(Formerly MENDELIANA)

BAG

Journal of Basic & Applied Genetics

December 2020
Volumen XXXI
No. 2
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.2020.31.02

Journal of the
Argentine Society
of Genetics

BAG | Journal of Basic & Applied Genetics

V. XXXI - No.2

December 2020

Included in:

Cited by:

Comité Editorial

Imágen de tapa:

Acca sellowiana, "falso guayabo" o "guayabo del país"

María Lis Echeverría

Contenidos

BAG | Journal of Basic & Applied Genetics

Vol XXXI (2): 7-11; December 2020

Journal of the Argentine Society of Genetics

FIRST KARYOTYPE REPORT ON *Colocasia oresbia*: A COMPARATIVE CYTOGENETIC STUDY BETWEEN TWO VARIETIES

PRIMER REPORTE DEL CARIOTIPO DE *Colocasia oresbia*: UN ESTUDIO CITOGENÉTICO COMPARATIVO ENTRE DOS VARIEDADES

Begum Kazi N.[1], Dash Chandan K.[2], Sultana Syeda S.[2]

[1] Department of Botany, Jagannath University, Dhaka-1100, Bangladesh.

[2] Cytogenetics Laboratory, Department of Botany, University of Dhaka, Dhaka-1000, Bangladesh.

Corresponding author:
Syeda Sharmeen Sultana
syedasharmeensultana@du.ac.bd

ORCID 0000-0002-6196-0844

Cite this article as:
Begum Kazi N., Dash Chandan K., Sultana Syeda S. 2020. FIRST KARYOTYPE REPORT ON *Colocasia oresbia*: A COMPARATIVE CYTOGENETIC STUDY BETWEEN TWO VARIETIES. BAG. Journal of Basic and Applied Genetics XXXI (2): 7-11.

Received: 03/15/2020
Revised version received: 06/06/2020
Accepted: 06/18/2020
General Editor: Elsa Camadro
DOI: 10.35407/bag.2020.31.02.01
ISSN online version: 1852-6233

ABSTRACT

Karyotypes of two *Colocasia oresbia* botanical varieties from Bangladesh were analyzed and compared with orcein, chromomycin A_3 (CMA) and 4´-6 diamidino-2-phenylindole (DAPI). Both varieties had 2n= 2x = 26 chromosomes (karyotypic formula: 20m+ 6sm) and a pair of satellites each. Total chromosome length was 144.18 ± 2.45 µm in *C. oresbia* var. *oresbia* and 133.02 ± 2.75 µm in *C. oresbia* var. *stolonifera*. The karyotype of *Colocasia oresbia* var. *oresbia* is 2A whereas that of *C. oresbia* var. *stolonifera* is 1A. Six CMA and four DAPI bands were observed in *C. oresbia* var. *oresbia* and eight CMA and six DAPI bands in *C. oresbia* var. *stolonifera*. However, in these two morphologically distinct *C. oresbia* varieties of two different ecological zones, the same somatic chromosome number, diversification in various karyotypic parameters and CMA/DAPI-banding patterns were observed. In addition to taxonomic characters, the studied karyotype features will contribute to the characterization of these two *C. oresbia* varieties and to establish a base for future research.

Key words: chromosome banding; CMA; DAPI; Karyotype.

RESUMEN

Se analizaron y compararon los cariotipos de dos variedades botánicas de *Colocasia oresbia* de Bangladesh con orceína, chromomicina A_3 (CMA) y 4-6 diamidino-2-phenilindol (DAPI). Ambas variedades presentaron 2n= 2x = 26 cromosomas (fórmula cariotípica: 20m+ 6sm) y un par de satélites cada una. La longitud total de cromosomas fue 144,18 ± 2,45 µm en *C. oresbia* var. *oresbia* y 133.02 ± 2.75 µm en *C. oresbia* var. *stolonifera*. El cariotipo de *Colocasia oresbia* var. *oresbia* es 2ª, y 1ª el de *C. oresbia* var. *stolonifera*. Se observaron seis bandas CMA y cuatro DAPI en *C. oresbia* var. *oresbia* y ocho bandas CMA y seis DAPI en *C. oresbia* var. *stolonifera*. Sin embargo, en estas dos variedades morfológicamente distintivas de *C. oresbia* de dos zonas ecológicas diferentes se observó el mismo número cromosómico somático, diversificación en varios parámetros cariotípicos y en patrones de bandeo CMA/DAPI. En adición a los caracteres taxonómicos, las características de los cariotipos estudiados contribuirán a la caracterización de estas dos variedades de *C. oresbia* y a establecer una base para futuras investigaciones.

Palabras clave: bandeo cromosómico; CMA; DAPI; cariotipo

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

INTRODUCTION

The genus *Colocasia* Schott belonging to the Araceae family, comprises about 20 species over the world (Li and Boyce, 2010). A total of nine of these species has been reported for Bangladesh so far, such as *C. affinis* Schott, *C. esculenta* (L.) Schott, *C. fallax* Schott, *C. gigantea* (Blume) Hook. f., *C. heterochroma* H. Li *et* Z.X. & Wei, *C. lihengiae* C.L. Long *et* K.M. Liu, *C. mannii* Hook. f., *C. oresbia* A. Hay and *C. virosa* Kunth. (Ara and Hassan, 2019). This genus is popular because it is edible and has medicinal, ornamental and cultural importance. Ara and Hassan (2019) reported and differentiated two varieties of *C. oresbia* from Bangladesh *viz. C. oresbia* A. Hay var. *stolonifera* H. Ara & M.A. Hassan, var. nov. and *C. oresbia* A. Hay var. *oresbia* based on several prominent morphological features. In fact, most species of this genus are morphologically distinct although the morphological features of a few of them are very confusing. In those cases, karyo-morphological information can open a new direction for evaluating the relationship among them. The nature and degree of karyotype differences obtained from conventional and fluorescent banding techniques could be useful to discuss plant phylogeny. In addition, the cytogenetical information will be useful for development of successful breeding programs in this crop. So far *C. oresbia* is unexplored cytogenetically. Therefore, in the present study, a combination of morphological and cytogenetical analyses with orcein, chromomycin A_3 (CMA) and 4´-6 diamidino-2-phenylindole (DAPI) were carried out for the first time to present karyotype data from two varieties of *C. oresbia viz. Colocasia oresbia* var. *stolonifera* and *Colocasia oresbia* var. *oresbia* to determine chromosomal relationships among them.

MATERIALS AND METHODS

Two varieties of *Colocasia oresbia viz. C. oresbia* A. Hay var. *stolonifera* H. Ara & M.A. Hassan, var. nov. and *C. oresbia* A. Hay var. *oresbia* were studied. *Colocasia oresbia* var. *oresbia* was collected from Chittagong, Cox's Bazar, Khagrachari, Moulvibazar, Rangamati, Kaptai, Rajbari area, Shubalong and Dhaka (flat regions) of Bangladesh whereas *C. oresbia* var. *stolonifera* was found and collected only from Rangamati district (hilly regions), Bangladesh. For cytogenetic investigation, healthy roots of ten individuals of each variety were collected and pretreated with 2 mM 8-hydroxyquinoline for 3 h at room temperature followed by 15 min fixation in 45% acetic acid at 4 °C, then hydrolyzed in a mixture of 1 N HCl and 45% acetic acid (2:1 v/v) at 60 °C for 3 min. The root tips were stained and squashed in 1% aceto-orcein. For CMA- and DAPI-banding, Alam and Kondo's (1995) method was used with slight modifications. Slides were observed under a Nikon (Eclipse 50i) fluorescent microscope with a blue violet (BV) filter cassette for CMA and an ultraviolet (UV) one for DAPI-banding. CMA binds with GC (Guanine-Cytosine)-rich repetitive sequences of the genome expressing yellow fluorescence, and DAPI binds to AT (Adenine-Thymine)-rich repeats giving a characteristic blue color (Schweizer, 1976).

For every staining, at least 50 cells were observed in each variety. The idiograms were made on the basis of chromosome size in decreasing order. Levan *et al.* (1964) was followed for determining centromeric type of chromosomes. Karyotype asymmetry index (AI) was also calculated to determine the degree of karyotype heterogeneity (Paszko, 2006).

RESULTS AND DISCUSSION

Morphological investigation

The two studied varieties of *Colocasia oresbia* show some prominent morphological dissimilarities. *Colocasia oresbia* var. *stolonifera* has stolons, which are absent in *C. oresbia* var. *oresbia*. They also show differences in inflorescence formation: the inflorescence of *C. oresbia* var. *stolonifera* is normally formed in group of up to 3 but in *C. oresbia* var. *oresbia* inflorescence occurred in group of up to 8 (never less than 4).

Somatic chromosome number and karyotype analysis

This present study provides detailed chromosomal information of *C. oresbia* for the first time. The two varieties are found to possess 2n=26 chromosomes (Figure 1A, B; Table 1). Somatic chromosome numbers 2n=28 and 42 have been reported for most of the studied species of this genus. Besides, some infrequent records such as 2n=26 in *C. gigantea* and *C. esculenta*, 2n=38 in *C. antiquorum*, and 2n=56 (tetraploid) in *C. esculenta* have also been reported (Wang *et al.*, 2017). Previous literature has stated the basic chromosome number of *Colocasia* is x=14 since most of the species belonging to this genus have 2n=28 chromosomes (Yang *et al.*, 2003). Other researchers have suggested that chromosomal variation regarding ploidy levels and aneuploidy occurred frequently in this genus (Fedorov, 1974; Kumar and Subramanian, 1979; Cao and Long, 2004; Huang *et al.*, 2012). Moreover, the presence of euploid and aneuploid cytotypes in different species represents inconstancy in the basic chromosome number. The reported basic chromosome numbers are x=13, 14, 19, present in 2x, 3x, and 4x cytotypes (Wang *et al.*, 2017). Previous studies concerning genus *Colocasia* showed that x=14 should be considered as ancestral basic chromosome number (Yang *et al.*, 2003; Wang *et al.*, 2017). In two varieties of *C. oresbia* of the present study, the basic chromosome

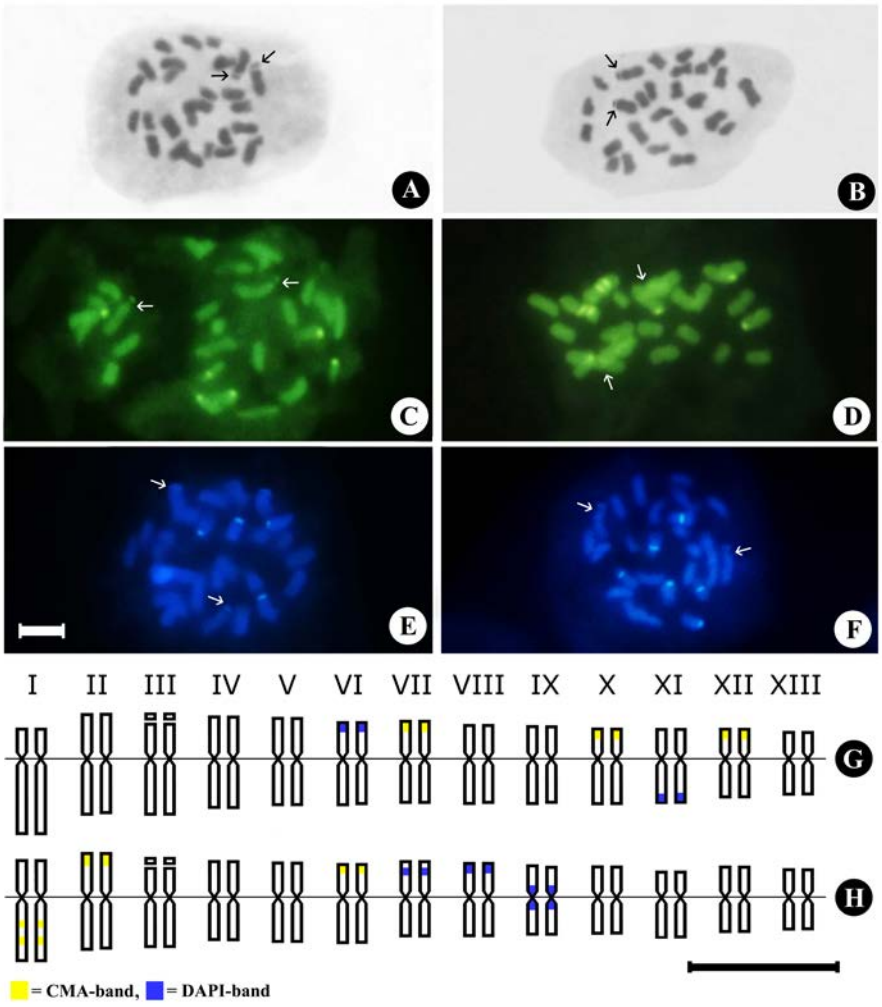

Figure 1. Metaphase chromosomes and idiograms of two *Colocasia oresbia* varieties. A. Orcein-stained mitotic metaphase of *C. oresbia* var. *oresbia*, B. Orcein-stained mitotic metaphase of C. *oresbia* var. *stolonifera*, C. CMA-stained mitotic metaphase of C. *oresbia* var. *oresbia*, D. CMA-stained mitotic metaphase of C. *oresbia* var. *stolonifera*, E. DAPI-stained mitotic metaphase of C. *oresbia* var. *oresbia*, F. DAPI-stained mitotic metaphase of C. *oresbia* var. *stolonifera*, G. Idiogram of C. *oresbia* var. *oresbia*, H. Idiogram of C. *oresbia* var. *stolonifera*. Arrows indicate satellites. Bars=10 µm.

number is x=13. Other previously reported basic chromosome numbers of x=13 and x=19 indicate that these two basic numbers probably originated from x=14 by secondary modifications (Leong-Škorničková *et al.*, 2007).

Both varieties of *C. oresbia* display relatively homogeneous karyotype arrangement with metacentric and submetacentric chromosomes with a KF of 20m + 6sm, and have one pair of satellites in chromosome pair III (Figure 1G, H). However, these two varieties show differences in other karyotype parameters. *Colocasia oresbia* var. *oresbia* and *C. oresbia* var. *stolonifera* has TCL of 144.18 ± 2.45 µm and 133.02 ± 2.75 µm, respectively (Table 1). The ACL is lower in *C. oresbia* var. *stolonifera* (5.12 µm) than *C. oresbia* var. *oresbia* (5.55 µm). The RCL is 4.23-7.02 µm in *C. oresbia* var. *oresbia* and 4.05-6.75 µm in *C. oresbia* var. *stolonifera*. The RL is 2.93-4.87% in *C. oresbia* var. *oresbia* whereas 3.04-5.07% in *C. oresbia* var. *stolonifera*.

When evolutionary positions are taken into consideration in relation to the karyotypic nature, symmetric karyotypes are usually regarded as primitive and asymmetrical as advanced, since karyotype asymmetry can be considered to be the dynamic force behind speciation (Stebbins, 1971). Furthermore, a higher AI value represents more asymmetric karyotypes (Paszko, 2006). The studied asymmetry index of karyotype reveals that the karyotype of *C. oresbia* var. *oresbia* is more asymmetric than the karyotype of *C. oresbia* var. *stolonifera*. Thus, *C. oresbia* var. *oresbia* is more advanced from an evolutionary point of view. Chromosome number and size along with karyotypic

Table 1. Comparative cytogenetical analysis of two *Colocasia oresbia* varieties.

Features	C. *oresbia* var. *oresbia*	C. *oresbia* var. *stolonifera*
Orcein staining		
2n	26	26
No. of satellites	2	2
KF	20m+6sm	20m+6sm
TCL (µm)	144.18±2.45	133.02±2.75
RCL (µm)	4.23–7.02	4.05–6.75
RL (%)	2.93–4.87	3.04–5.07
ACL (µm)	5.55	5.12
AI	1.65	1.20
Karyotype category	2A	1A
CMA		
No. of bands	6	8
Total banded region (µm)	7.65±0.53	9.03±0.74
Banded region (%)	5.31	6.79
DAPI		
No. of bands	4	6
Total banded region (µm)	4.50±0.78	7.20±0.68
Banded region (%)	3.12	5.41

2n=Somatic chromosome number; KF=Karyotypic formula; TCL=Total chromosome length; RCL=Range of chromosomal length; RL=Relative length of chromosome; ACL=Average chromosome length; AI=Asymmetry index of karyotype.

features are subjected to evolutionary change (Lavia *et al.*, 2009). Chromosome evolution can take place either by increasing or decreasing chromosomal length (Brandham and Doherty, 1998; Martel *et al.*, 2004). In this case, the total length of the chromosome complements increase in the course of evolution, since both varieties have similar 2n numbers and karyotype formula. *Colocasia oresbia* var. *oresbia* and *C. oresbia* var. *stolonifera* have 2A and 1A karyotypes, respectively, which also correlate with the asymmetric index (Table 1).

Fluorescent banding

Each variety exhibited distinct CMA-banding pattern (Figure 1C, H; Table 1). Six and eight CMA-bands were found in *C. oresbia* var. *oresbia* and *C. oresbia* var. *stolonifera*, respectively, with 5.31% GC-rich repeats in *C. oresbia* var. *oresbia* and 6.79 % in *C. oresbia* var. *stolonifera*. Six chromosomes (pairs VII, X and XII) of *C. oresbia* var. *oresbia* and four chromosomes (pairs II and VI) of *C. oresbia* var. *stolonifera* exhibited terminal CMA-bands. In addition, two chromosomes (pair I) of *C. oresbia* var. *stolonifera* had a peculiar CMA-banding pattern. In this variety, two chromosomes possess a pair of interstitial bands that may be used as chromosome markers. Four and six DAPI-bands were observed in *C. oresbia* var. *oresbia* and *C. oresbia* var. *stolonifera*, respectively. The DAPI-banded regions are 3.12% and 5.41% of the total chromosome complements in *C. oresbia* var. *oresbia* and *C. oresbia* var. *stolonifera*, respectively. Four terminal DAPI-bands (pairs VI and XI) in *C. oresbia* var. *oresbia* and two terminal DAPI-bands (pair VIII) in *C. oresbia* var. *stolonifera* were found. In addition, two centromeric (pair IX) and two intercalary DAPI-bands (pair VII) were also observed in *C. oresbia* var. *stolonifera* (Figure 1G, H). The mentioned findings suggest that each variety has a characteristic CMA and DAPI banded pattern with different number, location, total banded regions and percentage of GC- and AT-rich segments. Most of the bands are present at the terminal regions of the short arms of the respective chromosomes (Figure 1C, H). The presence of terminal bands indicated the tendency of accumulating GC- and AT-rich repetitive sequences at the chromosomal ends. Even though both varieties have the same chromosome number, diversification in karyotypic features and reshuffling of GC- and AT-rich

banded regions were observed. The variation in karyotype indices and fluorescence banding patterns may be the result of inversions, deletions or unequal translocations, among other chromosomal aberrations. The diversity in karyotypes of these two varieties may have arisen due to the exposure to different environmental conditions.

This research is the first cytogenetical report for *C. oresbia*. The findings of the present study would be useful for future breeding programs and a contribution to the systematics of the species.

REFERENCES

Alam S.S., Kondo K. (1995) Differential staining with orcein, Giemsa, CMA and DAPI for comparative chromosome study of 12 species of Australian *Drosera* (Droseraceae). American J. Bot. 82: 1278-1286.

Ara H., Hassan M.A. (2019) Four new varieties of the family Araceae from Bangladesh. Bangladesh J. Plant Taxon. 26: 13-28.

Brandham P.E., Doherty M.J. (1998) Genome size variation in the Aloaceae, an angiosperm family displaying karyotypic orthoselection. Ann. Bot. 82: 67-73.

Cao L.M., Long C.L. (2004) Chromosome numbers of eight *Colocasia* taxa and karyotypes of five species occurring in China. Acta Bot. Yunnan 26: 310-316.

Fedorov A.N.A. (1974) Chromosome Numbers of Flowering Plants. Otto Koettz Science Publishers, N-624 W, Germany.

Huang X.F., Ke W.D., Liu Y.M., Ye Y.Y., Li S.M., Peng J., Liu Y.P., Li F. (2012) Chromosomal ploidy identification of Taro (*Colocasia*) germplasm resources. China Veget. 6: 42-46.

Kumar V., Subramaniam B. (1987) Chromosome atlas of flowering plants of the Indian subcontinent: Dicotyledons. Botanical Survey of India, Kolkata.

Lavia G.I., Ortiz A.M., Fernández A. (2009) Karyotypic studies in wild germplasm of *Arachis* (Leguminosae). Genet. Resour. Crop. Evol. 56: 755-764.

Leong-Škorničková J., Šída O., Jarolímová V., Sabu M., Fér T., Trávníček P., Suda J. (2007) Chromosome numbers and genome size variation in Indian species of *Curcuma* (Zingiberaceae). Ann. Bot. 100: 505-526.

Levan A., Fredga K., Sandberg A.A. (1964) Nomenclature for centromeric position on chromosomes. Hereditas 52: 201-220.

Li H., Boyce P.C. (2010) *Colocasia*. In: Wu Z.Y., Raven P.H., Hong D.Y. (Eds.) Flora of China. Vol. 23. Science Press, St. Louis, pp. 73-75.

Martel E., Poncet V., Lamy F., Siljak-Yakovlev S., Lejeune B., Sarr A. (2004) Chromosome evolution of *Pennisetum* species (Poaceae): Implications of ITS phylogeny. Plant Syst. Evol. 249: 139-149.

Paszko B. (2006) A critical review and a new proposal of karyotype asymmetry indices. Plant Syst. Evol. 258: 39-48.

Schweizer D. (1976) Reverse fluorescent chromosome banding with Chromomycin and DAPI. Chromosoma 58: 307-324.

Stebbins G.L. (1971) Chromosomal Evolution in Higher Plants. University Park Press, Baltimore.

Wang G.Y., Zhang X.M., Qian M., Hu X.Y., Yang Y.P. (2017) Chromosome number and genome size variation in *Colocasia* (Araceae) from China. J. Plant Res. 130: 989-997.

Yang Z.Y., Yi T.S., Li H., Gong X. (2003) A cytological study on three species of *Colocasia* (Araceae) from Yunnan. Caryologia 56: 323-327.

ACKNOWLEDGEMENT

The authors are grateful to Dr. Hosene Ara, Ex-Director (In-charge), Bangladesh National Herbarium, Chiriakhana Road, Mirpur-1, Dhaka-1216, for providing the research materials and necessary information about *Colocasia oresbia*.

BAG | Journal of Basic & Applied Genetics

Vol XXXI (2): 13-25; December 2020

Journal of the Argentine Society of Genetics

SEQUENCE ANALYSIS SUGGESTS POSITIVE SELECTION ON THE BOVINE PRODYNORPHIN GENE

ANÁLISIS DE SECUENCIAS GENÓMICAS SUGIEREN QUE EL GEN DE LA PRODINORFINA ESTÁ BAJO SELECCIÓN POSITIVA EN BOVINOS

Suqueli García M.F.[1], Castellote M.A.[2], Corva P.M.[1]

[1] Facultad de Ciencias Agrarias, Universidad Nacional de Mar del Plata, Unidad Integrada Balcarce, C.C. 276, 7620 Balcarce, Argentina.

[2] Laboratorio de Agrobiotecnología, EEA Balcarce. Instituto Nacional de Tecnología Agropecuaria, Unidad Integrada Balcarce, C.C. 276, 7620 Balcarce, Argentina.

Corresponding author:
Pablo Marcelo Corva
corva.pablo@inta.gob.ar

ORCID 0000-0002-0660-3115

Cite this article as:
Suqueli García M.F., Castellote M.A., Corva P.M. 2020. SEQUENCE ANALYSIS SUGGESTS POSITIVE SELECTION ON THE BOVINE PRODYNORPHIN GENE. BAG. Journal of Basic and Applied Genetics XXXI (2): 13-25.

Received: 05/05/2020
Revised version received: 07/15/2020
Accepted: 07/24/2020
General Editor: Elsa Camadro
DOI: 10.35407/bag.2020.31.02.02
ISSN online version: 1852-6233

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

ABSTRACT

Dynorphin A is an endogenous opioid peptide that is part of the KNDy system in the hypothalamus of mammals. This peptide acts as an inhibitor of the GnRH pulse generation, thus regulating the onset of puberty and reproductive cycles. The *PDYN* gene encodes the propeptide Prodynorphin, the precursor of Dynorphin A. Despite its physiological relevance, *PDYN* has not emerged as a candidate gene associated with puberty in genomic association studies conducted in cattle. The present work aimed to search for signatures of selection on the *PDYN* gene among cattle breeds. To this, the whole genome sequences from 57 samples of ten cattle breeds were used. The samples were grouped based on breed selection history and their productive differences, particularly in terms of sexual precocity. The population structure was analyzed using Principal Component Analyses. To evidence recent selection processes, neutrality tests, such as Tajima's D and Fu & Li's F* and D* were performed in defined functional regions of *PDYN*. The putative promoter of *PDYN* showed a population structure that is in agreement with the criteria considered to make the groups. In that region, neutrality tests were consistently negative and resulted in statistically significant for the dairy breeds. Also, these breeds exhibited less variability in the haplotype analyses than the others. The results presented here suggest that regulatory regions of *PDYN* could be under positive selection, particularly in dairy breeds.

Key words: reproduction; KNDy neurons; Dynorphin; signatures of selection.

RESUMEN

Dinorfina A es un péptido opioide endógeno que forma parte del sistema KNDy en el hipotálamo de mamíferos. Este péptido actúa como inhibidor de la generación de los pulsos de GnRH, regulando así el inicio de la pubertad y los ciclos reproductivos. El gen *PDYN* codifica el propéptido Prodinorfina, precursor de Dinorfina A. A pesar de su relevancia fisiológica, *PDYN* no ha surgido como gen candidato asociado a pubertad en estudios de asociación genómicos en bovinos. El presente trabajo tuvo como objetivo buscar huellas de selección en el gen *PDYN* entre diferentes razas bovinas. Para alcanzarlo se utilizaron secuencias genómicas de 57 muestras de diez razas bovinas. Las muestras fueron agrupadas considerando la historia de selección y las diferencias productivas entre razas, particularmente en términos de precocidad sexual. La estructura poblacional fue analizada usando análisis de componentes principales. Para evidenciar procesos de selección recientes se realizaron pruebas de neutralidad, tales como D de Tajima y F* y D* de Fu & Li, en diferentes regiones funcionales de *PDYN*. El promotor putativo de *PDYN* mostró una estructura poblacional que es consistente con los criterios usados para agrupar las razas. En esa región, las pruebas de neutralidad fueron consistentemente negativas y estadísticamente significativas en las razas lecheras. Además, estas razas también exhibieron menor variabilidad en los análisis de haplotipos que las demás razas. Los resultados presentados aquí sugieren que regiones regulatorias de *PDYN* estarían bajo selección positiva, particularmente en razas bovinas lecheras.

