

Principios básicos para la utilización de los fenotipos en genética molecular

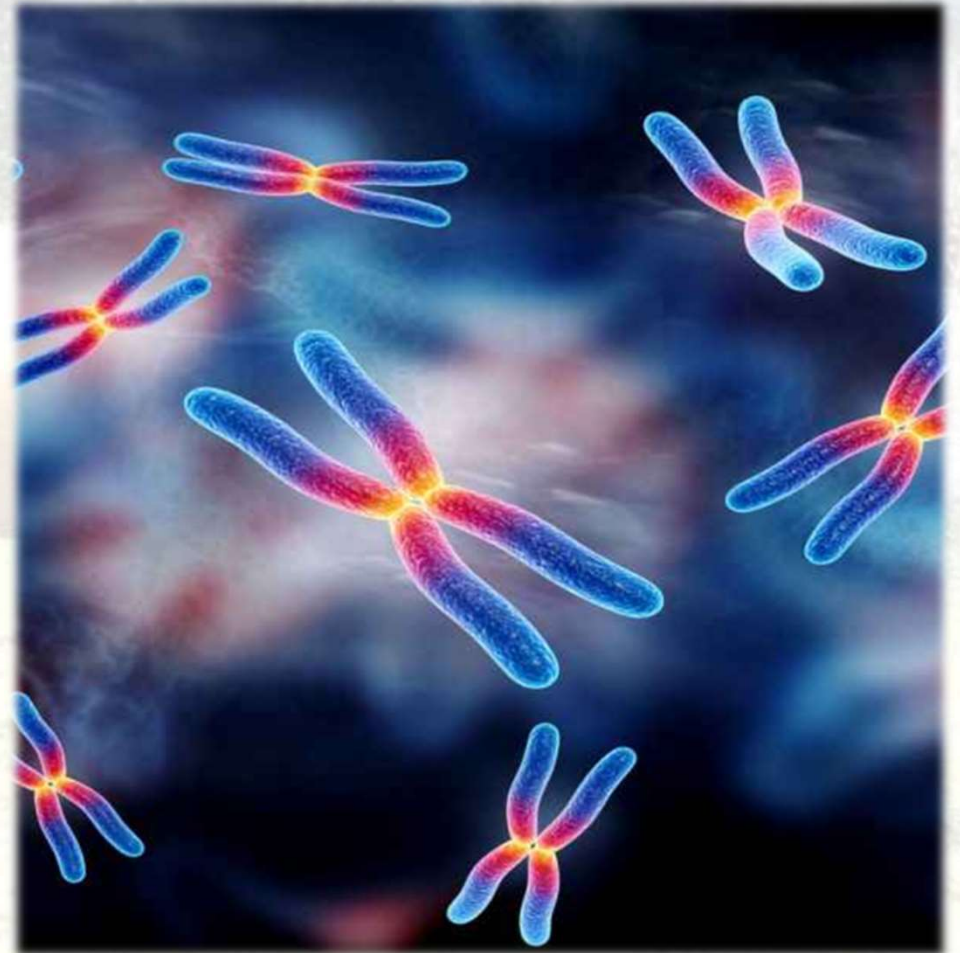




Breve introducción a la Genética Molecular y sus aplicaciones en ganadería

La genética molecular es el campo de la biología que estudia la estructura y la función de los genes a nivel molecular.

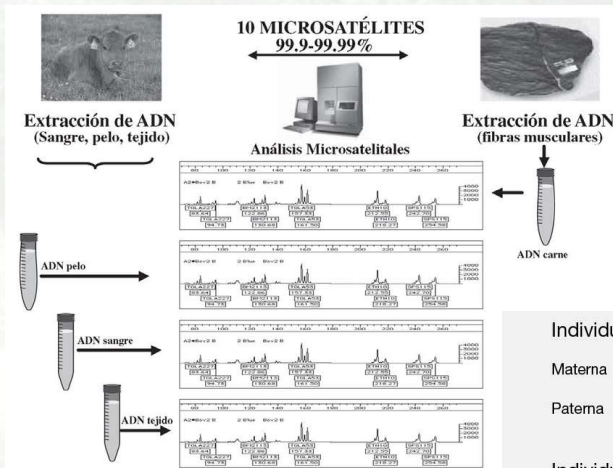
La genética molecular emplea los métodos de la genética y la biología molecular.



Aplicaciones de la Genética Molecular

- **Identificación Individual y Pruebas de Paternidades.**
- **Elaboración de Mapas Genéticos y Genómica Comparativa**
- **Estudios de Genética Poblacional**
- **Asignación de Individuos a una Raza**
- **Estudios sobre biodiversidad**
- **Estudios Forenses**

Marcadores Moleculares



Individuo 1

Materna ...CGATATTCC**T**ATCGAATGTC...

Paterna ...CGATATTCC**C**ATCGAATGTC...

Individuo 2

Materna ...CGATATTCC**C**ATCGAATGTC...

Paterna ...CGATATTCC**C**ATCGAATGTC...

Individuo 3

Materna ...CGATATTCC**T**ATCGAATGTC...

Paterna ...CGATATTCC**T**ATCGAATGTC...

Individuo 4

Materna ...CGATATTCC**C**ATCGAATGTC...

Paterna ...CGATATTCC**T**ATCGAATGTC...

Primer 1

CACCTGATATCTGGTA
GTGGACTATAGACCAT---ACACACACACACAC---GCTGTGATGGTCTAC

Microsatélite

CACCTGATATCTGGTA---TGTTGTGTGTGTGTG---CGACACTACCAGATG
GCTGTGATGGTCTAC

Primer 2

Son regiones específicas que “marcan” o sirven de referencia para detectar variaciones que pueden asociarse positiva o negativamente con un rasgo productivo en cualquier especie animal.

Existe gran diversidad de marcadores



RFLP (Restriction Fragment Length Polymorphism o Polimorfismo en la longitud de fragmentos de restricción)

SSLP (Simple sequence length polymorphism o Polimorfismo en la longitud de secuencias simples)

AFLP (Amplified fragment length polymorphism o Polimorfismo en la longitud de fragmentos amplificados)

RAPD (Random amplification of polymorphic DNA o Amplificación aleatoria de ADN polimórfico)

VNTR (Variable number tandem repeat o Número variable de repeticiones en tándem. Minisatélite)

SSR (Simple sequence repeat o Repetición de secuencia simple. Microsatélite)

STR (Short tandem repeat o Repeticiones cortas en tándem. Microsatélite)

SNP (Single nucleotide polymorphism o Polimorfismo de nucleótido simple)

SFP (Single feature polymorphism o Polimorfismos de Característica Única)

TRAPs (Target Region Amplification Polymorphism, en español Polimorfismos para la amplificación de regiones blanco)

DArT (Diversity Arrays Technology, es español Tecnología de Vectores, o Matrices, de Diversidad)

RAD (Restriction site associated DNA markers o marcadores de ADN asociados a sitios de restricción)

Genetic characterization of Latin-American Creole cattle using microsatellite markers

J. V. Delgado¹, A. M. Martínez¹, A. Acosta², L. A. Álvarez³, E. Armstrong⁴, E. Camacho⁵, J. Cañón⁶, O. Cortés⁶, S. Dunner⁶, V. Landi¹, J. R. Marques⁷, I. Martín-Burriel⁸, O. R. Martínez^{9,10}, R. D. Martínez¹¹, L. Melucci^{12,13}, J. E. Muñoz², M. C. T. Penedo¹⁴, A. Postiglioni⁴, J. Quiróz¹⁵, C. Rodellar⁸, P. Sponenberg¹⁶, O. Uffo², R. Ulloa-Arvizu¹⁷, J. L. Vega-Pla¹⁸, A. Villalobos¹⁹, D. Zambrano²⁰, P. Zaragoza⁸, L. T. Gama²¹ and C. Ginja^{14,21,22}

¹Departamento de Genética, Universidad de Córdoba, Campus Rabanales Edificio Gregor Mendel, 14071-Córdoba, Spain. ²Centro Nacional de Sanidad Agropecuaria (CENSA), San José de las Lajas, Apdo. 10, 32700-La Habana, Cuba. ³Universidad Nacional de Colombia, Sede Palmira, Cra. 32 No 12-00, Palmira, Valle del Cauca, Colombia. ⁴Área Genética, Departamento de Genética y Mejora Animal, Facultad de Veterinaria-U de la R, Montevideo, Uruguay. ⁵IFAPA centro Alameda del Obispo, Avda. Menéndez Pidal s/n, 14004-Córdoba, Spain. ⁶Departamento de Producción Animal, Facultad de Veterinaria, Universidad Complutense de Madrid, Avda. Puerta de Hierro, s/n 28040-Madrid, Spain. ⁷EMBRAPA Amazônia Oriental, Embrapa Amazônia Oriental, Trav. Dr. Enéas Pinheiro s/no. Caixa Postal, 48 Belém, Pará, Brazil. ⁸Laboratorio de Genética Bioquímica, Facultad de Veterinaria, Universidad de Zaragoza, Miguel Servet, 177, 20013-Zaragoza, Spain. ⁹Universidad Federal Rural de Pernambuco, Rua Dom Manoel de Medeiros, s/n, Dois Irmãos - CEP: 52171-900 - Recife/PE, Brasil. ¹⁰Centro Multidisciplinario de Investigaciones Tecnológicas, Dirección General de Investigación Científica y Tecnológica, Universidad Nacional de Asunción, Km. 11 - Campus San Lorenzo, Paraguay. ¹¹Genética Animal, Facultad de Ciencias Agrarias, Universidad Nacional de Lomas de Zamora, Ruta 4 - Km. 2 - Llavallol (CP 1836), Argentina. ¹²Facultad Ciencias Agrarias, Universidad Nacional de Mar del Plata, Ruta 226 Km 73,5 (7620) Balcarce, Argentina. ¹³Estación Experimental Agropecuaria Balcarce, Instituto Nacional de Tecnología Agropecuaria, Ruta 226 Km 73, 5 (7620) Balcarce, Argentina. ¹⁴Veterinary Genetics Laboratory, University of California, One Shields Avenue, Davis, CA 95616, USA. ¹⁵Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Av. Progreso 5 Col. Barrio de Santa Catarina, Coyoacán, México D.F. C.P. 04010, México. ¹⁶Virginia-Maryland Regional College of Veterinary Medicine, Virginia Tech, Duck Pond Drive, 0442, Blacksburg, VA 24061, USA. ¹⁷Facultad de Medicina Veterinaria y Zootecnia, Universidad Nacional Autónoma de México, Av. Universidad 3000, Ciudad Universitaria, México D.F. C.P. 04510, México. ¹⁸Laboratorio de Investigación Aplicada, Cría Caballar de las Fuerzas Armadas, Apartado de Correos 2087, 14080-Córdoba, Spain. ¹⁹Instituto de Investigación Agropecuaria, Estación Experimental El Ejido, Los Santos, Panamá. ²⁰Universidad Técnica Estatal de Quevedo, Quevedo, Ecuador. ²¹Departamento de Genética e Melhoramento Animal, Instituto Nacional dos Recursos Biológicos, Fonte Boa, 2005-048 Vale de Santarém, Portugal. ²²Instituto Superior de Agronomia, Tapada da Ajuda, 1349-017 Lisboa, Portugal.

Summary

Genetic diversity in and relationships among 26 Creole cattle breeds from 10 American countries were assessed using 19 microsatellites. Heterozygosity, *F*-statistics estimates, genetic distances, multivariate analyses and assignment tests were performed. The levels of within-breed diversity detected in Creole cattle were considerable and higher than those previously reported for European breeds, but similar to those found in other Latin American breeds. Differences among breeds accounted for 8.4% of the total genetic variability. Most breeds clustered separately when the number of pre-defined populations was 21 (the most probable *K* value), with the exception of some closely related breeds that shared the same cluster and others that were admixed. Despite the high genetic diversity detected, significant inbreeding was also observed within some breeds, and heterozygote excess was detected in

Genetic Footprints of Iberian Cattle in America 500 Years after the Arrival of Columbus

Amparo M. Martínez¹, Luis T. Gama^{2,3}, Javier Cañón⁴, Catarina Ginja⁵, Juan V. Delgado¹, Susana Dunner⁴, Vincenzo Landi¹, Inmaculada Martín-Burriel⁶, M. Cecilia T. Penedo⁷, Clementina Rodellar⁶, Jose Luis Vega-Pla^{8*}, Atzel Acosta⁹, Luz A. Álvarez¹⁰, Esperanza Camacho¹¹, Oscar Cortés⁴, Jose R. Marques¹², Roberto Martínez¹³, Ruben D. Martínez¹⁴, Lilia Melucci^{15,16}, Guillermo Martínez-Velázquez¹⁷, Jaime E. Muñoz¹⁰, Alicia Postiglioni¹⁸, Jorge Quiróz¹⁷, Philip Sponenberg¹⁹, Odalys Uffo⁹, Axel Villalobos²⁰, Delsito Zambrano²¹, Pilar Zaragoza⁶

¹Departamento de Genética, Universidad de Córdoba, Córdoba, Spain. ²L-INIA, Instituto Nacional dos Recursos Biológicos, Fonte Boa, Vale de Santarém, Portugal. ³CIISA - Faculdade de Medicina Veterinária, Universidade Técnica de Lisboa, Lisboa, Portugal. ⁴Departamento de Produção Animal, Faculdade de Veterinária, Universidade Complutense de Madrid, Madrid, Spain. ⁵Centre for Environmental Biology, Faculty of Sciences, University of Lisbon & Molecular Biology Group, Instituto Nacional de Recursos Biológicos, INIA, Lisbon, Portugal. ⁶Laboratório de Genética Bioquímica, Faculdade de Veterinária, Universidade de Zaragoza, Zaragoza, Spain. ⁷Veterinary Genetics Laboratory, University of California Davis, Davis, California, United States of America. ⁸Laboratório de Investigação Aplicada, Cría Caballar de las Fuerzas Armadas, Córdoba, Spain. ⁹Centro Nacional de Sanidad Agropecuaria, San José de las Lajas, La Habana, Cuba. ¹⁰Universidad Nacional de Colombia, Sede Palmira, Valle del Cauca, Colombia. ¹¹IFAPA, Centro Alameda del Obispo, Córdoba, Spain. ¹²EMBRAPA Amazônia Oriental, Belém, Pará, Brazil. ¹³Centro Multidisciplinario de Investigaciones Tecnológicas, Dirección General de Investigación Científica y Tecnológica, Universidad Nacional de Asunción, San Lorenzo, Paraguay. ¹⁴Genética Animal, Facultad de Ciencias Agrarias, Universidad Nacional de Lomas de Zamora, Lomas de Zamora, Argentina. ¹⁵Facultad Ciencias Agrarias, Universidad Nacional de Mar del Plata, Balcarce, Argentina. ¹⁶Estación Experimental Agropecuaria Balcarce, Instituto Nacional de Tecnología Agropecuaria, Balcarce, Argentina. ¹⁷Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias, Coyoacán, México. ¹⁸Área Genética, Departamento de Genética y Mejora Animal, Facultad de Veterinaria, Universidad de la República, Montevideo, Uruguay. ¹⁹Virginia-Maryland Regional College of Veterinary Medicine, Virginia Tech, Blacksburg, Virginia, United States of America. ²⁰Instituto de Investigación Agropecuaria, Estación Experimental El Ejido, Los Santos, Panamá. ²¹Universidad Técnica Estatal de Quevedo, Quevedo, Ecuador

Abstract

Background: American Creole cattle presumably descend from animals imported from the Iberian Peninsula during the period of colonization and settlement, through different migration routes, and may have also suffered the influence of cattle directly imported from Africa. The introduction of European cattle, which began in the 18th century, and later of Zebu from India, has threatened the survival of Creole populations, some of which have nearly disappeared or were admixed with exotic breeds. Assessment of the genetic status of Creole cattle is essential for the establishment of conservation programs of these historical resources.

Methodology/Principal Findings: We sampled 27 Creole populations, 39 Iberian, 9 European and 6 Zebu breeds. We used microsatellite markers to assess the origins of Creole cattle, and to investigate the influence of different breeds on their genetic make-up. The major ancestral contributions are from breeds of southern Spain and Portugal, in agreement with the historical ports of departure of ships sailing towards the Western Hemisphere. This Iberian contribution to Creoles may also include some African influence, given the influential role that African cattle have had in the development of Iberian breeds, but the possibility of a direct influence on Creoles of African cattle imported to America can not be discarded. In addition to the Iberian influence, the admixture with other European breeds was minor. The Creoles from tropical areas, especially those from the Caribbean, show clear signs of admixture with Zebu.

Conclusions/Significance: Nearly five centuries since cattle were first brought to the Americas, Creoles still show a strong and predominant signature of their Iberian ancestors. Creole breeds differ widely from each other, both in genetic structure and influences from other breeds. Efforts are needed to avoid their extinction or further genetic erosion, which would compromise centuries of selective adaptation to a wide range of environmental conditions.

Citation: Martínez AM, Gama LT, Cañón J, Ginja C, Delgado JV, et al. (2012) Genetic Footprints of Iberian Cattle in America 500 Years after the Arrival of Columbus. PLoS ONE 7(11): e49066. doi:10.1371/journal.pone.0049066

Editor: Sergios-Orestis Kolokotronis, Fordham University, United States of America

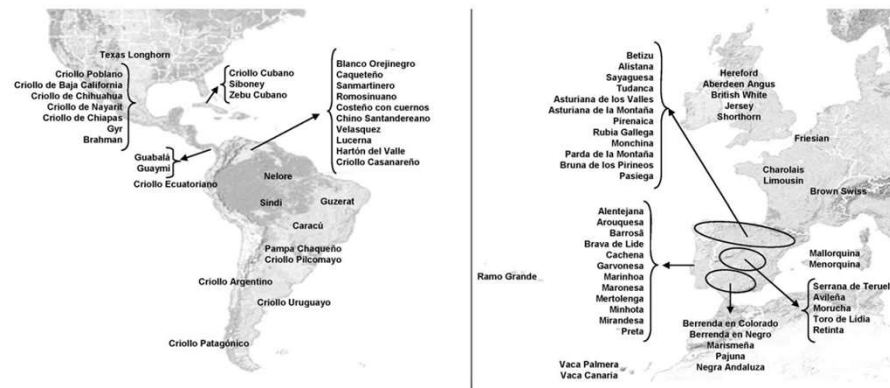
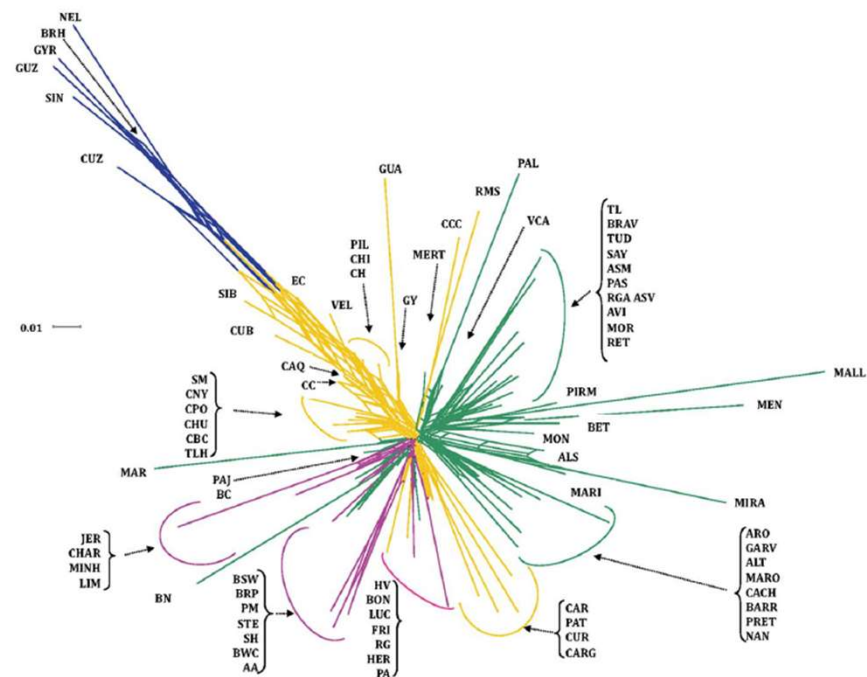
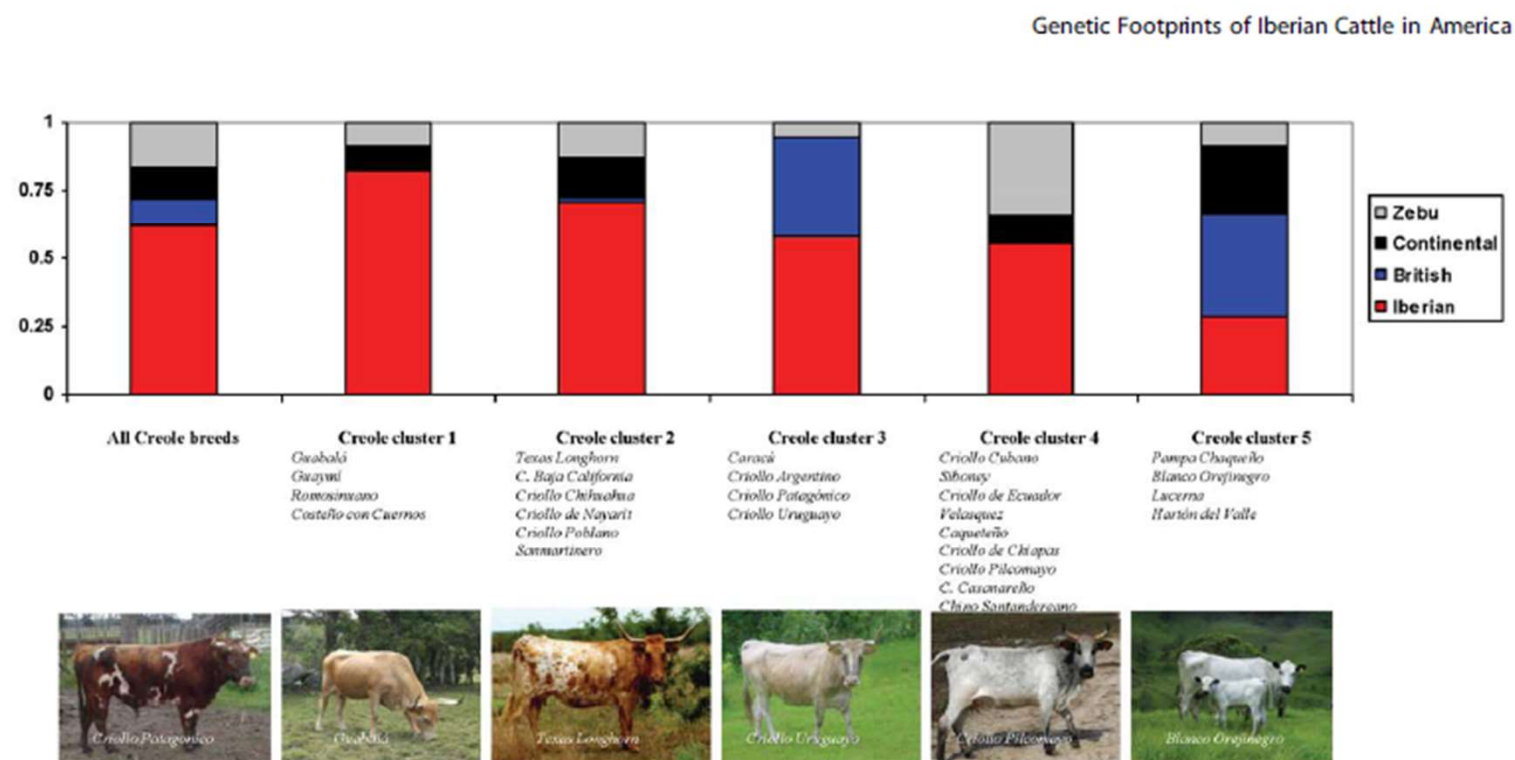
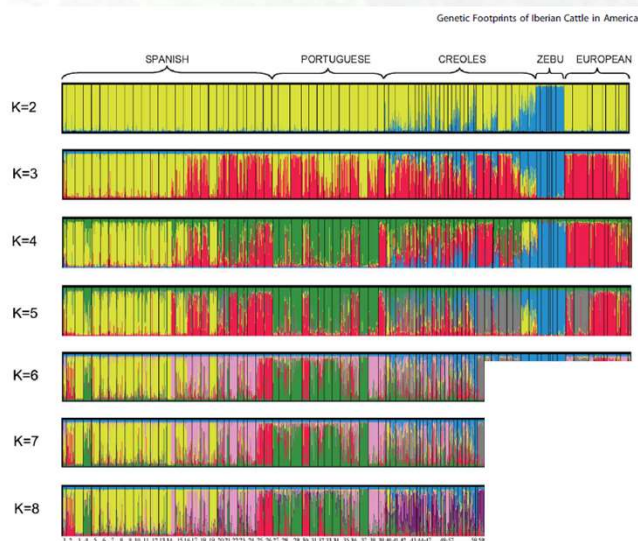


