



## **Generation Challenge Program**

**2005 Annual Report for the Executive Council of the CGIAR  
4 September 2006**

Submitted by Jean-Marcel Ribaut, Director, and  
Jenny Nelson, Communications Manager

Generation Challenge Program  
c/o CIMMYT  
Apdo. Postal 6-641  
06600 Mexico, D.F., Mexico  
+52 55 5804 2004

## **Contents**

<b>1. Executive Summary</b>	<b>3</b>
<b>2. Background</b>	<b>4</b>
2.1 Program objectives and structure	
2.2 Research strategy and priorities	
<b>3. Research Accomplishments</b>	<b>7</b>
3.1 Overview	
3.2 Technical Report and Outputs by Subprogram	
<b>4. Progress on other CP Activities</b>	<b>15</b>
4.1 Capacity Building	
4.2 Data Management	
4.3 Communications/Public awareness	
4.4 Others	
<b>5 Governance and Management</b>	<b>18</b>
5.1 Program Steering Committee	
5.2 Program Management Team	
5.3 Operational Issues and Challenges	
<b>6. Finance</b>	<b>19</b>
6.1 Financial Objectives and Outcomes	
6.2 Schedule of Contributions Received (by CP donor and amount)	
6.3 Schedule of Disbursements to Partners (CGIAR and outside)	
6.4 Resource Allocation/Expenditure (by subprogram)	
6.5 Other Issues on Financial Management	
<b>7. Lessons Learned</b>	<b>21</b>
<b>Appendix A. 2005 GCP Projects and 2005 Subprogram Technical Reports</b>	<b>23</b>
<b>Appendix B. Schedule of Contribution Received and Schedule of Disbursements to Partners</b>	<b>49</b>
<b>Appendix C: Resource Allocation/Expenditure</b>	<b>51</b>
<b>Appendix D. Products and Potential Users Table (for use by competitive grants proposal submitters)</b>	<b>52</b>

## 1. Executive Summary

The completion of the Generation Challenge Program's (GCP) second year was marked by a definitive shift from creating the underpinnings of the organization to implementing activities. In 2004, operating procedures were defined, partnerships were formed, the vision and mission of the fledgling program were articulated, and the path was formulated for meeting the GCP's objectives. In 2005, significant strides were made toward executing those well-laid plans.

Research activities commenced in January 2005 on the first round of competitive research grants—17 three-year projects of approximately US\$ 1 million each. Early 2005 also saw the work begin on a fresh round of commissioned grants, which will serve as the basis of the GCP platform of tools and technologies for genetic studies and applications. In total, the GCP initiated 70 competitive and commissioned research projects and capacity building activities in 2005.

The first products of the GCP are now available. For example, 21 of the 22 mandate crops<sup>1</sup> of the GCP have been genotyped and analyzed,<sup>2</sup> and reference collections have been derived for each of them. The GCP also established a genotyping support service in 2005 to aid national agricultural research systems (NARS) in characterizing their germplasm and comparing it with the GCP reference collections. The reference collections and genotyping support service have direct and important applications for plant breeders, who are critical users of GCP products. A summary of all GCP 2005 research outputs can be found at the end of this Executive Summary.

Implementation of research and capacity building activities can also be seen in the healthy growth of the GCP's assemblage of partners, which now includes more than 30 national programs from developing countries and over 25 advanced research institutes, in addition to the 18 GCP consortium members (nine Future Harvest Centers of the CGIAR, five advanced research institutions, and four national agricultural research systems [NARS]). Private sector involvement is also increasing. Hundreds of scientists around the world are now engaged in or affiliated with GCP research. In 2005, new mechanisms for interaction and project management were developed to support the new collaborations that the GCP incubated, thus enhancing the effective flow of funds, reporting, data, and project outputs.

Exemplifying the shift from planning to action in 2005 was the development of the GCP's delivery strategy<sup>3</sup>. While in 2004 the GCP focused on crafting an effective scientific strategy, in 2005 effort was directed to detailing how the program's science will substantially help plant breeders to improve germplasm to ultimately benefit the resource poor. Full integration of the scientific and delivery strategies is now underway in the form of the overall GCP Strategy. In 2006, the GCP is finalizing its strategy and identifying research priorities for the coming years. The strategy is not a departure from the establishment documents of the GCP but rather a refinement of them. The GCP expects to publish its strategy and priorities by the end of 2006, following approval by the Program Steering Committee.

---

<sup>1</sup> Andean roots and tubers, barley, cassava, chickpea, coconut, cowpea, finger millet, forages, groundnut, lentil, maize, *Musa*, pearl millet, *Phaseolus*, pigeon pea, potato, rice, sorghum, soybean, sweet potato, wheat, and yam.

<sup>2</sup> Up to 3,000 accessions with 50 SSR markers.

<sup>3</sup> See [http://www.generationcp.org/capcomer/Final\\_Delivery\\_Strategy.pdf](http://www.generationcp.org/capcomer/Final_Delivery_Strategy.pdf).

The GCP's second Medium Term Plan (MTP), articulated in mid-2005, further refined how the GCP organises its research toward medium- and long-term objectives. The 'projects' identified in the 2006-2008 Medium Term Plan under each subprogram (SP) reflect a maturation of the objectives of each SP and Generation as a whole. 2005 marked the first year of both the competitive grants and the Management Team-led commissioned projects (the 2004 portfolio was determined at the Technical Planning Meeting in Wageningen in 2003, before the Subprogram Leaders were on board), and so provided the first opportunity to map out the GCP portfolio in the 2006-2008 MTP in a clear and coherent way. Also reflected in the plan is a profound commitment to ensuring that the GCP's research outputs can be accessed by the wider research community, especially scientists in developing countries. Producing the Medium Term Plan has proven to be a very useful annual exercise for the GCP Management Team, and it has become a crucial tool for planning, reporting, and identifying areas for improvement.

The 2005 Annual Research Meeting in Rome was another important milestone. The ARM gives GCP scientists and partners the opportunity to hear firsthand about the research progress made by their colleagues during the preceding year. Discussions and Q&A sessions were notably lively following the numerous research presentations and plenaries. Partners were also availed many opportunities for face-to-face discussions to coordinate proposal development for new commissioned research projects, to start in January 2006. For the first time, mid-year reports of all ongoing projects were compiled and published as the ARM proceedings.

Finally, 2005 was also a year of administrative transition for Generation. Founding director Robert Zeigler, who oversaw the launch and establishment of the GCP, left in March to assume the position of Director General of the International Rice Research Institute (IRRI). Subprogram 4 leader Theo van Hintum served as interim director until July 2005, when Jean-Marcel Ribaut, formerly Biotechnology Group Leader at the International Maize and Wheat Improvement Center (CIMMYT), was appointed GCP Director. In addition, Subprogram 3 Leader, Jonathan Crouch, who was recently appointed head of the CIMMYT Genetic Resources Unit, departed the GCP. Recruitment for a new Subprogram 3 Leader began in early 2006.

This Annual Report, which is comprised of excerpts from the GCP's 2005 Annual Report and 2006 Workplan, was approved by the GCP's Program Steering Committee at its November 2005 meeting.

Table 1. Major Achievements in 2005

- |   |
|---|
| <ul style="list-style-type: none"> <li>• Reference sets of germplasm produced for 21 crops and significant progress made in marker development and application</li> <li>• GCP phenotyping capacity enhanced and refined</li> <li>• Genetic base of rice increased through systematic introgression of chromosome segments from related species into cultivated rice</li> <li>• Conserved orthologous markers linked on different species maps, providing the dual benefits of integrating functionality and genomic positions (at least within a crop group) and serves as an important long-term tool for using comparative genomics to identify common genome regions controlling target traits.</li> <li>• New candidate genes and gene-based markers for abiotic stress tolerance (low-P, aluminium) revealed through genome-wide expression analysis and isolation of QTL</li> </ul> |
|---|

- Inclusive teams established to develop molecular breeding systems, and new molecular breeding systems pilot-tested and refined
- Low-cost assay technologies developed for NARS and small- and medium-size breeding programs
- GCP ‘knowledge base’ under construction, including data templates and repositories, tools for data analysis, and access points for databases
- Web services implemented at several GCP institutions, allowing scientists to access and share data with the global research community
- Software development standards created and GCPWiki launched, now serving as a platform for development of software and sharing of bioinformatics tools and resources
- Capacity building accelerated to full speed: six GCP training courses and numerous other project workshops organised to train NARS scientists in analysis and application of plant genetic diversity, association genetics, DNA extraction methods and data mining tool, project proposal development, and others. Fellowships (8) and travel grants (30) awarded to support NARS scientists research in the themes of the GCP
- Microsatellite marker kits developed for 7 crops, allowing researchers at any lab anywhere in the world to compare the genetic diversity of their germplasm collection to a GCP reference sample.
- “Genetic Resource Policies and the Generation Challenge Programme” published, outlining the international and national policy context within which the GCP functions
- GCP Delivery Strategy developed, defining “users” and “products” in the GCP context and establishing how the GCP will ensure impacts

## 2. Background

### 2.1 Program Objectives and Structure

Created in 2003, the Generation Challenge Program is one of the three time-bound, independently governed challenge programs launched by the Consultative Group on International Agricultural Research (CGIAR) with the specific objective of catalyzing high-impact research that contributes substantially and rapidly to global development goals.

The mission of the Generation Challenge Program is to serve as a research and capacity building network that uses plant genetic diversity, advanced genomic science, and comparative biology to develop tools and technologies that enable plant breeders in the developing world to produce better crop varieties for resource-poor farmers. In line with this mission, by 2013 Generation is expected to contribute to the following objectives:

- Provide access to and promote the use of genetic diversity in plant improvement programs.
- Develop a public platform of genetic and genomic resources and tools, and support a global community that can use them.
- Generate and apply knowledge across crops, and demonstrate the potential of comparative genomics to impact plant improvement programs.

- Use genetic diversity and advanced science to develop products for plant breeding programs to improve the livelihoods of resource-poor farmers in marginal, drought-prone environments.

To realize these objectives, Generation activities are organized in five subprograms: (1) Genetic Diversity of Global Genetic Resources; (2) Comparative Genomics for Gene Discovery; (3) Trait Capture for Crop Improvement; (4) Genetic Resources, Genomic and Crop Information Systems, and Bioinformatics; and (5) Capacity Building and Enabling Delivery.

A unique and critical role that the Generation Challenge Program plays is as a global platform for understanding, accessing, and applying crop genetic diversity. Many national programs and other research institutions, particularly the CGIAR centers, have genebanks where they conserve, in trust for humanity, hundreds of thousands of crop accessions. Generation is the largest effort to date to systematically study those collections to better understand their potential for breeding improved varieties. All 22 “mandate” crops of the CGIAR research centers are included in the Generation mandate: Andean roots and tubers, barley, cassava, chickpea, coconut, cowpea, finger millet, forages, groundnut, lentil, maize, *Musa*, pearl millet, *Phaseolus*, pigeon pea, potato, rice, sorghum, soybean, sweet potato, wheat, and yam.

Many research institutions work in the same scientific domains as the Generation Challenge Program. By bringing together leading players in modern biology, however, and linking them to national research programs, Generation represents a new type of collaboration, both in magnitude and strategic focus, to fill the gap between biotechnology and practical applications that will benefit the world’s poor farmers. The Generation network currently consists of 22 consortium members (9 CGIAR centers, 7 national research systems, and 6 advanced research institutions). The network has expanded year by year. In 2005, in addition to the consortium members, more than 25 advanced research institutions from the North and South and 30 national research institutes were active partners in Generation research projects. A broad cross-section of stakeholders is represented and participates in Generation governance.

The Generation Challenge Program employs two funding schemes to carry out its research: competitive and commissioned grants. These complementary schemes provide flexibility to adjust the research portfolio to capture emerging opportunities, promote innovative partnerships, develop appropriate product delivery schemes, and balance resources across mandated activities. Currently the Program supports 17 competitive projects and 53 commissioned projects, representing US\$12.7 million in research investment in 2006 alone. The second call for competitive GCP grants, which will total US\$2 million per year over two years, will open in early 2006; projects will begin in 2007.

Generation’s fundamental efforts to develop new knowledge and products support the United Nations Millennium Development Goal of halving, by 2015, the number of hungry people and those living on less than a dollar a day. If these efforts result in the development of superior varieties that are used by farmers, they offer potential for poor farm households to improve food and nutritional security and income. They also offer the prospect of cheaper food for poor consumers. Most of the world’s poorest people continue to live in rural areas, where the production and consumption of staple food crops remain critical to livelihoods and food security.

Generation activities also contribute to the CGIAR System Priorities, especially Priority Area 2: *Producing more and better food at lower cost through genetic improvement*.

## **2.2 Research Strategy and Priorities**

The Generation Challenge Program was founded to harness the intersecting revolutions in molecular biology, bioinformatics, and communications, and to bring together diverse institutions to nourish the biological knowledge base of genetic diversity and genomics. From that, the GCP develops practical applications for plant breeding for marginal environments. Improving genetic gain under drought conditions has been identified as the area of focus for the GCP.

The GCP was created to address a broad mandate: develop a platform of genetic and genomic resources for plant breeding, to contribute to improved livelihoods among resource-poor farmers. The GCP's research strategy must balance the dual responsibility of both creating knowledge for the global research community and developing useful products for plant breeders. The GCP conceptualizes its research as a matrix of 'horizontal' and 'vertical' activities. Horizontal activities are those which generate knowledge, methodologies, and resources for use by the global research community. These activities address cross-cutting biological questions at different levels of plant architecture and across a broad set of crops, to further the research community's understanding of gene function and stress tolerance mechanisms. Vertical activities target a limited number of traits in as few as one crop and are aimed at producing tools or other outputs that can be directly applied in plant breeding.

The development of the GCP's delivery strategy in 2005 was a critical turning point for the GCP. Research planning and implementation of administrative structures had occupied the GCP during its first year, but now it was time to begin thinking about how GCP outputs are going to be delivered to breeding programs and be developed into useful crop varieties. The GCP is expected to make impacts on resource-poor farmers; therefore, we must ensure that our research outputs are delivered to users who can apply them in the development of improved varieties for farmers. The GCP delivery strategy defines what GCP products are and who the users of those products are. The delivery strategy requires commitment by the GCP to identify and match specific products to specific users, thus ensuring that (i) projects are conceived and planned with a product and user in mind, and (ii) that the designated users contribute to the overall development and delivery pipeline by using GCP products to develop other products geared more closely to farmers' needs.

In the delivery strategy, the GCP also to building capacity for the first tier of users of GCP products, as those individuals often need training before they can effectively utilise GCP products. Although the GCP will not provide finished varieties to farmers, through capacity building we strengthen our capability to ultimately deliver benefits to farmers.

## **3. Research Accomplishments**

### **3.1 Overview of GCP Progress in 2005 and First Products**

In 2005, research activities accelerated greatly. Early in the year, letters of award for the 17 competitive grants and roughly 50 commissioned projects were issued to the principal investigating institutions. Initiating GCP research projects, with the many new partnerships and

collaborations that were formed, was a considerable undertaking. Some projects got off to late or slow starts due to a range of problems: funds arriving late as a result of delays in delivering institutional financial reports, contracts, or subcontracts; changes in designated partners; or because 2005 work relied on as-yet-uncompleted 2004 work. A number of project management policies were developed or revised throughout the year to respond to these issues. Despite the unforeseen problems, the majority of projects started on schedule, and significant results were reported at the Annual Research Meeting in September 2005.

The GCP's competitive grants scheme is intended to attract the best science to the program, and in the first round of competitive grants this was achieved through a broad call for proposals that encouraged diverse ideas for scientific innovation. The second round of GCP competitive grants will open on 15 February 2006. The new call for proposals for competitive grants was developed with the principles of the GCP Delivery Strategy in mind, namely the ideas that GCP research should be targeted to developing country breeding programs and should result in products that can be directly applied by them in their crop improvement efforts. In addition, full proposal submitters are expected to submit a 'Products and Users' table (see Appendix D) that details exactly what their proposed project will produce and the names of the scientists, particularly those at NARS institutions, who the products will be delivered to. This table will be used as the basis for the project delivery plans, to be developed at a kick-off meeting for each new competitive project.

In accord with Generation's efforts to streamline its research portfolio, the second call for proposals is targeted toward five "thematic research areas": (1) association studies to validate candidate polymorphisms/genes, (2) refinement of drought tolerance through the documentation of new traits, and (3) development and application of gene-based markers; (4) the generation of new populations for introgression of new alleles from wide hybridization; and (5) the advancement of improved populations through molecular breeding. Restricting the call to these five thematic areas should still attract the best science and scientists to the program—now that the GCP has established its reputation for quality—but will also narrow the GCP's research focus to a more manageable scope considering resources available. As mentioned above, each new competitive grant will be required, with the support of the GCP and external experts, to develop a detailed delivery plan, including capacity building activities needed for product delivery to the users.