Palabras clave: reproducción; neuronas KNDy; Dinorfina; huellas de selección.

INTRODUCTION

Reproductive efficiency is one of the most valued aspects of animal production. In the particular case of beef and dairy cattle, reproduction-associated traits, especially of females, are included in the selection objective of the most modern breeds and they have received much attention in genetic and genomic studies.

Starting with puberty and then cyclically repeated along productive life, the reproductive process in females requires the release of the gonadotropin-releasing hormone (GnRH) from the hypothalamus, which in turn is needed for the secretion of gonadotropins from the anterior pituitary gland (Amstalden and Williams, 2015; Atkins *et al.*, 2013). Research devoted to understanding the mechanisms underlying the onset of puberty and the regulation of subsequent reproductive cycles in different mammal species contributed to building the "KNDy hypothesis" (Smith *et al.*, 2014). According to this hypothesis three peptides, Kisspeptin, Neurokinin B, and Dynorphin A, are co-expressed in the same group of neurons in specific regions of the hypothalamus (Goodman *et al.*, 2007). These peptides would have the role of a "pulse generator" that drives GnRH secretion. Neurokinin B and Dynorphin A have stimulatory and inhibitory effects on Kisspeptin expression, respectively. Moreover, these neurons also express the Neurokinin B receptor and the Dynorphin A receptor, but not the Kisspeptin receptor that is expressed in GnRH neurons (Weems *et al.*, 2018). The regulatory system involving the KNDy neurons would be sensitive to sex steroid feedback and different regulatory cues (*e.g.* energy balance, lactation, photoperiod, stress) but the specific mechanisms involved are not clear (Lehman *et al.*, 2010).

The role of Dynorphin A in the KNDy neurons is consistent with previous evidence showing an inhibitory effect of endogenous opioid peptides on reproduction (Malven, 1986). Suckling increases levels of endogenous opioid peptides in the brain, making the hypothalamus more sensitive to the negative feedback from estrogen and decreasing the production of GnRH (Squires, 2010).

Despite the role of Kisspeptin, Neurokinin B and Dynorphin A in the regulation of the reproductive cycle, the corresponding genes (*KISS1*, *TAC3*, and *PDYN*) and their receptors (*KISS1R*, *TACR3*, and *OPRK1*) have not been analyzed as functional or positional candidates in cattle. The only known exception is the Dynorphin receptor (*OPRK1*), proposed as a candidate gene associated with sexual precocity in a GWAS for sexual precocity in Nellore cattle from Brazil (Irano *et al.*, 2016).

From the standpoint of animal breeding, the dissection of genetic variation at the level of the KNDy system could be of help to understand differences in reproductive performance among cattle breeds, and in turn, to improve management strategies. As an example, we have evaluated the reproductive efficiency of Angus and Argentinean Creole females (Corva *et al.*, 1995). The Creole is a local beef breed that has been much less selected than Angus, but it is appreciated due to its rusticity and some other favorable traits such as calving ease. The superiority of Angus over Creole in reproductive performance was noticed starting with age at puberty (Pardo *et al.*, 2018) and other differences between the two breeds were detected in subsequent reproductive cycles (Corva *et al.*, 1995). While body weight change during the breeding season was a key factor defining pregnancy rate in Angus females, lactation and the interaction with the offspring was the most limiting factor in Creole females. The marked difference between European (*Bos taurus*) and Indicine (*Bos indicus*) breeds in sexual precocity is another well-known example of genetic variation in reproductive performance (Freetly *et al.*, 2011; Laster *et al.*, 1979) even when were evaluated in similar restrictive environmental conditions (Meirelles *et al.*, 1994; Ferraz Jr. *et al.*, 2018).

Taking together the evidences about the inhibitory effect of endogenous opioid peptides on reproduction and the recently proposed role of Dynorphin A in KNDy neurons, we hypothesized that *PDYN*, the gene that codes for the propeptide Prodynorphin, which is cleaved to produce several opioid peptides including Dynorphin A (Day et al., 1998, Figure 1), could be under selection in cattle. This selection probably resulting in the relaxation of the effects of different cues that convey to the hypothalamus to regulate the onset of reproductive activity. To test this hypothesis, we compared the sequence of the *PDYN* gene from different cattle breeds, taking advantage of the availability of whole genome sequences.

MATERIALS AND METHODS

Retrieval and processing of PDYN genomic sequences

The genomic sequence of the Bovine Reference Genome (ARS-UCD1.2) corresponding to a region of BTA13 spanning the *PDYN* gene (GenBank accession: NC_037340.1: 53,130,235 - 53,148,819 bp) was downloaded from the NCBI site (http://www.ncbi.nlm.nih.gov). This DNA sequence was used as a reference to retrieve the corresponding genomic sequence of *Bos indicus* (GenBank accession: NC_032662.1: 53,688,047 - 53,707,634 bp) and also the genomic sequences of a panel of individuals from selected bovine breeds as it is described below. The sequenced genomes of selected animals had already been aligned to the Bovine Reference Genome UMD 3.1.1 (Merchant *et al.*, 2014; Zimin *et al.*, 2009). Sequence alignment confirmed that apart from a change in relative coordinates, the region of interest resulted in identical between UMD 3.1.1 and ARS-UCD1.2, the latest version of the Bovine Reference Genome (wwww.bovinegenome.org).

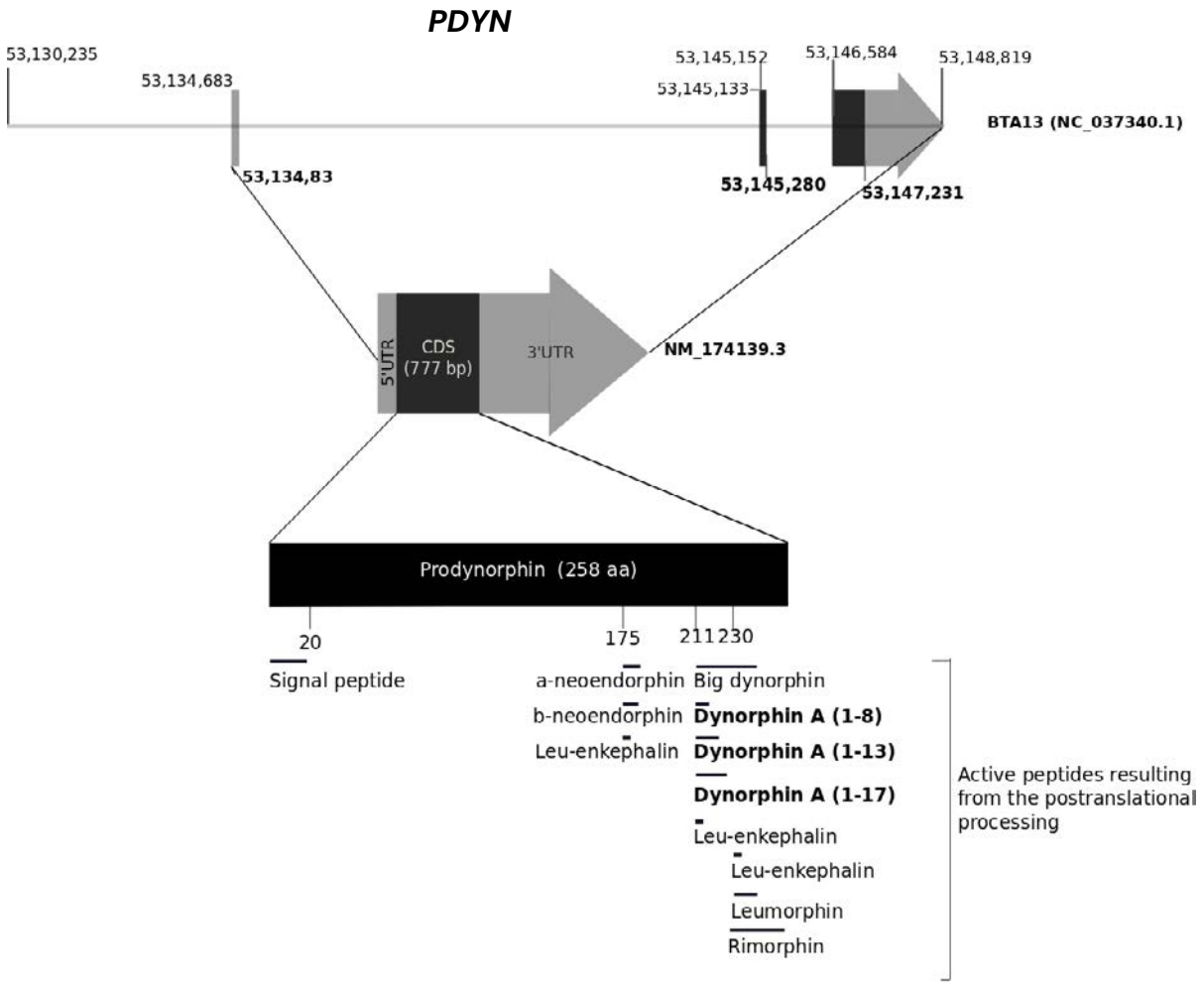

Figure 1. Structure of the bovine *PDYN* gene (Top), the corresponding transcript (Genbank sequence: NM_174139.3) (Middle) and the translated propeptide (Bottom). The different functional peptides resulting from proteolytic cleavage by the enzyme Prohormone convertase 2 are also indicated in the figure.

The whole panel of genomic sequences used for the analyses included: The reference *Bos taurus* and *Bos indicus* genomes, corresponding to the Hereford and Nellore breeds respectively; one individual representing extinct wild cattle (Auroch, *Bos primigenius*, Park *et al.*, 2015) and individuals from the "1000 Bull Genomes Project" (http://www.1000bullgenomes.com/) or the U.S. Meat Animal Research Center (USMARC) Beef Cattle Diversity Panel 2.9 (MBCDPv2.9, Heaton *et al.*, 2016). Selected breeds were: Holstein (n=13), Jersey (n=11), Angus (n=10), Limousine (n=6), Corriente (n=4), Longhorn (n=3) and Brahman (n=7). In the first experiment, the average sequencing depth was 8.3X (Daetwyler *et al.*, 2014) while in the experiment of Heaton *et al.* (2016) it was 14.8X. Aligned genomic sequences (in bam format) were downloaded from the Sequence Read Archive (SRA) division of NCBI (http://www.ncbi.nlm.nih.gov/sra) and the USMARC website (https://www.ars.usda.gov/plains-area/clay-center-ne/marc/wgs/bovref/), respectively. A detailed list of these genomic sequences is presented in Table 1. Based on breed selection history and well known productive differences among bovine biotypes (particularly in terms of sexual precocity; Diskin and Kenny, 2014; Sartori *et al.*, 2016) four contrasting groups were defined: highly selected *Bos taurus* dairy (Holstein, Jersey, n=24) and beef breeds (Angus, Limousine, Hereford, n=17); less selected -hardy- *Bos taurus* breeds (Corriente, Longhorn, Auroch, n=8) and zebu breeds (Brahman, Nellore, n=8). These groups are hereinafter identified as "Dairy", "Beef", "Less selected" and "Zebu", respectively.

Each genomic sequence in bam format was searched for variant sites using BCFtools software v1.5 (Li, 2011) with the command: "bcftools mpileup -f sequence.fasta input.bam | bcftools call -m -Ob | tabix | bcftools query -f '%POS\t%QUAL\t%ALT[\t%IUPACGT]\n' –o output.txt" where "sequence.fasta" was the sequence of the Bovine Reference Genome UMD 3.1.1 used as a reference in the original alignment. Variants with a Phred quality score below 20 (base call accuracy above 99%) were manually discarded and a consensus sequence was obtained for each sample in fasta format using UNIX

Table 1. Genomic sequences used to characterize the bovine Prodynorphin gene. Each retrieved sequence was aligned to the region of BTA13 spanning the *PDYN* gene (Bovine Reference Genome ARS-UCD1.2, GenBank accession: NC_037340.1: 53,129,235-53,148,819).

ID	Group	Breed	Whole genome coverage	*PDYN* region coverage
NC_037340.1	Beef	Hereford[1]		
SRR1262614	Beef	Angus[2]	10.0X	6.7X
SRR1262615	Beef	Angus[2]	10.0X	7.5X
SRR1262616	Beef	Angus[2]	9.5X	6.6X
SRR1262617	Beef	Angus[2]	9.8X	6.9X
SRR1262618	Beef	Angus[2]	10.0X	7.3X
SRR1262619	Beef	Angus[2]	10.0X	7.0X
SRR1262620	Beef	Angus[2]	11.0X	7.5X
SRR1262621	Beef	Angus[2]	6.7X	4.7X
SRR1262622	Beef	Angus[2]	8.0X	5.4X
SRR1262623	Beef	Angus[2]	8.5X	6.5X
19989863	Beef	Limousine[3]	10.7X	12.8X
19999833	Beef	Limousine[3]	11.9X	12.9X
19999834	Beef	Limousine[3]	10.6X	10.9X
19999835	Beef	Limousine[3]	11.5X	10.5X
19999837	Beef	Limousine[3]	11.2X	8.7X
19999838	Beef	Limousine[3]	10.4X	8.8X
SRR1262783	Dairy	Holstein[2]	5.7X	3.6X
SRR1188706	Dairy	Holstein[2]	16.0X	12.8X
SRR1262533	Dairy	Holstein[2]	6.6X	3.4X
SRR1262536	Dairy	Holstein[2]	6.4X	4.5X
SRR1262538	Dairy	Holstein[2]	6.3X	4.6X
SRR1262539	Dairy	Holstein[2]	6.1X	3.2X
SRR1262660	Dairy	Holstein[2]	5.6X	4.1X
SRR1262661	Dairy	Holstein[2]	13.0X	12.5X
SRR1262662	Dairy	Holstein[2]	5.7X	5.24X
SRR1262663	Dairy	Holstein[2]	10.0X	8.1X
SRR1262664	Dairy	Holstein[2]	14X	10.7X
SRR1262782	Dairy	Holstein[2]	12.0X	9.3X
SRR1262785	Dairy	Holstein[2]	15.0X	12.4X

ID: File identification in the original database. Dairy: Highly selected dairy breeds (*Bos taurus*); Beef: Highly selected beef breeds (*Bos taurus*); Less selected: less selected breeds (*Bos taurus*), includes the Auroch (Bos primigenius); Zebu: zebu breeds (*Bos indicus*). All samples were aligned to the genomic region of BTA13 spanning the *PDYN* gene and only this region of each sample was used for the analyses. [1]Bovine Reference Genome ARS-UCD1.2, GenBank accession: NC_037340.1: 53,130,235-53,148,819). [2]"1000 Bull Genome Project" Run 2, (http://www.1000bullgenomes.com/). [3]U.S. Meat Animal Research Center (USMARC) Beef Cattle Diversity Panel 2.9 (MBCDPv2.9, Heaton et al., 2016). [4]*Bos primigenius* archeological sample (Park et al., 2015). [5]*Bos indicus* Reference Genome Bos_indicus_1.0 (GenBank accession: NC_032662.1). *Average depth of coverage obtained from their respective original database. #Average depth of coverage calculated by using Mosdepth software v 0.2.9 (Pedersen and Quinlan, 2018).

Table 1. (cont.) Genomic sequences used to characterize the bovine Prodynorphin gene. Each retrieved sequence was aligned to the region of BTA13 spanning the *PDYN* gene (Bovine Reference Genome ARS-UCD1.2, GenBank accession: NC_037340.1: 53,129,235-53,148,819).

ID	Group	Breed	Whole genome coverage*	*PDYN* region coverage#
SRR1262789	Dairy	Jersey[2]	8.8X	6.0X
SRR1262790	Dairy	Jersey[2]	8.3X	6.2X
SRR1262791	Dairy	Jersey[2]	9.3X	7.0X
SRR1262792	Dairy	Jersey[2]	6.9X	5.4X
SRR1262793	Dairy	Jersey[2]	14.0X	10.7X
SRR1262794	Dairy	Jersey[2]	6.4X	5.1X
SRR1262795	Dairy	Jersey[2]	7.1X	6.5X
SRR1262796	Dairy	Jersey[2]	6.1X	3.7X
SRR1262797	Dairy	Jersey[2]	11.0X	8.9X
SRR1262798	Dairy	Jersey[2]	13.0X	11.3X
SRR1262803	Dairy	Jersey[2]	12.0X	12.0X
SRR2465682	Less selected	Auroch[4]	8.8X	6.9X
19202900	Less selected	Corriente[3]	11.3X	10.1X
19202901	Less selected	Corriente[3]	10.1X	10.1X
19202902	Less selected	Corriente[3]	12.9X	11.9X
19202903	Less selected	Corriente[3]	10.5X	11.2X
19999870	Less selected	Longhorn[3]	10.5X	10.1X
19999871	Less selected	Longhorn[3]	10.2X	11.1X
19999872	Less selected	Longhorn[3]	10.0X	9.2X
NC_032662.1	Zebu	Nellore[5]		
19919842	Zebu	Brahman[3]	19.6X	21.5X
19999811	Zebu	Brahman[3]	19.6X	21.5X
19999812	Zebu	Brahman[3]	10.0X	10.2X
19999813	Zebu	Brahman[3]	12.6X	9.3X
19999819	Zebu	Brahman[3]	12.1X	12.3X
SRR3659494	Zebu	Brahman[3]	6.0X	4.2X
SRR3659495	Zebu	Brahman[3]	5.3X	3.9X

ID: File identification in the original database. Dairy: Highly selected dairy breeds (*Bos taurus*); Beef: Highly selected beef breeds (*Bos taurus*); Less selected: less selected breeds (*Bos taurus*), includes the Auroch (Bos primigenius); Zebu: zebu breeds (*Bos indicus*). All samples were aligned to the genomic region of BTA13 spanning the *PDYN* gene and only this region of each sample was used for the analyses. [1]Bovine Reference Genome ARS-UCD1.2, GenBank accession: NC_037340.1: 53,130,235-53,148,819). [2]"1000 Bull Genome Project" Run 2, (http://www.1000bullgenomes.com/). [3]U.S. Meat Animal Research Center (USMARC) Beef Cattle Diversity Panel 2.9 (MBCDPv2.9, Heaton et al., 2016). [4]*Bos primigenius* archeological sample (Park et al., 2015). [5]*Bos indicus* Reference Genome Bos_indicus_1.0 (GenBank accession: NC_032662.1). *Average depth of coverage obtained from their respective original database. #Average depth of coverage calculated by using Mosdepth software v 0.2.9 (Pedersen and Quinlan, 2018).

commands. The average depth of coverage of each sample for the whole genome sequence was retrieved from the original database. On the other hand, the average depth of coverage of each sample for the *PDYN* gene region was calculated using Mosdepth software v 0.2.9 (Pedersen and Quinlan, 2018) (Table 1).

Haplotype reconstruction

The Prodynorphin cDNA was cloned by Jiang *et al.* (1997); Genbank accession U58500.1. This sequence agrees with accession NM_174139.3 from the annotation of the last version of the Bovine Reference Genome (ARS-UCD1.2), and therefore it was used in the present work as a reference of gene organization. In the comparative sequence analyses, the whole sequence and four functional regions of the mRNA were considered: a 1,000 bp fragment spanning the putative proximal promoter; the 5' untranslated region (5'UTR, 174 bp); the coding sequence (CDS, 777 bp); and the 3' untranslated region (3'UTR, 1,588 bp) (Figure 1).

The sequences of the breed panel were aligned online with the MAFFT software version 7 (https://mafft.cbrc.jp/alignment/server/) and the resulting alignment was manually edited using the Aliview software version 1.23 (Larsson, 2014). Then, haplotype phases from sequences with heterozygous sites were reconstructed with DnaSP version 6.10.03 (Rozas *et al.*, 2017) using the option fastPHASE (Scheet and Stephens, 2006) with default parameters. Haplotypes were reconstructed for the whole sequence and then trimmed into the functional regions as defined above.

The phylogenetic relationships among the haplotypes identified in the putative proximal promoter (1,000 bp) and in the coding sequence (CDS) of the *PDYN* gene, were inferred through a median-joining network analysis (Bandelt *et al.*, 1999) using the PopART software version 1.7 (Leigh and Bryant, 2015).

Principal component analyses (PCA)

To evaluate population structure and the genetic relationships among the four breed groups ("Dairy", "Beef", "Less selected" and "Zebu"), principal component analyses (PCA) were computed on haplotype frequencies of both, the putative proximal promoter region (1,000 bp) and the cDNA (complementary DNA, 2,539 bp), of *PDYN* using the 'prcomp' function of Rstudio v.3.4.4 (R Development Core Team, 2018). PCA results were plotted using the R package 'ggplot2' (Wickham, 2016).

Signatures of selection

To detect departures from the expectations of the neutral theory of evolution (Kimura, 1968) the neutrality tests Tajima's D (Tajima, 1989) and Fu & Li's F* and D* (Fu and Li, 1993) were conducted. Tajima's D compares the estimate of DNA sequence variation based on the average pairwise distance between all sequences in the sample (φ), to the estimate of DNA sequence variation based on the observed number of segregating sites and the number of chromosomes in a sample (θ). Under neutrality, the means θ and φ should be approximately equal to each other. Therefore, the expected value of Tajima's D for populations conforming to a standard neutral model is zero. Significant deviations from zero indicate a skew in the allele frequency distribution relative to neutral expectations. Positive values of Tajima's D arise from an excess of intermediate frequency alleles and can result from population bottlenecks, structure and/or balancing selection. Negative values of Tajima's D indicate an excess of low frequency alleles and can result from population expansions or positive selection (Tajima, 1989; Biswas and Akey, 2006). On the other hand, Fu & Li´s test makes the distinction between old and recent mutations as determined by where they occur on the branches of genealogies. The D* and F* statistics compare an estimate of the population mutation rate based on the number of derived variants seen only once in a sample (referred to as singletons) with φ or θ, respectively. Fu & Li´s F* and D* values are negative when there is an excess of recent mutations and will be taken as evidence against the neutrality of mutations (Fu and Li, 1993; Biswas and Akey, 2006). These tests were conducted in the four defined cattle groups for the whole *PDYN* gene and also for each functional region defined above, using DnaSP version 6.10.03 (Rozas *et al.*, 2017).

RESULTS

Principal component analyses

The PCA conducted on the haplotype frequencies of the cDNA and the putative promoter of *PDYN* are presented in Figure 2. It can be seen that the distribution of the breed groups does not agree between both gene regions. In the case of the promoter, PC1 alone explains 79.7% of the variability. Breeds are distributed along this axis, in agreement with the criteria that were originally considered to make the groups (selection history and, particularly, sexual precocity) suggesting a population structure defined not only by breed isolation and genetic drift but probably also by selection. On the contrary, when haplotype frequencies of the cDNA are considered, PC1 and PC2 explain 37.1% and 32.4% of the variation, respectively, and breed distribution do not match grouping criteria.

Signatures of selection

Departures from neutral theory expectations were analyzed on each of the four defined breed groups ("Dairy", "Beef", "Less selected" and "Zebu") (Table 4). In the "Dairy" cattle group, these tests were

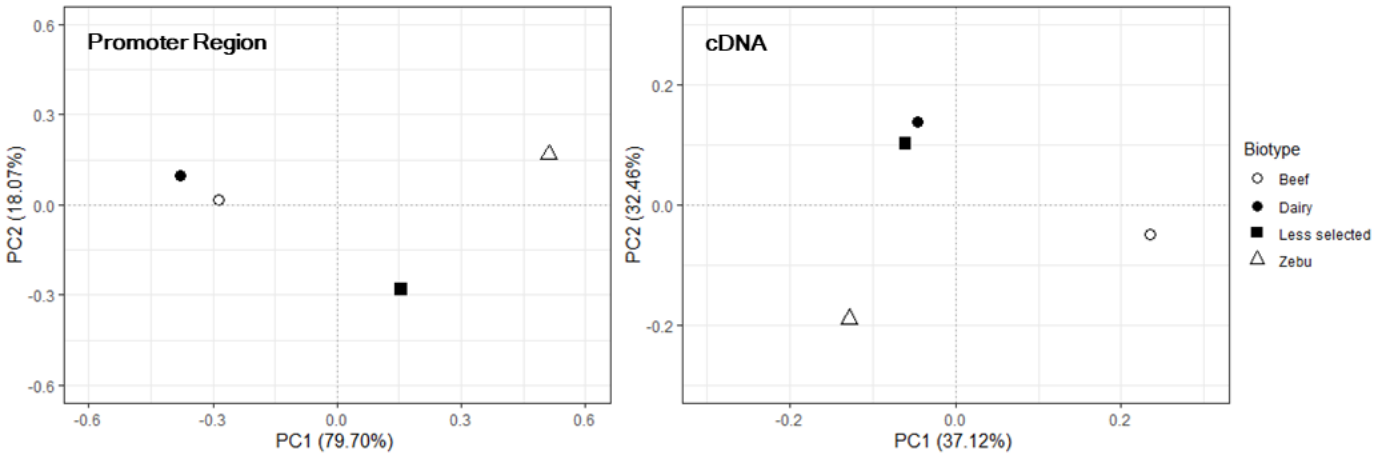

Figure 2. Bidimensional plot illustrating the results of the Principal Component Analysis performed on haplotype frequencies of the putative proximal promoter (left) and the cDNA (NM_174139.3) (right) of the bovine *PDYN* gene. The percentages of total variability explained by each PC are included between brackets. Dairy: Highly selected dairy breeds (*Bos taurus*); Beef: Highly selected beef breeds (*Bos taurus*); Less selected: Less selected breeds (*Bos taurus*), included the Auroch (*Bos primigenius*); Zebu: Zebu breeds (*Bos indicus*).

consistently negative and resulted to be statistically significant for the Genomic sequence as well as for the putative promoter region (Table 2). These results suggest that this gene and particularly the promoter region could be under positive selection. Moreover, these results pointed to a potential effect on regulatory features of *PDYN* gene expression in Dairy cattle. The other groups, on the contrary, had much higher nucleotide and haplotype diversity in the promoter and CDS than the Dairy breeds, and none of their neutrality tests were statistically significant (Table 2).