Figure 1. Geographic distribution of the 81 cattle breeds from America and Europe.
doi:10.1371/journal.pone.0049066.g001

Genetic Footprints of Iberian Cattle in America





**114 RAZAS,
40 CRIOLLAS,**

**49 INVESTIGADORES,
39 INSTITUCIONES**

**SCIENTIFIC
REPORTS**
nature research

OPEN

The genetic ancestry of American Creole cattle inferred from uniparental and autosomal genetic markers

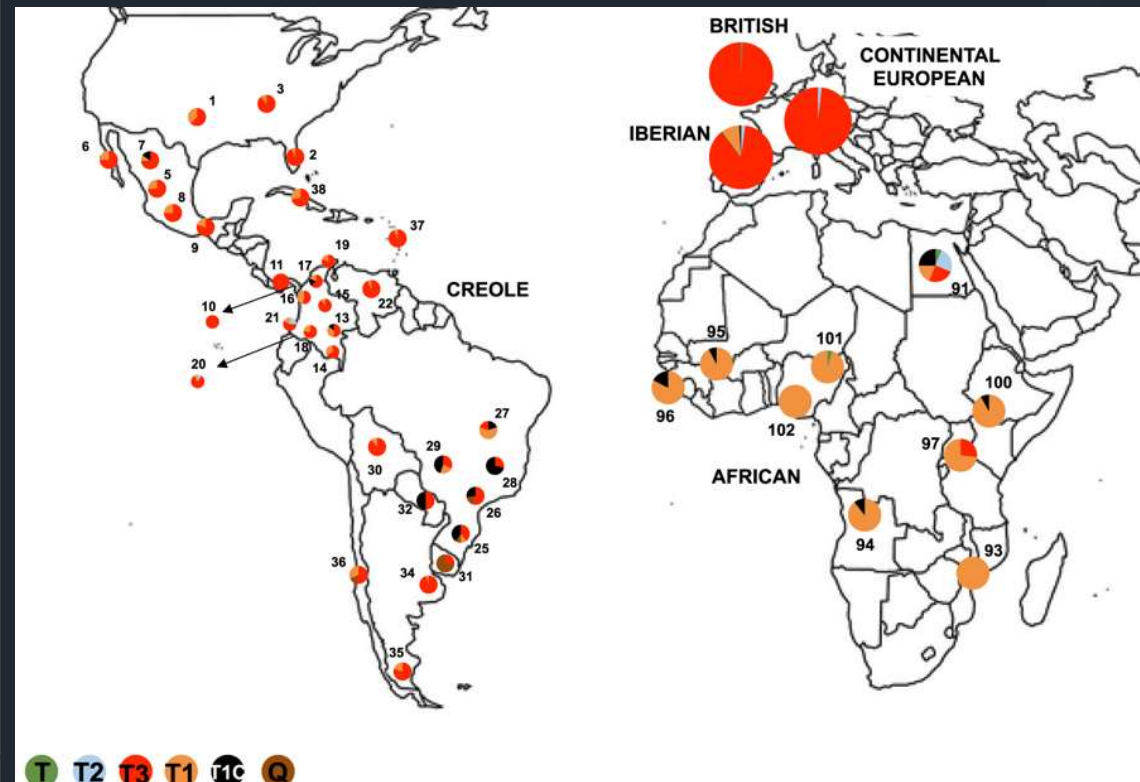
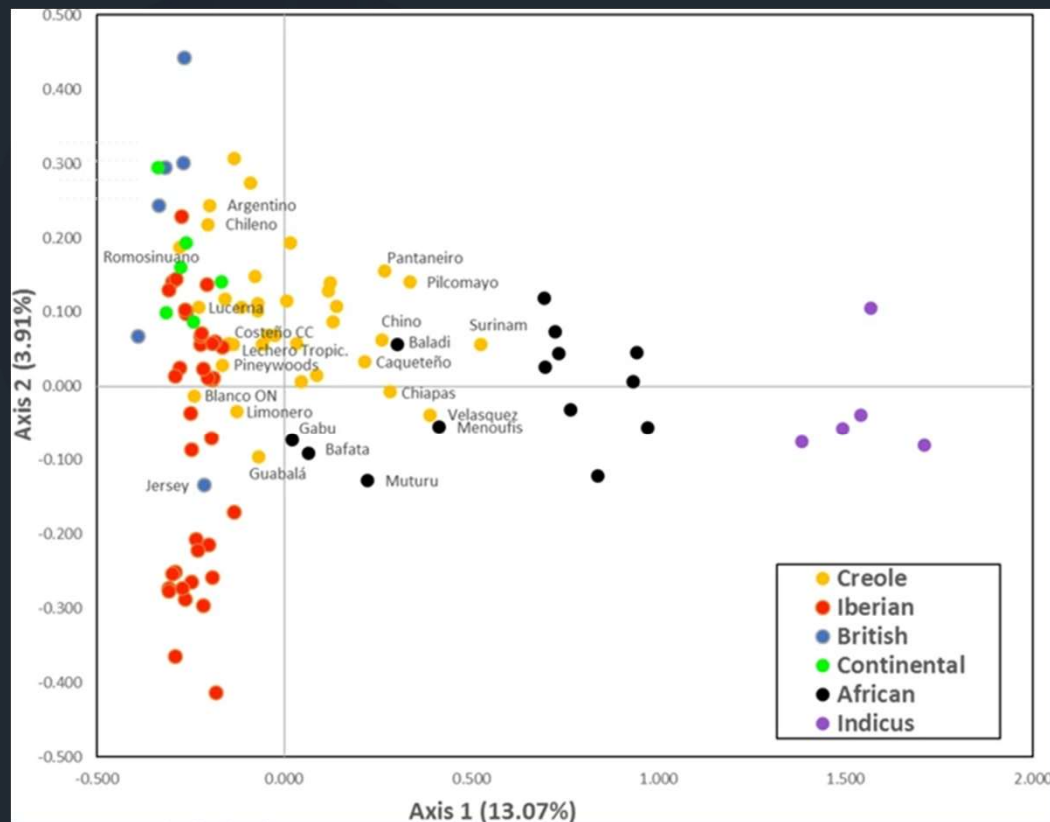
Received: 20 February 2019
Accepted: 16 July 2019
Published online: 07 August 2019

7 de agosto 2019

Catarina Ginja¹, Luis Telo Gama², Oscar Cortés³, Inmaculada Martín Burriel⁴, Jose Luis Vega-Pla⁵, Cecilia Penedo⁶, Phil Sponenberg⁷, Javier Cañón³, Arianne Sanz⁴, Andrea Alves do Egito⁸, Luz Angela Alvarez⁹, Guillermo Giovambattista¹⁰, Saif Agha¹¹, Andrés Rogberg-Muñoz¹², Maria Aparecida Cassiano Lara¹³, BioBovis Consortium^{*}, Juan Vicente Delgado¹⁴ & Amparo Martínez^{14,15}

Cattle imported from the Iberian Peninsula spread throughout America in the early years of discovery and colonization to originate Creole breeds, which adapted to a wide diversity of environments and later received influences from other origins, including zebu cattle in more recent years. We analyzed uniparental genetic markers and autosomal microsatellites in DNA samples from 114 cattle breeds distributed worldwide, including 40 Creole breeds representing the whole American continent, and samples from the Iberian Peninsula, British islands, Continental Europe, Africa and American zebu. We show that Creole breeds differ considerably from each other, and most have their own identity or group with others from neighboring regions. Results with mtDNA indicate that T1c-lineages are rare in Iberia but common in Africa and are well represented in Creoles from Brazil and Colombia, lending support to a direct African influence on Creoles. This is reinforced by the sharing of a unique Y-haplotype between cattle from Mozambique and Creoles from Argentina. Autosomal microsatellites indicate that Creoles occupy an intermediate position between African and European breeds, and some Creoles show a clear Iberian signature. Our results confirm the mixed ancestry of American Creole cattle and the role that African cattle have played in their development.

Estudio publicado en el año 2019 mediante marcadores uniparentales confirmaron la herencia ibérica (España y Portugal) y el aporte de razas africanas como la **N'Gabou** y **Bafatá** de Guinea-Bissau la **Muturú** de Nigeria y las razas egipcias **Baladí** y **Menoufis**





De la genética a la genómica

La primer secuenciación Bovina a una vaca Hereford (Bos taurus).

El genoma de una vaca Hereford hembra fue publicado en 2009. Fue secuenciado por el Consorcio de Secuenciación y Análisis del Genoma Bovino, un equipo de investigadores liderado por los Institutos Nacionales de Salud y el Departamento de Agricultura de los Estados Unidos.



Una vaca llamada L1 Dominette, originaria de Inglaterra, una raza caracterizada por la ausencia de cuernos, ayudará a comprender la evolución de los mamíferos y mejorará la eficiencia de la explotación ganadera.

Have you explored our new Comparative Genome Viewer (CGV) yet?



Genome Data Viewer

Home

Share this page

Reset All

More Tools

More Info

Bos taurus
(cattle)

Assembly: ARS-UCD1.3 (GCF_002263795.2)

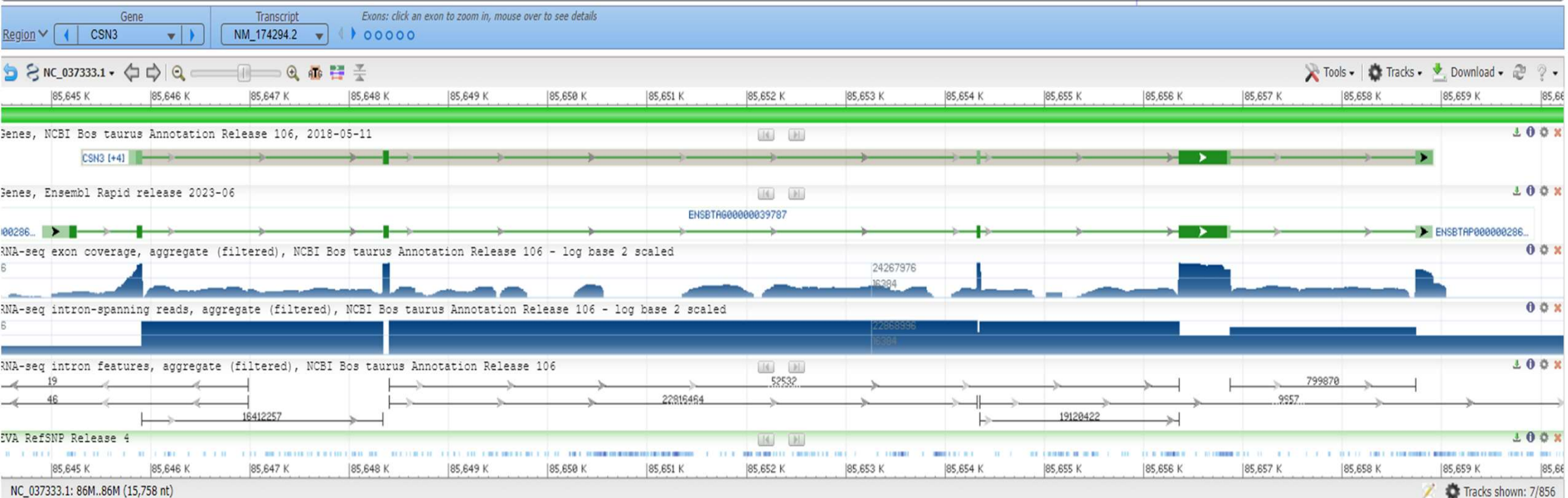
Chr 6 (NC_037333.1)

Search assembly

CSN3

Examples

>> NC_037333.1: 85,644,465 - 85,660,222



Comparación de genomas de referencia en el tiempo según CSN3

Genoma	Año	Localización	Método	Cobertura	Remitente
Btau_4.6.1 (bosTau7)	2011	88,470,896 - 88,483,957	Sanger	7X	Cattle Genome Sequencing International Consortium
UMD_3.1.1 (bosTau8)	2014	87,390,197 - 87,392,750	Sanger	9X	Center for Bioinformatics and Computational Biology, University of Maryland
Btau_5.0.1	2015	87,737,368 - 87,739,921	Sanger; PacBio RS II	19X	Cattle Genome Sequencing International Consortium
ARS-UCD1.3 (bosTau9)	2018	85,645,780 - 85,658,911	PacBio; Illumina NextSeq 500; Illumina HiSeq; Illumina GAI	80x	USDA ARS

Usos de la Genómica en la Ganadería

- Identificación de progenitores. Incluye el procedimiento de asignación de identidad, verificación o asignación de padres.
- Detección de individuos portadores de defectos genéticos, como es el caso de algunas enfermedades.
- Detección de genes que favorecen o no la producción (carne, leche, lana, huevo, etc).
- Selección de individuos sobresalientes.

“Pero se requiere la información fenotípica para poder realizar un análisis más preciso”

Frecuencias Alélicas de Variantes Polimórficas de Genes Asociados a Variables Ambientales de las Razas Guaymí y Guabalá en Panamá.

MARCADOR	Locus	Alelo	GUA	GUY	MARCADOR	Locus	Alelo	GUA	GUY
MED12L	rs41580133	C	0.167	0.605	TSNARE1	rs109875744	C	0.933	0.500
		T	0.833	0.395			T	0.067	0.500
		A	0.000	0.237			A	0.067	0.158
HSF2BP	rs41643488	G	1.000	0.763	PREX2	rs109169231	G	0.933	0.842
		A	0.000	0.132			C	0.400	0.974
		G	1.000	0.868			T	0.600	0.026
ADGRL2	rs42482471	A	1.000	0.921	RALYL	rs29010281	A	0.800	0.500
		G	0.000	0.079			G	0.200	0.500
		A	0.500	0.447			A	0.200	0.263
SPAG17	rs41624677	G	0.500	0.553		rs110400380	G	0.800	0.737
		G	0.900	0.632			A	0.500	0.025
		T	0.100	0.368			C	0.500	0.975
CTNNA2	rs41565994	G	0.933	0.289	SMYD3	rs41799745	A	0.467	0.000
		T	0.067	0.711			G	0.533	1.000
		A	0.600	0.579			A	0.633	0.632
	rs211690801	G	0.400	0.421		rs41799830	G	0.367	0.368
		G	1.000	0.737			C	0.533	1.000
		T	0.000	0.263			T	0.467	0.000
	rs29013419	C	0.000	0.026		rs41799658	A	0.633	0.053
		T	1.000	0.974			G	0.367	0.947
		C	1.000	0.763			A	0.533	0.421
	rs41668252	T	0.000	0.237		rs42383968	G	0.467	0.579
		A	0.033	0.053			A	0.400	0.289
		G	0.967	0.947			G	0.600	0.711
	rs41663399	C	0.800	0.842	LAMC1	rs29013977	G	1.000	0.947
		T	0.200	0.158			T	0.000	0.053
		G	0.800	0.474			A	0.467	0.368
	rs41584429	T	0.200	0.526	SUZ12	rs29019767	G	0.533	0.632
		A	0.100	0.789			A	0.567	0.579
		G	0.900	0.211			G	0.433	0.421
HSPH1	rs109165924	C	0.867	0.816	ZKSCAN7	rs42001169	A	1.000	0.553
		T	0.133	0.184			G	0.000	0.447
FAM107B	rs29017684	C	0.867	0.816	NRG1	rs42127055	A	1.000	0.553
		T	0.133	0.184			G	0.000	0.447

Loci Polimórficos Guaymí fueron similares a ganado bovino de Etiopía (83.4%) Edea et al. (2012) Menor al del ganado Hanwoo de Korea (94%).

Monorfismo observado en el marcador del gen PRLR (C) 20:39,136,666



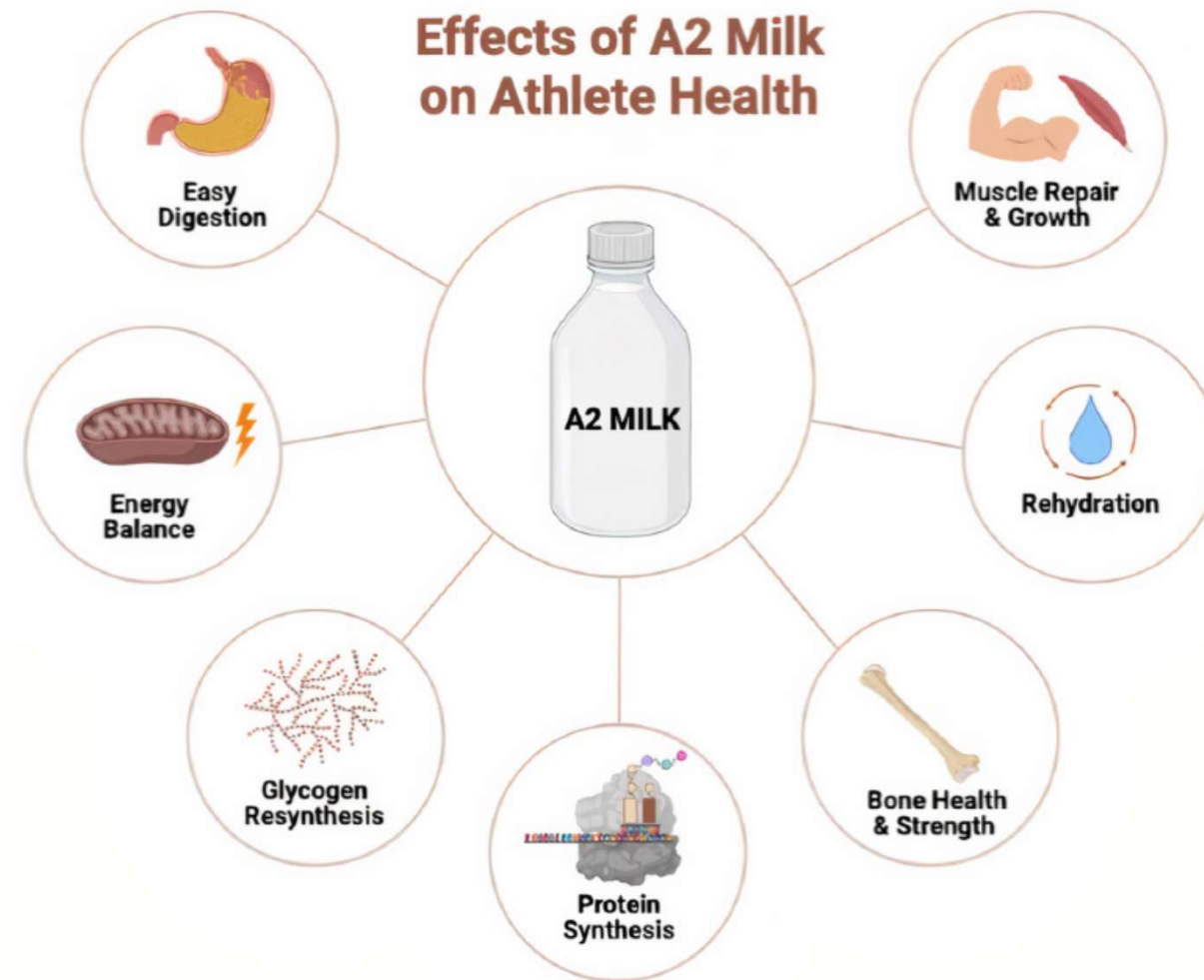
Media de Número efectivo de alelos (Ne), Índice de Shannon (I), Heterocigosis observada (Ho) y Heterocigosis esperada (He) de razas Guaymí y Guabalá para los alelos asociados a calidad de carne.

