Food, livelihood and health

The Generation Challenge Program

“Imagine this conference table is covered in sand,” says Merideth Bonierbale. “And every grain is a piece of information. Across the surface, small mounds are rising as scientists bring together the information that makes up their area of research. Occasionally they

may scan the horizon from the summit of their own mound and get a very general view of what is

may scan the horizon from the summit of their own mound and get a very general view of what is going on elsewhere, but they will rarely have an opportunity of delving into the detail of other mounds. Yet as time and knowledge progress we are evermore certain that information useful to us is hidden there…"

Information bottleneck

Dr. Bonierbale is the head of CIP's Germplasm Enhancement and Crop Improvement Division and she is speaking of a major problem facing today's crop improvement fraternity. Investigations into the genetic characteristics of food plants are advancing so rapidly that traditional methods of disseminating new findings are not enough to make sure that information is available to everyone who could use it. Not every result merits publication in a refereed journal, yet the unpublished research of a laboratory on one side of the globe might apply directly to the work of scientists on the other. How can they locate everything that might be relevant to their own investigations?

Thirty years ago the Green Revolution dramatically increased farm production through the spread of new plant types, irrigation and fertilizer. Today, the emerging Genomics Revolution (see Box 1) can bring new science to bear on problems encountered by the resource-poor farmers who derived little or no benefit from the earlier wave of innovation. But first, there has to be a way of collating, exploring and comparing the accumulating information – effectively and on a reasonable timescale.

Generation Challenge Program

During 2004, CIP's Information Technology Unit has made a significant contribution to alleviating this problem, not simply for the benefit of CIP's own in-house scientists, but also for the benefit of a much wider network of research institutions. CIP's information technology expertise is a valuable component of the Generation Challenge Program (see Box 2). This is an initiative that brings together three sets of partners – CGIAR, advanced research institutes (ARIs) and national agricultural research systems (NARS) in developing countries – with the aim of delivering the fruits of the Genomics Revolution to resource-poor farmers.

"Seventy-five per cent of the world's poorest people live in rural areas and rely on agriculture for food and income, and a great many of them live on marginal land, where modern domesticated varieties do not always fare well," points out Robert Zeigler, former Director of the Generation Challenge Program. "Crop varieties that can withstand harsh conditions and survive on few inputs could greatly help these people in their interminable struggle with food insecurity and poverty."

"Most developing countries do not have the scientific infrastructure or trained staff necessary to apply advanced genomics," explained Zeigler. "Unfortunately, this means that the scientific institutions in closest proximity to

Box 1

The Genomics Revolution

The Genomics Revolution is producing plants specifically designed and bred to overcome the difficult conditions found in smallholders' fields and marginal environments and so improve the quality and quantity of these farmers' yields. It has been fueled by technological advances that have allowed scientists to unravel the genetic code of an organism and draw up maps of the entire genome, with landmarks locating genes associated with particular characteristics, such as adaptability, disease resistance or environmental tolerance. And as the genetic maps are filled with ever-finer detail, breeders are cross-referencing the maps of different species (see illustrations on opposite page and following). In this way, the genetics of whole families of plants can be compared. Thus the existence of genetic information on a particular trait can be predicted in one species, from the detailed genetic studies of another. This is a major advance, making it possible to explore the relationship of gene sequence to function across species, further expanding the opportunities for significant breakthroughs and delivering the promise of the Genomics Revolution.

Up to now, plant breeding has been as much an art as a science. A highly successful art with an honorable pedigree, true, but always a subjective exercise, based on the experience and skill of the plant breeder to choose parents for crosses and to select out, either visually or by means of empirical tests, improved individuals from among the progeny of those crosses. There are notable exceptions, in which a few identified genes have had a dramatic effect when bred into other varieties (for example, the directed introduction of dwarfing genes in wheat and rice was responsible for the increased productivity of the Green Revolution), but comparative genomics has the capacity to make these exceptions the rule and, ultimately, to change the plant breeders' art into an objectively based science.

struggling farmers and poor rural areas are the least equipped to take advantage of the technological revolution that may help those people. To make a lasting impact," he continues, "the Generation Challenge Program must create a global environment that promotes scientific innovation by NARS in those countries."