Haplotype description

Table 3 shows the mutations identified in the promoter region of the *PDYN* gene that had evidence of positive selection according to the neutrality tests, and in the CDS. These mutations define the haplotypes reported in Table 4.

The combination of 14 SNPs defined the haplotypes of the promoter region (1,000 bp) (Table 4). In this region, a total of 32 haplotypes were found among the 114 chromosomes (57 samples by 2 chromosomes each) analyzed. Sixteen of them were found in at least two chromosomes and were represented by 98 chromosomes (Table 4). The most abundant haplotype (p-Hap_1, ref; n= 56) was found in the Bovine Reference Genome ARS-UCD1.2 but it was absent in the "Zebu" cattle group. In the CDS (777 bp), a total of 28 haplotypes defined by 10 SNPs were found among the 114 chromosomes analyzed. Twelve of them were found in at least two chromosomes and were represented by 98 chromosomes (Table 3). In this region, the haplotype found in the Bovine Reference Genome ARS-UCD1.2 (c-Hap_1, ref; n=15) was not the most abundant and was the only one with the "A" allele in the SNP rs42388967 (SNP 3 in the CDS, Table 4). Although "Beef" and "Dairy" groups shared their most frequent haplotype (p-Hap_9) in the promoter that was not the case for the CDS (c-Hap_9 and c-Hap_1 for "Dairy" and "Beef" groups, respectively).

The promoter of *PDYN* had five different haplotypes in dairy breeds but with little variability, given that 34 out of 40 chromosomes shared p-Hap_1. This same haplotype was present in 20 out of 33 chromosomes of selected beef breeds, but the rest of the chromosomes of this group had seven different haplotypes in low frequencies each (Table 4). Moreover, in the group of dairy breeds, 40 out of 44 haplotypes of de CDS share the same six alleles in their 5' end. This same trend was detected in the promoter, where 37 out of 40 haplotypes share the same seven alleles in their 3' end. This pattern is not seen in any other breed group (Table 4). Zebu breeds had four haplotypes with a predominance of p-Hap_11 in the promoter region and showed little coincidence in haplotype distribution with the other breed groups (Table 4). Interestingly, differences between breed groups were smoothed when the CDS was considered and even the "Zebu" group showed more coincidence with the other *Bos taurus* groups. Nevertheless, for the CDS the "Dairy" group showed again the lowest variability (30 out of 44 chromosomes had Hap_9; Table 4).

The "Median Joining" network constructed using the haplotypes identified in the promoter region suggested the existence of two clades separated by three mutational steps (Figure 3A). In this network, the haplotypes found in the "Dairy" cattle group could be differentiated from those of the "Zebu" cattle group. Noteworthy, the "Dairy" group exhibited few haplotypes and they formed a star-like configuration with the reference haplotype (p-Hap_1) as central haplotype in the clade

Table 2. Diversity measures and neutrality tests. Results corresponding to the whole genomic sequence of *PDYN* and each of the different functional regions of transcript NM_174139.3 are presented for each group of breeds.

Region	Group	n	h	Hd	S	π	θ	Tajima's D	Fu and Li's D*	Fu and Li's F*
PDYN										
Genomic sequence	**Dairy**	**48**	**45**	**0.997 ± 0.004**	**188**	**0.00107**	**0.00216**	**-1.83842 ***	**-4.69877 ****	**-4.33775 ****
	Beef	34	31	0.989 ± 0.012	142	0.00172	0.00177	-0.12272	0.48822	0.32942
	Less selected	16	16	1.000 ± 0.022	225	0.00282	0.00346	-0.81225	-0.90965	-1.02013
	Zebu	16	16	1.000 ± 0.022	270	0.00381	0.00381	-0.36633	-0.14531	-0.24128
PDYN mRNA (NM_174139.3)										
Promoter (1,000bp)	**Dairy**	**48**	**11**	**0.469 ± 0.090**	**13**	**0.00070**	**0.00293**	**-2.30211 ****	**-4.04418 ****	**-4.09018 ****
	Beef	34	7	0.608 ± 0.092	6	0.00143	0.00147	-0.08484	1.20449	0.94936
	Less selected	16	9	0.900 ± 0.050	9	0.00249	0.00272	-0.31003	-1.01097	-0.94107
	Zebu	16	7	0.750 ± 0.107	10	0.00158	0.00302	-1.78041 #	-1.66704	-1.95299
5'UTR	Dairy	48	2	0.042 ± 0.039	1	0.00024	0.00129	-1.10686	-1.82907	-1.87498
	Beef	34	2	0.337 ± 0.083	1	0.00194	0.00141	0.56557	0.58040	0.66397
	Less selected	16	2	0.500 ± 0.074	1	0.00287	0.00173	1.30896	0.68829	0.96505
	Zebu	16	2	0.125 ± 0.106	1	0.00072	0.00173	-1.16221	-1.45287	-1.56820
CDS	Dairy	48	10	0.598 ± 0.079	9	0.00157	0.00261	-1.12637	-1.91413	-1.95186 #
	Beef	34	15	0.836 ± 0.056	8	0.00296	0.00252	0.52289	1.31840	1.25513
	Less selected	16	11	0.950 ± 0.036	9	0.00326	0.00349	-0.24232	0.42855	0.28108
	Zebu	16	10	0.942 ± 0.036	7	0.00347	0.00272	0.98621	1.31791	1.40942
3'UTR	Dairy	48	26	0.926 ± 0.028	35	0.00472	0.00497	-0.17413	-1.84513	-1.48846
	Beef	34	16	0.775 ± 0.074	18	0.00292	0.00278	0.16929	0.90055	0.78173
	Less selected	16	15	0.992 ± 0.025	22	0.00539	0.00477	0.28355	0.46251	0.75674
	Zebu	16	14	0.975 ± 0.035	18	0.00491	0.00342	1.73346	1.01041	1.39731

n: number of chromosomes; h: number of haplotypes; Hd: haplotype diversity; S: number of segregating sites; π: nucleotide diversity (per site); θ= 4Nμ (N and μ are the effective population size, and the mutation rate per DNA sequence per generation, respectively); Dairy: Highly selected dairy breeds (*Bos taurus*); Beef: Highly selected beef breeds (*Bos taurus*); Less selected: less selected breeds (*Bos taurus*), includes the Auroch (*Bos primigenius*); Zebu: zebu breeds (*Bos indicus*). #0.1>p>0.05; *p<0.05; **p<0.01. A region of the gene was considered under selection when both Tajima's and Fu & Li's neutrality tests produced significant results. These regions are shown in bold. 5'UTR and 3'UTR: untranslated regions; CDS: coding sequence.

(Figure 3A). In contrast, the haplotypes from the other groups were sparsely distributed. On the other hand, the "Median Joining" network constructed using the haplotypes identified in the CDS did not exhibit a clear pattern of variation and contrary to what happened in the promoter region, the haplotypes found in the CDS in the different cattle groups were not separated in this network (Figure 3B).

The 28 haplotypes identified in the CDS (Table 4) could be translated into six propeptide sequences, which had no more than four changes compared to the translated reference sequence. The four breed groups coincidentally had the same translated sequence (translated sequence of c-Hap_9) as the most abundant (Table 4). Interestingly, none of the aminoacid substitutions affected the sequence of the functional peptides that are processed from the translated propeptide (Table 3, Figure 1).

Table 3. List of mutations identified in two regions of the *PDYN* gene (promoter and coding sequence, CDS).

Region	Position	ID	Mutation	aa position	aa substitution	Coordinates (ARS-UCD 1.2)
***PDYN* mRNA promoter (1,000 bp)**	1	rs483143590	C>T			53,133,835
	2	rs133702134	C>T			53,133,850
	3		T>C			53,133,873
	4		->T			53,133,905
	5		G>A			53,133,933
	6		C>G			53,133,938
	7		G>T			53,133,945
	8		T>-			53,134,299
	9		C>-			53,134,300
	10		C>T			53,134,467
	11	rs480238520	C>A			53,134,474
	12	rs445119268	->C			53,134,494
	13		A>C			53,134,596
	14		T>G			53,134,604
CDS (777 bp)	1	rs518871097	C>T	15		53,146,628
	2	rs211042415	T>C	52		53,146,739
	3	rs42388967	A>G	75	R>G	53,146,806
	4	rs1116570653	C>A	76	L>I	53,146,809
	5	rs473280776	A>G	114	N>S	53,146,924
	6	rs718114917	C>T	127	A>V	53,146,963
	7	rs132886325	G>A	127		53,146,964
	8	rs524857811	C>T	154		53,147,045
	9	rs134167025	G>A	160	A>T	53,147,061
	10	rs42388969	A>C	177		53,147,114

The "Position" column indicates the position in the corresponding haplotypes described in Table 4.

DISCUSSION

The involvement of Endogenous Opioid Peptides (EOP) as negative regulators of the reproductive cycle has been known for a long time, but their role has not been completely elucidated. In cattle, EOP received attention as factors consistently involved in the inhibitory effects of suckling on reproductive activity (Malven *et al.*, 1986). Frequent suckling or milking increases levels of EOP in the brain, which in turn increases the sensitivity of the hypothalamus to the negative feedback from estrogen (Squires, 2010). This experimental evidence confirms the role of Dynorphin A and EOP in general, as inhibitors of reproductive activity in mammals, and justify the search of selection signatures conducted in the present work to explain the well-known variability in reproductive performance among cattle breeds, that is noticeable since the onset of puberty (Diskin and Kenny, 2014).

In the present experiment, we analyzed an annotated sequence of *PDYN*, which encoded, among others, the endogenous opioid peptide Dynorphin A (Figure 1), from the last version of the Bovine Reference Genome (Genbank accession NM_174139.3). This sequence has a genomic organization that is consistent with the cDNA originally cloned by Jiang *et al.* (1997); Genbank accession U58500.1).

The dataset analyzed here included the only available sample from the Auroch (*Bos primigenius*, Park *et al.*, 2015) that were grouped with less selected breeds. This sample did not show marked differences with other groups. However, it must be noted that a single sample is not enough to identify signatures of selection. Thus,

Table 4. List of haplotypes. List of haplotypes defined by mutations identified in the coding sequence (CDS) and in the promoter region of the *PDYN*. Only haplotypes found in at least two chromosomes are shown.

Region	Haplotype	n	D	B	L_S	Z	n changes	Description
***PDYN* mRNA promoter (1,000bp)**	p-Hap_1	56	34	20	2		reference	CCT-ACTTCCC-AT
	p-Hap_11	9			1	8	4	C**T**T-**G**CT--CC-AT
	p-Hap_7	5		2	3		2	C**T**T-**G**CTTCCC-AT
	p-Hap_8	4	1	3			1	CCT-AC**G**TCCC-AT
	p-Hap_2	2		2			7	C**TTTG**CT--C**A**-A**G**
	p-Hap_3	2		1	1		3	CCT-AC**G**--CC-AT
	p-Hap_4	2		1	1		2	CCT-ACT--CC-AT
	p-Hap_5	2		2			3	C**T**T-**G**CTTCC**A**-AT
	p-Hap_9	2		2			2	CCT-AC**G**TC**T**C-AT
	p-Hap_12	2			1	1	4	CCT-AC**G**--**T**C-AT
	p-Hap_16	2	1		1		1	CCT-ACTTC**T**C-AT
	p-Hap_17	2			2		2	CCT-AC**G**TC**T**C**C**AT
	p-Hap_18	2	2				1	CCT-ACTTCCC-**C**T
	p-Hap_23	2	2				1	**T**CT-ACTTCCC-AT
	p-Hap_30	2				2	5	C**TC**-**G**CT--CC-AT
	p-Hap_31	2				2	5	C**T**T-**GG**T--CC-AT
	Totals	98	40	33	12	13		
CDS (777 bp)	c-Hap_1	15	1	13	1		reference	CTACACGCGA
	c-Hap_9	36	30	2	2	2	3	CT***G***C***G***CGCG**C**
	c-Hap_2	11	3	5		3	4	C**C*G***C***G***CGCG**C**
	c-Hap_3	10	4	3	3		2	CT***G***C***G***CGCGA
	c-Hap_12	9	5	1	1	2	4	CT***G***C***G***C**A**C**A**A
	cHap_15	4	1	1	2		3	CT***G***C***G***C**A**CGA
	c-Hap_16	3			1	2	3	CT***G***C***G***CGC**A**A
	c-Hap_10	2		1	1		1	CT***G***CACGCGA
	c-Hap_13	2		1		1	3	C**C*G***C***G***CGCGA
	c-Hap_17	2			1	1	6	**TT*G***C***GT***GC**AC**
	c-Hap_18	2			2		5	CT***GA***G**A**C**A**A
	c-Hap_28	2				2	5	C**C*G***C***G***CGT**GC**
	Totals	98	44	27	14	13		

The total number of identified chromosomes (n) are discriminated by breed group. Dairy: Highly selected dairy breeds (*Bos taurus*); Beef: Highly selected beef breeds (*Bos taurus*); Less selected: less selected breeds (*Bos taurus*), includes the Auroch (*Bos primigenius*); Zebu: zebu breeds (*Bos indicus*). n changes: number of changes respect to the bovine reference genome ARS-UCD1.2. Differences respect to the reference genome are shown in bold. The positions of SNPs that generate aa substitutions in the CDS are in bold italic font.

more genomes from this group should be sequenced to be more conclusive about the effects of domestication (Orlando, 2015).

The PCA results presented in Figure 2, showing sample substructure differed depending in the analyzed regions of *PDYN* (promoter and cDNA) indicates that the observed differences are not only justified by the well-known genetic distance between *Bos indicus* and *Bos taurus* breeds, product of sub-speciation, and justify the definition of four separated groups used in the other analyses. The comparison of haplotype frequencies among breed groups for both genomic regions (Table 4, Figure 3) also supports the existence of directional selection on regulatory regions of the gene and not only random genetic drift associated with the breed or subspecies formation in the definition of sequence variation.

It could be argued that low genetic variability is the result of reduced effective population size (N_e), a feature common to selected populations in most livestock species (Taylor *et al.*, 2016). In this case, the group of dairy breeds was integrated by two different breeds, Holstein and Jersey that share many selection objectives and also a common production system. Also, this group had the largest sample size.

The methods that were used in the present work are appropriate to identify recent selection (Biswas and Akey, 2006). Indeed, the results of the Tajima's D (Tajima, 1989) and Fu & Li D* and F* (Fu and Li, 1993) exhibited statistically significant negative values in the "Dairy" group (Table 2) which pointed out to a recent selection process in modern, highly specialized breeds. This evidence of recent selection is consistent with the

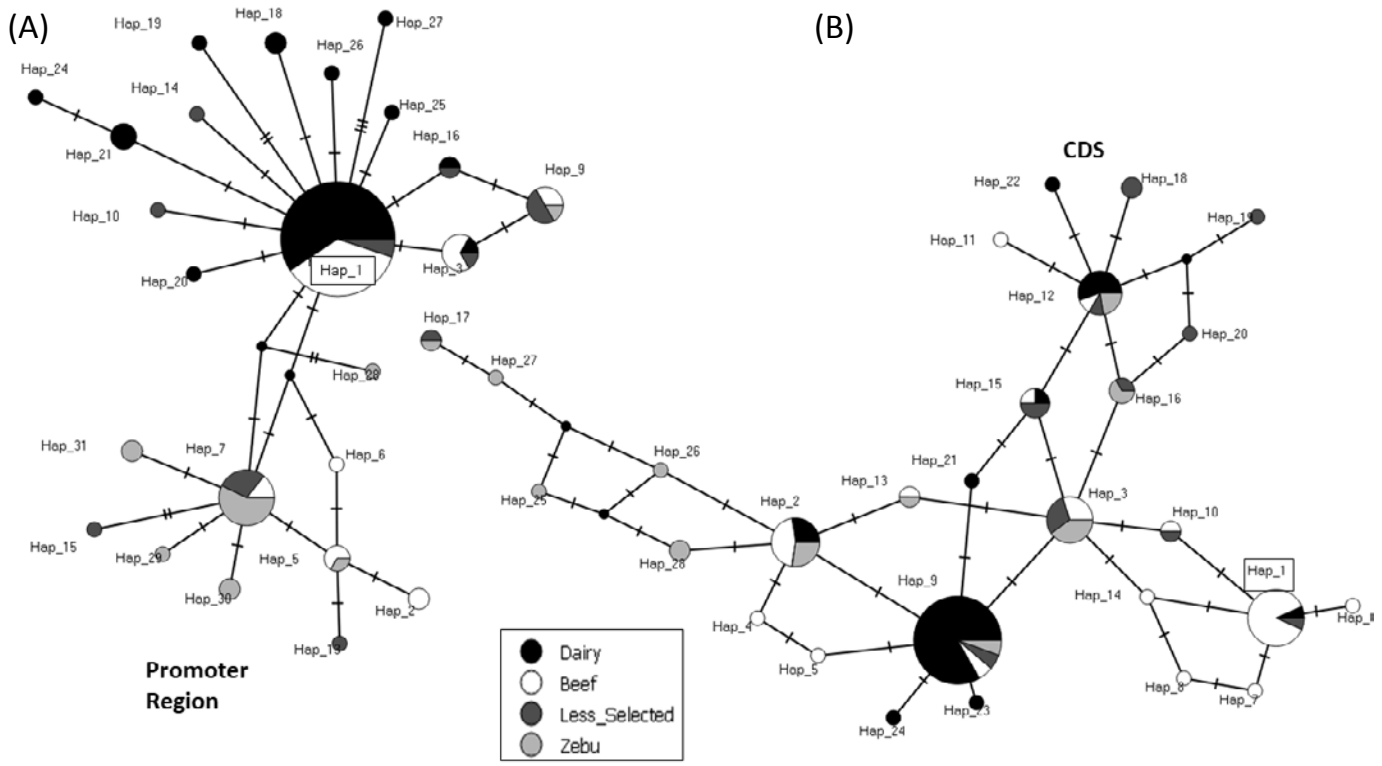

Figure 3. Median Joining network constructed with the haplotypes corresponding to (A) the putative proximal promoter (1,000 bp) and (B) the coding sequence (CDS) of *PDYN* gene. The size of each circle is proportional to the corresponding chromosome frequency. Slashes represent the number of mutations separating each haplotype. In each network Hap_1 (box) corresponds to the Bovine Reference Genome ARS-UCD1.2. Dairy: Highly selected dairy breeds (*Bos taurus*); Beef: Highly selected beef breeds (*Bos taurus*); Less selected: Less selected breeds (*Bos taurus*), included the Auroch (*Bos primigenius*); Zebu: Zebu breeds (*Bos indicus*).

emphasis of intense artificial selection on production traits, particularly those related to reproductive efficiency. A compelling argument in favor of positive selection on a gene such as *PDYN* comes from the intrinsic characteristics of different livestock production systems. Under extensive conditions, the cow not only nurses but also protects the progeny, and the maternal bond exerts a strong inhibiting effect on reproductive activity (Williams, 1990). In more intensified systems such as those for dairy production the calf is separated from the mother hours after parturition. Moreover, dairy cows are usually expected to get pregnant under a strongly negative energy balance (Butler and Smith, 1989; Beam and Butler, 1999; Vercouteren *et al.*, 2015). Clearly, in these extreme cases, the physiological and environmental cues that have to be processed at the hypothalamic level to regulate reproductive cycles are very different. It could be hypothesized that in females from more intensified production systems, the effects of some inhibitory mechanisms acting on the regulation of reproduction are relaxed, favoring the selection response for reproductive efficiency. The results reported here showed that the "Dairy" group presented fewer haplotypes than the other cattle groups in the promoter region and that these haplotypes exhibit low differences among them (Table 4, Figure 3A). A detailed description on the epigenetic regulation of the expression on *KISS1* and *TAC3* by implementing histone modifications was recently made by Toro *et al.* (2018). However, it does not seem to be the case of *PDYN* gene, the third member of the triad of the "KNDy hypothesis" (Smith *et al.*, 2014). The results presented here warrants further investigation on the mechanisms underlying the *PDYN* transcriptional regulation.

The results presented here suggest that there is selection pressure acting on putative regulatory regions on the proximal promoter of *PDYN* (Table 2), which seems to be a common feature of selection for quantitative or complex traits. Modifications of the coding sequence are more common in the case of genes underlying mendelian traits, in which the mutation has a strong effect on the phenotype (Boyle *et al.*, 2017). Although aminoacid substitutions on the propeptide were detected in the present analysis, they do not seem to affect the sequence of functional peptides, stressing their biological importance (Table 3, Figure 1). Also, a not clear pattern of differentiation was observed among the haplotypes of the four cattle groups in the coding

sequence (Figure 3B). Selection on regulatory regions of *PDYN* would affect the expression of all the different peptides coded by the gene. Given the physiological role of Dynorphin A, the interest to improve reproductive efficiency in highly specialized cattle breeds provides a very reasonable explanation to justify selection on *PDYN*. Nevertheless, at this point, the relevance of the other peptides cannot be ruled out. Further research should clarify the role of EOP, particularly Dynorphin A, in the regulation of reproduction both within and between cattle breeds.

The usefulness of QTN identification (Quantitative Trait Nucleotides underlying Quantitative Trait Loci) in animal breeding is a matter of debate. Present genomic selection in cattle is mostly based on dense arrays of anonymous markers; however, methods are being developed to include QTN information to improve prediction accuracy (Fragomeni *et al.*, 2017). Besides, there are initiatives to enhance progress in animal selection trough the combination of conventional breeding and gene editing (*e.g.*: Promotion of Alleles by Genome Editing; Jenko *et al.*, 2015). However, there are doubts about the efficacy of this approach due to the intrinsic polygenic basis of quantitative variation (Simianer, 2018). Independently of potential direct applications of the knowledge about *PDYN* variation among cattle breeds, the present work contributed with elements for a deeper understanding of the complex interaction among genetic variation, artificial selection, and environmental effects.

BIBLIOGRAPHY

Amstalden M., Williams G.L. (2015) Neuroendocrine Control of Estrus and Ovulation. In: Hopper R.M. (Ed.) Bovine Reproduction. Wiley Blackwell, Starkville, Mississippi, USA, pp. 203-219.

Atkins J.A., Pohler K.G., Smith M.F. (2013) Physiology and Endocrinology of Puberty in Heifers. Vet. Clin. North. Am. Food. Anim. Pract. 29: 479-492.

Bandelt H.J., Forster P., Röhl A. (1999) Median-Joining Networks for Inferring Intraspecific Phylogenies. Mol. Biol. Evol. 16: 37-48.

Beam S.W., Butler W.R. (1999) Effects of energy balance on follicular development and first ovulation in postpartum dairy cows. J. Reprod. Fertil. Suppl. 54: 411-424.

Biswas S., Akey J.M. (2006) Genomic insights into positive selection. Trends Genet. 22: 437-446.

Boyle E.A., Li Y.I., Pritchard J.K. (2017) An expanded view of complex traits: from polygenic to omnigenic. Cell. 169: 1177-1186.

Butler W.R., Smith R.D. (1989) Interrelationships Between Energy Balance and Postpartum Reproductive Function in Dairy Cattle. J. Dairy Sci. 72: 767-783.

Corva P.M., Villarreal E.L., Mezzadra C.A. (1995) Reproductive traits of Angus, Criollo and reciprocal crossbred females in the temperate area of Argentina. Anim. Sci. 61: 241-249.

Daetwyler H., Capitan A., Pausch H. (2014) Whole-genome sequencing of 234 bulls facilitates mapping of monogenic and complex traits in cattle. Nat. Genet. 46: 858-865.

Day R., Lazure C., Basak A. (1998) Prodynorphin Processing by Proprotein Convertase 2. J. Biol. Chem. 273: 829-836.

Diskin M.G., Kenny D.A. (2014) Optimising reproductive performance of beef cows and replacement heifers. Animal 8: 27-39.

Ferraz Jr. M.V.C., Pires A.V., Santos M.H. (2018) A combination of nutrition and genetics is able to reduce age at puberty in Nelore heifers to below 18 months. Animal 12: 569-574.

Fragomeni B.O., Lourenco D.A.L., Masuda Y. (2017) Incorporation of causative quantitative trait nucleotides in single-step GBLUP. Genet. Sel. Evol. 49: 59.

Freetly H.C., Kuehn L.A., Cundiff L.V. (2011) Growth curves of crossbred cows sired by Hereford, Angus, Belgian Blue, Brahman, Boran, and Tuli bulls, and the fraction of mature body weight and height at puberty. J. Anim. Sci. 89: 2373-2379.

Fu Y.X., Li W.H. (1993) Statistical tests of neutrality of mutations. Genetics 133: 693-709.

Goodman R.L., Lehman M.N., Smith J.T. (2007) Kisspeptin neurons in the arcuate nucleus of the ewe express both dynorphin A and neurokinin B. Endocrinology 148: 5752-5760.

Heaton M.P., Smith T.P.L., Carnahan J.K. (2016) Using diverse U.S. beef cattle genomes to identify missense mutations in EPAS1, a gene associated with high-altitude pulmonary hypertension. F1000Research 5: 2003.

Irano N., De Camargo G.M.F., Costa R.B. (2016) Genome-wide association study for indicator traits of sexual precocity in Nellore cattle. PLoS One 11: 1-14.

Jenko J., Gorjanc G., Cleveland M.A. (2015) Potential of promotion of alleles by genome editing to improve quantitative traits in livestock breeding programs. Genet. Sel. Evol. 47: 55.

Jiang H., Weesner G.D., Malven P.V. (1997) cDNA sequence and expression of bovine prodynorphin. Gene 186: 279-283.

Kimura M. (1968) Evolutionary Rate at the Molecular Level. Nature 217: 624-626.

Larsson A. (2014) AliView: a fast and lightweight alignment viewer and editor for large datasets. Bioinformatics 30: 3276-3278.

Laster D.B., Smith G.M., Cundiff L.V. (1979) Characterization of biological types of cattle (Cycle II). II Postweaning growth and puberty of heifers. J. Anim. Sci. 48: 500-508.