Gen	Variante	Ne	Guabalá			Guaymí			
			I	Ho	He	Ne	I	Ho	He
CAST	rs109804679	1.557	0.543	0.467	0.370	1.362	0.436	0.316	0.273
CAST	rs109677393	1.991	0.691	0.533	0.515	1.870	0.658	0.526	0.478
CAST	rs109354718	1.991	0.691	0.533	0.515	1.994	0.692	0.632	0.512
CAPN3	rs109425380	1.991	0.691	0.400	0.515	1.362	0.436	0.316	0.273
CAPN13	rs108960548	1.724	0.611	0.600	0.434	1.819	0.642	0.368	0.462
CAPN5	rs41772701	1.301	0.393	0.267	0.239	1.699	0.602	0.474	0.422
CAPN8	rs109316815	1.000	0.000	0.000	0.000	1.498	0.515	0.316	0.341
CAPN1	rs17872000	1.000	0.000	0.000	0.000	1.978	0.688	0.474	0.508
CAPN1	rs17871058	1.867	0.657	0.467	0.480	1.870	0.658	0.421	0.478
CAPN1	rs17872050	1.000	0.000	0.000	0.000	1.819	0.642	0.474	0.462

Frecuencias alélicas de variantes polimórficas de genes de caseínas de razas Guaymí y Guabalá (Genoma de referencia UMD 3.1.1)

Gen/Alelos	RefSeq	Guabalá		Guaymí	
		Frecuencia alélica		Frecuencia alélica	
CSN1S1	rs133474041	0.800(G)	0.200 (A)	0.684 (G)	0.316 (A)
CSN2	rs109299401	1.000(T)	0.000 (G)	0.789 (T)	0.211(G)
CSN2	rs43703011*	1.000(C)	0.000(A)	0.842 (C)	0.158 (A)
CSN1S2	rs441966828	1.000(C)	0.000(T)	0.842 (C)	0.158 (T)
CSN3	rs450402006	0.733 (C)	0.267 (T)	0.658 (C)	0.342 (T)
CSN3	rs43703015	0.433 (T)	0.567 (C)	0.500 (T)	0.500 (C)
CSN3	rs43703016	0.429 (C)	0.571(A)	0.500(C)	0.500 (A)
CSN3	rs439304887	1.000(A)	0.000(G)	0.842 (A)	0.158 (G)
CSN3	rs110014544	0.433 (G)	0.567 (A)	0.528 (G)	0.472 (A)





Visible effects of the consumption of A2 milk on athlete health.



A microscopic image of a plant stem cross-section, stained with a blue and purple dye. The image shows a dense arrangement of small, circular cells forming the cortex and pith. Several large, circular vascular bundles are visible, each containing a central xylem region and an outer phloem region. The bundles are arranged in a ring-like pattern around a central pith.

LOS FENOTIPOS

In the age of the genotype.....



Im Zeitalter des Genotyps ist der Phänotyp König

Fenotype blijft de koning

#PHENOTYPE IS KING!



في عصر التركيب الجيني
البيانات المظهرية هي الملك

En la era del genotipo ...
¡El fenotipo es el rey!

Την εποχή του γονοτύπου, ο φαινότυπος είναι βασιλιάς!

SIMPOSIO OBJETIVOS DE SELECCIÓN Y METODOLOGÍAS EN
MEJORAMIENTO GENÉTICO ANIMAL EN LA ERA POST-GENÓMICA

فینوٹائپ بادشاہ ہے



Michael Coffey

Scotland's Rural College

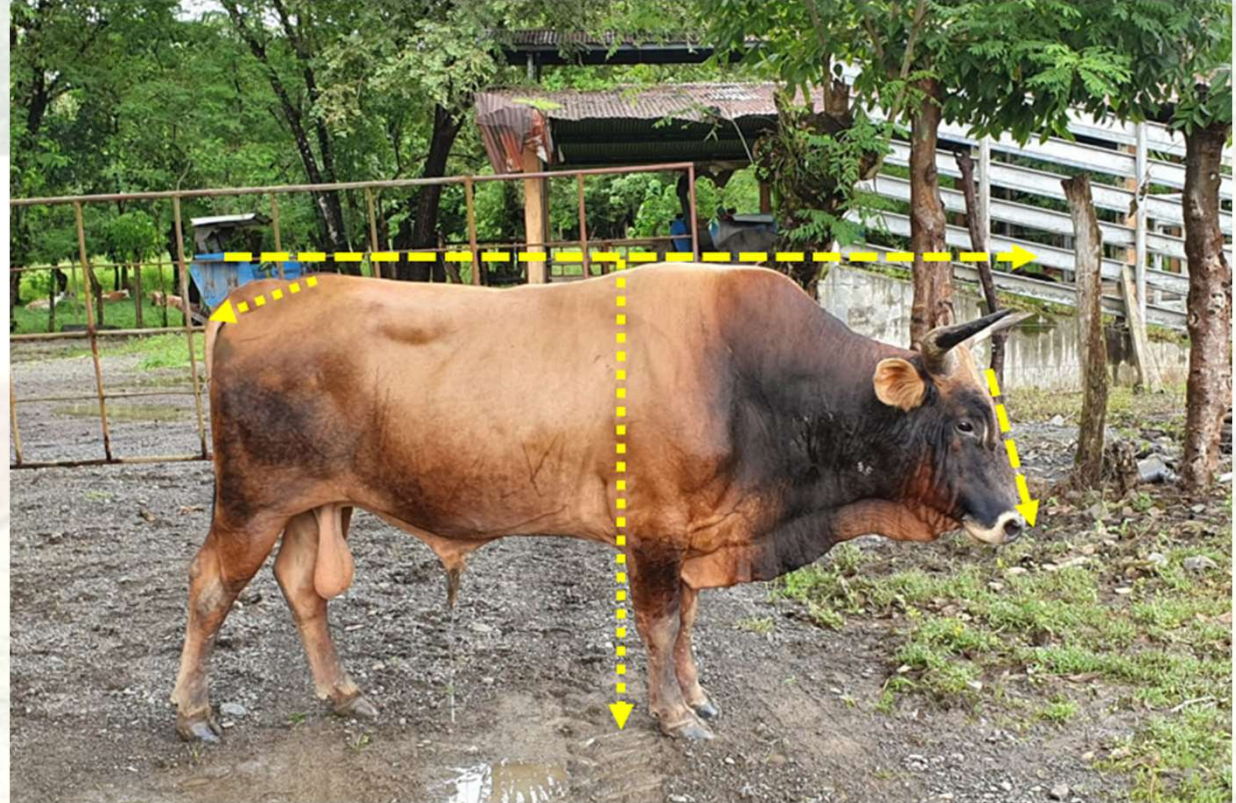
TODAVÍA ESTA VIGENTE ESTA ECUACIÓN



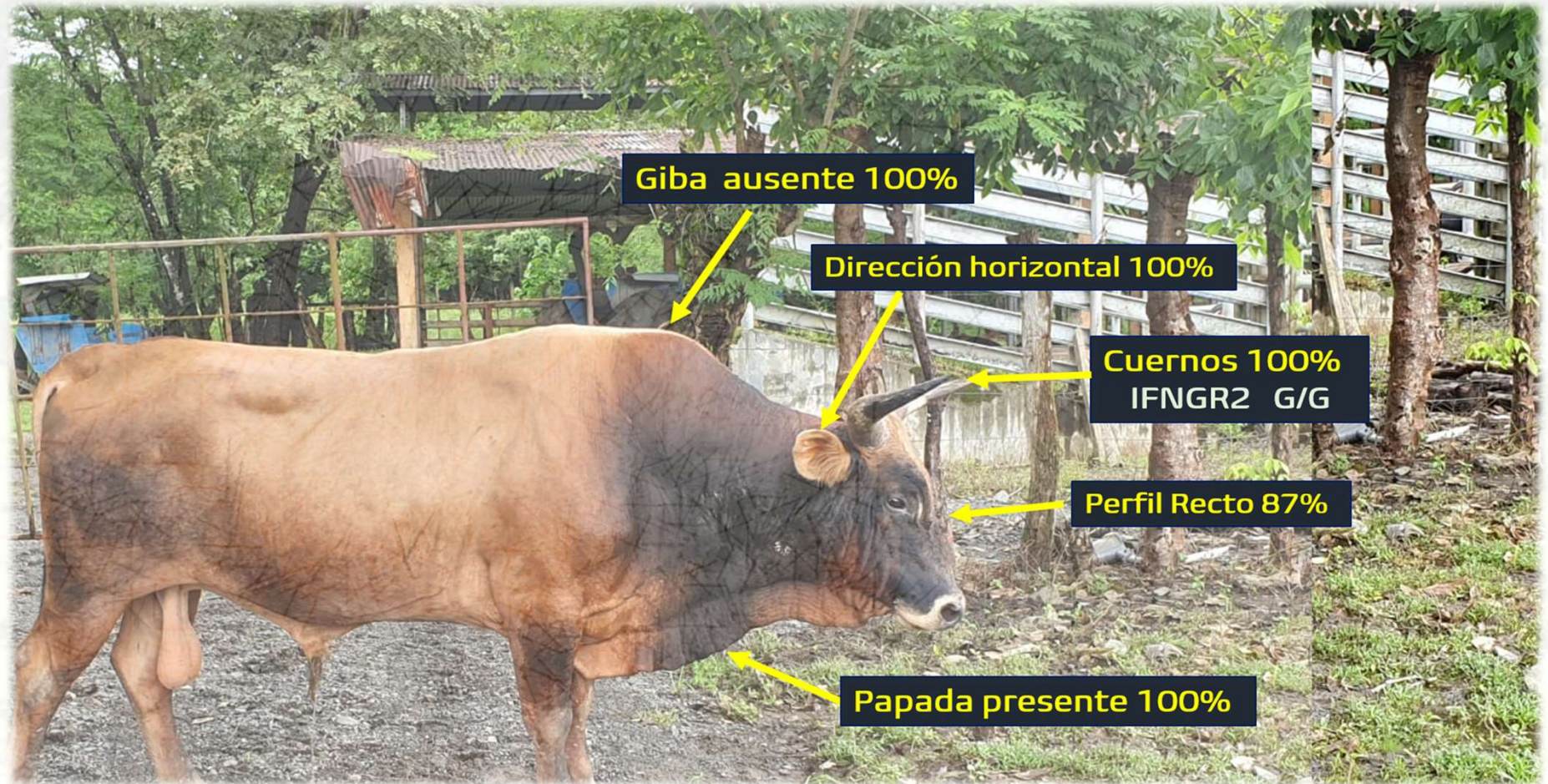
FENOTIPO = GENOTIPO + AMBIENTE + 2COV (G, A) + G x A

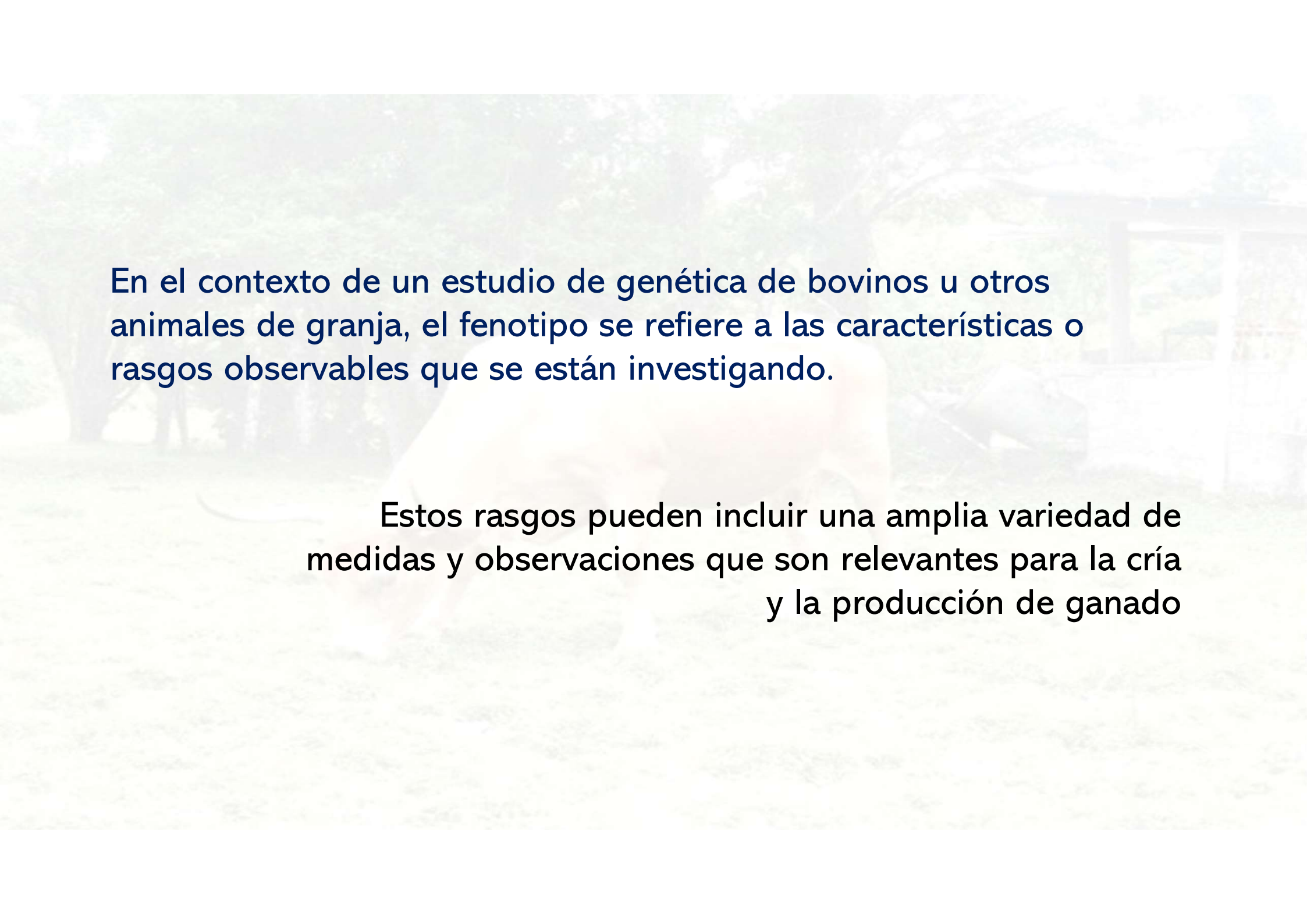
$$\sigma_F = \sigma_G + \sigma_A + 2\text{COV}(G,A) + \sigma_{GA}$$

Los fenotipos son las características observables y medibles de un organismo, que resultan de la interacción entre su genotipo (su información genética) y el ambiente en el que se desarrolla.



Estas características pueden ser muy variadas y van desde rasgos físicos y morfológicos hasta aspectos relacionados con el rendimiento y la salud.





En el contexto de un estudio de genética de bovinos u otros animales de granja, el fenotipo se refiere a las características o rasgos observables que se están investigando.

Estos rasgos pueden incluir una amplia variedad de medidas y observaciones que son relevantes para la cría y la producción de ganado

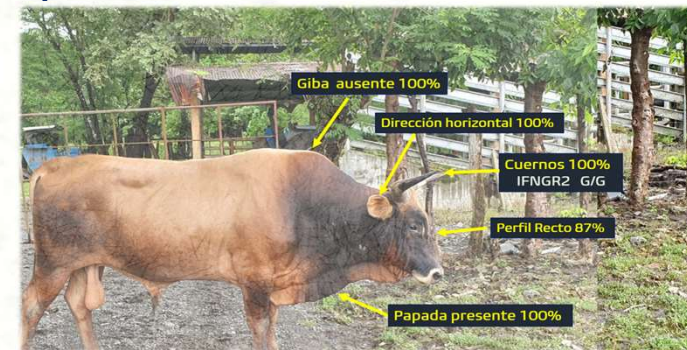
ALGUNOS EJEMPLOS DE FENOTIPOS EN BOVINOS Y ANIMALES DE GRANJA PODRÍAN SER:

1. **Peso al destete:** El peso del ternero cuando es destetado de la madre.
2. **Tasa de crecimiento diario:** La velocidad a la que el animal gana peso durante un período específico.
3. **Tamaño del músculo:** Medidas de la masa muscular en áreas específicas del cuerpo.
4. **Producción de leche:** La cantidad de leche que una vaca produce en un período de tiempo determinado.
5. **Resistencia a enfermedades:** La capacidad del animal para resistir enfermedades comunes.
6. **Calidad de la carne o leche:** Características como el contenido de grasa, la ternura de la carne o la calidad de la leche.

Fenotipos Morfológicos:

Estos se refieren a las características físicas y estructurales de un animal. En la ganadería bovina, algunos ejemplos de fenotipos morfológicos incluyen:

- **Color del pelaje:** Algunas razas de ganado bovino tienen pelajes de diferentes colores, como el Guaymí o el Hereford rojo.
- **Tamaño y peso corporal:** El tamaño y el peso de un bovino pueden variar considerablemente según la raza y el linaje.
- **Longitud de cuerno:** Algunas razas tienen cuernos largos, mientras que otras son deshornadas.
- **Conformación del cuerpo:** Esto se refiere a la estructura del cuerpo del animal, como la forma de la cabeza, la espalda, la grupa, etc.





Fenotipos de Rendimiento: Estos se centran en las características que afectan la producción y el rendimiento del ganado. Ejemplos de fenotipos de rendimiento incluyen:

- **Producción de leche:** La cantidad de leche que una vaca puede producir en un período determinado.
- **Ganancia de peso diaria:** La velocidad a la que un bovino aumenta su peso durante la fase de engorde.



TRABAJOS SOBRE ASPECTOS FENOTÍPICOS

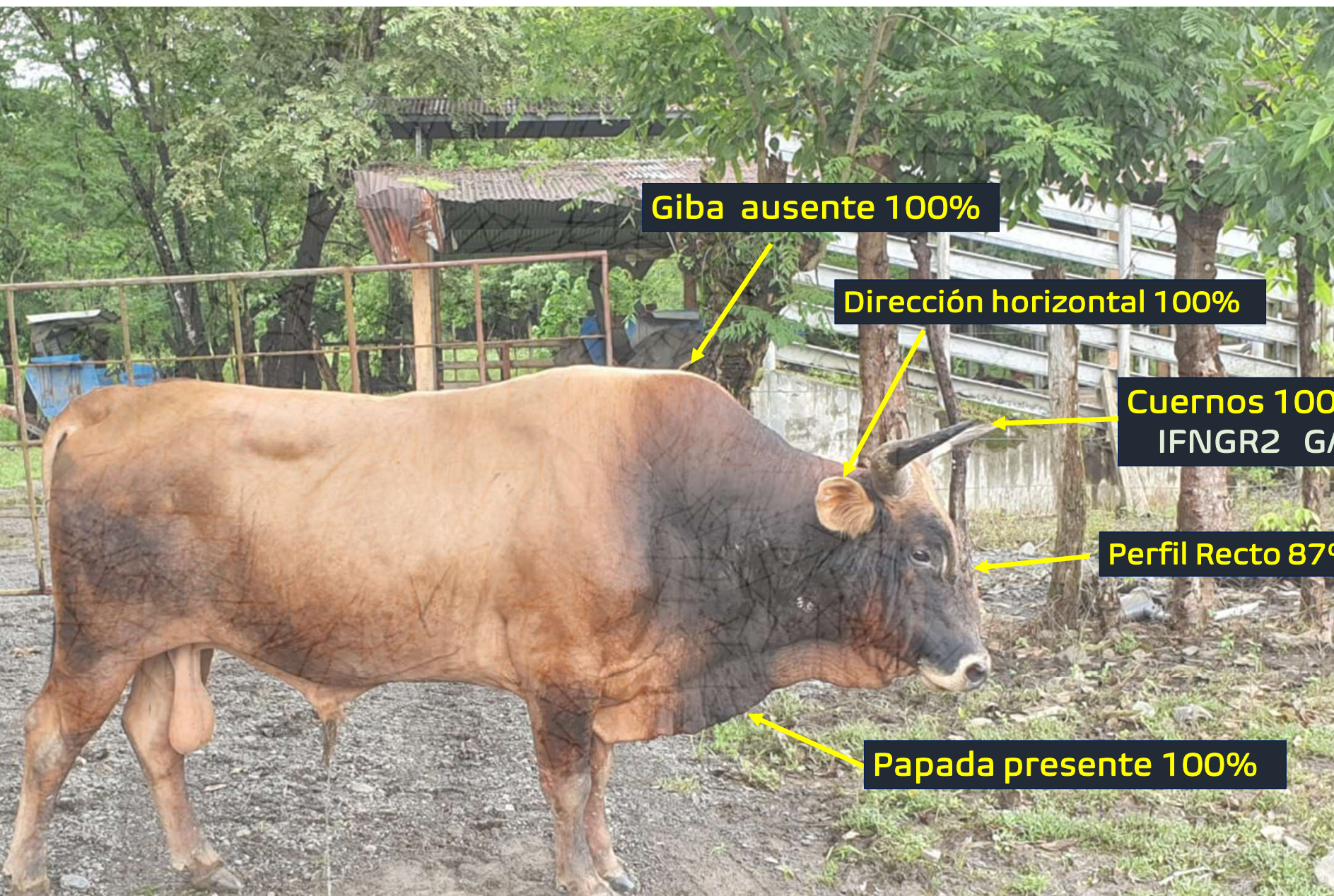
*C4R4C73R1Z4C1ÓN MORFOLÓGICA Y
FAN3RÓP7ICA D3 L4 R4Z4 GU4YMÍ3N
C3N7R05 D3 CON53RV4CIÓN*



La apreciación externa de una población como un grupo racial bajo un enfoque general es lo primero que ejecuta de manera natural el ganadero en su finca o rancho.



