



Hunt Institute for Botanical Documentation
5th Floor, Hunt Library
Carnegie Mellon University
4909 Frew Street
Pittsburgh, PA 15213-3890
Telephone: 412-268-2434
Email: huntinst@andrew.cmu.edu
Web site: www.huntbotanical.org

The Hunt Institute is committed to making its collections accessible for research. We are pleased to offer this digitized item.

Usage guidelines

We have provided this low-resolution, digitized version for research purposes. To inquire about publishing any images from this item, please contact the Institute.

Statement on harmful and offensive content

The Hunt Institute Archives contains hundreds of thousands of pages of historical content, writing and images, created by thousands of individuals connected to the botanical sciences. Due to the wide range of time and social context in which these materials were created, some of the collections contain material that reflect outdated, biased, offensive and possibly violent views, opinions and actions. The Hunt Institute for Botanical Documentation does not endorse the views expressed in these materials, which are inconsistent with our dedication to creating an inclusive, accessible and anti-discriminatory research environment. Archival records are historical documents, and the Hunt Institute keeps such records unaltered to maintain their integrity and to foster accountability for the actions and views of the collections' creators.

Many of the historical collections in the Hunt Institute Archives contain personal correspondence, notes, recollections and opinions, which may contain language, ideas or stereotypes that are offensive or harmful to others. These collections are maintained as records of the individuals involved and do not reflect the views or values of the Hunt Institute for Botanical Documentation or those of Carnegie Mellon University.

About the Institute

The Hunt Institute for Botanical Documentation, a research division of Carnegie Mellon University, specializes in the history of botany and all aspects of plant science and serves the international scientific community through research and documentation. To this end, the Institute acquires and maintains authoritative collections of books, plant images, manuscripts, portraits and data files, and provides publications and other modes of information service. The Institute meets the reference needs of botanists, biologists, historians, conservationists, librarians, bibliographers and the public at large, especially those concerned with any aspect of the North American flora.

Hunt Institute was dedicated in 1961 as the Rachel McMasters Miller Hunt Botanical Library, an international center for bibliographical research and service in the interests of botany and horticulture, as well as a center for the study of all aspects of the history of the plant sciences. By 1971 the Library's activities had so diversified that the name was changed to Hunt Institute for Botanical Documentation. Growth in collections and research projects led to the establishment of four programmatic departments: Archives, Art, Bibliography and the Library.

Doc. dr WŁADYSŁAW KULPA

Instytut Hodowli Ras
i Nasilenności WSR

Lublin, ul. Akademicka 15, tel. 300-21
Polska — Poland

Institute of Plant Breeding,
College of Agric.

LEE KEHAYAS

director

STEVENSON & KELLOGG LTD.

management consultants

150 Eglinton Ave. East, Toronto, Canada, 483-4313

Colloquium
On
Plant Genetic Resources

Prof. D.J. ROGERS

14—15 APRIL 1972 IZMIR—TURKEY

DIPLOMGÄRTNER
LOTHAR SEIDEWITZ
DOCUMENTATION OFFICER

ARBEITSGRUPPE GENBANK
WORKING GROUP GENETIC RESOURCES CENTER
INSTITUT FÜR PFLANZENBAU UND SAATGUTFORSCHUNG

D-33 BHAUNNSCHWEIG-FAL
F.R. GERMANY

BUNDESALLEE 50
TELEFON 6531/596-1

DOCTOR
ING DUMITRU TEACI

- SEF SECTIE -

ROMANIA - BUCURESTI IV
STR. C-TIN NOTTARA, 7
TEL. 61.15.46

INSTITUTUL DE CERCETARI
PENTRU PEDOLOGIE
TEL. 17.21.80

K. Z. BUDIN
Dr. Sci., Prof.

N. I. Vavilov All-Union
Institute of Plant Industry

Herzen Str. 44
Leningrad, USSR

THE CONCEPTS AND APPLICABILITY OF TAXIR FOR STORAGE AND RETRIEVAL OF GENE
BANK DATA

Outline of presentation by David J. Rogers at EUCARPIA Meeting

IZMIR, TURKEY

April 11, 1972

"The taxonomist should be in a position to provide a system in which all further information can be accommodated and which has a high predictive value. Such a system should act as an efficient data storage and retrieval framework."

J. G. Hawkes in Frankel and Bennett, Genetic Resources in Plants--Their Exploration and Conservation, Blackwell, 1970, page 83.

1. The above quotation is prophetic. The same spirit guided the development of TAXIR--TAXonomic Information Retrieval.
2. The portions of taxonomy that serve as a model.
 - A. The typical reporting methods of taxonomy as an effective information retrieval system.
 - B. Taxonomic terms reinterpreted for information retrieval purposes.
 - 1) Key: A binary searching and identification routine.
 - 2) Taxonomic Character: A single basis for description, defined over the set of objects to be described. (We use the word descriptor to mean the same as character).
 - 3) Character State: For the individual specimen being described, a precise term is given. This becomes a descriptor state.
 - 4) Taxon: An item forming the basis for description.
 - 5) Species (Taxon) Description: A condensed abstract of knowledge (using descriptors) about an item.
 - 6) Hierarchical Classification: A method of organizing data banks with varying levels of detailed description.
 - C. Data, when interpreted (or placed into a specific context) becomes information.
 - D. TAXIR will process either data or information.

3. Description of the TAXIR System, using an example of an actual gene bank situation. The Plant Introduction Station, USDA, Pullman, Washington, Dr. Sam Dietz, Director, for data on Phaseolus species accessioned (approximately 6,000 accessions).

FIRST STAGE--Pre-TAXIR Activities

Director of gene bank wishes to install a system to store and retrieve data on accessions of genetic materials.

QUESTION: Can we afford it?--See attached comparison of manual versus TAXIR Systems.

QUESTION: How many accessions in bank at present?

QUESTION: What is yearly average accession rate?

QUESTION: Is a new program of exploration and introduction anticipated?

QUESTION: If a new system is to be installed requiring computing facilities, how do we arrange for them?

QUESTION: How much descriptive information (how many descriptors?) necessary for crop material? Refer to Professor Konzak's Standardized Format for Small Grains.

Assume that all above questions answered satisfactorily.

SECOND STAGE--Pre-TAXIR Activities

Development of necessary descriptive information about accessions.

1. Develop descriptor list: The following is an extracted example:

A. Identification of each accession

- a. Introduction number
- b. Scientific name
- c. Common name(s)
- d. Cultivar name
- e. Date of initial introduction
- f. Date of sending out propagation materials to specialists., etc.

B. Locating information

- a. Country of origin
- b. State/Province of origin
- c. Precise locality
- d. Collector
- e. Collector number
- f. Date of collection
- g. Type of propagule, etc.

C. Morphological descriptors

- a. Habit
- b. Stems
- c. Leaves
- d. Flowers
- e. Fruits
- etc.

D. Genetic/plant breeding data

- a. Appropriate descriptors for crop, etc.

E. Cultivation requirements

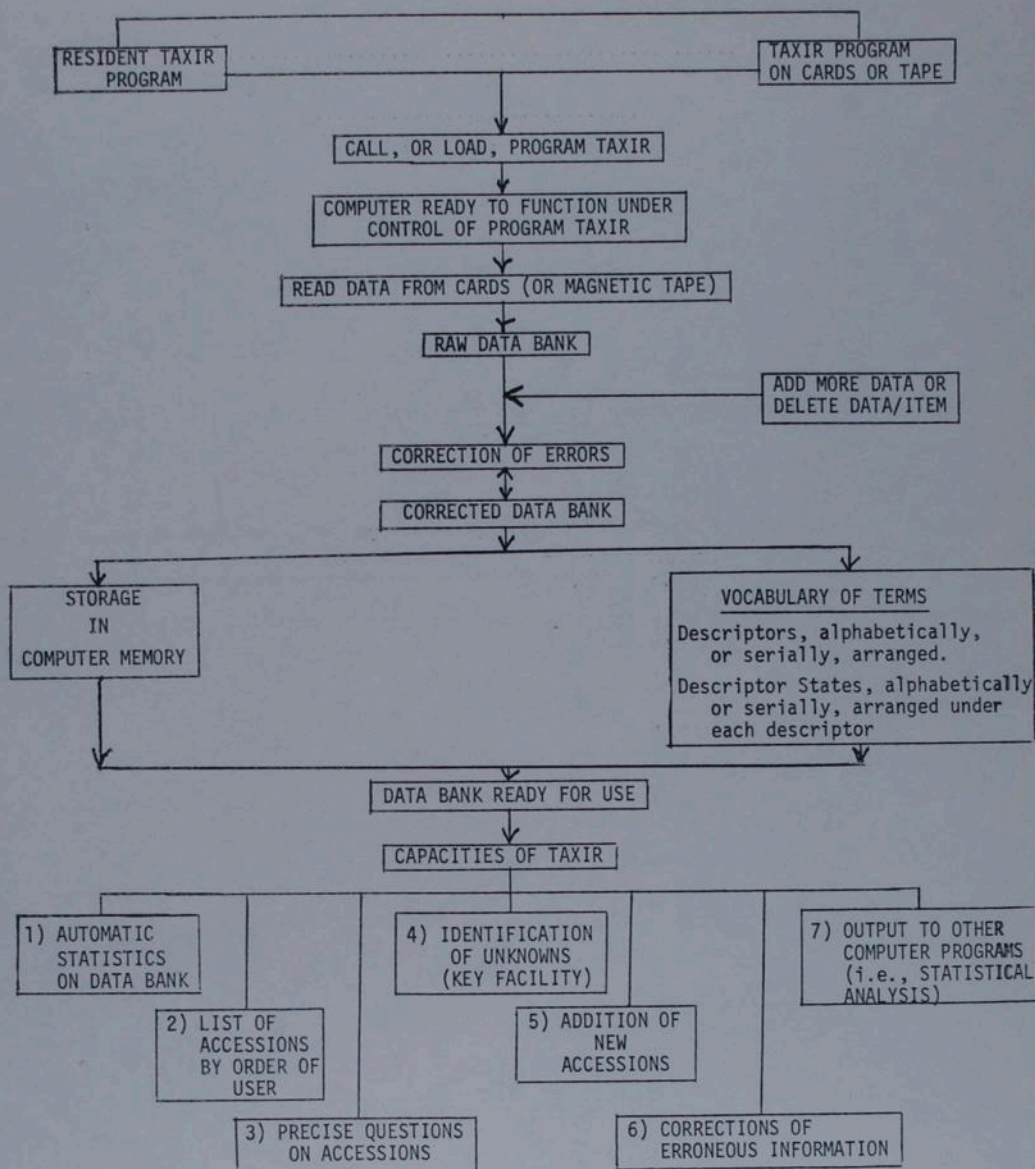
- a. Appropriate descriptors for crop, etc.

NOTE: Each lower case letter above indicates one descriptor.

THIRD STAGE--Data Capturing (or preparation)

1. Discover most efficient data gathering procedures.
 - A. Preprinted, tabular format?
 - B. Prepunched (Port-A-Punch) IBM cards?
 - C. Standardized notebook.
2. Have accession numbers been assigned to individual items in gene bank?
3. Are the individual accessions to be the items described?
4. Are data-gathering personnel trained in descriptive methods?
5. Has descriptor list been checked with TAXIR specialist.
6. Have facilities to key-punch data on IBM cards been established?
 - A. Rental basis
 - B. Use of other organizational facilities.
7. Are TAXIR computer programs installed in resident or rented computers?

FOURTH STAGE--TAXIR Processing



- 1) Automatic Statistics on Data Bank: The system can printout statistics on accessions such as total items in the data bank, number of accessions from a given geographical area, number of accessions belonging to a given taxon, etc., etc.

- 2) Lists of Accessions by Order of User: The system can generate catalogs, (or lists), employing data organization capability of TAXIR.

EXAMPLE: All the germ plasm accessions in the bank from a specified geographical area can be selectively retrieved from the bank using Boolean capability of the system, and the information on those accessions (such as accession number, taxonomic name, bio-chemical properties, etc., etc.) can be organized and printed out in any desired format.

- 3) Precise Questions on Accessions: Through the use of the selective retrieval capability of the TAXIR system, germ plasm possessing a specified combination of qualities can be searched and located.

EXAMPLE: The plant breeder desires to have germ plasm of Phaseolus species which gives white seeds with bushy plant habit adapted to warm dry climates.

- 4) Identification of Unknowns (Key Facility): The phenotypic data on the accessions within TAXIR bank allows taxonomic identification of the unknowns. The use of the TAXIR system as a simultaneous key obviates the difficulties imposed by the conventional dichotomous indented keys, in identifications.

- 5) Addition of New Accessions: The information bank remains open ended, and the TAXIR system can add new information on accessions already in the bank (e.g., use of a given donor parent in a plant breeding program which will generate more cytogenetic data) on that accession. This information can be easily incorporated into the TAXIR bank. New accessions may be added as information becomes available.

- 6) Corrections of Erroneous Information: The TAXIR system can be commanded to search and correct data in the bank. This is a type of editing routine.

- 7) Output to Other Computer Programs (i.e., Statistical Analyses): Output to other programs such as Chi Square Regression Analysis, Cluster Analysis, Character Analysis, etc.

SUMMARY OF TAXIR

The TAXIR (TAXonomic Information and Retrieval) program is a part of a series of functions, analytical and synthetic programs incorporated in the design of a communication and information system. It embraces a set of 18 independent control statements, which when used in a selected series, perform functions useful in structuring, archiving, organizing, manipulating, changing, and creating data to the desired specifications of the user. The usefulness of TAXIR as a module of a data management system may be expressed as follows:

- o TAXIR is organized and designed to take advantage of the logical capabilities of Boolean Algebra thus allowing for efficient use of intricate and involved inference and demonstration criteria.
- o TAXIR stores the data bank as basic characteristic codes or functions. This technique enables information to be compressed logarithmically into bit representations in the computer and also facilitates Boolean expressions of search criteria to be executed through calculations rather than traditional file searching. This method of storage and information retrieval is a great saving in computer memory and time.

TAXIR COMMANDS

The following brief discussion of TAXIR commands is intended to explain some of the data maintenance, conversion and executive features of the TAXIR computer program.

Any "thing" that may be detailed, particularized, or described, will be called an item with the understanding that a descriptor is a basis or a structure for comparing these items. Specific statements regarding any item are states assignable to a descriptor.

Descriptor states for each item are input sequentially, separated by commas, in any of the mediums (cards, magnetic tape, paper tapes, teletype, etc.) acceptable to computers. The command: DEFINE ITEMS will read in, store, order, etc., the raw data into the computer in a format suitable for manipulation. The input data are screened for certain types of errors of omission or commission and these items will be automatically excluded from the data bank. The user is advised of this for each exclusion. A print-out of the entire input bank is provided if the user desires.

COMMAND DEFINE DESCRIPTORS

This command with its accompanying definitions enables the user to design and build his information bank to his specifications. For instance, if a descriptor is specified as a "name" descriptor, the descriptor states are likely to be nouns or adjectives in some written language. TAXIR will construct the appropriate files for the dictionaries and the administration of these dictionaries (ordering, synonymy, etc.). If a descriptor is described as "coded", it will expect something of the nature of 1, 2, 3, etc., as input for a particular description with the instruction to permute and replace the 1, 2, 3, etc., by red, white, blue, etc. A "FROM-TO" type descriptor allows ranges of numbers to be specified such as all flocks of birds between 1,000 and 4,000, to temperature between 25°F and 25.8°F. The EQUALS facility enables the user to use the same dictionary for more than one descriptor.

The command HOW MANY simply prints out the count of the number of items that satisfy the query.

The command PRINT may be formulated so that a single item, all the items, and/or a series of complex restriction conditions employing the AND, OR, NOT and the parentheses of Boolean expressions will be printed out on any of a number of peripheral processors (printers, CRT tubes, etc.).

GENERATE is a TAXIR command that will printout on a magnetic tape the selected information for input in 15 or so statistical programs.

A series of commands are available for the maintenance of a data bank such as CORRECTION, DELETE STATE, DELETE ITEMS and DEFINE MORE DESCRIPTORS. It is also possible to create new descriptions from old information already input to the bank.

A number of other commands (a total of 18) are useful in data management such as MEMO, TIME, ETC.

A series of error messages (78 in number) are also incorporated in the program. These inform the user of errors of commitment on his part or of errors in the data and are useful in editing the information in the bank.

COMPARISON BETWEEN MANUAL DATA FILING SYSTEM AND
TAXIR SYSTEM

MANUAL

TAXIR

- | | | |
|--|-----------|---|
| <p>1. Data to be filed arranged under limited number of prechosen index headings:</p> <ul style="list-style-type: none"> 1. Alphabetic 2. Geographic 3. Taxonomic classification 4. Subject headings 5. Numerical | <p>←→</p> | <p>1. Data to be filed arranged under any number of index headings. Gene bank director and scientists choose all desirable headings. Additional file headings may be added at will. All files automatically cross-referenced.</p> |
| <p>} All require cross-referencing system.</p> | | |
| <p>2. Each file requires a separate physical file, thus, much duplication of each item added to the files.</p> | <p>←→</p> | <p>2. Each datum recorded but once, eliminating need for duplicate effort.</p> |
| <p>3. The separate files listed above cannot be ordered into one overall file.</p> | <p>←→</p> | <p>3. All files ordered together in one operation, all under one program command statement.</p> |
| <p>4. Great difficulty extracting additional data which does not fall into a major file heading.</p> | <p>←→</p> | <p>4. Each datum individually addressed not dependent on a major file heading.</p> |
| <p>5. Once established, reorganizing files into a different filing scheme is costly and time consuming.</p> | <p>←→</p> | <p>5. Any desired changes in file organization accomplished with one computer program instruction.</p> |
| <p>6. No possibility of guaranteeing precise and accurate responses when a question is asked.</p> | <p>←→</p> | <p>6. Automatic searching of data file extracts all pertinent responses, without missing any one.</p> |
| <p>7. Time to search complete file is prohibitive.</p> | <p>←→</p> | <p>7. Time to search complete file measured in seconds.</p> |
| <p>8. Little possibility of keeping statistics on file contents.</p> | <p>←→</p> | <p>8. Constant count of data items in file.</p> |
| <p>9. Difficult correction, deletion, or updating files.</p> | <p>←→</p> | <p>9. All corrections, deletions, updating of files made accurately and instantly with one program instruction for correction, one for deletion, and one for updating.</p> |

ID TAXIR *

SAMPLE QUESTION FROM PHASEOLUS GENE BANK, USDA REGIONAL PLANT INTRODUCTION STATION, PULLMAN WASHINGTON

QUERY FOR 809 MCMILLAN U OF FLA LIST ITEMS WITH SEED COLOR, B*

RESPONSE:

109860	136744	140311	150400	150404	150408	150413	150417	150957	152326
155213	163556	163566	163583	164616	164739	164777	165417	165419	165420
165422	165433	165434	165435	165453	165455	165461	165462	165466	167107
169730	169758	169767	172020	173041	173046	174908	175257	175261	175849
175859	177051	181784	181796	182983	186502	186504	186505	189011	189014
189414	189568	190070	190072	190073	190076	194573	194579	194530	194581
194582	194583	194587	195347	195354	195355	195361	195364	195368	195369
195370	195371	195374	195384	195391	195397	195398	195799	195802	196927
197446	197967	199038	200940	200947	200956	200959	200962	200967	200968
200969	200974	200999	201300	201324	201326	201333	201365	201378	201483
201493	201495	202081	203916	203918	203923	203924	203927	203958	204963
205361	207141	207175	207186	207198	207226	207227	207228	207229	207231
207233	207276	207283	207284	207299	207331	207372	207373	207391	207408
207409	207410	207422	207586	207832	208768	208769	209461	209465	209472
209475	209496	209621	209801	209811	212021	212024	224714	224717	224729
224743	226916	230364	251049	268108	268110	276327	278669	280655	281594
282101	282108	282117	287536	288013	288014	288015	288016	288017	288018
288019	288605	289339	289383	289409	289414	289422	289455	289470	298111
300655	300656	300657	300658	300660	300662	300663	300664	300665	300666
300668	300669	300672	300673	300674	300675	300676	300677	300679	300680
300681	300682	300685	302530	303639	304108	304109	304111	304112	304116
304117	306154	306155	306158	306159	306160	306161	306162	306163	306164
306165	306167	306171	207435	249553	269325	311175			

NO. OF ITEMS IN QUERY RESPONSE = 227

END*

ID TAXID *

SAMPLE QUESTION FROM PHASEOLUS GENE BANK, USDA REGIONAL PLANT INTRODUCTION STATION, PULMAN, WASHINGTON

QUERY FOR DAVID L ROGERS LIST ITEMS WITH PLANT HABIT, B AND POD LENGTH, FROM
5 TO 8 AND SEED COLOR, W*

RESPONSE:

137117	164926	169838	169843	169848	169869	169920	174996	179440	180732
180748	180751	180752	180754	180755	180759	180761	181919	197225	205213
206986	212025	212026	226931	226937	271994	271995	271997	272004	285693
289380	289423	289424	289426	289430	289435	289438	289441	289443	289450
289451	289464	289473	289475	302539					

NO. OF ITEMS IN QUERY RESPONSE = 45

SAMPLE QUESTION FROM PHASEOLUS GENE BANK, USDA REGIONAL PLANT INTRODUCTION STATION,
PULLMAN, WASHINGTON

ID TAXID *

QUERY PRINT: COUNTRY OF ORIGIN AND PLANT HABIT AND DAYS OF MATURITY FOR
ITEMS WHICH HAVE SEED COLOR, A*

RESPONSE:
ITEM NO.

REQUESTED STATES

109860 VNZ . SV. M

136744 CDA . R . M

140311 IRN . SV. UNKNOWN

*An easy way to pick up blanks is
our data*

150400 ELS . V . M

150404 ELS . SV. M

150408 ELS . SV. M

150413 ELS . R . M

150417 ELS . SV. M

150957 MEX . SV. L

152326 ECD . SV. M

155213 PRG . SV. UNKNOWN

163556 GTA . V . M

163566 GTA . R . M

163583 GTA . B . M

164616 IND . SV. M

Salem, Madras

164739 IND . SV. M

Mumbai, Bombay

164777 IND . SV. M

Punjab, India

165417 MEX . SV. M

165419 MEX . SV. M

165420 MEX . SV. M

165422 MEX . SV. M

165433 MEX . SV. M

165436 MEX . SV. M

165435 MEX . SV. M

165453 MEX . V . M

165455 MEX . SV. M

SAMPLE QUESTION FROM PHASEOLUS GENE BANK,
USDA REGIONAL PLANT INTRODUCTION STATION,
PULLMAN, WASHINGTON
ID 13314

QUERY HYLAND WEST MALAYSIA 4 27 71 PRINT: COUNTRY OF ORIGIN FOR ITEMS WITH
PLANT ERECT, S AND
SEED SIZE, W AND
DAYS OF MATURITY, E*

RESPONSE:
ITEM NO. REQUESTED STATES

128513	USA
136659	CEA
136724	CEA
136745	CDA
209806	KYA
223359	IRN
226931	GMN
226938	GMN
271958	SPN

NO. OF ITEMS IN QUERY RESPONSE = 9

END*

Computerisation of Grass Information in Kenya

Urban Gullberg

UNDP/FAO Range Management Project

Nairobi, Kenya

In the exploration and evaluation of genetic resources of grass in Kenya, one of the present activities is to computerise the existing information. FAO has therefore, in cooperation with the University of Nairobi set up an Information Storage and Retrieval System, considering and utilising the work done by other groups in this field. Some of the operations in this system will be briefly presented here.

The operations we think are of special interest are related to the input of information that already has been recorded. In our situation they are of primary importance since so much valuable information exists at a couple of the breeding stations in the country. In the processing of information one man functions as if he was a 'documentation man' at a genetic resources centre. He has easy access to an ICL-1900 computer, and he has, together with programmers, worked out the computer programs that are used. The raw data, however, has not been brought to him. It has been on hand written cards at the breeding station at Kitale, 400 km away from Nairobi.

As a basis for the computerisation work, we have set up a data base, or a format, that shows how the information related to one accession will be subdivided into descriptors. The data base we work with consists of descriptors relating to the identification, maintenance and origin of the accession. There also are descriptors relating to observations. The computer program we work with allows us to add or delete new descriptors to our data base. For the recording and storage of information, we have allocated a maximum length to each of the descriptors, the shortest ones corresponding to one alphanumeric character, the longest to more than 200 characters.

The computerisation has been made in three Steps. First the information from the breeding stations records has been transferred to a recording sheet. Secondly the sheets have been sent to the 'documentation Centre', where they have been checked and prepared for key-punching, and at last the sheets have been key punched so that the information can be transferred to magnetic tape.

To prepare the recording sheets that have been used in all three steps, the structure and maintenance of the original information at the breeding station was carefully checked. We found that information on one accession was recorded at up till five different places on Introduction cards, Field cards, Seed exchange cards, living nursery map and on herbaria specimens. The descriptors for which any of the accessions had information were taken out from our data base, and they were arranged so that the information easily could be transferred from one of the five sources at a time. This was possible to do, since all the descriptors in the data base have an identification number that goes with them and any subset of the data base could therefore function as a record and the descriptors could be put in any order. The computer will rearrange them to the order they have in the data base. The transfer of data to the recording sheet is kindly done by the staff at the breeding station in Kitale.

When all the information from one accession has been transferred from the records in the breeding station to the recording sheet, the sheets are sent to the 'documentation centre'. There the data is checked and transferred so that it is ready for key punching. In this step we code the information under some of the descriptors. This we do gradually that is, for each successive accession we check the scores one has recorded. That scores that have occurred on previous accessions are recorded on a set of vocabulary cards one for each descriptor, and if the scores on the

accession under treatment already are recorded, the code in the vocabulary is written at the appropriate place on the sheet. If the score not is recorded in the vocabulary a new code is given to it. This code is added to the vocabulary and it is recorded on the recording sheet.

This gradual building up of the vocabulary, has been considered the most efficient way of doing the scoring in our situation. We are not familiar enough with the material we record information on, to on before hand be able to make a comprehensive scoring of all our descriptors. We therefore take the new scores as they come, and when we have recorded a sufficient amount of accessions, we will check our vocabulary and if necessary change the scoring and the codes of the scores, so that we come in agreement with standardired scoring. The changes in codes and scores will be made by the computer.

When the recording sheets are checked and coded, they are given to the key puncher. The key punching is made on paper-tape, mainly since most of the incoming information is recorded under a small fraction of the descriptors that the whole data base constitutes.

The estimates of the time needed for the steps involved show, that the first two steps are much more time consuming than the key punching. One person transfers 1,000 to 2,000 characters from the original records to a key-punch-prepared sheet in one hour. Up till 8,000 characters can be key punched during the same time.

The Documentation System of the German Genebank and
its Relation to International Genebank Standardization

A new concept for international standardization of genebank documentation has been developed by co-operation of the German genebank at Braunschweig-Völkenrode and Professor Konzak of the Department of Agronomy, Washington State University, Pullman, USA. Principles of which are mainly based on former documentation forms developed by the joint JAEA-FAO division in 1966. Furthermore principles for documentation on genetic material developed for the European Common Market and for the European Breeder's Association (UPOV, Geneva) have been considered.

Standardization of genebank documentation is regarded to be based on a vocabulary that as a thesaurus has to be issued by an internationally recognized co-ordinating office. Corresponding discussion at the world's genebanks have to confirm that descriptors will be usable in one of the most convenient languages, while descriptors in different languages correspond with each other synonymously.

The degree of descriptors' expression is coded in a very simple and most commonly understandable form using figures from 1 to 9 wherever coding is advisable. Here figure 1 stands for the least and 9 for the strongest expression of a character.

Absolute figures should be employed in these cases where a classification expressing "very low" or "very high" can be misunderstood. Here the specification of a key location or circumstances for the observation of a character can be included.

./...

The thesaurus will comprise non-crop specific and crop specific descriptors. As non-crop specific descriptors those are understood that are used for identification and for the records on transfer and origin (collection; selection or bred accession; induced mutation). Crop specific descriptors will cover all kind of expressions describing an accession phenotypically (morphology, composition and quality, reaction to pests, diseases, environment) and genotypically as far as the latter is interesting.

Documents are thought to be composed due to individual needs and corresponding to the availability of information. In this way only the construction of data banks by aid of computers can be guaranteed. A set of data or a data bank can be placed on tapes in a space saving way. While only disposing of concentrated information sets a retrieval is consequently possible in a highly efficient and rapid way. ~~The intention is to~~ have a special layout any time a new data set has to be keypunched where the 80 column punch card serves as a porter of data. In the possible free field option the only restriction to be made refers to the sequence of information that has to be kept unchanged for an entire set of data.

The computer technical background for the plant genetic documentation and information system presented here is based on the so-called TAXIR system that was originally developed by the Colorado State University. It is operated on an IBM Computer 360/65 or 67 and its applicability on another computer system (German SIEMENS 4004) is actually under investigation.

