

Plant Genetic Resources Passport Information Management System (PGRPIMS)

R.C. Agrawal and Innaiah Pudota
National Bureau of Plant Genetic Resources, New Delhi
(Received : October, 2005)

SUMMARY

A software PGRPIMS (Plant Genetic Resources Passport Information Management System) has been developed for managing the information collected by the exploration staff working in various Indian Council of Agricultural Research Institutes, State Agricultural Universities, Non Governmental Organizations, and many other organizations in India. This software is in use at National Bureau of Plant Genetic Resources and is having data related to more than 3,00,000 accessions collected by various organizations. The software is developed using the visual basic as the front end and with SQL Server as the RDBMS in the back end. This software is based on the client server technology. The design and description mentioned about the PGRPIMS software in this paper may be used by various organizations working in the area of plant genetic resources and maintaining the passport data related to plant genetic resources.

Key words : Passport information, Plant genetic resources, Visual Basic, SQL Server, Accessioning, Database, Descriptors.

1. INTRODUCTION

In plant genetic resources (PGR), documentation of information is imperative for planning and implementing activities, related to their conservation, sustainable utilization and sharing of benefits accrued from their use. The need for countries to develop, maintain and exchange such information is specifically recognized in the Articles 7d, 17 of the Convention of Biological Diversity (CBD 1993), and the priority activities 17 and 18 of the Global Plan of Action (FAO 1996). High-quality information helps the genebanks in current resources utilization, reduced duplication of effort and vision the future direction. Also full and proper documentation shall help to authenticate the information on origin, development and maintenance of the germplasm accessions, thus, preventing unauthorized use and protection of genetic resources through Intellectual Property Rights systems, without commensurate benefit sharing. Earlier, the information on PGR was documented through the collection tour reports, memoirs of collection expeditions and the catalogues published after collection and/or evaluation of germplasm. The developments in computation technology made

information management in PGR through databases as very important at the national, regional and global levels using standard descriptors for effective exchange of information (Lipman *et al.* 1997). Presently, a good number of genebanks have developed their own information system, fitting to their requirements and based on the availability of the computer system, and tailored Database Management System software. The Germplasm Resources Information Network (GRIN) system is effectively being used to monitor information on world's largest collection at the National Center for Genetic Resources Preservation (NCGRP), Fort Collins and the cooperating institutions within the United States Department of Agriculture (USDA) research system in USA (<http://www.ars-grin.gov/>). The International Plant Genetic Resources Institute and USDA have developed a genebank management software known as pcGRIN, which is a stand-alone version of the GRIN database that runs on a personal computer and contains data for individual crops (http://www.biodiversityinternational.org/Themes/Germplasm_Documentation/index.asp). The System-wide Information Network for Genetic Resources (SINGER) has been established as the genetic resources information exchange network for nearly 0.7

million of the Future Harvest Centres of the Consultative Group on International Agricultural Research (CGIAR) and associated partners (<http://singer.grinfo.net/>). In addition, the International Agricultural Research Institutes have developed their own information systems for handling the germplasm in their respective mandate crops. The Federal Information System Genetic Resources (BIG), Germany, integrates databases on PGR *ex situ*, (e.g. collections of botanical gardens and accessions of German gene banks) and *in situ*, including the wild flora of Germany (<http://www.big-flora.de/>).

The National Bureau of Plant Genetic Resources (NBPGR) has been entrusted with the responsibilities to plan, conduct, promote, coordinate and take lead in activities concerning germplasm collection, conservation, evaluation, introduction, exchange, documentation and sustainable management of diverse germplasm of crop plants and their wild relatives with a view to ensure their availability for use to the researchers. The Indian Plant Genetic Resources Management System (IPGeRMS) under the aegis of the Indian Council of Agricultural Research (ICAR), spearheaded by the NBPGR, is emerging as a dynamic system which holds prominent place among the global genebanks. The Bureau operates with a network of ten regional stations spread over different phyto-geographical zones in the country. The Bureau also maintains linkages with over 40 National Active Germplasm Sites, ICAR Institutes, State Agricultural Universities and general Universities to operate IPGeRMS.

