Colocar el titulo del TFM aquí

Esta tesis se escribio usando los paquetes de R (R) Markdown, LaTeX , `bookdown` y `amsterdown`.

Colocar el titulo del TFM aquí

Tesis académica para obtener
el grado de Máster en Mejora Genética y
Biotecnología de la Reproducción bajo la
dirección del prof. dr. Miguel Pérez Enciso
ante una comisión constituida por la Junta del Máster,
para ser defendido en publico el
Colocar aquí la fecha de la defensa, a las colocar la hora aquí

Jorge Leonardo López Martínez

Dirección:

Director: prof. dr. M. Pérez-Enciso Centre for Research in Agricultural Genomics

Índice general

Capítulo 1

Revisión de literatura

1.1.

Capítulo 2

Titulo

Resumen
Insert abstract.

Possibly insert citation here.

2.1. Introducción

La teoría de la genética en el estudio de caracteres cuantitativos se estableció hace más de un siglo cuando Ronald Fisher presentó un documento (Fisher 1918) donde dio a conocer el desarrollo de la teoría del modelo infinitesimal, permitiendo con ello unificar dos de las escuelas de pensamiento que para ese entonces estaban en constante debate: la escuela de pensamiento Mendeliano, cuyo objetivo consistía en localizar y caracterizar factores de herencia, y la escuela de pensamiento biométrico, cuyo origen se remonta a Galton quien buscaba aplicar modelos biométricos con el fin de estudiar las relaciones entre parientes (Nelson, Pettersson, y Carlborg 2012; Blasco y Toro 2014).

La teoría del modelo infinitesimal desarrollado por Fisher establece que la varianza genética de un carácter esta determinado por un gran número de factores Mendelianos, cada uno de los cuales tiene una pequeña contribución aditiva al fenotipo de dicho carácter (Nelson, Pettersson, y Carlborg 2012; Turelli 2017). Naturalmente, los modelos usados en estudios de mejoramiento genético han sido concebidos en base a esta teoría (Villemereuil et al. 2016; Pérez-Enciso 2017), siendo ejemplo de ello el mejor predictor lineal insesgado (BLUP) y el mejor predictor lineal insesgado genómico (GBLUP).

En las ciencias animales, el valor de cría estimado (EBV) se suele predecir en función de un conjunto de modelos que relacionan el fenotipo de una población con la información del pedigrí, mediante el uso del BLUP. No obstante, este método no es factible para poblaciones sin información de pedigrí o con una estructura poblacional compleja, como suele ser el caso de las plantas (Nakaya y Isobe 2012; Tong y Nikoloski 2021). Para el año 2001, Meuwissen, Hayes y Goddard propusieron un método innovador para predecir los valores de cría basado en marcadores de ADN (GEBV), denominándose tiempo después como selección genómica (Nakaya y Isobe 2012; Blasco y Toro 2014), el cual permitió también superar las limitaciones que suponia el uso del BLUP para predecir los valores de cría en plantas.

Hoy en día, la selección genómica se considera como un método potencial para el mejoramiento genético en plantas (Nakaya y Isobe 2012), ya que sus ciclos reproductivos suelen ser prolongados, por lo cual con el uso de la selección genómica es posible acelerar dichos ciclos reproductivos con el beneficio adicional de mejorar la tasa de ganancia genética anual por unidad de tiempo y costo (Desta y Ortiz 2014; Jurcic et al. 2021). Además, los datos sobre marcadores de ADN en todo el genoma están cada vez más disponibles para cultivos de relevancia agronómica (Tong y Nikoloski 2021).

El GBLUP es uno de los métodos más comunes de selección genómica (Jurcic et al. 2021). De hecho, es el método más popular debido a su simplicidad al sustituir la matriz de relación de parentesco basado en pedigríes (Wright 1922) por una matriz de relación basada en marcadores de ADN (Hayes, Visscher, y Goddard 2009). Así mismo, el GBLUP predice con may-

or precisión los GEBV en comparación a los EBV del BLUP, debido a que con el primero se estima mejor las relaciones entre individuos (Misztal, Aggrrey, y Muir 2012), por lo cual la matriz de las relaciones genómicas suele verse como un estimador mejorado de las relaciones basadas en marcadores en lugar de pedigríes (Legarra et al. 2014).

En términos generales, la selección genómica es un proceso de tres pasos en el que los individuos, sobre la base de su información fenotípica y de pedigrí, son evaluados inicialmente mediante una evaluación genética tradicional por medio del BLUP, y posteriormente a partir de los fenotipos corregidos o pseudo-fenotipos resultantes de esta evaluación genética inicial, es llevado a cabo un análisis genómico de los individuos genotipados mediante el GBLUP. Por último y en base a la información generada, se calculan los GEBV por medio de un índice de selección (Legarra, Aguilar, y Misztal 2009; Misztal, Legarra, y Aguilar 2009; Misztal, Aggrrey, y Muir 2012; Legarra et al. 2014; Misztal, Lourenco, y Legarra 2020).

Como no todos los individuos pueden genotiparse, la selección genómica se lleva a cabo a partir del proceso anterior de tres pasos (Legarra, Aguilar, y Misztal 2009). Sin embargo, este proceso es tendente a cometer errores (Misztal, Aggrrey, y Muir 2012), además de presentar inconvenientes como son la perdida de información y la difícultad de generalizarse a caracteres múltiples y maternos (Legarra, Aguilar, y Misztal 2009; Legarra et al. 2014). Conscientes de esto, Legarra, Aguilar, y Misztal (2009) simplificaron el proceso de varios pasos al desarrollar un método de selección genómica, en el que los fenotipos de los individuos genotipados y no genotipados se analizan conjuntamente para predecir sus valores de cría (Imai et al. 2019; Jurcic et al. 2021), método el cual se denomino como mejor predictor lineal insesgado genómico de un solo paso (ssGBLUP).

En el ssGBLUP se dispone de una matriz de parentesco genómica global de individuos genotipados y no genotipados, denominada como matriz de relación combinada o matriz H. Esta matriz se obtiene combinando información de la relación genómica entre individuos genotipados, e información de pedigrí entre individuos genotipados y no genotipados (Imai et al. 2019). Con ello, el proceso anterior de tres pasos tiende a simplificarse al incorporar la información genómica desde el primer paso (Legarra et al. 2014; Misztal, Legarra, y Aguilar 2009), sin la necesidad del calculo posterior de fenotipos corregidos y la construcción del índice de selección mencionado previamente (Misztal, Lourenco, y Legarra 2020).

