



Management system of an active genebank of common bean

Marlon Mathias Dacal Coan^{1*}, Diego Stähelin¹, Jefferson Luis Meirelles Coimbra¹, Sílvio Luís Rafaeli Neto¹, Altamir Frederico Guidolin¹, and Tomas Pellizzaro Pereira¹

Received 20 July 2009

Accepted 03 October 2009

ABSTRACT - *Genetic resources are conserved ex situ in genebanks, which are used as a source of useful alleles. The information generated by the Active Common Bean Genebank (ACBG) is stored in a Database Management system which ensures a quick and safe data recovery. Using Microsoft Access, a database system was developed to store and search bases of accession combinations or directly among morphological, physiological and yield descriptors of common bean. The system called Common Bean Genebank (CBGB) provides a graphical interface for the operation in the database, facilitating and optimizing the search time in germplasm catalogues.*

Key words: database, genetic resources, germplasm, accessions.

INTRODUCTION

Common bean is a legume grown in many countries and the genetic variability within the species (*Phaseolus vulgaris* L.) is great, which is exploited in the search for desirable agronomic traits in breeding.

Currently there is a great pressure for the selection of more uniform and more productive cultivars, which can lead to the loss of genetic variability in traits such as yield stability, disease resistance, nutritional and technological grain quality, among others. These traits are more likely to be found in accessions due to the great genetic diversity represented in a genebank (Loarce et al. 1996, Carbonell et al. 2003, Rodrigues et al. 2002).

Information about variability can be found in genebanks by the evaluation and characterization of accessions for breeding programs. According to Chiorato et al. (2005), common bean breeding involves countless procedures until a cultivar can be

recommended. All information must be considered of equal importance for future use.

The importance and usefulness of a variety description is related to the goals of its users. Data of botanical traits are required for genetic and evolutionary studies that are carried out in genebanks, while the variety description used for breeding includes agricultural and commercial traits that are relevant for farmers and in agribusiness (Muñoz et al. 1993).

The characterization and evaluation of bean germplasm generate an enormous amount of information making queries a laborious task for a number of activities, e.g., search for target information, listing, updating, multiplication and regeneration of accessions. Consequently, the different stages of germplasm conservation includes a range of activities that require or generate information (Paiting et al. 1995).

Data storage in static systems is problematic, for example, in electronic charts that indicate no

¹ Universidade do Estado de Santa Catarina, Centro de Ciências Agroveterinárias, Instituto de Melhoramento e Genética Molecular. Av. Camões 2090, 88520-000, Lages, SC, Brazil. *E-mail: marloncoan@gmail.com

relationships between information from different tables. This makes data queries difficult, when a large volume of information is involved. The solution to this problem would be to export these data to a database system where the data are related in a system, which would, if properly designed, ensure the access to updated and accurate information.

A Database Management System (DBMS) consists of a set of associated data with the main objective of providing an environment for the efficient recovery and storage of agronomic information in the database (Silberschatz et al. 1999). The database provides a graphical interface with operations within the base and ensures that the data set is inter-related. The relationship between tables is essential, because several tables may be included, each containing data about a particular entity, but containing one or more data in common (Matos 2004).

By the DBMS information can also be found and accessed quickly and a considerable amount of data can be stored. Moreover, such systems occupy little space and the data they contain can be duplicated as a safety measure (Santos and Bettencourt 2001).

The objective if this research was to develop a Database Management System (DBMS), based on a conceptual model that relates to real-world objects, in this case the traits of the existing accessions called related entities in the project, to spare time and work required in the labor-intensive tasks related to genebanks, as for recording, grouping, updating, and a systematic and detailed organization of the data accession for the work with genetic resources.

