted to the experimental environment could be excluded at the investigators discretion. The combination of all these criteria requires development of simple, easy-to-use software for elaborating the set of materials for any new experiment.

#### *Phenotyping framework for the GCP*

It is widely recognized that there must be highly reliable, consistent, and interpretable phenotype data associated with the germplasm that has been genotyped for the volumes of genotype data to ultimately relate to plant and crop performance and be useful in crop improvement programs. Therefore the GCP has devoted, and will continue to devote, considerable resources towards establishing GCP-wide norms and standards. The first round of intensive consultation was organized in early July in Montpellier, involving a well-balanced proportion of 40 scientists from the CG centers, ARIs, and NARS (40, 30, and 30% respectively). The following positions were important outputs:

- A draft minimum set of environmental data which should characterise the environment of all experiments of the GCP in field, greenhouse, or growth chamber experiments has been drawn up. The next step will be to determine how these will be incorporated into individual projects and in the construction of the GCP database.

- Methods have been proposed to combine detailed measurements on a very small number of test genotypes and a very small number of cheap and rapid measurements on all genotypes. These will address an important set of environmental data concerns: (i) the perception of the environment by each studied individual genotype; and, (ii) the synchrony of the phenology of individual genotypes with environmental stressing events.
- It is not desirable that the GCP selects a set of relevant traits which would be common to all projects. In contrast, elements for evaluating the “quality” of traits have been proposed. A database to guide trait characterisation is under development.
- Some decisions, such as the choice of traits, the control and manipulation of environmental variables or the respective roles of genetic, modelling, physiology, and transcript/protein analyses are of the responsibility of individual projects. They are an essential characteristic of the scientific quality of those projects. In this respect, the GCP scientists and external advisors aimed to help and propose, but not to fix, general rules.
- Genetic variation within the samples (especially in cases of sexually propagated [partial] outbreeders) is considered incompatible with association studies. Simplification of the genetic constitution of the samples, with or without selection pressure, is considered a prerequisite for treating this kind of material.
- During the course of the workshop, it emerged how important it is to use specific characters among the materials that are to be compared in a phenotypic evaluation, such as phenology and gross morphology. It is likely that the phenotypic characterization of reference samples will require a season of gross field observation in the target environment in order to select those accessions that are most comparable and that will yield the most meaningful comparisons; this can also serve to increase seed for subsequent phenotyping experiments.

One of the working hypotheses that emerged from the phenotyping consultation was that phenotyping is not a series of repetitive trait measurements aimed to a standard characterization of the genetic material, but a creative activity based on scientific hypotheses. It requires a combination of skills including physical and physiological concepts and methods, modelling approaches, and genetic strategies.

Considerable potential exists for the GCP to promote and support integration of modelling approaches and physical or physiological concepts in the phenotyping process. Specific activities will include exchange of scientists, PhD theses elaborated in common, and development of a modelling course.

As observed in other subprograms, the involvement of non-consortium experts has proven very useful and promising for the future of the GCP.

#### *Policy position and white papers*

The following policy position papers are under development: Working definition of ‘resource-poor’; Humanitarian licenses; Liability & Stewardship; Opportunities for freeing up IP constraints, concentrating on the BIOS and PIPRA initiatives; Benefit sharing, including an analysis of the relevance for the CP for accessing materials, both

from countries with a national access law and countries without; and a case study on special issues with groundnut (non-annex 1 crop in the International Treaty on Biodiversity). They will be presented in draft form for discussion by the broader GCP in the Brisbane annual research meeting.

### ***Subprogram 2***

The Year 1 Work Plan of Subprogram 2 is organized in four work clusters designed to gather germplasm and genomic resources for identifying and validating stress tolerance genes with an emphasis on drought. Overall, we have developed a coordinated strategy among researchers in using diverse genetic materials that exhibit certain attributes of drought tolerance. In coordination with SP1, a common phenotyping framework of techniques (though not necessarily the same methodologies), plant developmental stages, and parameters has been developed to enable cross-species comparison. Each participating team has planted selected stocks for detailed drought phenotyping in the field studies (results expected toward the end of 2004). Each research team on drought has also expanded characterization of QTL in selected species. For consensus markers across species, we have initiated marker design using orthologous sequences across species. Genetic materials (RNA) from different genotypes and species have been prepared and shared among laboratories for testing with different gene array platforms (actual hybridization data expected in July and August). We have expanded and improved the characterization of EST libraries of *Musa* and cassava.

Below, we highlight specific achievements to illustrate the tangible results and collaborative nature of the work cluster activities:

- **A set of germplasm with drought-tolerant attributes or representative of donor gene pools has been selected from individual crops (e.g., sorghum, barley, wheat, rice, maize, chickpea, common beans, cassava, and potato) for detailed phenotypic analysis under water stress conditions.** In coordination with Subprogram 3, recurrent parents for backcrossing have been identified in different crops. At ICRISAT, for example, a sorghum variety susceptible to end-of-season terminal drought stress but well-adapted to the drought-prone environments in India was selected as the key recurrent parent for backcrossing of the stay-green component of terminal drought tolerance. Thirty-five additional recurrent parents were selected on the basis of their adaptation to other important drought-prone sorghum production environments.
- **Phenotyping protocols for assessing on drought tolerance are being implemented in selected crops and field methodologies for drought tolerance are under testing using advanced breeding lines.** For example, common protocols for the accurate characterization of silk and leaf elongation across temperate and tropical maize and have been developed and a set of CIMMYT lines are currently being screened. A first set of field methodologies for drought tolerance for use in the GCP have been tested using advanced breeding lines. In sorghum, replicated field experiments were designed to assess effects of putative drought tolerance QTLs at ICRISAT. Twenty-one QTL introgression lines and their common recurrent parent were evaluated, but none of the early-generation (BC<sub>3</sub>F<sub>3</sub>) stay-green QTL introgression lines were more

productive than the recurrent parent in a mild terminal drought stress environment. In potato, methodologies for field and greenhouse experiments were evaluated to identify key parameters for monitoring drought response. The method involves the establishment of drought conditions at tuber onset in the stressed group and continued watering in the control group. Experiments are being planned with 20 native Andean potato genotypes (selected from highlands field trials) under controlled greenhouse conditions.

- **Conserved orthologous markers (COS) are under design and evaluation in monocots and dicots.** CAAS has defined more than 1,700 markers from conserved cereal ESTs with potential for producing cross-species markers for wheat, rice, maize, and barley. A subset of these primers (aliquots) will be sent to IRRI to test across monocots (rice, maize, sorghum, wheat, and banana) to define conserved orthologous genes useful across a broad range of species. IRRI has selected genes with supporting evidence from rice and other species for their involvement in drought. Sequences across the monocots were retrieved from GenBank. Multiple sequence alignments were accomplished, and 119 primers with or without degeneracy were designed for conserved domains. Conditions for product amplification across selected monocots are currently being optimized. Preliminary results indicate amplification is possible among the monocots with genomes the size of maize or smaller (including *Musa*). ICARDA, in collaboration with Montana State University, developed COS markers by targeting exonic sequences flanking introns based on the rice whole-genome sequence database. A total of 136 primers (targeting 50 contigs) have been tested on durum, dicoccoides, and barley genotypes. The majority of the primers showed high specificity (one to two developed fragments). Sequencing has been conducted on some fragments from different genotypes to validate the targeted exonic region.

Through the coordination of CIP and Cornell, a list of potato tentative consensus sequences exhibiting a high degree of similarity to genes of *Arabidopsis*, rice, and tomato has been defined. Currently, the list comprises 27 potato orthologs of drought responsive genes from tomato or rice. Nine of these sequences could be derived from genes belonging to gene families and thus may be only of limited applicability as COS markers. Preliminary results suggest that COS from tomato-*Arabidopsis*-coffee (made available by Cornell) are expected to provide 'universal' markers for Solanaceae, and possibly for dicots. The most promising COS identified across Solanaceae will be recommended for assessment in other dicots. For other dicots, a list of 287 candidate genes with possible involvement in drought tolerance has been compiled from *Arabidopsis* and soybean sequence databases. IITA, in collaboration with CIAT and University of California-Riverside, is using sequences from soybean databases and is currently designing putative COS primers for cowpea.

- **Algorithms for extraction of candidate sequences for conserved genes from EST databases are being tested and shared.** A computational approach under development at Cornell was shared with CIP for testing COS markers in potato toward general use in dicots in the GCP (Feinan Wu and Steve Tanksley, paper in preparation). Multiple species representing three plant families were used reciprocally

as query and subject by BLAST, such that by screening tomato, potato, *Arabidopsis*, and coffee databases 12 query-subject pairs were obtained. For each species, a different number of best-matched groups were obtained, after which the common groups shared by the four sets of best matched groups were identified. Consensus sequences were obtained by first aligning DNA and inferred protein sequences using a multiple sequence alignment tool T\_COFFEE, followed by manual editing.

- **Progress in producing and evaluating gene expression datasets for cross-species comparative analyses.** Research teams within SP 2 are exploring the use of different gene array platforms for gene expression analysis. The rice 22K oligo chip (a product of NIAS-Agilent collaboration) is being evaluated for its utility as a common expression platform for several cereals (rice, maize, wheat, barely at IRRI, CIMMYT, and NIAS). RNA from the reproductive stage (panicle) of rice genotypes under contrasting water regimes (drought stress vs control) was prepared and used for hybridization with the 22K chip at NIAS. Also, RNA of parental lines of a rice recombinant inbred population segregating for stress tolerance (biotic and abiotic stresses) was used to hybridize to the 22K chip in preparation for genome-wide segregational analysis (IRRI, NIAS). In all cases, high quality hybridization was obtained and data are being analyzed and will be presented in Brisbane meeting (IRRI, NIAS).

In collaboration with Pioneer, CIMMYT has identified an initial set of genes (in thousands) through comparative microarray profiling across parental materials. A second set of analysis is in progress to compare contrasting genotypes from a segregating population. These are being studied in detail using more tissues and/or different time points with real time PCR. CIMMYT has also purchased publicly available maize microarrays and contracted the service facilities of the Mexican National University (UNAM) have been contracted for the screening of materials (ears, silks and leaves) from different populations. A comparison across experiments should allow identification of a core set of genes for further study taking into account the pathways of importance and map position of QTLs identified for drought.

At CIAT, a search for genes related to stress inducible traits (e.g. stomata closure) and for constitutive traits (e.g. root growth) from public databases was completed, leading to the identification of about 300 genes. About 150 genes were selected as potential key genes for each trait. Primers for these genes have been designed and ordered for gene expression analysis using real time PCR. The list of the genes has been made available to IITA for mapping experiments.

CIP has defined genes putatively involved in drought stress tolerance pathways of potato. These pathways include stress signaling, carbon metabolism, adaptation to osmotic stress, antioxidant production, and detoxification. The selection of pathways will be reviewed when data from high throughput gene expression experiments is available. RNA has been prepared from selected potato genotypes and quantitative PCR experiments are being performed to detect expression changes in candidate genes in leaves of stressed and non-stressed native potato varieties. The gene

functions tested comprise drought stress adaptation (e.g., dehydrins), antioxidant biosynthesis, and heat shock genes. Microarray experiments are planned in collaboration with Virginia Polytechnic Institute and TIGR.

- **EST libraries expanded for several species under stress conditions.** Two SSH libraries of *Musa* have been constructed starting from RNA from cryopreserved cellular embryogenic cell suspension. Two libraries (i.e., reciprocal subtraction) of 2000 cDNA each are now available (INIBAP). IITA will begin construction of cDNA library shortly using mRNA from seedlings of drought sensitive but high water use efficiency (WUE) *Musa acuminata* (AA). Drought tolerant *Musa balbisiana* (BB genome) accessions and a *Musa balbisiana*/*Musa textilis* interspecific hybrid have been water-stressed and RNA prepared for subtractive cDNA library construction.
- **Mapping population characterization for drought QTL analysis.** A variety of population development and mapping experiments are in progress in multiple crops (e.g., sorghum, maize, potato, rice, *Musa*). These experiments provide the materials to apply COS markers to develop consensus QTL maps. In addition, the phenotypically-selected lines and advanced backcross progeny provide the materials for expression analysis or association tests with genetic variation (in coordination with SP1).

### ***Subprogram 3***

The primary focus of SP3 lies in the application of genomics tools and the development of products based on outputs from the other subprograms. For this reason, first year activities in SP3 have been largely limited to the validation of linked markers for drought tolerance as a means of establishing an effective operational framework in molecular breeding across each crop group. These foundation building R&D activities have been supplemented by the development of a database of elite varieties that are considered the most appropriate candidate background genotypes for MAS or transgene introgression. In addition, a number of small activities were commissioned for only the first year to help synergize appropriate population development and QTL mapping of component traits of drought tolerance in a range of lesser-studied crops.

Two major community capacity building activities were also commissioned during the first year. A workshop was convened to synthesize the GCP research strategy and policy development regarding the application of transgenic solutions for complex traits such as drought tolerance. This has already resulted in several consortium-based proposals while a synthesis and strategy paper will be released soon.

An intensive three-week molecular breeding training course will be carried out at the end of the year to launch a molecular breeding community of practice in Africa. The course aims to provide training that is tailored to the needs of the participants and effectively becomes a part of their research or breeding programs. Participants will have a working knowledge of DNA marker genotyping techniques. The training course will then aim to intensify and broaden the expertise of those participants to directly improve the efficiency and impact of their current activities.

It is our intention to convene this type of training course on an annual basis in East, Southern, and West Africa as well as key locations in Asia and Latin America. This first course in Kenya has three broad goals:

- Provide advanced training to current genomics scientists in the region
- Establish functional relationships for technical backstopping and trouble shooting
- Empower practitioners in the region to make best use of the BECA hub

It is envisaged that participation in this course will include members of pre-existing marker-assisted selection networks coordinated by the Rockefeller Foundation in East and Southern Africa, and the Kirkhouse Trust in West Africa.

Below, some specific highlights from first year activities:

- **Elite varieties identified, agronomically evaluated, and market appraised.** A database of elite varieties has been created with an initial focus on the following crops but this will ultimately be extended to include all crops: rice, maize, wheat, barley, sorghum, pearl millet, cassava, potato, plantain/banana, common bean, cowpea, and chickpea. The database will include all available characterization, evaluation, pedigree, and synonym data as well as information related to ownership, usage and production profile. This is an important milestone towards ensuring that all necessary information is collated, collected, and/or generated to enable the most appropriate focus varieties to be selected in terms of agronomic performance, market preference, and trading potential. The next step in this area will be to document appropriate baseline information for those varieties that will be used in subsequent impact assessment studies. These activities will be carried out in collaboration with the Harvest Plus Challenge Program with the aim of jointly developing seed-based products combining improved resilience and enhanced nutritional value.
- **Gene and trait introgression technologies developed, validated, and refined.** Gene-based markers for blast disease resistance in rice have been used for the proof-of-concept pyramiding of a substantial number of QTL (up to 16) in a single genotype (in this case, up to 10 lines with good agronomic backgrounds and drought tolerance). These introgression lines are currently being evaluated under diverse water regimes at IRRI and at diverse national program locations in India, Philippines, and Indonesia.