Al ser una forma de BLUP en el que la matriz de relación de parentesco es sustituida por la matriz de relación combinada (Legarra, Aguilar, y Misztal 2009; Legarra et al. 2014; Blasco 2021), el ssGBLUP se puede adecuar con facilidad a caracteres múltiples y maternos (Blasco 2021), además se adapta también a las herramientas informaticas ya desarrolladas en base al BLUP (Lourenco et al. 2020). Este hecho hace del ssGBLUP un método de uso rutinario para la evaluación genómica en animales, donde ha demostrado que

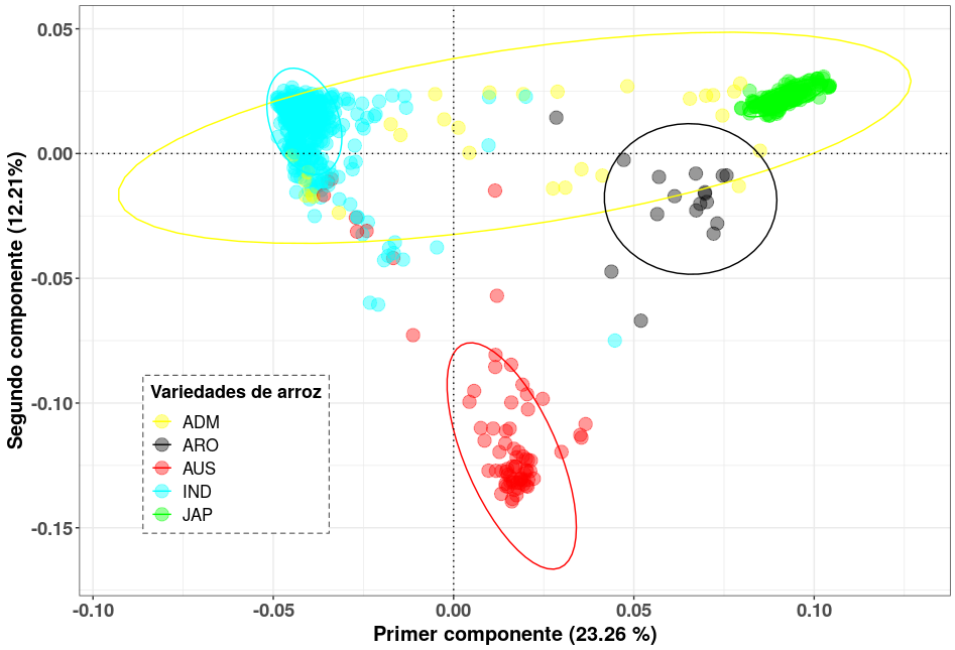
produce una predicción más precisa de los valores de cría en comparación a los métodos BLUP y GBLUP antes mencionados (Misztal, Aggrrey, y Muir 2012; Pérez-Rodríguez et al. 2017; Misztal, Lourenco, y Legarra 2020). No obstante, el uso del ssGBLUP para la selección genómica en plantas es más reciente y escaso (Pérez-Rodríguez et al. 2017; Jurcic et al. 2021). En consecuencia, el **objetivo**

2.2. Métodos

2.2.1. Recurso vegetal y datos fenotípicos

Los conjuntos de datos se obtuvieron del Rice SNP-Seek Database[1], el cual es un cibersitio con información sobre datos de genotipado de SNPs y de fenotipos de distintas variedades de arroz (*Oryza sativa L.*). Posteriormente, dichos conjuntos de datos fueron usados por Vourlaki et al. (s. f.), quienes sometieron los datos de genotipado de SNPs a procedimientos de control de calidad, en los que fueron eliminados loci de SNPs con una frecuencia del alelo menor de menos de 0.01 y con una tasa de ausencia mayor a 0.01.

Mediante un análisis de componentes principales realizado sobre los datos de genotipado de SNPs (Figura 2.1) se observaron diferentes grupos varietales de arroz, de los cuales la variedad indica fue seleccionada para llevar a cabo este estudio una vez la misma era el grupo varietal con mayor número de individuos genotipados (451 individuos de un total de 738).



[1] https://snp-seek.irri.org/index.zul;jsessionid=DD991975FDC4F320BE3C33ED056D0363

Figura 2.1: Análisis de componentes principales en datos de arroz. Los puntos y las circuferencias de color representan distintos grupos varietales: tipo intermedio o mezclado (ADM), aromático (ARO), aus (AUS), indica (IND) y japónica (JAP).

En relación a los datos de fenotipo, el conjunto de datos proporciono información sobre distintos caracteres fenotípicos de relevancia agronómica como son la trillabilidad de la panícula, el peso del grano, la fuerza del culmo, entre otros (Figura 2.2), siendo seleccionada para este estudio el carácter tiempo de floración ya que en este se obervo suficiente variación fenotípica.

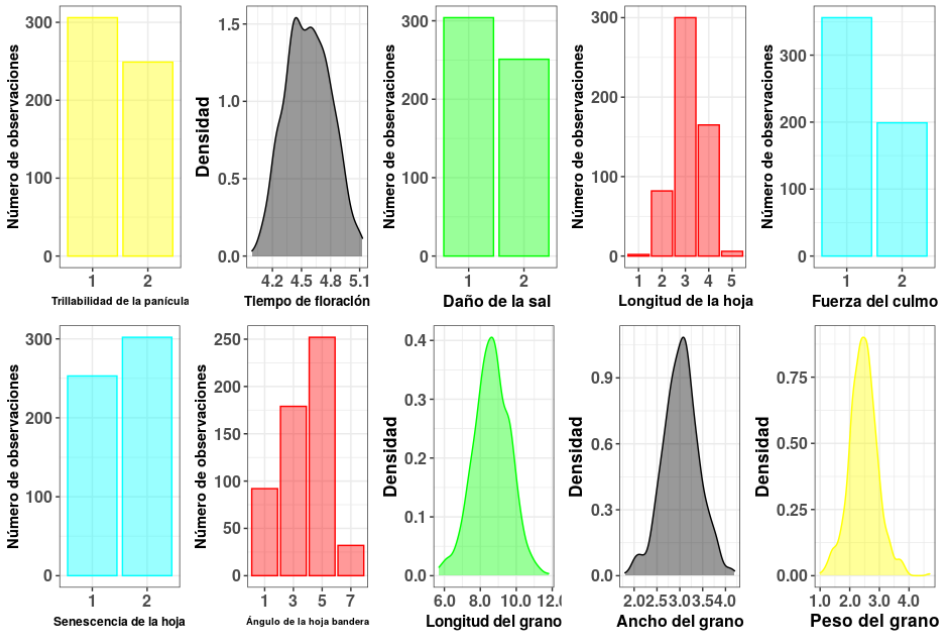


Figura 2.2: Distribución de cada uno de los caracteres del conjunto de datos fenotípicos de arroz.

En lo que respecta a la información de pedigrí, esta no estaba disponible. Por ello, se utilizó la metodología implementada en el software MOLCOANC (Fernández y Toro 2006) con el fin de contar con esta información. Este software . (Figura 2.3).

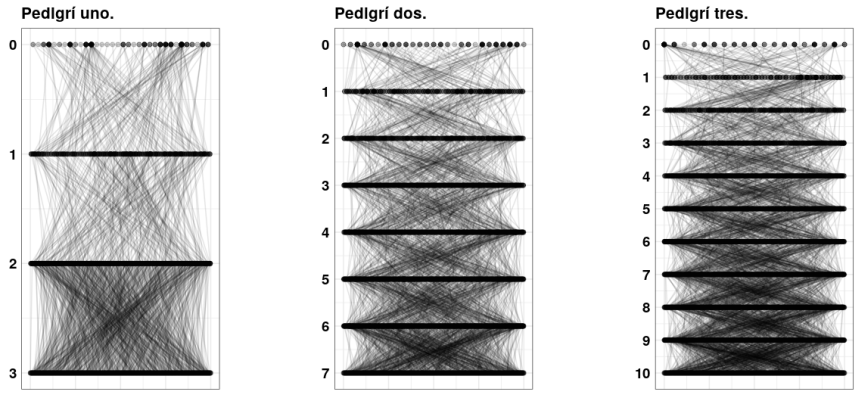


Figura 2.3: .

