THE ROLE OF THE INTERNATIONAL COCOA GERMPLASM DATABASE AND THE INTERNATIONAL COCOA QUARANTINE CENTRE IN INFORMATION MANAGEMENT AND DISTRIBUTION OF COCOA GENETIC RESOURCES

C.J. Turnbull, A.J. Daymond, H. Lake, B.E. Main, B. Radha,
N.C. Cryer, M.J. End * and P. Hadley

School of Biological Science, Harborne Building, University of Reading, RG6 6AS, UK
* Cocoa Research Association Ltd., Knightlands, Wickford, SS12 9JR, UK

SUMMARY

The University of Reading has been a key player in the cocoa research community for over 20 years, supplying disease-free germplasm through the International Cocoa Quarantine Centre (ICQC,R), as well as providing a single point of reference for information about cocoa clones through the International Cocoa Germplasm Database (ICGD).

ICGD is the primary source of information on the characteristics, origin and location of cocoa germplasm resources worldwide. The database provides access to data and photographs from a wide variety of sources, all of which are fully referenced. Since 1989, the ICGD has been available to research institutes and genebanks around the world free of charge. The contribution of data and input into the development of the database by cocoa research institutes and individuals has added considerably to the success and longevity of the project.

ICGD has recently undergone a complete transformation as a result of the shift in focus towards online development (www.icgd.reading.ac.uk), which allows more rapid dissemination of new information. This has provided the opportunity to take a more user-friendly approach to presenting ICGD. It also means that the database is now well suited to involvement in international collaborations and networks, such as CocoaGenDB (in collaboration with CIRAD, France and USDA) and CacaoNet (the global network for cacao genetic resources), in which ICGD is one of the central components of the simple yet robust Information Management System (IMS) being developed.

The quarantine centre at Reading will also become a key resource within the CacaoNet initiative, since it provides the primary route for safe international movement of cocoa germplasm. Procedures employed within ICQC,R follow international guidelines and include extensive virus indexing and observations for pests and fungal diseases.

The ICQC,R collection currently contains over 300 post-quarantine accessions of diverse origins, which have been supplied to institutes in over 20 cocoa-growing countries, including both established and newly emerging producers. Several collections also supply new germplasm, which is selected on the basis of resistance to economically important diseases and other key agronomic and quality traits.

Here we illustrate the evolution of the work on cocoa genetic resources at Reading over the past 20 years. We discuss how ICGD can continue to develop with a greater focus on global networks while still maintaining the traditional values of universal accessibility, and we will explore how cocoa germplasm movement via ICQC,R has impacted the global cocoa research community.

INTRODUCTION

Over the past 100 years or so, samples from thousands of different cocoa plants have been collected from the rainforest, farmers' fields and historic breeding programmes and are now grown as living trees ('accessions') in genebanks, usually under the management of national research institutes in cocoa producing countries. Many of these are currently held in the two international cocoa genebanks: around 3000 at the International Cocoa Genebank, Trinidad (ICG,T), curated by the Cocoa Research Unit (CRU) of the University of the West Indies; about 1,100 at the International Cacao Collection at CATIE (IC3), Costa Rica, managed by the Centro Agronómico Tropical de Investigación y Enseñanza (CATIE). Samples of this material are freely available to *bona fide* users, such as cocoa breeders and other scientists, once it has passed the appropriate quarantine measures to prevent the spread of pests and diseases. Several of the genebanks, including the international genebanks, have programmes to collect data on the flower, leaf and pod/bean characteristics of their material. This information, together with data

from genetic fingerprinting projects, can be used to study the variability of the material, with the aim of maximising the genetic diversity they hold whilst eliminating unnecessary duplication. Many of the research centres are also evaluating their material so that they can provide information of immediate value to breeders, such as reaction to local pests/diseases, flavour, cocoa butter content and yield characteristics. The ICGD's role is to collate all this information and make it easily accessible to the research community who must provide the cocoa farmers with the improved planting materials and associated agronomical techniques to ensure a sustainable future for the whole cocoa economy.