Lehman M.N., Coolen L.M., Goodman R.L. (2010) Minireview: Kisspeptin/neurokinin B/dynorphin (KNDy) cells of the arcuate nucleus: A central node in the control of gonadotropin-releasing hormone secretion. Endocrinology 151: 3479-3489.

Leigh J., Bryant D. (2015) PopART: Full-feature software for haplotype network construction. Methods Ecol. Evol. 6: 1110-1116. Available at: http://popart.otago.ac.nz.

Li H. (2011) A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27: 2987-2993.

Malven P.V. (1986) Inhibition of pituitary LH release resulting from endogenous opioid peptides. Domest. Anim. Endocrinol. 3: 135-144.

Malven P.V., Parfet J.R., Gregg D.W. (1986) Relationships among Concentrations of Four Opioid Neuropeptides and Luteinizing

Hormone-Releasing Hormone in Neural Tissues of Beef Cows Following Early Weaning. J. Anim. Sci. 62: 723-733.

Meirelles C.F., Abdalla A.L., Vitti D.M.S.S. (1994) The effect of feed supplementation on the onset of puberty in Brazilian dairy heifers. Sci. Agric. 51: 374-380.

Merchant S., Wood D.E., Salzberg S.L. (2014) Unexpected cross-species contamination in genome sequencing projects. Peer J. 2: e675.

Orlando L. (2015) The first aurochs genome reveals the breeding history of British and European cattle. Genome Biol. 16: 225.

Pardo A.M., Villarreal E.L., Papaleo Mazzucco J. (2018) Sexual precocity and productivity of beef cattle female under grazing conditions. Anim. Prod. Sci. 59: 757-766.

Park S.D.E., Magee D.A., McGettigan P.A. (2015) Genome sequencing of the extinct Eurasian wild aurochs, *Bos primigenius*, illuminates the phylogeography and evolution of cattle. Genome Biol. 16: 234.

Pedersen B.S., Quinlan A.R. (2018) Mosdepth: quick coverage calculation for genomes and exomes. Bioinformatics 34: 867-868.

R Development Core Team (2018) R: a language and environment for statistical computing Vienna, Austria: The R Foundation for Statistical Computing. Available at: http://www.r-project.org/.

Rozas J., Ferrer-Mata A., Sanchez DelBarrio J.C. (2017) DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol. Biol. Evol. 34: 3299-3302.

Sartori R., Gimenes L.U., Monteiro Jr. P.L.J. (2016) Metabolic and endocrine differences between *Bos taurus* and *Bos indicus* females that impact the interaction of nutrition with reproduction. Theriogenology 86: 32-40.

Scheet P., Stephens M. (2006) A Fast and Flexible Statistical Model for Large-Scale Population Genotype Data: Applications to Inferring Missing Genotypes and Haplotypic Phase. Am. J. Hum. Genet. 78: 629-644.

Simianer H. (2018) Of Cows and Cars. J. Anim. Breed. Genet. 135: 249-250.

Smith M.F., Hawken P.A.R., Lehman M.N. (2014) The role of kisspeptin in reproductive function in the ewe. In: Juengel J.L., Miyamoto A., Price C., Reynolds L.P., Smith M.F., Web R. (Eds.) Reproduction in Domestic Ruminants VIII. Context, UK, pp. 105-116.

Squires E.J. (2010) Applied Animal Endocrinology. 2nd Ed. CABI Publishing, Guelph, Ontario, Canada.

Tajima F. (1989) Statistical Method for Testing the Neutral Mutation. Genetics 123: 585-595.

Taylor J.F., Taylor K.H., Decker J.E. (2016) Holsteins are the genomic selection poster cows. Proc. Natl. Acad. Sci. 113: 7690-7692.

Toro C.A., Wright H., Aylwin C.F. (2018) Trithorax dependent changes in chromatin landscape at enhancer and promoter regions drive female puberty. Nat. Commun. 9: 57.

Vercouteren M.M.A.A., Bittar J.H.J., Pinedo P.J. (2015) Factors associated with early cyclicity in postpartum dairy cows. J. Dairy Sci. 98: 229-239.

Weems P.W., Lehman M.N., Coolen L.M. (2018) The Roles of Neurokinins and Endogenous Opioid Peptides in Control of Pulsatile LH Secretion. In: Litwack G. (Ed.) Vitamins and Hormones. Elsevier Inc., Academic Press, pp. 89-135.

Wickham H. (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag, New York.

Williams G.L. (1990) Suckling as a regulator of postpartum rebreeding in cattle: a review. J. Anim. Sci. 68: 831-852.

Zimin A.V., Delcher A.L., Florea L. (2009) A whole-genome assembly of the domestic cow, *Bos taurus*. Genome Biol. 10: R42.

BAG | Journal of Basic & Applied Genetics

Vol XXXI (2): 27-38; December 2020

Journal of the Argentine Society of Genetics

CHROMOSOME NUMBER VARIATION IN PART OF THE FLORA OF PROTECTED WILD AREAS IN THE ARAUCANIA REGION OF SOUTHERN CHILE

VARIACIÓN DEL NÚMERO CROMOSÓMICO EN PARTE DE LA FLORA DE ÁREAS SILVESTRES PROTEGIDAS EN LA REGIÓN DE LA ARAUCANÍA, SUR DE CHILE

Jara-Seguel P.[1,2], Urrutia-Estrada J.[3], Vallejos N.[1], Andrade E.[4], Jara M.[5]

[1] Departamento de Ciencias Biológicas y Químicas, Universidad Católica de Temuco, Chile.

[2] Núcleo de Estudios Ambientales (NEA), Facultad de Recursos Naturales, Universidad Católica de Temuco, Chile.

[3] Laboratorio de Invasiones Biológicas, Universidad de Concepción, Chile

[4] Programa de Doctorado en Educación, Facultad de Educación, Universidad Católica de Temuco, Chile.

[5] Department of Population Health and Pathobiology, North Carolina State University, USA.

Corresponding author:
Pedro Jara-Seguel
pjara@uct.cl

ORCID 0000-0002-9340-1780

Cite this article as:
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 ARAUCANÍA REGION OF SOUTHERN CHILE. BAG. Journal of Basic and Applied Genetics XXXI (2): 27-38.

Received: 03/16/2020
Revised version received: 08/04/2020
Accepted: 08/31/2020

General Editor: Elsa Camadro

DOI: 10.35407/bag.2020.31.02.03

ISSN online version: 1852-6233

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

ABSTRACT

An analysis was made of the correspondence between species diversity and chromosome number (CN) diversity across 13 Protected Wild Areas (PWA) in the Araucanía Region of southern Chile, encompassing 84 plant species with available cytogenetic data. Our aim was to establish whether higher species diversity within a PWA entails higher CN variation as based on the index of chromosome number heterogeneity (ICNH). The CN data were extracted from databases for Chilean plants, and the ICNH for the flora of each PWA was calculated. Results showed that in nine PWA the species diversity clearly correlates with CN diversity. However, four PWA do not fit this trend. The percentage of species with CN data varied between 9.6% and 24.5% among PWA, with 11 PWA presenting percentages higher than 11%. A 27.3% of the Chilean vascular plant species with available cytogenetic data were studied here for the 13 PWA. The results obtained by studying one part of the flora with available CN data suggest that the PWA could be an important reservoir of genetic diversity at a chromosome level, thus justifying the protective role of the PWA as biodiversity conservation sites.

Key words: Chromosome number heterogeneity; floristic diversity; Chilean flora.

RESUMEN

Se realizó un análisis de la correspondencia entre la diversidad de especies y la diversidad de números cromosómicos (CN) en 13 Áreas Silvestres Protegidas (PWA) en la Región de La Araucanía en el sur de Chile, incluyendo 84 especies de plantas con datos citogenéticos disponibles. El objetivo fue establecer si una mayor diversidad de especies dentro de un PWA implica una mayor diversidad en CN expresado en base al Índice de Heterogeneidad Cromosómica (ICNH). Los CN de cada especie se extrajeron de bases de datos para plantas chilenas y se calculó el ICNH para la flora de cada PWA. Los resultados mostraron que en nueve PWA la diversidad de especies se correlaciona claramente con la diversidad de CN. Sin embargo, cuatro PWA no se ajustan a esta tendencia. El porcentaje de especies con datos de CN varió entre 9,6% y 24,5% entre PWA, con 11 PWA presentando porcentajes superiores al 11%. Un 27,3% de las especies de plantas vasculares chilenas con datos citogenéticos disponibles fueron estudiadas para las 13 PWA. Los resultados obtenidos al estudiar parte de la flora sugieren que las PWA serían un reservorio importante de diversidad genética a nivel cromosómico como se muestra aquí, justificando así el papel protector de las PWA como sitios de conservación de la biodiversidad.

Palabras clave: Heterogeneidad del número cromosómico; diversidad florística; flora chilena.

INTRODUCTION

The Chilean Protected Wild Areas (PWA) system started up in 1984 as a dependent institution of SNASPE (National System of State Protected Wild Areas) (Pauchard and Villarroel, 2002) encompassing 105 terrestrial PWA throughout Chile which are currently managed by the National Forest Corporation (Corporación Nacional Forestal-CONAF). Since their inception the PWA have been understood to be high biodiversity sites along the length of the Chilean territory and many of them are relics of extensive old forests, including taxa from multiple geographical origins (Troncoso *et al.*, 1980; van der Hammen and Cleef, 1983; Villagrán and Hinojosa, 1997; Moreira Muñoz, 2011; Armesto *et al.*, 2010; Scherson *et al.*, 2017). The vascular flora of the PWA is known to be one of the most visible forms of life in the forests that contain them and plant species are vertically organised as herbaceous, shrub and arboreal strata (Smith, 1973; Ramírez *et al.*, 1990), giving rise to environments that harbour an important diversity of organisms belonging to different kingdoms (Smith Ramírez *et al.*, 2007; Marín *et al.*, 2017).

The genetic diversity of the Chilean flora is a heritage that is important to conserve, and for this reason its study requires the use of multiple tools to facilitate its description (Jara Seguel and Urrutia, 2012). The genetic diversity of Chilean plants was initially analysed using isozyme electrophoresis and later on, with the advent of DNA technologies, fingerprint profiling was conducted in populations of single species (Premoli, 1997; Premoli *et al.*, 2000; 2012; Torres Díaz *et al.*, 2007; Premoli and Mathiasen, 2011; García Gonzales *et al.*, 2008; Martin *et al.*, 2014; Bastías *et al.*, 2016). Other studies have used DNA sequences focused on performing phylogenetic reconstructions including species of various families and orders (Aagesen and Sanso, 2003; Davis *et al.*, 2004; Chacon *et al.*, 2012a; 2012b; Delaveau *et al.*, 2013; Jara Arancio *et al.*, 2013; Givnish *et al.*, 2016). In this context, comprehensive work was carried out reconstructing a spatial phylogenetic tree that included 756 native genera of vascular plants (*ca.* 87% of the total in Chile) thus evaluating both the phylogenetic diversity and endemism of Chilean flora (Scherson *et al.*, 2017). The Scherson study on Chilean flora does not specify whether the sampled specimens were taken within the PWA mentioned as such by us. However, the geographic coordinates of various sites that they describe coincide with the PWA located in the Andean range, Central valley and Nahuelbuta range (see appendix A with supplementary material in Scherson *et al.*, 2017). An additional organisational level to analyse genetic diversity is the chromosomal (Stebbins, 1971; Levin, 2002; Windham and Yatzkievych, 2003; Severns and Liston, 2008; Peruzzi *et al.*, 2012; Morero *et al.*, 2015). The chromosome set represented by the CN accounts for the complete genome in addition to the chromosome morphology, thus determining a nuclear architecture that is unique to each species. This nuclei ordering is key to understand the organisation and functionality of the plant genomes both in interphase processes and in cell division (Schneider and Grosscheldl, 2007; Heslop Harrison and Schwarzacher, 2011). Specifically, gene expression depends on the ordering of multiple chromosome domains within the nucleus (Fernández Donoso and Berríos, 1985; Gregory, 2001). For decades many studies described the CN independently of the chromosome morphology, using it as a basic genetic character to analyse similarity or variation between species (Peruzzi *et al.*, 2012; 2014). Currently CN data are available for *ca.* 307 species of Chilean vascular plants, which represent 135 genera and 60 families (*ca.* 6.6% of the total; Jara Seguel and Urrutia Estrada, 2020) many of them inhabiting in PWA throughout the continental territory. Based on these data, a high CN variation is observed within Chilean vascular plants along the continent and in insular areas.

The CN could be a good marker to evaluate genetic variation in the flora of the PWA as a whole and not just based on single species, making it possible to overlay the species diversity. Thus, two matrices analysing diversity –the floristic and the chromosomal– can be superimposed. Quantitative analyses based on the index of chromosome number heterogeneity (ICNH) have been recently proposed to compare CN variation among different plant or animal taxa (Peruzzi *et al.*, 2014), which could be used to determine quantitative CN diversity in areas harbouring different species of native plants such as occurs in the PWA.

In the Araucanía Region of southern Chile (from 37º to 39º S) there are 13 terrestrial PWA. Ten PWA are located in the Andean range forming the Araucarias Biosphere Reserve, whereas one PWA is located in the Central valley and two PWA are located in the Nahuelbuta range near the Pacific coast (CONAF, 2013). A recent cadastre carried out only for the 10 PWA of the Biosphere Reserve (Natural Reserve (RN) Malleco, RN Las Nalcas, RN Malalcahuello, RN Alto Biobío, RN China Muerta, RN Villarrica, National Park (PN) Tolhuaca, PN Conguillío, PN Villarrica and PN Huerquehue) recognised 829 species present in these areas (Hauenstein and Saavedra, 2019). Nevertheless, the PWA located in the Central valley [Natural Monument (MN) Cerro Ñielol] and Nahuelbuta range (PN Nahuelbuta and MN Contulmo) (Baeza *et al.*, 1999; Arriagada, 2002; Saavedra and Morales, 2008), have been described as presenting high floristic diversity but it have not yet been considered a biosphere reserve. In this context, the high species diversity could be correlated with high genetic diversity, such as was discussed theoretically by Vellend (2005), by simulating correlation models

between both levels, and could also correspond to a CN variation as a genetic character in PWA. At present, no accurate cadastre of number of species with CN data has been published for the 13 PWA from the Araucanía Region in Chile, although information is available in electronic databases as well as in various printed sources. However, the species diversity for these areas represented by the number of species is well known and has been documented in several cadastres, leading us to query: i) does higher species diversity within a PWA entail higher CN diversity? and ii) are there appreciable differences in mean CN between PWA? In this study, we supply evidence based on the index of chromosome number heterogeneity (ICNH) calculations to provide a partial answer to these questions, using the CN data available for species inhabiting these PWA.

MATERIALS AND METHODS

Study areas

In this study, we evaluated the 13 PWA present in the Araucanía Region of southern Chile. The Araucanía Region is located between 37° 35´ and 39° 37´ S and from 70° 50´ W to the coast of the Pacific Ocean. Boundaries for each PWA are given here only as a reference since they depend on government definitions rather than geographical vegetation units (Table 1).

Floristic and cytogenetic data

Plant species in each PWA were obtained from floristic cadastres (e.g., Baeza *et al.*, 1999; Finckh *et al.*, 1995; Arriagada, 2002; Sepúlveda, 2004; Cortés, 2005; CONAF, 2009; Saavedra and Morales, 2008; Saavedra, 2009a; 2009b; 2009c; Saavedra and Hauenstein, 2010a; 2010b; Hauenstein, 2011a; 2011b). So for each PWA the species were listed and their respective CN were looked up in the CPCD (Chilean Plants Cytogenetic Database; Jara Seguel and Urrutia Estrada, 2020, with cytogenetic data for 402 species). Data on the geographical location (Region and Province) of each species were also obtained from CPCD, as well as the original source where cytogenetic data were published. As a criterion to determine high or low floristic diversity in each PWA we calculated the mean species diversity (±SD) for all 13 PWA. Thus, values that are above the mean will have high floristic diversity, and values under the mean will have low floristic diversity.

Chromosome number variation

To quantify the CN variation of the species within the PWA we followed all the steps proposed by Peruzzi *et al.* (2014). We calculated mean CN, standard deviation (SD) and frequency of B and Odd chromosomes. To quantify the variation of CN we calculated the square root of the

Table 1. Parameters studied for each PWA from the Araucanía Region, Southern Chile. TFD, total floristic diversity (number of species); NSC, number of species with CN data; CSS, cytogenetically studied species (%); ICNH, index of chromosome number heterogeneity (ICNH values are scored as low <30, high >30, and very high >40 as per Peruzzi et al. 2014). PN, National Park; RN, Natural Reserve; MN, Natural Monument.

Area Number	PWA	Surface (Km²)	TFD	NSC	CSS (%)	Mean CN	SD CN	fB+fOCN	ICNH
1	RN Las Nalcas	175.3	137	22	16.0	46.7	50.3	0	30.7
2	RN China Muerta	111.7	102	25	24.5	44.6	47.4	0	30.2
3	PN Villarrica	610.0	199	31	15.5	45.2	43.3	0.03	29.1
4	RN Malalcahuello	127.9	283	35	12.3	56.4	69.6	0	41.2
5	PN Tolhuaca	633.7	324	41	12.6	50.6	47.2	0.02	32.1
6	PN Conguillío	608.0	359	44	12.2	53.6	62.8	0	38.1
7	PN Huerquehue	125.0	245	29	11.8	55.4	54.2	0	36.0
8	RN Malleco	166.2	289	35	12.1	48.2	45.0	0.03	30.8
9	PN Nahuelbuta	68.3	311	31	9.96	41.5	40.0	0.03	26.8
10	RN Villarrica	613.5	96	16	16.6	42.8	31.0	0	24.0
11	MN Contulmo	820.0	239	33	13.8	51.7	46.9	0.03	32.4
12	MN Cerro Ñielol	844.0	165	34	20.6	55.7	61.6	0.03	38.5
13	RN Alto Biobío	330.0	119	14	11.7	33.1	38.3	0	23.4

area of the ideal triangle built in a three-variable radar plot, where the vertices of the triangle are defined by mean CN, SD, and % (*f*B + *f*OCN). The triangle gives a graphic representation of CN variation in a group (PWA in our study) and its area can be easily seen as the sum of the three areas subtended by the smaller triangles set along the plot axes. We defined this value as the Index of Chromosome Number Heterogeneity (ICNH), which was calculated according to the formula:

$$ICNH = \sqrt{\sin\left(\frac{2\pi}{3}\right)\frac{ab + ac + bc}{2}}$$

where ***a*** is the mean chromosome number (CN), ***b*** is the standard deviation (SD) of CN, and ***c*** is % (*f*B chromosomes + *f*Odd CN). The resulting value can vary from 0, if no variation occurs in a group, to +∞, although very high values can only be reached theoretically.

The mean CN±SD calculated for each PWA were then statistically compared between areas. Statistical pre-tests based on Kolmogorov-Smirnov and Shapiro-Wilk (using the same dataset) suggest the use of non-parametric analyses, given that the chromosome numbers showed an abnormal distribution ($p>0.05$). Thus, mean CN across PWA were compared using the Kruskal-Wallis test.

Correlation coefficients were calculated for mean CN and SD and for *f*B and *f*Odd CN, and grouped in three levels; weak (up to 0.3), moderate (0.4-0.7), and strong (>0.7) (Peruzzi *et al.*, 2014).

To determine the floristic similarity among the different PWA, a cluster analysis was carried out using SIMPROF (Similarity Profile) ($p<0.05$). This analysis is based on the conformation of a matrix consisting of the presence or absence of species, for which the Jaccard similarity index was calculated (Pielou, 1975).

Spatial analysis

The identification of geographic locations among the 13 PWA was carried out through spatial analysis while ArcGIS 10.3, Datum WGS 84, and Time Zone 18 S were used for the cartographic analysis. PWA species information was supplied by the Chilean National Forest Corporation (CONAF). The resulting cartography was contrasted with base lines, floristic cadastres, and available literature regarding PWA in order to determine the floristic diversity and relationship with the CN variation of the species obtained from the Chilean Plants Cytogenetic Database and other articles as cited above.

RESULTS

The ICNH values and other quantitative parameters for PWA are shown in Table 1. In Figure 1 plots are shown with the respective triangle of three PWA with higher ICNH and three PWA with lower ICNH. In total, we found CN data for 84 plant species present in the 13 PWA, encompassing 57 genera and 36 families belonging to Pteridophytes, Gymnosperms, and Angiosperms (Appendix 1 includes a list of studied species with CN and the PWA where each species occur). The number of species analysed here represent 27.3% of the total Chilean flora with available CN data (ca. 307 species studied to date), covering 12.3% of the total terrestrial PWA established in Chile. In addition, depending on the PWA, different percentages of species have been studied, ranging from 9.9% in PN Nahuelbuta to 24.5% in RN China Muerta.

According to floristic cadastres, PN Conguillío showed the highest floristic diversity among PWA with 359 species, while RN Villarrica was among the lowest with only 96 species. With regard to the number of species with cytogenetic data, the more extreme values were found in PN Conguillío with 44 species and RN Alto Biobío with only 14 species. The percentage of species with CN data (NSC – number of species with CN data) varied between 9.6% (PN Nahuelbuta) and 24.5% (RN China Muerta), although 11 PWA showed percentages higher than 11%. The mean species diversity estimated here for the 13 PWA was 220.6±90.3 species. Thus, seven PWA were higher than this mean value and six PWA were lower.

Only RN Malalcahuello presented very high ICNH (>40) and eight PWA presented high ICNH (>30). In turn, four PWA displayed low ICNH (<30). The coverage of the triangles shows clear differences between the higher and lower ICNH values obtained according to the score. The correlation found between mean CN and SD for the overall dataset was strong and positive ($r=0.80$), while the correlation between *f*B and *f*OCN was not estimated since *f*OCN was zero in all 13 PWA. Mean CN did not show significant differences between the 13 PWA when compared by means of the Kruskall-Wallis test (with significance at a 0.8011 level).

The floristic similarity cluster presented four significantly different groups with a resemblance higher than 60%: i) RN Alto Biobío, ii) MN Contulmo and MN Cerro Ñielol, iii) RN Villarrica, and iv) the remaining nine PWA (Figure 2A). Thus, the more related PWA groups shared a higher number of species among them. As a reference, only one species (*Nothofagus dombeyi*) is shared by all 13 PWA, whereas 24 species are shared by seven PWA, *i.e.* around 50% of the total PWA analysed here.

It is worth noting that, within our dataset, the diploid CN 16, 18, 22, 26, and 28 were observed in all studied PWA, whereas the polyploid CN shared among some PWA were 28, 116, 144, 164, 216, and 328, depending on the species. In total, 28 different CN were found across all 13 PWA, including the 84 species. Additionally, the

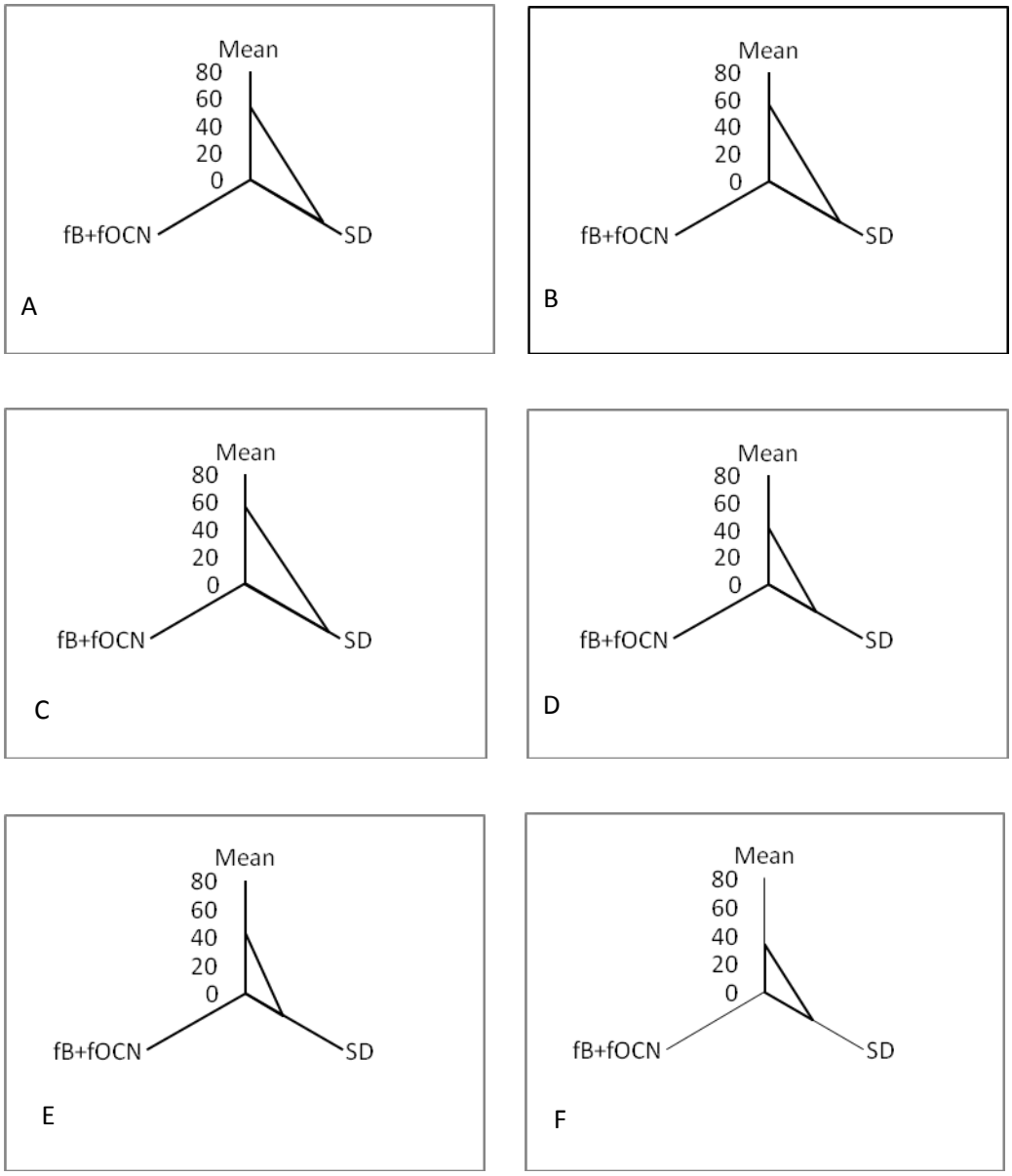

Figure 1. Comparison of three-variable (mean CN, *f*B+*f*OCN, and SD) radar plots between PWA with very high (> 40), high (>30) and low (<30) ICNH values. Note triangle coverage differences between PWA. Very high ICNH: RN Malalcahuello (A); high ICNH: MN Cerro Ñielol (B), and PN Conguillío (C); and low ICNH: PN Nahuelbuta (D), RN Villarrica (E) and RN Alto Biobío (F). ICNH values of each PWA are shown in Table 1.

modal CN for 12 PWA was 26, with the exception of MN Contulmo, which presented a modal number of 22.