Individual and flanking linked SSR markers for at least 12 genomic regions associated with drought tolerance in maize are being tested in 6 diverse breeding populations by CIMMYT, INRA, and IITA based on comparison of phenotypic performance under glasshouse and field conditions with genotypic profiles of both breeding lines and their selfed progeny. Similarly, single and flanking linked SSR markers for QTL contributing to terminal drought tolerance in sorghum and pearl millet are being tested through the phenotypic validation of products of MAS. In the case of sorghum, two different sources of the staygreen component of drought tolerance are being introgressed into elite varieties of economic importance in Latin America, Africa, and Asia. For pearl millet, QTL are being introgressed into a variety of popular hybrid variety maintainer lines and evaluated in diverse moisture regimes.

SSR markers in common bean linked to important QTL for drought tolerance from four diverse sources are being tested through the phenotypic validation in Mexico and Brazil of multiple MAS backcross generations for three economically important varieties.

The development of effective molecular breeding tools for components of drought tolerance in legumes is being synergized by fostering large scale mapping efforts in common bean, cowpea, and chickpea. Similarly, intensive population development efforts are ongoing in cassava by IITA and CIAT to establish breeding populations for the effective pyramiding of drought tolerance with pest and disease resistance traits.

- **Communities of practice established, supported, and synergized.** To launch and sustain the molecular breeding community of practice in East Africa, the GCP together with the NEPAD Biosciences Facility for Eastern and Central Africa is holding an intensive annual three week training course in molecular breeding. This training program has three broad goals:
  - Provide advanced training to current users in the region
  - Establish functional relationships for technical backstopping and trouble shooting
  - Empower practitioners in the region to make best use of the biosciences hub

The course aims to provide training that is tailored to the needs of the participants and effectively becomes a part of their research or breeding programs. In this way, the training program and community of practice will be initiated through participants already involved in DNA marker research or breeding applications. Thus, the training course will intensify and broaden the expertise of those participants to directly improve the efficiency and impact of their current activities. The first few days of the course will be held in conjunction with a short sensitization workshop for plant breeders and science managers.

Towards the end of the course participants will also be provided with an appreciation of the high throughput genotyping facilities available at Biosciences and how to access and utilize these facilities for the maximum benefit of their research and breeding activities. This is a particularly important component for establishing a functional community of practice. Periodic access to a high throughput genotyping facility will help NARS breeding programs achieve cost effective molecular breeding success stories which should in turn synergize sustainable national investments for infrastructural and capacity development in this area. Similarly, occasional access to a state-of-the-art genomics facility may have dramatic impacts on the pace and impact of national research projects.

#### ***Subprogram 4***

The first six months of 2004 have been a very active and fruitful time for SP4. Herewith some highlights of first year SP4 activities:

- **Management, structure, and design for SP4 established.** Shortly after the appointment of Theo van Hintum as Subprogram Leader, a Consultation Workshop was organized in Rome, where about 50 participants, both GCP partners and invited



experts, discussed the content and planning for SP4. As input for this meeting, eight white papers were produced describing the general global status of following topics: ‘Germplasm Information Systems;’ ‘Fingerprinting and Allele Data Systems;’ ‘Mapping Data and Analysis Systems;’ ‘Functional Genomic Information and Analysis System;’ ‘Laboratory Information Management Systems (LIMS);’ ‘Central Registry and User Needs;’ ‘Interoperability and Infrastructure;’ and ‘GRID Computing.’ These white papers are being updated and most will be made public later this year. Based on the discussions at the meeting, the foreseen activities were regrouped and reformulated, and task-leaders were appointed to act as contact to the SP4 leader. Based on inputs from the task-leaders, a detailed Workplan was compiled. A virtual workspace was commissioned and is now beginning to be used for exchange of documents and as a discussion platform.

From May 31<sup>st</sup> through June 4<sup>th</sup>, a combined GCP SP4 system design workshop was convened at CIMMYT, involving a large range of Consortium partners with the addition of a significant number of non-Consortium invited bioinformatics experts. A smaller, follow-up workshop was convened at IRRI in July to undertake more detailed design and some prototyping of GCP systems.

- **Integrated germplasm and crop information systems and LIMS reviewed.**

Institutes involved in the development of the GCP Information System have reviewed existing systems and identified components which will be reused for CP data as well as components which must be developed or adopted. All institutes, IRRI, CIP, ICARDA, IITA, ICRISAT, CIAT, and IPGRI-INIBAP have suitable genetic resources information management systems, but most need to develop components to handle marker and genotype data. IRRI, CIP and INIBAP are investigating the GERMINATE schema from the Scottish Crops Research Institute (SCRI), ICARDA is testing the Gene management System of the International Crop Information System (ICIS), IITA, ICRISAT and CIAT are extending existing in-house systems. A comparison of the GERMINATE and ICIS-DMS schemata at an IRRI-hosted implementation workshop in July has demonstrated remarkable similarities between the two and IRRI is now attempting to merge the best ideas of both into ICIS-DMS and then develop a marker/allele management system to link with genotype data in the DMS. IRRI also has an ongoing effort to port ICIS to a new Java language based architecture that includes advanced technologies such as life sciences identifiers (LSID), controlled vocabularies and ontologies; and web services, as discussed within the SP4 design discussions.

Concerning the development and deployment of LIMS, momentum has been created and various activities are taking place at various institutions ranging from an inventory of requirements to inclusion of LIMS output into the breeding database. A LIMS system is being commissioned at IRRI to capture SP2 activities for universal marker development.

- **Plans for interoperability, infrastructure, and a central registry developed.**  
Collaborations between the GCP and a number of pertinent international initiatives were established to apply advanced data interoperability technology to the needs of the GCP:
  - SP4 was formally acknowledged as Plant Ontology Consortium Collaborator
  - Gene ontology consortium interactions were established
  - The principal BioMoby research scientists Mark Wilkinson (Genome Canada), Martin Senger (EBI/MyGrid), Gary Schiltz (NCGR) and Rebecca Ernst (PlaNet/MIPS) were directly involved in GCP design & implementation meetings (BioMoby web services technology is being incorporated into the ICIS Java platform by IRRI)
  - TAIR was visited in May by the task leader to discuss collaboration.
  - MaizeGDB was invited to the GCP design meeting.
  - Some IRRI discussions with Gramene are ongoing concerning bilateral web services integration of germplasm and crop genomic information.

The design, implementation, and deployment of the first generation infrastructure for expert curation of data standards (models, controlled vocabularies and ontology) was started by selecting the Generic Model Organism Database ([www.gmod.org](http://www.gmod.org)) “Chado” schema based database for commissioning a prototype GCP CVO management framework. This schema is somewhat inspired by the designs of other public ontology development efforts like the Gene Ontology and Plant Ontology Consortia. In addition, the “Protégé” ontology management system is being assessed for its utility in representing semantic information in the GCP. A community process was initiated to extend the inventory of existing public data exchange standards to cover gaps relative to CP platform needs and requirements starting the compilation of public CVO’s (e.g., Gene Ontology, Sequence Ontology) into the “Chado” schema.

Finally, BioMoby was reviewed and adopted at the CIMMYT design workshop as primary network integration technology and initiatives to implement this technology were taken. Pilot projects involving SINGER-Eurisco and IRRI-Gramene were proposed for near-term implementation. Further design discussions about BioMoby were undertaken at the IRRI-hosted workshop in July.

For the short term solutions, criterion and priority issues of IP related to the Genetic Mapping Data Repository Task have been identified, discussed and developed, particularly in relation to the use of Virtual Workspace Designs. The inventory listings of software have been broken down into those of three groups: GNU public license, author restricted but open license, and commercial license. And finally, an extensible and flexible structural and functional repository design that can provide interoperability has been designed and will be implemented shortly.

For the central registry, IPGRI has prototyped database mirroring software for semi-automatic archiving of GCP data sets to a central location.

- **Tools and databases to support SP1, 2, and 3 assessed and under development.**  
In the framework of support to the functional genomics activities (SP2): stress candidate gene discovery and characterization schema options were reviewed, and it was decided to use the Generic Model Organism Database (GMOD; [www.gmod.org](http://www.gmod.org))

curation tools and (“Chado”) schema as a starting point for building a comparative gene catalogue. A prototype database is being commissioned this summer. An IRRI-hosted GCP site for SP2 data sets is under construction at <http://www.iris.irri.org/generation>.

The Stanford Microarray Database “Longhorn” open source microarray database system is being deployed as a repository for comparative gene expression profile data to capture year 1 gene expression data (from SP2 experiments being undertaken at NIAS in Japan).

CIP, ICRISAT, ILRI ,and IRRI have purchased Paracel HPC clusters cross-linked into a global grid facility. These tools will be documented with basic bioinformatics help/tutorial instructions.

Furthermore, tools are under development at EMBRAPA based on CORBA to experiment with the universal adaptors of well-known public domain packages that might be put together to store, visualize, and analyze genomic sequences and ESTs. This adaptation allowed the addition of new parts and evolution of the system.

### ***Subprogram 5***

This Subprogram was the latest addition to the Challenge Program and, as such, was not discussed as a separate subprogram in the planning meeting held in Wageningen in August 2003. Because of that, a strategy document was prepared this year by the Capacity Building Coordinator, Carmen de Vicente, to describe the objectives of SP5 as well as the underlying principles that will guide the preparation of the work plan for the following years. The strategy document details:

- Who the beneficiaries of capacity building activities in the GCP are
- Capacity building schemes that will be used in the GCP
- Why needs assessment activities are the basis of our work plan
- Partnerships and alliances that will be sought to guarantee wider impact in our endeavor
- Coordination mechanisms that will be put in place to ensure coherence in capacity building activities throughout the GCP
- Need for monitoring and measuring impact
- Plans for mobilizing extra resources to fulfill our objectives in the subprogram.

During the Wageningen meeting, a few activities were embedded in Subprogram 1 to be carried out in year 1 as a means to prepare the development of a thorough capacity building work plan for the duration of the GCP. One of these activities was the organization of a workshop to assess the capacity/training needs of candidate partners of the GCP, representing a wide array of developing country institutions (NARS and universities mainly). Potential participants to the workshop have been selected based on nominations submitted by GCP members. The workshop, with approximately 35 participants, will be held the week of August 2<sup>nd</sup> in Costa Rica.

In addition to that, different capacity building activities were included in the year 1 work plan of the thematic subprograms and have been already conducted in the course of 2004. A workshop was held in Nairobi (ILRI) during the last week of April entitled “CAGT –

Crops with Appropriate Gene Technologies.” The workshop brought together for the first time all of the genetic transformation specialists in the CGIAR as well as several outside experts to discuss common approaches, challenges, and opportunities. The occasion was used to discuss good ideas for joint preparation of pre-proposals. The Generation Challenge Program Information Systems and Network Design Workshop (SP4, Mexico 31 May-4 June) gathered around 20 experts from within and outside the GCP to develop a comprehensive architectural blueprint for the GCP platform, network, and data registry. Following, a Paracel training course for the CGIAR GCP group took place in Pasadena, CA from June 7<sup>th</sup> to 9<sup>th</sup>. Another workshop took place during the week of the 21<sup>st</sup> June in Zaragoza, Spain. Recommendations for genotyping data analysis and definition of guidelines for the selection of germplasm to advance to SP2 and SP3 were the two basic outputs of the workshop. A second workshop was held focusing on phenotyping for drought stress tolerance during the week of July 5<sup>th</sup> in Montpellier, France. The workshop brought together scientists from the CGs, ARIs, and NARS (40, 30 and 30%). In addition to the technical conclusions, the workshop indicated opportunities for progress including exchange of scientists, PhD theses, and preparation of training courses.

At the outset, we recognized that training and capacity building needs go beyond the transfer of tools and knowledge to NARS. The capacity building activities of the GCP should involve expanding the expertise of CGIAR Center researchers and enhancing the understanding by ARIs of the practical challenges and limitations confronting NARS and CGIAR staff. Thus, we have built in Year 1 activities to implement this “reciprocal” training process. An example of this activity is the on-going collaboration on gene expression between NIAS and IRRI. Experiments are designed and being executed to take advantage of the extensive experience of NIAS in gene chip technology. Three researchers from IRRI are currently working at NIAS (for 2-4 weeks each) to generate data on drought-response gene expression and test the utility of rice gene chips for heterologous hybridization. Such a shuttle research arrangement is being planned for other GCP partners to work at NIAS as well.

Still to come, an Intensive Training Program in Molecular Breeding will be conducted late 2004 at the Biosciences Center in ILRI (Nairobi) in collaboration between the GCP and NEPAD. The course aims at providing training tailored to the needs of the participants and effectively becoming a part of their research or breeding programs. It is hoped that the training program will become intimately linked with competitive research grant programs to ensure a continuity of applications and to support a step-wise increase in national capacity in this area. In conjunction with this activity, SP3 will design a syllabus for a course on marker assisted selection and breeding, which will be the basis for the training plan in this subprogram.

In SP4, several capacity building activities are pending for 2004:

- A survey to be carried out among the Consortium Member institutions to gather information concerning:
  - the definition of areas of expertise related to SP4 (data-basing, bioinformatics, platforms, analysis software, etc...)
  - existing tools (training materials, etc...) and delivery mechanisms available, both within and outside the CP Consortium

- the available expertise within the CP Consortium members for each of the areas of expertise identified
  - the existing gaps in expertise within the CP Consortium as stated by the member's focal points
  - the identification of sources of expertise available outside the CP Consortium.
- Online training for GCP partners who did not attend the meeting in California who are interested in Paracel clusters and cross integration.

Finally, a number of scientists from NARS outside of the Consortium will actively participate in the research activities planned for year 1 in the different subprograms. While most of them are ongoing, an extract of activities may be found in the work plan of year 1 pages 12 to 15. A similar scheme will continue for the rest of the GCP, as the participation of NARS was introduced as a basic criterion for the selection of pre-proposals submitted to the first competitive call launched in April 2004.

Results of the needs assessment workshop to be carried out in Costa Rica in August 2004 will provide a strong base to move the capacity building program forward. A concrete work plan will be put together to ensure that all different aspects of capacity building are included: availability of training materials, basic curricula for training courses, establishment of fellowship and travel grant programs supporting ongoing research projects, establishment of reference centers for particular topics, and approaches to link with existing capacity building institutions, among others.

## **Administration and Governance**

It is essential that there be an effective management and governance system in place to assure that the GCP funds are managed efficiently and effectively, that they are allocated in a transparent and orderly fashion, and that the research and capacity building agendas are executed in a high quality way. This is a particular challenge for virtual programs like the GCP, since they do not have legal status, nor do they have physical infrastructure in which to execute their research agendas. Thus, in this first MTP it is appropriate to summarize briefly our progress in establishing the mechanisms that will assure a smoothly functioning Challenge Program. The major accomplishments are summarized below.

