

The International Cocoa Germplasm Database

Michelle J. End, R.M. Wadsworth & P. Hadley

School of Plant Sciences, University of Reading,
Whiteknights, P.O. Box 221, Reading RG6 2AS, UK.

The International Cocoa Germplasm Database (ICGD) has been compiled to make information on the origins, present location and characteristics of cocoa clones more widely available. The database is supplied in the form of a self-contained package for use on IBM PC and compatible computers. A printed version containing tables of information extracted from the database is now available which will be particularly useful to those who do not have access to a suitable computer. The importance of building and maintaining good links with cocoa genebanks and research centres is emphasized.

Introduction

Information on the origins and characteristics of accessions is essential for both the efficient management and utilization of germplasm resources. Such information is needed to avoid problems of misidentification and duplication in genebanks and to target future collecting expeditions by identifying areas from which little material has been collected and areas where there is high genetic variability. This information is also important for the utilization of germplasm since it can assist research centres in deciding which clones should be added to their collection for evaluation, experimentation and inclusion in breeding programmes.

Although a considerable amount of information on cocoa germplasm exists, it is fragmented, much of it has had a very limited distribution and some has not been available outside the research centre from where it was collected. The International Cocoa Germplasm Database (ICGD) has been compiled to make this information more widely available to all cocoa research centres. Since the database contains information primarily on one species, *Theobroma cacao*, with limited data for other *Theobroma* and *Herrania* species, it is relatively small and can be managed using software developed for Personal Computers. This is in contrast to many of the large germplasm information databases such as GRIN (Genetic Resources Information Network) (Perry, 1988) which are managed using centralized mainframe and minicomputers and rely on telecommunication links to exchange information with research centres. The ICGD is supplied to research centres in the form of a self-contained package for use on IBM PC and compatible computers. This package only requires a Personal Computer and does not need additional software or telecommunication links and so can be used by research centres in developing countries where computing and telecommunication facilities may be limited. A printed version is now available which contains tables of information extracted from the database. This printed version will compliment the computerized version and will be particularly useful to those who do not have access to a suitable computer. The ICGD project also offers an 'information service' and can produce customized reports and searches if they are required.

Information Contained in the ICGD

Over 10,000 clone names have been encountered during the compilation of the database and programs have been designed so that synonyms and alternative formats of clone names can be recognized. The ICGD contains information on the origins, present location and characteristics of cocoa clones (Table 1). The information has been compiled from a range of sources which are acknowledged in the Bibliography. Data from the ICGD suggest that over 8800 different cocoa clones exist in genebanks around the world (Table 2) with information on

Table 1 The types of information that can be found in the ICGD

Clone Names	Information on clone names and synonyms encountered including the meaning and/or origin of names.	
Origins	Wild Collections: details of the collection site of each accession taken from the wild.	
	Selections: details of the origins of selections and hybrids.	
Present Location	Information on the clones held in germplasm collections.	
Morphological, Agronomic and Quality Characteristics	BAL Plantations, Malaysia	butter fat content. 126 clones [F4] ^I
	Brazilian/International Expedition to Brazil, 1965	field measurements. 55 accessions [M14, D1, A1]
	CATIE, Costa Rica	two sources 1. 372 clones [M42, B1] 2. 65 clones [M58, B1, F1, D5]
	CEPEC, Brazil	three sources 1. 564 clones [B1, M4]. 2. SIC and SIAL accessions. 26 clones [M26] 3. field measurements of RO and AM (CAB) accessions. 115 clones [M18]
	IBTA/IBPGR Expedition to Bolivia, 1987	field measurements. 22 clones [M5]
	ICGT, Trinidad	two sources 1. leaf and flower characteristics. 247 clones [M32] 2. fruit characteristics. 38 clones [M31]
	IRCC/CIRAD, Guyane	field measurements of GU accessions. 248 clones [M13]
	LCTAP, Ecuador	field measurements of LCT-EEN and EBC accessions. 356 clones [M29]
	MAC, Caucagua, Venezuela	two sources 1. 60 clones [A8, B4, D7, M25] 2. cocoa butter characteristics. 13 clones [A4, F3]
	Prang Besar, Malaysia	two sources. PBC accessions 1. 10 clones [A10, D3, F1, M13] 2. 14 clones [A1, D3, F1, M6]
	Stress and Disease Reactions	results from various collections. Information mainly from published review articles. 318 clones [D7].
Bibliography	Contains the full reference for all the sources of information used in the compilation of the ICGD.	