DISCUSSION

Correlation analyses between species diversity and genetic diversity as a whole (as a community) have not been previously reported for Chilean plants. We studied the correspondence between both levels -species diversity and genetic diversity- but using the CN diversity as a genetic character with data available up to date in Chilean plants.

Chromosome number diversity

The analyses carried out in this study showed that six PWA with high species diversity (≥220 species) have a clear correlation with very high (>40) and high (>30) ICNH values, according to the scores obtained herein. In turn, three PWA with low species diversity (<220 species) also have a clear correlation with low ICNH values (<30) (See Table 1 and Figure 1 showing PWA with higher and lower ICNH). However, some PWA do not fit into this trend: for example, PN Nahuelbuta has high species diversity (311 species) but low ICNH (value of 26.8), whereas all three, MN Cerro Ñielol,

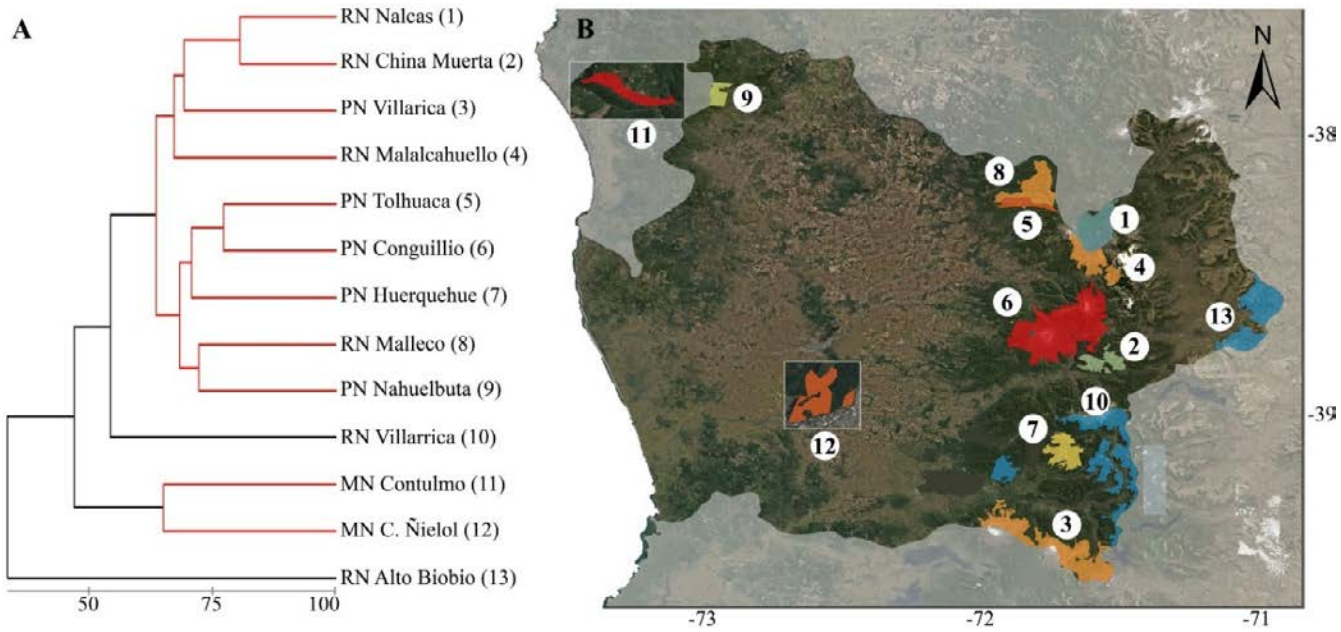

Figure 2. Cluster of floristic similarity for all 13 PWA (A). Coloured lines indicate groups without significant differences. PN, National Park; RN, National Reserve; MN, Natural Monument. Cartography showing the location of 13 PWA analysed in this study (B). Numbers represent each PWA as described in Table 1.

RN Las Nalcas and RN China Muerta, have low species diversity (165 species, 137 species, and 102 species respectively) but high ICNH (value of 38.5, 30.7 and 30.2 respectively). It is important to remark that the flora of MN Cerro Ñielol, a PWA located within the urban radius, is made up mainly of native species remaining from the original forest, but a high number of native species have also been introduced from other nearby areas as a conservation tool (Saavedra and Morales, 2008). Many of these species have CN data available in the databases (Jara Seguel and Urrutia, 2020), thus increasing the CN diversity of the MN Cerro Ñielol flora.

The very high and high ICNH values observed in six PWA is indicative of high CN diversity among them. Nevertheless, all 13 PWA studied here showed higher ICNH than was previously estimated for 243 Chilean species of vascular plants with an ICNH of 22.4 (Peruzzi *et al.*, 2014), but until that date, the databases did not include several polyploid species of pteridophytes which were added in our study. The ICNH for the six PWA mentioned above could be explained by the presence of polyploid species -25 in total- representing four genera of ferns and 12 genera of angiosperms with variable CN, ranging from 28 to 328, many of them being tetraploid, hexaploid or octoploid (Jara Seguel *et al.*, 2006; Jara Seguel and Urrutia, 2012; Jara Seguel and Urrutia Estrada, 2018; Morero *et al.*, 2015) (Figure 1A, B, and C). It is worth noting that diploid angiosperm species are the predominant plant group within the dataset (49 species in total), although they present a lower CN ranging from 8 to 32 as compared to polyploid taxa. An explanation for all ICNH >30 found in nine PWA (independent of the floristic diversity) may be related to the growth form of the plants. This is so because chromosome evolution proceeds much faster in herbs than in angiosperm shrubs and trees, as well as in conifers, as discussed by Levin and Wilson (1996), who described a net increase in the diversity of chromosome numbers and species numbers over time. A similar relationship may have occurred during the evolution of the flora in Chilean forests. As shown in our results, RN Malalcahuello, MN Cerro Ñielol, and PN Conguillio with higher ICNH (>30) present a high percentage of herbs (between 70% and 85% of herbs including ferns) *vis-à-vis* shrubs and trees (including conifers) within their flora with available CN data (NSC). In contrast, PN Nahuelbuta, RN Villarrica and RN Alto Biobio with lower ICNH have percentages of herbs lower than 67% (between 43% and 67%). All remaining PWA not mentioned above with ICNH >30 have a percentage of herbs of between 62% and 78%. Thus our findings show a clear relationship between high number of herb species and CN diversity.

All these aspects related to genetic variation are decisive for conservation biology (Severns and Liston, 2008) and represent a primary objective pursued by the PWA system. However, the presence of scant polyploid species with known CN data within the chromosome dataset from PN Nahuelbuta, RN Villarrica, RN Alto Biobío, and PN Villarrica could explain their lower ICNH (Figure 2D, E, and F). In these PWA only a few genera of pteridophytes and angiosperms with polyploidy records are present (Arriagada, 2002; CONAF, 2009) but unfortunately no chromosome counts are available for the species that inhabit these areas.

Other relevant observations are related to the

geographical location of the PWA along three separate longitudinal strips represented by the Andean mountain range, the Central valley, and the Nahuelbuta mountain range. None of these strips show a clear correspondence with high or low ICNH values, despite the differences in their climatic, geographic and geological conditions, which undoubtedly affect the flora (Montgomery *et al.*, 2001; Moreira Muñoz, 2011). As discussed by Levin and Wilson (1996), rates of evolution at both the karyotypic and organismal levels are related to the breeding structure of species and to environmental predictability. Using this reasoning to understand the cytoevolutionary process in PWA, it is remarkable to observe that different environments exist along their expanse. As such, small populations in variable habitats experiencing fluctuations in habitat hospitality are the most conducive to the fixation of chromosomal novelties (changes in CN). On the contrary, large continuous populations -where climatic and biotic pressures are stable over time- are likely to be more conservative in terms of chromosomal structure (stable CN). These cytogenetic aspects have not been studied for a large part of continental Chilean plants and cytoevolutionary mechanisms have only been described in some detail for some genera of herbaceous species (*e.g.*, *Chaetanthera*, *Alstroemeria*; Baeza *et al.*, 2009; 2015; 2018).

With regard to *f*B chromosomes, their occurrence in Chilean plants is very low (Jara Seguel and Urrutia, 2012) and their contribution to the ICNH values is negligible. Only one species within the studied dataset (*Lapageria rosea*) has been described as having B chromosomes (Jara Seguel and Zúñiga, 2004) in six PWA. Moreover, values of *f*OCN=0 were obtained in all PWA, since there were no species with an odd number of chromosomes across the dataset.

We also observed that various PWA share several species and therefore their respective CN. This may explain the non-significant differences in mean CN observed among PWA despite the CN heterogeneity described when estimating the correlation between mean CN and standard deviation. For example, seven PWA -*ca.* 50% of the total areas studied here share the presence of 25 species. The CN shared among PWA were 16, 18, 22, 26, 28, 116, 144, 164, 216 and 328 of a total of 28 different numbers. Similarity analysis showing the relation among PWA support this observation, where areas forming the same group share a higher number of species and therefore their respective CN (Figure 2A). Other PWA share a lower number of species both with other areas and when separated as a single branch in the analysis (RN Alto Biobío and RN Villarrica). In addition, a few species of different genera (taxonomically unrelated) or divisions present in various PWA share the same CN which may additionally support the non-significant differences observed in mean CN. For example, a modal CN of 26 was found in 12 PWA, appearing in three genera of three different divisions within the dataset [*e.g.*, the angiosperm genus *Nothofagus* (*N. alpina*, *N. antarctica*, *N. dombeyi*, *N. obliqua* and *N. pumilio*), the gymnosperm *Araucaria araucana*, and the pteridophyta *Hymenophyllum dentatum* and *H. tunbrigense* (Jara Seguel and Urrutia Estrada, 2020)]. An exception was PN Contulmo (group two in Figure 2A), presenting a high ICNH (>30) with a modal CN of 22, which was predominant over species with 26 chromosomes. The taxonomic composition in PN Contulmo showed a predominance of species with CN=22 (*e.g.*, one species for each of the *Ugni*, *Luma*, *Galium*, *Podanthus*, and *Chaetanthera* genera; Jara Seguel and Urrutia Estrada, 2020), unlike other PWA. Biogeographically, it is worth noting that most of the genera shared among the PWA have a Gondwanean distribution, which is present in southern South America and Oceania (Jara Seguel *et al.*, 2006; 2010; 2014; Chacon *et al.*, 2012a; Morero *et al.*, 2015). Some of these genera are part of the paleo-endemic flora, mainly ferns and conifers, displaying a recurrent presence in the region that includes the PWA studied here (Southern region according to Scherson *et al.*, 2017). Thus, we suggest that many of the CN found here could be plesiomorphic features within some families, specifically those containing species representative of paleo-endemic flora. Many of these CN, as well as the genera and family that contain them, are also shared with New Zealand and Australian vascular flora (Jara Seguel *et al.*, 2006; 2010; 2014; Morero *et al.*, 2015).

As mentioned above, CN data are available for *ca.* 6.6% of Chilean vascular plant species of which only 84 species present in the 13 PWA studied here have available CN data. Based on our results, it is possible that the real CN variation in these PWA might be even greater than the one estimated in this study, because a vast part of native species has not yet been cytogenetically studied, *i.e.* between 75% and 90%, depending on the PWA (CSS in Table 1). This CN diversity could also be superimposed on gene variation among species, thus adding a new matrix of analyses (molecular) to the two described here. In this way, previous studies of the vascular flora of Chile studied as a community, *e.g.* Scherson *et al.* (2017), based on DNA sequences and the present work based on CN diversity, provide a robust framework to continue studying the correlation between floristic diversity and genetic diversity at various levels, thus highlighting the genetic diversity present in the Chilean flora and also justifying its protection where PWA have a crucial role.

Appendix 1. Total species analysed in 13 PWA from Araucanía Region, Southern Chile (84 species). CN, chromosome number (2*n*). Numbers represent each PWA as described in Table 1.

Family	Species	CN	PWA where each species occur
Alstroemeriaceae	*Alstroemeria aurea*	16	1-2-3-4-5-6-7-8-9-10
	Alstroemeria ligtu	16	6-11
	Alstroemeria patagonica	16	4
	Alstroemeria pulchra	16	12
	Bomarea salsilla	18	11-12
	Luzuriaga radicans	20	5-6-7-8-9-11-12
Amaryllidaceae	*Phycella ignea*	16	9
	Rhodophiala advena	18	4-8
	Rhodophiala andicola	16	1-2-3-4-6-13
	Rhodophiala montana	18	2-13
Araucariaceae	*Araucaria araucana*	26	1-2-3-4-5-6-7-8-9-10-12-13
Aspleniaceae	*Asplenium dareoides*	144	1-2-3-4-5-6-7-8-10-11
Asteraceae	*Baccharis patagonica*	18	4-5-6-7-9
	Chaetanthera chilensis	22	1-4-5
	Chaetanthera elegans	22	11
	Erigeron andicola	36	5-13
	Gnaphalium viravira	28	6
	Haplopappus glutinosus	10	6-7
	Haplopappus grindelioides	10	4
	Hypochaeris acaulis	8	1-4-13
	Hypochaeris tenuifolia	8	3-4-6
	Lagenophora hariotii	18	2-7
	Podanthus ovatifolius	22	11
	Senecio subumbellatus	80	4
Berberidaceae	*Berberis empetrifolia*	14	1-2-3-4-5-6-10-13
	Berberis ilicifolia	28	8
	Berberis microphylla	28	1-2-3-4-5-6-7-8-9-10-12-13
Berberidopsidaceae	*Berberidopsis corallina*	42	11
Blechnaceae	*Blechnum chilense*	66	4-5-6-7-8-9-10-12
	Blechnum hastatum	66	1-2-3-4-5-6-7-8-9-10-11-12
	Blechnum mochaenum	66	5-6-11-12
	Blechnum penna-marina	66	1-2-3-4-5-6-7-8-9-12
Celastraceae	*Azara serrata*	18	9
Cupressaceae	*Austrocedrus chilensis*	22	2-5-6-8-13
Dryopteridaceae	*Megalastrum spectabile*	82	6-7-9-11-12
	Polystichum andinum	164	13
	Polystichum chilense	164	3-4-5-6-7-8-11
	Polystichum plicatum	164	3-4-5-6-7
	Polystichum subintegerrimum	328	4-6-12

Appendix 1. Continuation. Total species analysed in 13 PWA from Araucanía Region, Southern Chile (84 species). CN, chromosome number (2*n*). Numbers represent each PWA as described in Table 1.

Family	Species	CN	PWA where each species occur
Equisetaceae	*Equisetum bogotense*	216	1-2-4-5-6-7-8-9-11-12
Eremolepidaceae	*Lepidoceras chilense*	72	5-6-7
Fabaceae	*Adesmia boronioides*	20	5
Grossulariaceae	*Ribes magellanicum*	32	1-2-3-4-5-6-8-9-13
Hymenophyllaceae	*Hymenoglossum cruentum*	72	11-12
	Hymenophyllum caudiculatum	72	6-7-12
	Hymenophyllum dentatum	26	3-6-11-12
	Hymenophyllum ferrugineum	72	5
	Hymenophyllum tunbrigense	26	11
Iridaceae	*Herbertia lahue*	42	11-12
	Libertia chilensis	72	5-6-8-9-11-12
Lardizabalaceae	*Lardizabala biternata*	28	3-8-12
Lauraceae	*Persea lingue*	24	3-4-5-7-9-12
Lentibulariaceae	*Pinguicula antarctica*	16	9
Loasaceae	*Loasa acanthifolia*	38	2-3-5-6-8-11-12
Loranthaceae	*Desmaria mutabilis*	32	5-6-8-9
	Notanthera heterophylla	24	12
	Tristerix corymbosus	24	2-5-6-8-9-12
Monimiaceae	*Peumus boldus*	78	3-11-12
Myrtaceae	*Luma apiculata*	22	3-4-5-6-7-8-9-11-12
	Myrteola nummularia	44	5-6
	Ugni molinae	22	3-8-11-12
Nothofagaceae	*Nothofagus alpina*	26	1-2-3-4-5-6-7-8-9-10-11-12
	Nothofagus antarctica	26	1-2-3-4-5-6-7-8-9-10-13
	Nothofagus dombeyi	26	1-2-3-4-5-6-7-8-9-10-11-12-13
	Nothofagus obliqua	26	1-2-3-4-5-6-8-9-10-11-12
	Nothofagus pumilio	26	1-2-3-4-5-6-7-8-9-10-13
Onagraceae	*Fuchsia magellanica*	44	1-2-3-4-5-6-8-9-10-12
Philesiaceae	*Lapageria rosea*	30+2B	3-5-8-9-11-12
	Philesia magellanica	30	8
Plantaginaceae	*Ourisia coccinea*	16	3-6-7
Poaceae	*Danthonia araucana*	24	9
	Danthonia malacantha	48	11
	Festuca gracillima	42	10
Proteaceae	*Orites myrtoidea*	28	4-5
Pteridaceae	*Adiantum chilense*	116	1-2-3-4-5-6-7-8-9-11-12
Rubiaceae	*Galium araucanum*	22	11
	Galium hypocarpium	22	1-2-3-5-6-7-8-9-11-12-13
	Nertera granadensis	44	1-2-3-4-5-6-7-8-9-10-11-12

Appendix 1. Continuation. Total species analysed in 13 PWA from Araucanía Region, Southern Chile (84 species). CN, chromosome number (2*n*). Numbers represent each PWA as described in Table 1.

Family	Species	CN	PWA where each species occur
Rutaceae	*Pitavia punctata*	36	8
Solanaceae	*Solanum crispum*	24	8-10-11
	Solanum etuberosum	24	5
	Solanum tuberosum	24	4-6
Tecophilaeaceae	*Conanthera bifolia*	28	11
Verbenaceae	*Rhaphithamnus spinosus*	18	1-3-4-5-6-7-8-9-11-12

BIBLIOGRAPHY

Aagesen L., Sanso M. (2003) The phylogeny of the Alstroemeriaceae, based on morphology, *rps*16 Intron, and *rbc*L sequence data. Syst. Bot. 28 (1): 47-69.

Armesto J., Manuschevich D., Mora A., Smith Ramírez C., Rozzi R., Abarzúa A., Marquet P. (2010) From the Holocene to the Anthropocene: a historical framework for land cover change in southwestern South America in the past 15,000 years. Land Use Policy 27: 148-160.

Arriagada C. (2002) Estudio de la línea de base de flora del Parque Nacional Nahuelbuta. Temuco, Chile. Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Baeza M., Marticorena C., Rodríguez R. (1999) Catálogo de la flora vascular del Monumento Natural Contulmo. Gayana Bot. 56: 125-135.

Baeza C., Ruiz E., Negritto M. (2009) Importancia del cariotipo en la taxonomía y evolución del género *Chaetanthera* (Asteraceae): Evidencias preliminares para especies que crecen en Chile. Gayana Bot. 66 (1): 50-57.

Baeza C., Finot V., Ruiz E. (2015) Comparative karyotype analysis of populations in the *Alstroemeria presliana* Herbert (Alstroemeriaceae) complex in Chile. Gent. Mol. Biol. 38 (2): 199-204.

Baeza C., Finot V., Ruiz E., Carrasco P., Novoa P., Rosas M., Toro-Núñez O. (2018) Cytotaxonomic study of the Chilean endemic complex *Alstroemeria magnifica* Herb. (Alstroemeriaceae). Genetics and Molecular Biology: DOI:10.1590/1678-4685-GMB-2017-0157.

Bastías A., Correa F., Rojas P., Almada R., Muñoz C., Sagredo B. (2016) Identification and characterization of microsatellite loci in Maqui (*Aristotelia chilensis* [Molina] Stunz) using next generation sequencing (NGS). Plos One 11 (17): e0159825.doi:10.137/journal.pone.0159825.

Chacón J., Camargo de Asis M., Merrow A., Renner S. (2012a) From east Gondwana to Central America: historical biogeography of Alstroemeriaceae. J. Biogeogr. 39: 1806-1818.

Chacón J., Souza A., Baeza C., Renner S. (2012b) Ribosomal DNA distribution a genus wide phylogeny reveal patterns of chromosomal evolution in *Alstroemeria* (Alstroemeriaceae). Am. J. Bot. 99 (9): 1501-1512.

CONAF, Corporación Nacional Forestal (2009) Catálogo florístico Reserva Nacional Alto Biobío. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

CONAF, Corporación Nacional Forestal (2013) CONAF en las Áreas Silvestres Protegidas del Estado: conservando la flora y fauna amenazada. Santiago de Chile: Corporación Nacional Forestal.

Cortés M. (2005) Componente flora silvestre y vegetación para la elaboración del plan de manejo del Parque Nacional Villarrica. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Davis J., Stevenson D., Petersen G., Seberg O., Campbell L. Freudenstein, J., Goldman, D., Hardy, C., Michelangeli, F., Simmons, M., Specht, C., Vergara-Silva F., Gandolfo M. (2004) A phylogeny of the monocots, as inferred from *rbc*L and *atp*A sequence variation, and a comparison of methods for calculating jackknife and bootstrap values. Syst. Bot. 29: 467–510.

Delaveau C., Fuentes Arce G., Ruiz E., Hasbún R., Uribe M., Valenzuela S. (2013) Variabilidad genética mediante AFLP en tres relictos de *Gomortega keule* (Molina) Baillon: especie endémica chilena en peligro de extinción. Gayana Bot. 70: 188-194.

Fernández Donoso R., Berríos S. (1985) La arquitectura nuclear y su injerencia en la variabilidad del cariotipo. In: Fernández Donoso R. (Ed.) El núcleo, los cromosomas y la evolución. UNESCO, pp. 68-113.

Finckh M., Frank D., Moller A., Paulsch A., Thomas S. (1995) Análisis de las comunidades boscosas del Parque Nacional Villarrica (IX Región, Chile) y su conservación para el patrimonio natural. Temuco, Chile: Universidad de Bayreuth, Corporación Nacional Forestal.

García Gonzáles R., Carrasco B., Peñailillo P., Letelier L., Herrera R., Lavandero B., Moya M., Caligari P. (2008) Genetic variability and structure of *Gomortega keule* (Molina) Baillon (Gomortegaceae) relict populations: geographical and genetic fragmentation and its implications for conservation. Botany 86: 1299-1310.

Givnish T., Zuluaga A., Marques I., Lam V., Soto Gómez M., Iles W., Ames M., Spalink D., Moeller J., Briggs B., Lyon S., Stevenson D., Zomlefer W., Graham S. (2016) Phylogenomics and historical biogeography of the monocot order Liliales: out of Australia and through Antarctica. Cladistic 0: 1-25.

Gregory T. (2001) Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. Biol. Rev. 76: 65-101.

Hauenstein E. (2011a) Catastro, descripción y evaluación de asociaciones vegetacionales en dos unidades del SNASPE. Parque Nacional Huerquehue. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Hauenstein E. (2011b) Catastro, descripción y evaluación de asociaciones vegetacionales en dos unidades del SNASPE. Reserva Nacional Malalcahuello. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Hauenstein E., Saavedra M. (2019) Caracterización de la flora vascular presente en la Reserva de la Biosfera Araucarias. Temuco: Corporación Nacional Forestal. Reserva de la Biosfera Araucarias: un territorio para Chile y el Mundo, pp. 60-90.

Heslop-Harrison J., Schwarzacher T. (2011) Organisation of the plant genome in chromosomes. Plant J. 66:18-33

Jara Arancio P., Arroyo M., Guerrero P., Hinojosa L., Arancio G., Mendez M. (2013) Phylogenetic perspectives on biome shifts in *Leucocoryne* (Alliaceae) in relation to climatic niche evolution in western South America. J. Biogeogr. 41: 328-338.

Jara Seguel P., Zúñiga C. (2004) El cariotipo de *Lapageria rosea* Ruiz et Pav. (Liliales: Philesiaceae). Gayana Bot. 61: 76-78.

Jara Seguel P., Romero Mieres M., Palma Rojas C. (2006) Chromosome numbers of Chilean Pteridophytes: first contribution. Gayana Bot. 63: 115-118.

Jara Seguel P., Zúñiga C., Romero Mieres M., Palma Rojas C., von Brand E. (2010) Karyotype study in *Luzuriaga radicans* (Liliales: Luzuriagaeae). Biología 65: 813-816.

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

Jara Seguel P., Cárcamo Fincheira P., Palma Rojas C., von Brand E. (2014) Chromosome number of two Chilean species of *Nothofagus* (Nothofagaceae). Gayana Bot. 71 (2): 287-289.

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

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

Levin D., Wilson A. (1996) Rates of evolution in seeds plants: Net increase in diversity of chromosome numbers and species numbers through time. Proc. Natl. Acad. Sci. USA 73 (8): 2086-2090.

Levin D. (2002) The role of chromosomal change in plant evolution. Oxford University Press, New York.

Marín C., Aguilera P., Fritz O., Godoy R. (2017) Factors affecting arbuscular mycorrhizal fungi of Chilean temperate rainforests. J. Soil Sci. Plant Nut. 17 (4): 966-984.

Martín M., Mattioni C., Lusini I., Molina J., Cherubini M., Drake F., Herrera M., Villani F., Martín L. (2014) New insights into the genetic structure of *Araucaria araucana* forests based on molecular and historic evidences. Tree Genet. Genomes 10: 839-851.

Montgomery D., Balco G., Willett S. (2001) Climate, tectonics, and the morphology of the Andes. Geology 29: 579-582.

Moreira Muñoz A. (2011) Plant geography of Chile. Series Plant and Vegetation. Dordrecht, Springer.

Morero R., Chiarini F., Urdampilleta J., Barboza G., Barrington D. (2015) Cytological study of *Polystichum* (Dryopteridaceae) species from southern South America. Aust. J. Bot. 63: 403-414.

