



UNIVERSIDAD NACIONAL DE LA PLATA



FACULTAD DE CIENCIAS VETERINARIAS

Trabajo de tesis realizado como requisito para optar al título de
DOCTORA EN CIENCIAS VETERINARIAS

**Caracterización genética del caballo Polo Argentino
mediante el análisis de datos genealógicos y de
polimorfismos de nucleótido simple**

AUTORA: Med.Vet. Mgtr. AZCONA, Florencia

DIRECTOR: Dr. DEMYDA PEYRÁS, Sebastián.

CODIRECTORA: Dra. PERAL GARCÍA, Pilar.

LUGAR DE TRABAJO

Instituto de Genética Veterinaria "Ingeniero Fernando Noel Dulout" (IGEVET)

MIEMBROS DEL JURADO

Dr. CORVA, Pablo Marcelo

Dra. KIENAST, Mariana

Dr. TRIGO, Pablo Ignacio

2023

Agradecimientos

A la Facultad de Ciencias Veterinarias de la Universidad Nacional de La Plata por brindarme educación gratuita y de calidad durante todos estos años.

Al Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET) por su programa de becas doctorales que me permitieron realizar la carrera de doctorado.

A mi director de tesis y codirector de beca Sebastián Demyda Peyrás, por la formación brindada y la constante supervisión durante todo el proceso de investigación y elaboración de la tesis.

A mi codirectora de tesis y directora de beca Pilar Peral García, por estar siempre predispuesta a escuchar mis inquietudes y a colaborar en todo lo posible.

Al IGEVET por abrirme sus puertas y poner a disposición sus recursos e instalaciones para desarrollar este trabajo.

A la AACCP y a todos los criadores y propietarios de caballos de polo que accedieron a participar del proyecto y aportaron muestras de sus animales.

Al grupo de Investigación MERAGEM de la Universidad de Córdoba (España), por compartir parte de los datos de su banco genético, y por las acertadas intervenciones de Antonio Molina y Mercedes Valera durante el análisis de los datos del Capítulo I.

A mis compañeros de la Cátedra de Medicina Equina y del Hospital de Equinos de la Facultad, por permitirme formar parte y por el apoyo en todos estos años; y en particular a Ramón López que desinteresadamente facilitó parte de las muestras utilizadas en esta tesis.

A todos mis compañeros y amigos del Becariado, los que estuvieron desde el principio y los que se sumaron en el camino, por los consejos y el aliento, por las juntadas, los mates, los asados y todos los momentos compartidos.

A mis amigas vetes de la carrera, por toda la trayectoria recorrida; a mis amigos de Olavarría, Juli, Seba y Manu, por estar ahí incondicionalmente desde siempre.

Muy especialmente a mi familia, por el acompañamiento y el enorme esfuerzo que hicieron siempre, a pesar de todo, para darme la oportunidad de estudiar lo que había elegido.

Finalmente, a todas las personas que de una u otra manera colaboraron y sus aportes hicieron posible la concreción de esta tesis.

Citas bibliográficas correspondientes a publicaciones parciales de la tesis

Azcona, F.; Molina Alcalá, A.; Valera, V.; Dorado, J.; Demyda Peyrás, S. (2018). Impact of the use of large-scale embryo transfer programs in the increase of inbreeding and relativeness in the Argentinean Polo horse. *Reproduction in Domestic Animals*. <https://onlinelibrary.wiley.com/toc/14390531/2018/53/S2>. 22 Annual Conference of the European Society for Domestic Animal Reproduction. European Society for Domestic Animal Reproduction (ESDAR). España, Córdoba.

Azcona, F; Valera, M; Molina, A; Trigo, P; Peral García, P; Solé, M; Demyda-Peyrás, S. (2020). Impact of reproductive biotechnologies on genetic variability of Argentine Polo horses. *Livestock Science*, 231, 103848. <http://dx.doi.org/10.1016/j.livsci.2019.103848>

Azcona, F; Molina, A; Peral García, P; Demyda Peyrás, S. (2021). Genomic data reveals a serious underestimation of pedigree inbreeding levels in Polo Argentino horses. *Proceedings of the 38th International Conference on Animal Genetics*. https://www.isag.us/Docs/Proceedings/ISAG2021_Proceedings.pdf. 38th International Society for Animal Genetics Conference. Sudáfrica.

Azcona, F; Molina, A; Piroso, Y; Valera, M; Peral García, P; Trigo, P; Demyda-Peyrás, S. Long term effect of large-scale embryo transfer programs in the genetic structure and variability of Argentinean Polo horses: a 30 years retrospective (2021). *Journal of Equine Veterinary Science*. <https://www.sciencedirect.com/journal/journal-of-equine-veterinary->

[science/vol/89/suppl/C](#). International Symposium on Equine Embryo Transfer and Technologies (ISEET). Italia, Pisa.

Azcona, F; Molina, A; Peral García, P; Demyda-Peyrás, S. (2021). Does morphology make the athlete?: Genetic analysis of Argentine polo horses lines. *Book of Abstracts of the 72nd Annual Meeting of the European Federation of Animal Science*. <https://www.wageningenacademic.com/doi/book/10.3920/978-90-8686-918-3>. 72nd Annual Meeting of the European Federation of Animal. Suiza, Davos.

Azcona, F; Karlau, A; Trigo, P; Alvarez, R; Demyda-Peyrás, S. (2023). Is the Argentinean Polo pony a horse breed?: genomic characterization and comparison with Thoroughbreds using SNP-array data. 39th International Society for Animal Genetics Conference. *Aceptado para su presentación y publicación en el congreso.*

ÍNDICE DE CONTENIDOS

<i>Sección</i>	<i>Página</i>
Abreviaturas y símbolos.....	XV
Resumen	1
Abstract	3
INTRODUCCIÓN	5
1. La producción equina en Argentina	5
1.1. Origen y difusión del polo.....	6
1.2. El caballo de polo en Argentina.....	7
2. Caracterización genética.....	11
2.1. Diversidad genética	12
2.2. Consanguinidad y depresión endogámica.....	14
2.3. Métodos de caracterización genética.....	18
3. Justificación	20
HIPÓTESIS.....	21
OBJETIVO GENERAL	22
Objetivos específicos.....	22
CAPÍTULO I: Análisis de los datos genealógicos en el caballo Polo Argentino: caracterización genética e impacto de las biotecnologías reproductivas sobre la variabilidad.	24
1. Introducción	24
2. Materiales y métodos	26

2.1. Datos genealógicos	26
2.2. Análisis demográfico	26
2.3. Parámetros de variabilidad genética	27
2.4. Análisis de fundadores y ancestros	28
2.5. Estructura poblacional	29
3. Resultados.....	30
3.1. Análisis demográfico	30
3.2. Parámetros de variabilidad genética	33
3.3. Análisis de fundadores y ancestros	35
3.4. Estructura poblacional	36
4. Discusión	36
5. Material Suplementario	47
CAPÍTULO II: Análisis de la variabilidad genética en el Caballo Polo Argentino mediante datos genómicos: determinación del número y distribución de bloques de homocigosidad (ROH).	
1. Introducción	49
2. Materiales y métodos	54
2.1. Material animal	54
2.2. Obtención de muestras y extracción de ADN.....	55
2.3. Genotipificación con microarray de SNP	55
2.3.1. Control de calidad de los genotipos.....	56
2.4. Análisis genómicos.....	56
2.4.1. Heterocigosidad y equilibrio Hardy Weinberg.....	56

2.4.2. Determinación de ROH y coeficiente de endogamia.....	58
2.4.3. Identificación de Islas ROH (ROHi).....	60
3. Resultados.....	61
3.1. Heterocigosis, frecuencias alélicas y Hardy Weinberg.....	61
3.2. Determinación de ROH y coeficientes de endogamia	63
3.3. Islas ROH.....	65
4. Discusión	67
5. Material Suplementario	80
CAPÍTULO III: Análisis de la estructura genética del caballo Polo Argentino, su relación con otras razas e identificación de huellas de selección.	83
1. Introducción	83
2. Materiales y métodos	85
2.1. Material animal	85
2.2. Estandarización de los datos genómicos.....	87
2.3. Estructuración genética y diferenciación entre razas	88
2.3.1. Análisis de la varianza molecular (AMOVA)	88
2.3.2. Análisis de componentes principales (PCA).....	88
2.3.3. Análisis de componente genética (admixture).....	89
2.3.4. Índice F_{ST}	90
3. Resultados.....	91
3.1. Análisis de la Varianza Molecular (AMOVA).....	91
3.2. Análisis de componentes principales (PCA).....	92

3.3. Análisis de componente genética (admixture).....	93
3.4. Índice F_{ST}	95
4. Discusión	97
5. Material Suplementario	109
CAPÍTULO IV: Caracterización genética de marcadores asociados a caracteres fenotípicos con impacto en la aptitud deportiva en el caballo Polo Argentino. .	118
1. Introducción	118
2. Materiales y métodos	122
2.1. Material animal y genotipado	122
2.2. Marcadores genéticos	123
2.3. Frecuencias genotípicas y alélicas de los marcadores	123
3. Resultados.....	124
4. Discusión	130
CONCLUSIONES	137
REFERENCIAS BIBLIOGRÁFICAS	142

ÍNDICE DE IMÁGENES

Imagen 1. Equino de raza Polo Argentino. Gentileza del criador Francisco Bargallo (Genética Santa Agueda).....	8
Imagen 2. Yegua Polo Argentino disputando un partido de polo. Gentileza del criador Francisco Bargallo (Genética Santa Agueda).....	9

ÍNDICE DE FIGURAS

CAPÍTULO I

Figura 1.1. Evolución del número de ejemplares registrados por año de nacimiento desde la apertura del libro genealógico	31
Figura 1.2. Porcentaje de individuos nacidos por transferencia embrionaria y registrados en el libro genealógico.....	31
Figura 1.3. Evolución del coeficiente de endogamia durante los últimos 16 años.....	33

CAPÍTULO II

Figura 2.1. Gráfico Manhattan con los valores-p del test de permutaciones de los SNP dentro de bloques de homocigosidad.....	66
--	----

CAPÍTULO III

Figura 3.1. PCA entre las distintas razas equinas.....	92
Figura 3.2. PCA entre Polo Argentino (PA) y Sangre Pura de Carrera (SPC).....	93
Figura 3.3. Análisis Admixture entre las distintas razas estudiadas.....	94
Figura 3.4. $-\log_{10}$ del valor-q de los índices F_{ST} por SNP y por cromosoma entre Polo Argentino y SPC.....	96

CAPÍTULO IV

Figura 4.1. Evolución de las frecuencias genotípicas de los marcadores con mayor polimorfismo en los caballos Polo Argentino, durante las últimas 3 generaciones.....	129
---	-----

FIGURAS SUPLEMENTARIAS

CAPÍTULO III

Figura S3.1. Error de validación cruzada (error cv) de los diferentes K en el análisis admixture, entre todas las razas (A) y entre Polo Argentino y SPC (B)..... 109

ÍNDICE DE TABLAS

CAPÍTULO I

Tabla 1.1. Intervalo generacional en el caballo Polo Argentino.....	32
Tabla 1.2. Análisis de fundadores y ancestros.....	35
Tabla 1.3. Ancestros más contribuyentes de cada grupo.....	37

CAPÍTULO II

Tabla 2.1. Promedio del número de marcadores (nSNP), frecuencia del alelo menor (MAF), heterocigosidad observada (H_o) y esperada (H_e), y valor-p del equilibrio Hardy Weinberg de cada locus por cromosoma en el caballo Polo Argentino.....	62
Tabla 2.2. Cantidad y longitud promedio de ROH en la población Polo Argentino.....	66
Tabla 2.3. Promedio de los coeficientes de endogamia calculados. Correlación de Spearman (ρ) entre el F_{PED} , F_{PED3} , F_{IS} y los F_{ROH}	64
Tabla 2.4. Promedio de los coeficientes de endogamia F_{ROH} por cromosoma en la población.....	65
Tabla 2.5. Regiones genómicas seleccionadas como potenciales ROHi.....	66

CAPÍTULO III

Tabla 3.1. Origen de las muestras utilizadas para el análisis.....	87
Tabla 3.2. AMOVA entre razas (n=15).....	91
Tabla 3.3. Valores de F_{ST} entre las distintas razas analizadas.....	96

CAPÍTULO IV

Tabla 4.1. Marcadores SNP asociados a caracteres de interés en la especie equina genotipados con el Microarray Illumina Equine SNP70 array (Illumina, Inc.) (Neogen, 2023).....	124
Tabla 4.2. Marcadores SNPs polimórficos en la raza Polo Argentino.....	127
Tabla 4.3. Frecuencias genotípicas (A) y alélicas (B) de los SNP polimórficos asociados a caracteres morfológicos, velocidad y temperamento en los distintos grupos.....	127

TABLAS SUPLEMENTARIAS

CAPÍTULO I

Tabla S1.1. Evolución de los distintos parámetros genéticos analizados por generaciones.	47
Tabla S1.2. Promedio del coeficiente de endogamia en individuos endogámicos ($F_{PED}>0$) y proporción de individuos endogámicos por grupo y por categoría de F_{PED}	48

CAPÍTULO II

Tabla S2.1. Genes encontrados dentro de cada ROHi.....	80
Tabla S2.2. Genes encontrados dentro de cada ROHi y su respectiva clasificación de ontología génica (GO term), según DAVID.....	81

CAPÍTULO III

Tabla S3.1. Análisis admixture. Promedio de los porcentajes de asignación a cada población para todas las razas (K-10) (A) y para Polo Argentino y SPC (K-2) (B).....	109
Tabla S3.2. Regiones genómicas con mayor diferenciación genética entre Polo Argentino y SPC, de acuerdo con el índice F_{ST}	110
Tabla S3.3 Genes contenidos dentro de las regiones genómicas con alta diferenciación genética entre Polo Argentino y SPC.....	113
Tabla S3.4. Categorías GO (Gene Ontology) de los genes encontrados en el análisis funcional de las regiones de alta diferenciación genética entre Polo Argentino y SPC, extraído de la plataforma DAVID.....	116

Abreviaturas y símbolos

ΔF : Incremento de la endogamia.

Δf : incremento de la coascendencia.

AACCP: Asociación Argentina de Criadores de Caballos de Polo.

ADN: Ácido desoxirribonucleico.

AMOVA: Análisis de la varianza molecular.

AR: Coeficiente de parentesco (“Average relatedness”).

cM: Centimorgan.

DE: Desvío estándar.

ECA: Cromosoma *Equus Caballus*.

ECG: número de generaciones completas equivalentes.

EBV: Valores de cría (“Estimated Breeding Values”).

F: coeficiente de endogamia individual.

f: coeficiente de coascendencia.

F_A: número efectivo de ancestros.

FAO: Organización de las Naciones Unidas para la Alimentación y Agricultura.

F_E: número efectivo de fundadores.

F_G: número efectivo de genomas fundadores.

F_{PED}: Coeficiente de endogamia calculado con datos de pedigrí.

F_{PED3} : Coeficiente de endogamia calculado con datos de pedigrí de las últimas 3 generaciones.

F_R : Tasa de endogamia.

F_{ROH} : Coeficiente de endogamia calculado a partir de ROH.

F_{ROH_3G} : Coeficiente de endogamia calculado a partir de ROH $\geq 16,6$ Mb.

F_{ROH_6G} : Coeficiente de endogamia calculado a partir de ROH $\geq 8,3$ Mb.

F_{ROH_9G} : Coeficiente de endogamia calculado a partir de ROH $\geq 5,5$ Mb.

F_{ROH_ANC} : Coeficiente de endogamia calculado a partir de ROH $< 5,5$ Mb.

GO: Ontología genética.

He: Heterocigosidad esperada.

Ho: Heterocigosidad observada.

H-W: Hardy Weinberg.

IBD: Identidad por descendencia.

IG: intervalo generacional.

ISAG: International Society of Animal Genetics.

LD: desequilibrio de ligamiento ("Linkage Disequilibrium").

Mb: Megabase.

N_e : Tamaño efectivo de la población.

N_E : Tamaño efectivo de la población estimado vía incremento individual de la endogamia.

N_{EC} : Tamaño efectivo de la población estimado vía incremento individual de la coascendencia.

Pb: Pares de bases.

PCA: Análisis de componentes principales.

QTL: Loci de caracteres cuantitativos (“Quantitative Traits Loci”)

ROH: bloques de homocigosidad (“Runs of Homozygosity”).

ROHi: Islas ROH

SNP: Polimorfismo de nucleótido simple (“Single Nucleotide Polimorfism”).

SPC: Sangre Pura de Carrera.

STR: microsatélites (“Short tandem repeats”).

TE: transferencia embrionaria.

Caracterización genética del caballo Polo Argentino mediante el análisis de datos genealógicos y de polimorfismos de nucleótido simple

Resumen

El objetivo general de esta tesis fue evaluar la variabilidad y estructuración genética de la raza equina Polo Argentino, utilizando datos genealógicos y genotípicos obtenidos con un microarray de SNP.

Se utilizaron los datos de pedigrí para calcular distintos parámetros genéticos poblacionales. Se seleccionaron 385 individuos Polo Argentino para ser genotipados con el microarray Illumina Equine GGP array (70k). Se realizó una caracterización de bloques de homocigosidad (ROH) y de la endogamia a nivel molecular. Se incorporaron datos genotípicos de otras 14 razas equinas para evaluar la estructuración, composición y diferenciación genética.

El análisis genealógico reveló que en la última generación más del 70% de los individuos nacieron por transferencia embrionaria. Esto se asoció a una reducción del intervalo generacional y mayor flujo genético entre criadores, pero también al aumento de la endogamia promedio y una contribución genética desigual de ancestros (principalmente Pura Sangre de Carrera - SPC).

El coeficiente de endogamia molecular fue mayor que el calculado con datos genealógicos (14% vs. 0,8%, valor- $p < 0,001$) y estuvo caracterizado por la presencia mayoritaria de ROH cortos (<5,5Mb) lo cual demuestra endogamia ancestral. Se identificaron 4 islas ROH ubicadas en los cromosomas 1, 7 y 17, las cuales contenían 25 genes con distintas funciones reportadas. El SPC fue la principal influencia genética de la raza, aunque se observó una leve e incipiente

diferenciación genética. Varias regiones genómicas presentaron elevados índices F_{ST} entre el Polo Argentino y SPC, conteniendo genes con funciones variadas, las cuales podrían constituir huellas de selección. Finalmente, algunos marcadores asociados a caracteres fenotípicos vinculados a la aptitud deportiva (velocidad y tamaño corporal) demostraron estar bajo selección en el caballo Polo Argentino.

El conocimiento de la situación genética de la raza Polo Argentino resulta de especial interés para guiar las decisiones de cría y contribuir con el progreso genético.

Palabras clave: caballo Polo Argentino, variabilidad genética, endogamia, bloques de homocigosidad, huellas de selección.

Genetic characterization of Argentine Polo horse using genealogical and single nucleotide polymorphisms data.

Abstract

The aim of this thesis was to assess the genetic variability and structure of the Argentine Polo horse using genealogical and genotypic data obtained with a SNP microarray. Pedigree data were used to calculate different population genetic parameters. A total of 385 Polo Argentino individuals were selected for genotyping with the Illumina Equine GGP array (70k). A characterization of runs of homozygosity (ROH) and molecular inbreeding was performed. Genotypic data from 14 other equine breeds were incorporated to evaluate genetic structure, composition, and differentiation.

Analysis of pedigree data revealed that over 70% of the individuals in the last generation were born by embryo transfer. This phenomenon was associated with a reduction in the generation interval and increased gene flow between breeders. However, it also resulted higher average inbreeding and an unequal genetic contribution of ancestors, mainly Thoroughbred.

The molecular inbreeding coefficient was higher than the coefficient calculated with genealogical data (14% vs. 0.8%, p -value <0.001). It was characterized by the presence of mostly short runs of homozygosity (<5.5 Mb), indicating ancestral inbreeding. Four ROH islands located on chromosomes 1, 7 and 17 were identified, which contained 25 genes with different reported functions. The Thoroughbred was the main genetic influence in the breed, although slight and incipient genetic differentiation was observed. Several genomic regions presented high F_{ST} indices between Argentine Polo and Thoroughbred, containing genes with diverse functions, which could constitute putative selection sweeps. Finally, some markers

associated with phenotypic traits linked to sporting ability, such as speed and body size, were found to be under selection in Argentine Polo horse.

Knowledge of the genetic status of the Argentine Polo horse is of special interest to guide breeding decisions and contribute to genetic progress.

Keywords: Argentine Polo horse, Genetic variability, inbreeding, runs of homozygosity, selection sweeps.

INTRODUCCIÓN

1. La producción equina en Argentina

Argentina es un país con una fuerte tradición en la producción equina, asociada principalmente al trabajo rural y al deporte. Esto ha incentivado la cría de caballos de diferentes razas y con distintas finalidades. Actualmente se crían en nuestro país más de 20 razas, que se registran en las 3 entidades de registros genealógicos reconocidas por el Ministerio de Agricultura, Ganadería y Pesca (MAGyP): Studbook Argentino, Asociación Argentina de Fomento Equino y Sociedad Rural Argentina.

La producción equina en Argentina no sólo se ve incentivada por el rol que el caballo ha ocupado en la cultura campestre, sino también por las ventajas geográficas que el entorno le confiere. Según datos del Ministerio, en 2022 existían alrededor de 2,4 millones de equinos, alojados en 209.679 establecimientos, entre los que se incluyen haras, hipódromos, clubes hípicos, centros tradicionalistas, establecimientos productivos, entre otros (MAGyP, 2023). Alrededor de 20 mil ejemplares de pura raza se inscriben anualmente en promedio en sus respectivos libros genealógicos. La industria hípica genera miles de puestos de trabajo, tanto directos como indirectos, incluyendo peones, petiseros, vareadores, domadores, herradores, veterinarios y agrónomos (entre tantos otros).

El desarrollo que ha alcanzado este sector en la Argentina ha generado un gran reconocimiento a nivel internacional. En este sentido, nuestro país está dentro de los principales productores de caballos Pura Sangre de Carrera, y es el principal productor y exportador de caballos Polo Argentino a nivel mundial, sin dudas uno de los estandartes de la cría hípica nacional (MAGyP, 2023).

1.1. Origen y difusión del polo

El polo es un deporte ecuestre muy antiguo cuyo origen se remonta al siglo III. Se cree que una versión primitiva de este juego era practicada por los pueblos nómades de Asia central y posteriormente en la antigua Persia, desde donde se dispersó a través de la Ruta de la Seda hacia otros países del continente asiático, como China e India (MOP, 2018). En estos países alcanzó cierta popularidad entre la realeza y los líderes militares de alto rango, quienes practicaban el deporte de manera asidua, en parte también como una forma de entrenamiento militar. La difusión del polo hacia occidente tuvo lugar a mediados del siglo XIX, cuando Gran Bretaña asume el control del subcontinente indio. El retorno hacia las Islas de los militares de caballería británicos que habían sido destinados al territorio indio propició el desarrollo del polo en Reino Unido, y su práctica se extendió entre las clases aristocráticas y oficiales de alto rango. Durante la segunda mitad del siglo XIX se establecen numerosos equipos y clubes de polo, entre ellos el de Hurlingham. Es precisamente en este contexto cuando se establecen las reglas básicas que rigen el polo como lo conocemos hoy en día (Parrish, 2018).

La difusión del polo también se dio hacia el continente americano, donde alcanzó una importante magnitud en Estados Unidos y en Argentina. En nuestro país, la llegada del polo se dio en el contexto del arribo de inmigrantes británicos hacia fines del siglo XIX, ligados a la producción agropecuaria y a la tenencia de grandes extensiones de tierra, lo que facilitó las condiciones para la práctica de este deporte. El establecimiento de esta nueva aristocracia británica en la Argentina favoreció la creación de numerosos clubes y la organización de competencias. Sin embargo, más allá del impulso dado por la influencia extranjera, la popularidad y el nivel de perfeccionamiento alcanzado por el polo en nuestro país se vio también

muy favorecida por la conexión con la cultura gaucha y la vinculación de los criollos con el caballo (Parrish, 2018).

1.2. El caballo de polo en Argentina

Si bien el polo ha ido evolucionando a través de los años, ciertas características se mantienen y hacen que el caballo utilizado para este deporte tenga cualidades particulares. Antiguamente se limitaba la alzada de los caballos que podían utilizarse para el juego (razón por la cual aún se suelen denominar a estos caballos como “ponis” o “petisos”), aunque actualmente ya no existen tales restricciones (MOP, 2018). El polo es un deporte muy dinámico que requiere de un caballo con mucha velocidad y potencia, con la agilidad necesaria para tolerar los constantes cambios de marcha y dirección. Pero, a su vez, debe tener un temperamento equilibrado, capaz de resistir la presión y la brusquedad del juego, ser dócil y fácil de manejar por el jinete. A pesar de que cada jugador utiliza varios caballos durante un partido, se requiere de animales con resistencia (*stamina*) capaces de recuperarse rápidamente para poder continuar con el juego.

En nuestro país, si bien en sus comienzos los caballos más utilizados para jugar al polo eran en general los mestizos de caballos Criollos utilizados para el trabajo en las estancias, el perfeccionamiento del deporte hizo necesaria la incorporación de animales más finos y veloces, lo que se logró con el cruzamiento con caballos Pura Sangre de Carrera (AACCP, 2023). La importancia de contar con caballos de calidad, los cuales hacen una gran diferencia dentro del campo de juego, llevó a la necesidad de formalizar la cría del caballo de polo, un tipo de animal que reúna las condiciones ideales para la práctica específica de este deporte. En este contexto, se crea en el año 1984 la Asociación Argentina de Criadores de Caballos de Polo, con el fin de organizar el registro genealógico de los ahora llamados

caballos Polo Argentino y fomentar su cría y mejoramiento (AACCP, 2023). El estándar racial del caballo de polo es flexible, pero se describen como animales equilibrados, con una alzada media a la cruz de 156 cm y un peso que oscila entre 400 y 500 kg. Su contextura es fuerte, con un cuerpo musculoso y perfil rectilíneo, y de temperamento sanguíneo (Imagen 1). Estas características acompañan una funcionalidad vigorosa y rápida, de acción baja y elasticidad en sus movimientos (Imagen 2, (AACCP, 2023)). El libro genealógico se mantiene abierto hasta hoy en día, permitiendo la incorporación de ejemplares externos que cumplan ciertos criterios de aptitud y tipo definidos por la propia asociación.



Imagen 1. Equino de raza Polo Argentino. Gentileza del criador Francisco Bargallo (Genética Santa Agueda).

Argentina es considerada cuna de los mejores polistas y líder en la cría y entrenamiento del caballo de polo a nivel mundial. Nuestro país se destaca por la calidad de los productos Polo Argentino, siendo animales muy demandados en países donde se juega polo de alto Hándicap. En la actualidad, la cabaña equina nacional está conformada por numerosas razas, entre las que se destaca el caballo Polo Argentino, no sólo por la mencionada calidad de los ejemplares, sino

también por la cantidad de productos inscriptos año a año. Junto con el caballo Pura Sangre de Carrera y el Criollo Argentino, constituyen las tres principales razas con mayor cantidad de nacimientos por año. Más aún, nuestro país es reconocido como el primer productor de caballos de polo en el mundo, alcanzando más del 70% del total de equinos exportados de Argentina (MAGyP, 2023).



Imagen 2. Yegua Polo Argentino disputando un partido de polo. Gentileza del criador Francisco Bargallo (Genética Santa Agueda).

A su vez, la cría del Polo Argentino se vale de la utilización comercial de biotecnologías reproductivas, en parte gracias a la laxitud de la Asociación de Criadores en la reglamentación y los criterios de inscripción de sus productos. De esta forma, la raza se caracteriza por el amplio uso de la inseminación artificial y la transferencia embrionaria, así como también una incipiente producción de embriones sexados (Herrera y col., 2014). Además, se destacan últimamente la clonación de individuos a escala comercial (Maserati y Mutto, 2016), posicionando a la raza a la vanguardia en términos de métodos de cría.

La transferencia embrionaria es una biotecnología reproductiva que comenzó a utilizarse en la década de los 80' en la especie equina, y consiste en la colecta de los embriones producidos por una yegua donante en estadio de blastocisto, para ser trasferidos a una yegua receptora, que finalmente llevará a cabo la gestación. Esto ha facilitado la reproducción en yeguas gerontes o subfértiles, así como en potrancas jóvenes o que estén en competición. Precisamente estas características han sido de sumo interés para la raza Polo Argentino, donde las hembras (por sus características de docilidad y facilidad de entrenamiento) son ampliamente preferidas para el juego lo que, a su vez, hace que su carrera deportiva sea prolongada, impidiendo su reproducción natural. De esta forma, yeguas que aún están siendo utilizadas para el deporte, pueden destinarse como donantes de embriones en programas de transferencia embrionaria, o incluso pueden obtenerse una mayor cantidad de potrillos por yegua en una temporada. Pero, además, desde el punto de vista genético, el uso de transferencia de embriones, así como la inseminación artificial, permite incrementar la presión de selección por vía materna o paterna, respectivamente, aumentando el progreso genético (Squires y col., 1999; Campbell, 2014).

Con la creación de la AACCP y la gran demanda de caballos de polo de élite, la raza ha atravesado una gran expansión y crecimiento en los últimos años. Esto ha derivado en una rápida evolución fenotípica y un perfeccionamiento de los ejemplares, producto de la cuidadosa selección realizada por los criadores. Si bien no existe un programa de mejora formalmente instaurado, la utilización de biotecnologías reproductivas (como la inseminación artificial y la transferencia embrionaria) ha permitido la rápida dispersión de genética considerada "superior" o mejoradora, acompañada por la constitución del libro genealógico en sí mismo. Sin embargo, este rápido e intenso proceso selectivo podría generar rápidos

cambios en la variabilidad genética, producto de la aparición de cuellos de botella poblacionales o niveles de endogamia elevados.

2. Caracterización genética

La caracterización de los recursos zoogenéticos incluye al conjunto de actividades tendientes a la descripción y conocimiento de su estado de situación, como así también de la tendencia seguida durante un período de tiempo dado (FAO, 2010). La caracterización genética es una herramienta fundamental para guiar las decisiones de cruzamiento y selección en una población. Pero también permite conocer la variabilidad genética existente en una población, así como su estructura, lo cual aporta conocimientos sobre los fenómenos a los que ha sido sometida y permite detectar eventos que puedan afectarla (Valera y col., 2005).

Según la Organización de las Naciones Unidas para la Alimentación y Agricultura (FAO), una raza puede definirse como un grupo homogéneo y subespecífico de animales domésticos con características externas definibles e identificables que permiten distinguirlo de otros grupos dentro de la misma especie. Sin embargo, este término muchas veces cobra un sentido cultural que no se ajusta perfectamente a lo técnico, y se utiliza más bien para definir un grupo para el cual la separación geográfica y/o cultural de grupos fenotípicamente semejantes ha llevado a la aceptación de su identidad por separado (FAO, 2021). Es por esto por lo que una raza puede entenderse como una población en términos genéticos, donde los individuos se reproducen entre sí, manteniendo y transmitiendo un acervo génico particular.

El conocimiento del pool genético es un punto de partida fundamental en la gestión de una raza, así como para el diseño de estrategias de mejoramiento. El manejo

efectivo de los recursos genéticos de los animales domésticos requiere un exhaustivo conocimiento de las características de las razas, lo que incluye datos demográficos, de rendimiento, de adaptación al ambiente y de diversidad genética dentro y entre razas (Groeneveld y col., 2010). En este sentido, el análisis de datos demográficos es un punto clave para la determinación del estado de situación de una raza en particular. Uno de los parámetros más frecuentemente utilizados y recomendados por FAO para el monitoreo de poblaciones es el tamaño efectivo de la población (N_e). Si bien existen múltiples formas de cálculo del N_e , algunas de ellas permiten indirectamente estimar la tasa de endogamia, pudiendo predecir así el nivel de riesgo, independientemente del tamaño real de dicha población (Groeneveld y col., 2010).

2.1. Diversidad genética

La diversidad genética puede ser entendida como el espectro de diferencias o variantes genéticas (polimorfismos) presentes en una población, sea ésta una especie, raza o grupo particular de animales. Esta variabilidad es fruto de miles de años de evolución, y se remonta en general al origen y desarrollo de la especie. En el caballo, las diferencias ambientales, así como el propio proceso de domesticación y selección hecho por el humano, han derivado en la formación de múltiples razas genéticamente definidas que mantienen la diversidad dentro de la especie (Groeneveld y col., 2010). Sin embargo, en el último siglo se puede observar una erosión general de los recursos genéticos, producto de la implementación de estrategias de mejoramiento para maximizar el rendimiento (FAO, 2010).

En términos generales, una menor diversidad genética ha sido asociada a una menor adaptabilidad biológica, en múltiples especies, incluida la equina. Sin bien

existen excepciones, el aumento de la tasa de endogamia (producto de la pérdida de variabilidad genética) se vincula con una baja en la fertilidad, una mayor susceptibilidad a enfermedades, una menor adaptabilidad al medio, y un menor rendimiento, entre otros (Charlesworth y Willis, 2009). Por otro lado, una menor diversidad provoca una menor respuesta a la selección, ya que se reduce la varianza genética y las diferencias observadas entre los individuos, las cuales pasan a ser explicadas en mayor medida por factores no-genéticos.

Dentro de la dinámica de las poblaciones, son múltiples los eventos que pueden desencadenar cambios más o menos abruptos en la variabilidad genética. La pérdida de variantes alélicas por fenómenos estocásticos como la deriva génica y la presencia de cuellos de botella poblacionales pueden ser un desencadenante, principalmente en poblaciones pequeñas y cerradas (Falconer, 1989). En cambio, los cruzamientos dirigidos, el aislamiento y la selección juegan un rol fundamental en las razas domésticas, porque pueden determinar la rápida dispersión y fijación de caracteres (deseables o no), pero también la pérdida abrupta de variantes en pocas generaciones (Groeneveld y col., 2010). A su vez, la intensidad de selección puede ser mayor con la implementación de biotecnologías reproductivas (como la inseminación artificial y la transferencia embrionaria, ampliamente difundidas en la especie equina), generando un mayor número de descendientes por cada reproductor de lo que sería posible por métodos naturales. Este hecho además producirá, en la mayoría de los casos, un incremento del parentesco entre los individuos de la población, haciendo que se generen cruzamientos consanguíneos (Nicholas, 1996; Gengler y Druet, 2001).

2.2. Consanguinidad y depresión endogámica

La consanguinidad o endogamia (“inbreeding”) es el efecto genético fruto del cruzamiento entre individuos emparentados. En términos de magnitud, su importancia puede ser medida a través de un coeficiente (coeficiente de endogamia, F), que determina la probabilidad que un individuo tenga en un locus dado dos alelos idénticos por descendencia. Esto implica que, para cada locus, un individuo herede dos copias iguales de un alelo derivadas de un ancestro en común (Wright, 1931). Es por esto por lo que la endogamia es un reflejo de la homocigosidad que presenta un animal o una población determinada.

En una población cerrada, es inevitable que la selección produzca un incremento de la endogamia a través del tiempo. Si analizamos diferentes razas, estas pueden comportarse como una población más o menos cerrada, dependiendo de la flexibilidad de las asociaciones que la gestionan en permitir el cruzamiento con individuos de otras razas. Sin embargo, en la actualidad, la mayoría de las razas equinas pueden considerarse “cerradas” en función de la contribución desigual de unos pocos reproductores a la siguiente generación. Tal es así que muchas veces se recurre al “inbreeding” como una estrategia de cruzamiento, con el objetivo de fijar caracteres deseables (relacionados al rendimiento), aumentando la uniformidad en la población. A su vez, el incremento de la utilización de animales de élite (mejorantes) contribuye al aumento de las tasas de endogamia, lo que recae en una disminución del N_e . Valores de N_e bajos pueden acrecentar los efectos de la depresión endogámica y la deriva génica, lo cual podría comprometer la viabilidad de una población a largo plazo (Peripolli y col., 2016).

La endogamia disminuye la media poblacional fenotípica de caracteres relacionados con la aptitud biológica, fenómeno conocido como depresión endogámica (Charlesworth y Willis, 2009). Esto puede evidenciarse por la

aparición de fenotipos anormales (muchos de los cuales pueden ser letales en etapas tempranas de la gestación) y enfermedades de base genética, así como por una reducción en la supervivencia, la fertilidad o las tasas de crecimiento de la descendencia de individuos muy emparentados. Tradicionalmente se ha relacionado a la depresión endogámica con caracteres de supervivencia y aptitud biológica, pero sus efectos también han sido documentados en otros caracteres, muchos de los cuales son de importancia productiva (Leroy, 2014). Sin embargo, la expresión de la depresión endogámica puede ser variable, por lo que entender la base genética por la cual se produce resulta clave. Se han descrito tres mecanismos genéticos responsables de este fenómeno: dominancia parcial (aumento en la expresión de alelos recesivos deletéreos), sobredominancia (superioridad fenotípica de los heterocigotas, en comparación a los homocigotas) y epistasis (mayor probabilidad de combinaciones genéticas favorables en los heterocigotas) (Leroy, 2014). Sin embargo, el peso relativo de un mecanismo u otro sobre la aparición de la depresión endogámica es variable en cada población, y difícil de cuantificar.

Los efectos causados por depresión endogámica han sido documentados también en la especie equina. Laseca y col. (2020) estudiaron el efecto de la endogamia sobre los valores de cría estimados para el número de partos en yeguas Pura Raza Española, encontrando una correlación negativa entre ambas variables, fundamentalmente al analizar la homocigosidad presente en el cromosoma X. Pirosanto y col. (2020) estudiaron la calidad espermática de sementales de la misma raza, encontrando un efecto negativo de la endogamia sobre el volumen del eyaculado y la motilidad espermática progresiva. También se han asociado a la consanguinidad ciertas condiciones patológicas como la retención placentaria en yeguas Frisonas (Sevinga y col., 2004) y deformidades flexurales de los

miembros por retracción del tendón flexor digital profundo (*Club Foot*) en caballos Árabes (Comparini y col., 2019). Por otro lado, se ha descrito la presencia de depresión endogámica en diferentes caracteres morfológicos (como alzada a la cruz, altura al pecho, longitud del cuerpo, entre muchos otros) en el caballo Pura Raza Español (Gómez y col., 2009) y Campolina de Brasil (Bussiman y col., 2018). Estos hallazgos cobran una especial relevancia, dada la importancia que tiene una correcta conformación en el desempeño deportivo de los individuos. Estudios recientes han evaluado el efecto del inbreeding sobre la performance de caballos Pura Sangre de Carrera, sugiriendo la persistencia de alelos potencialmente deletéreos en la población (carga genética) a pesar de la continua selección por rendimiento (Todd y col., 2018).

El fenómeno de la endogamia y sus consecuencias es complejo y variable entre razas, por lo que su análisis resulta clave. En una población existen alelos recesivos con efecto deletéreo que se mantienen en equilibrio selección-mutación, y que la mayoría se encuentran ocultos en los genotipos heterocigotas (Nicholas, 1987). La consanguinidad, al aumentar la proporción de loci en homocigosis, favorece la expresión de dichos alelos recesivos. Bajo la hipótesis de la dominancia parcial, es posible que la endogamia y la selección permitan la eliminación paulatina de los alelos deletéreos, fenómeno conocido como purga de la carga genética indeseable (alelos con efecto deletéreo o que produzcan una desventaja visible sobre el fenotipo de un individuo (Hedrick y Garcia-Dorado, 2016)). Sin embargo, esta purga es más probable que se produzca bajo una fuerte presión de selección y bajas tasas de endogamia.

La selección en contra de los alelos recesivos puede ser más o menos eficaz, dependiendo que la intensidad de selección sea mayor que la deriva génica. En aquellas poblaciones bajo una fuerte selección, es probable se produzca una

disminución de la frecuencia de los alelos indeseables a través de las generaciones, aunque difícilmente sea absoluta. Esto no sólo depende de las estrategias de cría que lleven a cabo los criadores, sino también de la magnitud del efecto que tengan dichas variantes sobre el fenotipo, así como del tamaño de la población (Boakes y col., 2007). Aquellos alelos que tengan un efecto fenotípico importante serán más fácilmente eliminados de la población que los de efecto pequeño (por ejemplo, aquellos relacionados a la supervivencia vs. el rendimiento deportivo). Sin embargo, en las poblaciones domésticas debe considerarse la presencia de flujo genético, a través de la incorporación de nuevos reproductores, así como la posibilidad de ligamiento con regiones afectadas a la sobredominancia o barrido selectivo, lo cual dificulta el proceso de purga. Por otro lado, un incremento abrupto y marcado de la endogamia puede lograr que los alelos se fijen más rápidamente de lo que la selección podría eliminarlos.

La heterogeneidad de la depresión endogámica es otro factor para considerar, que ha sido descrito dentro y entre poblaciones. Como es de suponer, no todos los individuos portan en la misma medida los alelos potencialmente deletéreos, por lo que cada reproductor puede contribuir de manera diferente a la depresión endogámica de una población (Lacy y col., 1996; Baumung y col., 2015).

Por todo lo expuesto, se entiende que el efecto de la depresión endogámica depende en gran medida de la dinámica de la población estudiada, en la cual intervienen factores relacionados a la presión de selección, la tasa de incremento de la endogamia, la presencia de cuellos de botella, la posibilidad de flujo genético y la utilización de biotecnologías reproductivas, entre muchos otros. Es importante considerar estos aspectos para definir el estado de situación de una raza y planificar su futura evolución, con relación a programas de cría y mejoramiento.

2.3. Métodos de caracterización genética

Los registros genealógicos son datos útiles en la caracterización genética de una raza, ya que permiten estimar una serie de parámetros de importancia en la genética poblacional. Sin embargo, esto es posible solamente en aquellas razas formalmente establecidas que cuentan con una asociación de criadores que llevan adelante el registro y filiación de sus animales. Aun así, la calidad y profundidad del pedigrí tienen un rol central, ya que de ellos dependerá en buena parte la correcta interpretación de los resultados (Kardos y col., 2015; Perdomo-González y col., 2022).

La caracterización genética también puede realizarse a escala genómica, a partir de la determinación de distintas variantes presentes en el ADN de los individuos de una población. La determinación del genotipo de un individuo para una serie de locus dados también permite determinar la existencia de polimorfismos que pueden ser utilizados para su identificación o para el estudio de diversos caracteres fenotípicos asociados a las regiones del genoma en la que se encuentran. A su vez, el análisis de la variabilidad existente dentro de una población permite caracterizar genéticamente a las especies o razas y establecer diferencias entre las poblaciones e individuos (Peral García y col., 1995). Los estudios de caracterización han ido evolucionando conforme se han desarrollado nuevas técnicas de genética molecular. Para ellos se han estudiado desde polimorfismos bioquímicos, como los grupos sanguíneos y proteínas séricas (Peral García, 1994), hasta la detección de marcadores microsatélites (STR), los cuales siguen utilizándose hasta hoy en día para el control de filiación, e inclusive para caracterizar genéticamente las poblaciones (Delgado y col., 2014).

Sin embargo, existe una tendencia a reemplazar el uso de marcadores STR por SNP (polimorfismos de nucleótido simple), los cuales son una fuente de variación

genética de mucha mayor escala. Si bien su variabilidad está limitada a una mutación puntual dentro de la secuencia de ADN, su abundancia dentro del genoma los convierte en la fuente más frecuente de cambios polimórficos de los seres vivos. Además, hoy en día se cuenta con tecnologías que permiten obtener una gran cantidad de información genética acerca de cada individuo a un costo mucho menor y de manera automatizada (Schaefer y McCue, 2020), mediante el empleo de microarrays de SNP (también denominados chips o microarreglos). Estos permiten analizar la presencia de decenas de miles de variantes genéticas a partir de una simple muestra de ADN. En términos técnicos, cada microarray contiene miles de sondas de ADN conocidas que están diseñadas para detectar los alelos existentes en miles de SNP (previamente determinados). Mediante la hibridación específica con las regiones complementarias de cada una de las posibles variantes de cada loci se determina el genotipo de la muestra analizada.

Los microarrays de SNP tienen múltiples aplicaciones, como el mapeo de caracteres de interés, la caracterización genómica y la comparación entre poblaciones (Schaefer y McCue, 2020). En la especie equina son ampliamente utilizados en estudios poblacionales, incluyendo el cálculo de la consanguinidad (Todd y col., 2018), la detección de huellas de selección (Santos y col., 2021), la determinación de la variabilidad y diferenciación genética entre grupos o poblaciones (Gurgul y col., 2014; Schurink y col., 2019), entre otros. Esta metodología es ampliamente ventajosa desde el punto de vista económico, ya que permite el análisis simultáneo de una gran cantidad de marcadores a partir de una muestra de ADN, obteniendo así un perfil genético global del individuo.

Existe un creciente interés en la genotipificación de razas equinas mediante técnicas de genotipado masivo para su caracterización, ya que permitirá una mejor comprensión del estado de situación, la relación existente con otras razas y sus

orígenes, así como la asociación con enfermedades o caracteres de interés para el rendimiento deportivo (Petersen y col., 2013a; Petersen y col., 2013b). Esto constituirá el puntapié inicial para el posterior desarrollo de estudios de asociación y planes de mejoramiento acordes a las necesidades del sector, en particular en la Argentina, donde este tipo de estudios es escaso.