L. SEIDEWITZ

HANDLING of INFORMATION on GENETIC RESOURCES
at the ITALIAN GERM-PLASM STATION

In order to record information in a standardized form a computer programme has been set up at the Bari Germ-plasm Laboratory. In preparing it, we had in mind the following conditions:

1. To establish an I/R system agronomic, ecological and genetic characters, which permits separate analysis for different species.
2. To set up a documentation programme which provides information of general interest, on the nature and origin if the material on morphological and physiological characters, as well as information about the value of material from the breeding point of view.
3. To store all the information in a form that can save space in the computer without missing any information.
4. To provide all the information in different languages. In agreement with these points a programme has been developed. The I/R system used is a partial version of TAXIR. That version was prepared by Dr. SOUSA at Washington State University in Pullman. It provides a maximum of 36 retrievable characters. An example of input, dictionary and output is given in appendix tables.

The documentation programme consists of three points:

- a) General information b) origin of material c) special part
- different according to the crop. The information is provided in agreement with the forms edited by FAO in 1969.

An example of the output in Italian language is provided in appendix tables.

E. PORCEDDU

THE USE OF FIELD PUNCHED CARDS FOR INFORMATION MANAGEMENT
OF SMALL COLLECTIONS

A. C. ZEVEN

Institute of Plant Breeding (I.V.P.), Wageningen, The Netherlands

When the programme of this conference was heralded I observed that Part 2 'Information Storage and Retrieval' only concerns papers on computerized systems. Many keepers of collections - most of them plant breeders - only maintain small ones. Therefore, they cannot easily use a computer to store the information gathered, unless they have one near at hand. I suggested to Prof. J. G. Hawkes that at least one lecture on a hand system should be presented.

Most people are acquainted with some type of hand system. They use them, for instance, for a literature index. The simplest system is an ordinary card index. In most cases the cards are arranged in a fixed system for instance alphabetically on first author's name, chronologically or according to the U.D.C. Another system is one of cards with coloured tabs, or one with needle punched cards. The last can be divided into edge-notched cards and field punched cards. I am using the field punched cards for my Triticinae collection. The picture shows the holes allotted to the characteristics recorded. In most cases I used three holes for a characteristic, viz. 1-absent, 2-present but intermediate, 3-present, or Ne-genotype $1=Ne_1Ne_1Ne_2Ne_2$, $2=ne_1ne_1Ne_2Ne_2$ and $3=ne_1ne_1ne_2ne_2$. I have not used all holes, because new interests may emerge and these can then be included by using some of the spare holes.

By inserting one or more needles in the holes of the desired characteristic(s) it is possible to select the collection numbers answering the desired requirements.

It is an easy system and quite useful for small collections. It takes little time to punch the card. The retrieval of a certain one is quite fast.

CONFERENCE on EUROPEAN and REGIONAL GENE BANKS
Technical Aspects of Tissue Culture Storage for Gene Banks

by

Dr. G.G. Henshaw

Department of Botany, University of Birmingham, U.K.

Tissue culture storage for gene bank purposes could, in theory, provide a useful alternative to the more orthodox storage methods where these have been found to present technical problems. It is now well known that cells from a wide range of plant tissue cultures are demonstrably totipotent with the result that they can be used to regenerate normal plants. This potential has already been exploited commercially for propagation purposes and for the production of virus free plants.

From the gene bank point of view difficulties would undoubtedly arise because of the genetic instability which is a characteristic of many tissue cultures. It might, however, be possible to stabilize the cultures, either by careful control of the growth conditions, or by storage of frozen or freeze-dried tissue.

The technical aspects of tissue culture storage are discussed and in particular five major areas presenting substantial problems are examined, viz.:

- I) Isolation of tissue or organ culture:
- II) Maintenance of culture avoiding genetical change or loss of morphogenetic potential.
- III) Freeze-storage of cultures.
- IV) Freeze-dry storage of cultures.
- V) Regeneration of plants from cultures.

The World Collections of Sorghum and Millets

Teland R. House

The topic assigned to me is "Evaluation and Utilization of Plant Genetic Resources". Most of what I prepared to say was very well covered in the morning session. It appears that it would be more useful for me to outline more specifically what is happening with collections of Sorghum and Millets, both from an informative point of view and to present a few new thoughts.

Dr.

Dr. K. O. Rastice shifted from Mexico to India in 1957 to participate in the Indian Agricultural Programme of the Rockefeller Foundation. He brought with him a collection of sorghums. During the period 1959-1962 he gave leadership to all India collections of sorghum and several of the millets. He also received collections organized by breeders primarily in Africa, and the U.S.A. The accumulation of existing collections was further strengthened by requests made by the National Science Foundation of the U.S.A. This activity is still going on, for example, we have recently received a large collection from the Cameroons, and have requested a collection from Rwanda. The present situation with respect to these collections is as follows:

<u>Scientific Name</u>	<u>Common Name</u>	<u>Number of Collections</u>
<i>Sorghum bicolor</i>	Sorghum	12000
<i>Pennisetum typhoides</i>	Pearl millet	1500
<i>Eleusine coracana</i>	Finger millet	900
<i>Setaria italica</i>	Foxtail millet	530
<i>Panicum miliare</i>	Common millet	615
<i>Panicum miliaceum</i>	Proso	485
<i>Paspalum scrobiculatum</i>	Kodo	485
<i>Echinochloa colona</i>	-	285

./...

These collections were initially established in support of the crop improvement programmes in India and expanded to world collections. The sorghum and pearl millet collections have been most extensively used, and these have been screened in several parts of the world. A number of measures of morphological characters have been made in the sorghum collection and these have been published as a special publication of the Indian Society of Genetics and Plant Breeding.^{1/} A sizeable portion of the collection has been evaluated for tolerance to the shoot fly Atherigone varia soccata and the stem borer, Chile zonellus (partellus) and reported as Memoir No. 7 of the Indian Entomological Society.

Other data have been taken indicating entries showing resistance to a number of diseases: Rust, Puccinia purpurea, Downy mildew, Sclerospora sorghi, Sugary disease, Sphacelia sorghi, Leaf blight, Helminthosporium turcicum, Grey leaf spot, Cercospora sorghi.

Information about the collection has been and is being accumulated at other locations. Drs. Webster and Futrell collected and published data indicating resistance to downy mildew and sooty stripe (Ramulispora sorghi)^{1/} in a substantial portion of the world collection.

Dr. R. C. Pickett at Purdue University is screening the world collection for protein content, lysine and several other amino acids.

It is likely that data will accumulate as more and more scientists in different parts of the world use the collection. The relevant results of most of this research will appear in published form; some will appear in projected reports, etc. These publications and
./...

-
- ^{1/} 1. Catalogue of World Collection of Sorghum and Pennisetum. Indian J. Genet., 27 (Spl. No.): 1-394, Indian Society of Genetics and Plant Breeding, New Delhi-12, (Price \$6.00)
2. Resistance to stem borer Chile zonellus (Swinhoe) and shoot fly Atherigone varia soccata Rond. In the World Sorghum Collection in India. Memoire of the Entomological Society of India No. 7.
3. Futrell, M.C. and O.J. Webster, New Sources of Resistance to the Downy mildew disease of sorghum. Plant Disease Reporter 50, 641, 1966.
4. Futrell, M.C. and O.J. Webster. Races of sorghums resistant to sooty stripe disease. Plant Disease Reporter 50, 606, 1966.

reports will normally be read by others interested in the same general problems. The use of existing publications and reports, then, has been and continues to be an important vehicle for the spread of information about the collection. It is important, as part of the collection function, that such published information be accumulated. This information would be useful reference material.

The collection provided a broad germplasm base on which to build the breeding programme. Interestingly, only one section of it, the yellow endosperm types were extensively used. The yellow endosperm types when crossed with combine Kafir 60, a cytoplasmic male-sterile seed parent, resulted in a hybrid that yielded well, had a grain quality approaching local preference, and threshed easily. Recently other entries in the collection have been more extensively used but primarily as parents in hybrids from which new lines have been selected. The whole collection has been useful as a source of insect and disease resistance and screening for protein quality.

The collection has been rigorously screened for tolerance to the shoot fly, Atherigone varia succata. Several forms of resistance have been identified, antibiosis, recover resistance (tiller survives after main culms killed), and oviposition preference; however, none of these forms of resistance is sufficiently good from an economic point of view. Consequently, thought is now being given to intergenetic crossing and it may be desirable to collect species of these genera for which crosses with sorghum have already been reported.

Several problems have been realized and several comments can be made which are of interest:

1. The whole collection of sorghum and of the various millets has been treated as a working collection and frequently increased. These collections are now in long term storage in the USDA storage facility at Fort Collins, but the collections in India must still be

./...

placed in long term storage and a smaller working collection identified.

2. Some enteries have been difficult to maintain at Hyderabad, India, making it difficult to build up a large supply of seed thus enhancing the need for frequent multiplication. It is known, for example, that the vigor of the Ethiopian collections is very much reduced under Hyderabad conditions. Such problems support the concept that collections should be increased in or close to the area of origination so that the collection is maintained in adequate quantity and in its original form.

3. Pearl millet is a cross pollinated crop. Sample size is a problem and careful head bagging at flowering time is essential. The nature of the pearl millet collection has changed even over the relatively short period of its existence. Dr. Ilenn Burton has suggested that a collection of pearl millet should be of about one kilogram in quantity and that this sample be placed in good long term storage. Numbers of samples would be reduced and each sample would have to be carefully made preferably represent the subrace or type of concern.

4. The lists accompanying seed requests have been time consuming to organize and incomplete. There has been an interest in the accessions pedigree and origin. This interest has now been answered by mimeographing and distributing the accessions books.

5. To adequately describe the agronomic characteristics of a collection is difficult. It has been useful to invite scientists from different parts of India to visit the collection when it was being increased or screened and make their own selections. Scientists in the U.S.A. have been able to see the collection in Puerto Rico.

./...

6. There has been a decided lack of good records about origin of collections, and in many instances, the information is not available from the breeder who made the collection. Wild and weedy relatives are also virtually absent. This indicates the need for some training of collectors before collections are made.

7. The maintenance and distribution of collections is expensive. The crop improvement projects are pressed for adequate funds and facilities and it is difficult to append the cost of collections. Certainly, it is not possible to simply add this burden to the existing crop improvement programmes. This seems to be a very real opportunity for international agencies to support collection activities as they are of international interest and frequently difficult to append to country programmes.

8. It has been considered important that these collections be kept in long term storage, in at least two places in different parts of the world to ensure their safety. The collections could be lost if kept at one location which was somehow destroyed. The project in India has been cooperating with the USDA and as mentioned earlier, the collections are now at Fort Collins, Colorado, and at Hyderabad, India.

Sorghum is a tropical grass, it has been adapted to grow at latitudes as much as 45° away from the equator. However, because of photoperiod problems, many collections adapted in the tropics will not flower at the higher latitudes. The USDA and the Texas Agricultural Experiment Station, as well as several commercial companies in the U.S.A., have begun to 'convert' these tropical types to shorter and less photoperiod sensitive types. The conversion is primarily a backcrossing process where the exotic type is the recurrent parent, and a short photo-insensitive line is the non-recurrent parent. Crossing and backcrossing is undertaken in a tropical location; the Federal Station, Mayaguez, Puerto Rico, (17°N) while the segregating generations are

./...

grown at a more northern location; the Plains Research and Extension Station of the Texas Experiment Station, Lubbock, Texas (34°N). (The Texas USDA Programme) The short early types are sent back to Mayaguez for the next cycle of backcrossing and selfing (the whole cycle takes one year). After three or four cycles of backcrossing the exotic type is recovered but is short and essentially photo-insensitive. Breeders in the U.S.A. have been keenly interested in this conversion programme because of the tremendous array of new genotypes available for their use. Entries from the conversion programme are also of interest to workers in the tropics, for the same reasons, or in a partially converted condition, from which a new line might be selected.

The pedigree system used is beginning to have an international structure. Indian sorghum (IS) numbers were assigned to sorghum accessions; Indian Pennisetum (IP) numbers to the pearl millet collections, etc. Within India, the IS number was retained as a pedigree by any station using the entry, and the location and season was added as an origin which in fact, made a complete pedigree. IS 534 sown at Hyderabad during the Khasif (summer) season of 1970, in row 25 would have the following pedigree: IS 534 Hyd 71 K 25. The same entry sown at Coimbatore would have a pedigree: IS 534 CO 71 K 105. The USDA has agreed to use IS numbers rather than PI numbers as related to the sorghum collection only. Entries from the conversion programme retain the IS number but add a "C"; i.e., IS534C. The IS number will be retained as these collections are used in the Near East and North Africa with appropriate symbols to indicate location and season. This represents a level of progress aimed at avoiding a proliferation of numbers associated with the same entry.

A number of years ago, the Rockefeller Foundation organized several committees to study the problem of germplasm collection and preservation of several important crops. Dr. J. R. Harlan is serving as chairman of the committee concerned with sorghum and

./...

millet. I am a member of this committee and would like to share with you the recommendations of the committee for the types of collections which should be developed as well as a few other ideas of interest.

1. Accessions collection - A collection in which seed of every collection is saved. This reduces the chance for genetic loss by discarding "apparent duplicates" and by bulking (drop in gene frequency).
2. Basic collection - A working collection developed by what might be taken as a two way stratification: (a) race, subrace, geographic distribution and adaptation; and (b) traits of economic importance (insect and disease resistance, combining ability, nitrogen responsiveness, etc.). It would be necessary to carefully evaluate the collection, likely drawing on the experience of workers in various parts of the world in developing the basic collection.
3. Spontaneous collection - This collection would consist of wild and weedy races and maintained separately primarily because of special management practices required and because some types are a nuisance as weeds.
4. Bulk collections - A series of bulks should be established compositing like materials. It is necessary to be careful that valuable traits, such as resistance to an insect pest, are not lost in the bulking process. Special bulks might be created of varieties with some special attribute. Entries in any one bulk should be similar in origin, height, maturity and adaptation.
5. Named Variety collection - This collection would consist of all varieties (not hybrids) named and released by private and public institutions. At some point in time these varieties were considered outstanding. Such elite germplasm might well have continuing or repeated value.

./...

6. Genetic Stocks collection - A collection of stocks with known genetic factors, translocations, inversions, or other special traits; also, lines with resistance to particular races of pathogenes, insects, etc. This would be a kind of gene bank.

7. Population collection - Populations should be created from carefully selected entries to conserve germplasm but also to be improved by selection with the objective of providing base material for breeding programmes of the future.

The basic collection as well as the bulk and population collections should be stored at several locations and could be kept in conjunction with a major crop improvement programme. Quarantine regulations gradually become more restrictive making it more and more difficult to freely moved seeds. It would be useful if these types of collections could be maintained at least on the various continents where the crop is of interest. This would not only help solve the quarantine problem but would reduce the burden on any one location for the supply of seeds. It would also provide a location for the increase of that part of the collection adapted in that area.

The maintenance of a voucher sample for entries in the collection is important. This would be particularly true for the basic collection. Every time the collection is increased, there is a chance for errors due to handling and even a change over time due to such problems as sample size and environmental influences. A herbarium of voucher samples would be useful to assist in maintaining type. The best voucher sample should include the head at flowering time, the head at maturity, the flag leaf, and at least one other leaf including leaf sheath and two nodes (communication with J.R. Harlan). If it is not possible to develop voucher samples as elaborate as this, it would be most helpful to have the head at maturity. Frequently, a given sample is not uniform and several heads may be required to represent the collection. The voucher samples can easily be taken from the first increase following collection.

./...

This committee has identified weaknesses in the collection and recommended areas of priority concern for additional collecting activity so that the germplasm of these crops is adequately protected.

Several international institutes now exist and a new one (ICRISAT) is just developing near Hyderabad, India, which will serve as a world centre for sorghum and pearl millet as well as chickpea and pigeon pea. These institutes and their outreach opportunities provide an excellent base to centre collection activities. Activities related to collections are of primary concern to ICRISAT, and it is likely that this existing germplasm of their crops will be adequately preserved before existing types are replaced by higher yielding varieties. This institute might logically coordinate activities related to collections of crops of interest to it. This is logical because it, like the other international institutes, will have contacts with many countries which would facilitate collection organization, the maintenance of working collections in several areas of the world, and increase in or near areas of origination. Its staff should be knowledgeable about the collection and be in an ideal situation to send appropriate entries in response to requests.

ALAD The Ford Foundation
P. O. Box 2379
Beirut, Lebanon

SUMMARY

A computer-based recording system for Pisum accession and genetic data.

Brian Snoad

John Innes Institute, Norwich, U.K.

April, 1972

The necessity for a more efficient method of recording and storing data has steadily become obvious as the size of the Pisum collection at the John Innes Institute has increased. Too much data was available for storage and retrieval using edge-punched cards and so an effort has been made to produce a computerized system.

The recording system relies upon standardized forms, each one based upon an 80 column IBM card. These forms allow for four categories of information: accession, genotype, phenotype and resistance to disease. The accession data recording sheets are based upon those originally purposed for an FAO/IAEA scheme but modified so that they cater more specifically for peas. The remaining three sections are new in concept and the recording forms permit simple, rapid and efficient scoring from the growing plants.. Transfer of data from the recording sheets to the punched cards is equally quick and efficient.

The accession data forms allow for the specification of the material, its accession record and details of its origin as an induced mutation, a selection or a collection. In addition details of any chromosomal re-arrangements can be noted. The remaining forms have space for the recording of over 400 genes (some 250 spaces having already been allocated), about 200 phenotypic descriptions if necessary and an equally large number of disease resistance reactions.

At present the data is stored on disk in an IBM 1130 computer. Programmes have been written by Dr. M.W. Johnson and Mr. R. Stanfield which allow the initial creation of the file and any subsequent additions or deletions of whole or part records. They also allow any individual item of data or any combination of items to be requested and for the appropriate information to be listed. A dictionary set has been created which allows output of the genotype, phenotype and disease assistance data to be printed in human language. Similar dictionary sets could of course, be provided for all sections should this be thought necessary.

The system works well and it allows rapid retrieval of all types of data collected from 900 accessions of peas. System such as this could be of value in those areas where extensive computing facilities do not exist and, providing the output is readily understandable, there would be no problems in the transfer of data to more sophisticated systems as and when they become available. Putting this system into operation has provided valuable information about many of the associated requirements and problems and has demonstrated that practice is far more rewarding than theory in initiating an efficient scheme.

Activities and programmes of the Italian

Germ plasm Laboratory

Prof. G.I. Scarascia Mugnozza, Bari, Italy

The establishment in Southern Italy of a germ plasm laboratory was decided by the National Research Council at the end of 1969.

Bari was chosen since its region (Fuglia) is for geographical and ecological conditions very representative of the Mediterranean zone. Bari was also selected because of the presence of a Faculty of Agricultural Sciences and of other Institutions dealing with problems of Mediterranean agriculture; therefore the Laboratory can make profitable use of competence and experience already gained as well as of laboratory and experimental field facilities already set up. Since the germ plasm laboratory is envisaged to take part in national and international programmes as well, it has been established as an autonomous body. It is also the centre of a national network of working collections. The laboratory acts as centre responsible for the maintenance of genetic stocks, for the promotion of exploration and introduction, for the storage of documentation in computerized form and for exchange and distribution of materials and informations. The working collections are entrusted to research stations which will characterize and evaluate accessions as part of their own activities.

Present activities concentrate in consequence of a recognized emergency situation, upon the following aspects: a) the establishment of a large collections of Triticum durum and related tetraploid species; b) the rescue, preservation and identification of local varieties lines, ecotypes and primitive cultivars of grain legumes, forage legumes, grasses and vegetables, since long time cultivated in Southern Italy and now rapidly disappearing.

Expeditions have been already carried out (Sicily, Sardinia, Southern Apennines) or are being programmed laboratory ~~researches~~ and studies have also been started namely: seed physiology in different storage conditions, cytogenetical analyses of a collection of Vicia species; in vitro culture of meristems of artichokes.

The essential aspects of the future programmes can be outlined as follows:

- 1) Development of collections, through expeditions also in other regions, of genetic resources of seed-propagated crops of relevant importance for the Mediterranean agriculture.
- 2) Establishment of collections of vegetatively propagated plants of Mediterranean origin, since resources are continuously dwindling.
- 3) Collection of primitive cultivars or ecotypes of species, which although of minor agricultural importance or not originated in the Mediterranean region, have had, nevertheless, a process of natural man-controlled adaptation in areas of this region.

In conclusion; the future programme of the Bari Germ Plasm Laboratory may be envisaged as a two-fold programme: from one side, enlargement of the crop collections in number and size and increase of the Italian network of working collections, on the other side insertion of the Laboratory in a larger (European and International) framework, since the very nature of collecting, evaluating, exchanging and exploiting genetic resources requires international co-operation.

Seed Storage for Conservation of Plant Gene Resources

James F. Harrington, University of Calif., Davis

The messages I am going to bring you today are simple but extremely important for gene bank seed storage.

Storage of seed for maximum longevity does not involve only the gene bank storage facility. Seeds begin deteriorating from the moment they are mature. Weathering on the plant, poor moisture and temperature conditions between collection and storage, and clearing injury can all hasten the loss of viability of the seed after it is placed in the seed bank. Do not store poor germinating low vigor lots.

The two most important causes of rapid deterioration in seeds are high moisture and high temperature. Two rules-of-thumb are useful guides. 1. A decrease of 1 % in seed moisture content doubles the life of the seed. 2. A 5 °C decrease in storage temperature doubles the life of the seed. Rule one is applicable between 12 % and 5 % seed moisture. Above 12 % storage fungi attack seed and below 5 % free radicle formation may decrease seed longevity. Rule two applies at least to 0 °C and if seeds are dried below 10% moisture, it may apply at below freezing temperatures. There is danger that if the seeds are not in sealed containers at subfreezing temperatures, they may gain moisture with ice crystal formation in cells and death of the seed. Seed moisture is a function of the relative humidity of the air. Below 65% RH, or about 12% seed moisture, no molds grow. Thus RH below 65% is safe for short time storage. But the ideal RH is 20-25% or in equilibrium with 5-7% seed moisture. If seeds are to be dried, the RH of the air must be less than the equilibrium moisture of the seed, so moisture will move from the seed to the air. Heating air reduces the RH of the air. A psychrometric chart can be used to determine the air

RH at different temperatures. However, heating to too high a temperature will injure seeds, shortening their storage life or even killing them if the temperature is high enough. A temperature of 35°C is best for drying many seeds and 40°C should be the maximum drying temperature. People from hot countries will say "But our air temperatures go above 40°C ". Our experience in California where this is often true, is that seeds maturing on the plants or dried in the sun are indeed injured at temperatures above 40°C .

In a warm climate of high humidity drying cannot be done by heating air alone but must be done using dehumidification as well. Either a desiccant such as silica gel or refrigeration can be used for dehumidification at warm temperatures.

Having dried the seeds to 5-7% moisture, their moisture content must be kept this low throughout their storage life. If stored in cotton or paper bags in an ordinary or high humidity storage, they will quickly regain moisture and deteriorate more rapidly. To keep them dry, a moisture-proof container is essential. This can range from an aluminum foil pouch, through a tin can, to gasketed metal bins, to a moisture-proof storage room.

A germ bank seed storage should be both moisture-proof and well insulated for efficient maintenance of a cold temperature. To construct such a storage, the room should have no windows and only one door. The door should be a gasketed refrigeration door. Preferably there should be an anteroom between two doors to enter. On the room side of the structural floor, walls, and ceiling a continuous moisture-proof barrier must be installed. This can be a 3 mm. thick layer of asphalt, 10 mil polyethylene sheets overlapped and sealed, or aluminum foil laminated paper tarred at the overlaps. The insulation is placed inside the

moisture barrier. It can be of such insulation material as styrofoam, or glass wool. Finally an inside coreving is used to finish the room and protect the insulation and moisture barrier.

A refrigeration unit must be of the proper size for the room and contents. If it is, a heating unit on the refrigeration cooling coils can periodically defrost the coils and achieve some dehumidification. But to maintain a temperature of 0° or 10°C and a RH of 25% a dehumidifier is necessary. There are three approaches. All can be successful if properly installed and maintained. A silicagel unit with the dehumidifier on the outside coupled with refrigeration is one successful method. The dehumidifier should be on the outside of the room because in reactivating the silicagel considerable heat is generated. A refrigerated lithium chloride spray unit is effectively working at the Japanese gene bank seed storage. Care is needed with this system as lithium chloride is corrosive to the pumping equipment and the solution is toxic to humans. The third method is to dry the seed to 5-7% moisture and store it in moisture proof containers. If there are several subsamples, gasketed containers in which indicator silicagel is included can be used. If the silicagel changes from blue to pink it can be replaced and redried, thus keeping the seeds in the container safely dry. This method allows quick removal of a subsample, yet keeps the remaining subsamples dry.

Whatever the method of storing the germ plasm seed lots, it is essential to have a rigidly adhered to maintenance and control system. Humidity and temperature records must be kept. Germination tests must be made systematically, at least every five years. If a decline in germination occurs, or even before that, if a decline in vigor occurs, there must be the growing out of another generation to minimize

genetic drift and mutations occurring in aged seed, but especially to prevent loss of the germ plasm by death of all the seeds.

In summary, to maintain gene bank seed moisture and temperature must be low. A seed moisture content of 5-7% and a temperature of 0°C appears to be best for most seeds with our present knowledge. The storage should be insulated and moisture-proofed. Constant maintenance and control checks are essential. If these principles are followed, longevity of seed of most species should exceed 30 years without loss in germination.

New Physiological and Biochemical Techniques for Obtaining
Maximum Longevity of Seeds in Gene Banks

J.F. Harrington, University of California, Davis

What do we know about obtaining maximum longevity of seeds in gene banks? We now know much about the proper environment. We know that the two most deleterious factors affecting seed longevity are high seed moisture and high seed temperature. Other environmental factors which may be important in gene bank storage of seed are oxygen, light, and radiation. An important aspect of seed storage that must be mentioned is the fact that a small percentage of plant species have seeds that cannot be dried without killing the seed. The effect of each one of these factors on seed longevity will be briefly discussed in relation to the physiology and biochemistry of the seed.

Seed Moisture

As is well known, at seed moisture contents above 30-40% (fresh weight basis) non-dormant seeds can germinate. Between this moisture range and 18-20% moisture, heating can occur due to seed and microorganism respiration. Such a rapid respiration requires oxygen and kills the seed quickly. From 18-20% to 12-14% seed moisture species of storage fungi, which are capable of growing at low relative humidities (65-95%) attack seed in storage. In storage then, sprouting, heating, and fungal attack all quickly destroy the viability of seed. Therefore stored seed should always have a moisture content less than 12%.

Yet several experiments have shown that seeds of some species can survive in the soil fully imbibed (above 40% moisture) for many, even hundreds of years, yet will germinate quickly when given proper germinating conditions. Unfortunately, we do not know the percentage of

seeds that survived in the soil this many years. The point I wish to make here is that it is possible that environments other than extreme dryness may be best for maximum longevity for some species. Little is known of the physiology and biochemistry of longevity under moist storage except that dormancy is involved. A high concentration of CO_2 or ethylene or a low concentration of oxygen may be involved. Still other compounds may be involved. For example, seeds of witchweed, a parasitic plant, remains imbibed in the soil for many years. The seeds do not germinate until a cytokinin leaches from the roots of the host plant which stimulates the witch weed seed to germinate. Moist storage of seeds needs research. It may be an economical means of long term storage of certain species included in seed germ banks.

So far, however, for most seeds the greatest longevity is achieved by drying the seeds to moisture contents below 12%. In fact there is increasing evidence that about 5-6% seed moisture is ideal for maximum longevity. By analogy with research in the food dehydration industry, it appears probable that the moisture of the seed at about 5% has been reduced to a monomolecular layer of water around the macromolecules. Macromolecules in a seed include starch, proteins such as enzymes, and nucleic acids in chromosomes. When the seed moisture is reduced below 5%, breaks occur in this monomolecular water layer. As these breaks become wider, free radicals, produced non-enzymotically by lipid autoxidation can more easily with the exposed portions of the proteins and nucleic acids, destroying their functioning. Most plant and animal tissues produce tocopherols which react with the free radicals, rendering them harmless. But tocopherols are produced by enzymes. Enzymes do not function in seeds at moisture contents below 10%. Therefore the tocopherols in the seed are tied up by the free radicals. New free radicals then destroy the functioning of other cell compounds. Feeding cows tocopherols has increased the tocopherol content of their milk reducing autoxidation of the lipids in milk. It may be possible to treat the mother plant to increase the seed tocopherol content

therefore increasing the storage life of seed. On an imbibition treatment of the seed with a tocopherol solution and redrying the seed might be effective. The production of free radicals is accelerated by high temperature and by a high oxygen atmosphere. As we know both these environmental factors hasten seed again.

Several workers have suggested that in seeds of low moisture, the activity of only some of the enzymes may produce toxic substances which accumulate. None have been specifically identified, but Floris transplanted aged wheat embryos onto young endosperms and young embryos onto aged endosperms. She obtained a decline in germination and vigor in both these transplants as compared to the control young on young transplants. This strongly indicates an accumulation of toxic metabolites although other explanations are possible. Research needs to be continued to determine if toxic substances do accumulate in aged seed. Then further research might find substances which could be incorporated into seeds which would neutralize these toxic substances and increase seed longevity.