One of the main objectives of the NBPGR is to plan, coordinate and conduct explorations for collecting germplasm. India, being the centre of diversity for several crops, holds a rich wealth of useful plants and their wild relatives. The Bureau has so far conducted more than 2000 explorations in diverse agro-ecological regions and habitats representing different phyto-geographical zones of the country. The explorations are undertaken from the headquarters, the regional stations and base centres which has helped in assembling more than 2,30,000 accessions of agri-horticultural crops/ genera. NBPGR has initiated a special drive under the National Agricultural Technology Project (NATP) to collect the plant diversity in India by involving several ICAR institutes, SAUs, NGOs and government departments.

Plant Genetic Resources Passport Information Management System (PGRPIMS) is basically a system

to maintain a database of information on indigenous plant germplasm collected from various parts of the country by Agriculture Scientists and other researchers. Information contains the details about the place of collection, mission details, and other crop details e.g. status, source, frequency, land aspect, material, soil texture, soil color etc. The system is in operation at NBPGR and all the passport information is maintained using this system. An extensive requirement analysis for the development of this PGRPIMS was done during a NATP project on “Integrated National Agricultural Research Information System” (Agrawal 2004). The design and description mentioned about the PGRPIMS software in this paper may be used by various organizations involved in PGR activities for maintaining the passport data details of all the plant genetic resources collected and maintained.

2. METHODOLOGY

2.1 Passport Data Sheet

The software has been designed to accommodate all the information recorded by an explorer on a passport data sheet provided by NBPGR and some additional information like date of start of the survey, name and address of all the collaborating institutes, total number of sampling sites, type of storage for the collected material, availability of the evaluation/ characterization data, availability of the herbarium etc. which are required to administer the collected material has also been incorporated in this software. A sample of the data sheet being used for the recording of the passport data is given below.

Details of the institute depositing the material (also mention the name of the person)

Date (of collection).....Collector No.....

Accession No. (To be assigned by NBPGR).....

Species Name.....Common Name

Cultivar/ Vernacular Name.....

Regions Explored

Village/ Block..... District.....

State.....

Latitude.....NLongitude.....E Altitude.....m

Temp..... Rainfall.....

Source	1. Natural wild 2. Disturbed wild 3. Farmer's field 4. Threshing yard 5. Fallow 6. Farm store 7. Market 8. Garden 9. Institute 10.
Status	1. Wild 2. Weed 3. Landrace 4. Primitive cultivar 5. Breeder's line
Frequency	1. Abundant 2. Frequent 3. Occasional 4. Rare
Material	1. Seed 2. Fruits 3. Inflorescence 4. Roots 5. Tubers 6. Rhizomes 7. Suckers 8. Live plants 9. Herbarium 10.
Sample Type	1. Population 2. Pure line 3. Individual plant
Sample Method	1. Bulk 2. Random 3. Selective (non-random)
Habitat	1. Cultivated 2. Disturbed 3. Partly disturbed 4.
Disease Symptoms	1. Susceptible 2. Mildly susceptible 3. Tolerant 4. Resistant 5. Immune
Insect/ Pest/ Nematode Infection	1. Mild 2. Moderate 3. High
Cultural Practices	1. Irrigated 2. Rainfed 3. Arid 4. Wet 5.
Season	1. Kharif 2. Rabi 3. Spring – summer Approx. Sowing Date
Associated Crop	1. Sole 2. Mixed with
Soil Colour	1. Black 2. Yellow 3. Red 4. Brown 5.
Soil Texture	1. Sandy 2. Sandy Loam 3. Loam 4. Silt Loam 5. Clay 6. Silt
Stoniness	1. Stony 2. Pulverized 3.
Land Aspect	1. Level 2. Crest 3. Escarpment 4. Rounded summit 5. Upper summit 6.
Slope	1. Terrace 2. Lower slope 3. Open depression 4. Closed depression
Topography	1. Swamp 2. Flood plain 3. Level 4. Undulating 5. Hilly dissected 6. Steeply dissected 7. Mountainous 8. Valley
Agronomic Scope	1. Very poor 2. Poor 3. Average 4. Good 5. Very good