3. Justificación

A pesar de la importancia que tiene la cría caballar en nuestro país, la producción se realiza desde antaño de manera artesanal, siendo las decisiones de mejora genética tomadas de manera intuitiva por los productores. Tampoco existe un programa de mejora que permita seleccionar animales de calidad, en base a sus características morfológicas y funcionales acordes a los posibles objetivos de selección para la disciplina deportiva en que participan los animales (en este caso el polo). No obstante, los criadores han realizado una fuerte selección en la cría del Polo Argentino, logrando un caballo con una aptitud muy codiciada para la práctica del deporte, no sólo en Argentina sino también en el resto del mundo. La evolución fenotípica y funcional de la raza, acompañada del amplio uso de biotecnologías reproductivas (transferencia embrionaria, inseminación artificial, etc.) y de la clonación, resultan eventos de interés para el estudio de la variabilidad y estructuración genética, así como de su relación con otras razas.

Por lo tanto, en esta tesis se realizará una caracterización genética del caballo Polo Argentino, evaluando su variabilidad y estructura genética, así como la relación con otras razas equinas, utilizando datos genealógicos y genómicos.

HIPÓTESIS

En esta tesis se plantean tres hipótesis de trabajo:

- El uso de marcadores moleculares de tipo SNP permite discriminar a los caballos Polo Argentino de ejemplares pertenecientes a otras razas equinas.
- La alta presión de selección y el amplio uso de biotecnologías reproductivas está produciendo un estrechamiento de la variabilidad genética del caballo Polo Argentino y un aumento de la endogamia de la población.
- La distribución y cantidad de bloques de homocigosidad en el genoma pueden utilizarse para inferir genes y/o rutas metabólicas potencialmente afectadas por la endogamia.

OBJETIVO GENERAL

El objetivo general del presente trabajo de tesis fue determinar la variabilidad y estructuración genética de la raza equina Polo Argentino, por medio de la aplicación de técnicas de genotipado masivo, empleando microarrays de SNPs.

Objetivos específicos

La tesis contempla los siguientes objetivos específicos:

- Realizar la caracterización genética preliminar de la raza Polo Argentino mediante el análisis de los datos de pedigrí.
- Seleccionar un grupo de animales específicos en base al análisis genealógico que representen la variabilidad existente de la raza, para su caracterización utilizando técnicas moleculares de genotipado masivo.
- Determinar las frecuencias alélicas de los marcadores analizados y compararlas con animales pertenecientes a otras razas equinas de importancia con el fin de comprender mejor la estructuración genética de la población estudiada.
- Establecer los niveles de endogamia de la población estudiada mediante estudios moleculares y compararlos con los obtenidos mediante los registros de pedigrí, determinado de manera real la variabilidad genética poblacional e individual, así como su variación a lo largo del tiempo.

- Determinar las posibles implicancias genéticas de la endogamia mediante el análisis de las regiones de homocigosidad (ROH) y su efecto genómico-funcional utilizando técnicas bioinformáticas de ontología génica.

CAPÍTULO I: Análisis de los datos genealógicos en el caballo Polo Argentino: caracterización genética e impacto de las biotecnologías reproductivas sobre la variabilidad.

1. Introducción

El estudio de la estructura genética y demográfica de las poblaciones animales es una herramienta valiosa para analizar circunstancias o eventos importantes que afectan a su historia (Valera y col., 2005). Históricamente, esto se han realizado analizando datos de pedigrí o genealógicos en razas equinas de todo el mundo (Maccluer y col., 1983; Poncet y col., 2006; Bartolomé y col., 2011; Pjontek y col., 2012; Vicente y col., 2012; Borowska y Szwaczkowski, 2014; Petersen y col., 2014). Además de su uso en la caracterización genética poblacional, los resultados de estos análisis son especialmente interesantes como herramienta auxiliar en las decisiones de cría, ya que muchas veces las mismas son realizadas por los criadores de manera meramente artesanal (Medeiros y col., 2014; Giontella y col., 2019). Sin embargo, la existencia de registros genealógicos detallados y consistentes es el primer requerimiento para obtener resultados fiables, así como para implementar programas de mejora genética basados en metodologías científicas (Maccluer y col., 1983; Cervantes y col., 2008).

El caballo Polo Argentino es muy popular y codiciado entre los jugadores de polo de élite a nivel mundial (Maserati y Mutto, 2016). Si bien lleva cuatro décadas de registro, se mantiene un libro genealógico abierto, permitiendo la incorporación de individuos de cualquier otra raza, siempre y cuando cumplan ciertos requisitos de aptitud y biotipo. Adicionalmente, la Sociedad Rural Argentina lleva adelante el registro genealógico de la raza realizando la identificación individual y el control

de filiación mediante la genotipificación de un panel de marcadores microsatélites recomendados por la ISAG.

Un aspecto interesante a destacar es que, desde su creación, la AACCP permite (e incluso fomenta) el uso de biotecnologías reproductivas avanzadas que estén disponibles en la especie equina. Es por esto que la raza Polo Argentino se caracteriza por la existencia de grandes programas de transferencia embrionaria (TE), como así también por el uso de biotecnologías de última generación, como el sexado de embriones (Herrera y col., 2014), la transferencia de ovocitos (Riera y col., 2016) y la clonación (Maserati y Mutto, 2016; Gambini y Maserati, 2018). Esto ha posicionado al caballo Polo Argentino como líder en términos de tecnologías reproductivas. A su vez, el caballo de polo es en la actualidad uno de los caballos argentinos más exportado y distribuido alrededor del mundo (MAGyP, 2023). Todas estas características convierten a la raza en un interesante modelo para analizar los efectos reales que ha tenido la implementación a escala comercial de las biotecnologías reproductivas, y en particular la TE. Si bien en otras especies ganaderas se ha demostrado que esta técnica produce una mejora genética y una mayor presión de selección (Mota y col., 2013), también se la ha asociado con la disminución de la variabilidad genética, tanto intra como entre poblaciones.

A pesar de la importancia cultural y económica de la raza, los estudios genéticos y/o reproductivos en el caballo de polo argentino son prácticamente nulos. Es por ello que este capítulo tiene por objetivos:

- Realizar la caracterización genética preliminar de la raza Polo Argentino mediante el análisis de los datos de pedigrí.
- Evaluar el efecto del uso de los programas de TE a gran escala sobre la variabilidad genética.

2. Materiales y métodos

2.1. Datos genealógicos

En este estudio se utilizó la información de registros de pedigrí de los caballos Polo Argentino existentes en las bases de datos disponibles en las páginas web de la AACCP y de la Sociedad Rural Argentina. Adicionalmente, se incluyeron datos obtenidos del Studbook Argentino (registros de caballos Pura Sangre de Carrera), con el objetivo de reconstruir las genealogías de los individuos fundadores más importantes. La base de datos final incluyó 115.615 ejemplares registrados hasta el año 2021. Con el objeto de mejorar la precisión de los parámetros calculados, sólo los individuos con al menos tres generaciones completas equivalentes (ECG) (Maignel y col., 1996) y nacidos entre los años 2006 y 2021 (últimas dos generaciones, aproximadamente) fueron usados como población de referencia (grupo PR; n=41.873). Finalmente, con fines experimentales, este grupo fue dividido en dos subgrupos: individuos nacidos por transferencia embrionaria (grupo TE; n=32.731) e individuos nacidos por gestación por su madre biológica (grupo NO-TE; n=9.142). Cabe aclarar que aquellos ejemplares clonados que tuvieran hijos fueron reemplazados por el animal fundador (original) en la base de datos para los cálculos posteriores (excepto para el ítem 2.2 Análisis demográfico).

2.2. Análisis demográfico

El análisis inicial incluyó una descripción del número de animales registrados por año, los individuos utilizados como sementales y yeguas madres, así como la proporción de sexos nacidos y registrados. Se calculó el intervalo generacional (IG), definido como la edad media de los padres al nacimiento de su descendencia que será destinada a la reproducción (James, 1977), por las cuatro vías posibles

(padre-hijo, padre-hija, madre-hijo y madre-hija). Las comparaciones entre los grupos nacidos por transferencia embrionaria y por gestación natural de esta sección incluyen a todos los individuos registrados (y no sólo a los de la últimas dos generaciones), por lo que se referirá como TE_{TOTAL} y $NO-TE_{TOTAL}$. Las diferencias estadísticas entre la proporción de sexos, el número de hijos e IG por grupos fueron evaluadas mediante la prueba de proporciones Z y prueba T de Student.

2.3. Parámetros de variabilidad genética

La variabilidad genética fue evaluada mediante los siguientes parámetros:

- I. Coeficiente de endogamia individual (F_{PED}), definido como la probabilidad de que un individuo posea dos alelos idénticos por descendencia (Wright, 1931).
- II. Tasa de endogamia entre grupos (F_R), la cual cuantifica la proporción del F_{PED} del grupo ET comparado con NO-ET, determinando la dinámica de la variación del inbreeding entre grupos. Este parámetro fue calculado como:

$$F_R = \frac{F_{TE} - F_{NO-TE}}{F_{NO-TE}}$$

Donde F_{TE} y F_{NO-TE} son la media de los F_{PED} de los grupos TE y NO-TE, respectivamente.

- III. Coeficiente de parentesco medio (AR – “average relatedness”), definido como la probabilidad de que un alelo tomado al azar de la población provenga de un animal determinado (Gutierrez y Goyache, 2005). Cuanto mayor sea este parámetro, mayor es la representatividad que tiene un

individuo en la población y mayor es la participación que tiene un animal fundador en el origen de la raza.

- IV. Coeficiente de coascendencia para cada par de individuos (f), estimado como la probabilidad de que dos individuos compartan alelos idénticos por descendencia (Malécot, 1948).
- V. Tamaño efectivo de la población, definido como el número de animales que mantendría el actual aumento de la endogamia, si contribuyeran equitativamente a la siguiente generación. Este parámetro fue estimado tanto por vía incremento individual de la endogamia (N_E) (Gutiérrez y col., 2009) como de la coascendencia (N_{EC}) (Cervantes y col., 2011).

Las diferencias entre los grupos se determinaron usando la prueba de Mann-Whitney, ya que los datos no presentaron distribución normal (prueba de Kolmogorov Smirnov, valor- $p < 0,01$).

2.4. Análisis de fundadores y ancestros

Se calculó el número efectivo de fundadores (F_E), definido como el número de fundadores que produciría la misma variabilidad genética observada en la población si contribuyeran equitativamente a la misma; y el número efectivo de ancestros (F_A), definido como el número mínimo de individuos (no necesariamente fundadores) que explican la variabilidad genética de la raza. De acuerdo con Boichard y col. (1997), este enfoque dual permite determinar la existencia de cuellos de botella, calculando la proporción F_E/F_A , la cual es mayor a 1 cuando la contribución de los ancestros es desigual. Adicionalmente, se determinó el número de genomas fundadores equivalentes (F_G), definido como el número de fundadores que se espera que produzcan la misma diversidad genética que en la población estudiada, si los fundadores estuvieran igualmente representados y no

se produjera ninguna pérdida de alelos (Lacy, 1989). Este parámetro tiene en cuenta específicamente la pérdida aleatoria de alelos durante los cuellos de botella genéticos producidos por las contribuciones desequilibradas de los fundadores. Por último, se evaluó la composición racial de los diez ejemplares que más contribuyeron a la variabilidad genética, para determinar la influencia de las razas foráneas en la constitución actual del Polo Argentino.

2.5. Estructura poblacional

Se evaluó la posible estructuración de la raza en subpoblaciones parcialmente aisladas entre criadores, a través del cálculo del índice de fijación F_{ST} , siguiendo la metodología descrita por Caballero y Toro (2002). Este índice se calculó entre los criadores que cumplían dos requisitos: (1) criar tanto animales TE y NO-TE y (2) tener un tamaño mínimo de animales registrados (establecido en 40 ejemplares). En total se seleccionaron 83 criadores, con los que se obtuvieron dos índices F_{ST} entre criadores, uno representó la distancia genética entre los productos nacidos por TE y el otro la distancia entre los animales NO-TE. Entre ambos índices se calculó el coeficiente de correlación de Spearman. Este parámetro permitió determinar si la similitud genética entre criadores se veía afectado por el sistema de cría.

Para identificar subdivisiones menos evidentes resultantes de estructuraciones parciales, se realizó un análisis adicional de acuerdo con la metodología descrita por Cervantes y col. (2011). Ésta se basa en considerar la relación entre el tamaño efectivo de la población calculado vía incremento individual de la coascendencia y la endogamia entre pares (N_{EC}/N_E).

Todos los parámetros fueron calculados utilizando el software *ENDOG v.4.8.* (Gutierrez y Goyache, 2005). El análisis estadístico se realizó con el software *STATISTICA v.12.* (StatSoft, 2012).

3. Resultados

3.1. Análisis demográfico

El número de ejemplares nacidos e inscritos anualmente desde la creación de la AACCP en 1984 aumentó considerablemente hasta 2009 (Figura 1.1), con aproximadamente el 67% de la población total registrada durante las últimas dos generaciones (grupo PR). Durante el período 2006-2021, la proporción de hembras/machos se mantuvo relativamente estable (cerca de 1,5:1), pero comenzó a aumentar en el año 2010, alcanzando tasas cercanas a 4,75:1 en 2019 (Figura 1.1). La integridad del pedigrí mostró una media de $29,8 \pm 2,4$ generaciones máximas y $5,1 \pm 1,5$ ECG, lo que constituyen valores robustos en comparación con estudios similares (Bartolomé y col., 2011; Medeiros y col., 2014). Asimismo, el ECG fue de $5,3 \pm 1,5$ y $4,4 \pm 1,2$ en los grupos TE y NO-TE, respectivamente.

Los individuos nacidos por transferencia embrionaria comienzan a registrarse en el año 1990, mostrando un aumento significativo desde 2004, alcanzando porcentajes superiores al 70% del total de caballos inscritos en el AACCP durante los últimos nueve años evaluados (Figura 1.2).

Al momento de realizar este estudio, sólo el 12% de los machos y el 28% de las hembras de toda la población se convirtieron en sementales y yeguas madres, produciendo una media de 28,7 y 4,2 potros por individuo, respectivamente.

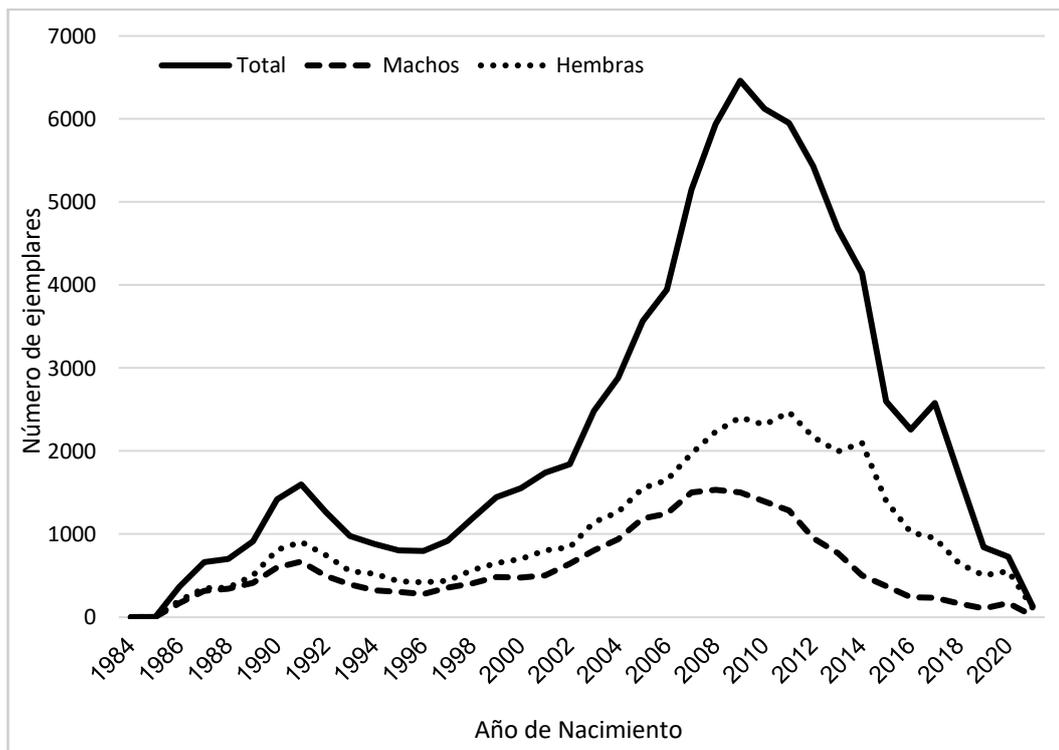


Figura 1.1. Evolución del número de ejemplares registrados por año de nacimiento desde la apertura del libro genealógico

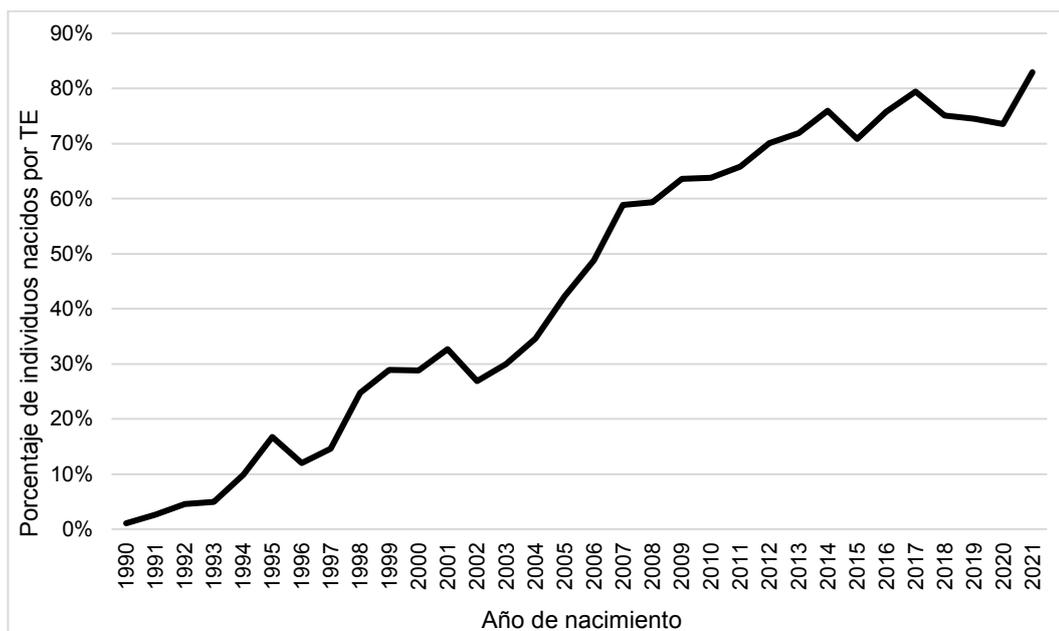


Figura 1.2. Porcentaje de individuos nacidos por transferencia embrionaria y registrados en el libro genealógico.

Sin embargo, hubo un claro desequilibrio en la contribución individual, ya que 36 sementales (1,2% del total) registraron más de 300 potros cada uno, lo que representa el 18% del total de ejemplares registrados en la base de datos.

Además, un mayor porcentaje de caballos fueron utilizados como sementales en TE_{TOTAL} (14,9% frente a 9,2% en NO-TE; valor-p<0,01), produciendo también en promedio 12 potros más cada uno (36,8 en TE_{TOTAL} frente a 24 en NO-TE_{TOTAL}). Por el contrario, el porcentaje de hembras utilizadas como yeguas madres fue significativamente menor en TE_{TOTAL} (22,3%) comparado con NO-TE_{TOTAL} (36,5%, valor-p<0,01). La cantidad promedio de crías nacidas por transferencia embrionaria por yegua y por año fue de $2 \pm 0,27$.

El IG fue de 7,8 años en promedio para toda la población (Tabla 1.1). Todas las vías fueron considerablemente menores en los caballos del grupo TE_{TOTAL} en comparación con los caballos NO-TE_{TOTAL} (Tabla 1.1), con una diferencia de casi 2 años en promedio (7 frente a 8,9 años, respectivamente; valor-p < 0,01).

Tabla 1.1. Intervalo generacional en el caballo Polo Argentino

Vía	Intervalo generacional (años)		
	Total	TE _{TOTAL}	NO-ET _{TOTAL}
Padre-hijo	7,7 ± 4,2	7 ± 4,4 ^a	9 ± 4,1 ^b
Padre-hija	7 ± 3,5	6,4 ± 3,8 ^a	7,6 ± 3,5 ^b
Madre-hijo	8,1 ± 3,6	7,5 ± 3,4 ^a	10,4 ± 4 ^b
Madre-hija	8,1 ± 3,8	7 ± 3,5 ^a	9,7 ± 3,8 ^b
Promedio	7,8 ± 3,8	7 ± 3,6 ^a	8,9 ± 3,9 ^b

Valores expresados como media ± desvío estándar. Las comparaciones estadísticas se realizaron entre los grupos TE y NO-TE. Letras diferentes indican diferencias significativas (valor-p<0,01).

3.2. Parámetros de variabilidad genética

En líneas generales, los valores de consanguinidad fueron bajos ($F_{PED}=0,86\%$ en PR), pero mostraron una tendencia al alza durante los últimos años, con porcentajes más altos en el grupo TE comparado con NO-TE (0,9 vs 0,6%, respectivamente; valor- $p<0,001$; Figura 1.3). También se encontró un incremento del F_{PED} promedio por generaciones (Tabla Suplementaria S1.1). Aunque la F_R mostró una tendencia variable, fue predominantemente positiva (Figura 1.3), demostrando así un claro efecto del grupo TE en el aumento de la consanguinidad. Esto concuerda con el aumento observado en la proporción de animales consanguíneos ($F_{PED} > 0$), que se incrementó del 39% (período 2006-2013) al 61% (período 2014-2021, valor- $p<0,01$). Se observaron también diferencias entre la proporción de individuos endogámicos entre TE y NO-TE (50% vs. 29%, respectivamente; valor- $p<0,01$; Tabla Suplementaria S1.2).

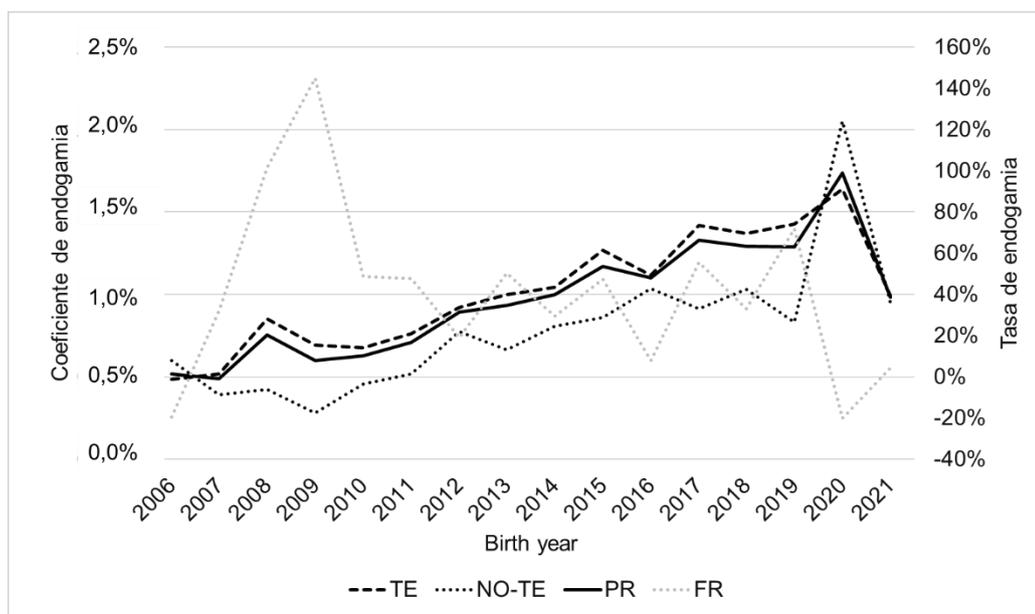


Figura 1.3. Evolución del coeficiente de endogamia durante los últimos 16 años.

Asimismo, la proporción de animales con alta consanguinidad (definida como $F_{PED} > 6,25\%$; equivalente a un cruzamiento entre primos hermanos) en el grupo PR fue de 5,6%, aumentando de 4,8% a 7,8% el período 2006-2013 a 2014-2021 (valor- $p < 0,01$). También se observaron valores significativamente más altos en TE (6%) que en NO-TE (4,7%) (valor- $p < 0,01$). Esta diferencia se manifestó también por la alta correlación observada entre F_{PED} y el porcentaje de animales nacidos por TE por año ($r = 0,72$; valor- $p < 0,01$). El promedio de F_{PED} en animales endogámicos ($F_{PED} > 0$) fue de 1,9%, siendo mayor en el grupo NO-TE respecto a TE (2,2% vs. 1,8%, valor- $p < 0,01$), probablemente debido a la mayor proporción de individuos con $F_{PED} \geq 0,125$ en el grupo NO-TE (ver Tabla Suplementaria S1.2). El coeficiente AR fue de 1,2% en PR, encontrándose diferencias significativas entre ambos grupos analizados (TE = 1,3% vs. NO-TE = 0,7%; valor- $p < 0,001$). Aunque los valores de AR son relativamente bajos, estos tiendieron a incrementarse progresivamente en el tiempo, mostrando un incremento del 27% en la última generación (Tabla Suplementaria S1.1). Con respecto al coeficiente f , se observaron valores significativamente mayores en el grupo TE (1,8% vs. 0,7%; valor- $p < 0,01$), evidenciando un mayor grado de parentesco entre los ejemplares de este grupo.

Finalmente, con el objetivo de facilitar el cómputo, el tamaño efectivo de la población se estimó únicamente con los individuos nacidos en la última generación (2014-2021); donde el grupo PR mostró valores acordes con el origen abierto de esta raza ($N_{EC} = 68,3 \pm 1,5$; $N_E = 233,9 \pm 34,8$). Se observó una leve diferencia entre los grupos TE ($N_{EC} = 62,9 \pm 1,3$; $N_E = 230,8 \pm 30,8$) y NO-TE ($N_{EC} = 79,6 \pm 1,9$; $N_E = 246,1 \pm 48,9$), sugiriendo la existencia de factores externos que influyen el proceso de selección. La evolución del N_E en las distintas generaciones puede observarse en la Tabla Suplementaria S1.1.

3.3. Análisis de fundadores y ancestros

El número de fundadores en la PR fue de 7.114 individuos, un valor elevado esperable para una raza de pedigrí abierto, en la que no hay limitaciones genealógicas para registrar un nuevo ejemplar (Tabla 1.2). Sin embargo, F_E fue de sólo 371 animales. Por otro lado, se encontraron claras diferencias entre grupos, donde el F_E y F_A fueron significativamente menores en TE en comparación a NO-TE. De manera similar, la relación F_E/F_A fue mayor y el F_G menor en el grupo TE (ver Tabla 1.2). Finalmente, se observó una proporción de fundadores cada vez menor en las distintas generaciones, así como también en la relación $F_E/\text{fundadores}$, la cual tuvo una fuerte correlación negativa (-0,95, valor- $p < 0,01$) con el porcentaje de individuos nacidos por transferencia embrionaria (Tabla Suplementaria S1.1).

Tabla 1.2. Análisis de fundadores y ancestros

Parámetros	PR	TE	NO-TE
Número de animales	41.873	32.731	9.142
Número de fundadores	7.114	4.518	4.654
Proporción de fundadores ^a	0,17	0,14	0,51
Número de ancestros contribuyentes	6.339	3.767	3.868
Ancestros 50%	25	18	84
F_E	371	303	680
F_A	52	42	114
F_G	35	28	72
Tasa F_E/F_A	7,1	7,2	6,0

PR: Población de referencia nacida durante las últimas dos generaciones; TE: individuos de PR nacidos por transferencia embrionaria; NO-ET: individuos de PR no nacidos por transferencia embrionaria. ^aLa proporción de fundadores fue calculada como la relación entre el número de fundadores y el número de animales dentro de cada grupo.

En la Tabla 1.3 se muestran los diez ancestros que más contribuyen a la variabilidad genética en la PR, así como en los grupos TE y NO-TE. La mayoría de esos ancestros son caballos SPC, muchos de ellos registrados en el Studbook Argentino, siendo otros mestizos de SPC, según se pudo reconstruir su

genealogía. Resultados similares se encontraron cuando el análisis se expandió a los 50 ancestros más contribuyentes (datos no mostrados), revelando una gran proporción de sangres SPC en el origen del caballo Polo Argentino. Cabe destacar que los mismos individuos fueron los ancestros más representativos en PR y TE, aunque mostraron un mayor porcentaje de contribución marginal en este último grupo (Tabla 1.3). Por otro lado, el grupo NO-TE mostró una menor influencia de los ancestros más representativos y una amplia diversidad genética.

3.4. Estructura poblacional

El promedio del índice F_{ST} (entre criadores) mostró valores más bajos en TE que en NO-TE (0,016 vs 0,026, respectivamente; valor- $p < 0,001$). El coeficiente de correlación de Spearman entre los índices F_{ST} fue muy bajo ($\rho = 0,26$; valor- $p < 0,01$). Estos datos revelan un incremento en la distancia genética entre criadores que no emplean la transferencia embrionaria, en comparación a los que sí lo hacen. El tamaño efectivo de la población fue levemente mayor en el grupo NO-TE respecto de TE, al igual que la relación N_{EC}/N_E (0,32 vs 0,27, respectivamente), aunque estos valores no indican subestructuración dentro de los grupos.

4. Discusión

La caracterización genética utilizando datos genealógicos juega un rol muy importante en la monitorización del desarrollo y de la posible pérdida de variabilidad en las poblaciones. Este análisis es particularmente valioso en razas donde las estrategias de cría y selección usualmente se basan únicamente en caracteres fenotípicos (Sánchez-Guerrero y col., 2017), pero también cuando el uso de técnicas de reproducción asistida a gran escala aumenta la intensidad de

Tabla 1.3. Ancestros más contribuyentes de cada grupo

Ancestro	Padre	Sexo	Año de nacimiento	Composición racial	Contribución genética marginal		
					PR	TE	NO-TE
SPORTIVO	Lode	M	1992	SPC	7,6% (1 ^{ro})	9,0% (1 ^{ro})	2,7% (4 ^{to})
PUCARA	Sequito	M	1981	SPC	6,9% (2 ^{do})	7,7% (2 ^{do})	4,0% (2 ^{do})
TOP SECRET	Ganelon	M	1972	SPC	4,3% (3 ^{ro})	4,6% (3 ^{ro})	
GAUCHA ACHA	Apricot	H	1973	SPC	3,4% (4 ^{to})	3,6% (5 ^{to})	
RAINBOW CORNER	Rainbow Quest	M	1993	SPC	3,3% (5 ^{to})	3,7% (4 ^{to})	2,1% (6 ^{to})
SIMPATICA	Vengo Tarde	H	1986	50-100% SPC	2,9% (6 ^{to})	3,4% (6 ^{to})	1,3% (10 ^{mo})
GUINDA	-	H	≈1980	-	2,4% (7 ^{mo})	2,8% (7 ^{mo})	
POLO NEVADITO^a	Polo Puripayo	M	1993	75% SPC	2,4% (8 ^{vo})	1,8% (9 ^{no})	4,3% (1 ^{ro})
CHUSMA	Worthy Too	H	-	SPC	1,6% (9 ^{no})	1,8% (10 ^{mo})	
LAMBADA	Sequito	H	≈1996	75-100% SPC	1,6% (10 ^{mo})	1,9% (8 ^{vo})	
EL SOL	Top Secret	M	1987	SPC			3,7% (3 ^{ro})
POLO PUREZA^a	High Host	H	1988	SPC			2,0% (5 ^{to})
PERUGINO	Southern Halo	M	1995	SPC			1,8% (7 ^{mo})
INDULTO	Liloy	M	1978	SPC			1,7% (8 ^{vo})
LA LUNA	Top Secret	H	≈1980	SPC			1,7% (9 ^{no})

^a No fundador en el libro genealógico del Polo Argentino. (-) Desconocido; M: macho; H, hembra; PR: Población de referencia nacida durante las últimas dos generaciones; TE: individuos de PR nacidos por transferencia embrionaria; NO-ET: individuos de PR no nacidos por transferencia embrionaria. En la contribución genética marginal se indica entre paréntesis el orden de mayor contribución dentro de cada grupo.

selección (Marinho y col., 2015). Ambas condiciones están presentes en el caballo Polo Argentino, el cual está fuertemente seleccionado en base a su rendimiento deportivo, y que a su vez basa su producción en la utilización de un número reducido de animales reproductores sometidos a grandes programas de transferencia embrionaria (Losinno y col., 2001).

En consecuencia, este estudio proporciona un enfoque interesante para determinar los efectos genéticos producidos por el uso de diferentes técnicas reproductivas asistidas como método de cría en una población comercial de caballos. En los equinos, la decisión de producir un cruzamiento en particular entre un semental y una yegua con el fin de obtener una característica dada en la progenie se toma muchas veces sin seguir ningún esquema basado en el uso de valores de cría (EBV), ya que estos no están disponibles en la mayoría de las razas (Gómez y col., 2010). Sin embargo, los criadores están acostumbrados a desarrollar sus propios esquemas de cruzamiento seleccionando los linajes maternos o paternos de su interés junto a un análisis subjetivo (no estandarizado) del rendimiento de la descendencia.

En las razas pequeñas, esto podría llevar a un uso desproporcionado de un número reducido de reproductores, con la consiguiente aparición de cuellos de botella genéticos y la pérdida de alelos o variantes valiosas en pocas generaciones (Janova y col., 2013). Según nuestros resultados, esto podría haber ocurrido de forma lenta pero constante en el caballo Polo Argentino, en particular durante los últimos años, donde se redujo el número de sementales empleados de manera significativa, pero más aún en el grupo de individuos criados por TE. Aunque los parámetros analizados muestran que la variabilidad genética sigue siendo elevada, probablemente porque es una raza relativamente nueva que además permanece con una política de registro abierto, nuestros resultados

sugieren la existencia de una tendencia hacia la pérdida de diversidad. Además, el aumento del parentesco entre los individuos, sobre todo durante las últimas dos generaciones, debería tenerse en cuenta en el mediano a largo plazo.

A pesar de su reciente creación, el censo del caballo Polo Argentino es uno de los más grandes entre los caballos de raza criados en Argentina, con alrededor de 15.000 individuos registrados en la última generación (Figura 1.1). Sin embargo, hay que considerar que el reducido número de animales registrados en los años 2019 a 2021 puede deberse a que el proceso de registro aún se encontraba abierto para los animales nacidos en esas fechas al momento de realizar este análisis.

El marcado incremento de la cantidad de registros de animales observado desde el año 2004 coincide con la irrupción de los programas de TE a gran escala, los cuales representan actualmente más del 70% de los productos nacidos registrados. Como ya ha sido descrito en otras especies, el uso de esta técnica de reproducción asistida podría producir un incremento en la ganancia genética en una población dada, al favorecer la distribución de una genética mejoradora pero también a través de una reducción del intervalo generacional, particularmente de los linajes maternos (Gengler y Druet, 2001). A pesar de que el progreso genético no ha podido ser apropiadamente determinado aun en el caballo Polo Argentino, debido a la ausencia de una caracterización fenotípica objetiva ni de estimaciones de parámetros genéticos, existe un consenso entre los criadores y jugadores de polo respecto de la evolución y la mejora en la performance deportiva que ha atravesado la raza, lo cual no puede ser ignorado. Prueba de esto es la presencia del Polo Argentino en los torneos más reconocidos y de mayor nivel alrededor del mundo. Por otro lado, esta política de cría podría causar una pérdida en la variabilidad genética lo que se observa por la disminución constante del porcentaje

de individuos, en particular las yeguas, empleados como reproductores. Este efecto ha sido también descrito por Sitzenstock y col. (2013) en un estudio de simulación de una población equina alemana, mostrando que la implementación de grandes programas de TE y el uso de unas pocas yeguas seleccionadas como reproductoras provocaría una gran reducción de la varianza genética y del tamaño efectivo poblacional. En este sentido, nuestro estudio es uno de los primeros en validar esta situación, utilizando un gran número de datos de registro de una raza equina de deporte, en la cual, además, más de la mitad de su población es producida por TE.

El intervalo generacional, el cual afecta el nivel de progreso genético por año en una población, fue menor que los reportados previamente en varias razas equinas como el SPC (IG = 10,6, Moureaux y col. (1996), IG = 10,62, Bokor y col. (2013)), Hannoveriano (IG entre 8,9 y 11,1, Hamann y Distl (2008)), Cuarto de Milla (IG = 10,5, Petersen y col. (2014)), Caballo de Deporte Español (IG = 10,8, Bartolomé y col. (2011)), Caballo de Deporte Francés (IG = 12, Dubois y Ricard (2007)), Silla Francés (IG = 11,7, Moureaux y col. (1996)) y Pura Raza Español (IG = 10,1, Valera y col. (2005)). Es bien sabido que las yeguas que participan en actividades deportivas o recreativas usualmente no son utilizadas como reproductoras de manera natural, disminuyendo así su vida reproductiva efectiva e incrementando el IG (Valera y col., 2005). Esta situación podría ser exacerbada por la larga vida deportiva que tiene el caballo Polo Argentino comparado con otras razas deportivas como el SPC. Basta considerar que sólo el proceso de entrenamiento de un caballo de polo lleva al menos cinco años, a los que incluso se suman un par de años más de adaptación antes que pueda empezar a jugar en torneos de alto hándicap. Por lo tanto, las yeguas madres pueden ser sólo destinadas a la reproducción por métodos naturales a edades avanzadas, lo cual se asocia con

baja fertilidad, bajas tasas de preñez y menor producción de potros por individuo (Marinone y col., 2015). Nuestros resultados demuestran que esta limitación puede ser superada por la TE ya que la gestación es llevada a cabo en una yegua receptora, liberando al animal de alto valor genético para su uso deportivo. Este hecho es particularmente importante en la raza, ya que la temporada de torneos de los animales de élite se limita generalmente a sólo seis meses por año, permitiendo el uso de las yeguas madres como donantes de embriones o de los machos como sementales durante el resto del año. Adicionalmente, las yeguas jugadoras pueden ser usadas como donantes de embriones a edades muy tempranas (incluso antes de los 4 años, de acuerdo con los registros de la base de datos de Polo), lo cual reduce significativamente el intervalo generacional.

El efecto reproductivo de la transferencia embrionaria ya ha sido asociado con la mejora de la ganancia genética en el ganado a través de tres posibles mecanismos: (1) aumento de la intensidad de selección (Nicholas, 1996), aumento en el número de crías por madre; (2) aumento de la precisión de los EBV, al incluir los datos de la TE en los modelos genéticos (Junqueira y col., 2018); y (3) reducción del intervalo entre generaciones (Mota y col., 2013). En nuestro caso, dos de estos mecanismos se observaron en el grupo TE (IG corto y mayor número de potros por yegua), lo que apoya el consenso general entre criadores y jugadores de que el progreso en el rendimiento del Polo Argentino durante las dos últimas décadas se ha visto muy favorecido por el uso de la TE. Sin embargo, este hecho no pudo determinarse en este estudio debido a la falta de un programa de evaluación sistemática y a la ausencia de una caracterización fenotípica de la población.

Una de las posibles desventajas del uso intensivo de las técnicas de reproducción asistida es la reducción de la variabilidad genética. En el presente estudio, la

consanguinidad media en el Polo Argentino (0,86% en PR) sigue siendo inferior que en otras razas como el Hannoveriano (1,33%, Hamann y Distl (2008)), Pura Raza Española (8,48%, Valera y col. (2005)), Lipizzano (10,8%, Zechner y col. (2002)), Haflinger (6,59%, Gandini y col. (1992)) y Pura Raza Árabe (7%, Moureaux y col. (1996)). Sin embargo, F_{PED} fue ligeramente superior a la reportada en el Caballo de Deporte Español (Bartolomé y col., 2011) y Brasileño (Medeiros y col., 2014) ($F_{PED} \approx 0,6\%$), los cuales también tienen un libro genealógico de reciente creación y sostienen una política de inscripción abierta. En el caso del Polo Argentino, se observó una tendencia al aumento de la endogamia promedio, sobre todo durante los últimos 15 años, lo cual coincide con la aparición comercial a gran escala de los programas de TE en la raza. Tal es así que la endogamia fue más alta en el grupo de TE durante la mayoría del período analizado, como puede verse en la Figura 1.3. A pesar de que la juventud de la raza (solo 40 años de historia) y su actual política de registro abierto puede explicar en parte los bajos valores medios de F_{PED} encontrados en este estudio, se observó un aumento del AR y de la endogamia, lo cual sugiere cierta cautela. Más aún, este efecto hubiese sido mayor si la población se hubiera gestionado como una raza cerrada, ya que no se hubiesen podido utilizar sementales externos a la raza que limitaran el incremento de ambos efectos genéticos. Sin embargo, el hecho de que en los últimos años se utilizan cada vez más como reproductores a los ejemplares ya registrados en lugar de nuevos individuos "fundadores" (datos no mostrados), podría derivar en la reducción a largo plazo de la diversidad genética.

Tanto F_{PED} como AR se incrementaron abruptamente (F_{PED} más del doble en animales endogámicos) durante las últimas dos generaciones. Esta tendencia fue más notoria en el grupo TE, ya que esta tecnología permite a los criadores utilizar intensamente unos pocos reproductores relevantes y destacados para la cría, para

obtener productos de mejor rendimiento y más homogéneos, de acuerdo con sus intereses comerciales. En un estudio reciente, Todd y col. (2018) encontraron una fuerte correlación negativa entre la endogamia y la performance de carrera en caballos SPC australianos después de analizar más de 150.000 resultados de carrera. De manera similar, Gómez y col. (2009) determinaron la presencia de depresión endogámica en varios caracteres morfológicos en caballos Andaluces, por lo que recomiendan la inclusión del F_{PED} en los modelos de evaluación genética, con el objetivo de mejorar la precisión de los EBV. Si bien no existe un valor dado a partir del cual los efectos de la depresión endogámica se hacen más notorios, y a pesar de que el promedio del F_{PED} es aún bajo en el caballo Polo Argentino, la tendencia ascendente observada durante los últimos años, en particular en los animales producidos por TE, debería ser considerada por los criadores al decidir los futuros cruzamientos, para evitar la aparición de depresión endogámica en los ejemplares.

El efecto de la TE también se ve reflejado en el promedio de la coascendencia de cada grupo (significativamente mayor en TE), posiblemente debido a una mayor intensidad de selección sobre las madres. Debido a que las yeguas producen un mayor número de potrillos durante su vida reproductiva, su representatividad en la población se incrementa, aumentando así los valores de AR y limitando aún más la diversidad genética. Para controlar esta situación, la Asociación de Criadores de Caballos Pura Raza Española ha desarrollado e implementado exitosamente enfoques sistemáticos para supervisar el apareamiento de los reproductores utilizando un software ad-hoc para poder determinar las relaciones entre todos los reproductores disponibles en una cabaña (Melgarejo y col., 2000). Usando este software, se logró disminuir sustancialmente la tasa de endogamia en un periodo de tiempo relativamente corto (Valera y col., 2005). Sin embargo,

este tipo de métodos no está disponible de manera sistemática en los caballos Polo Argentino. Dado que el uso de las técnicas de TE aumenta las posibilidades de apareamiento entre sementales y yeguas que se encuentran en diferentes rebaños y ubicaciones, la aplicación de este tipo de tecnologías podría ayudar a evitar el cruzamiento de animales muy emparentados.

El tamaño efectivo de la población es un indicador importante de la evolución de la variabilidad genética de una población. En PR, el N_{EC} fue menor en comparación a reportes publicados de otras razas compuestas, como el Caballo de Deporte Brasileño ($N_{EC}=188,5$; Medeiros y col. (2014)), el Anglo-árabe ($N_{EC}=291,2$; Cervantes y col. (2011)) y el Caballo de Deporte Español ($N_{EC}=1.046$; Bartolomé y col. (2011)), lo que sugiere una menor variabilidad probablemente debida a la estrategia de selección empleada. Cervantes y col. (2011) describieron que en una población no estructurada se espera que N_{EC} y N_E presenten valores muy similares entre sí, ya que el incremento de la endogamia (ΔF) tiende a ser mayor en una población subdividida, mientras que el aumento de la coascendencia (Δf) prácticamente no se ve afectado. En nuestro estudio, la relación N_{EC}/N_E en NO-TE fue apenas superior que en TE; de hecho, esta relación fue menor a 1, lo cual indica ausencia de estructuración, posiblemente debido al alto flujo génico entre ambos grupos. Sin embargo, los valores de F_{ST} fueron mayores entre los individuos del grupo NO-TE pertenecientes a distintos criadores. Esto podría explicarse porque los individuos criados por gestación natural están más emparentados dentro de los grupos, probablemente debido a la menor conexión entre cabañas, ya que las yeguas de cría normalmente permanecen con el mismo criador durante toda su vida reproductiva. Por consiguiente, el análisis sugiere que la TE permite un mayor intercambio de germoplasma seleccionado entre criadores, sobre todo a través de las líneas maternas, en comparación con NO-

TE, en la que se observa una mayor homogeneidad dentro de los grupos, más que entre ellos.