The converse of accumulation of toxic materials is the breakdown of triggering mechanisms required for germination. The need for gibberellic acid and cytokinins for germination is well documented. But no data is available on whether these hormones break down as seed age. This should be an interesting field of research. Takayanagi and I studied another hormone, ethylene. We used an ethylene enriched atmosphere to replace the decline in endogenous production of ethylene in partially aged rape seeds. The ethylene in the atmosphere restored these partially aged seeds so that normal seedlings were produced. However, ethylene had no effect on seeds more severely aged, we think because additional deterioration factors prevented normal germination. Future research may discover additional compounds lacking in aged seeds. These might also be added to the germination medium. This is, in a sense, prolonging seed viability.

Seed Temperature

Again within limits the lower the ~~seed~~ temperature, the seed will maintain viability. As with low moisture at cold temperatures (below 5°C) there is little or no respiration, fungus activity, or insect activity. If lipid autoxidation, accumulation of toxic materials, and breakdown of triggering mechanisms are involved in seed again, these processes will occur more slowly the colder the seed is stored. This is true at least down to the freezing point of water in the seed. It is believed that if the seed moisture content is below 10%, the water in the seed is bound and cannot crystallize. Therefore it is possible that storage of dried seed in subfreezing temperatures should prolong viability even further. Becquerel kept seeds in liquid hydrogen for a few months without injury. After this experimental he made the statement that seeds so stored were in a state of suspended animation and should remain viable forever. He may be right but it will take an expensive long term experimental to even partially prove him right.

Seeds of below 10% moisture stored at sub-freezing temperatures may not immediately form ice crystals destroying the cells and seed viability, but in analogy with frozen foods it is possible that after a long period of storage in this humid atmosphere ice crystals may form or may form on thawing. Thus more research needs to be done on long-term sub-freezing storage of seed. Moisture proof packaging with sub-freezing storage may be the answer.

Seed Oxygen

Oxygen is essential for aerobic respiration. As I have already said, it hasten lipid autoxidation. It is involved in many other enzymatic and non-enzymatic reactions in plant cells. Thus the exclusion of oxygen should prolong seed viability. However, experiments to date do not consistantly show an advantage to excluding oxygen from

dry seed in storage. One possible reason is that in a sealed container oxygen is reduced by seed respiration and non-enzymatic reactions such as the browning reaction. Therefore, most seed in long-term air storage comparisons with other atmospheres have actually been unintentionally under low oxygen storage. More thoughtful research on the role of oxygen on seed longevity needs to be done.

Radiation

Light, at least the ultraviolet part and other radiations do adversely affect seed longevity. Most of this affect has been observed on chromosomal abnormalities and mutations. I am sure Professor D'Amato will speak to this subject so I shall not.

Seeds which Cannot be Dried

In any germ plasma storage of seed, consideration must be given to how to maintain the species of whose seeds cannot be dried. Tissue culture, which Dr. Henshaw will discuss is one possibility. Some are perennials and gene bank nurseries can be used, although very expensive. To date, long-term storage has not been achieved for the seeds of any of these species in spite of a fair amount of research. The physiology and biochemistry of why seeds of most species can be dried without injury yet seeds of a few other species cannot is a field of research completely untapped so far as I know. Such research should be very fruitful and should add considerably to our understanding and control of seed longevity.

Conclusions

We know and can control several environmental factors which influence seed longevity. Relative humidity which controls seed moisture and temperature appear to be most important. We are just beginning to scratch the surface of the biochemistry and physiology of seed again. There are several possible physiological and biochemical

techniques which may increase seed longevity. In incorporation of tocopherols into the seed and enriching the germination atmosphere with ethylene appear promising approaches. Much research needs to be done in this interesting field.

Generalised Distribution Mapping System

Summary of Proposals

1. The area to be mapped shall be rectangular, areas of other shapes being presented within a suitable rectangular frame. The mesh of the reference grid shall be square and the north-south and east-west scales of the map equal. The final choice of size of the plotted map shall be made by the system with regard to the physical dimensions of the plotter to be used.
2. Within the above limitations, definitions of the area to be mapped, the mesh size of the reference grid, identification of entities mapped, observation codes used and symbols to be drawn, shall be provided by the user.
3. Data shall be prepared for the computer on standard 80-column punched cards, one observation to a card, in a format duplicating as closely as possible that of the field observation lists to minimise the labour of, and danger of error in, punching.
4. Off-line card handling equipment shall be used wherever possible in preference to either manual or computer operations.
5. As data are input to the computer system they shall be written to a file on magnetic tape, to provide a permanent record of the data and to ensure, so far as possible, that each card passes through the computer's card reader once only.
6. Maps shall be produced by programs taking their data from the magnetic tape file, not directly from cards. These programs shall be sufficiently versatile to permit the mapping of any selection and/or combination of data.

./...

- 7: The magnetic tape file shall be capable of acting as input to a suitable set of statistical routines, either directly or by way of a pre-processor program to form part of the suite of programs supplied.
- 8: To assist interchangeability of the system between different computer installations the major part shall be written in a generally available programming language (FORTRAN IV) avoiding the use of machine-dependent operations, e.g. PRINT, An character formats within greater than 1.
- 9: Where operations are required which cannot be readily achieved in FORTRAN IV (e.g. bit string handling) or which may differ from installation to installation (e.g. comparison of variables holding non-numeric information, resource interrogation, communication with a graph plotter or graphical output system) the implementor shall be at liberty to use Assembler Code and/or non-standard coding. Such coding shall be used only for purposes such as the above and not solely in the interest of object program efficiency, unless the FORTRAN-coded version is unacceptably slow. Coding other than standard FORTRAN IV shall appear only in small modules, each carrying out one single well-defined function, and such modules shall be particularly clearly identified and documented.
- 10: Sufficient clear documentation, both internal (comment cards) and external (on paper) shall be supplied to enable a programmer familiar with FORTRAN IV to carry out any modifications necessary to adapt the system to another installation or to make minor changes in its operation.

B.L. Kershaw
February, 1972

Generalised Distribution Mapping System

Preliminary Proposals

Introduction

The purpose of this document is to present preliminary proposals for a computer system for the mapping of the geographical distribution of entities. We first describe the general framework in which the system will be required to operate.

It will be assumed that an area of ground is covered by a reference grid, dividing it into squares, and that if one or more of the entities mapped has been observed in a particular square a symbol will be drawn by the computer, by means of some kind of graph platter, in the corresponding square of the map. No attempt will be made to distinguish position within the square, i.e. the symbols will be placed with an accuracy of plus or minus half a square.

The map produced will be rectangular, areas of other shapes being presented within a suitable rectangular frame. The system will not attempt to do anything other than occurrence symbols, and in particular the reference grid itself will not be drawn. It is appreciated that a set of symbols with no indication of how they are located as a whole within the area mapped is useless, so provision will be made for each map to contain some form of register marks for the positioning of a suitable transparent overlay or overprinted map giving geographical details of the region concerned.

Phases of operation and programs

The production of a map from a set of field observations involves two distinct steps:

- 1: Data input, validation and the preparation of a data file on
e.g. magnetic tape.
- 2: The drawing of maps by means of data extracted from this
data file.

Each of these two phases may advantageously be carried out by a separate program or set of programs. A particular set of observations need then be input to the system once only, then becoming available for mapping in many different ways. The data file could also become the input to a set of statistical routines, allowing the making of statistical deductions from the data.

Data input modes

The minimum amount of information needed to permit a symbol to be placed on a map is the grid reference at which an entity was observed. A map drawn from a data file containing only this minimum information would indicate the presence or absence of entities over the area surveyed, and this may be sufficient for some applications.

For others it may be necessary to place on the map not just a symbol indicating the presence of the observed entity, but one of a number of different symbols chosen on the basis of some further information contained in the observation record. In a botanical survey it might be required to differentiate the habitat in which a plant was observed and/or its frequency of occurrence, while a map of the distribution of immigrant families in a town might also show the type of accommodation occupied by each.

The full description of a general observation needs four items of information, the entity concerned, x- and y- coordinates and 'further information', as explained above, which will be referred to as 'type' information.

If the map is required to indicate only presence or absence of entities (Mode A) it would be wasteful to provide and store null type information for each observation record. A Mode B map, using different symbols for different record types, can be drawn only if type information is provided and stored for each record. We thus arrive at the concept of two modes for data input and file storage.

Observations can be stored in groups, all members of a group referring to the same entity, so entity identification need appear only once in a group header. This must also contain an indication of the mode of the group.

Data medium and format

The data medium chosen should be generally available, the equipment to prepare is easy to use, and it should permit the production of data listings for check-reading. It should be easy to edit, sort and merge data off-line. The ability to sort data off-line is very important, as presorting of data can save a great deal of computer time, particularly with large data volumes.

Punched cards satisfy all these requirements better than any other data medium. Ready availability of equipment indicates that standard 80-column cards should be used.

The format of the cards should duplicate as far as possible that of the input documents (observation lists) to reduce the labour of, and danger of error in, punching.

For complete freedom of off-line data handling one and only one observation code should appear on each card, and each observation

should be punched complete. It is unlikely that an observation code will fill a card completely. (In the Warwickshire survey the longest possible code used 5 characters of entity identification, 4 of positional information and 6 of type information, a total of 15 characters. Thus 5 such observations could have been punched on a card.) If both cards and computer time are chargeable, the extra time taken for computer sorting of observations will outweigh the extra cost of cards incurred by punching only one observation to a card, and even if computer time is free, the ease of off-line data handling probably more than compensates for the wasteful use of cards. We shall assume that only one observation code is punched per card.

The fact that the whole code must appear on the card does not necessarily mean that the whole of every field of the code must be keyed. Observation lists as they come in from field observers will consist of lists of entities observed in a particular region, and by the use of the program card facility of an IBM 029 card punch, or similar facilities on other punches, the location code can be transferred from card to card by gang-punching. Only when punching of a list from a new location begins need the location information be keyed.

The identification of entities on the card may be either text matter (e.g. the botanical name of a plant species) or a numeric code. Both have advantages and disadvantages.

A numeric code is shorter, easier to punch, off-line sorting on it is quicker (fewer columns and only one pass per column, compared with two for an alphabetic sort), it occupies less space in computer storage and can be checked by program more quickly. Its disadvantages are that it is not so meaningful during check-reading of the data and that at some stage before punching a code must be added to each observation record.

If text is used for identification, the entity is immediately obvious during check-reading and no work is required on the observation lists before punching. The disadvantages are the extra length of the identifier (this might be considered an advantage in one way, as making for more efficient use of cards) and the possibility of mis-punching. Entity names are not, of course, necessarily meaningful to punch operators. Botanical names are likely to appear to them as meaningless character strings, and even an English word may be misread for another, similar word. Verification reduces but does not eliminate the danger of mis-punching, as does check-reading.

For economy of storage and ease of program control numeric entity identifiers will be stored in the data file, so it is probably best to assign these before punching. In a survey planned with computer processing in mind, field observers could be asked to record numeric identification instead of, or in addition to, text identification. This is a point which should be borne in mind in planning future surveys.

The above considerations dictate a fixed record format on cards. The advantages of fixed length fields, sorting, merging and check-reading are such that the alternative, variable length fields, is unacceptable. Variable fields would be useful only if it was required to pack as many observations on each card as possible.

Stages of data preparation

It is highly desirable that numeric entity identifiers be supplied for each record before punching. This having been done, punching becomes an exercise in copy-typing from the observation lists. After verification, the cards are listed on a tabulator, the listings check-read against the original field lists and any errors corrected.

The cards are now sorted on entity identifier, bringing + together all records for the same entity. This sort replaces the writing of records on sets of cards, one set for eachity entity, which was done for the Warwickshire survey.

Note that at this stage the ability to check-read is lost, since the order of the cards no longer corresponds with that of the field lists.

It would be possible to sort further on positional information, but this is unnecessary. Each observation must be referred to its map position by the computer, and this will take the same time whether the observations are ordered or random.

Entity header cards

If only numeric entity identifiers appear on the cards, the text description of an entity is not available to the computer. To remedy this, each block of cards referring to one entity should be preceded by an entity header card containing the nimeric identifier and the text description. These cards can be used to generate headings for computer output. They can be inserted in the file either manually or by the use of a sorter or collator.

Data editing

The insertion of addional, or removal or correction of erroneous, observations is simple while the data are on cards. After they have been written to the computer file these operations become more difficult, and it seems reasonable to insist that all data editing must be complete before computer input begins.

The medium for the computer file

It is not easy to make a firm choice between the two possibilities for the file medium, magnetic tape and exchangeable disc. The random access capability of disc appears attractive, but mapping is not really a random access operation. The main activity of a mapping program will be the serial retrieving of a section of the file referring to a particular entity, random access being required only to locate the start of the file section. Thus the random access capability of disc would be largely wasted, access to the start of file sections being achieved almost as well by calculated skipping of magnetic tape.

Magnetic tape is not only cheaper than exchangeable disc, but is almost certain to be available on any computer of a size capable of handling a mapping application. Instructions are defined in FORTRAN for handling magnetic tape, but not for disc, so if disc were used a subroutine would have to be provided to carry out transfers, and this would be another area in which the program would be installation-dependent.

We thus opt, albeit not firmly, for magnetic tape as the medium for the main computer file.

The file index

Whatever the medium of the main file, it will be necessary to define a number of subsidiary files, the chief of which is an index to the main file to permit rapid location of any required entity.

It is difficult to hold and index, which would require frequent updating, on a magnetic tape file itself. It would be possible to hold the subsidiary files on another tape or an exchangeable disc, but as the largest of these, the index file, is quite small, two words for each entity (a total of 1444 words for the entire Warwickshire survey), these subsidiary files would not make too great a demand on space in the computer's on-line file store.

The file creation program

A major function of this program will be feasibility checking of the data and the generation of error messages. It will be necessary to supply details of acceptable data to enable it to carry out these checks.

The checking of numeric entity codes will probably not be feasible, particularly if these codes are sparse, but it is easy to check that the entity code of an observation corresponds to that of the preceding section header. This check, besides guarding against mixed cards, ensures that a block of cards not preceded by a header card will be rejection.

The location information can be checked to see whether each observation falls within the defined area. A data card must be supplied giving the acceptable limits of x and y . (Possibly the grid references of the South-west and North-east corners of the map area.) This card must precede the observation cards.

For a Mode A file this concludes data validation. For a Mode B file the type code must be checked, and data cards specifying acceptable type codes must precede the main body of the card deck.

The minimum amount of information to allow type codes to be a list of acceptable type codes.

For space economy and processing ease a numeric encoding of each type code (type number) will be written to the file rather than a copy of the code itself, and the input program must assign these type number.

It would be possible to assign type number starting from unity to the type codes in the order in which they occur in the input, but this is not entirely safe. If the type code cards were mixed between runs of the program, the type numbers in different parts of the file would be inconsistent.

Supplying a required type number on the same card as the type code, besides preventing this danger, also permits a large measure of control. It would be possible to assign the same type number to two or more type codes, but as a general principle this is deprecated. It is advisable to retain as much information as possible in the file, in this case by assigning a unique type number to every possible type code.

A more important type of control which can be exercised is the assigning of ordered numbers to sets of type codes which are themselves ordered, e.g. botanical observations in the same habitat but differing in frequency of occurrence. Suitable assignment of type numbers would assist subsequent ranking of observations.

For a run of the file creation program in Mode E the following data cards would thus be required:

A card giving the x and y limits of the area mapped.

One or more type cards, each giving a type code and the corresponding type number. This block of cards would be terminated by a blank card, i.e. one having a blank type code and zero type number.

For each entity would follow an entity header card and one or more observation cards, each giving the entity number, the grid reference of the observation and the type code.

The entire data deck would be terminated by a dummy entity header card with an entity number of zero.

For a Mode A file creation run no type cards would be required. If the x and y limits card was immediately followed by the blank card terminating the types card this would be sufficient information to switch the program into Mode A.

Action of the file creation program

Three points require consideration under this heading. The first is the format of grid references in the file. If references are given on the National Grid, they could be stored in their full form rather than converting them into a local system having its origin at the south-west corner of the area mapped. By doing this the possible future linking of a number of local surveys into a national survey would be facilitated.

It is difficult, if not impossible, to modify the contents of magnetic tape but easy to append additional information at the end of the reel. In practice amendment is done by copying the whole file to another reel of tape, incorporating changes as they occur. This introduces the concept of file generations, since after an amendment run two copies of the file exist, a latest generation containing the amendments and the previous generation without them.

If it can be guaranteed that the card data for an entity will be complete and correct at the time of file creation, so successive runs of the creation program do no more than add to the file sections for further entities, a simple 'appending' file creation system can be written. Such a system could in principle operate with only one copy of the file, although for safety it is advisable to keep a back-up copy.

If amendments to the body of the file may be required, a more complex 'amending' system will be needed using at least two generations of the file.

In practice three generations usually co-exist, the so-called 'grandfather-father-son' technique. Each reel of tape in turn occupies the position of latest generation ('son'), the tapes being used cyclically. The difficulty with such a system is keeping track of which reel is playing which role at a given moment. In principle the computer itself could do this, with the aid of a small on-line file, but with an operating system such as that on KDF9 this would result in an unacceptable time penalty while the required tapes were located and loaded. With a more sophisticated operating system on a multi-programming machine, such as the more advanced versions of GEORGE on the 1900 series, it may be possible for the program itself to decide which reels it wants today, ask the operators for them and then suspend itself until they are loaded.

With a three-taped system it is not important to keep back-up copies, since it is only necessary to repeat one run to recover the state of the files after an accident. The difficulty with operating in such a mode with an automatic reel-assignment system is that after an accident the on-line tape allocation control file will have been advanced, and will thus be wrong for the repeated run. This can be avoided by delaying advancing the tape allocations until the start of the next run of the program, and supplying a data card to specify either a new run or a repeat of a previous faulty run.

Continuation tape reels will probably not be needed for a normal survey (they were not in the Warwickshire survey), but if they are wanted assistance will presumably be available from the operating system (e.g. GEORGE) in assigning and selecting them.

The choice between an appending and an amending system must be made afresh for every application, in the light of the facts of each case. Whichever system is used a preliminary program will be required to create an empty file, set up an empty index and ^{an amending system} for an amending system.

set up the on-line file controlling tape allocation.

Which type of system is chosen will affect the third point to be considered, which is the action to be taken on a data error. In an amending system it will be possible to ignore the offending card and write the acceptable data to the file, leaving the wrong data to be inserted during a later run of the program. For an appending system it is essential that the data written to the file be right first time, so an error would mean the ignoring, and subsequent re-input, of the whole of the data for entity. (The input program for the Warwickshire survey, worked in this way.)

Requirements for the mapping program

The most obvious requirement is to map the distribution of an entity or a set of entities, one map per entity, using either a single symbol or a set of symbols to distinguish different type codes. It may also be required to map the co-distribution of two or more entities on a single map, using distinct symbols to show occurrences of different entities.

For maximum generality it will be left to the user to specify the symbols to be drawn. A required map is then specified by setting up a correspondance between a set of observation records (entity and/or type) and the symbols to be drawn to represent these observations.

Symbol specification

For symbol specification each square of the map will be considere to be covered by a sub-grid of 60 x 60 squares (this number is chosen because of its many divisors) and the computer instructed to form a particular symbol by joining together certain pairs of points on this sub-grid. To draw a vertical line across the centre of the square, being a type number and the second the number of the symbol to be drawn for that type of observation.

for example, the computer would join the points (30 , 60) and (30 , 0). The specifying of symbols formed of straight lines is simple by this technique. Circles and arcs are a different matter, and it may be desirable to provide a circular arc generator subroutine.

Each card specifying a symbol will contain a symbol number and one or more pairs of points to be joined together to form that symbol. An example may make this clearer. Suppose vertical, horizontal and diagonal lines right across the square. The general format of the symbol defining cards for these symbols would be something like:

- 1 (30, 60) TO (30, 0)
- 2 (0 ,30) TO (60, 30)
- 3 (0,60) TO (60, 0)
- 4 (0, 0) TO(60, 60)

The assigning of numbers to the symbols is for reference only and in no way implies any ordering of the symbols. As will become clear later, the correspondence between observation codes and symbols may be set up in any desired way.

Symbols consisting of more than one line or having detached parts could be defined just as easily. Suppose symbols 5, 6 and 7 were required to be respectively a small square, a plus cross, and an equals sign, they could be defined in some such way as;

- 5 (20, 20) TO (20, 40), (20, 40) TO (40, 40),
(40, 40) TO (40, 20), (40, 20) TO (20, 20)
- 6 (30, 20) TO (30, 40), (20, 30) TO (40, 30)
- 7 (20, 20) TO (40, 20), (20, 40) TO (40, 40)

Symbol design

Although not strictly concerned with system specification, a few words on the design of symbols may not be out of place.

Several symbols may appear in the same map square, so they must be designed not to obscure each other or give rise to misleading combinations. Thus if vertical and horizontal lines are used a plus cross cannot be.

Symbols depending for their interpretation on their position in the square are usually unsatisfactory.

Symbols extending right across the square and joining up with similar symbols in neighbouring squares give rise to recognisable areas of shading on the map where the entity concerned appears over an area of country.

The type of marking device in the plotter used affects the drawing of thick lines. With an ink pen, a thick line may consist of two lines separated by the thickness of the pen point. This trick will not work with a ball-point pen. Some photographic plotters allow programmed control of line thickness by varying the width of the light beam, and different strengths of line may be possible on a cathode ray tube plotter by varying the brightness of the spot.

Establishing code-to-symbol correspondences

It is necessary to establish a correspondence between codes which may appear in the file and the symbols to be drawn to represent them. A code consists of two parts, an entity number and a type number. The correspondence established between codes and symbols must be one-to-one, i.e. a particular code must correspond to only one symbol, but the correspondence between symbols and codes may be many-to-one, i.e. more than one code may call for the drawing of the same symbol.

The program may be required to operate in any one of three correspondence modes, which may be called distribution mapping, co-distribution mapping and (for want of a better word) complex mapping. In distribution mapping each map represents the distribution of a single entity, and different symbols show the occurrence of different type codes. In co-distribution mapping one map represents the distribution of two or more entities, and different symbols show the occurrence of the different entities. In complex mapping the symbol to be drawn is selected by a combination of the entity and type numbers. We consider each of these modes in turn.

Distribution mapping

One map is drawn per entity, i.e. file section. A data card must switch the program into distribution mode and this same card may contain a list of identifiers of entities to be mapped or a pair or a number of pairs of identifiers to request mapping of all entities which identifiers within these inclusive ranges.

If the whole file is to be mapped, the work being broken into separate computer runs only for convenience in fitting into the machine's schedule, it would be convenient to be able to specify "begin mapping with the first entity not yet already mapped" and "continue as long as possible without exceeding the program's allocation of computer resources (i.e. time)". These orders could be conveyed to the system by a first entity limit of zero and a last entity limit of zero respectively.

The principle of correspondence establishment

The correspondence between type numbers and symbols is established by a list of number pairs, the first number in each pair

Let us take an example to make this clearer and at the same time to give some idea of the control possible by this technique. Assume that the card data contained six possible type codes, AB, CD, EF, GH, JK, and LM, which have been assigned type numbers from 1 to 6 respectively.

First suppose we have defined six symbols, also numbered 1 to 6, and we require a simple one-to-one correspondence, type number 1 producing symbol 1, type 2 symbol 2 and so on. The six number pairs necessary for this would be :

(1,1), (2,2), (3,3), (4,4), (5,5), (6,6)

If the order of the symbols differed from that of the type codes this list could be modified to, say:

(1,3), (2,2), (3,5), (4,1), (5,6), (6,4)

So far we have been establishing a one-to-one correspondence between type numbers and symbols. Suppose now we had defined fewer symbols than type codes, say three, so that the required correspondence would be many-to-one. In particular, suppose we required type codes AB, CD, and EF to map into symbol 1, codes GH and JK into symbol 2 and code LM into symbol 3, we could achieve this by the following number pairs:

(1,1), (2,1), (3,1), (4,2), (5,2), (6,3)

If no correspondence is established for a particular type number, that number will not give rise to a symbol on the map. Suppose we wanted type codes AB and EF to map into symbol 1, codes GH and LM into symbol 2 and codes CD and JK to be ignored, we could use the number pairs :

(1,1), (3,1), (4,2), (6,2)

By a combination of the above techniques any required correspondence between types and symbols can be specified in distribution mapping. The same technique, with suitable modifications, is applied in the other mapping modes.

Co-distribution mapping

The first data card must switch the program into co-distribution mode but cannot contain any entity information. The first number in each pair of correspondence setters is now an entity number rather than a type number. Apart from this, the correspondences are set up exactly as before.

It will be necessary to tell the computer when to begin a new map. If the number of entities to be co-mapped is small enough, the start of a new data card may signal the start of a new map; otherwise, a continuation symbol must be used to link two or more cards into a single map specifier.

Complex mapping

The first data card must switch the program into complexmode and each map must be specified by a series of number triples in which the first is an entity number, the second a type number and the third the number of the symbol to be drawn when that combination of entity and type occurs. Again a one-card-per.map or a continuation symbol technique may be used for map delimitation.

Complex mapping is likely to involve a good deal of tape skipping and/or plotter pen movement, and may be very time consuming.

./...

Resource interrogation

All programs must have a running time limit, and may have to request the use of other computer resources such as mill units. It is easy to specify so much work to be done in a run that the program's allocation of one or more of its resources is exceeded, resulting in forcible program termination with possible corruption of files. To avoid this, the system must check on its use of resources and decide not to attempt another operation (input a batch of data or draw a map) if there appears to be a danger of exceeding any allocation.

Interrogation of remaining resources will be installation-dependent; this is one of the areas which will require recoding in transferring the system from one installation to another.

A routine "resources" (RESRCS) must be written for each installation on which the system is to run. This routine will be called before each operation is attempted, and will return a "go/no-go" signal. On receiving the "no-go" signal the system will enter a wind-up routine.

As failure to attempt an operation for which resources did in fact remain is less disastrous than forcible program termination with corruption of files, the "resources" routine must be pessimistic in its assessment of whether it is safe to proceed.

Communication with a plotting system

The system will require to communicate with either the plotter hardware or the computer's graphical output system at either four or five points:

./...

1. To set up a suitable correspondence between the notional map within the computer and the physical map drawn by the plotter.
2. To draw a straight line connecting two nominated points on the notional map.
3. To move on to another map ("frame") on the plotter.
4. To terminate output to the plotter or graphical output system.

The above points of contact are essential if the system is to operate at all. It will be convenient if a fifth point of contact can be provided:

5. To write a character string beginning at a nominated point on the notional map, for map headings.

These four or five points of contact can be achieved by four or five installation-dependent subroutines. The first affects the system as seen by the user. To permit the setting up of a suitable correspondence between the notional and physical maps the system must be supplied with a data card giving the x and y limits of the area mapped and either the size of square of the mapping grid (which need not correspond to the observation grid) or the number of mapping grid squares along one side of the map. This card, while necessary to permit the system to operate, provides still further control. By having the observation and mapping grids differing it becomes possible to map the whole of the area on a small scale and coarse grid, or (if the grid references of the observations are accurate enough) part of the area on a larger scale and a finer grid.

Overall view of the mapping program

So far we have introduced four types of data card, all of which must be present before a map can be drawn:

1. Mode selector card, which in the case of distribution mode may also carry entity selection information.
2. Symbol specifier card(s).
3. Symbol correspondence setting card(s).
4. Area and grid size setting card.

Although not strictly necessary it is desirable to specify a required order for these cards, to simplify manual checking that all required data are present. This will not preclude changing only part of the data for a second or subsequent maps. The most logical required order appears to be :

1. Area and grid size setting card.
2. Symbol specifier card(s).
3. Mode selector card.
4. Symbol correspondence setting card(s).

The area, grid size and symbols need be specified once only for each run. A change in these parameters does not necessarily require a new mode or symbol correspondences, but a new mode requires new correspondences, since these differ between the various modes.

The correspondence data actually demand the drawing of a map. When they are read, the computer can check the data for completeness, i.e. that all necessary data are present and that no correspondence has been specified to an undefined symbol.

The resources used for drawing a map will vary from map to map. Experience in the use of the system will show roughly how much work can be done by a program with a particular allocation of resources, and indicate a suitable balance of resources to be requested so that one type does not run out far ahead of others. The "resources" routine will act as a "back-stop" if too much is attempted, besides controlling the operation of the system in

Use of statistical routines

No details are available of suites of statistical routines at the time of writing, so it is not possible to state confidently what facilities will be available. An attempt will be made to interface the main data file with one or more suites of statistical routines, and no difficulty in doing this is anticipated in principle.

Overall view of the system

As seen by the user the system will consist of three parts, with a possible fourth. First he will have to run a preliminary program to create a number of files, assign magnetic tapes and set up an empty main data file.

The main operation of the system will consist of running a program to input and validate card data and write them to the main file, followed by one to use this file to draw maps. The running of both programs will probably be divided into fairly short periods for convenience of fitting into the computer's schedule, and it will be possible to alternate them provided no attempt is made to map an entity whose data have not yet been written to the file.