Pauchard A., Villarroel P. (2002) Protected areas in Chile: history, current status, and challenges. Nat. Area J. 22: 318-330.

Peruzzi L., Góralski G., Joachimiak A., Bedini G. (2012) Does actually mean chromosome number increase with latitude in vascular plants? An answer from the comparison of Italian, Slovak and Polish floras. Comp. Cytogenet. 6: 371-377.

Peruzzi L., Caparelli K., Bedini G. (2014) A new index for the quantification of chromosome number variation: an application to selected animal and plant groups. J. Theor. Biol. 353: 55-60.

Pielou E. (1975) Ecological diversity. John Wiley & Sons, New York.

Premoli A. (1997) Genetic variation in a geographically restricted and two widespread species of South American *Nothofagus*. J. Biogeogr. 24: 883-892.

Premoli A., Kitzberger T., Veblen T. (2000) Conservation genetics of the endangered conifer *Fitzroya cupressoides* in Chile and Argentina. Conserv. Genet. 1: 57-66.

Premoli A., Mathiasen P. (2011) Respuestas ecofisiológicas adaptativas y plásticas en ambientes secos de montaña: *Nothofagus pumilio*, el árbol que acaparó los Andes australes. Ecología Austral 21: 251-269.

Premoli A., Acosta M., Mathiasen M., Donoso C. (2012) Variación genética en *Nothofagus* (Subgénero *Nothofagus*). Bosque 33: 115-125.

Ramírez C., Labbe S., San Martín C., Figueroa H. (1990) Sinecología de los bosques de boldo (*Peumus boldus*) de la cuenca del Río Bueno, Chile. Bosque 11: 45-56.

Saavedra M. (2009a) Catastro de flora silvestre de la Reserva Nacional Malleco. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Saavedra M. (2009b) Línea de base de la flora de la Reserva Nacional Alto Biobío, para la elaboración de su plan de manejo. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Saavedra M. (2009c) Línea de base de la flora silvestre del Parque Nacional Tolhuaca. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Saavedra M., Hauenstein E. (2010a) Línea de base de flora y vegetación de la Reserva Nacional China Muerta. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Saavedra M., Hauenstein E. (2010b) Línea de base de la flora y vegetación de la Reserva Nacional Nalcas. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Saavedra M., Morales S. (2008) Línea de base de la flora del Monumento Natural Cerro Ñielol, para elaboración de su plan de manejo. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Scherson R., Thornhill A., Urbina Casanova R., Freyman W., Pliscoff P., Mishler B. (2017) Spatial phylogenetics of the vascular flora of Chile. Mol. Phylogenet. Evol. 112: 88-95.

Schneider R., Grosschedl R. (2007) Dynamics and interplay of nuclear architecture, genome organization and gene expression. Genes Dev. 21: 302-304.

Sepúlveda P. (2004) Línea de base, componente vegetacional Parque Nacional Conguillío. Temuco, Chile: Departamento de Áreas Protegidas y Patrimonio Silvestre, Corporación Nacional Forestal.

Severns P., Liston A. (2008) Intraspecific chromosome number variation: a neglected threat to the conservation of rare plants. Conserv. Biol. 22: 1641-1647.

Smith A. (1973) Stratification of temperate and tropical forests. Am. Nat. 107: 671-683.

Smith Ramírez C., Díaz I., Pliscoff P., Valdovinos C., Méndez M., Larraín J., Samaniego H. (2007) Distribution patterns of flora and fauna in southern Chilean Coastal rain forests: Integrating Natural History and GIS. Biodivers. Conserv. 16: 2627-2648.

Stebbins L. (1971) Chromosomal evolution

in higher plants. University Park Press, London.

Torres Díaz C., Ruiz E., González F., Fuentes G., Cavieres L. (2007) Genetic diversity in *Nothofagus alessandrii* (Fagaceae), an endangered endemic tree species of the coastal maulino forest of Central Chile. Ann. Bot. 100: 75-82.

Troncoso A., Villagrán C., Muñoz M. (1980) Una nueva hipótesis acerca del origen y edad del bosque de Fray Jorge (Coquimbo, Chile). Bol. Mus. Nac. Hist. Nat. 37: 117-152.

Van Der Hammen T., Cleef A. (1983) Datos para la historia de la flora andina. Rev. Chil. Hist. Nat. 56: 97-107.

Vellend M. (2005) Species diversity and genetic diversity: parallel processes and correlated patterns. Am. Nat. 66: 199-215.

Villagrán C., Hinojosa L. (1997) Historia de los bosques del sur de Sudamérica II: análisis fitogeográfico. Rev. Chil. Hist. Nat. 70: 241-267.

Windham M., Yatskievych G. (2003) Chromosome studies of Cheilanthoid ferns (Pteridaceae: Cheilanthoideae) from the Western United States and Mexico. Am. J. Bot. 90: 1788-1800.

ACKNOWLEDGEMENTS

The authors would like to express their appreciation to the Núcleo de Estudios Ambientales (NEA), Universidad Católica de Temuco, for funding the current implementation of the Chilean Plants Cytogenetic Database (CPCD). Our thanks also to Paola Jara Arancio and anonymous reviewers for reading the manuscript and providing valuable comments.

BAG | Journal of Basic & Applied Genetics

Vol XXXI (2): 39-44; December 2020

Journal of the Argentine Society of Genetics

CYSTIC HYGROMA AND THE IMPORTANCE OF THE PRENATAL DIAGNOSIS: ABOUT A CASE

HIGROMA QUÍSTICO Y LA IMPORTANCIA DE SU DIAGNÓSTICO PRENATAL: A PROPÓSITO DE UN CASO

Almonacid I.[1], Zuleta P.[2], Neri C.[2], Zannin A.[2], Garzón M.[2], Dorado J.[2], Medina F.[3], Celis L.G.[2]

[1] Departamento de Patología, Hospital Central de la Policía, Bogotá, Colombia;

[2] Facultad de Medicina, Universidad de la Sabana, Bogotá, Colombia;

[3] Departamento de Radiología, Hospital Militar Central, Bogotá, Colombia.

Corresponding author:
Luis Gustavo Celis
luis.celis@unisabana.edu.co

0000-0002-0338-6258

Cite this article as:
Almonacid I., Zuleta P., Neri C., Zannin A., Garzón M., Dorado J., Medina F., Celis L.G. 2020. CYSTIC HYGROMA AND THE IMPORTANCE OF THE PRENATAL DIAGNOSIS: ABOUT A CASE. BAG. Journal of Basic and Applied Genetics XXXI (2): 39-44..

Received: 04/14/2020
Revised version received: 09/19/2020
Accepted: 10/22/2020
General Editor: Elsa Camadro
DOI: 10.35407/bag.2020.31.02.04
ISSN online version: 1852-6233

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

ABSTRACT

The cystic hygroma is the malformation of the lymphatic system that is most frequently observed in the prenatal period and is located mainly in the neck and/or the nape of the neck. Its detection rate has increased since the implementation of fetal nuchal translucency (NT) in the first trimester of pregnancy and its presence has been associated with congenital abnormalities, aneuploidies, pregnancy loss, and developmental disorders. The aim of this case is to highlight the importance of antenatal diagnosis of cystic hygroma in order to perform early intervention and avoid fetal death. It is received, for anatomopathological study, a fetus of undetermined sex product of the first pregnancy of a 19 year-old mother without previous prenatal controls, with the presence of a large cystic mass that extends from the face to the neck. The histological study confirms the diagnosis of cystic hygroma. As there was no karyotype analysis, it was not possible to establish the preexistence of any genetic abnormality. Also known as cystic lymphangioma, is a benign vascular tumor whose antenatal diagnosis by ultrasonography is essential in the evolution and prognosis of the disease. Unfortunately in our case, the lack of prenatal controls and the absence of ultrasonographic studies that would allow knowing the characteristics of this lymphangioma, could significantly impact in the fatal outcome.

Key words: lymphangioma; prenatal diagnosis; fetal nuchal translucency.

RESUMEN

El higroma quístico es la malformación del sistema linfático que más frecuentemente se observa en el período prenatal y que se ubica principalmente en el cuello y/o la nuca. Su tasa de detección ha aumentado desde la implementación de la translucencia nucal fetal (TN) en el primer trimestre de embarazo, y su presencia se ha relacionado con anomalías congénitas, aneuploidías, pérdida del embarazo y trastornos en el desarrollo. El objetivo de la presentación de este caso es resaltar la importancia del diagnóstico antenatal del higroma quístico, con el fin de realizar una intervención precoz y evitar la muerte fetal. Se recibe para estudio anatomopatológico, feto de sexo indeterminado producto del primer embarazo de una madre de 19 años de edad sin previos controles prenatales, con presencia de una gran masa quística que se extiende desde el rostro hasta la nuca. Mediante el estudio histológico se confirma el diagnóstico de higroma quístico. Al carecer de análisis de cariotipo no fue posible establecer la preexistencia de alguna anomalía genética. El también conocido como linfangioma quístico, es un tumor vascular benigno cuyo diagnóstico antenatal mediante la ultrasonografía resulta fundamental en la evolución y pronóstico de la enfermedad. Desafortunadamente en nuestro caso, la falta de controles prenatales y la ausencia de estudios ultrasonográficos que permitieran conocer las características de este linfangioma, pudo impactar significativamente en el desenlace fatal.

Palabras clave: linfangioma; diagnóstico prenatal; translucencia nucal fetal.

INTRODUCCIÓN

El higroma quístico es una malformación del sistema linfático que se origina desde el desarrollo embrionario (Aymelek *et al.*, 2019), pudiendo ser de carácter transitoria o persistente (Özcan *et al.*, 2019). La ubicación más frecuente es a nivel cervical y se puede asociar a síndromes genéticos y anormalidades cromosómicas (Schreurs *et al.*, 2018). El diagnóstico del higroma quístico se hace a través de la ultrasonografía midiendo la translucencia nucal fetal en el primer trimestre del embarazo (Cicatiello *et al.*, 2019). Como exámenes complementarios, tenemos el análisis de cariotipo y el *microarray* cromosómico para confirmar la presencia o ausencia de aneuploidías (Schreurs *et al.*, 2018; Cicatiello *et al.*, 2019). Generalmente este tipo de malformación tiene pobre pronóstico, particularmente si se asocia a otras condiciones como la hidropesía fetal (Özcan *et al.*, 2019). El objetivo de la presentación de este caso es resaltar la importancia del diagnóstico antenatal del higroma quístico, con el fin de realizar una intervención precoz y evitar la muerte fetal.

MATERIALES Y MÉTODOS

Declaración ética, consentimiento y permisos

Este estudio se realizó bajo los lineamientos de la Declaración de Helsinki, así como la resolución N° 8430 de 1993 y N° 2378 de 2008 del Ministerio de Salud y Protección Social de Colombia, los cuales velan por los intereses de los sujetos estudiados. Se brindó la información necesaria a familiares, en este caso a los padres por ser los representantes legales, se aclararon dudas y posteriormente se firmó el consentimiento informado previo al inicio del estudio.

Metodología

Se recibe al producto del primer embarazo de una mujer de 19 años de edad, remitido a la Facultad de Medicina de la Universidad de La Sabana, para estudios anatomopatólogicos, genéticos e imagenológicos complementarios. Posteriormente, se realiza la búsqueda de literatura en bases de datos como PubMed, Clinicalkey, Science Direct y OMIM para presentar el caso clínico reportado y la información disponible actualizada.

Reporte de caso

Se recibe feto de sexo indeterminado producto de primer embarazo de madre de 19 años de edad, sin previos controles prenatales, quien al examen macroscópico presentaba una longitud cráneo-plantar de 21 centímetros y 7 centímetros de perímetro cefálico, con presencia de una gran masa quística de 7 cm de largo por 4 cm de ancho, que se extiende desde la región frontal hasta la nuca (Figura 1). Al corte impresiona una masa quística, unilocular, de superficie lisa y pared menor de 0,1 cm ocupada por contenido acuoso, sin conexión a meninges ni encéfalo, de aspecto normal y con severa autolisis (Figura 2).

Los huesos cráneo-faciales son de características usuales (Figura 3A). En cara, los ojos están presentes y normalmente localizados, la boca es normal, mientras que la piel tanto de la nariz como de la cara están severamente edematosas. Los miembros superiores e inferiores están presentes y poseen un marcado edema. En la descripción interna de los órganos al corte, los pulmones se encuentran presentes y atelectásicos. El corazón es pequeño, en el hígado hay autolisis y los riñones están presentes pero disminuidos en tamaño (Figura 3B). A su vez, la placenta presenta focos de consistencia firmes y en el cordón umbilical se observan dos vasos sanguíneos (Figura 3C).

RESULTADOS

Los cortes histológicos con Hematoxilina Eosina mostraron una dilatación de los conductos linfáticos asociados a espacios vasculares grandes e irregulares revestidos por células epiteliales aplanadas con un estroma fibroblástico con linfocitos en su interior, hallazgos que corresponden a un tumor vascular benigno de tipo linfangioma quístico (Figura 4). Esta malformación corresponde a una anomalía de los vasos linfáticos yugulares que impide el correcto drenaje de la linfa (Noia *et al.*, 2019). Un fallo en el canal comunicante o su bloqueo, permite la formación de quistes de retención en región cervical (Schreurs *et al.*, 2018). Asimismo, se asocia a anomalías cromosómicas tales como: Síndrome de Turner, Síndrome de Klinefelter y trisomías 13, 18 y 21 (Özcan *et al.*, 2019). Lamentablemente en este caso, la madre sufre la pérdida gestacional antes de la realización del análisis de cariotipo, además durante el examen postmortem, se tomaron muestras en sangre periférica y bazo buscando obtener células viables pero fue imposible puesto que la toma de muestra para cariotipo se realizó 48 horas posteriores al aborto. Por lo anterior, no se pudo establecer una asociación directa entre el higroma quístico y la presencia de alguna alteración cromosómica concomitante que contribuyera al desenlace final.

DISCUSIÓN

El higroma quístico, también conocido como linfangioma quístico, es una malformación vascular/linfática congénita caracterizada por la dilatación de los conductos y los sacos linfáticos debido a un fallo en la

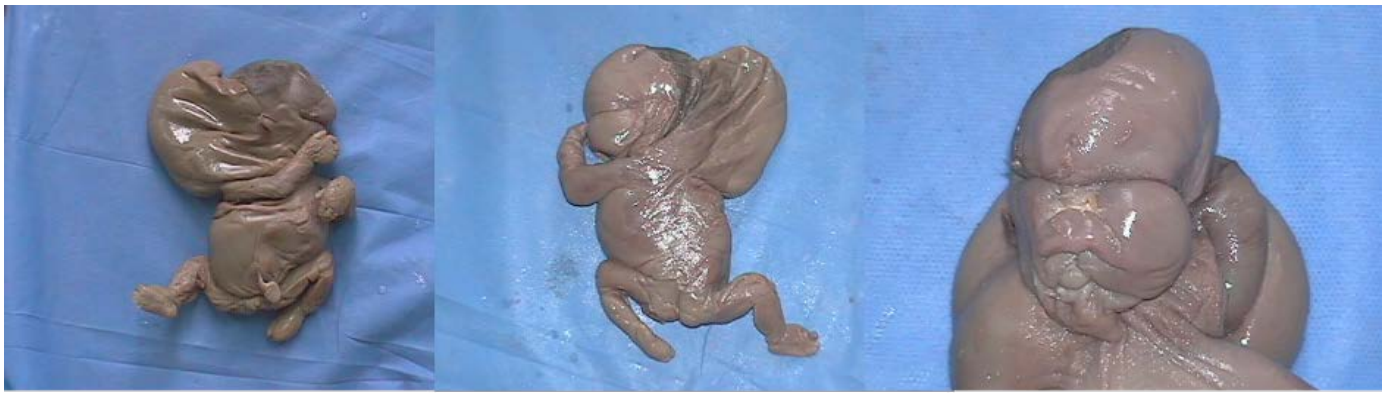

Figura 1. Feto de sexo indeterminado con presencia de una gran masa quística que abarca rostro, cuero cabelludo y nuca.

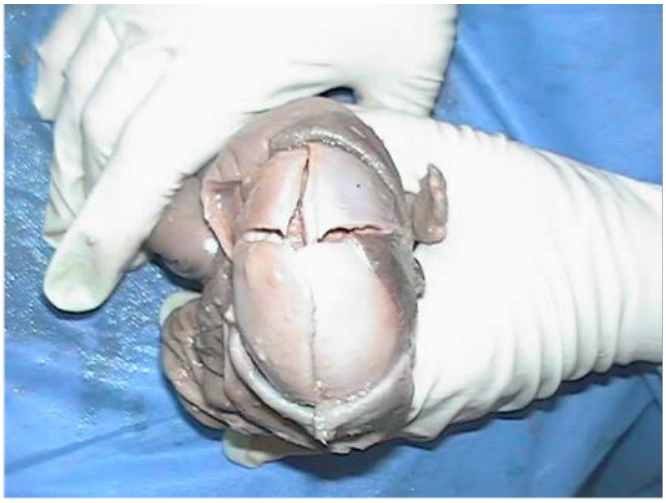

Figura 2. Al corte se observa una gran cavidad quística de aspecto normal y contenido acuoso, sin conexión a meninges ni encéfalo. Las estructuras óseas son de aspecto usual.

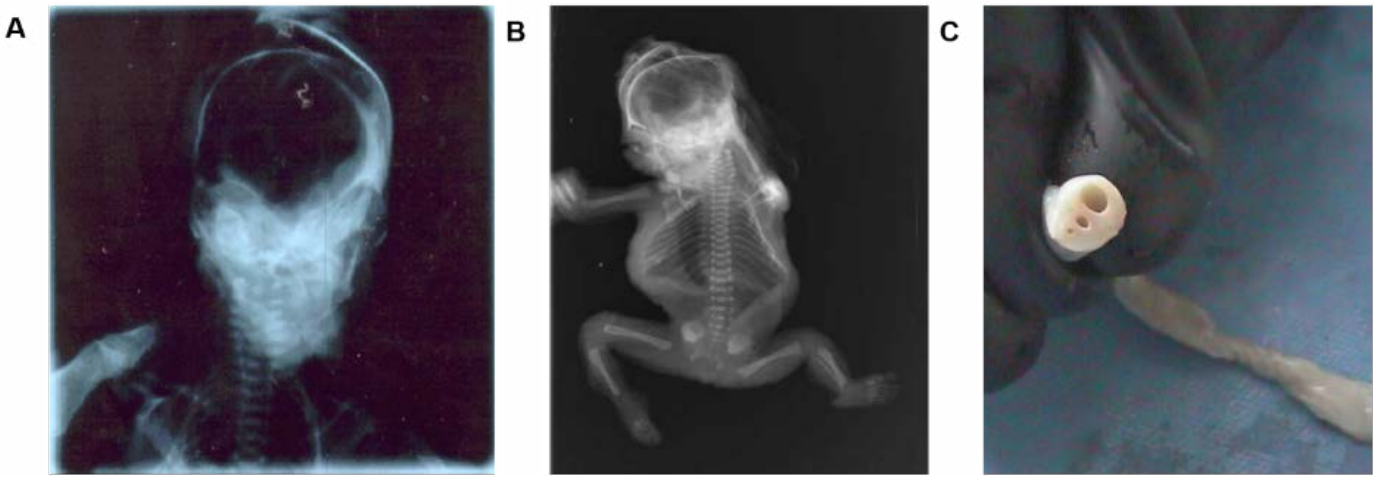

Figura 3. A) Los huesos de la cara son de características usuales; B) Se observa atelectasia pulmonar y disminución del tamaño del corazón y ambos riñones, el resto de órganos sin anomalías; C) En el cordón umbilical se observan dos vasos sanguíneos.

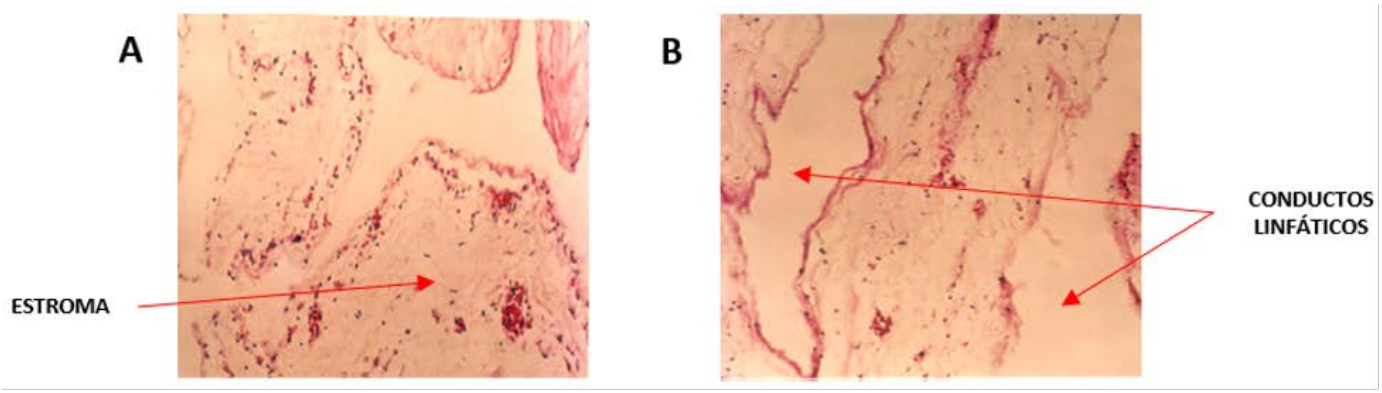

Figura 4. A) Estroma fibroblástico con linfocitos en su interior; B) Dilatación de los conductos linfáticos (H&E x 100)

conexión entre el sistema linfático y venoso (Schreurs *et al.*, 2018; Noia *et al.*, 2019). Esta malformación se encuentra frecuentemente ubicada en cuello/nuca (75%), axilas (20%), retroperitoneo (5%), extremidades (2%) y mediastino (1%) (Chen y Zheng, 2019; Chen *et al.*, 2017). Los higromas quísticos pueden clasificarse según el tamaño en microquísticos (<2 cm), macroquísticos (>2 cm) y mixtos (masa compuesta de múltiples quistes separados por septos) (Munteanu *et al.*, 2016). A su vez, también puede clasificarse en septado y no septado. La incidencia del linfangioma quístico se estima en 1/1.000-6.000 nacidos vivos y 1/750 abortos espontáneos (Yakıştıran *et al.*, 2019). En nuestro caso, se trató de un higroma macroquístico septado en nuca, que por sus características histopatológicas, prometía un final desalentador.

El higroma quístico en nuca se ha asociado desde un 50% a un 80% a aneuploidías cromosómicas (Yakıştıran *et al.*, 2019). Entre las aneuploidías más comúnmente observadas tenemos el Síndrome de Turner (60%), el Síndrome de Down, el Síndrome de Klinefelter y las trisomías 18 y 13 (Aymelek *et al.*, 2019; Kadam *et al.*, 2017). En un análisis realizado a 116 fetos con Síndrome de Turner se observó que las anomalías estructurales encontradas fueron, en orden de frecuencia, el higroma quístico (59,5%), la hidropesía fetal (19%), anomalías cardíacas congénitas (7,8%), defecto pulmonar (3,5%), defecto renal (2,6%) y anomalías en el sistema nervioso central (1,7%) (Chen *et al.*, 2018). Del mismo modo, este tumor puede encontrarse en el 22-32% de los fetos con cariotipo normal (Özcan *et al.*, 2019). Sin embargo, en el presente estudio, no fue posible la toma de cariotipo por lo que no se pudo establecer la preexistencia de alguna anomalía genética.

El patrón de oro para el diagnóstico antenatal del linfangioma quístico en nuca es la ultrasonografía, procedimiento no invasivo y costo-efectivo (Munteanu *et al.*, 2016). Al final de la quinta semana de gestación, inicia el desarrollo del sistema linfático en el feto (Schreurs *et al.*, 2018; Munteanu *et al.*, 2016). Para la octava semana del embarazo, ya se encuentran desarrollados los primeros seis sacos linfáticos e inicia la comunicación entre los conductos linfáticos yugulares izquierdos y derechos (Aymelek *et al.*, 2019). En el ultrasonido, la imagen del linfangioma quístico puede verse como un área de sonolucencia del tejido blando que consiste en una cavidad simétrica que puede o no estar separada por uno o varios septos (Chen *et al.*, 2017). De esta premisa parte la realización del examen de translucencia nucal fetal para el diagnóstico ultrasonográfico del higroma o linfangioma quístico, el cual cobra vital importancia en el momento de hacer una intervención precoz en estos pacientes y evitar la muerte fetal. Sin embargo, en el caso anteriormente descrito, la falta de controles prenatales no permitió la realización de ningún examen de ultrasonido por lo que no se pudo realizar el diagnóstico antenatal de dicha condición.

La medición de la translucencia nucal (TN) se ha convertido en una herramienta útil en el tamizaje de aneuploidías en el primer trimestre de gestación. De acuerdo con la Fundación de Medicina Fetal, la edad gestacional óptima para medir la TN es desde la semana 11 a la semana 13,6 del embarazo (Scholl y Chasen, 2016). Una medida superior al percentil 99 (>3,5 mm) se asocia a un riesgo aumentado de padecer más de 50 condiciones, entre ellas, las anomalías cromosómicas, las malformaciones fetales, el daño cardíaco, ciertos síndromes genéticos y hasta la muerte fetal (Cicatiello *et al.*, 2019). Cuando el higroma quístico es diagnosticado, el feto debe someterse a procedimientos más invasivos como el análisis de cariotipo (Yakıştıran *et al.*, 2019). En los casos donde haya un aumento de la TN con un cariotipo normal, se recomienda la realización del microarray cromosómico o cariotipo molecular (Schreurs *et al.*, 2018). Dicha técnica nos permite identificar deleciones o duplicaciones cromosómicas no diagnosticadas previamente. En un meta-análisis reciente practicado en fetos con aumento de la TN, el

microarray cromosómico detectó un 4% de variaciones en el número de copias que no fueron observadas en el cariotipo convencional (Cicatiello *et al.*, 2019).