Ethonobotanical Uses of Part(s)	1. Stem 2. Leaf 3. Root 4. Fruit 5. Flower 6. Whole plant 7. Seed 8. Others
Kind	1. Food 2. Medicine 3. Fibre 4. Timber 5. Fodder 6. Fuel 7. Insecticide/Pesticide 8. Other
Informant(s)	1. Local Vaidya 2. House-wife 3. Old folk 4. Graizer/Shepherd 5. Other

Farmer's/ Donor's Name Ethnic Group

2.2 Architecture of PGRPIMS

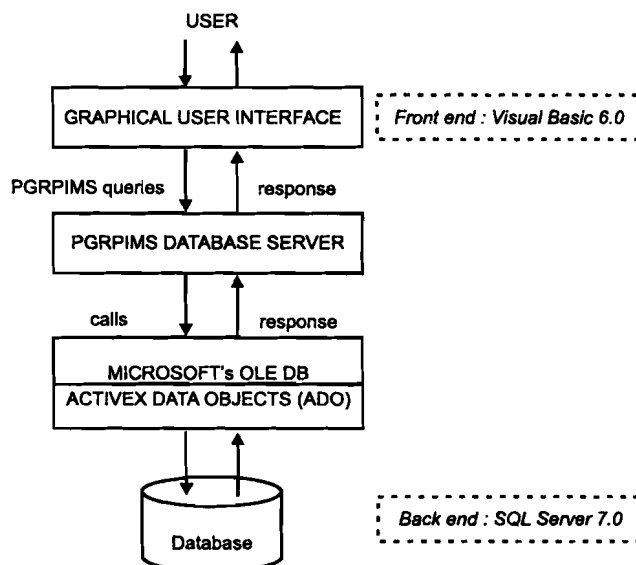


Fig. 1. Architecture of PGRPIMS

PGRPIMS is a relational information system with SQL Server ver 7.0 at the back-end and Visual Basic 6.0 as front-end (Fig. 1). PGRPIMS architecture is composed of two major components, namely, the PGRPIMS Database Server and the Graphical User Interface. The database server implements the relational model of PGRPIMS. The database server is also responsible for parsing and processing the queries that it receives from the user (Desai 1999). The user interacts with the database server through the graphical user interface. The GUI of PGRPIMS has been developed using Visual Basic 6 (Dan Fox 1999) and the reports are generated through Crystal Report.

Due to the reason of involvement of composite and overlapping keys in PGRPIMS, after completing the activity of identifying attributes, their aggregations,

associations, an intuitive normalization of relationships to third normal form was used (Elmasri and Navathe 2006).

2.3 Dataflow Diagram of the PGRPIMS

System analysts use process models (i.e. data flow diagrams, DFDs) to show information flow and processing in a system. The model usually starts with a context diagram showing the system bubble surrounded by the external environment identified by external entities (Barker 1991). Data flows bring information to and from the system process. In other words we can say that it is a description of data and the manual and machine processing performed on the data as it moves and changes from one stage to the next.

Below we provide two levels of data flow diagram of the PGRPIMS (Fig. 2 and 3).