MATERIAL AND METHODS

The information of the common bean database was provided by the Instituto de Melhoramento e Genética Molecular (IMEGEM), of the Universidade do Estado de Santa Catarina (UDESC) of the Active Common Bean genebank (ACBG). The accessions contained in this database had mostly been collected in southern Brazil, in the states of Santa Catarina, Parana, Rio Grande do Sul, as well as in other parts of the country. To characterize these data, most of the globally promoted minimal descriptors (IPGRI 2001) were used as well as descriptors of EMBRAPA (Silva 2005), with a passport, characterization and preliminary and subsequent evaluation of physiological, morphological

and yield traits. The descriptors of the International Center for Tropical Agriculture were used for disease evaluation (CIAT 1991).

The database was developed in the laboratory of geoprocessing of the Centro de Ciências Agroveterinárias –CAV/UDESC, in Lages, SC. The software Microsoft Access 2007 was used, due to its system of cardinality and graphical user interface, facilitating the operation. The conceptual design was a high-level model described by the Entity-Relationship (ER). According to Silberschatz et al. (1999) the E-R is based on the observation of the real world with a set of basic objects called entities and of the set of relationships between them.

The entities are represented in the bank by their properties that describe the appearance of the object. The E-R model follows an important rule of mapping (cardinality), in which the number of entities to which the other entity is related through that set of relationship is expressed (Takai et al. 2005). The E-R model is a logical structure of the CBGB database in the cardinality, entity and relationship diagram (illustrated in Figures 1 and 2). The accessions are identified as unique based on a number, representing the primary key, which is generated automatically when a new accession is inserted.

The key declarations and forms of relationship were used as integrity rules, ensuring consistency and semantics of the data (Figure 1 and 2). The primary key is a set of one or more properties which, considered collectively, allow the doubtless identification of an entity within the set, as well as a relationship within a set of relationships (Silberschatz et al. 1999).

Figure 3 shows the cardinality between two entities, whereas passport and collection show the relationship between the two entities. The original binary relationship Collection has a Passport and has the cardinality ratio 1: N. This means that the original collection entity may be related to numerous Passport entities (or one passport related to several collections).

RESULTS AND DISCUSSION

The Common Bean Genebank (CBGB) is easily accessible for system users, optimizing the time for information searching by standardized electronic data and organization of the Active Common Bean genebank (ACBG) accessions, providing accurate and reliable