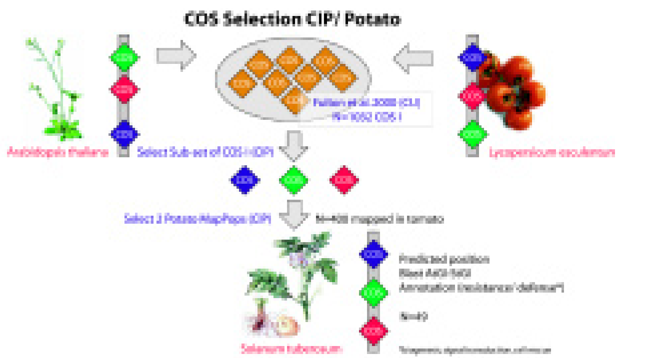
High Performance Computing system

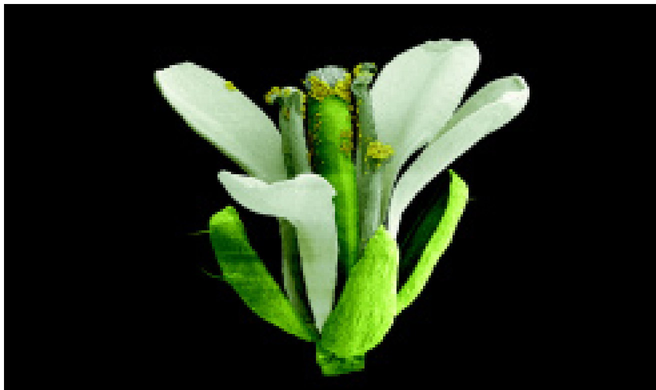
Challenge Program member scientists often quip that the fourth of the Program's five subprograms – Bioinformatics (SP4 see Box 2) – "is the lynchpin of the whole global undertaking." If SP4 is the lynchpin of the system, then CIP is responsible for ensuring that it functions smoothly and effectively. Managed by its Head, Anthony Collins, CIP's Information Technology Unit has set up a Paracel cluster computer system made up of four nodes (based at CIP and three other research establishments within the CGIAR network) with dual 64-bit processing units and a terabyte of database storage (with substantial capability for future expansion), which will facilitate the analysis of very large germplasm, molecular and functional genomics data sets. And this High Performance Computing (HPC) system is already enabling scientists to make progress on important challenges. “For instance, the HPC facility has greatly enhanced the application of comparative genomics to the search for drought-tolerant genes in potatoes,” says CIP biotechnologist Roland Schafleitner.

In a world confronting the prospects of climate change, drought is one of the most serious challenges faced by resource-poor farmers. Comparative genomics could provide the vital clues needed to enhance drought-tolerance in potatoes and other staple crops of developing countries. In this work, the drought-stress responses at the genetic level in a known or model plant are compared with those of potato genotypes exhibiting different levels of tolerance. These comparisons provide invaluable information on which genes are involved in drought-tolerance mechanisms, but an enormous amount of data has to be processed to arrive at the result.

The plant used for comparison is *Arabidopsis* – thale cress – a small weed whose entire genome has been sequenced. In recent years, scientists have shown that *Arabidopsis* has about 2,300 genes that are involved with its response to stresses such as drought, cold or high salinity. With this knowledge to hand, scientists tested potatoes for the presence of orthologous genes (meaning genes that have the same stress-responsive function as in *Arabidopsis*), and could begin evaluating the potential of those genes as candidates for drought tolerance research.

The first step in this procedure was to discover whether high-sequence similarity to stress-induced *Arabidopsis* genes also occurred in the potato. “Comparing the 2,300 *Arabidopsis* stress-responsive gene sequences that had been identified with the

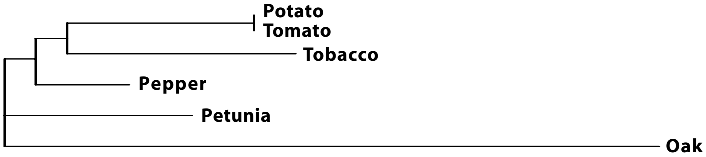




***Arabidopsis thaliana* wildtype flower**. Scanning electron microscopy image, artificially coloured. *Arabidopsis* is approximately 5mm in size