Both programs may detect data errors and generate messages to guide remedial action. So far as possible, data validation will ensure that no erroneous data are written to the file or useless maps produced. Data cannot, of course, be checked for correctness, only for feasibility.

The fourth part of the system will be the use of the data file as input to some suite of statistical routines to produce numerical output.

./...

The general principle throughout the design of the system will be to build in as few constraints as possible and to require the user to specify as many items as possible. While admittedly making the system harder to use, this approach confers great generality and flexibility. It is felt that such a system is preferable to a tailored one, as used for the Warwickshire survey, which would require extensive re-writing for another application or installation.

Programming standards

Although not strictly of interest to the user, it may be useful to mention the principles to be followed in program writing.

All programs shall be fully modularised, the module structure reflecting the logical structure of the program.

So far as possible the facilities of FORTRAN IV shall be used, avoiding installation-dependent coding. Where the use of installation-dependent coding is unavoidable, for example in bit string handling or communication with a plotting system, such coding shall be isolated in small modules which are easily re-written for another installation.

All coding shall be fully documented, both internally by comment cards and externally on paper. Documentation, both internal and external, shall be designed to enable a programmer, familiar with FORTRAN IV and a particular installation but unfamiliar with the system, readily to acquire such information as he needs to adapt the system to another installation or to make minor changes to its operation.

The documentation of a self-contained module shall itself be self-contained, so that a programmer is able to locate and understand a particular module he is required to modify without necessarily understanding the whole system in detail.

Results from the system

At the end of his use of the system, the user will possess maps, drawn in a format and using symbols specified by him, possible statistical deductions from the data, and a permanent record of the whole of the data, again in a format specified by him, and suitable for rapid input to a computer. This permanent record may be of great future value.

B. L. Kershaw

February, 1972

Need for genetic conservation

by

Dr. Gösta Julén, Svalöv, Sweden

Genetic variation is fundamental for any progress through plant breeding. Genetic variation can be increased by a) genic reorganization through hybridization, b) mutation and c) mutation spontaneous or induced.

Natural variation with regard to cultivated plants is represented in wild ancestors, in primitive cultivated forms and in improved varieties obtained through systematic breeding.

Primitive cultivated local varieties are now disappearing due to replacement by improved varieties. Wild species are eradicated through the ploughing up of uncultivated areas, through wide use of weedicides and through hard grazing. This valuable gene resource can be saved by collection of diverse material of these types and artificial preservation of them.

The importance of the preservation of this type of gene resources is illustrated in following examples:

A- A local variety collected in Turkey of no agronomic value in Europe was found to be resistant to all known races of mildew. It is now used as an important present material in the breeding of mildew resistant varieties.

B- A wild type of kale collected from the Mediterranean region was crossed with oil-turnip. After chromosome doubling some new rape types were obtained with higher degree of winterhardiness than in any previously known rape material.

C- A Polish local variety of rape, Bornovski, of little agronomic value due to low yield was found to have extraordinary low content of sulfor-glucosinulates. It is now used in the rape breeding work

aiming at production of a variety from which the rape-cake can be used as a valuable protein fodder.

D- In screening the world collection of barley for improved nutritive value one very poor type from Ethiopia turned out to have very high protein contents combined with high lysin content. This type "Hiproly" is now widely used in barley breeding.

E- Genes for resistance towards *Verticillium albo-atrum* (wilt~~ing~~ disease) in Lucern were found in a very low frequency in some population varieties. From those highly resistant varieties have now been developed.

A REVIEW OF EUROPEAN GENE BANKS

Dr. Gösta Julén

Swalöv, Sweden

Land races and local varieties with a very wide genetic variation previously cultivated in all countries has been the source from which the plant breeders have obtained their basic breeding material. These local varieties have however in many regions long ago been replaced by improved varieties and in other regions the same development is now under way with accelerating speed. Thanks to new uniform, high yielding varieties the production has been considerably increased but at the same time a lot of basic genetic material is getting lost which might be a great disadvantage for the future plant breeding.

To save as much as possible of this genetic material available in primitive land races and in wild material was the idea behind the planning of an introduction centre at İzmir started in FAO in 1960 and the organization of the FAO meeting on Plant Introduction and exploration held in Rome 1961. In this meeting plans were drawn up for a global network of collection and introduction centres. Some years later the İzmir project materialized and the institute here has now been working as a collection centre for a number of years.

In Europe Eucarpia through its president Prof. E. Åkerberg in 1966 made an enquiry among European plant breeding institutes exploring the interest for the erection of regional gene banks. The reaction was positive and at a meeting in Wageningen it was decided that Eucarpia should take further action for this purpose. An Eucarpia gene bank committee consisting of Dr. Line from Germany, Dr. Lamberts from the Netherlands and Dr. Ellerström from Sweden was established.

This committee met in Vienna in 1958 in order to discuss and to plan further activities of Eucarpia in this field. It was proposed that ^{regional} gene banks with long term storage facilities should be erected. These gene banks should be responsible for the storage of seed lots and for distribution of seed samples requested by plant breeders. The multiplication of seed material when needed should be assigned to those institutions which originally had supplied the seed. Such an arrangement should have the advantages that the costs for the preservation of such material should be considerably reduced and the multiplication of the material by the institutions by which it was previously raised would permit the best conditions for the material to maintain its specific properties. It was further suggested that in cooperation with FAO an "Index Seminum" would be erected in which the studied material would be described as exhaustively as possible. Four regional centres for Western Europe, Central and Eastern Europe, The Mediterranean area and Scandinavia respectively was suggested.

After this meeting a questionnaire was sent out to all Eucarpia members to find out the number of entries to be stored. The replies were too few to give a complete picture of the need.

A new meeting of the committee together with representatives for FAO and the Scandinavian gene bank committee was held at Svalöv, Sweden in 1969. At this meeting it was decided that the gene banks should arrange conservation of wild and primitive material, local and well adapted varieties, valuable advanced gene material which would be kept by the breeders and genetically and cytologically well defined material of cultivated plants. It was further concluded that three centres should be established, one for western and central Europe, one for Mediterranean area and one for the Scandinavian countries. The centre for the Mediterranean area should be organized in cooperation with FAO. A good cooperation should be arranged between the gene centres, FAO and Eucarpia should function as coordinators. Finally it was decided that all information should be recorded according to standardized methods of recording. Then information should be stored in a central information centre developed by FAO.

A second question ^{gene} was now sent to Eucarpia members with further information. This time the response was more complete and it was found that there was a need for storage of at least 105 000 entries.

A third meeting of the gene bank committee was held at Braunschweig-Völkenrode in 1970. It was reported that decisions were taken to establish regional gene banks at Bari in Italy and Völkenrode in Germany. Steps had also been taken in Scandinavia and although no final decision had been taken it seemed likely that a gene bank for this area should be established in Stockholm at the Seed Central Institute. Two working groups on seed physiology (Dr. Frahl Braunschweig, Völkenrode; Prof. Esbo, Stockholm and one representative from Bari) and documentation (Mr. Seidowitz, Braunschweig-Völkenrode and Dr. Gullberg, Stockholm) were established.

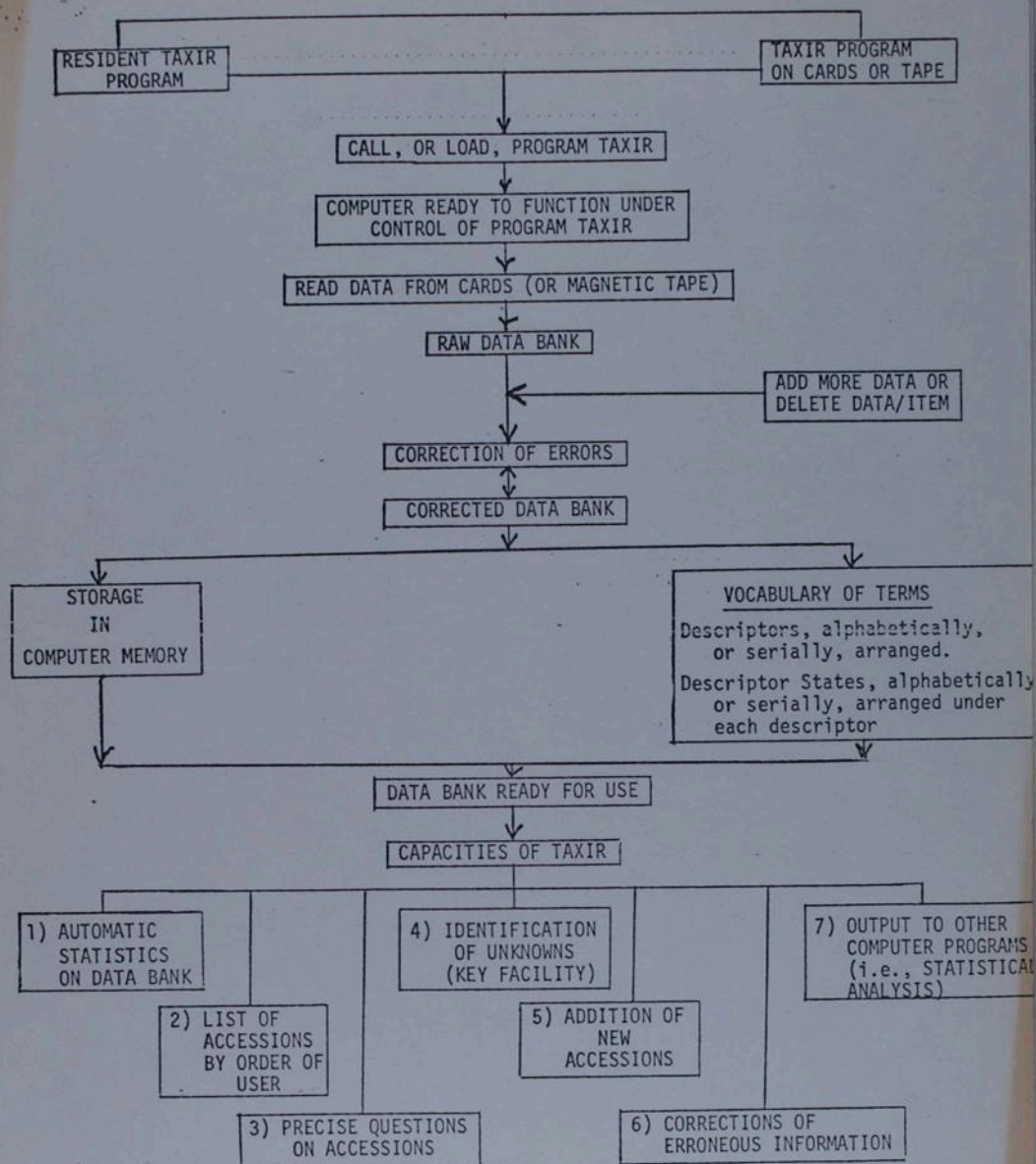
They should review the fields of problems and co-ordinate research in these fields avoiding duplications. They should also contact the group of specialists set up by FAO consisting of Dr. Edward James for long term seed storage, Prof. Hawkes for maintenance of vegetatively reproducing plants and Prof. Konzak for genetic information and data processing. It was agreed that efforts should be made to standardize a field record system so as to fit with the documentation system. In order to ensure necessary co-ordination the directors of the regional gene banks should meet annually.

Prof. Taygi from the Introduction Centre at Izmir had expressed a wish for a close co-operation between his institute and the Eucarpia gene banks activities.

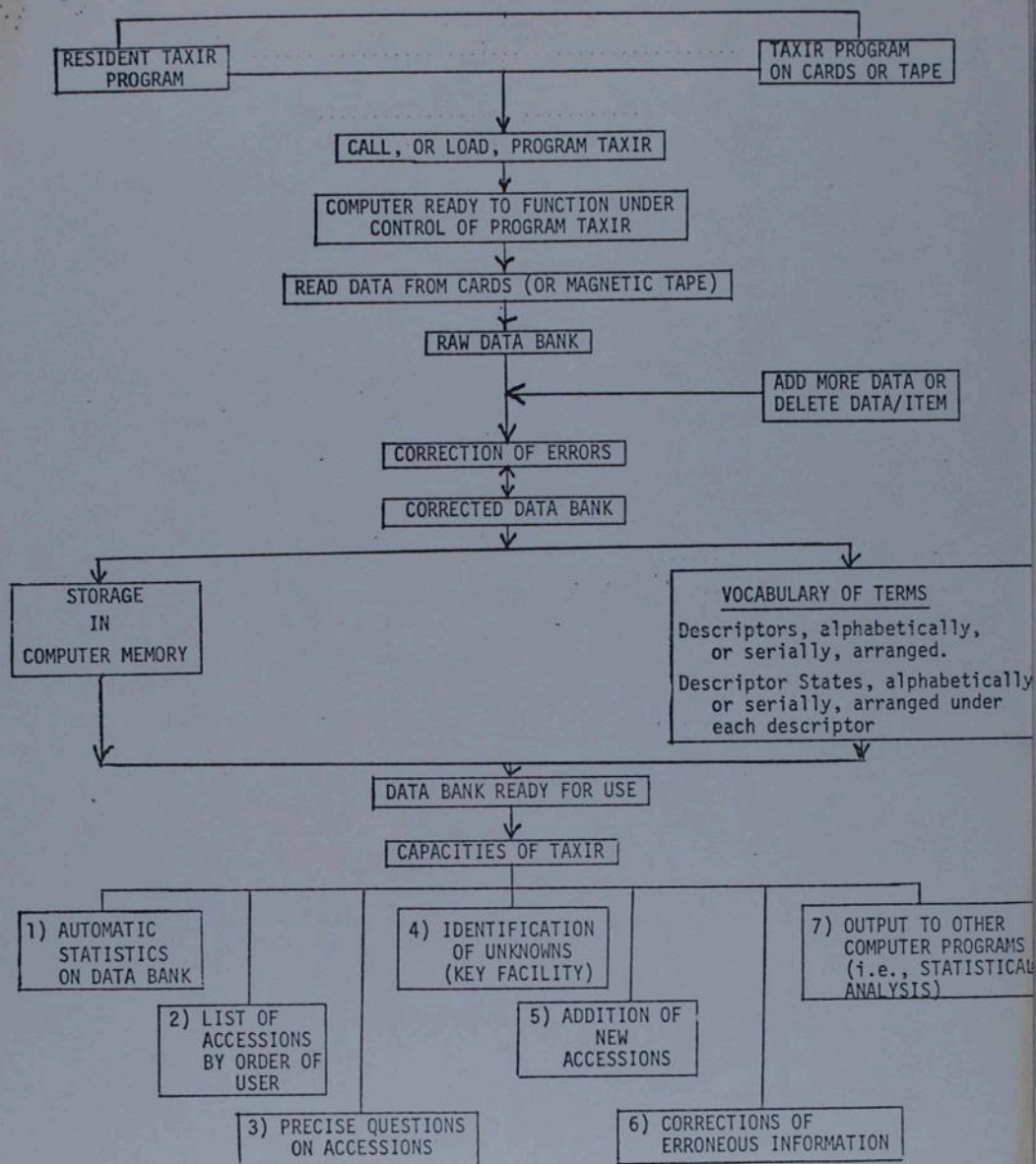
Since this meeting was held the gene banks in Bari and Braunschweig-Völkenrode has been established and have started their activities. Reports on these activities will be presented separately. Concerning the Scandinavian gene bank no final decision has been taken but the Ministry of agriculture is now nominating a governmental study group to draw up the plans for the future development in this field.

In this gene banks should be efficient full co-operation from the breeders is necessary. They should provide the material to be stored and also the information related to the material. It might be possible that breeders hesitate to cooperate with the gene banks due to commercial interest not least after the enforcement of the international convention on plant breeders right. Necessary arrangements to overcome this hesitation must be agreed upon in order to give full efficiency to the gene banks for the benefit future plant breeding work.

FOURTH STAGE--TAXIR Processing



FOURTH STAGE--TAXIR Processing



COMPARISON BETWEEN MANUAL DATA FILING SYSTEM AND TAXIR SYSTEM

MANUAL

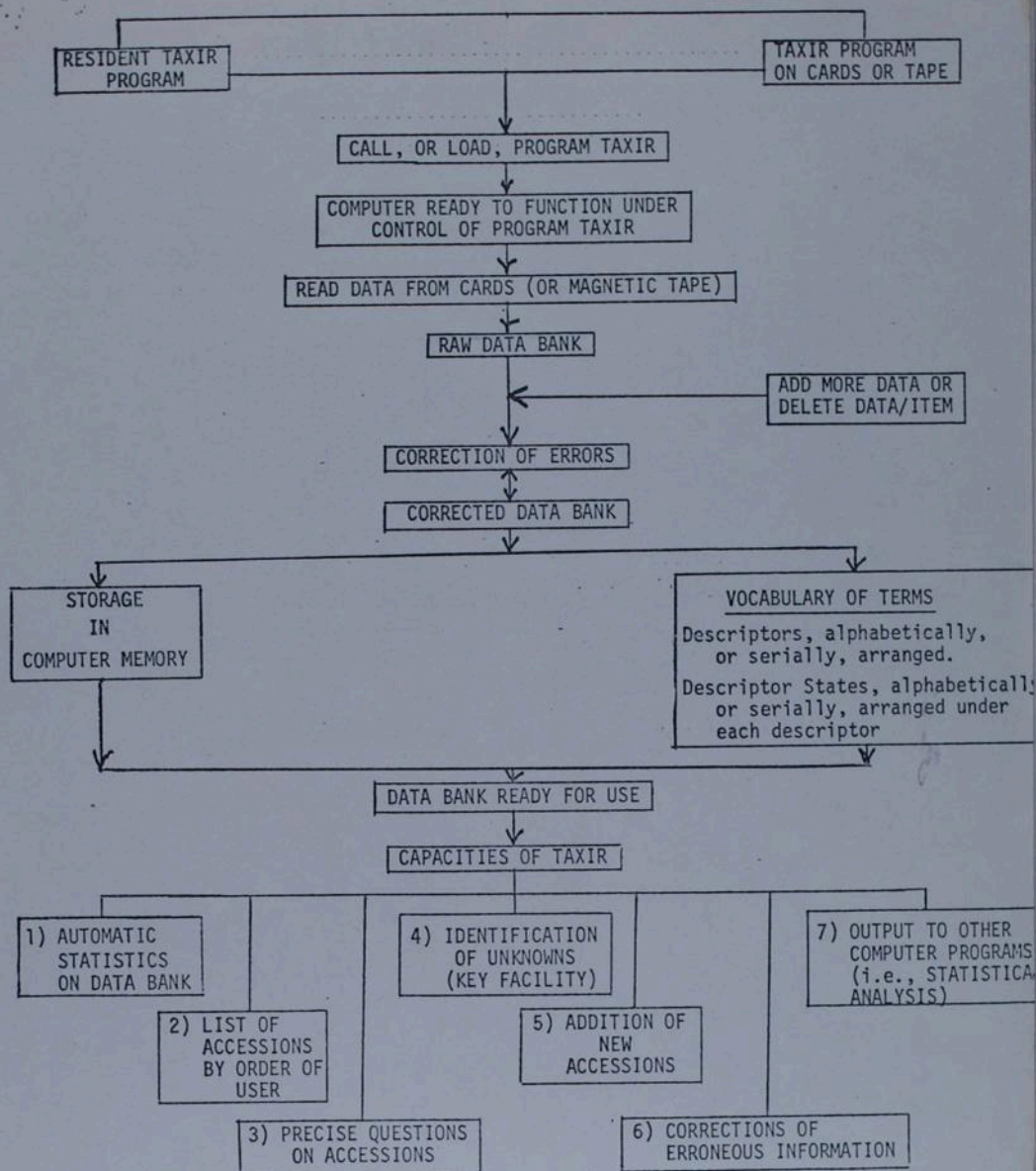
TAXIR

- | | | |
|--|----|--|
| 1. Data to be filed arranged under limited number of prechosen index headings:
1. Alphabetic
2. Geographic
3. Taxonomic classification
4. Subject headings
5. Numerical | ←→ | 1. Data to be filed arranged under any number of index headings. Gene bank director and scientists choose all desirable headings. Additional file headings may be added at will. All files automatically cross-referenced. |
| | | All require cross-referencing system. |
| 2. Each file requires a separate physical file, thus, much duplication of each item added to the files. | ←→ | 2. Each datum recorded but once, eliminating need for duplicate effort. |
| 3. The separate files listed above cannot be ordered into one over-all file. | ←→ | 3. All files ordered together in one operation, all under one program command statement. |
| 4. Great difficulty extracting additional data which does not fall into a major file heading. | ←→ | 4. Each datum individually addressed not dependent on a major file heading. |
| 5. Once established, reorganizing files into a different filing scheme is costly and time consuming. | ←→ | 5. Any desired changes in file organization accomplished with one computer program instruction. |
| 6. No possibility of guaranteeing precise and accurate responses when a question is asked. | ←→ | 6. Automatic searching of data file extracts all pertinent responses, without missing any one. |
| 7. Time to search complete file is prohibitive. | ←→ | 7. Time to search complete file measured in seconds. |
| 8. Little possibility of keeping statistics on file contents. | ←→ | 8. Constant count of data items in file. |
| 9. Difficult correction, deletion, or updating files. | ←→ | 9. All corrections, deletions, updating of files made accurately and instantly with one program instruction for correction, one for deletion, and one for updating. |

COMPARISON BETWEEN MANUAL DATA FILING SYSTEM AND
TAXIR SYSTEM

<u>MANUAL</u>		<u>TAXIR</u>
1. Data to be filed arranged under limited number of prechosen index headings: 1. Alphabetic 2. Geographic 3. Taxonomic classification 4. Subject headings 5. Numerical	← →	1. Data to be filed arranged under any number of index headings. Gene bank director and scientists choose all desirable headings. Additional file headings may be added at will. All files automatically cross-referenced.
	} All require cross-referencing system.	
2. Each file requires a separate physical file, thus, much duplication of each item added to the files.	← →	2. Each datum recorded but once, eliminating need for duplicate effort.
3. The separate files listed above cannot be ordered into one overall file.	← →	3. All files ordered together in one operation, all under one program command statement.
4. Great difficulty extracting additional data which does not fall into a major file heading.	← →	4. Each datum individually addressed not dependent on a major file heading.
5. Once established, reorganizing files into a different filing scheme is costly and time consuming.	← →	5. Any desired changes in file organization accomplished with one computer program instruction.
6. No possibility of guaranteeing precise and accurate responses when a question is asked.	← →	6. Automatic searching of data file extracts all pertinent responses, without missing any one.
7. Time to search complete file is prohibitive.	← →	7. Time to search complete file measured in seconds.
8. Little possibility of keeping statistics on file contents.	← →	8. Constant count of data items in file.
9. Difficult correction, deletion, or updating files.	← →	9. All corrections, deletions, updating of files made accurately and instantly with one program instruction for correction, one for deletion, and one for updating.

FOURTH STAGE--TAXIR Processing



COMPARISON BETWEEN MANUAL DATA FILING SYSTEM AND TAXIR SYSTEM

MANUAL

TAXIR

- | | | |
|---|---|--|
| 1. Data to be filed arranged under limited number of prechosen index headings: | ←→ | 1. Data to be filed arranged under any number of index headings. Gene bank director and scientists choose all desirable headings. Additional file headings may be added at will. All files automatically cross-referenced. |
| 1. Alphabetic
2. Geographic
3. Taxonomic classification
4. Subject headings
5. Numerical | } All require cross-referencing system. | |
| 2. Each file requires a separate physical file, thus, much duplication of each item added to the files. | ←→ | 2. Each datum recorded but once, eliminating need for duplicate effort. |
| 3. The separate files listed above cannot be ordered into one overall file. | ←→ | 3. All files ordered together in one operation, all under one program command statement. |
| 4. Great difficulty extracting additional data which does not fall into a major file heading. | ←→ | 4. Each datum individually addressed not dependent on a major file heading. |
| 5. Once established, reorganizing files into a different filing scheme is costly and time consuming. | ←→ | 5. Any desired changes in file organization accomplished with one computer program instruction. |
| 6. No possibility of guaranteeing precise and accurate responses when a question is asked. | ←→ | 6. Automatic searching of data file extracts all pertinent responses, without missing any one. |
| 7. Time to search complete file is prohibitive. | ←→ | 7. Time to search complete file measured in seconds. |
| 8. Little possibility of keeping statistics on file contents. | ←→ | 8. Constant count of data items in file. |
| 9. Difficult correction, deletion, or updating files. | ←→ | 9. All corrections, deletions, updating of files made accurately and instantly with one program instruction for correction, one for deletion, and one for updating. |

PARTICIPANTS

Eucarpia Conference on "European Gene Banks"

<u>Austria</u>	: Dr. K. Mikaelson
<u>Australia</u>	: Sir O. Frankel
<u>Belgium</u>	: Dr. V. Melard
<u>Canada</u>	: Dr. A.M. Galia
<u>England</u>	: Mr. B. Ford-Lloyd Mr. M.T. Jackson Dr. G.G. Henshaw Dr. D.A. Wilkins Prof. J.G. Hawkes Dr. B. Snod Dr. R.N.H. Whitehouse Dr. G.E. Dixon Dr. T.I. Emezc Mr. B.L. Kershaw Dr. P. Thompson
~~<u>F.A.C.</u>~~	: ~~Dr. J. Leon~~
<u>Japan</u>	: Dr. M. Matsuura
<u>Kenya</u>	: Dr. U. Gullberg
<u>Netherlands</u>	: Dr. H. Lamberts Dr. A.C. Zeven
<u>Pakistan</u>	: Mr. A.R. Rao
<u>Peru</u>	: Mr. A. Huaman
<u>Philippines</u>	: Dr. T.T. Chang
<u>Poland</u>	: Dr. W. Kulpa
<u>Spain</u>	: Prof. J. Miranda
<u>Sweden</u>	: Dr. G. Julén Prof. H. Esbo

Turkey

: Prof. V. Taygi
Prof. F. İncökara
Prof. M. Dokuzoğuz
Prof. Ş. Elçi
Mr. T. Aydın
Mr. A. Ürem
Dr. Ü. Emiroğlu
Dr. B.C. Wright
Dr. Ş. Emiroğlu
Dr. W. Hertzsch
Dr. M. Gökçöl
Mr. Ş. Bekiroğlu
Dr. A. Klatt
Dr. K.J. Neddenriep
Dr. R. Boye
Dr. Halidun Eresen
Dr. Basri Devecioğlu
Mr. V. Caethoven
Dr. J.M. Prescott
Mr. B. Doruk
Mr. S. Aksu
Mr. N. Dinçer
Dr. E. Kjellqvist
Miss H.A. Sencer
Dr. O. Manas
Mrs. A. Manas
Miss N. Aydın
Mrs. N. Settar
Mr. I. Özmat
Mr. K. Kayhan

U.S.A.

: Prof. J.F. Harrington
Prof. C.F. Konzak
Dr. D.J. Rogers

U.S.S.R.

: Prof. K. Budin

West Germany

: Prof. E. Knapp
Prof. H. Kuckuck
Prof. G. Ziegenbein
Prof. D. Bommer
Dr. K.D. Krolow
Dr. L. Seidewitz
Dr. A. Grahl

Italy

Prof. L. Quagliariotti
Prof. E. Porceddu
Prof. G.T. Scaranicia - Mugnozza

Planning meeting - procedures -

1. Identifying objectives - or problems

a) What ~~types~~ types of info. to be exchanged.

b) What are the audiences (or users)

c) What types of facilities (computer, ~~none~~ computer, key punch, etc.) needed -

d) methods of interchange - within a gene center, between gene centers.

Make flow chart of document center -

Read form → document center → cards - program

Other input
for other
purposes → " → "

Identification and

Development of different data banks -
such as

1. General introduction bank

2. Collector's data bank

3. Individual crop banks, ~~such as~~ ^{such as} genetic & pathologic data

4. Source of information on location of gene banks of particular holdings -

5. ~~Index~~

Each
type of
data
bank
requires
separate
types of
documentation.

Long-term gene bank storage - sep. data bank
versus

working collections. - ~~separate data bank~~

If you want but one accession no., what do you do with
when you get in collections from other banks already assigned.

Accession number assignment-

Use either numeric or alpha-numeric combinations-

Perhaps latter is better, because can use alphabetic notation to identify the particular gene bank, and numeric notation associated with the alpha, to identify the particular accession.

Or may be better to use a separate descriptor to identify the gene bank.

What is being done for documenting various wheat trials at Iqum?

Need a S

4 gene banks in Eur.

Ea. resp. for maintenance + multiplication -
" " " catalogues of their materials.

Differences between problems of gene banks + plant breeders

Multiplication of genetic material different from
regeneration of gene bank contents.

A newsletter proposed to inform others what ea gene bank
is doing - as a part of Bull. of Eucarpia.

Also proposed in a yearly meeting of gene bank ^{re} representatives
to inform on expeditions planned - another newsletter
proposed - and a committee to coordinate.

Rules requested ^{from gene banks} for bringing in material to the gene banks.

Problems of governmental relations - Eucarpia can help
escape the governing by national govt.