2.2.2. Modelo para la predicción genómica y habilidad predictiva

Para llevar a cabo la predicción genómica mediante el mejor predictor lineal insesgado genómico de un solo paso (ssGBLUP), se eliminaron los loci de SNPs con una frecuencia del alelo menor de menos de 0.05. La predicción genómica se realizó mediante el siguiente modelo con los datos descritos anteriormente:

$$y = Za + e, \tag{2.1}$$

donde y representa el valor del fenotipo a predecir (tiempo de floración) y Z es la matriz de incidencia que relaciona a con y. El vector a representa los valores genotípicos como se describen en el siguiente parrafo, y e es el vector de residuos con una distribución que se asume normal con media igual a 0 y matriz de covarianza $I\sigma_e^2$.

En la ecuación (1), a

Para identificar el efecto sobre la predictibilidad del tamaño de la muestra de entrenamiento, el número de datos de genotipado de SNPs y el número de individuos genotipados, se usaron diferentes subconjuntos de datos (Figura 2.4) con la siguientes características:

1. Diferente información de pedigrí:
2. Diferentes densidades de SNPs:
3. Distinta cantidad de individuos genotipados:

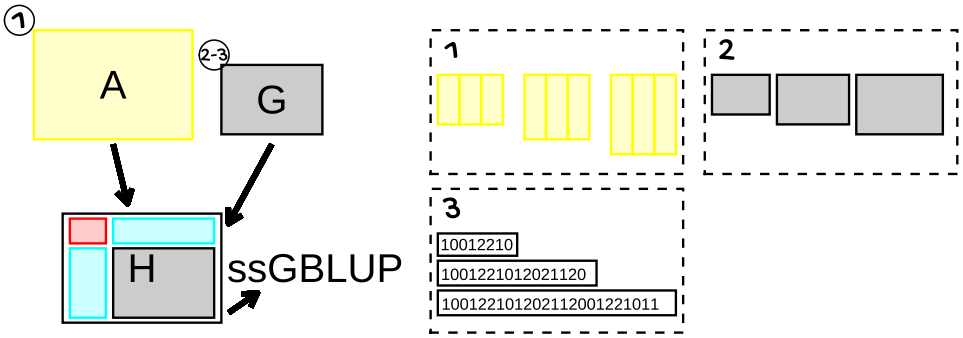
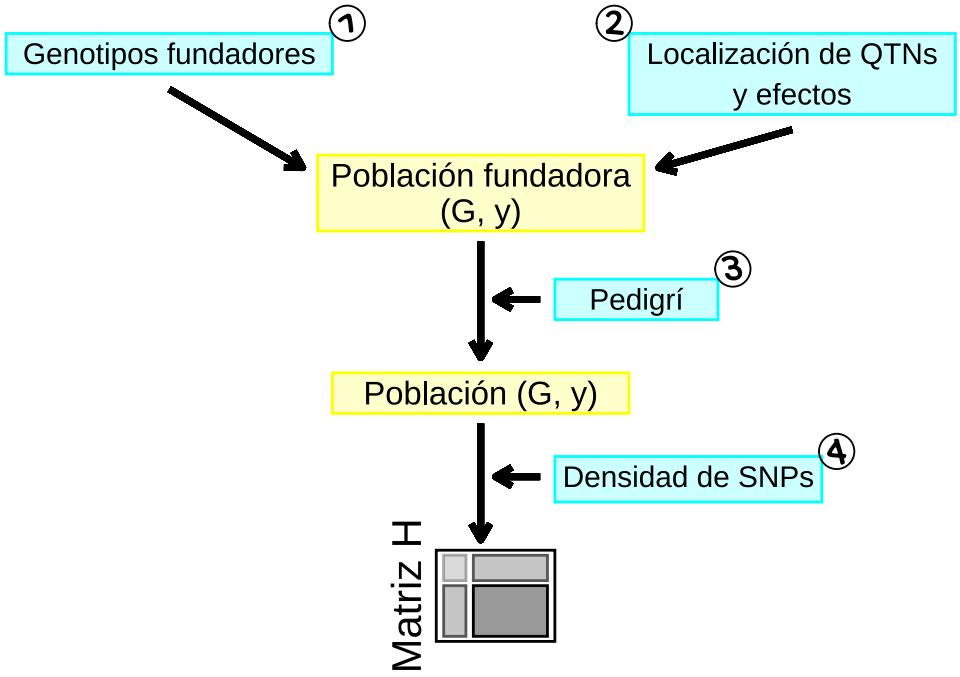


Figura 2.4: Esquema del calculo de la matriz H a partir de las matrices A y G, con base en diferentes subconjuntos de datos. El recuadro 1 representa los tres pedigríes con diferentes número de individuos y que posteriormente se usaron para el calculo de la matriz A. El recuadro 2 representa matrices G con distinta dimensión dado el número de individuos genotipados. El recuadro 3 representa diferentes densidades de SNPs.

Se uso el coeficiente de correlación entre los valores fenotípicos observados y predichos como medida de la predictibilidad. De acuerdo a Xua, Zhub, y Zhang (2014), la predictibilidad debe obtenerse usando una muestra de validación independiente o mediante validación cruzada donde los individuos predichos no deben contribuir a la estimación de parámetros. En este sentido, el valor fenotípico observado de 48 del total de 451 individuos de la variedad indica (que corresponde a los individuos clasificados como variedades mejoradas) se considero como faltante.

2.2.3. Simulación



Hola

2.3. Resultados

2.3.1. Fenotipo y heredabilidad

Tabla 2.1: Estimaciones de heredabilidad para el caracter tiempo de floración estimado por BLUP basado en el pedigrí.

	reml		
Parámetros	Pedigrí 1	Pedigrí 2	Pedigrí 3
Varianza aditiva	0.49	0.46	0.57
Varianza ambiental	0.11	0.17	0.13
Heredabilidad	0.82	0.73	0.81