Photograph © by Jürgen Berger, Max Plank institut for Developmental Botany, Germany

```
Potato   AGGAAGCTTTTAGCCTTTTCGACAAGGATGGCGATGGCTGTATTACTACCAAGGAGTTGGGAACAGTGATG
Tomato   AGGAAGCTTTTAGCCTTTTCGACAAGGATGGCGATGGCTGTATTACTACCAAGGAGTTGGGAACAGTGATG
Tobacco  AGGAGGCCTTTAGCCTTTTCGACAAGGACGGCGATGGCTGTATTACTACCAAGGAATTGGGAACAGTGATG
Pepper   AGGAGGCTTTTAGCCTTTTTGACAAGGACGGCGATGGCTGTATTACTACCAAGGAGTTGGGAACAGTGATG
Petunia  AGGAAGCTTTTAGCCTTTTTGACAAGGACGGTGACGGCTGTATTACTACCAAGGAGTTGGGAACAGTGATG
Oak      AGGAAGCCTTCAGCCTCTTTGACAAGGACGGCGATGGCTGCATCACTACCAAAGAGTTGGGAACAGTCATG
         **** ** ** ***** ** ******** ** ** ***** ** ******** ** *********** ***
```



Genetic maps. Genomics allows scientists to create genetic maps that allow cross referencngacross species. This exaple shows simil arities in part of a calmodulin gene in different Solanaceae species and in oak. Such sequence similarity among the same genes of different species is the basis of comparative genomics. This particular map is valuable because calmodulin proteins are cellular calcium sensors. These proteins interact at the same time with calcium ions and a wide range of other proteins, translating calcium signals into modulation of enzyme activity. Calmodulin-mediated signals are involved in a many important cellular functions, such as growth, development and stress response.

approximately 40,000 potato gene sequences available today, would have meant performing 2,300 searches on the potato material," Roland Schafleitner explains. "So we began by clustering the 2,300 stress-responsive *Arabidopsis* genes according to their spatial and temporal expression patterns, and then selected some 450 genes whose role was consistently significant. Then we had to search for orthologs to those genes in the potato material. With conventional equipment that would have taken days and driven us crazy," he adds. "But we had the HPC system. The results were available almost immediately as a sound foundation for further research into drought tolerance in potatoes."

SP4 an innovator in the field

"It is important to remember that CIP's HPC is part of a process, not just a facility," says Reinhard Simon, Head of CIP's Research Infomatics Unit, who coordinated the establishment of an

Box2 The Generation Challenge Program

CGIAR's Generation Challenge Program has five subprograms, within which scientists

1. explore the genetic diversity and potential of its 22 mandate crops*
2. focus on developing genomic technologies and approaches to advance understanding of genetic principles across significant crop species in developing countries
3. seek ways to increase the efficiency, speed and scope of plant breeding
4. build a viable bioinformatics platform
5. work to improve the research capabilities of developing countries.

* Andean root and tuber crops, barley, cassava, chickpea, coconut, cowpea, finger millet, forages, groundnut, lentil, maize, *Musa*, pearl millet, *Phaseoulus*, pigeon pea, potato, rice, sorghum, soybean, sweet potato, wheat and yam

infrastructure for building capacity, storing and exchanging data across the network in a standard form. "With genomics breaking down the knowledge boundaries between plant genomes, the informatics systems must work seamlessly across institutes and crops to meet the storage and analysis needs of this increasingly integrated world of plant genetics," he explains. "Every aspect of the system we are building in SP4 has to be agreed upon by every information technology manager in the network. It has to overcome the problems of using existing data, capturing all experimental data and then storing everything in a mutually comprehensible and interactive form. It's a challenge," he concludes, with characteristic understatement.

But it is a challenge that excites an enthusiastic response, not least in Theo van Hintum, a bioinformatics specialist at Wageningen University in the Netherlands, who heads SP4. He is optimistic about the potential of the SP4 platform. "There is a clear commitment among our partners to make the SP4 infrastructure easily accessible and to make our increasingly complex analyses easier," he says. "SP4 gets really interesting when we start talking about processing all the information that is available through the network. There is a whole range of applications that researchers have only dreamt about that this platform will enable us to do. These new applications and analyses will make SP4 an innovator in this field, and the Generation Challenge Program a leader in making high science work for the people who need it most."