Adicionalmente, existe una relación entre el tipo de aneuploidía y la edad gestacional en la que se diagnostica por primera vez el linfangioma quístico de nuca. Los higromas diagnosticados durante el primer trimestre de embarazo tienden a asociarse a las trisomías 21, 18 y 13 (Gezdirici *et al.*, 2017). Mientras que, en el segundo trimestre de embarazo, los higromas quísticos no septados se relacionan al Síndrome de Down, y los higromas quísticos septados al Síndrome de Turner (Orgul *et al.*, 2017). Como la gran mayoría de los higromas quísticos son diagnosticados en el primer trimestre de embarazo, es de suma importancia realizar un adecuado diagnóstico de los síndromes frecuentemente asociados. La prueba de ADN de células fetales libres en suero materno es un examen no invasivo con una excelente sensibilidad y especificidad en el diagnóstico de las trisomías 21, 18 y 13 (Cicatiello *et al.*, 2019; Reimers *et al.*, 2018). En un análisis reciente, se encontró que esta prueba tiene una tasa de detección para la trisomía 21 del 99% con una tasa de falsos positivos del 0,1%. Sin embargo, la prueba de ADN de células fetales libres es costosa por lo que aún no puede ser considerada como parte del tamizaje rutinario de aneuploidías (Vičić *et al.*, 2017).

En cuanto al pronóstico, en términos generales, los higromas quísticos diagnosticados en fases tempranas del embarazo, tienen peor pronóstico que aquellos diagnosticados en el tercer trimestre de gestación (Schreurs *et al.*, 2018). Otros factores que empeoran el pronóstico en pacientes con higroma quístico son: la presencia de anomalías cromosómicas, la hidropesía fetal, una translucencia nucal mayor a 6 mm, los higromas quísticos septados y la asociación con otra malformación mayor (Chen *et al.*, 2017). La combinación de linfangioma quístico e hidropesía fetal resulta en muerte fetal en un 96,5% de los casos (Özcan *et al.*, 2019). A pesar de lo anterior, se ha descripto que el 42% de los higromas quísticos que alcanzan el término del embarazo desaparecen al nacimiento (Kadam *et al.*, 2017). También se ha descripto la resolución espontánea del higroma antes de la semana 20 de gestación, lo que es considerado factor de buen pronóstico. Sin embargo, en estos casos, se recomienda continuar seguimiento estricto con ultrasonografía y cariotipo mensual (Özcan *et al.*, 2019).

Fetos con aumento de la translucencia nucal y cariotipo normal, sin otra anomalía congénita, tienen una tasa de mortalidad perinatal del 3,3% al 11,8% (Noia *et al.*, 2019). Pero, en caso de presentar un cariotipo anormal, la terminación del embarazo debe tomarse en consideración dado al pobre pronóstico (Munteanu *et al.*, 2016). En cuanto al manejo intrauterino, la escleroterapia debe considerarse en casos donde exista el diagnóstico de higroma quístico en ausencia de hidropesía fetal y anomalías cromosómicas y/o estructurales, o cuando se trate de un higroma macroquístico bajo las mismas condiciones (Waner, 2018). En nuestro caso, y suponiendo que el cariotipo era normal por tratarse de un higroma macroquístico, existía la posibilidad del manejo con agentes esclerosantes. A su vez, el manejo quirúrgico debe considerarse en los casos con obstrucción de la vía aérea, después de la escleroterapia no exitosa y en pacientes con enfermedad microquística (García *et al.*, 2018). Cuando el diagnóstico de higroma quístico se hace en el tercer trimestre de gestación y/o en embarazos a término, no existe un consenso sobre la modalidad de parto. Sin embargo, algunos especialistas prefieren el parto vaginal sobre la cesárea (Munteanu *et al.*, 2016). Debemos tener en cuenta que un gran número de higromas quísticos desaparecen al momento del nacimiento, mientras que otros persisten y deben ser manejados quirúrgicamente a futuro (Waner, 2018).

CONCLUSIÓN

El linfangioma quístico es una malformación congénita rara de pobre pronóstico; traemos a colación este caso para resaltar la importancia del diagnóstico prenatal temprano de las patologías congénitas, por lo que resulta fundamental la realización de los controles prenatales en todas las gestantes con el fin de identificar de manera oportuna alguna malformación que pudiese acarrear consigo desenlaces no favorables. La medición ultrasonográfica de la translucencia nucal fetal, es de gran utilidad en el diagnóstico temprano del higroma quístico. En el caso expuesto, al tratarse de una madre adolescente con nulos controles prenatales, fue imposible anticipar el diagnóstico de dicha malformación, lo que no permitió realizar ningún tipo de intervención y, desafortunadamente, culminó con la pérdida gestacional.

BIBLIOGRAFÍA

Aymelek H.S., Oğur G., Tosun M. (2019) Genetic Burden and Outcome of Cystic Hygromas Detected Antenatally: Results of 93 Pregnancies from a Single Center in the Northern Region of Turkey. J. Med. Ultrasound. 10; 27 (4): 181-186. [PMID: 31867191].

Chen C.P., Chang S.Y., Lau H.S. (2018) First-trimester cystic hygroma and omphalocele in a fetus with Turner syndrome. Taiwan J. Obstet. Gynecol. 57 (5): 763-764. [PMID: 30342669].

Chen Y.N., Chen C.P., Lin C.J. (2017) Prenatal Ultrasound Evaluation and Outcome of Pregnancy with Fetal Cystic Hygromas and Lymphangiomas. J. Med. Ultrasound. 25 (1): 12-15. [PMID: 30065449].

Chen H.Y., Zheng J.Q. (2019) A case report of Turner syndrome associated with fetal nuchal cystic hygroma and bilateral syndactyly of the hands and feet. Ital. J. Pediatr. 18; 45 (1): 85. [PMID: 31319890].

Cicatiello R., Pignataro P., Izzo A. (2019) Chromosomal Microarray Analysis versus Karyotyping in Fetuses with Increased Nuchal Translucency. Med. Sci. 7 (3). [PMID: 30818867].

García S., Padilla A., Martínez I. (2018) Diagnóstico y pronóstico prenatal de linfangiomas fetales. Reporte de dos casos. Ginecol. Obstet. Mex. 86 (12): 831-840. [DOI: https://doi.org/10.24245/gom.v86i12.2112].

Gezdirici A., Ekiz A., Güleç E.Y. (2017) How necessary is to analyze PTPN11 gene in fetuses with first trimester cystic hygroma and normal karyotype? J. Matern. Fetal Neonatal Med. 30 (8): 938-941. [PMID: 27193571].

Kadam R.M., Kumar A.N., Prasad V. (2017) Giant Cervico-Thoracic Cystic Hygroma in a Preterm: A Case Report. J. Neonatal. Surg. 10; 6 (3): 66. [PMID: 28920026].

Munteanu O., Cîrstoiu M.M., Filipoiu F.M. (2016) Morphological and ultrasonographic study of fetuses with cervical hygroma. A cases series. Rom. J. Morphol. Embryol. 57 (4): 1421-1427. [PMID: 28174814].

Noia G., Maltese P.E., Zampino G. (2019) Cystic Hygroma: A Preliminary Genetic Study and a Short Review from the Literature. Lymphat Res. Biol. 17 (1): 30-39. [PMID: 30475086].

Orgul G., Ozyuncu O., Oktem A. (2017) Management and outcomes of cystic hygromas: experience of a tertiary center. J. Ultrasound. 20 (2): 127-131. [PMID: 28593002].

Özcan H.Ç., Uğur M.G., Balat Ö. (2019) Analysis of cystic hygroma diagnosed in the prenatal period: 5-years' experience at a tertiary hospital in Southeastern Turkey. J. Matern. Fetal Neonatal Med. 32 (11): 1800-1805. [PMID: 29241391].

Reimers R.M., Mason Suares H., Little S.E. (2018) When ultrasound anomalies are present: An estimation of the frequency of chromosome abnormalities not detected by cell-free DNA aneuploidy screens. Prenat. Diagn. 38 (4): 250-257. [PMID: 29436713].

Scholl J., Chasen S.T. (2016) First trimester cystic hygroma: does early detection matter? Prenat. Diagn. 36 (5): 432-6. [PMID: 26934494].

Schreurs L., Lannoo L., De Catte L. (2018) First trimester cystic hygroma colli: retrospective analysis in a tertiary center. Eur. J. Obstet. Gynecol. Reprod. Biol. 231: 60-64. [PMID: 30321790].

Vičić A., Hafner T., Bekavac Vlatković I. (2017) Prenatal diagnosis of Down syndrome: A 13-year retrospective study. Taiwan J. Obstet. Gynecol. 56 (6): 731-735. [PMID: 29241910].

Waner M. (2018) Multidisciplinary Approach to the Management of Lymphatic Malformations of the Head and Neck. Otolaryngol. Clin. North Am. 51 (1): 159-172. [PMID: 29217060].

Yakıştıran B., Altınboğa O., Canpolat E. (2019) Analysis of cystic hygroma diagnosed in the first trimester: single-center experience. J. Turk. Ger. Gynecol. Assoc. [PMID: 31298510].

BAG | Journal of Basic & Applied Genetics

Vol XXXI (2): 45-53; December 2020

Journal of the Argentine Society of Genetics

ALLELES ASSOCIATED TO DISEASE SEVERITY INDEX OF MAL DE RÍO CUARTO DISEASE IN MAIZE EXOTIC GERMPLASM

ALELOS ASOCIADOS AL ÍNDICE DE SEVERIDAD DE LA ENFERMEDAD MAL DE RÍO CUARTO EN GERMOPLASMA EXÓTICO DE MAÍZ

Rossi E.A.[1,2], Ruiz M.[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:
Ezequiel Rossi
erossi@ayv.unrc.edu.ar

0000-0003-2043-6760

Cite this article as:
Rossi E.A., Ruiz M., Bonamico N.C., Balzarini M.G. 2020. ALLELES ASSOCIATED TO DISEASE SEVERITY INDEX OF MAL DE RÍO CUARTO DISEASE IN MAIZE EXOTIC GERMPLASM. BAG. Journal of Basic and Applied Genetics XXXI (2): 45-53.

Received: 05/12/2020
Revised version received: 09/15/2020
Accepted: 10/14/2020
General Editor: Elsa Camadro
DOI: 10.35407/bag.2020.31.02.05
ISSN online version: 1852-6233

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

ABSTRACT

Mal de Río Cuarto (MRC) is one of the most important viral diseases of maize in Argentina. The disease severity index (DSI) allows to combine the incidence and severity of a disease in a single metric. The genotypic reaction to MRC has been extensively studied in biparental populations. However, this complex trait has not been analyzed by genome-wide association studies (GWAS). The aim of this work is to identify new resistance alleles associated with DSI of MRC in an exotic germplasm from the International Maize and Wheat Improvement Center (CIMMYT). A population of maize lines from CIMMYT was phenotypically evaluated in environments in the area where the disease is endemic. The predictors of genetic effects (BLUP, best linear unbiased predictor) and 78,376 SNP markers (Single Nucleotide Polymorphism) were used to perform the GWAS in 186 maize lines. The values of variance components and mean-basis heritability suggest a wide genotypic variability in the population. The GWAS allowed to identify 11 putative QTL of resistance to MRC. The incorporation of exotic germplasm into local maize breeding programs could contribute favorably to the creation of hybrids with a higher level of resistance to MRC. The predictive ability of associated markers with MRC resistance indicates that marker-assisted selection is an advisable tool for selecting MRC resistant genotypes.

Key words: Disease severity index; genome-wide association study; QTL; SNP.

RESUMEN

El Mal de Río Cuarto (MRC) es una de las enfermedades virales más importantes del maíz en Argentina. El índice de severidad de enfermedad (ISE) permite combinar la incidencia y la severidad de una enfermedad en una métrica única. La reacción genotípica a MRC ha sido muy estudiada en poblaciones biparentales, sin embargo este carácter complejo no se ha analizado mediante estudios de mapeo por asociación. El objetivo del presente trabajo es identificar nuevos alelos de resistencia asociados con el ISE de la enfermedad MRC de maíz en un germoplasma exótico del Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT). Una población de líneas de maíz del CIMMYT se evaluó fenotípicamente en ambientes donde la enfermedad es endémica. Los predictores del efecto genotípico (BLUP, *best linear unbiased predictor*) del ISE de MRC y 78.376 marcadores SNP (*Single Nucleotide Polymorphism*) se usaron para realizar el mapeo por asociación en 186 líneas de maíz. Los componentes de varianza y los valores de heredabilidad sugieren una amplia variabilidad genotípica en la población de líneas. El mapeo por asociación permitió identificar 11 posibles QTL de resistencia a MRC. La incorporación de germoplasma exótico en los programas de mejoramiento de maíz locales podría contribuir favorablemente a la creación de genotipos híbridos con mayor nivel de resistencia a MRC. La capacidad predictiva de los marcadores asociados con la resistencia a MRC indican que la selección asistida por marcadores es una herramienta recomendable para seleccionar genotipos resistentes a MRC.

Palabras clave: índice de severidad de enfermedad; mapeo por asociación; QTL; SNP.

INTRODUCCIÓN

El maíz (*Zea mays* L.) es hospedante natural de más de 50 virus (Lapierre y Signoret, 2004), y una de las enfermedades virales más importantes en Argentina es el Mal de Río Cuarto (MRC) (Gimenez Pecci *et al.*, 2012). El agente causal de la enfermedad MRC es el *Mal de Río Cuarto virus* (MRCV) perteneciente a la familia Reoviridae, género *Fijivirus* (Distéfano *et al.*, 2002). Este virus es naturalmente transmitido en forma persistente, cíclica y propagativa por medio de insectos, siendo la chicharrita *Delphacodes kuscheli* Fennah el principal vector de la enfermedad (Remes Lenicov *et al.*, 1985). La forma de transmisión indica que el vector es reservorio natural del virus y la población de macrópteros migrantes constituye el principal inóculo de la enfermedad (Ornaghi *et al.*, 1993).

El principal síntoma de la enfermedad MRC es la presencia de enaciones o protuberancias sobre las nervaduras en el envés de las hojas. Las plantas pueden presentar otros síntomas como resultado de las modificaciones en los niveles hormonales endógenos, tales como tallos achatados, entrenudos cortos, hojas del tercio superior recortadas o reducidas a la vaina foliar, panojas atrofiadas de tamaño reducido y espigas múltiples con pocos o sin granos (Giménez Pecci *et al.*, 2012). Las siembras tempranas y el uso de agroquímicos son estrategias de manejo agronómico para reducir la enfermedad MRC mediante el control del insecto vector, pero la siembra de genotipos resistentes es el proceder recomendado (Di Renzo *et al.*, 2004).

La reacción a la enfermedad MRC se comporta como un carácter cuantitativo (Di Renzo *et al.*, 2002). Estudios moleculares de la reacción genotípica a MRC, han permitido identificar alelos de resistencia en poblaciones biparentales (Di Renzo *et al.*, 2004; Bonamico *et al.*, 2012; Rossi *et al.*, 2015). Sin embargo, en este tipo de mapeo pueden perderse alelos que no segregan entre los parentales de la población y alelos asociados a loci de efecto menor (Warburton *et al.*, 2015). La creciente disponibilidad de datos genómicos polimórficos y el potencial para explorar múltiples eventos de recombinación ocurridos en la historia evolutiva de un germoplasma específico, han convertido a los estudios de mapeo por asociación (GWAS, del inglés *genome wide association studies*) en una importante alternativa para estudiar caracteres de variación cuantitativa en plantas (Hao *et al.*, 2015).

La respuesta de los genotipos a enfermedades virales es comúnmente estimada mediante los caracteres incidencia y severidad (Rossi *et al.*, 2019b). Sin embargo, en MRC existen diversos síntomas que se manifiestan en distintos grados de severidad y con distinta incidencia en las poblaciones evaluadas. Para contemplar esta característica de la enfermedad se han realizado estudios donde los caracteres incidencia y severidad se han sintetizado o resumido en un índice, como es el caso del índice de severidad de enfermedad (ISE) (Shi *et al.*, 2012; Bonamico *et al.*, 2013; Chen *et al.*, 2015). Este índice constituye una media de la severidad de la reacción al MRC, ponderada por la incidencia de los distintos grados de severidad observada en un conjunto de plantas de cada genotipo. Si bien se han realizado análisis de QTL en poblaciones biparentales para el ISE de MRC en la zona de Argentina donde la enfermedad es endémica (Bonamico *et al.*, 2012; Di Renzo *et al.*, 2004), no se ha analizado este carácter en el contexto de GWAS, principalmente por la falta del fenotipado en poblaciones de gran tamaño extensamente genotipadas.

El Centro Internacional de Mejoramiento de Maíz y Trigo (CIMMYT) es una importante fuente de germoplasma para la incorporación de alelos exóticos en los programas de mejoramiento de todo el mundo. Las líneas de maíz de CIMMYT han sido extensivamente genotipadas mediante marcadores SNPs (Wu *et al.*, 2016) y son adecuadas para realizar estudios de mapeo por asociación en distintas partes del mundo donde puedan ser evaluadas fenotípicamente; algunas de estas líneas ya han mostrado buena adaptación en la zona de Argentina donde la enfermedad es endémica (Rossi *et al.*, 2019a). El objetivo del presente trabajo fue identificar nuevos alelos de resistencia asociados al ISE de la enfermedad MRC de maíz en un germoplasma exótico de CIMMYT.

MATERIALES Y MÉTODOS

Material vegetal y ensayo de campo

Una población de 210 líneas endocriadas de maíz proveniente del CIMMYT se evaluó fenotípicamente para la reacción a la enfermedad viral Mal de Río Cuarto en la zona donde la enfermedad es endémica. La población fue caracterizada por su adaptación a la región sur de Córdoba y su diversidad genotípica por Rossi *et al.* (2019a). Los ensayos de campo se realizaron en las localidades de Río Cuarto (64° 20′ W, 33° 8′ S) y Sampacho (64° 44′ W, 33° 20′ S), Córdoba, Argentina, bajo condiciones de transmisión natural. Los ensayos se establecieron en los ciclos agrícolas 2018-2019 y 2019-2020 en Río Cuarto y en el ciclo agrícola 2018-2019 en Sampacho. Cada combinación año-localidad se consideró un ambiente de análisis: Río Cuarto 2018-2019 (RC-18-19), Río Cuarto 2019-2020 (RC-19-20) y Sampacho 2018-2019 (SA-18-19). Cada genotipo se sembró en parcelas de un surco de 2,5 m de largo. En cada ambiente se utilizó un diseño parcialmente repetido (Cullis *et al.*, 2006) donde un 25% de los genotipos se repitió tres veces y el 75% de los genotipos restantes se sembró en una sola parcela, sin repeticiones. El 25% de los genotipos repetidos fue diferente en cada ambiente de acuerdo a la disponibilidad de semillas. Las plantas

de cada parcela se evaluaron individualmente según el grado de severidad propuesto por Ornaghi *et al.* (1993): 0=planta 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 cada parcela se estimó el índice de severidad de enfermedad (ISE) según Di Renzo *et al.* (2002), el cual involucra los caracteres incidencia y severidad de MRC. Este índice puede tomar valores de 0 a 100, donde 0 corresponde a genotipos sin síntomas y 100 a genotipos severamente afectados.

$$ISE = \frac{\sum(grado\ i \times plantas\ en\ el\ grado\ i)}{total\ de\ plantas \times 3} \times 100$$

Datos genotípicos

La caracterización molecular de la colección de líneas de maíz de CIMMYT fue realizada por Wu *et al.* (2016). Un total de 362.008 marcadores SNP, obtenidos mediante genotipado por secuenciación, está disponible públicamente en el repositorio de datos de CIMMYT (http://data.cimmyt.org/dvn). A partir del total de marcadores, se seleccionaron 78.376 SNP con una tasa de datos faltantes menor al 35% y una mínima frecuencia alélica superior a 0,05.

Análisis estadístico

Los datos fenotípicos multi-ambientales se analizaron con la función "mmer" del paquete "sommer" (Covarrubias Pazaran, 2016) en el software R (R Core Team 2016). Para obtener la mejor predicción lineal insesgada (BLUP) (West *et al.*, 2007) de los efectos de genotipos a través de ambientes, se ajustó un modelo lineal mixto (MLM) multi-ambiental. El modelo lineal mixto ajustado incluyó efectos de ambientes fijos, efectos de genotipos aleatorios y efectos de interacción genotipo-ambiente (GE) aleatorios:

$$\boldsymbol{y} = \boldsymbol{X\beta} + \boldsymbol{Z}\,\boldsymbol{u} + \boldsymbol{\varepsilon}$$

donde $\boldsymbol{y}$ es el vector de las observaciones fenotípicas, $\boldsymbol{X}$ y $\boldsymbol{Z}$ son las matrices de diseño. El vector $\boldsymbol{\beta}$ representa los efectos ambientales considerados como fijos, $\boldsymbol{u}$ representa los efectos aleatorios de genotipo y de la interacción GE y $\boldsymbol{\varepsilon}$ representa el error aleatorio. Los efectos genéticos aleatorios se asumieron normalmente distribuidos N $(0, \sigma_g^2)$, con la matriz de varianza-covarianza (**G**) con estructura diagonal. Los efectos de la interacción GE se asumieron normalmente distribuidos con media cero y matriz de varianza-covarianza no estructurada. Los errores se asumieron normalmente distribuidos *i.i.d.* N $(0, \sigma_e^2)$ e independientes del resto de efectos aleatorios. Ajustado el modelo se obtuvo el mejor predictor lineal insesgado (BLUP) del efecto de genotipo. El valor del BLUP para cada genotipo mide el efecto genético de la línea, descontando el efecto ambiental y la potencial interacción GE, y este valor fue usado como variable respuesta en el estudio de mapeo por asociación con el perfil molecular.

La heredabilidad del ISE, se estimó a partir de los componentes de varianza obtenidos del MLM ajustado de acuerdo a Holland *et al.* (2010).

En ambientes individuales:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\sigma_e^2/p\right)}$$

A través de ambientes:

$$H^2 = \frac{\sigma_g^2}{\sigma_g^2 + \left(\sigma_{ge}^2/e\right) + \left(\sigma_e^2/p\right)}$$

donde σ_g^2 es el componente de varianza genotípica, σ_{ge}^2 es el componente de varianza de la interacción GE, σ_e^2 es la varianza residual, *e* es el número de ambientes, y *p* es la media ponderada del número de repeticiones de cada genotipo a través de ambientes.

Mapeo por asociación

El software TASSEL 5.2.60 (Bradbury *et al.*, 2007) se usó para realizar la asociación entre los 78.376 SNPs y los BLUPs del ISE a través de ambientes para la población de líneas de maíz con información fenotípica y genotípica (n=186). Se compararon distintos modelos estadísticos de asociación (GWAS) para establecer el de mejor ajuste a los datos del estudio: 1) modelo lineal general propuesto por Pritchard *et al.* (2000), con la estructura poblacional determinada por la matriz Q con tres grupos (Wu *et al.*, 2016), como covariable de efecto fijo (modelo Q); 2) modelo lineal general propuesto por Price *et al.* (2006), con la estructura poblacional determinada mediante análisis de componentes principales, como covariable de efecto fijo (modelo PCA); 3) modelo lineal mixto propuesto por Parisseaux y Bernardo (2004), con la matriz de parentesco relativo entre individuos (*Kinship*) como efecto aleatorio (modelo K); 4) modelo lineal mixto con la matriz *kinship* como efecto aleatorio y la matriz Q como covariable de efecto fijo (modelo Q+K) (Yu *et al.*, 2006); 5) modelo lineal mixto con la matriz *kinship* como efecto aleatorio y las componentes principales como covariable de efecto fijo (modelo PCA+K) (Zhao *et al.*, 2007). La comparación entre modelos se realizó mediante el gráfico de distribución empírica (*Q-Q plot*) que permite evaluar la desviación de los valores de probabilidad ($-\log_{10}$ de valores-*P*) observados respecto

a los esperados bajo la hipótesis de asociación nula entre el marcador y la variable respuesta, *i.e.* el valor BLUP de la línea para el carácter ISE. Se utilizó el procedimiento propuesto por Li y Ji (2005) para corregir los valores-*P* y evitar el incremento de la tasa de falsos positivos debido a la multiplicidad de pruebas de hipótesis que se realizan al evaluar la asociación marcador por marcador. La significancia estadística, para declarar asociación significativa entre un marcador y el ISE, se definió en el valor umbral de $-\log_{10}$ (valor-*P*) >4 (valor *P*<0.0001). El gráfico de distribución empírica (*Q-Q plot*) y el gráfico *Manhattan plot*, donde se muestran los resultados de las pruebas de asociación marcador por marcador, se realizaron con el paquete "qqman" (Turner, 2018) del software R (R Core Team 2016) usando el modelo de asociación seleccionado.

Para predecir el efecto genotípico del ISE de MRC se ajustaron dos modelos de predicción genómica usando los méritos genéticos de las líneas como variable respuesta (*i.e.* BLUP de los efectos de genotipo del ISE). En el primer modelo, se consideraron como predictores todos los marcadores SNP utilizados en el mapeo por asociación. En el segundo modelo, solo se consideraron los marcadores SNP que resultaron estadísticamente asociados en el GWAS. El modelo de predicción genómica utilizado en ambos casos fue el modelo Bayes C propuesto por Habbier *et al.* (2011). La implementación de este modelo se realizó mediante el paquete BGLR (Pérez Rodríguez y de los Campos, 2014) del software R (R Core Team 2016). La eficiencia de la predicción se evaluó mediante validación cruzada. En cada repetición de la validación, los datos se dividieron aleatoriamente en un conjunto de entrenamiento (70% de las observaciones) y el conjunto complementario se usó para validar las predicciones del mérito genético, realizada a partir de uno u otro conjunto de marcadores (30% de las observaciones). La validación se repitió 50 veces con asignaciones aleatorias independientes de las líneas a los conjuntos de datos de entrenamiento y validación. La eficiencia de la predicción, con uno u otro conjunto de marcadores, se cuantificó utilizando la correlación entre los méritos genéticos de cada línea y los méritos predichos por el modelo de predicción genómica. La correlación se calculó en las 50 repeticiones aleatorias del proceso.