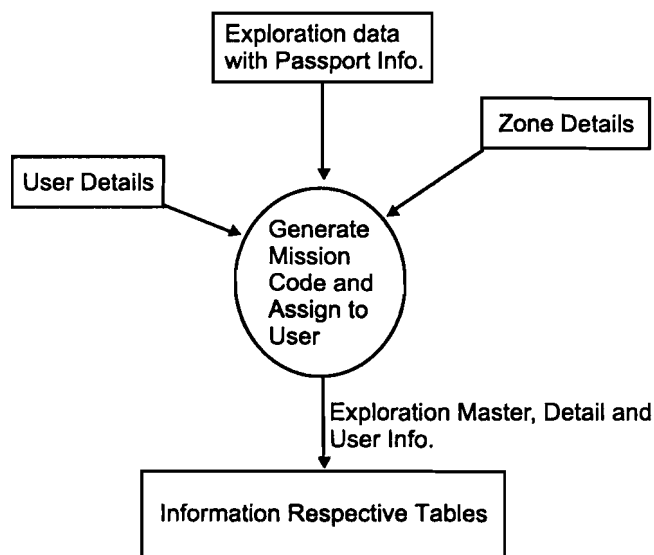


Fig. 2. First level of data flow diagram

2.4 User Management

Following two types of users have been defined in PGRPIMS

- a. Administrator - All rights like creating new users, deleting users and change user passwords, assigning accession numbers for users for passport & exploration details entry are available to this user category. Permission to update the data related to any table is also available to this user category.

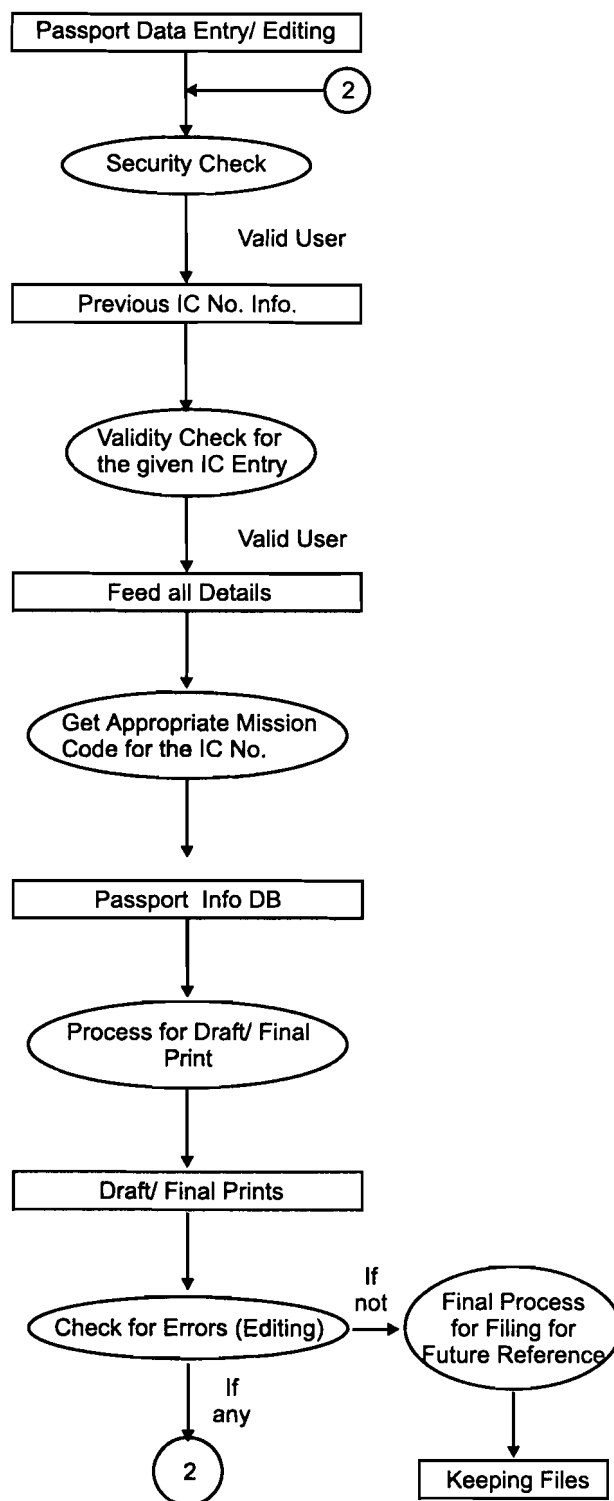


Fig. 3. Second level of data flow diagram

- b. Ordinary Users - Only data entry forms to update/ edit data into the database along with various types of reports/ queries are available for this type of user.

3. MODULES AVAILABLE IN THE SYSTEM

PGRPIMS is divided into 4 main modules viz. Data Entry, Admin Forms, Query Module and Report Module, which are discussed below.