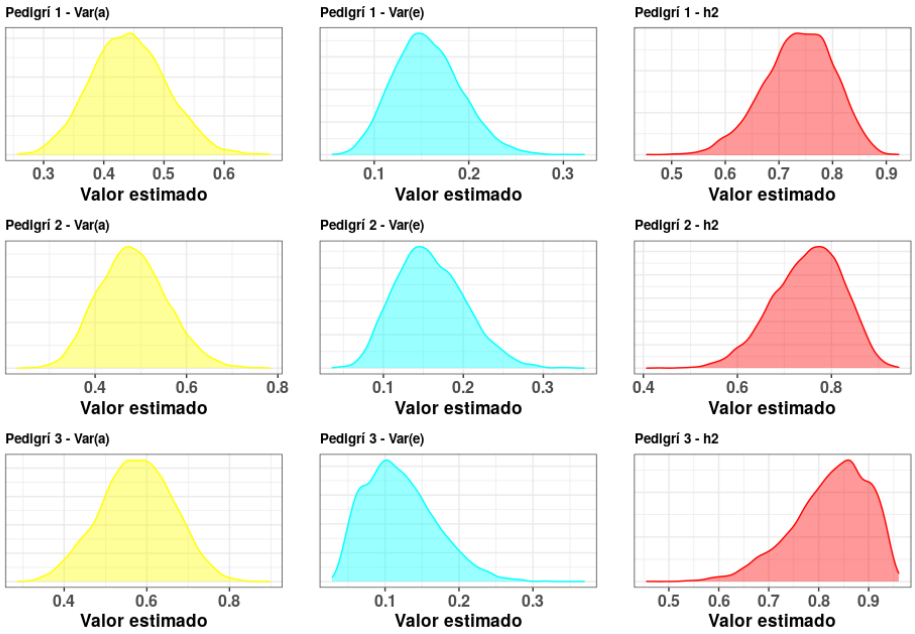
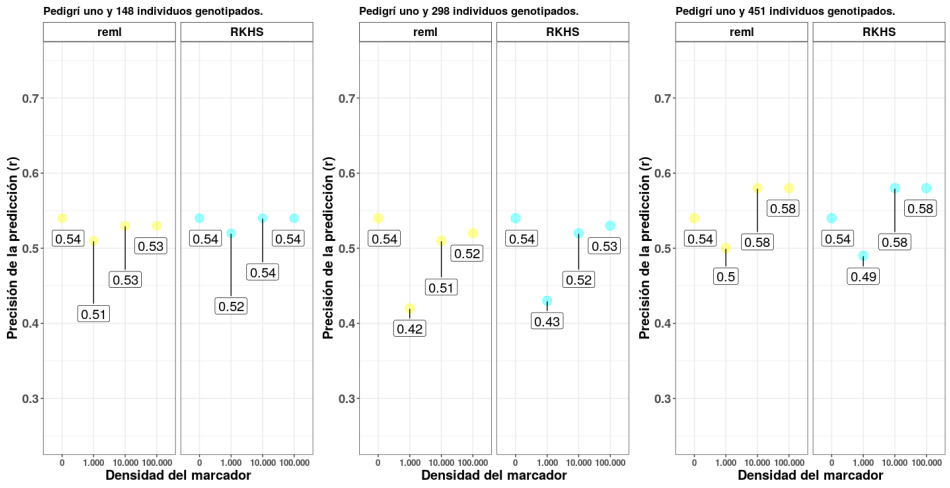


Figura 2.5: .

Los resultados del análisis de máxima verosimilitud restringida (REML) y... (RKHS) bajo el modelo aditivo se observan en la Figura 2.5.



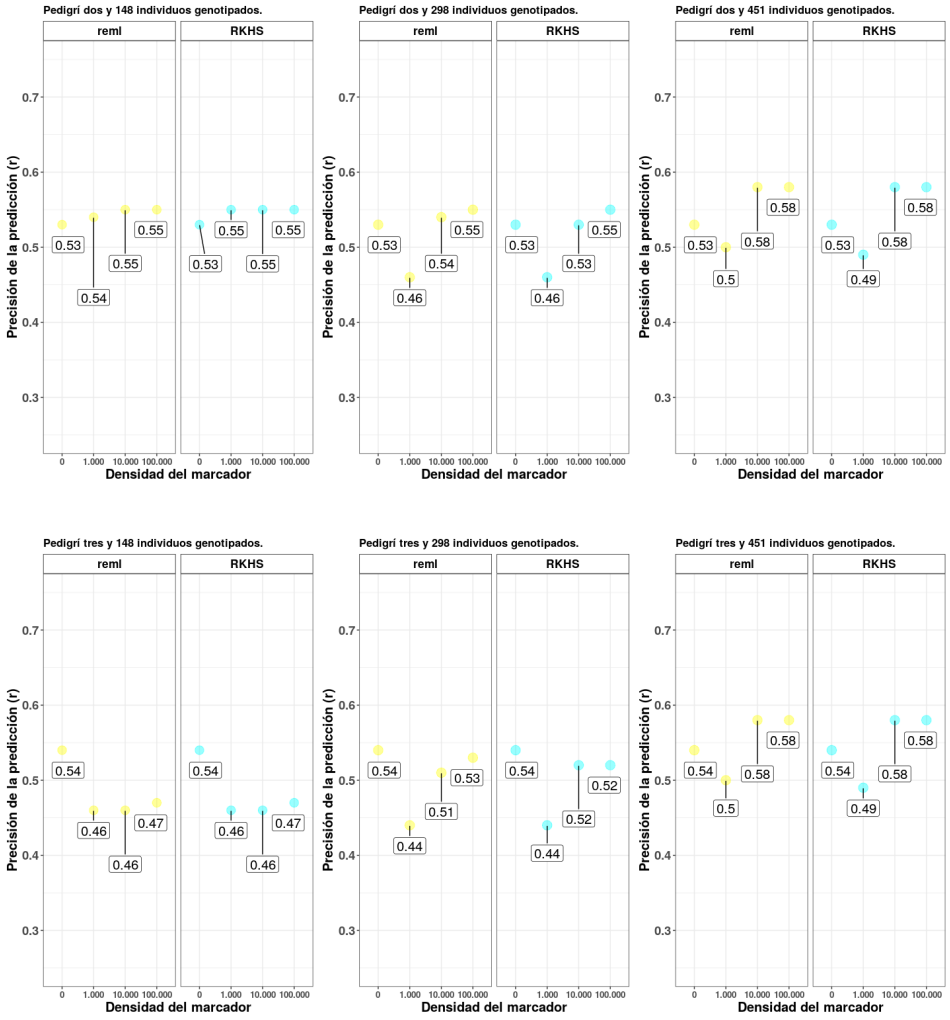


Figura 2.6: .