### ***Governance***

- Established a Program Steering Committee
  - Process in place for the orderly change in Chairman
- Recruited Program Advisory Committee recruited following consultation with Program Steering Committee
  - Terms of Reference (TOR) for the PAC developed
- Selected a Stakeholders' Committee with GFAR after detailed and transparent selection process led by GFAR
  - TOR for Stakeholders' Committee developed with GFAR

- First year work plan agreements
  - Include reporting requirements
- Developed Consortium Agreement after exhaustive consultation with Consortium members to reconcile needs and requirements of CGIAR and no-CGIAR members (and after detailed inputs from the private sector) which includes major sections on:
  - Responsibilities of Consortium membership
  - TOR for PSC
  - IP policy
- Quantified in-kind contributions of CGIAR centers to GCP
- Fund and resource mobilization yielded major investment by DFID, with strong expectations that Rockefeller Foundation and Pioneer Hybred/DuPont Company will contribute over the next year

### ***Research Management***

- A Director was internationally recruited and appointed, followed by appointment of a communications assistant and 50% time secretary
- The Directors office oversaw execution of first year work plan
- Funds for first year were disbursed following the development of terms and letters of agreement between GCP institutions and CIMMYT/GCP Director's office.
- Recruited and appointed research management team following a competitive, transparent search from within GCP consortium members (3 Subprogram Leaders from CGIAR centers and 2 from non-CGIAR GCP members)
- Designed and conducted a transparent competitive grants program
- Developed capacity building strategy
- Needs assessment consultation with NARS
- Established strategic linkages with other CPs and NEPAD Biosciences (BECA)
- Undertook a wide range of intensive consultations with GCP participating scientists to establish norms and standards for research approaches and protocols, and data management
- CIMMYT provides excellent accounting and other managerial support

### ***Communications***

- Developed communications strategy
- Created logo and developed new name and identity
- Established a web site as the GCP's face to the world ([www.generationcp.org](http://www.generationcp.org))
- Established a virtual work space
  - Archive and share GCP documents
  - Discussion platform
  - Edit major documents
  - Securely manage competitive grants program (post pre-proposals and full proposals for convenient access by reviewers)
- Developed communications materials
  - Brochure & poster
  - New look to Web site
- Developed a monthly e-newsletter
- Quarterly reporting to PSC

- GCP sessions and/or presentations in 6 major international meetings; Director and Subprogram Leaders presented GCP to scientific and donor community on over 25 occasions

### ***Linkages***

- Establishment a set of joint activities with Harvest Plus and Water for Food Challenge Programs
  - Regular consultations with both to identify areas of common interest
  - All CPs met in Rome for IP workshop with CAS
  - GCP and Harvest Plus have agreed to adopt a common end-user strategy and approach in target areas (initially east Africa and South Asia)
    - Harvest Plus to take lead in Africa and GCP in South Asia
  - GCP and Harvest Plus to develop a common crop (molecular marker) improvement capacity building and community of practice program for East Africa out of the NEPAD BECA facility at ILRI. In all ventures together, both GCP and Harvest Plus will commit resources.
- Formal agreement with the NEPAD BECA facility in ILRI to jointly develop a capacity building strategy and other activities of mutual interest
- Identification of a common need/interest with McKnight Foundation to assess and then design a new generation of crop improvement networks tailored to the specific nuances of comparative biology and genomics targeting plant breeding
- Have developed a funding and research strategy with Rockefeller Foundation to move the RF-funded Resilient Crops program into the GCP research arena over the next three to four years

### ***Intellectual Property***

- Worked with CAS and their legal consultants to develop draft IP policies for the GCP Consortium Agreement
- Have arrived at a broadly acceptable IP policy after several rounds of fruitful discussion
- Are in serious discussions with CAMBIA to have the GCP be an early and pilot adopter of the BIOS approach to open source biotechnology

### ***Private Sector***

- Regular interaction (in coordination with Harvest Plus) with the Private Sector Committee of the CGIAR
- Held intensive two-day consultations with Pioneer Hybred staff to discuss possible interactions with GCP. Developed simple test case projects involving maize.
- Requested and received detailed feedback from Pioneer on the private sector opinion of our IP policies. From those discussions, it seems likely that it will be possible to work out mutually acceptable terminology.
- Negotiated with Paracel Corporation very favorable terms for high performance cluster computer systems for four CGIAR centers, including access to their global Web-based supercomputing grid for all members of the GCP

## **Financial Highlights**

As stated in our CGIAR-approved founding document, we are committed to awarding research support in two ways: an externally peer-reviewed competitive grants program and a commissioned grants program. The former is intended to attract the highest quality science to key problems and opportunities presented by the GCP. The commissioned work is designed to assure that the basic technical platforms are in place to support the over all GCP mission, that there is a robust and coherent capacity building effort, and that critical research areas not covered by competitively awarded grants are supported. Both are open to non-consortium member participation, but at this time the lead PI and submitting institution must be Consortium members.

One consequence of the research funding approach for the GCP in its first year is that it is not in synchrony with the timing of the Science Council Review and ExCo assessments. The GCP management committee will determine the rankings of the competitive grant proposal based on external peer review and will solicit commissioned proposals for decision around the same time. These will be reviewed and approved by the Program Steering Committee in November. Thus, while we will have an excellent picture of the nature of the work that we will be conducting, as indicated in the outputs, activities, and timelines (Annex 1), the specifics of who (which CGIAR centers and non-CGIAR institutions) will be conducting the research cannot be known at the time of the writing of this MTP. Therefore, CGIAR centers will not be showing GCP allocations in this years' financing plans, nor will we be showing significant allocations to CGIAR institutions in our financing plans. However, we are prepared to issue an updated financing plan in the first two months of 2005 to reflect those allocations.

The following tables shows actual fund allocation for 2004. The funds expected for 2004 are highly likely (based on written communications from donors). The amounts for the out years reflect reasonable expectations based on discussions with new potential donors. For example, we have had very promising discussions with the Rockefeller Foundation regarding support of significant drought-related research programs for the next three years. This potential support is not shown in 2005 but is anticipated for 2006 and 2007. Likewise we will be actively seeking support for ex ante impact and policy analyses.

The CGIAR Secretariat instructed the Challenge Programs to roll our detailed financial tables into the financial tables of the host center. Therefore the details of GCP financial distributions across sector, commodity, and functional categories are embedded in the CIMMYT financial tables in their current MTP. We have prepared the following tables to indicate to the reader how our funds are distributed within the GCP, by CGIAR outputs, and CGIAR outputs by donor.



### Financial status and projections for the Generation Challenge Program

Subprogram	Actual 2004	Plan (US\$ '000)		
		2005	2006	2007
SP1: Genetic Resources Characterization	3,266	3,900	3,000	2,500
SP2: Gene Discovery	1,355	1,600	3,000	4,000
SP3: Trait Capture for Crop Improvement	795	1,600	3,000	4,000
SP4: Bioinformatics systems	2,550	3,900	3,500	3,500
SP5: Capacity Building	420	2,300	2,500	2,500
Total (100% financed by donors)	8,386	13,300	15,000	16,500
Documented In-Kind Contributions (CGIAR)	4,730	3,200	3,200	3,200
Documented In-Kind Contributions (ARIs)	1,028	2,000	2,000	2,000

### Allocation of 2005 GCP donor-supplied resources across CGIAR outputs.

GCP Subprogram	US\$ ('000)	Germplasm Collection & Conservation	Germplasm Improvement	Sustainable Production	Policy	Enhance NARS
SP1	3,900	0.9	0.1			
SP2	1,600	0.6	0.3	0.1		
SP3	1,600		0.7	0.2	0.1	
SP4	3,900	0.5	0.3		0.2	
SP5	2,300					1
Total	13,300 <sup>1</sup>	6420	3160	480	940	2300

<sup>1</sup>In kind contributions follow the same allocation pattern.

**Allocation of 2005 donor contributions across GCP projects. US\$ ('000)**

CGIAR Output	WB	DFID	EC	Sweden	Austria	Total
Germplasm Collection & Conservation	380	2,260	3,630	100	50	6,420
Germplasm Improvement	1,500	1,660	0			3,160
Sustainable Production	240	240				480
Policy	500	440				940
Enhancing NARS	380		1,920			2,300
<b>Total</b>	<b>3,000</b>	<b>4,600</b>	<b>5,550</b>	<b>100</b>	<b>50</b>	<b>13,300</b>

**Generation Challenge Program funds assigned to CGIAR centers for research purposes during calendar year 2004**

CGIAR Center	Amount (US\$)
CIAT	777,735
CIMMYT	770,696
CIP	440,492
ICARDA	287,669
ICRISAT	404,667
IITA	792,202
IPGRI (CAS & INIBAP)	1,037,001
IRRI	714,777
WARDA	22,219
<b>Total</b>	<b>5,247,468</b>

## **ANNEX 1. GCP SUBPROGRAMS AND PROJECTS (OUTPUTS)**

**Objective 1:** Develop a platform for, and conduct, analysis of genetic diversity in international crop genetic resources and apply this to improve major crops for drought tolerance and other related traits of importance to resource-poor farmers

### **Subprogram 1: Genetic Diversity of Global Genetic Resources**

#### **Rationale**

The broad genetic diversity of global genetic resources provides the materials for varietal improvement. The Generation Challenge Program undertakes in-depth studies to better understand the diversity of traits and modes of adaptation and to access the corresponding genetic factors for use in varietal improvement. Subprogram 1, focused on genetic diversity, is organized for identifying novel and diverse variants at genes involved in traits of interest in breeding, for use in SP3. In accordance with the global priority, particular emphasis is given to drought tolerance. All components involve the participation of scientists and students from national cooperating institutions and universities, under the orchestration of SP5.

The various partners hold and manage large germplasm collections, in addition to those held by other national programs. The CG center collections, held in trust under the auspices of FAO, are the largest, but gains can be made by complementing these from other sources. Out of the twenty two crops potentially concerned, several groups will be considered on the basis of the biology, the resources available and the spontaneous involvement of the community. Some will receive more attention: the "priority crops" will be used for the most advanced and integrated experiments (including accurate phenotyping), and the "representative crops" will be used for methodological developments.

The first step of the work plan aims, for a given crop, at identifying a sample derived from all sources available (the "composite set") to represent the diversity of the crop and its wild relatives. The information for this first stage of sampling consists of data available from standard management practices for the various contributor collections. Through this process, distinct groups of germplasm, such as wild populations, landraces, cultivars and breeding materials, can be represented.

The composite set must then be surveyed with anonymous molecular markers. These must be perfectly standardized, so that simple, high throughput analyses are possible for use in multiple laboratories, as well as anywhere in the future. This will set a standard for describing the structure of the crop germplasm, in relation to which any new sample can be located with a simple analysis. This structure represents predominant trait and marker combinations and helps identify diversity-rich components of the germplasm.

The information gathered on the composite set serves for selecting a "reference sample" that will be the preferred material for advanced characterization. This selection has to reconcile several objectives:

- Exploring all the components of the global diversity

- Promoting populations of materials with continuous variation for best highlighting functional relationships among polymorphisms
- Focusing on materials whose within-sample diversity enables accurate characterization at the molecular and phenotypic levels
- Encompassing an array of documented ecological adaptations for planning and conducting meaningful phenotyping experiments.

The reference sample represents the material for subsequent molecular and phenotyping characterization efforts. Integration of accumulated information will allow the detection of associations between molecular polymorphisms and important agronomic features and, in the longer term, models for crop adaptation for these traits can be developed.

Identification of molecular polymorphisms can focus on candidate genes (thought to be involved in reaction to drought) identified with SP2, so that their involvement can be tested and the most favorable alleles detected. Or, identification of molecular variants at sufficient density across the whole genome could highlight unsuspected regions explaining part of drought tolerance. In both cases, the cost and efficiency of these new techniques can be improved through these efforts.

Phenotyping experiments have to be designed that optimize the combination of germplasm surveyed and traits measured, taking into account the contributions to drought tolerance, the expected heritabilities, the breadth of genetic control, and the difficulty of the assessments. An expert group will be created, to define the ontology and formats for the description of traits, experiments, environments, etc., with the view to establish a drought tolerance database.

The analysis of associations will be assisted by early assessments of the level of linkage disequilibrium in crops spanning the breadth of the biology represented in the CP. Methodological developments by SP4 will adapt the association study software to the range of target situations. Specific populations will be developed for those crops that display important within-accession variation, for subsequent characterization and association analysis.

Protocols will be developed to ensure access to and benefit sharing for derivatives from accessions, in line with the Convention on Biological Diversity and the International Treaty on Plant Genetic Resources, on a sustainable basis beyond the life of the GCP.

**Output (1.1 in Logframe, Annex 2):** Organize a rational access to sources of genes and alleles involved in key agricultural traits on the basis of coordinated surveys of molecular and phenotypic variation among the accessible germplasm collections for CG mandate crops.

### **Impact**

- The discovery of original sources of diversity and novel alleles for crop improvement will be facilitated through linkages between various germplasm banks and explicit rules of access and benefit sharing.
- A synergistic interplay established between genomics and germplasm science based on identifying appropriate germplasm and refining choices in candidate genes for testing functional hypotheses.

- Long term integrated multidisciplinary characterization of key germplasm will be undertaken for a better understanding of trait and crop adaptation, leveraging comparative genomics among related groups of species.
- For each of the major world food crops, researchers will have access to a reference sample for capturing new sources of genetic variation and validating the effect of particular genes, alleles, and chromosome segments for inclusion in breeding programs.

### **Activities**

**Activity 1.1.1:** An improved understanding of the structure of the diversity for the major world food crops, diagnostic molecular markers for subsequent germplasm analysis identified, and a set of reference samples designed for integrated characterization

**Activity 1.1.2:** A range of techniques accessible in key laboratories applied for high throughput molecular characterization of germplasm

**Activity 1.1.3:** Establishment and implementation of a scientific and organizational framework to describe tolerance to drought

**Activity 1.1.4:** Molecular polymorphisms associated with higher tolerance to drought; integration of methodological improvements into these studies

**Activity 1.1.5:** Protocols established to allow germplasm exchange with proper access and benefit sharing from the derivatives of the program

### **Indicators (Duration/milestones)**

**Activity 1.1.1:** An improved understanding of the structure of the diversity and diagnostic molecular markers for the major world food crops

#### **2005**

- Publication of the description of molecular variation and germplasm classification for rice, wheat, cassava, cowpea, common bean, and *Musa*.
- Reference samples with molecular, passport, and general adaptation descriptors recommended for integrated phenotyping for rice, wheat, cassava, cowpea, common bean, and *Musa*.
- Guidelines and software for selecting materials for phenotyping experiments using the reference samples established.

#### **2006**

- Description on the GCP website of standard molecular methods for the characterization of germplasm in pearl millet, finger millet, sweet potato, yam, ground nut, lentil, pigeon pea, soybean, coconut, and selected forages.
- Publication of the description of molecular variation and germplasm classification for barley, maize, sorghum, potato, and chickpea.

- Reference samples with molecular, passport, and general adaptation descriptors recommended for integrated phenotyping for barley, maize, sorghum, potato, and chickpea.

**2007**

- Publication of the description of molecular variation and germplasm classification for pearl millet, finger millet, sweet potato, yam, ground nut, lentil, pigeon pea, soybean, coconut, and selected forages.
- Reference samples with molecular, passport, and general adaptation descriptors recommended for integrated phenotyping in pearl millet, finger millet, sweet potato, yam, ground nut, lentil, pigeon pea, soybean, and selected forages.