El F_E es otro parámetro importante para determinar si la contribución de los individuos fundadores está balanceada. Se espera que el número total de fundadores sea igual al F_E cuando todos contribuyen equitativamente (Lacy, 1989). Sin embargo, esto no es lo que suele ocurrir en las poblaciones domésticas de ganado, donde el F_E es generalmente más bajo, indicando una pérdida de diversidad genética proporcional a cuánto difieran esos valores. Este hecho ha sido reportado en varias razas equinas como el Caballo de Deporte Español (Bartolomé y col., 2011) y el Pura Raza Español (Valera y col., 2005). En nuestro estudio, se observa una clara diferencia en la proporción de fundadores entre los grupos TE y NO-TE. Además, F_E y F_A fueron sustancialmente más bajos que el número total de fundadores y ancestros en PR, mientras que la relación F_E/F_A sugiere la presencia de un cuello de botella. Esta tendencia se ve de forma más notoria en el grupo TE, en comparación con NO-TE, lo cual podría explicarse por el uso de un reducido número de sementales y yeguas madres, causando una gran presión de selección. Este hecho también explicaría las diferencias observadas en F_G entre ambos grupos. Dado que el criterio de agrupación empleado en los 40.000 caballos analizados en este trabajo incluyó únicamente el uso de técnicas de TE, podemos demostrar que la implementación de grandes programas de TE tienen efecto sobre la reducción de la variabilidad genética, aumentando la homogeneidad dentro de la población.

Finalmente, es importante destacar que los ancestros que más han contribuido al pool genético de la raza Polo Argentino fueron SPC. A pesar de que está permitida la incorporación de individuos de cualquier origen, la información actual confirma que en sus inicios el Polo Argentino estuvo altamente influenciado por esta raza.

Sin embargo, nuestro análisis también mostró que dicha influencia fue producida por un escaso número de caballos que fueron empleados ampliamente como reproductores durante un largo período de tiempo, más que por un uso inicial de SPC como reproductores. Esto podría haber afectado también a la variabilidad genética, como ya se ha descrito en el Caballo de Deporte Brasileño (Medeiros y col., 2014), el Asturcón (Álvarez y col., 2011) y tres razas de tiro checas (Vostrá-Vydrová y col., 2016). Más aun teniendo en cuenta que el SPC es una raza muy antigua con un registro cerrado, creada a partir de unos pocos sementales (Binns y col., 2012) con una limitada variabilidad genética (Cunningham y col., 2001; Bokor y col., 2013). Dado que esta raza está involucrada en la formación del Polo Argentino, es posible que su escasa diversidad podría ser arrastrada hacia la raza, afectando aún más su estatus genético. Sin embargo, a pesar de que los registros genealógicos fueron parcialmente reconstruidos con datos del Studbook Argentino para realizar este estudio, las limitaciones de los datos de pedigrí impiden evaluar esta hipótesis con precisión. Es por esto por lo que se requerirán estudios moleculares para una mejor aproximación a la variabilidad genética de la raza, fundamentalmente para considerar la presencia de endogamia ancestral.

5. Material Suplementario

Tabla S1.1. Evolución de los distintos parámetros genéticos analizados por generaciones.

Nacimiento	n	%TE	ECG	F _{PED} (%)	AR (%)	fun	fun/n	F _E /fun	F _E	F _A	F _E /F _A	N _E
Antes de 1990	141	0%	3,8 ± 0,69	0,49 ± 2,76	0,37 ± 0,36	474	3,36	0,28	134	16	8,4	485,8 ± 124,1
1990-1997	933	35%	4,08 ± 0,73	0,18 ± 1,25	0,64 ± 0,57	994	1,1	0,36	356	42	8,5	1406,2 ± 282,6
1998-2005	5319	66%	4,3 ± 1,13	0,4 ± 2,2	0,87 ± 0,53	2350	0,44	0,20	478	71	6,7	602,3 ± 146,2
2006-2013	28800	78%	4,8 ± 1,4	0,7 ± 2,3	1,09 ± 0,62	6231	0,22	0,07	416	62	6,7	363,4 ± 66,2
2014-2021	13077	79%	5,6 ± 1,51	1,2 ± 2,5	1,38 ± 0,63	3440	0,26	0,08	267	34	7,8	233,9 ± 34,8
Correlación con %TE				0,85	0,97**	0,84	-0,76	-0,95**	0,3	0,34	-0,55	-0,66

n: número de animales; %TE: porcentaje de individuos nacidos por transferencia embrionaria; ECG: generaciones completas equivalentes; fun: fundadores. ** valor-p<0,01.

Tabla S1.2. Promedio del coeficiente de endogamia en individuos endogámicos ($F_{PED} > 0$) y proporción de individuos endogámicos por grupo y por categoría de F_{PED} .

	PR	TE	NO-TE
n total	41.873	32.731	9.142
$F_{PED} > 0$			
n individuos	19.262	16.605	2.657
Proporción de individuos endogámicos	46%	50%**	29%**
F_{PED} media \pm DE	1,9% \pm 3%	1,8% \pm 3%**	2,2% \pm 4%**
$F_{PED} > 0$ a 0,06249			
N individuos	16.888	14.664	2.224
Proporción de individuos endogámicos	40%	45%**	24%**
F_{PED} media \pm DE	0,86% \pm 1%	0,89% \pm 1%**	0,73% \pm 1%**
F_{PED} 0,0625 a 0,1249			
N individuos	1.817	1.532	285
Proporción de individuos endogámicos	4%	4,7%**	3,1%**
F_{PED} media \pm DE	7,1% \pm 1%	7,1% \pm 1%	7,2% \pm 1%
$F_{PED} \geq 0,125$			
N individuos	557	381	176
Proporción de individuos endogámicos	1,3%	1,2%**	1,9%**
F_{PED} media \pm DE	15% \pm 4%	15% \pm 3%	15,2% \pm 5%

DE: desvío estándar. La proporción de individuos endogámicos fue calculada como la cantidad de individuos en cada categoría de F_{PED} (n individuos) sobre el número total de individuos (n total).

CAPÍTULO II: Análisis de la variabilidad genética en el Caballo Polo Argentino mediante datos genómicos: determinación del número y distribución de bloques de homocigosidad (ROH).

1. Introducción

El estudio de la estructura genética es un aspecto muy importante en la genética de poblaciones ya que permite comprender y evaluar la distribución de la diversidad dentro de las mismas. El monitoreo y mantenimiento de la diversidad genética es fundamental para alcanzar objetivos de mejora genética (Machmoum y col., 2020).

La endogamia es un fenómeno recurrente en la producción animal. Si bien muchas veces se recurre a ella como herramienta para lograr uniformidad en la población, uno de sus efectos negativos es el aumento de la frecuencia de alelos recesivos en estado homocigoto (Curik y col., 2014). Este incremento ha sido asociado a una disminución significativa en el rendimiento, la productividad y la fertilidad, así como al incremento en la prevalencia de enfermedades de origen genético, fenómeno conocido como depresión endogámica (Charlesworth y Willis, 2009). Si bien el incremento de la endogamia es un fenómeno común en los equinos, tiende a ser más preocupante en razas de censo reducido (por ejemplo, razas autóctonas) y/o con un libro genealógico cerrado (es decir, que no se permite el cruzamiento de individuos que no estén registrados). Sin embargo, y a pesar de que el Polo Argentino mantiene un libro genealógico abierto, según los resultados obtenidos en el Capítulo I, el amplio uso de las biotecnologías reproductivas y la alta presión de selección ha producido un aumento de la endogamia promedio,

reduciendo la variabilidad genética por debajo de la esperada para una población de este tipo. Adicionalmente, hay que destacar que la mayor influencia externa que posee el Caballo de Polo es la raza SPC, la cual tiene un registro genealógico cerrado que data del siglo XVIII (Bower y col., 2012).

El uso de datos moleculares para el cálculo de la variabilidad genética y la estimación de la endogamia es una valiosa herramienta que permite establecer el estado de situación de una población, sorteando los inconvenientes que pudiera plantear el uso de datos genealógicos (Hedrick y Garcia-Dorado, 2016). La disponibilidad de plataformas de genotipado con alta densidad de marcadores ha permitido desarrollar diferentes metodologías para estudiar la diversidad genética (McCue y col., 2012). En los últimos años, la aparición en el mercado de microarrays de SNP para diversas especies, incluida la equina, ha hecho que este tipo de estudios sea cada vez más frecuente. Actualmente existen decenas de publicaciones que analizan la variabilidad genética, la estructuración y el origen filogenético de numerosas razas equinas a partir de genotipos de miles de individuos (Todd y col., 2018; Schurink y col., 2019; Cosgrove y col., 2020; Poyato-Bonilla y col., 2021). Si bien el análisis de datos de pedigrí ha sido altamente informativo en el estudio de este tipo de poblaciones, la utilización de datos genómicos ha producido, sin lugar a duda, un salto de calidad en este tipos de estudios (Kardos y col., 2015).

Uno de los enfoques actualmente más utilizados para el estudio de la endogamia a nivel molecular es el análisis de la distribución y prevalencia en el genoma de bloques de homocigosidad (ROH – del inglés “runs of homozygosity”). Los ROH son segmentos continuos del genoma en los que un individuo es homocigoto en todos los loci (McQuillan y col., 2008; Curik y col., 2014). El origen de estas regiones puede estar asociado a la presencia de pequeñas inversiones que

suprimen la recombinación o eventos demográficos, como deriva génica o cuellos de botella (Gorssen y col., 2021). Sin embargo, la causa primaria es el cruzamiento entre individuos emparentados (Ferenčaković y col., 2013).

Un individuo que es fruto de un cruce consanguíneo tiene la posibilidad de heredar las dos variantes alélicas, para un *locus* dado, de un mismo ancestro en común. En conjunto, estos loci formarán haplotipos en estado homocigota (ROH), que además tenderán a ser más largos cuánto más reciente sea el cruzamiento endogámico (Howrigan y col., 2011). De acuerdo con la teoría propuesta por Fisher (1954), durante la formación de los gametos, el ADN se ve sometido a sucesivas meiosis y procesos de recombinación que tenderían a “romper” estos haplotipos, fragmentándolos en ROH más pequeños a través de las generaciones. Por lo tanto, se asume que la longitud de los ROH está asociada inversamente con la lejanía en que se haya producido el evento endogámico: los ROH más largos derivan de un ancestro en común reciente (endogamia reciente), mientras que los más cortos de uno más lejano (endogamia ancestral)(Curik y col., 2014; Poyato-Bonilla y col., 2021). Por tal motivo, el estudio de los patrones de distribución de estas regiones y sus características en una población permite inferir su historia demográfica (Ceballos y col., 2018). Pero también, el estudio de ROH asociados a caracteres o genes de interés permiten entender los procesos selectivos a los que ha sido sometida la población, más aun teniendo en cuenta que pueden considerarse reservorios de alelos potencialmente deletéreos (Peripolli y col., 2016).

El análisis de ROH permite también determinar el coeficiente de endogamia (F_{ROH}) de un individuo, como la proporción del genoma que se encuentra cubierta por dichos fragmentos (McQuillan y col., 2008). Pero también permite calcular un coeficiente F_{ROH} para cada cromosoma en particular, evidenciando su distribución

y tamaño a nivel cromosómico. Además, la depresión endogámica no solamente está vinculada al porcentaje del genoma que se encuentra en homocigosis, sino también a su distribución (Laseca y col., 2022). Esto puede explicar por qué individuos con el mismo coeficiente de consanguinidad tienen distinto grado de depresión endogámica (Perdomo-González y col., 2021). La identificación de regiones en homocigosis posibilita localizar los loci que están afectados, y qué genes o QTL se encuentran en ellas, lo cual permite además inferir parcialmente las bases genéticas que gobiernan un carácter dado.

Los ROH pueden distribuirse a través de todo el genoma, aunque son más comunes de encontrar en regiones con alto LD y baja tasa de recombinación (Ceballos y col., 2018). Sin embargo, existe una mayor abundancia de ROH en regiones específicas del genoma que están bajo selección positiva (Gorssen y col., 2021). Esto hace que la distribución real de los ROH de los distintos individuos en una población sea desigual, evidenciando segmentos cromosómicos con alta prevalencia (hotspot), denominados islas ROH (ROHi). En contraposición, las zonas con relativamente baja frecuencia de ROH (coldspot) en general se asocian con genes con funciones críticas, en las que la presencia de alelos deletéreos en homocigosis no sería viable para la especie (Curik y col., 2014).

La identificación de islas ROH es una de las metodologías utilizadas para detectar huellas de selección en estudios poblacionales (Saravanan y col., 2020), ya que dan cuenta de patrones genéticos característicos observados en regiones del genoma que están bajo selección intensa. Estos procesos de selección tienden a ocurrir en zonas específicas del genoma que controlan diferentes caracteres, como la conformación, reproducción, rendimiento, comportamiento, susceptibilidad a enfermedades, entre otros. Si bien se dan naturalmente con la evolución de las especies, en animales domésticos suelen estar asociados a la

selección artificial producida por los seres humanos durante su proceso de domesticación y mejoramiento (Nielsen, 2005). Por tal motivo, la identificación de huellas de selección en una población es importante para localizar genes relacionados a los caracteres bajo selección. Cuando se selecciona una variante genética determinada, se produce una reducción de la variabilidad en las regiones contiguas, por acción del desequilibrio de ligamiento y la tendencia a la fijación de la variante de interés a través de las generaciones. Este proceso de reducción de la variabilidad en zonas con efecto neutral adyacentes al locus “beneficioso” que está bajo selección, se lo conoce como barrido selectivo (en inglés, “hitchhiking”) (Saravanan y col., 2020).

Existen varios enfoques metodológicos para la detección de huellas de selección, cada uno de los cuales presenta algunas ventajas y desventajas (Saravanan y col., 2020). Entre ellos, la determinación de ROHi en el genoma es uno de los métodos basados en la reducción de la variabilidad genética local, que permite identificar estas huellas intra-raza (Grilz-Seger y col., 2019; Colpitts y col., 2022; Laseca y col., 2022). En la especie equina, se han publicado varios artículos que evalúan la presencia de huellas de selección bajo distintas metodologías y en distintas razas, como en caballos SPC (Moon y col., 2015), Pura Raza Español (Laseca y col., 2022), Warmblood de Suecia (Ablondi y col., 2019), Cuarto de Milla, razas de paso (gaited – breeds) y miniatura (Petersen y col., 2013b).

Por todo lo expuesto, el objetivo general de este capítulo fue evaluar la variabilidad genética del caballo Polo Argentino a nivel molecular, mediante la determinación de la heterocigosidad existente en la población y la caracterización de ROH a lo largo del genoma.

Los objetivos particulares de este capítulo fueron:

- Seleccionar un grupo de animales específicos en base al análisis genealógico que representen la variabilidad existente de la raza, para su caracterización utilizando técnicas moleculares de genotipado masivo.
- Establecer los niveles de endogamia de la población estudiada mediante estudios moleculares y compararlos con los obtenidos mediante los registros de pedigrí, determinado de manera real la variabilidad genética poblacional e individual, así como su variación a lo largo del tiempo.
- Determinar las posibles implicancias genéticas de la endogamia mediante el análisis de las regiones de homocigosidad (ROH) y su efecto genómico-funcional utilizando técnicas bioinformáticas de ontología génica.

2. Materiales y métodos

2.1. Material animal

Se seleccionaron 385 ejemplares Polo Argentino (65 machos y 320 hembras), provenientes de diferentes criadores (n= 61) a los que se les extrajo muestras de sangre o pelo, para la posterior extracción de ADN. Las muestras fueron tomadas en haras, centros de transferencia embrionaria o centros de entrenamiento localizados en distintas provincias del país. El criterio de selección de los ejemplares fue que estuvieran inscriptos en el libro genealógico de la raza, registrándose su nombre y SBA. Además, se procuró que el pool de animales muestreados represente las líneas genéticas más importantes del Polo Argentino, de acuerdo con los resultados obtenidos del análisis poblacional con datos genealógicos realizados en el Capítulo I de esta tesis. De esta manera, los animales elegidos fueron descendientes de varias de las principales líneas de ancestros más contribuyentes en la población, procurando así que la selección representara la variabilidad de la raza.

2.2. Obtención de muestras y extracción de ADN

Las muestras de pelo se obtuvieron por arrancamiento de al menos 30 pelos con bulbo de la cola, y se almacenaron en sobres debidamente rotulados. Las muestras de sangre se obtuvieron por punción de la vena yugular con aguja 21G acoplada a una jeringa de 10 ml, e inmediatamente preservada en un tubo rotulado y con anticoagulante EDTA. Ambos tipos de muestras se almacenaron en freezer a -80°C hasta su posterior procesamiento.

La extracción de ADN se realizó en el laboratorio del Servicio de Genética de Animales Domésticos del IGEVET con kits comerciales, siguiendo las recomendaciones del fabricante para cada tipo de tejido (sangre o pelo). Se utilizaron dos kits diferentes, según disponibilidad en el laboratorio: DNeasy Blood & Tissue (Qiagen, Valencia, CA) y HigherPurity™ Blood & Cell Culture DNA Extraction Mini Spin Kit (Canvax Biotech, Cordoba, Spain). La calidad del ADN extraído se comprobó mediante parámetros estándar como concentración (ng/dl) y pureza (absorbancias A260/280 y A260/230) utilizando un espectrofotómetro (Thermo Scientific™ Multiskan™ GO).

2.3. Genotipificación con microarray de SNP

Aquellas muestras con una calidad óptima (al menos 20 ng/ul de ADN, A260/280 > 1,5 y A260/230 > 0,5) fueron genotipadas utilizando el microarray Illumina Equine GGP array (Illumina, San Diego, USA) en las instalaciones de NEOGEN® (Ayr, Escocia). Este microarray, que cuenta con 71.590 SNPs distribuidos uniformemente a lo largo del genoma (Wu y col., 2016), fue creado a partir de los datos genómicos de numerosas razas equinas, entre las que se encuentra el SPC. Cuenta también con varios marcadores que han sido asociados a caracteres de interés en la especie equina, incluyendo rasgos comportamentales, morfológicos,

funcionales y asociados a diversas enfermedades de origen genético (Neogen, 2023).

2.3.1. Control de calidad de los genotipos

Los datos crudos de genotipado fueron recibidos en formato *final.report*. Dichos archivos fueron transformados al formato *PLINK long-format genotype file (lgen)* utilizando scripts propios en lenguaje R (R-Core-Team, 2023). Finalmente, los archivos *lgen* fueron transformados a los formatos binario (*bim*, *bed*, *fam*) y *ped-map* (dependiendo el análisis a realizar) mediante el paquete *PLINK* (Purcell y col., 2007).

Para este trabajo se utilizaron los SNPs localizados en todos los autosomas (ECA 1 a ECA 31) y en el ECA X, eliminándose aquellos marcadores mitocondriales, del cromosoma Y o no asignados a ningún cromosoma. Con el objetivo de garantizar una adecuada calidad de los genotipos se realizó el filtrado de los datos por “*call rate*” por SNP y por individuo, utilizando el programa *PLINK*. Sólo los SNPs que se genotiparon en al menos el 90% de las muestras fueron incluidos en los sucesivos análisis (función “*--geno 0,1*”). Por otro lado, sólo las muestras que tenían más del 95% de los SNPs genotipados fueron conservadas (función “*--mind 0,05*”). Como resultado, se mantuvieron 383 ejemplares Polo Argentino con 69.746 SNPs genotipados.

2.4. Análisis genómicos

2.4.1. Heterocigosidad y equilibrio Hardy Weinberg

Uno de los primeros cálculos realizados fue el de la heterocigosidad en la población, la cual es una medida de variabilidad que indica la proporción de

individuos que son heterocigotas para los loci analizados. Este parámetro puede descomponerse en heterocigosidad observada (H_o) y heterocigosidad esperada (H_e). La primera de ellas se calcula simplemente con las frecuencias observadas en la población, es decir, constituye la heterocigosis real de la muestra. La H_e , en cambio, se calcula como la proporción de heterocigotas considerando las frecuencias genotípicas esperadas. Estas se calculan a partir de las frecuencias alélicas, como $2pq$, donde p y q representan la frecuencia del alelo A1 y A2, respectivamente.

La diferenciación entre la H_o y H_e permite comprobar si una población se encuentra en equilibrio de acuerdo con la ley de Hardy Weinberg (H-W), que establece que las frecuencias genotípicas observadas en la población están dadas por las frecuencias alélicas de la generación anterior (parental), de modo que:

- La frecuencia de los homocigotas es igual al cuadrado de la frecuencia de cada alelo (p^2 y q^2).
- La frecuencia de los heterocigotas es igual al doble del producto de las frecuencias alélicas ($2pq$).

Sin embargo, para que esta ley se cumpla deben cumplirse una serie de supuestos, como por ejemplo: panmixia (apareamiento al azar), ausencia de selección, mutación, migración ni deriva génica (Nicholas, 1987). Por lo tanto, si la población no está en equilibrio H-W, puede suponerse que alguno de esos supuestos no se cumple.

Estos parámetros fueron estimados con el programa *PLINK* (Purcell y col., 2007) mediante la función "--hardy".

2.4.2. Determinación de ROH y coeficiente de endogamia

Para el análisis de la endogamia se determinó la presencia de ROH a lo largo de todo el genoma de cada individuo, que, como se mencionó anteriormente, son segmentos continuos de ADN de largo mínimo definido que se presentan en homocigosis (Meyermans y col., 2020). El análisis se realizó utilizando el método de ventana deslizante (sliding window) implementado en el paquete de R *DetectRUNS* (Biscarini y col., 2018). Esta metodología utiliza una ventana de SNPs que se desliza escaneando el genoma, puntuando cada SNP según la proporción de veces en que aparece en una ventana homocigota. Este procedimiento permite determinar si un SNP está o no dentro de un ROH, de acuerdo con una serie de parámetros que se especifican previamente:

- Tamaño de la ventana ("*windowSize*"): se define un número de SNPs que constituirán la ventana de deslizamiento o exploración.
- Umbral ("*threshold*"): número mínimo de veces que un SNP debe estar dentro de una ventana en homocigosis para considerarlo parte de un ROH.
- Número mínimo de SNP dentro de un ROH ("*minSNP*").
- Número máximo de SNP que se tolera que sean heterocigotas dentro de un ROH ("*maxOppWindow*").
- Número máximo de genotipos perdidos tolerado ("*maxMissWindow*").
- Distancia máxima entre dos SNPs consecutivos para ser considerados un potencial ROH ("*maxGap*").
- Longitud mínima que debe tener un ROH en pares de bases ("*minLengthBps*").

En nuestro análisis, los ROH se determinaron en los cromosomas autosómicos (todos los individuos) y el cromosoma X (sólo en las hembras), utilizando los

siguientes parámetros recomendados por Meyermans y col. (2020), basados en las mejores prácticas disponibles en la actualidad:

WindowSize = 30; *minSNP* = 10; *threshold* = 0,05; *maxGap* = 10^6 ; *minLengthBps* = 1000000, *maxOppWindow* = 0, *maxMissWindow* = 1.

Posteriormente, se determinó el coeficiente de endogamia de cada individuo (F_{ROH}), a nivel genómico y cromosómico, el cual se calculó como:

$$F_{ROH} = \frac{\sum LROH}{L_{genoma}}$$

Donde $\sum LROH$ es la sumatoria del largo total de todos lo ROH, y L_{genoma} es la longitud total del genoma o del cromosoma, en caso de que se calcule el F_{ROH} por cromosoma, utilizando la función “*summary.runs*” del paquete *DetectRUNS* (Biscarini y col., 2018).

A su vez, con el objetivo de evaluar la endogamia ancestral y reciente, se determinó el F_{ROH} de los individuos a partir de ROH de determinada longitud, estimando el número de generaciones hacia atrás en las que se produjo el cruce consanguíneo de acuerdo con la teoría postulada por Fisher (1954). Según esta teoría, la longitud promedio (en centimorgan, cM) de los fragmentos homocigotas de un ancestro común sigue una distribución de $100/2g$, siendo g el número de generaciones acaecidas desde el evento consanguíneo. Asumiendo que $1cM=1Mb$, los ROH $\geq 16,6$ Mb se habrán producido en las últimas 3 generaciones (endogamia reciente; F_{ROH_3G}); los ROH $\geq 8,3$ Mb dentro de las últimas 6 generaciones (F_{ROH_6G}); y ROH $\geq 5,5$ Mb en las últimas 9 generaciones (F_{ROH_9G}). La endogamia ancestral (F_{ROH_ANC}) fue estimada como: F_{ROH} total - F_{ROH_9G} , por lo tanto, indica eventos consanguíneos que se produjeron hace más de 9 generaciones y se corresponde con ROH $< 5,5$ Mb de largo.

También se calculó para cada individuo el coeficiente de endogamia con los datos de pedigrí (F_{PED})(Wright, 1931) y el coeficiente de endogamia basado en la diferencia entre el número observado y el esperado de genotipos homocigóticos (F_{IS}). F_{PED} fue calculado con el software *ENDOG* (Gutierrez y Goyache, 2005), así como el número de generaciones equivalentes y completas, tal como se detalla en el Capítulo I. F_{IS} se determinó usando el software *PLINK* (función "--het") (Purcell y col., 2007).

Se calculó la media de todos los parámetros mencionados en la población y la correlación por rangos de Spearman entre los distintos coeficientes de endogamia calculados. El F_{ROH} fue calculado además para cada cromosoma particular. El análisis estadístico fue realizado con el software *STATISTICA* (StatSoft, 2012) así como el paquete *Base* del entorno R (R-Core-Team, 2023).

2.4.3. Identificación de Islas ROH (ROHi)

Luego de calcular los ROH presentes en cada individuo, se buscaron las regiones genómicas más enriquecidas de ROH en toda la población, como potenciales islas ROH (ROHi). Para esto, por un lado, se identificó la incidencia de ROH en la población, utilizando nuevamente el paquete de R *DetectRUNS* (Biscarini y col., 2018) y de acuerdo a la metodología empleada por otros autores (Peripolli y col., 2018; Gorssen y col., 2021). Paralelamente, se determinó el porcentaje de veces que un SNP estaba incluido dentro de un ROH (incidencia por SNP) y se realizó un test de permutaciones (100.000 iteraciones por SNP) bajo la hipótesis nula de que los ROH están distribuidos uniformemente a través del genoma, de acuerdo con la metodología descrita por Goszczynski y col. (2018). Sólo los ROH presentes en más del 60% de la población y cuyos SNPs tuvieran un valor-p promedio significativo en el test de permutaciones (valor-p<0,01) fueron

seleccionados como potenciales ROHi (lo cual indica un incremento de 100 veces respecto de la probabilidad de aparición aleatoria de ROH en esa región).

Una vez identificadas las ROHi, se realizó un análisis de ontología genética para cada una de ellas. Cada segmento que constituía una ROHi fue intersectado con el genoma equino (versión EquCab 3.0) utilizando el paquete de R *HelloRanges* (Lawrence, 2021), lo cual permitió determinar qué genes se encontraban dentro de dichos intervalos genómicos. Finalmente, se realizó una caracterización de esos genes en las plataformas PANTHER (Mi y col., 2017) y DAVID (Huang da y col., 2009), así como una exhaustiva revisión bibliográfica en las principales bases de datos online (PubMed y Scopus), con el fin encontrar la función asociada a dichos genes. La búsqueda bibliográfica se realizó utilizando palabras como “*gen*” AND “horse”. Cuando el resultado fue negativo se buscó la función de dichos genes en otras especies como el bovino o el ratón. Se seleccionaron para su discusión aquellos que tenían una función relacionada a caracteres que pudieran influir en el rendimiento deportivo, el comportamiento y la conducta, el metabolismo energético y mitocondrial, la conformación músculo esquelética y la salud de los individuos.

3. Resultados

3.1. Heterocigosis, frecuencias alélicas y Hardy Weinberg

En la Tabla 2.1 se describen la cantidad de marcadores en cada cromosoma, la media de la H_o , H_e y el valor-p del equilibrio de Hardy Weinberg de los SNPs por cromosoma en la raza Polo Argentino. La heterocigosidad observada (H_o) tuvo una media de 0,324 en las muestras analizadas. Este valor fue similar en todos los cromosomas, siendo el ECA 1 el de menor heterocigosidad (0,305), mientras

que el ECA 26 tuvo el mayor valor (0,354). En promedio, los valores-p del equilibrio de Hardy Weinberg fueron no significativos.

Tabla 2.1. Promedio del número de marcadores (nSNP), frecuencia del alelo menor (MAF), heterocigosidad observada (Ho) y esperada (He), y valor-p del equilibrio Hardy Weinberg de cada locus por cromosoma en el caballo Polo Argentino.

ECA	nSNP	MAF	Ho	He	H-W Valor-p
ECA 1	5283	0,226	0,305	0,300	0,585
ECA 2	3678	0,244	0,324	0,322	0,576
ECA 3	3420	0,237	0,315	0,313	0,591
ECA 4	3156	0,255	0,332	0,328	0,596
ECA 5	3000	0,244	0,321	0,314	0,564
ECA 6	2592	0,237	0,316	0,312	0,584
ECA 7	2829	0,250	0,330	0,322	0,558
ECA 8	2879	0,251	0,332	0,328	0,570
ECA 9	2517	0,227	0,307	0,301	0,580
ECA 10	2507	0,258	0,338	0,332	0,572
ECA 11	1891	0,239	0,318	0,313	0,601
ECA 12	900	0,241	0,314	0,314	0,602
ECA 13	1214	0,237	0,320	0,310	0,560
ECA 14	2865	0,245	0,318	0,318	0,591
ECA 15	2751	0,253	0,332	0,328	0,587
ECA 16	2631	0,238	0,314	0,314	0,599
ECA 17	2398	0,260	0,341	0,337	0,564
ECA 18	2341	0,253	0,336	0,329	0,551
ECA 19	1827	0,254	0,328	0,327	0,628
ECA 20	1908	0,249	0,332	0,327	0,571
ECA 21	1763	0,243	0,321	0,317	0,573
ECA 22	1480	0,226	0,308	0,304	0,580
ECA 23	1614	0,235	0,322	0,315	0,600
ECA 24	1486	0,242	0,321	0,320	0,574
ECA 25	1148	0,251	0,323	0,324	0,598
ECA 26	1109	0,275	0,354	0,350	0,583
ECA 27	1184	0,252	0,336	0,327	0,569
ECA 28	1342	0,239	0,321	0,316	0,603
ECA 29	908	0,255	0,329	0,327	0,566
ECA 30	899	0,239	0,317	0,315	0,609
ECA 31	699	0,258	0,333	0,333	0,572
ECA X	3527	0,239	0,308	0,314	0,546
Media ± DE	2179	0,245 ± 0,158	0,324 ± 0,176	0,319 ± 0,171	0,581 ± 0,34

DE: Desvío estándar

3.2. Determinación de ROH y coeficientes de endogamia

En la Tabla 2.2. se describe la cantidad de ROH detectados en la población, así como su longitud promedio. En total se detectaron 42.727 ROH promediando 111 ROH por individuo. El promedio de largo de ROH fue de 3,5 Mb con un rango entre 1 Mb y 76,5 Mb, siendo la gran mayoría de una longitud menor a 5,5 Mb. Sin embargo, casi el 70% de los individuos presentaron al menos un ROH $\geq 16,6$ Mb, lo cual supone eventos endogámicos recientes. El análisis de la distribución a lo largo del genoma demostró que casi todos los cromosomas tuvieron ROH de todas las longitudes analizadas, excepto los ECA 12, 13, 24 y 31 que no presentaron ROH $\geq 16,6$ Mb.

En términos de endogamia, el promedio de F_{PED} para las muestras Polo Argentino fue de 0,8%, estimado con un promedio 4,8 generaciones completas equivalentes conocidas (Tabla 2.3). Sin embargo, el F_{ROH} fue claramente superior, con valores promedio de alrededor del 14%. La mayor proporción del F_{ROH} es debida a la presencia de ROH cortos, menores a 5,5 Mb (F_{ROH_ANC}). El F_{IS} promedio fue negativo (-0,013).

Tabla 2.2. Cantidad y longitud promedio de ROH en la población Polo Argentino.

Longitud ROH	Número de ROH	Porcentaje del total ROH	Longitud media \pm DE (Mb)	Porcentaje de individuos con ROH
<5,5 Mb	36.011	84%	2,3 \pm 1,2	100%
5,5-8,3 Mb	3.621	8%	6,7 \pm 0,8	100%
8,3-16,6 Mb	2.569	6%	11,1 \pm 2,2	99%
$\geq 16,6$ Mb	526	1%	22 \pm 6,6	69%
Total	42.727	100%	3,5 \pm 3,4	100%

DE: desvío estándar

La mayor correlación fue observada entre el F_{ROH} y F_{IS} ($\rho = 0,79$, Tabla 2.3). La correlación entre F_{PED} y F_{ROH} fue de sólo 0,28, siendo aún menor con el F_{ROH_ANC} ($\rho = 0,17$). El F_{PED3} , es decir, el F_{PED} producido en las últimas 3 generaciones, tuvo una mayor correlación con el F_{ROH_3G} , mientras que con el F_{ROH_ANC} su

correlación fue baja y no significativa. Cuando se consideró solamente a los individuos con al menos 3 generaciones completas conocidas de ancestros, la media del F_{PED} fue de 2% ($n= 101$). En este grupo de individuos las correlaciones fueron mayores, siendo entre F_{PED} y F_{ROH} fue de 0,49 ($p<0,01$) y entre el F_{PED} y F_{ROH_3G} fue de 0,39 ($p<0,01$). No se observó una tendencia en las variables de endogamia analizadas a través de los años, ni se hallaron correlaciones significativas entre dichas variables y la fecha de nacimiento de los individuos.

Tabla 2.3. Promedio de los coeficientes de endogamia calculados. Correlación de Spearman (ρ) entre el F_{PED} , F_{PED3} , F_{IS} y los F_{ROH} .

	Media \pm DE	Porcentaje del F_{ROH} total	Correlación F_{PED}	Correlación F_{PED3}	Correlación F_{IS}
F_{PED}	0,008 \pm 0,026	-	-	-	-
F_{PED3}	0,0025 \pm 0,023	-	0,38***	-	-
F_{IS}	-0,013 \pm 0,03	-	0,31***	0,22***	-
F_{ROH}	0,148 \pm 0,025	-	0,28***	0,19***	0,79***
F_{ROH_3G}	0,008 \pm 0,010	5%	0,14**	0,21***	0,55***
F_{ROH_6G}	0,032 \pm 0,017	22%	0,21***	0,21***	0,67***
F_{ROH_9G}	0,055 \pm 0,020	38%	0,25***	0,19***	0,72***
F_{ROH_ANC}	0,092 \pm 0,013	62%	0,17***	0,06	0,4***

Valor- $p<0,01$; *Valor- $p<0,001$. DE: desvío estándar.

En la Tabla 2.4 se muestra la media de los distintos coeficientes F_{ROH} en cada cromosoma. El porcentaje de endogamia (F_{ROH}) por cromosoma fue variable, con un rango entre 0,24 para el ECA 4 y 0,08 para el ECA 8. Sin embargo, los datos presentaron una gran dispersión, ya que algunos individuos presentaron una alta proporción de alguno de sus cromosomas cubiertos por ROH (incluso hasta valores de más del 80% en ciertos casos) mientras que otros no registraron ROH. El ECA 9 fue el que presentó mayor proporción de ROH mayores a 16,6 Mb ($F_{ROH_3G} = 0,07$).

Tabla 2.4. Promedio de los coeficientes de endogamia F_{ROH} por cromosoma en la población.

	F_{ROH}	F_{ROH_3G}	F_{ROH_6G}	F_{ROH_9G}	F_{ROH_ANC}
ECA 1	0,21 ± 0,08	0,03 ± 0,06	0,07 ± 0,08	0,11 ± 0,08	0,1 ± 0,09
ECA 2	0,18 ± 0,09	0,02 ± 0,07	0,04 ± 0,07	0,07 ± 0,08	0,11 ± 0,11
ECA 3	0,14 ± 0,09	0,02 ± 0,06	0,04 ± 0,07	0,07 ± 0,08	0,07 ± 0,07
ECA 4	0,24 ± 0,11	0,03 ± 0,08	0,07 ± 0,1	0,11 ± 0,1	0,13 ± 0,13
ECA 5	0,12 ± 0,09	0,01 ± 0,05	0,03 ± 0,06	0,05 ± 0,07	0,07 ± 0,07
ECA 6	0,16 ± 0,1	0,01 ± 0,04	0,03 ± 0,06	0,06 ± 0,08	0,1 ± 0,11
ECA 7	0,16 ± 0,09	0,01 ± 0,05	0,04 ± 0,07	0,07 ± 0,08	0,09 ± 0,09
ECA 8	0,08 ± 0,07	0,01 ± 0,04	0,02 ± 0,05	0,03 ± 0,06	0,05 ± 0,06
ECA 9	0,18 ± 0,12	0,07 ± 0,11	0,08 ± 0,11	0,1 ± 0,11	0,08 ± 0,08
ECA 10	0,14 ± 0,1	0,03 ± 0,08	0,05 ± 0,08	0,06 ± 0,09	0,08 ± 0,08
ECA 11	0,18 ± 0,11	0,01 ± 0,06	0,04 ± 0,1	0,07 ± 0,1	0,11 ± 0,11
ECA 12	0,11 ± 0,11	0 ± 0	0,01 ± 0,04	0,02 ± 0,07	0,09 ± 0,09
ECA 13	0,09 ± 0,09	0 ± 0	0 ± 0,03	0,02 ± 0,06	0,07 ± 0,07
ECA 14	0,18 ± 0,1	0,02 ± 0,08	0,05 ± 0,09	0,08 ± 0,1	0,1 ± 0,1
ECA 15	0,12 ± 0,09	0,01 ± 0,04	0,03 ± 0,07	0,05 ± 0,08	0,07 ± 0,07
ECA 16	0,12 ± 0,08	0,01 ± 0,05	0,03 ± 0,07	0,05 ± 0,07	0,07 ± 0,07
ECA 17	0,18 ± 0,1	0,02 ± 0,09	0,05 ± 0,1	0,07 ± 0,11	0,11 ± 0,11
ECA 18	0,19 ± 0,1	0,01 ± 0,06	0,05 ± 0,08	0,08 ± 0,1	0,11 ± 0,11
ECA 19	0,1 ± 0,09	0,003 ± 0,03	0,01 ± 0,05	0,03 ± 0,07	0,07 ± 0,07
ECA 20	0,12 ± 0,09	0,001 ± 0,02	0,01 ± 0,05	0,03 ± 0,06	0,09 ± 0,09
ECA 21	0,14 ± 0,11	0,01 ± 0,04	0,02 ± 0,07	0,05 ± 0,1	0,09 ± 0,09
ECA 22	0,18 ± 0,13	0,02 ± 0,08	0,06 ± 0,11	0,09 ± 0,12	0,09 ± 0,09
ECA 23	0,15 ± 0,11	0,01 ± 0,04	0,02 ± 0,07	0,04 ± 0,09	0,11 ± 0,11
ECA 24	0,13 ± 0,11	0 ± 0	0,02 ± 0,06	0,04 ± 0,08	0,09 ± 0,09
ECA 25	0,16 ± 0,13	0,002 ± 0,04	0,03 ± 0,09	0,05 ± 0,11	0,11 ± 0,11
ECA 26	0,16 ± 0,12	0,002 ± 0,03	0,02 ± 0,07	0,04 ± 0,09	0,11 ± 0,11
ECA 27	0,13 ± 0,11	0,002 ± 0,03	0,01 ± 0,05	0,03 ± 0,08	0,1 ± 0,1
ECA 28	0,13 ± 0,1	0,01 ± 0,05	0,02 ± 0,07	0,04 ± 0,09	0,09 ± 0,09
ECA 29	0,12 ± 0,11	0,003 ± 0,05	0,01 ± 0,07	0,03 ± 0,08	0,09 ± 0,09
ECA 30	0,17 ± 0,12	0,002 ± 0,04	0,01 ± 0,06	0,03 ± 0,09	0,14 ± 0,13
ECA 31	0,12 ± 0,11	0 ± 0	0,004 ± 0,04	0,01 ± 0,06	0,11 ± 0,11
ECA X	0,13 ± 0,15	0,02 ± 0,05	0,04 ± 0,1	0,07 ± 0,12	0,06 ± 0,05

Valores expresados como media ± desvío estándar

3.3. Islas ROH

Se detectaron 4 regiones enriquecidas con ROH, localizadas en tres cromosomas distintos, como potenciales ROHi en la población de Polo Argentino (Tabla 2.5).

Estas regiones estuvieron presentes en al menos el 60% de las muestras y el

promedio del valor-p de sus SNPs fue significativo ($p < 0,01$), de acuerdo con el criterio fijado previamente. Además, todas presentaron SNPs con valores-p altamente significativos ($p < 0,001$; Figura 2.1). La incidencia de SNPs (es decir, el porcentaje de veces que un SNP estuvo incluido dentro de un ROH en la población) fue en promedio de 15% (con un desvío estándar de 8,6%), y un valor máximo de 70%. En la Figura 2.1 se observa la distribución de los valores-p del test de permutaciones por SNP (expresado como $-\log_{10}(\text{valor-p})$).

Tabla 2.5. Regiones genómicas seleccionadas como potenciales ROHi.

ECA	N° SNPs	Inicio	Fin	Longitud (pb)	Valor-p promedio
1	27	21.872.021	22.733.810	861.789	0,0034
1	46	45.343.494	47.516.617	2.173.123	0,0016
7	64	40.712.529	43.841.379	3.128.850	0,0030
17	38	20.602.482	22.333.840	1.731.358	0,0033

Se detalla el inicio del ROHi y el final dentro del cromosoma, así como su longitud (en pares de bases). El valor-p promedio se calculó como la media de los valores-p de los SNPs contenidos dentro del segmento (obtenidos mediante el test de permutaciones).

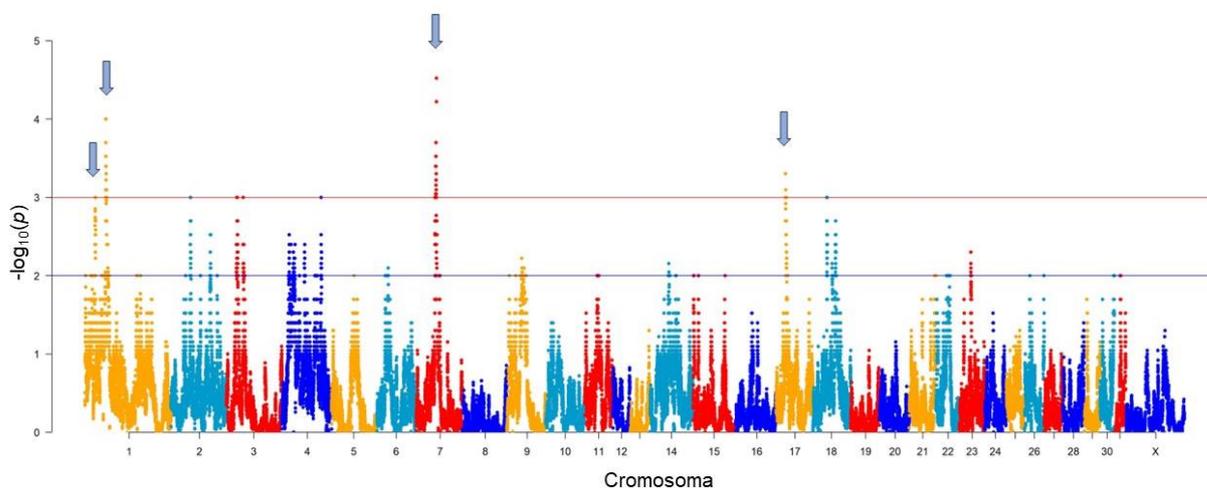


Figura 2.1. Gráfico Manhattan con los valores-p del test de permutaciones de los SNP dentro de bloques de homocigosidad. Las flechas azules indican la ubicación de las cuatro ROHi detectadas en la población Polo Argentino.

El análisis ontológico demostró que las 4 regiones (ROHi) contenían 25 genes diferentes: uno localizado en ECA 1 (*ZWINT*), 10 localizados en ECA 7 (*ACAD8*, *B3GAT1*, *GLB1L2*, *IGSF9B*, *JAM3*, *NCAPD3*, *NTM*, *OPCML*, *SPATA19*, *THYN1*) y 14 localizados en ECA 17 (*ARL11*, *CAB39L*, *CDADC1*, *CYSLTR2*, *EBPL*, *FNDC3A*, *KCNRG*, *KPNA3*, *MLNR*, *PHF11*, *RCBTB1*, *SETDB2*, *SPRYD7*, *TRIM13*). Los detalles de los genes se describen en la Tabla Suplementaria S2.1.

El análisis funcional reveló varias categorías GO (procesos biológicos) de los cuales carbohydrate metabolic process (GO: 0005975) e intracellular protein transport (GO: 0006886) fueron los más frecuentes (ver Tabla Suplementaria S2.2). La búsqueda bibliográfica en las bases de datos online permitió encontrar reportes sobre la función y asociación de dichos genes con distintos caracteres de interés, en la especie equina u otras especies (como modelos experimentales), muchos de los cuales fueron relacionados al comportamiento y funcionamiento neuronal, el metabolismo energético, el desarrollo músculo esquelético y el crecimiento, la reproducción, así como al desarrollo de algunas enfermedades específicas.