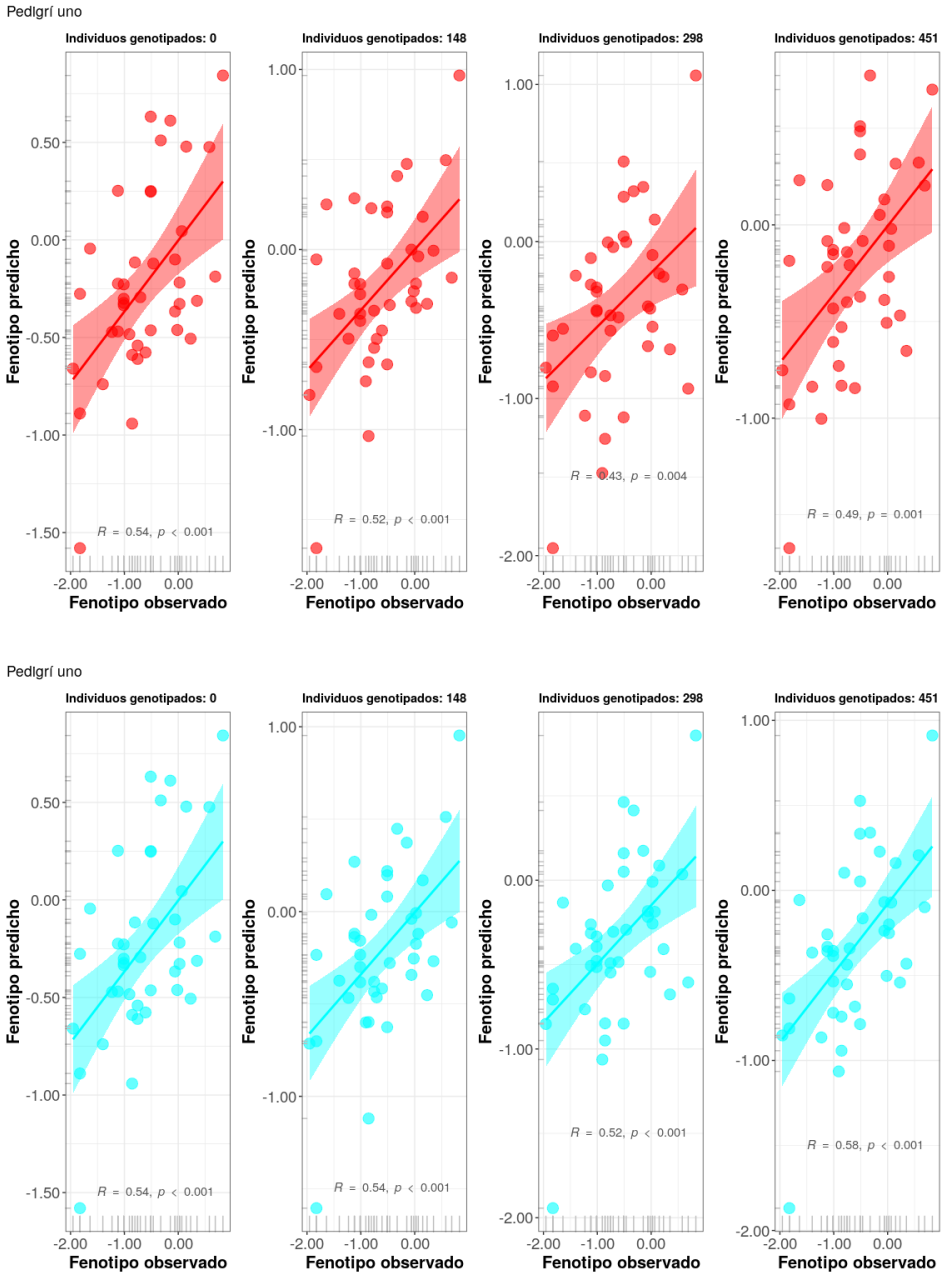
2.4. Discusión

Apéndice A

Anexo del capitulo 2

```
fn.mH <- function(ped, mG) { # Esta función recibe como argu-
                             # mentos los datos con estructura
                             # (id | sire | dam | Gen (TRUE/FALSE))
                             # y la matriz de relaciones genómicas.

  # 1. Se calcula la matriz de relaciones aditivas con base en
  # el pedigrí (A)

  ped_edit <- editPed( # Esta función ordena el pedigrí.
    sire = ped$sire,
    dam = ped$dam,
    label = ped$id
    )
  pedi <- pedigree( # Aquí se usa la salida anterior (ya orde-
                    # nado) y se crea un objeto de clase pedigree.
    sire = ped_edit$sire,
    dam = ped_edit$dam,
    label = ped_edit$label
    )
  Matrix_A <- getA(ped = pedi) # Esto dara la matriz de relaciones
                               # aditivas A.

  # 2. De lo anterior (Matriz_A) se extraen las partes correspon-
  # dientes a individuos no genotipados (1) y genotipados (2)

  # Individuos no genotipados:
  A_11 <- Matrix_A[ped$Genotiped != 1, ped$Genotiped != 1]
  # Individuos genotipados:
  A_22 <- Matrix_A[ped$Genotiped == 1, ped$Genotiped == 1]
```

```
# Individuos no genotipados (en filas) y genotipados (en
# columnas):
A_12 <- Matrix_A[ped$Genotiped != 1, ped$Genotiped == 1]
# Transpuesta de la anterior (individuos no genotipados en
# columnas y genotipados en filas):
A_21 <- t(A_12)

# 3. Se coloca el nombre de las filas y y de las columnas
# de la matriz G según los individuos genotipados

rownames(mG) <- ped$id[ped$Genotiped == 1]
colnames(mG) <- ped$id[ped$Genotiped == 1]

# 4. Teniendo todos los componentes de la matriz H, se pro-
# cede a su construcción y a calcular su inversa

H_11 <- A_11 -
  (A_12 %*% solve(A_22) %*% A_21) +
  (A_12 %*% solve(A_22) %*% mG %*% solve(A_22) %*% A_21)
H_12 <- A_12 %*% solve(A_22) %*% mG
H_21 <- t(H_12)
H_22 <- mG

H_11_H_12 <- cbind(H_11, H_12)
H_21_H_22 <- cbind(H_21, H_22)
mH <- rbind(H_11_H_12, H_21_H_22)

mH <- mH[order(as.numeric(rownames(mH))),
         order(as.numeric(colnames(mH)))]
mH <- Matrix(mH)

# 5. Finalmente se indica retornar la inversa de la ma-
# triz H (mH_1)

return(mH)
}
```


Pedlgrí uno
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.52, p < 0.001
R = 0.43, p = 0.004
R = 0.49, p = 0.001
Pedlgrí uno
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.54, p < 0.001
R = 0.52, p < 0.001
R = 0.58, p < 0.001

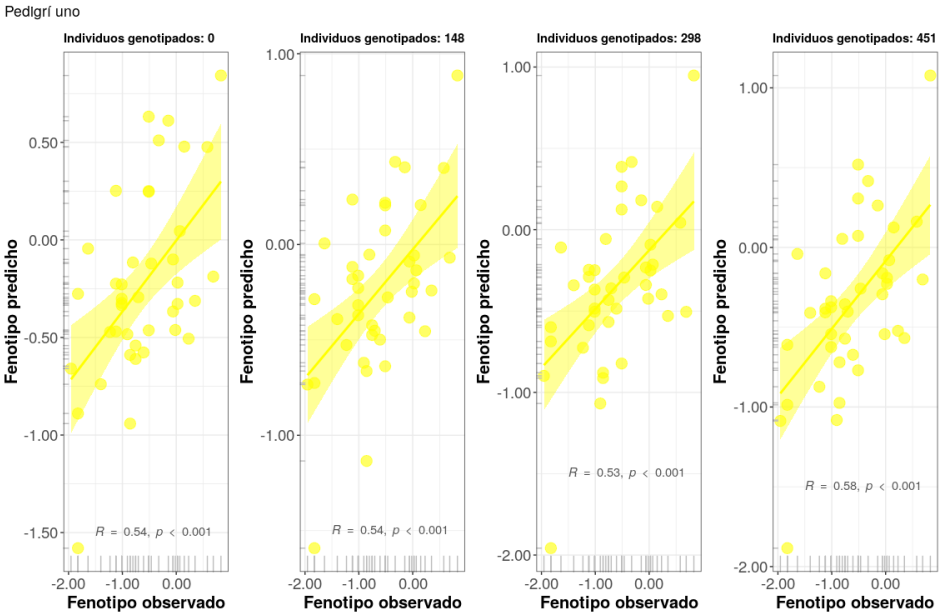
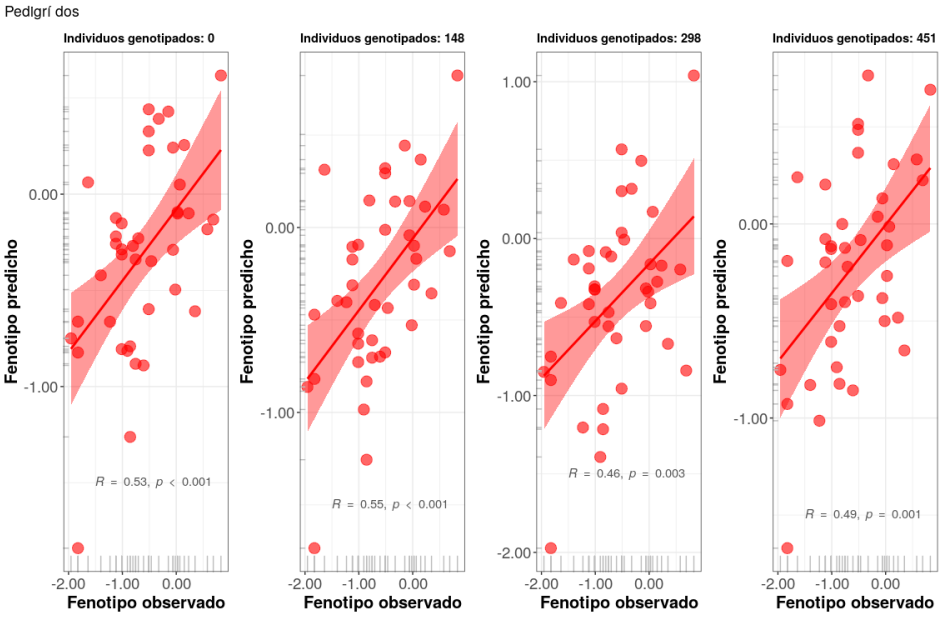