But on problems of exploration, keep in touch with FAO.

CROP SPECIFIC PROBLEMS

Gene Bank

1
Multiplication

work done in
this area
by contract
to others
or
by gene bank
itself

I CULTIVAR COLLECTION

II GERM PLASM

III SPECIAL "GENES"

1

TYPES OF DESCRIPTORS

1. MORPH.

2. DEGREE OF RESISTANCE to various diseases

Differentiation between vertical + horizontal resistance
Type of resistance (and degree of res.)

Information

1.

1. Identification data
 2. Collection data
a, b, c
 3. Storage data
 4. Distribution data
 5. Descriptive data
a, b, c
-

⇒ Problem of multiple language descriptions.

EUCARPIA

(European Association for Research
in Plant Breeding)

CONFERENCE

ON

EUROPEAN AND REGIONAL

GENE BANKS

10th-15th April, 1972

at

Izmir, Turkey

The meeting has been arranged by the Eucarpia Gene Bank Committee and the Eucarpia Section 'Wild Species and Primitive Forms'. It will take place at the Agricultural Research and Introduction Centre, Izmir, Turkey, by kind permission of the Director, Professor Vamik Taysi.

EUCARPIA

(European Association for Research in Plant Breeding)

CONFERENCE

ON

EUROPEAN AND REGIONAL

GENE BANKS

10 th-15 th April, 1972

at

Izmir, Turkey

The meeting has been arranged by the Eucarpia Gene Bank Committee and the Eucarpia Section «Wild Species and Primitive Forms». It will take place at the Agricultural Research and Introduction Centre, Izmir, Turkey, by kind permission of the Director, Professor Vamık Tayşi.

ACCESSION NUMBER		COLLECTION NUMBER	
CODE	NUMBER	TEAM	NUMBER
FAMILY		SITE SAMPLE NO.	
GENUS			
SPECIES			
SSP. OR BOT. VAR.			
LOCAL NAME			
EXPEDITION TEAM			
COUNTRY		EXPEDITION LEADER	
ORGANIZATION		LATITUDE	
LONGITUDE		ASPECT	
SLOPE		ALTITUDE	
COLLECTION		SITE DESCRIPTION	
SAMPLE		HERBARIUM	
1. SEED N-POP		1. V. SCARCE	
2. SEED B-POP		2. SCARCE	
3. POLLEN N-POP		3. INFREQUENT	
4. POLLEN B-POP		4. FREQUENT	
5. VES. N-POP		5. V. FREQUENT	
6. VES. B-POP		6. SOLID STAND	
DISTURBANCE FACTORS		SITE DESCRIPTION	
CODE:		LATITUDE	
1. NONE		LONGITUDE	
2. MARGINAL		ALTITUDE	
3. LOW		ASPECT	
4. MEDIUM		SLOPE	
5. INTENSE		OTHER, SPECIFY	
6. WIND EROSION		FALLOW	
7. WATER EROSION		CULTIVATION	
8. FLOODING		ABANDONED	
9. DRAINAGE		FERTILIZATION - CHEMICAL	
10. IRRIGATION		FERTILIZATION - ORGANIC	
11. FIRE		HERBICIDES	
12. GRAZING		OTHER, SPECIFY	
13. CUTTING		OTHER, SPECIFY	
SITE LOCATION			
OTHER INFORMATION (ADDITIONAL DATA, NOTABLE PATHOGENS AND PESTS OF/IN VICINITY OF COLLECTED SPECIES)			

THE DOCUMENTATION OF COLLECTIONS AND INFORMATION MANAGEMENT

BY GENETIC RESOURCE CENTRES¹

C. F. Konzak²

Documentary information always has been recognized as adding dimension to the worth of plant collections, be they herbarium specimens or living genetic resource materials. The documentation of specimens has made possible the development of concepts in taxonomy, ecology and genetics, extending from ancient history through the development of the major works of Linnaeus, Darwin and others. Vavilov recognized the great importance of documentary information to genetics and plant breeding in studies that led to the development of his concepts of the origins of cultivated plants in relation to the distribution of natural variation and in his law of homologous series (19).

Jack Harlan's recent treatise on the relation of centres of origin to genetic variability in crop plants has relied heavily not only on documentary evidence from plant explorations but also on documented discoveries of man and plants from the past (archeology, anthropology) and present (genetics, ecology) and from other sources (4). Though Harlan has clearly demolished Vavilov's suggestion that centres of origin were also primary centres of genetic diversity, Vavilov surely would have reached a similar conclusion had he lived to complete his research. However, he did manage to identify what he called secondary centres and these secondary centres are components among those Harlan now calls non-centres.

¹Scientific Paper No. _____. College of Agriculture Research Center, Washington State University, Pullman, Washington, U.S.A. Project Numbers 1568 and 1920.

²Professor of Agronomy & Genetics, Washington State University, Pullman Washington, U. S. A.

Harlan's recent analysis has a better genetic basis than the widely portrayed view of Vavilov. This new concept is of considerable importance in practice and particularly to plant exploration philosophy since it suggests that useful genetic variants may be found anywhere plants grow and particularly when they are exposed for many generations to either natural or man-made selective forces. Moreover, as we now know, certain naturally occurring genetic variations eventually may be concentrated by diseases endemic to certain regions and man's use, while the frequency at which other genetic variations are found may not be similarly affected. We can interpret this evidence as suggesting that the survey type of collecting which involves sampling plant populations rather than individuals on a grid basis, has the greatest chance of efficiently finding new genetic traits. It is possible that more concentrated sampling in areas subject to a disease may yield a relatively greater range of genetic variability for reaction to that disease. However, areas subject to a disease may not be the only origins for resistance to that disease. A similar situation may pertain for other traits, but unfortunately, we still do not know whether the example mentioned above for disease resistance is really true or whether it just seems to be true because we have a natural force working for us to "flag" useful genetic variants. These new developments are described here to show how important documentary evidence has been to the development of our concepts of the evolution of genetic resources of the plants we use. In turn, concepts of natural evolution influence plant breeding practice, which is in fact managed evolution. The further evolution of genetic variability in our crop species will certainly be more concentrated at the world's plant improvement centres. However,

since managed evolution cannot progress without genetic variability, we must preserve an adequate base of genetic resources.

Our concern at these meetings is with the aim to rescue and conserve the naturally evolved genetic variability (especially in our major crops) still available. This report will briefly describe proposed standard principles and procedures for documenting genetic resource materials and possible mechanisms to facilitate information management by Genetic Resource Centres.

RESULTS AND DISCUSSION

The recent IBP-FAO meetings on the conservation of genetic resources have helped bring into focus the importance of such advances as mentioned above that have been made in research on sampling methods, conservation practice, techniques of exploration, automation in information management and in the development of efficient methods to obtain better documentation (3).

My particular concern in this program is with the documentation of all genetic resources because I am convinced that efficient management of the documentary information will not only effect monetary savings to all users of genetic materials, aid in the conservation of these resources and facilitate their fullest and proper use in agricultural practice. It is perhaps a fortunate point in history that we have access to automatic data processing machines with the potential and capacity to manage and usefully analyze documentary information on the anticipated large number of new collections from explorations now being planned, let alone deal with the documentation on the many thousands of stocks already maintained at genetic resource centres.

Standardization

Experience has shown us that to fully employ the power of computers, we will need to record in standardized terms and in an orderly and systematic way whatever information we deem important. Neither of these requirements need be restrictive to individuality or initiative, but may in fact simplify and reduce the effort involved in recording data. However, we have few internationally standardized methods on which to build, and the development and wide acceptance of methods recommended for standardization can cost precious time. Therefore, it is of the utmost importance that rapid progress be made in the development and testing of methods amenable to both standardization and computer processing.

A number of initial steps have been taken toward international standardization of crop research data recording methodology. Some of these steps were taken with the encouragement and sponsorship of FAO and IAEA, and various phases of the work are now well underway. In spite of the funding impediments this project has had, important advances have been made, and components of a system are gradually (though slowly) being realized. Those components particularly worthy of mention in this context are the pedigree illustration method published by Purdy et al. (5), the guidelines for disease data recording as presented by Loegering (12), the work of Browder (1) on the same subject. McNeal et al. (13) have described an advanced data processing system built according to an unpublished model designed by Dr. Keith Finlay (now with CIMMYT) and myself. Other versions of this original model are already widely used over the world. Peterson et al. (14) have more recently published also a procedure for using the computer to assist the management of early generation materials in a breeding program and Hudson et al. (5) have just described the use of

TAXIR, a fast, efficient, inexpensive information retrieval system to meet some of the needs of a plant introduction program. Several other works in progress are yet unpublished. Among these, standardized plant growth stage scales are being developed by Zadoks and others (personal communication) as are standard guides to the description of plant characters and disease or physiological reactions (C. F. Konzak et al, unpublished). Potentially standard methods of recording information on the origin and history of genetic materials are currently being tested for collections from plant exploration here at ARIC (6) and for induced mutants by IAEA in Vienna (personal communication).

In addition, much useful related work has been published or is in progress. This work and ideas freely offered from many sources have contributed measurably to the development of a concept for an international genetic resources information network. One segment of a working model system within such a network will soon be tested at Washington State University. The model is being built from the base upward, rather than the other way around, as will be described later. The goal is not merely to gather information into data banks; rather, it is to employ the data banks not only to accumulate and store useful information but more importantly to transmit this information quickly and freely to users. Many of the users will, in turn, be prime contributors of information in cyclic fashion to the system. Thus, the purpose of the system must be to facilitate the feedback of information with all its components.

Basic Documentation

The kind of records to be developed and/or shared by Genetic Resource Information Banks may be somewhat variable depending on the local, national,

regional and world wide importance of a crop, which, in consequence, will affect the rate at which both information and resource materials are acquired. However, the basic documentation required for all genetic resource materials, irrespective of their origin, is much the same (Table 1). The documentation should include appropriate identification in terms of accession numbers, taxonomic nomenclature and common or vernacular names. Details about history, sources, etc. of an accession will differ depending on the origin of stocks. Three basic origins of genetic materials are now known, including collections from explorations (which are our major concern here) as well as recombined (cross-bred) and/or selected materials and natural or artificially induced mutations. All of these materials are important and potentially unique resources and all belong in germ plasm collections.

As mentioned earlier, standard methods for constructing accession numbers, brief designations and pedigrees have been proposed (13,14). The International Code of Botanical Nomenclature has been recommended as a guide to naming. Proposed standard methods of recording documentary information based on the different origins of materials are under development (2,11). One model for a "standard" plant field collection record form was developed and extensively tested last year in explorations by the ARIC. An improved version of this form is currently in process (6). Because of our mutual interest here, a few further statements about the newly designed forms and handbook for documentation of collections are appropriate. The collection record forms are greatly simplified in design, recording procedure is easy and systematic, thus, readily amenable to direct transfer to machine readable punch cards or tape. Guides to the entry of information are presented and the forms are

bound in a spiral pocket size booklet for ease of use in the field. These handbooks may prove useful in training new collectors and in assuring universal communicability of the recorded information.

A "Documentation Dictionary"

Once the identity as to nomenclature, history and origin are established, the further documentation of genetic resource materials is essentially crop specific. Even so, the general format may be applicable to all crops and, therefore, at least the principles could be standardized. We have on several occasions suggested ways in which the requirement of universal intelligibility of description and evaluation data might be met (6,7,8,9,10,11) and recently Lothar Seidewitz and I had the opportunity to work together on an approach that may have wide application in the development of standardized documentation procedures (18). This approach incorporates much of earlier ideas, but avoids some of the difficulties encountered in efforts to develop crop specific recording forms. Our recent approach to the problem is to develop crop specific illustrated "documentation dictionaries" incorporating as well all of the procedures and guides proposed for international standardized documentation (including identification, description and evaluation) for a particular crop or crop group. Eventually, other facets of documentation should be added as, for example, procedures for genetic terminology which, in principle, should follow a standardized approach. In the plan for the "documentation dictionary" we side-stepped, at least temporarily, the problem of developing forms for recording information. This was done to some extent because it was difficult to establish which data would be most often recorded. Moreover, complex forms are often burdensome to cooperators. Also, we found that it is often more convenient

to acquire information in smaller units, and to avoid manual transfers of data which could lead to inaccuracies. The method also allows complete flexibility in the construction of recording forms, the key requirements being the accession number and, if available, a name or designation for data recorded on each accession.

The essential features of the identification section in the proposed "documentation dictionaries" already have been discussed. The method of recording descriptive data is similar, in principle, to a method suggested earlier. For each descriptive trait presented in order of plant parts, there will be a page presenting the recommended standard codes, standard terms for each plant part or descriptor, and for each character or descriptor state and each will be illustrated by a photograph or drawing. The precisely equivalent terms for the descriptor state can be assured in translations to other languages both via the codes and the illustrations. An example of this approach is shown in Fig. 1.

We propose that evaluation data be handled in a similar manner, though for each type of evaluation a transformation from raw to classified data would be necessary. As an example, the principles of recording and classifying evaluation data are illustrated in Dr. James Hoffmann's classification of bunt resistance sources. The classification is a simple transformation from experimental data recorded in percent of spikes per plot infected with the disease. Disease reaction classes are based on their estimated usefulness in plant improvement and cultural practice, an approach that is consistent with procedures proposed earlier by Loegering (12). Thus, three classes of bunt reaction can be distinguished based on the raw data (Table 2).

Table 2

Bunt Reaction Classes (Hoffmann)

<u>Class</u>	<u>Highest Recorded Percent Infection</u>	<u>Use Category</u>
R (Resistant)	0-10	useful
I (Intermediate)	10-30	questionable
S (Susceptible)	30-99	not useful

We found, however, that a fourth disease reaction category was necessary in practice because test populations often were mixed, since both resistant and susceptible lines could be isolated from them. An X code was thus used to identify this classification of the original accession. Reselections were given new accession numbers in a different series. It is Dr. Hoffmann's view that the intermediate class may also represent a mixed or heterogeneous population with respect to its reaction to a particular *Tilletia* culture but the I (intermediate) class does fit the criterion of uncertain usefulness.

The procedures followed in processing the bunt disease data illustrate how the proposed information management mechanism is developed from the actual experimental data base (Fig. 2).

Teamwork by Drs. G. W. Bruehl, B. Devecioglu, J. A. Hoffmann, R. J. Metzger* has shown new ways to use the classified data by joining them

*Department of Plant Pathology, Washington State University, Pullman, Washington; Leader of Regional Spring Wheat Research, Agricultural Research and Introduction Centre, Karsiyaka-Izmir, Turkey; Plant Pathologist, United States Department of Agriculture, Agricultural Research Service, Pullman, Washington; and United States Department of Agriculture, Agricultural Research Service, Cereal Crops Research Branch, Oregon State University, Corvallis, Oregon; respectively.

with estimates of their geographical origin based on the documentary information in the USDA Plant Inventory records. A crude analysis showed that many of the bunt resistant wheats originated in Turkey. Maps were then constructed using the computer to show the location of origin of all resistant wheats so far tested among the 2500 Turkish wheats in the USDA Collection as of 1969. This first analysis indicated that wheats with some bunt resistance were well distributed over the country. However, when the sites of origin of only the few accessions resistant to dwarf bunt and to snow mold were plotted, it was found that these resistance sources came almost exclusively from Eastern Turkey (Fig. 3) and nearby areas in Iran, Afghanistan and the USSR. The few highly resistant accessions from Western Turkey may have been transported there. One collection of *T. monococcum* could have had a similar basic origin. The basis for the resistance of the accession from Southwestern Turkey may be a distinct source but the genetic analyses such as those conducted by Dr. R. J. Metzger and colleagues may clarify this point (20). What we see emerging for the first time is a picture of the evolution of disease resistance genes of great importance in one of our major crops.

A similar kind of study conducted at CIMMYT in Mexico on stem rust resistance in durum wheats was recently described by Rajaram, Devocioglu and others (16). This study showed that 21 out of 349 durum wheats from the USDA World Wheat Collection were highly resistant to four differential cultures of stem rust (*P. graminis*). Of these, 20 came from the provinces of Mersin, Adana and Hatay in Turkey, and one from nearby Cyprus. Should this rust resistance prove to be monogenic, as was suggested, the information could mean important savings in practical breeding. The breeder might choose as a parent the accession with the best combination of useful traits

unless, of course, the accessions are basically all alike.

This use of documentation also may have an impact on tactics of evaluation as well as of resource collection.

Collation of Useful Documentation by Genetic Resource Centres

The identification, description and classified evaluations for the most part constitute the documentary information of highest value to users of genetic resources. In the case of plant collections, the identification and descriptive information might best be accumulated by the Genetic Resource Centre in the region of their origin. Evaluations might be made anywhere and at any time and should be entered in the genetic resources information retrieval (IR) system as they become available. Since evaluations are only as useful as can be estimated in relation to other data, the assignment of an experienced scientist as coordinator for certain kinds of data appears desirable as an intermediate link to the IR system. This type of approach is successful as employed by the USDA. The coordinator accumulates and processes the experimental data to provide the IR data bank with classified data. The classified data in the genetic resources IR system serves as a guide to the actual data. Access to the data and their source is through the coordinator who would maintain the machine readable raw data bases for access on request or according to need. We have suggested a possible scheme for coordinating the acquisition of information by a Genetic Resource Centre as in the USA (Table 4). Certain components of this scheme now exist (c. f. Fig. 3) and already serve the designated function.

International Regional Genetic Resource Information Centres

The establishment of International Regional Centres to conserve, manage and accumulate information about genetic resources will be a major forward step in the development of what should become an internationally coordinated Genetic Resources Information System. The fast, efficient and inexpensive information retrieval system such as TAXIR described earlier by Rogers (17) and the reference type retrieval system SOLAR, developed at Washington State University for the U. S. Forest Service, are examples of IR systems which might meet the needs of this program. As yet, however, only a limited amount of testing has been done using TAXIR and only a start has been made to test SOLAR for our types of use. Both systems have features we need but neither system has all of the capability we may need. TAXIR has already been used effectively at MSU to assist the activities of the local Plant Introduction Station with its services to users of the USDA Phaseolus Collection. Processing of user questions against a data bank of describing 26 characteristics of over 3500 bean accessions can be obtained at little more cost than would be required for typing a letter (Table 3).

A similar data bank has been constructed for maize accessions at Iowa State University, for Pisum accessions at Pullman and for an artichoke collection at Bari, Italy. We are also in the process of developing a TAXIR data bank for wheat.

The information component just described fits into a structure for information management at each Genetic Resources Centre (Fig. 5). The structure should be designed to facilitate the varied roles of the Centre

in the conservation of genetic materials and in the acquisition of information about those materials. Interaction among regional and national centres in the accumulation and sharing of information and resource materials should be developed not only to better serve the needs of users of the resources but also to assure that no resource materials will ever again be lost.

We have here at Izmir the basis for a complete model of such a system and particularly so if the ARIC is developed as a Regional Genetic Resources Centre. A functioning complete model system is needed to learn and resolve the technical and other problems that must be met as a world wide information network is developed. With the additional Genetic Resource Centres in Bari, Italy and Braunschweig, Germany and proposed for several other locations in the world there is an obvious need and opportunity for joint participation in studies to develop communication systems. The Izmir Centre is currently limited somewhat because of the rather small computer at Ege University. This is important primarily to the efficient retrieval of information as with the TAXIR and SOLAR information systems but a factor which should be considered by the Turkish Government in relation to the future development of the Agricultural Research and Introduction Centre and Ege University.

I might add that I found the cooperation among units of the ARIC and the Computing Centre at Ege University to be excellent. Progress is being made in making machine readable the documentary information on the germ plasm collections at the Centre. However, I would like to see some increased emphasis given to this work in the next year since many technical

aspects of data management remain to be tested in practice.

SUMMARY AND CONCLUSIONS

Documentation increases manifold the value of genetic resource collections. Procedures for improving documentation should be adapted for computer processing and based on uniform principles for all crop species. International standardization of data recording procedures would assure universal intelligibility of genetic resource documentation. Considerable progress has already been made in the development of principles and methods for recording data and on the development of information management systems adapted for, but not requiring, computer processing. The merging of information from several sources may be appreciably facilitated by the use of computing machines; more important, the speedy processing and merging of information by computers can be of assistance in the planning of explorations for new resource materials, in the conservation of genetic resources and in facilitating their evaluation and fuller use in agricultural practice.

Plans for an international genetic resources information network were developed and small scale models have been tested in practice. Development of more complex systems will still be necessary to identify and resolve problems expected in an international system. The development of the ARIC at Izmir as a Regional Genetic Resource Centre with capability for computer assisted information management would provide the opportunity for further progress toward a viable international network of Genetic Resource Centres.

Table 1

Basic File	
IDENTIFICATION	- identification number, key taxonomy, name, cross-reference number, type of origin, etc.
DESCRIPTION	- morphology, key identification traits
EVALUATIONS	- classified disease response data, quality evaluation, etc.
Secondary File(s)	
GENETIC ORIGIN	collection in nature - collection record - breeding and selection - parent(s), pedigree(s), breeder(s) mutation - parent stock - originator - treatment, etc.
SOURCES	- other holders, synonyms
GENETICS	- genes identified or probable and reference number
GENETICS LIBRARY	- references + cross-index to stocks and information type, raw data sources (coordinators)
Special Duplicate Sub-Files With Complete Records	
CULTIVARS	- as per needs for variety identification
SPECIAL GENETIC AND BREEDING STOCKS	- elite disease and insect resistance sources, substitution lines, genetic testers, etc.
MAINTENANCE RECORDS	- quantities of stocks, viability records, reproduction records, distribution records, etc.

Table 3

Sample Output from Questions to TAXIR Data Bank on
USDA Phaseolus (Bean) Collection Maintained at Pullman, Washington

```

ID          TAXIR      *
READ DATA BANK 1 *
ID          TAXIR      *

QUERY WAYNE S. TATE, PULLMAN, WASH. PRINT:
PLANT HABIT, PLANT SIZE, PLANT STEM COLOR, FLOWER COLOR,
POD LENGTH, POD CURVE, POD CROSS SECTION,
POD FIBER, POD STRING, SEED COLOR FOR ITEMS
WITH POD COLOR, WX AND POD TYPE, SN AND SEED SIZE, M*
RESPONSE:
ITEM NO. REQUESTED STATES
110233 V UNKNOWN G MX 3 3 3 4 - RBR
136702 B M G PU PI 4 3 3 4 DT
169730 B V S G PU PU 4 3 3 4 DB
226936 B M S G PU PU 4 3 3 4 LT
226937 B M S G G W LA 4 2 2 3 W
226938 B M S G G W LA 4 4 4 3 UNKNOWN DT
251049 B B M G PU PU 4 4 4 4 B
285698 B M S G G W W 4 4 4 1 T
285699 B B S G G W W 4 5 3 3 LT
285700 B B S G G W W 4 5 3 3 LT
285701 B B S G G W W 4 5 3 3 LT
285702 B B S G G W W 4 5 3 3 LT
285703 B B S G G W W 4 5 3 3 LT
285704 B B S G G W W 4 5 3 3 LT
285705 B B S G G W W 4 5 3 3 LT
285706 B B S G G W W 4 5 3 3 LT
285707 B B S G G W W 4 5 3 3 LT
285708 B B S G G W W 4 5 3 3 LT
285709 B B S G G W W 4 5 3 3 LT
285710 B B S G G W W 4 5 3 3 LT
NO. OF ITEMS IN QUERY RESPONSE = 13
END*

```

```

ID          TAXIR      *
READ DATA BANK 1 *
ID          TAXIR      *

QUERY FOR C. F. KONZAK, PULLMAN, WASH. LIST ITEMS WITH
PLANT HABIT, B AND
PLANT SIZE, M AND
PLANT ERRECT, 4 AND
FLOWER COLOR, PI AND
POD LENGTH, 3 AND
POD CURVE, 2 AND
SEED COLOR, RT*
RESPONSE:
ITEM NO. REQUESTED STATES
172018 V S 3 G 2 3 2 M UNKNOWN
NO. OF ITEMS IN QUERY RESPONSE = 1
END*

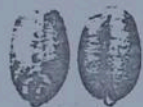
```

Table 4
POSSIBLE SCHEME FOR COMPILATION OF INFORMATION
ON WHEAT - (USA)

<u>INFORMATION TYPE</u>	<u>COORDINATOR</u>
IDENTIFICATION	MANAGER
DESCRIPTION	CULTIVARS, BY VARIETY REVIEW BOARD OTHER STOCKS BY MANAGER
EVALUATIONS	BY TYPE OF KEY CHARACTERISTIC
QUALITY	REGIONAL LABORATORIES
AGRONOMIC ADAPTABILITY	REGIONAL COORDINATORS
DISEASE AND PEST RESISTANCE	
RUSTS	IRN COORDINATOR, BELTSVILLE
MILDEW	REGIONAL CEREAL DISEASE LABORATORY
BUNT	REGIONAL CEREAL DISEASE LABORATORY
OTHER	AS APPROPRIATE
GENETICS	WHEAT GENETICS COORDINATOR(S) BY TRAIT OR CHROMOSOME ?
PHYSIOLOGICAL	?
ORIGIN	MANAGER OF STOCKS
STOCK SOURCES	MANAGER OF STOCKS

GRAIN SHAPE

DORSAL, VENTRAL VIEWS



ROUNDED

1



ROUNDED-OVATE

3



OVATE

5



OVOID ELONGATE

7



ELLIPTICAL

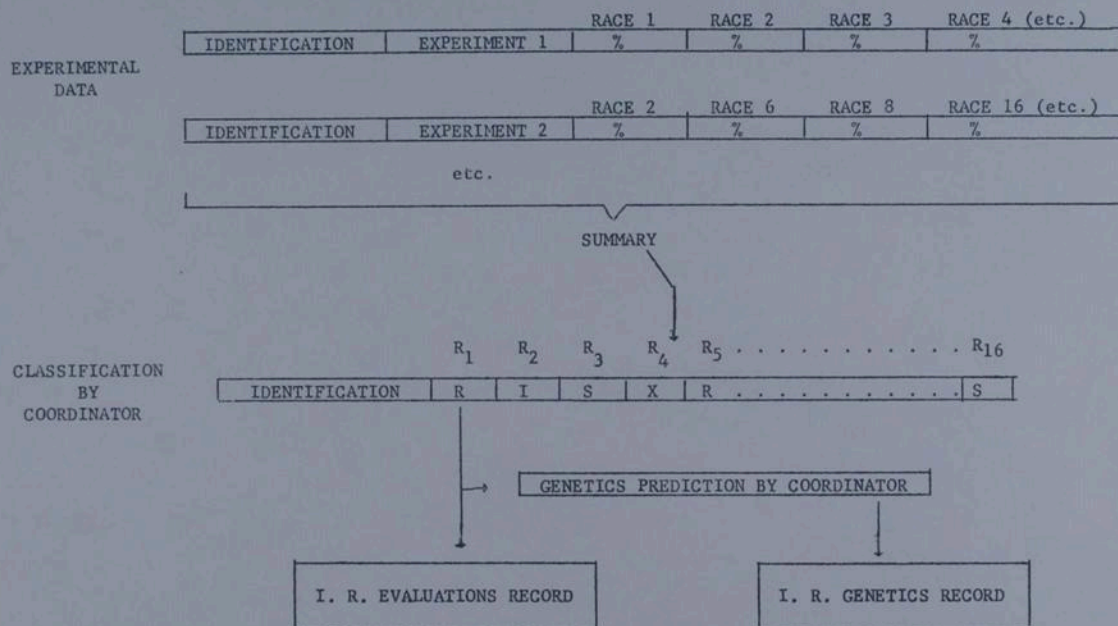
9

Code

Figure 1. Design of descriptor units in planned "documentation dictionary" in preparation.

Figure 2

DATA MANAGEMENT SCHEME
(EXAMPLE: BUNT DISEASE DATA)



TURKEY DWARF BUNT RESISTANCE

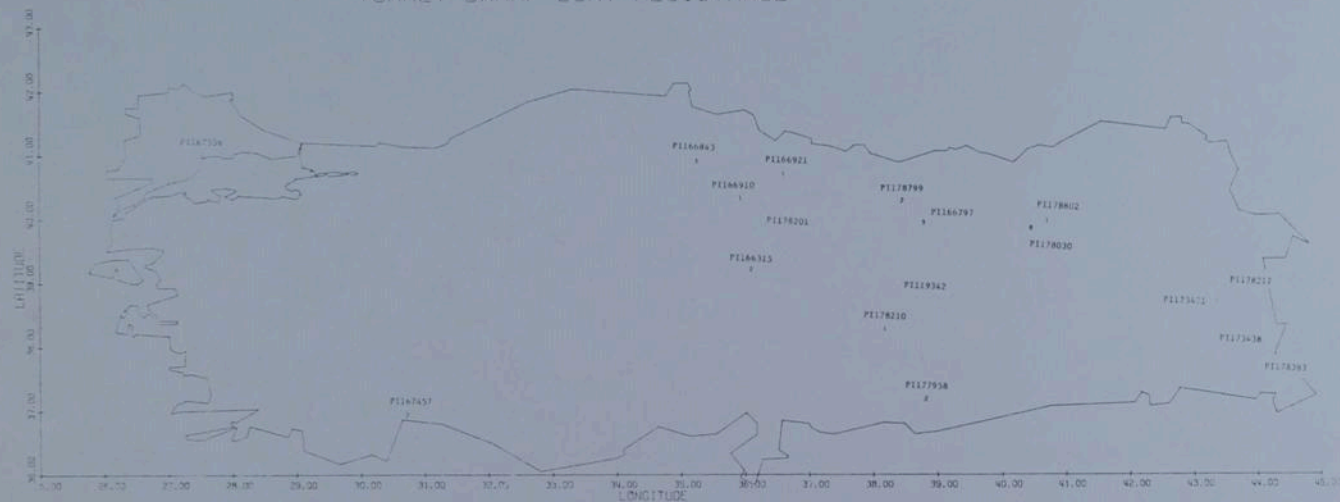


Figure 3. Computer plot showing the geographical distribution of dwarf bunt resistant PI wheat accessions among the USDA World Collection of Turkish Wheats, based on data of Hoffmann, 1972, unpublished.