4. Discusión

En este capítulo se ha caracterizado la variabilidad genética de la raza Polo Argentino utilizando datos genómicos. Este estudio se centró en la detección de regiones ROH, la cual constituye una de las metodologías más empleadas últimamente para este tipo de análisis. Además, este enfoque permite sortear las desventajas que pudieran surgir del análisis de datos genealógicos, como la fiabilidad y la profundidad de pedigrí. Si bien la fiabilidad es un aspecto resuelto en el Polo Argentino, ya que la SRA lleva adelante un test de filiación e identificación individual a cada animal anotado, la falta de profundidad del pedigrí

constituye uno de los ítems más importantes para tener en cuenta en los análisis poblacionales en este tipo de razas jóvenes, que tienen pocas generaciones inscriptas.

Los valores de heterocigosidad molecular (H_o y H_e) en la población indicaron niveles moderados de diversidad genética, en comparación a otras razas. Petersen y col. (2013a) encontró que la H_e en el SPC fue de 0,31. Corbi Botto (2020) analizó la variabilidad en caballos Criollo Argentino, y encontró valores mayores de H_e y H_o ($\approx 0,34$). Razas de menor censo o una base genética reducida como el Pura Raza Español (Poyato-Bonilla y col., 2021), Pura Raza Árabe y Noriker (Druml y col., 2018), presentaron valores de H_o menores (entre 0,23 y 0,28). A pesar de que el Polo Argentino es una raza oficialmente abierta, los niveles de heterocigosidad fueron similares a los reportados en razas de gran censo y de varias décadas de mantener un registro cerrado (como el SPC y el Criollo). Esto podría ser debido a que posee un flujo minoritario de individuos pertenecientes a otras razas o sin registro, como ha sido indicado por Martínez y col. (2021). Por otro lado, la relación entre H_o y H_e sugiere un exceso de heterocigotas y una desviación de la endogamia, lo cual se advierte también por la presencia de un F_{IS} promedio negativo en los ejemplares analizados. Estos resultados concuerdan con los presentados por Martínez y col. (2021), quienes analizaron la variabilidad genética de caballos Polo Argentino mediante un panel de 15 marcadores microsatélites (utilizados para la identificación individual y de filiación), encontrando también valores de F_{IS} negativos.

En promedio, el 14% del genoma de la población Polo Argentino analizada se encontró cubierta por ROH ($F_{ROH} = 0,14$). Niveles similares de endogamia han sido reportados en caballos de raza SPC (Cunningham y col., 2001; Todd y col., 2018), que son la principal influencia genética del Polo Argentino. Un estudio realizado

por Hill y col. (2022) con alrededor de 8.950 caballos SPC encontró que el promedio de F_{ROH} fue de 0,28. Por otro lado, Schurink y col. (2019) analizaron la diversidad genética en 9 razas equinas, encontrando valores de F_{ROH} que oscilaron entre 0,22 y 0,04, aunque las razas Warmblood con un pedigrí abierto y flujo génico con otras poblaciones tuvieron los valores más bajos (0,05 – 0,097). Sin embargo, es necesario destacar que la metodología empleada, y los parámetros utilizados pueden afectar ampliamente la detección de ROH en el genoma. En este sentido, Meyermans y col. (2020) demostraron que la densidad del genotipado empleada (mediana o alta densidad), la longitud mínima del segmento, la cantidad de heterocigotas permitidos, y el control de calidad previo de los datos, entre otros, son parámetros que pueden afectar de diferente forma la identificación de los ROH. Además, no existe en la actualidad un consenso claro que determine cuáles son los parámetros óptimos para cada raza y/o especie. Por lo tanto, si bien las comparaciones entre las tendencias obtenidas son altamente informativas, la comparación de valores absolutos entre estudios debe ser cautelosa, ya que no necesariamente comparten los mismos criterios de identificación de ROH.

La mayor proporción de F_{ROH} en el caballo Polo Argentino corresponde a ROH cortos, menores a 5 Mb, lo cual indica que los ancestros comunes se ubican varias generaciones atrás (más de 10). Esto es debido a que en un cruzamiento endogámico se generan ROH largos que se acortan debido a las sucesivas meiosis a los que son sometidos a través de las generaciones (Sumreddee y col., 2021). Por tal motivo, la identificación de estas regiones en homocigosis puede ayudar a comprender la potencial depresión endogámica en una población. Por ejemplo, un metaanálisis realizado a partir de 154 artículos publicados sobre la depresión endogámica en diversas especies ganaderas (incluida la equina) reveló

que, en promedio, por cada aumento del 1% en la consanguinidad hay una disminución del 0,13% en el valor medio del carácter seleccionado (Doekes y col., 2021). Sin embargo, a pesar de que la consanguinidad aumenta la presencia de alelos deletéreos recesivos en estado homocigota, la endogamia y la selección pueden contribuir a la purga de esta carga genética negativa. La importancia de descomponer el coeficiente de endogamia en "ancestral" y "reciente" radica, en parte, en su capacidad para discriminar los alelos en homocigosis que han atravesado procesos de selección (Baumung et al., 2015). En este sentido, los ROH cortos, que han atravesado varias generaciones en estado homocigota porque se han originado varias generaciones atrás, tienen menos probabilidad de contener alelos deletéreos que los ROH largos. Esto se explica porque si un alelo ha estado en homocigosis por varias generaciones, es posible que haya atravesado procesos de selección, por lo que es menos probable que tenga un efecto deletéreo sobre el fenotipo. Es decir, que si los alelos que contribuyen a la depresión endogámica son eliminados efectivamente en los ancestros por selección (cuando están en estado homocigota), la depresión endogámica debería disminuir a medida que la endogamia ancestral aumenta (equilibrio mutación - selección) (Charlesworth y Willis, 2009). Este conocimiento puede ser útil para entender la relación entre la endogamia y la carga de alelos deletéreos en poblaciones consanguíneas.

Hill y col. (2022) encontraron que la probabilidad de debutar en carreras disminuía con el aumento del F_{ROH} en caballos SPC. Este efecto se debía a la presencia de ROH largos en el genoma (> 5 Mb), mientras que los ROH cortos (< 5 Mb) no tuvieron ninguna incidencia sobre la variable analizada. A pesar de esto, otros reportes indican que la purga de los alelos deletéreos no siempre ocurre efectivamente a pesar de la selección. Por ejemplo, Zhang y col. (2015) estudiaron

la distribución de ROH y variantes deletéreas en ganado bovino, encontrando que las mismas se ubicaban fundamentalmente en ROH de corta y mediana longitud. Todd y col. (2018) hallaron una relación negativa entre F_{PED} y distintas medidas de rendimiento en carreras de caballos SPC, a pesar de que el coeficiente de historia ancestral de consanguinidad (es decir, el número de veces que un alelo ha sido idéntico por descendencia en el pedigrí de un individuo) tuvo una asociación positiva con el rendimiento. Los autores consideraron que, si bien la endogamia y la selección habían aumentado la frecuencia de alelos favorables, no había logrado eliminar eficazmente la carga genética en la población.

Aunque en el Polo Argentino no ha sido posible evaluar aun la relación entre la endogamia y el rendimiento, debido a la falta de valores fenotípicos específicos, podría considerarse que el aumento de la endogamia ancestral no necesariamente eliminaría por completo el efecto de la depresión endogámica. La presencia de cuellos de botella poblacionales, como han ocurrido no sólo en el surgimiento del SPC sino también en la actualidad del caballo Polo Argentino, puede fijar alelos indeseables en la población, incluso por barrido selectivo en regiones que están bajo selección positiva (Todd y col., 2018). Además, el tamaño del efecto de los alelos sobre los caracteres analizados, así como la presión de selección ejercida, juegan roles muy importantes en el proceso. La purga es mucho más efectiva para alelos que tienen un efecto importante en el fenotipo, por ejemplo, alelos letales, que aquellos que son parcialmente deletéreos (Hedrick y Garcia-Dorado, 2016). Estos últimos son los mayores responsables de la depresión endogámica observada en caracteres complejos como los relacionados a la producción y el rendimiento deportivo (Sumreddee y col., 2021). Por otro lado, las tasas de incremento de la endogamia demasiado altas pueden favorecer la fijación de alelos indeseables más rápidamente de lo que la selección podría

eliminarlos. Por tal motivo, la ausencia de ROH \geq 16,6 Mb en algunos cromosomas podría deberse a la presencia de genes claves para el desarrollo y la supervivencia, que no tolerarían la presencia de alelos deletéreos en estado homocigota (Hedrick y Garcia-Dorado, 2016).

El F_{ROH} promedio de la raza fue sustancialmente mayor al F_{PED} (0,14 vs. 0,008), mientras que el F_{PED} y el $F_{ROH_{3G}}$ tuvieron valores promedios iguales en la población (0,008). Esto podría deberse a que el F_{PED} en esta raza también indica niveles de endogamia relativamente recientes, dada la cantidad de generaciones conocidas del pedigrí de los individuos. La correlación entre el $F_{ROH_{3G}}$ y el F_{PED3} fue mayor que la encontrada entre $F_{ROH_{3G}}$ y el F_{PED} , ya que ambos ($F_{ROH_{3G}}$ y F_{PED3}) reflejan eventos endogámicos producidos en las últimas 3 generaciones. A pesar de ello, en ambos casos la correlación fue moderada a baja, probablemente debido a que los individuos poseían un pedigrí relativamente incompleto. De hecho, al considerar sólo a los individuos con al menos 3 generaciones completas, los valores de correlación entre F_{PED} y $F_{ROH_{3G}}$ ascendieron, al igual que el F_{PED} promedio. En el mismo sentido, la correlación entre los métodos moleculares de cálculo de la endogamia fue elevada (F_{ROH} y F_{IS}), pero el F_{PED} mostró una baja correlación con ambos (0,28 y 0,31, respectivamente). En ambos casos, las discrepancias observadas podrían deberse a la relativamente baja profundidad del pedigrí de los animales analizados (4,8 generaciones completas equivalentes en promedio), lo cual genera una subestimación de los valores reales de endogamia (F_{PED}) (Curik y col., 2014). Por el contrario, el F_{ROH} puede capturar segmentos de IBD que se hayan originado varias generaciones atrás, por lo que tiende a ser un indicador más fiable del nivel real de consanguinidad de un individuo.

Algunos estudios han encontrado una correlación moderada a alta entre las estimaciones de la endogamia a partir de datos de pedigrí y genómicos en

animales domésticos (Peripolli y col., 2016; Goszczynski y col., 2018; Perdomo-González y col., 2022). Sin embargo, otros trabajos realizados en equinos han reportado correlaciones moderadas a bajas para estos coeficientes. Todd y col. (2018) también encontraron correlaciones similares a las de nuestro trabajo en caballos SPC australianos, con valores de 0,19 entre el F_{PED} y F_{ROH} y de 0,35 entre F_{PED} y $F_{ROH} > 16$ Mb (equivalentes a los $F_{ROH_{3G}}$ calculado en este capítulo). Los autores hipotizaron que esta baja correlación podía ser atribuida a que una gran proporción de la endogamia era ancestral y podría ser mejor estimada a través de los datos de pedigrí, ya que el mismo contaba con 24 generaciones completas equivalentes. La capacidad de detección de la endogamia ancestral a nivel molecular, caracterizada por ROH cortos, en cambio, depende de la densidad de SNPs del array y de los parámetros utilizados para su detección, por lo que podría no equiparar a lo calculado a través de un pedigrí robusto como el de dicha población. En el Polo Argentino, en cambio, la situación podría ser la inversa, ya que el F_{PED} se basa en una población con pocas generaciones conocidas. Por otro lado, la raza tiene la particularidad de que la mayoría de sus ancestros más contribuyentes son individuos provenientes de una población cerrada desde hace décadas (SPC) por lo que la asunción de parentesco nulo entre ejemplares fundadores probablemente sea errónea. Esto hace que el F_{PED} subestime los valores de endogamia en una población en la que la mayor proporción se genera en generaciones lejanas de las que no se cuenta con información alguna (al menos en el libro genealógico del Polo Argentino). A su vez, el F_{PED} es un método probabilístico que no tiene en cuenta la naturaleza estocástica de la recombinación o los cambios de frecuencias alélicas debido a la selección, procesos que pueden alterar la asociación con la endogamia real de una población (Curik y col., 2014; Hedrick y Garcia-Dorado, 2016). Fawcett y col. (2019) describieron en caballos SPC japoneses (con un pedigrí menos robusto, similar al

del Polo Argentino) una correlación entre F_{ROH} y F_{PED} de entre 0,34 y 0,37, dependiendo la cantidad de generaciones consideradas. En caballos de raza Mangalarga Marchador (Bizarria Dos Santos y col., 2021) ha sido descrita una situación similar, en la que el F_{PED} subestima ampliamente los valores genómicos de endogamia (F_{ROH}). Incluso la correlación en esta raza entre el F_{PED} y el F_{ROH} fue de tan sólo 0,02, probablemente también debido al carácter incompleto del pedigrí.

Si bien los ROH se han encontrado distribuidos a lo largo de todo el genoma, han sido en ciertas regiones de los cromosomas 1, 7 y 17 donde han estado en mayor abundancia. En ellos se han encontrado 4 ROHi caracterizadas por su alta frecuencia relativa en la población. En el SPC han sido descritas varias regiones genómicas de baja diversidad alélica, dos de las cuales solapan con los ROHi del ECA 1 y ECA 7 (Fawcett y col., 2019). Además, la ROHi del ECA 17 coincide con un haplotipo de alta frecuencia que ha sido encontrado también en esta raza (Petersen y col., 2013b). Las ROHi constituyen patrones genómicos de reducción de variabilidad comunes en la población posiblemente debido a que contienen genes que han estado bajo selección o barrido selectivo (Saravanan y col., 2020). Es por ello por lo que hemos analizado aquellos genes contenidos dentro de dichas regiones, habiéndose encontrado 25 genes cuya función ha sido previamente descrita en animales.

Desde la domesticación hasta la creación de diferentes razas, los seres humanos han seleccionado características conductuales óptimas para cumplir con funciones específicas. En la especie equina, los rasgos de comportamiento son un criterio fundamental para la selección (Yokomori y col., 2023). Varios de los genes detectados dentro de regiones ROHi están relacionados con el funcionamiento neurológico. Por tal motivo, es posible que el proceso de selección haya dejado huellas en diferentes loci relacionados con el desarrollo del sistema

nervioso, la neurotransmisión y la inteligencia (Wickens y Brooks, 2020). Si bien la determinación de patrones conductuales en caballos de polo es escasa (Álvarez y col., 2022a; Álvarez y col., 2022b; Álvarez y col., 2022c) los resultados de nuestro análisis sugieren que esta estrategia permitiría identificar marcadores genéticos para ser empleados en los esquemas de selección por comportamiento.

Los genes *OPCML* y *NTM* localizados en el ECA 7 han sido asociados a caracteres de “durabilidad” en caballos de carrera. En un estudio de GWAS realizado por McGivney y col. (2019), ambos genes mostraron una asociación con el número de partidas a los 2 y 3 años, período más importante desde el punto de vista económico en la carrera deportiva de los SPC. La función de dichos genes ha sido vinculada con la neurobiología, el comportamiento y el desarrollo del sistema nervioso; en humanos se lo ha vinculado con la inteligencia y la cognición (Pan y col., 2011). Un estudio encontró que estos genes se encontraban en regiones con una alta diferenciación genética entre razas equinas pesadas y livianas, lo cual constituía una posible huella de selección (Gurgul y col., 2019). Los autores sugieren que la funcionalidad de estos genes podría potencialmente explicar las diferencias temperamentales típicas de ambos biotipos. El gen *KNCRG* codifica una proteína que regula la actividad de los canales de potasio dependientes de voltaje, y que está relacionado con la excitabilidad neuronal y el aprendizaje, ha sido asociado al mejor tiempo de carrera en caballos de trote noruegos (Velie y col., 2018). Estos hallazgos sostienen la idea de que los factores ligados a la conducta (temperamento, capacidad de aprendizaje, etc.) tienen incidencia directa sobre la disciplina y el rendimiento deportivo, y podrían tener un importante impacto en los caballos de polo.

Otros genes presentes en las ROHi se asocian con caracteres conductuales. Por ejemplo, el gen *IGSF9B* codifica una proteína de adhesión que regula circuitos

neuronales relevantes para el comportamiento, y se la ha asociado con caracteres de la conducta en ratones, como la ansiedad (Babaev y col., 2018), así como diferentes trastornos de la conducta e inteligencia en humanos (Coleman y col., 2019). El gen *B3GAT1* se expresa ampliamente en el sistema nervioso, se relaciona con el desarrollo neuronal, el aprendizaje y la memoria en ratas y se ha postulado que puede tener un rol preponderante en la gestión del estrés y en el proceso de domesticación de esta especie (Zeng y col., 2017). El gen *KPNA3* se ha relacionado a diferentes trastornos de la conducta en humanos y a los comportamientos motivacionales en estudios experimentales en ratones (Aomine y col., 2022).

Seis de los genes localizados en ROHi (*ACAD8*, *MLNR*, *NCAPD3*, *SPATA19*, *FNDC3A* y *ZWINT*) fueron previamente vinculados a funciones metabólicas y/o reproductivas. El gen *ACAD8* pertenece a una familia de flavoenzimas mitocondriales que catalizan la deshidrogenación alfa y beta de los ésteres de acil-CoA en el catabolismo de los ácidos grasos y aminoácidos de cadena ramificada. Polimorfismos de este gen se lo ha asociado a la tasa de crecimiento y calidad de la carne en la especie bovina (Li y col., 2007). *MLNR* codifica el receptor de la motilina, el cual interviene en el movimiento intestinal y la regulación del apetito, por lo que es un gen candidato para caracteres de crecimiento (Rosenfeld y Garthwaite, 1987; Takahashi y col., 2014). El gen *NCAPD3* codifica la subunidad D3 del complejo condensina II, la cual se localiza en las mitocondrias, y tiene un rol preponderante en la función mitocondrial y el estrés oxidativo (Deutschman y col., 2019). Además, este gen se ha encontrado sobre expresado en toros con alta fertilidad, por lo que podría tener un rol en la funcionalidad espermática (Muhammad Aslam y col., 2018). De forma similar, el gen *SPATA19* también ha sido asociado a la funcionalidad mitocondrial en la línea espermática en ratones

(Mi y col., 2015). Un estudio hecho en ratones encontró que el gen *FNDC3A* se expresa en las espermátidas e interviene en la adhesión con las células de Sertoli, por lo que también cumple un rol importante en la fertilidad del macho (Obholz y col., 2006). En la especie equina, un estudio describió que el volumen del eyaculado en caballos Franches-Montagnes estuvo sugestivamente asociado a un SNP en el gen *ZWINT*, el cual tienen función dentro del proceso de meiosis (Gmel y col., 2021). Este gen también ha estado dentro de una región de baja diversidad en caballos SPC (Fawcett y col., 2019). La identificación de genes o QTLs asociados a caracteres como el crecimiento y la reproducción son muy importantes en la producción equina, ya que tienen un impacto directo sobre la economía y la eficiencia del criador. Por otro lado, los genes que intervienen directamente sobre el metabolismo energético y la función mitocondrial también son candidatos para evaluar la actividad muscular, la producción de energía, la termorregulación, la utilización del oxígeno y la respuesta al estrés oxidativo. Todos estos factores son claves en la adaptación y el desempeño del animal durante el ejercicio (Littiere y col., 2020).

Por otro lado, los genes *PHF11*, *SETDB2* y *RCBTB1*, localizados en el ECA 17, han sido descritos como candidatos para la despigmentación tipo vitíligo (Druml y col., 2021). El vitíligo es un trastorno autoinmune que cursa con la destrucción de los melanocitos en la piel, y se da más frecuentemente en caballos de pelaje tordillo, presente en el Polo Argentino. Los genes *PHF11* y *SETDB2* se expresan en células del sistema inmune innato y están involucrados en los procesos de remodelación de la cromatina. El gen *RCBTB1*, en cambio, es supresor tumoral. También se ha observado una relación inversa entre esta entidad y la aparición de melanomas (frecuente también en ejemplares tordillos).

Otros genes han sido vinculados con funciones inmunes. El gen *ARL11* ha sido asociado a la activación de los macrófagos frente al estímulo de los lipopolisacáridos (Arya y col., 2018). El gen *CYSLTR2* codifica el receptor de leucotrienos, un derivado del ácido araquidónico liberado por leucocitos, cuya activación produce en los órganos blanco broncoconstricción, hiperreactividad de las vías respiratorias, aumento de la permeabilidad vascular y eosinofilia. Un estudio reciente describe la asociación entre polimorfismos del gen *CYSLTR2* y el asma en humanos (Dos y col., 2019), una condición de base alérgica que es frecuente también en equinos, sobre todo en aquellos que se mantienen estabulados y expuestos a diversos alérgenos ambientales. A su vez, el gen *JAM3* codifica una molécula de adhesión que se expresa en varios tejidos, entre ellos el músculo liso bronquial. Ben Hamouda y col. (2021) encontraron que este gen se encuentra sub expresado en el tejido muscular liso de los bronquios de caballos asmáticos, por lo que puede tener un rol importante en el desarrollo de esta enfermedad, y en los cambios de remodelación que la acompañan.

Finalmente, este estudio ha demostrado la potencialidad del análisis de ROHi como una herramienta importante para la comprensión de la estructura genética, la evolución y la selección de la raza. El rendimiento atlético de los caballos deportivos como el Polo Argentino se ve influenciado por diversos factores, ambientales y genéticos. Son múltiples los genes que podrían estar involucrados en la performance individual (Littiere y col., 2020), los cuales abarcan funciones variadas asociadas a la cognición, la conducta, el metabolismo energético, el crecimiento y la salud. La identificación de ROHi podría ser útil para localizar marcadores y genes asociados con estos caracteres, los cuales podrían ser valiosos no sólo para una mejor comprensión de los factores ligados al

rendimiento, sino también para ser aplicados directamente en la selección y el mejoramiento.

5. Material Suplementario

Tabla S2.1. Genes encontrados dentro de cada ROHi.

Ensembl ID	ECA	Inicio	Fin	Gen	Gen - nombre
ENSECAG00000000633	1	46780554	46783398	<i>ZWINT</i>	ZW10 interacting kinetochore protein
ENSECAG000000011007	7	43393063	43436204	<i>ACAD8</i>	acyl-CoA dehydrogenase family member 8
ENSECAG000000008023	7	43703357	43717469	<i>B3GAT1</i>	beta-1,3-glucuronyltransferase 1
ENSECAG000000024922	7	43577518	43698132	<i>GLB1L2</i>	galactosidase beta 1 like 2
ENSECAG000000012105	7	43108173	43156743	<i>IGSF9B</i>	immunoglobulin superfamily member 9B
ENSECAG000000015550	7	43261488	43324210	<i>JAM3</i>	junctional adhesion molecule 3
ENSECAG000000018370	7	43327359	43391974	<i>NCAPD3</i>	non-SMC condensin II complex subunit D3
ENSECAG000000007574	7	40744711	41648685	<i>NTM</i>	neurotrimin
ENSECAG000000009233	7	41739002	42744507	<i>OPCML</i>	opioid binding protein/cell adhesion molecule like
ENSECAG000000026872	7	43056449	43062642	<i>SPATA19</i>	spermatogenesis associated 19
ENSECAG000000001596	7	43415012	43420148	<i>THYN1</i>	thymocyte nuclear protein 1
ENSECAG000000004611	17	21404904	21408405	<i>ARL11</i>	ADP ribosylation factor like GTPase 11
ENSECAG000000000879	17	21647757	21721546	<i>CAB39L</i>	calcium binding protein 39 like
ENSECAG000000020688	17	21742207	21789523	<i>CDADC1</i>	cytidine and dCMP deaminase domain containing 1
ENSECAG000000004678	17	22242285	22243322	<i>CYSLTR2</i>	cysteinyl leukotriene receptor 2
ENSECAG000000021272	17	21338733	21379524	<i>EBPL</i>	EBP like
ENSECAG000000023697	17	21824152	21992965	<i>FNDC3A</i>	fibronectin type III domain containing 3A
ENSECAG000000008284	17	21068763	21073013	<i>KCNRG</i>	potassium channel regulator
ENSECAG000000015530	17	21239190	21330658	<i>KPNA3</i>	karyopherin subunit alpha 3
ENSECAG0000000035160	17	21808715	21812731	<i>MLNR</i>	motilin receptor
ENSECAG000000014972	17	21489896	21518337	<i>PHF11</i>	PHD finger protein 11
ENSECAG000000021791	17	21437926	21487762	<i>RCBTB1</i>	RCC1 and BTB domain containing protein 1
ENSECAG000000020357	17	21520962	21558652	<i>SETDB2</i>	SET domain bifurcated histone lysine methyltransferase 2
ENSECAG000000011472	17	21143804	21159229	<i>SPRYD7</i>	SPRY domain containing 7
ENSECAG000000011381	17	21074912	21076141	<i>TRIM13</i>	tripartite motif containing 13

El inicio y el fin corresponden a la ubicación genómica (en pares de bases) dentro de cada cromosoma.

Tabla S2.2. Genes encontrados dentro de cada ROHi y su respectiva clasificación de ontología génica (GO term), según DAVID.

ECA	Gen	GOTERM_BP_DIRECT	GOTERM_MF_DIRECT
1	ZWINT	GO:0000070~mitotic sister chromatid segregation,GO:0007094~mitotic spindle assembly checkpoint,GO:0051649~establishment of localization in cell,	GO:0047485~protein N-terminus binding,
7	THYN1	No asignado	No asignado
7	NTM	No asignado	No asignado
7	B3GAT1	GO:0005975~carbohydrate metabolic process,GO:0006486~protein glycosylation,GO:0050650~chondroitin sulfate proteoglycan biosynthetic process,	GO:0015018~galactosylgalactosylxylosylprotein 3-beta-glucuronosyltransferase activity,GO:0046872~metal ion binding,
7	OPCML	No asignado	No asignado
7	ACAD8	GO:0006629~lipid metabolic process,GO:0006886~intracellular protein transport,GO:0042147~retrograde transport, endosome to Golgi,	GO:0003995~acyl-CoA dehydrogenase activity,GO:0016627~oxidoreductase activity, acting on the CH-CH group of donors,GO:0050660~flavin adenine dinucleotide binding,
7	IGSF9B	No asignado	No asignado
7	JAM3	GO:0001525~angiogenesis,GO:0001780~neutrophil homeostasis,GO:0002250~adaptive immune response,GO:0002318~myeloid progenitor cell differentiation,GO:0002523~leukocyte migration involved in inflammatory response,GO:0002693~positive regulation of cellular extravasation,GO:0007155~cell adhesion,GO:0007160~cell-matrix adhesion,GO:0007286~spermatid development,GO:0019226~transmission of nerve impulse,GO:0030010~establishment of cell polarity,GO:0033624~negative regulation of integrin activation,GO:0033629~negative regulation of cell adhesion mediated by integrin,GO:0034113~heterotypic cell-cell adhesion,GO:0034333~adherens junction assembly,GO:0034394~protein localization to cell surface,GO:0042552~myelination,GO:0045176~apical protein localization,GO:0090022~regulation of neutrophil chemotaxis,GO:0090138~regulation of actin cytoskeleton organization by cell-cell adhesion,GO:0097241~hematopoietic stem cell migration to bone marrow,GO:0097530~granulocyte migration,GO:1902414~protein localization to cell junction,GO:1905710~positive regulation of membrane permeability,	GO:0005178~integrin binding,GO:0042803~protein homodimerization activity,GO:0046982~protein heterodimerization activity,GO:0098632~protein binding involved in cell-cell adhesion,
7	NCAPD3	GO:0007076~mitotic chromosome condensation,GO:0010032~meiotic chromosome condensation,GO:0051301~cell division,	GO:0003682~chromatin binding,GO:0035064~methylated histone binding,GO:0042393~histone binding,
7	GLB1L2	GO:0005975~carbohydrate metabolic process,	GO:0004565~beta-galactosidase activity,
7	SPATA19	No asignado	No asignado
17	CAB39L	GO:0035556~intracellular signal transduction,	GO:0043539~protein serine/threonine kinase activator activity,
17	ARL11	GO:0006886~intracellular protein transport,GO:0016192~vesicle-mediated transport,	GO:0003924~GTPase activity,GO:0005525~GTP binding,
17	CYSLTR2	GO:0007218~neuropeptide signaling pathway,	GO:0001631~cysteinyl leukotriene receptor activity,GO:0008528~G-protein coupled peptide receptor activity,
17	KCNRG	GO:0051260~protein homooligomerization,GO:1902260~negative regulation of delayed rectifier potassium channel activity,	GO:0042802~identical protein binding,

17	<i>TRIM13</i>	GO:0010332~response to gamma radiation,GO:0010942~positive regulation of cell death,GO:0016239~positive regulation of macroautophagy,GO:0016567~protein ubiquitination,GO:0030433~ubiquitin-dependent ERAD pathway,GO:0032897~negative regulation of viral transcription,GO:0043123~positive regulation of I-kappaB kinase/NF-kappaB signaling,GO:0045087~innate immune response,GO:0051092~positive regulation of NF-kappaB transcription factor activity,GO:0051865~protein autoubiquitination,GO:1902187~negative regulation of viral release from host cell,	GO:0003713~transcription coactivator activity,GO:0008270~zinc ion binding,GO:0061630~ubiquitin protein ligase activity,
17	<i>SPRYD7</i>	No asignado	No asignado
17	<i>PHF11</i>	GO:0000278~mitotic cell cycle,GO:0007059~chromosome segregation,GO:0010629~negative regulation of gene expression,GO:0051567~histone H3-K9 methylation,GO:0070828~heterochromatin organization,GO:0090309~positive regulation of methylation-dependent chromatin silencing,	GO:0003677~DNA binding,GO:0008270~zinc ion binding,GO:0018024~histone-lysine N-methyltransferase activity,GO:0046974~histone methyltransferase activity (H3-K9 specific),
17	<i>KPNA3</i>	GO:0006606~protein import into nucleus,GO:0006607~NLS-bearing protein import into nucleus,	GO:0008022~protein C-terminus binding,GO:0008139~nuclear localization sequence binding,GO:0061608~nuclear import signal receptor activity,
17	<i>SETDB2</i>	GO:0000278~mitotic cell cycle,GO:0007059~chromosome segregation,GO:0010629~negative regulation of gene expression,GO:0051567~histone H3-K9 methylation,GO:0070828~heterochromatin organization,GO:0090309~positive regulation of methylation-dependent chromatin silencing,	GO:0003677~DNA binding,GO:0008270~zinc ion binding,GO:0018024~histone-lysine N-methyltransferase activity,GO:0046974~histone methyltransferase activity (H3-K9 specific),
17	<i>CDADC1</i>	GO:0070383~DNA cytosine deamination,	GO:0004126~cytidine deaminase activity,GO:0008270~zinc ion binding,GO:0042803~protein homodimerization activity,GO:0061676~importin-alpha family protein binding,
17	<i>EBPL</i>	GO:0016125~sterol metabolic process,	GO:0047750~cholesterol delta-isomerase activity,
17	<i>RCBTB1</i>	No asignado	No asignado
17	<i>FNDC3A</i>	No asignado	No asignado
17	<i>MLNR</i>	No asignado	GO:0008528~G-protein coupled peptide receptor activity,GO:0042562~hormone binding,

CAPÍTULO III: Análisis de la estructura genética del caballo Polo Argentino, su relación con otras razas e identificación de huellas de selección.

1. Introducción

El caballo es un animal domesticado desde hace aproximadamente 5.500 años (Fawcett y col., 2019). A lo largo de la historia ha acompañado el desarrollo de la humanidad y ha sido seleccionado con diversos propósitos, como la producción cárnica, el tiro, el trabajo ganadero, el ocio y una amplia variedad de deportes de velocidad, potencia y/o resistencia (Schurink y col., 2019). Dichos fenotipos han sido evidentemente acompañados por una estructuración genética dentro de la especie, resultando en la formación de diversas razas con características distintivas. La presencia de subgrupos (razas) dentro de la población se caracteriza por una baja variabilidad dentro de cada subpoblación y una alta variabilidad si las comparamos entre si (McCue y col., 2012). Este proceso se ve generalmente acentuado por diversos factores, como el aislamiento geográfico o las distintas políticas de cría y registro de los ejemplares. Más allá de su origen, muchas de las razas que cuentan con una asociación de criadores establecen criterios propios de registro de los individuos, permitiendo o no el intercambio genético con otras razas.

Como se ha mencionado previamente en esta tesis, las estrategias de selección tienden a provocar modificaciones en las frecuencias alélicas de regiones genómicas específicas que controlan caracteres particulares de los individuos. Estos cambios generan patrones genéticos característicos en la población, denominados huellas o firmas de selección (Nielsen, 2005). En última instancia,

estos cambios a nivel genómico son los que han permitido la diferenciación y especialización de las razas para fines específicos, así como la adaptación a las condiciones ambientales en las cuales se han desarrollado.

Las huellas de selección pueden ser detectadas mediante diferentes metodologías genómicas. Una de ellas es el análisis de los patrones de ROH en la población (ROHi) (Saravanan y col., 2020), la cual fue utilizada en el Capítulo II de esta tesis, permitiendo identificar 4 ROHi en la población Polo Argentino. Otros enfoques se basan en las variaciones de las frecuencias génicas entre dos poblaciones, por ejemplo, mediante el cálculo del índice F_{ST} , el cual es una medida de distancia genética basada en la comparación de la heterocigosidad observada entre dos subpoblaciones. Este índice, que puede ser estimado para cada marcador genético, tiene un rango de variación entre 0 a 1, donde 0 indica ausencia de diferenciación, mientras que 1 refiere a una diferenciación máxima. Este enfoque permite identificar aquellos marcadores con mayor distancia genética entre dos grupos comparados, lo cual indica selección positiva (Zhao y col., 2015; Saravanan y col., 2020). Siendo que existe una correlación entre fenotipo y genotipo, es esperable que estos marcadores estén contenidos dentro de regiones que se asocien con caracteres de interés que expliquen las diferencias entre las razas.

El caballo Polo Argentino es una raza joven creada oficialmente en el año 1984, sin embargo, la cría del caballo “tipo polo” se remonta hacia fines de siglo XIX, cuando el deporte desembarca en Argentina (Parrish, 2018). Aunque las características particulares de esta raza han ido cambiando con el tiempo, se destaca la gran influencia genética del caballo de raza SPC en sus orígenes, con la cual mantiene un alto flujo genético actualmente, como fue descrito en el Capítulo I de esta tesis. Sin embargo, a pesar de esta estrecha relación, ambas

razas se utilizan para deportes diferentes para los que se requieren, a su vez, distintas aptitudes funcionales. Por lo tanto, sería esperable que el proceso de selección realizado sobre el caballo de polo haya producido la aparición de diferencias entre los animales a nivel del genoma. Pero, además, el carácter abierto de la política de registro del caballo Polo Argentino posibilita la introgresión de líneas genéticas provenientes de otras razas equinas que aporten caracteres deseables a la población. Tal es así, que se describe también la utilización de caballos Criollos o mestizos en los orígenes de la raza, ya que eran los tipos de caballos utilizados para el trabajo en las estancias (Laffaye, 2009; AACCP, 2023).

Por lo tanto, el objetivo de este capítulo fue determinar las frecuencias alélicas de los marcadores analizados y compararlas con animales pertenecientes a otras razas equinas de importancia con el fin de comprender mejor la estructuración genética del caballo Polo Argentino. Por otro lado, otro de los objetivos fue identificar huellas de selección a partir de la comparación de las frecuencias génicas con el caballo SPC, con el fin de localizar genes o loci asociados con caracteres que puedan explicar posibles diferencias entre ambas poblaciones.

2. Materiales y métodos

2.1. Material animal

En el presente capítulo se utilizaron las muestras de caballos Polo Argentino descritas en el Capítulo II (n=385), pertenecientes a distintos criadores del país. Adicionalmente, se seleccionaron 12 caballos de raza SPC (inscritos en Studbook Argentino) de distintas caballerizas y haras ubicados en la Ciudad de La Plata y alrededores, para la obtención de muestras de sangre. El criterio para la selección de estos animales fue que su única actividad deportiva haya sido el turf.

Adicionalmente, se tomaron muestras de 8 equinos de raza SPC que se utilizaban actualmente para la práctica del Polo en un club hípico español, no estando inscritos en el Libro Genealógico del Polo Argentino, los cuales se identificarán en el resto de la tesis como SPC – tipo polo. Con las muestras biológicas obtenidas se extrajo ADN y se realizó la genotipificación con el microarray de Illumina Equine GGP array (Illumina, San Diego, USA), de igual manera que para las muestras de ADN de los caballos Polo Argentino (ver sección 2. Materiales y métodos del Capítulo II).

Finalmente, se incluyeron datos genómicos de 25 equinos Peruano de Paso previamente genotipados por nuestro grupo de trabajo, 102 animales de 8 razas equinas pertenecientes al banco de datos genéticos del grupo MERAGEM (Mejora de Razas y Genética Molecular) de la Universidad de Córdoba (España) y de 405 caballos de 8 razas, publicados en un repositorio público por Momen y col. (2022) (Tabla 3.1). Siendo que los animales fueron genotipados utilizando diferentes plataformas (Axiom MNEC670K, MNEc2M e Illumina Equine GGP array), se creó una nueva base de datos definitiva manteniendo sólo los marcadores comunes a ambas tecnologías, compuesta por 42.570 SNPs autosómicos genotipados en 937 equinos de 15 razas diferentes.

Adicionalmente, con el objeto de evaluar si existían diferencias entre individuos con varias generaciones de ancestros inscritos en el libro genealógico y aquellos recientemente registrados, se determinó el número de generaciones completas que tenía cada ejemplar Polo Argentino. Posteriormente se conformaron dos grupos divergentes: Gen0 (aquellos individuos “base”, sin generaciones Polo) y Gen3 (aquellos que tenían al menos 3 generaciones completas de ancestros registrados en el libro genealógico).

Tabla 3.1. Origen de las muestras utilizadas para el análisis.

Raza	n animales	Fuente	Tipo de raza
Pura Raza Árabe	9	MERAGEM	Deporte
	36	Momem y col., (2022)	
Belga	22	Momem y col., (2022)	Tiro
Burguete	10	MERAGEM	Producción carne
Cuarto de milla	72	Momem y col., (2022)	Deporte / trabajo ganadero
Criollo Argentino	5	MERAGEM	Deporte / trabajo ganadero
Jaca Navarra	10	MERAGEM	Producción carne
Lusitano	3	MERAGEM	Deporte
	21	Momem y col., (2022)	
Menorquín	13	MERAGEM	Deporte / ocio
Peruano de Paso	25	Datos propios	Deporte / ocio
	162	Momem y col., (2022)	
Pura Raza Español	42	MERAGEM	Deporte
Polo Argentino	385	Datos propios	Deporte
Shetland pony	24	Momem y col., (2022)	Poni - Ocio
SPC	12	Datos propios	Deporte
	24	Momem y col., (2022)	
SPC - tipo polo	8	Datos propios	Deporte
Trotador Español	10	MERAGEM	Deporte
Welsh poni	44	Momem y col., (2022)	Poni Deporte / ocio

2.2. Estandarización de los datos genómicos

En primera instancia se realizó un control calidad de los genotipos mediante el filtrado de los datos por “call rate” por SNP y por individuo, utilizando el programa *PLINK* (Purcell y col., 2007). Al igual que en el Capítulo II, sólo los SNPs que se genotiparon en al menos el 90% de las muestras fueron incluidos en los sucesivos análisis (función “--geno 0,1”). Por otro lado, sólo las muestras que tenían más del 95% de los SNPs genotipados fueron conservadas (función “--mind 0,05”). Este conjunto de datos (SET_1) quedó conformado por 42.502 SNPs y 929 muestras (se eliminaron 2 muestras de Polo Argentino y 6 de PRE). Adicionalmente, se conformó un segundo conjunto de datos (SET_2) en el cual se eliminaron los marcadores con una frecuencia del alelo menor inferior al 1% (“--maf 0,01”) y se conservaron sólo los marcadores que no estuvieran en desequilibrio de ligamiento,

considerando un umbral de r^2 de 0,2 (“--indep-pairwise 50 5 0.2”), quedando conformado por 17.602 SNPs.

2.3. Estructuración genética y diferenciación entre razas

La estructuración genética y discriminación entre las razas analizadas fue realizada mediante 4 aproximaciones moleculares diferentes: Análisis de varianza molecular (AMOVA), análisis de componente principales (PCA), análisis de admixture y comparación de índices F_{ST} .

2.3.1. Análisis de la varianza molecular (AMOVA)

Las diferencias genómicas entre las distintas razas equinas fueron en principio evaluadas mediante un análisis de la varianza molecular (AMOVA) según la metodología descrita por Excoffier y col. (1992). El modelo incluyó el cálculo de la varianza entre razas y entre individuos dentro de cada raza. Para esto se utilizó el conjunto de datos SET_2 el cual primeramente convertido desde formato binario (*bed. bim. fam*) a objetos *genind* del entorno estadístico R mediante el paquete *Radiator* (Gosselin, 2020), con los cuales se realizó el AMOVA, implementado en el paquete de R *Poppr* (Kamvar y col., 2015).

2.3.2. Análisis de componentes principales (PCA)

Con el objetivo de evaluar la estructuración de la población Polo Argentino y su relación con otras razas equinas se realizó un análisis de componentes principales (PCA). El PCA realiza una reducción dimensional de los datos genómicos analizados a una serie de componentes principales que reflejen de manera óptima su variabilidad.

El análisis fue realizado con los datos del SET_2 mediante la función “*PCAir*” del paquete *Genesis* (Gogarten y col., 2019) en el entorno R (R-Core-Team, 2023). Esta función tiene la particularidad de que considera el parentesco molecular entre los individuos, y selecciona un subconjunto de individuos representativos no emparentados entre sí, sobre los cuales realiza un PCA estándar. Esto permite evitar el sesgo que pudiera producirse por la presencia de estructuras familiares (conocidas o no) dentro de la muestra. En este caso, el coeficiente de parentesco umbral para considerar a los individuos emparentados fue de 0,022. Los resultados fueron graficados utilizando el paquete *ggplot2* (Wickham, 2016).

2.3.3. Análisis de componente genética (*admixture*)

Para inferir la estructura de la población se realizó también un análisis *admixture*, el cual utiliza un enfoque bayesiano para estimar la ascendencia mediante un modelo de máxima verosimilitud (Alexander y col., 2009), asignando a cada individuo su proporción de pertenencia a un conglomerado determinado (K). El análisis fue realizado utilizando un modelo discriminante no supervisado en base a la información genómica del SET_2, con el programa *Admixture* (Alexander y col., 2009). El análisis fue implementado evaluando 24 *k rates* diferentes (K2-K25). Cabe aclarar que, con el fin de disminuir la diferencia en los tamaños muestrales entre las razas y minimizar una posible distorsión de los resultados (Puechmaille, 2016), se optó por analizar un subconjunto de individuos de las dos razas más numerosas: Polo Argentino (n = 60, pertenecientes a 24 criadores distintos) y Peruano de Paso (n = 62). El criterio de selección de estos animales se basó en el parentesco, ya que se incluyeron solo los individuos de estas razas con un umbral de coascendencia máximo de 0,022 (igual al utilizado para el PCA). Finalmente, el valor de K óptimo fue aquel que presentó un menor error en la

prueba de validación cruzada (cv). Los datos fueron graficados utilizando el programa *Structure plot* (Ramasamy y col., 2014).

2.3.4. Índice F_{ST}

Se realizó el cálculo del índice de fijación F_{ST} como medida cuantificable de distancia genética entre poblaciones. Este índice explica la variación genética que es debida a la presencia de estructuración dentro de la población; compara la variabilidad en las subpoblaciones en relación con la variabilidad total, a partir de la varianza de las frecuencias génicas. Sus valores toman un rango de 0 a 1, donde 0 indica que no existe diferenciación genética y las subpoblaciones comparten sus frecuencias génicas, mientras que un valor de 1 indica una estructuración total (Wright, 1951).

En este caso, el cálculo se realizó según la metodología descrita por Weir y Cockerham (1984), el cual corrige los valores según el tamaño de la muestra. Los índices F_{ST} fueron calculados de forma pareada entre cada raza utilizando los datos del SET_1, como una media de los valores de F_{ST} ponderados de cada SNP individualmente estimados con el paquete de R *OutFLANK* (Whitlock y Lotterhos, 2015). Además, con el objetivo de detectar posibles huellas de selección en la raza Polo Argentino, se identificaron los loci que mayor divergencia tenían con la raza SPC. Para ello, se utilizó también el paquete de R *OutFLANK*, el cual adopta la metodología descrita por Whitlock y Lotterhos (2015), en la que se detectan los loci con valores atípicos según la distribución inferida de los valores de F_{ST} de la población. Para este cálculo, se utilizaron los marcadores con una heterocigosidad mínima de 0,05, tal como recomiendan los autores, en los cuales se determinó la probabilidad de que las diferencias observadas entre poblaciones sean debidas al azar, es decir, su valor-p. Este valor fue posteriormente transformado a valor-q,

mediante una corrección por la proporción esperada de falsos positivos (FDR). Aquellos SNPs con un valor-q menor a 0,05 ($-\log(q) > 1,3$) fueron seleccionados como significativos. Finalmente, se consideraron como regiones genómicas afectadas a aquellas comprendidas entre ± 50.000 pares de bases de la posición del SNP significativo. Se seleccionó ese rango para asegurar un desequilibrio de ligamiento (r^2) de al menos 0,2, de acuerdo con lo reportado por McCue y col. (2012). Estas regiones fueron analizadas funcionalmente en base a la anotación más reciente del genoma equino (EquCab 3.0) siguiendo la metodología descrita para las ROHi en el Capítulo II.

