(Formerly MENDELIANA)

BAG

Journal of Basic & Applied Genetics

July 2021
Volume XXXII
Issue 1
E-ISNN: 1852-6322

Journal of the Argentine Society of Genetics
Revista de la Sociedad Argentina de Genética

www.sag.org.ar/jbag
Buenos Aires, Argentina

DOI: 10.35407/bag.2021.32.01

Journal of the
Argentine Society
of Genetics

BAG | Journal of Basic & Applied Genetics

V. XXXII - No.1

July 2021

Included in:

Cited by:

Comité Editorial

Imágen de tapa:

El poder de los colores de primavera.
Manzanilla común, *Chamaemelum nobile* L.
Javier Di Matteo

Contents

BAG | Journal of Basic & Applied Genetics

Vol XXXII (1): 7-10; July 2021

Journal of the Argentine Society of Genetics

MORE THAN A CENTURY OF CYTOGENETIC STUDIES IN CHILEAN PLANTS: HOW MUCH HAVE WE PROGRESSED?

MÁS DE UN SIGLO DE ESTUDIOS CITOGENÉTICOS EN PLANTAS CHILENAS: ¿CUÁNTO HEMOS PROGRESADO?

Jara Seguel P.[1,2], Palma Rojas C.[3]

[1] Departamento de Ciencias Biológicas y Químicas.

[2] Núcleo de Estudios Ambientales, Facultad de Recursos Naturales, Universidad Católica de Temuco, Casilla 15-D, Temuco, Chile.

[3] Departamento de Biología, Facultad de Ciencias, Universidad de La Serena, Casilla 599, La Serena, Chile.

Corresponding author:
P. Jara Seguel
pjara@uct.cl

ORCID 0000-0002-9340-1780

Cite this article as:
Jara Seguel P., Palma Rojas C. 2021. MORE THAN A CENTURY OF CYTOGENETIC STUDIES IN CHILEAN PLANTS: HOW MUCH HAVE WE PROGRESSED?. BAG. Journal of Basic and Applied Genetics XXXII (1): 7-10.

Received: 04/27/2021
Accepted: 04/30/2021
General Editor: Elsa Camadro
DOI: 10.35407/bag.2021.32.01.01
ISSN online version: 1852-6233

ABSTRACT

An overview is provided on the cytogenetic of Chilean plants, highlighting information gathered from more than a century of work carried out by foreign and national researchers who have contributed to the study of native species. We briefly present the progress made to date and also emphasize some strategies that, in our opinion, could spur further advances in this second century of cytogenetic studies in Chilean plants.

Key words: Cytogenetics, cytogenomics, Chilean plants.

RESUMEN

Se presenta una visión general de la citogenética de plantas chilenas, destacando información recopilada durante más de un siglo de trabajo realizado por investigadores nacionales y extranjeros que han contribuido al estudio de las especies nativas. Presentamos brevemente los progresos realizados hasta la fecha y también destacamos algunas estrategias que, en nuestra opinión, podrían impulsar mayores avances en este segundo siglo de estudios citogenéticos en plantas chilenas.

Palabras clave: Citogenética, citogenómica, plantas chilenas.

Available online at
www.sag.org.ar/jbag

INTRODUCTION

The cytogenetics of Chilean plants has had a fragmented development along its history, especially in its beginnings. In the last decades, however, it has made important contributions to the study of plant diversity, incorporating classical quantitative karyotype analysis and more recently modern cytogenomic methods. However, in the present, strategies to further progress have not yet been discussed among Chilean cytogeneticists.

The first study on the cytogenetics of Chilean plants reported the chromosome number of *Alstroemeria chilensis* Lem. (Syn. *Alstroemeria ligtu* L., Alstroemeriaceae), which was published by Strasburger (1882) almost at the end of the 19th century in the Archiv für Mikrobiologie und Anatomie in Germany. Later on, at the beginning of the 20th century, more studies on the cytogenetics of Chilean plants were published from 1929 onwards. Since then, relevant contributions have been made by foreign cytogeneticists such as Whyte (1929), Sato (1938, 1943), Goodspeed (1940), Titov de Tschicshow (1954) and Esponda (1970), who described chromosome number and morphology in species of several native genera (e.g., *Alstroemeria*, *Bomarea*, *Lapageria* and in back then *Hippeastrum*). At that same time, Sanz (1955, 1965, 1968, 1970), a Chilean cytogeneticist, made pioneering contributions applying cytogenetic methods to plants, focusing his work on native species of the genera *Alstroemeria*, *Calceolaria*, and *Leucocoryne*, among others. In later decades, a gradual increase in chromosome studies of four botanical divisions (Bryophyta, Pteridophyta, Pinophyta, and Magnoliophyta) including terrestrial and aquatic plants is observed, with major advances achieved since 2001 until now. Some reviews have discussed aspects on this subject which can be consulted for more details (Jara Seguel and Urrutia, 2012; Jara Seguel and Urrutia Estrada, 2018).

In this article we briefly present the progress on cytogenetic studies of Chilean plants made to date, and also suggest some strategies that, in our opinion, could spur further advances in this second century of cytogenetic studies.

HOW MUCH HAVE WE PROGRESSED?

At present, 122 publications on cytogenetics are available covering *ca.* 402 Chilean plant species (Jara Seguel and Urrutia Estrada, 2020). This number of studied species is equivalent to 6.5% of the total flora, according to statistics published in floristic reviews (*ca.* 6,103 land plant species; Villagrán, 2020). This percentage is alarmingly low compared to other regions around the world, with percentages ranging from 35.0% in Italy to 80.0% in New Zealand (Peruzzi *et al.* 2011). Unfortunately, in South America only Paraguay (with *ca.* 313 studied species; Jara Seguel and Urrutia, 2012) and Brazil (with *ca.* 699 species studied from the Cerrado Ecoregion; Roa and Telles, 2017) have estimations on the number of studied species, representing the only comparison parameter that we have to evaluate progress. The above paucity is coupled with scant funding for projects on this specific issue in Chile. Since 2007 only two government projects (FONDECYT) and one with academic funding were awarded to a research group of the Universidad de Concepción (Baeza C., Negritto M.; Repositorio ANID 2021). Our group (Jara Seguel P., Palma Rojas C.) receives financing annually from the Núcleo de Estudios Ambientales (project MECESUP UCT0804, 2011) of the Universidad Católica de Temuco, but the resources are mainly earmarked for operational expenses.

According to the number of species cytogenetically studied so far for Chilean plants, it is clear that in the initial century the progresses were few and intermittent (Jara Seguel and Urrutia 2012), experiencing difficulties such as the shortage of Chilean specialists, which led to a large part of the studies being carried out by foreign cytogeneticists. Publications recorded for the last decade show that two Chilean research groups (those mentioned above) maintain active productivity on the subject by focusing on various families and studying different cytogenetic features e.g., chromosome number, karyotype morphology, C-values, C and Ag-NOR banding, as well as cytogenomic markers e.g., 5S/45S rDNA localization through the application of fluorescent *in situ* hybridization (Baeza and Schrader, 2005; Baeza *et al.*, 2007; Cajas *et al.*, 2009; Jara Seguel *et al.*, 2012; Chacón *et al.*, 2012). Many of these chromosome data have been used to envisage phylogenetic and evolutionary hypotheses in some families (Chacón *et al.*, 2012; Jara Seguel *et al.*, 2021) with some species included in cytoevolutionary studies of global flora (Smarda *et al.*, 2014; Carta *et al.*, 2020). Plant taxonomy has also required cytogenetic support in the case of some families (Jara Seguel and Urrutia, 2012). Cytogeography is another incipient line of research in Chile which could be useful to understand patterns of distribution of cytogenetic diversity along the latitudinal and longitudinal gradients of the continent or in insular areas with different geographic locations and geological origins (Stuessy and Baeza, 2017; Jara Seguel *et al.*, 2020; 2021). Applications in conservation genetics can also be visualized as an interesting field of study in the near future (Jara Seguel and Urrutia, 2012; Jara Seguel *et al.*, 2020).

WHAT CAN WE DO IN THE FUTURE?

In this second century of cytogenetic studies in Chilean plants, much remains to be done and the challenges for the small Chilean community of cytogeneticists are great. According to the statistics *ca.* 93.5% of the Chilean native species have yet to be studied. In our opinion various strategies are required to make further progress, and we propose: i) the training of new specialists at the graduate and undergraduate level who are willing to address this discipline of genetics instead of other branches perhaps with a molecular approach, although obviously one does not exclude the other; ii) specialization may also be necessary for current researchers specifically in the fields of modern cytogenomics (Eykelenboom and Tanaka 2020) and cytoinformatics (e.g., chromEvol, Mayrose *et al.*, 2010), thus making it possible to increase the critical mass of highly specialized cytogeneticists; iii) the structuring of research groups necessarily has to be planned to ensure that they work in cooperation in order to streamline efforts and financial resources, either prioritizing endemic species (*ca.* 45% of the Chilean continental flora) or those with conservation problems (*ca.* 70 critically threatened species); iv) monetary resources could be raised from companies (mining, forestry, agricultural), charitable trusts, philanthropists and environmentalists, i.e. all those social groups that are linked to the direct use of native plants or that cause effects on them. The future advances in cytogenetic studies of Chilean plants require the contribution of various actors such as government, academia, research centers and economic groups. In this third millennium, in the middle of the post-genomic era, progress in cytogenetic studies is highly necessary for Chilean plants, especially with the advent of global climate change that is strongly affecting the flora and the ecosystems. In this scenario of biodiversity threat, it is necessary to understand evolutionary aspects of Chilean plants, allowing cytogenetics to contribute fundamental knowledge that could be included in modern phylogenomic studies (Posada, 2016), encompassing the analysis of different genome compartments (e.g., nuclear, plastidial, and mitochondrial).

ACKNOWLEDGEMENTS

We would like to express our thanks to Núcleo de Estudios Ambientales (NEA), Dirección General de Investigación of the Universidad Católica de Temuco, Chile.

BIBLIOGRAPHY

Baeza C., Schrader O. (2005) Comparative karyotype analysis in *Haplopappus* Cass. and *Grindelia* Willd. (Asteraceae) by double FISH with rDNA specific genes. Plant Syst. Evol. 251:161-172.

Baeza C., Schrader O., Budahn H. (2007) Characterization of geographically isolated accessions in five *Alstroemeria* L. species (Chile) using FISH of tamdely repeated DNA sequences and RAPD analysis. Plant Syst. Evol. 269:1-14.

Cajas D., Baeza C., Ruiz E., Negritto M. (2009) Análisis citogenético en poblaciones de *Alstroemeria hookeri* Lodd. ssp. *hookeri* (Alstroemeriaceae) en la Región del Bío-Bío, Chile. Gayana Bot. 66(2):117-126.

Carta A., Bedini G., Peruzzi L. (2020) A deep dive into the ancestral chromosome number and genome size of flowering plants. New Phytol. 228:1097–1106.

Chacón J., Sousa A., Baeza C., Renner S. (2012) Ribosomal DNA distribution and a genus-wide phylogeny reveal patterns of chromosomal evolution in *Alstroemeria* (Alstroemeriaceae). Am. J. Bot. 99:1501-1512.

Eykelenboom J., Tanaka T. (2020) Zooming in on chromosome dynamics. Cell Cycle 19(12):1422-1432.

Esponda P. (1970) Cytotaxonomy of two species of the genus *Hippeastrum* (Amarillidaceae) Cytologia 35(3):431- 433.

Goodspeed T. (1940) Amaryllidaceae of the University of California botanical expedition to the Andes. Herbertia 7:17-31.

Jara Seguel P., Palma Rojas C., Contreras J., von Brand E. (2012) Chromosome localisation of nucleolar organizer region in *Rhodophiala bagnoldii* (Herb.) Traub (Asparagales: Amaryllidaceae) determined by silver nitrate staining. Gayana Bot. 69(1): 201-203.

Jara Seguel P., Urrutia J. (2012) Cytogenetics of Chilean angiosperms: advances and prospects. Rev. Chil. Hist. Nat. 85:1-12.

Jara Seguel P., Urrutia Estrada J. (2018) Chilean Plants Cytogenetic Database: coverage, characters and usages. J. Basic Appl. Genet. 29:65-69.

Jara Seguel P., Urrutia Estrada J. (2020) Chilean Plants Cytogenetic Database. http://www.chileanpcd.com/ (accessed April 2021).

Jara Seguel P., Urrutia Estrada J., Vallejos N., Andrade E., Jara M. (2020) Chromosome number variation in part of the flora of protected wild areas in the Araucania Region of southern Chile. J. Basic. Appl. Genet. 31:27-38.

Jara Seguel P., Jara Arancio P., Andrade E., Urrutia Estrada J., Palma Rojas C., Araya Jaime C. (2021) Cytogenetics of wild species of the Alstroemeriaceae family (Liliales). Plant Syst. Evol. 307:34. https://doi.org/10.1007/s00606-021-01756-1.

Mayrose I., Barker M., Otto S. (2010) Probabilistic models of chromosome number evolution and the inference of polyploidy. Syst. Biol. 59:132–144.

Peruzzi L., Dawson M., Bedini G. (2011) Chromosome number variation in two antipodean floras. AoB Plants. doi:10.1093/aobpla/plr020.

Posada D. (2016) Phylogenomics for Systematic Biology. Syst. Biol. 65(3):353-356.

Proyecto MECESUP UCT0804 (2011) Fortalecimiento de las capacidades científicas y del post-grado a través del desarrollo de núcleos de investigación en las áreas prioritarias de la Universidad Católica de Temuco. Dirección General de Investigación y Posgrado, Universidad Católica de Temuco, Chile.

Repositorio ANID (2021) Agencia Nacional de Investigación y Desarrollo, Chile. http://repositorio.conicyt.cl/handle/10533/108045/browse?authority=f0574ffe-3fed-4a87-b37c-1be07497d275&type=author. (accessed April 2021)

Roa F., Telles M. (2017) The Cerrado (Brazil) plant cytogenetics database. Comp. Cytogen. 11(2): 285–297.

Sanz C. (1955) Observaciones cromosomales en plantas chilenas (II parte). Agricultura Técnica (currently Chil. J. Agric. Res.) 15(1):5-11.

Sanz C. (1965) Observaciones cromosomales en plantas chilenas (III parte). Agricultura Técnica (currently Chil. J. Agric. Res.) 25 (3): 124-127.

Sanz C. (1968) Determinación cromosomal en Huilli, *Leucocoryne ixioides* Lindl. Agricultura Técnica (currently Chil. J. Agric. Res.) 28(1):10.

Sanz C. (1970) Observaciones cromosomales en *Lapiedra chilensis*. Agricultura Técnica (currently Chil. J. Agric. Res.) 30(1):110.

Sato D. (1938) Karyotype alteration and phylogeny. IV. Karyotypes in Amaryllidaceae with special reference to sat-chromosomes. Cytologia 9:226-240.

Sato D. (1943) Karyotype alteration and phylogeny in Liliaceae and allied families. J. Jap. Bot. 12:57-161.

Smarda P., Bures P., Horová L., Leitch I., Mucina L., Pacini E., Tichý L., Grulich V., Rotreklová O. (2014) Ecological and evolutionary significance of genomic GC content diversity in monocots. PNAS 111(39):4096-4102.

Strasburger E. (1882) Über den Teilungsvorgang der Zellkerne und das Vehaltnis der Kernteilung zur Zellteilung. Arch. Mikrobiol. Anat. 21:476–596.

Stuessy T., Baeza M. (2017) Chromosome numbers. In: Stuessy T., Crawford D., López-Sepúlveda P., Baeza C., Ruiz E. (Eds.) Plants of oceanic islands. Evolution, biogeography, and conservation of the flora of the Juan Fernández (Robinson Crusoe) Archipelago. Cambridge University Press, London, pp. 165-187.

Titov de Tschicshow N. (1954) Estudios citológicos en *Lapageria rosea* Ruiz et Pavon. Boletín de la Sociedad de Biología de Concepción 29: 3-6.

Villagrán C. (2020) Historia biogeográfica de las briófitas de Chile. Gayana Bot. 77(2):73-114.

Whyte R. (1929) Chromosome studies. I. Relationships of the genera *Alstroemeria* and *Bomarea*. New Phytol. 28:319-344.

BAG | Journal of Basic & Applied Genetics

Vol XXXII (1): 11-24; July 2021

Journal of the Argentine Society of Genetics

Ex situ PLANT GERMPLASM CONSERVATION REVISED AT THE LIGHT OF MECHANISMS AND METHODS OF GENETICS

CONSERVACIÓN DE GERMPLASMA *ex situ* REVISADA A LA LUZ DE MECANISMOS Y MÉTODOS DE GENÉTICA

Camadro E.L.[1], **Rimieri P.**[2]

[1] Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) and Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata (FCA, UNMdP), Ruta Nacional 226 km 73.5, 7620 Balcarce, Argentina.

[2] Estación Experimental Agropecuaria Pergamino, Instituto Nacional de Tecnología Agropecuaria (INTA), Avenida Frondizi (Ruta 32) km 4.5, 2700 Pergamino, Argentina.

Corresponding author:
Elsa L. Camadro
camadro.elsa@inta.gob.ar

ORCID 0000-0002-1739-5059

Cite this article as:
Camadro E.L., Rimieri P. 2021. *Ex situ* PLANT GERMPLASM CONSERVATION REVISED AT THE LIGHT OF MECHANISMS AND METHODS OF GENETICS. BAG. Journal of Basic and Applied Genetics XXXII (1): 11-24.

Received: 06/15/2021
Accepted: 06/23/2021

General Editor: Elsa Camadro

DOI: 10.35407/bag.2021.32.01.02

ISSN online version: 1852-6233

Available online at
www.sag.org.ar/jbag

ABSTRACT

Plant genetic resources for food and agriculture are *ex situ* conserved in germplasm banks as samples (accessions) of natural or naturalized populations, either as the originally sampled propagules (mainly seeds) or their multiplications. The premises underlying *ex situ* conservation are that (a) it is the safest and cheapest alternative for germplasm preservation for future generations and (b) accessions are representative of the genetic diversity encountered in nature. In the past decades, ideas, alternatives and considerations have been put forward on the topic, and protocols have been devised for plant germplasm sampling, conservation and multiplication. However, limitations in the management efficiency of germplasm banks have been pointed out by international organizations. In our opinion, germplasm banks in general need to revise their functioning and management at the light of principles and methods of Genetics. To that end, it is necessary to consider the reproductive biology of higher plants -whose genetic consequences at both the individual plant and the population levels are not always either fully understood or taken into account in devising the protocols-, the genetic structures of wild and cultivated populations, and the course of the genetic material in the populations. In this paper, we discuss the three topics and provide an example of a national forage breeding program, from germplasm bank accessions as the germplasm of origin to the obtainment of commercial cultivars. Finally, we present a proposal as a base for discussion among curators, researchers and breeders.