PedIgrí uno
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.54, p < 0.001
R = 0.53, p < 0.001
R = 0.58, p < 0.001

Figura A.1:


PedIgrí dos
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.53, p < 0.001
R = 0.55, p < 0.001
R = 0.46, p = 0.003
R = 0.49, p = 0.001

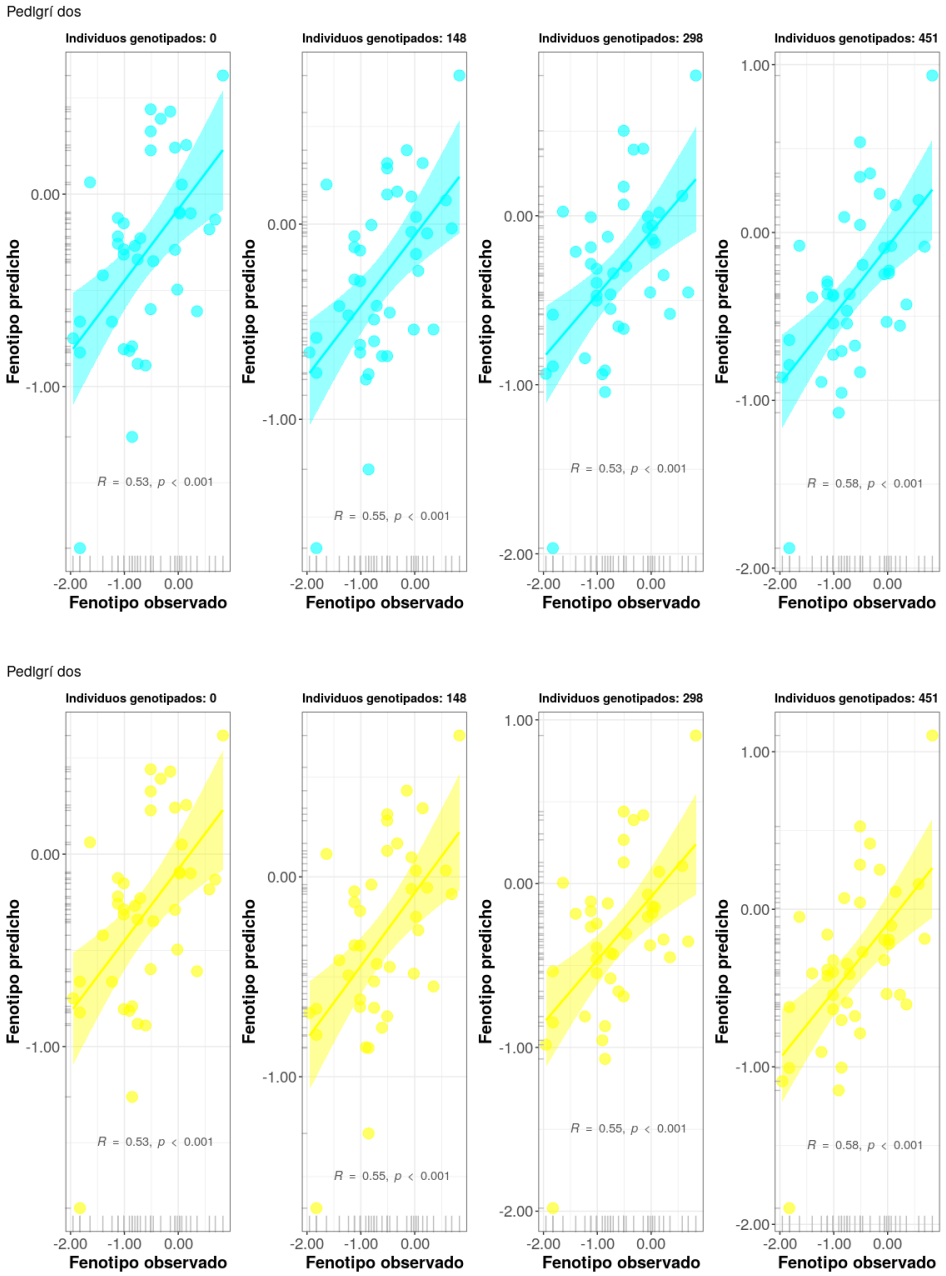
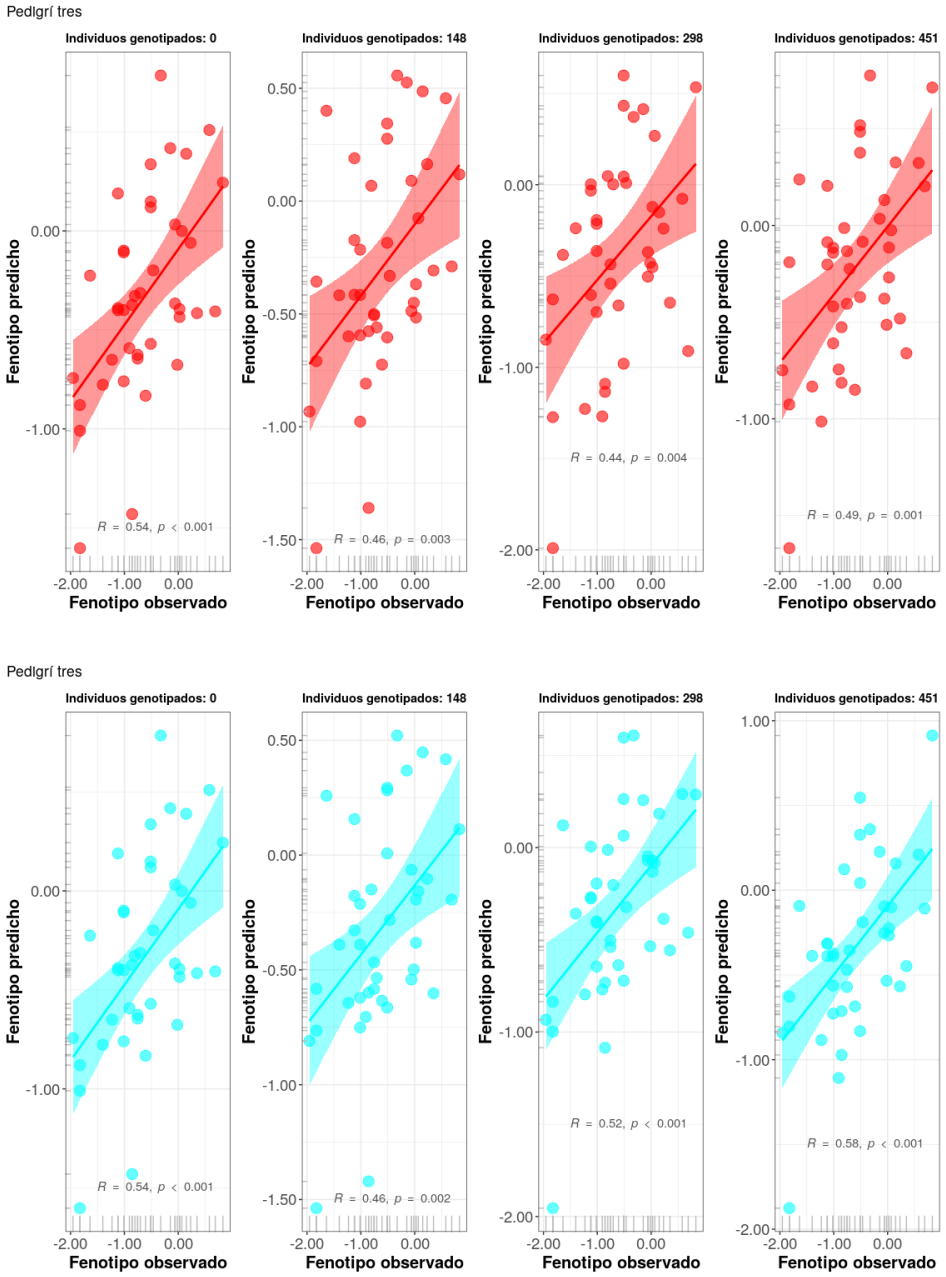