**Activity 1.1.2: A range of techniques accessible in reference laboratories applied for high throughput molecular characterization of germplasm**

**2005**

- Assessment of various genome-wide molecular characterization techniques by small scale comparative surveys among crops yet to be characterized with microsatellite and/or other markers.
- Selection of a range of representative crops for medium scale genome-wide characterization on the basis of existing linkage disequilibrium.
- Comparative assessment of targeted molecular characterization techniques (SNPs or integrated haplotypes in candidate genes) using several reference samples and several sets of orthologous drought tolerance candidate genes.

**2006**

- Assessment of genome-wide molecular characterization techniques (continued).
- Methodological analysis of medium scale genome-wide characterization experiments for representative crop reference samples.
- Assessment of targeted molecular characterization techniques (continued).
- Methodological analysis of medium scale targeted characterization for representative crop reference samples.

**2007**

- Methodological analysis of large scale genome-wide characterization experiments for priority crop reference samples.
- Identification of reference laboratories for genome-wide characterization.
- Methodological analysis of large scale characterization of priority crop diversity for drought tolerance candidate genes.
- Identification of reference laboratories for targeted molecular characterization.

**Activity 1.1.3: Establishment and implementation of a scientific and organizational framework to describe tolerance to drought**

**2005**

- Finalization of trait ontology for drought tolerance assessment.

- First testing (general adaptation, phenology, morphology) and seed multiplication of several reference samples in potential drought tolerance assessment locations.
- Elaboration of a global strategy for coordinated phenotyping, including selection of priority crops.

#### **2006**

- Upgrading of selected phenotyping facilities.
- First comprehensive assessment for priority crops.
- Establishment of a database for multi-location/year evaluations.
- Refinement of the strategy for coordinated phenotyping; identification of the first centers of excellence for phenotyping.

#### **2007**

- Additional upgrading of selected phenotyping facilities.
- Repetition of assessments for priority crops.
- Continued refinement of the strategy for coordinated phenotyping within a network of centers of excellence.

### **Activity 1.1.4: molecular polymorphisms associated with higher tolerance to drought, identified from integration of methodological progress in genotyping and phenotyping**

#### **2005**

- Identification of priority crops for linkage disequilibrium studies.
- Identification of priority candidate genes and their orthologs in representative species.
- Assessment of allelic diversity for priority candidate genes in representative species.
- Selection of experimental populations on the basis of adaptation, phenology and morphology among reference samples.

#### **2006**

- First medium-intensity genome-wide molecular characterization and assessment of the resolution power of linkage disequilibrium.
- First large-scale diversity assessment for an initial group of candidate genes in a few priority crops.
- First phenotyping experiments covering significant parts of reference samples for priority crops.

#### **2007**

- First association patterns between traits and molecular diversity.
- First assessment of linkage disequilibrium-based mapping approach in priority crops.
- First test of the involvement of candidate genes in the adaptation to drought tolerance.

**Activity 1.1.5: protocols to allow germplasm exchange and proper access and benefit sharing from the derivatives of the program**

**2005**

- Analysis of experience within the consortium and collaborating NARS for the circulation of germplasm for genotyping and phenotyping.
- Analysis of experience of collaboration with private partners.
- Improved protocols in line with the published policies of access and benefit sharing of the CBD and the FAO ITPGR.
- Amended consortium agreement.

**2006**

- Further analysis of international collaborations involving germplasm exchange.
- Protocols for public access to genotyping/phenotyping technologies developed within the GCP.

**2007**

- Protocols for public access to molecular tags of useful polymorphisms developed within the GCP.
- Synthesis and publication of the CP experience on international cooperation in genomics-based germplasm enhancement.

**Users/Beneficiaries**

- The GCP consortium members
- The NARS and ARIs genebanks and breeding programs

**Collaborators**

- The GCP consortium members and the collaborating NARS
- All parties involved in varietal improvement for the GCP crops
- All parties involved in plant genetic improvement for drought tolerance

**Collaborators (outside consortium)**

WARDA, CFP (Bolivia), CORPOICA (Colombia), Univ Nacional (Colombia), INTA (Argentina), Carbap (Cameroon), Ceraas (Senegal), IPK (Germany), SCRI (Scotland), University of Idaho (USA), University of Leicester (UK), DArT Pty Ltd (Australia), University of Western Australia (Australia)

<i>Linkages to CGIAR Research Output (Percent of Resources)</i>			
Germplasm collection and conservation (including evaluation and characterization)	80	Water and Food	0
Germplasm improvement		Harvest Plus	10
Sustainable production systems			
Policy research and development	10		
Enhancing NARES capacity	In SP5		
Financing Plan (US\$'000)	3,900		



## Subprogram 2: Comparative Genomics and Gene Discovery

### Rationale

This Subprogram focuses on genomic tools, technologies, and approaches to achieve an understanding of gene systems across many species of importance to developing country agriculture.

The CGIAR and NARS as a whole have the unique advantage to produce extensive phenotypic data across economically important species and relate such information to molecular variation by genetic mapping or association analyses. For almost all of the CGIAR mandate crops, quite extensive genetic maps within the species are already available. Comparative maps are also available for many CGIAR crops with their models. These are most advanced for the cereals (rice, sorghum, maize, barley, millet, wheat), but potato-tomato and some inter-legume comparisons are also available. Beyond mapping, an expanding amount of sequence information as well as improved informatic tools have enabled us to identify orthologous genes that are evolutionarily conserved in function. Thus, we will be increasingly using sequence information (gene content) to predict gene functions across large families of plant species.

The primary role of this subprogram is to discover and validate the function of key genes central to the practical objectives of the GCP. We will bring together an array of genetic resources and apply analytical tools to anchor gene expression polymorphism (transcripts, proteins, and metabolites) onto genetic variation (in coordination with SP1) and relate expression and genomic polymorphisms to phenotypic performance (in coordination with SP3). This will be achieved by conducting experiments across different crop groups to reveal gene function through comparative analyses. The primary activities will include:

- Examining gene expression (in a broad sense—including transcripts, proteins, and/or metabolites—depending on the robustness of the technology) using phenotypically-informative crops or genotypes within a crop
- Validating the causal relationship between expression and phenotype using appropriate genetic stocks (targeted mutations, over- and under-expression systems)
- Relating expression-phenotype relationship to genomic variation (with SP1)

With sequences of an entire genome such as *Arabidopsis* and rice, and extensive sets of expressed sequences for many other crops available, cDNA microarrays or oligo gene chips are now much more accessible than a few years ago. By examining the expression of many genes simultaneously under a specific condition, e.g., under drought stress, the complex interactions of different biochemical pathways can be understood and genes identified that are responsible for improving a complex trait. The same argument applies to other means of genome-wide expression analyses, including protein and metabolite profiling. The resulting information facilitates the dissection of genetic and metabolic systems of the organism. Structural and functional conservation in gene regulatory circuits is particularly relevant for identifying genes with large effects on phenotypes. The effect of transcriptional factors is well illustrated by the DREB (dehydration responsive element binding) gene family. Manipulation of such regulatory elements can

bring about dramatic changes in phenotypes that are often viewed as being controlled by many genes with minor effects.

Several recent advances are particularly relevant to this Subprogram. First, the power of SNP analysis in pinpointing the genetic control of complex traits has been demonstrated in medical genetics (e.g., use of SNP data to identify genes causing asthma and regulatory functions in disease immunity). Second, genome-wide segregational analysis has been used to determine the causal relationship between QTL and gene expression (Schadt et al 2003, *Nature* 422:297-302); such an approach could help alleviate the problem of limited predictive power of low resolution maps. Third, conserved co-expression in diverse biological systems has been used successfully to reveal gene function (Stuart et al 2003, *Science* 302:249-255). Subprogram 2 must therefore be positioned to capture these innovations and future conceptual advances and apply them to crop genomics.

Recognizing that funding decision has not yet been made at the time of MTP submission, our work plan is based on an analysis of the selected pre-proposals in the competitive grant program and an assessment of the critical areas that merit commissioned research. The work plan, though tentative in nature until funding decisions are made, attempts to harness the collective strength of participating researchers to achieve overall program objectives. The 3-year work plan will concentrate on the following medium-term objectives:

- Assemble genetic knowledge on stress tolerance (drought and others) through a combination of informatics analysis and empirical studies using advanced genetic stocks available to or developed by the GCP.
- Mobilize high-quality phenotyping of selected genetic materials for functional validation. This will include systematic phenotyping of mutants and evaluation of NIL with unique chromosomal segments for functional validation.
- Conduct parallel studies on stress tolerance using well-developed biological systems to yield near-term results. This will include drought as well as other stress tolerance traits as a basis to investigate drought tolerance.
- Investigate possible interactive effects (synergistic or antagonistic) of stress responses, recognizing that multiple stress tolerance is needed in delivering pre-breeding products to end-users.
- Contribute to the development of genomic tools and “designer” genetic stocks in selected crops to enable applications of successful approaches.
- Initiate SNP analysis in target genes (orthologs and paralogs) in multiple species.

### **Output (1.2 in Logframe, Annex 2)**

Genes conferring stress tolerance, with emphasis in drought, in multiple crop species identified through the use of cross-cutting tools developed for comparative genomics.

### **Impact**

- Genetic resources and tools available for trait discovery in multiple crop species
- Comparative genomics approaches developed for multiple species

- Improved understanding of complex traits by genome-wide approaches to assessing gene function and expression
- Identification of specific genes/markers with breeding values verified by multiple functional evidence
- Improved breeding efficiency by using a large pool of candidate genes identified by functional genomics
- NARES partners trained in applying functional genomic tools

### **Activities**

**Activity 1.2.1:** Assembly of genomics and germplasm resources through consolidating existing (and developing new) specialized genetic stocks and framework genetic markers for target crops.

**Activity 1.2.2:** Develop comparative maps *within* and *across* species and deploy comparative mapping tools to CP partners, linked to the CP consensus map repository and major international plant databases

**Activity 1.2.3:** Assign genes and pathways to putative phenotypes through the convergence of genome variation, expression patterns and phenotypic data

**Activity 1.2.4:** Validate genes and pathways through evaluation of under- or over expression constructs or variants (induced or natural) of the target genes

**Activity 1.2.5:** Genomic resources and information platform disseminated to NARES through networks and training (linked to subprogram 5)

### **Indicators (Duration/Milestones)**

**Activity 1.2.1:** Assembly of genomics and germplasm resources through consolidating existing (and developing new) specialized genetic stocks and framework genetic markers for target crops

**2005**

- Detailed phenotypic characterization of selected lines and populations in selected species to serve as the basis for genetic studies.
- Consensus developed for adoption of drought-related screening and evaluation protocols by participating research teams and standard conditions for phenotypic parameters determined.
- Set of drought phenotypic data for major cereals (maize, sorghum, pearl millet, and rice) established to enable identification of candidate genes using genomic sequences of rice and/or maize.
- Sets of putative orthologous candidate genes for drought resistance extracted from existing genome resources and assembled (with SP4).
- Network established for systematic phenotyping of rice mutants under multiple conditional stresses.
- Initiation of production of mutant stocks in selected crop genotypes (e.g., cowpea, sorghum, and other crop species amenable for mutagenesis) suitable for forward and reverse genetics (TILLING) as a long-term resource for gene identification and validation.

## 2006

- Phenotypes of sequence-indexed mutant lines determined for a set of stress tolerance genes.
- NIL established for detailed phenotypic and expression analysis of abiotic stress tolerance (e.g., drought, salinity, P-deficient, aluminum toxicity) in selected cereals.
- NIL with defined chromosomal QTL for biotic resistance produced for validating phenotypic and expression analyses.
- TILLING populations evaluated in selected crop species.

## 2007

- Assembly and documentation of candidate gene collections for drought resistance-related pathways.
- A large collection of induced allelic variants in different mutant stocks available in selected crops.

### **Activity 1.2.2: Develop comparative maps *within* and *across* species and deploy comparative mapping tools to CP partners, linked to the CP consensus map repository and major international plant databases**

## 2005

- Genetic maps based on populations segregating for drought tolerance of target crops completed.
- Comparative analysis of genotypic and phenotypic data to infer gene/QTL for drought tolerance positions on crop genetic maps completed.
- COS and other consensus markers tested and validated to develop and enhance comparative maps within and across crop families.
- Comparative maps of selected dicots and monocots created or enhanced with 100 COS or conserved EST/SSR markers.
- Increased precision of location of drought-related QTL for four target species (maize, sorghum, pearl millet, and rice).

## 2006

- Detailed comparative map of drought-QTL across maize, sorghum, pearl millet, and rice.
- Determination of synteny of common drought-responsive genes in selected tuber species.
- Data mining of the information publicly available about drought tolerance-QTL locations on different crop genetic maps.

## 2007

- Sequence variation of orthologs spanning taxonomic range of monocots determined (initially across rice and *Musa*).
- Cross-species maps anchored with orthologous genes with evidence of conferring stress tolerance in targeted crops.

**Activity 1.2.3: Assign genes and pathways to putative phenotypes through the convergence evidence of genome variation, expression patterns, and phenotypic data**

**2005**

- Available gene/sequence data for stress-response pathways across species consolidated.
- EST libraries enhanced in cowpea, *Phaseolus*, cassava, and millet and made available for sequence comparison and expression analysis.
- Gene expression profiles established for selected phenotypes and crop genotypes, (including rice, maize, wheat, and barley) and identification of common and unique genes correlated with phenotypic expression across species.
- Gene expression analysis of reproductive stage(s) (specifically relevant to drought tolerance) conducted to reveal subsets of genes correlated with drought tolerance.

**2006**

- Expression QTL (eQTL) and their interactions determined based on segregational analysis in well characterized segregating population (e.g., recombinant inbred lines segregating for multiple QTL) of selected species.
- Diagnostic subarrays with ~1,000 genes (including stress-responsive genes and regulatory factors) developed and evaluated for expression analysis and/or genotyping.
- Comparative datasets of plant genes involved in hormone metabolism and response categorized into gene families assigned to pathways and crop physiology models.
- Drought tolerance genes/pathways/models of other crop clusters (cereals, root crops, legumes) compared and differences and similarities in genes/pathways among crops determined.
- Co-regulated genes under drought stress in multiple species (wheat, rice, maize, barley) determined by comparative data analysis.
- Candidate genes conferring tolerance to soil toxicity/deficiency (e.g., Al, P) tolerance genes from sorghum, maize and wheat identified and characterized.

**2007**

- Determination of metabolites specifically linked to carbohydrate metabolism, signaling pathways, ABA regulatory pathways, and accumulation of osmolytes and antioxidance.
- SNP haplotypes (of maize lines determined and pattern of association between phenotypes, metabolites profile and SNP haplotype to reveal causal relationship
- Candidate genes controlling terminal drought tolerance (e.g., peduncle elongation and sterility) identified.
- Orthologs of genes conferring salinity and phosphorus-deficiency tolerance identified across species as a prelude of gene cloning in multiple species (legumes, cereals).
- A better understanding of the diversity and functioning of molecular and physiological mechanisms of crop tolerance to soil toxicity in drought-prone environments.

**Activity 1.2.4: Validate genes and pathways through evaluation of under- or over-expression constructs or variants (induced or natural) of the target genes**

**2005**

- Panel of candidate genes established for searching knockout mutants in rice and other mutant collections.
- Mutants with loss of QTL in disease resistance identified as a validation case study for other complex traits.
- Gene pyramids for target traits developed as a means for confirming phenotypic effects of alleles.