Key words: accessions, breeding, genetic resources, germplasm banks, population genetics

RESUMEN

Los recursos genéticos vegetales para la alimentación y la agricultura se conservan *ex situ* en bancos de germoplasma como muestras (introducciones) de poblaciones naturales o naturalizadas ya sea como propágulos originales (mayoritariamente semillas) o sus multiplicaciones. Las premisas subyacentes son que (a) es la alternativa más segura y barata de preservación de germoplasma para futuras generaciones y (b) las introducciones son representativas de la diversidad genética que se encuentra en la naturaleza En las últimas décadas, se han presentado ideas, alternativas y consideraciones sobre el tema y se han elaborado protocolos para el muestreo, conservación y multiplicación de germoplasma. Sin embargo, organizaciones internacionales han señalado limitaciones en la eficiencia del manejo de los bancos de germoplasma. En nuestra opinión, se necesita revisar el funcionamiento y manejo de dichos bancos en general a la luz de los principios y métodos de Genética. Para tal fin, es necesario considerar la biología reproductiva de las plantas superiores -cuyas consecuencias genéticas a nivel de planta individual y de población no se comprenden en su totalidad o no se consideran al idear los protocolos -, las estructuras genéticas de poblaciones naturales y cultivadas, y el curso del material genético en las poblaciones. En este trabajo discutimos los tres temas y proveemos un ejemplo de un programa nacional de mejoramiento de forrajeras, desde las introducciones como germoplasma de origen hasta la obtención de cultivares comerciales. Finalmente, presentamos una propuesta como base de discusión entre curadores, investigadores y mejoradores.

Palabras clave: introducciones, mejoramiento genético, recursos genéticos, bancos de germoplasma, genética de poblaciones

INTRODUCTION

With the aim of contributing to the development of coherent and effective strategies for conservation of plant genetic resources for food and agriculture, ideas, alternatives and considerations have been put forward over the years in many methodological publications. Limitations in the management efficiency of germplasm banks, not infrequently carried out without appropriate planning, were pointed out in "The State of the World´s Plant Genetic Resources for Food and Agriculture" (FAO, 1996). In that report, it was considered that over 65% of the worldwide *ex situ* conserved collections needed regeneration. Almost 10 years later, the logistics of germplasm banks was integrally analyzed in the last manual published by Biodiversity International (previously IBPGR or International Board for Plant Genetic Resources) (Engels and Visser, 2006). As judged by the magnitude of the advancements made over the previous decades at the global level, the authors recognized that the response of germplasm banks had been scarce regarding the utilization of the appropriate strategies for the *ex situ* conservation of collections. For curators, this manual constituted a guide for adopting a more critical, balanced and creative approach to germplasm conservation. Useful information was presented on various management aspects to solve frequently encountered operative problems with the incorporation of new and better technologies. In particular, important elements were analyzed and options were discussed to improve the efficiency and effectiveness of operations both according to costs and by taking into account genetic and economic implications for rationalization of the logistics. From a further analysis of the history and evolution of germplasm banks, it was concluded that these banks had gone through periods of questioning about their function or operativity. Among others, the following reasons were given: limited resources; excess or loss of accessions; lack of representativeness of the natural genetic diversity in the accessions, modifications in conservation and multiplication protocols, and changes in the conservation objectives due to the demands of breeding (development of commercial varieties) and agroecological programs (preservation of local varieties or landraces).

More than a decade has gone by since the publication of Engels and Visser´s (2006) document. However, in our opinion, there is still a need to revise the functioning and management of germplasm banks in general. We consider that it is timely to present an approach at the light of principles and methods of Genetics. In this regard, the principles and methods established and used at the individual level (cell, tissue, organ, organism) (e.g., what is the genetic material, how it is transmitted and arranged, how it changes and functions) are not the same as those established and used at the population level (which are related to the course of the genetic material in the populations). We consider that our proposal -based on considerations of the modes of reproduction and their genetic consequences, the genetic structures of wild and cultivated populations, and principles of population genetics- could serve as a base document for discussion among curators, researchers and breeders on the adequacy of the current protocols for *ex situ* conservation of the natural genetic diversity. To the best of our knowledge, this approach relating gametes, gene flow, fertilization and other biological phenomena that have important genetic components has not previously been integrally and routinely used. In this regard, there are many examples in the literature in which the "structure" of collections of wild or cultivated species has been assimilated to the "genetic structure" of populations or in which the term has been used in regard to the total genetic diversity and its partitioning at various levels by means of statistical analysis (ANOVA, AMOVA, STRUCTURE program), even though the definition of "genetic structure" in Genetics is clearly different, as it will be discussed. Moreover, for some statistical analyses (e.g., traditional cluster analysis) it has been considered appropriate to assume that sexual reproduction can occur either by autogamy or allogamy and, therefore, that a population of an autogamous species is genetically homogeneous and a population of an allogamous species is genetically heterogeneous. However, the variability that can be encountered in the genetic structure of a natural population at a given time would depend, among other factors, on the preponderant mode and type of reproduction of the population of origin, as it will be explained.

Ex situ CONSERVATION

Plant germplasm conservation is mainly carried out *ex situ* in the form of samples of propagules (accessions). These propagules can be either the originally sampled ones in natural or naturalized populations, or their regenerants obtained in the same bank or from inter-bank exchange. In the last decades, there has been a change in emphasis away from this type of conservation and towards the *in situ* conservation of locally adapted landraces and crop wild relatives (CWR) within or outside protected areas (Maxted *et al.*, 1997; Maxted *et al.*, 2016; FAO, 2017). However, *ex situ* conservation has advantages and disadvantages *per se* and in relation to other conservation methods (Kjaer *et al.*, 2001, in Hammer and Teklu, 2008); thus, the *ex situ* and *in situ* approaches are complementary, fulfilling different purposes.

Plant accessions are usually conserved under specific categories, mainly assigned according to

morphological phenotypes, with the relatively more recent incorporation of molecular tools (see Camadro, 2012). This type of classification into taxonomic or typological species (TS) responds to the Taxonomic Species Concept (TSC); according to this concept, species are immutable entities because they have reached the end of the evolutive process. Plants can also be classified as biological species (BS) on the basis of breeding relationships when the Biological Species Concept (BSC) is applied, regardless of their morphological phenotypes. TS and BS do not necessary overlap; thus, the use of the term ''species'' generates much confusion when the distinction between them is not clearly made (see Grant, 1981). Moreover, taxonomic categories are periodically subjected to revision because they are human constructions. Thus, taxonomic nomenclatures and "species" numbers in a given plant group can vary over the years according to the taxonomist(s) involved in the task. For example, the number of potato "species" (*Solanum* L. section Potato; Dicotyledoneae) has been reduced in the last 40 years from approximately 235 (seven of them cultivated and 228 wild) to 203, 189 and 111 (four of them cultivated and 107 wild) (in Poulsen Hornum and Camadro, 2021), whereas in brome grasses (*Bromus* L. section *Ceratochloa*), with approximately 160 recognized "species", the large morphological variation encountered in the section led Williams et al. (2011) to point out that "Hybridization is rife in this section, making species boundaries obscure and the taxonomy very difficult". Notwithstanding, and as previously stated, collections are assigned specific categories for their incorporation and conservation as accessions in germplasm banks, without specification of the concept (either TSC or BSC) used for their classification (see an example at http://www.ars-grin.gov/npgs/collections.html). The species concept employed in the taxonomy of a plant group, however, has genetic consequences for both conservation and seed regeneration and multiplication protocols (see Poulsen Hornum and Camadro, 2021).

Germplasm bank accessions can be composed of (a) seeds of sexually reproducing or apomictic plants; (b) plants derived from vegetative organs (e.g., tubers, stolons, corms, leaves) cultivated in the field, or plantlets cultivated *in vitro*; (c) pollen, embryos or tissues conserved in liquid nitrogen (FAO, 2017). This type of conservation is justified when: (a) natural or naturalized populations are subjected to -or at risk of being subjected to- genetic erosion, or are affected by the extinction of native or naturalized plant communities; (b) there is a need for developing or complementing breeding programs through pre-breeding in less domesticated species, or for complementing working collections in breeding programs of advanced-breeding species for transferring genes or gene combinations from unexploited sources; (c) there are lines, clones or compounds synthesizing general adaptation, agronomic aptitude and productive potential that have been discarded in breeding programs, or varieties of reference that have been replaced by new ones in the commercial circuit but that can eventually be of value in breeding; (d) there are landraces or old varieties, often linked with traditional food products and organoleptic properties, that have cultural or economic value (or both) for small farmers.

¿WHAT PART OF THE GENETIC DIVERSITY NEEDS TO BE PROTECTED?

Ex situ conservation steps from the premises that (a) this form of conservation is the safest and cheapest alternative for preserving plant genetic resources for forthcoming generations, and (b) accessions are representative of the diversity encountered in the environments from which they were sampled: spatial (landscape, plant communities), morphological, and molecular. The two premises -along with the provision of detailed passport information- are important. However, an approach is needed to ensure that accessions faithfully represent both the sampled populations and the portion of the genetic diversity that needs to be protected. It has to be taken into account that genetic drift can occur if, in planning the operations, there is not a strict consideration of a combination of various phenomena. These can span from manipulations at the sampling time to various aspects of reproductive genetics during seed regeneration or multiplication, including the possible action of internal crossing barriers within accessions, e.g., male sterility, pollen-pistil incompatibility, nuclear-cytoplasmic genome interactions, among other biological phenomena (see Camadro 2012; Poulsen Hornum and Camadro, 2021). Thus, the estimation of genetic diversity ought to be complemented with detailed information on the genetic structure and reproductive biology of the population at the sampling time and, fundamentally, during the *ex situ* regeneration or multiplication processes. This last concept, if not integrally applied, nullifies the premise of security, economics and representativeness of the accessions because duplicates would not be detected and some gene (allele) frequencies might be unknowingly increased, decreased or eroded during the multiplication process. In summary, the genetic diversity and variability represented by an accession could be unnoticedly changed during propagule regeneration or multiplication; as a consequence, the accession would no longer represent the actual diversity and variability of the sampled population (Hammer and Teklu, 2008; Erazzú *et al.*, 2009; Cadima *et al.*, 2017; Poulsen Hornum and Camadro, 2021).

GERMPLASM BANKS

Many germplasm banks had their origin in plant breeding and research programs and were not necessarily designed to assimilate genetical approaches for *in situ* and *ex situ* conservation. Thus, it is important to critically examine the precise objectives of germplasm banks to identify possible limitations in their functioning. If clear objectives are established, it would be feasible to plan what genetic resources should be conserved and to choose the most adequate protocols for that end, establishing priorities and recognizing limitations and the biological complexities of the species of interest, including the form of propagation., Frankel (1984) proposed to establish core collections to facilitate germplasm management and use after defining the objectives. Core collections are collections of limited size, with minimum similarity among the composing accessions and much smaller than the collection(s) from which they were derived. Or as defined by Johnson and Hodgkin (1999), a core collection is a subset of one collection that represents with minimum repetition the genetic diversity of a cultivated species and its wild relatives.

A CONSERVATION APPROACH BASED ON THE GENETIC STRUCTURE OF POPULATIONS

The main objective of *ex situ* conservation is to have the maximum genetic diversity of a species represented in the accessions, previous establishment of the necessity of conservation, the increment of the number of propagules, and the maintenance of this diversity for conservation and exchange. These aspects ought to be known to define the representativeness of the originally sampled population in the accession. As complements, gaps and priorities have to be identified in the collection for conservation of strategic genetic resources and the determination of their potential applied value.

GENETIC MAKEUP OF POPULATIONS AND INDIVIDUALS IN NATURE

From a biological perspective, a natural population is a community of potentially inter-breeding individuals growing at a given locality, which share a common gene pool and represents a dynamic panmictic unit (Johansen 1903 and Dobzhansky, 1935, in Rieger *et al.*, 1976). The largest group of potentially inter-breeding individuals is the species which, in turn, is composed of local populations, each of them inter-communicating and inter-grading with the others. The sum of all factors governing the pattern by which gametes of various individuals unite with each other during fertilization makes up the population structure which, in nature, is a consequence of gene flow rates and environmental heterogeneity (Gilmoure and Gregor, 1939, in Rieger *et al.*, 1976).

By extension, the genetic structure of a population, either natural or artificial, is the type, quantity and distribution of the genetic variation present in that population expressed in terms of gene (allele) or genotypic frequencies. Thus, the genetic structure of a population depends on the mode and type of reproduction of the plant group or species that conform it. In this regard, it has to be taken into account that higher plants can reproduce either sexually or asexually, or have both types of reproduction available to them; consequently, the genetic structure of a given population can vary over time.

MODES OF REPRODUCTION AND GENETIC CONSEQUENCES

Sexual Reproduction

The production of sexual propagules (sexual seeds) entails the formation of *n* megaspores and *n* microspores (pollen grains or male gametophytes) by meiosis, followed by the formation of *n* female gametes and *n* male gametes by post-meiotic mitosis. The double fertilization of the *n* egg cell and the binucleated (*n* + *n*) central cell of the female gametophyte (embryo sac), each by one of the two *n* male gametes carried by the microspore, originates one 2*n* cell and one 3*n* cell which, respectively, give rise to the 2*n* embryo and the 3*n* endosperm by mitosis (Dumas and Mogensen, 1993). The events involved in sexual reproduction allow for the occurrence of two rounds of genetic recombination: (1) at meiosis, by segregation of chromosomes and genes, and (2) at fertilization, by nuclear fusion of the uniting gametes. Therefore, each sexual cycle provides the opportunity for the formation of new genotypic combinations.

Autogamy and allogamy

There are two types of sexual reproduction: allogamy or cross-fertilization and autogamy or self-fertilization. Allogamy maintains heterozygosity at most loci if the breeding population is large enough, whereas strict self-fertilization leads to homozygosity in most loci and, eventually, to allele fixation.

Two main factors promote allogamy: spatial and temporal separation of sexual organs. Spatial separation can occur (a) within the plant itself, e.g. maize (*Zea mays* L.), which bears female and male inflorescences at different positions along the axis, and (b) between plants, e.g. asparagus (*Asparagus officinalis* L.), with individual plants bearing only one type of imperfect flowers,

either with stamens or pistils (occasionally, perfect flowers are formed in either type of plant, allowing self-fertilization). Temporal separation (dichogamy) is the result of differences in the maturation time of female and male reproductive organs (protogyny and protandry, respectively), which in a plant can occur in (a) flowers or inflorescences along the axis, e.g. maize, or (b) within an inflorescence, e.g. carrot (*Daucus carota* L.) and sunflower (*Helianthus annuus* L.). However, there could be simultaneous maturation (homogamy) without autogamy in the presence of other factors: (a) chasmogamy (the flower is open when pollen is shed and/or the stigma is receptive) in otherwise cleistogamous flowers (the pollen is shed and the stigma is receptive when the flower is closed), e.g. *Bromus* spp. section *Ceratochloa* (Wolff *et al.*, 1996; Langer and Wilson, 1965; Leofanti *et al.* 2013); (b) hercogamy (physiological barriers), in plants with genetically controlled self-incompatibility systems in which the flowers are either (b_1) homomorphic (of one morphological type), e.g. potatoes and tomatoes (*Solanum* L. spp.), stone fruits such as almonds and cherries (*Prunus* L. spp.), Crucifers (*Brassica* L. spp.) such as cabbage, colza and kale, among others, or (b_2) heteromorphic, e.g. common flax (*Linum usitatissimum* L.) and loosestrife (*Lythrium junceum* Banks & Sol.); and (c) sterility (being male sterility the most frequent type) due to malformations in the reproductive organs or abnormalities in meiosis that prevent either production of viable pollen or its release from the anthers and, thus, self-fertilization. Breakdown of hercogamy, dichogamy, or self-incompatibility precedes the shift of the breeding system from obligate outcrossing towards autogamy due to structural and positional changes in the hermaphrodite flower, bud pollination and, finally, cleistogamy (in Frankel and Galun, 1977).

Autogamy and allogamy have both specular positive and negative characteristics. The positive characteristics of autogamy *vs*. allogamy are: genotype fixation and genotype specialization, which result in thriving of adapted genotypes over time in stable environments; guaranteed fertilization with economy of pollen; and adaptation to long distance dispersal because only one seed can start a population. The negative characteristics of autogamy are the other face of the coin: genetic inflexibility due to a lower capacity of "genetic storage" (of alleles and intra-locus and inter-loci interactions) and, thus, inability of the population to cope over time with changing environments ("evolutive compression"); and unguaranteed fertilization with the consequent waste of pollen.

Asexual Reproduction

Asexual propagules can originate by means of (a) seeds (agamospermy) or (b) other structures (agamic or vegetative reproduction). In agamospermy, there could be morphological alternation of generations or not. There is morphological alternation of generations when diplosporous or aposporous 2*n* gametophytes are formed, respectively, from 2*n* archesporial or 2*n* somatic cells, and either the 2*n* egg or other 2*n* cell of the gametophyte develops parthenogenetically in a process accompanied by the development of the endosperm either after fertilization of the central cell (pseudogamy) or without fertilization of this cell. On the other hand, there is no alternation of generations if the 2*n* embryos develop by adventive embryony or sporophytic budding from cells of the nucellus or integuments of the ovule (somatic embryogenesis) (Asker, 1980; Burnham, 1980). In plants with agamospermous reproduction, embryos (a) can be clones of the mother plant if they originate by somatic embryogenesis, apospory, or diplospory with a modified meiosis genetically equivalent to a mitosis, or (b) can genetically differ from the mother plant if the modified meiosis in diplospory entails a certain amount of recombination. In plants with agamic or vegetative reproduction, propagules (bulbs, corms, tubers, stolons, or rhizomes, among other structures) are formed by mitosis in somatic tissues, thus, they are clones of the mother plant.

ARE THE MODES AND TYPES OF REPRODUCTION STRICT?

Higher plants may have more than one mode or type of reproduction as a result of genotype x environment interactions. Sexually reproducing plants can be (a) autogamous, e.g. wheat (*Triticum aestivum* L.), tobacco (*Nicotiana tabacum* L.), garden tomato (*Solanum lycopersicum* L.); allogamous, e.g. maize, carrot, garden asparagus; (b) autogamous with a percentage of allogamy, e.g. beans (*Phaseolus* L. spp.); (c) allogamous with a percentage of autogamy, e.g. maize, sunflower (*Helianthus annuus* L.), asparagus. Autogamous plants could be considered a prelude to evolutionary extinction if it were not for the fact that local differentiation in ecological niches maintains a massive storage of genetic diversity (in Frankel and Galun, 1977). Similarly, asexual reproduction is not strict; otherwise, it will also be an end road in evolution. It is frequently combined with sexual reproduction by allogamy, e.g. potatoes, grasses.

NATURAL AND NATURALIZED POPULATIONS

Sexually Reproducing Species

In autogamous species, individual plants with disomic inheritance (diploids and disomic polyploids, e.g. 2x *Triticum monococcum* L., 4x *T. turgidum* L., 6x *T.*

aestivum L.) are expected to be highly homozygous for one genetic combination (Fig. 1a) or more than one (Fig. 1b). Populations of autogamous species, however, can be genetically homogeneous to a greater or lesser extent depending on whether they have a percentage of allogamy or not. For example, the percentage of allogamy in Proso millet (*Panicum miliacium* L.), with wind-dispersed pollen, can be more that 10%, whereas in Lima beans (*Phaseolus lunatus* L.), with bees-dispersed pollen, this percentage can range from 0% to 80%. Moreover, the proportion of cleistogamous *vs.* chasmogamous flowers (e.g., in *Lespedeza* Michx. ssp.) could variably increase the percentage of allogamy in a given season (in Frankel and Galun, 1977). If individual plants have opportunities for hybridization even from time to time, the population can be composed of plants either homozygous for one genetic combination (Fig. 1a) or more than one (Fig. 1b), or heterozygous for one or more loci (Fig. 1c) because they might be F_1 hybrids, backcrosses to the homozygous parents, or advanced segregating generations. Therefore, populations can be either homogeneous or heterogeneous in various degrees. In inbreeding species, the variation among populations is expected to be larger than within populations in contrast with outbreeding species. In a review of experiments carried out with isozymes in autogamous and allogamous species, Schoen and Brown (1991) found that inbreeders exhibited markedly greater population variation than outbreeders according to Nei´s gene diversity statistics.