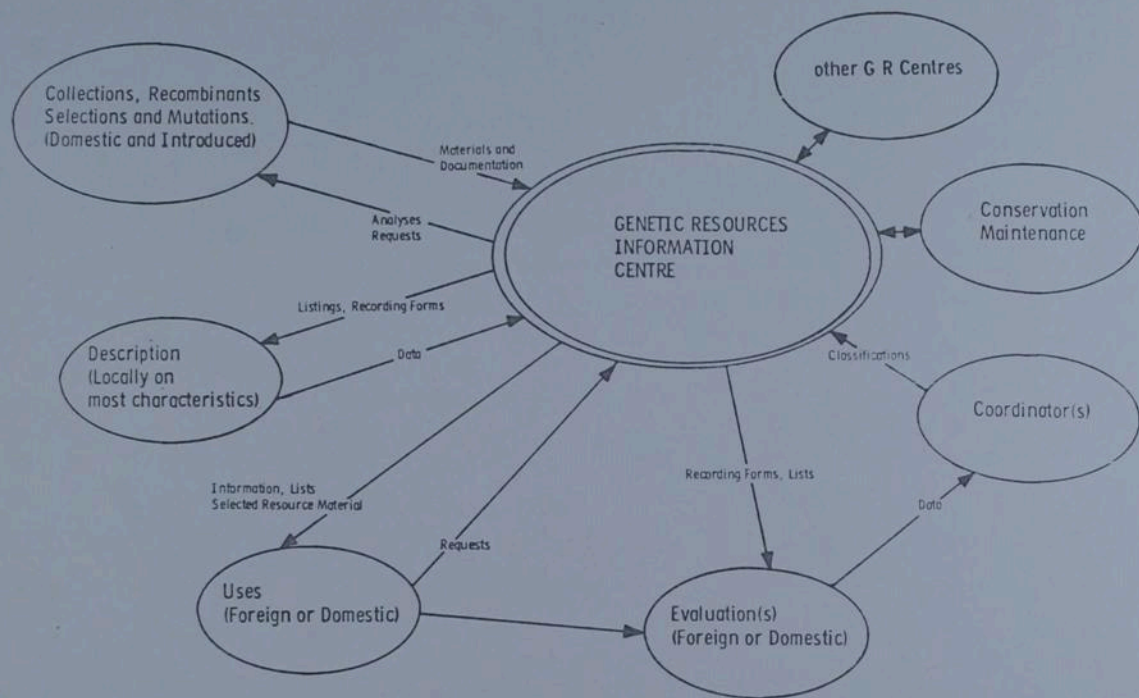


Figure 4. Flow chart illustrating cyclic flow of information and materials between component units associated with national and international regional Genetic Resource Centres.

REFERENCES

1. Browder, L. E. Pathogenic specialization in cereal rust fungi, especially *Puccinia recondita* f. sp. *tritici*: concepts, methods of study, and application. USDA Technical Bulletin 1432.
2. Finlay, K. W. and C. F. Konzak. 1970. "Information storage and retrieval". In Genetic Resources in Plants - Their Exploration and Conservation. O. H. Frankel and E. Bennett, eds. F. A. Davis Co., Philadelphia, Pennsylvania. pp. 461-465.
3. Frankel, O. H. and E. Bennett, eds. 1970. Genetic Resources in Plants - Their Exploration and Conservation. IBP Handbook No. 11. F. A. Davis Co., Philadelphia, Pennsylvania. 554 pp.
4. Harlan, J. R. 1971. Agricultural origins: centers and non-centers. Science 174:468-474.
5. Hudson, L. W., R. D. Dutton, Mary Massara Reynolds and W. E. Walden. 1971. TAXIR - A biologically oriented information retrieval system as an aid to plant introduction. Economic Botany 25(4):401-406.
6. Konzak, C. F. 1972. Conference on European and Regional Gene Banks, Izmir, Turkey (Eucarpia). "Documentation of collection and information management by genetic resource centers".
7. Konzak, C. F. 1966. Report of an expert group which met to consider international standardization, integration and mechanization of crop data recording and processing. In Mutations in Plant Breeding, International Atomic Energy Agency, Vienna.
8. Konzak, C. F., G. Delhove and B. Sigurbjörnsson. 1968. "A world plant genetic resources information system". In Proceedings of the XII International Congress of Genetics, Tokyo, Japan. (Abstract) Vol. I, p. 259.
9. Konzak, C. F. and S. M. Dietz. 1969. Documentation for the conservation, management, and use of plant genetic resources. Economic Botany 23(4):299-308.
10. Konzak, C. F. and Björn Sigurbjörnsson. 1966. "International cooperation in standardization of procedures in crop research data recording". In Proceedings of the Fifth Yugoslav Symposium on Research in Wheat, Novi Sad, IAEA, 691-696.
11. Konzak, C. F., W. E. Walden and F. P. Sousa. 1970. Problems and progress in the management of information on genetic resources. Seiken Zihō 22:91-97.

12. Loegering, W. Q. 1968. Problems of storing plant disease data for retrieval by automatic data processing machines. Washington Agricultural Experiment Station, College of Agriculture, Washington State University, Bulletin 705, 6 pp.
13. McNeal, F. H., C. F. Konzak, E. P. Smith, W. S. Tate and T. S. Russell. 1971. A uniform system for recording and processing cereal research data. USDA Bulletin ARS 34-121. 42 pp.
14. Peterson, C. J., O. A. Vogel, R. E. Allan and W. S. Tate. 1971. Systematic management of early generation progenies. Crop Science 11:304-305.
15. Purdy, Laurence H., W. Q. Loegering, C. F. Konzak, C. J. Peterson and R. E. Allan. 1968. A proposed standard method for illustrating pedigrees of small grain varieties. Crop Science 8:405-406.
16. Rajaram, S., B. Devecioglu, N. E. Borlaug, K. W. Finlay, G. Varughese and E. Rodriguez. 1970. A survey on stem rust resistance in the USDA Durum Collection and in CIMMYT Durum Breeding Lines. CIMMYT Research Bulletin 14:1-25, International Maize and Wheat Improvement Center, Mexico, D. F.
17. Rogers, D. J. 1972. Conference on European and Regional Gene Banks, Izmir, Turkey (Eucarpia). "The concepts of applicability of TAXIR for storage and retrieval of gene bank data".
18. Seidewitz, L. 1972. Conference on European and Regional Gene Banks, Izmir, Turkey (Eucarpia). "The documentation system of the German Gene Bank in relation to international standardization".
19. Vavilov, N. I. 1949/50. The Origin, Variation, Immunity and Breeding of Cultivated Plants. Translated by K. S. Chester. Chronica Botanica, Waltham, Massachusetts. Vol. 13, 364 pp.
20. Waud, J. L. and R. J. Metzger. 1970. Inheritance of a new factor (Bt8) for resistance to common bunt in wheat. Crop Science 10:703-704.

STANDARDIZED DOCUMENTATION PROCEDURES FOR GERM PLASM COLLECTIONS

C. F. Konzak

INTRODUCTION

In December 1965, a group of consultants were called together in Vienna under the joint sponsorship of the FAO and IAEA to consider "International standardization, integration and mechanization of crop data recording and processing."

As a result of their deliberations these consultants recommended that an international working group be established to investigate areas of crop research data recording in which international standardization of methods might be encouraged. The original consultants group set down a number of guidelines to be followed, and a Working Group on International Standardization in Crop Research Data Recording was initiated under the joint sponsorship of FAO and IAEA in early 1966 with this author as coordinator (3, 4).

Among the several general areas recommended for international standardization was that concerning the documentation and description of genetic resource materials. The group envisioned that the use of computers might provide the basis for a major breakthrough in the management of information on genetic resource materials--a breakthrough that could lead to greater crop research efficiency and international cooperation, and if appropriate actions were taken, greater standardization might be achieved by recommending data recording methods adapted to automated processing.

When one considers the enormity of the task to which the working group was dedicated, it is perhaps little wonder that just those few years ago it was considered by many to be far too complex a subject to bring about change.

However, as many of you now know, changes are occurring and computer processing of crop research data is already common in many parts of the

world and expanding rapidly. In fact, many of the guidelines established by the 1965 Consultants Group and by the later sessions of members of the working group have been applied in the development of automated data management systems.

Concepts of an international network for the accumulation and exchange of information on genetic resources have developed as the result of much sorting and testing of ideas. The foundation now appears established for initiating activities using automated data management both on a regional and international basis in conjunction with programs to systematically collect, conserve, evaluate and more efficiently utilize genetic resources.

Although there are a number of areas in which studies of potentially standard methods are under investigation, I would like to place emphasis here on the basis for and the development of methods for standardized documentation of germ plasm materials collected from their natural habitats.

RESULTS AND DISCUSSION

There are many reasons for encouraging the standardization of documentation on plant germ plasm collections. A number of these have been discussed in recent publications (1, 2, 5, 6, 7, 8, 9, 10). It may be sufficient to point out that without a measure of standardization, much of the documentation on germ plasm collections may be hardly worth the effort if the information is not accessible, and to make the documentary information efficiently accessible, as via computer retrieval, a measure of standardization is imperative. Standardization might help in the training especially of local plant explorers via FAO, UNDP, IBP and other programs. Through these developing programs in plant exploration we also have the opportunity to acquire quantities of data from which some meaningful analyses can be derived, and

through use of automated data processing initially gathered information might be used to guide further explorations. Furthermore, the documentation need no longer be laid to rest in archives, but could be put to use to aid and expedite the evaluation and full utilization of the new collections. At the least, the documentation helps to prevent loss of identity, and provides the user with some fragments of information by which he may select and sample materials for evaluation, when, as is usually the case, his resources are limited.

To illustrate the kind of use that might be made of documentary information, when Dr. Devocioglu was an FAO Fellow at WSU, he and Drs. Hoffman and Bruehl constructed a map showing the collection sites of wheats carrying high resistance to dwarf bunt. The map was transmitted to the ARIC in Turkey, and proved useful to stimulate further collections of wheats in the critical areas. I understand that as a direct result, many new wheat collections were recently sent to the U. S. These will be entered into disease screening trials, and while there may still be time for even further collecting, it may be possible for the pathologists to develop a better concept of the evolution of genes for resistance to bunt. The results of these studies are also of considerable value to Turkey, in part because of the identification of useful resistance sources, but also because breeders use these resistance sources and commonly make available their recombinants to assist the Turkish variety development programs.

Much of the important documentation of germ plasm collections must be made in the field. To do this efficiently, a standard field recording form not only facilitates data collection but also facilitates the processing of both the collections and the documents. In fact, from the standpoint of processing, variability in the composition as well as in the order with

which documentary information is presented is a nightmare if you are attempting to make it machine readable.

Our research to develop a standard form for use in the field by plant explorers to record important documentary information on germ plasm collections was first initiated in 1966. Since then there has been input of ideas and experiences from a great many interested and renowned scientists. A first attempt at the design and construction of proposed standard field collection forms adapted for easy transfer of information to machine readable records was made at Pullman in (1969-70) with the collaboration of Dr. Basri Devecioglu. Reviews of this form were mixed. Many plant collectors liked the general idea, shape and other features of the forms but several of the experienced collectors objected to some of the detail incorporated as per suggestions collated from earlier Working Group meetings. Although many of you probably have seen those forms it is doubtful many ever were tested in the field. A revision of that form was made in June 1971, at Izmir. To illustrate for you the effectiveness of the teamwork between staff of the Agricultural Research and Introduction Centre and of Ege University, it is significant that the drawings were completed, photographed and the forms printed all within a period of 10 days. Dr. Ebbe Kjellqvist was able to conduct a preliminary test for collecting in the field in the Istanbul area during the following two weeks. The forms were used exclusively during explorations to Eastern Turkey in the fall of 1971 and proved to be quite satisfactory. However, we have since collected additional suggestions to increase the acceptability of the forms to experienced collectors and some adjustments would also further facilitate transfer of the information to machine readable records. These points will be described later. Although one does not like to deal with frequent changes,

we believe that something of real value is evolving as we make improvements to the forms. Besides, each advance is an improvement considering the highly variable records on most earlier collections.

Since our aim in this project is to develop a standard record form which may also prove useful in the IBP sponsored training programs for plant collectors, we will describe the proposed standard field collection record form with the planned modifications. Suggestions for further improvements will be appreciated.

The overall basic composition of the proposed standard documentation form has not changed since the first construction. What has changed has been the order in which different items of information are to be recorded, the detail to be recorded and the completeness of instructions. Figure 1a shows the identification section, including the collection and assigned accession numbers, taxonomic and common names. In this section, space is reserved for an eight space coded gene bank (GB) accession number, as is space for an eight digit uniform (UN) number. Neither of these will be entered by the plant explorer, the assignment is made at the conservation center. The uniform accession number is intended to provide for coordination between gene banks to avoid unintended duplications. The spaces for the collection number, however, are filled in by the plant explorer. The collection number, based on a design used extensively by FAO and UNDP personnel, is composed of a code (usually two initials of the explorer), the date of collection, and a number constructed of a sequential site number for that date and a sequential collection number among collections made at that site. The method is flexible enough to accommodate the preference of those experienced collectors who may wish to use their own sequential collection numbering system in place of the site and sample numbers. Using

the sequential numbering method, a collector could number his record pages and thus his collections ahead of time.

Some problems have been encountered in assigning adequate but reasonable amounts of space for taxonomic names, allowing for flexibility without overburdening the collector at the same time. We first considered coding plant names, but then decided that approach was too complicated and required too much of the collector. With advances being made in computer storage capacity, it is likely that most coding will be unnecessary, even if expedient for efficient use with certain information retrieval systems. Besides, we want the computer to work for us, not the other way around! The only important criterion is to provide the vehicle for complete identification. This probably can be done conveniently by abbreviating family and genus to their first 8 letters, but for species and subspecies or botanical varieties it appears that compound names are common so up to 16 spaces may be needed for a positive identification. Even so, all the collector need write down is enough of the name to permit either the computer or the documentation staff to identify it. While we did not think it necessary to require the collector to enter the authority for the species names, extra space is available should he wish to do so. Most likely each collector will follow one or few taxonomic references. In cases where positive identification is not possible the collector should provide pertinent information in the text section discussed later.

Cultivar names, local designations, are also useful identification and 24 spaces is allowed as a reasonable compromise in space allocation. Longer names are to be abbreviated and presented in complete form in the text section described later, as are additional names or synonyms. The cultivar

name and accession number are useful in many of the data management operations that follow in the maintenance and evaluation of germ plasm stocks, thus a reasonable space limit for these terms is advantageous.

The second section of the record form identifies the collection in a general way, the kind of material collected and whether its population basis is genetically narrow or broad, i.e. whether a single or few plants were collected, whether an attempt was made to obtain a broad based representative sample, or if the population base is unknown, as might be the case for materials collected at bazaars, for example.

If herbarium specimens or photos are taken these are to be noted, and it is useful to know how abundant the material was at a given site. The months in which growth begins and stops is of interest particularly in woody species, which often tend to retain their growth season even when moved to another environment. A quick general check to describe the source or place the actual sample was obtained has been added in order to meet a variety of situations. In many cases, the record may sometimes obviate recording some of the detail described later. The identity of the particular expedition and team as well as the year and expedition leader (Fig.1b) when applicable would facilitate efforts to recognize and justify the value of plant explorations, by way of collected materials used later in plant improvement programs. The practical significance of this use should not be underestimated. However, that section and all following sections carrying information about the site and locale need be entered only once since the computer can more easily associate the records of all samples collected at a given site. With the proposed standard system of collection numbering the processing is easier and more straight forward.

The next section, dealing with the site description and location, is of real importance for relocating collection sites and for mapping and survey analyses of collected areas. This section includes a site description in terms of latitude, longitude in compass degrees, and altitude as well as the aspect or direction of slope and the slope in degrees from horizontal. Most collectors will carry a compass and altimeter and it is easy to estimate the slope with sufficient accuracy. The latitude and longitude or map location can be entered at base camp at the end of the day or at the laboratory. Aspect and slope may provide useful information on which to guess the ability of the collected material to prosper under perhaps adverse conditions of exposure, for example.

To supplement the map location data, most collectors like to record the country, province, place, distance to the site, and other pertinent information. A computer retrievable section was added for that purpose. The final section allows for entry of additional information the collector considers might be pertinent and useful. This completes the first page and most essential part of the documentation.

The reverse side of the documentation record (Fig.2a) carries information describing the topography of the area and the edaphic character of the collection site, a record of notable disturbance factors, the vegetation of the surrounding area, and a description of the plant community in terms of additional species in the vicinity of the collection site.

This part of the documentation record has been worked over several times, mainly to simplify the recording and make the job easier for the collector. Many accepted internationally standard systems were found to be too complex and would consequently be ignored by the collector. Ways must

be found to make the task simple and interesting because if we are ever going to be able to associate plant adaptation with edaphic or environmental conditions, sufficient data will need to be collected to permit a meaningful study. As the information need be recorded only once per site and the categories are presented systematically and in simplified form, most collectors may be willing to provide the more complete documentation.

Among the edaphic features considered important as accessory information is a description of the soil on which the particular collection was growing. Several approaches were suggested, and an effort to apply an international soil composition chart proved reasonably acceptable in the test but some other collectors still thought it to be too complicated. We propose, now, a greatly simplified system of soil textural classes commonly employed by soil conservation service personnel. The depth of the "A" horizon and estimated total depth of the soil profile are considered important parameters and the collector should have no difficulty with them. The stoniness categories are also among the international soil description standards and should cause no difficulty. Facility to note the presence of large rocks was recently added since quite fertile soils may not be cultivated because rocks make cultivation impractical. The drainage situation at the collection site may greatly affect plant adaptation and should be recorded. Soil pH may be more difficult to estimate, but accuracy is not needed. Some collectors may bring along litmus but the intent of the record is mainly to distinguish unusual situations usually evident in the growth and communities of plants,

in the occurrence of minerals, and sometimes in soil color. The collector usually will know the general situation before he collects. The same considerations are true for the record on salinity. Most often salinity is virtually absent or low but medium to high saline conditions would be useful notes. Soil color description is useful information. The greatly simplified recording procedure should not burden the collector. We might hasten to add that the collector's notes are not intended to suffice for a climatological and ecological survey, but often the local conditions noted by the collector are not identified in any survey records.

In the next section, a number of common disturbance factors are listed along with a code number to indicate the intensity of disturbance recognized by the collector. Disturbance factors need be checked only if evident to the collector.

The last two sections (Fig.2b) permit the collector to describe in somewhat greater detail the ecological situation at the collection site and surrounding area. A good simple classification of climates and vegetation types has been hard to find. We tested what was a suitable classification for the Mediterranean area but found that the terms are at least not applicable in other areas.

To meet the more general need we are suggesting a modification of the systems developed by Trewartha and Küchler as variations of the earlier Köppen classification. In these systems, general climate types and subtypes are listed and within each, the main vegetation types and subtypes can be described rather simply. Guides for recording will be provided with the instructions.

In the final section, space is allocated for the recording of species recognized in plant communities of the collection site and adjacent areas.

We felt that this approach avoided the necessity for the collector to know the terms for particular plant successions, etc., and in fact, a knowledge of the exact species in association with the material collected would have more real value. The computer could be used to categorize types of plant communities.

Those collectors who like to keep complete documentation with their samples may use carbon paper between two record sheets, tearing off one to go with the sample. However, separate small tabs for the collection and herbarium specimens are provided in perforated sheets included either in the collecting record book or in a separate book (Fig. 5). The separation of tabs and record forms often is desirable in case an assistant is available to fill out the tabs while the collector records the documentation.

SUMMARY AND CONCLUSIONS

This report describes a proposed standard field collector's recording form which if more generally used, should help to improve the quantity and uniformity of the documentation of collected materials. But the information recorded by the plant explorer is only part of the documentation considered essential for genetic resource materials. Sufficient descriptive (morphological and other) information must also be acquired to help the genetic resource center to maintain the identity of the stocks. With increasing numbers of stocks, it will become increasingly more difficult to avoid mixups, possibly resulting in the loss of valuable materials. Unfortunately, few genetic resource materials are yet adequately described and standardized guides to facilitate the description of these materials are urgently needed. Cooperative work at our laboratory, in Germany, and elsewhere is proceeding toward that end with already good agreement on basic principles.

Also required is a certain amount of "housekeeping" documentation at each of the centers to assure that the materials do not lose viability or are otherwise lost and for a similar reason it will be important not only to have at least duplicated preservation of resource materials, but also to readily establish where materials are duplicated in order to facilitate the correction of mistakes and to reestablish stocks of materials inadvertently lost.

Computers can prove extremely valuable in these efforts to manage documentary information on genetic resources. Standardization is the key to the potential of this powerful tool.

ACCESSION UN NUMBER		COLLECTION NUMBER	
(OFFICE USE)		TEAM NUMBER	
CODE NUMBER		DATE DAY MO YR SITE SAMPLE	
FAMILY		GENUS	
SPECIES			
SSP. OR BOT. VAR.			
LOCAL NAME			
COLLECTION			
SAMPLE		HERBARIUM	ABUNDANCE
1 SEED N-POP 2 SEED B-POP 3 SEED U-POP 4 POLLEN N-POP 5 POLLEN B-POP 6 POLLEN U-POP 7 VEG N-POP 8 VEG B-POP 9 VEG U-POP 1 SEED 2 POLLEN 3 VEG 4 FLOWERS 5 PHOTO 6 SOLID STAND		1 V SCARCE 2 SCARCE 3 INFREQUENT 4 FREQUENT 5 V FREQUENT 6 SOLID STAND GROWTH INITIATION M1 MATURATION M2	
		SOURCE	
		1 NAT. HABITAT 2 ROAD SIDE 3 FARM STORAGE 4 MARKET 5 EXPR STATION 6 "WASTE" LAND 7 CULTIV LAND 8 OTHER	
EXPEDITION TEAM			
COUNTRY		ORGANIZATION	YEAR EXPEDITION LEADER
SITE DESCRIPTION			
LATITUDE	LONGITUDE	ALTITUDE	ASPECT SLOPE
DEGREES MIN	DEGREES MIN	METERS	DEGREES DEGREES
SITE LOCATION			
COUNTRY	PROVINCE		
PLACE	OTHER INFO:		
OTHER INFORMATION (ADDITIONAL DATA, NOTABLE PATHOGENS AND PESTS OF/IN VICINITY OF COLLECTED SPECIES)			

Figure 2. (a) upper and (b) lower parts of back side of proposed standard collection record form.

COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME	COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME
COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME	COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME
COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME	COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME
COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME	COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME
COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME	COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME
COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME	COLLECTION NUMBER TEAM _____ NUMBER _____ CODE DAY MO YEAR SITE SAMPLE _____ SCIENTIFIC NAME

Figure 3. Sheet of tear-tabs for placing with collections, herbarium specimens, photos, etc.

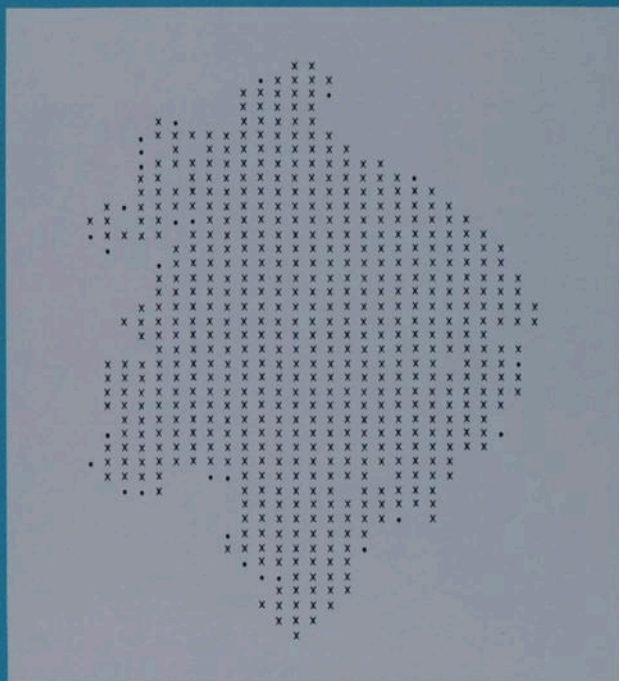
REFERENCES

1. Hudson, L. W., R. D. Dutton, Mary Massara Reynolds, and W. E. Walden. 1971. TAXIR - A biologically oriented information retrieval system as an aid to plant introduction. *Economic Botany* 25(4): 401-406.
2. Finlay, K. W. and C. F. Konzak. 1970. "Information storage and retrieval" *In* Genetic resources in plants - their exploration and conservation. O. H. Frankel and E. Bennett, eds. F. A. Davis Co., Philadelphia, Pennsylvania. pp.461-465.
3. Konzak, C. F. Report of an expert group which met to consider international standardization, integration and mechanization of crop data recording and processing. "Mutations in plant breeding", International Atomic Energy Agency, Vienna, 1966.
4. Konzak, C. F. and Björn Sigurbjörnsson. International cooperation in standardization of procedures in crop research data recording. Fifth Yugoslav Symposium on Research in Wheat, Novi Sad 12-18 VI 1966, IAEA, 691-696.
5. Konzak, C. F., G. Delhove and B. Sigurbjörnsson. A world plant genetic resources information system. XII International Congress of Genetics. August, 1968. Tokyo, Japan. (Abstract) Vol.I, p.259.
6. Konzak, C. F. and S. M. Dietz. 1969. Documentation for the conservation, management, and use of plant genetic resources. *Economic Botany* 23(4): 299-308.
7. Konzak, C. F., W. E. Walden, and F. P. Sousa. 1970. Problems and progress in the management of information on genetic resources. *Seiken Zihō* 22: 91-97.
8. Loegering, W. Q. 1968. Problems of storing plant disease data for retrieval by automatic data processing machines. Washington Agricultural Experiment Station, College of Agriculture, Washington State University, Bulletin 705, 6 pp.
9. McNeal, F. H., C. F. Konzak, E. P. Smith, W. S. Tate, and T. S. Russell. 1971. A uniform system for recording and processing cereal research data. *USDA Bulletin ARS* 34-121. 42 pp.
10. Purdy, Laurence H., W. Q. Loegering, C. F. Konzak, C. J. Peterson, and R. E. Allan. 1968. A proposed standard method for illustrating pedigrees of small grain varieties. *Crop Science* 8: 405-406.

A Computer - Mapped Flora

A Study of the County of Warwickshire

D. A. Cadbury, J. G. Hawkes and R. C. Readett



Published for the Birmingham Natural
History Society by Academic Press

A progress report on computerisation of grass information
in Kenya

by

U.H. Gullberg
Assoc. Expert Geneticist
UNDP/FAO Range Management Project
Kenya

In planning for future projects in agriculture, the Kenyan government considered the problem of genetic resources in general, and more specifically has planned exploration activities for grasses and legumes. This is a natural follow-up of the work that already has been done in exploration, evaluation and utilisation of genetic resources of plants in the country. The work on computerisation of grass information that is presented here is one of the activities now going on in this field.

Ideas on storage and retrieval of information related to genetic resources have been evolved in other countries, and it has been considered advisable to introduce and apply those ideas in Kenya. The work done on genetic resources information in IAEA, at the Washington State University and at FAO has thus taken into consideration when this system of handling grass information has been set up.

The system built up is an information storage and retrieval system, based on computer processing, and its information will mainly be variety data that is needed in the work with genetic resources. More specifically, the users of the information are considered to be plant explorers, genetic resources centres or equivalent institutes and plant breeders. Initially it is Panicum maximum information that has been processed, but the system could be applied on any crop. In order to include as much as possible of existing and future information on varieties one of the main objectives has been to make the system flexible.

It is envisaged that one body, here called a documentation centre, will execute the different operations involved in the processing of the information. It will handle the input of information, keep it up-to-date, and help the users with their requests. The documentation centre will furthermore keep contacts with other documentation centres for the exchange of information and for the co-ordination of activities.