**2006**

- Function of disease defense candidate genes within chromosomal segments responsible for QTL tested by over-expression or knock-out mutants.
- Function of the identified candidates for salinity and phosphorus-deficiency tolerance confirmed by complementation test.
- Function of drought-tolerance candidate genes evaluated in gene pyramids, knockout, or over-expression lines in target crops.
- Interaction of biotic and abiotic stress tolerance genes determined in selected case studies.

**2007**

- Predicted phenotypic value of alleles confirmed by gene pyramid in elite lines for selected traits (biotic and abiotic stresses).
- Databases established based on gene array and/or proteomic analyses (linked with SP4).
- Allele-indexed, stress-tolerance breeding lines available to SP3 for implementation of MAS (e.g., combining P-deficiency tolerance with drought/disease resistance lines).
- Orthologous stress-tolerance genes identified in target crops (cereals, legumes, and tubers) supported with experimental data.

**Activity 1.2.5: Genomic resources and information platform disseminated to NARES through Networks and training** (linked to subprogram 5)

- See SP5 milestones

**Users/Beneficiaries**

The outputs of this project will provide methodologies to discover a large suite of genes, particularly in stress tolerance, useful for crop improvement, hence serving a wide range of users. The genetic stocks produced will be a permanent resource for researchers worldwide. The genes and traits discovered will be used by a global community of researchers and their collaborators in NARES and ARIs and breeders from the public and private sector.

**Collaborators**

CGIAR Centers and other public research initiatives and institutes, advanced research institutes and NARS inside the consortium. Beijing Genomics Institute, Kansas State

University, Virginia Tech, Montana State University, University of California-Riverside, University of California-Davis, Scottish Crops Research Institute (SCRI), International Center for Genetic Engineering and Biotechnology (New Delhi), African Centre for Gene Technologies (Pretoria), University of Adelaide, and CSIRO (Australia), among others.

Linkages to CGIAR Research Outputs (Percent of Resources)	
Germplasm collection and conservation (including evaluation and characterization)	60
Germplasm improvement	30
Sustainable production systems	10
Policy research and development	
Enhancing NARES capacity	In SP5
Financing Plan (US\$'000)	1,600

### **Subprogram 3: Trait Capture for Crop Improvement**

#### **Rationale**

This subprogram focuses on the validation and refinement of molecular breeding systems and the resultant enhanced germplasm with the primary purpose of increasing the efficiency, speed, and scope of plant breeding gains. This includes a substantial commitment to create appropriate technologies for application of marker-assisted selection in national breeding programs, to provide technical assistance for the rapid and effective uptake of molecular breeding in tropical staple crops and to foster the development of communities of practice supported by regional centers of excellence and state of the art technologies and approaches.

The development of effective systems for breeding complex traits such as drought tolerance has eluded most practitioners despite a great deal of R&D investment which for some crops has spanned more than 50 years. The recent developments in genomics and bioinformatics offer a real opportunity for dissecting drought tolerance into component traits and developing tools to manipulate the underlying genes. However, reconstructing effective drought traits will require considerable advances in whole plant physiology modeling as well as in the gene and trait manipulation technologies required for impact at the scale of international germplasm collections and modern plant breeding programs. The GCP proposes to operate at all these R&D levels in the innovation to impact continuum.

The comparative genomics and biology theme of the GCP provides an operational structure for priority setting and focusing research activities within three crop groups: cereals, legumes, and clonal crops. Inevitably, global research progress in most of the cereals is sufficient for the development and application of gene-based marker systems for components of tolerance to drought and other abiotic stresses. Additional targeted investments will be required in millet (the most drought tolerant but least studied of the major cereal crops), though. Conversely, the critical mass of genomics researchers and resources in the legume and clonal crops is much less well developed. For this reason,

Careful prioritization of crop focuses will be applied to ensure rapid and compelling proof-of-concept in key representatives of these crop groups. However, significant direct spillovers from sequence, gene, and trait analyses are expected to significantly impact progress in most other crops in each group. In addition, all crops are likely to be impacted by advances in generic facilitating technologies including: advances in the development of standardized phenotyping protocols, whole plant physiology modeling, molecular breeding simulation studies and decision support tools, as well as procedures for creating low cost trait diagnostics and high throughput array-based genotyping systems. Most of these activities will be carried out through intensive collaboration with scientists in SP1, SP2, and SP4.

The selection of appropriate background genotypes is a critically important process for molecular breeding programs to ensure widespread impact of new genes and traits. Thus, we will ensure that all necessary information is collated, collected, and/or generated to enable the most appropriate varieties and breeding lines to be selected based on agronomic performance in diverse environments plus farmer, processor, and consumer preference and trading potential. In addition, we are documenting appropriate baseline information for those varieties (including production and constraint mapping) that will be used in subsequent impact assessment studies. In this way, the GCP is emerging as a product driven initiative with a strong value chain-based approach to product development and delivery. We strive to move beyond just developing more productive crops to creating new varieties with enhanced stability (reduced vulnerability) and improved value (increased profitability), thus providing a real and sustainable impact on improving the livelihoods of our poorest stakeholders. This will inevitably demand the development of strong alliances with a much broader range of partners including NGOs and the indigenous private sector. Implicit in this is a movement away from linear technology hand-over to a systemic integration between those who need the knowledge and those who supply, validate, and refine it. In turn, this requires our capacity building activities to move beyond just providing expertise and knowledge to also building skills and systems.

Many activities in this subprogram are highly dependent on an effective consortium approach: for example, dealing with the challenges of genotype-by-environment interaction and whole plant physiology modeling by their very nature require coordinated input from many scientists of different disciplines, eco-regions, and types of institution. At the same time, many allied activities in this subprogram can capture substantial economies in time, cost, and efficiency through following a community-based approach. For example, centralized validation and refinement of new technologies for routine application in national breeding programs as well as community support labs offering low cost high throughput genotyping services based on technologies beyond the reach of most national breeding programs. Finally, the creation of effective systemically integrated communities of practice offers excellent opportunities for capturing interdisciplinary synergies and end-user feedback on priorities and outputs. In addition, such communities foster strong technology uptake and product delivery pathways.

The following work plan represents a tentative attempt to articulate the specific details of this operational framework whilst recognizing that funding decisions for the 2005-2007 period have yet to be finalized at the time of writing. There are clearly a number of



recurrent themes that will almost certainly be represented in our short-term portfolio that offer opportunities for dramatic progress in a four primary areas:

- **Validation of gene-based molecular breeding in leading cereal crops:** proof-of-concept pyramiding of a large number of QTL influencing disease resistance, and development and validation of gene-based markers for abiotic stress tolerance traits with a less complex genetic basis than drought tolerance, as well as development of new high throughput genotyping technologies for gene-based markers
- **Application of drought tolerance MAS with linked markers in lesser-studied crops:** validation of SSR-based markers for MAS of components of drought tolerance in a number of representatives of all three crop groups, and development of low cost technologies to enable routine application of marker-assisted selection in diverse national breeding programs
- **Optimization of knowledge-led (gene and component trait) breeding systems:** based on whole plant physiology modeling and simulation analysis of alternative breeding approaches, and development of decision support tools to enable rapid development and application of such systems
- **Establishment of effective product development and delivery pathways:** through the creation and support of a wide range of systemically integrated networks of alliances

In the medium-term, the pipeline of new gene-based tools for manipulating drought tolerance will be flowing out of SP1 and SP2 for application across a wide range of crops in SP3. Finally, in the longer-term we envisage a continuous and iterative process that both adopts new technological advances and seeks maximum synergy from optimum combinations of the old and the new.

### **Output (1.3 in Logframe, Annex 2)**

- Confirm the value of gene-based markers for molecular breeding of complex abiotic stress tolerance traits with particular emphasis on drought tolerance.
- Develop tools for large-scale cost-effective genotyping at centers of excellence serving regional molecular breeding communities.
- Establish simulation and modeling approaches for manipulating whole plant physiology and optimizing breeding systems for abiotic stress enhancement.
- Implement integrated molecular breeding systems for rapid and efficient generation of resilient, productive, and profitable tropical staple crops.
- Generate tools for low tech, low cost routine marker-assisted selection in national breeding programs.

### **Impact**

- A product-driven approach established for seed-based technologies with enhanced resilience, productivity, and profitability.
- Improved breeding systems developed capable of rapidly and efficiently responding to farmer, market, and consumer demands.
- National and regional capacity in modern plant breeding enhanced.
- A pipeline of new varieties developed, deployed, and diffused.

## **Activities**

**Activity 1.3.1:** Elite varieties identified, agronomically evaluated, and market appraised

**Activity 1.3.2:** Gene and trait introgression technologies developed, validated, and refined

**Activity 1.3.3:** Molecular breeding programs designed, implemented, and improved

**Activity 1.3.4:** Communities of practice established, supported, and synergized

**Activity 1.3.5:** Improved seed-products developed, deployed, and impact assessed

## **Indicators (Duration/milestones)**

**Activity 1.3.1:** Elite varieties identified, agronomically evaluated, and market appraised

### **2005**

- Database of elite varieties created for rice, maize, wheat, barley, sorghum, pearl millet, cassava, potato, plantain/banana, common bean, cowpea, and chickpea including all available characterization, evaluation, and pedigree data.
- Detailed agronomic evaluation and market appraisal data generated and/or collated from replicated multilocation trials across major target cropping regions for elite varieties of rice, maize, wheat, barley, sorghum, pearl millet, cassava, potato, plantain/banana, common bean, cowpea, and chickpea.
- Market profiles and opportunities synthesized for selected crops in target regions.

### **2006**

- Consensus developed on varieties to be targeted by the GCP and varieties to be subsequently used by national breeding programs for rice, maize, wheat, barley, sorghum, pearl millet, cassava, potato, plantain/banana, common bean, cowpea, and chickpea.
- Database of elite varieties extended to include finger millet, sweet potato, yam, coconut, soybean, groundnut, lentil, pigeonpea, and forages including all available characterization, evaluation, and pedigree data.
- Production and constraints profile mapping data collected and/or collated for the selected varieties and value chain analysis completed.

### **2007**

- Detailed agronomic evaluation and market appraisal data generated and/or collated from replicated multilocation trials across major target cropping regions. for elite varieties of finger millet, sweet potato, yam, coconut, soybean, groundnut, lentil, pigeonpea, and forages.
- Consensus developed on varieties to be targeted by the GCP and varieties to be subsequently used by national breeding programs for finger millet, sweet potato, yam, coconut, soybean, groundnut, lentil, pigeonpea, and forages.
- Production and constraints profile mapping data collected and/or collated for the selected varieties and value chain analysis completed.

### **Activity 1.3.2: Gene and trait introgression technologies developed, validated, and refined**

#### ***2005***

- Pyramiding of a substantial number of blast disease resistance QTL in a single rice genotype and performance validation in multiple stress environments and locations.
- Validation of linked-markers for components of drought tolerance in maize, sorghum, pearl millet, and common bean.
- Development of low-cost technologies for pyramiding important agronomic traits in priority cereal and legume crops such as rice, maize, soybean, and common bean.
- Evaluate transgene stacking technology using multiple disease resistance genes in rice.
- Preliminary freedom to operate analysis completed for a selection of candidate technologies aimed at use in product development or direct deployment.

#### ***2006***

- Testing of functionally validated gene-based markers for salinity and phosphorus deficiency stresses in rice.
- Validation of candidate gene-based markers for components of drought tolerance in priority cereals (rice, maize, and sorghum) and legumes (common bean, cowpea, and chickpea).
- Development of low-cost technologies for pyramiding important agronomic traits in clonal crops including potato, cassava, sweet potato, and plantain/banana.
- Evaluate transgene stacking technology in legume (groundnut) and clonal crop (cassava) representatives with high transformation efficiencies model crops.
- Freedom to operate analysis completed for all technologies to be used in product development or direct deployment during 2007-2010.

#### ***2007***

- Testing of gene-based marker-assisted selection for abiotic stress tolerance in key representatives of each crop group.
- Validation of candidate gene-based markers for components of drought tolerance in selected complex polyploid crops such as wheat, groundnut, and potato.
- Evaluation of low cost genotype arrays for high throughput MAS in each crop group.
- Field evaluation of products from transgene stacking experiments in cereals, legumes, and clonal crops.

### **Activity 1.3.3: Molecular breeding programs designed, implemented and improved**

#### ***2005***

- Develop consensus QTL analysis model for drought tolerance in maize.
- Generate whole plant physiology models of drought tolerance in cereal crops including wheat and sorghum.
- Collate and interface or create computational molecular breeding decision support tools.

- Intensive multilocal physiological evaluation under various glasshouse water stress regimes of DREB transgenes in all available priority crop backgrounds including rice, maize, and groundnut.

#### **2006**

- Determine optimum marker-assisted selection system for drought tolerance in maize.
- Generate whole plant physiology models of drought tolerance in legumes such as common bean and chickpea.
- Utilize simulation analysis to determine optimum molecular breeding systems for enhancing drought tolerance in cereals such as wheat and sorghum.
- Apply molecular breeding techniques to broaden the genetic base of selected crops such as rice and groundnut.
- Intensive multilocal physiological evaluation under various field water stress regimes of DREB transgenes in all available priority crop backgrounds including rice, maize, and groundnut.

#### **2007**

- Generate whole plant physiology models of drought tolerance in complex polyploid crops such as groundnut, potato, and plantain/banana.
- Utilize simulation analysis to determine optimum molecular breeding systems for enhancing drought tolerance in legume crops such as common bean and chickpea.
- Implement and evaluate optimum molecular breeding systems in cereal crops.

### **Activity 1.3.4: Communities of practice established, supported and synergized**

#### **2005**

- Develop a molecular breeding community of practice in East Africa, with a service, support and resource facility at the NEPAD Biosciences center, Nairobi.
- Establish a molecular breeding training program for Africa in collaboration with the Harvest Plus Challenge Program and NEPAD's Biosciences initiatives.
- Participatory varietal selection networks engaged in priority setting discussions through the establishment of regional consultative foresight processes.
- Foster IP management and innovation policy development capacity in key regional hubs.

#### **2006**

- Develop molecular breeding communities of practice in Latin America, South Asia and Southeast Asia with support units and training programs created in collaboration with key relevant regional initiatives.
- Foster cost effective large-scale shuttle genotyping facilities in key regional hubs with open access to NARS, NGOs, and SMEs breeding programs.
- New molecular breeding technologies and products evaluated and diffused across crops and crop groups.

*2007*

- Establish molecular breeding communities of practice in additional target regions.
- New transgenic technologies and products evaluated and diffused across crops and crop groups.
- Foster biosafety testing policy development and implementation capacity in key regional hubs.

**Activity 1.3.5: Improved seed-products developed, deployed and impact assessed**

*2005*

- Product delivery pathway networks (including NARS, NGOs, SMEs, and legislators) engaged in priority setting discussions through the establishment of regional consultative foresight processes in Africa.
- Establish seeds system development program for key common commodities in collaboration with Harvest Plus Challenge Program and key regional partners.
- Foster capacity for cost-benefit analysis, value chain analysis, and impact assessment in partner institutions.