The 2005 commissioned research portfolio provided cohesion to the GCP's agenda of scientific activities. Many of the new projects for 2006 grew out of the 2005 projects, and thus ensure continuity in the research while building linkages among projects. New 2006 commissioned projects include important and exciting research in the fields of allele mining, genomics, and gene discovery. The list of 2006 projects may be found in the appendices of this document. Scientific results and progress achieved in 2005 in the competitive and commissioned research projects is presented in detail in the Subprogram Updates section of this report.

Capacity building embedded in the research activities is strongly encouraged in both the competitive and commissioned project schemes. A system for tracking the beneficiaries of project-based capacity building was implemented in 2005 in Subprogram 5, and the fellowship and travel grant schemes of the GCP will target supporting more of such activities in 2006. The GCP Training Program also began in earnest this year, with courses on plant genetic diversity and marker-assisted breeding offered in three regions (Asia, Africa, and Latin America) to NARS scientists, along with courses on project proposal development, also offered in the three regions.

Only two years into this Challenge Program, the first GCP products are being delivered. The GCP has 22 mandate crops and, thanks to having established a network of committed partners, researchers now have access to the major sources of genetic diversity for all of them. The genotyping and analysis of these crops (up to 3,000 accessions of 21 crops were analyzed with different sets of about 50 SSR markers) was a central GCP activity during 2004 and 2005, which produced a massive set of data representing the first large-scale look at the genetic diversity of so many staple crops. The data was generated using standardised protocols and the same set of markers for each crop, which the GCP calls 'Microsatellite Kits.' Once published (in 2006, online and in hard copy), the Microsatellite Kits will allow researchers at any institution anywhere in the world to evaluate the genetic diversity of their germplasm collections or breeding material and reference it with that of respective collections in CGIAR and other national gene banks. This has direct and important applications for plant breeding.

But even given the information now available from the GCP's work, many institutions that would like to better characterise their germplasm do not have access to the equipment or expertise needed for such endeavours. The GCP established a genotyping support service in late 2005 to facilitate this by providing training in genotyping methods and protocols and assistance with data production and analysis. Institutions using this service can evaluate their germplasm against the GCP reference germplasm, which yields benefits for them and the GCP: they can see where their collections are situated within the global scope of diversity for the crop; in turn, the data from the outside collections can help verify the diversity captured in the GCP global reference sets. Taken together, the Microsatellite Kits and the genotyping support service demonstrate that some GCP 'products' can have direct application to the selection of parental lines for new crosses, thereby impacting breeding activities within and outside the GCP.

### **3.2 Technical Report and Outputs by Subprogram**

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

Subprogram 1 (SP1) focuses on identifying novel, diversified, and superior variants of genes involved in target traits. SP1 activities are closely to Subprogram 2 (SP2) outputs, and like SP2, SP1 outputs are aimed at direct applications in Subprogram 3. In accordance with the GCP's global priorities, particular emphasis is given to drought tolerance.

Access to source of genetic diversity that may supply genes and alleles involved in key agricultural traits is essential to the mission and objectives of the GCP. Through the GCP's network of consortium members and partners, vast germplasm collections are available to the program, but to unlock the genetic diversity present in those collections, the structure of the collections must be understood through coordinated surveys of molecular and phenotypic variation. With hundreds of thousands of accessions across many different crops collected in gene banks around the world, high-throughput molecular screening techniques must be developed to genotype the collections for fast results. Appropriate screening techniques to phenotype the collections for traits such as drought tolerance also need to be developed and applied to obtain reliable and analyzable phenotype descriptions accompanied by the associated descriptions of environmental and weather conditions. Once the genotypes and phenotypes have been established, association studies and/or population genetic approaches must be conducted to illustrate their interactions.

General objectives and outputs for 2005 in Subprogram 1 are as follows. The full technical report on all projects and outputs in Subprogram 1 can be found in Appendix A:

***Create an improved understanding of the structure of diversity for the world's major food crops***

- Progress made in genotyping, selection of core samples, marker development and application, and evaluation of whole-genome survey methods: Subprogram 1 is charged with the characterization of genetic resources for the purpose of describing the global genetic diversity in staple crops, knowledge which is critical to crop improvement programs. Significant progress has been made in this activity, which involves first the collation of information on the various germplasm collections (the “composite set”) and then extraction of a representative sample (“core sample”) of 200-3,000 accessions that best represent the diversity available in each crop. In 2005, core samples were finalized for 21 crops. Major progress was also recorded in systematic efforts to use molecular markers to structure the diversity in germplasm collections that support active crop breeding programs in the CGIAR Centers. The array of molecular markers that have been mastered and implemented is widening. The throughput of genotyping with SSR markers is increasing; several partner laboratories have improved local organization since the beginning of the activity in 2004. Complementary methods with proven efficiency are now available: DArT as a whole-genome genotyping method is remarkable for its high throughput and low cost, and EcoTILLing is impressive for its simplicity with the use of agarose gels.
- Genetic diversity characterized using molecular methods: This activity has been conducted within several competitive grant projects. For example, a project led by CIMMYT seeks to develop informative DNA markers through association mapping in maize to improve drought tolerance in cereals. The project focuses on candidate genes involved in carbohydrate and ABA pathways, and plant phenotyping will be conducted in 5 environments across 3 continents. Another project, led by CIAT, explores natural genetic variation; it is developing genomic resources and introgression lines for four AA genome rice relatives and uses SSRs and SNPs to monitor interspecific introgression. The CIAT project is an example of how additional gene sources are mobilised by broadening the genetic base of the pools accessible for recombination and breeding.

***Establish and implement a scientific and organizational framework to describe tolerance to drought***

- Phenotyping capacity of GCP enhanced and refined: The phenotyping network under development at Embrapa (Brazil) has considerably optimized access to efficient, coordinated multilocational phenotyping platforms, supporting the evaluation by environment descriptions and drawing experience from advanced physiological characterization performed in crop-specific projects. This network now works in coordination with the whole plant modelling project led by CIRAD, enhancing capacity in both projects.

***Develop novel materials with new genes and alleles for genetic studies and plant improvement***

- Genetic base in rice increased: Complementary to the typical crop reference samples under development, the current germplasm base in rice is being widened by the production of

novel materials through systematic introgression of chromosome segments from related species into cultivated rice.

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

Plant traits for adaptation to environmental stresses are often controlled by complex genetic systems subject to influence by genotype x environment interactions. To effectively combine the right complements of genes and alleles in a breeding program, we need to have an adequate understanding of the genetic mechanisms underlying adaptive processes. Such an understanding is particularly important in cases such as drought tolerance, where the genetic effects are often small and the phenotypes are difficult to measure. Advances in genomic tools and knowledge from model organisms provide exciting opportunities to dissect the genetic control of complex traits and identify potentially useful genes. Yet, practical applications of the new tools for agronomic improvement require a level of integration that is often difficult to implement by individual disciplines alone.

To achieve an understanding of the genetic mechanisms underlying the adaptive processes, a scientific and collaborative environment to enable gene discovery as well as applications is needed. To achieve this, cross-cutting research platforms for effective applications of genomic tools and knowledge to decipher genetic control of complex traits must be established. Using these platforms, the genes to alleviate the targeted problems can be identified efficiently by pooling resources and expertise. To realize the potential of these approaches, however, capacity to apply new tools must be enhanced and a pipeline to move results into practice must be developed. Demonstrating the success of these approaches in a few targeted cases in the short- to medium-term is important to help lay the road map for broad applications of these new areas of science.

To meet these objectives, Subprogram 2 is designed to maximize the use of genomic and genetic resources available in the research community. We support the production of specialised stocks that will elevate the level of genetic research in different crops, and we apply comparative approaches to leverage genetic knowledge from multiple plant species to investigate and validate gene functions important for stress tolerance.

General objectives and outputs for 2005 in Subprogram 2 are as follows. The full technical report on all projects and outputs in Subprogram 2 can be found in Appendix A:

#### ***Assemble genomics and germplasm resources through consolidating and developing specialized genetic stocks and framework genetic markers***

- Progress made in assembling genetic and genomic stocks, and resources successfully leveraged from various partners: The assembly and development of specialised genetic and genomic resources have encouraged many players to contribute to this effort, which has mobilised resources from leading institutions and stimulated partners to create and utilise specialised genetic stocks for gene discovery. For example, the rice mutant network has linked multiple laboratories to apply the mutant resources in systematic phenotyping. Currently, the OryGenes DB (<http://orygenesdb.cirad.fr>) developed at CIRAD (France) has about 56,000 publicly available insertion sequences tagged by FST sequence on the rice genome sequence as annotated by TIGR. As of mid-2005, the OryGenes DB was used to search for knockout inserts in candidate stress-associated genes and had a 22% success rate, providing a treasure trove for extracting mutants of many plant genes.

***Develop comparative maps within and across species and framework genetic markers for target crops***

- Conserved markers developed to link different maps: The use of orthologous markers to link different species maps has the dual benefits of integrating functionality and genomic positions (at least within a crop group) and will be an important long-term tool for using comparative genomics to identify common genome regions controlling target traits. Despite initial difficulties, positive progress is seen in the development of conserved markers to link different species (within the dicots and the monocots). We expect progress from the marker projects through use of information being generated by the research community at large and through more extensive international linkages.

***Conduct genome-wide expression analyses to reveal new candidate genes***

- Positive results are seen from experiments using genome-wide expression as a means to identify candidate genes to contribute to target phenotypes. A large gene expression dataset has been generated using available rice oligo chips from NIAS (Japan) (22K) and the Beijing Genomics Institute (60K) to correlate gene expression with genetic regions expressing QTL or mutant phenotypes. Integration of QTL and expression analyses can be used to narrow the choice of candidate genes and determine the causal relationships between expression and phenotypes.

***Derive gene-based markers through isolation of large QTL***

- Progress in the fine mapping and near-isolation of tolerance genes to aluminum toxicity, phosphorus deficiency, and salinity are encouraging. The cloning of the aluminum toxicity tolerance gene in sorghum will hasten the development of elite breeding lines. Markers tightly linked to phosphorus uptake efficiency are being used in a selection program to combine traits suitable for the drought-prone environments in Indonesia. We can fast-track the delivery of useful genes to breeding programs to address specific problems arising from stresses that are often associated with drought-prone environments. The availability of such markers will also help build capacity in breeding programs to apply marker-aided selection.

**Subprogram3: Trait Capture for Crop Improvement**

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. However, the recent developments in genomics, computational systems, and biometrics offer a real opportunity for simultaneously manipulating the component traits of drought tolerance. Yet the greater challenge remains to use this knowledge and skill to develop products that will have significant impact on the livelihoods of farmers in resource-poor cropping systems. To create such projects will require a substantial change in how public sector scientists operate within multidisciplinary teams and across organizations.

Global research progress in many of the cereals is sufficient to begin the development and application of gene-based marker systems for components of tolerance to drought and other abiotic stresses (rice, maize, sorghum, wheat, barley). Thus, emphasis in these crops is more on the translation and/or application of preexisting research outputs. However, additional targeted

investments are required for example in pearl millet (the most drought tolerant but least studied of the major cereal crops).

Conversely, the global genomics researchers and resources in the legume and clonal crops are still well below critical mass. Unfortunately, the resources currently available to the GCP are insufficient to support a comprehensive program in all crops. For this reason, careful prioritisation of crop focuses will be applied to ensure rapid and compelling proof-of-concept in key representative crop: for example cowpea (legumes), ground nut (oilseeds), and cassava (clonal crops).

Subprogram 3 strives to create product-driven teams that span all R&D levels in the innovation to impact continuum. In particular, we aim to bridge the gap between research outputs and product delivery. Alliances with the private sector (both multi-national corporations and small- to medium-size enterprises) will be especially critical to achieve these goals. In addition, significant direct spillovers from sequence, gene, and trait analyses in model species are expected to substantially impact progress in those closely related crops with a minimum of genomics resources and critical mass expertise. All crops are likely to benefit from generic advances in genomic platform technologies, low cost marker screening technologies, and molecular breeding simulation and decision-support systems.

General objectives and outputs for 2005 in Subprogram 3 are as follows. The full technical report on all projects and outputs in Subprogram 3 can be found in Appendix A:

***Establish multi-institute, multi-disciplinary molecular breeding teams***

- Subprogram 3 projects have established strong collaboration across disciplines, crops, and types of institution. For example, the project ‘Development of Low-Cost Technologies for Pyramiding Useful Genes from Wild Relatives of Cassava into Elite Progenitors’ brings together scientists in a number of African and Latin American NARS, while the project ‘Development of Low-tech, Gene-based Trait Assay Technologies in Rice and Wheat’ brings together scientists in a number of African and Asian NARS.

***Forge new paradigms in plant breeding in collaboration with users (CGIAR, NARS, university, and SME breeding programs)***

- GCP investments are already driving exciting new approaches to breeding lesser studied crops. For example, the project ‘Unlocking the Genetic Diversity in Peanut's Wild Relatives with Genomic and Genetic Tools’ uses wide crosses and molecular marker analysis to drive a new paradigm in groundnut breeding based on the successful use of synthetics in breeding of other complex polyploid crops, such as wheat and canola. Based on the narrow genetic base of groundnut varieties, this approach is likely to have large impacts on groundnut breeding gains.

***Test molecular breeding systems with simply inherited traits***

- It is important for less-studied crops that we move ahead with whatever technologies are available. For this reason, several SP3 projects focus on simply inherited traits pending the availability of resources for drought tolerance. Thus the proof-of-concept for using synthetic germplasm in groundnut breeding is being carried out via multiple disease resistance, whereas pest resistance is the focus for deploying MAS in cowpea. In both

projects, efforts are simultaneously being made to generate the necessary resources for mapping and MAS of drought tolerance.

***Develop low-cost assay technologies for NARS and small- and medium-size breeding programs***

- Good progress is being made in the development of low-cost assay technologies for gene-based MAS of disease resistance in rice and of grain quality in maize, and for linked markers for pest and disease resistance in cassava. These proof-of-concept activities provide essential methodological insights for routine, large-scale marker conversion activities, once GCP gene-based technologies for drought tolerance emerge.

**Subprogram 4: Bioinformatics**

Subprogramme 4 (SP4) addresses the challenge of linking and integrating the GCP information components, and analysis tools into a coherent information gateway, and supporting GCP activities in terms of bioinformatics tools.

SP4 aims at implementing a strategy that allows all Generation CP data to be accessed and shared by the consortium and by the rest of the world. A second objective is to make sure that the components of the resulting platform are of sufficient quality, both in terms of the actual data and the tools to analyse and manage that data. A third and final objective is to create the necessary tools, and to generate the knowledge to support the research in the first three subprogrammes.

General objectives and outputs for 2005 in Subprogram 4 are as follows. The full technical report on all projects and outputs in Subprogram 4 can be found in Appendix A:

***Establish the GCP information platform for information exchange and improve its components***

- GCP ‘knowledge base’ under development: The GCPWiki serves as an example of the type of infrastructure that is being developed through SP4. GCPWiki provides an environment in which users contribute to the content of the web pages. As such, it allows the joint development of documents and, in the process, the creation of a common knowledge base. Since the system also stores older documents, allows users to compare versions, and easily see which documents have been changed recently, users can have full control over the evolving content. This model has been followed in several SP4 activities and is being promoted in the other Subprograms.
- Software development standards developed: To support all software development, a recommended platform has been selected with respect to language, software development tools, and type of environment (open source). On this basis, a large number of projects have developed applications that will be integrated into a common client next year.
- Web services implemented: A package for installation within the CGIAR Centers participating in the GCP was developed, training has been provided to technical staff, and many of the Centers have now installed their first GCP web services. These Centers are now ready to make the GCP data available to the GCP and the wider bioinformatics community—a fundamental part of the GCP mission.

## 4. Progress on Other CP Activities

### 4.1 Capacity Building

Subprogram 5 is responsible for capacity building in the Generation Challenge Program. Though it is managed in its own Subprogram, capacity building is considered integral to the success of the GCP, and thus also integral to the research activities underway.

Objectives and outputs for 2005 in Subprogram 5 are as follows. The detailed technical report for Subprogram 5 can be found in Appendix A:

#### *Build capacity of NARS researchers*

- **Training:** More than 150 NARS scientists participated in GCP training courses or hands-on training within GCP-funded projects. Emphasis was given to building linkages with a selected community of NARS scientists (for example, to provide training to accommodate a specific role in GCP research and delivery) and to strengthening links among partners. Training events were used to find out about the capacities of participants and their organizations to contribute to the aims of the GCP, especially to identify possible roles for them in the delivery chain of GCP products. These opportunities were also used to identify additional training and capacity building needs that the GCP could support under its mandate. As well as fulfilling expectations for training per se, courses and workshops played a significant role in raising general awareness of the GCP, the value of broad research partnerships, and the need to promote tighter linkages between laboratory and field scientists.
- **Fellowships, travel grants, and other activities:** Eight fellowships were awarded for short periods of research at GCP organizations, and one PhD fellowship was granted. Approximately 30 travel grants facilitated attendance at GCP-related conferences or helped to build linkages among Consortium scientists. Other activities were conducted in collaboration with other Subprograms to promote sharing of technical knowledge generated in the GCP with national program researchers outside the Consortium, so that the benefits could extend to their institutions.