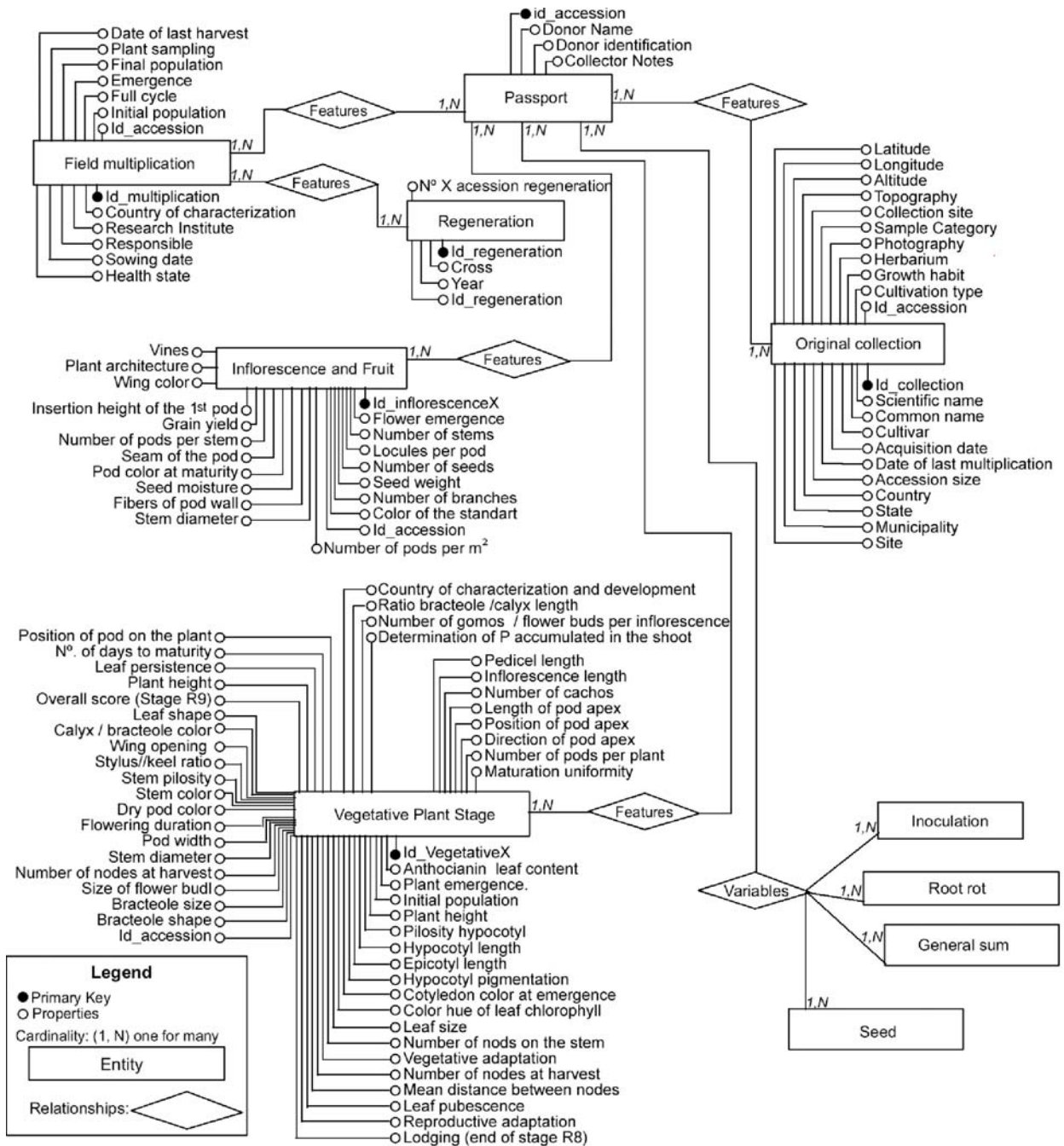


Figure 1. Mapping (Cardinality) and diagram of the Entity-Relationship

information and allowing data import and export in various formats by direct mail, and even exporting according to the standards used in statistical programs.

The MS Access system implements the generation of graphics in different ways using the records from the database. The operations in the database provide an

electronic environment that is useful for the curator's activities by using forms by which the data can be queried, viewed and edited in one or more tables, since some forms can be created that contain data from more than one table. The proper forms allow the maintenance of database with the options of inserting new data, alterations or updating