On the other hand, allogamy is obligate only in monoecious species with strict self-incompatibility systems, and in dioecious species. The spatial and temporal separation of the reproductive organs, as previously explained, promotes but does not force this type of sexual reproduction. In individual plants of both diploid and polyploid allogamous species, most loci are expected to be in heterozygosity, although there could also be loci in homozygosity. Natural populations are expected to be highly heterogeneous (Fig. 1d), being the genetic diversity higher within than between populations as demonstrated, for example, in wild potatoes (Bedonni and Camadro, 2009; Erazzú *et al*. 2009).

Asexually reproducing species

A few higher plants exhibit only asexual reproduction (e.g. garlic, *Allium sativum* L.) but most plants with this mode of reproduction can also reproduce sexually under certain environmental conditions (see Frankel and Galun, 1977). The environmental conditions can modify not only the proportion of allogamy in sexual reproducing plants, as previously explained, but the preponderant mode of reproduction of a given population as well. Examples can be found in apomictic grasses (Knox, 1967; Quarin, 1986; Rebozzi *et al*., 2011) and wild potatoes (Leofanti *et al*., 2019), among other plant groups. It is a common mistake to consider that natural populations of asexually reproducing plants are genetically homogeneous. On the contrary, these populations can be composed of plants of either the same genotype (one clone; Fig. 1e and 1g) or different genotypes (more than one clone; Fig. 1f and 1h) because asexual reproduction is usually combined with sexual reproduction by allogamy. Therefore, a population with the two alternative modes of reproduction can be a mix of clones as a result of either hybridization followed by vegetative reproduction in the subsequent generations or facultative apomixis. Individual plants of asexually reproducing species can be highly heterozygous, but some loci can be in homozygosis. Populations with asexual reproduction can be either homogeneous or heterogeneous in various degrees (see Ellstrand and Roose, 1987).

Summarizing, a thorough knowledge and understanding of the reproductive biology and genetics of the species of interest is needed in order to (1) develop the appropriate sampling and regeneration protocols to try to capture an important amount of the genetic diversity present in a population, and (2) avoid or minimize gene (allele) erosion during seed regeneration. Moreover, and given that the types and modes of reproduction are not necessarily strict in a given plant group and a given environment, it is: (a) inappropriate to carry out statistical analyses under the assumption that populations have only one type of reproduction (e.g. for sexually reproducing species, either autogamy or allogamy) and, therefore, that they there are genetically either homogeneous or heterogeneous, and (b) advisable to resample the populations in environmentally contrasting years, whenever possible. In this regard, samples of a given population taken in different moments should be used to conform the accession (instead of naming each sample as a new accession) to maximize the amount of the captured natural genetic diversity at a given site. It is our opinion that no specific guidelines should be given for curators. Instead, and based on the knowledge of the reproductive biology and genetics of the plant species or group of interest, the principles and methods of population genetics should be applied to prevent or reduce gene erosion in the conserved germplasm.

BREEDING POPULATIONS

Genetic makeup

Rimieri (2017) has pointed out that it is necessary to differentiate *ex situ* and *in situ* conserved plant genetic resources from those plant resources collected, maintained and utilized for human subsistence, which

are the result of the application of selection or breeding methods. According to this approach, the protection of the biodiversity and the application of mutagenic, biochemical, molecular and genetic engineering tools are compatible and complementary.

Plant breeding is the heritable improvement of plants, usually acknowledged as a combination of art and science. Approximately 11,000 years ago, domestication of plants and animals evolved from the hunter-gatherer lifestyle. But it was in the 20th century, with the rediscovery of Mendel´s laws of inheritance, that plant breeding became an applied discipline, which makes use of principles from a variety of other disciplines to improve the genetic potential of plants cultivated for food, feed, and/or metabolites of interest, among others. Plant breeders make use of conventional methods (parental selection, controlled crosses, progeny selection) to introduce desirable traits to their object of improvement (Gallais, 1990; Allard, 1999) with the relatively more recent aid of biotechnologies, e.g., transgenesis, cisgenesis, intragenesis, and gene edition (Al-Khayri *et al*. 2015; Cardi, 2016). In spite of the advancements in genome manipulation, plant breeding remains a high time- and resource-consuming process, particularly in crop species with narrow genetic bases.

The final products of plant breeding are cultivated varieties or *cultivars* (a term coined by contracting the two previous terms to establish a difference with *botanical varieties*, which correspond to a taxonomic rank between subspecies and form). Cultivars are obtained in usually long processes, essentially Mendelian in nature and probabilistic. They are classified into five types according to the reproductive system of the target species and the genetic structure of the artificial populations: (1) *lines* or *line cultivar*, generally of only one genotype (*pure line*; Fig. 1a); (2) F_1 *hybrid* or *hybrid cultivar*, of only one genotype resulting from a cross between two pure lines, with heterotic effects, represented in Fig. 1e with two loci in heterozygosity (hybrid vigor) and one locus in homozygosity (overdominance), and in Fig. 1g with three loci in heterozygosity; variants of this type of cultivar are named *semi-hybrid cultivars*; (3*) population or population cultivar*, a mixture of genotypes of either autogamous (Fig. 1b), allogamous (Fig. 1d), or apomictic plants. In forage crops, a population cultivar composed of practically isogenic pure lines, similar in phenology and morphological type, is known as a *multiline cultivar*; in allogamous species, this type of cultivar is a population of wide genetic base resulting, in general, from mass selection (Gallais and Bannerot, 1992); (4) *synthetics* or *synthetic cultivar*, similar to population cultivars but only for allogamous species, with paternal control of the origin (*polycross*) (Fig. 1d), or *hybrids* with low vigor depression in F_2; (5) *clones* or *clone cultivar*, composed of only one genotype (Fig. 1e), or two or more genotypes, e.g., clonal hybrids of dioecious species such as asparagus (Fig. 1f) and scions and grafts of fruit trees and ornamentals (Fig. 1h), selected from any structure or obtained by mutagenesis and either macro- or micropropagated (Rimieri, 2017). The subject of the plant protection system -that will be further explained- is a variety (cultivar), that is, a plant grouping within a single botanical taxon of the lowest known rank. Such grouping is defined by the expression of the characteristics resulting from either a given genotype (e.g. one clone, line, or F_1 hybrid) or a combination of genotypes (e.g., a complex hybrid or synthetic variety) (UPOV, 2002).

INTELLECTUAL RIGHTS PROPERTY

The conservation and utilization of plant genetic resources have always required the consideration of diverse factors beyond the biological diversity itself. Among others, the following can be mentioned: genetic transformation technologies, technologies of information and communication (TICs), linked to an increasing world recognition of the value of these resources (Visser and Nap, 2002), and intellectual rights property of both genetic resources and breeding products (Gepts, 2006).

The International Union for the Protection of New Varieties of Plants (UPOV) was created in 1961 to provide and promote an effective system of plant variety protection, with the objective of encouraging the development new plant varieties in its numerous member countries (UPOV, 2020). However, with the advent of plant biotechnologies, patent rights began to affect the access to both genetic resources and commercial varieties. In contrast to the breeder´s rights, patent rights limit the access of third parties to patented genes, with the consequent negative effect on the use of genetic resources. As Eriksson et al. (2020) have discussed, different legal frameworks applicable to the use of the genetic resources have been developed. With the scientific and technical progress in research and breeding achieved in the past few decades, these frameworks have become increasingly complex. Notwithstanding, the Convention on Biological Diversity (CBD, 2020) in its art. 13, recognizes the sovereign rights of the states on the genetic resources located within their frontiers. Based on the principles contained in the CBD and the 2011 Nagoya Protocol plus the decisions of the Parties, international goals on access and benefit-sharing have been established (see Sirakaya, 2019).

UPOV is only concerned with protected plant varieties. However, there is a spectrum of plant genetic resources that does not fall into this category: populations of CWR, landraces, and unprotected plant varieties. These genetic resources are not affected by UPOV or plant breeders´ rights, but they may be regulated by other

treaties or schemes, e.g., the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRF), the previously mentioned CBD, and seed marketing regulations (UPOV, 2016).

FROM GERMPLASM BANK ACCESSIONS TO COMMERCIAL CULTIVARS

The potential utilization of *ex situ* conserved germplasm responds to specific needs of broadening the genetic variability or the gene pool of the breeders´ working collections, particularly in crop species in which the advancements by selection are slow. From this germplasm, new genotypes or gene combinations can be developed for incorporation into breeding programs (Cooper *et al.*, 2002; Rimieri and Wolff, 2010).

One proposal to combine a more efficient conservation of the genetic diversity present in the accessions and to utilize part of the genetic variability of this germplasm in plant breeding is the development of the previously mentioned core collections. The establishment of core collections, which concentrate high genetic diversity in a small number of samples with the avoidance of duplicates, can contribute to the utilization of germplasm in research and pre-breeding, and to the increase of the efficiency of germplasm bank management and inter-bank exchange. Furthermore, with the complement of molecular biology tools, genetic engineering and geographic information systems (GIS), the efficiency and sustainable conservation of plant genetic resources advocated by FAO (1996) would be likely incremented.

GENETIC RESOURCES, POPULATION STRUCTURE AND OBTAINMENT OF COMMERCIAL CULTIVARS

The expansion of the genetic base and pre-breeding shortens the gap between basic germplasm and crop genotypes. However, plant breeders seem to be reluctant to employ plant materials coming directly from germplasm banks because these materials lack, in general, adaptation for their use in breeding. The lack of adaptation is a consequence of the cultivation environment of the crop species and the agronomic management practices, plus the genetic structures of commercial cultivars and the compatibility and interactions of the wild germplasm with the genetic background of the breeder´s elite collection. Notwithstanding, the three elements -genetic resources, population structure, and commercial cultivar development- can be combined. Following, an example is given of forage breeding program to illustrate the close inter-disciplinary relationship between the use of germplasm from working collections and germplasm banks and the application of methods and tools of commercial cultivar development.

In forage crops in general, cultivars are populations, lines and genotypes adapted to the environmental and agronomic conditions of a growing region. They may have their origin in one or more of the following: (a) working collections of research groups involved in population evaluation and selection, (b) foreign cultivars, (c) cultivars adapted to cultivation conditions and animal utilization but no longer available in the market, (d) breeders´ own collections obtained from native and naturalized populations or from old implanted fields, and (e) selected samples -according to previously defined criteria- from national and international collections of botanical gardens, introduction and acclimatization gardens, and germplasm banks. It is, therefore, necessary to remark that the decision on the germplasm to be conserved and its possible utilization in breeding programs has to be based on (1) the initial germplasm, obtained by collection or exchange, with special emphasis in its representativeness of the genetic diversity of the species and the adaptation to the environment and cultivation; and (2) consideration of (a) agronomic and genetic parameters in the original samples and in the subsequent characterization, (b) the predominant mode of reproduction, for propagule multiplication, and (c) the predominant or more representative genetic structures, also for propagule multiplication or the development of core collections, pre-breeding, or commercial cultivar breeding.

TALL FESCUE AS AN EXAMPLE

Tall fescue (*Festuca arundinacea* Schreb.) is a perennial forage grass of temperate climate, of utmost importance and diffusion in Argentina. This species is allogamous, with cleistogamous and chasmogamous flowers, of hexaploid origin and with disomic inheritance. The breeding program carried out at the Pergamino Experimental Station (Exp. Stn.), National Institute of Agropecuarian Technology (INTA), in the Pampas region of central Argentina, is succinctly described in Table 1. It is proposed as an integral model for germplasm management and utilization in general.

The needs of initiating a tall fescue breeding program and of creating a forage germplasm bank in the country stepped from the following:

(1) agroecological conditions: (a) there were no native forage species adapted to cattle grazing, and (b) the forage production of native and naturalized forage species subjected to intensive grazing was low.

(2) technological situation: (a) there were no forage germplasm banks, and (b) the grasslands were subjected to intensive grazing.

In response to this situation:

(1) Temperate forages species with high forage production and adapted to intensive grazing were introduced, characterized and evaluated in agronomic, biological, genetical and animal production studies.

(2) Populations and ecotypes for planting and grazing were selected; cultivars were created, released and disseminated in the region (the area of cultivated pastures was increased with the local cultivar *Pergamino El Palenque MAG*); adaptation and production were evaluated. This germplasm became part of both the working collection of the forage breeding program and the germplasm bank of Pergamino Exp. Stn.

(3) Foreign cultivars were introduced to widen the genetic base of tall fescue in Argentina but, in general, they had poor agroecological and grazing adaptation.

(4) The need of exploring the available global germplasm was established. A forage germplasm bank was created with the adapted local germplasm and the world collection. Collections were evaluated and characterized; protocols were applied to maintain the genetic diversity; core collections were created.

(5) Pre-breeding was initiated for other traits (adaptation to saline soils, forage nutritional value, etc.).

(6) Selected genotypes continued to be incorporated into the germplasm bank.

(7) Animal production was increased in the region.

Summarizing, steps and protocols were followed in tall fescue to integrate objectives of introduction of forage species for intensive grazing, obtainment of populations and ecotypes for germplasm management and utilization in integrated crop-livestock systems, adoption of modern cultivars, pre-breeding for other traits, enhancement of the germplasm bank and increase of animal production. We consider that the Argentinian tall fescue breeding program is a good example of FAO´s proposition (FAO, 1996) on the association and complementation of germplasm banks with breeding programs.

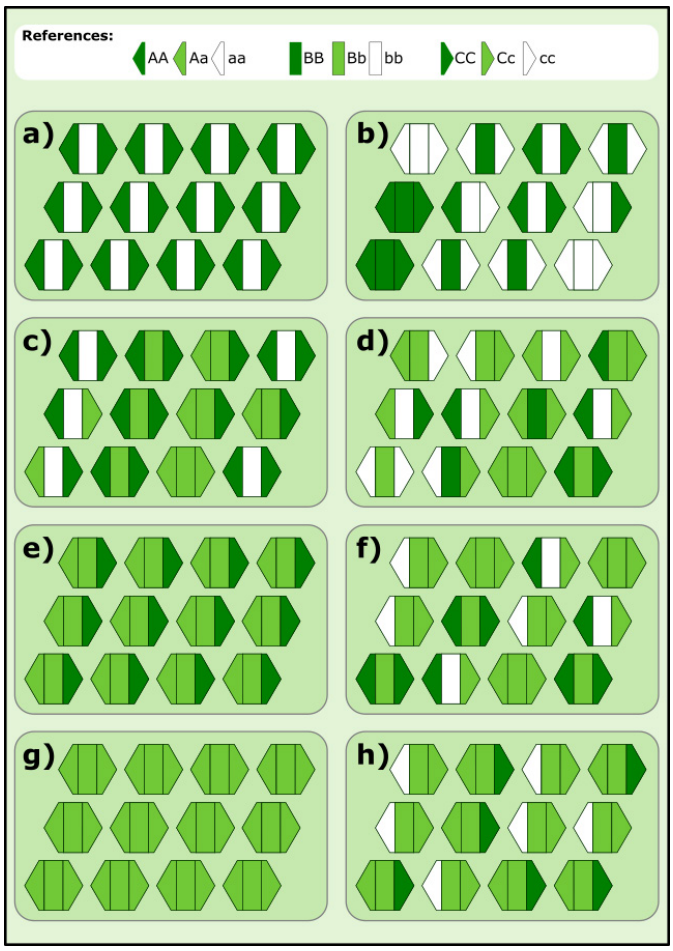


Figure 1. Genetic structure of natural *(NP)* and breeding *(BP)* populations according to modes and types of reproduction. *NP*: (a), (b) and (c) *autogamous diploids and disomic polyploids* (a) homogeneous, with all loci in homozygosity in one combination, (b) heterogeneous, with all loci in homozygosity in various combinations, (c) with a percentage of allogamy; (d) *allogamous diploids*, heterogeneous, with loci in homozygosity and heterozygosity; (e) to (h) *clones*, homogeneous, with either loci in homozygosity and heterozygosity (e) or all loci in heterozygosity (g) for one combination, or heterogeneous with more than one genotype (f) and (h). *BP*: *lines*, homogeneous, with all loci in homozygosity (a); F_1 *hybrids*, homogeneous, with two loci in heterozygosity (hybrid vigor) and one in homozygosity (overdominance) (e) and with all loci in heterozygosity (g); *populations*, heterogeneous, of autogamous (b) and allogamous (d) species; *synthetics*, heterogeneous, with loci in homozygosity and heterozygosity (d); *clones*, homogeneous with loci in homozygosity and heterozygosity in one combination (e) and (g), or heterogeneous, with loci in homozygosity or heterozygosity in more than one combination (f) and (h).

Table 1. Methods and achievements in the Argentinian tall fescue (*Festuca arundinacea* Schreb.) breeding program: from germplasm introduction and collection to obtainment of commercial cultivars.

Year	Institution	Methodology	Achievement	References
1940-	Pergamino Exp. Farm, Ministry of Agriculture (MAG)	Introduction of forage crop populations Studies of adaptation to edafoclimatic environments of the Pampas	Establishment of introduction and acclimatization gardens	Boelcke and Echeverría (1950)
1951-		Evaluation of populations, followed by 1-2 selection cycles or off-type plants roguing	Obtainment of phenotypically uniform populations according to species. Release of the first 38 Argentinian cultivars of 28 forage species, with wide adaptation and diffusion, validated over the next 30 years	Villar and Serrano (1963) Serrano (1985)
***ca*.1961**		Mass selection in introduced germplasm: Alta, Kentucky 31, Goar (records kept on materials and trials)	Release and diffusion of cultivar-population* *Pergamino El Palenque MAG* (being free of *Acremonium coenophialum*, it became a reference cultivar)	Maddaloni and Ferrari (2001)
1980-	Pergamino Exp. Stn. National Institute of Agropecuarian Technology (INTA)	Organization of a Forage Germplasm Bank	Forage Germplasm Bank established. Incorporation of (a) populations introduced from 1947 and on, selected for adaptation and persistence, (b) samples (accessions) from sown and naturalized Pampas populations	

Table 1 (continue). Methods and achievements in the Argentinian tall fescue (*Festuca arundinacea* Schreb.) breeding program: from germplasm introduction and collection to obtainment of commercial cultivars.