#### *Research policy issues of relevance to the GCP and propose solutions*

- Protocols were analyzed and developed to facilitate germplasm exchange and proper access and benefit-sharing of the derivatives of the Program, in line with the published policies of the Convention on Biological Diversity and the FAO International Treaty on Plant Genetic Resources. Reports were produced for generic questions of immediate importance to the GCP, and a seminar was conducted to call attention to policy issues influencing the operation of the GCP, especially those related to access to genetic resources and intellectual property rights.

#### *Develop a delivery strategy to ensure delivery of GCP products to users*

- The newly completed Delivery Strategy,<sup>4</sup> a major milestone for the GCP, describes mechanisms that the Program will employ to ensure its products reach their intended users

---

<sup>4</sup> See [http://www.generationcp.org/capcomer/Final\\_Delivery\\_Strategy.pdf](http://www.generationcp.org/capcomer/Final_Delivery_Strategy.pdf).

(most often ‘intermediate’ users in the larger value chain of product development). The document defines the users of the outputs of GCP projects (who they are and how they will use these outputs) and establishes that every GCP project will have a product delivery plan in place, developed in coordination with the appropriate set of users of the project outputs. The aim of the Delivery Strategy is to ensure—from conception through implementation to completion of research projects—that GCP products reach the next level of users, who will in turn be able to produce another product in the value chain, linking laboratories to breeding programs to farmers’ fields.

## **4.2 Data Management**

Subprogram 4, which is devoted to the development and maintenance of genetic, genomic, and crop information systems and bioinformatics, is largely responsible for data management issues in the Generation Challenge Program (see section Research Accomplishments section above for specific outputs of Subprogram 4 in 2005). However, the issue of data release and availability is also of serious concern to the GCP, and some important steps were taken in 2005 to deal with this issue. Starting in 2006, all project proposals are required to include a section outlining what data is to be produced, in what format, and where it will be released upon the end of the project. To enforce this, a new clause was added to contractual agreements requiring project principal investigators to produce and share data as stated in the approved project proposal.

## **4.3 Communications/Public Awareness**

Communications plays a critical role in the operation and growth of the Generation Challenge Program. A major communications objective is that all members and affiliates of the GCP contribute to establishing Generation as an important player and partner-of-choice in the global research community. To that end, the GCP sponsors a session at the Plant and Animal Genome (PAG) Conference every year and supports key symposia and workshops that relate to GCP areas of interest. The GCP director and subprogram leaders participate in various international and national conferences, in addition to PAG, as GCP emissaries.

In 2005, the communications unit served as the project office of the GCP, drafting and distributing contracts, fielding queries from project offices at partner institutions, and proposing project management policies as needed. In 2006, these activities have been assigned to the new Project Officer.

Objectives and outputs for 2005 in GCP communications are as follows:

### ***Improve and manage information flow within and outside the GCP:***

- E-newsletter: a periodic compilation of announcements of interest, GCP events, training opportunities, etc., was distributed to all GCP member scientists, partners, and many stakeholders (distribution list contains over 1,200 contacts)
- Contacts database: The communications unit developed and maintains a growing database of over 1,200 GCP affiliates, partners, potential collaborators, and others who have expressed an interest in the program.
- Point of contact: with few full-time staff at GCP headquarters, the communications unit is the logical entry point for newcomers as well as GCP members.

### ***Coordinate public awareness efforts for target audiences:***

- GCP Website: the communications unit is responsible for the development and maintenance of the Generation website ([www.generationcp.org](http://www.generationcp.org)), which serves as the virtual library for GCP official documents as well as the public face of the GCP. The site receives over 4,000 hits per month.
- GCP in the news: the communications unit responds to requests for news bulletins, etc., and pursues opportunities for GCP features, interviews, and other mentions in various media outlets.
- GCP public awareness materials: the GCP brochure, folder, poster, bound reports, and other items were sent to thousands of people in 2005.
- Targeted public awareness materials for international conferences, specific campaigns, etc.: the communications unit provides the service of developing materials about GCP themes or activities for outside events.

***Package and communicate GCP outputs:***

- GCP products: the communications unit constructs web pages and publications to distribute and publicize GCP products such as informatics tools, training materials, protocols, and other useful information.
- Program reports: the development of the Annual Report, Medium Term Plan, donor proposals and reports, and all other official GCP documents are coordinated and edited and/or written by the communications unit.

#### **4.4 Project and Product Management Policies**

Several key achievements in the area of IP and policy were made in 2005:

***Reporting requirements and templates established for all projects:*** The GCP reporting schedule and templates for reports were developed and institutionalized in 2005. All projects must submit two technical updates per year (15 May and 15 October are the deadlines) for approval by the relevant subprogram leader and must submit a final technical report upon completion of the project. All projects over \$250,000 per year will also submit a yearly substantial technical report (31 December deadline). Financial reports are due up to 45 days after 31 December each year. Project funds for subsequent years will not be disbursed until technical and financial reports have been approved by the subprogram leader.

***GCP Review and Advisory Panel established:*** To enhance quality control of the science conducted within the GCP, a consultative team of external reviewers was appointed to review commissioned research and reinforce the Management Team's strategic decisions. The Review and Advisory Panel (RAP) is an important mechanism for the Management Team to bring in respected peers with outside perspectives to the strategic planning process and to employ them as reviewers of the commissioned project proposals.

***Humanitarian use agreement developed:*** The GCP consortium agreement provides that intellectual property developed by a GCP consortium member or supporting participant is the property of the institute that developed the IP. This is standard practice in joint research arrangements. However, in order to ensure that IP resulting from such research can be used for the purpose for which the work was funded, it is also a standard practice for the parties to agree that the resulting technology can be used for certain, specified purposes. Therefore, each Consortium

Member has a non-exclusive, royalty-free right to use Challenge Program IP for the activities with the aim to provide technology and products to the resource-poor on a royalty-free basis. The GCP Humanitarian Use agreement, approved by the Program Steering Committee in 2005, will be used in developing new agreements with non-consortium members.

***Contractual agreements streamlined:*** A contract was developed to allow the GCP to enter into agreements with non-consortium members, which was necessitated by the second call for competitive grants that will allow non-GCP consortium member institutions to serve as principal investigator on projects. Subcontract templates that GCP members and non-members can use for their subcontracting needs were also developed, one for small amounts of funds (for “services”) and one for larger amounts, both of which require the subcontracting party to adhere to the GCP’s IP policies.

## **5. Governance and Management**

### **5.1 Program Steering Committee**

The Program Steering Committee governs the GCP on the highest level, with input from the Program Advisory Committee and the Stakeholders Committee. The Program Steering Committee currently has 20 voting members; it consists of one representative of each consortium member institution (Director General or nominee), an Independent Chairperson, the GCP Director, and anyone else approved by the PSC. A representative of the Global Forum for Agricultural Research (GFAR) also sits on the PSC as an observer (non-voting member). In 2005, the Program Steering Committee offered provisional membership to four new organizations at its November 2005 meeting: INRA-Morocco, CINVESTAV-Mexico, BIOTEC-Thailand, and the Istituto Agronomico d’Oltremare (IAO) of Italy. Provisional membership provides all benefits of full membership except for membership on the Program Steering Committee, because the GCP Consortium Agreement is under review and there is no document currently for new members to sign. Once any changes to the Consortium Agreement are ratified, the provisional members will be offered full membership and will then be able to participate fully in GCP activities, including PSC membership.

The PSC meets once annually, where it makes most of its decisions and plans for the following year. At its 2005 meeting, in addition to the approval of the 4 new provisional members, the PSC also appointed a Task Force to review GCP governance issues and another Task Force to provide advice on management issues. The Task Force on governance will present its findings to the PSC during 2006 or at the 2006 meeting in December in Washington DC.

### **5.2 Program Management Team**

The GCP Management Team, who handles the day-to-day operation of the GCP, consists of the Director and five Subprogram Leaders (SPL), one for each subprogram. Within the Management Team, the Director handles as much of the administrative affairs as possible, including the donor contacts, etc. This allows the SPLs to concentrate on the content of their subprograms. To make sure that the SPLs stay rooted in their respective organizations, they are appointed for 50% of their time, with additional financial support to allow continuation of their ‘local’ research agendas. This implies that they are all employed and hosted by different institutions, all GCP Consortium

Members. This geographical dispersal of the managers has large implications on the communication between them. Therefore the Management Team tries to convene four times per year in conjunction with other meetings, such as the Annual Research Meeting and the Program Steering Committee Meeting. Phone conferences are held at least once per month, and, of course, there is a significant amount of email traffic between the Management Team.

In 2006, a new Subprogram 3 leader will be recruited (to be consistent with the executive summary. The position will be announced in the relevant fora (GCP members list, CIMMYT website and Human Resources, CGIAR distribution lists, etc.) as well as advertised in appropriate media outlets (New Scientist, Economist).

### **5.3 Operational Issues and Challenges**

Compared to the many achievements the GCP has made since its inception, the operational challenges are few but notable. Data sharing and management, as mentioned above, are important issues in the GCP, and important steps are being taken to deal with them. Another challenge is that the GCP is not an institution but rather a network of many institutions, and despite the numerous mechanisms we have put in place to minimize transaction costs, nothing beats face-to-face communication. The GCP website and various specialized portals serve as a main hub for information exchange within the GCP, and SP4 has pioneered a host of tools and software options for making virtual collaborations more efficient. The GCP Annual Research Meeting, however, remains the premier opportunity for communication each year in the GCP, when all of the project principal investigators come together to discuss their progress and other scientific issues. As transactions costs go, the cost of this meeting is relatively low considering how fruitful it is for GCP management and member scientists.

The major operational challenge the GCP is currently facing is that of insufficient staff levels to handle the mounting responsibilities of a \$13 million per year program. The Management Team of the GCP is composed of 5 scientists, who manage research within the GCP at 50% of their time and conduct research at their home institutions the other 50% of the time. While having a Management Team composed of active scientists is a point of pride in the GCP, this also means that the GCP does not have full time managers to monitor and evaluate research progress on the 70 projects which are currently running in the GCP. In addition, the GCP functions with a very small support staff of 4 people, who are responsible for the project management, finance, communications, web development, event planning, and many other duties in the GCP. The staffing issue will need to be resolved over the coming year.

A final issue is that of GCP governance. The Program Steering Committee of the GCP, which is composed of one representative per consortium member institution (of which there are currently 19) has recognized its cumbersome nature and has nominated a Task Force to recommend a revised structure.

## **6. Finance<sup>5</sup>**

---

<sup>5</sup> For easy reference on details and tables, please see GCP 2005 Annual Report and Year 3 Workplan.

The GCP has maintained a healthy financial condition since its inception, thanks in large part to donor community who generously support the GCP. Our major donors are the European Commission (EC), the UK Department for International Development (DFID), and the World Bank (WB), contributing about 90% to our total income.

Compared to 2004 (US\$ 10.9 M), our income increased in 2005 by about 28% (US\$ 14 M). This overall increase was the result of growth in the size of the contributions from both the EC and the World Bank. The level of contribution from our major donors is anticipated to remain at the same level in foreign currency terms for 2006. However, income for 2006 is projected to decline to US\$ 12.3M, reflecting the strengthening US dollar value relative to the currencies of our two major European donors. Budget projections have been based on a conservative exchange rate (5% below the lowest in December 2005) to anticipate some potential further decline in the US dollar value relative to European currencies during 2006.

A total of \$6.2 M represented the carryover from 2004 to 2005, which reflected the delay in dispersing some funds due to a number of factors common to new programs and the timing of some donor contributions that bridged years (e.g., DFID arrives in July and is to cover the March-April fiscal year). Resource allocation for research in 2005 and 2006 has deliberately exceeded income to help reduce and absorb this carryover. At the end of 2006 the total carry-forward is projected to be US\$ 2.1 M, with a net carryover of US\$ 0.8 M once 2006 remaining commitments (20% of the 2006 commissioned funds) have been removed. From 2007 on, the GCP will plan research activities according to our income status.

To clearly distinguish income versus expenditure and GCP assets, the presentation of the general financial tables in this report have been adjusted. The summary financial reports (income and expenses plus net assets) for 2005 and 2006 are shown in Tables 1 and 2. Details of 2005 and 2006 expenses are shown in Appendices D and E. Financial information presented for 2005 is based on actual year-end financial reports while the figure for 2006 is a projection based on anticipated income and expenditure. It is clear that by far the largest portion of our funds go to directly support the research and capacity building efforts of the GCP and its partners.

In 2005, an additional US\$ 0.5 M was moved into the GCP reserves, which now total of US\$1 M. This reserve is shown as “Contingency Reserve” in the net assets section of the summary financial tables (Tables 1 and 2).

Although interactions with potential new donors were limited this year, due in large part to the transitions the GCP underwent, timely reporting to and good communication with our current donors has been maintained. Special attention will be dedicated during 2006 to diversifying the GCP’s funding base to bring in additional funds for the purpose of further consolidating the research agenda and implementing GCP product delivery.

## **6.1 Financial Objectives and Outcomes** – Same as in MTP

The GCP financial approach on the 2005 MTP is as follows:

### **Income**

Our total actual 2005 income was \$14.193m, including other income (interest) of \$186k, which was \$342k more than our MTP projected revenues for 2005 of \$13.851.

Grant income amounted to \$14.0m and was contributed by our major donors in the following percentage terms: 43% from European Commission; 31.5% from United Kingdom (DFID); 17.8% from World Bank; 6.4% from Rockefeller Foundation; and, 1.4% from Sweden (SIDA).

### **Expenditure**

2005 actual expenditure amounted to \$15.560m which, when compared with the MTP expenses for 2005 of \$16.503m, reflects a difference of \$943k. This was the result of some reduction in planned activities due to the transition of the new GCP Director and some 2005 commitments which were transferred to 2006.

The excess of expenditure over income (\$1.367m) incurred during 2005 was covered by the 2004 carryover of \$6.2m that the GCP commenced with 2005, as shown in the Statements of Financial Position.

**6.2 Schedule of Contributions Received (by CP donor and amount)** – See Appendix B.

**6.3 Schedule of Disbursements to Partners (CGIAR and outside)** – See Appendix B.

**6.4 Resource Allocation/Expenditure (by subprogram)** – See Appendix C.

**6.5 Other Issues on Financial Management** – None

## **7. Lessons Learned**

Most of the lessons learned in the GCP thus far mirror the operational challenges elaborated above: we need to build a data management system that encourages researchers to share their data more readily; we need good communication, collaboration, and project management tools to make the innovative partnerships of the GCP work most efficiently; we need more staff to manage the GCP effectively as it grows in terms of projects, partners, and outputs; and our current governance structure is not ideal for the needs of the GCP. Reforms are underway to address all of these issues.

Another lesson we've learned is that this program needs clear priorities to better determine resource allocation across its set of 22 mandate crops and multiple research disciplines. Towards this end, a priority setting exercise was undertaken in early 2006, and the results will be presented to the PSC for their approval in the next few months. The refined priorities will ensure that the GCP is able to make impacts in both the short- and long-term.

We have also learned that despite the best planning efforts, there is much work to be done in enforcing the GCP Consortium Agreement. Data sharing issues, as mentioned above, were the first to arise, but now that products are being released, questions of intellectual property are becoming more and more frequent. It will be important to have staff or consultants who are familiar with the content of the GCP who can address these legal aspects.