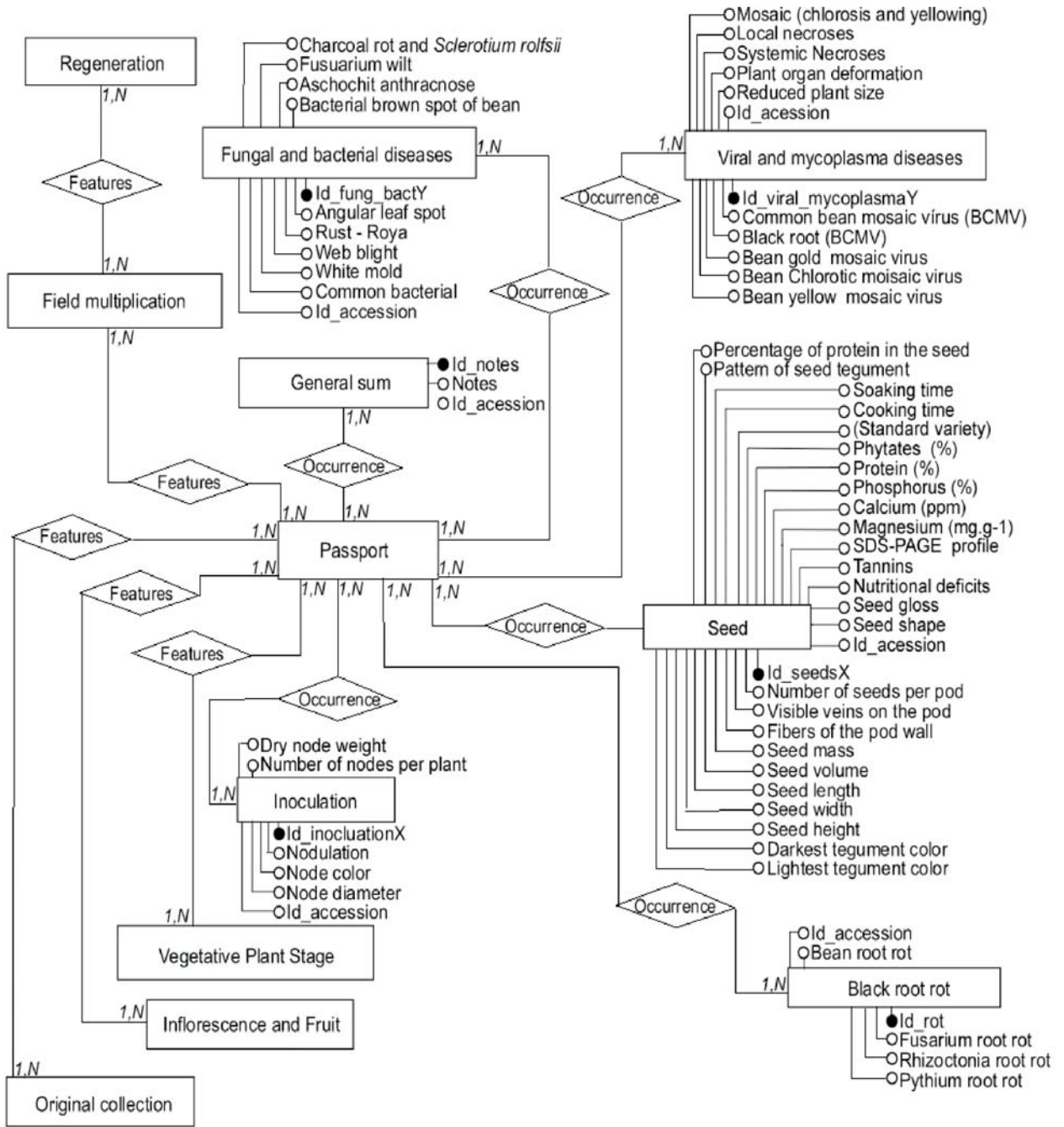


Figure 2. Mapping (Cardinality) and diagram of the Entity-Relationship



Figure 3. Cardinality between the two entities, passport and collection

of descriptors of germplasm assessment, usually used for data that are frequently changed and the exclusion of a germplasm or its characteristics. Figure 4 shows an example of a collection form.

The visualizations are based on forms and reports; the reports are generated by the forms and can perform a set of various inter-related forms by selecting the data relevant for each table, creating combinations to display these data in the reports. The records can be organized in a particular order and classified according to the objective of data visualization.

The accessions in the MS Access system recover, filter, sort and join data in one command, combine data from multiple tables in a single view. The queries find, display and process the data according to the user's instructions. This computerized search environment replaces the manual scanning in the CBGB catalog of target accessions.

Queries can also be performed by a SQL (Structured Query Language) that provides a powerful declarative query language. According to Silberschatz et al. (1999) the SQL uses a combination of constructors in algebra and relational calculus, is used so that queries written in them can be automatically optimized and executed efficiently - providing full power to the programming language for automatic optimization and to overcome limitations.

The common bean germplasm is available at www.imegem.cav.udesc.br/bdf.rar with the structure of tables, their relationships and cardinalities without the properties but only with the entities, which allows the insertion of specific data of accessions characterized by the responsible institution.