3. Resultados

3.1. Análisis de la Varianza Molecular (AMOVA)

El porcentaje de la varianza explicado por la raza de procedencia fue de 9,55%, mientras que la variabilidad entre individuos dentro de cada raza fue baja (Tabla 3.2). Sin embargo, en el análisis realizado entre el Polo Argentino y el SPC, el porcentaje de la varianza explicado por la raza fue menor (4,3%).

Tabla 3.2. AMOVA entre razas (n=15).

Componente de la varianza	GL	Suma cuadrados	Cuadrados medios	Sigma	% varianza explicada
Raza	15	886365,7	59091	561,2	9,55
Individuos dentro de raza	913	4868180,2	5332,1	17,6	0,3
Individuos	929	4920839,7	5296,9	5226,9	90,15
Total	1857	10675385,6	5748,7	5875,7	100

GL: grados de libertad

Dentro de las muestras analizadas en Polo Argentino, el criador explicó el 5,8 % de la varianza, mientras que el número de generaciones en cada individuo (comparación entre los grupos Gen0 y Gen3) fue de sólo 2,9 %.

3.2. Análisis de componentes principales (PCA)

En la Figura 3.1 se pueden observar la agrupación de los individuos de las distintas razas de acuerdo con los primeros dos componentes, los cuales explicaron un 25,8% de la variabilidad. El Polo Argentino conforma un grupo relativamente homogéneo con el SPC y los SPC – tipo Polo, así como con algunos individuos Cuartos de Milla. Los PRE, Lusitanos, Menorquín y Criollos también se agrupan entre sí, mientras que los caballos Peruano de Paso forman un grupo bien separado del resto.

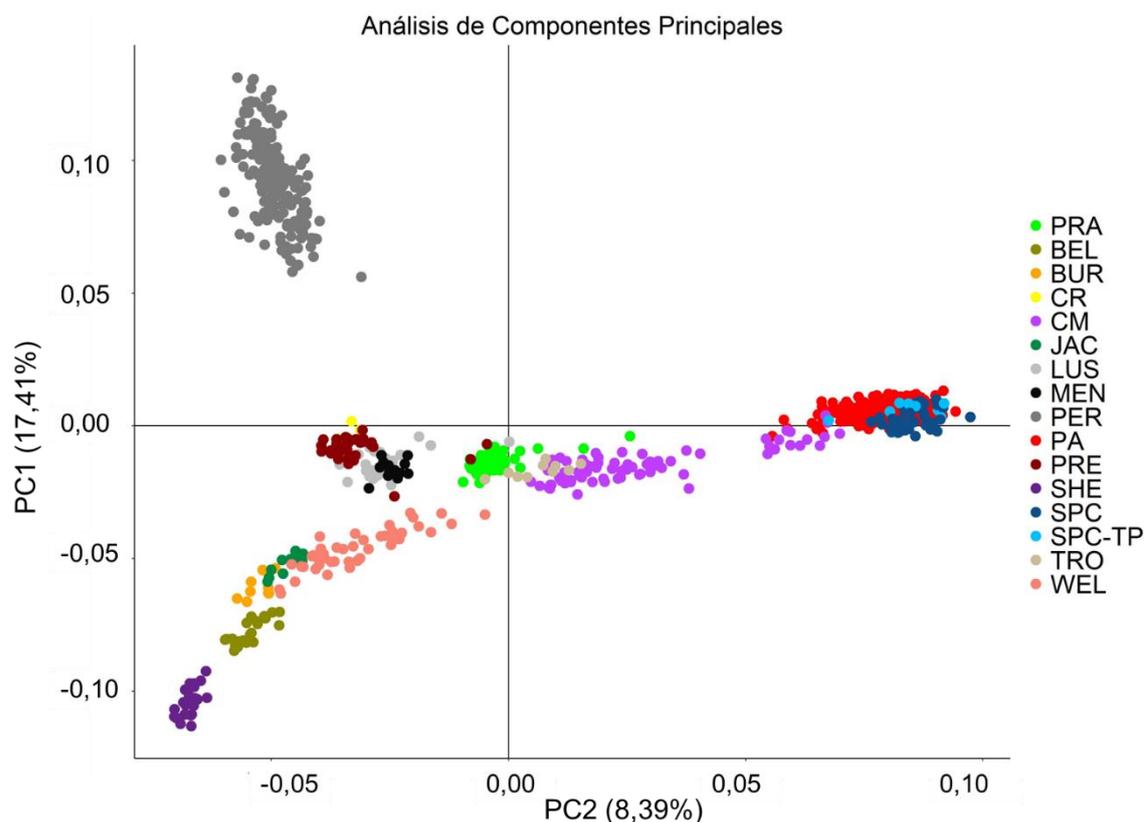


Figura 3.1. PCA entre las distintas razas equinas. Los ejes corresponden al componente 1 (PC1) y componente 2 (PC2) del análisis. PRA: Pura Raza Árabe; BEL: Belga; BUR: Burguete; JAC: Jaca Navarra; CR: Criollo; CM: Cuarto de Milla; PER: Peruano; LUS: Lusitano; MEN: Menorquín; PRE: Pura Raza Español; SPC: Sangre Pura de Carrera; SPC-TP: SPC – tipo Polo; PA: Polo Argentino; TRO: Trotador Español; WEL: Welsh; SHE: Shetland.

Con el fin de evaluar mejor la estructuración genética del Polo Argentino y su relación con el SPC, se realizó un PCA considerando únicamente estas dos razas

(Figura 3.2). En la Figura 3.2.A se observa una leve diferenciación entre los individuos Polo Argentino y los SPC, aunque el porcentaje de varianza explicado por ambos componentes fue bajo, si bien es esperable debido a la incipiente estructuración entre ambas. Esta diferencia fue levemente mayor cuando se consideró únicamente ejemplares Polo Argentino con 3 o más generaciones completas (Gen3) en su pedigrí (Figura 3.2.B y C).

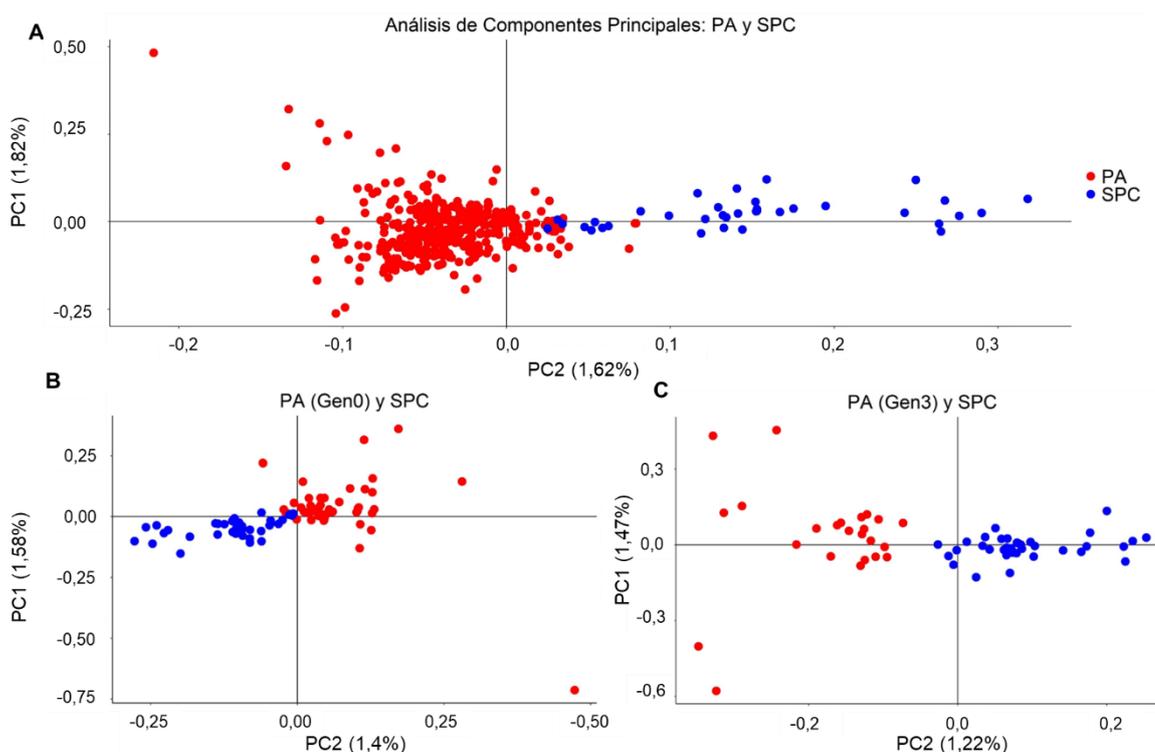


Figura 3.2. PCA entre Polo Argentino (PA) y Sangre Pura de Carrera (SPC). Gen0 y Gen3 corresponden a grupos de Polo Argentino con cero y tres o más generaciones Polo en su libro genealógico, respectivamente. Los ejes corresponden al componente 1 (PC1) y componente 2 (PC2) del análisis.

3.3. Análisis de componente genética (admixture)

En la Figura 3.3A se muestran la asignación de cada individuo a una población (K). Se encontraron resultados similares al PCA, donde K-10 fue seleccionado como el valor de K óptimo cuando se analizaron todas las razas, ya que presentó

el menor error de validación cruzada (Ver Figura S3.1). Este clúster K permitió asignar la mayoría de las razas a una población, aunque algunas de ellas presentaron una ascendencia común. El Polo Argentino y SPC conformaron una población relativamente homogénea (K-10), con más del 90% de asignación a la población 1 (Tabla Suplementaria S3.1). Las razas con asignaciones más heterogéneas fueron el Criollo Argentino y el Trotador Español. El mayor porcentaje de la composición del Polo Argentino fue compartido con el SPC (Figura 3.3).

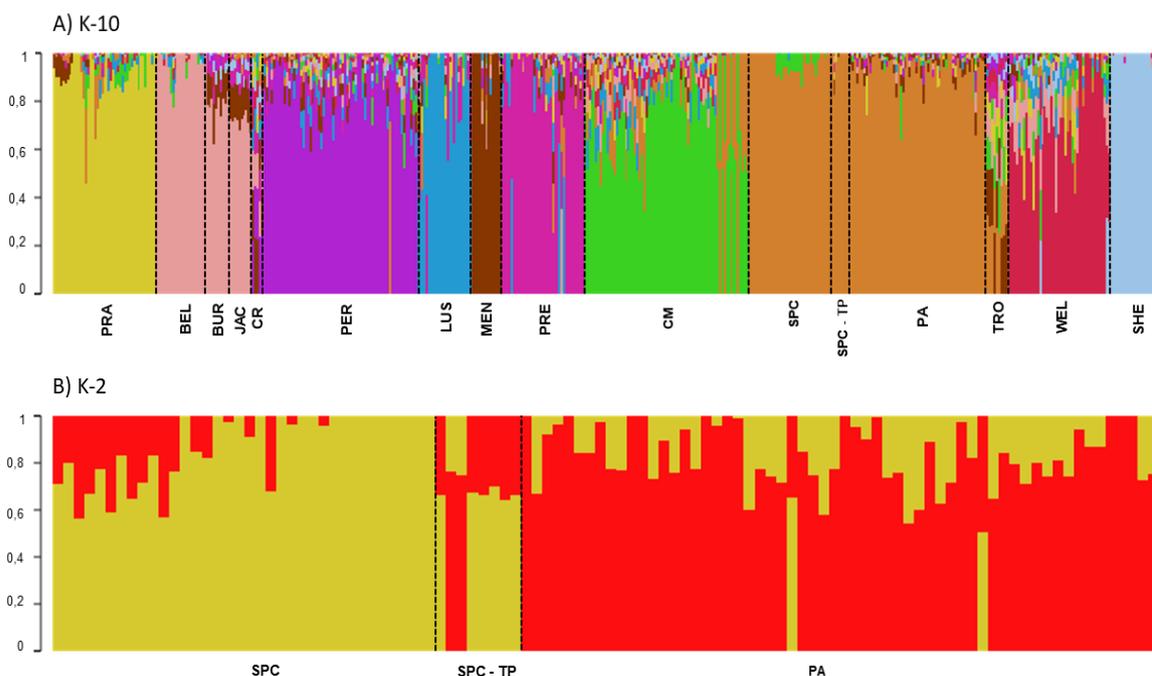


Figura 3.3. Análisis Admixture entre las distintas razas estudiadas. Cada columna del gráfico corresponde a un individuo. PRA: Pura Raza Árabe; BEL: Belga; BUR: Burguete; JAC: Jaca Navarra; CR: Criollo; PER: Peruano; LUS: Lusitano; MEN: Menorquín; PRE: Pura Raza Español; CM: Cuarto de Milla; SPC: Sangre Pura de Carrera; SPC-TP: SPC – tipo Polo; PA: Polo Argentino; TRO: Trotador Español; WEL: Welsh; SHE: Shetland. **A)** K-10 entre las 15 razas; **B)** K-2 entre SPC, SPC – tipo Polo y Polo Argentino. Las primeras 12 columnas de la raza SPC corresponden a los ejemplares muestreados en Argentina.

Los resultados del análisis admixture entre Polo Argentino y SPC reveló que el valor óptimo de K fue de 2 (Figura 3.3B). Los individuos Polo Argentino mostraron

una composición más heterogénea, aunque en promedio fueron asignados en su mayor proporción a la población 1 (82% - Tabla Suplementaria S3.1). La composición de los SPC, en cambio, fue en su mayoría de la población 2. Los ejemplares SPC – tipo Polo tuvieron una composición variable, mientras que los SPC muestreados en Argentina mostraron una proporción mayor de la población 1 en relación con el resto de los de la raza.

3.4. Índice F_{ST}

La diferenciación genética entre las razas también se cuantificó mediante el índice F_{ST} (Tabla 3.3). Los F_{ST} entre los grupos analizados tuvieron valores entre 0,004 y 0,189. Las razas menos distantes al Polo Argentino fueron el SPC (incluyendo al SPC – tipo Polo) y el Cuarto de Milla. Mientras que las razas más diferenciadas fueron el Shetland, Burguete y Jaca Navarra. Los caballos SPC – tipo Polo analizados en este estudio tuvieron una mayor distancia genética con el Polo Argentino (0,007) que con el SPC (0,004), si bien estos valores fueron bajos en comparación al resto (valor- $p < 0,01$). Por otro lado, los caballos Polo Argentino con 3 o más generaciones polo mostraron un mayor índice F_{ST} con los SPC que aquellos que tenían 0 generaciones (Gen3 = 0,025 vs. Gen0 = 0,011; valor- $p < 0,01$).

Los resultados del índice F_{ST} entre el Polo Argentino y SPC por pares de SNP individuales se presentan en la Figura 3.4. En total, 289 SNPs tuvieron un valor- $q < 0,05$ distribuidos en todos los cromosomas autosómicos (excepto el ECA 13), con valores de F_{ST} que oscilaron entre 0,17 y 0,92. Estos marcadores se agruparon en 263 regiones candidatas para ser huellas de selección en el Polo Argentino, e incluyeron a 108 genes de función conocida (Tablas Suplementarias S3.2 y S3.3). El análisis funcional de las regiones candidatas demostró que los

genes afectados están relacionados a una amplia variedad de procesos biológicos y funciones moleculares (Ver Tabla suplementaria S3.4). De ellos, 41 fueron seleccionados como genes candidatos de acuerdo con su función putativa luego de una revisión de la bibliografía existente.

Tabla 3.3. Valores de F_{ST} entre las distintas razas analizadas.

Raza	PRA	BEL	BUR	CR	QM	JAC	LUS	MEN	PER	PRE	PA	SHE	SPC	SPC-TP	TRO	WEL
BEL	0,099															
BUR	0,095	0,029														
CR	0,078	0,083	0,072													
QM	0,057	0,077	0,072	0,056												
JAC	0,096	0,047	0,035	0,080	0,074											
LUS	0,060	0,070	0,064	0,050	0,049	0,067										
MEN	0,084	0,085	0,074	0,071	0,067	0,080	0,060									
PER	0,079	0,076	0,068	0,050	0,068	0,071	0,056	0,070								
PRE	0,075	0,082	0,071	0,061	0,066	0,075	0,033	0,068	0,063							
PA	0,106	0,143	0,146	0,122	0,050	0,144	0,113	0,129	0,111	0,125						
SHE	0,127	0,102	0,105	0,131	0,108	0,109	0,106	0,122	0,098	0,110	0,163					
SPC	0,103	0,137	0,135	0,121	0,042	0,135	0,103	0,121	0,122	0,117	0,015	0,167				
SPC-TP	0,104	0,146	0,134	0,110	0,037	0,137	0,099	0,122	0,119	0,121	0,007	0,189	0,004			
TRO	0,083	0,102	0,091	0,080	0,053	0,096	0,068	0,086	0,085	0,083	0,101	0,141	0,096	0,088		
WEL	0,067	0,054	0,049	0,053	0,052	0,054	0,048	0,065	0,059	0,061	0,113	0,086	0,107	0,108	0,072	

F_{ST}

<0,05
0,05 a 0,099
0,1 a 0,13
>0,13

PRA: Pura Raza Árabe; BEL: Belga; BUR: Burguete; CR: Criollo; CM: Cuarto de Milla; JAC: Jaca Navarra; LUS: Lusitano; MEN: Menorquín; PER: Peruano de Paso; PRE: Pura Raza Español; PA: Polo Argentino; SHE: Shetland; SPC: Sangre Pura de Carrera; SPC-TP: SPC – tipo Polo; TRO: Trotador Español; WEL: Welsh.

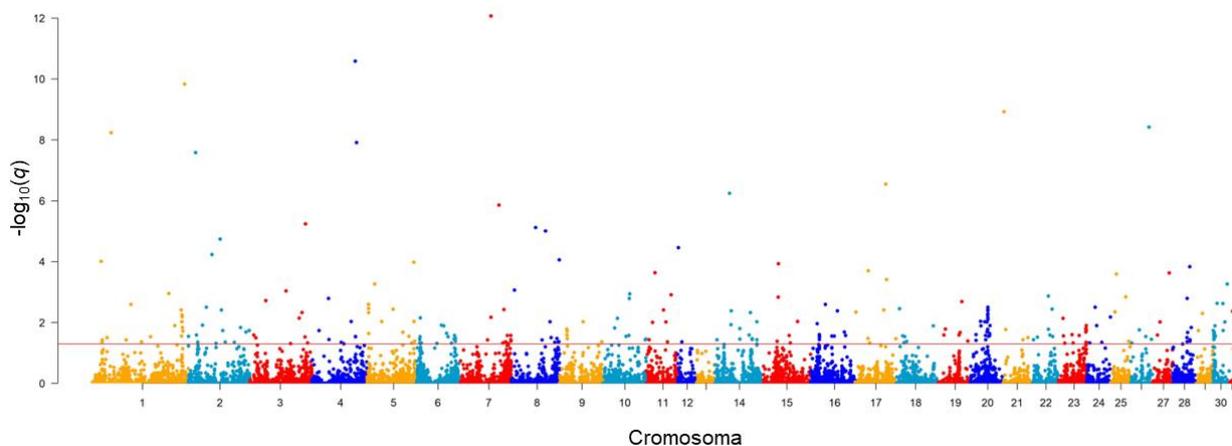


Figura 3.4. $-\log_{10}$ del valor-q de los índices F_{ST} por SNP y por cromosoma entre Polo Argentino y SPC. La línea roja indica el umbral de significancia ($-\log_{10}(q) > 1,3$ o valor-q < 0,05).

4. Discusión

En este capítulo se analizó la relación del caballo Polo Argentino con otras razas equinas de variada procedencia y utilizadas para diversos propósitos, como el deporte, el trabajo o incluso la producción cárnica. Para ello se utilizaron diferentes metodologías que permitieron evidenciar la distancia genética entre las razas o poblaciones a partir de los datos de genotipado con microarrays de SNP.

El AMOVA mostró que la mayor proporción de la varianza explicada por la estructuración racial fue menor al 10%, siendo la variación dentro de razas un porcentaje minoritario. En la bibliografía consultada es posible encontrar valores similares de varianza atribuida a la asignación racial, dependiendo el grado de cercanía entre las mismas. Por ejemplo, Petersen y col. (2013a) analizaron la diversidad genética de 36 razas equinas y la varianza genética entre las mismas fue de 10%, mientras que la varianza dentro de cada raza también fue baja (0,53%). A pesar de la cercanía entre el Polo Argentino y el SPC, la raza explicó alrededor de un 4% de la variabilidad encontrada. Algunos trabajos que analizan la presencia de subpoblaciones dentro de determinadas razas han encontrado porcentajes de varianza similares, a pesar de muchas de ellas presentan una estructura genética que se ha ido estableciendo a lo largo de su historia, la cual es mucho más antigua que la del Polo Argentino. Tal es así que Cosgrove y col. (2020) al estudiar la estructura genética del caballo Pura Raza Árabe encontraron que los linajes dentro de la raza explicaban el 5,5% de la varianza genética. Mientras que Poyato-Bonilla y col. (2021) evaluaron la diferenciación genética de la estirpe Cartujana en el Pura Raza Español, encontrando que la presencia de dos subpoblaciones en la raza explicaba un 5,2% de la varianza genética.

El bajo porcentaje de la varianza explicado por el criador dentro de la raza Polo Argentino podría deberse a la amplia conexión que existe entre los criadores,

donde existe un alto flujo génico e intercambio de reproductores entre los mismos, los cuales se ven favorecidos por la utilización de biotecnologías de la reproducción como la transferencia embrionaria y la inseminación artificial. Esto se condice con los resultados obtenidos en el Capítulo I de esta tesis, donde dicha interconexión entre criadores pudo constatararse con los datos de registro del libro genealógico de la raza.

Los análisis PCA, F_{ST} y admixture permitieron cuantificar y visualizar la distancia genética entre las poblaciones, mostrando resultados consistentes entre ellos. En el PCA se observó que las razas Polo Argentino, SPC y algunos individuos Cuarto de Milla formaron un conglomerado claramente separado del resto. La cercana relación con el Cuarto de Milla se explica por la gran influencia que posee del caballo SPC (Avila y col., 2018; Fawcett y col., 2019). Esta raza estadounidense se cría para múltiples disciplinas, como la rienda, el trabajo con ganado y las carreras de corta distancia (de allí deriva su nombre), por lo que es esperable que se presente cierta heterogeneidad dentro de la población. En un estudio realizado por Petersen y col. (2014) en el que se estudió la diversidad genética del Cuarto de Milla, se encontró que la raza presentaba un alto grado de estructuración en función de la actividad ecuestre para la que habían sido seleccionados los animales. Más aun, los individuos utilizados para las carreras tenían una mayor influencia del caballo SPC que los que participaban en otras disciplinas ecuestres. Si bien no se pudo identificar la actividad para la que habían sido seleccionados los animales que se utilizaron en este estudio, es posible que los que se encuentran más cercanos al SPC pertenezcan a líneas genéticas de carrera. Por otro lado, el Pura Raza Árabe formó un grupo separado del SPC, tal como se ha reportado en estudios previos (Cosgrove y col., 2020), con valores de F_{ST}

moderados entre ambas poblaciones, a pesar de la narración histórica que indica la estrecha relación en el origen entre ambas razas el cual se remonta siglos atrás.

El PCA realizado entre el Polo Argentino y el SPC separadamente, demostró una mayor diferenciación, a pesar del bajo porcentaje de variación explicado por los primeros dos componentes (PC1 y PC2). El Polo Argentino mostró una distribución homogénea, sin observarse estructuración evidente entre los individuos. El mayor grado de solapamiento entre individuos de distintas razas en el PCA también se ha observado entre aquellas poblaciones que por su historia evolutiva se conoce que están más estrechamente relacionadas. Por ejemplo, otro de los conglomerados evidentes estuvo conformado por las razas Pura Raza Español, Lusitano, Menorquín y Criollo Argentino. Las tres primeras son de origen ibérico y de similar biotipo, utilizados para distintas disciplinas hípcas como el rejoneo, la doma clásica y de alta escuela. El Criollo Argentino es una raza autóctona de nuestro país y la región, cuyo origen se remonta a los caballos traídos por los conquistadores españoles en el siglo XV, por lo que su ascendencia comparte posiblemente linajes con dichas razas ibéricas (Karlau y col., 2023). El Jaca Navarra y el Burguete también se mostraron agrupados, siendo ambas razas pesadas autóctonas de España, cuyo principal uso es la producción de carne (MAPAMA, 2023). El Peruano de Paso fue la raza más claramente individualizable del resto. A pesar de que su origen también se remonta a los caballos traídos por los conquistadores desde la península ibérica, presenta rasgos únicos relacionados a su locomoción, caracterizada por el paso amplado en cuatro tiempos (en bípedos laterales) (Barraza y col., 2018).

El análisis pareado del índice F_{ST} entre el Polo Argentino y el resto de las razas tuvo valores moderados a altos, con excepción de aquellas que en el PCA ya habían demostrado una mayor cercanía (SPC y Cuarto de Milla,

fundamentalmente). Al igual que lo visualizado en el PCA, la mayor diferenciación genética fue observada por un lado con la raza Shetland (raza de pony de las más pequeñas que existen) y con las razas pesadas (de tiro o de producción cárnica) como Belga, Burguete y Jaca Navarra ($F_{ST} > 0,13$). Este grado de diferenciación podría ser explicado por los diferentes orígenes y objetivos de selección de cada una de estas razas, lo cual ha aportado características morfo funcionales diferentes a cada una de ellas. Por otro lado, las razas deportivas como el Pura Raza Árabe, Trotador Español, Pura Raza Español y Welsh tuvieron un índice F_{ST} moderado con el Polo Argentino. La menor diferenciación genética fue observada con la raza SPC, con la cual tuvo un valor de $F_{ST}=0,015$ lo cual que podría considerarse relativamente bajo (Schurink y col., 2019). Similares valores de F_{ST} se han encontrado, por ejemplo, entre 4 razas Warmblood neerlandesas que, aunque estaban seleccionadas para diferentes propósitos, tenían un pedigrí abierto y un alto flujo genético entre sí (Schurink y col., 2019). Una situación parecida ocurrió entre las distintas subpoblaciones analizadas del caballo Pantaneiro de Brasil, donde se analizó el impacto de las barreras geográficas en la estructuración de la raza (Nogueira y col., 2021).

El SPC presentó también bajos valores de diferenciación genética con el Cuarto de Milla, aunque casi tres veces mayor que la observada con el Polo Argentino (0,042 vs. 0,015). Interesantemente, fue posible observar una menor distancia genética entre el Polo Argentino y el SPC – tipo Polo que entre el Polo Argentino y el SPC, posiblemente debido a que los caballos SPC seleccionados para la práctica del polo compartan características importantes para el deporte. De la misma manera, los caballos con 3 o más generaciones de ancestros inscriptos en el libro genealógico del Polo Argentino (Gen3) también tuvieron una mayor distancia genética con el SPC ($F_{ST}=0,025$), similares a los encontrados, por

ejemplo, entre el Burguete y el Belga o entre el PRE y el Lusitano. Esto podría indicar que los individuos que han sido seleccionados por varias generaciones para un deporte diferente (en este caso, el polo) presentan una incipiente diferenciación genética con la población de origen (el SPC) producto de la selección y/o fijación de características genéticas distintivas deseables dentro de la raza y al efecto fundador de algunas líneas genéticas específicas de SPC que han sido seleccionadas para un deporte distinto al turf. Por otro lado, la introgresión constante de individuos sin ancestros inscriptos en el libro genealógico (Gen0) confirma la existencia de un alto flujo genético con la raza SPC que se mantiene hoy en día.

El análisis admixture también evidencia la relación ancestral entre las distintas razas y particularmente la estrecha vinculación entre el Polo Argentino y el SPC. Si bien el número de K seleccionado como óptimo permitió explicar la estructuración genética de las distintas razas en función de su historia y evolución, en el caso de caballos Polo Argentino y SPC esto no fue posible, ya que conformaron un grupo con una composición relativamente homogénea común. Sin embargo, hay que destacar que si bien el Polo Argentino se mantiene como un población abierta, no se evidencia una heterogeneidad mayor que la observada en otras poblaciones cerradas, probablemente debido a que el mayor flujo genético lo mantiene con el SPC, una raza que mantiene su política de registro cerrada desde finales del siglo XVIII.

La influencia del SPC en el caballo de polo es esencial porque la velocidad es una de las características más buscadas en el polo moderno (Laffaye, 2009). Sin embargo, un caballo de polo no sólo debe ser veloz dentro del campo de juego, sino que también debe ser capaz de frenar, cambiar de dirección y velocidad bruscamente, poseer suficiente docilidad y concentración para entender y

responder a las ayudas del jinete con el dinamismo apropiado que el deporte demanda (MOP, 2018). Aunque el intercambio genético es constante entre el Polo Argentino y el SPC (unidireccionalmente), los ejemplares SPC fundadores posiblemente reúnan ciertas características fenotípicas (y genéticas) que les otorguen la aptitud necesaria para el juego. Es por ello por lo que, dentro del espectro de variabilidad genética que posee el SPC, probablemente el Polo Argentino haya heredado ciertos patrones específicos que responden a los criterios de selección aplicados. De hecho, Laffaye (2009) ha mencionado que la simple adquisición de ejemplares de carrera con un desempeño regular en los hipódromos para convertirlos en jugadores de polo, no ha dado tan buenos resultados como la selección y cría específica de ejemplares SPC para este deporte. Además, la amplia utilización de reproductores “famosos”, potenciada por la posibilidad de empelar biotecnologías de la reproducción sin restricciones (algo no permitido en la cría del SPC) podrían haber contribuido a establecer ese efecto fundador rápidamente.

Huellas de selección en el Polo Argentino

El análisis de las diferencias en las frecuencias alélicas entre los marcadores observada entre el Polo Argentino y el SPC permitió identificar regiones genómicas que podrían constituir incipientes huellas de selección. En este sentido, los marcadores con índices F_{ST} inusualmente altos se encuentran (generalmente) sometidos a una alta presión de selección en una de las razas, mientras que los que componen el centro de la distribución de F_{ST} es más factible que tengan un efecto neutral (Whitlock y Lotterhos, 2015). Por tal motivo, el análisis de los genes contenidos y/o próximos a dichas regiones podrían ayudar a comprender las características fenotípicas asociadas que están bajo selección positiva en la raza.

Se han publicado varios artículos que describen huellas de selección en la raza SPC, mediante distintos enfoques metodológicos, los cuales se basan en identificar las variantes genéticas que son más frecuentes entre los caballos SPC que entre los individuos de otras razas (Bailey y col., 2022). La mayoría de ellos se centra en encontrar los genes o regiones genómicas que subyacen a la adaptación al rendimiento deportivo en esta raza, la cual ha sido seleccionada desde hace siglos para las carreras (Petersen y col., 2013b; Moon y col., 2015; Fawcett y col., 2019).

Varios de los genes identificados dentro de las regiones con altos valores de F_{ST} se relacionaron con la funcionalidad y desarrollo del sistema nervioso, así como a diferentes características conductuales, en diversas especies. Uno de ellos es el gen *CDH13*, que ha sido identificado como un gen candidato para controlar los rasgos de personalidad en equinos SPC, en los cuales también se han reportado mutaciones no sinónimas que podrían dar lugar a cambios en la estructura proteica de los productos de dicho gen (Yokomori y col., 2023). El gen *EXOC4* ha sido implicado en una amplia gama de fenotipos conductuales y funciones cognitivas complejas. Un estudio realizado en bovinos de raza Brahman encontró que ciertos polimorfismos en regiones intrónicas de dicho gen se asociaban con la velocidad de huida (Paredes-Sánchez y col., 2020). Los genes *EPHA5* (Mamiya y col., 2008), *MACROD2* (Crawford y col., 2021), *NEGR1* (Söker y col., 2008) y *EYA3* (Noh y col., 2019) han sido estudiados en modelos murinos, donde se han asociado a diferentes conductas como la agresividad, el comportamiento exploratorio, la ansiedad e hiperactividad. Varios de los genes encontrados se encontraron asociados a las sinapsis neuronales y neuromusculares, como los genes *KCND2* (Cheng y col., 2021), *MCF2L* (Hayashi y col., 2013), *VT11A* (Tang y col., 2022) y *STXBP5* (Geerts y col., 2015). Algunos genes han sido asociados

a diversas habilidades cognitivas en ratones, entre las que se encuentran el aprendizaje y la memoria, como el gen *ST3GAL3* (Aboul-Naga y col., 2022), *CPEB3*, *VPS13B*, *CTNND2* (Kim y col., 2019; Qu y col., 2020; Wang y col., 2021) y *PTPRD* (Uhl y Martinez, 2019). Por último, fueron varios los genes vinculados con el neurodesarrollo, como *EPHA5* (Cooper y col., 2009), *PBX3* (Rhee y col., 2004), *SDCCAG8* (Insolera y col., 2014), *CDH13* (Ziegler y col., 2021), *GLS2* (Campos-Sandoval y col., 2015), *MACF1* (Moffat y col., 2017), *NAV2* (McNeill y col., 2011), *NEGR1* (Noh y col., 2019), *NRP1* (Martín-Fernández y col., 2022) y *SERPINI1* (Han y col., 2021).

Es importante remarcar que los genes relacionados a la estructura y funcionamiento neurológico constituyen genes candidatos para caracteres relacionados al temperamento y el comportamiento. Estos caracteres son claves en las especies domésticas, y más aún en la especie equina, en la cual diferentes rasgos del carácter son y han sido de los principales criterios de selección (Yokomori y col., 2023). De hecho, se han encontrado huellas de selección conteniendo múltiples genes relacionados al desarrollo cerebral y el comportamiento en el caballo, lo cual probablemente haya contribuido a generar los cambios cognitivos y sociales asociados a la domesticación (Librado y col., 2016). En este sentido, el caballo SPC se destaca por poseer un temperamento sanguíneo e hiperactivo, lo que hace que algunos caballos sean conflictivos y difíciles de manejar durante su etapa de entrenamiento y estabulación, constituyendo una causa importante de retiro de ejemplares de la actividad (Thomson y col., 2014; Holtby y col., 2023). El temperamento y la personalidad de los equinos son críticos a la hora de seleccionar a los ejemplares para una actividad determinada. Considerando las características conductuales del SPC, Yokomori y col. (2023) han destacado el condicionamiento del temperamento en

esta raza para la elección de su destino después de terminar la carrera deportiva asociada al turf.

Si bien no se han evaluado objetivamente las diferencias comportamentales entre caballos SPC y Polo Argentino, es sabido que entre las cualidades más valoradas de estos últimos destacan las relacionadas al temperamento, como la docilidad e inteligencia, pero también el brío suficiente como para cooperar con el jinete y desenvolverse de la mejor manera posible en un deporte tan dinámico como es el polo. Holtby y col. (2023) encontraron varios genes asociados a la adaptación al estrés en caballos SPC jóvenes, lo cual podría contribuir al desarrollo de estrategias de selección para mejorar el rendimiento de los equinos y las condiciones de bienestar animal.

Se han encontrado una serie de genes con funciones relacionadas al metabolismo, los cuales pueden ser claves para el ejercicio y la utilización de los sustratos energéticos. En el caballo SPC, que ha sido criado y seleccionado por más de tres siglos para la actividad atlética, se han encontrado huellas de selección conteniendo varios genes con funciones directas sobre el metabolismo energético (Gu y col., 2009). En nuestro trabajo se han encontrado genes que intervienen en el metabolismo lipídico y la adipogénesis (*GPAM* y *OSBPL10*, (Fairn y McMaster, 2008) (Yu y col., 2017; Yu y col., 2021), el funcionamiento tiroideo (*PLCD4*, (Stuczyńska y col., 2018), el ritmo circadiano (*ARNTL*, (Murphy y col., 2014), la síntesis del grupo hemo (*FECH*, (Ferreira, 1999), así como la acumulación de especies reactivas del oxígeno (*AKT3*, (Polytarchou y col., 2020). Los genes que participan en el metabolismo lipídico pueden tener un rol clave en la utilización de los mismos como sustratos energéticos durante el ejercicio, sobre todo en aquellos con un mayor componente aeróbico (Littiere y col., 2020). Por otro lado, el ejercicio intenso induce la formación excesiva de especies reactivas

del oxígeno al aumentar el consumo de oxígeno (Boffi, 2006). Es por esto por lo que el estrés oxidativo es un aspecto ampliamente abordado en la fisiología del ejercicio, por el potencial daño que el mismo desencadena sobre las moléculas del organismo (Littiere y col., 2020). Cosgrove y col. (2020) encontraron una huella de selección en el ECA 16 de caballos Pura Raza Árabe, que contenía el gen *GPX1*, el cual codifica la enzima glutatión peroxidasa, con función antioxidante. El aumento de la actividad de esta enzima, por ejemplo, mediante la suplementación con selenio, puede reducir el daño oxidativo de los tejidos y mejorar potencialmente el rendimiento (White y col., 2016). En general se considera que las adaptaciones fisiológicas de un animal para cumplir con las demandas metabólicas exigidas durante el ejercicio son fundamentales para su correcto desempeño atlético (Littiere y col., 2020).

Varios de los genes encontrado están ligados al funcionamiento y desarrollo del sistema músculo esquelético. Entre ellos se encuentran los genes *PDZRN3* y *PPP4R2*; el primero interviene en el desarrollo morfológico embrionario y la diferenciación de los mioblastos (Ko y col., 2006), mientras que la función de *PPP4R2* se relaciona con la diferenciación y supervivencia de las neuronas motoras (Bosio y col., 2012). En un GWAS realizado por Shin y col. (2015) en caballos de carrera se encontró que ambos genes presentaron polimorfismos asociados con el rendimiento deportivo (tiempo de carrera). Por otro lado, el gen *SH3RF2*, el cual codifica una ubiquitina-proteína ligasa E3 que regula los procesos de apoptosis en el tejido muscular (efecto anti-apoptótico), cumple un rol preponderante en el desarrollo de las fibras y el crecimiento del tejido. Un estudio realizado con caballos Pura Raza Árabe encontró que los niveles de expresión de este gen diferían de acuerdo con el nivel de entrenamiento de los animales (Ropka-Molik y col., 2018). Más aún, se analizó una mutación puntual del gen

(*SH3RF2:c.796T>C*) y su asociación con los resultados en las carreras, encontrando que el genotipo C/C afectaba negativamente la probabilidad de ganar carreras, el número de carreras en las que participó el caballo y el valor económico de los premios en la raza Árabe. Por lo tanto, el rol de este gen sobre la regeneración de las fibras musculares podría tener un impacto considerable en la adaptación al ejercicio, incluso también en caballos que participen en otras disciplinas como el polo.

Los genes *CNTNAP5*, *PCDH7* y *CTNND*, estuvieron asociados a diversos caracteres conformacionales en otras especies domésticas como bovinos, caprinos, ovinos y cerdos, como por ejemplo la altura y longitud del cuerpo (Vanvanhossou y col., 2020), el diámetro bicostal (Rahmatalla y col., 2018), la conformación de los miembros (Abdalla y col., 2021), el tamaño corporal (Liu y col., 2021) y el crecimiento (Zhang y col., 2016; Easa y col., 2022). Los genes *UNC45B*, *COL23A*, *LTBP2* y *CLEC3B* intervienen de distintas formas durante la formación y desarrollo del músculo y el sistema osteoarticular (Goessler y col., 2005; Qanbari y col., 2011; Buzanskas y col., 2014; Myhre y col., 2014). Estos genes podrían tener injerencia en el crecimiento y el tamaño corporal, aunque no se han encontrado reportados en la especie equina.

El gen *ATF1* y el *SNX10* intervienen en la formación y función de los osteoclastos, los cuales participan en la resorción y homeostasis del hueso (Sultana y col., 2020; Liu y col., 2022). Los genes *UST* y *CHST11* codifican enzimas que participan en la sulfatación de la condroitina, un proteoglicano importante para el cartílago articular. Las alteraciones osteoarticulares pueden condicionar y limitar la vida deportiva, por lo que constituyen factores cruciales para el bienestar y el rendimiento deportivo de los equinos (Littiere y col., 2020). Estas alteraciones son más o menos frecuentes de acuerdo con la disciplina hípica y el tipo de ejercicio

que se practica, ya que cada deporte tiene sus características particulares que hace que las exigencias sobre el sistema osteoarticular sean diferentes y los animales se vean expuestos a diversos factores de riesgo que los predispongan a lesiones específicas (Harvey y col., 2012; Inness y Morgan, 2015; Parkes y Witte, 2015).

El rendimiento atlético en los equinos es un carácter complejo y que se encuentra regulado por varios genes (Littiere y col., 2020). La identificación de huellas de selección en el Polo Argentino aporta información valiosa que contribuye a comprender la genética de los caracteres que intervienen en el rendimiento deportivo. El principal objetivo a futuro será poder aplicar directamente estos conocimientos en el desarrollo de herramientas genéticas para la mejora en la producción. Se requerirán futuros estudios para determinar la magnitud de dichas diferencias desde un punto de vista fenotípico y validar la funcionalidad de los genes encontrados en la raza.

5. Material Suplementario

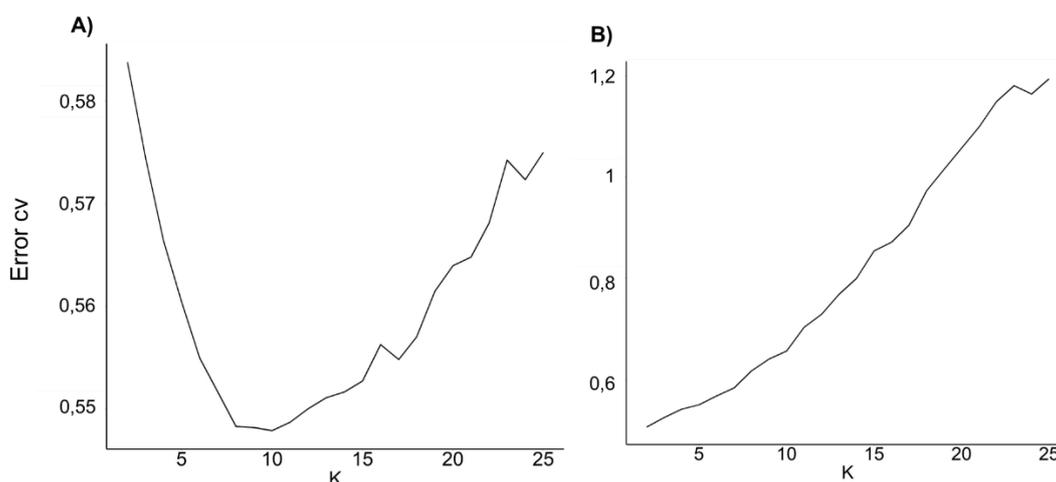


Figura S3.1. Error de validación cruzada (error cv) de los diferentes K en el análisis admixture, entre todas las razas (A) y entre Polo Argentino y SPC (B). El valor de K óptimo fue seleccionado como aquel que presentó el error cv más bajo entre los K analizados.

Tabla S3.1. Análisis admixture. Promedio de los porcentajes de asignación a cada población para todas las razas (K-10) (A) y para Polo Argentino y SPC (K-2) (B).

A)

	n	K-1	K-2	K-3	K-4	K-5	K-6	K-7	K-8	K-9	K-10
Pura Raza Árabe	45	0,03	0,90	0,02	0,02	0,00	0,00	0,01	0,00	0,00	0,02
Belga	22	0,00	0,00	0,01	0,01	0,00	0,00	0,01	0,96	0,00	0,00
Burguete	10	0,00	0,00	0,00	0,00	0,02	0,05	0,01	0,77	0,03	0,12
Jaca Navarra	10	0,02	0,00	0,00	0,00	0,02	0,06	0,01	0,72	0,04	0,14
Criollo	5	0,03	0,06	0,02	0,08	0,23	0,11	0,07	0,14	0,06	0,20
Peruano	67	0,02	0,01	0,02	0,03	0,82	0,02	0,02	0,02	0,02	0,03
Lusitano	24	0,02	0,02	0,01	0,79	0,01	0,11	0,01	0,01	0,00	0,01
Menorquín	13	0,01	0,01	0,00	0,01	0,01	0,01	0,00	0,02	0,00	0,93
Pura Raza Español	36	0,03	0,01	0,00	0,08	0,01	0,81	0,00	0,02	0,00	0,03
Cuarto de Milla	72	0,14	0,02	0,69	0,05	0,01	0,01	0,03	0,04	0,01	0,01
SPC	36	0,96	0,00	0,03	0,00	0,00	0,00	0,00	0,00	0,00	0,00
SPC – tipo Polo	8	0,97	0,01	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,01
Polo Argentino	60	0,92	0,01	0,01	0,00	0,00	0,01	0,00	0,01	0,00	0,03
Trotador	10	0,25	0,10	0,14	0,01	0,04	0,06	0,04	0,11	0,03	0,23
Welsh	44	0,03	0,04	0,04	0,05	0,01	0,01	0,68	0,10	0,04	0,01
Shetland	24	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,99	0,00

B)

	n	K-1	K-2
SPC	36	0,12	0,88
SPC - Tipo Polo	8	0,44	0,56
Polo Argentino	60	0,82	0,18

En color rojo se señala el mayor porcentaje de cada población asignado a cada raza.

Tabla S3.2. Regiones genómicas con mayor diferenciación genética entre Polo Argentino y SPC, de acuerdo con el índice F_{ST} .