## Appendix A. 2005 GCP projects and 2005 Subprogram Technical Reports

#	Type	List 2005	Subprogram	GCP Code	Project Title	Status
1	Competitive	2005	1	7	Measuring linkage disequilibrium across three genomic regions in rice	open
2	Competitive	2005	1	10	Exploring Natural Genetic Variation: Developing Genomic Resources and Introgression Lines for FourAA Genome Rice Relatives	ongoing
3	Competitive	2005	1	13	Development of informative DNA markers through association mapping in maize to improve drought tolerance in cereals	ongoing
4	Competitive	2005	1	14	Characterization of genetic diversity of maize populations: Documenting global maize migration from the center of origin	ongoing
5	Competitive	2005	1	17	Allele Mining Based on Non-Coding Regulatory SNPs in barley germplasm	ongoing
6	Commissioned	2005	1	2005-01	Completing genotyping	ongoing
7	Commissioned	2005	1	2005-02	Distribution of reference germplasm	ongoing
8	Commissioned	2005	1	2005-03	Molecular characterization of Tier 2 (Orphan) crops	ongoing
9	Commissioned	2005	1	2005-04	Assessing DArTs as a genome-wide scanning technology	completed
10	Commissioned	2005	1	2005-05	Assessing Ecotilling as a methodology for targeted genotyping and SNP discovery	ongoing
11	Commissioned	2005	1	2005-06	Supporting emergence of reference drought tolerance phenotyping centers	ongoing
12	Commissioned	2005	1	2005-07	Whole-plant modeling	ongoing
13	Commissioned	2005	1	2005-08	Population structure, phenotypic information and association studies in long-generation crops	ongoing
14	Competitive	2005	2	1	Identifying Genes Responsible for Failure of Grain Formation in Rice and Wheat under Drought	ongoing
15	Competitive	2005	2	2	Revitalizing Marginal Lands: Discovery of Genes for Tolerance of Saline and Phosphorus Deficient Soils to Enhance and Sustain Productivity	ongoing
16	Competitive	2005	2	8	Targeted discovery of superior disease QTL, alleles in the maize and rice genomes	ongoing
17	Competitive	2005	2	11	Functional genomics of cross-species resistance to fungal diseases in rice and wheat	ongoing
18	Competitive	2005	2	15	Determination of a common genetic basis for tissue growth rate under water-limited conditions across plant organs and genomes	ongoing
19	Competitive	2005	2	16	Isolation and Characterization of Aluminum Tolerance Genes in the Cereals: An Integrated Functional Genomic, Molecular Genetic and Physiological Analysis	ongoing
20	Commissioned	2005	2	2005-09	Systematic evaluation of rice mutant collections for conditional phenotypes with emphasis on stress tolerance	ongoing

21	Commissioned	2005	2	2005-10	Collection, distribution, phenotyping, and genotyping directed towards utilization of existing wheat genetic stocks to enhance tolerance/resistance of wheat cultivars to abiotic and biotic stresses with emphasis on drought	ongoing
22	Commissioned	2005	2	2005-11	Legume mutant resource development	ongoing
23	Commissioned	2005	2	2005-12	A saturated potato mutant population for functional genomics among Solanaceae and tuber crops	ongoing
24	Commissioned	2005	2	2005-13	Crop gene expression profiles and stress -gene arrays	ongoing
25	Commissioned	2005	2	2005-14	Stress-response enriched EST resources for target species	ongoing
26	Commissioned	2005	2	2005-15	Musa genome frame-map construction and connection with the rice sequence	ongoing
27	Commissioned	2005	2	2005-16	Validation of conserved orthologous markers	ongoing
28	Commissioned	2005	2	2005-17	Comparative QTL mapping for drought tolerance	ongoing
29	Commissioned	2005	2	2005-35	Sequencing Multiple and Diverse Rice Varieties	ongoing
30	Competitive	2005	3	3	Identifying the physiological and genetic traits that make cassava one of the most drought tolerant crops	ongoing
31	Competitive	2005	3	5	Unlocking the genetic diversity in peanut's wild relatives with genomic and genetic tools	ongoing
32	Competitive	2005	3	6	Marker Development and Marker-Assisted Selection for Striga Resistance in Cowpea	ongoing
33	Competitive	2005	3	9	Development of Low-Cost Technologies for Pyramiding Useful Genes From Wild Relatives of Cassava into Elite Progenitors	ongoing
34	Competitive	2005	3	12	Drought Tolerant Rice Cultivars for North China and South/Southeast Asia by Highly Efficient Pyramiding of QTL's from Diverse Origins	ongoing
35	Commissioned	2005	3	2005-18	Development of low-cost gene-based trait assay technologies in cereals	ongoing
36	Commissioned	2005	3	2005-19	Evaluation and deployment of transgenic drought tolerant varieties	ongoing
37	Commissioned	2005	3	2005-20	Optimizing marker-assisted breeding systems for drought tolerance in cereals through linkage of physiological and genetic models	ongoing
38	Commissioned	2005	3	2005-21	Planning for effective product development, delivery, and use	ongoing
39	Competitive	2005	4	4	An eco-physiological - statistical framework for the analysis of GxE and QTLxE as occurring in abiotic stress trials, with applications to the CIMMYT drought stress programs in tropical maize and bread wheat	ongoing
40	Commissioned	2005	4	2005-22	Development of GenerationCP domain (data) models	ongoing
41	Commissioned	2005	4	2005-23	Implementation of web services technology in GenerationCP Consortium	ongoing
42	Commissioned	2005	4	2005-24	Application and development of web services technology	ongoing
43	Commissioned	2005	4	2005-25	Creation and maintenance of data templates	ongoing
44	Commissioned	2005	4	2005-26	Management of GenerationCP Central Registry	ongoing
45	Commissioned	2005	4	2005-27	Integration of the High Performance Computing (HPC)-facilities in the GenerationCP toolbox	ongoing
46	Commissioned	2005	4	2005-28	Development of an integrated info. platform	ongoing

47	Commissioned	2005	4	2005-29	Creation of institutional bioinformatics capacity	completed
48	Commissioned	2005	4	2005-30	Development of decision support systems for sampling germplasm	completed
49	Commissioned	2005	4	2005-31	Development of ortholog-function display tools	ongoing
50	Commissioned	2005	4	2005-32	Development of crop gene expression database and data mining tools	ongoing
51	Commissioned	2005	4	2005-33	Development of an integrated decision support system for MAS and MAB	ongoing
52	Commissioned	2005	4	2005-34	GCP software engineering and collaboration platform	ongoing
53	Commissioned	2005	5	2005-CB01	Training materials for genetic diversity analysis of germplasm	ongoing
54	Commissioned	2005	5	2005-CB02	Development of training materials for a course in genomics and comparative genomics, and design of course curriculum	ongoing
55	Commissioned	2005	5	2005-CB03	Gathering/Development of training materials for a course in marker-assisted selection and breeding and design of course curriculum	ongoing
56	Commissioned	2005	5	2005-CB04	Development of training materials for a course in bioinformatics and design of course curriculum	ongoing
57	Commissioned	2005	5	2005-CB05a	Reference microsat kits --rice	ongoing
58	Commissioned	2005	5	2005-CB07	QTL design course	completed
59	Commissioned	2005	5	2005-CB09	Genotyping/data analysis workshop	completed
60	Commissioned	2005	5	2005-CB10	Policy/IPR seminar at CAAS conference	completed
61	Commissioned	2005	5	2005-CB11a	Diversity/Breeding courses	completed
62	Commissioned	2005	5	2005-CB12	Project Proposal workshops	completed
63	Commissioned	2005	5	2005-CB13	The Institute for genomic diversity's interactive resource center	ongoing
64	Commissioned	2005	5	2005-CB14	Regional PGR Courses	ongoing
65	Commissioned	2005	5	2005-CB15	Distance learning module for scientists on Genetic Resource Policies and their implications for Freedom -to-Operate.	ongoing
66	Commissioned	2005	5	2005-CB16	Intellectual Property and Access&Benefit Sharing-helpdesk and On-line Resource for the GCP Community, Partners, and Stakeholders.	ongoing
67	Commissioned	2005	5	2005-CB17	Reporting for Product Distribution: An asset inventory system for the GCP	ongoing
68	Commissioned	2005	5	2005-CB18	Expert Group--Delivery Strategy Workshop	completed
69	Commissioned	2005	5	2005-CB21	Fellowships and Travel Grants	completed
70	Commissioned	2005	5	2005-CB23	Genotyping support service	ongoing

## 2005 TECHNICAL REPORT BY SUBPROGRAM

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

Detailed objectives and outputs for 2005 in Subprogram 1 are as follows:

#### *1. Create an improved understanding of the structure of diversity for the world's major food crops*

SP1 has undertaken systematic work on the crops where the CGIAR has active breeding programs. The work started in 2004 for eleven crops and in 2005 for another seven. The main initial undertaking in SP1 has been the genotyping with SSR markers of representative germplasm samples, which laid the foundation for the whole GCP. The data serve for identifying reference samples which will be the materials of choice for integrating further molecular and phenotypic characterizations.

For each crop, a first step consisted in collating information on various existing collections in order to apply a simple rationale for extracting a representative sample (the “composite set”); this is coordinated by the International Agricultural Research Centres in charge of each crop and has been completed for all 18 crops. The size of the sets depends on the number of resources globally available in the collections, ranging from several hundred accessions to a maximum of 3,000 for the most important crops.

A second step consists in characterising the composite set with molecular markers in order to reveal the structure of the diversity and to extract a reduced sample (the “reference sample”) that will be made available for additional characterization and evaluation in order to reveal functional correlations. This is in progress for 18 crops (see Table 2 below).

Table 2. Molecular characterization of composite sets in GCP crops

Crop	Lead institution and partners	year start – end	Core sample genotyping target N acc. x N loci	Genotyping % nov-2005
Rice	IRRI-CAAS-CIAT-WARDA -Agropolis	2004-2005	3000 x 50	70
Maize	CIMMYT-CAAS-IITA -Agropolis	2004-2005	1775 x 50	70
Wheat	CIMMYT-CAAS-ICARDA-Agropolis	2004-2006	(2600 + 400) x 50	60
Sorghum	ICRISAT-Agropolis-CAAS	2004-2006	(700 + 2300) x 50	50
Barley	ICARDA-CAAS	2004-2006	(500 + 2500) x 50	50
Common bean	CIAT-Embrapa	2004-2005	3000 x 50	60
Cowpea	IITA	2004-2005	2000 x 50	30
Chickpea	ICRISAT-ICARDA	2004-2006	286 + 2714	60
Cassava	CIAT-Embrapa-IITA	2004-2005	3000 x 36	60
Potato	CIP	2004-2005	1000 x 50	80
Musa	IPGRI-IITA-Agropolis	2004-2005	960 x 50	30
Finger millet	ICRISAT	2005-2006	1000 x 20	20
Groundnut	ICRISAT-Embrapa	2005-2006	1000 x 20	50
Pigeon pea	ICRISAT	2005-2006	1000 x 20	20
Lentil	ICARDA	2005-2006	1000 x 30	20
Yam	IITA	2005-2006	350 x 20	20
Coconut	Agropolis	2005-2006	1000 x 22	80
Sweet potato	CIP	2005-2006	500 x 50	20
Pearl millet	ICRISAT	2006	to be determined	-

Given the variability of steps in the task, the heterogeneity of progress between crops, and the implications of including diverse partners, reporting on these activities has been a bit disorganised. It is planned to have a final report for all crops in the form of a collective publication with detailed standardised descriptions of polymorphism and its implications. This will be an important landmark and asset. Despite the shared difficulties to follow the planned calendar, very large amounts of data have been produced and will be sufficient for refining smaller germplasm samples with enhanced representativeness. These reference samples will be made available and distributed to collaborators worldwide.

Besides the above mainstream activity, molecular characterization of genetic diversity is being practised within several competitive projects. One example is that of a project led by CIMMYT called *Development of informative DNA markers through association mapping in maize to improve drought tolerance in cereals*, where both neutral markers (SSRs and SNPs) and candidate gene allelic sequences are being used on a set of tropical maize inbred lines. Another example is that of the project led by CIAT titled *Exploring natural genetic variation: developing genomic resources and introgression lines for four AA genome rice relatives*, which uses both SSRs and SNPs for monitoring interspecific introgression. The latter is an example of additional mobilisation of gene sources by broadening the genetic base of the pools accessible to recombination and breeding.

## **2. Assess molecular characterization methods**

Other technologies that compliment standard molecular characterization for elucidating germplasm structure (such as SSRs) will be very useful if proven efficient. Single Nucleotide Polymorphisms (SNPs) are the markers of choice for those crops where massive sequence data are available, such as ESTs from diverse germplasm. The quick evolution of SNP detection technologies and the present relative efficiency of resequencing vs setting up SNP detection suggests that it is wise that the GCP remain an observer of the technology for the time being. Contacts are being made to try and build alliance with high capacity partners, both for allele resequencing (SNP discovery) and for SNP detection (both as neutral markers or within candidate genes). Several competitive projects include work on SNP discovery and detection, both in rice, led by CIAT for monitoring interspecific introgression, and for maize, led by CIMMYT for developing association studies using candidate genes.

Along another line, work is ongoing in barley within a competitive project led by ICARDA called *Allele mining based on non-coding regulatory SNPs*. Reciprocal F1 hybrids are being produced to serve for testing the efficiency to identify cis-regulatory elements through allelic imbalance assays. A set of 70 stress responsive ESTs were identified and primers designed. SNPs were already identified and verified in 12 of them.

For the crops without much sequence information, SP1 is assessing EcoTILLing and Diversity Arrays Technology (DArT). EcoTILLing is used to identify SNPs in diverse germplasm pools without the need to sequence many genotypes. The pilot project is run by IRRI, which has advanced experience in rice, and adaptation work is being done at Agropolis for sorghum and banana (triploid). Primers were designed for a set of 16 tentative candidate genes given in the proposal, unlabelled primers have been obtained and the amplification efficiency was confirmed for 12 of them. Labeled primers for use on the LiCor have been obtained and verified for 6

genes. A modified procedure has been established that uses agarose gels and is therefore much more easily transferable. Ecotilling data have been generated for rice on 10 genes in over 330 accessions. Unfortunately the efficiency of the rice primers for amplification on sorghum (despite favorable homology) and *Musa* (no homology information) is poor, which will force the project partners to define species-specific primers.

DArT represents a potential platform for whole genome profiling in orphan crops. The work has several components, which have been completed at the end of 2005. A hundred accessions have been selected for each of sorghum, rice, and wheat, and existing DArT arrays are being hybridized. This is being analysed for testing the comparison of the diversity revealed with DArT to that revealed with SSRs; the data produced for sorghum and rice consist in matrices of 92 accessions and 516 polymorphisms out of 6144 clones spotted and 90 accessions x 519 polymorphic markers for rice also out of 6144 clones spotted, respectively. In cassava, the DArT technology has been proven efficient to reveal numerous polymorphic markers using cultivars and wild relatives (Xia et al, 2005). A scientist from Thailand has started a nine months stay at DArT Pty Ltd in Canberra, Australia, for preparing new libraries and extending the analysis to a large number of cultivars, focusing on accessions with large variation for dry matter content (DMC). Libraries are being expanded. The potential of the method is being explored on coconut and banana in collaboration with Agropolis, where a scientist from Sri Lanka (Coconut Research Institute) will spend six months conducting the testing. Several complexity reduction methods were tested and one promising method has been identified (PstI/BstNI). The work on banana started early September when a scientist from France joined DArT for a three-month stay. Two arrays bearing 9216 clones have been built for coconut and four arrays bearing 3072 clones for banana. Initial diversity surveys for coconut (31 acc x 128 polymorphisms) and for banana (48 x 332) suggest high concordance with other types of markers.

### ***3. Establish and implement a scientific and organizational framework to describe tolerance to drought***

The backbone of this activity consists of optimising access to efficient coordinated multilocal phenotyping platforms, supporting the evaluation by environment descriptions and whole plant modelling activities, and drawing experience from advanced physiological characterization performed in crop-specific projects. A *Drought Tolerance Phenotyping Network*, coordinated by Embrapa, is being put in place and reinforced for integrating materials of international origin for evaluation. The initial phase of this project was used for purchasing or upgrading of various requisite equipment. The two groups involved in the Drought Phenotyping Network (coordinated by Embrapa) and in Whole Plant Modelling (coordinated by Cirad) met in May 2005 at Embrapa/Rice and Bean in Goiania, Brazil, in order to set the planned collaborations and exchanged further at the Annual Research Meeting in September. Priority crops considered by the project are upland rice, maize, and sorghum. The regions considered by the project will be: (i) in Brazil, Piaui, Tocantins, Goias, Minas Gerais, and possibly Mato Grosso ; (ii) in West Africa, sub-saharan region for sorghum and possibly upland rice crops ; (iii) Central America for maize crop (to be confirmed). A target population of environments analysis will be carried out in all these regions and crops, provided that the required data sets are available. A scientist from Embrapa is in charge and will collect necessary meteorological and soil data with support of all institutions participating, according to site and country they are stored. He will work in Cirad, Montpellier, from September 2005 to August 2006. He will spend two months at CSIRO in

Australia in 2006 in order to be trained on Apsim model and clustering tools. Another scientist from Embrapa will be trained at Cirad on SarraH model in November 2005 and will build up specific model for maize. Another scientist from Embrapa Meio Norte will also be trained in late 2005 or 2006, for specific adaptation of the model to cowpea and legume. At the end of the project, a general training course for Embrapa scientists will be organised by CIRAD, CSIRO and Embrapa.