Sistema gerenciador de banco ativo de germoplasma de feijão

RESUMO – Os recursos genéticos são conservados *ex situ* em bancos de germoplasma, os quais são utilizados como fonte de alelos úteis. Os sistemas de Gerenciamento de Banco de Dados proporcionam um ambiente eficiente de armazenamento das informações geradas pelos Bancos Ativos de Feijão (BAF), por exemplo, promovendo uma recuperação destes dados de forma rápida e segura. Foi Desenvolvido em Microsoft Access, um sistema de banco de dados para armazenar e pesquisar em bases de combinação de acessos ou diretamente em descritores morfológicos, fisiológicos e de produção na cultura do feijão. O sistema denominado de Banco de Dados de Feijão (BDF) apresenta interface gráfica para operação na base de dados, facilitando e otimizando o tempo de busca nos catálogos de germoplasma.

Palavras chave: banco de dados, recursos genéticos, germoplasma, acessos.

Figure 4. Collection form and passport

Therefore, computer data bases advantageously replace traditional means of data storage. Based on the development and the results, characterized by flexibility, availability of information in the system and time savings it was concluded that the DBMS Microsoft Accession meets the demands of the active common bean germplasm bank *ex situ*, is practical and fast, and contributes significantly to the organization and safety of genetic resources conserved in genebanks.

ACKNOWLEDGEMENTS

We thank the UDESC, the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and the Programa Institucional de Bolsas de Iniciação Científica (PIBIC), for granting a scholarship and financial support of this research.

REFERENCES

- Carbonell SAM, Carvalho CRI and Pereira VR (2003) Qualidade tecnológica de grãos de genótipos de feijoeiro cultivados em diferentes ambientes. **Bragantia** **62**: 369-379.
- Chiorato AF, Carbonell SAM, Colombo CA, Dias LAS and Ito MF (2005) Genetic diversity of common bean accessions in the germplasm bank of the Instituto Agronômico – IAC. **Crop Breeding and Applied Biotechnology** **5**: 1-9.
- CIAT (1991) **Sistema Estudar para La Evaluación de Germoplasma de Frijol**. CIAT, Cali, 56p.
- IPGRI (2001) **Descritores para (*Phaseolus vulgaris* L.)**. IPGRI, Rome, 45p.
- Loarce Y, Gallego R and Ferrer EA (1996) A comparative analysis of the genetic relationship between rye cultivars using RFLP and RAPD markers. **Euphytica** **88**: 107-115.
- Matos L (2004) **Desvendando o Access**. Digerati Books, São Paulo, 96p.
- Muñoz G, Giraldo G and Fernández de Soto J (1993) **Descritores varietales: arroz, frijol, maíz, sorgo**. CIAT, Cali, 169p.
- Painting KA, Perry MC, Denning RA and Ayad WG (1995) **Guidebook for genetic resources documentation**. IPGRI, Rome, 317p.
- Rodrigues LS, Antunes IF, Teixeira MG and Silva JB (2002) Divergência genética entre cultivares locais e cultivares melhoradas de feijão. **Pesquisa Agropecuária Brasileira** **37**: 1275-1284.
- Santos E and Bettencourt E (2001) **Manual de apoio à formação e treino em Conservação *ex situ* de Recursos Fitogenéticos**. INIA e IPGRI-SSA, Nairobi, 221p.
- Silberschatz A, Korth HF and Sudarshan S (1999) **Sistema de banco de dados** 3ed, Pearson Makron Books, São Paulo, 778p.
- Silva HT (2005) **Descritores mínimos indicados para caracterizar cultivares/variedades de feijão comum (*Phaseolus vulgaris* L.)**. EMBRAPA Arroz e Feijão, Santo Antonio de Goiás, 32p.
- Takai OK, Italiano IC and Ferreira JE (2005) **Introdução a banco de dados**. DCC, IME-USP, Editoração Eletrônica, São Paulo, 124p. Available at: www.ime.usp.br/~jef/apostila.pdf. Accessed in January 08, 2009.