THE INTERNATIONAL COCOA QUARANTINE CENTRE, READING

Project Rationale

The International Cocoa Quarantine Centre (ICQC,R) is an intermediate quarantine operation that serves to facilitate the safe movement of cocoa germplasm from one cocoa growing region to another. Such safe movement is particularly important in cocoa given that many pests and diseases are confined to particular geographical regions (e.g. cocoa swollen shoot virus in West Africa, witches' broom disease in parts of South America, and cocoa pod borer in S.E. Asia). ICQC,R is ideally located in a non-producing country, which minimises the risk of endemic cocoa diseases entering the facility.

Development of the ICQC,R and Current Procedures

ICQC,R started at the University of Reading in 1985 when 90 accessions were transferred from the Royal Botanic Gardens at Kew after they took the decision to no longer offer a quarantine service for cocoa. Initially the centre only distributed material from these accessions, until full quarantine operations started two years later. The physical capacity of the centre has increased in a number of increments over time such that it now occupies over 1000 m^2 of greenhouse space. Efficient maintenance is achieved through the use of hydroponic watering systems. Recently the greenhouses have been fitted with moveable shade screens that prevent excessive light entering the growing environment and also have an energy saving value.

The quarantine procedures operated by the ICQC,R are in accordance with international guidelines (Frison *et al*., 1999), which are currently in the process of being updated (Engels *et al*., 2009). The procedures include a virus indexing process that lasts for two years. ICQC,R is also working towards supplying a Standard Material Transfer Agreement (SMTA) with each shipment of material. This will help establish the access and benefit conditions as laid down in the International Treaty on Plant Genetic Resources for Food and Agriculture (IT) and means that the centre will operate within the guidelines of the Multilateral System [ftp://ftp.fao.org/ag/agp/planttreaty/agreements/smta/SMTAe.pdf].

Impact of the ICQC,R

The ICQC,R currently holds around 300 different cocoa genotypes with approximately a further 100 passing through the quarantine process. The full list can be found at 'http://www.icgd.reading.ac.uk/accession.php'. A particular genotype is held for as long as there is a demand for it.

The ICQC,R has distributed cocoa genotypes (usually in the form of budwood cuttings) to over 20 institutes in cocoa growing countries. These include all the major producing countries, such as Côte d'Ivoire, Ghana, Indonesia and Brazil, as well as emerging countries, such as Vietnam and Australia. The ICQC,R, along with CIRAD (Centre de Coopération Internationale en Recherche Agronomique pour le Développement, France), also provided material for the CFC/ICCO/IPGRI Clonal Trial (Eskes, 2001); a trial replicated in 10 different countries.

In a survey of recipients of germplasm from ICQC,R, of the 14 respondents, 7 cited breeding as a use for material received and 12 of the institutes said that they used the material for local evaluations and general enrichment of available genetic resources.

Prioritisation of Genetic Resources

Most new material imported into the ICQC,R comes from ICG,T, although material has also been received from other genebanks, including IC3. In recent years, the CFC/ICCO/IPGRI project collection (Sounigo *et al*., 2006) has provided a focus for new material being imported into the ICQC,R. Furthermore, cocoa genotypes at CRU that have been selected/screened for black pod resistance and

witches' broom disease respectively have also been imported. In order to guide future imports, germplasm recipients were asked to score the importance of particular traits and to list specific clones or groups that are not currently held in ICQC,R and that they might be interested in obtaining. Traits which were universally considered important included high yield potential, resistance to black pod and bean quality traits (Table 1).

Table 1. The top ranked cocoa character traits in a survey of recipients of germplasm from ICQC,R.

Trait	Average score *
Resistance to black pod	4.8
High yield potential	4.6
Improving genetic diversity	4.6
Large bean size	4.3
High bean lipid content	4.0

* Respondents were asked to score traits on a scale of 1 to 5, with 5 being the most desirable.