For fine phenotyping to be carried out, five accessions from the GCP reference samples will be selected. Different cases are identified:

- upland rice: complementary experiments conducted in Porangatu, Terezina and Goiania during rainy and dry seasons
- maize : possibly (to be confirmed) complementary experiments in Terezina, Sete Lagoas, and Janauba
- sorghum: complementary experiments conducted in Janauba, Sete Lagoas, Goiania and/or Porangatu during rainy and dry seasons. Existing data for West Africa will be exploited.

Numerous trials are going on and materials of the GCP reference samples can be integrated as they are identified and made available.

Meanwhile, a pioneering competitive project led by CIMMYT has started a large scale evaluation of over 400 hybrids of maize in relation to drought stress in diverse environments. In addition to the practical experience of widespread exchange and characterization, the group performs assays of glucose and sucrose, ABA and ABA-glucose ester as well as phaseic acid and dehydrin on leaf samples and silk samples.

#### ***4. Develop methodologies for relating genotype to phenotype***

Relating genotype at molecular markers (i.e., a molecular polymorphism) to phenotype for important traits can rest on direct functional involvement of the molecular polymorphism in the trait or on indirect association through linkage disequilibrium (LD) between the marker and the causal functional polymorphism. Therefore the extent of LD in a crop is essential for interpreting the meaning of statistical associations and for determining which populations can be used to map/reveal favorable genes and alleles through association studies. It is being monitored at the whole species level in two contrasting species, namely the annual autogamous sorghum and the perennial allogamous coconut, as well as in specific populations of rice in Indonesia. The study of 12 RFLP marker loci in a single region of about 5 cM in sorghum has revealed significant LD spanning 2 to 3 cM. It also exemplified how important it is to refine the reference sample before undertaking association studies. In coconut, a method has been refined to assess the level of LD in this highly heterozygous species. Two closely linked SSR markers (ca. 1 cM) display strong LD in populations (namely Mozambique, Panama, Vanuatu and Brazil) where unlinked markers displayed none; data are being produced for another eight couples of SSR loci linked at 0 to 7 cM. In rice, the study focusses on three regions bearing disease resistance genes. A researcher from Indonesia came to Cornell for a training of 4 months (April-July 2005). A total of 250 PCR primers have been designed, 60 in the *Xa7* region on chromosome 6, 60 around *Xa13* on chromosome 8, and 130 around the cluster *Xa4/Xa22/Xa26* on chromosome 11. Sequences from the 8 diverse Indonesian rice accessions have been generated for 215 primer pairs. Seventy five

of the sequences from the 8 rice accessions have been aligned and SNPs have been called; 41 SNP detection primers have been designed.

Simultaneously, as an attempt to integrate association studies in the course of ongoing characterization and breeding activities, similar exercises of LD assessment are ongoing for additional cases where accurate phenotypic data are available for materials amenable to LD-based mapping. These are the cassava breeding program at CIAT; the potato breeding program at CIP; the banana breeding program at CARBAP, Cameroon; the coconut breeding program at VARTC, Vanuatu; and the yam (*Dioscorea alata*) germplasm evaluation program at VARTC, Vanuatu. In the former three species, the possibilities for choosing the best materials have been analysed. Various options still exist; the assessment of the level of LD will be of primary importance. The case is more clearly defined in coconut, with materials consisting of 200 trees representing four generations of breeding materials. A total of 219 trees have been sampled in VARTC and DNA has been extracted. They are being analysed by a scientist from Sri Lanka during a training session in Montpellier in September-November 2005. A total of 31 loci will be surveyed, including 13 international reference markers and 9 couples of linked markers. In addition, a coconut breeder from Vanuatu will visit CIRAD in November for a three week training session on DNA extraction, principles and applications of molecular breeding and data management. For yam, the idea is that the insular history may have involved bottlenecks that have established strong LD in the well characterised populations of Vanuatu; a mapping exercise is undertaken to identify linked markers among the already used AFLPs, for quick assessment of LD.

## **Subprogram2: Comparative Genomics for Gene Discovery**

Detailed objectives and outputs for 2005 in Subprogram 2 are as follows:

### ***1. Assembly of genomics and germplasm resources through consolidating existing (and developing new) specialised genetic stocks and framework genetic markers for target crops***

Projects supported under this Project focus on adding value to existing genetic and genomic resources and creating new ones when such investment would open new approaches and leverage collaboration. A project was initiated to assemble special wheat genetic stocks relevant for gene discovery and make them available for systematic phenotyping. A Wheat Genetic Stocks Utilisation Workshop was held in 5-7 April 2005 at CIMMYT. A preliminary list of wheat genetic stock candidates (with over 3,300 lines identified), available for distribution and characterization by the workshop participants, was prepared. It was agreed that stocks currently available (30-50g of seed) will be sent to CIMMYT immediately for further multiplication and phenotypic characterization

A network of seven laboratories around the world (Wageningen University, CIRAD, NIAS, IRRI, CAAS, Huazhong Agricultural University, and CIAT) was formed to characterize rice mutants with insertions in candidate stress response genes. This network of laboratories collectively has produced the largest collection of rice mutants in the world providing flanking sequence tag (FST) databases for query of mutations in target genes. Currently, the OryGenes DB (<http://orygenesdb.cirad.fr>) developed at CIRAD has about 56,000 publicly available insertion sequences tagged by FST sequence on the rice genome sequence as annotated by TIGR.

As of mid-2005, this DB was used to search for knockout inserts in candidate stress-associated genes and provided about 22% insertion coverage. These insertion lines have been categorised for phenotyping at CIAT.

Because of the growing evidence of potential roles of non-coding small RNA in regulation of many plant processes (including stress response), the group began a screen for genes controlling small RNA metabolism. Rice orthologs of seven *Arabidopsis* genes involved in the biogenesis of microRNA (miRNA) and short interfering RNA (siRNA) were identified. Insertion lines in these candidate genes were identified in Hirochika's Tos17 mutant collection in Japan as well as in Gyn An's T-DNA insertion collection in Korea. These lines have been assembled for phenotyping under different stress conditions. Using a population of EMS-induced IR64 mutants (about 2,200 lines), DNA pools (8 lines/pool) were produced to screen for mutations by TILLING. So far, from screening four small RNA related genes in 800 plants, a putative variant in *SGS3* (*Suppressor of Gene Silencing*) has been identified.

Field-screening of random T-DNA insertion lines under drought stress conditions is underway at Huazhong Agricultural University. Drought stress was applied to rice plants at the vegetative stage (30-45 days after germination) in a sandy field. Within half a year, a total of 4,006 rice mutant families (20 plants each) were screened for drought tolerance at the vegetative stage in the field. Among them, 29 families showed segregation of drought sensitivity and one family showed segregation of drought tolerance. PCR analysis using the hygromycin resistance gene suggested that the mutant plants were positive for T-DNA insertion for all the 30 families. So far, flanking sequences of insertions were available for 10 mutant families. These putative mutants are now subject to a second round of evaluation of drought tolerance at both the vegetative and reproductive stages under well-controlled conditions.

Because plant growth hormones abscisic acid (ABA) and gibberellic acid (GA) play a major role in regulating gene expression during plant development and response to environmental stresses, IRRI has systematically identified mutants showing altered peduncle elongation, dormancy, and sensitivity to ABA during germination. So far 10 mutants with Elongated Uppermost Internode (*eui*) have been identified; the peduncles of these mutants elongate at the rate of 8-12 cm per day, compared to 4-5 cm per day in the wild type. The screen for the dormancy mutants yielded 11 mutants expressing a high level of dormancy (<50% seed germination compared to wild type). From a screen of 1,500 mutant lines for ABA insensitivity, 25 mutants were found to exhibit significant difference in germination, radicle, and shoot growth relative to the wild type. These mutants are now available for more detailed physiological analysis of drought-response traits in other GCP projects

In addition to rice, bean and a true-seed *Solanum* species were selected to produce mutant populations as a permanent resource for gene identification and validation. Multiple accessions of *S. verrucosum* available at CIP and the Scottish Crop Research Institute were examined for purity and homozygosity using molecular markers. The least heterozygous accessions were identified for seed increase and mutagenesis at CIP.

Bean genotype BAT93 was selected for mutagenesis. Pure seeds were produced at CIAT and provided to collaborators at Univ. of Geneva to conduct EMS mutagenesis. Mutagenesis protocol

has been worked out at Univ. of Geneva and seed of M1 families were sent to CIAT. The first set of 800 M1:2 plants and the second set of 1,000 M1:2 plants have or are being screened for phenotypic differences compared to the parental line BAT93. Phenotypic mutants (dwarfing, leaf fasciation, leaf variegation, spindly growth, etc.) have been documented and photographed. A database is being constructed with these photographs and characterization data.

We expanded available gene/sequence information for stress-response pathways across selected species. Construction of EST libraries for millet (IRISAT) and cowpea (IITA) was initiated to expand the genomic resources available for sequence comparison and expression analysis. Four cowpea lines, Vu7778 (drought susceptible), Tvu11986 (type I drought tolerance), Dan Ila (type II drought tolerance: stay green) and 12008D (fodder type), were subjected to drought stress and multiple tissue samples (root, stem, and leaf) were harvested for RNA extraction. Collaboration was extended to ILRI to expand the size of the EST dataset. For pearl millet, RNA was isolated from panicles and flag leaves of stressed and non-stressed plants of the two millet hybrids at nearly identical developmental stages. Stressed plants of the two genotypes had comparable levels of transpiration relative to their non-stressed counterparts. Differential EST library development is now underway.

The GCP supports a consortium effort to generate SNP data of multiple rice varieties. The project involves collaboration with Perlegen Sciences, a private company pioneering the use of chip-based technology to “re-sequence” whole genomes of human and other mammalian species. The first stage of this project targets 100 Mb of the rice genome across 15 diverse genotypes. Funding is being sought through partnerships under the International Rice Functional Genomics Consortium to complete the genome coverage and to expand the set of rice varieties to 20-30. As an example, Dr. Jan Leach at Colorado State University, together with TIGR and IRRI, has recently secured a USDA competitive grant to contribute to this effort, demonstrating the power of resource leveraging. All data from this project will be in the public domain. Similar to the HapMap project (The International HapMap Consortium. 2005. *Nature* 437: 1299-1320), the SNP database will enable the application of association genetics in rice as well as other plant species.

## ***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***

Several groups (CIP, Cornell, CAAS, ICARDA, INIBAP, and the Musa Genomics Consortium) worked together on the development of conserved orthologous markers (COS) to facilitate comparative mapping and assignment of function to conserved genes across species. At CIP, 100 COS II (second generation COS markers) primers for candidate genes for disease resistance and drought tolerance were designed with aid from Cornell’s SGN database. Primers were screened for polymorphism in subsets of 3 mapping populations of which 24 were polymorphic in at least one cross. DNA of bean (from CIAT) and sweet potato (CIP) were included in the screening and 30% gave clear PCR products. Genotypes of four wild and one cultivated potato species of known resistance phenotype were assembled toward application of COS to understand diversity in germplasm. Mapping and sequencing of products are in progress to validate the identity of the amplicons.

COS identification strategies were reviewed with support from SP4, using the custom BLAST from Paracel and genome annotation tool on the HPC. Further links with SP4 are established via collaboration on the common functional gene catalog. Available annotation tools for metabolic pathways and integration with sequence information were identified (BioCyc software) and are scheduled for evaluation and in second half of 2005. Documentation on COS annotation strategies was initiated for publication on a website. In collaboration with Agriculture & Agri-Food Canada, SNPs have been identified in potato DNA sequences corresponding to 31 tomato-Arabidopsis COS based on comparisons with publicly available sequences. Six have detected polymorphism among potato genotypes.

At ICARDA, 50 abiotic stress-(drought, cold, and ABA) induced gene sequences from microarray experiments of Arabidopsis were used to identify putative orthologues and develop COS markers for legumes. Sequences were aligned with EST sequences of *Medicago*, soybean, and *Lotus*. Primer design at conserved regions is in progress. Amplification products of 6 COS markers with six legumes of ICARDA' interest (faba bean, lentil, chickpea, and grasspea) are being investigated further by sequencing.

CAAS focused on the development of disease resistance EST-SSR markers across monocots, including wheat, rice, maize, and barley. From the public databases of NCBI and TIGR, a total of 48 SSR-containing ESTs were identified and 61 EST-SSR primers were designed based on wheat ESTs, of which 21 ESTs contain LRR or NBS-LRR domain, 21 contain PK or LRR-PK domain, the other 6 ESTs were related to disease resistance. So far, 24 primers produced expected PCR products, of which 12 showed polymorphism in wheat varieties and have been used for genetic diversity analysis, and 6 primers used for genetic mapping in three mapping populations. Of 15 primers tested, 12 produced strong bands across the four monocots, and the other 3 primers did not get amplicons in maize.

Results from COS marker work on *Musa* coordinated by INIBAP suggested relatively few COS primers from the existing Cornell COS database could amplify orthologous products for *Musa*. Refinement of primer design, ideally based on EST comparisons, is required to obtain more useable primers for targeted genes. Using both ESTs from *Musa* and/or genes from other species could be generally effective in finding the genes conserved in banana.

The experience gained from the development of COS markers suggests that sequencing of orthologous genes is important to understanding the variation present between different species. While conservation of genes between species can be detected by bioinformatic approaches at the DNA or protein level, the identification of conserved regions for primer design and the amplification of the targeted regions have proven to be more difficult than expected. It is therefore important to share experience between groups to improve primer design, and to understand realistically to what extent conserved gene primers can be developed across all species.

A project on targeted sequencing of the *Musa* genome and frame map construction was launched to expand on the genomic resources for *Musa*. The project is led by NIAS and INIBAP and involves several labs from the *Musa* research community. *Musa* sequence and map data produced by this project will enable detailed comparative analysis with rice to span the range of

monocots. Within half a year, the group has gathered around 32,000 EST clone sets from Syngenta and deposited them at the *Musa* Genome Resource Center (MGRC). A *Musa acuminata* cv. Tuu Gia BIBAC library was deposited at the MGRC. Three mapping populations containing at least 150 individuals were assembled. Diversity analysis was conducted using the six parents generating the three mapping populations. About 600 repetitive DNA sequences were isolated from *Musa acuminata* cv. Calcutta 4.

### ***3. Assign genes and pathways to putative phenotypes through the convergence evidence of genome variation, expression patterns, and phenotypic data***

Activities under this Project consider target traits of common concern across more than a single species. The studies involve comparative analysis of phenotypes and genetic mapping, and often adopt a genome-wide approach to identify a pool of candidate genes for testing additional hypotheses and to validate gene function.

The project on *Determination of a common genetic basis for tissue growth rate under water-limited conditions across plant organs and genomes* (led by CIMMYT) emphasises precise phenotyping of tissue response across maize, rice, and wheat under drought stress. The phenotypic data will be co-analysed with expression data to identify common (or distinct) genes responsible for tissue growth under stress conditions.

In the first year, most of the work revolved around detailed characterization of tissue/organ response in three crops. In maize, refined phenotyping was undertaken by the collaborative work between INRA-Agropolis and CIMMYT. Phenotyping methods were developed for analysing the responses of both leaves' and silks' growth to water deficit. The elongation rate of the sixth leaf of 120 RILs and environmental conditions were recorded in a greenhouse and a growth chamber over a period of drying. QTLs were detected for maximum growth under non stress conditions, as was response to soil water potential. These QTL colocalised with previous QTL for ASI under stress in the field, suggesting a common genetic mechanism for growth of silks and leaves (INRA).

CIMMYT's effort concentrates on characterization of genetic variability for leaf elongation, leaf development under water stress, yields, and related traits in maize in field conditions. A population of RILs (P1xP2 - 220 genotypes) was planted December 2004 and evaluated for a variety of characters under well-watered and stressed conditions. A parallel trial with two parental lines and six contrasting genotypes (three good and three bad) from the same segregating population was also evaluated to provide material for gene expression profiling.

For wheat, 100 lines were evaluated in the field under stress conditions and trait data collected. A subset was selected based on 2004 commissioned research and data for leaf and stem extension rate under stress and irrigated conditions collected. A subset of six pot-grown contrasting lines are currently being evaluated for leaf extension rate at known soil water potentials in controlled conditions. For rice, leaf emergence and elongation rate is being evaluated using 150 BC lines (Vandana x Moroberekan), under well-watered conditions in the field, and NILs (IR64 x Azucena) under well-watered and stressed conditions in the greenhouse (rainy season). These two populations will also be evaluated in the field during the dry season (January 2006).