3.1 Data Entry Module

3.1.1 Exploration details

The user with administrative permissions can feed the data in the exploration form, fields like mission code, donor code, regions explored, date of start of the survey, number of sampling sites, number of crop specific collections, number of other collections, total number of days of survey etc. For each exploration details, the code of the explorers (collectors) containing their details like name and addresses is also stored in this form. The mission code is automatically generated in this form, which is 12 letter code and is based on the zone (there are 10 zones for the collection of passport information) and year. The details of the mission code are

- 1st character : A/E (A-Augmented, E-Exploration)
- Next 4 characters : Year of exploration (Character 2-5)
- Next 4 characters : Exploration number undertaken by a particular zone in a year. This number is automatically increments by 1 for every exploration in a zone and starts from 1 every year. (Character 6-9)
- Next 3 characters : Starts with Z and has another 2 digit length for a zone (Character 10-12)

For example, for an exploration undertaken during year 2000 by the NBPGR Regional Station, Hyderabad (Zone-10), the Mission code E20000001Z10 will be generated (In this 'E' stands for Exploration, 2000 is year, 0001 is the exploration number undertaken in a particular year, and Z10 is zone).

There can be multiple collectors for one exploration. Separate primary tables for the address of the institute of the collectors and the concerned officials working in those institute is maintained.

Fig. 4. Passport information details for an accession. In this form, a particular accession can also be searched, added and modified

3.1.2 Details of passport information

In the passport data entry form (Fig. 4), the data can be added/updated. The information in this form is linked on the primary key “Mission code” which is generated by the Administrator. Most of the information in this form is updated with the use of either combo or list boxes. If the Status (Biological) of a seed sample is selected as “released”, then this form displays another form for the updating related to released varieties information which includes pedigree, year of release of the variety, and important traits.

3.2 Admin Forms Module

3.2.1 Assigning of a range of accession numbers for data entry

Data entry is done from many terminals and by various data entry operators. This requires a complete check on the specific range of contiguous accession numbers allotted to a particular data entry operator.

After filling all the exploration details, Administrator can assign the permission for the data entry of a particular exploration to any authorized data entry operator which helps in keeping complete record of the data entry work done by different data entry operators.

3.2.2 Modifying combo box options

There are several options in the passport data sheet which may need modification as per the requirement of the exploration staff. A provision has been incorporated for such modification and the short cut keys have also been defined for this purpose.

3.3 Query Module

Query module (Fig. 5) mainly contains three options like

1. Simple: It uses single condition.
2. Advanced: It uses multiple conditions like multi joins and multi field values.
3. Custom: In this one can select the required fields and use different group by and order by clauses.

The Query module can be used for customized queries like “how many accessions (IC numbers) are there in the database which are having the botanical name ‘Oryza sativa’” (Fig. 5). For this simple query option select the botanical name and just give ‘Ory’ which will provide list of all those accessions which are existing with ‘Ory’ botanical name. A SQL query is generated based on the required fields and the conditions which can be viewed and edited for customized results.