THE INTERNATIONAL COCOA GERMPLASM DATABASE

Background

ICGD is the primary source of information on the characteristics, origin and location of cocoa genetic resources worldwide, and is about to celebrate its 20th anniversary. The database aims to support the continuing breeding efforts required to maintain sustainable cocoa production in the face of increasing pressures from pest and diseases, current low yields, and the uncertainties posed by global climate change, whilst addressing the increasing need for environmental and social responsibility. ICGD provides the information needed by cocoa breeders and is widely used by national research institutes to select material for breeding programs that supply planting material for local farmers. Since 1989, the ICGD has been available to research institutes and genebanks around the world free of charge. Initially distributed on floppy disk and then CD-ROM (e.g. Wadsworth *et al*., 2004), the database has also been produced in printed form (e.g. Wadsworth *et al*., 1997). An online version (www.icgd.reading.ac.uk) was first released in 2002 and is now the primary focal point of ICGD's development, though a CD-ROM version will continue to be available for the foreseeable future. The contribution of data and input into the development of the database by cocoa research institutes and individuals has added considerably to the success and longevity of the project.

Developing ICGD as an online database

Recently, the majority of data from the original offline dBASE database have been transferred to the new online MySQL system. This process has included a complete restructuring of the database into a small number of trait-based tables, which makes updating and searching the database quicker and simpler than before.

As part of the reprogramming, a considerable effort has been made to improve the user interface of ICGD and make it more intuitive. Although relatively complex searches can already be made directly from the search page, the option to search a subset of clones from a previous search allows more complex searches to be built up. Several improvements have been made to the clone name searching modules, which take into account the renaming of mislabelled accessions (Turnbull *et al*., 2004) as well as synonyms, and a name checking facility has also been added. The information for each clone is displayed by character type and traits are grouped by reference, with relevant data displayed together.

JavaScript has been incorporated into some of the search and display routines, allowing the visible text on these pages to be better targeted to each user's requirements. For example, the dynamic and more intuitive trait-based search options available on the main search page remove the need to display the enormous list of every possible search option. The output display for clone details also includes the option to expand and contract the data presented on a page, with information on the traits used in the search expanded and other available data minimised by default. This enables a simple and efficient appearance of both the input and output screens.

Continued provision of a CD-ROM version of ICGD

The MicroWeb program (http://www.indigostar.com/microweb.htm) is now used to allow the website to be run locally on a PC, using a CD-ROM, hard disk drive or flash memory device (e.g. USB memory stick). This has allowed the development of the CD-ROM to run in parallel with ICGD Online instead of requiring separate programming. Beta versions of the CD-ROM have been successfully tested on a wide range of computers of different ages and speeds, running Windows 95 to Vista, and using the most common web browsers (Internet Explorer, Firefox, Opera and Safari).

Collaborations

Restructuring the database, transferring it to MySQL and hosting the main version online (using PHP scripting language), means that ICGD is much better suited to involvement in international collaborations and networks than in the past.

CocoaGen DB: The University of Reading has continued its close collaboration with CIRAD and USDA (United States Department of Agriculture) on CocoaGenDB, which combines the phenotypic information in ICGD with the genetic/genomic data in CIRAD's TropGENE database. A web services approach is being developed to allow complex searches of the information contained in both parent databases to continue, but without the need for CocoaGenDB to be maintained as a separate database that duplicates the original datasets.

CacaoNet: The participation of ICGD in CacaoNet has developed over the last few years, with the database becoming a key resource in the development and implementation of the CacaoNet Global Strategy for the Conservation and Utilization of Cocoa Genetic Resources (Engels *et al*., 2009). CacaoNet aims to create a 'virtual' genebank system, where accessions internationally agreed as having the highest priority for long term conservation and/or immediate value to breeders will be managed centrally under CacaoNet, though they will be physically located around the world in existing national and international genebanks.