Mediante las variables morfológicas se puede determinar el grado de homogeneidad o heterogeneidad que presentan los individuos dentro de una población o una raza.



Giba ausente 100%

Dirección horizontal 100%

Cuernos 100%
IFNGR2 G/G

Perfil Recto 87%

Papada presente 100%

Color de la capa

Overo

52

Entremezclado

22

Bayo

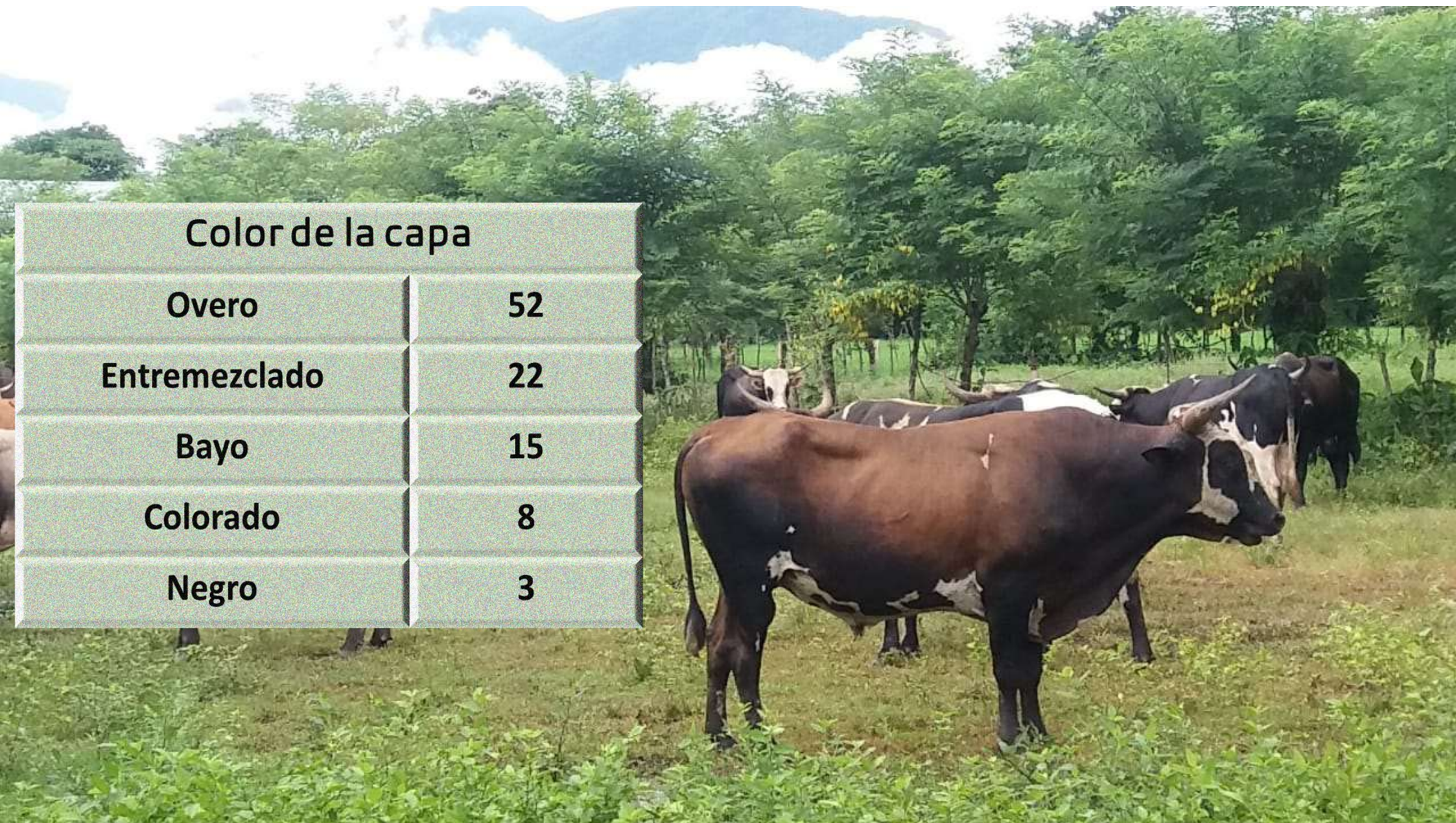
15

Colorado

8

Negro

3





Posición de Cuernos	Procero	1
	Ortocero	37
	Opistocero	62
Tipo de Gancho	Alto	56
	Medio	24
	Bajo	20

Fenotipos de Salud: Estos se relacionan con la salud y la resistencia a enfermedades de los animales. Ejemplos de fenotipos de salud en la ganadería bovina son:

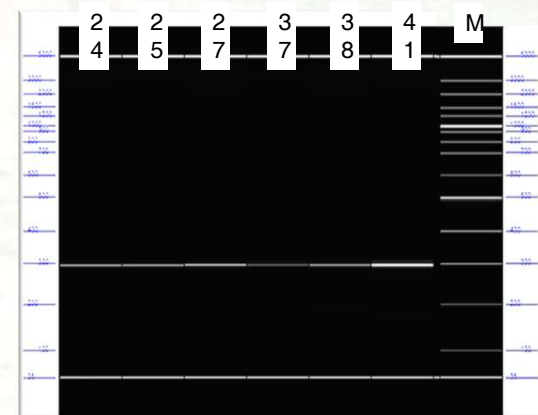
- **Resistencia a enfermedades específicas:** La capacidad de un bovino para resistir enfermedades comunes como la fiebre aftosa o la brucelosis.
- **Tasa de mortalidad:** La cantidad de animales que sobreviven a una enfermedad o problema de salud.



RAZA GUABALÁ	bp	Valor Máximo	e-value	% Identidad
*0101	181	276	1.0E-70	95
*0201	176	224	5.0E-55	91
*1801	196	302	2.0E-76	96
*3001	206	300	1.0E-77	96
*R-08	197	278	4.0E-71	93
*R-121	211	302	3.0E-78	93
*R-142	208	329	1.0E-86	95
*R-177	209	311	5.0E-81	94
*R-194	205	309	2.0E-80	95
*R-21	210	324	6.0E-85	95
*R-73	212	302	3.3E-67	93
RAZA GUAYMÍ				
*0101	208	321	6.7E-81	95
*1101	219	294	5.0E-76	93
*1104	148	165	3.0E-37	87
*3601	215	285	3.0E-73	93
*R-09	224	278	5.0E-71	91
*R-73	195	281	3.0E-72	95

EL GEN BoLA-DRB3.2 EN BOVINOS CRIOLLOS GUAYMÍ Y GUABALÁ

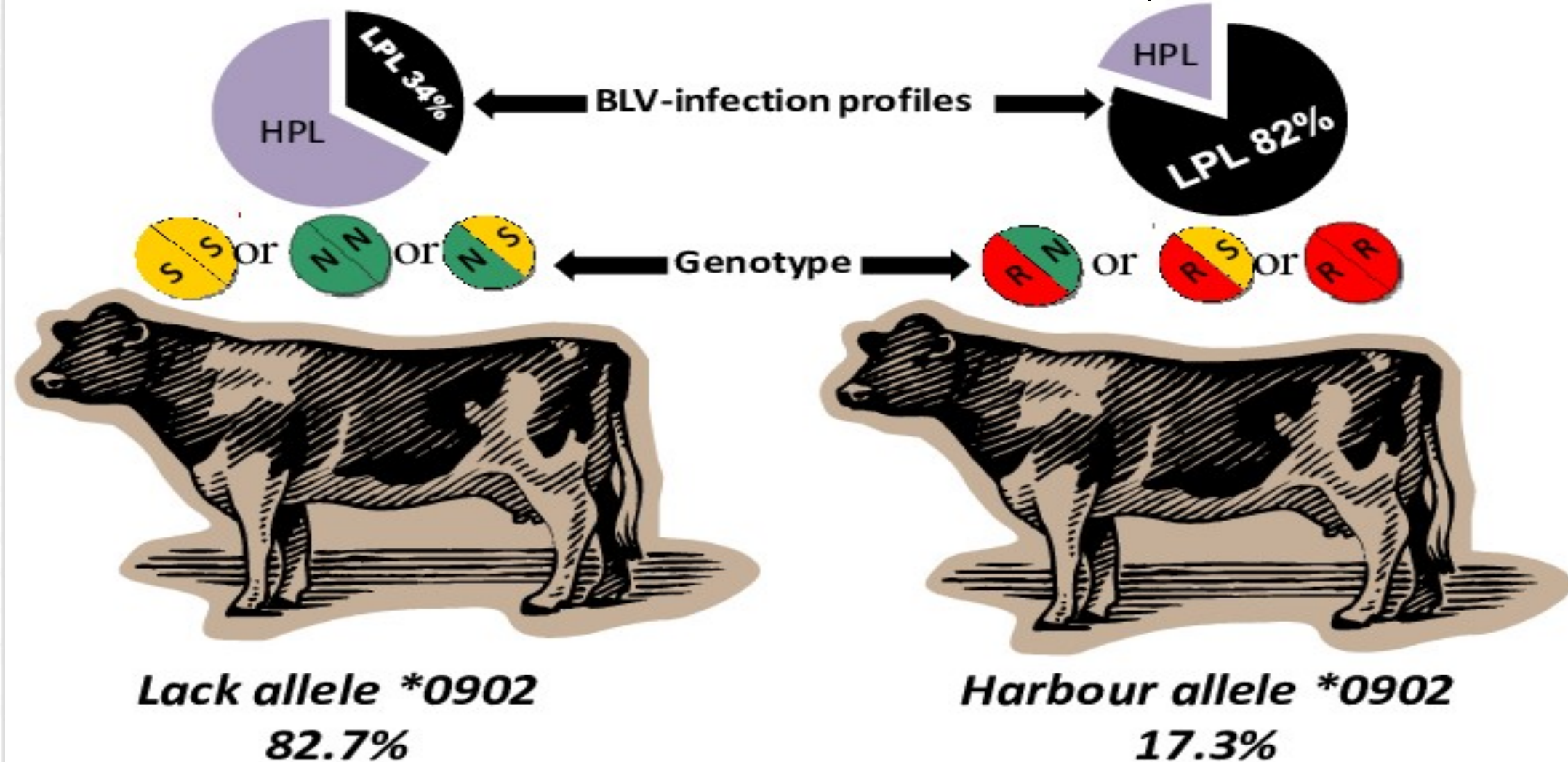
```
>|c|CRIOLLO_S11_1_2902
CTGCAGCACAT CCTGGAGTATGCTACGAGCGAGTGTCA
TCAACGGGACCGAGCGGGTGCGGT CCTGCACAGGTACT
AATGGAGAAGAGT CGTGCGCT CGACAGCGACTGGGGC
CCGGGCGGTGACCGAGCTGGGGCAGCGGGTCGCCGAGC
GAACAGCCAGAAGGACACCCTGGAGCGGGAGCGGGCCT
GACACGTACTGCAGACACAACCTACGGGGTCGGTGAGAGT
```

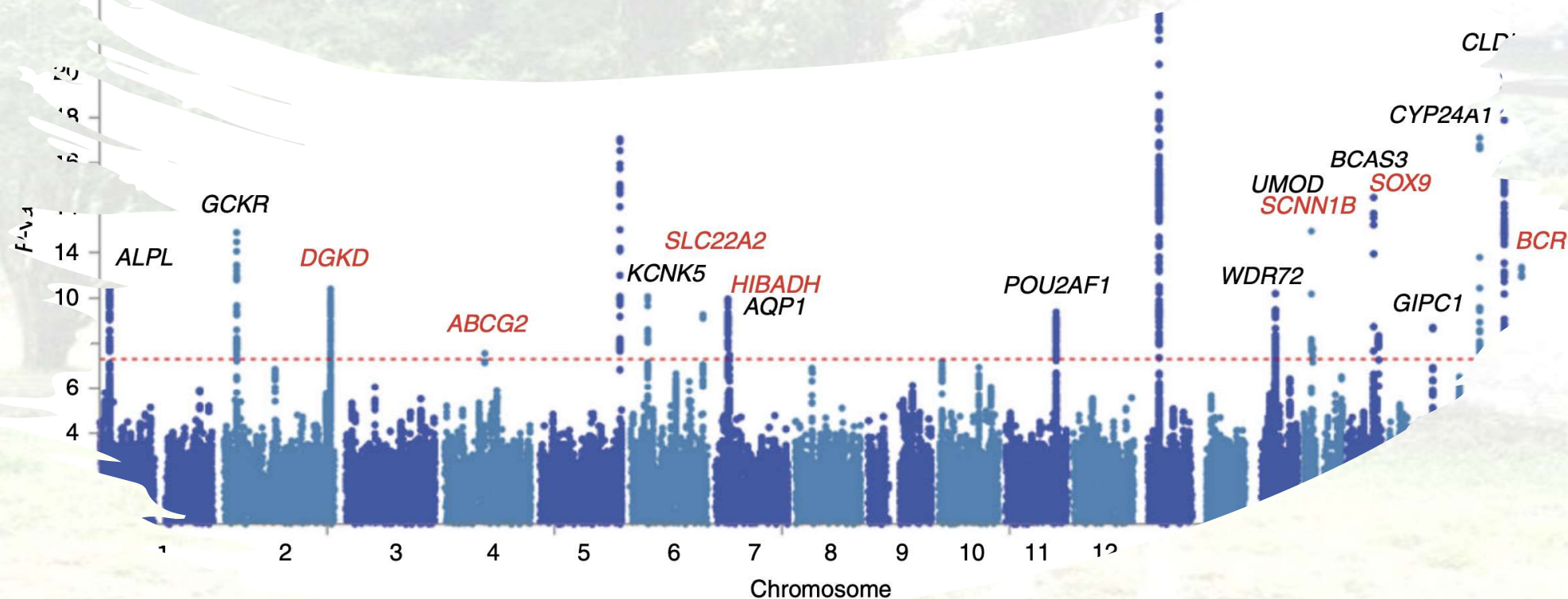


Villalobos-Cortés, A y González Herrera, R.
2018. Ciencia Agropecuaria no. 28:22-36.

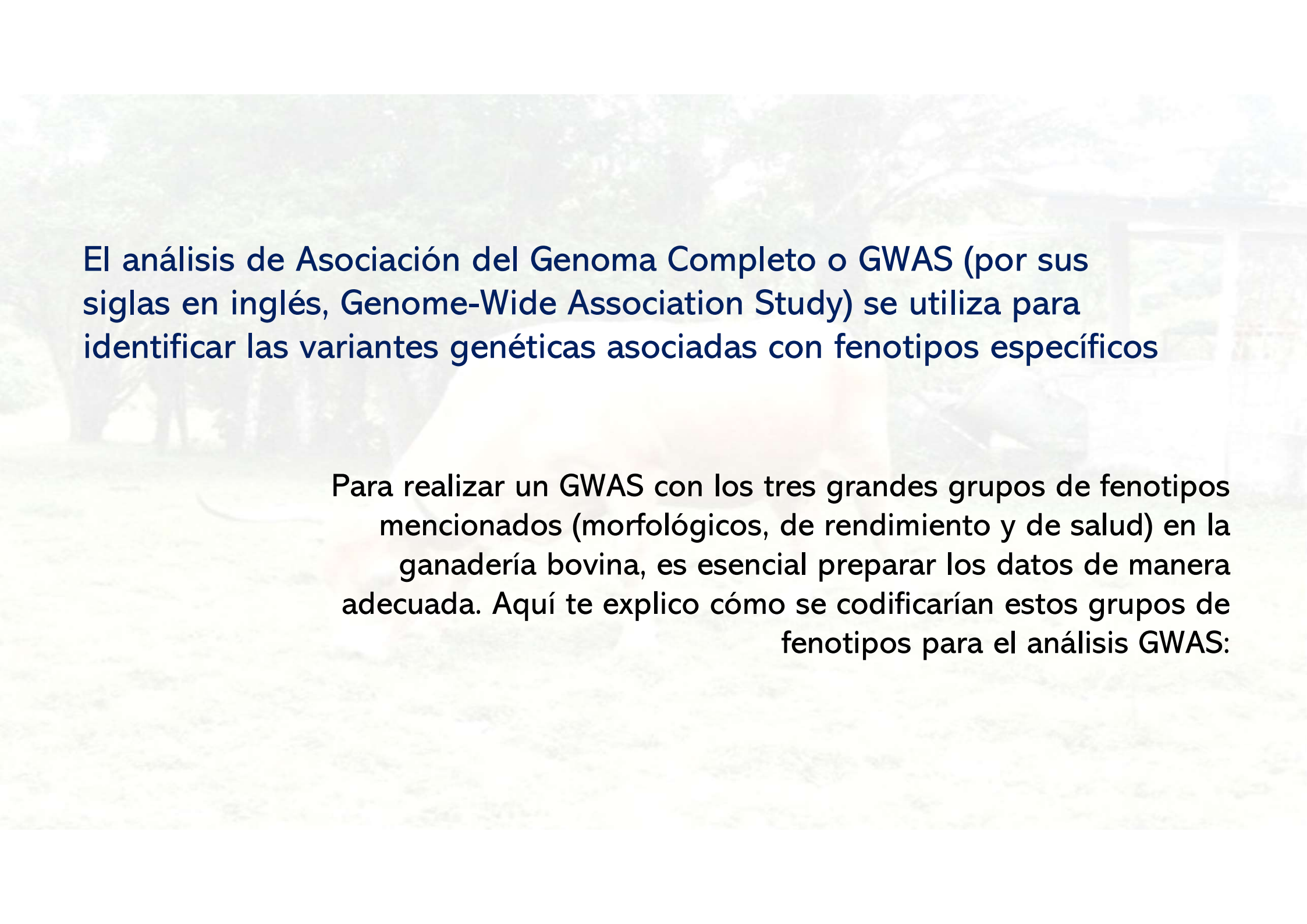
SECUENCIAS DEL 2° EXÓN DEL GEN BoLA-DRB3.2

Esteban et al., 2009. *Animal Genetics*





Ejemplos de codificación para análisis GWAS



El análisis de Asociación del Genoma Completo o GWAS (por sus siglas en inglés, Genome-Wide Association Study) se utiliza para identificar las variantes genéticas asociadas con fenotipos específicos

Para realizar un GWAS con los tres grandes grupos de fenotipos mencionados (morfológicos, de rendimiento y de salud) en la ganadería bovina, es esencial preparar los datos de manera adecuada. Aquí te explico cómo se codificarían estos grupos de fenotipos para el análisis GWAS:

Fenotipos Morfológicos:

Para los fenotipos morfológicos, como el color del pelaje o la longitud de cuerno, se puede codificar de manera categórica. Cada categoría representaría una variante del fenotipo.

Color del Pelaje: 1 = Negro, 2 = Rojo, 3 = Pardo, etc.

Longitud de Cuerno: 1 = Largos, 2 = Cortos, 3 = Deshornados, etc.