Two projects involved comparative analysis disease resistance. The CerealImmunity project was initiated to identify genes responsible for cross-species, non-host resistance. The team conducted artificial inoculations under controlled greenhouse conditions to identify non-host and host specific isolates of blast fungus of rice and wheat. Fourteen wheat cultivars, one rice cultivar and one barley cultivar were tested against 14 isolates of blast fungus collected from wheat, two from grasses (*Digitaria horizontalis* and *Eleusine indica*), and one each from barley and rice. All test isolates were virulent to wheat cultivars, including the rice, barley, and grass isolates, but showed differences in aggressiveness. The grass isolate from *D. horizontalis* was least aggressive on wheat cultivars. None of the isolates were virulent to the rice cultivar Bonança except the rice isolate. The most aggressive isolate (Py 5996) on wheat cultivars was selected for further studies on cytological characterization of non-host interactions. It is proposed to test eight rice cultivars including IR64 with *P. recondita* races. The screening of disease susceptible rice mutants could provide useful materials for identifying non-host resistance. Linkage between the CerealImmunity project and the rice mutant network were discussed during the latest GCP Annual Research Meeting.

Working on disease resistance in maize, researchers at Cornell and North Carolina State University developed heterogeneous inbred families (HIF) as genetic materials to serve both purposes of gene discovery and practical breeding. F5 progeny from several maize crosses were tested by molecular markers to identify pairs of near-isogenic lines (with respect to specific candidate genes or chromosomal regions). The families will be analysed for Northern Corn Leaf Blight resistance and with SSRs to identify chromosomal regions conferring disease resistance. This approach is considered a means to achieve the dual purposes of gene identification and development of diverse breeding lines in local breeding programs. The group also applied genome-wide expression analysis of rice advanced breeding lines with broad-spectrum resistance to reveal novel genes that mapped to chromosomal regions with significant disease QTL effects. Maize orthologs of these rice genes were identified from a maize sequence database to design markers for mapping in maize.

A series of gene expression experiments were conducted for selected phenotypes and crop genotypes to identify common and unique genes correlated with phenotypic expression across species. The newly released Affymetrix Barley array chips (22,840 elements/chip) were used to explore barley drought tolerance genes between drought tolerance cultivars ‘Tadmor’ and sensitive cultivar ‘WI2291.’ The preliminary result showed that gene categories, including mainly the kinases, transcription, signal transduction, photosynthesis, heat shock, pathogen resistance-related, and transporters were putatively involved in the drought tolerance during the heading stage. The approach will be extended to additional barley germplasm and also to wheat using the recently available wheat Affymetrix chip. Comparative analysis will enable the identification of common or distinct genes correlated with similar or different drought tolerance mechanisms in different pedigrees.

NIAS, IRRI, and the Beijing Genomics Institute (BGI) collaborated to evaluate the utility of rice gene chips for analysing drought response in cereals. NIAS and IRRI first tested the Agilent rice gene chips for cross hybridization with wheat and maize DNA. Moderate cross hybridization was observed with wheat but not with maize. Gene expression during vegetative drought stress was further investigated in rice and wheat using the 60K oligo chips from BGI. Rice and wheat

varieties of different levels of drought tolerance were used. Drought stress was imposed during the vegetative stage by withholding water until 20% field capacity or until the development of severe leaf wilting symptom. RNA were isolated from rice and wheat genotypes under different water stress regimes and hybridized to the BGI chip. Approximately 60% of features with significant hybridization to rice also hybridized to wheat. By contrasting rice and wheat genotypes with different levels of tolerance to vegetative drought stress, we identified approximately 50 genes found in common rice and wheat that are correlated with drought tolerance phenotypes. Of these, 12 genes showed a common expression patterns in both rice and wheat. A similar proportion of genes on the 22K Agilent chips showed significant hybridization with wheat cRNA. The correspondence between the genes identified in Agilent and BGI chips is being examined. These experiments suggest that valuable information could be gained from using a common chip platform for wheat and rice.

#### ***4. Validate genes and pathways through evaluation of under- or over-expression constructs or variants (induced or natural) of the target genes***

Activities under this Project have research components that are close to identifying specific candidate genes that can be subject to validation. Clearly, the distinction between the genome-wide discovery phase and gene function validation can be somewhat arbitrary. As more results are generated in other SP2 projects, more candidate genes will be available for functional validation.

The project on “Identifying genes responsible for failure of grain formation in rice and wheat under drought” aims at identifying opportunities to enhance reproductive-stage drought tolerance in rice and wheat through physiological, genetic, and molecular analyses of two yield determinants that are highly sensitive to field-level stress—panicle exertion and floret fertility. The team applied proteomics and microarray analysis to characterise the signal transduction pathways by which GA and ABA exert their effects on peduncle elongation. Chips from Agilent (with ~22K genes) and the Beijing Genomics Institute (with ~60K genes) have been hybridized with rice and wheat RNA. Hybridizations of the 22K array with three biological replications of peduncle RNA from well-watered, drought-stressed, and re-watered IR64 plants and well-watered plants of IR64 mutant (*eui10*) showing rapid peduncle elongation have been completed. Analysis of the data is underway. Detailed RT-PCR examination of the ABA-GA signal transduction pathways in peduncles has also been initiated, through examination of gene families of the ABRE-binding transcription factors and their post-synthetic modulators (the PKABA1-like protein kinases and protein phosphatases 2C). Of particular interest is the behavior of these genes at the base of the peduncle where cell division and elongation occur.

In the investigation of floret fertility, the team focuses on carbohydrate allocation to develop floral organs and cell- and tissue-types in wheat and rice. As in the peduncle, drought stress down-regulates all cell-wall invertases of the anthers and therefore disrupts the flow of carbon to anthers and pollen grains. RNA *in situ* hybridization is used to identify which cell-wall invertases and hexose transporters operate in each of the tissues of rice and wheat showing drought-sensitive development. A segregating population of ~900 lines has been developed to the F4 stage for the cross IR64 x Moroberekan. These two parents differ significantly in their tolerance to drought at heading. In particular, under drought stress in the IRRI Phytotron, Moroberekan shows greater floret fertility in the top four rachis branches than IR64. This is

correlated with clumping of the pollen in IR64. Pollen clumping can reduce the number of pollen grains released onto stigmatic surfaces for fertilisation or the ability of pollen to germinate on the stigmas. The segregating population is being used to test the hypothesis that the accumulation and breakdown of a particular anther glycoprotein governs the self-adhesion and stress tolerance of the floret.

Two projects are targeting traits controlled by QTL with large effects. The first project (led by IRRI) focused on the cloning of major QTL controlling tolerance to phosphorus-deficiency and salinity in rice. Through recombination analysis, the chromosomal region containing the *Pup1* gene (phosphorus uptake) was finely mapped, leading to the identification of two putative candidates for further analysis. More recently, comparison of genome sequences indicated genomic rearrangements between *indica* and *japonica* genotypes in *Pup1* region. The BAC clones in *Pup1* region have now been identified in Kasalath, the original donor of the *Pup1* allele. Additional candidate genes in the Kasalath BAC clones are being examined. For salinity tolerance, a number of QTL derived from a salt-tolerant variety Pokkali has been finely mapped, including the QTL *Salto1* on chromosome I. Using information available in the databases, all the genes located in the *Salto1* region and the borders regions limiting *Salto1* in chromosome I have been identified and classified according to their putative functions. The group plans to apply genome-wide expression analysis to identify differentially expressed genes in contrasting advanced backcross lines or near-isogenic lines containing different QTL for salinity tolerance.

The second project (led by Cornell) aims at cloning aluminum toxicity tolerance in sorghum and subsequently in related cereals. High-resolution mapping led to the identification of two markers that flanked the aluminum tolerance (*Alt<sub>SB</sub>*) and defined a 27 kb interval that spanned only three candidate ORFs. One of these ORFs encodes for a transporter-like protein that is implicated in organic acid efflux, and thus is a strong candidate for *Alt<sub>SB</sub>*. Furthermore, elite Al tolerant sorghum hybrids have been developed from the Embrapa breeding program.

Finally, we expect to make extensive use of available mutant populations as a community resource for gene function validation. The rice mutant network led by WUR has compiled a stress-associated gene (SAG) database based on several approaches, including expression induced by various abiotic stresses and overexpression of stress resistance genes. A panel of candidate genes has been established for searching knockout mutants in rice and other mutant collections. This is comprised of the following:

- 400 abiotic stress associated genes identified from publications
- 116 disease stress associated genes obtained from publications
- 260 (>1.5 fold) and 104 (>2 fold) BTH induced disease stress associated genes
- 48 putative rice orthologs of Arabidopsis genes revealed in a drought stress regulon
- 16 putative rice orthologs of Arabidopsis stress related RNAi mechanism genes.

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

Detailed objectives and outputs for 2005 in Subprogram 3 are as follows:

#### ***1. Develop drought tolerant rice Cultivars for North China and South/Southeast Asia by highly efficient pyramiding of QTL's from diverse origins***

This project started out with progeny testing of selected drought tolerant introgression lines (ILs) under stress and non-stress conditions and development of intercross populations for pyramiding QTL. The parental genotypes of these ILs have been screened with around 600 SSR markers and 100-200 of which are now being screened across the entire populations, while detailed phenotyping will be carried out during the coming season. Around 30 drought tolerant lines have already been identified that pyramid 30 or more drought tolerant QTLs. These will now be used for a large-scale backcross programme with a high yielding restorer line.

## **2. Identify the physiological and genetic traits that make cassava one of the most drought tolerant crops**

This project aims to find the best biological traits for improving drought tolerance by MAS in cassava whilst at the same time contributing to the improved understanding of the genetics and physiology of drought tolerance in cassava. Thus, this project aims to identify trait-marker associations for the development of a more cost-effective breeding process for drought tolerance in cassava. In particular, the project will assess the effect of the leaf retention gene for improving drought tolerance. Finally, the project will establish a strong network of institutions involved in the molecular breeding of drought tolerance in cassava.

Embrapa (Brazil) and CIAT (Colombia) breeding programmes identified 40 varieties with contrasting drought response phenotypes (broadly, 28 tolerant and 12 susceptible). These varieties were selected on the basis of 15 years of evaluation data and originated from Brazil, Colombia, Thailand and Venezuela. These varieties have been multiplied up through *in vitro* micropropagation at CIAT for detailed evaluation at Embrapa and University of Cornell.

## **3. Unlock the genetic diversity in peanut's wild relatives with genomic and genetic tools**

Based on prior research, this group has published the first diploid A genome map of *Arachis* (Moretzsohn et al 2005). However, through this project it has been possible to augment this map and consolidate some linkage groups in order to reach the desired number of linkage groups. Using this map it has also been possible to generate the first preliminary comparative map between *Arachis* and the model system *Lotus japonicus*. Meanwhile, efforts are ongoing to finalise a B genome map and an amphidiploid AABB genome map. Finally several repetitive element fragments from *Arachis* have been isolated and used to construct a pseudo-contig for fluorescent *in-situ* hybridization. These maps and genomic tools will be especially valuable for marker-assisted introgression breeding using synthetic amphidiploids that facilitates the introgression of vast sources of genetic diversity from diploid species.

Root-knot nematodes are known to increase water stress in infected field conditions and, therefore, represent an important component of a holistic approach to developing drought tolerance in groundnut. Four thousand single read ESTs were produced from roots inoculated with root-knot nematodes (*Meloidogyne arenaria* race1) and from non-inoculated roots. Assembly of the ESTs produced 963 contigs and 2537 singlets. Homologues to transcripts involved in responses to biotic and abiotic stresses have been identified and candidate gene-based marker development is underway.

Based on prior research, a range of synthetic amphidiploids have been created with putative resistance to diseases. This material will be used for proof-of-concept while new synthetics are

being generated and/or identified with putative drought tolerance. Thus, the project is currently screening the drought tolerance of a range of diploid species in Brazil and India.

#### **4. Conduct marker development and marker-assisted selection (for *Striga* resistance) in cowpea**

Cowpea is a critically important source of protein-rich food and feed in the drier parts of Africa. Amongst the major legume crops, cowpea has the highest levels drought tolerant germplasm yet national cowpea yields remain very low. This is partly due to the problem of retaining high levels of drought tolerance in breeding populations. But this is also due to the fact that where drought is a problem, so is *Striga*, insect pests and other abiotic stresses. This project focuses on the development of efficient MAS systems for *Striga* resistance, which also offers a simpler trait for proof-of-concept.

The ability of cowpea genotypes to resist *Striga* parasitism depends on the geographic origin of the parasite. Based on the differential resistance reaction exhibited by various cowpea genotypes, six different races have been identified. Initial efforts have focused on the development of allele-specific (co-dominant) molecular markers for two sources of race-specific resistance. Two SCAR markers have been identified that are suitable for high-throughput screening. While mapping of a third source of resistance to *Striga* has been initiated using AFLP and SSR markers and mapping of resistance to a second parasitic weed (*Alectra*) has also been initiated. Studies are also underway to determine whether the SCAR and AFLP markers retain their selective power when applied in a different population. Most recently around 1000 new cowpea genomic sequences were generated from which additional SSR markers will be generated. This project is also initiating some drought tolerance mapping studies.

This project has also identified a series of AFLP markers that discriminates among the different races of *Striga* that are parasitic on cowpea. These AFLP markers are now being converted into SCAR markers to facilitate rapid identification of pathogen diversity in the field.

#### **5. Develop low-cost technologies for pyramiding useful genes from wild relatives of cassava into elite progenitors**

Wild *Manihot* germplasm are a wealth of useful genes for the cultivated species *M. esculenta* but their use in regular breeding programmes is restricted due to the long reproductive breeding cycle of the crop and the deleterious linkage drag associated with the use of wild relatives in crop improvement. This project seeks to identify useful genes for pest and disease resistance, and post-harvest deterioration in cassava and to develop low cost marker tools for their rapid introgression into cassava.

Previous research had identified a RAPD and SSR markers linked to resistance to the cassava mosaic disease (CMD). During the first year, this project has successfully converted the RAPD marker into a SCAR marker for high throughput low cost MAS applications. The project has now shifted attention to the development of markers for resistance to pests and diseases with phenotyping evaluation of mapping populations at NARS in Brazil, Uganda, Ghana, and Nigeria.

#### **6. Develop low tech gene-based trait assay technologies in rice and wheat**

For marker-assisted improvement of rice for bacterial blight (BB) resistance, NARES collaborators from China, the Philippines, Indonesia, India and Africa were surveyed for the germplasm materials serving as recipients in their national programme for improving resistance to BB. In each case, target disease resistant loci were sequenced in all germplasm.

For the parallel situation in maize for quality protein maize (QPM), eight QPM donor sources from diverse agroecological zones (mainly tropical lowland and subtropical) and nine non-QPM recipient sources from tropical highland, tropical lowland and sub-tropical regions have been selected for sequence comparison of the alleles at the *opaque2* locus.

This project has also initiated the development of an allele-specific dot-blot (gel-free) assay, using the rice bacterial blight system as a pilot test. Unfortunately initial tests were confounded by the high degree of sequence similarity between resistant and susceptible alleles. However, this will be resolved when sequence data from all germplasm is available. A similar process has now also been initiated for the QPM system.

### ***7. Evaluate and deploy transgenic drought tolerant varieties***

This project has brought together physiologists from across the CG system to refine a standard methodology for agronomic evaluation of DREB transgenics for drought tolerance. This includes, rice (CIAT), wheat (CIMMYT), groundnut (ICRISAT), potato (Univ Tsukuba/JIRCAS). Next steps in the project will include multilocational field evaluation and assessment of background genotype effect (through multiple genotype transformations or backcross programmes).

### ***8. Simulate marker-assisted selection strategies for optimising molecular breeding systems for drought tolerance in cereals***

This project has been designed to present examples of many of the available options in utilising simulation to improve MAS. It is not a comprehensive study of any single approach, but general guidelines are also being developed. In particular, the aim is to develop some simulation examples that could be extended to other crops and different genetic models. The initial work is based on a study of an existing case study (wheat breeding at CSIRO) to combine known genes (using 'perfect' or near-perfect markers) into single genotypes for use as parents or further field screening. A population genetics model has been developed focused on the efficient use of marker-based selection in plant breeding and a case study in wheat has been completed and a publication has been drafted. In addition, the QUCim breeding module is being modified to realise the linkage with physiological models. Future activities, will include building case studies for rice, and investigating options for maize.

### ***9. Create product development plans***

The competitive grant projects that are involved in this study are: *Unlocking the genetic diversity in peanut's wild relatives with genomic and genetic tools/Embrapa*; *Drought Tolerant Rice Cultivars for North China and South/Southeast Asia by Highly Efficient Pyramiding of QTL's from Diverse Origins/CAAS*, and *Revitalising Marginal Lands: Discovery of Genes for Tolerance of Saline and Phosphorous Deficient Soils to Enhance and Sustain Productivity/IRRI*. The commissioned grant projects that are involved in this study are *Development of low tech gene-based trait assay technologies in rice and wheat/IRRI and CIMMYT*; *Development of low*

*tech gene-based trait assay technologies in rice and wheat/IRRI; and Simulation of marker-assisted selection strategies for optimising molecular breeding systems for drought tolerance in cereals/CSIRO.*

It is hoped that this project will provide baseline examples for the GCP community to enable an efficient transition to all future proposals including a preliminary product development plan.