* Information is given separately for each data set and unless otherwise stated, measurements were made on material growing at the research station.

^I Indicates the number of the descriptors recorded in each of the following categories; A = agronomic characteristics, B = breeding information, D = disease reaction, F = fat and quality characteristics, M = morphological descriptors.

Table 2 The germplasm collections for which a list of accessions is included in the ICGD.

	Germplasm Collection	Total No. of Clones	Unique Clones	Clones existing in 1 or 2 other collections
Inter-national Genebanks	CATIE, Costa Rica (1992)	730	355	130
	ICGT, Trinidad (1992)	2451	1854	293
Quarantine Stations	France - IRCC Montpellier (1989)	79	37	40
	UK - University of Reading (1992)	119	3	37
	USA - USDA Mayaguez (1992)	306	56	59
	USA - USDA Miami (1992)	179	30	36
South America	Brazil - CEPEC (1990)	560	240	97
	Brazil - CEPLAC (1987)	872	776	48
	Colombia - CNI Palmira (1987)	418	295	41
	Ecuador - EE Napo, San Carlos (1990)	253	194	53
	Ecuador - EET Pichilingue (1987)	465	263	77
	French Guiana - IRCC/CIRAD Combi (1992)	365	303	14
	Surinam - Agricultural Experiment Station (1971)	9	0	0
	Venezuela - MAC Caucagua (1973) ¹	164	59	21
Central America and the Caribbean	Grenada - Ministry of Agriculture, Lands & Fisheries (1971)	4	0	0
	Guatemala - EE Los Brillantes (1989)	116	18	31
	Honduras - CEDEC, La Masica (1992)	137	19	41
	Jamaica - Cocoa Industry Board (1971)	11	0	1
West Africa	Cameroon - IRCC (1985)	200	139	25
	Cameroon - N'koomvone Research Station (1971)	6	0	0
	Gabon - SOCAGAB (1992)	35	0	15
	Ghana - Cocoa Research Institute (1992)	743	237	190
	Ivory Coast - IRCC Bingerville (1989)	670	513	60
	Nigeria - Cocoa Research Institute (1971)	124	15	51
	Togo - IRCC Togo (1991)	153	76	30
Asia	India - CPCRI, Kerala (1992)	99	61	15
	India - Kerala Agricultural University (1991)	140	111	12
	Malaysia - Agricultural Research Centre, Sabah (1990)	608	129	190
	Malaysia - BAL Plantations Sdn. Bhd. (1991)	134	5	6
	Malaysia - Felda (1991)	287	15	43
	Malaysia - Golden Hope Plantations (1991)	301	26	75
	Malaysia - MARDI (1989) ¹	12	1	2
	Malaysia - United Plantations (1991)	192	1	45
	Sri Lanka - CARI (1971)	13	1	0
	Thailand - HRCC (1988)	22	2	0
Oceania	Fiji (1992)	40	14	5
Total		11017	5848	1783

¹ Updated clone list should be available in the near future

characteristics available for approximately 20% of these clones. Many research centres, which have programmes to characterize and evaluate their germplasm, are continuing to supply data to the ICGD as they become available. It is believed that further standardisation of the methods used in these characterization and evaluation studies will make this information even more valuable to germplasm managers and plant breeders.

The ICGD has been designed to allow users to access information in a variety of ways. Some of the types of questions that can be answered using the ICGD include:

What does the clone name P stand for, does it have any synonyms and have more than one group of clones been given this name?