Year	Institution	Methodology	Achievement	References
1990-		Selection with methodologies according to demands of a competitive cultivar market. Complementary germplasm studies	Initiation of (a) a breeding program for obtainment of superior synthetic cultivars, (b) characterization of accessions	
1995-		Obtainment of the first Argentinian synthetic cultivar** *Palenque Plus INTA*		Rimieri (1995)
		120 derived genotypes selected and evaluated		
2000-		Incorporation of a large part of the fescue world collection	350 accessions introduced	Rosso *et al.* (2001)
		Morphological and agronomic characterization	36 selected populations	Rimieri andWolff (2010)
		Evaluation for forage nutritive value		
		Molecular (SSR) characterization	Core collection established	Cuyeu *et al.* (2013)

CONCLUSIONS

The premises of this paper are that *ex situ* conservation of the genetic diversity contained in CWR and the utilization of the natural genetic variability in cultivar breeding require the application of reproduction and population genetics concepts in order to choose or develop the appropriate criteria and experimental strategies.

An important fact that needs to be taken into consideration for devising germplasm collection and *ex situ* conservation strategies is that the modes and types of reproduction have different genetic consequences for the following generation. Natural or naturalized populations, even those of autogamous species, can be heterogeneous, and the predominant mode and type of reproduction of a given species can vary according to environmental conditions during the growing cycle.

Biological systems, particularly plant systems, are very complex, thus, assumptions are usually made in an attempt to investigate them. Since discrepancies between "reality" and "assumptions" can be large, the conclusions withdrawn from experimental works need to be adjusted to the plant materials and methods of study to have scientific support. In this regard, there are many reports in the literature on plant and crop physiology of the main food crops (e.g., wheat, maize, sunflower, soybeans) and the "genetic progress" or "genetic gain" that has been achieved in commercial cultivar breeding over the past decades (see Lo Valvo et al. 2018 as an example). However, their potential contribution in crop breeding needs to be ascertained by making focus on the analysis of the genetic structure of populations and the sources of genetic variability available to the breeder (commercial cultivars, land races, CWR). The genetic structure has to be related to the main methods used in those studies and others of related disciplines for the interpretation of the results in the frame of their eventual application in crop management or breeding.

PROPOSAL

We consider that the following information is needed as a basic input to start the analysis of the current germplasm bank protocols at the light of the principles and methods of Genetics:

(a) Genus (or genera) and species of accessions in the germplasm bank

(b) Preponderant mode(s) and type(s) of reproduction

(c) Geographic distribution and sampled areas

(d) Sampling strategies

(e) Passport data of collections in general, from the oldest to the newest

(f) *Ex situ* regeneration/multiplication protocols

(g) Characterization type (morphological, genetic, molecular, agronomic), if any.

This information would allow the evaluation in the *ex situ* collections of:

(a) Representativeness of the collections, geographical and environmental (at macro- and micro- levels).

(b) Adequacy of strategies and protocols for collection and regeneration or multiplication of accessions to the principles of population genetics: population reproductive size (*N*= actual number of plants in the population, and *Ne*= effective number of plants, which contribute alleles to the next generation), population genetic structure, gene (allele) frequencies, processes that can alter gene frequencies.

(c) Representativeness of the natural genetic diversity in the collections.

(d) Necessity of carrying out new collections in the already sampled areas or in as yet unexplored ones.

Furthermore, to ascertain if wild germplasm conservation and commercial breeding converge at some point, the following questions should be addressed:

(1) In pre-breeding:

(a) Is pre-breeding an objective of germplasm banks?

(b) What is considered to be more important in the germplasm bank, the representativeness of the natural genetic diversity in the accessions or the likely immediate use of the conserved germplasm?

(2) In breeding:

(a) Is it considered that the collections can be directly used in breeding programs or that pre-breeding is required as a first step?

(b) Is it known which is the genetic background of populations or genotypes adapted to cultivation that has to be maintained or recovered after manipulations to incorporate new germplasm in the cultivated pool (e. g. hybridizations, backcrosses or other techniques or methods)?

As a first step in this direction, we will coordinate a workshop which is part of the program of ALAG 2021 (XVIII Latin American Congress of Genetics; alagenet.org/alag2021/en/scientific-program/#talleres). In advance, the invited researchers and curators will provide in written response to the formulated questions. The discussion and analysis of the responses will be carried out at the light of the principles and methods of Genetics during the event. The expected final product is a document on the current managing practices in germplasm banks of seven participating countries; if appropriate, the document will also contain propositions for the eventual modifications of protocols.

Finally, as Maxted and Kell (2009) have pointed out, there is a need for CWR characterization and evaluation, development of genomic databases of known useful genes from these sources, and improvement of gene transfer techniques from wild to cultivated species, among others. Notwithstanding, we consider that a previous basic requirement for successful conservation and utilization of the natural genetic diversity and genetic variability is the application of strategies and protocols based on the principles and methods of population genetics, modes of reproduction and genetic structures of CWR populations.

ACKNOWLEDGEMENTS

To former students and colleagues that, throughout the years, contributed to the generation of knowledges and the discussion of ideas on germplasm conservation and use. To the public institutions that provided financial support for the authors´s lines of investigations and graduate students´ scholarhips, mainly Instituto Nacional de Tecnología Agropecuaria (INTA); Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET); Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata (UNMdP); and Facultad de Ciencias Agrarias, Universidad Nacional de Rosario (UNR). To Dr. J. Federico Maune, who kindly diagrammed Figure 1.

REFERENCES

Allard R.W. (1999) Principles of Plant Breeding (2nd ed.). John Wiley & Sons, Hoboken, New Jersey.

Al-Khayri J.A., Jain S.M., Johnson D.V. (2015) Advances in Plant Breeding Strategies: Breeding, Biotechnology and Molecular Tools, Vol. 1. Springer International Publishing, Switzerland.

Asker S. (1980) Gametophytic apomixis: elements and genetic regulation. Hereditas 9: 277-293. Doi: 10.1111/j.1601-5223.1980.tb01367.x.

Bedonni M.C., Camadro E.L. (2009) Morphological and molecular evidence of natural interspecific hybridization in the diploid potato *Solanum kurtzianum* from Argentina.

Botany 87: 78-87. Doi: 10.1139/B08-116.

Boelcke O., Echeverría I. (1950) Valor y comportamiento de mezclas forrajeras comerciales para praderas permanentes en la región de Pergamino. Publicación No. 26. INTA Estación Experimental Pergamino, Argentina.

Burnham C.R. (1980) Discussions in Cytogenetics (6th Ed.). University of Minnesota, St. Paul.

Cadima Fuentes X., van Treuren R., Hoekstra R., van den Berg R.G., Sosef M.S.M. (2017) Genetic diversity of Bolivian wild potato germplasm: changes during *ex situ* conservation management and comparisons with resampled *in situ* populations. Genet. Resour. Crop Evol. 64: 331-344. Doi: 10.1007/s10722-015-0357-9.

Camadro E.L. (2012) Relevance of the genetic structure of natural populations, and sampling and classification approaches for conservation and use of wild crop relatives: potato as an example. Botany 90: 1065-1072. Doi: 10.1139/b2012-090.

Cardi T. (2016) Cisgenesis and genome editing: combining efforts for a smarter use of genetic resources in crop breeding. Plant Breeding 135: 139-147. Doi: 10.1111/pbr.12345

CBD (2020) Convention on Biological Diversity. https://www.cbd.int (accessed November 2020)

Cooper HD, Spillane C, Hodgkin T (2000) Broadening the genetic base of crop production. CABI, Wallingford; FAO and IPGRI, Rome. http://www. bioversityinternational. org/Publications/pubfile.asp?ID_PUB=618. (accessed June 2020).

Cuyeu A.R., Rosso B., Pagano E., Soto G., Fox R., Ayub N.D. (2013) Genetic diversity in a world germplasm collection of tall fescue. Genet. Mol. Biol. 36: 237-242. Doi: 10.1590/S1415-47572013005000021.

Dumas C., Mogensen H.L. (1993) Maize as a model system for experimental embryogenesis in flowering plants. Plant Cell 5: 1337-1348. Doi: 10.1105/tpc.5.10.1337.

Ellstrand N.C., Roose M.L. (1987) Patterns of genotypic diversity in clonal plant species. Am. J. Bot. 74: 123-131. Doi: 10.1002/j.1537-2197.1987.tb08586.x.

Engels J.M.M., Visser, L. (2006) Guide to effective management of germplasm collections. IPGRI Handbooks for Germplasm Banks 6. IPGRI, Rome.

Erazzú L.E., Camadro E.L., Clausen A.M. (2009) Persistence over time, overlapping distribution and molecular indications of interspecific hybridization in wild potato populations of Northwest Argentina. Euphytica 168: 249–262. Doi: 10.1007/s10681-009-9938-z.

Eriksson D., Ortiz R., Visser R.G.F., Vives-Vallés J.A., Prieto H. (2020) Editorial: Leeway to operate with plant genetic resources. Front. Plant Sci 11: 911. Doi: 10.3389/fpls.2020.00911.

FAO (1996) The State of the World's Plant Genetic Resources for Food and Agriculture. FAO, Rome.

FAO (2017) Voluntary guidelines for the conservation and sustainable use of crop wild relatives and wild food plants. FAO, Rome. http://www.fao.org/3/a-i7788e.pdf. (accessed June 2020).

Frankel O.H. (1984) Genetic perspectives of germplasm conservation. In: Arber W., Llimensee K., Peacock W.J., Starlinger P. (Eds.) Genetic Manipulation: Impact on Man and Society. Cambridge University Press, Cambridge.

Frankel R., Galun E. (1977) Pollination Mechanisms, Reproduction and Plant Breeding. Springer-Verlag, Berlin, Heidelberg, New York.

Gallais A. (1990) Théorie de la sélection en amélioration des plantes. Masson, Paris.

Gallais A., Bannerot H. (1992) Améloration des espéces végétales cultivées. Objetives et critéers de sélection. INRA, Paris.

Gepts P. (2006) Plant Genetic Resources Conservation and Utilization: The Accomplishments and Future of a Societal Insurance Policy. Crop Sci. 46: 2278-2292. Doi: 10.2135/cropsci2006.03.0169gas.

Grant V. (1981) Plant Speciation. Columbia University Press, New York and London.

Hammer K., Teklu Y. (2008) Plant Genetic Resources: Selected Issues from Genetic Erosion to Genetic Engineering. J. Agric. Rural Dev. Trop. Subtrop. 109: 15–50.

Johnson R.C., Hodgkin T. (1999) Core collections for today and tomorrow. IPGRI, Rome. http://www.bioversityinternational.org/publications/Web%5Fversion/43/. (accessed June 2020).

Knox R.B. (1967) Apomixis: Seasonal and Population Differences in a Grass. Science 157: 325-326. Doi: 10.1126/science.157.3786.325.

Langer R.H.M., Wilson D. (1965). Environmental control of cleistogamy in prairie grass (*Bromus unioloides* H.B.K.). New Phytol. 64: 80-85. Doi: 10.1111/j.1469-8137.1965.tb05377.

Leofanti G.A., Camadro E.L., Erazzú, L.E. (2019) Variation over time in morphological phenotypes and reproductive behavior in a natural wild potato population from Tucumán, Argentina. Genet. Resour. Crop Evol. 67: 139-161. Doi: 10.1007/s10722-019-00858-7.

Leofanti G.A., Camadro E.L., Echeverría M.M., Alonso S.I. (2013) Anormalidades meióticas en especies nativas del género *Bromus* (Secc. Ceratochloa) de la Argentina. BAG. J. Basic Appl. Genet. 23 (Suppl. 1): 78.

Lo Valvo P.J., Miralles D.J., Serrago R.A. (2018) Genetic Progress in Argentine bread wheat varieties release between 1918 and 2011: Changes in physiological and numerical yield components. Field Crop Res. 21: 314-321. Doi: 10.1016/bs.agron.2021.02.005.

Maddaloni J., Ferrari L. (2001) Forrajeras y pasturas del ecosistema templado húmedo de la Argentina. Facultad de Cs. Agrarias, Universidad Nacional de Lomas de Zamora, Lomas de Zamora.

Maxted N., Kell S. (2009) Establishment of a Global Network for the Ex Situ Conservation of Crop Wild Relatives: Status and Needs. FAO Commission on Genetic Resources for Food and Agriculture. FAO, Rome.

Maxted N., Ford-Lloyd B.V., Hawkes J.G. (1997) Complementary Conservation Strategies. In: Maxted N., Ford-Lloyd B.V., Hawkes J.G. (Eds.) Plant Genetic Conservation: The In Situ Approach. Chapman and Hall, London, pp. 15-40.

Maxted N., Castañeda Álvarez N.P., Vincent H.A., Magos Brehm J. (2012) Gap analysis: a tool for genetic conservation. In: Guarino L., Ramanatha Rao V., Goldberg E. (Eds.) Collecting Plant Genetic Diversity: Technical Guidelines, 2011 update. Biodiversity International, Rome, pp. 1-17.

Maxted N., Amri A., Castañeda-Alvarez N.P., Dias S., Dullo M.E., Fielder H., Ford-Lloyd B. V., Iriondo J.M.. Magos Brehm J., Nilsen L.B., Thormann I., Vincent H., Kell S.P. (2016) Joining up the dots: a systematic perspective of crop wild relative conservation and use.

In: Maxted N. et al. (Eds.) Enhancing crop gene pool use: capturing wild relative and landrace diversity for crop improvement. Cambridge University Press, Cambridge, pp. 87-124.

Poulsen Hornum A., Camadro E.L. (2021) Expression of internal reproductive barriers in a germplasm bank accession of the wild potato *Solanum chacoense* Bitter in three *ex situ* regeneration cycles. Genet. Resour. Crop Evol. 68:915-938. Doi: 10.1007/s10722-020-01034-y.

Quarin C.L. (1986) Seasonal changes in the incidence of apomixis of diploid, triploid, and tetraploid *Paspalum cromyorrhizon*. Euphytica 35: 515-522. Doi: 10.1007/BF00021860.

Rebozzio R., Sartor M., Quarin C.L., Espinoza F. (2011) Residual sexuality and its seasonal variation in natural apomictic *Paspalum notatum* accessions. Biol. Plantarum 55: 391-395. Doi: 10.1007/s10535-011-0062-2.

Rieger R., Michaelis A., Green M.M. (1976) A Glossary of Genetics, Classical and Molecular (5th ed.). Springer-Verlag, Berlin, Heidelberg, New York.

Rimieri P. (1995) *Palenque Plus INTA*. Expte. INASE No. 3780 (05/29/95) https://inta.gob.ar/variedades/palenque-plus-inta. (accessed June 2020).

Rimieri P. (2017) La diversidad genética y la variabilidad genética: dos conceptos diferentes asociados al germoplasma y al mejoramiento genético vegetal. BAG. J. Basic Appl. Genet. 28 (2): 7-13.

Rimieri P., Wolff R. (2010) La genética y el estado actual de la obtención y adopción de cultivares forrajeros en Argentina. BAG. J. Basic Appl. Genet. 1 (2): 1-7.

Rosso B., Pagano E., Rimieri P. (2001) Evaluation and utilization of the tall fescue germplasm collection at Pergamino INTA, Argentina. Proceedings of the XIX International Grassland Congress. Sao Paulo, Brazil, p 504. https://www.internationalgrasslands.org/publications. (accessed July 2020).

Schoen D.J., Brown A.H.D. (1991) Intraspecific variation in population gene diversity and effective population size correlates with the mating system in plants. P. Natl. Acad. Sci. USA 88: 4494–4497. Doi: 10.1073/pnas.88.10.4494.

Serrano H. (1985) Premio Academia Nacional de Agronomía y Veterinaria. http://bibliotecadigital.bolsadecereales.com.ar/greenstone/collect/pubper/index/assoc/HASH1b0b/d2948809.dir/Ejercicio%201986.pdf. (accessed July 2020)

Sirakaya A. (2019) Balanced Options for Access and Benefit-Sharing: Stakeholder Insights on Provider Country Legislation. Front. Plant Sci. 10: 1175. Doi: 10.3389/fpls.2019.01175.

UPOV (2002) Council Nineteenth Extraordinary Session Geneva, April 19, 2002. https://www.upov.int/edocs/infdocs/en/0_c_extr_19_2_rev.pdf. (accessed August 2020).

UPOV (2016) Symposium on possible interrelations between the International Treaty on Plant Genetic Resources for Food and Agriculture (ITPGRFA) and the International Convention for the Protection of New Varieties of Plants (UPOV Convention). https://www.upov.int/edocs/mdocs/upov/en/upov_itpgrfa_sym_ge_16/upov_itpgrfa_sym_ge_16_2_proceedings.pdf. (accessed August 2020).

Villar A.D., Serrano H. (1963) Praderas permanentes para la región pampeana húmeda. Boletín Técnico No. 2, INTA Rafaela Experimental Station, Rafaela.

Visser B., Nap J.P. (2002) Biotechnology and agrobiodiversity in Biological and Medical Sciences. Encyclopedia of Life Support Systems (EOLSS). Eolss Publishers, Oxford. http://www.eolss.net

Williams W.M., Stewart A.V., Williamson M.L. (2011) *Bromus*. In: Kole C. (Ed.) Wild Crop Relatives: Genomic and Breeding Resources, Millets and Grasses. Springer Verlag, Berlin, pp. 15-30.

Wolff R., Abbott L., Pistorale S. (1996) Reproductive behavior of *Bromus catharticus* Vahl. (Cebadilla criolla) in natural and cultivated populations. J Genet Breed 50: 121-128.

BAG | Journal of Basic & Applied Genetics

Vol XXXII (1): 25-33; July 2021

Journal of the Argentine Society of Genetics

MULTI-TRAIT MODELS FOR GENOMIC REGIONS ASSOCIATED WITH MAL DE RÍO CUARTO AND BACTERIAL DISEASE IN MAIZE

MODELOS MULTIVARIADOS EN LA BÚSQUEDA DE REGIONES GENÓMICAS PARA RESISTENCIA A MAL DE RÍO CUARTO Y BACTERIOSIS EN MAÍZ

Ruiz M.[1,2], Rossi E.A.[1,2], Bonamico N.C.[1,2], Balzarini M.G.[3,4]

[1] INIAB-Instituto de Investigaciones Agrobiotecnológicas (CONICET-UNRC).

[2] Facultad de Agronomía y Veterinaria, Universidad Nacional de Río Cuarto.

[3] CONICET-Consejo Nacional de Investigaciones Científicas y Técnicas.

[4] Facultad de Ciencias Agropecuarias, Universidad Nacional de Córdoba.

Corresponding author:
Marcos Ruiz
mruiz@ayv.unrc.edu.ar

ORCID 0000-0001-5624-0183

Cite this article as:
Ruiz M., Rossi E.A., Bonamico N.C., Balzarini M.G. 2021. MULTI-TRAIT MODELS FOR GENOMIC REGIONS ASSOCIATED WITH MAL DE RÍO CUARTO AND BACTERIAL DISEASE IN MAIZE. BAG. Journal of Basic and Applied Genetics XXXII (1): 25-33.