#### **Subprogram4: Bioinformatics**

Detailed objectives and outputs for 2005 in Subprogram 4 are as follows:

##### ***1. Establish the GCP ‘information platform’***

This is the largest project in SP4 and aims at actually creating the platform needed for effective information exchange. The approach that was decided on in the GCPs first year, 2004, is that of web services. Web services allow wrapping local databases and software tools in such a way that they appear to be part of one system: they can be approached via a common protocol and a common language via the internet. The protocol can be interpreted by the wrapper, and the language can be translated by the wrapper in the language understood by the local database or software tool, which answers the query or does the analysis. The output is then translated back to the common language sent back to the one requesting it using the common protocol.

Implementation of this approach has a number of components: the technology (software and protocol) should be available and the common language developed; the technology should be implemented by the consortium members requiring training; and the technology should be used. And since this can not be accomplished immediately, some short term solutions should be created that are compatible with the final solutions. In 2005 some major steps towards this goal have been realised. This involved a number of activities, each with their own outputs.

The first output is the development of the common language, based on models of the information-domain. To make sure that these models will form a good basis not only for data exchange but also for software development, it was decided to take a solid approach. In 2005 methodology has been selected for the modelling, and the format for recording the models has been selected (Omondo EclipseUML). Editorial teams for the different sub-domains have been formed (“Generic Core Models”, “Germplasm/Phenotype/Genotype Models”, “Passport Models”, “Functional Genomics Model”). These teams did the modeling resulting in version 1 models for all sub-domains. These models were consolidated into a single comprehensive model and posted into the GCP Middleware CropForge project at <http://cropforge.irri.org/projects/gcpmiddleware/>.

The second output is concerned with the training of staff and implementation of web services technology in the institutions. Several approaches and technologies for an easy deployment at the institutes have been considered. In June a workshop on “Web Services, its technical fundamentals and future implementations” was held to bring together experts of the different technologies that are to be used (DiGIR, BioCASE and BioMOBY) with members of the Generation CP Consortium. This meeting aimed at providing a better understanding and promotion of web service technologies, but, most of all, it had to provide a clear implementation

plan for the deployment and implementation of web services. A tool (Model Mapper Toolkit) for generating web services on site is expected to be ready before the end of the year.

The third output concerns the creation of a registry for links to the datasets (yellow pages) or datasets themselves that are not available as web services yet. For this purpose an inventory of all available data sets generated in GCP projects has been made and the PIs of these projects have been approached asking for the data sets or information about the availability. A website has been created that will make this information available, including links to the actual data sets (if available). This site will also allow GCP management to monitor the availability of data.

The fourth output aims at improving the web services technology and applying it in a number of show cases, to establish the GCP as a relevant player in the international arena. For this purpose several software tools have been developed. An important example is the new “MOBY Services Support (“MOSES”)” framework to accelerate Java web services development. A number of GCP partners are now using this tool kit to develop MOBY web services in their projects. Also a “MOBY Dashboard” graphical user interface to facilitate the specification, registration, code generation and testing of web services is being developed. GCP partners continue to register additional web service data types and services in BioMOBY central.

The fifth output concerns a solution to the short term issues that are not resolved yet by applying the web services technology, by creating small applications for uploading and centrally storing data sets. For this purpose, templates for GCP passport data and fingerprinting data (SSR) have been developed and published. These templates are based on the GCP domain models (as described above). The mapping template is expected to be finalised by the end of November. This will follow as much as possible the popular and well defined CMAP format. Phase 2 of the project is now expected to start in December with the development of templates for phenotyping, SNP and DArT fingerprinting data. Tools to convert data in the templates to other popular formats, such as input files for analysis tools, are being developed.

The sixth and final output aims at properly capturing and analysing the wishes of the users and providing a platform for simultaneous software development, both facilitating effective software development. A web-based system for collaborative software development (<http://cropforge.irri.org>) has been created. Currently, there are 43 hosted projects and 93 registered users of this system, many, but not all of the projects currently hosted on CropForge are GCP related. Each project has its own administrator who is responsible for the project content, project resources configuration and project team management. A web-based system for collaborative development of textual content, CGP Wiki (<http://cropwiki.irri.org/gcp>) is in use since February 2005. Currently, there are about 280 pages that can be considered proper content pages. There have been a total of 23,689 page views, and 2,974 page edits since the Wiki was setup. There are 123 registered users, of which 12 users from 7 different institutions have administrator roles. Currently, most of the content of the site is related to SP4 according to the project proposal, but other subprograms have indicated their interest to use this platform in the near future.

## ***2. Improve the GCP information platform components***

The second project (in MTP terminology) has to do with the quality of the components that are to be part of the platform created in the first project. To ensure this quality of local curation of data, a number of issues have to be considered. There is the issue of institutional capacity needed to act as data supplier to the GCP, the issue of the quality of the data that are supplied, and finally the capacity needed by GCP scientists that can not be supplied by one single center.

Concerning institutional capacity, the CGP has the policy that consortium members are responsible themselves for creating appropriate capacity. However, where the GCP can create synergies, these opportunities will be used. In addition, the GCP aims to kick-start the development of institutional bioinformatics capacity by supporting the building of that capacity. The financial support provided by the GCP for this purpose is phased out from 50k\$ in 2004 via 33k\$ in 2005 to 17k\$ next year. It is provided to all eight CGIAR institutions that have been in the consortium since the beginning (CIAT, CIMMYT, CIP, ICARDA, IITA, IPGRI and IRRI). The funds have been used for a variety of activities, ranging from hiring and training bioinformatics staff, to installing and/or improving the LIMS in the institute.

Concerning the data quality of data and the systems they are managed and analysed with, a number of activities have been deployed. It became clear that the data quality issue required different approaches from the platform issue. In 2006 these activities will be separated. Guidelines were established during a SP4 meeting in February for conducting base-line quality surveys of data in GCP repositories. These guidelines were used as input of a meeting Data Quality Workshop in August where base line quality surveys were reviewed from IPGRI-INIBAP, IRRI, CIP, ICRISAT, ICARDA and IITA. The development of a Data Quality Strategy for the GCP was started. Concerning the platform development a different approach was followed. Nine separate activities were organised ranging from the definition of a general platform architecture, to implementation of domain models in GCP middleware, further development of aspects of LIMS, data warehousing and adaptation of functional genomics tools. These relatively small activities will for the building blocks in terms of components and experiences that will be used to create a versatile platform in the coming years.

The final element in improving the components of the GCP Information Platform is the creation of new joint institutional capacity, capacity that could not be established without the collaboration in the GCP. The creation, implementation, and integration into GCP toolbox of a high performance computing (HPC) facility, last year, was the first step in that direction. In 2005 the primary emphasis of work has been upon the development of use cases by CIP, ICRISAT and IRRI teams as the basis to facilitate and stimulate use of the HPCs by all GCP collaborators. Use cases implemented this year include automatic gene annotation including COS using BLAST with custom scripts, simulation tools for linear models (statistics) implemented in R, association tests and population sub-structure analysis based on molecular markers (using Structure), construction of a pipeline of public domain tools for sequence assembly, SNP detection and visualisation. Further, a large number open source software tools have been structured and compiled for the cluster and have been installed on the Paracel HPC (these tools include MegaBlast from TGICL for EST clustering; PCAP and cap3, implemented using MPI and PBS for job scheduling.) Web service access has been created to these and other tools (including Bio-Mirror of public sequence databases, EMBOSS, TIGR microarray analysis tools, MAANOVA R Statistics package).

### ***3. Create software in support of GCP activities***

The third component of the SP4 activities is the one that will grow in importance once the first two have established themselves: the direct support to the GCP activities in terms of software tools and algorithms. The first three Subprograms have data collection, curation, and analysis needs that SP4 must address for optimal deployment of GCP outputs. Based on the needs formulated by the other SPs, five activities were articulated whose outputs would address the needs.

The first activity aims at facilitating germplasm sampling based on all available passport, phenotype, and genotype data, depending on the nature of the sample required. Algorithms have been developed for representative sub sampling (cluster analysis with dependent data and several approaches for joint analysis of molecular and phenotypic data), and ‘structural’ LD free sub sampling. The wide range of strategies that were developed and tested was implemented in an existing software package (DarWin). Next year, capacity will be made available to support the use of these algorithms by SP1 scientists.

The second activity aims at creating access to gene orthology relationships across species and related paralogy relationships within gene families. This is a two year project that started in 2005 with some delay due to problems with the recruitment of staff. A GMOD Chado database is being deployed for the stress catalog. The “Apollo” genome sequence browser, now known to include the “JalView” multiple sequence and phylogenetic tree viewer, is being considered for GCP platform deployment in the activity. Next year there will be more results to report!

The third activity in this component of SP4 will create a crop gene expression database that will allow scientists to easily find and compare expression data across species. This is also a two year project, but large achievements have already been made in the first year. A prototype of the crop gene expression database was constructed and will be made available in December 2005. It is based on the Rice expression data base (<http://red.dna.affrc.go.jp/RED>). The unified Transcription Unit-based system (UTUS) was introduced to allow coverage of a wide spectrum of probes. A total 62,617 TU were determined by mapping genomic (Predicted CDS) and transcript-based nucleotide sequences onto the rice genome (pseudomolecule rel. 3). For the description of the microarray experiments the MIAME-plant standard is used. A tool for the analysis of the clustered gene set and cis elements appearing in their promoter regions has been developed. The new database has been connected to the BioMOBY central registry and several web services have been started (including gene search). We expect that we currently have access to more than 200 microarray data sets (these will be mounted after publication).

The fourth activity aims at allowing breeders to more efficiently use markers in breeding programs by integrating existing software in an integrated platform (iMAS). This two year project has seen a very good start. Software for inclusion in the platform has been selected and largely tested. This involves following eight packages: IRRISTAT, GMendel (and possibly MapDisto), PlabQTL and Win QTL-Cartographer, Tassel, PopMin, and GGT. The Java-based development of the system is in progress according to the planning, a GUI has been developed and successfully tested. All software packages have been incorporated using Java-based open-source tools.

The fifth and final activity will create an eco-physiological – statistical framework for the analysis of GxE and QTLxE as occurring in abiotic stress trials, facilitating a better understanding of results of phenotyping experiments. This is the only ‘competitive grant’ project in SP4, and it started with a delay of five months. The work on wheat and maize has started. For wheat the work has focused so far on phenotypic and molecular data collation, preliminary analyses of phenotypic data, and the generation of the molecular map, using various methods. For maize, a single-trait multi-environment mixed model was developed and evaluated using the historical data on a population for which alternative analyses have been performed in the past. The mixed model methodology will in the coming months be extended to allow for multi-trait multi-environment analysis.

### **Subprogram 5: Capacity Building and Enabling Delivery**

Detailed objectives and outputs for 2005 in Subprogram 5 are as follows:

#### ***1. Establish a Training Program***

Three training courses on Project Proposal Development were conducted in Africa (Cotonou, Benin), Asia (Kuala Lumpur, Malaysia), and Latin America (Quito, Ecuador). The objective of these courses was to increase the capacity of the GCP African, Latin American, and Asian institutions as well as that of potential partners in developing high-quality project proposals, leading to more effective distribution of research outputs, results, and an increased fundraising ability. As a result of the courses, 69 scientists from NARS and CGIAR institutions and potential partners to the GCP were trained in planning and writing quality project proposals. They also acquired knowledge on donors’ funding criteria (especially those for social impact and sustainability) and on how to respect them in the design of concept notes and project proposals. The participants represented 45 NARS institutions from 33 countries (9 countries in Africa, 11 in Asia, and 13 in Latin America).

Some exciting outcomes of the courses are: 1) several participants already used the course materials to teach colleagues upon return to their home institution, distributed the course CD, and already started its translation to Spanish; 2) a few participants working in the same crop decided to write a collaborative project proposal; 3) an interview of the course organisers was broadcasted in Radio Cotonou (Benin), and one column in the Espresso de Guayaquil was dedicated to the importance of the workshop in Ecuador.

Three regional training courses on Genetic Diversity and Molecular Breeding were conducted at University of Pretoria-FABI in South Africa (Africa), Kasetsart University in Thailand (Asia) and INIA-La Platina in Chile (Latin America). The total number of participants that benefited from the course is 59. They represented 41 NARS institutions from 27 countries (10 in Africa, 10 in Asia, and 7 in Latin America). The objectives of this workshop were to provide both conceptual and hands-on training in the use of plant genetic diversity and molecular marker-assisted breeding, with emphasis on practical applied usage and improving the links between plant breeding, germplasm management and utilisation, and molecular biology methods. Participants’ average evaluation scores of the courses ranged from 3.94 to 4.54 (0-5 scale).

Selected highlights of the courses are: 1) the selection of participants for the course in South Africa prioritised teams of molecular geneticists and breeders from the same institution; 2) instructors for courses included GCP Consortium scientists as well as regional scientists, and the evaluation of the latter were particularly high. This feature was deemed as extremely valuable to ensure good communication between participants and lecturers, helping the creation of links among scientists and promoting regional sustainability; 3) left-over funds from the training events were distributed among outstanding participants selected by the organisers to advance research in their home institutions in line with the learning experience gained during the course, 4) The participants in the course in Thailand formed AgBioAsia, as a Yahoo Group ([www.yahogroups.com/groups/AgBioAsia](http://www.yahogroups.com/groups/AgBioAsia)), to foster collaborations with themselves and with the GCP organisers as a mechanism for technical backstopping, re-training, and problem solving in the Asia region.

A special partnership was established with the International Foundation for Science (IFS) to link technical courses on molecular genetics/breeding organised by SP5 with their call for research grants for young scientists in developing countries. Selected participants for courses were encouraged to apply to the IFS call and course instructors offered guidance for the preparation of proposals to ensure competitiveness. IFS earmarked two grants for each of the GCP courses. A pre-selection of suitable candidates will be made by the GCP followed by another evaluation by the standard IFS selection panel.

Linked to Subprogram 1, a workshop titled Molecular Markers for Allele Mining was organised by IPGRI based on results from the GCP's year 1 genotyping efforts. The workshop, conducted at the MS Swaminathan Research Foundation (MSSRF) in Chennai, India, gathered participants representing: the System-Wide Genetic Resources Program (SGRP, 5) of the CGIAR, the crops included in the SP1 genotyping activities of year 1 (10), crops involved in a SP1 commissioned project on association genetics (5), a number of experts in germplasm characterization (3), NARS from developed countries (4) and NARS from developing countries (8, plus 6 Indian scientists). It was an opportunity for scientists involved in SP1 to exchange their experience, successes, and difficulties, but also to have top scientists from NARS, advisers from advanced laboratories, and germplasm curators provide valuable ideas for improving activities. As a result of the workshop, the idea developed to open the platform of advanced laboratories to those NARS partners with information-rich germplasm. This could provide novel genetic information from *ad hoc* molecular marker genotyping.

Linked to Subprogram 4, a workshop was co-organised with the Instituto Agronomico Mediterraneo de Zaragoza (IAMZ, Spain) on Design and Analysis of Multi-Environmental Trials: Conventional and QTL-based methods. The objective of the course was to provide the participants with working knowledge of statistical tools to be applied in breeding programs requiring the use of multi-environmental trials to assess the responses of genotypes and their dependence on the environment. The program covered the design of multi-environmental trials, the analysis and interpretation of genotype by environment interactions, the use of appropriate software, and the development of appropriate breeding strategies to better predict the genotypic responses. The course was attended by plant breeders, geneticists, and applied statisticians, all already involved in the subject matter presented. In total, there were 30 participants from 16 countries.

In most workshops, a questionnaire was circulated to assess the capacity building needs of potential GCP partners, requesting their perspectives on strengths and capacities of the institution to which they belong, their potential role in and contribution to the GCP, and the expectations they might have from the GCP. Common needs were the following: 1) to continue with the GCP fellowship program and extend it for PhD candidates; 2) to further molecular breeding training, 3) to promote a small grant program for NARS so that they can contribute to research in their own countries; 4) to upgrade a selected number of research facilities in the South; and 5) to train technical scientific staff in national programs. A common wish was to host GCP activities in their countries and institutions.