Color de la capa

Overo

1

Entremezclado

2

Bayo

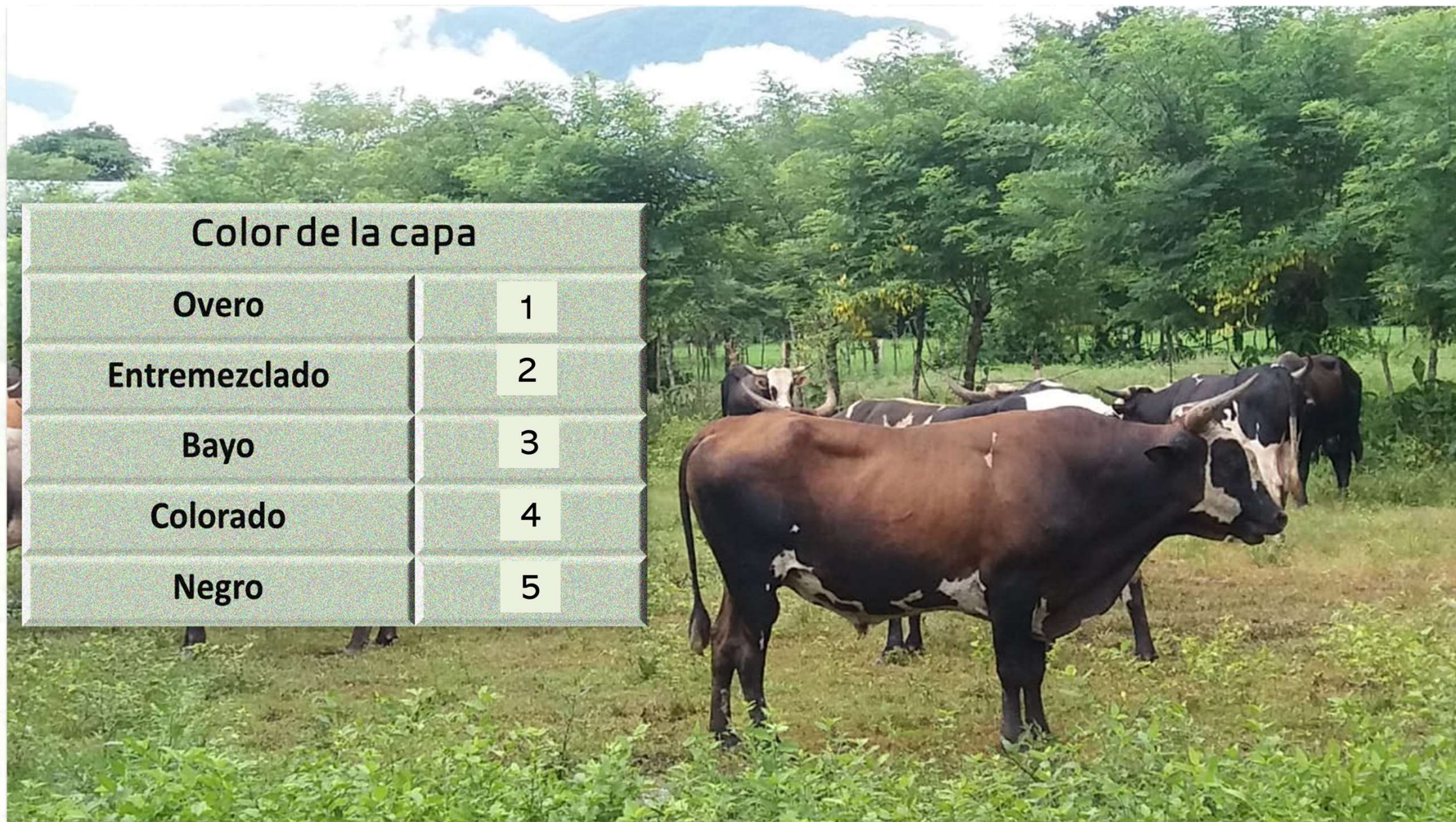
3

Colorado

4

Negro

5



Fenotipos de Rendimiento:

Los fenotipos de rendimiento, como la producción de leche o la ganancia de peso diaria, se pueden codificar como valores numéricos continuos. Cada animal tendría un valor que representa su rendimiento en ese aspecto.

Producción de Leche (litros por día).

Ganancia de Peso Diaria (kilogramos por día).



- Fenotipos de Salud:

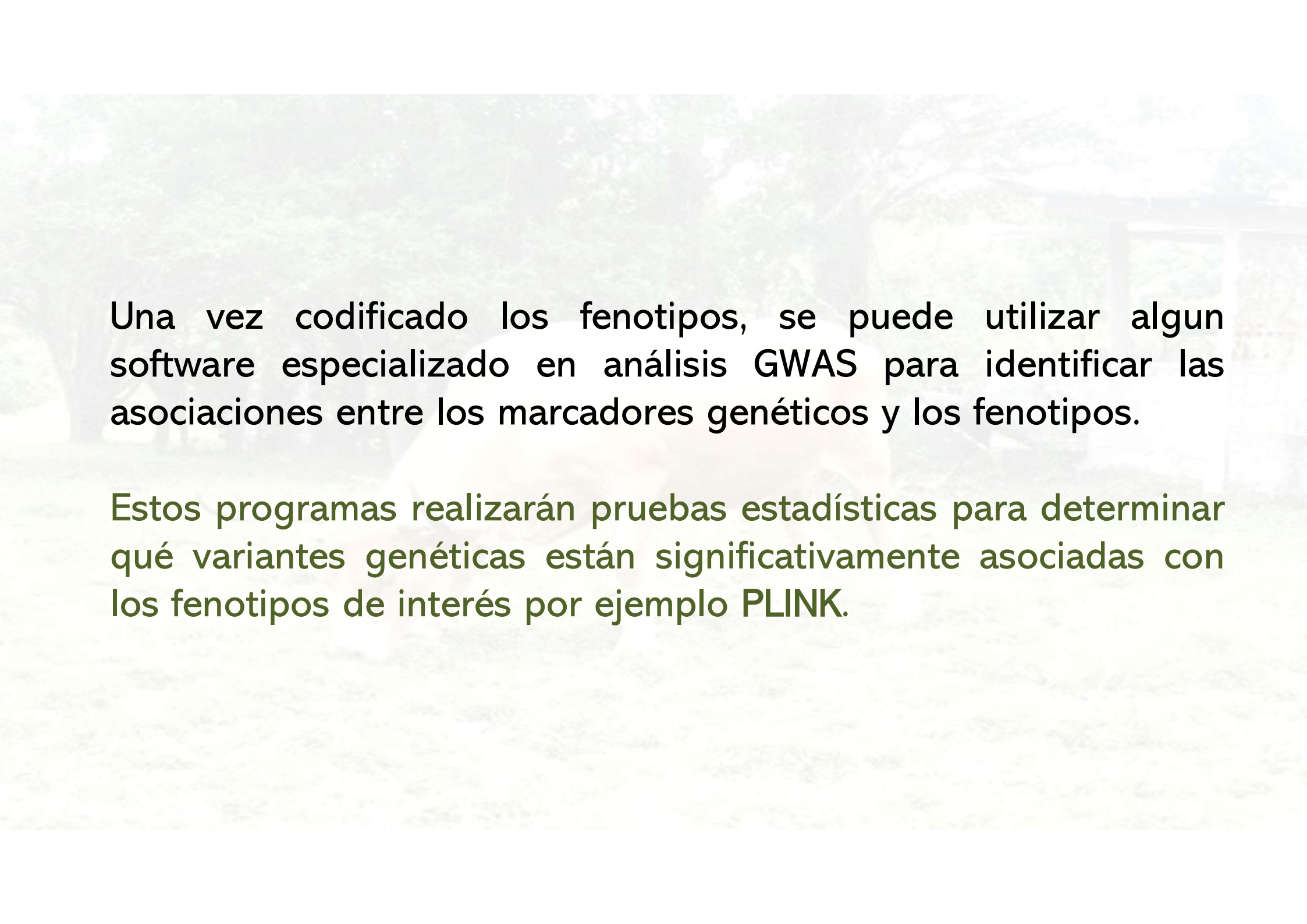
Los fenotipos de salud, como la resistencia a enfermedades o la tasa de mortalidad, también pueden ser valores numéricos.

- Resistencia a Enfermedades:

- 1 = Alta resistencia,
- 2 = Moderada resistencia,
- 3 = Baja resistencia.

- Tasa de Mortalidad (%).





Una vez codificado los fenotipos, se puede utilizar algun software especializado en análisis GWAS para identificar las asociaciones entre los marcadores genéticos y los fenotipos.

Estos programas realizarán pruebas estadísticas para determinar qué variantes genéticas están significativamente asociadas con los fenotipos de interés por ejemplo PLINK.

IID	FIID	Padre	Madre	Sex	Color	Codigo	Cuerno	Codigo	Prod de Leche (litros/día)	GDP (kg/día)	Resistencia	Código	Mortalidad (%)	SNP1	SNP2
1	1	0	0	0	Negro	1	Largos	1	12.5	0.85	Moderada	3	5	AA	GG
2	1	0	0	0	Rojo	2	Cortos	2	14.2	0.92	Alta	1	2	AG	CC
3	2	0	0	0	Pardo	3	Deshornados	3	11.8	0.78	Moderada	2	3	GG	GC
4	2	0	0	0	Negro	1	Cortos	2	13.1	0.88	Alta	1	4	AG	CC
5	3	0	0	0	Bayo	4	Cortos	2	11.6	0.90	Alta	1	4	AA	GC
6	3	0	0	0	Rojo	2	Deshornados	3	12.9	0.88	Moderada	2	5	GG	GC
7	4	0	0	0	Negro	1	Largos	1	15.1	0.60	Baja	3	3	AG	CC
8	4	0	0	0	Negro	1	Cortos	2	11.1	0.78	Baja	3	3	AA	CC

Estos datos se utilizarían en un análisis GWAS para identificar asociaciones entre los genotipos (SNP) y los fenotipos de interés (Color del Pelaje, Longitud de Cuerno, Producción de Leche, Ganancia de Peso Diaria, Resistencia a Enfermedades y Tasa de Mortalidad). El análisis evaluaría qué variantes genéticas están relacionadas de manera significativa con los fenotipos.



Archivos tipo PED
y MAP

Archivo tipo PED Plain Text Pedigree

El archivo PED contiene información sobre los genotipos y los fenotipos de los individuos cada fila representa un individuo y contiene información sobre su familia, padre, madre, género y genotipos en los marcadores genéticos (SNP).

Los fenotipos también se incluyen en las columnas correspondientes, como "ColorCuerno," "ProduccionLeche," "GananciaPeso," "Resistencia," "Enfermedad," y "TasaMortalidad."

1	1	0	0	0	1	1	12.5	0.85	3	5	AA	GG
2	1	0	0	0	2	2	14.2	0.92	1	2	AG	CC
3	2	0	0	0	3	3	11.8	0.78	2	3	GG	GC
4	2	0	0	0	1	2	13.1	0.88	1	4	AG	CC
5	3	0	0	0	4	2	11.6	0.90	1	4	AA	GC
6	3	0	0	0	2	3	12.9	0.88	2	5	GG	GC
7	4	0	0	0	1	1	15.1	0.60	3	3	AG	CC
8	4	0	0	0	1	2	11.1	0.78	3	3	AA	CC

El archivo MAP proporciona información sobre la ubicación y el nombre de los marcadores genéticos utilizados en el análisis

Cada fila representa un marcador genético y contiene tres columnas:

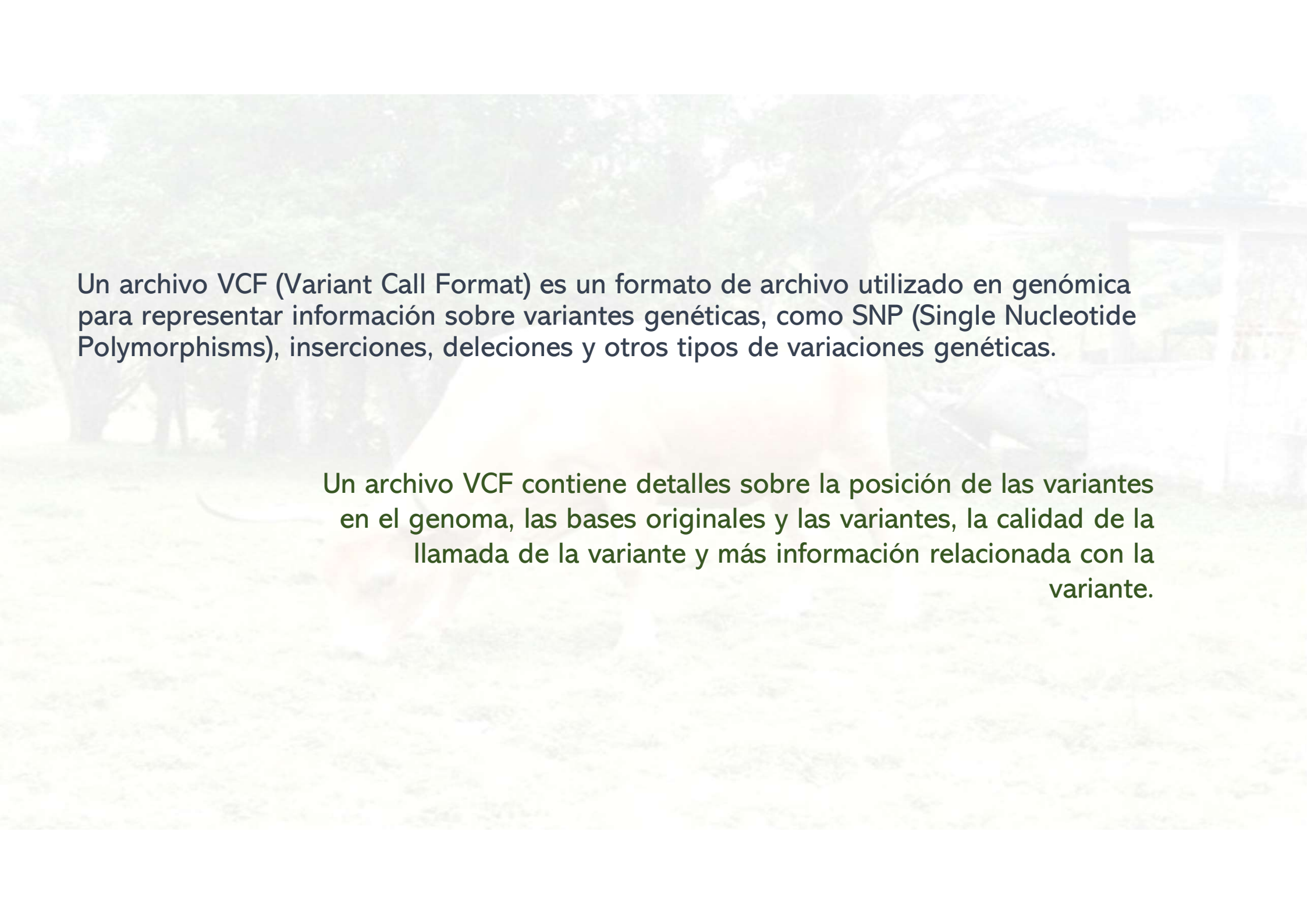
- Nombre del marcador (por ejemplo, "SNP1"),
- Número del cromosoma (por ejemplo, "1" para el cromosoma 1)
- Ubicación física del marcador en el genoma (por ejemplo, "0.0").

Archivo tipo MAP

Marker Allele Frequency

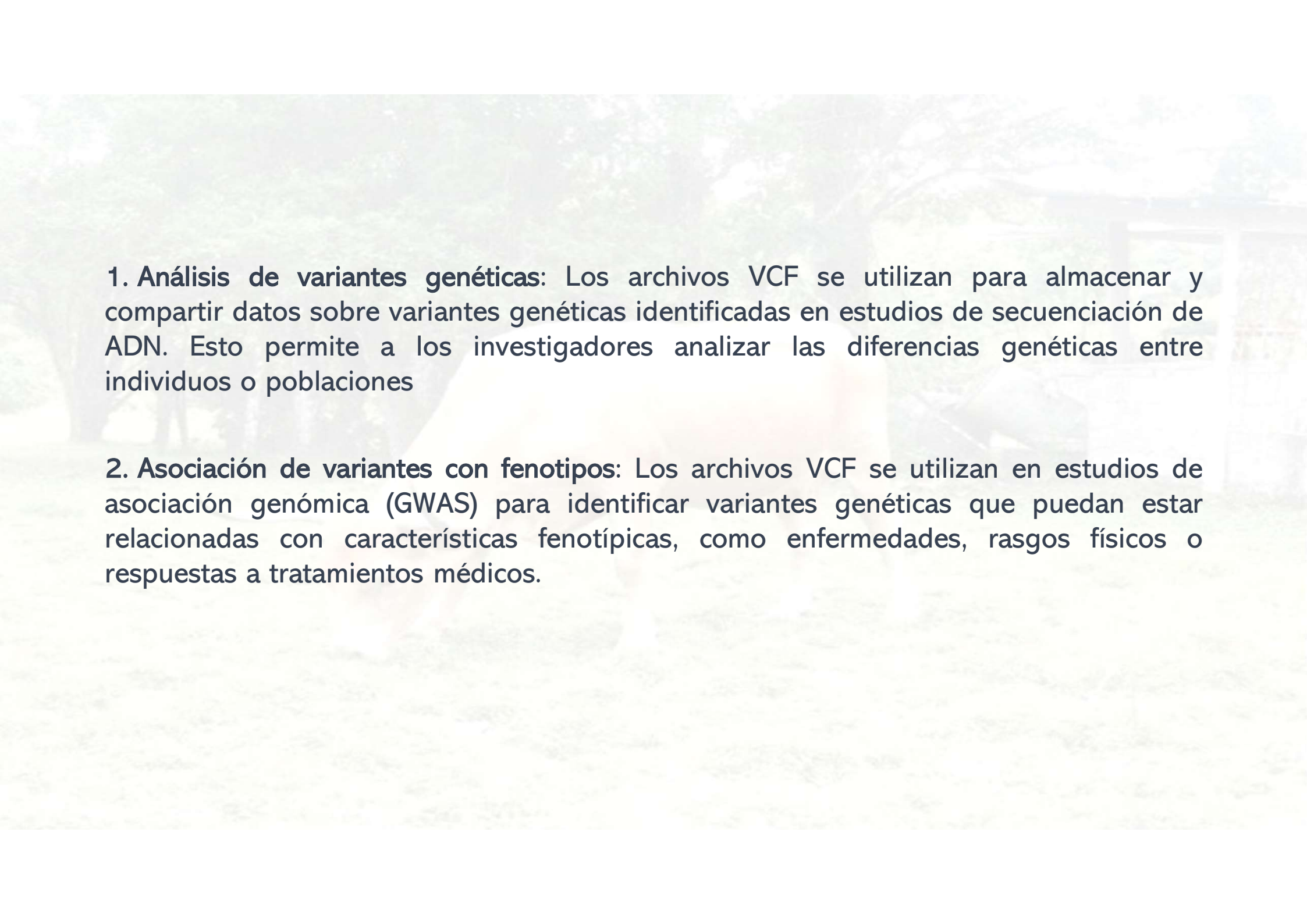
SNP1	1	0.0
SNP2	1	0.1





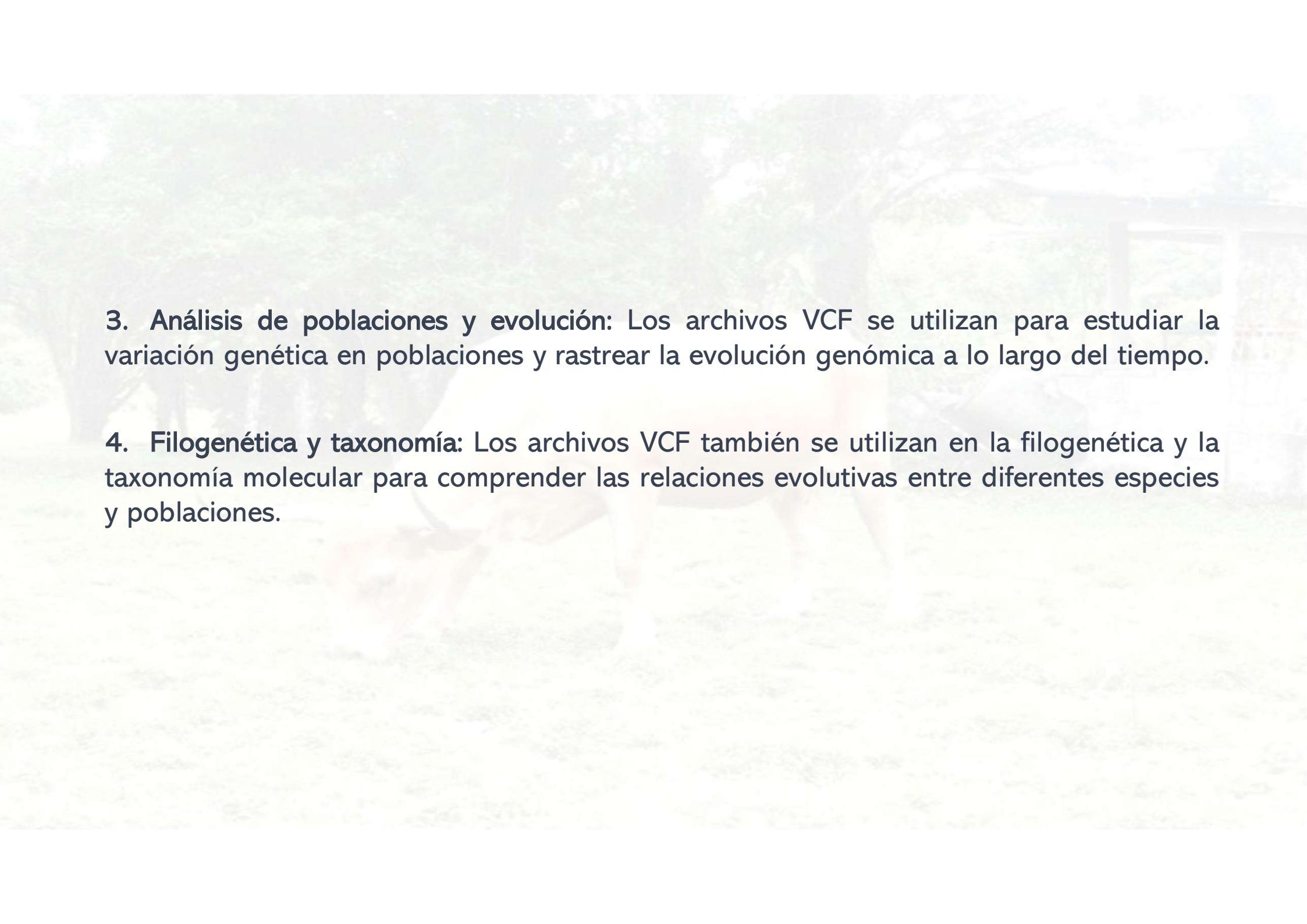
Un archivo VCF (Variant Call Format) es un formato de archivo utilizado en genómica para representar información sobre variantes genéticas, como SNP (Single Nucleotide Polymorphisms), inserciones, deleciones y otros tipos de variaciones genéticas.

Un archivo VCF contiene detalles sobre la posición de las variantes en el genoma, las bases originales y las variantes, la calidad de la llamada de la variante y más información relacionada con la variante.