Región	ECA	Inicio	Fin	Valor-q promedio	Región	ECA	Inicio	Fin	Valor-q promedio
1	1	17398962	17498962	0,0001	133	12	1076531	1176531	0,0000
2	1	19152673	19252673	0,0443	134	12	7982345	8082345	0,0432
3	1	19455842	19555842	0,0376	135	14	3351121	3451121	0,0370
4	1	28801844	28901844	0,0315	136	14	27956526	28056526	0,0000
5	1	36772376	36872376	0,0000	137	14	30923133	31023133	0,0119
6	1	57112433	57212433	0,0000	138	14	31324156	31424156	0,0042
7	1	66538044	66638044	0,0378	139	14	48091649	48191649	0,0350
8	1	75672489	75772489	0,0025	140	14	48204900	48304900	0,0160
9	1	96314222	96414222	0,0432	141	14	66212658	66312658	0,0260
10	1	96523983	96623983	0,0443	142	14	68923298	69023298	0,0048
11	1	114082029	114182029	0,0295	143	14	72056277	72203990	0,0354
12	1	150095001	150195001	0,0011	144	14	72323680	72423719	0,0342
13	1	161502161	161602161	0,0127	145	14	81449389	81549389	0,0095
14	1	174006048	174106048	0,0039	146	14	81549881	81649881	0,0370
15	1	175234869	175334869	0,0282	147	15	26846500	26946500	0,0418
16	1	175342146	175442146	0,0095	148	15	28990026	29090026	0,0015
17	1	175996094	176096173	0,0076	149	15	29244238	29344238	0,0001
18	1	176264941	176364941	0,0065	150	15	50632122	50732122	0,0269
19	1	176931720	177031720	0,0146	151	15	52614101	52714101	0,0472
20	1	177622634	177722634	0,0194	152	15	66995471	67095471	0,0093
21	1	181109105	181209105	0,0000	153	16	14146240	14246240	0,0110
22	2	406433	506433	0,0285	154	16	18359167	18528058	0,0274
23	2	14418041	14518041	0,0000	155	16	18831645	18931645	0,0297
24	2	16227335	16327335	0,0258	156	16	29985890	30085890	0,0025
25	2	18792233	18892233	0,0443	157	16	43045895	43145895	0,0281
26	2	27810816	27910816	0,0123	158	16	47617846	47717846	0,0281
27	2	35442273	35542273	0,0032	159	16	53932510	54032510	0,0041
28	2	46617965	46717965	0,0001	160	16	67543113	67643113	0,0206
29	2	54650173	54750173	0,0491	161	16	70089677	70193577	0,0267
30	2	62862130	62962130	0,0000	162	17	1326819	1426819	0,0046
31	2	65376155	65476155	0,0039	163	17	25052154	25152154	0,0336
32	2	67784169	67884169	0,0184	164	17	25479322	25579322	0,0002
33	2	72685326	72785326	0,0443	165	17	28456048	28556048	0,0472
34	2	90410089	90510089	0,0447	166	17	55843564	55943564	0,0039
35	2	100252846	100352846	0,0000	167	17	59807916	59907916	0,0000
36	2	102869745	102969745	0,0146	168	17	61331160	61431160	0,0004
37	2	113322772	113422772	0,0208	169	17	79758605	79858605	0,0342
38	2	119636894	119736894	0,0184	170	18	6237540	6337540	0,0035
39	3	7312314	7412314	0,0258	171	18	12770807	12870807	0,0276

40	3	11839910	11939910	0,0321	172	18	15063196	15163196	0,0461
41	3	31200432	31300432	0,0019	173	18	16511718	16611718	0,0291
42	3	31823055	31923055	0,0000	174	18	20142818	20242818	0,0423
43	3	70795695	70895695	0,0009	175	18	72752756	72852756	0,0130
44	3	79677377	79777377	0,0495	176	19	10828431	10928431	0,0260
45	3	96768824	96868824	0,0072	177	19	13743876	13843876	0,0164
46	3	102046154	102146154	0,0048	178	19	40424216	40524216	0,0260
47	3	108100647	108200647	0,0299	179	19	42328996	42428996	0,0213
48	3	108685883	108785883	0,0000	180	19	45896771	45996771	0,0021
49	3	113254843	113354843	0,0472	181	19	57910343	58010343	0,0409
50	4	14021152	14121152	0,0184	182	20	11881254	11985456	0,0319
51	4	32758114	32858114	0,0016	183	20	20846774	20946774	0,0092
52	4	33416995	33516995	0,0367	184	20	21660337	21760337	0,0130
53	4	57629728	57729728	0,0449	185	20	28842781	28942781	0,0095
54	4	62672658	62772658	0,0501	186	20	30632452	30732452	0,0267
55	4	77159807	77259807	0,0093	187	20	30956494	31056494	0,0370
56	4	85148739	85248739	0,0000	188	20	33460205	33560205	0,0096
57	4	85675918	85776103	0,0294	189	20	33537124	33637124	0,0203
58	4	87616481	87716481	0,0000	190	20	34633633	34734652	0,0177
59	5	2423513	2523513	0,0025	191	20	34897368	34997368	0,0057
60	5	2493134	2593134	0,0025	192	20	34997804	35161406	0,0057
61	5	2569779	2669779	0,0025	193	20	35126300	35234791	0,0042
62	5	2651643	2751643	0,0047	194	20	36362175	36462175	0,0285
63	5	2783054	2883054	0,0035	195	20	38326845	38426845	0,0260
64	5	2960287	3060287	0,0232	196	20	38378957	38519492	0,0251
65	5	14490442	14590442	0,0006	197	20	38482274	38582274	0,0271
66	5	28346660	28446660	0,0093	198	20	39443954	39543954	0,0127
67	5	51029460	51129460	0,0037	199	21	2330591	2430591	0,0000
68	5	70026848	70126848	0,0210	200	21	5641381	5741381	0,0169
69	5	81788243	81888243	0,0285	201	21	39657288	39757288	0,0373
70	5	87185581	87285581	0,0472	202	21	49399182	49499182	0,0315
71	5	91528324	91628324	0,0001	203	22	1446448	1546448	0,0370
72	5	91771629	91871629	0,0093	204	22	9047534	9147534	0,0304
73	5	93267204	93367424	0,0418	205	22	30712165	30812165	0,0014
74	6	6667788	6767788	0,0360	206	22	30991782	31091782	0,0168
75	6	6724154	6824154	0,0472	207	22	33948294	34048294	0,0261
76	6	7223130	7323130	0,0295	208	22	37888805	37988805	0,0037
77	6	7603709	7703709	0,0071	209	22	39436049	39536049	0,0292
78	6	7927926	8027926	0,0492	210	23	8863490	8963490	0,0074
79	6	9623184	9723184	0,0472	211	23	9878006	9978006	0,0276
80	6	9776838	9876838	0,0472	212	23	12267174	12367174	0,0479
81	6	40260819	40399034	0,0246	213	23	26858714	26958714	0,0252
82	6	49235317	49335317	0,0123	214	23	28837886	28937886	0,0472

83	6	53620372	53720372	0,0130	215	23	40460573	40560573	0,0269
84	6	59422319	59522319	0,0227	216	23	53906527	54077526	0,0184
85	6	62226849	62326849	0,0295	217	23	54336962	54436962	0,0451
86	6	69114378	69214378	0,0482	218	23	54388878	54488878	0,0448
87	6	74064503	74164503	0,0350	219	23	54643757	54768534	0,0185
88	6	74917530	75017530	0,0269	220	23	54747926	54847926	0,0127
89	6	76606696	76740048	0,0373	221	24	6137549	6237549	0,0472
90	7	54369009	54469009	0,0376	222	24	17051900	17151900	0,0032
91	7	61008996	61108996	0,0067	223	24	19950796	20050796	0,0127
92	7	61227461	61327461	0,0000	224	24	30084040	30184040	0,0447
93	7	76718901	76818901	0,0000	225	24	46792486	46892486	0,0067
94	7	84284363	84384363	0,0472	226	25	8935452	9035452	0,0045
95	7	86433920	86533920	0,0038	227	25	11817842	11917842	0,0003
96	7	89936375	90036375	0,0436	228	25	30048239	30148239	0,0015
97	7	93116003	93216003	0,0267	229	25	35845053	35945053	0,0418
98	7	94657089	94757089	0,0493	230	26	2508918	2608918	0,0472
99	7	10008211	100108211	0,0370	231	26	16941244	17041244	0,0175
100	7	100062081	100162081	0,0269	232	26	29720617	29820617	0,0289
101	7	100283180	100383180	0,0406	233	26	36354508	36454508	0,0000
102	8	8520728	8620728	0,0009	234	26	40860502	40960502	0,0360
103	8	50147053	50247053	0,0000	235	27	10343983	10443983	0,0260
104	8	62629159	62729159	0,0376	236	27	15276488	15376488	0,0097
105	8	69727032	69827032	0,0000	237	27	33331769	33431769	0,0002
106	8	78360899	78460899	0,0095	238	28	28967194	29067194	0,0315
107	8	80428515	80528515	0,0315	239	28	29028863	29128863	0,0016
108	8	87764839	87864839	0,0449	240	28	29087174	29187174	0,0429
109	8	91968039	92068039	0,0338	241	28	30207728	30307728	0,0213
110	8	94449865	94549865	0,0432	242	28	34070560	34170560	0,0001
111	8	95700307	95800307	0,0425	243	28	34826505	34926660	0,0394
112	8	96636858	96736858	0,0001	244	28	37272053	37372592	0,0146
113	9	14664600	14764600	0,0168	245	29	5748972	5848972	0,0184
114	9	15078016	15181732	0,0225	246	29	13683130	13783130	0,0050
115	9	15321007	15421007	0,0342	247	30	169588	269588	0,0184
116	9	46720050	46820050	0,0095	248	30	1281504	1381504	0,0282
117	9	83079383	83179383	0,0436	249	30	1334992	1434992	0,0282
118	10	22939603	23039603	0,0152	250	30	1398757	1542276	0,0003
119	10	28100305	28200305	0,0074	251	30	1551842	1651842	0,0350
120	10	44959998	45059998	0,0285	252	30	1715866	1815866	0,0216
121	10	50627034	50727034	0,0267	253	30	1825424	1925424	0,0305
122	10	51348579	51448579	0,0016	254	30	2413088	2513088	0,0210
123	10	52328696	52428696	0,0011	255	30	3773672	3873965	0,0130
124	10	82575251	82675251	0,0370	256	30	3896072	3996072	0,0269
125	10	84249473	84349473	0,0354	257	30	3947088	4047088	0,0297

126	11	5456808	5556808	0,0370	258	30	7705960	7805960	0,0023
127	11	12011934	12111934	0,0100	259	30	19042761	19142761	0,0024
128	11	16702634	16802634	0,0002	260	30	19811461	19911461	0,0184
129	11	33620635	33720635	0,0039	261	30	22815130	22915130	0,0096
130	11	38042380	38142380	0,0096	262	30	27602870	27702870	0,0006
131	11	41072737	41172737	0,0438	263	31	17786493	17886493	0,0043
132	11	48394840	48494840	0,0012					

Tabla S3.3 Genes contenidos dentro de las regiones genómicas con alta diferenciación genética entre Polo Argentino y SPC. Los genes marcados en letra negra constituyen genes candidatos para caracteres de interés en la raza, de acuerdo con la revisión bibliográfica realizada.

Ensembl ID	ECA	Inicio	Fin	Gen	Nombre del gen
ENSECAG00000024113	1	17246345	17532536	<i>ABLIM1</i>	Actin binding LIM protein 1
ENSECAG00000012950	1	18958087	19296736	<i>VT11A</i>	Vesicle transport through interaction with t-SNAREs 1A
ENSECAG00000014742	1	19495382	19553229	<i>GPAM</i>	Glycerol-3-phosphate acyltransferase 1, mitochondrial
ENSECAG00000022888	1	28628298	28883830	<i>BTRC</i>	Beta-transducin repeat containing E3 ubiquitin protein ligase
ENSECAG00000000077	1	36793230	37029591	<i>CPEB3</i>	Cytoplasmic polyadenylation element binding protein 3
ENSECAG00000007190	1	57158901	57164014	<i>HNRNPH3</i>	Heterogeneous nuclear ribonucleoprotein H3
ENSECAG00000021661	1	75693237	75791736	<i>MTR</i>	Methionine synthase
ENSECAG00000024912	1	150139309	150300428	<i>FSIP1</i>	Fibrous sheath-interacting protein 1
ENSECAG00000002597	1	175151384	175591666	<i>SLC25A21</i>	Solute carrier family 25 member 21
ENSECAG00000010115	1	175970916	176160771	<i>TTC6</i>	Probable UDP-N-acetylglucosamine-peptide N-acetylglucosaminyltransferase SPINDLY
ENSECAG00000014310	1	176957271	177056623	<i>SEC23A</i>	Protein transport protein SEC23
ENSECAG00000012021	2	14450864	14620335	<i>ST3GAL3</i>	ST3 beta-galactoside alpha-2,3-sialyltransferase 3
ENSECAG00000013366	2	18559232	18877920	<i>MACF1</i>	Microtubule actin crosslinking factor 1
ENSECAG00000017176	2	27859290	28030594	<i>EYA3</i>	Eyes absent homolog
ENSECAG000000000771	2	46609438	46673741	<i>TP73</i>	Tumor protein p73
ENSECAG00000024339	2	113097371	113387502	<i>CAMK2D</i>	Calcium/calmodulin-dependent protein kinase
ENSECAG00000010911	2	119670631	119694557	<i>INTS12</i>	Integrator complex subunit 12
ENSECAG00000021162	3	30849654	31850887	<i>CDH13</i>	Cadherin-13
ENSECAG00000000530	3	70757826	71114534	<i>EPHA5</i>	Receptor protein-tyrosine kinase
ENSECAG00000000599	3	96815625	97212288	<i>PCDH7</i>	Protocadherin 7
ENSECAG00000010537	4	13981090	14374324	<i>HECW1</i>	HECT-type E3 ubiquitin transferase
ENSECAG00000013745	4	57618965	57703870	<i>SNX10</i>	Sorting nexin 10
ENSECAG00000021914	4	62299006	62805826	<i>PDE1C</i>	Phosphodiesterase
ENSECAG00000002703	4	76883151	77335350	<i>KCND2</i>	Potassium voltage-gated channel subfamily D member 2
ENSECAG00000013579	4	85165569	85280962	<i>COPG2</i>	Coatomer subunit gamma
ENSECAG00000024017	4	85165569	85289548	<i>TSGA13</i>	Testis specific 13
ENSECAG00000024078	4	85670418	85960277	<i>MKLN1</i>	Muskelin 1
ENSECAG00000004434	4	87471964	88207600	<i>EXOC4</i>	Exocyst complex component Sec8

ENSECAG00000014082	5	2506321	2740264	<i>SRGAP2</i>	SLIT-ROBO Rho GTPase activating protein 2
ENSECAG00000019451	5	2779506	2845391	<i>RASSF5</i>	Ras association domain family member 5
ENSECAG00000015180	5	81836976	81879187	<i>NEXN</i>	Nexilin F-actin binding protein
ENSECAG00000014027	5	86621481	87495679	<i>NEGR1</i>	Neuronal growth regulator 1
ENSECAG00000011003	5	91570529	91961725	<i>PDE4B</i>	Phosphodiesterase
ENSECAG00000012582	6	7962229	8016648	<i>PLCD4</i>	Phosphoinositide phospholipase C
ENSECAG00000038335	6	40329625	40329831	<i>SMIM10L1</i>	Small integral membrane protein 10 like 1
ENSECAG00000036288	6	40344710	40348201	<i>TAS2R42</i>	Taste receptor type 2
ENSECAG00000023690	6	53491416	53690283	<i>STK38L</i>	Serine/threonine kinase 38 like
ENSECAG00000016445	6	69153819	69191499	<i>ATF1</i>	Activating transcription factor 1
ENSECAG00000008933	6	74959219	74972534	<i>GLS2</i>	Glutaminase
ENSECAG00000030065	7	61048628	61062565	<i>FZD4</i>	Frizzled-4
ENSECAG00000012118	7	84304636	84341389	<i>ARNTL</i>	Aryl hydrocarbon receptor nuclear translocator-like protein 1
ENSECAG00000017807	7	89626174	90023454	<i>NAV2</i>	Neuron navigator 2
ENSECAG00000022068	7	99811304	100064824	<i>RCN1</i>	Reticulocalbin 1
ENSECAG00000022308	7	100303122	100348001	<i>WT1</i>	WT1 transcription factor
ENSECAG00000012137	8	78397158	78430399	<i>FECH</i>	Ferrochelatase
ENSECAG00000002784	8	87810676	88108730	<i>CCDC102B</i>	Coiled-coil domain containing 102B
ENSECAG00000016377	8	91970419	92115390	<i>FBXO15</i>	F-box protein 15
ENSECAG00000005459	8	94388158	94501052	<i>ZNF236</i>	Zinc finger protein 236
ENSECAG00000008636	9	14941295	15249938	<i>EYA1</i>	Eyes absent homolog
ENSECAG00000019852	9	46664879	47441220	<i>VPS13B</i>	Vacuolar protein sorting 13 homolog B
ENSECAG00000032503	10	22727442	23022867	<i>LOC111775446</i>	Zinc finger protein 677-like
ENSECAG00000015552	11	33631559	33673667	<i>GDPD1</i>	Glycerophosphodiester phosphodiesterase domain containing 1
ENSECAG0000001034	11	38087233	38115481	<i>UNC45B</i>	Protein unc-45 homolog B
ENSECAG00000024961	11	48428126	48447217	<i>XAF1</i>	XIAP associated factor 1
ENSECAG00000016829	12	1000405	1167325	<i>ABTB2</i>	Ankyrin repeat and BTB domain containing 2
ENSECAG00000041924	14	3267668	3582764	<i>COL23A1</i>	Collagen type XXIII alpha 1 chain
ENSECAG00000024467	14	30950659	31073558	<i>SH3RF2</i>	SH3 domain containing ring finger 2
ENSECAG00000013676	14	71927655	72174893	<i>CAST</i>	Calpain inhibitor
ENSECAG00000020422	14	72360726	72405039	<i>PCSK1</i>	Neuroendocrine convertase 1
ENSECAG00000024042	15	26648772	27462993	<i>LRRTM4</i>	Leucine rich repeat transmembrane neuronal 4
ENSECAG00000017150	15	28905424	29056711	<i>TACR1</i>	NK-1 receptor
ENSECAG00000040062	15	52502287	52677226	<i>KCNK12</i>	Potassium two pore domain channel subfamily K member 12
ENSECAG00000015422	16	14185353	14210518	<i>CRBN</i>	Protein cereblon
ENSECAG00000014864	16	18387021	18613380	<i>PDZRN3</i>	PDZ domain containing ring finger 3
ENSECAG00000000689	16	18864765	18911420	<i>PPP4R2</i>	Protein phosphatase 4 regulatory subunit 2
ENSECAG00000032582	16	29762287	30541116	<i>FHIT</i>	Bis(5'-adenosyl)-triphosphatase
ENSECAG00000006251	16	43086077	43096083	<i>CLEC3B</i>	C-type lectin domain family 3 member B
ENSECAG00000023266	16	53702077	53987335	<i>OSBPL10</i>	Oxysterol-binding protein
ENSECAG00000021178	16	70011758	70192968	<i>UBA5</i>	UFM1-activating enzyme

ENSECAG00000016662	17	79641189	79810330	MCF2L	MCF.2 cell line derived transforming sequence like
ENSECAG00000017885	18	5864335	6638949	CNTNAP5	Contactin associated protein family member 5
ENSECAG0000000161	18	14774893	16004989	DPP10	Dipeptidyl peptidase like 10
ENSECAG00000021657	18	72794287	72830055	SF3B1	Splicing factor 3b subunit 1
ENSECAG00000009903	19	10855276	10880573	SERPINI1	Serpin family I member 1
ENSECAG00000018307	19	13686949	14067089	TNIK	TRAF2 and NCK interacting kinase
ENSECAG00000010725	19	40411393	40759469	STXBP5L	Syntaxin binding protein 5 like
ENSECAG00000015662	19	57897040	58258243	CMSS1	Cms1 ribosomal small subunit homolog
ENSECAG00000016973	19	57956981	58250918	FILIP1L	Filamin A interacting protein 1 like
ENSECAG00000004794	20	34551922	34733636	SLC39A7	Solute carrier family 39 member 7
ENSECAG00000007328	20	34856435	35017180	ITPR3	Inositol 1,4,5-trisphosphate receptor type 3
ENSECAG00000009417	20	34945196	34956988	IP6K3	Kinase
ENSECAG00000017336	20	38442723	38751281	ZFAND3	Zinc finger AN1-type containing 3
ENSECAG00000012713	21	2374146	2391294	CHERP	Calcium homeostasis endoplasmic reticulum protein
ENSECAG00000010901	21	49170984	50051293	CTNND2	Catenin delta 2
ENSECAG00000000170	22	8656455	10521112	MACROD2	Mono-ADP ribosylhydrolase 2
ENSECAG00000005925	22	33965752	34005068	MYBL2	MYB proto-oncogene like 2
ENSECAG00000014676	22	37897424	38079142	PREX1	Phosphatidylinositol-3,4,5-trisphosphate dependent Rac exchange factor 1
ENSECAG00000019772	22	39442194	39520318	PTPN1	Protein-tyrosine-phosphatase
ENSECAG00000009864	23	28744226	29126165	PTPRD	Protein-tyrosine-phosphatase
ENSECAG00000040430	23	40477341	40610953	LOC100065392	S-methyl-5'-thioadenosine phosphorylase
ENSECAG00000012044	23	54009915	54069828	PTPDC1	Protein tyrosine phosphatase domain containing 1
ENSECAG00000017930	23	54316105	54398935	PHF2	PHD finger protein 2
ENSECAG00000011076	23	54409688	54700087	FAM120A	Family with sequence similarity 120A
ENSECAG00000019990	23	54581208	54695376	WNK2	Non-specific serine/threonine protein kinase
ENSECAG00000018631	24	19962114	20126585	LTBP2	Latent transforming growth factor beta binding protein 2
ENSECAG00000016942	24	46839572	46847449	NUDT14	Nudix hydrolase 14
ENSECAG00000032033	25	8924474	8985510	LOC100062753	Cyclicin-2
ENSECAG00000016804	25	30050681	30254088	PBX3	PBX homeobox 3
ENSECAG00000008057	26	40846456	41121670	TSPEAR	Thrombospondin type laminin G domain and EAR repeats
ENSECAG00000023430	26	40866225	40924064	TRPM2	Transient receptor potential cation channel subfamily M member 2
ENSECAG00000019904	28	28881986	29132639	CHST11	Carbohydrate sulfotransferase
ENSECAG00000002145	28	30221475	30290960	NUAK1	NUAK family kinase 1
ENSECAG00000017245	28	37299581	37372179	TAB1	Synaptogyrin 1
ENSECAG00000020753	29	5714325	5858785	NRP1	Neuropilin
ENSECAG00000019722	30	1460985	1774820	AKT3	Non-specific serine/threonine protein kinase
ENSECAG00000008477	30	1780137	2002320	SDCCAG8	SHH signaling and ciliogenesis regulator SDCCAG8
ENSECAG00000009422	30	3595984	4081736	RGS7	Regulator of G protein signaling 7
ENSECAG00000021865	31	17581381	17931722	UST	Uronyl 2-sulfotransferase

Tabla S3.4. Categorías GO (Gene Ontology) de los genes encontrados en el análisis funcional de las regiones de alta diferenciación genética entre Polo Argentino y SPC, extraído de la plataforma DAVID.

Categoría	Término	n	%	Valor-p	Genes	Enriquecimiento	Corrección Bonferroni
GOTERM_BP_DIRECT	GO:0035556~intracellular signal transduction	10	8,5	0,000403	ENSECAG00000009422, ENSECAG00000016662, ENSECAG00000019990, ENSECAG00000019451, ENSECAG00000023690, ENSECAG00000002145, ENSECAG00000012582, ENSECAG00000019722, ENSECAG00000014676, ENSECAG00000018307	4,4	0,16434219
GOTERM_BP_DIRECT	GO:0031532~actin cytoskeleton reorganization	4	3,4	0,005092	ENSECAG00000019772, ENSECAG000000024078, ENSECAG00000018307, ENSECAG000000020753	11,4	0,89685205
GOTERM_BP_DIRECT	GO:0006470~protein dephosphorylation	4	3,4	0,007384	ENSECAG00000009864, ENSECAG00000000689, ENSECAG00000017245, ENSECAG000000022888	9,9	0,96305307
GOTERM_BP_DIRECT	GO:0051209~release of sequestered calcium ion into cytosol	3	2,5	0,008153	ENSECAG00000012713, ENSECAG00000023430, ENSECAG00000007328	21,7	0,97382847
GOTERM_BP_DIRECT	GO:0016576~histone dephosphorylation	2	1,7	0,024644	ENSECAG00000017176, ENSECAG00000008636	79,6	0,99998495
GOTERM_BP_DIRECT	GO:0042753~positive regulation of circadian rhythm	2	1,7	0,036740	ENSECAG00000012118, ENSECAG000000022888	53,1	0,99999994
GOTERM_BP_DIRECT	GO:0006468~protein phosphorylation	4	3,4	0,058706	ENSECAG00000019990, ENSECAG00000002145, ENSECAG00000009417, ENSECAG00000018307	4,5	1
GOTERM_BP_DIRECT	GO:0008285~negative regulation of cell proliferation	5	4,2	0,060392	ENSECAG00000012713, ENSECAG00000021162, ENSECAG00000019990, ENSECAG00000019451, ENSECAG000000022308	3,4	1
GOTERM_BP_DIRECT	GO:2000772~regulation of cellular senescence	2	1,7	0,066334	ENSECAG00000012118, ENSECAG00000002145	28,9	1
GOTERM_BP_DIRECT	GO:0071577~zinc II ion transmembrane transport	2	1,7	0,072144	ENSECAG00000023430, ENSECAG000000004794	26,5	1
GOTERM_BP_DIRECT	GO:0060828~regulation of canonical Wnt signaling pathway	2	1,7	0,089359	ENSECAG00000010901, ENSECAG000000022888	21,2	1
GOTERM_BP_DIRECT	GO:1990830~cellular response to leukemia inhibitory factor	3	2,5	0,092352	ENSECAG00000030065, ENSECAG00000013745, ENSECAG00000005925	5,8	1
GOTERM_BP_DIRECT	GO:0098609~cell-cell adhesion	3	2,5	0,092352	ENSECAG00000010901, ENSECAG00000021162, ENSECAG00000014027	5,8	1
GOTERM_BP_DIRECT	GO:0007585~respiratory gaseous exchange	2	1,7	0,095027	ENSECAG00000016804, ENSECAG00000019904	19,9	1
GOTERM_MF_DIRECT	GO:0046872~metal ion binding	18	15,2	0,003398	ENSECAG00000010911, ENSECAG00000015422, ENSECAG00000011003, ENSECAG00000032503, ENSECAG00000016942, ENSECAG00000002703, ENSECAG000000021178, ENSECAG00000017176, ENSECAG000000000771,	2,1	0,42188013

					ENSECAG00000021914, ENSECAG00000022308, ENSECAG00000024113, ENSECAG00000024467, ENSECAG00000008636, ENSECAG00000020753		
GOTERM_MF_DIRECT	GO:0004725~protein tyrosine phosphatase activity	5	4,2	0,004095	ENSECAG00000019772, ENSECAG00000009864, ENSECAG00000017176, ENSECAG00000012044, ENSECAG00000008636	7,6	0,48345868
GOTERM_MF_DIRECT	GO:0005509~calcium ion binding	9	7,6	0,031902	ENSECAG00000022068, ENSECAG00000018631, ENSECAG00000023430, ENSECAG00000006251, ENSECAG00000021162, ENSECAG00000007328, ENSECAG000000012582, ENSECAG00000013366, ENSECAG00000000599	2,4	0,9945919
GOTERM_MF_DIRECT	GO:0047631~ADP-ribose diphosphatase activity	2	1,7	0,034019	ENSECAG00000023430, ENSECAG00000016942	57,3	0,99619866
GOTERM_MF_DIRECT	GO:0000981~RNA polymerase II transcription factor activity, sequence-specific DNA binding	11	9,3	0,055966	ENSECAG00000012118, ENSECAG00000016804, ENSECAG00000005925, ENSECAG00000016445, ENSECAG00000005459, ENSECAG00000000771, ENSECAG00000022308, ENSECAG00000032503	1,9	0,99990603
GOTERM_MF_DIRECT	GO:0000978~RNA polymerase II core promoter proximal region sequence-specific DNA binding	11	9,3	0,068619	ENSECAG00000012118, ENSECAG00000016804, ENSECAG00000005925, ENSECAG00000016445, ENSECAG00000005459, ENSECAG00000000771, ENSECAG00000022308, ENSECAG00000032503	1,8	0,9999893
GOTERM_MF_DIRECT	GO:0044325~ion channel binding	3	2,5	0,087713	ENSECAG00000012713, ENSECAG0000000161, ENSECAG00000011003	6,0	0,99999962
GOTERM_MF_DIRECT	GO:0005096~GTPase activator activity	4	3,4	0,098815	ENSECAG00000014082, ENSECAG00000010725, ENSECAG00000014310, ENSECAG00000020753	3,6	0,99999995

CAPÍTULO IV: Caracterización genética de marcadores asociados a caracteres fenotípicos con impacto en la aptitud deportiva en el caballo Polo Argentino.

1. Introducción

Los avances en el estudio de la genómica han despertado un gran interés por la identificación de variantes genéticas que expliquen rasgos de importancia para la cría equina. Tal es así que existen cada vez más estudios de asociación de genoma completo en esta especie que tienen por objetivo determinar a nivel molecular cuáles son los genes y variantes que determinan el fenotipo (Littiere y col., 2020). El conocimiento generado ha posibilitado el desarrollo de test genéticos los cuales pueden implementarse en los esquemas de selección para predecir la probabilidad de obtener ciertas características fenotípicas deseadas en la progenie con el objetivo de incrementar su rendimiento en la disciplina en la que se desempeñe. Hasta la actualidad, se han identificado varios marcadores genéticos relacionados con caracteres asociados al rendimiento deportivo en los equinos. Estos abarcan rasgos muy variados ligados tanto al comportamiento como la velocidad, el tipo de locomoción y el tamaño corporal (Bailey y Brooks, 2013; Brooks, 2021).

El caballo es una de las especies en las que el temperamento tiene una mayor importancia, al punto de ser un objetivo de selección clave. El temperamento refiere a la reacción de un animal ante una situación y es un componente estable de la personalidad del individuo a lo largo del tiempo. Si bien la personalidad de un caballo se considera en parte influenciada por las experiencias que haya atravesado en su vida, existe evidencia de que parte del temperamento está

regulado por la composición genética del individuo. Ciertos rasgos de comportamiento observables y cuantificables tienen una base genética y se han asociado con diferentes polimorfismos genéticos (Wickens y Brooks, 2020). Las variantes que se presentan en genes que codifican neurotransmisores u hormonas influyen en las diferencias individuales de rasgos de la personalidad (Inoue-Murayama, 2009). Las vías neurotransmisoras alteran el comportamiento modulando las propiedades de transporte, oxidación y unión al receptor de la cascada de señalización. Es por esto por lo que los genes involucrados en estas vías son candidatos para explicar las conductas observadas en los individuos (Wickens y Brooks, 2020).

La dopamina es uno de los neurotransmisores más importantes que puede modular el comportamiento en diferentes especies. Se conocen varios factores que regulan los niveles de dopamina en el individuo entre los que se encuentra el genotipo (Kim y col., 2022). En 2005, Momozawa y col. (2005b) estudiaron varios polimorfismos del gen del receptor D4 de la dopamina (*DRD4*) en el caballo SPC, logrando asociar una mutación puntual con características de comportamiento, a través de un análisis de encuestas. Su trabajo permitió el desarrollo de un marcador localizado en dicho gen cuya variabilidad alélica se asociaba significativamente a las puntuaciones de "curiosidad" y de "vigilancia" que los cuidadores caracterizaban de sus caballos. El término "curiosidad" se definió como el interés que presentó el animal por objetos novedosos y su acercamiento a ellos, por lo que podría considerarse un análogo de la "búsqueda de novedades". Mientras que el término "vigilancia" se refiere a la tendencia de un caballo a permanecer alerta y examinar cuidadosamente su entorno (Momozawa y col., 2005a). Hoy por hoy, este test está disponible comercialmente y es utilizado en programas de cría.

Por otro lado, Hill y col. (2010b) encontraron que ciertas mutaciones en el gen de la miostatina (*MSTN*) se relacionan con la mejor aptitud para diferentes distancias de carrera en caballos SPC. Este gen está altamente conservado y se expresa en el músculo esquelético maduro y en desarrollo (Bellinge y col., 2005). La miostatina es una proteína que pertenece a la super familia del factor de crecimiento transformante β (TGF β), y actúa modulando el desarrollo de la masa muscular. Se han descrito varios polimorfismos asociados a este gen, entre los que se destaca una inserción SINE (“short interspersed nuclear element”) en el promotor y dos SNP, g.66493737C>T y BIEC2-438999 (Chr18:g.67027762), los cuales se encuentran en desequilibrio de ligamiento y tienen gran influencia en la determinación de la distancia de carrera (corta, mediana o larga) en la cual un individuo SPC tendrá un rendimiento más adecuado (Hill y col., 2019). Por lo tanto, la determinación del genotipo permite elegir con mayor precisión las carreras que se adapten mejor al potencial genético de un individuo.

El tamaño corporal y las distintas medidas morfométricas tienen una amplia variabilidad en la especie equina, en función de la raza, el biotipo y la utilidad para la cual se seleccionen los animales. Estos caracteres han sido ampliamente estudiados y tienen un impacto notable en las diferentes disciplinas ecuestres, debido a que la conformación tiene un efecto directo sobre la biomecánica y la locomoción (Sole y col., 2014). Existen varias medidas que en su conjunto determinan el “tamaño” de un ejemplar, siendo la alzada a la cruz una de las más comunes de evaluar y que se describe en gran parte de los estándares raciales. La heredabilidad de este carácter varía dependiendo la raza analizada, pero en términos generales es moderada-alta; por ejemplo, en el caballo Pura Raza Español se cita una heredabilidad de 0,8 (Sánchez-Guerrero y col., 2016) y en el Lipizzano de 0,52 (Zechner y col., 2001). Desde el punto de vista genético, se han identificado varios loci que desempeñan un papel clave en la determinación del

tamaño corporal, en particular la alzada a la cruz. Makvandi-Nejad y col. (2012) encontraron 4 loci ubicados en el ECA 3, 6, 9 y 11 que explicaban más del 80% de la variación genética del tamaño corporal en un estudio realizado con 16 razas equinas. Metzger y col. (2018), por otro lado, encontraron otras variantes genéticas en el ECA 1, 6, 11 y 19 que explicaban el 72% del tamaño corporal en razas miniatura como el Pony Shetland. Estas variantes se encuentran próximas a genes con funciones relacionadas al crecimiento y desarrollo que permiten, parcialmente, predecir la altura potencial de un individuo.

Muchos de los estudios de asociación realizados para la detección de marcadores con funciones predictivas han sido realizados con caballos SPC, ya que es una de las razas de mayor importancia económica que se encuentra ampliamente distribuida por todo el mundo (Yoon y col., 2018). Si bien el caballo Polo Argentino es una raza joven, cuya base genética se origina mayormente en ejemplares SPC que reúnen características adecuadas para el polo (Laffaye, 2009), el proceso selectivo al cual ha sido sometida la raza podría generar cambios en las frecuencias génicas de marcadores específicos que determinan caracteres de importancia para esta disciplina, como ser su altura, capacidad atlética y temperamento.

Es por esto por lo que el objetivo de este capítulo fue describir la variabilidad de marcadores genéticos asociados a caracteres con influencia en la aptitud deportiva en caballos Polo Argentino. Adicionalmente, considerando su estrecha relación con el SPC, se buscó determinar si la selección para una disciplina deportiva diferente (turf vs. polo) ha modificado el perfil genético de los marcadores analizados en la raza Polo Argentino.

2. Materiales y métodos

2.1. Material animal y genotipado

Se utilizaron datos genotípicos de caballos Polo Argentino (n = 385) y SPC (n = 25), obtenidos utilizando el microarray Illumina Equine GGP array 70k (Illumina, Inc.) como fue descrito en la sección de Materiales y métodos de los Capítulos II y III de esta tesis. Los caballos Polo Argentino analizados pertenecen a distintos criadores del país, estando inscriptos todos ellos en el libro genealógico de la raza. Los caballos SPC analizados pertenecían a diferentes caballerizas de la provincia de Buenos Aires, habiéndose utilizado únicamente para competencias de turf.

Con el objetivo de evaluar las diferencias entre los individuos recién inscriptos en el libro genealógico y aquellos con mayor cantidad de ancestros registrados, se determinó el número de generaciones completas de los 385 caballos Polo Argentino analizados, utilizando los registros de pedigrí de la raza (~115.000). Posteriormente, se dividieron los animales en dos grupos divergentes: aquellos con cero generaciones completas (grupo Gen0, n = 45) y los que tenían al menos tres generaciones completas (grupo Gen3, n = 24). Paralelamente, se determinó cada uno de los animales que había participado en alguno de los torneos integrantes de la Triple Corona de alto Hándicap de Argentina (Campeonato Abierto de Tortugas Country Club, Campeonato Abierto de Hurlingham y Campeonato Argentino Abierto de Palermo), empleando los registros de la AACCP hasta la temporada 2022. Aquellos individuos que habían participado en al menos uno de estos torneos se incluyeron en el grupo Élite (n = 83), mientras que el resto constituyó el grupo No-élite (n = 302). Finalmente, los animales fueron también agrupados según habían recibido o no algún premio en un torneo o exposición organizado por la AACCP (grupos Premiado, n = 27 y No-premiado, n = 358). Cabe aclarar que se incluyeron en la evaluación aquellos premios recibidos

en las exposiciones (además de los recibidos en partidos) ya que las mismas incluyen una evaluación funcional y morfológica.

2.2. Marcadores genéticos

Del total de SNPs genotipados se filtraron aquellos asociados a caracteres fenotípicos de interés para la práctica del polo. En total se analizaron 9 SNPs asociados a 4 caracteres que pudieran influir directamente sobre el desempeño atlético: temperamento (curiosidad y vigilancia – asociados a polimorfismos del gen *DRD4*), velocidad (gen *MSTN* – speed gene) y diversos marcadores asociados al tamaño corporal descritos tanto en razas de alzada estándar como en ponis (Tabla 4.1).

2.3. Frecuencias genotípicas y alélicas de los marcadores

Se calcularon las frecuencias genotípicas y alélicas de cada marcador para cada raza y cada subgrupo. También se determinó si los locus se encontraban en equilibrio de Hardy Weinberg, utilizando el programa *PLINK* (Purcell y col., 2007). Las diferencias en las frecuencias genotípicas entre cada grupo fueron evaluadas mediante un test de proporciones Z con el paquete *Base* de R (R-Core-Team, 2023).

Adicionalmente se calculó el índice F_{ST} como medida de distancia genética para cada marcador entre la población Polo Argentino y SPC, utilizando el paquete de R *OutFLANK* (Whitlock y Lotterhos, 2015). Este paquete de R implementa una metodología que permite detectar los loci con valores atípicos según la distribución inferida de los valores de F_{ST} de la población, y su correspondiente valor-p. Para ello se utilizó el total de marcadores SNP del microarray, con una tasa de

genotipado de al menos 90% y se utilizaron las muestras que tenían más del 95% de los SNPs genotipados. Estos filtros se aplicaron con el software *PLINK* (Purcell y col., 2007), conformándose el conjunto de datos final por 66.265 SNPs y 408 muestras (383 Polo Argentino y 25 SPC).

Tabla 4.1. Marcadores SNP asociados a caracteres de interés en la especie equina genotipados con el Microarray Illumina Equine SNP70 array (Illumina, Inc.) (Neogen, 2023).

Caracteres	SNP	Gen	Alelos	Fenotipo asociado
Curiosidad y vigilancia (factor de riesgo) (Momozawa y col., 2005b)	Chr12:g.35662671	<i>DRD4</i>	A	Baja curiosidad y alta vigilancia (A/A y A/G)
			G	Alta curiosidad y baja vigilancia (G/G)
Distancia de carrera (Hill y col., 2019)	Chr18:g.67027762	<i>MSTN</i>	A	Distancia de carrera corta (potencia)
			G	Distancia de carrera larga (resistencia)
Tamaño corporal (Makvandi-Nejad y col., 2012)	Chr3:g.107374136	<i>LCORL</i>	T	Tamaño corporal pequeño
			C	Tamaño corporal grande
	Chr6:g.81481064	<i>HMG2</i>	C	Tamaño corporal pequeño
			T	Tamaño corporal grande
	Chr9:g.77656570	<i>ZFAT</i>	C	Tamaño corporal pequeño
			T	Tamaño corporal grande
Chr11:g.23334511	<i>LASP1</i>	G	Tamaño corporal pequeño	
		A	Tamaño corporal grande	
Tamaño corporal –ponis (Metzger y col., 2018)	Chr1:g.106168922	<i>ADAMTS17</i>	A	Tamaño corporal pequeño
			C	Tamaño corporal grande
	Chr19:g.30992750	<i>OSTN</i>	A	Tamaño corporal pequeño
			G	Tamaño corporal grande
Chr11:g.15494088	<i>GH1</i>	T	Tamaño corporal pequeño	
		C	Tamaño corporal grande	

El nombre del SNP se identifica por su ubicación en cromosoma y posición genómica dentro del mismo, de acuerdo con la versión del genoma equino EquCab3.0

3. Resultados

Del total de SNP analizados (n = 9), 6 resultaron polimórficos en el caballo Polo Argentino (Tabla 4.2). Estos marcadores corresponden a los asociados a los genes *DRD4* (Chr12:g.35662671), *MSTN* (Chr18:g.67027762), *LCORL*

(Chr3:g.107374136), *ZFAT* (Chr9:g.77656570), *LASP1* (Chr11:g.23334511) y *OSTN* (Chr19:g.30992750). En los caballos SPC los mismos marcadores presentaron variabilidad, con excepción del marcador Chr19:g.30992750, el cual se encontró fijado en la población. Dos de los SNPs analizados tuvieron valores significativamente elevados de F_{ST} , mostrando una alta diferenciación genética entre las dos poblaciones (Tabla 4.2).

En la Tabla 4.3 se muestran las frecuencias genotípicas (A) y alélicas (B) en los grupos analizados. La frecuencia del genotipo G/G del marcador Chr12:g.35662671 (gen *DRD4*) fue de 33% en el Polo Argentino, mientras que los genotipos A/A y A/G, los cuales se asocian a un fenotipo de baja curiosidad y alta vigilancia, fueron mayoritarios. Resultados similares fueron encontrados en la raza SPC, así como entre los distintos grupos analizados (Tabla 4.3A). Solo se encontraron diferencias significativas (valor- $p < 0,05$) entre los genotipos A/A de este marcador entre Polo Argentino y SPC (18% vs. 32%). La frecuencia del alelo A fue mayor en el SPC que en el Polo Argentino (52% vs. 43%), aunque esta diferencia no resultó estadísticamente significativa (Tabla 4.3B).

El marcador Chr18:g.67027762 (gen *MSTN*) se encontró en desequilibrio de Hardy Weinberg en la raza Polo Argentino (Tabla 4.2), por lo que podría encontrarse bajo selección. La frecuencia del alelo G fue mayor en esta raza respecto del SPC (69% vs. 40% valor- $p < 0,001$). La mayor diferencia entre ambas razas se encontró entre los genotipos homocigotas, ya que en el Polo Argentino se observó una menor proporción de homocigotas A/A (valor- $p < 0,001$) y una mayor proporción de homocigotas G/G (valor- $p < 0,05$). Estas diferencias concuerdan con el índice F_{ST} significativamente alto encontrado entre ambas razas ($F_{ST}=0,15$; valor- $p=0,01$). La mayoría de los ejemplares fueron heterocigotas en todos los grupos de la raza Polo Argentino analizados, excepto para los del

grupo Premiado, donde los ejemplares G/G fueron mayoría, y ninguno de ellos fue homocigota A/A. (Tabla 4.3A). Los caballos del grupo Gen3 mostraron una mayor frecuencia del alelo G y del genotipo G/G en comparación al grupo Gen0, aunque dichas diferencias no resultaron estadísticamente significativas (Tabla 4.3A y B).

Los marcadores Chr6:g.81481064 (*HMGA2*), Chr1:g.106168922 (*ADMTS17*) y Chr11:g.15494088 (*GH1*) se encontraron fijados en el Polo Argentino para los alelos asociados a un fenotipo de tamaño corporal grande. Sin embargo, el marcador Chr3:g.107374136 presentó una baja variabilidad en el Polo Argentino, siendo la frecuencia del alelo T (asociado al tamaño corporal pequeño) de 99,6%, ya que sólo 3 individuos fueron heterocigotas C/T (~1%) y ninguno homocigota C/C. En el SPC, en cambio, se observó una mayor frecuencia del alelo C (22% vs. 0,4%, valor-p= $2,2 \times 10^{-16}$) fundamentalmente a expensas de los genotipos heterocigotas. El índice F_{ST} entre las razas para este marcador fue elevado ($F_{ST}=0,64$; valor-p= $1,6 \times 10^{-08}$) demostrando una alta diferenciación genética. No hubo diferencias significativas entre ambas razas para el resto de los marcadores de tamaño corporal. No obstante, se observaron diferencias en las frecuencias alélicas y genotípicas para el marcador Chr11:g.23334511 (*LASP1*) entre los grupos Gen0 y Gen3, y entre Élite y No-élite (Tabla 4.3A y B). Un único ejemplar Polo Argentino resultó heterocigota para el SNP Chr19:g.30992750 (*OSTN*) el cual se encuentra asociado al tamaño corporal en razas de ponis, el resto fue homocigota para el alelo G.