Already an important feature of ICGD, the location of accessions will be a vital component of CacaoNet, both in terms of developing a conservation strategy based on existing collections and the future management of the 'virtual global collection'. ICGD is also one of the central components of the Information Management System (IMS) being developed in collaboration with Bioversity International and CIRAD, which is required to link specific, high quality data to the individual accessions that will make up CacaoNet's large Global Strategic Base and Active Collections (Engels *et al*., 2009). Web services will enable ICGD and CocoaGenDB to integrate into the CacaoNet IMS very efficiently, reducing the data collation and programming effort while ensuring the component databases maintain their integrity.

A database for the accessions in the CacaoNet collection, named CANGIS (CacaoNet Germplasm Information System), has been developed in collaboration with Bioversity International (Montpellier). The prototype version is using information from ICGD on International Clone Trial (ICT) accessions (Sounigo *et al*., 2006) held in the international collections maintained by CRU and CATIE. The system is modelled on the new version of MGIS (*Musa* Germplasm Information System) that is currently being developed, but uses the minimum set of cocoa descriptors set out in the CFC/ICCO/IPGRI project 'Cocoa Germplasm Utilization and Conservation, a Global Approach' (Bekele and Butler, 2000).

Germplasm Management Tool

The Online Germplasm Management Tool was initially a small project designed to help the ICQC,R prioritise the accessions held in the post-quarantine collection for replication and distribution using data held in the ICGD. However, it was realised that the flexible and dynamic web-based management tool that was created had a much broader appeal (Turnbull *et al*., 2005). Following further development, the program can impartially and unambiguously rank accessions, based on information on the traits individuals possess and the genetic diversity they represent. The program is flexible in the number of accessions and traits that can be input and weighting of individual traits is user defined. This allows users to generate lists of accessions which are prioritised according to their needs, whether they are seeking material with complementary resistance mechanisms to a particular disease, large beans with high fat content, material that will augment the genetic diversity they already hold in their genebank, or any combination of these and other traits.

Future Plans

Close collaboration with CacaoNet and CIRAD will continue, with further programming required to develop web services in ICGD for CANGIS and CocoaGenDB, though developing these links between databases will ensure information is easily accessible with the minimum amount of duplication of both development time and data.

When restructuring and basic programming have been finished, work will concentrate on requesting and entering new data (including that required for CacaoNet). The continued provision of information will be addressed by developing a collaborative network of ICGD/CacaoNet data providers. Those researchers, breeders and genebank curators that are actively characterising and evaluating cocoa clones will be invited to join the network, with the aim of making the process of sharing information as simple as possible and encouraging people to do so.

The ICGD aims to make user's requirements its priority and would welcome any comments or suggestions, which can be sent to the address given at the top of this paper or e-mailed to 'icgd@reading.ac.uk'. These addresses can also be used to send any data (including accessions lists, morphological characteristics or agronomic studies) that organisations or individuals may wish to submit; all entries are fully referenced to their original authors.

CONCLUDING REMARKS

The University of Reading has a long history in the field of cocoa genetic resources, distributing disease-free material to breeders, genebank curators and researchers from the ICQC,R and providing free access to information on the characteristics, origin and location of cocoa germplasm resources worldwide through ICGD.

The conservation and utilisation of Cocoa Genetic Resources are vital components of a sustainable global cocoa economy and both projects continue to look towards the future, following global trends being adopted by the cocoa community, as well as other crop groups. ICQC,R has started to take a more pro-active approach to selecting accessions to be maintained in the post-quarantine collection and is moving towards including an SMTA with each consignment that is sent out. Although still available on CD-ROM, ICGD is now a predominately online database, exploiting widely-used, freely-available software to produce a user-friendly and intuitive interface. These developments also present an opportunity to develop web services, allowing information to be more easily integrated with that of other databases.

ICQC,R and ICGD both have key roles in international projects and networks, such as the CFC/ICCO/Bioversity project and CacaoNet, which will help breeders and researchers to meet the future demands of cocoa production for higher yields, increased bean quality and improved resistance to pests and diseases, while addressing the uncertainties of global climate change and the need for greater environmental and social responsibilities.