1. Análisis de variantes genéticas: Los archivos VCF se utilizan para almacenar y compartir datos sobre variantes genéticas identificadas en estudios de secuenciación de ADN. Esto permite a los investigadores analizar las diferencias genéticas entre individuos o poblaciones

2. Asociación de variantes con fenotipos: Los archivos VCF se utilizan en estudios de asociación genómica (GWAS) para identificar variantes genéticas que puedan estar relacionadas con características fenotípicas, como enfermedades, rasgos físicos o respuestas a tratamientos médicos.



3. Análisis de poblaciones y evolución: Los archivos VCF se utilizan para estudiar la variación genética en poblaciones y rastrear la evolución genómica a lo largo del tiempo.

4. Filogenética y taxonomía: Los archivos VCF también se utilizan en la filogenética y la taxonomía molecular para comprender las relaciones evolutivas entre diferentes especies y poblaciones.

Ejemplo de archivo VCF

##fileformat=VCFv4.2

##INFO=<ID=DP,Number=1,Type=Integer,Description="Depth">

##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">

#CHROM	POS	ID	REF	ALT	QUAL	FILTER	INFO	FORMAT	Sample1	Sample2	Sample3	Sample4	Sample5
1	1001	SNP1	A	G	.	PASS	DP=50	GT	0/0	0/1	1/1	0/0	1/0
1	2002	SNP2	C	T	.	PASS	DP=45	GT	1/1	0/0	0/1	1/0	0/0
1	3003	SNP3	G	A	.	PASS	DP=60	GT	0/1	0/1	0/0	1/1	0/0
1	4004	SNP4	T	C	.	PASS	DP=55	GT	0/0	0/0	1/1	0/0	1/1
1	5005	SNP5	A	T	.	PASS	DP=48	GT	0/1	0/0	0/1	1/0	0/1


```
##fileformat=VCFv4.2
##FILTER=<ID=PASS,Description="All filters passed">
##fileDate=20200504
##source=apt-format-result:2.11.0
##reference=NCBI Bos_taurus_UMD_3.1.1
##reference=genome-version-ncbi(Bos_taurus_UMD_3.1.1)
##reference=genome-version-ucsc(bosTau8)
##reference=netaffx-version(35.r2.a1)
##phasing=none
##CHROM=<ID=UNKNOWNCHR,Number=1,Type=String,Description="chromosome or contig not available">
##INFO=<ID=UNKNOWNPOSITION,Number=0,Type=Flag,Description="Unknown position">
##INFO=<ID=probeset_id,Number=1,Type=String,Description="Measured probeset">
##INFO=<ID=RSID,Number=1,Type=String,Description="dbSNP RS ID">
##INFO=<ID=CR,Number=1,Type=Float,Description="Call rate">
##INFO=<ID=ConversionType,Number=1,Type=String,Description="Probeset category">
##FILTER=<ID=FAIL,Description="Annotated by SNP QC as not the best and recommended probeset for the marker.">
##contig=<ID=1>
##contig=<ID=2>
##contig=<ID=3>
##contig=<ID=4>
##contig=<ID=5>
##INFO=<ID=AC,Number=A,Type=Integer,Description="Allele count in genotypes">
##INFO=<ID=AN,Number=1,Type=Integer,Description="Total number of alleles in called genotypes">
##bcftools_viewVersion=1.9+htslib-1.9
##bcftools_viewCommand=view -S listKeep merged_ok_SNPs.vcf; Date=Fri May 29 23:09:24 2020
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT GUA1 GUA2 GUA3 GUA4 GUA6 GUA7 GUA8 GUA9 GUA10 GUA11 GUA12
1 135098 Affx-106525819 A G . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs41609588;AC=48;AN=68 GT 1/1 0/0
1 267940 Affx-106521083 C T . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs108982244;AC=21;AN=68 GT 1/1 0/0
1 393248 Affx-106517367 C A . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs43703977;AC=40;AN=68 GT 1/1 0/0
1 471078 Affx-106514379 C T . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs110887329;AC=29;AN=68 GT 1/1 1/1
1 516404 Affx-113716754 G A . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs29015852;AC=25;AN=68 GT 0/0 0/0
1 883895 Affx-106497142 C T . PASS CR=98;ConversionType=PolyHighResolution;RSID=rs110467572;AC=29;AN=68 GT 1/1 0/0
1 929617 Affx-113714070 A G . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs109605346;AC=3;AN=68 GT 0/0 0/0
1 1009504 Affx-106525870 A C . PASS CR=97;ConversionType=PolyHighResolution;RSID=rs110064393;AC=33;AN=68 GT 0/0 1/1
1 1359951 Affx-106506942 C T . PASS CR=98;ConversionType=PolyHighResolution;RSID=rs41255293;AC=44;AN=64 GT ./ 1/1
1 1390292 Affx-257081007 G A . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs110490165;AC=30;AN=68 GT 0/0 0/0
1 1768587 Affx-115869978 C A . PASS CR=100;ConversionType=MonoHighResolution;RSID=rs210350155;AC=0;AN=68 GT 0/0 0/0
1 2049400 Affx-113744659 A G . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs110875985;AC=49;AN=68 GT 1/1 1/1
1 2211567 Affx-113731149 G A . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs109393868;AC=25;AN=68 GT 0/0 1/1
1 2313042 Affx-106533345 T C . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs110490165;AC=30;AN=68 GT 1/1 1/1
1 2415018 Affx-113737642 A C . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs108978478;AC=51;AN=68 GT 1/1 1/1
1 3079342 Affx-113738133 A G . PASS CR=97;ConversionType=PolyHighResolution;RSID=rs43180934;AC=33;AN=68 GT 0/0 1/1
1 3197378 Affx-113747314 C T . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs42381983;AC=32;AN=68 GT 0/0 1/1
1 3249057 Affx-106544056 G T . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs29012842;AC=46;AN=68 GT 1/1 1/1
1 4052161 Affx-115865059 T C . PASS CR=99;ConversionType=PolyHighResolution;RSID=rs41633082;AC=32;AN=68 GT 0/0 0/0
1 4311365 Affx-106496760 T C . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs43204112;AC=42;AN=68 GT 1/1 1/1
1 5179453 Affx-106546631 T C . PASS CR=99;ConversionType=PolyHighResolution;RSID=rs42845302;AC=39;AN=66 GT 0/0 1/1
1 5519845 Affx-106542484 C T . PASS CR=100;ConversionType=PolyHighResolution;RSID=rs43210840;AC=17;AN=68 GT 0/0 0/0
```

EMPLEO DE DATOS PARA ANÁLISIS DE ASOCIACIÓN FENOTIPO-GENOTIPO



Paso 1: Crear un archivo de configuración PLINK (plink.ped y plink.map)

Primero, debes tener dos archivos: uno para los datos fenotípicos y genotípicos (plink.ped) y otro para la información de los marcadores (plink.map). Aquí está un ejemplo de cómo podrían ser estos archivos:

plink.ped	FID	IID	PID	MID	SEX	PHE	GENOTIPE DATA									
	1	1	0	0	1	1	A	A	G	G	C	C	T	T	A	A
	2	1	0	0	2	2	C	C	T	T	G	G	T	T	G	G
	3	1	0	0	1	2	A	A	T	T	C	C	C	C	T	T
	4	2	0	0	2	1	G	G	C	C	G	G	G	G	A	A
	5	2	0	0	1	2	C	C	G	G	T	T	C	C	G	G
	6	1	0	0	2	1	T	T	G	G	T	T	C	C	T	T
	7	2	0	0	1	1	A	A	G	G	C	C	C	C	G	G
	8	1	0	0	2	2	C	C	T	T	G	G	C	C	G	G
	9	2	0	0	1	2	G	G	C	C	G	G	T	T	C	C
10	2	0	0	2	1	T	T	G	G	C	C	G	G	A	A	

En el archivo PLINK .ped, cada columna representa información específica sobre los individuos en tu estudio. Aquí está la interpretación de cada columna en el archivo PLINK .ped:

1. **Family ID (FID):** Identificación de la familia o grupo al que pertenece el individuo.
2. **Individual ID (IID):** Identificación única de cada individuo en el estudio.
3. **Paternal ID (PID):** Identificación del padre biológico del individuo
4. **Maternal ID (MID):** Identificación de la madre biológica del individuo.
5. **Sex (SEX):** Género del individuo. Generalmente, se utiliza una convención numérica, donde 1 representa a los hombres y 2 a las mujeres.
6. **Phenotype (PHE):** El valor fenotípico del individuo para el rasgo o característica de interés en el estudio. Puede ser una puntuación, un diagnóstico binario, una categoría, etc.
7. **Genotype Data:** A partir de la columna 7 en adelante, se proporcionan los genotipos de los marcadores genómicos para cada individuo. En el ejemplo, se usan letras para representar los alelos marcadores

El archivo .map contiene información sobre cada marcador (por lo general SNPs) que se genotipó en el archivo .ped. Cada línea del archivo representa un marcador y las columnas están organizadas de la siguiente manera:

plink.map

1 SNP1 0 1

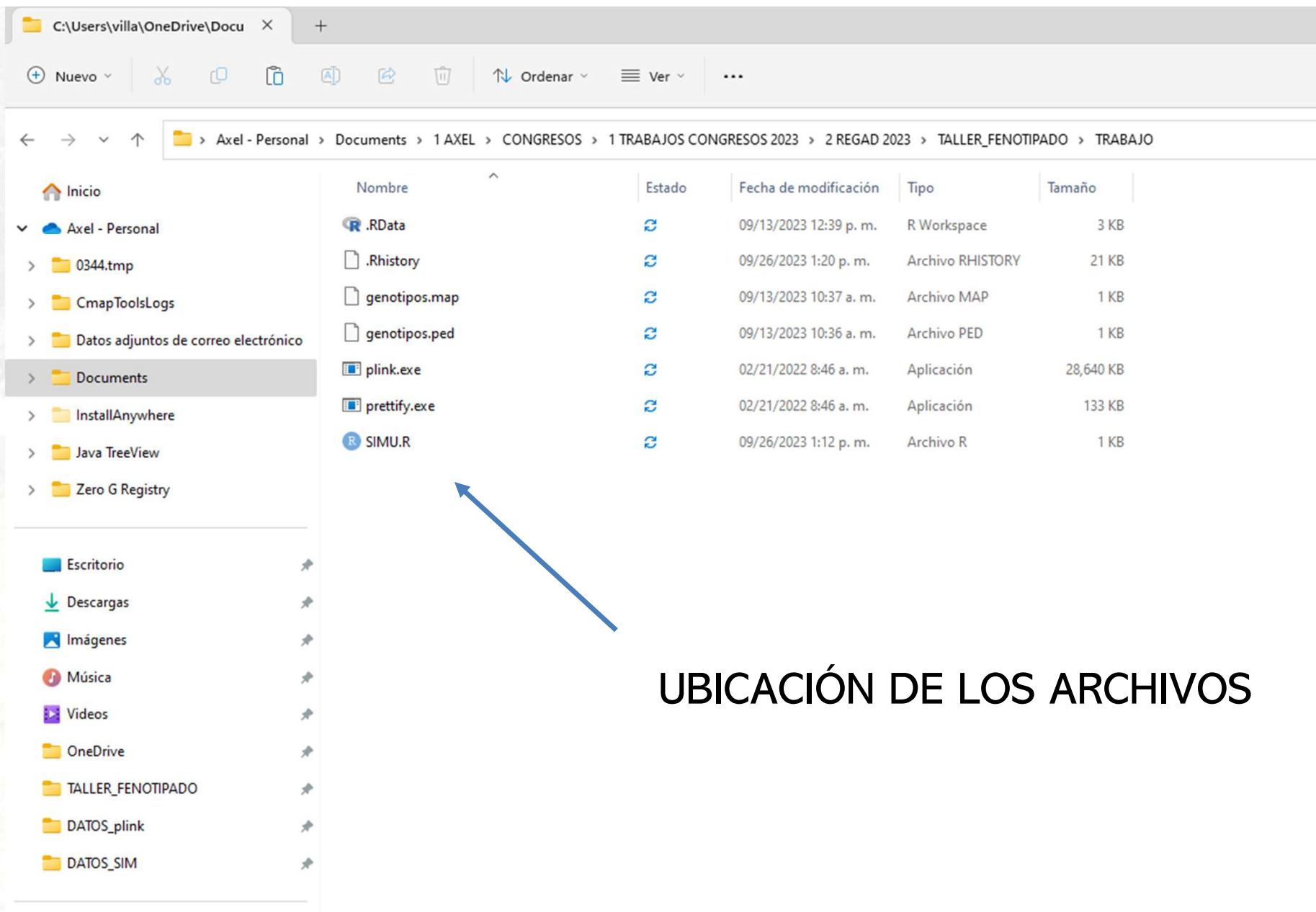
1 SNP2 0 2

1 SNP3 0 3

1 SNP4 0 4

1 SNP5 0 5

1. **Cromosoma:** Número del cromosoma o designación del cromosoma.
2. **ID del marcador:** Identificación única para el SNP/marcador.
3. **Distancia genética:** Distancia desde el marcador anterior (generalmente en morgans o centimorgans). En muchos archivos .map, este valor puede ser 0 si no se dispone de datos de mapa genético.
4. **Posición física:** Posición del marcador/SNP en el cromosoma (por lo general en bases/pares de bases).



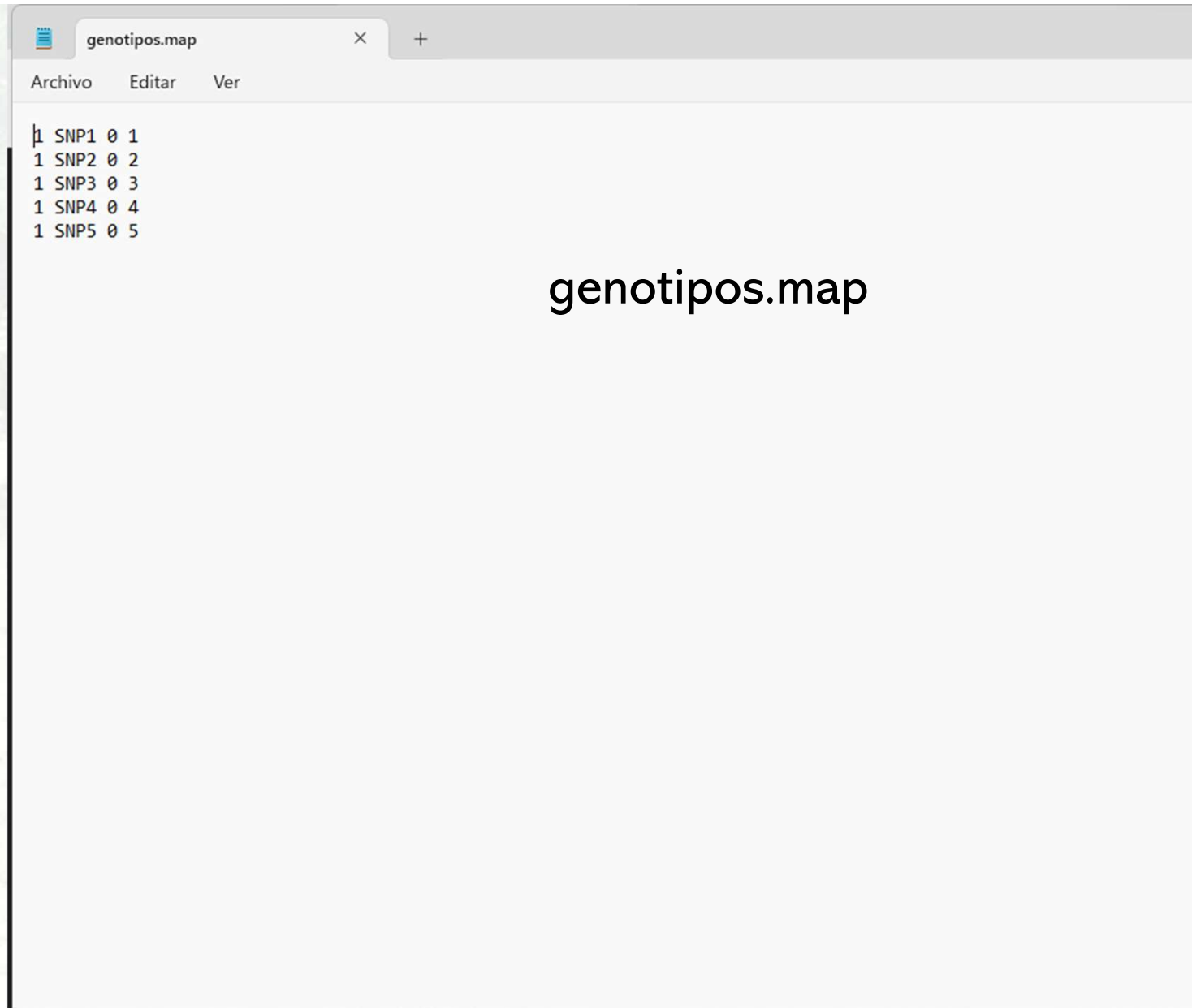
UBICACIÓN DE LOS ARCHIVOS

genotipos.ped

Archivo Editar Ver

```
1 1 0 0 1 1 A A G G C C T T A A
2 1 0 0 2 2 C C T T G G T T G G
3 1 0 0 1 2 A A T T C C C C T T
4 2 0 0 2 1 G G C C G G G G A A
5 2 0 0 1 2 C C G G T T C C G G
6 1 0 0 2 1 T T G G T T C C T T
7 2 0 0 1 1 A A G G C C C C G G
8 1 0 0 2 2 C C T T G G C C G G
9 2 0 0 1 2 G G C C G G T T C C
10 2 0 0 2 1 T T G G C C G G A A
```

genotipos.ped



genotipos.map

Ejemplo simplificado de unificación de Archivos VCF y Fenotipos

Archivo de fenotipos se presenta en archivo tipo CSV

ID,Phenotype

1,0.75

2,0.60

3,0.90

4,0.45

5,0.80

Donde:

ID: identificador del individuo

Phenotype: fenotipo de interés.

Contiene información para cinco individuos.

```
##fileformat=VCFv4.2
##INFO=<ID=DP,Number=1,Type=Integer,Description="Depth">
##FORMAT=<ID=GT,Number=1,Type=String,Description="Genotype">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT Sample1 Sample2
Sample3 Sample4 Sample5
1 1001 SNP1 A G . PASS DP=50 GT 0/0 0/1 1/1 0/0 1/0
1 2002 SNP2 C T . PASS DP=45 GT 1/1 0/0 0/1 1/0 0/0
1 3003 SNP3 G A . PASS DP=60 GT 0/1 0/1 0/0 1/1 0/0
1 4004 SNP4 T C . PASS DP=55 GT 0/0 0/0 1/1 0/0 1/1
1 5005 SNP5 A T . PASS DP=48 GT 0/1 0/0 0/1 1/0 0/1
```

Este archivo VCF de ejemplo contiene información ficticia de cinco SNPs y cinco individuos. Los genotipos están representados en el formato "**GT**" (genotype), donde "**0/0**" indica homocigoto para el alelo de referencia, "**0/1**" indica heterocigoto y "**1/1**" indica homocigoto

Ahora, para unificar el archivo VCF y el archivo CSV de fenotipos en PLINK 1.9, puedes utilizar el siguiente script:

Crear un archivo .ped desde el archivo CSV de fenotipos

plink --vcf ejemplo.vcf --pheno fenotipos.csv --make-bed --out resultado

Unificar los archivos de genotipos y fenotipos

plink --bfile resultado --merge-list lista-archivos.txt --out unificado

Donde "lista-archivos.txt" es un archivo de texto que contiene la lista de archivos VCF a unificar:

ejemplo.vcf

(otros archivos VCF si es necesario)

Finalmente, puedes utilizar el archivo unificado (unificado.bed, unificado.bim, unificado.fam) para tus análisis en PLINK 1.9.

A white bull with a red collar is grazing in a grassy field. In the background, there are green trees and a wooden building. The text "Análisis de Datos con PLINK 1.9" is overlaid on the image.

Análisis de Datos con PLINK 1.9

SOFTWARES REQUERIDOS

- [The R Project for Statistical Computing](#)
- RStudio Desktop
- [PLINK 1.9 - cog-genomics.org](#)

C:\Users\villa\OneDrive\Docu X +

Nuevo ▾

Ordenar ▾ Ver ▾ ...

← → ▾ ↑

Inicio

▼ Axel - Personal

> 0344.tmp

> CmapToolsLogs

> Datos adjuntos de correo electrónico

> Documents

> InstallAnywhere

> Java TreeView

> Zero G Registry

Escritorio

Descargas

Imágenes

Música

Videos

OneDrive

TALLER_FENOTIPADO

DATOS_plink

DATOS_SIM

Nombre	Estado	Fecha de modificación	Tipo	Tamaño
.RData	↻	09/13/2023 12:39 p. m.	R Workspace	3 KB
.Rhstory	↻	09/26/2023 1:20 p. m.	Archivo RHISTORY	21 KB
genotipos.map	↻	09/13/2023 10:37 a. m.	Archivo MAP	1 KB
genotipos.ped	↻	09/13/2023 10:36 a. m.	Archivo PED	1 KB
plink.exe	↻	02/21/2022 8:46 a. m.	Aplicación	28,640 KB
prettify.exe	↻	02/21/2022 8:46 a. m.	Aplicación	133 KB
SIMU.R	↻	09/26/2023 1:12 p. m.	Archivo R	1 KB

ANÁLISIS CON ARCHIVOS MAP Y PED UTILIZANDO PLINK 1.9

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins Project: (None)

SIMU.R*

Source on Save Run Source

```
1 # clear workplace
2 rm(list = ls())
3
4 # Load the qqman library
5 library(qqman)
6
7 # Analisis de asociacion
8 system("plink.exe --file genotipos --assoc --out gwas_results")
9
10 # Read the GWAS results
11 results <- read.table("gwas_results.assoc", header = TRUE)
12
13 # Crear Tabla de Resultados en Excel
14 library(openxlsx)
15 write.xlsx(results, "asoc.xlsx") # ver por genotipo
16
17 # Create a Manhattan plot without labels first
18 manhattan_plot <- manhattan(results, main = "Manhattan Plot")
19
20 # Customize the Manhattan plot by adding labels to specific SNPs
21 text(x = results$BP, y = -log10(results$P), labels = results$SNP,
22      pos = 3, cex = 0.9, col = "blue")
23
24 # Show the modified Manhattan plot
25 print(manhattan_plot)
```

Limpiamos área de trabajo

Cargamos Librería de análisis

Environment History Console

Global Environment

Environment is empty

Files Plots Packages Help

TALLER_FENOTIPADO > TRABAJO

Name

- ..
- SIMU.R
- .RData
- plink.exe
- prettify.exe
- genotipos.map
- genotipos.ped

Console Terminal Background Jobs

R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/2 REGAD 2023/TALLER_FENOTIPADO/TRABAJO/

results <- read.table("gwas_results.assoc", header = TRUE)

View(results)

RStudio interface showing an R script and its execution output.