Pedlgrí dos
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.53, p < 0.001
R = 0.55, p < 0.001
R = 0.53, p < 0.001
R = 0.58, p < 0.001
Pedlgrí dos
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.53, p < 0.001
R = 0.55, p < 0.001
R = 0.55, p < 0.001
R = 0.58, p < 0.001

Figura A.2:


Pedlgrí tres
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.46, p = 0.003
R = 0.44, p = 0.004
R = 0.49, p = 0.001
Pedlgrí tres
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.46, p = 0.002
R = 0.52, p < 0.001
R = 0.58, p < 0.001

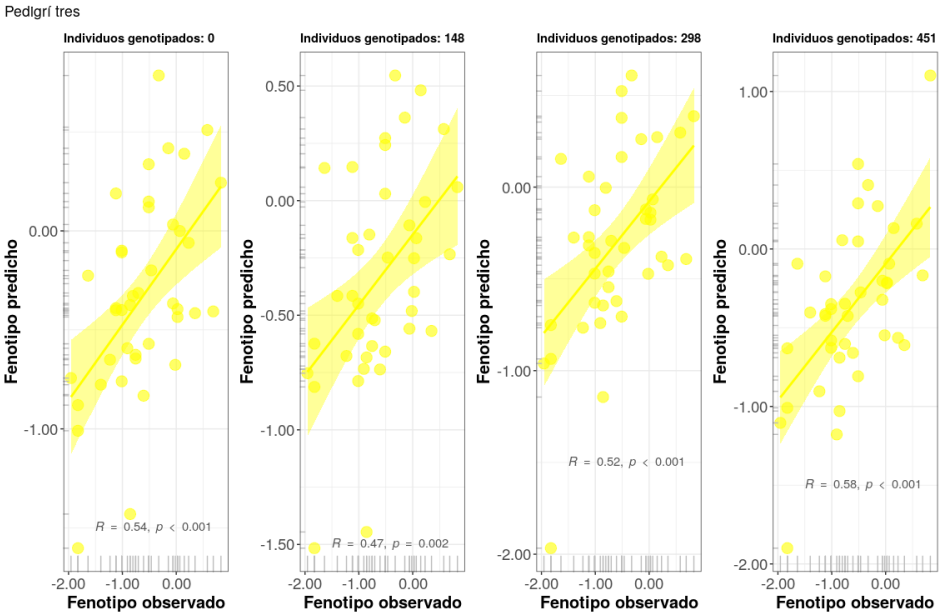
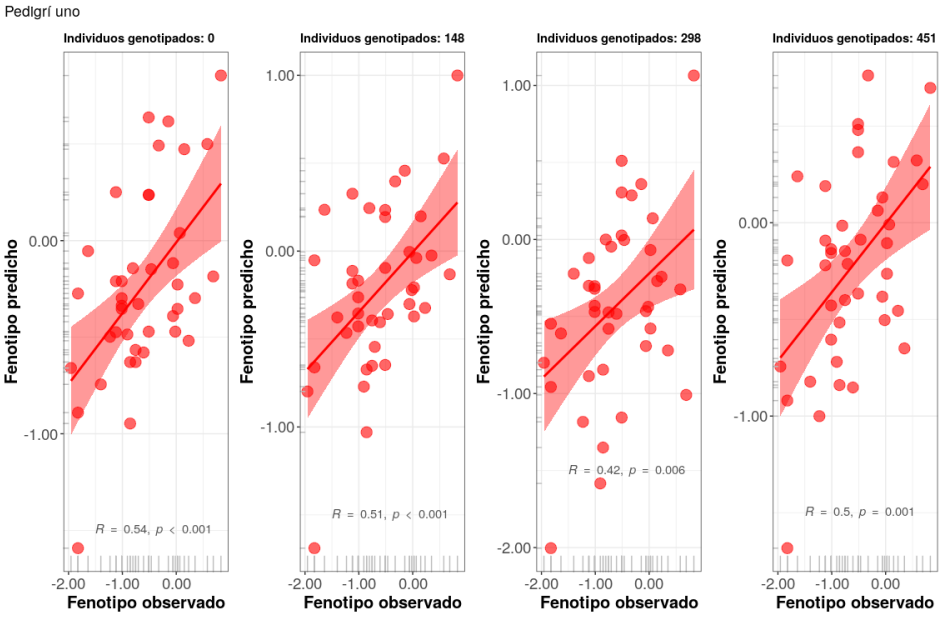

Pedigrí tres
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.47, p = 0.002
R = 0.52, p < 0.001
R = 0.58, p < 0.001

Figura A.3:


Pedigrí uno
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.51, p < 0.001
R = 0.42, p = 0.006
R = 0.5, p = 0.001