Received: 06/25/2020
Revised version received: 10/30/2020
Accepted: 11/10/2020
General Editor: Elsa Camadro
DOI: 10.35407/bag.2021.32.01.03
ISSN online version: 1852-6233

ABSTRACT

Maize (*Zea Mays* L.) production has been greatly benefited from the improvement of inbred lines in regard to the resistance to diseases. However, the absence of resistant genotypes to bacteriosis is remarkable. The aim of the study was to identify genomic regions for resistance to Mal de Río Cuarto (MRC) and to bacterial disease (BD) in a diverse maize germplasm evaluated in the Argentinian region where MRC virus is endemic. A maize diverse population was assessed for both diseases during the 2019-2020 crop season. Incidence and severity of MRC and BD were estimated for each line and a genome wide association study (GWAS) was conducted with 78,376 SNP markers. A multi-trait mixed linear model was used for simultaneous evaluation of resistance to MRC and BD in the scored lines. The germplasm showed high genetic variability for both MRC and BD resistance. No significant genetic correlation was observed between the response to both diseases. Promising genomic regions for resistance to MRC and BD were identified and will be confirmed in further trials.

Key words: maize disease; genome wide association study; SNP; multi-trait model.

RESUMEN

La producción de maíz (*Zea Mays* L.) ha sido ampliamente beneficiada con la mejora de líneas endocriadas respecto a la resistencia a enfermedades causadas por virus y hongos. Sin embargo, es notable la ausencia de genotipos resistentes a bacteriosis. El objetivo del presente estudio fue identificar regiones genómicas para la mejora de resistencia a Mal de Río Cuarto (MRC) y a bacteriosis (BD) en un germoplasma diverso de maíz. Se evaluó, para ambas enfermedades, una población diversa de líneas de maíz en el ciclo de cultivo 2019-2020 en la región argentina donde la virosis MRC es endémica. Se estimó incidencia y severidad de MRC y BD en cada línea y se realizó un estudio de mapeo por asociación (GWAS) con 78.376 marcadores SNPs. Un modelo multicarácter se utilizó para evaluar simultáneamente la resistencia a MRC y BD en las líneas evaluadas. El germoplasma evidenció alta variabilidad genética tanto para la mejora de la resistencia a MRC como a BD, pero no se observó correlación genética significativa entre la respuesta a ambas enfermedades. Se identificaron regiones genómicas promisorias para resistencia a MRC y a BD, que serán confirmadas en evaluaciones en nuevos ambientes.

Palabras clave: enfermedad en maíz; mapeo por asociación; SNP; modelo multivariado.

Available online at
www.sag.org.ar/jbag

INTRODUCCIÓN

El maíz es un importante cultivo a nivel mundial con una producción aproximada de 1.148 millones de toneladas. Argentina es el cuarto productor con un total de 43,5 millones de toneladas cosechadas en el ciclo agrícola 2018/19. Entre otros factores, su producción es afectada por la presencia de enfermedades que pueden amenazar el cultivo y consecuentemente la seguridad alimentaria y la sustentabilidad agrícola (Nelson *et al.*, 2018). Las enfermedades más comunes en maíz son causadas por virus y hongos (Agrios, 2005) aunque entre las patologías emergentes, se encuentran las producidas por bacterias. Las enfermedades producidas por bacterias patógenas en maíz han incrementado su prevalencia en Argentina (Plazas, 2018), debido posiblemente a la masiva adopción de la siembra directa. Es notable la ausencia de genotipos resistentes a este tipo de enfermedades (Gurr y Rushton, 2005). La base genética de la respuesta a la infección por bacterias o bacteriosis (BD) ha sido significativamente menos investigada que aquélla relacionada con enfermedades fúngicas y virales en maíz (Rossi *et al.*, 2019).

Entre las enfermedades virales, el Mal de Río Cuarto (MRC) juega un rol prevalente para el cultivo de maíz en Argentina ya que ha causado severas pérdidas de rendimiento con diferentes valores de incidencia y severidad a través de los años (Giménez Pecci *et al.*, 2012). El agente causal es el *Mal de Río Cuarto virus* (MRCV), el cual se clasifica como un miembro del género *Fijivirus*, familia *Reoviridae* (King *et al.*, 2012) y se transmite de manera persistente propagativa por insectos vectores, principalmente por la chicharrita *Delphacodes kuscheli* Fennah (Ornaghi *et al.*, 1993).

La resistencia genética es el método más eficiente y efectivo para el control de MRC y por ello se han realizado numerosas evaluaciones de material genético en la zona donde la enfermedad es endémica (Di Renzo *et al.*, 2004; Bonamico *et al.*, 2012; Rossi *et al.*, 2015). Numerosos QTL (*Quantitative trait loci*) para resistencia a MRC se han identificado a través del fenotipado de poblaciones biparentales generadas localmente. El mapeo por asociación (GWAS, *genome wide association study*) también es usado para la identificación de variantes alélicas específicas que incrementan la resistencia a enfermedades (Zila *et al.*, 2014). Para GWAS, es esencial contar con poblaciones compuestas por individuos genotípica y fenotípicamente diversos, que presenten una alta densidad de polimorfismos en la secuencia de ADN (Yan *et al.*, 2011).

Germoplasmas diversos de especies alógamas como el maíz demandan alta densidad de marcadores. Esto se debe a que el desequilibrio de ligamiento puede disminuir entre dos sitios polimórficos a una distancia corta (unos pocos miles pares de bases) a causa de la alta frecuencia de recombinación (Remington *et al.*, 2001). Las líneas de maíz del Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT), desarrolladas durante los últimos 25 años, se han convertido en la principal fuente pública de germoplasma diverso de maíz (Chen *et al.*, 2016). La genotipificación densa de este germoplasma de alta diversidad genómica aporta información para la implementación de GWAS para identificar *loci* de resistencia a enfermedades presentes en la región maicera de Argentina.

Si bien es común que los mejoradores evalúen múltiples caracteres en sus esquemas de selección (Malosetti *et al.*, 2008) en programas de mejoramiento genético, la evaluación de resistencia a MRC no se ha realizado simultáneamente con BD. El enfoque multicarácter incrementa el poder de detección de aquellos QTL que afectan a más de un carácter simultáneamente (Knott y Haley, 2000). El ajuste de un modelo que permita caracterizar genotipos considerando más de un carácter y sus potenciales correlaciones (Covarrubias Pazaran, 2016) representa una forma novedosa de abordar la resistencia a enfermedades. El objetivo del presente estudio fue identificar regiones genómicas de maíz promisorias para la mejora de la resistencia a la enfermedad MRC y a BD en un germoplasma diverso de maíz.

MATERIALES Y MÉTODOS

Material vegetal y ensayo de campo

Se sembró una población diversa de 185 líneas de maíz del Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT) para determinar su resistencia a Mal de Río Cuarto (MRC) y a bacteriosis (BD) durante el ciclo de cultivo 2019-2020 en Río Cuarto, Córdoba, Argentina (64° 20′ W, 33° 8′ S). Se utilizó un diseño parcialmente repetido (p-rep) (Cullis *et al.*, 2006). El diseño implicó el uso de tres repeticiones en 50 líneas (de modo que p= 27 %) y parcelas individuales en las 135 líneas restantes. Cada línea se sembró en una parcela que consistió de un surco de 2,5 m de longitud y 0,52 m de ancho. Se sembró a doble densidad y se ralearon plantas, tres semanas posteriores a la emergencia para obtener 10 plantas por parcela. La infección de ambas enfermedades ocurrió de manera natural. Líneas experimentales susceptibles se evaluaron conjuntamente en el ensayo para verificar la ocurrencia de MRC.

Ambas enfermedades se evaluaron mediante la observación de síntomas en la etapa de floración (95-100 días post siembra). La incidencia de MRC y BD se estimó como la proporción de plantas que presentaron síntomas sobre el total de plantas de cada parcela (INC-MRC y INC-BD, respectivamente). La severidad en ambas enfermedades (SEV-MRC y SEV-BD) se evaluó utilizando las escalas propuestas por Ornaghi *et al.* (1999) y Schuelter *et al.* (2003) para MRC y BD, respectivamente.

Para SEV-MRC, cada planta se clasificó por el grado de severidad: 0= sin síntomas; 1= presencia de enaciones; 2= presencia de enaciones + acortamiento de entrenudos + láminas foliares atrofiadas en el tercio superior; 3= máximo desarrollo de la enfermedad con enaciones + acortamiento de entrenudos + láminas foliares atrofiadas en el tercio superior + espigas pequeñas, múltiples y sin granos. Para SEV-BD, las plantas de cada parcela se evaluaron visualmente por la lesión foliar, siendo 1= sin lesiones; 2= lesiones dispersas; 3= hasta el 50 % de las hojas con lesiones, con lesiones graves en el 25 % de las hojas inferiores; 4= hasta el 75 % de las hojas con lesiones, con lesiones graves en el 50 % de las hojas inferiores; 5= 100 % de las hojas con lesiones, con lesiones graves en el 75 % de las hojas inferiores; 6= planta muerta. Para el análisis estadístico se usó la severidad promedio por parcela.

Datos genómicos

La caracterización genotípica realizada por Wu *et al.* (2016) con marcadores moleculares del tipo SNP utilizada en este trabajo se encuentra disponible en http://data.cimmyt.org/dvn. De un total de 362.008 SNPs se seleccionaron 78.376 marcadores distribuidos en los 10 cromosomas de maíz. La selección de marcadores se basó en la calidad, y en un primer paso, se eliminaron marcadores con una frecuencia alélica menor a 0,05 o con errores de secuenciación. Posteriormente, se eliminaron los marcadores con una tasa de datos faltantes superior al 35 %. La base de datos utilizada en este trabajo está disponible en https://github.com/PlantbreedingUNRC/GWAS-MRC-BD.

Análisis estadístico

Los datos fenotípicos se analizaron utilizando un modelo lineal mixto (MLM) multicarácter (Maier *et al.*, 2015). La base de datos utilizada en el modelo consta de 185 líneas de maíz evaluadas mediante cuatro caracteres (INC-MRC, SEV-MRC, INC-BD y SEV-BD). Se usó un modelo lineal mixto multivariado para obtener el predictor lineal del efecto aleatorio de genotipo, es decir un BLUP multivariado de efectos de genotipo que considera los cuatro caracteres evaluados simultáneamente. La observación y_{ijk} para el carácter k es escrito como una función lineal del efecto del bloque desde donde es observada y del efecto de genotipo,

$$y_{ijk} = \mu + \beta_i + G_j + \varepsilon_{ijk} \text{ para el carácter } k$$

donde y_{ijk} es la observación fenotípica para el carácter k=INC-MRC, SEV-MRC, INC-BD, SEV-BD, μ es la media general, β_i es el efecto aleatorio deli-ésimobloque $\beta_i \sim N(0, \sigma^2_b)$, G_j es el efecto aleatorio del j-ésimogenotipo y ε_{ijk} es el componente de error aleatorio independiente que se supone normal con media y varianza constante.

El MLM multivariado es un modelo para el vector de observaciones **Y** donde se consideran las observaciones de los k caracteres y que estima no sólo varianzas de los caracteres sino también covarianzas entre ellos. En este trabajo se ajustó utilizando la función "mmer" del paquete "sommer" (CovarrubiasPazaran, 2016), software R (R Core Team, 2016).

Los componentes de varianza obtenidos del MLM multivariado fueron usados para estimar la heredabilidad del carácter k, tal como lo propusieron Hallauer y Miranda (1988).

$$H^2 = \frac{\sigma^2_{gk}}{\sigma^2_{gk} + \left(\sigma^2_{ek}/p\right)}$$

donde σ^2_{gk} es la varianza genotípica para el carácter k-ésimo, σ^2_{ek} es la varianza residual asociada a las observaciones de ese carácter, y p es una media ponderada del número de repeticiones (Holland *et al.*, 2003). Las varianzas genéticas estimadas para cada carácter y las covarianzas genética entre pares de caracteres, obtenidas del ajuste del MLM multivariado, permitieron estimar la correlación genética (r_g) entre los caracteres l y m:

$$r_{g(l,m)} = \frac{\sigma^2_{g(l,m)}}{\sqrt{\sigma^2_{gl}\,\sigma^2_{gm}}}$$

donde $\sigma^2_{g(l,m)}$ es la covarianza genética entre los caracteres l y m, y los elementos del denominador corresponden a las varianzas genéticas de ambos caracteres.

Asociación entre variación fenotípica y variación genotípica

El software Tassel 5.2.60 (Bradbury *et al.*, 2007) se utilizó para realizar el análisis de asociación entre variación fenotípica y variación genotípica. El GWAS se llevó a cabo con 78.376 SNPs. La incidencia y la severidad de MRC y BD fueron usadas como variables dependientes.

Para llevar a cabo la asociación se ajustaron seis modelos. Los modelos fueron: 1) el modelo *Naive*, que realiza la asociación de la información genotípica y fenotípica sin tener en cuenta la posible existencia de estructura genética en la población en estudio; 2) el modelo Q, que utiliza la matriz de estructura poblacional definida con el programa Structure (Pritchard, 2000) considerando la correlación genética entre líneas; 3) el modelo PCA, el cual modela la estructura poblacional subyacente en la población de líneas de maíz, mediante la incorporación de cinco componentes principales (Price *et al.*, 2006). También se ajustaron modelos

que incorporaron la matriz de *Kinship* para modelar la relación genética entre dos líneas de la población en estudio, 4) el modelo K (Parisseaux y Bernardo, 2004); 5) el modelo Q + K (Yu *et al.*, 2006); y 6) el modelo PCA + K (Zhao *et al.*, 2007). A partir del gráfico de quantiles-quantiles (*Q-Q plot*) se seleccionaron los modelos apropiados para cada carácter. Este gráfico compara los valores observados del -log10 (*valor-p*) para cada marcador con los valores esperados del -log10 (*valor-p*) bajo la hipótesis nula de no asociación entre marcadores moleculares y los caracteres en estudio. El procedimiento de Li y Ji (2005) se utilizó para realizar correcciones de los valores *p* por multiplicidad. Los marcadores SNPs que superaron el umbral de -log10 (*valor-p*) >4 ($p<0,0001$), se consideraron asociados de manera significativa con el carácter. Los gráficos *Q-Q plot* y *Manhattan plot* se realizaron con el paquete "qqman" (Turner, 2018) del software R (R Core Team, 2016) utilizando las salidas de GWAS del software Tassel.

RESULTADOS

Para ambas enfermedades se observaron tanto valores mínimos como valores máximos de las respectivas escalas de evaluación de síntomas. El valor medio de INC-MRC fue 43% y de SEV-MRC 0,88, mientras que para INC-BD y SEV-BD fueron 83% y 2,5, respectivamente. El valor estimado de heredabilidad varió entre 0,35 (INC-BD) y 0,7 (SEV-BD) (Tabla 1).

Tabla 1. Medidas resumen y parámetros genéticos estimados para incidencia (INC) y severidad (SEV) de Mal de Río Cuarto (MRC) y de bacteriosis (BD) en una población diversa de 185 líneas de maíz evaluadas en Río Cuarto, Córdoba, Argentina, durante el ciclo de cultivo 2019/2020.

Carácter	Media ± EE	Mínimo	Máximo	σ_g^2	H^2
INC-MRC	43,25±2,13	0	100	367,16	0,47
SEV-MRC	0,88±0,05	0	3	0,25	0,55
INC-BD	82,96±1,87	0	100	216,09	0,35
SEV-BD	2,53±0,07	1	6	0,59	0,69

EE= error experimental; σ_g^2 = varianza genotípica; H^2= Heredabilidad

El MLM multivariado reveló que la correlación genética entre caracteres es positiva y significativa cuando se trata de INC y SEV de una misma enfermedad, mientras que la correlación entre caracteres de distinta enfermedad no resultó estadísticamente significativa (Tabla 2).

Tabla 2. Correlación entre los caracteres incidencia (INC) y severidad (SEV) de Mal de Río Cuarto (MRC) y de bacteriosis (BD) en una población diversa de 185 líneas de maíz evaluadas en Río Cuarto, Córdoba, Argentina, en el ciclo de cultivo 2019/2020.

Carácter	INC-MRC	SEV-MRC	INC-BD	SEV-BD
INC-MRC	1	0,89*	-0,23	0,03
SEV-MRC	-	1	-0,25	0,01
INC-BD	-	-	1	0,94*
SEV-BD	-	-	-	1

*correlación significativa ($p < 0,0001$)

INC-MRC: Incidencia de Mal de Río Cuarto, SEV-MRC: Severidad de Mal de Río Cuarto, INC-BD: Incidencia de bacteriosis, SEV-BD: severidad de bacteriosis.

A partir de la frecuencia alélica de los loci de cada línea, se clasificaron con el software Structure los genotipos del panel en tres grupos. El primer grupo estuvo compuesto por cuatro líneas, el segundo por 108 líneas y el tercer grupo por 73 líneas. Consecuentemente, la estructura genética que determina el agrupamiento de líneas fue contemplada en los modelos de asociación entre la variación fenotípica y la genotípica ajustadas carácter por carácter. El modelo Q (modelo que considera la estructura de grupos sugerida por el software Structure) fue el modelo de mejor ajuste para INC-MRC y SEV-MRC. Sin embargo, para INC-BD y SEV-BD la estructura de grupos fue menos discreta y quedó mejor representada por el modelo de asociación PCA + K, modelo que consideró cinco componentes del análisis de componentes principales de los datos moleculares como covariable y que además incorpora la correlación genética entre pares de líneas mediante la matriz de parentesco K (Figura 1).

El mapeo asociativo utilizando los 78.376 SNPs para GWAS, se realizó con el modelo más apropiado para cada carácter (Figura 1). Cinco regiones fueron declaradas como estadísticamente significativas tanto en el carácter INC-MRC como para SEV-MRC. Se identificaron otras 15 regiones como aportantes de alelos promisorios para INC-BD y SEV-BD. No se encontraron regiones genómicas que fuesen promisorias para mejorar la resistencia de ambas enfermedades simultáneamente. La proporción de variación fenotípica explicada por cada región cromosómica identificada como de alto potencial para la selección por resistencia a MRC osciló entre 0,11 y 0,19; valores similares fueron observados para BD (Tabla 3 y 4).