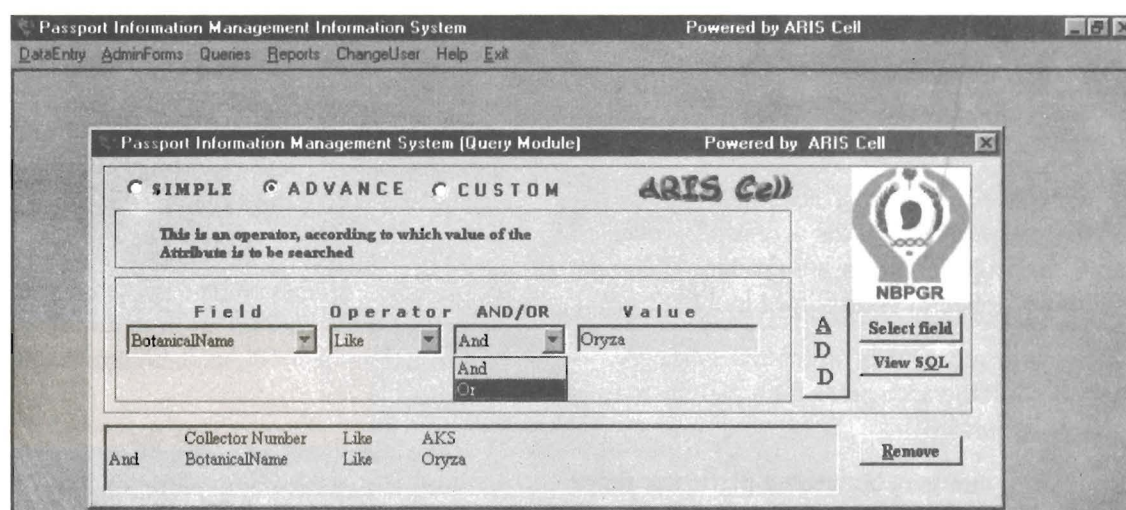


Fig. 5. Options available in the advanced query module

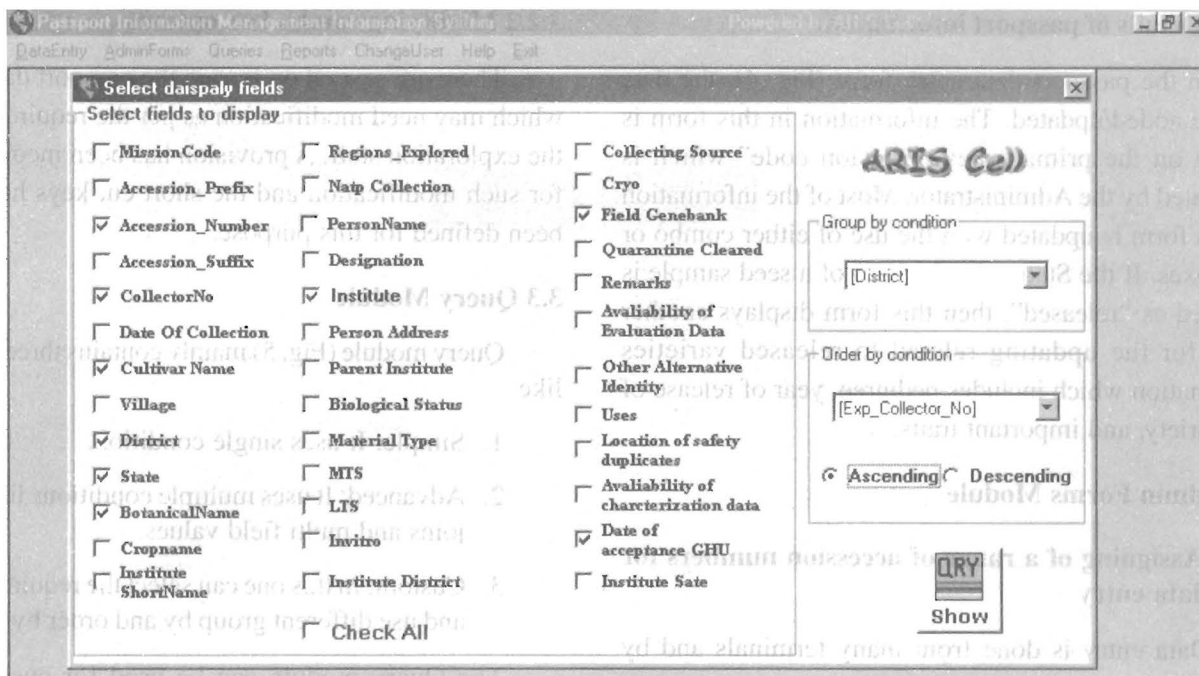


Fig. 6. Field names can be selected for the query output and a option for the group by and order by can also be selected

In the query output, total number of records count and all the selected field headings are displayed at the top. The fields to be included in the query result (Fig. 6) can be incorporated by clicking on the "Select field" (Fig. 5).