En la Figura 4.1 se muestra la evolución de las frecuencias genotípicas de los marcadores con mayor polimorfismo en el Polo Argentino, durante las últimas 3 generaciones según la fecha de nacimiento de los ejemplares (períodos 1997-2004, 2005-2012 y 2013-2020 de acuerdo con el intervalo generacional de Polo Argentino estimado en 8 años – ver Capítulo I). Las frecuencias genotípicas del

Tabla 4.2. Marcadores SNPs polimórficos en la raza Polo Argentino.

Carácter y gen asociado	SNP	Alelos		Frecuencias genotípicas			Frecuencias alélicas		HWE	F _{ST}
		A1	A2	A1/A1	A1/A2	A2/A2	A1	A2	Valor - p	
Curiosidad y vigilancia (<i>DRD4</i>)	Chr12:g.35662671	A	G	18%	49%	33%	43%	57%	1	-0,003
Velocidad – distancia Carrera (<i>MSTN</i>)	Chr18:g.67027762	A	G	6%	50%	44%	31%	69%	0,0005702	0,15 *
Tamaño corporal (<i>LCORL</i>)	Chr3:g.107374136	C	T	0%	1%	99%	0,4%	99,6%	1	0,64 ***
Tamaño corporal (<i>HMGA2</i>)	Chr6:g.81481064	C	T	0%	0%	100%	0%	100%	1	-
Tamaño corporal (<i>ZFAT</i>)	Chr9:g.77656570	T	C	1,5%	27%	71,5%	15%	85%	0,4207	0,002
Tamaño corporal (<i>LASP1</i>)	Chr11:g.23334511	G	A	16%	45%	39%	38%	62%	0,4515	0,004
Tamaño corporal (<i>ADAMTS17</i>)	Chr1:g.106168922	A	C	0%	0%	100%	0%	100%	1	-
Tamaño corporal (<i>OSTN</i>)	Chr19:g.30992750	A	G	0%	0,26%	99,74%	0,13%	99,87%	1	-0,021
Tamaño corporal (<i>GH1</i>)	Chr11:g.15494088	T	C	0%	0%	100%	0%	100%	1	-

Tabla de frecuencias genotípicas y alélicas en el caballo Polo Argentino. El F_{ST} fue calculado entre la población Polo Argentino y SPC. *Valor-p < 0,05, *** Valor-p < 0,001

Tabla 4.3. Frecuencias genotípicas (A) y alélicas (B) de los SNP polimórficos asociados a caracteres morfológicos, velocidad y temperamento en los distintos grupos.

A)

Frecuencias genotípicas																		
Grupos	Curiosidad y vigilancia (<i>DRD4</i>)		Velocidad – distancia Carrera (<i>MSTN</i>)			Tamaño corporal (<i>LCORL</i> , <i>ZFAT</i> , <i>LASP1</i> , <i>OSTN</i>)												
	Chr12:g.35662671		Chr18:g.67027762			Chr3:g.107374136			Chr9:g.77656570			Chr11:g.23334511			Chr19:g.30992750			
	A/A+A/G	G/G	A/A	A/G	G/G	C/C	C/T	T/T	T/T	T/C	C/C	G/G	G/A	A/A	A/A	A/G	G/G	
Polo Argentino	67%	33%	6%**	50%	44%*	0%*	1%**	99%**	1,5%	27%	71,5%	16%	45%	39%	0%	0,26%	99,74%	
SPC	72%	28%	40%**	40%	20%*	4%*	36%**	60%**	4%	28%	68%	24%	44%	32%	0%	0%	100%	
Gen0	71%	29%	7%	55%	38%	0%	0%	100%	2%	25%	73%	13%	31%	56%*	0%	0%	100%	

Gen3	67%	33%	4%	42%	54%	0%	0%	100%	4%	33%	63%	25%	42%	33%*	0%	0%	100%
Elite	71%	29%	5%	54%	41%	0%	2%	98%	2%	30%	68%	23%*	40%	37%	0%	0%	100%
No-élite	67%	33%	6%	49%	45%	0%	0%	100%	1%	26%	73%	14%*	47%	39%	0%	0,33%	99,67%
Premiado	74%	26%	0%	44%	56%	0%	4%	96%	4%	26%	70%	26%	41%	33%	0%	0%	100%
No-premiado	67%	33%	6%	51%	43%	0%	1%	99%	1%	27%	72%	15%	46%	39%	0%	0,28%	99,72%

Las diferencias estadísticas fueron calculadas con una prueba de proporciones Z. *Valor-p < 0,05; ** Valor-p < 0,001

B)

Frecuencias alélicas												
Grupo	Curiosidad y vigilancia (<i>DRD4</i>)		Velocidad – distancia Carrera (<i>MSTN</i>)		Tamaño corporal (<i>LCORL</i> , <i>ZFAT</i> , <i>LASP1</i> , <i>OSTN</i>)							
	Chr12:g.35662671		Chr18:g.67027762		Chr3:g.107374136		Chr9:g.77656570		Chr11:g.23334511		Chr19:g.30992750	
	A	G	A	G	C	T	T	C	G	A	A	G
Polo Argentino	43%	57%	31%**	69%**	0,4%**	99,6%**	15%	85%	38%	62%	0,13%	99,87%
SPC	52%	48%	60%**	40%**	22%**	78%**	18%	82%	46%	54%	0%	100%
Gen0	43%	57%	34%	66%	0%	100%	14%	86%	29%*	71%*	0%	100%
Gen3	44%	56%	25%	75%	0%	100%	21%	79%	46%*	54%*	0%	100%
Élite	43%	57%	32%	68%	1,2%	98,8%	18%	82%	43%	57%	0%	100%
No-élite	43%	57%	31%	69%	0,17%	99,8%	14%	86%	37%	63%	0,17%	99,83%
Premiado	46%	54%	22%	78%	1,8%	98,2%	16%	84%	46%	54%	0%	100%
No-premiado	43%	57%	32%	68%	0,3%	99,7%	15%	85%	38%	62%	0,14%	99,86%

Las diferencias estadísticas fueron calculadas con una prueba de proporciones Z. Valor-p < 0,05; ** Valor-p < 0,001

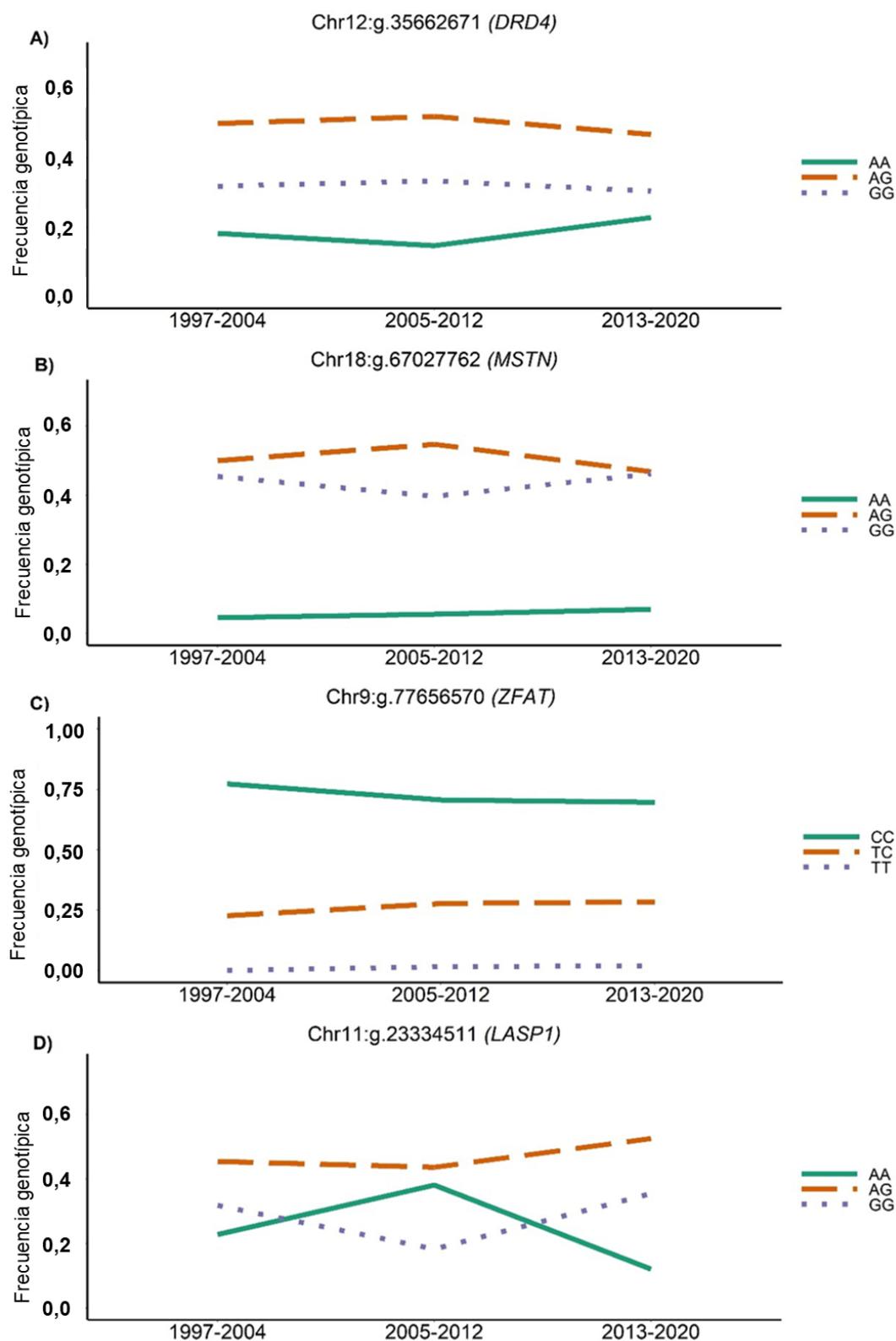


Figura 4.1. Evolución de las frecuencias genotípicas de los marcadores con mayor polimorfismo en los caballos Polo Argentino, durante las últimas 3 generaciones. Generación 1997-2004 (n= 22); 2005-2012 (n= 126); 2013-2020 (n= 158). Ejemplares sin fecha de nacimiento conocida no fueron incluidos en el gráfico.

marcador Chr12:g.35662671 (*DRD4*) se mantuvieron relativamente estables a lo largo del tiempo, aunque hubo una ligera tendencia hacia un aumento de homocigotas A/A y disminución de G/G. En cuanto al marcador Chr18:g.67027762 (*MSTN*), se pudo observar que la frecuencia de homocigotas G/G ha aumentado en la última generación, igualando a la frecuencia de heterocigotas. Finalmente, el marcador Chr11:g.23334511 (*LASP1*) tuvo una reducción en la frecuencia de homocigotas A/A durante la última generación, en favor de los otros dos genotipos.

4. Discusión

En este capítulo se analizó por primera vez la variabilidad genética en 9 loci diferentes asociados a caracteres de interés en la especie equina en una población altamente representativa del caballo Polo Argentino. Los resultados obtenidos demuestran que existen ciertos genotipos que son más comunes en la raza en comparación con el caballo SPC, reforzando la idea de la existencia de una incipiente huella de selección en el caballo de polo que impulsa una identidad genética propia.

El comportamiento es uno de los caracteres más valorados en las especies domésticas, teniendo una importancia aún mayor en los equinos, ya que la interacción y cooperación entre el caballo y el humano es la base de las actividades ecuestres. Este aspecto no sólo influye sobre el desempeño deportivo, sea cual sea la disciplina, sino también sobre la seguridad tanto para el caballo como para las personas que interactúan con él (Hausberger y col., 2008).

Los resultados obtenidos a partir del análisis del SNP Chr12:g.35662671 demuestran que el genotipo más frecuente en el Polo Argentino (así como en el SPC) es el asociado a una baja curiosidad y alta vigilancia. Este SNP está

localizado en el exón 3 del gen *DRD4* y se trata de una sustitución A-G (mutación no sinónima), que causa un cambio de aminoácido (asparagina a ácido aspártico) lo cual ha sido asociado a algunos aspectos del temperamento en el equino (Hori y col., 2013). En nuestro estudio, si bien las frecuencias génicas fueron similares entre ambas razas, el alelo G fue el de mayor frecuencia en Polo Argentino, mientras que en el SPC lo fue el alelo A, probablemente debido a las diferencias observadas en los genotipos A/A. Algunos estudios realizados en SPC muestran resultados similares, con frecuencias del alelo A entre 0,53 y 0,62 (Momozawa y col., 2005a; Hori y col., 2013). Un estudio que comparó las frecuencias alélicas de esta mutación entre diferentes razas nativas de Japón y SPC, encontró que los animales pertenecientes a las razas nativas Hokkaido y Taishu tenían frecuencias del alelo A muy bajas en comparación al caballo de carrera (entre 0,03 y 0,08; (Hori y col., 2013)). Los autores explicaron que las diferencias podrían deberse a que el SPC es una raza que se ha seleccionado por décadas casi exclusivamente para el rendimiento en carreras, y presenta un temperamento sanguíneo e hiperactivo. Esta característica muchas veces constituye una causa de retiro de animales de la actividad e incluso puede ser un obstáculo a la hora de reinsertar a los ejemplares retirados del hipódromo en otras actividades hípcas (Holtby y col., 2023).

Los procesos de selección o los cuellos de botella poblacionales ocurren naturalmente en la formación de una raza, estableciéndose diferencias genéticas, sean estas funcionales o no (efecto "hitchhicking"). El caballo Polo Argentino está estrechamente relacionado con el SPC, pero las disciplinas deportivas para las que son seleccionados (polo y turf, respectivamente) son completamente diferentes y requieren distintas habilidades, incluidas las relacionadas al comportamiento. A pesar de ello, si bien se encontraron diferencias significativas en los genotipos A/A del gen *DDR4* entre Polo Argentino y SPC, no hubo

diferencias en las frecuencias de los fenotipos asociados (genotipos A/A + A/G vs. G/G). Además, de acuerdo con las muestras analizadas, la proporción de los distintos genotipos se mantiene relativamente estable a lo largo de los años en la población Polo Argentino. En ellos predominan los genotipos A/A y A/G, los cuales se asocian a un temperamento catalogado como de baja curiosidad y alta vigilancia. A pesar de que en estos caballos se buscan cualidades de mansedumbre y docilidad, también es bien valorado por los criadores y jinetes su temperamento vivaz y sanguíneo, propio de las razas de “sangre caliente” como es el SPC. Además de los fenotipos de curiosidad y vigilancia antes mencionados, Ninomiya et al. (2013) evaluó la asociación entre el SNP Chr12:g.35662671 y la respuesta a la frustración en caballos estabulados. Los autores encontraron que los individuos sin el alelo A (genotipos G/G) destinaban mayor tiempo a las respuestas conductuales a la frustración (estereotipias) que aquellos con el alelo A (genotipos A/A y A/G). Si bien las respuestas conductuales estereotipadas son producto de situaciones de estrés para el animal, la identificación de una base genética que explique las diferencias fenotípicas observables podría ser de utilidad para la selección de individuos que toleren mejor las condiciones medioambientales a las que son expuestos durante su vida, mejorando las pautas de bienestar animal (Hausberger y col., 2008).

En nuestro trabajo se observó que la mayor proporción de caballos Polo Argentino presentó un genotipo heterocigota para el marcador Chr18:g.67027762 (*MSTN*), lo cual se relaciona con características intermedias de potencia y resistencia. De acuerdo con el desequilibrio de Hardy-Weinberg observado en la población es altamente probable que se encuentre bajo selección. La miostatina es una proteína que ha sido ampliamente caracterizada en diversas especies domésticas, como porcinos (Stratil y Kopecný, 1999), caninos (Mosher y col., 2007), bovinos (Grobet y col., 1997), ovinos (Boman y col., 2009) y equinos (Hill y col., 2010a).

Una inserción SINE en el promotor del gen *MSTN* afecta los niveles de expresión limitando la producción de miostatina (Rooney y col., 2018). El SNP g.66493737C>T (incluido en un haplotipo que contiene el gen *MSTN*) se encuentra fuertemente ligado a la inserción SINE, con una concordancia del 100% en esta raza (Hill et al., 2010b). El SNP analizado en nuestro estudio (BIEC2-438999 o Chr18:g.67027762) se encuentra también fuertemente asociado con g.66493737C>T (desequilibrio de ligamiento $r^2=0,93$) (Hill et al., 2019). De acuerdo con esta asociación, los caballos con el genotipo A/A para Chr18:g.67027762 presentan un mejor rendimiento en distancias cortas (<1600m), y fenotípicamente presentan una musculatura más desarrollada y compacta. El aumento de la musculatura se produce a expensas de las fibras glucolíticas rápidas de tipo IIB, caracterizadas por un mayor diámetro, baja densidad mitocondrial y capilar, una mayor velocidad de contracción y una baja resistencia a la fatiga, por lo que intervienen en ejercicios que requieren potencia por cortos períodos (Boffi, 2006; Binns y col., 2010). Por el contrario, los caballos G/G son mejores para carreras de larga distancia (>1600m), lo cual requiere una mayor resistencia, mientras que los heterocigotos muestran un fenotipo intermedio (Hill y col., 2010a).

Los caballos con el genotipo A/G (Chr18:g.67027762), como la mayoría de los analizados en este estudio, muestran una proporción de fibras musculares tipo I y IIA mayor que los homocigotas A/A, aunque menor que los G/G (Rooney y col., 2017). Por lo tanto, aquellos individuos con una mayor proporción de fibras tipo I tendrán una mayor capacidad oxidativa y resistencia a la fatiga que los de mayor proporción de fibras tipo IIX. A pesar de que el polo es un deporte que requiere movimientos rápidos y cambios bruscos de aceleración y desaceleración, la resistencia es también una característica importante para soportar la intensidad de un chukker. Si bien la dinámica del juego ha cambiado en los últimos años, ya

que existe una tendencia que busca potenciar las reacciones explosivas del caballo dentro del campo, jugando intensamente unos pocos minutos cada vez, esto no se ha traducido en un aumento de los genotipos A/A en las muestras analizadas. Tampoco se encontraron diferencias significativas en las frecuencias de genotipos entre los jugadores de Élite y No-élite. Por el contrario, los resultados muestran que se seleccionan preferentemente fenotipos intermedios o de resistencia, lo cual podría ser debido a las múltiples variables que intervienen en la selección de un ejemplar para un deporte tan complejo como el polo.

Los caracteres morfológicos son uno de los principales criterios de selección en la cría equina. La conformación de un ejemplar es muy importante ya que está en estrecha relación con la funcionalidad, el rendimiento deportivo, la salud y la longevidad (Sánchez-Guerrero y col., 2016). Por otro lado, este tipo de caracteres suelen tener una alta heredabilidad (es decir, un alto porcentaje de la varianza que es explicada por factores genéticos aditivos) por lo que la selección sobre ellos suele ser efectiva en el corto plazo (Bailey y Brooks, 2013).

Los resultados en este trabajo mostraron que los individuos de la raza Polo Argentino muestreados tienen alelos asociados tanto a una baja como una alta alzada a la cruz, lo cual es consistente con un caballo de mediano tamaño. A diferencia de los humanos, en los caballos el tamaño corporal está determinado por un pequeño número de loci con un gran efecto (Makvandi-Nejad y col., 2012). Los SNPs de tamaño corporal analizados en este estudio explican una gran proporción de la variación observada y se encuentran próximos a genes que se ha descrito que afectan al tamaño corporal en varias especies.

El marcador Chr3:g.107374136 (*LCORL*) fue uno de los que presentó una amplia y significativa diferenciación genética entre el Polo Argentino y el SPC ($F_{ST}=0,64$). Este marcador es el que mejor se asocia al tamaño corporal, siendo el alelo T el de mayor frecuencia en razas de pequeño tamaño (incluidas varias razas de ponis

como el Shetland, Falabella, Caspian, etc.) y el alelo C el más común en razas grandes (Hannoveriano, Belga, Clydesdale, Shire, etc.). Además, al explicar en gran parte la diferencia de tamaño corporal dentro de la raza SPC resulta de especial importancia para evaluar este carácter en el Polo Argentino. También se ha encontrado que dicha mutación se asociaba a otras medidas morfométricas como la longitud del cuerpo, la circunferencia del pecho y de la caña (He y col., 2015; Tozaki y col., 2016). En el Polo Argentino este locus tuvo una variabilidad muy reducida, ya que, excepto tres ejemplares heterocigotas, todos fueron homocigotas T/T, genotipo que se relaciona a un menor tamaño corporal. Esta diferencia podría suponer una selección positiva en el caballo de polo por ejemplares con una menor alzada dentro de la variabilidad que presenta el caballo de carrera. Esto probablemente se relacione con las preferencias de los jinetes y criadores, y con las ventajas que un caballo más pequeño y ágil tenga en el campo de juego. El biotipo considerado “correcto” o más apto para el polo actualmente es el de un animal de medidas intermedias, de movimientos bajos, con una alzada que ronda 1,56 m, siendo lo suficientemente liviano para mantener la agilidad, pero sin perder potencia y fortaleza (AACCP, 2023).

Todas las muestras que se analizaron en nuestro trabajo tuvieron fijado el alelo T para el locus Chr6:g.81481064 (*HMG2*). Los marcadores Chr9:g.77656570 (*ZFAT*) y Chr11:g.23334511 (*LASP1*) presentaron mayor variabilidad que los anteriores en ambas razas. La menor frecuencia de genotipos A/A (mayor tamaño) del marcador Chr11:g.23334511 (*LASP1*) entre individuos Gen3 y Gen0, así como la mayor frecuencia del genotipo G/G (pequeño tamaño) en los individuos de Élite podría suponer que existe cierta preferencia y selección hacia individuos de menor alzada, aunque la evolución observada en los últimos 20 años sugiere una tendencia hacia el aumento de los heterocigotas.

Los marcadores de tamaño corporal en ponis se encuentran asociados a distintos genes que tienen función sobre el crecimiento y desarrollo muscular esquelético (Metzger y col., 2018). Las mutaciones en estos marcadores han sido asociadas a la reducción de la alzada en ponies de tamaño miniatura. El marcador Chr19:g.30992750 (*OSTN*) es uno de los que más ampliamente contribuye a la disminución del tamaño corporal en las razas miniatura, con un efecto más notorio en los homocigotas (baja dominancia) (Metzger y col., 2018). En el Polo Argentino es destacable la presencia de un ejemplar con la mutación del marcador Chr19:g.30992750 (*OSTN*), ya que el mismo no ha sido descrito en razas de tamaño estándar, aunque sí en ponis de mayor alzada, como el Lewitzer, el cual tiene una estatura más variable. El ejemplar Polo Argentino heterocigota para la mutación tiene una generación de padres registrados, por lo que es posible que la frecuencia de dicho alelo sea mayor en la población. Sin embargo, es necesario destacar que en las razas miniatura se ha observado un efecto sinérgico entre los 4 loci, por lo que la presencia de un solo alelo mutado puede que no represente un efecto significativo sobre el fenotipo (Metzger y col., 2018).

Finalmente, cabe destacar que la posibilidad de disponer de información genómica tiene un impacto directo sobre la cría y el mejoramiento. La identificación de variantes genéticas que están asociadas con caracteres específicos, como la velocidad, el temperamento y la conformación constituye una poderosa herramienta, ya que posibilitará que los criadores dispongan de información objetiva para la selección de sus ejemplares y la mejora del rendimiento y la capacidad atlética (Bellone y Avila, 2020).

CONCLUSIONES

En este trabajo de tesis se ha evaluado y caracterizado la variabilidad genética poblacional del caballo Polo Argentino, utilizando datos genealógicos y genómicos.

En el Capítulo I se analizó la influencia del uso de programas de transferencia embrionaria a gran escala en la variabilidad genética de la raza mediante la utilización de datos de pedigree de alrededor de 115 mil individuos Polo Argentino. Los resultados demostraron que estas técnicas han acortado el intervalo generacional, favoreciendo la distribución de ciertas líneas genéticas (probablemente “superiores” o mejoradoras), pero también podrían reducir la variabilidad genética a mediano/largo plazo.

El análisis genealógico muestra que, aunque el parentesco medio y la endogamia han aumentado en la población durante los últimos años, la situación no es crítica, probablemente debido a la existencia de un libro genealógico abierto. Sin embargo, el estudio confirma que la principal influencia genética de la raza es de caballos Sangre Pura de Carrera, y que la contribución desigual (aumentada) de ciertos reproductores ha producido un efecto de cuello de botella en la población.

La pérdida de variabilidad genética se asoció a los programas de transferencia embrionaria. Por lo tanto, su uso debería complementarse con estrategias para mitigar el potencial efecto de la depresión endogámica para las futuras generaciones.

En el Capítulo II se estudió la variabilidad genética de la raza Polo Argentino a partir de datos genómicos obtenidos con el microarray de SNP Illumina Equine GGP array (70k), complementando los estudios genéticos realizados en el Capítulo I. Se empleó un enfoque basado en la determinación de bloques de

homocigosidad en el genoma, y la estimación del coeficiente de endogamia genómico (F_{ROH}), el cual en promedio fue notablemente superior al F_{PED} en las muestras analizadas, lo cual sugiere que el caballo de polo ha heredado una carga endogámica de sus antecesores SPC. Estos bloques de homocigosidad fueron mayoritariamente de tamaño corto (<5,5 Mb), lo cual refuerza esta teoría. A pesar de esto, el 99% de los individuos muestreados presentó ROH de tamaño medio (entre 8,3 y 16,6 Mb) y casi el 70% ROH de tamaño grande ($\geq 16,6$ Mb), lo cual demuestra también la existencia de eventos endogámicos recientes.

También quedó evidenciado que el uso de datos moleculares fue más apropiado que las estimaciones basadas en el pedigrí para evaluar la endogamia en el caballo Polo Argentino, debido a las características particulares de la raza relacionadas a su origen y evolución.

Si bien las estimaciones moleculares de la endogamia tuvieron una alta correlación (F_{IS} y F_{ROH}), el F_{ROH} permite considerar la longitud y ubicación de los segmentos en homocigosis y, de esta manera, aportar información relativa a la historia demográfica que puede resultar útil para evaluar el impacto de la endogamia y la purga de los alelos deletéreos en la población. A pesar de ello, el real efecto de la depresión endogámica no ha podido ser valorado hasta el momento, debido a la ausencia de datos fenotípicos en el caballo Polo Argentino.

El análisis de la distribución y frecuencia de los bloques de homocigosidad permitió la detección de 4 regiones del genoma con una alta frecuencia en la población (ROHi), las cuales contenían genes relacionados a caracteres de comportamiento, funciones metabólicas, reproductivas y enfermedades específicas. La detección de ROHi constituye una valiosa herramienta para evaluar la historia y selección de la raza, y para la identificación de genes con asociados a caracteres de interés en la especie equina.

En el Capítulo III se analizó la estructura y diferenciación genética del caballo Polo Argentino con 14 razas equinas, utilizando los datos de genotipado con el microarray de SNP. Los resultados reafirman que la raza Polo Argentino posee una estrecha relación con el Pura Sangre de Carrera, no habiéndose evidenciado una relación importante ni introgresión significativa de material genético de otras razas. Aun así, fue posible observar una leve e incipiente diferenciación genética entre el Polo Argentino y el Pura Sangre de Carrera, similar a la encontrada entre otras razas altamente emparentadas. Esto podría ser fruto de la selección diferencial entre ambas poblaciones que, si bien comparten características relacionadas a la aptitud atlética, se utilizan con propósitos diferentes. Es posible que el efecto fundador de ciertas líneas genéticas de Pura Sangre de Carrera seleccionadas específicamente para el polo le otorgue una incipiente identidad propia. A pesar de que el Polo Argentino es una raza relativamente joven, la intensidad de selección, la contribución genética desigual de los ancestros y el acortamiento de los intervalos generacionales (favorecido por el uso de biotecnologías reproductivas) podría ayudar a acelerar este proceso.

Adicionalmente, se encontraron marcadores genéticos con una diferenciación inusualmente alta respecto del Pura Sangre de Carrera, distribuidos por todo el genoma, lo que permitió identificar regiones que podrían constituir incipientes huellas de selección propias del Polo Argentino. Dichas regiones se encontraron próximas o dentro de genes con variadas funciones biológicas, muchas de ellas descritas en genes homólogos de otras especies animales, las cuales podrían explicar diferencias fenotípicas entre estas razas y asociarse a la aptitud deportiva de los ejemplares. Estos resultados podrían contribuir al conocimiento de la genética del rendimiento deportivo en el equino en general, y el caballo de polo en particular.

Finalmente, en el Capítulo IV se analizó la frecuencia de variantes genéticas asociadas a diferentes caracteres de interés en la raza Polo Argentino, realizándose además un análisis comparativo con el caballo Pura Sangre de Carrera, su principal influencia genética. Estas variantes estuvieron asociadas a rasgos del temperamento (curiosidad y vigilancia), la velocidad (distancia de carrera) y la conformación (tamaño corporal).

El marcador asociado al gen *DRD4*, el cual se vincula con características comportamentales, presentó variabilidad genética en el Polo Argentino, cobrando especial interés dada la importancia del temperamento como criterio de selección en la especie equina y en la raza en particular. A pesar de ello, no hubo diferencias estadísticamente significativas entre las dos razas estudiadas.

Por el contrario, las variantes asociadas a la velocidad - distancia de carrera y el tamaño corporal sí tuvieron diferencias entre el Polo Argentino y el Sangre Pura de Carrera, indicando selección genética sobre caracteres asociados al rendimiento deportivo.

En el Polo Argentino el marcador asociado al gen *MSTN* mostró una mayor prevalencia de la variante asociada a una mayor aptitud para las carreras largas y la resistencia, con predominio de los genotipos heterocigotas, y una fuerte evidencia de que este locus se encuentra bajo selección.

En cuanto al tamaño corporal, si bien el Polo Argentino posee variantes que determinan un fenotipo intermedio, se destaca la alta frecuencia del alelo T del marcador asociado al gen *LCORL* (99%) y la elevada diferenciación genética de este locus con el Sangre Pura de Carrera. Este hallazgo, junto con las diferencias observadas en el marcador asociado al gen *LASP1* en los individuos de Élite y en aquellos con mayor cantidad de ancestros registrados, supone una tendencia

hacia la selección de individuos con un tamaño más pequeño, en concordancia con la preferencia de los criadores y jinetes.

Si bien no se ha evaluado fenotípicamente al caballo Polo Argentino, las variantes analizadas en este trabajo hacen suponer que existe una selección de ejemplares Pura Sangre de Carrera que reúnen características óptimas para el polo, y dicha selección tiende a generar diferencias genéticas entre ambas poblaciones. Serán necesarios estudios adicionales para evaluar, desde una perspectiva fenotípica y funcional, la influencia que estas variantes pueden tener en el desempeño de los caballos en el deporte del polo.

El conocimiento de la situación genética de la raza resulta de especial interés para mantener el bienestar, guiar las decisiones de cría y contribuir con el progreso genético. Más aun considerando que la principal afluencia genética es una población cerrada altamente seleccionada desde el siglo XVIII: el Pura Sangre de Carrera. La alta presión de selección implementada por los criadores, de la mano de la utilización de biotecnologías de la reproducción y de la clonación equina a escala comercial, son factores que generan rápidos cambios en la variabilidad genética de la raza. Esto podría traducirse en mejora genética, aunque también en la fijación de caracteres indeseables o en una disminución del potencial de progreso genético a largo plazo. En este sentido, el carácter abierto de la política de registro actual del Polo Argentino podría considerarse una ventaja, ya que permite la introgresión de nuevas variantes provenientes de otras líneas o razas, aunque el alcance dependerá de las decisiones adoptadas por los criadores.

REFERENCIAS BIBLIOGRÁFICAS

1. AACCP. Asociacion Argentina de Criadores de caballos de Polo. <http://www.criapoloargentino.com.ar/>.
2. Abdalla I M, Lu X, Nazar M, Arbab A A I, Xu T, Yousif M H, Mao Y, Yang Z. Genome-Wide Association Study Identifies Candidate Genes Associated with Feet and Leg Conformation Traits in Chinese Holstein Cattle. *Animals*. 2021. 11: e2259.
3. Ablondi M, Viklund Å, Lindgren G, Eriksson S, Mikko S. Signatures of selection in the genome of Swedish warmblood horses selected for sport performance. *BMC Genomics*. 2019. 20: 717.
4. Aboul-Naga A M, Alsamman A M, El Allali A, Elshafie M H, Abdelal E S, Abdelkhalek T M, Abdelsabour T H, Mohamed L G, Hamwieh A. Genome-wide analysis identified candidate variants and genes associated with heat stress adaptation in Egyptian sheep breeds. *Front. Genet*. 2022. 13: 898522.
5. Alexander D H, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res*. 2009. 19: 1655-64.
6. Álvarez I, Royo L J, Pérez-Pardal L, Fernández I, Lorenzo L, Goyache F. Assessing diversity losses due to selection for coat colour in the endangered bay-Asturcón pony using microsatellites. *Livest. Sci*. 2011. 135: 199-204.
7. Álvarez R P, Demyda Peyrás S, Prado Silva R H, Arroyo P, Trigo P I. Análisis de componentes principales en las etapas clasificatorias de una prueba de doma. XXXIII Conferencias Internacionales de Veterinaria Equina. 2022a.
8. Álvarez R P, Marcuzzi O, Demyda Peyrás S, Prado Silva R H, Trigo P I. Análise preliminar das etapas de uma prova de doma. 39º Encontro Anual de Etologia da Sociedade Brasileira de Etologia (SBEt). 2022b.
9. Álvarez R P, Marcuzzi O, Demyda Peyrás S, Prado Silva R H, Trigo P I. Autoavaliação versus avaliação dos jurados em prova de doma de equinos. 39º Encontro Anual de Etologia da Sociedade Brasileira de Etologia (SBEt). 2022c.
10. Aomine Y, Sakurai K, Macpherson T, Ozawa T, Miyamoto Y, Yoneda Y, Oka M, Hikida T. Importin $\alpha 3$ (KPNA3) Deficiency Augments Effortful Reward-Seeking Behavior in Mice. *Front. Neurosci*. 2022. 16: 905991.
11. Arya S B, Kumar G, Kaur H, Kaur A, Tuli A. ARL11 regulates lipopolysaccharide-stimulated macrophage activation by promoting mitogen-activated protein kinase (MAPK) signaling. *J. Biol. Chem*. 2018. 293: 9892-9909.
12. Avila F, Mickelson J R, Schaefer R J, McCue M E. Genome-Wide Signatures of Selection Reveal Genes Associated With Performance in American Quarter Horse Subpopulations. *Front. Genet*. 2018. 9: 249.
13. Babaev O, Cruces-Solis H, Piletti Chatain C, Hammer M, Wenger S, Ali H, Karalis N, de Hoz L, Schlüter O M, Yanagawa Y, Ehrenreich H, Taschenberger H, Brose N, Krueger-Burg D. IgSF9b regulates anxiety behaviors through effects on centromedial amygdala inhibitory synapses. *Nat. Commun*. 2018. 9: 5400.
14. Bailey E, Brooks S A. Genetics of Performance. En: *Horse Genetics*. 2nd. CABI, 2013.

15. Bailey E, Petersen J L, Kalbfleisch T S. Genetics of Thoroughbred Racehorse Performance. *Annu. Rev. Anim. Biosci.* 2022. 10: 131-150.
16. Barraza J A M, Castillo J, Ruiz E. Diferenciación hipométrica entre el caballo peruano de paso del norte y centro sur del Perú. *Anales científicos. Universidad Nacional Agraria La Molina*, 2018.
17. Bartolomé E, Cervantes I, Valera M, Gutiérrez J P. Influence of foreign breeds on the genetic structure of the Spanish Sport Horse population. *Livest. Sci.* 2011. 142: 70-79.
18. Baumung R, Farkas J, Boichard D, Mészáros G, Sölkner J, Curik I. grain: A computer program to calculate ancestral and partial inbreeding coefficients using a gene dropping approach. *J. Anim. Breed. Genet.* 2015. 132: 100-108.
19. Bellinge R H, Liberles D A, Iaschi S P, O'Brien P A, Tay G K. Myostatin and its implications on animal breeding: a review. *Anim. Genet.* 2005. 36: 1-6.
20. Bellone R R, Avila F. Genetic Testing in the Horse. *Vet. Clin. North Am. Equine Pract.* 2020. 36: 211-234.
21. Ben Hamouda S, Miglino M A, de Sá Schiavo Matias G, Beauchamp G, Lavoie J P. Asthmatic Bronchial Matrices Determine the Gene Expression and Behavior of Smooth Muscle Cells in a 3D Culture Model. *Front. Allergy.* 2021. 2: 762026.
22. Binns M M, Boehler D A, Bailey E, Lear T L, Cardwell J M, Lambert D H. Inbreeding in the Thoroughbred horse. *Anim. Genet.* 2012. 43: 340-342.
23. Binns M M, Boehler D A, Lambert D H. Identification of the myostatin locus (MSTN) as having a major effect on optimum racing distance in the Thoroughbred horse in the USA. *Anim. Genet.* 2010. 41 Suppl 2: 154-8.
24. Biscarini F, Cozzi P, Gaspa G, Marras G. detectRUNS: An R Package to Detect Runs of Homozygosity and Heterozygosity in Diploid Genomes. 2018.
25. Bizarria Dos Santos W, Pimenta Schettini G, Fonseca M G, Pereira G L, Loyola Chardulo L A, Rodrigues Machado Neto O, Baldassini W A, Nunes de Oliveira H, Abdallah Curi R. Fine-scale estimation of inbreeding rates, runs of homozygosity and genome-wide heterozygosity levels in the Mangalarga Marchador horse breed. *J. Anim. Breed. Genet.* 2021. 138: 161-173.
26. Boakes E H, Wang J, Amos W. An investigation of inbreeding depression and purging in captive pedigreed populations. *Heredity.* 2007. 98: 172-82.
27. Boffi F M. Fisiología del ejercicio en equinos. *Intermédica*, 2006.
28. Boichard D, Maignel L, Verrier É. The value of using probabilities of gene origin to measure genetic variability in a population. *Genet. Sel. Evol.* 1997. 29: 5-23.
29. Bokor Á, Jónás D, Ducro B, Nagy I, Bokor J, Szabari M. Pedigree analysis of the Hungarian Thoroughbred population. *Livest. Sci.* 2013. 151: 1-10.
30. Boman I A, Klemetsdal G, Blichfeldt T, Nafstad O, Våge D I. A frameshift mutation in the coding region of the myostatin gene (MSTN) affects carcass conformation and fatness in Norwegian White Sheep (*Ovis aries*). *Anim. Genet.* 2009. 40: 418-22.
31. Borowska A, Szwaczkowski T. Pedigree analysis of Polish warmblood horses participating in riding performance tests. *Can. J. Anim. Sci.* 2014. 95: 21-29.

32. Bosio Y, Berto G, Camera P, Bianchi F, Ambrogio C, Claus P, Di Cunto F. PPP4R2 regulates neuronal cell differentiation and survival, functionally cooperating with SMN. *Eur. J. Cell Biol.* 2012. 91: 662-674.
33. Bower M A, McGivney B A, Campana M G, Gu J, Andersson L S, Barrett E, Davis C R, Mikko S, Stock F, Voronkova V, Bradley D G, Fahey A G, Lindgren G, MacHugh D E, Sulimova G, Hill E W. The genetic origin and history of speed in the Thoroughbred racehorse. *Nat. Commun.* 2012. 3: 643.
34. Brooks S A. Genomics in the Horse Industry: Discovering New Questions at Every Turn. *J Equine Vet Sci.* 2021. 100: 103456.
35. Bussiman F D O, Perez B, Ventura R, Peixoto M, Curi R, Balieiro J. Pedigree analysis and inbreeding effects over morphological traits in Campolina horse population. *Animal.* 2018. 12: 1-10.
36. Buzanskas M E, Grossi D A, Ventura R V, Schenkel F S, Sargolzaei M, Meirelles S L, Mokry F B, Higa R H, Mudadu M A, da Silva M V, Niciura S C, Torres R A, Jr., Alencar M M, Regitano L C, Munari D P. Genome-wide association for growth traits in Canchim beef cattle. *PLoS One.* 2014. 9: e94802.
37. Caballero A, Toro M A. Analysis of genetic diversity for the management of conserved subdivided populations. *Conserv. Genet.* 2002. 3: 289-299.
38. Campbell M L H. Embryo transfer in competition horses: Managing mares and expectations. *Equine Vet. Ed.* 2014. 26: 322-327.
39. Campos-Sandoval J A, Martín-Rufián M, Cardona C, Lobo C, Peñalver A, Márquez J. Glutaminases in brain: Multiple isoforms for many purposes. *Neurochem. Int.* 2015. 88: 1-5.
40. Ceballos F C, Joshi P K, Clark D W, Ramsay M, Wilson J F. Runs of homozygosity: windows into population history and trait architecture. *Nat. Rev. Genet.* 2018. 19: 220-234.
41. Cervantes I, Goyache F, Molina A, Valera M, Gutierrez J P. Estimation of effective population size from the rate of coancestry in pedigreed populations. *J. Anim. Breed. Genet.* 2011. 128: 56-63.
42. Cervantes I, Molina A, Goyache F, Gutiérrez J P, Valera M. Population history and genetic variability in the Spanish Arab Horse assessed via pedigree analysis. *Livest. Sci.* 2008. 113: 24-33.
43. Coleman J R I, Bryois J, Gaspar H A, Jansen P R, Savage J E, Skene N, Plomin R, Muñoz-Manchado A B, Linnarsson S, Crawford G, Hjerling-Leffler J, Sullivan P F, Posthuma D, Breen G. Biological annotation of genetic loci associated with intelligence in a meta-analysis of 87,740 individuals. *Mol. Psychiatry.* 2019. 24: 182-197.
44. Colpitts J, McLoughlin P D, Poissant J. Runs of homozygosity in Sable Island feral horses reveal the genomic consequences of inbreeding and divergence from domestic breeds. *BMC Genomics.* 2022. 23: 501.
45. Comparini L, Podestà A, Russo C, Cecchi F. Effect of inbreeding on the "Club Foot" disorder in Arabian Pureblood horses reared in Italy. *Open Vet. J.* 2019. 9: 273-280.
46. Cooper M A, Crockett D P, Nowakowski R S, Gale N W, Zhou R. Distribution of EphA5 receptor protein in the developing and adult mouse nervous system. *J. Comp. Neurol.* 2009. 514: 310-28.