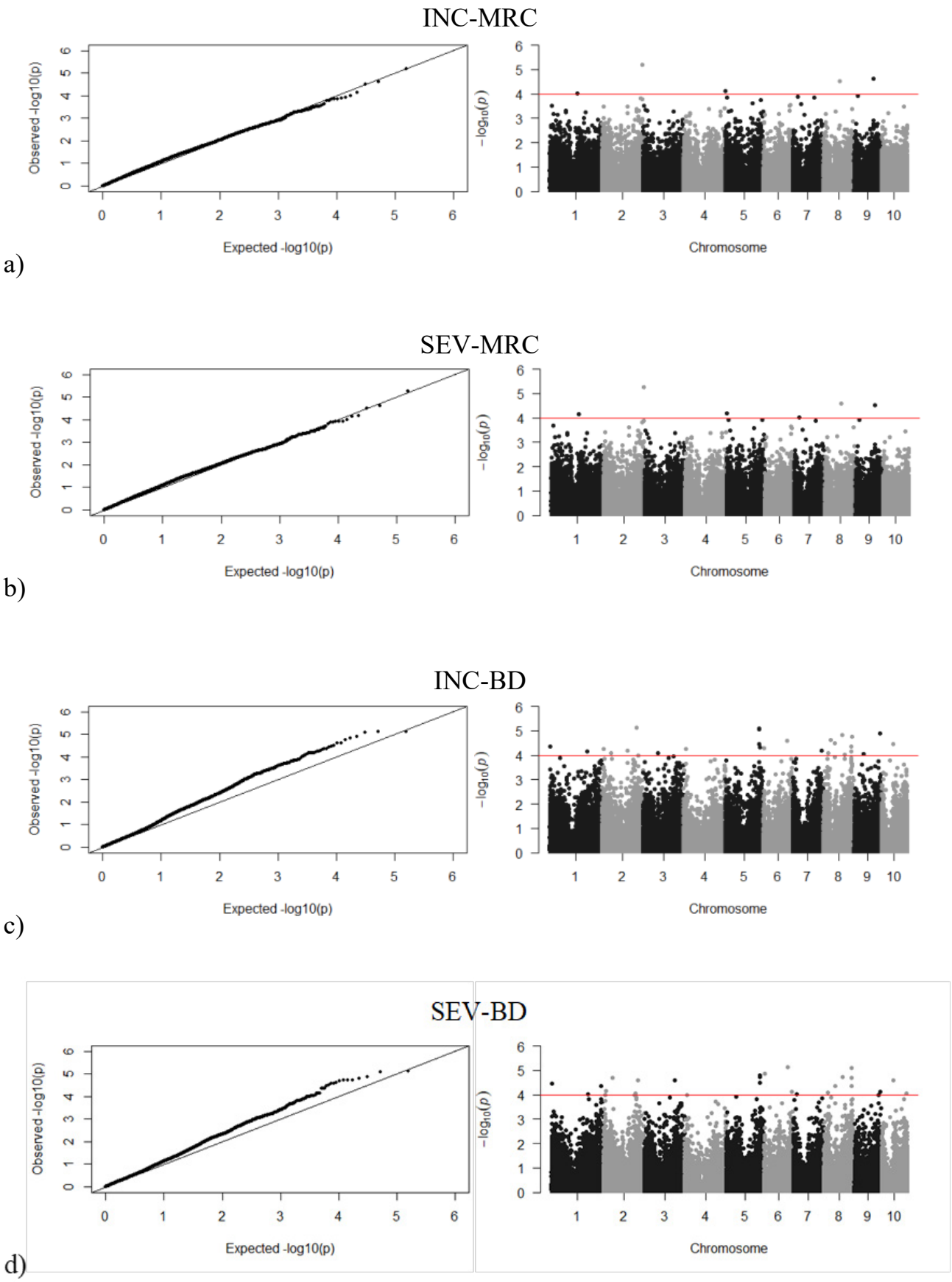


Figura 1. Gráficos *Q-Q plot* y *Manhattan plot* para el GWAS de los caracteres incidencia y severidad de Mal de Río Cuarto y de bacteriosis en una población diversa de 185 líneas de maíz. a) INC-MRC, b) SEV-MRC, c) INC-BD, d) SEV-BD. Umbral -log10 (0,0001).

Tabla 3. Marcadores moleculares asociados con incidencia (INC) y severidad (SEV) de Mal de Río Cuarto en una población diversa de 185 líneas de maíz evaluadas en Río Cuarto, Córdoba, Argentina, en el ciclo de cultivo 2019/2020.

Región genómica	Carácter	Marcador	Cromosoma	Bin	Alelo	*p*	R^2
1	INC	S1_158003836	1	1,05	A/T	$9{,}70\times10^{-5}$	0,11
	SEV					$7{,}01\times10^{-5}$	0,12
2	INC	S2_234192015	2	2,09	G/A	$6{,}40\times10^{-6}$	0,19
	SEV					$5{,}25\times10^{-5}$	0,19
3	INC	S5_1726939	5	5.00	C/A	$7{,}46\times10^{-5}$	0,12
	SEV					$6{,}40\times10^{-5}$	0,12
4	INC	S8_93112262	8	8,03	G/A	$3{,}05\times10^{-5}$	0,12
	SEV					$2{,}46\times10^{-5}$	0,12
5	INC	S9_112644510	9	9,04	T/A	$2{,}39\times10^{-5}$	0,18
	SEV					$3{,}07\times10^{-5}$	0,17

Umbral de significancia: valor-p < 0,0001.

DISCUSIÓN

La población diversa de líneas de maíz del CIMMYT evaluada en Río Cuarto, presentó amplia variabilidad fenotípica y genotípica respecto a la reacción de las líneas para resistencia a Mal de Río Cuarto (MRC) y a bacteriosis (BD). Los modelos mixtos ofrecen un enfoque apropiado para analizar conjuntamente los caracteres evaluados para la identificación de mejores genotipos considerando la posible correlación entre éstos (Malosetti *et al.*, 2008). La correlación entre caracteres puede ser positiva, negativa o nula. Wisser *et al.* (2011) y López Zuniga *et al.* (2019) observaron correlación positiva en caracteres de resistencia medido para tres enfermedades causadas por *Cochliobolus heterostrophus*, *Setosphaeria turcica*, y *Cercosporazeae maydis* en maíz. Sin embargo, en nuestro estudio no se observó correlación entre la respuesta al MRC y a la BD. Consecuentemente, el MLM multivariado constituye una herramienta útil para evaluar simultáneamente INC y SEV de MRC o de BD, pero no para el tratamiento de ambas enfermedades ya que la correlación genética entre resistencia a MRC y a BD fue baja. Los síntomas de BD se desarrollan principalmente en las hojas, mientras que, los síntomas de MRC afectan toda la planta cuando el ataque es severo (Abdala *et al.*, 1997). La reducción de la lámina foliar puede, consecuentemente, dificultar la evaluación de la resistencia a bacteriosis produciendo subestimaciones en las correlaciones genéticas entre caracteres.

Respecto a la estructura genética poblacional detectada por Structure, observamos alta consistencia con el agrupamiento de líneas como Lowland, Subtropical y Highland propuesto por Wu *et al.* (2016), quienes trabajaron con un set de líneas en el cual están contenidas las líneas del presente estudio. Al igual que lo observado por Gutiérrez *et al.* (2015), no hubo un único modelo GWAS que ajustara adecuadamente para todos los caracteres. El modelo que mejor ajustó para los caracteres INC-MRC y SEV-MRC fue el modelo Q (Pritchard, 2000); mientras que para INC-BD y SEV-BD ajustó el modelo PCA + K (Zhao *et al.*, 2007) que sugiere un agrupamiento más difuso y más dependiente de las relaciones de pares de líneas que el modelo Q.

El análisis de GWAS permitió identificar regiones genómicas previamente reportadas para la presencia de genes de resistencia a enfermedades. Di Renzo *et al.* (2004) y Bonamico *et al.* (2012) detectaron una región genómica que confiere resistencia a MRC en el bin 8.03 en una población segregante $F_{2:3}$ y en una población de RIL, ambas derivadas del cruzamiento entre las líneas de maíz M017 (susceptible) y BLS14 (resistente). Gowda *et al.* (2015), detectó una región genómica en el bin 1.05 para resistencia a *Maize lethal necrosis disease* en un germoplasma tropical de maíz. Gomes de Paula Lana *et al.* (2017), evaluaron la resistencia a bacteriosis, *Maize white spot*, en una población biparental y al igual que en el presente estudio detectaron regiones genómicas asociadas para resistencia en los bin 2.07 y 8.03. Rossi *et al.* (2020) informaron dos regiones genómicas asociadas para resistencia a *Maize white spot* en los bin1.01 y 8.03 a partir de una población de líneas tropicales de maíz. Coincidentemente, estas dos regiones fueron asociadas significativamente con resistencia a bacteriosis en el presente estudio donde se utilizó una población diversa de diferente procedencia.

Tabla 4. Marcadores moleculares asociados con incidencia (INC) y severidad (SEV) de bacteriosis en una población diversa de 185 líneas de maíz evaluadas en Río Cuarto, Córdoba, Argentina, en el ciclo de cultivo 2019/2020.

Región genómica	Carácter	Marcador	Cromosoma	Bin	Alelo	*p*	R^2
1	INC	S1_5246419	1	1,01	G/T	$4,39x10^{-5}$	0,18
	SEV					$3,35x10^{-5}$	0,18
2	INC	S2_10495415		2,02	C/A	$5,44x10^{-5}$	0,13
	SEV					$9,91x10^{-5}$	0,12
3	INC	S2_54359257	2	2,04	T/A	$8,37x10^{-5}$	0,15
	SEV					$1,97x10^{-5}$	0,18
4	INC	S2_200456510		2,07	A/G	$7,30x10^{-6}$	0,17
	SEV					$2,46x10^{-5}$	0,15
5	INC	S5_193547866			C/A	$7,84x10^{-6}$	0,18
	SEV					$1,55x10^{-5}$	0,17
6	INC	S5_193554917	5	5,05	G/C	$8,42\ x10^{-6}$	0,19
	SEV					$1,90x10^{-5}$	0,18
7	INC	S5_193649629			C/A	$3,58x10^{-5}$	0,15
	SEV					$3,11x10^{-5}$	0,15
8	INC	S6_1960822		6.00	G/A	$5,14x10^{-5}$	0,13
	SEV		6			$1,39x10^{-5}$	0,15
9	INC	S6_137143684		6,05	C/T	$2,45x10^{-5}$	0,15
	SEV					$7,56x10^{-6}$	0,17
10	INC	S8_105846679			A/G	$1,48x10^{-5}$	0,17
	SEV			8,03		$1,93x10^{-5}$	0,17
11	INC	S8_63553514			C/G	$3,12x10^{-5}$	0,15
	SEV		8			$4,49x10^{-5}$	0,14
12	INC	S8_161634845			T/C	$4,26x10^{-5}$	0,14
	SEV			8,06		$7,89x10^{-6}$	0,17
13	INC	S8_162228387			T/C	$6,84x10^{-5}$	0,15
	SEV					$4,50x10^{-5}$	0,15
14	INC	S9_151998388	9	9,07	T/G	$1,23x10^{-5}$	0,18
	SEV					$7,31x10^{-5}$	0,15
15	INC	S10_67227666	10	10,03	A/G	$3,55x10^{-5}$	0,17
	SEV					$2,55x10^{-5}$	0,18

Umbral de significancia: valor-p < 0,0001.

La variabilidad fenotípica y genotípica existente en el germoplasma diverso constituido por las líneas de maíz de CIMMYT resulta importante para apoyar programas de mejoramiento genético de maíz locales. El GWAS detectó regiones genómicas con alelos promisorios tanto para MRC como para BD, que individualmente explican una proporción de la variación fenotípica que oscila entre 10 y 20 % de la variación total, pero no se identificaron regiones genómicas comunes a ambas enfermedades. Estudios multiambientales permitirán confirmar si estas regiones genómicas resultan promisorias para incorporar resistencia a Mal de Río Cuarto y a bacteriosis en programas locales de mejoramiento genético de maíz.

BIBLIOGRAFÍA

Abdala G., Vigliocco A., Boito G., Lorenzo E. (1997) Dwarfism in Mal de Río Cuarto disease. Histology of maize stems and endogenous gibberellin levels. Biocell. 20: 211-220.

Agrios G. (2005) Plant pathology. Academic Press.

Bonamico N.C., Di Renzo M.A., Ibañez M.A., Borghi M.L., Díaz D.G., Salerno J.C., Balzarini M.G. (2012) QTL analysis of resistance to Mal de Río Cuarto disease in maize using recombinant inbred lines. J. Agric. Sci. 150 (5): 619-629. https://doi.org/10.1017/S0021859611000943.

Bradbury P., Zhang Z., Kroon D., Casstevens T., Ramdoss Y., Buckler E. (2007) Tassel: Software for association mapping of complex traits in diverse samples. Bioinformatics 23: 2633-2635.

Chen J., Zavala C., Ortega N., Petroli C., Franco J., Burgueño J., Costich D., Hearne S.J. (2016) The development of quality control genotyping approaches: A case study using elite maize lines. PLoS ONE 11 (6): 1-17. https://doi.org/10.1371/journal.pone.0157236.

Covarrubias Pazaran G. (2016) Genome assisted prediction of quantitative traits using the R package sommer. PLoS ONE 11 (6): e0156744. https://doi:10.1371/journal.pone.0156744.

Cullis B.R., Smith A.B., Coombes N.E. (2006) On the design of early generation variety trials with correlated data. J. Agric. Biol. Environ. Stat. 11 (4): 381-393. https://doi.org/10.1198/108571106X154443.

Di Renzo M.A., Bonamico N.C., Díaz D.G., Ibañez M.A., Faricelli M.E., Balzarini M.G., Salerno J.C. (2004) Microsatellite markers linked to QTL for resistance to Mal de Río Cuarto disease in *Zea mays* L. J. Agric. Sci. 142 (3): 289-295. https://doi.org/10.1017/S0021859604004307.

Gimenez Pecci M. P. (2012) Mal de Río Cuarto del maíz. In: Gimenez Pecci M.P., Laguna I.G., Lenardón S.L. (Eds.) Enfermedades del maíz producidas por virus y mollicutes en Argentina. INTA, Buenos Aires, pp. 41-56.

Gomes de Paula Lana U., Prazeres de Souza I.R., Noda R.W., Pastina M.M., Vieira Magalhaes J., Teixeira Guimaraes C. (2017) Quantitative trait loci and resistance gene analogs associated with maize white spot resistance. Plant Dis. 101: 200-208. doi:10.1094/PDIS-06-16-0899-RE.

Gowda M., Das B., Makumbi D., Babu R., Semagn K., Mahuku G., Olsen M.S., Bright J.M., Beyene Y., Prasanna B.M. (2015) Genome wide association and genomic prediction of resistance to maize lethal necrosis disease in tropical maize germplasm. Theor. Appl. Genet. 128 (10): 1957-1968. https://doi.org/10.1007/s00122-015-2559-0.

Gurr S.J., Rushton P.J. (2005) Engineering plants with increased disease resistance: what are we going to express? Trends Biotechnol. 23: 275-282.

Gutiérrez L., Germán S., Pereyra S., Hayes P.M., Pérez C.A., Capettini F., Locatelli A., Berberian H.M., Falconi E.E., Estrada R., Fros D., Gonza V., Altamirano H., Huerta Espino J., Neyra E., Orjeda G., Sandoval Islas S., Sing R., Turkington K., Castro A.J. (2015) Multi-environment multi-QTL association mapping identifies disease resistance QTL in barley germplasm from Latin America. Theor. Appl. Genet. 128: 501-519.

Hallauer A.R., Miranda J.B. (1988) Quantitative genetics in maize breeding, 2nd Edn. Iowa State University Press, Ames, IA.

Holland J.B., Nyquist W.E., Cervantes Martinez C.T. (2003) Estimating and interpreting heritability for plant breeding: an update. Plant Breed. Rev. 22: 9-111.

King A.M.Q., Adams M.J., Carstens E.B., Lefkowitz E.J. (2012) Virus taxonomy: classification and nomenclature of viruses: Ninth Report of the International Committee on Taxonomy of Viruses. Elsevier Academic Press.

Knott S.A., Haley C.S. (2000) Multitrait least squares for quantitative trait loci detection. Genetics 156 (2): 899-911.

Li J., Ji L. (2005) Adjusting multiple testing in multilocus analyses using the eigen values of a correlation matrix. Heredity 95: 221-227. https://doi.org/10.1038/sj.hdy.6800717.

Lopez Zuniga L.O., Wolters P., Davis S., Weldekidan T., Kolkman J.M., Rebecca N., Hooda K.S., Rucker E., Thomason W., Wisser R., Balint Kurti P. (2019) Using maize chromosome segment substitution line populations for the identification of loci associated with multiple disease resistance. G3: Genes, Genom. Genet. 9 (1): 189-201. https://doi.org/10.1534/g3.118.200866.

Maier R., Moser G., Chen G.B., Ripke S., Coryell W., Potash J.B., Scheftner W.A., Shi J., Weissman M.M., Hultman C.M., Landén M., Levinson D.F., Kendler K.S., Smoller J.W., Wray N.R., Lee S.H. (2015) Joint Analysis of Psychiatric Disorders Increases Accuracy of Risk Prediction for Schizophrenia, Bipolar Disorder, and Major Depressive Disorder. Am. J. Hum. Genet. 96: 283-294.

Malosetti M., Ribaut J.M., Vargas M., Crossa J., Van Eeuwijk F.A. (2008) A Multi-trait multi-environment QTL mixed model with an application to drought and nitrogen stress trials in maize (*Zea Mays* L.). Euphytica 161 (1-2): 241-57. https://doi.org/10.1007/s10681-007-9594-0.

Nelson R., Wiesner Hanks T., Wisser R., Balint Kurti P. (2018) Navigating complexity to breed disease-resistant crops. Nat. Rev. Genet. 19: 21-33. doi:10.1038/nrg.2017.82.

Ornaghi J.A., Boito G., Sanchez G., March G., Beviacqua J.E. (1993) Studies on the populations of *Delphacodes kuscheli Fennah* in different years and agricultural areas. J. Plant Genet. Breed. 47: 277-282.

Ornaghi J.A., March G.J., Boito G.T., Marinelli A., Beviacqua J.E., Giuggia J., Lenardon S.L. (1999) Infectivity in natural populations of *Delphacodes kuscheli* vector of "Mal Río Cuarto" Virus. Maydica. 44: 219-223.

Parisseaux B., Bernardo R. (2004) In silico mapping of quantitative trait loci in maize. Theor. Appl. Genet. 109: 508-514.

Plazas M.C., De Rossi R.L., Brücher E., Guerra F.A., Vilaró M., Guerra G.D., Wu G., Ortiz Castro M.C., Broders K. (2018) First report of *Xanthomonas vasicola* pv. vasculorum causing bacteria leaf streak of maize (*Zea Mays*) in Argentina. Plant Dis. 102 (5): 1026. https://doi.org/10.1094/PDIS-10-17-1578-PDN.

Price A., Patterson N., Plenge R., Weinblatt M., Shadick N., Reich D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38: 904-909.

Pritchard J., Stephens M., Rosenberg N., Donnelly P. (2000) Association mapping in structured populations. Am. J. Hum. Genet. 67: 170-181.

R Core Team (2016) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Remington D.L., Thornsberry J.M., Matsuoka Y., Wilson L.M., Whitt S.R., Doebley J., Kresovich S., Goodman M.M., Buckler E.S. (2001) Structure of linkage disequilibrium and phenotypic associations in the maize genome. Proc. Natl. Acad. Sci. U.S.A. 98 (20): 11479-84. https://doi.org/10.1073/pnas.201394398.

Rossi E.A., Borghi M.L., Di Renzo M.A., Bonamico N.C. (2015) Quantitative trait loci (QTL) identification for resistance to *Mal de Río Cuarto virus* (MRCV) in maize based on segregate population. Open Agric. J. 9: 48-55. doi:10.2174/18 74331501509010048.

Rossi E.A., Ruiz M., Rueda Calderón M.A., Bruno C.I., Bonamico N.C., Balzarini M.G. (2019) Metaanalysis of loci for resistance to maize diseases. Crop Sci. 59: 1-15.

Rossi E.S., Kuki M.C., Pinto R.J.B., Scapim C.A., Faria M.V., De Leon N. (2020) Genomic-wide association study for white spot resistance in a tropical maize germplasm. Euphytica 216 (1). https://doi.org/10.1007/s10681-019-2550-y.