The documentation or information storage and retrieval system

For the processing of information a documentation system has been built up. It consists of a number of files each containing one type of information. The most important and extensive file contains variety information with one document for each accession or variety. This information is computerised. Information on station data, references and synonyms is kept in separate files, and is so far not computerised.

The structure of a document for one accession, and the type of information it includes, is shown by its data base (Appendix 1). As seen from the appendix the data base is

divided into headings: Identification (the first part), Maintenance, Origin and Observations. Under each heading there are a number of items or fields that take a certain type of information. The length of the information under one item must be less or equal to a given maximum size and it will be classified and described by descriptions represented by codes, or it will be in the form of self-explanatory text or as numerical values. The number of items in the present data base are 700, with maximum sizes varying from one to 210 characters. The total length of a document, if all items are of maximum length is 9542 characters.

It is considered most suitable to have one data base for all variety information, instead of using different data bases for different crops. The inherent characteristics of a crop specific data base that could be useful, that is, descriptors specifically suited for the crop and automatic sorting of information by crop, could easily be built up into a system with one data base. Crop specific descriptors could be established within this data base and the information could be sorted if necessary. The use of one data base will furthermore take away some of the disadvantages with crop specific data bases such as the tendency to get crop specific descriptors and codes when general solutions could do, or the necessity to store and retrieve information from one crop at a time.

The smaller separate files, mentioned above, have documents that are built up in the same way as an accession-document but so far they consist of very few items.

The routine work with the documentation system will be:

- a. to put information related to genetic resources into the system;
- b. to control this information and keep it up-to-date;
- c. to reply to requests made by users.

a. The input of information will be governed by the data base. It shows how the information related to a variety should be sub-divided into items, so that the same type of information is stored at one well-defined place for all varieties.

When the information has been collected and recorded independently of the data base, the documentation centre will check how it has been recorded and what type of information has been recorded before it is put into the documentation system. If necessary the information will be restructured in order to be compatible with the data base. An example of this procedure is the transferring of grass-information from the files at Kitale National Agricultural Research Station to the documentation system. The recorded information was checked, and it was considered necessary to restructure it before it could be transferred to the documentation system. It was decided that the information should be transferred in three steps to the system. First the staff at Kitale should restructure it by transferring all the information on one variety to one set of forms (Appendix 2) that were based on the items in the data base. Secondly the information on this form should be coded, if necessary, and transferred to the box with its respective number. Lastly the information in the boxes should be punched on paper tape and thereby made ready for computerisation.

Some information will be recorded on forms that are based on the data base and that are outlined before the information is collected. This has been the case for the information recorded when wild material of Panicum maximum has been collected (Appendix 3).

b. The incoming information has to be controlled for transcription errors and to some extent for quality, although a quality check in most cases may be very difficult. A more specific control, planned for in the data base, will be to check if the same accession has been described under more than one accession number. In working with varieties in plant breeding and plant introduction, the multiple accession numbers causes a lot of unnecessary work and it has therefore been considered worthwhile to try to minimize the occurrence of them. The items indicated by Trace in the 'use of information' column are thought to be useful for this task.

Regarding the amount of information kept on each accession, it is very important that it is as up-to-date and complete as possible. To cater for this, information on when accession information has been received and from where it has come will be recorded, as well as data on where material has been sent. These items are indicated by Update in the 'use of information' column.

c. In the data base a number of items are indicated by Key in the 'use of information' column. These items will be used for retrieval purpose and for each accession the information on these items is transferred to a retrieval file. By using a computerised information storage and retrieval system on this file, it will be possible to retrieve the subset of the accessions stored that fulfil specified requirements. This group of accessions will then be traced in the accession file, and the information the user wants will be extracted and printed out from the corresponding fields in the data base.

The quality of the data base

The quality of the data base is of great importance for the documentation system. It will affect the quality of the information that will be stored in the system and also the efficiency of the system.

In the building up of the data base it is considered that the best quality will be reached by a successive improvement and a cyclic method of improvement has therefore been planned for. As the first step a preliminary data base has been set up. It will be used for a period of time on incoming information, which will be recorded and accumulated in the system. Requests will be made on the stored information, and replies will be supplied. When it is convenient the data base will be revised by a series of measures that are described below, a new data base will be created and the existing and incoming information will be adapted to this new version of the data base.

The structure and content of the preliminary data base presented in Appendix 1 is governed by the ways it is envisaged that the documentation system will be used. Items that are useful independent of the crop described are compiled under the headings: Identification, Maintenance and Origin. Two of the ways of using the documentation system, Update and Trace, have already been mentioned. It will mainly be the documentation centre that uses the information in this way. The information thought to be used by

plant breeders and plant explorers is indicated by Userinf. Crop specific items are under the heading Observations, and are mainly based on the descriptions that have been used on the incoming information. A maximum of six different 'observations' will be recorded, where the term 'observation' covers information from one place during one or in some cases several growing seasons. If there are more than six 'observations' the least informative have to be deleted.

Some of the measures for revision and improvement of this preliminary data base are:

a. Revising of codes: For each variety information that will be recorded the already existing description or the information to be described will be adapted to the data base. That is, each type of information will be recorded in the corresponding field of the data base, by a code or text as indicated in Appendix 1. For the items that will be coded, each alternative description that is significantly different from the previous ones will be given a new code. All the codes with their explanations are successively recorded in a vocabulary under their respective items. The vocabulary for one item is exemplified by item 762j = spikelet colour. The different alternative descriptions that have been accumulated up till now come from Panicum maximum material and they are:

PU = purple

GP = green (with some)/and purple

GT = green occasionally tinged with golden purple

DP = dull purple

PG = purple or/with some green

GB = golden brown.

The codes and the items will not be species specific, but if a description applies to several species the same item number and code will be used for all of them. There will thus be one single data base for all species, but each species will only have information in a certain sub-set of all the items.

When the alternative descriptions established in the above way are revised, the aim is to compile a set of descriptions for each item that give a comprehensive and objective classification of the possible alternatives of the item. This may result in that some of the alternatives already coded will be grouped together, that some of them will be considered of minor importance and therefore deleted or that new alternative descriptions will be created as complements to those already recorded. The revision may also show that some items should be split up into several and, even the opposite, that several items could be combined into one.

b. The revision of 'other'-information: If the information to be recorded does not fit into any of the well-defined items in the data base, the recording will be solved in one of the two following ways. The information will be put under the appropriate item of 'other'-information (i.e. any of the items 268i, 286i, 290i, 459, 469, 519, 627, 729j, 799j, 819j, 829j, 859j, 899j, 919j, 989j) or if it represents a clearcut alternative to already existing items, a new item will be created in the data base where the information will be put. This new item will be revised like any other coded item. The 'other' information however will not be

classified and accordingly it will soon consist of a wide range of descriptions. The aim of the revision of these items is to empty them of information by creating new well-defined items that have a comprehensive and objective set of alternative descriptions. This will probably be possible when a considerable amount of information has been accumulated under one 'other'-item, since parts of it could easily be grouped and classified into one or several new items that are well defined.

c. Using users requests in the revision: Request from the users will also be of help in the improvement of the data base. Their way of asking questions will show which information is relevant and what one should try to describe for a plant variety.

The retrieval system

The quality of the retrieval system is also of great importance to the documentation system. How efficiently the stored information will be used is highly dependent on it. The quality is measured in terms of the following parameters.

- a. Frequency of accessions retrieved of all accessions that actually fulfil the requirements requested.
- b. Frequency of accessions retrieved that do not fulfil the requirements.

If a is small, it means that the user has just obtained information on a few accessions out of all the accessions recorded on which he wanted information. If b is big, it means that a major information is of no interest to him. Factors affecting a and b are the quality of the information, the quality of the retrieval system and the type of request made. The request could be very specific (narrow) or it could be more general (broad) and in the table below it is shown how these two types of requests affect a and b in combination with high/low quality of the information and the retrieval system.

Quality	low	high
Request		
narrow	a. very low b. rather low	a. rather high b. very low
broad	a. rather high b. very high	a. very high b. rather low

It is considered that the highest priority for the users of the system is to get information on as many of the accessions as possible which fulfil their requirements. Independent of the quality of the recorded information it will therefore be the documentation centres policy to make broad request, in spite of the high frequency of unwanted accessions they will cause. However, the more accessions which are recorded and stored, the more narrow requests have to be made, so that the absolute amount of retrieved accessions is kept within reasonable limits.

But with more information stored even narrow request will still give a high a, that is most of the wanted accessions will be retrieved, since with the stepwise improvement of the data base caused by increased amount of accessions stored, an improved retrieval system follows, with a higher quality of the stored information (see the table).

The computerisation of information

For the computerisation we have chosen to work with a variable length record. This means that the documents will be stored on records that have a physical length and structure which varies from one accession to another. The reason for this type of record is to make input and storage of information economic. A fixed record including the whole data base would consist of many empty spaces, since most accessions only have information under a few items.

The computerisation will be done by putting the information on input-forms of the types shown in appendices 2 and 3, transferring the information from the forms to paper tape, and from there to magnetic tape. Because of the variable length record, each item has to have an identification label, the item number. This number has the advantage of simplifying the input of information, since any set of items, together with their item numbers, could be combined on an input-form, without any concern about the rest of the data base. The items that are filled in with information on such a form will be ended by an asterisk showing that they should be transferred to tape.

Paper-tape has been chosen as an input-media in favour of 80-column cards, since it copes better with variable length records. When a typewriter console becomes available it could replace the paper tape. From the paper-tape the information is transferred to magnetic tape, so that it is easily manageable and safely kept. This operation is done on an ICL-1900 computer. The program for the storage of information is written in machine language, and the items of one record are stored in consecutive order. The program also allows to add and delete items or to change the information under any item of an already stored record.

For the retrieval of information we have planned to use ICL's FIND-2 program on the accession file. It handles records of up to 4096 characters, which is less than the maximum length of a record in the accession file. We will therefore create a retrieval file from the accession file consisting of all items indicated by Key in the data base. On this file we will be able to put inquiries using logical and/or statements, in order to make selected retrieval.

The economy of the documentation system

The reasons for building up a documentation system derive from the need for better preservation and accessibility of variety information for plant breeders, plant explorers and genetic resources activities. How the system fulfils these functions will be evaluated when it has been in use for some time, and at that time areas of responsibility could be better defined. At present the work and expenses in some of the operations could be estimated accurately enough to give a picture of the bottlenecks in the system.

Based on the work up till now, we can roughly estimate the costs for:

a. the manual recording of information to a standardised format: Regarding information already recorded 1,000 to 2,000 characters an hour could be transferred to the data base format. If, however, the information is recorded on forms corresponding to the data base the extra work compared with any other procedure will be very small. When it is at its greatest this extra work could be done at a speed of 5,000 to 10,000 characters an hour.

b. The computerisation of the information: 8,000 characters could be transferred from forms to paper tape every hour at an hour-cost similar to that under a. To transfer the information from paper tape to magnetic tape will be at a speed of more than one million characters per hour. With the present solution for computerisation the hourly cost of this last step is approximately ten-fold to the other steps.

c. The cost of output of information: The cost is very dependent on type of request, but it is obvious that it will be so much less expensive if the information is computerised, that the computerisation costs soon will be covered.

As is seen from the above the main part of the work and the expenses for information already recorded will be to transfer it to a standardised format. For new information to be adapted to the system this cost will be comparable to the cost of computerisation.

In order that recorded information can be standardised at low cost, ways should be sought to adapt existing information to the current documentation system, and future information on new varieties should be recorded in such a way that it can be transferred automatically to the present system. We think that the basic requirements for these two objectives are catered for with the wide and flexible data base set up, and with the possibility to make an input-form that consists of any subset of items.

March 1972

UHG/wp

Data Base for Panicum maximum, March 1972

Item number	Maximum size	Description of item	Type of (*) information	Use of (**) information
100	12	Recorder of identification information	Code S	Update
101	6	Date of recording of identification information	-	Update
102	12	References to identification information	Code R	Update + Userinf.
110	12	Documentation centres accession number	-	Unique number+Key
120	40	Botanical name	Text	Userinf+Key
130	70	Vernacular name(s)	Text	Userinf
140	140	Synonymal accession numbers and other coded designations	Code Sy	Userinf+ Trace
150	70	The 'cluster' to which the above acc. belongs	-	Userinf+ Trace
160	210	Accessions originating from the above acc.	-	Userinf+ Trace
<u>MAINTENANCE</u> (Space is given for a maximum of five different stations, i= 1,5)				
200i	12	Recorder of information from <u>m.s.</u> (maintenance station) i	Code S	Update
201i	6	Date of recording of information from <u>m.s.i.</u>	-	Update
202i	12	References to the accession at <u>m.s.i.</u>	Code R	Update+ Userinf.
210i	12	Maintenance station i	Code S	Userinf.+Key
220i	12	Donor station I of the accession to <u>m.s.i.</u>	Code S	Userinf.+ Trace
221i	6	Date received from donorstation I to <u>m.s.i</u>	-	Userinf + Trace

* Type of information recorded:

Code = Alphanumeric code which is explained in the vocabulary
 Code R = Code which refers to the reference file
 Code S = Code which refers to the station file
 Code Sy = Code which refers to the synonym file
 Num = Measurement with units indicated
 Text = Self-explanatory text or name
 Text Sy = Name which refers to the synonym file

** See explanations in the text

Item number	Maximum size	Description of item	Type of (*) information	Use of (**) information
222i	2	Type of material from donor station I to <u>m.s.i</u>	Code	Userinf + Trace
230i	12	Donorstation II of the accession to <u>m.s.i</u>	Code S	Userinf + Trace
231i	6	Date received from donorstation II to <u>m.s.i</u>	-	Userinf + Trace
232i	2	Type of material from donor station II to <u>m.s.i</u>	Code	Userinf + Trace
240i	16	Accession number used at <u>m.s.i</u>	Code Sy	Userinf
241i	40	Botanical name used at <u>m.s.</u> if if different from item 120	Text Sy	Userinf
242i	20	'Accession specific' vernacular name used at <u>m.s.i</u>	Text Sy	Userinf
250i	6	Material maintained at <u>m.s.i</u>	Code	Userinf + key
251i	5	Reasons for maintenance at <u>m.s.i</u>	Code	Userinf
260i	16	Seed moisture at <u>m.s.i</u>	Code or Num	Userinf
261i	16	Seed storage temp. at <u>m.s.i</u>	Code or Num	Userinf
268i	40	Other seed storage conditions at <u>m.s.i</u>	Text	Userinf
269i	70	History of seed maintenance at <u>m.s.i</u>	Text and/or Code S	Userinf + Trace
270i	2	The ways plants are maintained at <u>m.s.i</u>	Code	Userinf
277i	70	History of plant maintenance at <u>m.s.i</u>	Text and/or Code S	Userinf + Trace
280i	2	Number of herbaria specimen at <u>m.s.i</u>	Num	Userinf
281i	9	Plant parts represented by herbaria specimen at <u>m.s.i</u>	Code	Userinf
286i	40	Other herbaria specimen inform at <u>m.s.i</u>	Text and/or Code S	Userinf
287i	6	Date mounted at <u>m.s.i</u>	-	Userinf
290i	70	Other ways of maintenance at <u>m.s.i</u>	Text and/or Code S	Userinf
300i	12	Station I dispatched to from <u>m.s.i</u>	Code S	Userinf + Trace
301i	6	Date of dispatch to Station I	-	Userinf + Trace

Item number	Maximum size	Description of item	Type of (*) information	Use of (**) information
302i	2	Type of material dispatched to Station I	Code	Userinf + Trace
303i	1	Information returned from Station I	Code	Update
310i	12	Station II dispatched to from <u>m.s.i</u>	Code S	Userinf+ Trace
311i	6	Date of dispatch to Station II	-	Userinf + Trace
312i	2	Type of material dispatched to Station II	Code	Userinf + trace
313i	1	Information returned from Station II	Code	Update
320i	12	Station III dispatched to from <u>m.s.i.</u>	Code S	Userinf+ Trace
321i	6	Date of dispatch to Station III	-	Userinf+ Trace
322i	2	Type of material dispatched to Station III	Code	Userinf+ Trace
323i	1	Information returned from Station III	Code	Update
330i	12	Station IV dispatched to form <u>m.s.i.</u>	Code S	Userinf+ Trace
331i	6	Date of dispatch to Station IV	-	Userinf+ Trace
332i	2	Type of material dispatched to Station IV	Code	Userinf+ Trace
333i	1	Information returned from Station IV	Code	Update
340i	12	Station V dispatched to from <u>m.s.i.</u>	Code S	Userinf+ Trace
341i	6	Date of dispatch to Station V	-	Userinf+ Trace
342i	2	Type of material dispatched to Station V	Code	Userinf+ Trace
343i	1	Information returned from Station V	Code	Update
350i	2	Number of dispatches from <u>m.s.i.</u>	Num	Userinf
<u>ORIGIN</u>				
400	12	Recorder of origin information	Code S	Update
401	6	Date of recording of origin information	-	Update

Item number	Maximum size	Description of item	Type of information	Use of information
402	12	References related to the origin	Code R	Userinf
403	1	Type of origin	Code	Userinf+Key
410	12	Collection-expedition during which the accession have been collected	Code S	Userinf
411	20	Name of collector	Text	Userinf
420	12	Number of the collected sample	Code	Userinf+Key
421	2	Type of material collected	Code	Userinf
430	21	Longitude and Latitude of the place of origin	Num	Userinf+Key
431	5	Attitude of place of origin	Num	Userinf+Key
432	70	Narrative description of place of origin	Text and/or Code	Userinf
440	4	Aspect of collection site	Num <u>or</u> Code	Userinf
441	3	Slope of collection site	Num <u>or</u> Code	Userinf
442	2	Topography of collection site	Code	Userinf
443	2	Soil-texture at collection site	Code	Userinf+Key
444	5	Soil-stoniness at collection site	Code	Userinf
445	5	Soil-depth at collection site	Num <u>or</u> Code	Userinf
446	5	Soil-colour at collection site	Code	Userinf
447	5	Soil-pH at collection site	Num <u>or</u> Code	Userinf+Key
448	5	Soil-salinity at collection site	Code	Userinf+Key
449	2	Soil-drainage at collection site	Code	Userinf+Key
450	2	Cultivation at collection site	Code	Userinf
451	2	Irrigation at collection site	Code	Userinf
452	2	Fertilization at collection site	Code	Userinf
453	2	Erosion at collection site	Code	Userinf
459	40	Other physical information from collection site	Text and/or <u>or</u> Code	Userinf
460	25	Name and occurrence of Species I at coll. site	Text+Code	Userinf
461	25	Name and occurrence of Species II at coll. site	Text+Code	Userinf
462	25	Name and occurrence of Species III at coll. site	Text+Code	Userinf
463	25	Name and occurrence of Species IV at coll. site	Text+Code	Userinf

Item number	Maximum size	Description of item	Type of information	Use of information
464	25	Name and occurrence of Species V at coll. site	Text+Code	Userinf
465	25	Name and occurrence of Species VI at coll. site	Text+Code	Userinf
466	25	Name and occurrence of Species VII at coll. site	Text+Code	Userinf
468	40	Type of plant community at collection site	Text or Code R	Userinf+Key
469	70	Other information from collection site	Text and/or Code	Userinf
470	25	Disease or Pest I and its effect at coll. site	Text+Code	Userinf
471	25	Disease or Pest II and its effect at coll. site	Text+Code	Userinf
472	25	Disease or Pest III and its effect at coll. site	Text+Code	Userinf
500	12	Station where the accession was selected	Code S	Userinf
501	20	Breeder producing the selection	Text	Userinf
510	6	Date of selection	-	Userinf
511	25	Ancestor(s) to the accession	Doc.centres no.	Userinf
512	5	Method of selection	Code	Userinf
519	40	Other information related to the selection	Text and/or Code	Userinf
600	12	Station where mutation work was done	Code S	Userinf
601	20	Breeder producing the induced mutation	Text	Userinf
610	6	Date of selection of mutant	-	Userinf
611	25	Ancestor(s) to the accession	Doc.Centres no.	Userinf.
621	40	Information on treatment	Text and/or Code	Userinf.
622	5	Method of selection	Code	Userinf
627	40	Other information related to the selection	Text and/or Code	Userinf

Item number	Maximum size	Description of item	Type of information	Use of information
<u>OBSERVATIONS</u> (Space is allocated for a maximum of 6 different occasions of observation, j= 1, 6)				
700j	12	Recorder of information from observation j	Code S	Update
701j	6	Date of recording of information from <u>obs. j</u>	-	Update
702j	12	References about <u>obs. j</u>	Code R	Userinf
710j	12	Station or other observation site of <u>obs. j</u>	Codes or Coll. site	Userinf+Key
711j	6	Date of planting or seeding for <u>obs. j</u>	-	Userinf
712j	6	Date of <u>obs. j</u>	-	Userinf
713j	10	Unit of accession observed at <u>obs. j</u>	Text or Code	Userinf
720	2	Way of vegetative spreading at <u>obs j</u>	Code	Userinf+Key
721j	2	The way the stem grows at <u>obs j</u>	Code	Userinf+Key
722j	2	Plant vigour at <u>obs j</u>	Code	Userinf+Key
723j	2	Speed of growth at <u>obs j</u>	Code	Userinf+Key
724j	2	Time of production of flowering stems at <u>obs j</u>	Code	Userinf
729j	70	Other growth habit information at <u>obs j</u>	Text and/or Code	Userinf
741j	10	General shape of plant at <u>obs j</u>	Text or Code	Userinf
742j	8	Plant height at <u>obs j</u>	Num	Userinf+Key
743j	8	Stem thickness for the lowest internode at <u>obs j</u>	Num	Userinf
744j	8	Number of nodes at <u>obs j</u>	Num	Userinf
745j	2	Node form at <u>obs j</u>	Code	Userinf
746j	2	Node hairiness at <u>obs j</u>	Code	Userinf
747j	2	Sheath hairiness at <u>obs j</u>	Code	Userinf
748j	2	Sheath edge form at <u>obs j</u>	Code	Userinf
749j	8	Length of ligule rim at <u>obs j</u>	Num	Userinf
750j	8	Length of ligule hairs at <u>obs j</u>	Num	Userinf
751j	10	Ligule description at <u>obs j</u>	Text or code	Userinf

Item number	Maximum size	Description of item	Type of information	Use of information
752j	8	Length of blade at <u>obs j</u>	Num	Userinf+Key
753j	8	Width of blade at <u>obs j</u>	Num	Userinf+Key
754j	2	Shape of blade at <u>obs j</u>	Code	Userinf
755j	2	Hairiness of blade at <u>obs j</u>	Code	Userinf
756j	2	Blade edge form at <u>obs j</u>	Code	Userinf
757j	5	Blade colour at <u>obs j</u>	Code	Userinf
758j	8	Length of panicle at <u>obs j</u>	Num	Userinf+Key
759j	8	Width of panicle at <u>obs j</u>	Num	Userinf
760j	2	Panicle description at <u>obs j</u>	Code	Userinf
761j	8	Spikelet length at <u>obs j</u>	Num	Userinf
762j	5	Spikelet colour at <u>obs j</u>	Code	Userinf+Key
763j	2	Spikelet surface structure at <u>obs j</u>	Code	Userinf
764j	8	Lower floret length at <u>obs j</u>	Num	Userinf
765j	5	Lower floret colour at <u>obs j</u>	Code	Userinf
766j	2	Sex of lower floret at <u>obs j</u>	Code	Userinf+Key
768j	8	Time of anthesis at <u>obs j</u>	Num	Userinf+Key
769j	2	Leafiness at observation j	Code	Userinf
770j	8	Number of flowering shoots at <u>obs j</u>	Num	Userinf+Key
771j	8	Herbage height at <u>obs j</u>	Num	Userinf+Key
772j	8	Total number of shoots at <u>obs j</u>	Num	Userinf+Key
773j	8	Branching at <u>obs j</u>	Num or Code	Userinf
799j	70	Other morphologic information at <u>obs j</u>	Text and/or Code	Userinf
800j	2	Shade at <u>obs j</u>	Code	Userinf+Key
801j	3	Drought resistance at <u>obs j</u>	Code	Userinf
819j	40	Other adaptability information at <u>obs j</u>	Text and/or Code	Userinf
820j	25	Disease I: name and its effect on the accession at <u>obs j</u>	Text+Code	Userinf
821j	25	Disease II: name and its effect on the accession at <u>obs j</u>	Text+Code	Userinf

Item number	Maximum size	Description of item	Type of information	Use of information
822j	25	Disease III: name and its effect on the accession at <u>obs j</u>	Text+Code	Userinf
823j	25	Disease IV: name and its effect on the accession at <u>obs j</u>	Text+Code	Userinf
829j	40	Other disease information at <u>obs j</u>	Text and/or Code	Userinf
850j	8	Time from seeding to germination at <u>obs j</u>	Num	Userinf
851j	8	Time from seeding to ready seed at <u>obs j</u>	Num	Userinf
852j	3	Quality of germination at <u>obs j</u>	Code	Userinf+Key
853j	3	Amount of seed at <u>obs j</u>	Num or Code	Userinf+Key
859j	40	Other seed information	Text and/or Code	Userinf
860j	2	Number of grazings possible judged on <u>obs j</u>	Num	Userinf+Key
861j	2	Number of silage cuts possible judged on <u>obs j</u>	Num	Userinf+Key
899j	40	Other management information from <u>obs j</u>	Text and/or Code	Userinf
900j	2	General quality at <u>obs j</u>	Code	Userinf+Key
901j	2	Leaf quality at <u>obs j</u>	Code	Userinf+Key
902j	2	Seed quality at <u>obs j</u>	Code	Userinf+Key
903j	2	Stem quality at <u>obs j</u>	Code	Userinf+Key
919j	40	Other quality information at <u>obs j</u>	Text and/or Code	Userinf
929j	40	Trade information at <u>obs j</u>	Text and/or Code	Userinf
989j	40	Other information at <u>obs j</u>	Text and/or Code	Userinf
999	2	Number of 'observations' made on the accession	Num	Userinf

VARIETY INFORMATION COLLECTED AT KITALE

Appendix 2

11101 1173*

Botanical name:

1200 A N I C U P M A X I M U M D.
A C Q *

11301

11401

Synonymal accession numbers:

22201

Donor station of the accession:

Date received: D D M M Y Y
221

222 ☒ Type of material received

Kitale's accession number: 240 K 5 2 5 3 1 *

Vernacular name: 241

250 Y E S *

251 H *

270

277

Nursery plot kept:; It's condition:

280 I X Herbaria specimen kept (indicate by no.) ...;

Date when H.S. was mounted: D D M M Y Y
287 05 09 62 *

300 I V O I *

Station dispatched to: IVOI

Date of dispatch: 301 03 05 65 *
D D M M Y Y

310 F T H I *

Station dispatched to: FTHI

Date of dispatch: 311 16 11 62 *
D D M M Y Y

320 H A N I *

Station dispatched to: HANI

Date of dispatch: 321 13 11 69 *
D D M M Y Y

330 N I S *

Station dispatched to: NIS

Date of dispatch: 331 00 06 61 *
D D M M Y Y

350 G * How many dispatches totally: ... 6

41031C*

IF THE ACCESSION IS COLLECTED

Name of collector: 4111

Date of collection: 412100319511X

412112X Type of material collected: .. roots ..;

Geographic location,

Longitude, latitude: 41301
0 0 0 ° . E 0 0 ° . N
S

Altitude:

41311110011X

Narrative description:

413218160R, WEST SUNKX

414121

414131

414151

414161

414171

414181

414191

415101

415191

Physical information from the collection
site
.....
.....;

416101

416111

416121

416131

417181

Habitat information from the collection site
.....
.....;

IF THE ACCESSION IS SELECTED

Name of breeder: 51011

51110 Year of selection
D D M M Y Y

51111

Accession that this accession is selected from:.....;

51121 Method of selection:

5119

Other information related to the origin of the accession:

..... Smooth leaves, the hairless form not common?

711021

Date of planting:

71112518104512X

Date of observations:

711212111511X

7111251X

Growth habit: Very vigorous, semierect.....

7111251X

fairly dense.....

7111211

713921

7155241X

7169210X

Morphologic-Genetic characteristics: Stems.....

71422180CMX

4 ft. high, thick, Panicle large, hairless.....

7143271X

Herbage 2-3 ft. high.....

7160211X

7164211AX

7191211

717121105CMX

8111211

8111211

8119211

Adaptability to environment pressures:

8120211

8121211

8129211

Diseases or Pests and their effects:

8151211

Seed characteristics:

8151211

8159211

81602110X

Management - Establishment:

81612141X

Very vigorous, early, large, early.....

8191211

91002101X

Quality (palatable, digestible, yield): Good.....
productivity, Outstanding, in quality, vigorous and prod.