*2006*

- Product delivery pathway networks (including NARS, NGOs, SMEs, and legislators) engaged in priority setting discussions through the establishment of regional consultative foresight processes in Latin America and Asia.
- Coordination of delivery pathway for proof-of-concept products of marker-assisted selection in key target areas in collaboration with Harvest Plus Challenge Program and key regional partners.

*2007*

- Catalyze local, national, and international market development for proof-of-concept products.
- Coordination of delivery pathway for proof-of-concept products of transgenic technology in key target areas in collaboration with Harvest Plus Challenge Program and key regional players.

**Users and Beneficiaries**

The tools, systems, community resources and germplasm generated by this subprogram will be of direct value and use to germplasm enhancers and plant breeders in NARS, NGOs and SMEs. The application of these outputs will directly benefit farmers, processors, traders, and consumers while the impact from investments in this subprogram will be particularly targeted on improving the likelihoods of resource-poor farming families in harsh production environments.

**Collaborators\***

- NARS – Bangladesh (DU), Brazil (EMBRAPA, UCB, CNPMF, CNPAF), Chile (UC), China (CAAS), Colombia (CORPOICA), Ghana (CRI), India (CPRI, IGAU, IIPR, NDUAT, TNAU, UASB), Indonesia (ICABGRRD), Kenya (KARI), Malawi (CC), Mexico (UAC), Nigeria (NRCRI), Philippines, Senegal (CERAAS), Tanzania

(UCABREN), Thailand (NSFCFC), Uganda (NARO, PRAPACE), Zimbabwe (SIRDC)

- CGIAR centers – CAS-IPGRI, CIAT, CIMMYT, CIP, ICARDA, ICRISAT, IITA, IPGRI, and IRRI
- Regional Networks – NEPAD’s Biosciences East and Central Africa, Rockefeller Foundation East and Southern Africa Marker-assisted Selection Network, Kirkhouse Trust West Africa Legume Molecular Breeding Network, Harvest Plus Challenge Program East and Southern Africa
- Advanced Research Centers (public) – Australia (CSIRO), Belgium (UNU-TECH), France (INRA, IRD, CIRAD, UNIBAP), Japan (NIAS, JIRCAS), Netherlands (WAU), South Africa (ARC), UK (US, JIC, DU), USA (CU, UCF, UVA, KSU, PGEC, UF, UCDC, USDA)

**\*many of which only proposed at the time of writing while many others expected to join in due course, including representatives of networks in Asia and Latin America, NGO, and SME communities globally as well as alliances with private sector advanced research centers.**

Linkages to CGIAR Research Outputs (Percent of Resources)	
Germplasm collection and conservation (including evaluation and characterization)	
Germplasm improvement	70
Sustainable production systems	20
Policy research and development	10
Enhancing NARES capacity	In SP5
Financing Plan (US\$'000)	1,600

#### **Subprogram 4: Genetic Resources, Genomic, and Crop Information Systems and Bioinformatics**

##### **Rationale**

In the first three technical subprograms of the Generation Challenge Program, substantial amounts of information will be generated in numerous labs and fields scattered over the globe. The value of this information will be largely dependent on the way they are managed, analysed, and made accessible. The way they can be analysed will, in turn, be dependent on the way analysis tools and other information sources will be available. The challenge of linking and integrating these information components and analysis tools into a coherent information gateway will therefore play a central role within the GCP and forms the basis of Subprogram 4.

By forging a strategic alliance between bioinformatic, biometric, and advanced data management techniques, an integrated genetic resources, genomics, and crop improvement information network will be created that will be the basis of genetic resource management and use for the coming century. The GCP elements of this network will be fully integrated in the global network of resources in the field of genomic science

and genetic resources management. The GCP will contribute substantially and shape this global network, but it will be the network of the global community of which the GCP is glad to be part of. This implies that SP4 will participate in on-going initiatives of the global communities, will use, support, and where necessary participate in developing global standards and common technologies for data exchange.

Since data have been generated in the GCP from the very beginning, it will be necessary to develop and implement short-term solutions for all data that need to be shared, awaiting the 'global network' solutions. These short-term solutions will use simple approaches that need to allow storage of the data in such a way that they can be fully interpreted and converted to any other format, providing full access to any authorized user.

In addition, critical questions concerning the methodologies for linking gene discovery with genetic resource characterization and crop evaluation data will be addressed by improving existing and developing new analytical tools. The first three subprograms will define the specific requirements of these tools.

In all development of new, or adaptation of existing, technology a short cycle approach will be followed. This implies an iterative approach in which simple applications build on very few use cases that are improved based on feedback from the user. Using this experience more use cases are added, etc. Scalability of the applied technology and design is therefore of utmost importance. This approach guarantees not only frequent interaction with users, but also tangible output during the course of project execution.

Finally, SP4 will obviously not be involved in all ICT issues relevant to the GCP; the consortium members will keep responsibility for their own data curation as far as possible, based on the principle that data can best be curated as close to the source as possible. They can choose, and in most cases have already chosen, their own hardware and software solutions for the local capture, validation, analysis, and storage of information. However, as far as possible, SP4 will propose software and data structures to use for that purpose. Given its responsibility for access to the data, SP4 will determine and impose the way the data and tools should be made available to the consortium and the world, and will support the institutions to implement the necessary technology.

#### **Output (1.4 in Logframe, Annex 2)**

A network of genetic resource, genomic, and crop improvement information platforms supporting GCP research created and integrated with complementary resources in the global community and managed by a network of bioinformatics specialists in GCP members and their partners.

#### **Impact**

- The quality of the databases maintained by the GCP consortium members will improve as a result of optimized curation of, and access to, GCP data.
- Data generated by the GCP and other relevant public data will be cross-linked and made available to anyone who is authorized to have access to it.

- The interpretation of data will be improved by the availability of better analysis tools, allowing more efficient approaches to scientific discovery and crop improvement in the GCP.
- GCP-enhanced data sources, analysis tools, and interoperability technology will contribute considerably to capacity building and collaboration within the world scientific, crop improvement, and germplasm conservation communities. These contributions, plus the transparent approach will make the GCP, and thus its consortium members, full collaborators in the global efforts to create global public goods in the respective fields.

### **Activities**

**Activity 1.4.1:** Timely solutions for storage of, and access to, data generated in the GCP

**Activity 1.4.2:** Improvement of quality of existing databases and analysis tools

**Activity 1.4.3:** Information network of GCP data sources and analysis tools developed and integrated over internet with international bioinformatics resources

**Activity 1.4.4:** New data processing and analysis tools developed that serve the needs of Subprograms 1, 2, and 3

### **Indicators (Duration/milestones)**

**Activity 1.2.1:** Timely solutions for storage of, and access to, data generated in the GCP

#### **2005**

- Development of a first release “model driven architecture” of common data models for all data types being generated in the GCP. Initial refinement of these models will primarily focus on germplasm, passport data, molecular characterization, and phenotype data.
- Alignment of data models to information system database schemata, software (“middleware”) data structures, and internet data exchange protocols adapted for use in the GCP. Adaptation/extension of existing database schemata, software design, and internet protocols to accommodate new data models.
- GCP data loaded into information systems adapted to use the new data models.
- Data in data models published on GCP partner and/or central websites.
- Inventory of available data sources and tools available on the GCP website.

#### **2006**

- Based on project experience, refinement of previously formulated germplasm, passport data, marker, and phenotype data models of the model driven architecture. Expansion of the range of data models to mapping and general genomics data being generated in the GCP.
- Data that are not previously accessible in public databases or GCP web-services loaded into information systems instantiating the revised data models.
- Data in revised data models published on GCP public websites.
- Inventory of available GCP data sources and tools maintained and updated.



## 2007

- Expansion of the range of data models to be used for storage of data being generated in the GCP to novel data types such as proteomics and metabolomics, association genetics, gene transfer data, and crop modeling information.
- Data that are not previously accessible in public databases or GCP web-services loaded in information systems instantiating the revised data models.
- Data in revised data models available on GCP public websites.
- Inventory of available GCP data sources and maintained and updated.

### **Activity 1.4.2: Improvement of quality of existing databases and analysis tools**

## 2005

- Compilation of publicly available crop molecular variation (e.g., SNP) and gene expression data sets into a GCP integrated resource.
- Further development of reference GCP information platform based on open crop information system projects such as ICIS and Germinate; inclusion of modules for storage and analysis of molecular marker data, improvement of the ‘genebank documentation’ component.
- Deployment of initial set of GCP information platform components at several consortium member sites (by individual partner site initiative).
- Studies aimed at assessing and improving data quality and access in existing databases at GCP partner sites, including quality assurance validation of data sets, increasing the interpretability of datasets by adding meta-data or summarizing.
- Workshop on data quality and data curation for all GCP consortium members.
- Comparison of analysis tools used in the consortium, published on the Internet for user assessment and feedback.
- Full deployment of first generation Subprogram 2 genomic information data sets – gene catalog, universal marker set, and gene expression data – published to GCP web portals.

## 2006

- Further development of GCP platform towards functional genomics data cross-linked to molecular variation data and germplasm; adoption of emerging ontologies.
- Continued deployment of initial set of GCP platform components at additional consortium member sites (by individual partner site initiative).
- Development and application of data quality assessment tools.
- Improvement of quality of database management at GCP partners by upgrading the existing systems toward agreed ontologies.

## 2007

- Continued deployment of initial set of GCP platform components at additional consortium member sites (by individual partner site initiative).
- Improvement of quality of database management at GCP partners by adoption of elements of the GCP platform, or by upgrading the existing systems toward agreed ontologies.

**Activity 1.4.3: Information network of GCP data sources and analysis tools developed and integrated over internet with international bioinformatics resources**

**2005**

- Participation in standards (ontology) consortia such as the Plant Ontology Consortium, with initial ontologies incorporated into relevant applications.
- Development, in collaboration with relevant international partners, of extensible Mark-up Language document type definitions (XML-DTD) for most important data types specified by common GCP data models. Initially these will be schemas for germplasm, passport data, marker data, molecular characterization, and phenotype data.
- Publication of the developed schemas on the appropriate GCP websites.
- Training of staff in the GCP Consortium in the development and application of web-services technology.
- Application of web-services technology for wrapping at the GCP consortium members; resulting in IRRI, IPGRI, and CIMMYT having web-services providers installed on top of existing institutional databases.
- Policy plan for the IP management of SP4, covering issues concerning all data and software that is produced and/or used within the GCP.

**2006**

- Continued involvement in standards consortia.
- Continued development of XML schemas for all data models. Schemas for more complicated data types will be published.
- Continued application of internet technology for wrapping information systems of GCP consortium members; all CGIAR centers involved in the GCP have some web-services.

**2007**

- Continued involvement in standardizing consortia.
- Continued development of XML schemas for all primary GCP data. Schemas for the remaining data types will be published.
- Continued application of technology for wrapping systems at the GCP Consortium members to provide an integrated network of systems for all CP data. The number of web-services provided by GCP consortium members will increase.

**Activity 1.4.4: New data processing and analysis tools developed that serve the needs of Subprograms 1, 2, and 3**

**2005**

- Full integration achieved of the High Performance Computing (HPC)-facilities commissioned in 2004 in the GCP toolbox by deployment of genomics database mirrors and by making documentation, user guides, and applications available, with training of GCP scientists.
- (supporting SP1) Interfaces commissioned to diverse datasets (initially passport data, phenotype data and molecular markers), and decision support systems for the

sampling of germplasm based on diverse datasets accommodating concepts such as core collection and elite collection; development of prototype LIMS for gene-locus focused (sequence based) allele mining data (initially, for rice; later, for other CG crops).

- (supporting SP2) Bioinformatics facilities – databases and tools – fully commissioned to identify and manage information about orthologous gene and marker loci across multiple crops; to support comparative gene expression profiling; and to enable Expressed Sequence Tag (EST) annotation. Development of multi-trait QTL mapping software, allowing inclusion of the effects of environmental interactions and pleiotropy; initial curation of comparative genomic maps developed using universal orthologous markers, applied across phase I GCP crops.
- (supporting SP3) Initial prototyping of marker assisted selection (MAS) and marker assisted backcrossing (MAB) decision support tools, based on integration of existing software elements and packages into one package; enhanced phenotype data capture, analysis, and integration in the GCP platform, with associated training of plant breeder community.
- Collaborate with SP5 to develop training resources to assist effective deployment of new GCP databases and tools.
- Standing inventory of emerging needs from SP1, SP2, and SP3 for realization in 2006.

## 2006

- (supporting SP1) Adaptation and extension of public tools for linkage disequilibrium and association genetics analysis of GCP germplasm core set; enhancement of tools for biodiversity analysis of germplasm (e.g., integration of the CGIAR ICT/KM “DIVA” software to the GCP platform, for the geographical biodiversity analysis of the GCP core germplasm set); development of gene-locus focused (sequence based) allele mining data analysis and integration tools integrated with associated LIMS (e.g. for high throughput SNP detection).
- (supporting SP2) Extension of GCP genomic information platform to integration information from emerging crop model organisms – e.g. maize and Medicago genome projects – and to additional comparative functional genomics data sets (e.g., mutant stocks in various species; EST libraries from orphan crops; proteomics and metabolomics data sets); prototyping of “systems biology” databases and tools (i.e. extension of comparative gene catalog to comparative genetic pathways information).
- (supporting SP3) Specification and development of novel (public, “open source”) molecular breeding and (predictive) crop modeling tools integrated with GCP information platform.
- Additional products identified in 2005 as needed by in the research programs of SP1, SP2, and SP3.
- Collaborate with SP5 to develop training resources to assist effective deployment of new GCP databases and tools.
- Standing inventory of emerging needs from SP1, SP2, and SP3 for realization in 2007.

## 2007

- (supporting SP1) Deployment of a fully integrated GCP platform and network combining germplasm, environment, and phenotype data with genotype (molecular variation) information.
- (supporting SP2) Extension of GCP platform to newly available data sets, with full database “vertical” integration of gene locus molecular variation with functional (i.e. protein residue) determinants at molecular (biochemical/physiological) level; prototype integration of genomics information with crop physiological models for some model crop systems.
- (supporting SP3) Full integration of molecular breeding and predictive crop models with molecular variation and gene function databases, for decision support of GCP plant breeding.
- Additional products identified in 2006 as needed by in the research programs of SP1, SP2, and SP3.
- Collaborate with SP5 to develop training resources to assist effective deployment of new GCP databases and tools.
- Standing inventory of emerging needs from SP1, SP2, and SP3 for realization in 2008.

### Users/Beneficiaries

- The research of GCP consortium members will be facilitated by providing short term and long term solutions for storing, processing, analyzing, and publishing (online) the data that are generated in the project.
- The global crop research community in general will benefit from the newly developed global public crop bioinformatics goods such as tools, datasets, and contributions to methodology and standard development.

### Collaborators

Since the basis of the approach of SP4 is integration in the global genetic resources, genomic, and crop improvement information network, the number of collaborators is very large; all involved in the relevant consortia and networks are (potential) collaborators and will be invited to meetings, asked for opinions on initiatives, etc.