Schuelter A.R., Prazeres De Souza I.R., Tavares F., Dos Santos M.X., Oliveira E., Guimarães C.T. (2003) *Phaeosphaeria* genetic control of maize resistance to *Phaeosphaeria*. Rev. Bras. Milho e Sorgo 2 (1): 80-86.

Turner S.D. (2018) qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J. Open Source Softw.

Wisser R.J., Kolkman J.M., Patzoldt M.E., Holland J.B., Yu J., Krakowsky M., Nelson R.J., Balint Kurti P.J. (2011) Multivariate analysis of maize disease resistances suggests a pleiotropic genetic basis and implicates a GST gene. Proc. Natl. Acad. Sci. U.S.A. 108 (18): 7339-44. https://doi.org/10.1073/pnas.1011739108.

Wu Y., San Vicente F., Huang K., Dhliwayo T., Costich D.E., Semagn K., Sudha N., Olsen M., Prassana B.M., Zhang X., Babu R. (2016) Molecular characterization of CIMMYT maize inbred lines with genotyping by sequencing SNPs. Theor. Appl. Genet. 129 (4): 753-65. https://doi.org/10.1007/s00122-016-2664-8.

Yan J.B., Warburton M., Crouch J. (2011) Association mapping for enhancing maize (*Zea mays* L.) genetic improvement. Crop Sci. 51: 433-49. https://doi: 10.2135/cropsci2010.04.0233.

Yu J., Pressoir G., Briggs W., Bi I., Yamasaki M., Doebley J., McMullen M., Gaut B., Holland J., Kresovich S., Buckler E. (2006) A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38: 203-208.

Zhao K., Aranzana M., Kim S., Lister C., Shindo C., Tang C., Toomajian C., Zheng H., Dean C., Marjoram P., Nordborg M. (2007) An *Arabidopsis* example of association mapping in structured samples. PLoS Genet. 3: e4.

Zila C.T., Ogut F., Romay M.C., Gardner C.A., Buckler E.S., Holland J.B. (2014) Genome-wide association study of fusarium ear rot disease in the U.S.A. Maize inbred line collection. BMC Plant Biol.14 (1): 1-15. https://doi.org/10.1186/s12870-014-0372-6.

BAG | Journal of Basic & Applied Genetics

Vol XXXII (1): 35-42; July 2021

Journal of the Argentine Society of Genetics

ANALYSIS OF MIXED DATA TO SELECT BANANAS CLONES (*Musa* SPP.) TO BE INCLUDED IN A GERMPLASM BANK

ANÁLISIS DE DATOS MIXTOS PARA SELECCIONAR CLONES DE BANANA (*Musa* SPP.) A SER INCLUIDOS EN UN BANCO DE GERMOPLASMA

Del Medico A. P.[1], Tenaglia G.[2], Lavalle A. L.[3], Vitelleschi M. S.[4], Pratta G. R.[1,5]

[1] Instituto de Investigaciones en Ciencias Agrarias de Rosario, Consejo Nacional de Investigaciones Científicas y Técnicas.

[2] Centro de Investigación y Desarrollo Tecnológico para la Pequeña Agricultura Familiar, Instituto Nacional de Tecnología Agropecuaria.

[3] Departamento de Estadística, Universidad Nacional del Comahue.

[4] Consejo de Investigaciones de la Universidad Nacional de Rosario, Instituto de Investigaciones Teóricas y Aplicadas de la Escuela de Estadística, Facultad de Ciencias Económicas y Estadística, Universidad Nacional de Rosario.

[5] Facultad de Ciencias Agrarias, Universidad Nacional de Rosario.

Corresponding author:
Guillermo R. Pratta
gpratta@unr.edu.ar

ORCID 0000-0002-3682-0946

Cite this article as:
Del Medico A. P., Tenaglia G., Lavalle A. L., Vitelleschi M. S., Pratta G. R. 2021. ANALYSIS OF MIXED DATA TO SELECT BANANAS CLONES (*Musa* SPP.) TO BE INCLUDED IN A GERMPLASM BANK. BAG. Journal of Basic and Applied Genetics XXXII (1): 25-42.

Received: 12/19/2021
Revised version received: 05/03/2021
Accepted: 06/17/2021

General Editor: Elsa Camadro

DOI: 10.35407/bag.2021.32.01.04

ISSN online version: 1852-6233

Available online at
www.sag.org.ar/jbag

ABSTRACT

In an asexually reproducing hybrid such as banana (*Musa* spp.), the assessment of clones in the short term is limited because replications are frequently unavailable in the proper number. The aim of this work is to propose the Multiple Factor Analysis of Mixed Data (MFAmix) as a tool for establishing objective criteria to identify banana clones that preserve variability for qualitative and quantitative variables. In the long term, the aim is the development of a banana germplasm bank. MFAmix was applied on a population composed of 124 banana clones collected from different farmers' fields and four controls. Two groups of variables related to the agronomic aptitude of the clones were evaluated, one composed of nine quantitative variables, and the other, composed of three dichotomous qualitative variables. A Selection Index (SI) was built from the MFAmix coordinates in order to rank the clones and select a subset that allows to preserve the existing genetic variability. The first two axes of MFAmix explained a 49.47% of the total data variability. The set of the banana clones was successfully characterized based on quantitative and qualitative variables. In the long term, the creation of a banana germplasm bank should consider the height and diameter of the plant, the rachis bunch weight and the hands weight, and the qualitative variable plant leafiness.

Key words: asexual hybrid, collection of germplasm, multivariate analysis, Musaceae.

RESUMEN

En un híbrido de reproducción asexual como banana (*Musa* spp.), la evaluación de los clones en el corto plazo es limitada debido a que generalmente no se cuenta con el número adecuado de repeticiones. El objetivo de este trabajo es aplicar la técnica de Análisis Factorial Múltiple de Datos Mixtos (AFMmix) como una herramienta para establecer criterios objetivos de manera de identificar clones de banana que preserven la variabilidad de los caracteres cualitativos y cuantitativos. A largo plazo, el objetivo es desarrollar un banco de germoplasma de banana. Se aplicó el AFMmix a una población de 124 clones de banana recolectados de diferentes campos de productores y cuatro testigos comerciales. Se evaluaron dos grupos de variables relacionadas con la aptitud agronómica de los clones, uno compuesto por nueve caracteres cuantitativos, y el otro, por tres caracteres cualitativos dicotómicos. Se construyó un Índice de Selección (IS) a partir de las coordenadas del AFMmix de manera de ordenar a los clones de banana para seleccionar un subconjunto de ellos que permita conservar la variabilidad genética existente. Los dos primeros ejes del AFMmix explicaron un 49,47% de la variabilidad total de los datos. Se caracterizó satisfactoriamente al conjunto de clones de banana a través de las variables cuantitativas y cualitativas. A largo plazo, en la creación de un banco de germoplasma de banana se debe considerar a la altura y diámetro de la planta, al peso del raquis y peso de las manos, y al carácter cualitativo frondosidad de la planta.

Palabras clave: híbrido asexual, colección de germoplasma, análisis multivariado, Musaceae.

INTRODUCTION

Banana (*Musa* spp.) is a crop of fundamental importance for the economies of many developing countries. In terms of gross production value, it is the fourth most important food crop in the world, after rice, wheat and maize (Arias *et al.*, 2004). In northern Argentina, a sub-tropical area, the banana crop suffers from suboptimal climate conditions, affecting diversity. Thus, there are genotypes which are adapted to environments which are less favorable for traditional production. Therefore, in Argentina there is a crop diversity which is unique in the world (Ermini *et al.*, 2013, 2016). The banana is an asexual reproduction hybrid, the selection of clones in the short term, is limited due to the lack of the appropriate number of repetitions.

A germplasm bank is a collection of live plant material which aims to preserve the genetic variability existing in one or more species of interest. Germplasm banks are the main means to protect the plant diversity of the different crop species, and identify accessions for breeding programs, basic researches, and production. Agronomy problems derived from the excessive uniformity of the crops can be solved by introducing local varieties. Hence both the conservation and use of the genetic variability available in germplasm banks are presently revaluated (Defacio, 2009).

Some authors reported the use of Principal Components Analysis (PCA) (Defacio, 2009) and Generalized Procrustes Analysis (GPA) (Bramardi, 2005) to identify varieties to conserve in a germplasm bank. The disadvantages of using these methodologies are as follows. Through PCA, it is not possible to analyze qualitative variables as active variables; i.e. they can only be introduced as supplementary variables, not intervening in calculating the coordinates of individuals and variables. Through GPA, it is possible to work with the synthetic variables resulting from applying PCA on the quantitative variables and Principal Coordinate Analysis (PCoA) on the qualitative variables. On this occasion, it is not possible to determine which variables are the ones which contribute the most to the formation of the axes or factors, making it difficult to characterize the cultivars through the evaluated variables. There is a relatively new methodology, the Multiple Factor Analysis of Mixed Data (MFAmix) (Pagès, 2002, 2014) that allows the characterization of the cultivars according to the quantitative and qualitative variables simultaneously and determines which variables mostly contribute to the total variability. MFAmix provides equations that are linear combinations of the original variables and form axes of highest variation that allow to differentiate cultivars. Therefore, the aim of this report was to propose the MFAmix as a tool for establishing objective criteria to identify banana clones that preserve variability for qualitative and quantitative variables. In the long term, the objective is to form a banana germplasm bank representing most of the plant diversity available in the northeastern region of Argentina.

MATERIALS AND METHODS

Plant material

The study population was composed of 124 banana clones collected from different producers' fields in the province of Formosa, Argentina (Ermini *et al.*, 2018) and four commercial varieties extensively used in the world production that were the experimental controls: Williams (Control 1), Jaffa (Control 2), Gal Azul (Control 3) and Gran Enanao (Control 4) (Figure 1).

Figure 1. Trees and bunches of banana fruit that correspond to clones collected from different farmers' fields.

An augmented block design (Cotes and Ñústez, 2001) was carried out with 15 blocks of 14 plants each, where only the controls have repetitions. It was accomplished at the experimental field of INTA Formosa which is located in northern Argentina (26°11'31.8"S, 58°12'22.4"W), during the 2016-2017 crop season.

Two groups of phenotypic variables related to the agronomic aptitude of the clones were evaluated, one

composed of nine continuous quantitative variables, and the other, composed of three dichotomous qualitative variables.

The quantitative variables were plant height (m), plant diameter (cm), rachis bunch weight (kg), hands weight (kg), second hand diameter (cm), last hand diameter (cm), second hand length (cm), last hand length (cm) and peel thickness (mm). In a previous communication (Del Medico *et al.*, 2018a), the existence of genetic variability for these quantitative variables was verified by a method originally developed to take into account the lack of genotypic replications for clones due to the experimental design used. The qualitative variables were plant leafiness (low or high), bunch size (small or big) and prolificity of tier bunch (low or high).

Statistical methodology

The MFAmix (Pagès, 2002, 2014) was applied. This methodology allows the analysis of data tables in which the same group of individuals is described through a group of variables, evaluated in different conditions, moments or places. Variables can be quantitative or qualitative, with the only restriction that the nature of these variables must be the same within each group. MFAmix provides a similar weighting to both kinds of variables. In general terms, the MFAmix algorithm consists of two stages. The *preliminary stage (separate analysis)*, in which each group of variables is analyzed separately, in the case of quantitative variables set, through a PCA and for qualitative variables set, through a Multiple Correspondence Analysis (MCA). The first eigenvalue from each of these analyses will be used in the subsequent step. The *main stage (global analysis)* consists in performing a PCA on the whole data resulting from the juxtaposition of the configurations obtained in the separate analysis, which are weighted by the inverse of the corresponding first eigenvalue. This weighting maintains the structure of each matrix and manages to balance the influence of the different groups of variables. The objective of the MFAmix technique is to highlight the main variability of individuals, the latter being balanced by the various groups of variables.

A global measurement of the relation between the configurations of both groups of variables defined for the same individuals could be calculated through the RV coefficient (Abdi, 2007). The RV coefficient takes values between zero (the configurations are orthogonal) and one (the configurations are homothetic).

A Selection Index (SI) was built from the coordinates of the quantitative variables obtained through MFAmix. SI is a linear combination of the standardized quantitative variables whose weights are their coordinates obtained through MFAmix, multiplied by the inertia explained in that factor. The quantitative variables involved in its construction were those whose contribution to each factor exceeded 2/3 of the corresponding maximum coefficient in absolute value. The construction of this SI was based on Del Medico *et al.* (2018b). The FactoMineR package of R statistical software was used to accomplish this analysis.

RESULTS

Multiple Factor Analysis of Mixed Data

The first PCA factor was moderately correlated with the first MCA factor (-0.61.) The rest of the correlations between the factors of the separate analyzes were low (Table 1). These correlation coefficients indicated that the intensity of the relationship between the two analyses was slight, being only linked on the first factor. Two factors were retained from the MFAmix, which explain 49.47% of the total data inertia (Table 2). No rotation method was used. The first two factors of the MFAmix were quite close to the factors of the same rank in the separate analyses, except the second factor of the quantitative group. Therefore, the use of MFAmix properly balanced the contribution of these two types of variables (Figure 2). The quantitative attributes which most contributed to the formation of the first factor were hands weight, rachis bunch weight, and diameter, width and height of the plant. In the second factor, no considerable contributions from the quantitative values were observed (Table 3).

The qualitative variables mostly contributing to the first factor were bunch size and prolificity of tier bunch. On the second factor, the largest contribution was made by the qualitative variable plant leafiness, followed by bunch size (Table 3). In the representation of both groups of variables on the first factor plane, there were no observable differences on the first derived factor. However, both groups showed differences on the second factor (Figure 3).

Table 1. Correlations between factors obtained in the preliminary stage (separate analysis) of Multiple Factorial Analysis of Mixed Data. The study population was composed of 124 banana clones and four controls. Two groups of variables related to the agronomic aptitude of the clones were evaluated, one composed of nine quantitative variables, and the other, composed of three dichotomous qualitative variables. Each group of variables was analyzed separately through a Principal Components Analysis (PCA) or Multiple Correspondence Analysis (MCA) as appropriate.

		Group 1 (PCA)		
		Factor 1	**Factor 2**	**Factor 3**
Group 2 (MCA)	**Factor 1**	-0.61	0.28	0.30
	Factor 2	0.00	-0.06	-0.09
	Factor 3	0.16	0.06	0.01

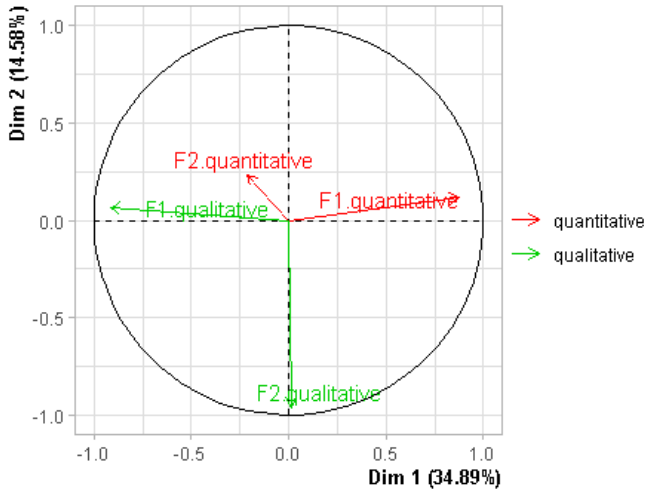


Figure 2. Factors of the separate analyzes on the first two axes of Multiple Factorial Analysis of Mixed Data (MFAmix). The study population was composed of 124 banana clones and four controls. Two groups of variables related to the agronomic aptitude of the clones were evaluated, one composed of nine quantitative variables, and the other composed of three dichotomous qualitative variables. Dim 1 and Dim 2 correspond to the Factor 1 and 2, respectively, obtained in the MFAmix.

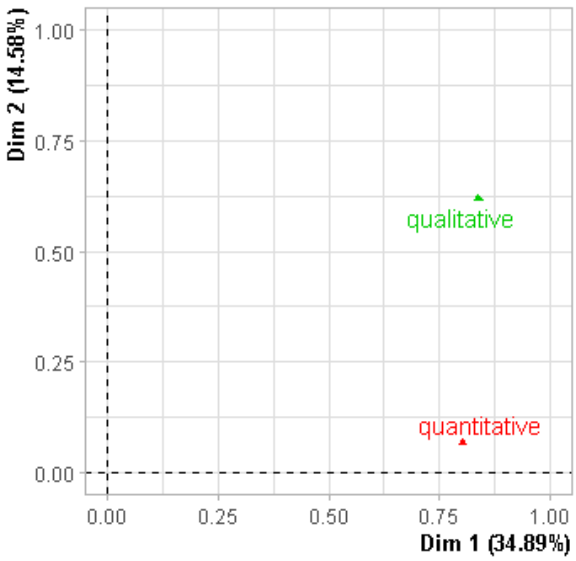


Figure 3. Representation of the groups on the first two factors of Multiple Factorial Analysis of Mixed Data (MFAmix). The study population was composed of 124 banana clones and four controls. Two groups of variables related to the agronomic aptitude of the clones were evaluated, one composed of nine quantitative variables, and the other composed of three dichotomous qualitative variables. Dim 1 and Dim 2 correspond to the Factor 1 and 2, respectively, obtained in the MFAmix.

Table 2. Decomposition of total inertia by factor obtained through Multiple Factor Analysis of Mixed Data. The study population was composed of 124 banana clones and four controls. Two groups of variables related to the agronomic aptitude of the clones were evaluated, one composed of nine quantitative variables, and the other, composed of three dichotomous qualitative variables.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Inertia %	34.89	14.58	12.46	8.45	8.24
Eigenvalues	1.64	0.69	0.59	0.40	0.39

The RV calculated between both groups of variables was equal to 0.30, indicating that the relationship between the configurations corresponding to both groups of variables under study was low, i.e., its information regarding total variability was complimentary. Considering that the RV obtained was low, that the discrepancies between the groups of variables appeared on the second factor, and that the evaluated variable which mostly contributes to the construction of such a factor was the qualitative variable plant leafiness, the banana clones were classified according to the aforementioned variable. Hence, two groups of clones were formed, one composed of Control 3, and four clones corresponding to plants with low plant leafiness, and the other one composed of Control 1, Control 2 and Control 4, and 78 clones corresponding to plants with high plant leafiness. For this reason, individuals were represented in the first principal plane of MFAmix, according to the qualitative variable plant leafiness.

The first factor orders the individuals according to the quantitative variables. The second factor separates the individuals according to the qualitative variable plant leafiness. The clones corresponding to plants with high plant leafiness were found in the superior part (in red) and plants with low plant leafiness in the inferior part (in green) (Figure 4).

Selection Index

Only the first factor was included in the construction of the SI, given that the quantitative variables did not present considerable contributions to the second MFAmix factor (Table 4) Based on data presented in Table 2 and Table 3, the SI constructed is:

$$SI = 1.64\,(0.62\,\textit{plant height} + 0.77\,\textit{plant diameter} + \\ + 0.82\,\textit{rachis bunch weigh} + 0.83\,\textit{hands weight})$$

The banana clones were arranged according to this SI in each of the two groups previously determined according to the plant leafiness (high or low) (Table 4).