47. Corbi Botto C. Análisis de la variabilidad de genes implicados en la respuesta inmune en poblaciones equinas de argentina mediante detección de polimorfismos de nucleótido simple. PhD. Universidad Nacional de La Plata. 2020.
48. Cosgrove E J, Sadeghi R, Schlamp F, Holl H M, Moradi-Shahrbabak M, Miraei-Ashtiani S R, Abdalla S, Shykind B, Troedsson M, Stefaniuk-Szmukier M, Prabhu A, Bucca S, Bugno-Poniewierska M, Wallner B, Malek J, Miller D C, Clark A G, Antczak D F, Brooks S A. Genome Diversity and the Origin of the Arabian Horse. *Sci. Rep.* 2020. 10: 9702.
49. Crawford K, Oliver P L, Agnew T, Hunn B H M, Ahel I. Behavioural Characterisation of MacroD1 and MacroD2 Knockout Mice. *Cells.* 2021. 10: 368.
50. Cunningham E P, Dooley J J, Splan R K, Bradley D G. Microsatellite diversity, pedigree relatedness and the contributions of founder lineages to thoroughbred horses. *Anim. Genet.* 2001. 32: 360-364.
51. Curik I, Ferenčaković M, Sölkner J. Inbreeding and runs of homozygosity: A possible solution to an old problem. *Livest. Sci.* 2014. 166: 26-34.
52. Charlesworth D, Willis J H. The genetics of inbreeding depression. *Nat. Rev. Genet.* 2009. 10: 783-96.
53. Cheng P, Qiu Z, Du Y. Potassium channels and autism spectrum disorder: An overview. *Int. J. Dev. Neurosci.* 2021. 81: 479-491.
54. Delgado J F, De Andrés N, Valera M, Gutiérrez J P, Cervantes I. Assessment of population structure depending on breeding objectives in Spanish Arabian horse by genealogical and molecular information. *Livest. Sci.* 2014. 168: 9-16.
55. Deutschman E, Ward J R, Kumar A, Ray G, Welch N, Lemieux M E, Dasarathy S, Longworth M S. Condensin II protein dysfunction impacts mitochondrial respiration and mitochondrial oxidative stress responses. *J. Cell Sci.* 2019. 132: jcs233783.
56. Doekes H P, Bijma P, Windig J J. How Depressing Is Inbreeding? A Meta-Analysis of 30 Years of Research on the Effects of Inbreeding in Livestock. *Genes.* 2021. 12: 926.
57. Dos S J T, Dos S C R, Alcântara-Neves N M, Barreto M L, Figueiredo C A. Variants in the CYSLTR2 are associated with asthma, atopy markers and helminths infections in the Brazilian population. *Prostaglandins Leukot. Essent. Fatty Acids.* 2019. 145: 15-22.
58. Druml T, Brem G, Velie B, Lindgren G, Horna M, Ricard A, Grilz-Seger G. Equine vitiligo-like depigmentation in grey horses is related to genes involved in immune response and tumor metastasis. *BMC Vet. Res.* 2021. 17: 336.
59. Druml T, Neuditschko M, Grilz-Seger G, Horna M, Ricard A, Mesaric M, Cotman M, Pausch H, Brem G. Population Networks Associated with Runs of Homozygosity Reveal New Insights into the Breeding History of the Haflinger Horse. *J. Hered.* 2018. 109: 384-392.
60. Dubois C, Ricard A. Efficiency of past selection of the French Sport Horse: Selle Français breed and suggestions for the future. *Livest. Sci.* 2007. 112: 161-171.
61. Easa A A, Selionova M, Aibazov M, Mamontova T, Sermyagin A, Belous A, Abdelmanova A, Deniskova T, Zinovieva N. Identification of Genomic Regions

- and Candidate Genes Associated with Body Weight and Body Conformation Traits in Karachai Goats. *Genes*. 2022. 13: 1773.
62. Excoffier L, Smouse P E, Quattro J M. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*. 1992. 131: 479-91.
 63. Fairn G D, McMaster C R. Emerging roles of the oxysterol-binding protein family in metabolism, transport, and signaling. *Cell. Mol. Life Sci*. 2008. 65: 228-36.
 64. Falconer D. Quantitative genetics. 3rd edition. England, Longman Scientific & Technical, 1989.
 65. FAO. Estado de la cuestión en la gestión de los recursos zoogenético. En: La situación de los recursos zoogenéticos mundiales para la alimentación y la agricultura. 2010.
 66. FAO. Food and Agriculture Organization of the United Nations. <http://www.fao.org>.
 67. Fawcett J A, Sato F, Sakamoto T, Iwasaki W M, Tozaki T, Innan H. Genome-wide SNP analysis of Japanese Thoroughbred racehorses. *PLoS One*. 2019. 14: e0218407.
 68. Ferencaković M, Sölkner J, Curik I. Estimating autozygosity from high-throughput information: Effects of SNP density and genotyping errors. *Genet. Sel. Evol*. 2013. 45: 42.
 69. Ferreira G C. Ferrochelataze. *The international journal of biochemistry & cell biology*. 1999. 31: 995-1000.
 70. Fisher R A. A fuller theory of “Junctions” in inbreeding. *Heredity*. 1954. 8: 187-197.
 71. Gambini A, Maserati M. A journey through horse cloning. *Reprod. Fertil. Dev*. 2018. 30: 8-17.
 72. Gandini G C, Bagnato A, Miglior F, Pagnacco G. Inbreeding in the Italian Haflinger horse. *J. Anim. Breed. Genet*. 1992. 109: 433-443.
 73. Geerts C J, Plomp J J, Koopmans B, Loos M, van der Pijl E M, van der Valk M A, Verhage M, Groffen A J. Tomosyn-2 is required for normal motor performance in mice and sustains neurotransmission at motor endplates. *Brain Struct. Funct*. 2015. 220: 1971-82.
 74. Gengler N, Druet T. Impact of biotechnology on animal breeding and genetics progress. En: *Biotechnology in Animal Husbandry. Focus on Biotechnology*. Springer, 2001, p. 33-45.
 75. Giontella A, Pieramati C, Silvestrelli M, Sarti F M. Analysis of founders and performance test effects on an autochthonous horse population through pedigree analysis: structure, genetic variability and inbreeding. *Animal*. 2019. 13: 15-24.
 76. Gmel A I, Burger D, Neuditschko M, 2021. A Novel QTL and a Candidate Gene Are Associated with the Progressive Motility of Franches-Montagnes Stallion Spermatozoa after Thaw. *Genes*, p. 1501.
 77. Goessler U R, Bugert P, Bieback K, Deml M, Sadick H, Hormann K, Riedel F. In-vitro analysis of the expression of TGFbeta -superfamily-members during

- chondrogenic differentiation of mesenchymal stem cells and chondrocytes during dedifferentiation in cell culture. *Cell. Mol. Biol. Lett.* 2005. 10: 345-62.
78. Gogarten S M, Sofer T, Chen H, Yu C, Brody J A, Thornton T A, Rice K M, Conomos M P. Genetic association testing using the GENESIS R/Bioconductor package. *Bioinformatics.* 2019. 35: 5346-5348.
 79. Gómez M D, Valera M, Molina A. Genetic analysis of racing performance of trotter horses in Spain. *Livest. Sci.* 2010. 127: 197-204.
 80. Gómez M D, Valera M, Molina A, Gutiérrez J P, Goyache F. Assessment of inbreeding depression for body measurements in Spanish Purebred (Andalusian) horses. *Livest. Sci.* 2009. 122: 149-155.
 81. Gorssen W, Meyermans R, Janssens S, Buys N. A publicly available repository of ROH islands reveals signatures of selection in different livestock and pet species. *Genet. Sel. Evol.* 2021. 53: 2.
 82. Gosselin T. radiator: RADseq Data Exploration, Manipulation and Visualization using R. 2020. R.p.v. 1.1.9.
 83. Goszczynski D, Molina A, Terán E, Morales-Durand H, Ross P, Cheng H, Giovambattista G, Demyda-Peyrás S. Runs of homozygosity in a selected cattle population with extremely inbred bulls: Descriptive and functional analyses revealed highly variable patterns. *PLoS One.* 2018. 13: e0200069.
 84. Grilz-Seger G, Druml T, Neuditschko M, Mesarič M, Cotman M, Brem G. Analysis of ROH patterns in the Noriker horse breed reveals signatures of selection for coat color and body size. *Anim. Genet.* 2019. 50: 334-346.
 85. Grobet L, Martin L J, Poncelet D, Pirottin D, Brouwers B, Riquet J, Schoeberlein A, Dunner S, Ménissier F, Massabanda J, Fries R, Hanset R, Georges M. A deletion in the bovine myostatin gene causes the double-muscling phenotype in cattle. *Nat. Genet.* 1997. 17: 71-4.
 86. Groeneveld L F, Lenstra J A, Eding H, Toro M A, Scherf B, Pilling D, Negrini R, Finlay E K, Jianlin H, Groeneveld E, Weigend S. Genetic diversity in farm animals - A review. *Anim. Genet.* 2010. 41: 6-31.
 87. Gu J, Orr N, Park S D, Katz L M, Sulimova G, MacHugh D E, Hill E W. A genome scan for positive selection in thoroughbred horses. *PLoS One.* 2009. 4: e5767.
 88. Gurgul A, Jasielczuk I, Semik-Gurgul E, Pawlina-Tyszko K, Stefaniuk-Szmukier M, Szmatoła T, Polak G, Tomczyk-Wrona I, Bugno-Poniewierska M. A genome-wide scan for diversifying selection signatures in selected horse breeds. *PLoS One.* 2019. 14: e0210751.
 89. Gurgul A, Semik E, Pawlina K, Szmatoła T, Jasielczuk I, Bugno-Poniewierska M. The application of genome-wide SNP genotyping methods in studies on livestock genomes. *J. Appl. Genet.* 2014. 55: 197-208.
 90. Gutiérrez J, Cervantes I, Goyache F. Improving the estimation of realized effective population sizes in farm animals. *J. Anim. Breed. Genet.* 2009. 126: 327-32.
 91. Gutierrez J P, Goyache F. A note on ENDOG: a computer program for analysing pedigree information. *J. Anim. Breed. Genet.* 2005. 122: 172-6.
 92. Hamann H, Distl O. Genetic variability in Hanoverian warmblood horses using pedigree analysis. *J. Anim. Sci.* 2008. 86: 1503-1513.

93. Han S, Fei F, Sun S, Zhang D, Dong Q, Wang X, Wang L. Increased anxiety was found in *serpini1* knockout zebrafish larval. *Biochem. Biophys. Res. Commun.* 2021. 534: 1013-1019.
94. Harvey A M, Williams S B, Singer E R. The effect of lateral heel studs on the kinematics of the equine digit while cantering on grass. *Vet. J.* 2012. 192: 217-21.
95. Hausberger M, Roche H, Henry S, Visser E K. A review of the human–horse relationship. *Appl. Anim. Behav. Sci.* 2008. 109: 1-24.
96. Hayashi T, Yoshida T, Ra M, Taguchi R, Mishina M. IL1RAPL1 associated with mental retardation and autism regulates the formation and stabilization of glutamatergic synapses of cortical neurons through RhoA signaling pathway. *PLoS One.* 2013. 8: e66254.
97. He S, Zhang L, Li W, Liu M. BIEC2-808543 SNP in the LCORL Gene is Associated with Body Conformation in the Yili Horse. *Anim. Biotechnol.* 2015. 26: 289-91.
98. Hedrick P W, Garcia-Dorado A. Understanding Inbreeding Depression, Purging, and Genetic Rescue. *Trends Ecol. Evol.* 2016. 31: 940-952.
99. Herrera C, Morikawa M I, Bello M B, Von Meyeren M, Eusebio Centeno J, Dufourq P, Martinez M M, Llorente J. Setting up equine embryo gender determination by preimplantation genetic diagnosis in a commercial embryo transfer program. *Theriogenology.* 2014. 81: 758-763.
100. Hill E W, Gu J, Eivers S S, Fonseca R G, McGivney B A, Govindarajan P, Orr N, Katz L M, MacHugh D. A Sequence Polymorphism in MSTN Predicts Sprinting Ability and Racing Stamina in Thoroughbred Horses. *PLoS One.* 2010a. 5: e8645.
101. Hill E W, McGivney B A, Gu J, Whiston R, MacHugh D E. A genome-wide SNP-association study confirms a sequence variant (g.66493737C>T) in the equine myostatin (MSTN) gene as the most powerful predictor of optimum racing distance for Thoroughbred racehorses. *BMC Genomics.* 2010b. 11: 552.
102. Hill E W, McGivney B A, Rooney M F, Katz L M, Parnell A, MacHugh D E. The contribution of myostatin (MSTN) and additional modifying genetic loci to race distance aptitude in Thoroughbred horses racing in different geographic regions. *Equine Vet. J.* 2019. 51: 625-633.
103. Hill E W, Stoffel M A, McGivney B A, MacHugh D E, Pemberton J M. Inbreeding depression and the probability of racing in the Thoroughbred horse. *Proc. Biol. Sci.* 2022. 289: 20220487.
104. Holtby A R, Hall T J, McGivney B A, Han H, Murphy K J, MacHugh D E, Katz L M, Hill E W. Integrative genomics analysis highlights functionally relevant genes for equine behaviour. *Anim. Genet.* 2023.
105. Hori Y, Ozaki T, Yamada Y, Tozaki T, Kim H-S, Takimoto A, Endo M, Manabe N, Inoue-Murayama M, Fujita K. Breed Differences in Dopamine Receptor D4 Gene (DRD4) in Horses. *J. Equine Sci.* 2013. 24: 31-6.
106. Howrigan D P, Simonson M A, Keller M C. Detecting autozygosity through runs of homozygosity: A comparison of three autozygosity detection algorithms. *BMC Genomics.* 2011. 12: 460.

107. Huang da W, Sherman B T, Lempicki R A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 2009. 4: 44-57.
108. Inness C M, Morgan K L. Polo pony injuries: player-owner reported risk, perception, mitigation and risk factors. *Equine Vet. J.* 2015. 47: 422-7.
109. Inoue-Murayama M. Genetic polymorphism as a background of animal behavior. *Anim. Sci. J.* 2009. 80: 113-20.
110. Insolera R, Shao W, Airik R, Hildebrandt F, Shi S H. SDCCAG8 regulates pericentriolar material recruitment and neuronal migration in the developing cortex. *Neuron.* 2014. 83: 805-22.
111. James J W. A note on selection differential and generation length when generations overlap. *Anim. Sci.* 1977. 24: 109-112.
112. Janova E, Futas J, Klumplerova M, Putnova L, Vrtkova I, Vyskocil M, Frolkova P, Horin P. Genetic diversity and conservation in a small endangered horse population. *J. Appl. Genet.* 2013. 54: 285-92.
113. Junqueira V S, Lopes P S, Resende M, Silva F F, Lourenço D, Yokoo M, Cardoso F. Impact of embryo transfer phenotypic records on large-scale beef cattle genetic evaluations. *Rev. Bras. Zootec.* 2018.
114. Kamvar Z N, Brooks J C, Grünwald N J. Novel R tools for analysis of genome-wide population genetic data with emphasis on clonality. *Front. Genet.* 2015. 6: 208.
115. Kardos M, Luikart G, Allendorf F W. Measuring individual inbreeding in the age of genomics: marker-based measures are better than pedigrees. *Heredity.* 2015. 115: 63-72.
116. Karlau A, Molina A, Antonini A, Demyda Peyrás S. The influence of foreign lineages in the genetic component of reproductive traits in Criollo Argentino mares: A 30-year study. *Livest. Sci.* 2023. 267: 105153.
117. Kim J, Jung H, Choi J Y, Lee J W, Yoon M. Plasma concentration of dopamine varies depending on breed, sex, and the genotype of DRD4 in horses. *J. Anim. Sci. Technol.* 2022. 64: 792-799.
118. Kim M J, Lee R U, Oh J, Choi J E, Kim H, Lee K, Hwang S K, Lee J H, Lee J A, Kaang B K, Lim C S, Lee Y S. Spatial Learning and Motor Deficits in Vacuolar Protein Sorting-associated Protein 13b (Vps13b) Mutant Mouse. *Exp. Neurobiol.* 2019. 28: 485-494.
119. Ko J A, Kimura Y, Matsuura K, Yamamoto H, Gondo T, Inui M. PDZRN3 (LNX3, SEMCAP3) is required for the differentiation of C2C12 myoblasts into myotubes. *J. Cell Sci.* 2006. 119: 5106-13.
120. Lacy R. Analysis of Founder Representation in Pedigrees - Founder Equivalents and Founder Genome Equivalents. *Zoobiology.* 1989. 8: 111-123.
121. Lacy R C, Alaks G, Walsh A. Hierarchical analysis of inbreeding depression in *Peromyscus polionotus*. *Evolution.* 1996. 50: 2187-2200.
122. Laffaye H A. *The Evolution of Polo.* McFarland, 2009.
123. Laseca N, Molina A, Ramón M, Valera M, Azcona F, Encina A, Demyda-Peyrás S. Fine-Scale Analysis of Runs of Homozygosity Islands Affecting Fertility in Mares. *Front. Vet. Sci.* 2022. 9: 754028.

124. Laseca N, Perdomo-González D I, Valera M, Molina Alcalá A, Sánchez Guerrero M, Azcona F, Pirosanto Y, Demyda Peyrás S. Foaling number are highly affected by the genomic homozygosity in the Pura Raza Spanish mares. *J Equine Vet Sci.* 2020. 89: 103098.
125. Lawrence M. HelloRanges: Introduce *Ranges to bedtools users. 2021.
126. Leroy G. Inbreeding depression in livestock species: review and meta-analysis. *Anim. Genet.* 2014. 45: 618-28.
127. Li H, Xu S, Gao X, Ren H. Structure of the bovine ACAD8 gene and the association of its polymorphism with the production traits. *J. Genet. Genomics.* 2007. 34: 315-20.
128. Librado P, Fages A, Gaunitz C, Leonardi M, Wagner S, Khan N, Hanghøj K, Alquraishi S A, Alfarhan A H, Al-Rasheid K A, Der Sarkissian C, Schubert M, Orlando L. The Evolutionary Origin and Genetic Makeup of Domestic Horses. *Genetics.* 2016. 204: 423-434.
129. Littiere T O, Castro G H F, Rodriguez M d P R, Bonafé C M, Magalhães A F B, Faleiros R R, Vieira J I G, Santos C G, Verardo L L. Identification and Functional Annotation of Genes Related to Horses' Performance: From GWAS to Post-GWAS. *Animals.* 2020. 10: 1173.
130. Liu H, Song H, Jiang Y, Jiang Y, Zhang F, Liu Y, Shi Y, Ding X, Wang C. A Single-Step Genome Wide Association Study on Body Size Traits Using Imputation-Based Whole-Genome Sequence Data in Yorkshire Pigs. *Front. Genet.* 2021. 12: e629049.
131. Liu L L, Xiao Y S, Huang W M, Liu S, Huang L X, Zhong J H, Jia P, Liu W Y. ATF1/miR-214-5p/ITGA7 axis promotes osteoclastogenesis to alter OVX-induced bone absorption. *Mol. Med.* 2022. 28: 56.
132. Losinno L, Aguilar J, Lisa H. Impact of multiple ovulations in a commercial equine embryo transfer programme. *Havemeyer Found. Monogr. Ser.* 2001. 3: 81-83.
133. Maccluer J, J. Boyce A, Dyke B, R. Weitkamp L, W. Pfenning D, J. Parsons C. Inbreeding and pedigree structure in Standardbred horses. *J. Hered.* 1983. 74: 394-399.
134. Machmoum M, Boujenane I, Azelhak R, Badaoui B, Petit D, Piro M. Genetic Diversity and Population Structure of Arabian Horse Populations Using Microsatellite Markers. *J Equine Vet Sci.* 2020. 93: 103200.
135. MAGyP. Ministerio de Economía Argentina. Secretaría de Agricultura, Ganadería y Pesca. Informe sobre existencias. Estadísticas de producción equina. https://www.magyp.gob.ar/sitio/areas/equinos/informacion_estadistica/.
136. Maignel L, Boichard D, Verrier E. Genetic variability of French dairy breeds estimated from pedigree information. *Interbull Bulletin.* 1996. 14: 49-54.
137. Makvandi-Nejad S, Hoffman G E, Allen J J, Chu E, Gu E, Chandler A M, Loredó A I, Bellone R R, Mezey J G, Brooks S A, Sutter N B. Four loci explain 83% of size variation in the horse. *PLoS One.* 2012. 7: e39929.
138. Malécot G. *Les mathematiques de l'Heredité.* Paris, Masson et Cie, 1948.

139. Mamiya P C, Hennesy Z, Zhou R, Wagner G C. Changes in attack behavior and activity in EphA5 knockout mice. *Brain Res.* 2008. 1205: 91-9.
140. MAPAMA. Catálogo Oficial de Razas. <http://www.mapama.gob.es/es/ganaderia/temas/zootecnia/razas-ganaderas/razas/catalogo/>.
141. Marinho L, Lunardelli P, Oliveira Rosa C, Sarapião F, Seneda M. Large-scale in vitro embryo production as a tool to improve animal breeding. En: *Advances in Medicine and Biology.* Nova Science Publishers, 2015, p. 93-102.
142. Marinone A I, Losinno L, Fumuso E, Rodríguez E M, Redolatti C, Cantatore S, Cuervo-Arango J. The effect of mare's age on multiple ovulation rate, embryo recovery, post-transfer pregnancy rate, and interovulatory interval in a commercial embryo transfer program in Argentina. *Anim. Reprod. Sci.* 2015. 158: 53-59.
143. Martín-Fernández F, Bermejo-Santos A, Bragg-Gonzalo L, Briz C G, Serrano-Saiz E, Nieto M. Role of Nrp1 in controlling cortical inter-hemispheric circuits. *eLife.* 2022. 11: e69776.
144. Martinez M M, Costa M, Corva P M. Analysis of Genetic Variability in the Argentine Polo Horse With a Panel of Microsatellite Markers. *J Equine Vet Sci.* 2021. 96: 103320.
145. Maserati M, Mutto A. In Vitro Production of Equine Embryos and Cloning: Today's Status. *J Equine Vet Sci.* 2016. 41: 42-50.
146. McCue M E, Bannasch D L, Petersen J L, Gurr J, Bailey E, Binns M M, Distl O, Guérin G, Hasegawa T, Hill E W, Leeb T, Lindgren G, Penedo M C T, Røed K H, Ryder O A, Swinburne J E, Tozaki T, Valberg S J, Vaudin M, Lindblad-Toh K, Wade C M, Mickelson J R. A high density SNP array for the domestic horse and extant Perissodactyla: Utility for association mapping, genetic diversity, and phylogeny studies. *PLoS Genet.* 2012. 8: e1002451.
147. McGivney B A, Hernandez B, Katz L M, MacHugh D E, McGovern S P, Parnell A C, Wiencko H L, Hill E W. A genomic prediction model for racecourse starts in the Thoroughbred horse. *Anim. Genet.* 2019. 50: 347-357.
148. McNeill E M, Klöckner-Bormann M, Roesler E C, Talton L E, Moechars D, Clagett-Dame M. Nav2 hypomorphic mutant mice are ataxic and exhibit abnormalities in cerebellar development. *Dev. Biol.* 2011. 353: 331-43.
149. McQuillan R, Leutenegger A L, Abdel-Rahman R, Franklin C S, Pericic M, Barac-Lauc L, Smolej-Narancic N, Janicijevic B, Polasek O, Tenesa A, MacLeod A K, Farrington S M, Rudan P, Hayward C, Vitart V, Rudan I, Wild S H, Dunlop M G, Wright A F, Campbell H, Wilson J F. Runs of Homozygosity in European Populations. *Am. J. Hum. Genet.* 2008. 83: 359-372.
150. Medeiros B R, Bertoli C D, Garbade P, McManus C M. Brazilian Sport Horse: Pedigree Analysis of the Brasileiro de Hipismo Breed. *Ital. J. Anim. Sci.* 2014. 13: 3146.
151. Melgarejo I, Valera M, Molina A, Rodero A. GESCAB: Software para la cría del caballo Español. Universidad de Córdoba, Servicio de Publicaciones Archivos de zootecnia 2000. 49
152. Metzger J, Rau J, Naccache F, Bas Conn L, Lindgren G, Distl O. Genome data uncover four synergistic key regulators for extremely small body size in horses. *BMC Genomics.* 2018. 19: 492.

153. Meyermans R, Gorssen W, Buys N, Janssens S. How to study runs of homozygosity using PLINK? A guide for analyzing medium density SNP data in livestock and pet species. *BMC Genomics*. 2020. 21:
154. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, Thomas P D. PANTHER version 11: expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. *Nucleic Acids Res*. 2017. 45: D183-d189.
155. Mi Y, Shi Z, Li J. Spata19 is critical for sperm mitochondrial function and male fertility. *Mol. Reprod. Dev*. 2015. 82: 907-13.
156. Moffat J J, Ka M, Jung E M, Smith A L, Kim W Y. The role of MACF1 in nervous system development and maintenance. *Semin. Cell Dev. Biol*. 2017. 69: 9-17.
157. Momen M, Brounts S H, Binversie E E, Sample S J, Rosa G J M, Davis B W, Muir P. Selection signature analyses and genome-wide association reveal genomic hotspot regions that reflect differences between breeds of horse with contrasting risk of degenerative suspensory ligament desmitis. *G3 (Bethesda)*. 2022. 12: 179.
158. Momozawa Y, Kusunose R, Kikusui T, Takeuchi Y, Mori Y. Assessment of equine temperament questionnaire by comparing factor structure between two separate surveys. *Appl. Anim. Behav. Sci*. 2005a. 92: 77-84.
159. Momozawa Y, Takeuchi Y, Kusunose R, Kikusui T, Mori Y. Association between equine temperament and polymorphisms in dopamine D4 receptor gene. *Mamm. Genome*. 2005b. 16: 538-544.
160. Moon S, Lee J W, Shin D, Shin K Y, Kim J, Choi I Y, Kim J, Kim H. A Genome-wide scan for selective sweeps in racing horses. *Asian-Australas. J. Anim. Sci*. 2015. 28: 1525-1531.
161. MOP. Museum of Polo - History of Polo. <http://www.polomuseum.com/sport-polo/history-polo>
162. Mosher D S, Quignon P, Bustamante C D, Sutter N B, Mellersh C S, Parker H G, Ostrander E A. A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. *PLoS Genet*. 2007. 3: e79.
163. Mota R R, Lopes P S, Marques L F A, da Silva L P, de Resende M D V, Torres R A. The influence of animals from embryo transfer on the genetic evaluation of growth in Simmental beef cattle by using multi-trait models. *Genet. Mol. Biol*. 2013. 36: 43-49.
164. Moureaux S, Verrier É, Ricard A, Mériaux J C. Genetic variability within French race and riding horse breeds from genealogical data and blood marker polymorphisms. *Genet. Sel. Evol*. 1996. 28: 83-102.
165. Muhammad Aslam M K, Sharma V K, Pandey S, Kumaresan A, Srinivasan A, Datta T K, Mohanty T K, Yadav S. Identification of biomarker candidates for fertility in spermatozoa of crossbred bulls through comparative proteomics. *Theriogenology*. 2018. 119: 43-51.
166. Murphy B A, Wagner A L, McGlynn O F, Kharazyan F, Browne J A, Elliott J A. Exercise influences circadian gene expression in equine skeletal muscle. *Vet. J*. 2014. 201: 39-45.

167. Myhre J L, Hills J A, Jean F, Pilgrim D B. Unc45b is essential for early myofibrillogenesis and costamere formation in zebrafish. *Dev. Biol.* 2014. 390: 26-40.
168. Neogen. GGP Equine. <https://www.neogen.com/es/categories/genotyping-arrays/ggp-equine/>.
169. Nicholas F W. *Genética Veterinaria*. Acribia, S.A., 1987.
170. Nicholas F W. Genetic improvement through technology reproductive Anim. *Reprod. Sci.* 1996. 42: 205-214.
171. Nielsen R. Molecular signatures of natural selection. *Annu. Rev. Genet.* 2005. 39: 197-218.
172. Nogueira M B, McManus C, de Faria D A, Santos S A O, Ianella P, Paiva S R. Fine-scale genetic diversity of the Brazilian Pantaneiro horse breed adapted to flooded regions. *Trop. Anim. Health Prod.* 2021. 53: 525.
173. Noh K, Lee H, Choi T Y, Joo Y, Kim S J, Kim H, Kim J Y, Jahng J W, Lee S, Choi S Y, Lee S J. Negr1 controls adult hippocampal neurogenesis and affective behaviors. *Mol. Psychiatry.* 2019. 24: 1189-1205.
174. Obholz K L, Akopyan A, Waymire K G, MacGregor G R. FNDC3A is required for adhesion between spermatids and Sertoli cells. *Dev. Biol.* 2006. 298: 498-513.
175. Pan Y, Wang K S, Aragam N. NTM and NR3C2 polymorphisms influencing intelligence: family-based association studies. *Prog. Neuro-Psychopharmacol. Biol. Psychiatry.* 2011. 35: 154-60.
176. Paredes-Sánchez F A, Sifuentes-Rincón A M, Casas E, Arellano-Vera W, Parra-Bracamonte G M, Riley D G, Welsh T H, Jr., Randel R D. Novel genes involved in the genetic architecture of temperament in Brahman cattle. *PLoS One.* 2020. 15: e0237825.
177. Parkes R S, Witte T H. The foot-surface interaction and its impact on musculoskeletal adaptation and injury risk in the horse. *Equine Vet. J.* 2015. 47: 519-25.
178. Parrish C. Origin, diffusion and development of polo: an East to West cultural flow. *Sport Soc.* 2018. 21: 1158-1169.
179. Peral García P. Caracterización racial de equinos de pura raza criolla: Su comparación con los equinos de pura raza española mediante el análisis de sus polimorfismos genéticos. Universidad Nacional de La Plata - Facultad de Ciencias Naturales y Museo 1994.
180. Peral García P, Kienast M, Gortari C, Díaz S, Maderna R, Duluot F N. Marcadores Genéticos en Equinos: sus Aplicaciones. *Analecta Vet.* 1995. XV: 11-21.
181. Perdomo-González D I, Laseca N, Demyda-Peyrás S, Valera M, Cervantes I, Molina A. Fine-tuning genomic and pedigree inbreeding rates in equine population with a deep and reliable stud book: the case of the Pura Raza Española horse. *J. Animal Sci. Biotechnol.* 2022. 13: 127.
182. Perdomo-González D I, Molina A, Sánchez-Guerrero M J, Bartolomé E, Varona L, Valera M. Genetic inbreeding depression load for fertility traits in Pura Raza Española mares. *J. Anim. Sci.* 2021. 99: skab316.

183. Peripolli E, Munari D P, Silva M V G B, Lima A L F, Irgang R, Baldi F. Runs of homozygosity: current knowledge and applications in livestock. *Anim. Genet.* 2016. 48: 255-271.
184. Peripolli E, Stafuzza N B, Munari D P, Lima A L F, Irgang R, Machado M A, Panetto J, Ventura R V, Baldi F, da Silva M. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle. *BMC Genomics.* 2018. 19: 34.
185. Petersen J L, Mickelson J R, Cleary K D, McCue M E. The american quarter horse: Population structure and relationship to the thoroughbred. *J. Hered.* 2014. 105: 148-162.
186. Petersen J L, Mickelson J R, Cothran E G, Andersson L S, Axelsson J, Bailey E, Bannasch D, Binns M M, Borges A S, Brama P, da Câmara Machado A, Distl O, Felicetti M, Fox-Clipsham L, Graves K T, Guérin G, Haase B, Hasegawa T, Hemmann K, Hill E W, Leeb T, Lindgren G, Lohi H, Lopes M S, McGivney B A, Mikko S, Orr N, Penedo M C T, Piercy R J, Raekallio M, Rieder S, Røed K H, Silvestrelli M, Swinburne J, Tozaki T, Vaudin M, Wade C, McCue M E. Genetic Diversity in the Modern Horse Illustrated from Genome-Wide SNP Data. *PLoS One.* 2013a. 8: e54997.
187. Petersen J L, Mickelson J R, Rendahl A K, Valberg S J, Andersson L S, Axelsson J, Bailey E, Bannasch D, Binns M M, Borges A S, Brama P, da Câmara Machado A, Capomaccio S, Cappelli K, Cothran E G, Distl O, Fox-Clipsham L, Graves K T, Guérin G, Haase B, Hasegawa T, Hemmann K, Hill E W, Leeb T, Lindgren G, Lohi H, Lopes M S, McGivney B A, Mikko S, Orr N, Penedo M C T, Piercy R J, Raekallio M, Rieder S, Røed K H, Swinburne J, Tozaki T, Vaudin M, Wade C M, McCue M E. Genome-Wide Analysis Reveals Selection for Important Traits in Domestic Horse Breeds. *PLoS Genet.* 2013b. 9: e1003211.
188. Pirosanto Y, Valera M, Molina Alcalá A, Dorado J, Demyda Peyrás S. 23 Sperm quality of Pure Spanish stallions is affected by inbreeding coefficient and age. *Reprod. Fertil. Dev.* 2020. 32: 137.
189. Pjontek J, Kadlečík O, Kasarda R, Horný M. Pedigree analysis in four Slovak endangered horse breeds. *Czech J. Anim. Sci.* 2012. 57: 54-64.
190. Polytaichou C, Hatziapostolou M, Yau T O, Christodoulou N, Hinds P W, Kottakis F, Sanidas I, Tsihchlis P N. Akt3 induces oxidative stress and DNA damage by activating the NADPH oxidase via phosphorylation of p47(phox). *P Natl Acad Sci USA.* 2020. 117: 28806-28815.
191. Poncet P A, Pfister W, Muntwyler J, Glowatzki-Mullis M L, Gaillard C. Analysis of pedigree and conformation data to explain genetic variability of the horse breed Franches-Montagnes. *J. Anim. Breed. Genet.* 2006. 123: 114-21.
192. Poyato-Bonilla J, Laseca N, Demyda Peyrás S, Molina Alcalá A, Valera M. 500 years of breeding in the Carthusian Strain of Pura Raza Español horse: An evolutionary analysis using genealogical and genomic data. *J. Anim. Breed. Genet.* 2021. 139: 84-99.
193. Puechmaille S J. The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Mol. Ecol. Resour.* 2016. 16: 608-27.
194. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira M A R, Bender D, Maller J, Sklar P, De Bakker P I W, Daly M J, Sham P C. PLINK: A tool set for whole-

- genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 2007. 81: 559-575.
195. Qanbari S, Gianola D, Hayes B, Schenkel F, Miller S, Moore S, Thaller G, Simianer H. Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. *BMC Genomics.* 2011. 12: 318.
 196. Qu W R, Sun Q H, Liu Q Q, Jin H J, Cui R J, Yang W, Song B, Li B J. Role of CPEB3 protein in learning and memory: new insights from synaptic plasticity. *Aging.* 2020. 12: 15169-15182.
 197. R-Core-Team. R: A language and environment for statistical computing. 2023. V. R Foundation for Statistical Computing, Austria.
 198. Rahmatalla S A, Arends D, Reissmann M, Wimmers K, Reyer H, Brockmann G A. Genome-wide association study of body morphological traits in Sudanese goats. *Anim. Genet.* 2018. 49: 478-482.
 199. Ramasamy R K, Ramasamy S, Bindroo B B, Naik V G. STRUCTURE PLOT: a program for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus.* 2014. 3: 431.
 200. Rhee J W, Arata A, Selleri L, Jacobs Y, Arata S, Onimaru H, Cleary M L. Pbx3 deficiency results in central hypoventilation. *Am. J. Pathol.* 2004. 165: 1343-50.
 201. Riera F L, Roldán J E, Gomez J, Hinrichs K. Factors affecting the efficiency of foal production in a commercial oocyte transfer program. *Theriogenology.* 2016. 85: 1053-1062.
 202. Rooney M F, Hill E W, Kelly V P, Porter R K. The "speed gene" effect of myostatin arises in Thoroughbred horses due to a promoter proximal SINE insertion. *PLoS One.* 2018. 13: e0205664.
 203. Rooney M F, Porter R K, Katz L M, Hill E W. Skeletal muscle mitochondrial bioenergetics and associations with myostatin genotypes in the Thoroughbred horse. *PLoS One.* 2017. 12: e0186247.
 204. Ropka-Molik K, Stefaniuk-Szmukier M, Piórkowska K, Szmatoła T, Bugno-Poniewierska M. Molecular characterization of the apoptosis-related SH3RF1 and SH3RF2 genes and their association with exercise performance in Arabian horses. *BMC Vet. Res.* 2018. 14: 237.
 205. Rosenfeld D J, Garthwaite T L. Central administration of motilin stimulates feeding in rats. *Physiol. Behav.* 1987. 39: 753-6.
 206. Sánchez-Guerrero M J, Cervantes I, Molina A, Gutiérrez J P, Valera M. Designing an early selection morphological linear traits index for dressage in the Pura Raza Español horse. *Animal.* 2017. 11: 948-957.
 207. Sánchez-Guerrero M J, Molina A, Gómez M D, Peña F, Valera M. Relationship between morphology and performance: Signature of mass-selection in Pura Raza Español horse. *Livest. Sci.* 2016. 185: 148-155.
 208. Santos W B, Schettini G P, Maiorano A M, Bussiman F O, Balieiro J C C, Ferraz G C, Pereira G L, Baldassini W A, Neto O R M, Oliveira H N, Curi R A. Genome-wide scans for signatures of selection in Mangalarga Marchador horses using high-throughput SNP genotyping. *BMC Genomics.* 2021. 22: 737.

209. Saravanan K A, Panigrahi M, Kumar H, Bhushan B, Dutt T, Mishra B P. Selection signatures in livestock genome: A review of concepts, approaches and applications. *Livest. Sci.* 2020. 241: 104257.
210. Schaefer R J, McCue M E. Equine Genotyping Arrays. *Vet. Clin. North Am. Equine Pract.* 2020. 36: 183-193.
211. Schurink A, Shrestha M, Eriksson S, Bosse M, Bovenhuis H, Back W, Johansson A M, Ducro B J. The Genomic Makeup of Nine Horse Populations Sampled in the Netherlands. *Genes.* 2019. 10: 480.
212. Sevinga M, Vrijenhoek T, Hesselinks J W, Barkema H W, Groen A F. Effect of inbreeding on the incidence of retained placenta in Friesian horses. *J. Anim. Sci.* 2004. 82: 982-6.
213. Shin D H, Lee J W, Park J E, Choi I Y, Oh H S, Kim H J, Kim H. Multiple genes related to muscle identified through a joint analysis of a two-stage genome-wide association study for racing performance of 1,156 Thoroughbreds. *Asian-Australas. J. Anim. Sci.* 2015. 28: 771-781.
214. Sitzenstock F, Rathke I, Ytournel F, Simianer H. The potential of embryo transfer in a German horse-breeding programme. *J. Anim. Breed. Genet.* 2013. 130: 199-208.
215. Sole M, Cervantes I, Gutierrez J P, Gomez M D, Valera M. Estimation of genetic parameters for morphological and functional traits in a Menorca horse population. *Span. J. Agric. Res.* 2014. 12: 125-132.
216. Squires E L, McCue P M, Vanderwall D. The current status of equine embryo transfer. *Theriogenology.* 1999. 51: 91-104.
217. StatSoft. Statistica software. 2012.
218. Stratil A, Kopečný M. Genomic organization, sequence and polymorphism of the porcine myostatin (GDF8; MSTN) gene. *Anim. Genet.* 1999. 30: 468-70.
219. Stuczyńska A, Piórkowska K, Tyra M, Żukowski K. The effect of QTL-rich region polymorphisms identified by targeted DNA-seq on pig production traits. *Mol. Biol. Rep.* 2018. 45: 361-371.
220. Sultana F, Morse L R, Picotto G, Liu W, Jha P K, Odgren P R, Battaglino R A. Snx10 and PIKfyve are required for lysosome formation in osteoclasts. *J. Cell. Biochem.* 2020. 121: 2927-2937.
221. Sumreddee P, Hay E H, Toghiani S, Roberts A, Aggrey S E, Rekaya R. Grid search approach to discriminate between old and recent inbreeding using phenotypic, pedigree and genomic information. *BMC Genomics.* 2021. 22: 538.
222. Takahashi H, Rikimaru K, Komatsu M, Uemoto Y, Suzuki K. Association between Motilin Receptor Gene Haplotypes and Growth Traits in Japanese Hinai-dori Crossbred Chickens. *Asian-Australas. J. Anim. Sci.* 2014. 27: 316-23.
223. Tang F, Fan J, Zhang X, Zou Z, Xiao D, Li X. The Role of Vti1a in Biological Functions and Its Possible Role in Nervous System Disorders. *Front. Mol. Neurosci.* 2022. 15: e918664.

224. Thomson P C, Hayek A R, Jones B, Evans D L, McGreevy P D. Number, causes and destinations of horses leaving the Australian Thoroughbred and Standardbred racing industries. *Aust Vet J.* 2014. 92: 303-11.
225. Todd E, Ho S, Thomson P, Ang R, Velie B, Hamilton N. Founder-specific inbreeding depression affects racing performance in Thoroughbred horses. *Sci. Rep.* 2018. 8: 6167.
226. Tozaki T, Sato F, Ishimaru M, Kikuchi M, Kakoi H, Hirota K I, Nagata S I. Sequence variants of BIEC2-808543 near LCORL are associated with body composition in Thoroughbreds under training. *J. Equine Sci.* 2016. 27: 107-114.
227. Uhl G R, Martinez M J. PTPRD: neurobiology, genetics, and initial pharmacology of a pleiotropic contributor to brain phenotypes. *Ann. N. Y. Acad. Sci.* 2019. 1451: 112-129.
228. Valera M, Molina A, Gutiérrez J P, Gómez J, Goyache F. Pedigree analysis in the Andalusian horse: Population structure, genetic variability and influence of the Carthusian strain. *Livest. Prod. Sci.* 2005. 95: 57-66.
229. Vanvanhossou S F U, Scheper C, Dossa L H, Yin T, Brügemann K, König S. A multi-breed GWAS for morphometric traits in four Beninese indigenous cattle breeds reveals loci associated with conformation, carcass and adaptive traits. *BMC Genomics.* 2020. 21: 783.
230. Velie B D, Fegraeus K J, Solé M, Rosengren M K, Røed K H, Ihler C-F, Strand E, Lindgren G. A genome-wide association study for harness racing success in the Norwegian-Swedish coldblooded trotter reveals genes for learning and energy metabolism. *BMC Genet.* 2018. 19: 80.
231. Vicente A A, Carolino N, Gama L T. Genetic diversity in the Lusitano horse breed assessed by pedigree analysis. *Livest. Sci.* 2012. 148: 16-25.
232. Vostrá-Vydrová H, Vostrý L, Hofmanová B, Krupa E, Veselá Z, Schmidová J. Genetic diversity within and gene flow between three draught horse breeds using genealogical information. *Czech J. Anim. Sci.* 2016. 61: 462-472.
233. Wang X, Xu M, Xu Q, Yang F, Tang H, Shao C, Wang L, Wang Y, Deng J, Wang S. Rictor is involved in Ctnnd2 deletion-induced impairment of spatial learning and memory but not autism-like behaviors. *Front. Biosci.* 2021. 26: 335-346.
234. Weir B S, Cockerham C C. Estimating F-statistics for the analysis of population structure. *Evolution.* 1984. 38: 1358-1370.
235. White S H, Johnson S E, Bobel J M, Warren L K. Dietary selenium and prolonged exercise alter gene expression and activity of antioxidant enzymes in equine skeletal muscle. *J. Anim. Sci.* 2016. 94: 2867-78.
236. Whitlock M, Lotterhos K. Reliable Detection of Loci Responsible for Local Adaptation: Inference of a Null Model through Trimming the Distribution of FST. *Am. Nat.* 2015. 186: 24-36.
237. Wickens C, Brooks S A. Genetics of Equine Behavioral Traits. *Vet. Clin. North Am. Equine Pract.* 2020. 36: 411-424.
238. Wickham H. *ggplot2: Elegant Graphics for Data Analysis.* Springer-Verlag New York. 2016.
239. Wright S. Evolution in Mendelian Populations. *Genetics.* 1931. 16: 97-159.

240. Wright S. The genetical structure of populations. *Ann. Eugen.* 1951. 15: 323-54.
241. Wu X-L, Xu J, Feng G, Wiggans G R, Taylor J F, He J, Qian C, Qiu J, Simpson B, Walker J, Bauck S. Optimal Design of Low-Density SNP Arrays for Genomic Prediction: Algorithm and Applications. *PLoS One.* 2016. 11: e0161719.
242. Yokomori T, Ohnuma A, Tozaki T, Segawa T, Ito T. Identification of Personality-Related Candidate Genes in Thoroughbred Racehorses Using a Bioinformatics-Based Approach Involving Functionally Annotated Human Genes. *Animals.* 2023. 13: 769.
243. Yoon S H, Lee W, Ahn H, Caetano-Anolles K, Park K D, Kim H. Origin and spread of Thoroughbred racehorses inferred from complete mitochondrial genome sequences: Phylogenomic and Bayesian coalescent perspectives. *PLoS One.* 2018. 13: e0203917.
244. Yu H, Zhao Y, Iqbal A, Xia L, Bai Z, Sun H, Fang X, Yang R, Zhao Z. Effects of polymorphism of the GPAM gene on milk quality traits and its relation to triglyceride metabolism in bovine mammary epithelial cells of dairy cattle. *Arch. Anim. Breed.* 2021. 64: 35-44.
245. Yu H, Zhao Z, Yu X, Li J, Lu C, Yang R. Bovine lipid metabolism related gene GPAM: Molecular characterization, function identification, and association analysis with fat deposition traits. *Gene.* 2017. 609: 9-18.
246. Zechner P, Sölkner J, Bodo I, Druml T, Baumung R, Achmann R, Marti E, Habe F, Brem G. Analysis of diversity and population structure in the Lipizzan horse breed based on pedigree information. *Livest. Prod. Sci.* 2002. 77: 137-146.
247. Zechner P, Zohman F, Sölkner J, Bodo I, Habe F, Marti E, Brem G. Morphological description of the Lipizzan horse population. *Livest. Prod. Sci.* 2001. 69: 163-177.
248. Zeng L, Ming C, Li Y, Su L Y, Su Y H, Otecko N O, Liu H Q, Wang M S, Yao Y G, Li H P, Wu D D, Zhang Y P. Rapid Evolution of Genes Involved in Learning and Energy Metabolism for Domestication of the Laboratory Rat. *Mol. Biol. Evol.* 2017. 34: 3148-3153.
249. Zhang L, Ma X, Xuan J, Wang H, Yuan Z, Wu M, Liu R, Zhu C, Wei C, Zhao F, Du L. Identification of MEF2B and TRHDE Gene Polymorphisms Related to Growth Traits in a New Ujumqin Sheep Population. *PLoS One.* 2016. 11: e0159504.
250. Zhang Q, Guldbrandtsen B, Bosse M, Lund M S, Sahana G. Runs of homozygosity and distribution of functional variants in the cattle genome. *BMC Genomics.* 2015. 16: 542.
251. Zhao F, McParland S, Kearney F, Du L, Berry D P. Detection of selection signatures in dairy and beef cattle using high-density genomic information. *Genet. Sel. Evol.* 2015. 47: 49.
252. Ziegler G C, Ehlis A C, Weber H, Vitale M R, Zöllner J E M, Ku H P, Schiele M A, Kürbitz L I, Romanos M, Pauli P, Kalisch R, Zwanzger P, Domschke K, Fallgatter A J, Reif A, Lesch K P. A Common CDH13 Variant Is Associated with Low Agreeableness and Neural Responses to Working Memory Tasks in ADHD. *Genes.* 2021. 12: 1356.