Major international bioinformatics collaborators are: GRAMENE, BioMOBY, S-MOBY, TIGR, NCGR, GERMINATE, Maize GDB, and Genoplante, among others.

Linkages to CGIAR Research Outputs (Percent of Resources)	
Germplasm collection and conservation (including evaluation and characterization)	50
Germplasm improvement	30
Sustainable production systems	
Policy research and development	20
Enhancing NARES capacity	In SP5
Financing Plan (US\$'000)	3,900

**Objective 2: Strengthen the capacity of NARS and GCP scientists to apply the tools of genomics, molecular biology, and bioinformatics to the analysis of genetic diversity held in germplasm collections, and to use this knowledge to improve crop breeding programs and to develop new stress-tolerant varieties.**

### **Subprogram 5: Capacity Building (Output 2.1 in Logframe Annex 2)**

#### **Rationale**

The Generation Challenge Programme (GCP) is designed to bridge gaps that prevent technological advances from being applied in agricultural research to target resource poor farmers. One important gap is between the advanced research laboratories in developed countries and international research programs located in the “South.” A second major gap is between the knowledge required to apply the fruits of this new knowledge and the existing capacity of scientists in developing countries. Bridging both these gaps will require a sustained effort in capacity building by participants of the GCP.

Capacity building is the focus of Subprogram 5 and is a major cross-sectional theme of the GCP. It has two dimensions: one is to better enable GCP members to carry out the cutting edge research agenda, and the second is to empower national program scientists to participate in GCP activities. This puts in place mechanisms by which GCP products can reach crop improvement programs and farmers.

In order to access the diversity held in NARS and CGIAR germplasm collections (wild species, landraces, and advanced cultivated genotypes), programs must be able to apply the most modern genomics and molecular biology tools and technologies. In certain cases, conventional crop improvement might not be viable for the crop and traits of interest. NARS scientists may then wish to consider genetic transformation as the avenue of choice for development of varieties targeted to their local needs. However, the knowledge base and technical capacity to use these tools must be acquired and kept up to date by the NARS. Capacity is also needed to build appropriate databases that contain traditional and molecular data on germplasm and to browse multi-disciplinary data repositories and to access and make use of data about traits, genes, and sequences. The GCP has a major role to play in establishing and maintaining this expertise.

In addition to the focus on NARS, it is understood that different consortium members have themselves different capacities, and that the rate of technological innovation is so rapid that capacity upgrading will be required for GCP members as well. Thus, Subprogram 5 will also take into consideration these needs of its member institutions.

Capacity building activities will be undertaken following two broad modalities. First, most activities supported by the GCP will have an embedded capacity building component. These activities will involve activities such as graduate and postdoctoral research, sabbatical leaves, internships, and other types of exchanges of scientists among participating institutions in the GCP. The second modality will be the development and execution of specific courses, workshops, on-line distance learning, and more. In all, the needs for capacity enhancement will be formulated after a thoughtful assessment of the

requirements of NARS, as well as GCP consortium members, and will form the basis for an organized work plan.

### **Impact**

- GCP consortium member scientists gain knowledge in the different scientific aspects of the thematic subprograms.
- A significant number of scientists in developing countries engaged in GCP research activities in the different subprograms.
- Capacity built in developing countries through the participation of a significant number of scientists in training events (online or face-to-face courses).
- A significant number of training materials identified and made available or newly developed to support sustainable learning and research in less developed countries.
- Scientists in developing countries able to benefit from the materials and technologies for application to research and applied plant breeding.
- Sustainable plant research and development in place in developing countries as a result of increased capacity of their national scientists as well as of appropriate partnerships established through the GCP activities .

### **Activities and Outputs**

*Output 2.1 NARS scientists with capacity to for full participation in CP*

**Activity 2.1.1:** Needs assessment conducted

**Activity 2.1.2:** NARS capacity building components built into commissioned research programs

**Activity 2.1.3:** NARS capacity building built into competitive grants program

**Activity 2.1.4:** Training courses or workshops on specific technologies and areas conducted

*Output 2.2 NARS Research Centers serves as Regional Research Hubs*

**Activity 2.2.1:** Specific needs of NARS RRHs addressed

*Output 2.3 Crop improvement networks in place to incorporate novel alleles/traits*

**Activity 2.3.1:** Meta analysis of existing and recent crop improvement networks with McKnight CCR and Harvest Plus

**Activity 2.3.2:** Development of regional improvement networks for multiple crops (in close cooperation with HarvestPlus, McKnight Foundation and NGOs)

**Activity 2.3.3:** Ex ante and longitudinal assessments of activities to maintain to relevance and focus on needs of resource poor farmers (with IFPRI)

### **Indicators (Duration/milestones)**

**Activity 2.1.1:** Needs assessment conducted.

**2005**

- A survey conducted to assess and analyze gaps in knowledge and skills required to meet the goals of the GCP among the consortium members for matters covering the four thematic subprograms.

- A survey circulated among all NARS involved in the competitive proposals as well as those that attended the Needs Assessment workshop in 2004 to check capacity building advancement.
- A monitoring mechanism devised to request feedback from GCP capacity building activities beneficiaries as well as from users of the compiled training materials available through the GCP website.
- A Community of Practice of capacity building team members from the GCP Consortium as well as from the other CPs established.

### **2006**

- A questionnaire prepared and circulated among additional potential partners/beneficiaries, 20 per region, to guarantee the increased impact of the GCP-generated knowledge and products. Replies to the questionnaire used to update training materials and appropriately refocus mechanisms and criteria on which capacity building activities were originally based.

### **2007**

- A questionnaire prepared and circulated among additional potential partners/beneficiaries, 20 per region, to guarantee the increased impact of the GCP generated knowledge and products. Replies to the questionnaire used to update training materials and appropriately refocus mechanisms and criteria on which capacity building activities were originally based.

### **Activity 2.1.2: NARS capacity building components built into commissioned research programs.**

#### **2005**

- At least eleven NARS scientists are involved in the development of reference molecular marker kits to analyze diversity of germplasm for the year 1 GCP crops at the most appropriate consortium member institution (link with SP1).
- A set of training materials for a course in genetic diversity analysis of germplasm is gathered and/or developed (link with SP1).
- A training manual for the proper use of sensors and simple models is developed for scientists involved in phenotyping of complex traits (link with SP1).
- A training program is designed for genetic diversity analysis of germplasm and for a course on complex traits phenotyping, including a training manual for the proper use of sensors and simple models for use by scientists involved in phenotyping of complex traits (link with SP1).
- A set of training materials for a course in genomics and comparative genomics is gathered and/or developed and a training curriculum is designed (link with SP1).
- A compilation is made of training materials for marker assisted-selection and breeding (link with SP3).
- Training manuals and programs are designed for scientists engaged in SP4 tool and technology development and administration and others for SP4 users, i.e. scientists engaged in SP1 and SP2 research projects.

- A repository for GCP training materials is designed and made available to GCP consortium members and partners outside the consortium (link with SP4).

#### **2006**

- At least eleven NARS scientists develop reference molecular marker kits to analyze diversity of germplasm for relevant crops in an appropriate consortium member institution (link with SP1).
- Training materials newly developed in 2005 for the different subprograms translated to Spanish and French.
- Training courses conducted in at least one Francophone African country and one Spanish speaking country in LAC.

#### **2007**

- At least 10 NARS scientists develop reference molecular marker kits to analyze diversity of germplasm for relevant crops in an appropriate consortium member institution (link with SP1).

### **Activity 2.1.3: NARS capacity building built into competitive grants program.**

#### **2005**

- At least fifteen capacity building activities included all together in the competitive projects awarded.
- At least fifteen NARS institutions and scientists engaged with GCP research activities through the approved competitive projects.
- A fellowship and travel grant scheme set up and implemented.
- At least two fellowships per subprogram are awarded to developing country nationals from research centers involved in the GCP.
- At least two travel grants per subprogram are awarded to developing country nationals from research centers involved in the GCP to attend a relevant conference or workshop.

#### **2006**

- At least fifteen capacity building activities included all together in the competitive projects awarded.
- At least fifteen NARS institutions and scientists engaged with GCP research activities through the approved competitive projects.
- At least two fellowships per subprogram are awarded to developing country nationals from research centers involved in the GCP.
- At least two travel grants per subprogram are awarded to developing country nationals from research centers involved in the GCP to attend a relevant conference or workshop.

#### **2007**

- At least fifteen capacity building activities included all together in the competitive projects awarded.



- At least fifteen NARS institutions and scientists engaged with GCP research activities through the approved competitive projects.
- At least two fellowships per subprogram are awarded to developing country nationals from research centers involved in the GCP.
- At least two travel grants per subprogram are awarded to developing country nationals from research centers involved in the GCP to attend a relevant conference or workshop.

**Activity 2.1.4: Training courses or workshops on specific technologies and areas conducted.**

**2005**

- A workshop conducted based on results of the year 1 work plan (genotyping) including significant participation of germplasm managers/curators from the CG (12) and NARS (24), representing at least 24 developing countries (link with SP1).
- A seminar is conducted to provide a basic level of understanding on policy, intellectual property issues, and access and benefit sharing. Fourteen participants from the consortium and at least a similar number from developing country GCP partners are anticipated.
- Three courses on marker-assisted selection and breeding are held in Asia/Pacific, Africa and Latin America respectively with the participation of 12 regional scientists, representing a minimum of 6 countries, per course (link with SP3).
- A platform technology (e.g., web-services) tutorial course held for GCP developers/administrators involved in SP4.
- A cluster/grid computing facilities and year 1 GCP platform use training course held for GCP scientists (link with SP4).
- A workshop is held on project proposal design and development, in collaboration with the other CPs, in Asia/Pacific, Africa, and Latin America. Altogether a minimum of 14 consortium member scientists and at least 12 developing country national scientists per region involved.

**2006**

- Three courses are organized for the use of molecular markers in genetic diversity assessment and germplasm management in Asia/Pacific, Africa, and Latin America respectively with the participation of 12 regional scientists, representing a minimum of 6 countries, per course (link with SP1).
- At least 10 developed and 10 developing country scientists attend an expert workshop to examine the minimum set of environmental data to characterize environments in all GCP experiments (link with SP1).
- A training course on phenotyping of complex traits is organized. Attendants include GCP scientists (14) and at least 14 scientists from developing country institutions involved in the GCP research.
- A seminar is conducted to provide a basic level of understanding on policy, intellectual property issues, and access and benefit sharing. Fourteen participants from the GCP and at least a similar number from developing country GCP partners are anticipated (link with SP1).

- Three courses on genomics and comparative genomics are held in Asia, Africa, and Latin America respectively with the participation of 12 regional scientists, representing a minimum of 6 countries, per course (link with SP2).
- Three courses on marker assisted selection and breeding are held in Asia/Pacific, Africa, and Latin America respectively with the participation of 12 regional scientists, representing a minimum of 6 countries, per course (link with SP3).
- A workshop is held on project proposal design and development, in collaboration with other CPs, in Asia/Pacific, Africa, and Latin America. Altogether a minimum of 14 consortium member scientists and at least 12 developing country national scientists per region involved.

### **2007**

- Three courses are organized for the use of molecular markers in genetic diversity assessment and germplasm management in Asia/Pacific, Africa, and Latin America respectively with the participation of 12 regional scientists, representing a minimum of 6 countries, per course (link with SP1).
- A training course on phenotyping of complex traits is organized. Attendants include GCP scientists (14) and at least 14 scientists from developing country institutions involved in GCP research (link with SP1).
- A seminar is conducted to provide a basic level of understanding regarding policy, intellectual property issues, and access and benefit sharing. Fourteen participants from the consortium and at least a similar number from developing country GCP partners are anticipated (link with SP1).
- Three courses on genomics and comparative genomics are held in Asia, Africa, and Latin America respectively with the participation of 12 regional scientists, representing a minimum of 6 countries, per course (link with SP2).
- Three courses on marker-assisted selection and breeding are held in Asia/Pacific, Africa, and Latin America respectively with the participation of 12 regional scientists, representing a minimum of 6 countries, per course (link with SP3).
- A workshop is held on project proposal design and development, in collaboration with other CPs, in Asia/Pacific, Africa, and Latin America. Altogether a minimum of 14 consortium member scientists and at least 12 developing country national scientists per region involved.

### **Activity 2.2.1: Specific needs of NARS RRHs addressed.**

### **2005**

- Regional research hubs in the South are identified through a systematic approach for the main subjects concerning each of the subprograms. At least two RRH are identified per region.
- Equipment, facilities, and capacity building needs of RRH identified.
- A consultation is made among the identified RRH to guarantee the appropriateness of the work plan as defined in 2004. The SP5 work plan is refined to reflect the outcome of the consultation.

- A telecommunications- and web-supported help desk is set up at a consortium member as a mechanism to ensure adequate and sustainable support to those capacity recipients in all its forms.

#### **2006**

- A consultation is made among RRH partners to refine work plan and adapt activities and mechanisms.
- RRHs facilities upgraded as needed to allow full participation in GCP and, possibly, regional crop improvement networks (see Activity 2.3.2).
- A pilot help desk is set up at a selected RRH per region to test the viability of devolving this function to research centers in the South.

#### **2007**

- A consultation is made among RRH partners to refine work plan and adapt activities and mechanisms.

### **Activity 2.3.1: Meta analysis of existing and recent crop improvement networks with McKnight CCR and Harvest Plus**

#### **2005**

- An evaluation is made of existing and recent networks including plant genetic resources, genomic, crop and bioinformatics networks
- Stakeholders and participants in networks interviewed to assess validity of initial conclusions
- Strengths and weaknesses of these networks collated and used as contribution to design regional, possibly multicrop, crop improvement networks

### **Activity 2.3.2: Development of regional improvement networks for multiple crops (in close collaboration with Harvest Plus, McKnight Foundation, and NGOs)**

#### **2005**

- An evaluation is made of existing and recent networks including plant genetic resources, genomic, crop, and bioinformatics networks.
- Strengths and weaknesses of these networks collated and used as contribution to design regional, possibly multicrop, crop improvement networks.
- Proposals prepared and submitted (with Harvest Plus) to granting agencies for support of regional networks in South and Southeast Asia, African and Latin America; RRHs from Activity 2.2.1 included as lead institutions.

#### **2006**

- Implementation workshops conducted with breeders and other stakeholders for those regional networks that secured funding.
- Minimum environmental data standards adopted.
- Facilities for network members to collect and report minimum data sets installed.

- First regional germplasm evaluation nurseries conducted.
- Mechanisms for monitoring progress articulated and implemented.

#### *2007*

- Data from first nurseries collected and analyzed.
- Second series of nurseries established.
- NARS regional network scientists participate in GCP annual research meeting.

#### **Activity 2.3.3: Ex ante and longitudinal assessments of activities to maintain to relevance and focus on needs of resource poor farmers (with IFPRI).**

#### *2005*

- Characterizing and modeling impacts of improved drought tolerance.
- A set of documented drought resistance strategies for a set of globally important food crops developed.
- A set of crop simulation models calibrated to simulate the specific plant-level drought resistant strategies.
- A set of look-up tables that defines the likely drought resistance “yield dividends” implied by different drought tolerance strategies, over a feasible range of drought tolerances, over a feasible range of soil conditions, over a feasible range of climate patterns.
- Report on the methods, results, and implications of this analysis.