Highlighted numbers in Table 4 identify the clones selected for the construction of the germplasm bank. 40 clones were selected to create the germplasm bank, which represent approximately 30% of the total banana clones studied in this research. The selected number of clones in each group was proportional to their size. It is recommended, in order to preserve the existing genetic

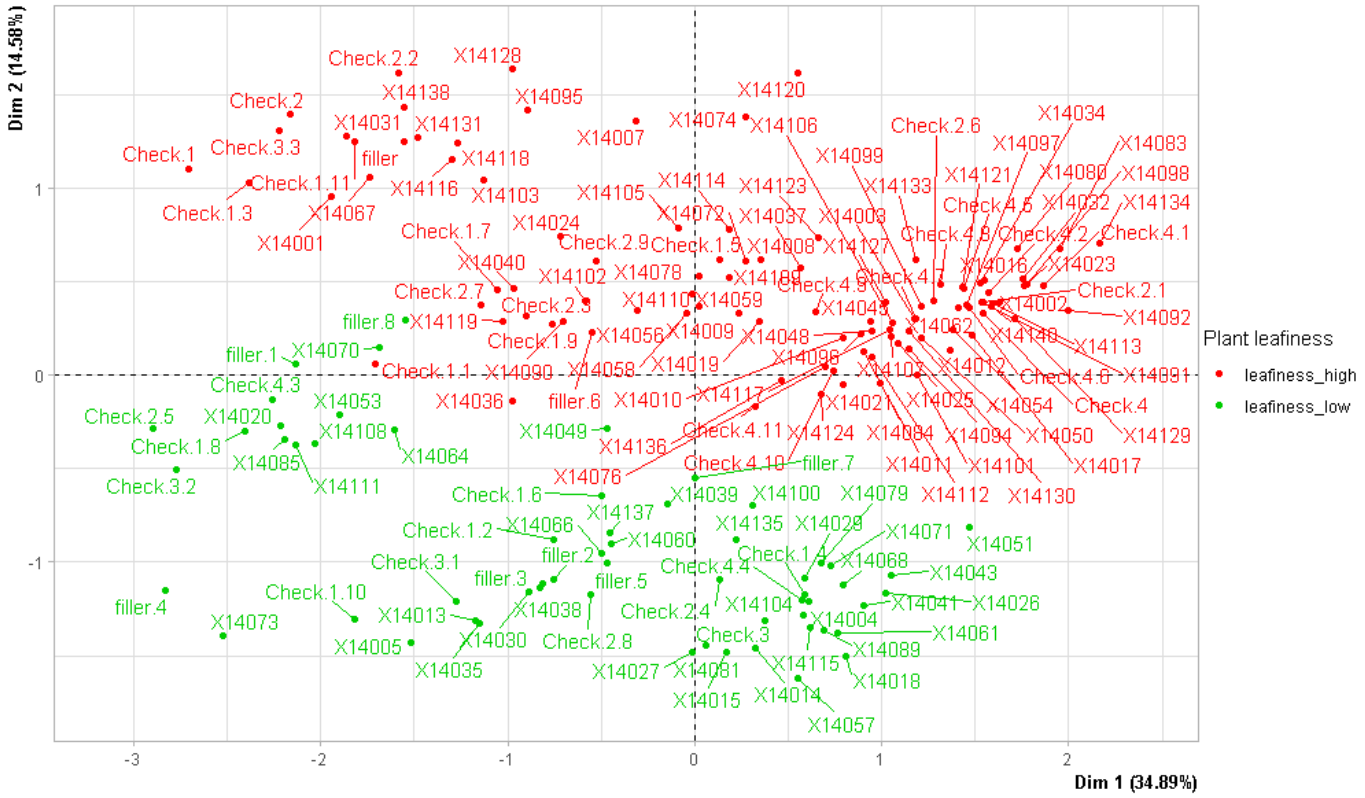


Figure 4. Representation of banana clones on the first two factors of Multiple Factorial Analysis of Mixed Data (MFAmix), according to the qualitative variable plant leafiness. The study population was composed of 124 banana clones and four controls. Two groups of variables related to the agronomic aptitude of the clones were evaluated, one composed of nine quantitative variables, and the other composed of three dichotomous qualitative variables. Dim 1 and Dim 2 correspond to the Factor 1 and 2, respectively, obtained in the MFAmix

variability, to select clones with high, moderate and low SI inside each group.

DISCUSSION

Adequate classification and conservation of the variability present in the crops and their relatives are essential for the conformation of germplasm banks, which results critical for future breeding programs (Fundora Mayor *et al.*, 2004). The abundance of material to evaluate, the handling limitations and the fact that, in general, many variables are studied jointly, make the conformation of a germplasm bank more difficult.

The use of quantitative and qualitative variables allows the characterization of crops in a different and complementary manner. For this reason, it is important to use an analysis technique which gets a consensus between both types of variables (Defacio, 2009). For example, Bramardi *et al.* (2005) evaluated cucumber cultivars for agronomic variables of qualitative and quantitative classes, using the GPA technique for the joint analysis, and Defacio (2016) evaluated local maize populations by GPA technique with the aim of simultaneously analyzing the quantitative and qualitative variables. This methodology is used in order to deal jointly with both kinds of variables. In those cases, more numerous groups of cultivars were obtained using each kind of variable separately. However, through GPA, it is not possible to determine which variables are the most contributing to the formation of the axes or factors, which makes it difficult to characterize the cultivars through the evaluated variables.

The benefit of MFAmix over other existent methodologies is that it assigns equal importance to both groups of variables. Additionally, it allows the characterization of the individuals according to the quantitative and qualitative variables, and thus to form a subset which presents the greater diversity.

The MFAmix is a technique that allows deciding a selection criterion that involves variables of different nature. Therefore, this methodology is an appropriate tool for establishing objective criteria through the construction of a SI for identifying banana clones that represent the plant diversity available in the Argentinian Northeast.

Table 4. Selection Index (SI) of banana clones classified according to plant leafiness (high or low). SI was built from the coordinates of the quantitative variables obtained through Multiple Factorial Analysis of Mixed Data (MFAmix). SI is a linear combination of the standardized quantitative variables whose weights are their coordinates obtained through MFAmix, multiplied by the inertia explained in that factor.

Plant leafiness	Clone	SI	Plant leafiness	Clone	SI
High	14040	-6.98	Low	**filler 5**	**-16.13**
	14031	**-6.22**		14073	-12.06
	14119	-5.91		**14020**	**-7.16**
	14024	**-5.84**		filler 2	-7.07
	14138	-5.45		14085	-5.59
	Test 1	-5.41		14005	-5.45
	filler 1	-4.77		Test 3	-5.21
	14067	-4.18		**14070**	**-5.18**
	Test 2	-4.18		filler 9	-4.99
	14131	-4.06		14013	-4.94
	14090	-3.97		14053	-4.89
	14001	**-3.95**		**14111**	**-4.83**
	14036	-3.89		**14108**	**-4.31**
	14007	-3.76		14035	-3.93
	14019	**-3.69**		14030	-2.96
	14102	**-3.42**		14038	-2.72
	filler 7	-2.62		**filler 4**	**-2.51**
	14128	-2.29		filler 3	-2.44
	14105	-2.17		**14064**	**-1.82**
	14118	**-2.04**		14027	-1.71
	14116	-1.86		14066	-1.69
	14136	-1.46		14081	-1.65
	14056	**-1.11**		**14060**	**-1.17**
	14103	-0.79		14137	-1.09
	14133	-0.74		14015	-0.94
	14095	**-0.41**		filler 6	-0.72
	14124	-0.15		14049	-0.72
	14010	**-0.10**		14039	-0.07
	14078	0.17		**14079**	**0.06**
	14074	0.23		14014	0.54
	14009	0.44		**14029**	**0.81**
	14058	**0.73**		14104	1.03
	14008	0.76		**filler 8**	**1.16**
	14076	**0.76**		14068	1.19
	14096	0.85		14004	1.38
	14110	0.93		14071	1.62
	14127	0.99		**14115**	**1.85**
	14072	1.04		14043	2.25
	14045	**1.13**		14057	2.57
	14084	1.26		**14089**	**2.63**
	14106	1.29		14061	2.70
	14112	1.39		14041	3.03

Table 4 (continue). Selection Index (SI) of banana clones classified according to plant leafiness (high or low). SI was built from the coordinates of the quantitative variables obtained through Multiple Factorial Analysis of Mixed Data (MFAmix). SI is a linear combination of the standardized quantitative variables whose weights are their coordinates obtained through MFAmix, multiplied by the inertia explained in that factor.

Plant leafiness	Clone	SI	Plant leafiness	Clone	SI
	14107	**1.42**		14100	3.43
	14025	1.45		**14018**	**3.86**
	14048	**1.58**		14135	4.13
	14120	1.70		14026	4.28
	Test 4	1.74		**14051**	**5.77**
	14117	**2.02**		-	-
	14011	2.23		-	-
	14114	**2.24**		-	-
	14021	2.28		-	-
	14109	2.34		-	-
	14130	**2.41**		-	-
	14003	2.45		-	-
	14094	2.81		-	-
	14050	3.03		-	-
	14059	3.08		-	-
	14121	**3.11**		-	-
	14062	3.17		-	-
	14054	3.66		-	-
	14099	3.68		-	-
	14101	**3.78**		-	-
	14097	3.82		-	-
	14016	**3.92**		-	-
	14091	4.00		-	-
	14037	4.15		-	-
	14123	4.34		-	-
	14080	4.35		-	-
	14034	**4.36**		-	-
	14012	4.66		-	-
	14002	4.67		-	-
	14113	**4.82**		-	-
	14140	5.11		-	-
	14032	5.31		-	-
	14023	**5.55**		-	-
	14017	5.61		-	-
	14083	5.79		-	-
	14098	5.97		-	-
	14129	**6.74**		-	-
	14092	**8.01**		-	-
	14134	8.02		-	-

CONCLUSION

In the present study, through the MFAmix technique, associations between both groups of variables were detected, and the characterization of banana clones according to quantitative and qualitative variables was successful. In the long term, the creation of a banana germplasm bank should consider the quantitative variables plant height and plant diameter, rachis bunch weight and hands weight, as well as the qualitative variable plant leafiness.

REFERENCES

Abdi H. (2007) RV coefficient and congruence coefficient. In: Salkind N.J. (Eds.) Encyclopedia of measurement and statistics. SAGE Publications, Inc., Thousand Oaks, California, pp. 849-853.

Arias P., Dankers C., Liu P., Pilkauskas P. (2004) La economía mundial del banano 1985-2002. Food & Agriculture Organization.

Bramardi S.J., Bernet G.P., Asíns M.J., Carbonell E.A. (2005) Simultaneous agronomic and molecular characterization of genotypes via the generalised procrustes analysis: an application to cucumber. Crop Science. 45: 1603-1609.

Cotes J.M., Ñústez C.E. (2001) Propuesta para el análisis de diseños aumentados en fitomejoramiento: Un caso en papa. Revista Latinoamericana de la Papa. 12: 15-34.

Defacio R.A. (2009) Caracterización y evaluación de la variabilidad genética en poblaciones nativas de maíz (*Zea mays* L.) de la provincia de Buenos Aires en base a descriptores morfológicos y agronómicos. Master thesis, Universidad Nacional de Rosario, Argentina.

Defacio R.A. (2016) Evaluación comparativa de distintas estrategias de análisis de datos para la caracterización y ordenamiento de la variabilidad genética de poblaciones locales de maíz (*Zea mays* L.). Doctoral thesis, Universidad Nacional de Rosario, Argentina.

Del Medico A.P., Ermini J.L., Tenaglia G.C., Vitelleschi M.S., Lavalle A.L., Pratta G.R. (2018a) Propuesta de un enfoque estadístico para seleccionar clones de banana a partir de un diseño aumentado teniendo en cuenta la variabilidad genética. Journal of Basic and Applied Genetics. 14: 51-57.

Del Medico A.P., Ermini J.L., Tenaglia G.C., Vitelleschi M.S., Lavalle A.L., Pratta G.R. (2018b) Índice de selección para múltiples caracteres en una población de banana (*Musa* spp.). XXIII Reunión Científica del Grupo Argentino de Biometría. II Encuentro Argentino - Chileno de Biometría. Octubre 2018, Neuquén, Argentina, pp. 53

Ermini J.L., Pantuso F.S., Tenaglia G.C., Pratta G.R. (2013) Marcadores de AFLP en el cultivo de banana: selección de combinaciones de cebadores y caracterización de la biodiversidad. Revista de la Facultad de Ciencias Exactas, Químicas y Naturales, Universidad de Morón. 11: 83-110.

Ermini, J.L., Tenaglia, G.C., Pratta, G.R. (2016) Genetic diversity, ancestry relationships and consensus among phenotype and genotype in banana (*Musa acuminata*) clones from Formosa (Argentina) farmers. Plant Cell Biotechnology and Molecular Biology. 17: 267-278.

Ermini, J.L., Tenaglia, G.C., Pratta, G.R. (2018) Molecular diversity in selected banana clones (Musa AAA "Cavendish"): adapted to the subtropical environment of Formosa Province (Argentina). American Journal of Plant Science. 9: 2504-2513.

Fundora Mayor Z., Hernandez M., López R., Fernández L., Sánchez A., López J., Ravelo E.L. (2004) Analysis of the variability in collected peanut (*Arachis hypogaea* L.) cultivars for the establishment of core collections. Plant Genetic Resources Newsletter. 137: 9-13.

Pagès J. Analyse factorielle de données mixtes. (2002). Revue de Statistique appliquée. 52: 93-111.

Pagès J. (2014). Multiple factor analysis by example using R. CRC Press, Boca Raton.

BAG | Journal of Basic & Applied Genetics
Vol XXXII (1): 43-43; July 2021
Journal of the Argentine Society of Genetics

OBITUARIO - OBITUARY

DRA. ANA LÍA VARGAS

26/08/1950 - 14/02/2021

Quien haya conocido a Ana Lía Vargas sabe sobre el amor que tenía por la docencia y la pasión con la que la ejercía. La conocí como alumna y puedo recordar su sonrisa y la paciencia al enseñar, que siguió cultivando con el correr de los años.

Dotada de una inteligencia privilegiada nunca se detuvo en seguir su formación en docencia y completó la especialidad en Docencia Universitaria, la Diplomatura en Educación Médica de la Universidad de Tucumán, así como becas y pasantías internacionales en el tema.

Estrechó vínculos con expertos en educación de la Escuela de Medicina de la Universidad de Harvard, así como de la Fundación Internacional para el Avance e Investigación sobre Educación Médica (FAIMER) realizando numerosas Jornadas y Talleres en nuestra Facultad.

Ejerció la Genética Médica con una especial predilección por la Citogenética y Medicina Fetal en el Instituto de Genética de la Facultad de Ciencias Médicas y en hospitales de nuestro medio, como el Hospital Luis Lagomaggiore y el Hospital Universitario, lugares donde se ganó el cariño de los que la conocieron.

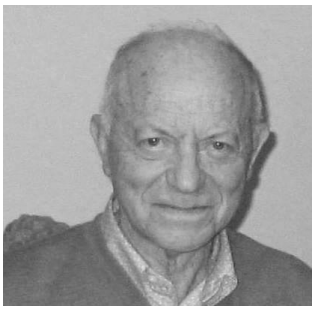
Su calidad como profesional y sobre todo los valores personales dejan en los que la conocimos y compartimos tiempo con ella una huella difícil de borrar. Su fallecimiento provocó numerosas expresiones de afecto de colegas médicos, docentes y alumnos, en el ámbito universitario y de salud de nuestro medio.

Alejandra Mampel

Especialista en Genética Médica
Universidad Nacional de Cuyo
Mendoza, Argentina

Available online at
www.sag.org.ar/jbag

BAG | Journal of Basic & Applied Genetics
Vol XXXII (1): 45-46; July 2021
Journal of the Argentine Society of Genetics



OBITUARIO - OBITUARY

ING. AGR. MIGUEL JACINTO ARTURI

La Plata, Buenos Aires
1929-2021

El Ing. Miguel J. Arturi completó sus estudios universitarios en la Facultad de Ciencias Agrarias y Forestales (FCAyF) de la Universidad Nacional de La Plata (UNLP), en 1959. En 1967 ganó una beca de la Organización de los Estados Americanos (OEA) para completar su formación en la Universidad de California donde obtuvo el grado de *Master of Science* (1969).

Se destacó en el área de la genética cuantitativa y del fitomejoramiento ocupando cargos como Director de la Estación Experimental Agropecuaria INTA La Banda (Santiago del Estero) y Director del Instituto Fitotécnico de Santa Catalina (FCAyF-UNLP). Su labor docente fue incansable, viéndose reflejada en la formación de recursos humanos a lo largo de su trayectoria. Fue docente de grado y posgrado en las áreas de Genética y de Mejoramiento Genético Animal y Vegetal, desempeñándose en la Universidad Nacional de Luján, la Universidad Nacional de Lomas de Zamora y la UNLP. En esta última Universidad fue distinguido como Profesor Extraordinario en la Categoría de Consulto. Fue autor de diversas creaciones fitogenéticas de algodón y cebadilla criolla, entre ellas el algodón "Quichua INTA" y las cebadillas criollas "Ñandú" y "Copetona". En 1984 publicó el libro de gran difusión: "El Algodón. Mejoramiento genético y técnica de su cultivo". Publicó numerosos trabajos científicos, tecnológicos y comunicaciones. En 1978, por su trabajo en el desarrollo de variedades de algodón para las áreas de regadío en Argentina, recibió el premio Fundación Ceres de la Academia Nacional de Agronomía y Veterinaria.

El Ing. Arturi fue miembro de la Sociedad Latinoamericana de Genética, de la Sociedad Argentina de Genética, del Grupo Argentino de Biometría, de *International Biometric Society*, de Comisiones Asesoras pertenecientes a la Comisión de Investigaciones Científicas de la Provincia de Buenos Aires (CIC-PBA) y al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET). En el ámbito de la FCAyF (UNLP) fue Consejero Superior, Director del Departamento de Biología y Ecología, Miembro del Comité Editorial de la Revista de la Facultad de Agronomía de la UNLP, Integrante y Coordinador de la Comisión de Grado Académico y Evaluador externo de proyectos de investigación para numerosas Universidades Nacionales.

Available online at
www.sag.org.ar/jbag

El Ing. Arturi predicaba con su ejemplo la honestidad, la humildad, el compromiso con la Institución y la hombría de bien. Como su discípula recordaré su respeto y dedicación a mi formación, sus meticulosos análisis estadísticos y sus devoluciones escritas en lápiz. Además de su confianza en sus becarias y en el trabajo por ellas realizado, deja el recuerdo como platense de pura cepa, de un hombre amoroso con su esposa e hijo; me quedan sus charlas futboleras al vaivén del desempeño de su querido Gimnasia y Esgrima de La Plata, charlas compartidas con mis colegas Mónica Collado y Mónica Aulicino con quienes también compartí sus minuciosos y detallados cronogramas al inicio de cada cursada. Perdurará en nosotras el recuerdo de su gusto por la ópera y su refinado sentido del humor. Hombre de pocas palabras, pero siempre las palabras precisas y necesarias, de enorme generosidad y pensamientos claros. Siempre lo vamos a recordar y a extrañar.

Con todo mi respeto, afecto e infinito agradecimiento,

María Victoria García

BAG | Journal of Basic & Applied Genetics