Was clone LCT-EEN 46 collected from the wild or selected in a breeding programme, and what is known of its history?

What material has survived from the Anglo-Colombian Expedition in 1952 and has any other material been collected from that region?

What is known about the characteristics of the EET clones?

Are there any clones which have been recorded as having large, white seeds and resistance to Black Pod disease?

Which clones are present in the genebank at CATIE, Costa Rica which are not held in the International Cocoa Genebank, Trinidad?

Development of Links with Genebanks and Research Centres

It is hoped that the ICGD will continue to provide useful information for cocoa research workers and growers. It is therefore essential that good links are established between the ICGD and the research centres to ensure that the ICGD meets their needs and contains accurate and up-to-date information. Research centres have a central role in the supply of information to the ICGD, not only in terms of clone lists and characterization data, but also in relation to their extensive knowledge of the germplasm. Such centres can also use the information contained in the ICGD to support their research activities. Version 2.1 of the ICGD, released in December 1991, has been distributed to over 80 organizations in 27 countries. Many of these organizations have responded by sending current accession lists and other data for inclusion in the database (Table 2).

Currently much of the information contained in the ICGD has been extracted from published reports and catalogues, but it is hoped that in the future much of the data will be supplied on computer diskette. This will facilitate the production of updates, enable automated data checking and minimize the introduction of errors. Although data can be extracted from various types of database, word processor or spreadsheet files, data is best exchanged in a standard format. In consultation with the genebanks and research centres, it is hoped that data can be exchanged in standard Dbase compatible files or in the forms being agreed by the International Union of Biological Societies Commission for Plant Taxonomic Databases. The ICGD should have a role in publicizing and encouraging the adoption of these standards.

A network of regional coordinators for the ICGD is established, which will have an increasingly important role in the future development of the ICGD. Regional coordinators will encourage research centres in their area to collect and donate information to the database and report back on how the ICGD is being used and how it can be improved for the users in their region.

Future Development of the ICGD

We have already received useful comments on how the database can be improved to better serve the needs of the users both in terms of the data it contains and in the way the information is accessed. We are continuing to improve the efficiency and performance of the database and the data that have been included. The ICGD has been designed so that most types of information can be incorporated. In this respect, it is similar to the International Legume Database and Information Service (ILDIS) which has been designed so that modules containing information from areas such as phytochemistry, biotechnology and germplasm resources can be attached to the central taxonomic database (Zarucchi *et al.*, 1993). The next version of the ICGD will include data from isozyme and RAPD studies, disease and pest screening experiments and yield potential trials, and other sorts of data are expected.

Information management software is developing rapidly and new hardware is becoming available, which is capable of handling different types of information, including images, more efficiently. For example, computerized floristic databases incorporating images of herbarium specimens and living plants have been produced to provide support for research at local institutions in the tropics (Gomez-Pompa, 1993). We recognize the potential value of such products in making cocoa germplasm information more accessible. The ICGD will continue to develop with guidance from the Regional Coordinators, taking into account not only the equipment of potential users, but also the information available at present and in the future.

References

- Gomez-Pompa, A. (1993) A futuristic view of floristic databases in the tropics, In *Proc. Designs for a Global Plant Species Information System Symposium*, Delphi, 1990", Systematics Association Special Volume, Oxford University Press (in press)
- Perry, M. (1988) The Germplasm Resources Information Network. A genetic resources information management system for the United States, *FAO/IBPGR Plant Genetic Resources Newsletter* **73/74** 21-26
- Zarucchi, J.L, Winfield, P.J., Polhill, R.M., Hollis, S., Bisby, F.A. and Allkin, R. (1993) The ILDIS project on the world's legume species diversity, In *Proc. Designs for a Global Plant Species Information System Symposium*, Delphi, 1990, Systematics Association Special Volume, Oxford University Press (in press)