RESULTADOS

El ISE de MRC de las líneas de maíz, al considerar los BLUPs a través de ambientes, sugiere una amplia variabilidad genotípica en el grupo de líneas evaluadas (Figura 1). La Tabla 1 muestra los valores medios del ISE en los tres ambientes de evaluación y a través de ambientes. Los tres ambientes de evaluación presentaron un adecuado nivel de transmisión natural del virus. El ensayo realizado en la localidad de Sampacho en el ciclo agrícola 2018-2019 (SA-18-19) fue el que presentó el mayor valor medio para el ISE. Los componentes de varianza genotípica fueron altamente significativos en los tres ambientes individuales y a través de ambientes. El componente de varianza de la interacción genotipo × ambiente también fue significativo pero pequeño respecto de la varianza genotípica a través de ambientes (Tabla 1). La heredabilidad osciló entre 0,46 y 0,60 en ambientes individuales y fue de 0,70 a través de ambientes.

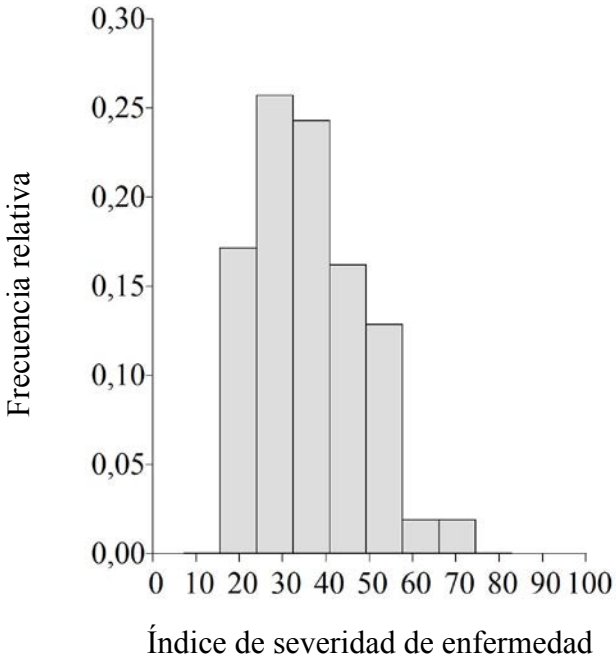

Figura 1. Frecuencia de distribución relativa para el índice de severidad de enfermedad (ISE) de Mal de Río Cuarto a través de ambientes.

Los gráficos de distribución empírica (*Q-Q plot*) permitieron comparar entre los distintos modelos de mapeo por asociación utilizados. El modelo Q+K presentó una distribución sesgada hacia los valores significativos. Los modelos K, PCA y PCA+K mostraron algunos puntos por debajo de la diagonal, indicando una posible sobre-parametrización (datos no mostrados). El modelo Q, que considera una estructura genética de la población de tres grupos, fue el que mejor ajustó para realizar la asociación entre los SNP y los BLUP del ISE (Figura 2).

Un total de 14 SNP resultaron estadísticamente asociados con el ISE a través de ambientes con un valor umbral de $-\log_{10}$ valor *P*>4,00 (valor *P*<0,0001) (Figura 2, Tabla 2). Según la posición física de los SNP, estos se agruparon en 11 posibles QTL de resistencia a MRC ubicados en los cromosomas 1, 2, 4, 6, 7, 8 y 9. El porcentaje de variación fenotípica explicada en forma

Tabla 1. Medias, mínimo, máximo, componentes de varianza y heredabilidad (H^2) del índice de severidad de enfermedad (ISE) de Mal de Río Cuarto (MRC) estimados a partir de 210 líneas de maíz en ambientes individuales y a través de ambientes de la zona donde la enfermedad es endémica.

Ambiente	Índice de severidad de enfermedad (ISE)					
	Media	Mínimo	Máximo	σ_g^2	σ_{ge}^2	H^2
RC-18-19	34,8	0,0	100,0	429,6***	-	0,54
RC-19-20	28,8	0,0	100,0	256,7***	-	0,46
SA-18-19	58,3	0,0	100,0	736,8***	-	0,60
BLUPs	36,0	10,0	80,0	397,2***	73,24*	0,70

RC-18-19: Río Cuarto 2018-2019; RC-19-20: Río Cuarto 2019-2020; SA-18-19: Sampacho 2018-2019.

σ_g^2 componente de varianza genotípica;

σ_{ge}^2 componente de varianza de la interacción genotipo×ambiente.

*** <0,0001;

*<0,05. Rango de ISE: 0-100%.

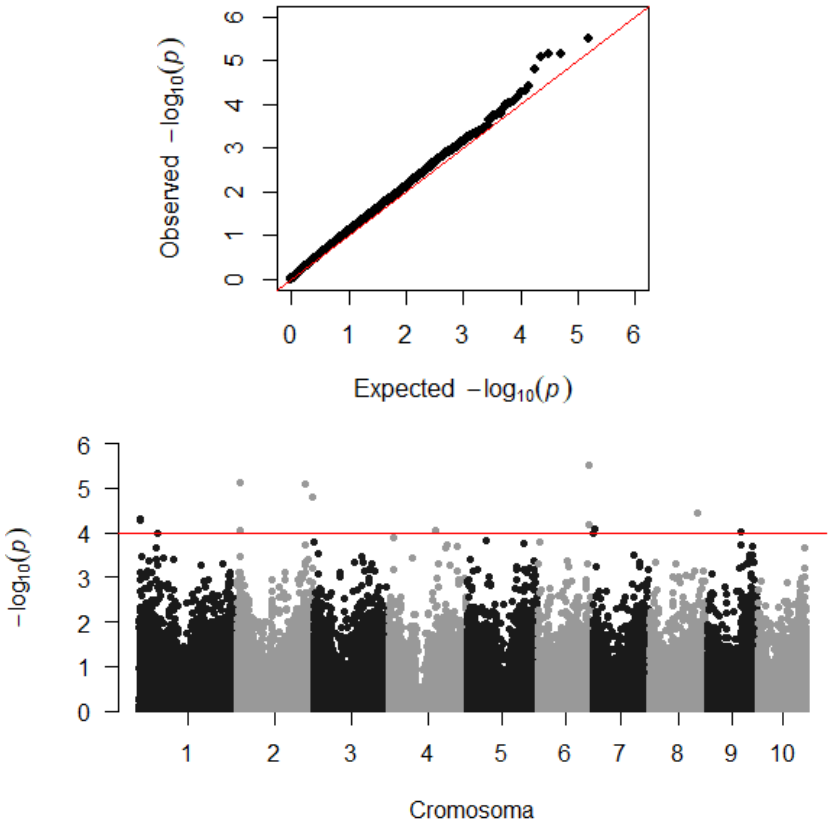

Figura 2. Distribución empírica *(Q-Q plot)* (arriba) y pruebas de asociación marcador por marcador (*Manhattan plot)* (abajo), de los resultados del mapeo por asociación para el índice de severidad de la enfermedad (ISE) de Mal de Río Cuarto a través de ambientes.

individual por cada QTL osciló entre 10 y 16% (Tabla 1). Los efectos del QTL 2 (bin 1.04) y del QTL 6 (bin 2.09) resultaron estadísticamente no significativos. Los QTL 9, 4 y 11, ubicados en los bin 7.01, 2.02 y 9.04 respectivamente, fueron los que presentaron mayor efecto genético aditivo (Tabla 2).

La eficiencia de la predicción fue en promedio de r=0,30 para el modelo donde se incluyeron todos los SNPs y aumentó a r=0,60 cuando sólo se incluyeron en el modelo el conjunto reducido de SNPs conformado con aquellos marcadores estadísticamente asociados al ISE de MRC según los resultados del GWAS (Figura 3).

Tabla 2. Detalles de los marcadores SNP asociados con el índice de severidad de enfermedad (ISE) de Mal de Río Cuarto.

QTL	Marcador	Cromosoma	Bin	Alelos	Efecto aditivo		Valor *P*	R^2
1	S1_2467319	1	1,01	A/T	-2,28	*	4,75x10-5	0,12
	S1_2467320			T/C			5,23x10-5	0,12
2	S1_56399197	1	1,04	T/C	-0,61	ns	9,90x10-5	0,13
3	S2_9278082	2	2,02	G/A	-2,17	**	8,80x10-5	0,12
4	S2_12441417	2	2,02	T/A	-4,07	***	7,22x10-6	0,15
	S2_12441430			A/C			7,22x10-6	0,15
5	S2_211077915	2	2,08	C/T	-2,78	**	8,29x10-6	0,12
6	S2_234192035	2	2,09	C/G	0,44	ns	1,57x10-5	0,16
7	S4_143512347	4	4,05	T/G	-2,02	*	8,92x10-5	0,10
8	S6_158274989	6	6,06	G/C	1,76	*	6,53x10-5	0,11
	S6_158275013			G/A			3,02x10-6	0,16
9	S7_7160192	7	7,01	A/T	-6,02	***	7,93x10-5	0,15
10	S8_146017787	8	8,05	G/T	-2,21	**	3,70x10-5	0,12
11	S9_102698811	9	9,04	C/T	-3,87	*	9,56x10-5	0,13

R^2: variación fenotípica explicada por cada marcador. La posición física exacta de cada SNP se puede inferir a partir del nombre del marcador, por ejemplo, S9_102698811: cromosoma 9; 102.698.811 pb. El alelo subrayado indica el alelo de resistencia para cada marcador.

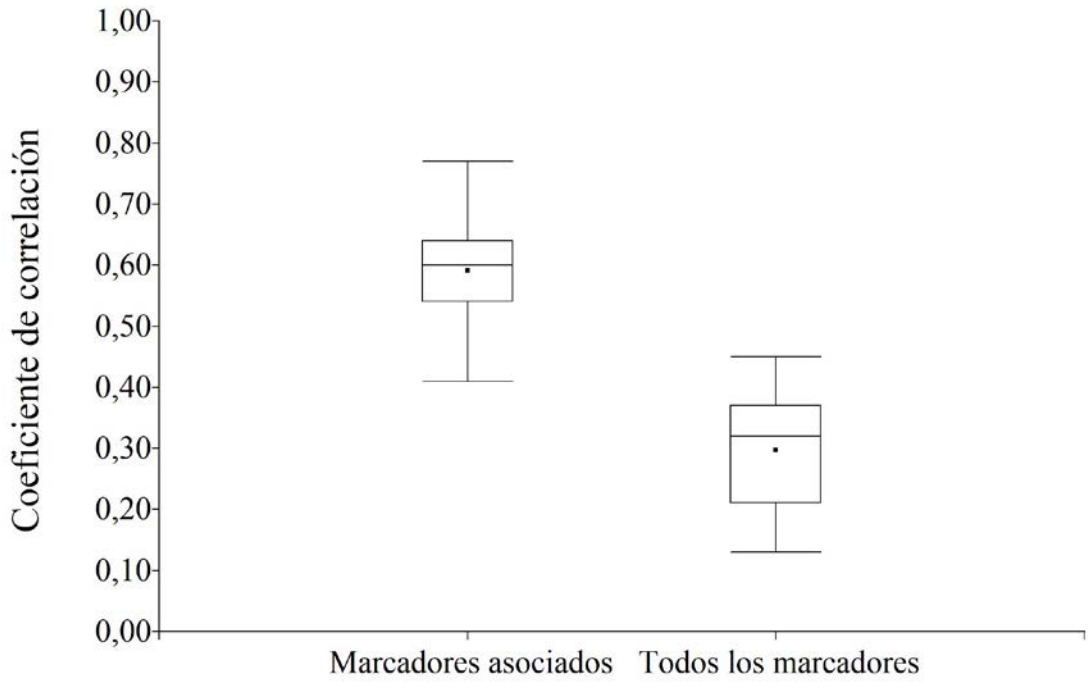

Figura 3. Distribución de la correlación entre los valores observados de los BLUP del índice de severidad de la enfermedad (ISE) Mal de Río Cuarto y los valores predichos por el modelo de predicción construido a partir de los marcadores SNPs identificados como significativamente asociados en el mapeo y a partir de todo el conjunto de SNPs disponibles.

DISCUSIÓN

La evaluación fenotípica precisa y confiable es uno de los principales requisitos para realizar estudios de mapeo por asociación (GWAS) (Yu *et al.*, 2008). En maíz, la enfermedad viral MRC es transmitida solamente por insectos, siendo el principal vector la chicharrita *Delphacodes kuscheli*, por lo que la evaluación fenotípica se debe realizar bajo condiciones de transmisión natural de la enfermedad. La variación de los valores extremos del ISE de MRC es común entre ambientes de evaluación. En el presente estudio, todos los ambientes registraron valores extremos del ISE y por lo tanto permitieron diferenciar a los genotipos. De todos modos, para evitar posibles efectos ambientales, se decidió utilizar los BLUP a través de ambientes como variable indicadora de la *performance* de cada línea durante el análisis de asociación con el perfil molecular.

Los componentes de varianza genotípica y los valores de heredabilidad estimados en el presente estudio indican una amplia variabilidad genotípica en la población de líneas de maíz proveniente del CIMMYT. Los valores de heredabilidad del ISE, estimados en ambientes individuales y a través de ambientes, son superiores a los reportados por Di Renzo *et al.* (2002), quienes evaluaron una población de familias $F_{2:3}$ en tres ambientes de la zona donde la enfermedad MRC es endémica.

En estudios de mapeo por asociación, la elección del modelo a utilizar es un paso importante para considerar los efectos que pueden tener la estructura poblacional presente y las relaciones de parentesco entre los genotipos en las asociaciones (Gutiérrez *et al.*, 2011). En nuestro trabajo, el modelo que mejor ajustó fue el que considera la estructura poblacional determinada por la matriz Q como covariable de efecto fijo. En este caso la matriz Q incluida se compone de tres grupos de acuerdo a la adaptación ambiental de las líneas de maíz. Esta estructura de la población fue propuesta por Wu *et al.* (2016), quienes caracterizaron molecularmente la colección completa de 538 líneas de maíz de CIMMYT que incluye a la población evaluada en el presente estudio.

La integración de datos fenotípicos y genotípicos podría aumentar potencialmente la eficiencia del mejoramiento de caracteres complejos (Guo *et al.*, 2020) como la resistencia a MRC. Si bien las bases genéticas de la enfermedad MRC han sido ampliamente estudiadas en poblaciones biparentales (Di Renzo *et al.*, 2004; Bonamico *et al.*, 2012), la evaluación de poblaciones diversas de germoplasma exótico ha sido muy poco explorada. En el presente estudio, el mapeo por asociación permitió identificar 11 posibles QTL para resistencia a MRC. Al comparar los QTL identificados para el ISE con los identificados en estudios previos, tanto para el ISE como para los caracteres incidencia y severidad, se observan coincidencias en las posiciones físicas y en la proporción de la variación fenotípica explicada. Por ejemplo, en el bin 4.05, donde nosotros identificamos un SNP asociado al ISE, Bonamico *et al.* (2013) identificaron un marcador microsatélite (SSR) para el ISE de MRC, al evaluar una población de líneas endocriadas recombinantes (RILs). En el bin 1.01, Bonamico *et al.* (2012) identificaron un marcador SSR asociado con incidencia y con severidad de MRC. En el mismo bin, en nuestro trabajo identificamos un SNP asociado al ISE de la enfermedad. Redinbaugh *et al.* (2018) combinaron resultados de estudios donde se identificaban QTL para resistencia a enfermedades virales en maíz e identificaron nueve regiones del genoma donde se agrupaban loci de resistencia (*clusters*). Uno de esos *clusters* se ubica en el bin 2.08 y reporta QTL de resistencia para tres enfermedades virales. En el mismo bin y en una posición física (pares de bases) muy próxima, identificamos un marcador asociado con el ISE de MRC. Los QTL identificados en este estudio, en los bin 7.01 y 9.04 podrían considerarse como nuevos alelos de resistencia aportados por el germoplasma exótico de CIMMYT debido a que en estudios previos no han sido reportados QTL para MRC en esas posiciones.

La relación entre incidencia y severidad de enfermedades vegetales es epidemiológicamente importante (Seem, 1984). Esta relación puede ser estudiada según la correlación genética entre ambos caracteres como se realizó en el trabajo de Rossi *et al.* (2020). Otra manera de estudiar la relación es a través de la estimación del ISE, el cual combina la incidencia y la severidad de una enfermedad en un único índice. Esto explica que los posibles QTL 1, 4, 8 y 9, identificados en este estudio, fueron también identificados por Rossi *et al.* (2020) al evaluar la reacción a MRC mediante los caracteres incidencia y severidad individualmente. Según Dintinger *et al.* (2005) los QTL asociados a incidencia y severidad podrían involucrar mecanismos de resistencia que reducen la evolución de síntomas severos en la planta, y mecanismos de resistencia que reducen la probabilidad que una planta manifieste síntomas.

La validación cruzada del modelo de predicción genómica demostró la capacidad predictiva de los marcadores asociados al ISE de MRC, identificados en el GWAS. La menor capacidad predictiva del modelo que incluye todos los marcadores podría estar relacionada a correlaciones entre los marcadores y de éstos con el efecto ambiente, los cuales podrían enmascarar el efecto de otros marcadores. Si bien se requieren numerosos estudios para avalar el desempeño de la predicción genómica para la resistencia a MRC en maíz, los resultados indican que la selección asistida por marcadores basada en los SNPs que se encontraron asociados al ISE de MRC podría ser una herramienta muy útil para la selección. El uso del genotipado mediante PCR específica de alelos (KASP, *Kompetitive Allele Specific PCR*), técnica de laboratorio que permite genotipar

alelos específicos de unos pocos marcadores, combinado con selección asistida por marcadores, representa una alternativa más económica y rápida que la selección genómica.

Los resultados sugieren que el germoplasma exótico de CIMMYT tiene amplia variabilidad genética para la resistencia a la enfermedad MRC. La evaluación de un germoplasma exótico para la reacción a una enfermedad endémica local contribuye a validar regiones genómicas identificadas en estudios previos e identificar nuevos alelos de resistencia. En la industria semillera argentina, la importación de paquetes tecnológicos "cerrados" es difícil porque se requiere de adaptación a las condiciones agroecológicas de cada región del país. Por lo tanto, la incorporación de germoplasma exótico adaptado a las condiciones agroecológicas locales, podría favorecer la creación de genotipos híbridos con mayor nivel de resistencia a MRC en los programas de mejoramiento de maíz. La eficiencia predictiva de los marcadores asociados con la resistencia a MRC indican que la selección asistida por marcadores constituye una herramienta promisoria para seleccionar genotipos resistentes a MRC.

BIBLIOGRAFÍA

Bonamico N.C., Di Renzo M.A., Borghi M.L., Ibañez M.A., Díaz D.G., Salerno J.C., Balzarini M.G. (2013) Mapeo de QTL para una medida multivariada de la reacción al virus del Mal de Río Cuarto. J. Basic Appl. Genet. 24: 11-21.

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: 619-629. https://doi.org/10.1017/S0021859611000943.

Bradbury P.J., Zhang Z., Kroon D.E., Casstevens T.M., Ramdoss Y., Buckler E.S. (2007) TASSEL: Software for association mapping of complex traits in diverse samples. Bioinformatics. https://doi.org/10.1093/bioinformatics/btm308.

Chen G., Wang X., Hao J., Yan J., Ding J. (2015) Genome-wide association implicates candidate genes conferring resistance to maize rough dwarf disease in maize. PLoS ONE 10: 1-13. https://doi.org/10.1371/journal.pone.0142001.

Covarrubias Pazaran G. (2016) Genome-Assisted prediction of quantitative traits using the r package sommer. PLoS ONE 11: 1-15. https://doi.org/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: 381-393. https://doi.org/10.1198/108571106X154443.

Dintinger J., Verger D., Caiveau S., Risterucci A.M., Gilles J., Chiroleu F., Hamon P. (2005) Genetic mapping of maize stripe disease resistance from the Mascarene source. Theor. Appl. Genet. 111: 347-359. https://doi.org/10.1007/s00122-005-2027-3.

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: 289-295. https://doi.org/10.1017/S0021859604004307.

Di Renzo M.A., Bonamico N.C., Díaz D.G., Salerno J.C., Ibañez M.A., Gesumaria J.J. (2002) Inheritance of resistance to Mal de Río Cuarto (MRC) disease in *Zea mays* (L.) J. Agric. Sci. 139: 47-53. https://DOI: 10.1017/S0021859602002241.

Distéfano A.J., Conci L.R., Muñoz Hidalgo M., Guzmán F.A., Hopp H.E., del Vas M. (2002) Sequence analysis of genome segments S4 and S8 of Mal de Río Cuarto virus (MRCV): evidence that the virus should be a separate *Fijivirus* species. Arch. Virol. 147: 1699-1709.

Gimenez Pecci M.P. (2012) Mal de Río Cuarto del maíz. En: 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.

Guo J., Pradhan S., Shahi D., Khan J., Mcbreen J., Bai G.J., Murphy P., Babar M.A. (2020) Increased prediction accuracy using combined genomic information and physiological traits in a soft wheat panel evaluated in multi-environments. Sci. Rep. 10: 7023. https://doi.org/10.1038/s41598-020-63919-3.

Gutierrez L., Cuesta Marcos A., Castro A.J., von Zitzewitz J., Schmitt M., Hayes P.M. (2011) Association mapping of malting quality quantitative trait loci in winter barley: positive signals from small germplasm arrays. Plant Genome 4: 256-272.

Habier D., Fernando R., Kizilkaya K., Garrick D. (2011) Extension of the Bayesian alphabet for genomic selection. BMC Bioinformatics 12: 186.

Hao D., Cheng Y., Chen G., Lu H., Shi M., Zhang Z., Huang X., Mao Y., Xue L. (2015) Identification of significant single nucleotide polymorphisms for resistance to maize rough dwarf disease in elite maize (*Zea mays* L.) inbred lines. Euphytica 203: 109-20.

Holland J.B., Nyquist W.E., Cervantes Martínez C.T. (2010) Estimating and Interpreting Heritability for Plant Breeding: An Update. Plant Breed. Rev. https://doi.org/10.1002/9780470650202.ch2.

Lapierre H., Signoret P.A. (2004) Viruses and virus diseases of *Poaceae* (Gramineae). INRA ED, Paris.

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.

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.

Parisseaux B., Bernardo R. (2004) In silico mapping of quantitative trait loci in maize. Theor. Appl. Genet. 109: 508-514. https://doi.org/10.1007/s00122-004-1666-0.

Perez Rodriguez P., de los Campos G. (2014) Genome-Wide Regression and Prediction with the BGLR Statistical Package. Genetics 198 (2): 483-495.

Price A.L., Patterson N.J., Plenge R.M., Weinblatt M.E., Shadick N.A., Reich D. (2006) Principal components analysis corrects for stratification in genome-wide association studies. Nat. Genet. 38: 904-909. https://doi.org/10.1038/ng1847.

Pritchard J.K., Stephens M., Rosenberg N.A., Donnelly P. (2000) Association mapping in structured populations. Am. J. Hum. Genet. 67: 170-181. https://doi.org/10.1086/302959.

R Core Team (2016) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, URL (http://www.R-project.org/).

Redinbaugh M.G., Lübberstedt T., Leng P., Xu M. (2018) The genetics and genomics of virus resistance in maize. In: Bennetzen J., Flint Garcia

S., Hirsch C., Tuberosa R. (Eds.) The Maize Genome. Compendium of Plant Genomes. Springer, Cham.

Remes Lenicov A.M.M., Tesón A., Dagoberto E., Huguet N. (1985) Hallazgo de uno de los vectores del Mal de Río Cuarto en maíz. Gaceta Agronómica 5: 251-258.

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.

Rossi E.A., Ruiz M., Bonamico N.C., Balzarini M.G. (2020) Genome-wide association study of resistance to Mal de Río Cuarto disease in maize. Agronomy J. https://doi.org/10.1002/agj2.20448.

Rossi E.A., Ruiz M., Di Renzo M., Bonamico N.C. (2019a) Genotypic diversity in 291 maize lines from CIMMYT and phenotypic characterization in southern Cordoba, Argentina. J. Basic Appl. Genet. 30: 25-33.

Rossi E.A., Ruiz M., Rueda Calderón M.A., Bruno C.I., Bonamico N.C., Balzarini M.G. (2019b) Meta-analysis of QTL studies for resistance to fungi and viruses in maize. Crop Sci. 59: 125-139. https://doi.org/10.2135/cropsci2018.05.0330.

Seem R.C. (1984) Disease Incidence and Severity Relationships. Annu. Rev. Phytopathol. 22: 133-150. https://doi.org/10.1146/annurev.py.22.090184.001025.

Shi L., Hao Z.F., Weng J.F., Xie C.X., Liu C.L., Zhang D., Zhang S. (2012) Identification of a major quantitative trait locus for resistance to maize rough dwarf virus in a Chinese maize inbred line X178 using a linkage map based on 514 gene-derived single nucleotide polymorphisms. Mol. Breed. 30: 615-625. https://doi.org/10.1007/s11032-011-9652-0.

Turner D.S. (2018) qqman: an R package for visualizing GWAS results using Q-Q and manhattan plots. J. Open Source Softw. 3: 731. https://doi.org/10.21105/joss.00731.

Warburton M., Tang J., Windham G., Hawkins L., Murray S., Xu W., Boykin D., Perkins A., Williams W. (2015) Genome-Wide Association Mapping of *Aspergillus flavus* and Aflatoxin Accumulation Resistance in Maize. Crop Sci. 55: 1-11.

West B., Welch K.B., Galecki A.T. (2007) Linear mixed models: A practical guide using statistical software. Chapman & Hall, Boca Raton, FL.

Wu Y., San Vicente F., Huang K., Dhliwayo T., Costich D.E., Semagn K., Babu R. (2016) Molecular characterization of CIMMYT maize inbred lines with genotyping by sequencing SNPs. Theor. Appl. Genet. 129: 753-765. https://doi.org/10.1007/s00122-016-2664-8.

Yu J., Holland J.B., McMullen M.D., Buckler E.S. (2008) Genetic design and statistical power of nested association mapping in maize. Genetics 178: 539-551. https://doi.org/10.1534/genetics.107.074245.

Yu J., Pressoir G., Briggs W.H., Bi I.V., Yamasaki M., Doebley J.F., Buckler E.S. (2006) A unified mixed model method for association mapping that accounts for multiple levels of relatedness. Nat. Genet. 38: 203-208. https://doi.org/10.1038/ng1702.

Zhao K., Aranzana M.J., Kim S., Lister C., Shindo C., Tang C., Nordborg M. (2007) An Arabidopsis example of association mapping in structured samples. PLoS Genet. 3: 71-82. https://doi.org/10.1371/journal.pgen.0030004.

BAG | Journal of Basic & Applied Genetics