3.4 Reports Module

Various preformatted reports available in the software are listed below.

1. Report of all accessions (IC Numbers):- This produces the list of all the accessions numbers which are there in the whole database.
2. Name wise retrieval of accession numbers:- Passport details for all the accessions collected by a particular collector are produced. The list of all the collector is displayed in a list.
3. Zone wise retrieval of accession numbers:- A report of all the accessions for a particular zone can be generated.
4. Accession numbers between a particular range produces options for two types of print output like Draft and Final Print: The draft print includes all the fields in the report and is sent for editing. The final print contains information on some important fields like the name of

collectors, mission code, collector number, village, district, state name, biological status and important traits and is used for the distribution to concerned collector for further reference.

4. RESULT AND DISCUSSIONS

The PGRPIMS is being effectively used by the researchers at NBPGR for the day-to-day management. This software does not require any specific training and helps in effective data capture and its retrieval in the desired format. It has been customized so that it can, not only respond to specific requirements in terms of querying, sharing and reporting, but also is flexible to change and eventually expand whenever the need arises.

Provisions have been made for built-in database level constraints and client level validation of attributes and the master tables for information such as taxonomic identity, addresses of the collaborating organizations and biological status.

The allotment of a set of contiguous accession numbers to a particular exploration helps in monitoring the day-to-day status of the data entry and processing of data sheets received. The reports according to various criteria like zone wise and for a particular set of accession number helps NBPGR to timely communicate the national identity to the concerned explorer. This national

identity is required for further processing of the germplasm accession like storage in the genebank and field evaluation. The research managers can query on any of the chosen attributes which helps for planning of future explorations and acquisition of germplasm from the gap areas identified.

5. CONCLUSIONS

The software "Plant Genetic Resources Passport Information Management System" has been developed at National Bureau of Plant Genetic Resources to replace the manual maintenance of all the passport data in registers. This has also helped in quick and efficient accession number allotment to various collected material and has helped in maintaining a national database of all plant genetic resources collected in India. For a general user, a separate web based query system based on the some minimum general fields has been provided on the NBPGR intranet which helps the user in knowing the status of the allotment of IC numbers (national identity provided to any accession). The design and description mentioned above may be adopted by various organizations working in the area of plant genetic resources and maintaining the passport information of plant genetic resources.

ACKNOWLEDGEMENTS

We wish to thank Dr. BS Dhillon, Dr. SS Malik, Mr. Rajeev Gambhir, Mr. Vijay Mandal, and all the staff of

ARIS and Exploration at NBPGR for their critical input, and discussion throughout the development of this PGRPIMS software.

REFERENCES

- Agrawal, R.C. (2004). Technical Report of the "Integrated National Agricultural Research Information System (Plant Genetic Resources)". NBPGR, New Delhi, 158.
- Barker R. (1991). *CASE*METHOD Entity Relationship Modeling*. Addison-Wesley Publishing Company.
- CBD (1993). Convention on Biological Diversity. Full text available at <http://www.biodiv.org/convention/articles.asp>
- Desai, B.C. (1999). *Introduction to Data Base System* (First Edition). Galgotia Publication, New Delhi.
- Dan Fox (1999). *Pure Visual Basic: A Code-Intensive Premium Reference*. Techmedia.
- Elmasri, Ramez and Navathe, B. Shamkant (2006). *Fundamentals of Database Systems*. Addison-Wesley Publishing Company, 335-379.
- FAO (1996). *Global for the Conservation and Sustainable Utilization of Plant Genetic Resources for Food and Agriculture: Leipzig Declaration*. International Technical Conference on Plant Genetic Resources Leipzig, Germany, June 17-23, Food and Agriculture Organization, Rome, Italy, 63.
- Lipman, E., Jongen, M.W.M, Hintum van, Th.J.L., Gass, T. and Maggioni, L. (1997). *Central Crop Databases: Tools for Plant Genetic Resources Management*. International Plant Genetic Resources Institute, Rome, Italy/CGN, Wageningen, The Netherlands.