Script Editor (SIMU.R):

```
1 # clear workplace
2 rm(list = ls())
3
4 # Load the qqman library
5 library(qqman)
6
7 # Analisis de asociacion
8 system("plink.exe --file genotipos --assoc --out gwas_results")
9
10 # Read the GWAS results
11 results <- read.table("gwas_results.assoc", header = TRUE)
12
13 # Crear Tabla de Resultados en Excel
14 library(openxlsx)
15 write.xlsx(results, "asoc.xlsx") # ver por genotipo
16
17 # Create a Manhattan plot without labels first
18 manhattan_plot <- manhattan(results, main = "Manhattan Plot")
```

Console:

```
R 4.3.0 ~ C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/2 REGAD 2023/TALLER_FENOTIPADO/TRABAJO/
writing c/c --assoc report to gwas_results.assoc ... done.
Warning: Variant 1 quadallelic; setting rarest alleles missing.
Warning: Variant 2 triallelic; setting rarest alleles missing.
Warning: Variant 3 triallelic; setting rarest alleles missing.
Warning: Variant 4 triallelic; setting rarest alleles missing.
Warning: Variant 5 quadallelic; setting rarest alleles missing.
[1] 0
>
```

Environment Panel:

Environment is empty

Files Panel:

- ..
- SIMU.R
- .RData
- plink.exe
- prettify.exe
- genotipos.map
- genotipos.ped
- gwas_results.log
- gwas_results.assoc

SIMU.R* x gwas_results.assoc x

	CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	1	SNP1	1	C	0.75	0	A	6	0.01431	NA
2	1	SNP2	2	T	0.75	0	G	9.6	0.001946	NA
3	1	SNP3	3	G	0.75	0.25	C	4	0.0455	9
4	1	SNP4	4	T	0.4	0.3333	C	0.07111	0.7897	1.333
5	1	SNP5	5	A	0	0.75	G	7.875	0.005012	0

1:1 Text file

Console Terminal x Background Jobs x

R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/2 REGAD 2023/TALLER_FENOTIPADO/TRABAJO/

Among remaining phenotypes, 5 are cases and 5 are controls.
Writing c/c --assoc report to gwas_results.assoc ... done.
Warning: variant 1 quadallelic; setting rarest alleles missing.
Warning: variant 2 triallelic; setting rarest alleles missing.
Warning: variant 3 triallelic; setting rarest alleles missing.
Warning: variant 4 triallelic; setting rarest alleles missing.
Warning: variant 5 quadallelic; setting rarest alleles missing.
[1] 0
> |

Environment History Con 128 MiB Global Environment

Environment is empty

Files Plots Packages H

TALLER_FENOTIPADO > TRABAJO

	Name
<input type="checkbox"/>	SIMU.R
<input type="checkbox"/>	.RData
<input type="checkbox"/>	.Rhistory
<input type="checkbox"/>	plink.exe
<input type="checkbox"/>	prettify.exe
<input type="checkbox"/>	genotipos.map
<input type="checkbox"/>	genotipos.ped
<input type="checkbox"/>	gwas_results.log
<input type="checkbox"/>	gwas_results.assoc

RStudio interface showing an R script and its execution results.

Script Editor (SIMU.R):

```
1 # clear workplace
2 rm(list = ls())
3
4 # Load the qqman library
5 library(qqman)
6
7 # Analisis de asociacion
8 system("plink.exe --file genotipos --assoc --out gwas_results")
9
10 # Read the GWAS results
11 results <- read.table("gwas_results.assoc", header = TRUE)
12
13 # Crear Tabla de Resultados en Excel
14 library(openxlsx)
15 write.xlsx(results, "asoc.xlsx") # ver por genotipo
16
17 # Create a Manhattan plot without labels first
18 manhattan_plot <- manhattan(results, main = "Manhattan Plot")
```

Environment Panel:

- Global Environment
- Data: results (5 obs. of 10 v...)

Files Panel:

- 23 > TALLER_FENOTIPADO > TRABAJO
- Files: .., SIMU.R, .RData, plink.exe, prettify.exe, genotipos.map, genotipos.ped, gwas_results.log, gwas_results.assoc

Console:

```
R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/2 REGAD 2023/TALLER_FENOTIPADO/TRABAJO/
Warning: Variant 2 triallelic; setting rarest alleles missing.
Warning: Variant 3 triallelic; setting rarest alleles missing.
Warning: Variant 4 triallelic; setting rarest alleles missing.
Warning: Variant 5 quadallelic; setting rarest alleles missing.
[1] 0
> # Read the GWAS results
> results <- read.table("gwas_results.assoc", header = TRUE)
>
```


RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

Environment History Connect

139 MiB Global Environment

Data

results 5 obs. of 10 v...

	CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	1	SNP1	1	C	0.75	0.0000	A	6.00000	0.014310	NA
2	1	SNP2	2	T	0.75	0.0000	G	9.60000	0.001946	NA
3	1	SNP3	3	G	0.75	0.2500	C	4.00000	0.045500	9.000
4	1	SNP4	4	T	0.40	0.3333	C	0.07111	0.789700	1.333
5	1	SNP5	5	A	0.00	0.7500	G	7.87500	0.005012	0.000

[No Title]

Showing 1 to 5 of 5 entries, 10 total columns

Console Terminal Background Jobs

```
R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/2 REGAD 2023/TALLER_FENOTIPADO/TRABAJO/  
warning: variant 3 triallelic; setting rarest alleles missing.  
warning: variant 4 triallelic; setting rarest alleles missing.  
warning: variant 5 quadallelic; setting rarest alleles missing.  
[1] 0  
> # Read the GWAS results  
> results <- read.table("gwas_results.assoc", header = TRUE)  
> view(results)  
>
```

Files Plots Packages Help

23 > TALLER_FENOTIPADO > TRABAJO

Name

- ..
- SIMU.R
- .RData
- plink.exe
- prettify.exe
- genotipos.map
- genotipos.ped
- gwas_results.log
- gwas_results.assoc

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

Environment History Connections

120 MiB

R Global Environment

Data

results 5 obs. of 10 v...

Files Plots Packages Help

23 > TALLER_FENOTIPADO > TRABAJO

Name

- SIMU.R
- .RData
- plink.exe
- prettify.exe
- genotipos.map
- genotipos.ped
- gwas_results.log
- gwas_results.assoc
- asoc.xlsx










```
3
4 # Load the qqman library
5 library(qqman)
6
7 #Análisis de asociación
8 system("plink.exe --file genotipos --assoc --out gwas_results")
9
10 # Read the GWAS results
11 results <- read.table("gwas_results.assoc", header = TRUE)
12
13 #Crear Tabla de Resultados en Excel
14 library(openxlsx)
15 write.xlsx(results,"asoc.xlsx")#ver tabla
16
17 # Create a Manhattan plot without labels first
18 manhattan_plot <- manhattan(results, main = "Manhattan Plot")
19
20 # Customize the Manhattan plot by adding labels to specific SNPs
```

17:1 (Top Level) R Script

Console Terminal Background Jobs

R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/2 REGAD 2023/TALLER_FENOTIPADO/TRABAJO/

```
> # Read the GWAS results
> results <- read.table("gwas_results.assoc", header = TRUE)
> view(results)
> view(results)
> #Crear Tabla de Resultados en Excel
> library(openxlsx)
> write.xlsx(results,"asoc.xlsx")#ver tabla
>
```


Nombre	Estado	Fecha de modificación	Tipo	Tamaño
 .RData	✓	09/13/2023 12:39 p. m.	R Workspace	3 KB
 asoc.xlsx ←	↻	09/28/2023 12:16 p. m.	Microsoft Excel W...	7 KB
 genotipos.map	✓	09/13/2023 10:37 a. m.	Archivo MAP	1 KB
 genotipos.ped	✓	09/13/2023 10:36 a. m.	Archivo PED	1 KB
 gwas_results.assoc ←	↻	09/28/2023 12:07 p. m.	Archivo ASSOC	1 KB
 gwas_results.log	↻	09/28/2023 12:07 p. m.	Documento de te...	2 KB
 plink.exe	✓	02/21/2022 8:46 a. m.	Aplicación	28,640 KB
 prettify.exe	✓	02/21/2022 8:46 a. m.	Aplicación	133 KB
 SIMU.R	✓	09/26/2023 1:12 p. m.	Archivo R	1 KB

File Home Insert Page Layout Formulas Data Review View Help Power Pivot

Clipboard: Paste, Cut, Copy, Format Painter

Font: Calibri, 11, Bold, Italic, Underline, Text Color, Background Color

Alignment: Left, Center, Right, Indent, Wrap Text, Merge & Center

Number: General, Percentage, Decimal, Fraction, More Options

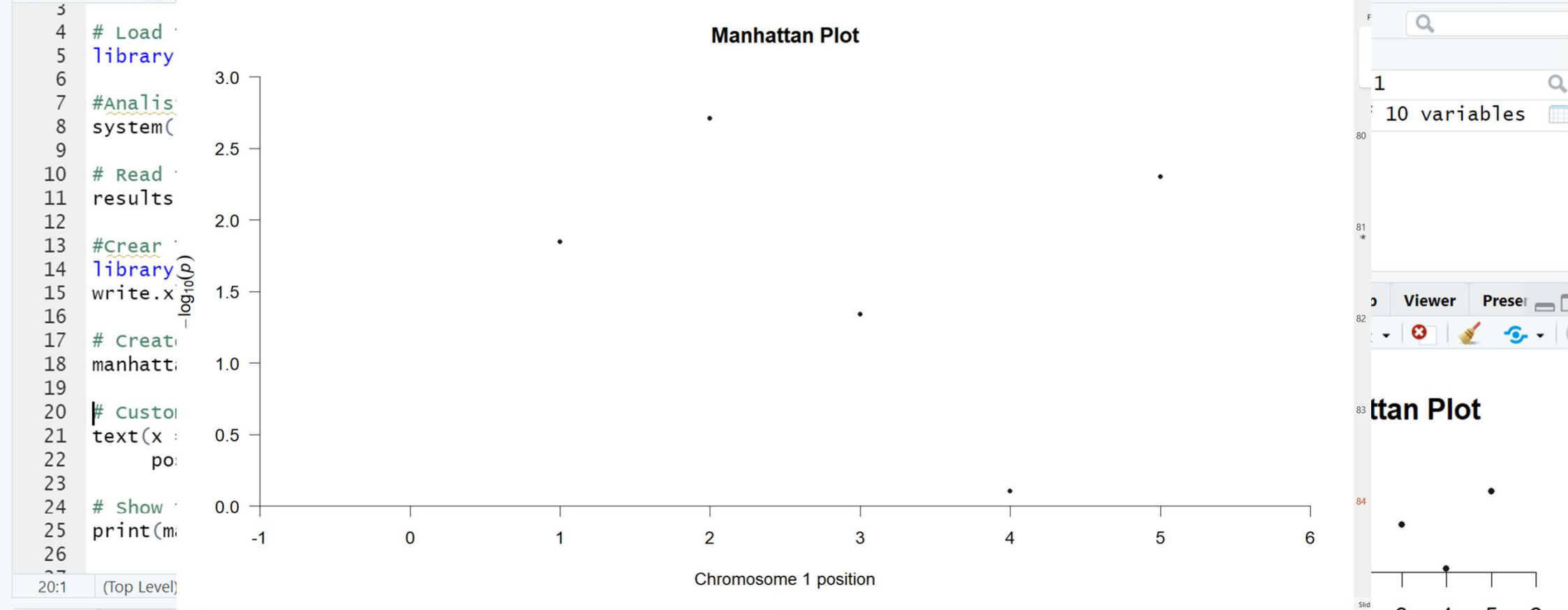
Styles: Normal, Bad, Good, Neutral, Calculation, Check Cell

Cells: Insert, Delete, Format

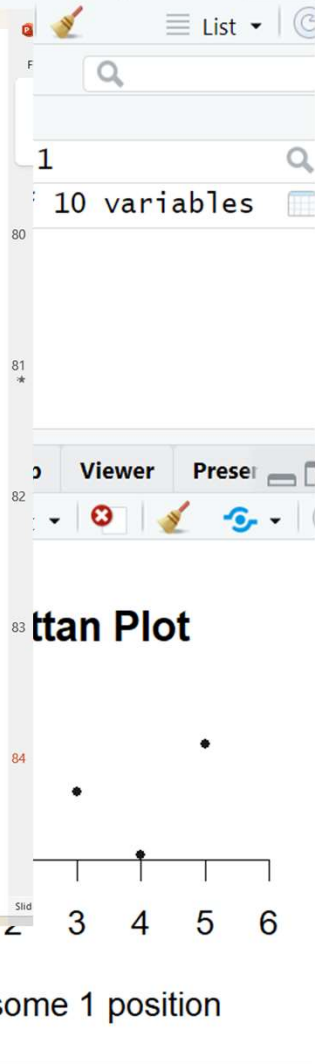
Editing: AutoSum, Fill, Clear, Sort & Filter, Find & Select

Analysis: Analyze Data, Add-ins

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
1	CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR													
2	1	SNP1	1	C	0.75	0	A	6	0.01431	NA													
3	1	SNP2	2	T	0.75	0	G	9.6	0.001946	NA													
4	1	SNP3	3	G	0.75	0.25	C	4	0.0455	9													
5	1	SNP4	4	T	0.4	0.3333	C	0.07111	0.7897	1.333													
6	1	SNP5	5	A	0	0.75	G	7.875	0.005012	0													
7																							
8																							
9																							
10																							
11																							
12																							
13																							
14																							
15																							
16																							
17																							
18																							
19																							
20																							
21																							
22																							
23																							
24																							
25																							
26																							
27																							
28																							
29																							
30																							
31																							
32																							
33																							
34																							
35																							
36																							
37																							



R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/2 REGAD 2023/TALLER_FENOTIPADO/TRABA
 > manhattan_plot <- manhattan(results, main = "Manhattan Plot")
 >



RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

SIMU.R*

Source on Save Run

```
4 # Load the qqman library
5 library(qqman)
6
7 #Análisis de asociación
8 system("plink.exe --file genotipos --assoc --out gwas_results")
9
10 # Read the GWAS results
11 results <- read.table("gwas_results.assoc", header = TRUE)
12
13 #Crear Tabla de Resultados en Excel
14 library(openxlsx)
15 write.xlsx(results, "asoc.xlsx")#ver tabla
16
17 # Create a Manhattan plot without labels first
18 manhattan_plot <- manhattan(results, main = "Manhattan Plot")
19
20 # Customize the Manhattan plot by adding labels to specific SNPs
21 text(x = results$BP, y = -log10(results$P), labels = results$SNP,
22      pos = 3, cex = 0.9, col = "blue")
23
24 # Show the modified Manhattan plot
25 print(manhattan_plot)
26
27
```

24:1 (Top Level) R Script

Console Terminal Background Jobs

R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/2 REGAD 2023/TALLER_FENOTIPADO/TRABA

pos = 3, cex = 0.9, col = blue

Project: (None)

Environment History Connections Tutorial

115 MiB

List

R Global Environment

Data

manhattan...	List of 1
results	5 obs. of 10 variables

Files Plots Packages Help Viewer Present

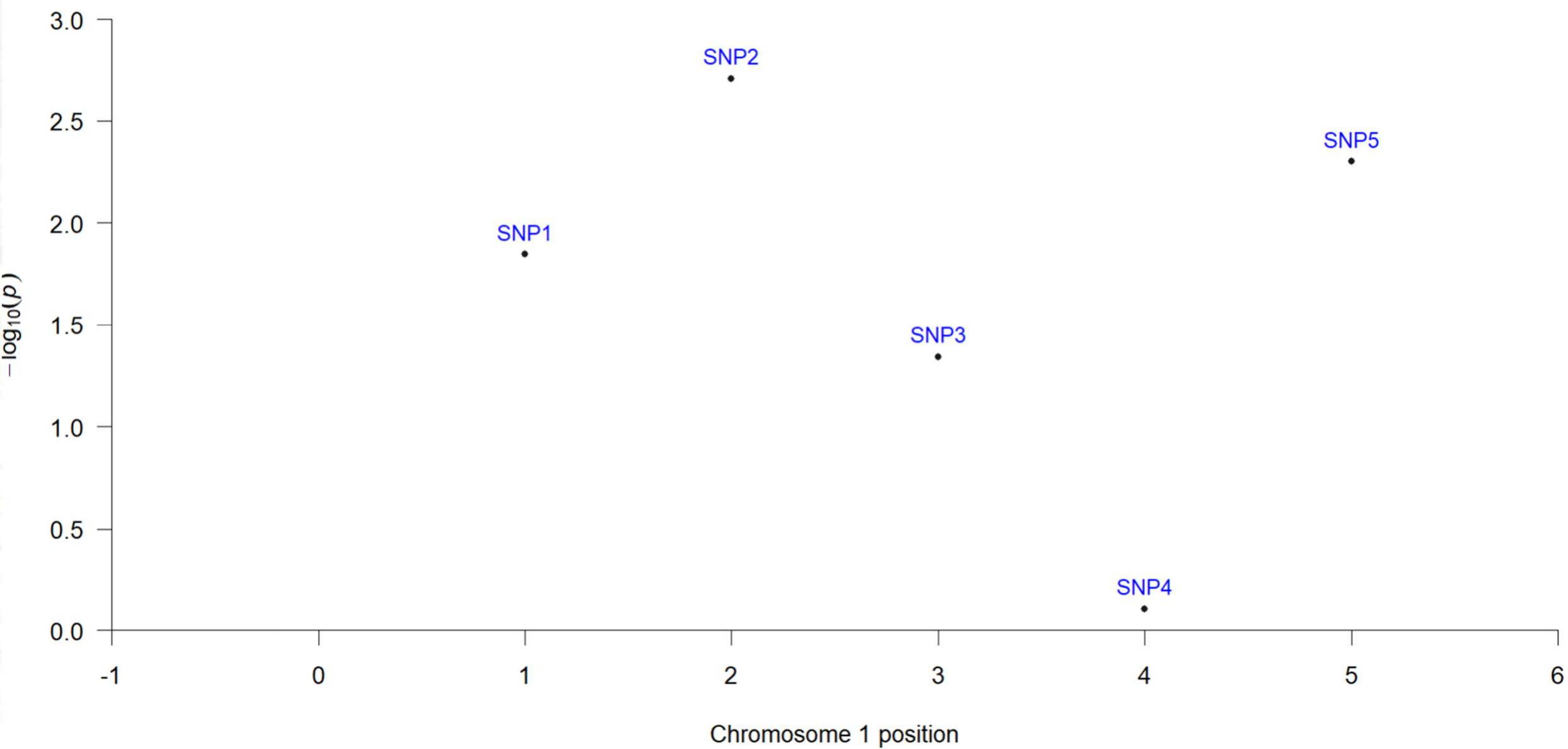
Zoom Export

Manhattan Plot

Chromosome 1 position

SNP	Chromosome 1 position	$-\log_{10}(p)$
SNP1	1	2.0
SNP3	3	1.5
SNP4	4	0.5
SNP5	5	2.5

Manhattan Plot





Evaluación de resultados con PLINK 1.9

CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	SNP1	1	C	0.75	0.00	A	6.00	0.014	NA
1	SNP2	2	T	0.75	0.00	G	9.60	0.002	NA
1	SNP3	3	G	0.75	0.25	C	4.00	0.046	9.00
1	SNP4	4	T	0.40	0.33	C	0.07	0.790	1.33
1	SNP5	5	A	0.00	0.75	G	7.88	0.005	0.00

1. **CHR**: Esta columna muestra el número de cromosoma al que pertenece el marcador genómico.
2. **SNP**: Aquí se muestra el nombre o identificador del marcador genómico.
3. **BP**: Indica la posición física del marcador en la secuencia del genoma humano.
4. **A1**: Muestra el alelo mayoritario (alfabeto) en la muestra.
5. **F_A**: Proporción de alelo mayoritario en casos (fenotipo positivo).
6. **F_U**: Proporción de alelo mayoritario en controles (fenotipo negativo).
7. **A2**: Muestra el alelo minoritario (alfabeto) en la muestra.
8. **CHISQ**: Estadística chi-cuadrado que evalúa la asociación entre el marcador y el fenotipo.
9. **p**: El valor p asociado a la estadística chi-cuadrado, que indica la significancia estadística de la asociación.
10. **OR**: La razón de probabilidades (odds ratio) que indica la fuerza y dirección de la asociación.

CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	SNP1	1	C	0.75	0.00	A	6.00	0.014	NA

Este resultado indica que el SNP1 en el cromosoma 1 está asociado de manera significativa con el fenotipo de interés. La proporción del alelo mayoritario en casos es alta en comparación con los controles, y el valor **p es menor que 0.05** (umbral típico de significancia). Sin embargo, no se proporciona el valor del odds ratio (OR), por lo que **no se puede determinar la fuerza y la dirección de la asociación**.

CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	SNP2	2	T	0.75	0.00	G	9.60	0.002	NA

Similar a SNP1, SNP2 en el cromosoma 1 muestra una fuerte asociación con el fenotipo de interés, con un valor p muy significativo. Sin embargo, nuevamente, **no se proporciona el odds ratio (OR)**.

CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	SNP3	3	G	0.75	0.25	C	4.00	0.046	9.00

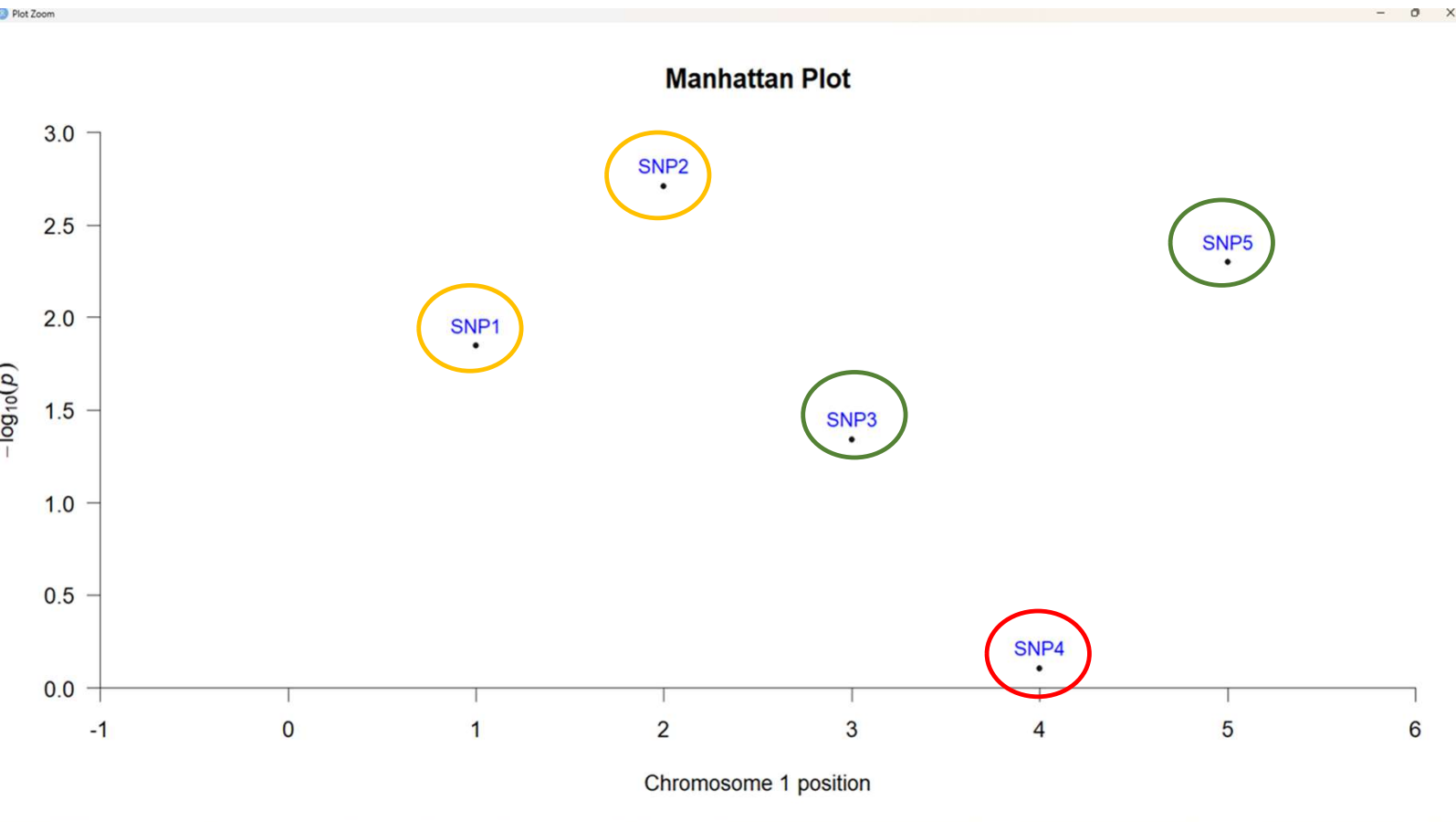
SNP3 muestra una asociación significativa con el fenotipo de interés. La proporción del alelo mayoritario en casos es alta, mientras que en controles es intermedia, lo que sugiere una asociación positiva. El odds ratio es de 9.00, lo que indica un mayor riesgo asociado con el alelo mayoritario G.

CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	SNP4	4	T	0.40	0.33	C	0.07	0.790	1.33

SNP4 muestra una asociación débil y no significativa con el fenotipo. La proporción de alelo mayoritario en casos y controles es similar, y el odds ratio está cerca de 1, lo que sugiere que no hay una asociación fuerte.

CHR	SNP	BP	A1	F_A	F_U	A2	CHISQ	P	OR
1	SNP5	5	A	0.00	0.75	G	7.88	0.005	0.00

SNP5 muestra una asociación significativa con el fenotipo, pero curiosamente, la proporción del alelo mayoritario en casos es 0.00, lo que sugiere que este alelo puede estar asociado con un menor riesgo de tener el fenotipo.



SNP3 y SNP5 muestran asociaciones significativas.

SNP1 y SNP2 también son prometedores, pero requieren una mayor investigación.

SNP4 no parece estar asociado de manera significativa con el fenotipo.



Análisis con Variant Calling Format



A diferencia de los archivos ped y map, donde están los fenotipos incluidos, este caso se deben unificar los archivos

C:\Users\villa\OneDrive\Docur x +

Sincronizando > Documents > 1 AXEL > CONGRESOS > 1 TRABAJOS CONGRESOS 2023 > 2 REGAD 2

Nuevo > Ordenar > Ver > ...

Nombre	Estado	Fecha de modificación	Tipo	Tamaño
.RData	✓	09/27/2023 1:53 p. m.	R Workspace	3 KB
.Rhistory	↻	09/28/2023 12:47 a. m.	Archivo RHISTORY	11 KB
fenotipos.txt	↻	09/28/2023 12:17 a. m.	Documento de te...	1 KB
fenotipos.xlsx	✓	09/27/2023 10:38 p. m.	Microsoft Excel W...	11 KB
genotipos.vcf	↻	09/27/2023 7:58 p. m.	Archivo VCF	1 KB
plink.exe	✓	02/21/2022 8:46 a. m.	Aplicación	28,640 KB
prettify.exe	✓	02/21/2022 8:46 a. m.	Aplicación	133 KB
UNIFICAR.R	↻	09/28/2023 12:47 a. m.	Archivo R	1 KB

Inicio

Galería

Axel - Personal

- 0344.tmp
- CmapToolsLogs
- Datos adjuntos de correo electrónico
- Documents
- InstallAnywhere
- Java TreeView
- Zero G Registry

Escritorio

Descargas

Imágenes

Música

Videos

ARCHIVO_VCF

borradores Tesis

TALLER_FENOTIPADO

TRABAJO

Archivos de Trabajo

genotipos.vcffenotipos.txt

ArchivoEditarVer

```
##fileformat=VCFv4.2
##INFO=<ID=DP,Number=1,Type=Integer,Description="Depth of coverage">
##INFO=<ID=AF,Number=A,Type=Float,Description="Allele Frequency">
#CHROM POS ID REF ALT QUAL FILTER INFO FORMAT Sample1 Sample2 Sample3 Sample4 Sample5 Sample6 Sample7 Sample8 Sample9 Sample10
1 1001 SNP1 A G 100 PASS DP=30;AF=0.6 GT 0/1 0/0 1/1 0/1 0/0 1/1 0/1 0/0 1/1 0/1
1 2001 SNP2 C T 150 PASS DP=40;AF=0.4 GT 0/1 1/1 0/1 0/0 1/1 0/1 1/1 0/1 0/0 0/1
1 3001 SNP3 G A 200 PASS DP=45;AF=0.7 GT 1/1 0/1 0/0 1/1 0/1 1/1 0/0 1/1 0/1 0/0
1 4001 SNP4 T C 80 PASS DP=25;AF=0.3 GT 0/1 0/0 0/1 0/1 0/0 0/1 0/0 0/1 0/0 0/1
1 5001 SNP5 G A 120 PASS DP=35;AF=0.5 GT 1/1 0/1 1/1 0/0 0/1 1/1 0/0 0/1 1/1 0/0
```

genotipos.vcf fenotipos.txt

ArchivoEditarVer

```
FID IID Phenotype1
1 1 180
1 2 195
1 3 175
2 4 185
2 5 200
2 6 170
1 7 150
1 8 200
1 9 201
1 10 202
```


Programacion básica para trabajar

```
# clear workplace  
rm(list =ls())
```

```
system("plink.exe --vcf genotipos.vcf --make-bed --allow-no-sex --out output_prefix")
```

```
system("plink.exe --bfile output_prefix --pheno fenotipos.txt --make-bed --allow-no-sex --out output_final")
```

```
system("plink.exe --bfile output_final --pheno fenotipos.txt --pheno-name Phenotype1 --allow-no-sex --assoc --out  
resultado_asociacion")
```

```
system("plink.exe --bfile output_final --pheno fenotipos.txt --pheno-name Phenotype1 --make-bed --allow-no-sex --  
out resultado_asociacion --linear")
```

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

UNIFICAR.R*

Source on Save Run

```
1 # clear workplace
2 rm(list = ls())
3
4 system("plink.exe --vcf genotipos.vcf --make-bed --allow-no-sex --out output_prefix")
5
6 system("plink.exe --bfile output_prefix --pheno fenotipos.txt --make-bed --allow-no-sex --
7
8 system("plink.exe --bfile output_final --pheno fenotipos.txt --pheno-name Phenotype1 --all
9
10 system("plink.exe --bfile output_final --pheno fenotipos.txt --pheno-name Phenotype1 --mak
11
12 |
```

12:1 (Top Level) R Script

Environment History Connections

79 MiB

Global Environment

Environment is empty

Files Plots Packages Help Viewer

Zoom Export

R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/2 REGAD 2023/TALLER_FENOTIPADO/ARCHIVO_VCF

14188 MB RAM detected; reserving 7094 MB for main workspace.
5 variants loaded from .bim file.
10 people (0 males, 0 females, 10 ambiguous) loaded from .fam.
Ambiguous sex IDs written to resultado_asociacion.nosex .
0 phenotype values present after --pheno.
Error: --linear without --all-pheno requires a scalar phenotype.
[1] 5
> # clear workplace
> rm(list = ls())
> |

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

UNIFICAR.R

Source on Save Run

```
1 # clear workplace
2 rm(list = ls())
3 system("plink.exe --vcf genotipos.vcf --make-bed --allow-no-sex --out output_prefix")
4
5 system("plink.exe --bfile output_prefix --pheno fenotipos.txt --make-bed --allow-no-sex --
6
7 system("plink.exe --bfile output_final --pheno fenotipos.txt --pheno-name Phenotype1 --all
8
9 system("plink.exe --bfile output_final --pheno fenotipos.txt --pheno-name Phenotype1 --mak
10
11
```

Environment History Connections Tut

145 MiB

Global Environment

Environment is empty

Files Plots Packages Help Viewer

AD 2023 > TALLER_FENOTIPADO > ARCHIVO_VCF

	Name	Size
<input type="checkbox"/>	plink.exe	28 MB
<input type="checkbox"/>	prettify.exe	132.8
<input type="checkbox"/>	ARCHIVO_FAM.txt	111 B
<input type="checkbox"/>	fenotipos.txt	111 B
<input type="checkbox"/>	genotipos.vcf	686 B
<input type="checkbox"/>	fenotipos.xlsx	10.5 K
<input type="checkbox"/>	output_prefix.log	1.1 KB
<input type="checkbox"/>	output_prefix.nosex	172 B
<input type="checkbox"/>	output_prefix.bed	18 B
<input type="checkbox"/>	output_prefix.fam	262 B
<input type="checkbox"/>	output_prefix.bim	95 B

Console Terminal Background Jobs

R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/1. REGAD 2023/TALLER_FENOTIPADO/ARCHIVO_VCF

using 1 thread (no multithreaded calculations invoked).

Before main variant filters, 10 founders and 0 nonfounders present.

calculating allele frequencies... done.

Total genotyping rate is exactly 1.

5 variants and 10 people pass filters and QC.

Note: No phenotypes present.

--make-bed to output_prefix.bed + output_prefix.bim + output_prefix.fam ...

done.

[1] 0

> |

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function

Addins

Project: (None)

UNIFICAR.R x output_prefix.fam x

1 sample1 sample1 0 0 0 -9
2 sample2 sample2 0 0 0 -9
3 sample3 sample3 0 0 0 -9
4 sample4 sample4 0 0 0 -9
5 sample5 sample5 0 0 0 -9
6 sample6 sample6 0 0 0 -9
7 sample7 sample7 0 0 0 -9
8 sample8 sample8 0 0 0 -9
9 sample9 sample9 0 0 0 -9
10 sample10 sample10 0 0 0 -9
11

Environment History Connections Tut

R Global Environment

Environment is empty

Files Plots Packages Help Viewer

AD 2023 > TALLER_FENOTIPADO > ARCHIVO_VCF

	Name	Size
	plink.exe	28 MB
	prettify.exe	132.8
	ARCHIVO_FAM.txt	111 B
	fenotipos.txt	111 B
	genotipos.vcf	686 B
	fenotipos.xlsx	10.5 K
	output_prefix.log	1.1 KB
	output_prefix.nosex	172 B
	output_prefix.bed	18 B
	output_prefix.fam	262 B
	output_prefix.bim	95 B

Console Terminal x Background Jobs x

R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/1. REGAD 2023/TALLER_FENOTIPADO/ARCHIVO_VCF

Using 1 thread (no multithreaded calculations invoked).

Before main variant filters, 10 founders and 0 nonfounders present.

Calculating allele frequencies... done.

Total genotyping rate is exactly 1.

5 variants and 10 people pass filters and QC.

Note: No phenotypes present.

--make-bed to output_prefix.bed + output_prefix.bim + output_prefix.fam ...

done.

[1] 0

>

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

UNIFICAR.R x output_final.fam x

```
1 1 1 0 0 0 180
2 1 2 0 0 0 195
3 1 3 0 0 0 175
4 2 4 0 0 0 185
5 2 5 0 0 0 200
6 2 6 0 0 0 170
7 1 7 0 0 0 150
8 1 8 0 0 0 200
9 1 9 0 0 0 201
10 1 10 0 0 0 202
11
```

1:1 Text file

Console Terminal x Background Jobs x

R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/1. REGAD 2023/TALLER_FENOTIPADO/ARCHIVO_VCF

```
10 phenotype values present after --pheno.
Using 1 thread (no multithreaded calculations invoked).
Before main variant filters, 10 founders and 0 nonfounders present.
calculating allele frequencies... done.
Total genotyping rate is exactly 1.
5 variants and 10 people pass filters and QC.
Phenotype data is quantitative.
--make-bed to output_final.bed + output_final.bim + output_final.fam ... done.
[1] 0
>
```

Project: (None)

Environment History Connections Tutor

144 MiB

Global Environment

Environment is empty

Files Plots Packages Help Viewer

AD 2023 > TALLER_FENOTIPADO > ARCHIVO_VCF

	Name	Size
	..	
	UNIFICAR.R	528 B
	.RData	2.5 KB
	.Rhistory	10.7 K
	plink.exe	28 MB
	prettify.exe	132.8 K
	ARCHIVO_FAM.txt	141 B
	fenotipos.txt	111 B
	genotipos.vcf	686 B
	fenotipos.xlsx	10.5 K
	outout_prefix.log	1.1 KB

RStudio

File Edit Code View Plots Session Build Debug Profile Tools Help

Go to file/function Addins

Project: (None)

Environment History Connections Tuts

145 MiB

Global Environment

Environment is empty

UNIFICAR.R

```
1 # clear workspace
```

```
no-sex --out output_prefix")
```

```
.txt --make-bed --allow-no-sex --out output_final")
```

```
txt --pheno-name Phenotype1 --allow-no-sex --assoc --out resultado_asociacion")
```

```
txt --pheno-name Phenotype1 --make-bed --allow-no-sex --linear --out resultado_asociacion ")
```

Console Terminal Background Jobs

R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/1. REGAD 2023/TALLER_FENOTIPADO/ARCHIVO_VC

Using 1 thread (no multithreaded calculations invoked).

Before main variant filters, 10 founders and 0 nonfounders present.

calculating allele frequencies... done.

Total genotyping rate is exactly 1.

5 variants and 10 people pass filters and QC.

Note: No phenotypes present.

```
--make-bed to output_prefix.bed + output_prefix.bim + output_prefix.fam ...
```

```
done.
```

```
[1] 0
```

```
>
```

Packages Help Viewer

TALLER_FENOTIPADO > ARCHIVO_VCF

	Size
pretify.exe	132.8
ARCHIVO_FAM.txt	111 B
fenotipos.txt	111 B
genotipos.vcf	686 B
fenotipos.xlsx	10.5 K
output_prefix.log	1.1 KB
output_prefix.nosex	172 B
output_prefix.bed	18 B
output_prefix.fam	262 B
output_prefix.bim	95 B

The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains a table with 10 columns: CHR, SNP, BP, NMISS, BETA, SE, R2, T, and P. The data is as follows:

	CHR	SNP	BP	NMISS	BETA	SE	R2	T	P
1	1	SNP1	1001	10	-8.167	6.867	0.1502	-1.189	0.2684
2	1	SNP2	2001	10	5.469	8.013	0.05503	0.6826	0.5142
3	1	SNP3	3001	10	-3.217	6.853	0.02682	-0.4695	0.6512
4	1	SNP4	4001	10	-1.167	11.77	0.001226	-0.09911	0.9235
5	1	SNP5	5001	10	-0.4638	6.945	0.0005572	-0.06678	0.9484
- Console:** Shows the execution of a script with the following output:


```

R 4.3.0 · C:/Users/villa/OneDrive/Documents/1 AXEL/CONGRESOS/1 TRABAJOS CONGRESOS 2023/1. REGAD 2023/TALLER_FENOTIPADO/ARCHIVO_VCF
Calculating allele frequencies... done.
Total genotyping rate is exactly 1.
5 variants and 10 people pass filters and QC.
Phenotype data is quantitative.
--make-bed to resultado_asociacion.bed + resultado_asociacion.bim +
resultado_asociacion.fam ... done.
writing linear model association results to resultado_asociacion.assoc.linear
... done.
[1] 0
>
      
```
- Environment:** Shows the Global Environment with the message "Environment is empty".
- Files:** Displays a file explorer view for the directory AD 2023 > TALLER_FENOTIPADO > ARCHIVO_VCF. The files listed are:
 - output_final.nosex (51 B)
 - output_final.bed (18 B)
 - output_final.fam (151 B)
 - output_final.bim (95 B)
 - resultado_asociacion.log (1.2 K)
 - resultado_asociacion.nosex (51 B)
 - resultado_asociacion.qassoc (522 B) - highlighted with a red arrow
 - resultado_asociacion.bed (18 B)
 - resultado_asociacion.fam (151 B)
 - resultado_asociacion.bim (95 B)
 - resultado_asociacion.assoc.linear (505 B)

- 1.**CHR** - Cromosoma
- 2.**SNP** - Identificación del SNP
- 3.**BP** - Posición del SNP en la base
- 4.**NMISS** - Número de datos no perdidos
- 5.**BETA** - Estimación del efecto
- 6.**SE** - Error estándar
- 7.**R2** - Coeficiente de determinación
- 8.**T** - Estadística T
- 9.**P** - Valor P

Análisis Descriptivo:

Resumen de la Tabla:

Revisar los valores mínimos, máximos y medios de cada columna numérica.

Contar el número de SNPs evaluados.

Calidad de Datos:

Evaluar NMISS para asegurarse de que tienes suficientes datos en cada SNP.

Efecto del SNP:

Los coeficientes BETA te dirán el tamaño del efecto que cada SNP tiene sobre el fenotipo. Un BETA positivo sugiere un efecto positivo en el fenotipo, mientras que un BETA negativo sugiere un efecto negativo.

Significancia Estadística:

Incluso con un valor P bajo, considera el contexto biológico y clínico de los SNPs y los genes relacionados.

A veces, los resultados estadísticamente significativos pueden no ser biológicamente significativos o clínicamente relevantes.

$$\text{Porcentaje de Datos Presentes} = (\text{NMISS} / \text{Tamaño de Muestra Total}) * 100 \quad (95-99\%)$$

Ejemplo Práctico:


En la tabla, el SNP1 tiene un **BETA** de -8.167, lo que sugiere un efecto negativo sustancial en el fenotipo por cada alelo adicional del SNP1.

Sin embargo, el valor p asociado (0.2684) no es estadísticamente significativo (si consideras un umbral de 0.05), lo que implica que debes ser cauteloso al interpretar este resultado.

Reflexión Final:

El valor **BETA** debe ser interpretado siempre en conjunto con otros parámetros y en el contexto específico de tu estudio y fenotipo de interés.

Además, realizar validaciones y replicaciones del estudio puede ayudarte a confirmar y entender mejor la importancia y relevancia de tus hallazgos.



El único medio de alcanzar el éxito y lograr todos nuestros objetivos, es seguir el camino correcto, el camino del honor, la amistad y el trabajo en equipo"

Precepto Samurai