9111211

9118211

9111211

9111211

9198211

[illegible]

710 60L SITE

713 Unit observed

7211ERX;

Growth habit

72

Growth habit

Rooting from some of the lower nodes

741

General shape

7421210±10cmX Plant height

74315-6cmX

Thickness at the bottom

74415X Number of nodes

745

Node form

746

Node surface

747

Sheath surface

748

Sheath edge

749 Length of sim

750 Length of hair

751

Ligule description

7521100±15cmX Length of blade

75313C±3mmX Width of blade

754

Shape of blade

755

Blade surface

756

Blade edge

757

Blade colour

758145cmX Length of panicle

759135cmX Width of panicle

760

Panicle branches

Lower branches
Semierect

761 Length of spikelet

762

Spikelet colour

763

Spikelet surface

76 Length of lower flower (764)/Length of bristles (767)

765

Colour of lower flower

766

Sex of lower flower

Distribution Mapping by Computer

B.L. Kershaw

Summary

Between 1950 and 1964 the County of Warwickshire, an area of country some 60 x 90 km. almost exactly in the centre of England, was intensively surveyed by amateur and professional botanists under the joint guidance of the Birmingham Natural History Society and the Department of Botany of the University of Birmingham. The distribution of over 700 species of vascular plants and some 200 species of bryophytes was recorded and subsequently published in "A Computer-mapped Flora; a study of the County of Warwickshire", Cadbury, Hawkes and Readett (1971).

Most records of field observations were stored in computer-readable form, statistical analyses and distribution maps being produced for incorporation in the published Flora. The systems analysis and programming for the production of the distribution maps, either by means of an incremental plotter or on the line printer of the computer, will be described.

From experience gained in this work it has proved possible to identify the tasks to be carried out by a computer system to draw distribution maps for any type of entity, whether botanical or not, and the development of such a system is at present in its early stages. The proposals for this system, as it will appear to the potential user, will be outlined.

EUCARPIA

(European Association for Research in Plant Breeding)

CONFERENCE ON EUROPEAN AND REGIONAL GENE BANKS

10th-15th April 1972 at Izmir, Turkey

11.30 a.m., Wednesday, 12th April (Morning session on
Conservation of Seeds, Pollen and Tissue Cultures)

Technical aspects of tissue culture storage for gene banks

by

Dr. G. G. Henshaw

Department of Botany, University of Birmingham, U.K.

Tissue culture storage for gene bank purposes could, in theory, provide a useful alternative to the more orthodox storage methods where these have been found to present technical problems. It is now well known that cells from a wide range of plant tissue cultures are demonstrably totipotent with the result that they can be used to regenerate normal plants. This potential has already been exploited commercially for propagation purposes and for the production of virus-free plants.

From the gene bank point of view difficulties would undoubtedly arise because of the genetic instability which is a characteristic of many tissue cultures. It might, however, be possible to stabilise the cultures, either by careful control of the growth conditions, or by storage of frozen or freeze-dried tissue.

The technical aspects of tissue culture storage are discussed and in particular five major areas presenting substantial problems are examined, viz:

- i) Isolation of tissue or organ culture.
- ii) Maintenance of culture avoiding genetical change or loss of morphogenetic potential.
- iii) Freeze-storage of cultures.
- iv) Freeze-dry storage of cultures.
- v) Regeneration of plants from cultures.

INTERNATIONAL COOPERATION IN CONSERVING, EVALUATING AND
UTILIZING RICE GERM PLASM RESOURCES

Te-Tzu Chang, Reynaldo L. Villareal and Genoveva C. Loresto ^{1/}

Abstract

The identification of outstanding sources of yielding potential and wide adaptability, resistance to major rice diseases or insects or both, and of high nutritional quality from the International Rice Research Institute germ plasm bank illustrates the contribution that useful genes from diverse sources and different species can make toward to rapid progress in rice breeding. The 16,000-accession bank also represents a successful example of international cooperation in conserving and utilizing commercially important cultivars as well as obsolete varieties and wild forms. Forty-two items are recorded and described for each of the 15,000 cultivars and breeding lines in the IRRI collection (Fig. 1).

Agronomically valuable genes are also found in a smaller number of exotic stocks, breeding lines, types reported to have special merit, primitive cultivars, and wild taxa. The identification of useful genes by systematically screening large collections will be continued on a multidisciplinary basis. Several national centres have recently initiated similar evaluation programs.

^{2/} The International Rice Research Institute, Los Baños, Philippines.

Table 1. Varieties in the IRRI collection compared with national collections and estimated number of indigenous varieties existing in Asian countries.

Country	Accessions maintained in		Estimated number vrs.
	IRRI collection ^{a/}	National collec. of indigenous	
Burma	146	200	1,000-
Cambodia	65	<u>b/</u>	<u>b/</u>
Ceylon	965	881	<u>b/</u>
China (mainland)	1,417	<u>b/</u>	<u>b/</u>
China (Taiwan)	766	1,216 ^{c/}	1,700
India	1,780	20,000 ^{a/e/}	<u>b/</u>
Indonesia	708	1,800	<u>b/</u>
Japan	755	1,804	2,000-
Korea	294	1,674	<u>b/</u>
Laos	587	250	<u>b/</u>
Malaysia (incl.Brunei)	504	1,244-	<u>b/</u>
Nepal	957	906	<u>b/</u>
Pakistan (East)	804	2,600	4,000
Pakistan (West)	70	<u>b/</u>	<u>b/</u>
Philippines	1,059	800	<u>b/</u>
Thailand	162	1,200-	3,500-
Vietnam (North)	9	<u>b/</u>	<u>b/</u>
Vietnam (South)	300	387 ^{d/}	940-

^{a/} Including a number of duplicate samples of the same variety from a different source.

^{b/} No information. ^{c/} Including 682 mainland varieties. ^{d/} Including 16 varieties from North Vietnam. ^{e/} Including collections in 14 states

Common features in the composition of several major rice collections (Table 1) point to the urgent need for national agencies in tropical Asia to immediately collect and evaluate indigenous germ plasm which now faces the threat of extinction because of rapid advances in varietal improvement and seed distribution. Further international collaboration is needed to facilitate surveys of existing collections, to plan and implement efficient field collection and evaluation, to improve seed preservation and exchange operations, to standardize documentation, and to facilitate retrieval and exchange.

S U M M A R Y

D. F. R. Bommer:

The Genebank at Braunschweig-Völkenrode

The genebank at Braunschweig-Völkenrode began its activity at the end of 1970 as a division in the Institute of Crop Science and Seed Research of the Federal Agricultural Research Center at Braunschweig, Germany. The genebank has storage facilities of 125 cbm, 3/4 of which are controlled at 3°C and 30% relative humidity and 1/4 at -20°C.

The overall activities center in 2 main aspects:

1. To serve plant breeders in our country and if possible in middle Europe in providing information and stocks of specific interest to their breeding programs and
2. to co-operate in an international framework for collection, documentation and conservation of plant germplasm resources.

Present activities concentrate upon the following topics:

The search for an internationally adapted documentation system for plant germplasm.

A survey of valuable germplasm in Germany and the offer to take this material in long term storage with the necessary documentation for retrieval.

To procure germplasm by international exchange for specific breeding programs.

To improve conditions for long term storage of seeds.

To inform plant breeders of new interesting plant germplasm.

The genebank at Braunschweig-Völkenrode is open to any ^{kind} friend of international co-operation and offers partnership in the collection and evaluation of wild and primitive forms of crop plants.

The Documentation System of the German Genebank and
its Relation to International Genebank Standardization

S U M M A R Y

A new concept for international standardization of gene bank documentation has been developed by co-operation of the German gene bank at Braunschweig-Völkenrode and Professor Konzak of the Department of Agronomy, Washington State University, Pullman, USA. Principles of which are mainly based on former documentation forms developed by the joint JAEA-FAO-division in 1966. Furthermore principles for documentation on genetic material developed for the European Common Market and for the European Breeder's Association (UPOV, Geneva) have been considered.

Standardization of gene bank documentation is regarded to be based on a vocabulary that as a thesaurus has to be issued by an internationally recognized co-ordinating office. Corresponding discussions at the world's gene banks have to confirm that descriptors will be usable in one of the most convenient languages, while descriptors in different languages correspond with each other synonymously.

The degree of descriptors' expression is coded in a very simple and most commonly understandable form using figures from 1 to 9 wherever coding is advisable. Here figure 1 stands for the least and 9 for the strongest expression of a character.

Absolute figures should be employed in those cases where a classification expressing "very low" or "very high" can be misunderstood. Here the specification of a key location or circumstances for the observation of a character can be included.

The thesaurus will comprise non-crop specific- and crop specific descriptors. As non-crop specific descriptors those are understood that are used for identification and for the records on transfer and origin (collection; selection or bred accession; induced mutation). Crop specific descriptors will cover all kind of expressions describing an accession phenotypically (morphology,

composition and quality, reaction to pests, diseases, environment) and genotypically as far as the latter is interesting.

Documents are thought to be composed due to individual needs and corresponding to the availability of information. In this way only the construction of data banks by aid of computers can be guaranteed. A set of data or a data bank can be placed on tapes in a space saving way. While only disposing of concentrated information sets a retrieval is consequently possible in a highly efficient and rapid way. The intention is to have a special layout any time a new data set has to be keypunched where the 80 column punch card serves as a porter of data. In the possible free field option the only restriction to be made refers to the sequence of information that has to be kept unchanged for an entire set of data.

The computer technical background for the plant genetic documentation and information system presented here is based on the so-called TAXIR system that was originally developed by the Colorado State University. It is operated on an IBM computer 360/65 or 67 and its applicability on another computer system (German SIEMENS 4004) is actually under investigation.

L. Seidewitz

Republic of Turkey - Ministry of Agriculture

Agricultural Research

and

Introduction Centre

COOLOQUUM

on

Plant Genetic Resources

14th-15th April, 1972

at

Izmir — Turkey

<u>Austria</u>	: Dr. K. Mikaelson
<u>Belgium</u>	: Dr. V. Melard
<u>Canada</u>	: Dr. A.N. Galia
<u>England</u>	: Mr. B. Ford-Lloyd Mr. M.T. Jackson Dr. G.G. Henshaw Dr. D.A. Wilkins Prof. J.G. Hawkes Dr. B. Snoad Dr. R.N.H. Whitehouse Dr. G.E. Dixon Dr. T.I. Emeccz Mr. B.L. Kershaw Dr. P. Thompson
<u>Italy</u>	: Prof. L. Quagliotti Prof. E. Porceddu Prof. G.T. Scarascia-Mugnozza
<u>Japan</u>	: Dr. M. Matsuura
<u>Kenya</u>	: Dr. U. Gullberg
<u>Netherlands</u>	: Dr. H. Lamberts Dr. A.C. Zeven
<u>Peru</u>	: Mr. Z. Huaman
<u>Philippines</u>	: Dr. T.T. Chang
<u>Poland</u>	: Dr. W. Kulpa
<u>Spain</u>	: Prof. J. Miranda
<u>Sweden</u>	: Dr. G. Julén Prof. H. Esbo

Turkey : Prof. Dr. Vamık Tayşi
Prof. Dr. Fethi İncekara
Mr. T. Aydın
Mr. A. Ürem
Dr. Ü. Emiroğlu
Dr. B.C. Wright
Dr. Ş. Emiroğlu
Dr. W. Hertzsch
Dr. M. Gökçöl
Mr. Ş. Bekişoğlu
Dr. A. Klatt
Dr. K.J. Neddenriep
Dr. R. Boye
Dr. Halidun Eresen
Dr. Basri Devecioğlu
Mr. V. Caethoven
Dr. J.M. Prescott
Mr. B. Doruk
Mr. S. Aksu
Mr. N. Dinger
Dr. E. Kjellqvist
Miss H.A. Sencer
Dr. O. Manas
Mrs. A. Manas
Miss N. Aydın
Mrs. N. Settar
Mr. I. Özmat
Mr. K. Kayhan

U.S.A. : Prof. J.F. Harrington
Prof. C.F. Konzak
Dr. D.J. Rogers

U.S.S.R. : Prof. K. Budin

West Germany : Prof. Dr. E. Knapp
Prof. Dr. H. Kuckuck
Dr. K.D. Krolow
Prof. Dr. G. Ziegenbein
Mr. L. Seidewitz
Dr. A. Grahel
Prof. Kammer



UNIVERSITY OF BIRMINGHAM

POSTGRADUATE COURSE IN
**Conservation and Utilization
of
Plant Genetic Resources**

*leading, in accordance with the general regulations of the
University, to the degree of M.Sc. by examination*

DEPARTMENT OF BOTANY

Hambes- Fri, 14 April

Techniques of coll.

- " " documenting
- " " storing

Training courses

1 short 2-3 months

Practical try for person already in the field
Given in the gene bank center.

Topics

1. Instr. in collecting various types
2. Recording of field data
3. How to preserve propagules for storing
4. How to transfer data for machine processing
5. What are problems of regeneration.

2. Longer training programs -

1. Necessary to enable to make effort in long term.
2. Botanical, genetic, ecologic education.

Course at Birmingham MSc

9 mos of lects seminars

3 mos res work

short theses

oral exam

Practical work at Izmir & Bari -

Details of theory

Global problems, FAO functions, introduction
to present problems

Evolution of cult. plants

Patterns of variation in wild & cult. sp.

Computer methods, methods of data processing.

Climatology. Plant productivity (physiol.)

over
→

Plant pathology-

Standardization

Use of germ plasma - pl. genetics - most pl. breeding stat.

Agricul. systems -

Exploration techniques - sampling, equipment,
how to set up expedition

Glasshouse + cult. techniques, watering, fertilizing, -
field trial techniques.

The research projects - what are suitable?

Taxonomy, cytologic or physiologic -

Tissue cult.

get paper describing courses

do you use visiting instructors?

UNIVERSITY OF BIRMINGHAM
DEPARTMENT OF BOTANY

Postgraduate M.Sc. Course in
Conservation and Utilization of Plant Genetic Resources

SYLLABUS

1. Introductory Lectures (Professor J. G. Hawkes) 2 lectures

Aims and methods of the M.Sc. Course. General review of world studies in exploration, evaluation, utilization and conservation of plant genetic resources. International collaboration and co-ordination. The role of FAO and IBP. Establishment of national, regional and international germplasm banks. Training of personnel.

2. Origin of Cultivated Plants (Professor J. G. Hawkes) 26 lectures
22 three-hour practicals

Taxonomy, origins and evolution under domestication of cultivated plants. Historical review: de Candolle, Darwin and Vavilov. Methods of study; botanical, cytogenetical, biochemical, archaeological, linguistic and historical techniques. Role of wild and weed species in agricultural origins and subsequent evolutionary processes. Detailed studies: (1) Cereals (2) Roots and tubers (3) Sugar plants, (4) Protein sources (5) Oil plants (6) Vegetables (7) Fruits (8) Spices and condiments (9) Stimulants (10) Fibres (11) Ornamentals. General conclusions.

3. Taxonomic Methods (Dr. J. T. Williams and Dr. R. N. Lester)
21 lectures
Some practicals

Advanced considerations of concept of characters; relation to adaptation, acclimatisation and polymorphism. The conflict of categories. Morphological criteria; phenotypic plasticity. Anatomical criteria; advancement indices. Palynology; diversity of derived types, relation to phylogeny. Comparative embryology and taxonomy. Cytological characters. Numerical taxonomic methods. Nomenclature.

General principles of Biochemical Systematics. Secondary compounds as characters revealed by chromatography; validation of hybrids; taxonomic importance of biochemical pathways. DNA and proteins as molecules with high information potential. Analysis of proteins by electrophoretic and immuno-chemical methods, to yield taxonomic information.

4. Genecology (Dr. D. A. Wilkins) 15 lectures

Morphological and physiological variation within plant species. Adaptation to differences of climate and soil. Experimental study of adaptive differences. Clinal and ecotypic variation. Micro-evolution and natural selection

5. Population Genetics (Professor J. Jinks and Dr. D. A. Jones,
Department of Genetics) 20 lectures

Elementary principles of Population Genetics; Hardy-Weinberg and its elaborations to include effects of mutation, selection, migration and population size.

Sources of Variation; point mutation, numerical and structural changes in the chromosome complement. Types of Variation; continuous and discontinuous, their properties and the methods appropriate for their study.

Breeding Systems and their effects on the maintenance and release of variation; outbreeding, inbreeding, apomixis, asexual systems and incompatibility systems. Types of selection and their consequences; stabilizing, directional and disruptive selection. Polymorphism. The Nature of the differences between species and higher taxa. Speciation. Evolution in action in present day wild populations.

6. Biological Statistics (Dr. B. W. Barnes, Department of Genetics)
17 lectures

An introductory course covering those aspects of statistics which are particularly useful in Biology. The emphasis of the course is on the practical application of statistics: 1. The use of statistics in summarising data, 2. An introduction to the concept of probability, 3. The binomial, poisson and normal distribution, 4. The concept of a test of significance. This will be illustrated by the χ^2 test, t-test and the analysis of variance, 5. The principles of the design of experiments, 6. Regression and correlation analysis.

7. Variation in wild species (Dr. J. T. Williams) 9 lectures

Hybridization: introgression, stabilization of segregates by amphidiploidy and introgression; Clausen's work on Potentilla and Nicotiana.

Breeding systems: Allogamous, autogamous sexual reproduction and mixed outbreeding and inbreeding. Asexual reproduction; agamospermy by advanced embryony and gamatophytic apomixis. Facultative apomixis; isolating mechanisms.

Advanced considerations of polyploidy: Taxonomic and geographical distribution. Taxonomic problems caused by polyploidy; cryptic polyploids; delimitation of genera. Polyploidy and incompatibility relationships. Polyploid permanent hybrids. Practical significance of polyploidy.

Advanced considerations of hybridity: Analysis of hybridity. Hybrid sterility. Heterosis. Differentiation - hybridization cycles.

8. Plant Pathology (Dr. G. M. Butler) 6 lectures

A review of current methods of disease control, their history, problems and potential.

Recognition of host resistance and pathogenic virulence, (a) Vertical resistance (often monogenic) (b) Horizontal resistance (polygenic). The importance of environment and inoculum dose in disease expression. Genetics of resistance and virulence; gene for gene relationship; variation in

pathogens. Disease in host populations; the growth of an epidemic and the different effects of horizontal and vertical resistance. The effects of selection pressures on populations of pathogens; the consequences of host monoculture compared with mixed culture; errors in field experiments.

9. Climatology (Mr. B. D. Giles, Department of Geography) 9 lectures

A consideration of climatic classifications; microclimatology (the climate of the air layer near the ground). Effects of various types of plant life on the climate of this layer and a consideration of larger areas such as towns and valleys.

10. Data Processing (Mr. B. L. Kershaw, Computer Centre) 5 lectures
5 seminars

A brief, non-mathematical introduction to computer programming. No previous knowledge will be assumed, and the emphasis will be on such topics as Data Encoding, Storage and Retrieval rather than on the more traditional mathematical applications of computers. The intention is to make students aware of the capabilities and limitations of computers in the biological sciences and to give them an idea of the type of work needed to implement useful systems, rather than to train them as programmers. The programming language used will be Egtran, the local dialect of Fortran IV. An optional course on computer programming is available.

11. Economic Plants (Dr. R. N. Lester) 10 lectures

Natural products used for pharmaceutical and other purposes. Collection and utilization of wild herbs for folk medicine; cultivation on a limited scale for officinal purposes. Important plant families (e.g. Solanaceae) and classes of chemicals (e.g. alkaloids). New crop plants and potential crop plants.

12. Productivity (Dr. D. B. Idle, Dr. D. A. Wilkins) 8 lectures

Environmental physiology of individual plants and communities, in both natural and agricultural situations; interaction of plants with physical components of environment; relation between structure and physical properties of leaves, plants and whole communities; applied aspects of energy acquisition and utilization in ecosystems.

13. Plant exploration (Professor J. G. Hawkes) 8 lectures

Patterns of variability of cultivated plants and wild species. Population structure and sampling methods. Mode of reproduction in relation to plant exploration.

Definition of purpose. Tactics of exploration. Methods of travel. Collaboration with local botanists and agronomists. Field data; provenance and agro-ecological information. Photographic records; herbarium collections; handling and preservation of material. Follow-up studies.

14. Conservation (Professor J. G. Hawkes)

8 lectures

Basic problems of genetic conservation. Gene banks, introduction stations, botanic gardens. Seed storage and renewal. Conservation of asexually propagated material. Pollen storage. National and international collections. Documentation, description and evaluation of material. Data storage. Quarantine practice; national and international requirements. Screening for pathogen resistance.

General conclusions and discussions. Future perspectives.

15. Plant Breeding (Dr. G. E. Dixon, National Vegetable Research Station, Wellesbourne)

6 lectures

Concepts and methods in relation to present genetical understandings. Defining breeding objectives. Formulating and developing a breeding programme. Production and maintenance of new cultivars.

16. Agricultural Systems (Dr. R. Moss, Department of Geography)

10 lectures

Ecological and economic implication of agricultural systems in the use of biological resources; significance of soil and ecological limitations; problems of land evaluation; role of weeds, pests, diseases, fertilizers and herbicides.

A special course of between five and ten seminar discussions are given each year by a member of the Crop Ecology and Genetic Resources Branch of the Plant Production and Protection Division, FAO, Rome, through the kind permission of the Branch Director, Dr. J. León. Other lectures by visiting speakers are arranged on an ad hoc basis.

Visits to plant breeding stations are arranged during the summer term.

JGH/PML

17/12/71

Sir Otto Frankel - 4/15/72

1

Global project for genetic resources.

All exist on essentially same species - none peculiar to one nation?

Evolution has produced what we have

Genetic assemblies must be preserved.

Responsibility to the future -

~~Genetic~~ Gene erosion in centers of origin -

Effort transcends single nation capabilities.

Sense of urgency compels global effort.

Expertise & facilities not equally distributed & must be shared.

Develop centers in areas of diversity using all necessary personnel.

Proposals -

Personnel needs -

Distribution of materials necessary -

Quarantine requires international agreements

Evaluations - plant breeder tools, and other specialists.

Seed storage methodologies & standards

Regeneration of stored seeds

Ecologically similar areas - global - should be recognized, and used.

Vegetative propagation - tissue culture.

Need large laboratories

Global coop. for safety sake - duplication of storage to prevent loss if 1 center is lost.

Inter rather than intranational conservation.

Antarctica proposed as internationally acceptable
storage area.

2

Nature of cooperation -

Ex. Australia - received wheat vars from all regions of
world - the Villmorán collection + the U.S.

USDA distributed much

Vanilow. " " - USSR Inst. Pl. Genetics.

Private corporations distribute K's

Plant Introd. Newsletter of FAO. informal info center.

1961 - 1st Intern. conf. on gene pools.

Harlan (senior) 1936 - USDA yearbook - threats to gene pools.

1967 IBP conf. in Rome - Frankel & Bennett.

1973 conf. in area of population problems.

Mar. 1973.

Institutions -

IRRI, IRI, CIMMYT,

FAO encourages exploration -

7 limited expeditions - freeze bank -
reorganized newsletter.

Cont → P.3

Frankel - Page 3

Methods of development -

No way of using FAO. Structure limited by national contributions & ~~desires~~ desires.

UNDP - also not effective agencies - because the country application do not assign top priorities to gene bank development.

UN Conf on Environment

Section on gene pool conservation -

Conto. submitted papers - FAO, USDA, etc.

Papers generally in agreement, so the resolution to adopt will probably pass.

Hope to emerge a fund like UNDP.

100 M \$ over 5 years.

another development

UN agency - World Bank, with 12 donor countries - donating a fund for agricultural research -

Tech. Advisory Committee - chmn. Sir John Crawford, Australian economist.

2 Reports before it.

1. Swaminathan of India (with Jorgensen) -

Chmn of TAG - Peter O'Brien - ~~from FAO~~ ^{neediness was too much regionalism, not enough coordination -}

meeting at Beltsville last month - wide representative status & recommendations.

1. Center within regions of genetic diversity
8 or 10 recommended - "center" is preferable -

4 centers regionally representative institutions & ecological areas.

Places: S. Asia - India Ag. Res. Inst. N. Delhi

Bogor

Mediterranean - Izmir

- Ethiopia

Africa - IITA

S. Am. - Tropical

Gene Bank

Support + budget accepted for center.

Motor vehicles, glasshouses, storage facilities.

Required center $\frac{1}{2}$ M/year - \$200,000 non-renewable.

2. Considered needs centrally located.

Experts, training, etc. organized through central org. + financed through.

3. Coordinating body of 9-10 people, 6 representatives of genetic resources center, 3 independent selected by TAC + 1 independent chosen with a staff

3 biologists - 1 director + 2 others.

1 legal advisory committee -

where to attach this group??

but FAO

Leon recommended a trust fund independent under direction of coordinating -

but with FAO logistics + administrative backup.

FAO - provides necessary global info. backup.

Documentation was not included, but not forgotten. Must be included as addition to report.

L. R. House - 4/14/72 -

Rockefeller program in millets & sorghum in
Hydrabad, India -

12000 collections - no computer processing -
has a kind of reporting system similar to
Sam Hietz's old mines sheets.

~~very few~~

Pedigree information system
recording a collection grown in
many localities.

| Has started a jerry-built info system
without any advice from ^{system} ~~computer~~ experts -

Has different types of collections

1. named variety collection -
2. genetic stocks "

3. Population " (with gene combinations)

Another institute by Rockefeller is
Internatl semi arid land res. center.

| Has TAC got Rockefeller turned in
for documentation of the global system

HANDLING OF INFORMATION ON GENETIC RESOURCES
AT THE ITALIAN GERM-PLASM STATION

In order to record information in a standardized form a computer programme has been set up at the Bari Germplasm Laboratory. In preparing it we had in mind the following conditions:

1. To establish an I/R system agronomic, ecological and genetical characters, which permits separate analysis for different species.
2. To set up a documentation programme which provides information of general interest, on the nature and the origin of the material on morphological and physiological characters, as well as information about the value of material from the breeding point of view.
3. To store all the information in a form that can save space in the computer without missing any information.
4. To provide all the information in different languages.

In agreement with these points a programme has been developed. The I/R system used is a partial version of TAXIR. That version was prepared by Dr. SOUSA at Washington State University in Pullman. It provides a maximum of 36 retrievable characters. An example of input, dictionary and output is given in appendix tables.

The documentation programme consists of three points:

a) General information b) origin of material c) special part
different according to the crop. The information/^{is} provided in agreement with the forms edited by F.A.O. in 1969.

An example of the output in Italian language is provided in appendix tables.

E. Porceddu

THE USE OF FIELD PUNCHED CARDS FOR INFORMATION MANAGEMENT OF
SMALL COLLECTIONS

A.C. ZEVEN

Institute of Plant Breeding (I.V.P.), Wageningen, The Netherlands

When the program of this conference was heralded I observed that part 2 'Information storage and retrieval' only concerns papers on computerized systems. Many keepers of collections - most of them plant breeders - only maintain small ones. Therefore, they cannot easily use a computer to store the information gathered, unless they have one near at hand. I suggested to Prof. J.G. Hawkes that at least one lecture on a hand system should be presented.

Most people are acquainted with some type of hand system. They use them, for instance, for a literature index. The simplest system is an ordinary card index. In most cases the cards are arranged in a fixed system for instance alphabetically on first author's name, chronologically or according to the U.D.C. Another system is one of cards with coloured tabs, or one with needle punched cards. The last can be divided into edge-notched cards and field punched cards. I am using the field punched cards for my *Triticinae* collection. The picture shows the holes allotted to the characteristics recorded. In most cases I used three holes for a characteristic, viz. 1 = absent, 2 = present but intermediate, 3 = present, or No-gonotype $1-ne_1-ne_1-ne_2-ne_2$, $2 = ne_1-ne_1-ne_2-ne_2$ and $3 = ne_1-ne_1-ne_2-ne_2$. I have not used all holes, because, new interests may emerge and these can then be included by using some of the spare holes.

By inserting one or more needles in the holes of the desired characteristics(s) it is possible to select the collection numbers answering the desired requirements.

It is an easy system and quite useful for small collections. It takes little time to punch the card. The retrieval of a certain card is quite fast.

----- EUCARPIA -----

EUROPEAN ASSOCIATION FOR RESEARCH ON PLANT BREEDING
EUROPAISCHE GESELLSCHAFT FÜR ZÜCHTUNGSFORSCHUNG
ASSOCIATION EUROPÉENNE POUR L'AMÉLIORATION DES PLANTES

P.O. Box 128
Wageningen
Holland

Eintragungsformular

..... den

Den Sekretariat des Verwaltungs-
rates EUCARPIA

1. Name

2. Norname

3. Nationalität

4. Wohnort

~~5. Privatadresse~~

6. Geburtsdatum

~~7. Titel und Eigenschaften~~

.....

8. Dienststelle

.....

9. Name des Instituts, des Betriebes oder der Gesellschaft:

.....

Sitz und Anschrift

.....

10. Unterzeichnete(r) Herr/Frau/Fräulein

bittet als Mitglied der obengenannten Gesellschaft eingetragen zu werden,
und wünscht teilzunehmen an folgenden Sektionen:

a. Kartoffeln

b. Getreide

c. Futterpflanzen

~~d. Gartenbaugewächse (aufgehoben)~~

e. Methodik und Technik

f. Physiologie

g. Mutationen

h. Wildarten und Primitivformen

i. Mais

j. Gemüse

Datum

k. Obstarten

Unterschrift

l. Zierpflanzen

Paten:

1. 2. 3.