#### *2006*

- Spatial patterns and welfare impacts of improved drought tolerance.
- Spatially-explicit global databases of environmental and climatic parameters, spatial distribution of productivity of target crops, and the spatial distribution of rural and urban populations, focusing in particular on identifying areas of poverty and malnourishment.
- Spatially-explicit global databases and maps of yields and yield changes under different drought tolerance strategies for at least four globally important food crops (e.g. wheat, maize, soybean, rice).
- Potential changes in gross revenue, calories per capita, and national food self-sufficiency.
- Prototype estimates of potential impacts on yield/production changes at the (sub-) national level on food self-sufficiency during normal and drought years.

#### *2007*

- Modeling of future impacts of drought tolerance on food markets and water resources.
- Select established and agricultural systems models integrated into IFPRI IMPACT-WATER model
- Increased spatial and commodity resolution of IMPACT-WATER.

- New models tested, calibrated, then utilized to assess alternative scenarios for drought tolerance in different crop varieties.

**Users/Beneficiaries**

- Scientists at GCP member institutes and in developing country national programmes, research institutions, and universities
- Research institutions, mainly in developing countries
- Regional and international networks of scientists
- Resource-poor farmers and consumers

**Collaborators**

GCP consortium members, ARIs, NARS attending the Needs Assessment workshop in year 1 (Aug. 04), NARS already engaged or to be engaged in commissioned and competitive proposals and pre-proposals, educational/instructional institutions in developed and developing countries, private seed sector, and networks for genetic resources, genomic, crop improvement, and bioinformatics. Also, Harvest Plus and McKnight Foundation.

Linkages to CGIAR Research Outputs (Percent of Resources)	
Germplasm collection and conservation (including evaluation and characterization)	
Germplasm improvement	
Sustainable production systems	
Policy research and development	
Enhancing NARES capacity	100
Financing Plan (US\$'000)	2,300

## ANNEX 2. GENERATION CHALLENGE PROGRAM LOGICAL FRAMEWORK

Narrative Summary	Verifiable indicators	Means of verification	Assumptions
<p><b>GCP Goal:</b> Improve livelihoods of resource-poor farmers and consumers in developing countries by accessing and utilizing the genetic potential of public genetic resources in plant breeding programs through the concerted generation, management, dissemination, and application of comparative biological knowledge</p>	<p>Per capita income and production increases in targeted countries/regions and crops; nutritional status of low income population segments; NARS and CGIAR center annual reports; GCP-identified novel alleles and traits in breeding programs and in popular crop varieties under production</p>	<p>Reports from international agencies (e.g. FAO, UNDP, WHO); breeding program records of partner NARS, national production statistics; molecular genotypes of parental, segregating, elite breeding materials, and adopted varieties</p>	<p>Assumptions relate to lower level results and activities</p>
<p><b>Objective 1:</b> Develop a platform for, and conduct, analysis of genetic diversity in international crop genetic resources and apply this to improve major crops for drought tolerance and other related traits of importance to resource-poor farmers</p>	<p>Tools and techniques for high throughput analyses, descriptions of diversity and structure of populations (collections) of crop species, management of data, association of drought tolerance and related traits with this diversity, and tools available to breeding programs.</p>	<p>Publications in literature and annual reports, presentations at scientific meetings, publication on web sites</p>	<p>Free access to and sharing of germplasm</p>
<p><b>Output (Sub-Program 1) 1.1:</b> Organize a rational access to sources of genes and alleles involved in key agricultural traits on the basis of coordinated surveys of molecular and phenotypic variation among the accessible germplasm collections for CG mandate crops.</p>	<p>Tools and techniques for high throughput analyses, descriptions of diversity and structure of populations (collections) of crop species</p>	<p>Publications in literature and annual reports, presentations at scientific meetings, publication on web sites</p>	<p>Free access to and sharing of germplasm</p>
<p><b>Activity 1.1.1</b> Creating an improved understanding of the structure of the diversity for the major world food crops, diagnostic molecular markers for subsequent germplasm analysis identified, and a set of reference samples designed for integrated characterization</p>	<p>See Annex 1 for specific milestones, indicators and time frames</p>	<p>Genetic structures of collections published in refereed literature, annual GCP reports, presented at international meetings; markers and lists of reference germplasm available on GCP web sites</p>	<p>GCP scientists have access to the full range of germplasm collections and can freely exchange material (at least DNA)</p>
<p><b>Activity 1.1.2</b> Develop a range of techniques accessible in key laboratories applied for high throughput molecular characterization of germplasm</p>	<p>See Annex 1 for specific milestones, indicators and time frames</p>	<p>Techniques published on GCP Websites and in literature</p>	<p>The laboratories have access to germplasm</p>

<b>Activity 1.1.3:</b> Establish and implement a scientific and organizational framework to describe tolerance to drought	See Annex 1 for specific milestones, indicators and time frames	Published protocols and practice; clear instructions for experimental practices accompany every GCP funded project	Agreement on protocols achieved within crop species
<b>Activity 1.1.4:</b> Associate molecular polymorphisms with higher tolerance to drought and integrate methodological improvements into these studies	See Annex 1 for specific milestones, indicators and time frames	Associations published in literature and Web sites	Association genetics that is well worked out in maize and humans can be applied routinely to a wide range of crops with different reproductive strategies
<b>Activity 1.1.5:</b> Establish protocols and policy analyses to allow germplasm exchange with proper access and benefit sharing from the derivatives of the program	See Annex 1 for specific milestones, indicators and time frames	Protocols and analyses published in literature and available on GCP Web sites	No perceived obstacle to the <i>analyses</i> ; though germplasm access and exchange may be problematic in some cases
<b>Output (Sub-Program 2) 1.2:</b> Genes conferring stress tolerance, with emphasis in drought, in multiple crop species identified through the use of cross-cutting tools developed for comparative genomics and biology		GCP annual reports and progress summaries on GCP Web site; publication of results in refereed journals, presented at conferences.	
<b>Activity 1.2.1:</b> Assembly of genomics and germplasm resources through consolidating existing (and developing new) specialized genetic stocks and framework genetic markers for target crops	See Annex 1 for specific milestones, indicators and time frames	Publication as above; availability of stocks and markers via GCP Web site and Sub-Program Leaders	Phenotyping protocols and facilities developed by SP 1 in place; efficient transformation protocols developed or available for target species
<b>Activity 1.2.2:</b> Develop comparative maps within and across species and deploy comparative mapping tools to CP partners, linked to the CP consensus map repository and major international plant databases	See Annex 1 for specific milestones, indicators and time frames	Maps available for consultation on GCP Website and published in refereed journals; record of Web-based mapping tools accessed; hyperlinks to international databases on GCP Web sites/services	Conserved ortholog sequence approach is feasible for monocots
<b>Activity 1.2.3:</b> Assign genes and pathways to putative phenotypes through the convergence evidence of genome variation, expression patterns and phenotypic data	See Annex 1 for specific milestones, indicators and time frames	GCP annual reports and progress summaries on GCP Web site; publication of results in refereed journals, presented at conferences.	Partners with adequate array spotting and analysis facilities and biochemistry/physiology expertise join GCP activities;

<b>Activity 1.2.4:</b> Validate genes and pathways through evaluation of under or over-expression constructs or variants (induced or natural) of the target genes	See Annex 1 for specific milestones, indicators and time frames	GCP annual reports and progress summaries on GCP Web site; publication of results in refereed journals, presented at conferences.	Efficient and "clean" gene stacking protocols available to GCP; permits to undertake field trials of transgenics are obtained and all biosafety requirements can be met
<b>Activity 1.2.5:</b> Disseminate genomic resources and information platform to NARES through networks and training (Links to Objective 2)	See Annex 1 for specific milestones, indicators and time frames	Agendas, programs, summaries and proceedings of network workshops, training and other events	Additional resources are obtained to support networks
<b>Output (Subprogram 3) 1.3:</b> Validation and application on elite varieties of gene-based markers for molecular-assisted breeding for tolerance to complex stresses (with emphasis on drought) in national and regional centers of excellence		Targeted NARS adoption of MAS for selection of complex traits evidenced by their annual reports	
<b>Activity 1.3.1:</b> Identification, agronomic evaluation and market appraisal of elite varieties suitable for conversion using marker assisted selection	See Annex 1 for specific milestones, indicators and time frames	GCP annual reports; data bases on GCP Web site	The elite varieties will remain preferred by farmers during life of GCP and beyond
<b>Activity 1.3.2:</b> Development, validation and refinement of gene and trait introgression technologies	See Annex 1 for specific milestones, indicators and time frames	GCP annual reports and published comparisons of efficiencies and cost-effectiveness; external or commissioned reviews	NARS and GCP consortium members and partners have adequate facilities to apply MAS
<b>Activity 1.3.3:</b> Design, implementation and improvement of molecular breeding programs in NARS and GCP Consortium members	See Annex 1 for specific milestones, indicators and time frames	NARS annual reports and other NARS publications; external reviewers; predictions of models reflects field experience	NARS and GCP consortium members and partners have adequate facilities to apply MAS; adequate grid computational power to run complex models
<b>Activity 1.3.4:</b> Establishment, support and promotion of molecular breeding communities of practice	See Annex 1 for specific milestones, indicators and time frames	Summaries of periodic meetings or other interactions among members of the community of practice	Institutional commitments of NARS and other GCP partners to sustain participation in communities
<b>Activity 1.3.5:</b> Development, deployment and impact assessment of improved seed-products	See Annex 1 for specific milestones, indicators and time frames	External or commissioned reviews and analyses of impact of varietal adoption	Harvest Plus GCP partnership is successful in developing and deploying an innovative end-user model



<b>Output (Sub-Program) 1.4:</b> A network of genetic resource, genomic and crop improvement information platforms supporting GCP research created and integrated with complementary resources in the global community and managed by a network of bioinformatics specialists in GCP members and their partners		GCP Web sites and Web Services offerings, proceedings of international bioinformatics meetings, membership rosters of international standards organizations, GCP annual reports	GCP members and their partners have or acquire adequate hardware, software, personnel, and band width to access GCP resources
<b>Activity 1.4.1</b> Timely solutions for storage of, and access to, data currently generated in the GCP	See Annex 1 for specific milestones, indicators and time frames	GCP Web Site offerings	Data generated and uploaded in a reasonable time frame
<b>Activity 1.4.2:</b> Improvement of quality of existing databases and analysis tools	See Annex 1 for specific milestones, indicators and time frames	GCP annual reports external or commissioned reviews	Adequate high quality bioinformatics and programming expertise can be contracted by GCP and its partners
<b>Activity 1.4.3:</b> Development of an information network of GCP data sources and analysis tools over the Internet and integration with international bioinformatics resources	See Annex 1 for specific milestones, indicators and time frames	GCP Web site and Web Service offerings	Adequate high quality bioinformatics and programming expertise can be contracted by GCP and its partners; International bioinformatics resources retain their support
<b>Activity 1.4.4:</b> Develop new data processing and analysis tools that serve the needs of Sub Programmes 1, 2 and 3.	See Annex 1 for specific milestones, indicators and time frames	GCP Web site and Web Service offerings	Adequate high quality bioinformatics and programming expertise can be contracted by GCP and its partners
<b>OBJECTIVE 2:</b> Strengthen the capacity of NARS and GCP consortium scientists to apply the tools of genomics, molecular biology and bioinformatics to the analysis of genetic diversity held in germplasm collections, and to use this knowledge to improve crop breeding programs and to develop new stress-tolerant varieties		Participant lists and presentation topics in GCP sponsored events and lists of GCP sponsored participants and their presentation titles/abstracts in international or national research meetings	
<b>Output 2.1:</b> NARS scientists with capacity for full participation in CP	See Annex 1 for specific milestones, indicators and time frames	Lists of GCP-funded research partners	NARS make their scientists available to participate in GCP
<b>Activity 2.1.1:</b> Needs assessment conducted	See Annex 1 for specific milestones, indicators and time frames	Reports of needs assessment workshops and other consultations	NARS complete questionnaires and follow up activities
<b>Activity 2.1.2:</b> NARS capacity building components built into commissioned research programs	See Annex 1 for specific milestones, indicators and time frames	Research plans and reports of commissioned research	NARS make their scientists available to participate in GCP

<b>Activity 2.1.3:</b> NARS capacity building built into competitive grants program	See Annex 1 for specific milestones, indicators and time frames	Research plans and reports of competitive grant proposals	NARS make their scientists available to participate in GCP
<b>Activity 2.1.4:</b> Training courses or workshops on specific technologies and areas conducted	See Annex 1 for specific milestones, indicators and time frames	Programs and summaries of training courses	NARS make their scientists available to participate in GCP; training sites with adequate bandwidth and other support facilities identified in each region
<b>Output 2.2:</b> NARS Research Centers serves as Regional Research Hubs		Records of resources allocated for this purpose; GCP and NARS annual reports	
<b>Activity 2.2.1:</b> Specific needs of NARS RRHs addressed	See Annex 1 for specific milestones, indicators and time frames	Correspondence between needs assessment reports and actual expenditures	NARS willing to make facilities and personnel available for regional/international program
<b>Output 2.3:</b> Crop improvement networks in place to incorporate novel alleles/traits		Funding allocations and expenditure reports for crop improvement networks, annual reports of network coordinators, annual reports of participating NARS, and GCP annual reports and annual meeting summaries	
<b>Activity 2.3.1:</b> Meta analysis of existing and recent crop improvement networks with McKnight CCR and Harvest Plus	See Annex 1 for specific milestones, indicators and time frames	Report on the meta analysis	Reports of past reviews of networks are available, complete and accurate
<b>Activity 2.3.2:</b> Development of regional improvement networks for multiple crops (in close cooperation with HarvestPlus and NGOs)	See Annex 1 for specific milestones, indicators and time frames	Funding allocations and expenditure reports for crop improvement networks, annual reports of network coordinators, annual reports of participating NARS, and GCP annual reports and annual meeting summaries	NARS willing to make facilities and personnel available for regional/international program; additional resources procured for this purpose
<b>Activity 2.3.3:</b> Ex ante and longitudinal assessments of activities to maintain to relevance and focus on needs of resource poor farmers	See Annex 1 for specific milestones, indicators and time frames	Published and internal GCP reports of the analyses	Additional resources procured with IFPRI and ILAC to support activities; crop modelers and economists can maintain an effective working relationship

### **ANNEX 3. GENERATION CHALLENGE PROGRAM CONSORTIUM MEMBERS**

Agropolis

Brazilian Agricultural Research Corporation (Embrapa)

Chinese Academy of Agricultural Sciences (CAAS)

Cornell University

International Center for Tropical Agriculture (CIAT)

International Maize and Wheat Improvement Center (CIMMYT)

International Potato Center (CIP)

International Center for Agricultural Research in the Dry Areas (ICARDA)

International Crops Research Institute for the Semi-Arid Tropics (ICRISAT)

International Institute for Tropical Agriculture (IITA)

International Plant Genetic Resources Institute (IPGRI)

International Rice Research Institute (IRRI)

John Innes Centre

National Institute of Agrobiological Sciences (NIAS-Japan)

Wageningen University