ACKNOWLEDGEMENTS

ICGD would like to thank all those institutions and people who have helped the project: those who have supplied information, who have advised us, and who have made comments and suggestions that have helped us to improve the database.

The ICGD project is supported by NYSE Liffe and administered by the Cocoa Research Association Ltd. (CRA Ltd.); a UK-based organisation managing scientific cocoa research on behalf of Cadbury, Mars, Kraft Foods and members of NYSE Liffe.

ICQC,R would like to thank all the institutes that participated in the survey on prioritisation and utilisation of material received from Reading.

ICQC,R is funded by CRA Ltd. and the USDA.

ICQC,R would like to acknowledge the financial support of the Netherlands Ministry of Agriculture, Nature and Food Quality (LNV) for the energy saving measures.

REFERENCES

BEKELE, F. and BUTLER, D.R. (2000). Proposed list of cocoa descriptors for characterisation. In: Eskes, A.B., Engels, J.M.M. and Lass, R.A. (editors) (2000). Working procedures for cocoa germplasm evaluation and selection. Proceedings of the CFC/ICCO/IPGRI Project Workshop, 1-6 February 1998, Montpellier, France. International Plant Genetic Resources Institute, Rome, Italy. pp. 41-48

ENGELS, J.M.M. , END, M.J., TURNBULL, C.J. , DAYMOND, A.J , CRYER, N.C. , RUAS, M. and ROUARD, M. (2009). A global strategy for the conservation and utilization of cacao genetic resources. Proceedings of the 16th International Cocoa Research Conference. In Press.

ESKES, A. B. (2001). Collaborative activities on cocoa germplasm utilization and conservation supported by the CFC/ICCO/IPGRI project. 13th International Cocoa Research Conference, Kota Kinabalu, Malaysia. pp 23-30.

FRISON, E.A., DIEKMANN M. and NOWELL, D. (1999). FAO/IPGRI Technical Guidelines for the Safe Movement of Germplasm. No. 20. Cacao (1st revision). Food and Agriculture Organization of the United Nations, Rome/International Plant Genetic Resources Institute, Rome.

SOUNIGO, O., BEKELE, F.L., IWARO, A.D., THÉVENIN, J.-M., BIDAISEE, G., UMAHARAN, R., SANKAR, D., MOTILAL, L., BUTLER, D.R. and ESKES, A.B. (2006). Description of cocoa clones proposed for the "CFC/ICCO/IPGRI Project Collection". In: Eskes, A.B. and Efron, Y. (editors) (2006). Global Approaches to Cocoa Germplasm Utilization and Conservation. Final report of the CFC/ICCO/IPGRI project on “Cocoa Germplasm Utilization and Conservation: a Global Approach” (1998-2004). CFC, Amsterdam, The Netherlands/ICCO, London, UK/IPGRI, Rome, Italy. pp 67-81.

TURNBULL, C.J., DAYMOND, A.J., CRYER, N.C., HADLEY, P. and WILKINSON, M.J. (2005). A Management Tool for the Prioritisation of Cocoa Germplasm Resources. Proceedings of the Malaysian International Cocoa Conference, 18-19 July.

TURNBULL, C.J., BUTLER, D.R., CRYER, N.C., ZHANG, D., LANAUD, C., DAYMOND, A.J., FORD, C.S., WILKINSON, M.J. and HADLEY, P. (2004). Tackling Mislabelling in Cocoa Germplasm Collections. INGENIC Newsletter, 9: pp 8-11.

WADSWORTH, R.M., FORD, C.S., TURNBULL, C.J. and HADLEY, P. (2004). ICGD 2003: International Cocoa Germplasm Database Version 5.2. CD-ROM. Euronext.liffe/University of Reading.

WADSWORTH, R.M., FORD, C.S., END, M.J. and HADLEY, P. (1997). International Cocoa Germplasm Database, Printed Version (Ed. 2). The London International Financial Futures and Options Exchange (LIFFE) /The University of Reading, U.K.