## ***2. Establish a Fellowships scheme***

A first call for the GCP fellowships was opened and 18 candidate applications were received. In a second call, 5 applicants competed. The full GCP Management Team evaluated and selected the winners as follows:

1. Chiedozi Egesi from the National Root Crops Research Institute (NRCRI) in Nigeria, to work on the project “Genetic mapping of resistance genes to major arthropod pests and delayed postharvest physiological deterioration (PPD) in Cassava” at the International Center for Tropical Agriculture (CIAT).
2. Daniel Fonceka from the Centre d'Etude Régional pour l'Amélioration de l'Adaptation à la Sécheresse (CERAAS) in Senegal went to the Centre de Coopération Internationale en Recherche Agronomique pour le Développement (CIRAD) to work on “Drought tolerance related candidate genes in sorghum and their allelic diversity as revealed by Ecotilling.”
3. Heru Kuswantoro, from the Indonesian Legume and Tuber Crops Research Institute (ILETRI), is at Cornell University working in the project “Characterising the genetic diversity of Indonesian soybean and groundnut landraces using microsatellite and SNP markers.”
4. Kameswara Rao Kottapalli, from the Directorate of Oilseeds Development in India, went to the National Institute of Agrobiological Sciences (NIAS) to work on “Functional genome analysis for candidate gene discovery from a chromosomal region having mapped QTLs/loci for broad-spectrum bacterial blight disease resistance and submergence tolerance in rice”.
5. Luis Carlos Rodríguez Zapata from the Centro de Investigación Científica de Yucatán (Mexico) is working on the development of COS markers for drought tolerance for Musa germplasm, in collaboration with the International Plant Genetic Resources Institute (IPGRI)/International Network for the Improvement of Banana and Plantain (INIBAP).
6. Matthews M. Dida, from Maseno University in Kenya, is working on the project “From rice to finger millet: Comparative mapping of blast resistance genes” at Cornell University and the University of Georgia.
7. Suresh Kumar Sampath from Bharathiar University (India) works in “Enabling Biological Databases Interoperability to create an on-line integrated information resource on Banana and Plantain a worldwide access” at IPGRI/INIBAP.
8. Syed Sarfraz Hussein from the National Centre of Excellence in Molecular Biology at the University of the Punjab (Pakistan) is at the Australian Centre for Plant Functional Genomics (ACPF) working on “Comparative Genomics of Drought Tolerance in Wheat.”

The Pioneer-GCP Fellowship towards a PhD in a US University was also opened for applications in early 2005. A team composed of the GCP Director, SP3, and SP5 leaders, as well as the

training officer at Pioneer, evaluated 12 applications. The award was given to Vinod Jakkula from India to conduct his dissertation research at the University of Georgia working on “Comparative and functional analysis of genes affecting plant height.”

### ***3. Establish a Travel grants scheme***

A travel grant program to support NARS' visits to GCP member institutions or GCP-related meetings was initiated. Fourteen standard travel grants were awarded to scientists from the National Agricultural Research Organisation (NARO, Uganda), the Ethiopian Agricultural Research Organisation (EARO), the Rwanda National Agricultural Research Organisation (ISAR), the Kenya Agricultural Research Institute (KARI), Moi University (Kenya), the Hanoi University of Technology (Vietnam), Huazhong Agricultural University (China), the University of Namibia, the Indian Agricultural Research Institute (IARI), Tamil Nadu Agricultural University (India) and the Mtskheta Breeding Station (Georgia). Visits were paid to the National Institute of Agrobiological Sciences (NIAS, Japan), IRRI (The Philippines), Cornell University (USA), ICRISAT (India), CIP (Peru), ICARDA (Syria), CIRAD (France), and the African Centre for Gene Technologies (ACGT)/ FABI (South Africa). Other grantees will attend the 7th International Wheat Conference (Mar del Plata, Argentina) or the 5th International Rice Genetics Symposium and 3rd International Rice Functional Genomics Symposium at IRRI (The Philippines). In addition, 14 awards were given to scientists to attend InterDrought II in Rome (Italy). They represented national research organizations of Senegal, Russia, Uzbekistan, Iran, Bangladesh, Brazil, Turkey, Egypt, India, Pakistan, Tunisia, and China.

### ***4. Develop learning materials or other resources relevant to the subject matter of the GCP***

In 2005, the development of a series of training materials also got underway in the areas: genetic diversity analysis, plant genomics, molecular breeding, and bioinformatics. These are in addition to those prepared *ad hoc* to provide instruction in diverse workshops, either as part of the training program (project proposal development and genetic diversity/molecular breeding) or organised in the context of research projects. Each training workshop produced a CD-Rom with the course contents, presentations of instructors, software if applicable, and reference publications in PDF format. The basic set of learning materials will be finalised and ready in 2006, and will be stored in a training material repository accessible through the GCP web site.

An Interactive Resource Center to offer online information and sustainable support to capacity recipients (<http://irc.igd.cornell.edu>) was also created. The site offers links to plant databases, laboratory protocols, freely available literature, and funding and training opportunities. It is still under development and soon is expected to contain lists of suppliers of reagents, primers, other laboratory supplies, databases made for specific needs, workshops -real and virtual, and online tutorials.

A proposal was prepared in collaboration with BECA (“Tapping Crop Biodiversity for the Resource Poor in East and Central Africa”) and submitted to the Rockefeller Foundation to match funds. The goals of the project are: a) to enable national researchers to apply genomics tools for characterization and enhancement of major food crops in the ASARECA region, b) to empower a wider spectrum of stakeholders to link into regional and international research activities, c) to provide an operational framework for trainees from the GCP-BECA Molecular Breeding Training Program to continue and intensify their training, d) to generate an active

network of molecular breeders effectively using the BECA facility, and e) to develop an operational framework for collaboration on comparative genomics in Africa fostering multidisciplinary teams working across three African crops (sorghum, cassava and bean).

### ***5. Monitor capacity building activities within research projects***

Providing hands-on capacity for NARS scientists outside the Consortium was a requirement of proposals submitted to the competitive call and highly encouraged in commissioned research activities in the technical subprograms. Mechanisms to follow-up on such activities were put in place in Subprogram 5 in 2005. Beneficiaries of capacity building activities were requested to fill out a questionnaire to inform the GCP Management about progress, rating the learning experience, applicability to their home institutions, and suitability of the program to fulfill their needs. Results are used to inform the placement of future trainees to research teams and to help better shape the activities. An interesting achievement of this endeavour has already been the increase of awareness of Consortium scientists of the importance and relevance of linking with outside NARS to reach the goals of the GCP. Up to date, ongoing research projects awarded through the competitive call in 2004 have engaged 23 scientists from developing country NARS in hands-on research. In addition, several training workshops have been organised with participation of as many as 138 scientists from 68 national institutions in 47 countries worldwide.

An activity closely related to SP1 was the development of microsatellite marker kits (reference molecular marker subsets to analyse diversity of germplasm and allow comparability across institutions and germplasm collections) for the crops genotyped in that subprogram in year 1 of the GCP. The conversion into a capacity building activity was completed by means of selected NARS scientists to participate in the development of the kits. Ten scientists from nine countries started work at a Consortium institution to carry out this activity: barley (Chao Lu, Yangzhou University, China), cassava (Luis Rodolfo Montes, Universidad de San Carlos, Guatemala), chickpea (SL Dwivedi, ICRISAT, India), common bean (Sandra Lorigados, INCA, Cuba), maize (Chaba Jampatong, National Corn and Sorghum Research Center, Thailand), Musa (Kouassi Koffi Simplicite, CNRA, Ivory Coast), potato (Eliana Alba Alba, PROINPA, Bolivia), rice (Reflinur S.P., CABGRRD, Indonesia), sorghum (Mbaye Ndoeye Sall, CERAAS, Senegal), and wheat (Genying Li, Shandong Academy of Agricultural Sciences, China).

### ***6. Conduct research on policy issues of relevance to the GCP***

Some achievements in the arena of policy research in 2005 are:

- A Policy session was organised in the framework of the Symposium on Genomics-based Plant Germplasm Research, held in Beijing, China (25-28 April). The objectives of the session were two-fold: 1) to raise awareness of policy issues, particularly access and benefit sharing under the Convention of Biological Diversity and the International Treaty on Plant Genetic Resources, and the ir implications on the technical work of the Generation Challenge Program, and 2) to get advice from National Program officials and members of the Consortium on their vision of access to genetic resources relevant to the GCP. The session was also used to give partners a chance to participate in the conference. The session was well attended and had high-level participants. As a result of the session, there was clear advice from the Chinese participants on the processes for getting access to germplasm and materials. The publication "Access to plant genetic resources for genomic

research for the poor: from global policies to target-oriented rules” by Niels P. Louwaars, Eva Thörn, José Esquinas-Alcazar, Shumin Wang, and Abebe Demissie was submitted for publication to “Genetic Resources Conservation and Utilisation.”

- A publication compiling a series of white papers on policy issues relevant for the Generation CP was prepared and published: Genetic Resource Policies and the Generation Challenge Program. It contains the following sections: a) the policy environment of the Generation CP regarding rights on biological materials, technologies and knowledge; b) humanitarian licenses; c) the evolving international regime of liability and redress relating to the use of genetically modified organisms; d) open source mechanisms: the example of BIOS; e) issues on access to genetic resources; and f) impacts of strengthened intellectual property rights regimes on the plant breeding industry in developing countries.
- A number of projects were commissioned as follows: 1) Distant Policies is a distant learning module for scientists on Genetic Resources Policies and their implications for ‘Freedom-to-Operate.’ The objective of this product is to provide a basic and practical tool to help scientists of Generation Challenge Program projects to understand the importance of rights associated with the access and use of plant genetic resources and tools, methods, and products protected by intellectual property rights (IPRs) and contracts when the results of research are to be used freely by smallholder farmers. It is addressed to GCP scientists who operate at a considerable distance from (inter)national PGR regulations and IPR, but also to scientists at NARS and other persons receiving GCP-products that need to be aware of possible strings attached to these (intermediary) products. 2) An Asset Inventory System for the Generation Challenge Program to provide a service function to GCP scientists and administration, in order to facilitate product delivery, distribution, and uptake (technology transfer). It is developing e-versions of asset/product identification and 3<sup>rd</sup> party materials-reporting forms, as well as an inventory database of GCP Products and 3<sup>rd</sup> party materials. 3) IP Matters is an Intellectual Property/Access and Benefit Sharing-Helpdesk, designed as an on-line resource for the GCP community, its partners and stakeholders. It provides a practical on-line service desk for assistance, clearing-house activity, and feedback on topics concerned with intellectual property matters in the broadest sense. 4) Regional PGR courses, which will prepare learning materials and a course curriculum for NARS, to be tested in 2006 in a face-to-face workshop.

### ***7. Develop a GCP Delivery Strategy***

The understanding of the specific role of the GCP in delivering its products has significantly evolved since the GCP was established in 2003. With the progress of the Program during its first year, it seemed clear that the GCP should develop a delivery strategy to ensure that indeed the GCP products reach the intended users and make impacts on the poor. Thus, an electronic forum was conducted among a number of experts (15) in relevant fields (agricultural economics, social sciences, intellectual property rights, bio-policies, agricultural geography, taxonomy, research planning, food technology, plant pathology, crop evolution, genetic resources, genetics-plant breeding) to the subject. Experts were drawn from within and outside the GCP Consortium, and included the public and the private sectors. The forum was driven by selected questions

addressing the significance of the involvement of the GCP in product delivery, the definition of intended users and GCP products, the mechanisms possible in the areas where the resource-poor live, and indicators of impact, among others.

Following, a face-to-face workshop was conducted to brainstorm about the delivery strategy in general and to discuss the important principles on which to base the drafting of the strategy document. It involved 17 participants: plant breeders; social scientists with backgrounds in rural innovation, agricultural economists with expertise in impact assessment, agricultural anthropologists; an intellectual property attorney, and farmer leaders from the three regions (Africa, Asia, and Latin America). The group represented perspectives of the CG Centres, NARS, ARIs, Harvest+ CP, donors, and stakeholders.

As a result, the GCP Management Team produced a final delivery strategy document with a supplementary text with details for its implementation.

The main thrust of this strategy is the adoption by the GCP of a value-chain based approach in which it aims to catalyse the various players needed to bridge the gap between upstream strategic research in advanced laboratories and target user communities. In order to make it real, the GCP will require a delivery plan for every project proposal submitted and approved, and it will play an oversight role by ensuring that its products are inserted into existing delivery systems.

#### **Appendix B: Schedule of Contribution Received and Schedule of Disbursements to Partners**

**GENERATION CHALLENGE PROGRAM**  
**Schedule of transactions from July 2003 through December 31,**  
**2005**  
**(USD)**

<u>Cash Receipts</u>	2003	2004	2005	Total
Austria	54,482			54,482
DFID/UK	-	4,675,625	4,417,000	9,092,625
EC	-	5,224,850	6,027,334	11,252,184
Kirkhouse	-	-	15,000	15,000
Pioneer found	-	50,000	20,000	70,000
Rockefeller Foundation	-	-	837,931	837,931
Sweden/SIDA	107,013	-	189,495	296,509
Syngenta	-	15,000	-	15,000
World Bank	<u>3,000,000</u>	<u>1,000,000</u>	<u>2,500,000</u>	<u>6,500,000</u>
<b>Total</b>	3,161,495	10,965,475	14,006,760	28,133,730
Interest income	-	-	186,361	186,361
<b>Total income</b>	<b>3,161,495</b>	<b>10,965,475</b>	<b>14,193,121</b>	<b>28,320,091</b>
<u>Disbursements</u>				

<b>GC Centers</b>				
CIAT	-	622,188	1,082,236	1,704,424
CIMMYT	-	770,696	1,886,879	2,657,575
CIP	-	352,394	455,661	808,054
ICARDA	-	230,135	546,555	776,690
ICRISAT	-	479,213	743,406	1,222,619
IITA	-	634,462	706,013	1,340,474
ILRI	-	50,000	44,346	94,346
IPGRI	-	897,201	1,486,943	2,384,144
IRRI	-	742,204	2,252,171	2,994,375
WARDA	-	22,219	6,099	28,318
<b>Total CG Centers</b>	-	<b>4,800,710</b>	<b>9,210,309</b>	<b>14,011,019</b>
<b>ARIs Institutes</b>				
Agropolis/Cirad	-	514,509	1,325,912	1,840,421
Cornell University	-	50,150	821,991	872,141
WUR (University of Wageningen)	-	211,508	544,308	755,816
<b>Total ARIs</b>	-	<b>776,167</b>	<b>2,692,211</b>	<b>3,468,378</b>
<b>NARS Institutes</b>				
CAAS	-	245,687	574,006	819,693
EMBRAPA	-	3,225	1,002,136	1,005,361
NIAS	-	34,832	256,196	291,028
Others (Fellowships & Travel Grants)	-	-	243,769	243,769
<b>Total NARS</b>	-	<b>283,744</b>	<b>2,076,108</b>	<b>2,359,852</b>
<b>Total disbursements for research <sup>1/</sup></b>	-	<b>5,860,621</b>	<b># 13,978,627</b>	<b># 19,839,249</b>
<b>Total disbursements for Program Management</b>	<b>500,888</b>	<b>1,545,515</b>	<b>1,581,003</b>	<b>3,627,407</b>
<b>Total disbursements</b>	<b>500,888</b>	<b>7,406,137</b>	<b>15,559,631</b>	<b>23,466,655</b>
<b>Surplus/(Deficit) for year</b>	<b><u>2,660,607</u></b>	<b><u>3,559,338</u></b>	<b><u>(1,366,510)</u></b>	<b><u>4,853,436</u></b>

1/ The figures reflected in the total CG/ARIs/NARS disbursements does not reflect subcontracts to other institutions. The actual breakdown of GCP funds for research is approximately 50/25/25 percent for the CG/ARIs/NARS respectively.

Note: GCP Financial figures are based on Schedule included in Cimmyt's Audited Financial

Statements (2003-2005)

## Appendix C: Resource Allocation/Expenditure

### GENERATION CHALLENGE PROGRAM

Allocation/Expenditure from July 2003 through December 31, 2005

(USD)

	2003 Expenditure	2004 Expenditure	2005 Expenditure
<b>Research</b>			
Collaboration/Partnerships	-	5,860,621	13,978,627
<b>Total</b>	-	5,860,621 #	13,978,627
<b>Program Management</b>			
Personnel	15,827	252,934	262,090
Supplies and Services	404,156	1,221,657	1,310,066
Operational Travel	80,905	70,925	8,848
<b>Total</b>	<b>500,888</b>	<b>1,545,516</b>	<b>1,581,004</b>
<b>TOTAL</b>	<b>500,888</b>	<b>7,406,137</b>	<b>15,559,631</b>

**Appendix D. Products and Potential Users Table (for use by competitive grants proposal submitters)**

Activity 1=	Expected products*	Potential users**	Endorsement of Institutions and scientist?	Who in the project has already collaborated with user Institution and scientist?	Foreseen limitations of Institution and scientist to use the product?
	<i>Please name and describe</i>	<i>Institution and scientist name</i>	<i>yes or no***</i>		
Product 1					
Product 2					
Product 3					
Product 4					
Activity 2=					
Product 1					
Product 2					
Activity n=					

\* They can be markers, sequences, genes, populations, lines, databases, protocols, varieties, technologies, etc.

\*\* For each product, there can be more than one potential user. List all.

\*\*\* This means that contact with the potential user scientist or institution has already taken place. Written endorsement from scientist/institution needed.