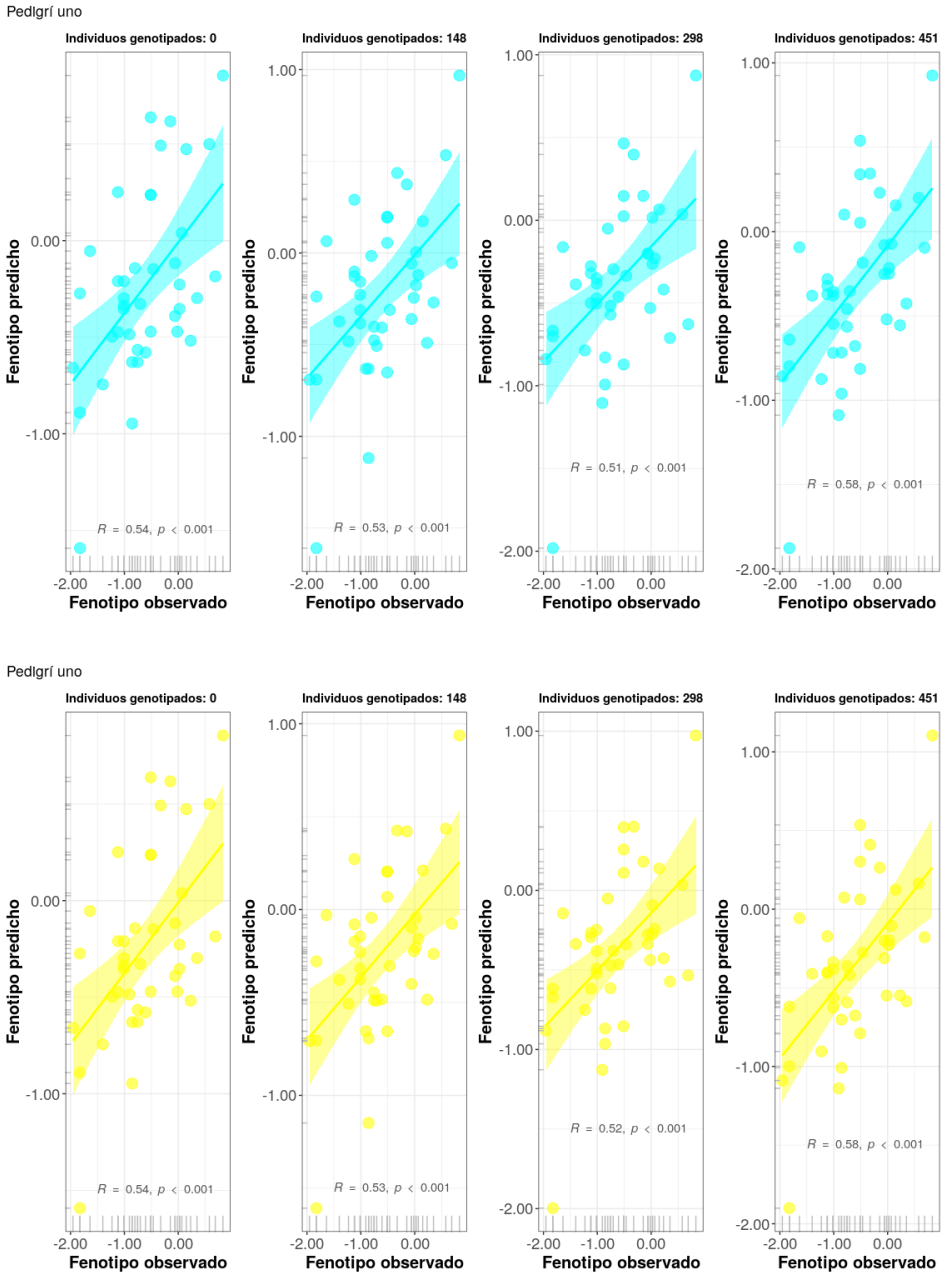
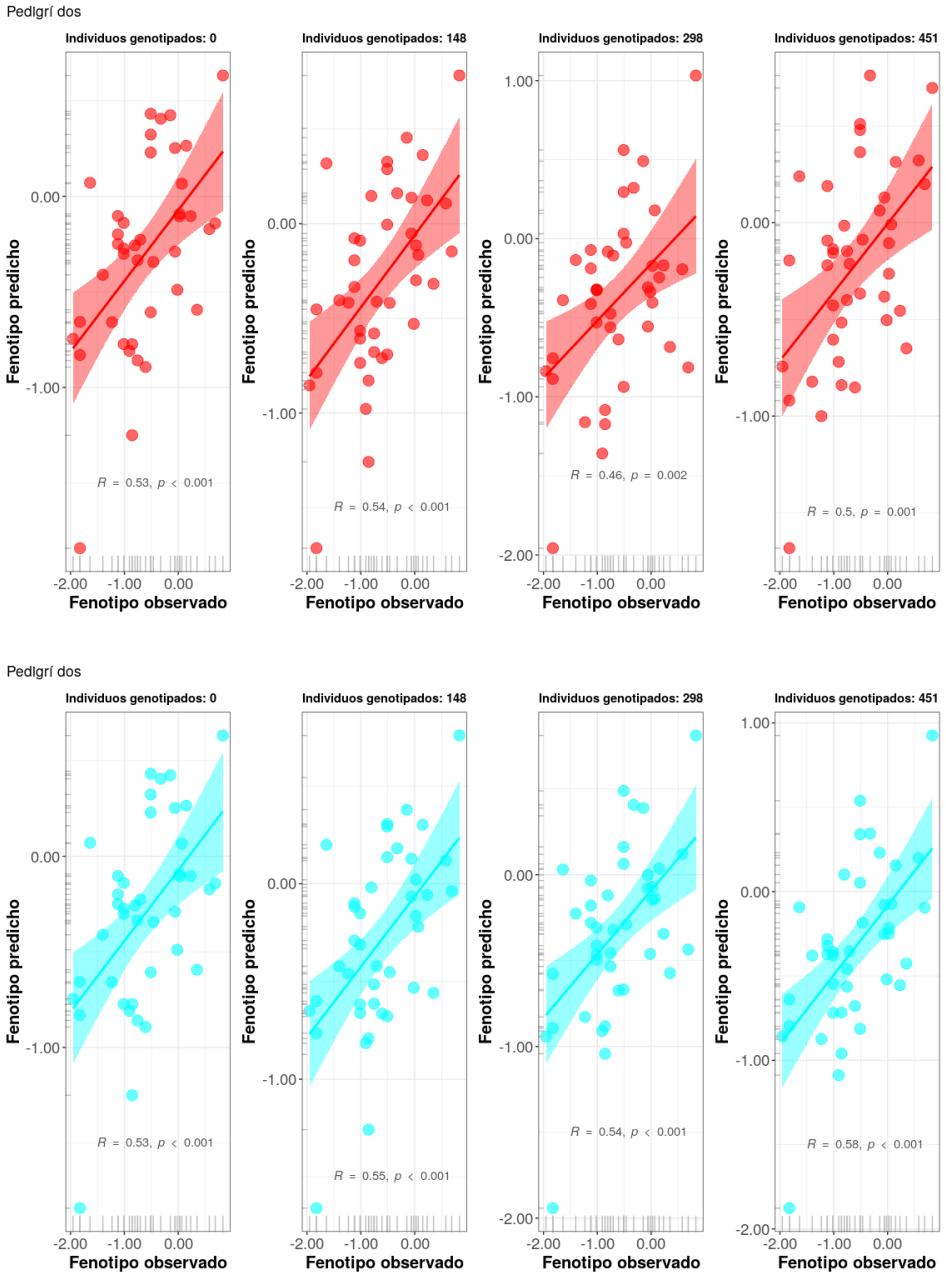

PedIgrí uno
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.53, p < 0.001
R = 0.51, p < 0.001
R = 0.58, p < 0.001
PedIgrí uno
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.53, p < 0.001
R = 0.52, p < 0.001
R = 0.58, p < 0.001

Figura A.4:


Pedlgrí dos
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.53, p < 0.001
R = 0.54, p < 0.001
R = 0.46, p = 0.002
R = 0.5, p = 0.001
Pedlgrí dos
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.53, p < 0.001
R = 0.55, p < 0.001
R = 0.54, p < 0.001
R = 0.58, p < 0.001

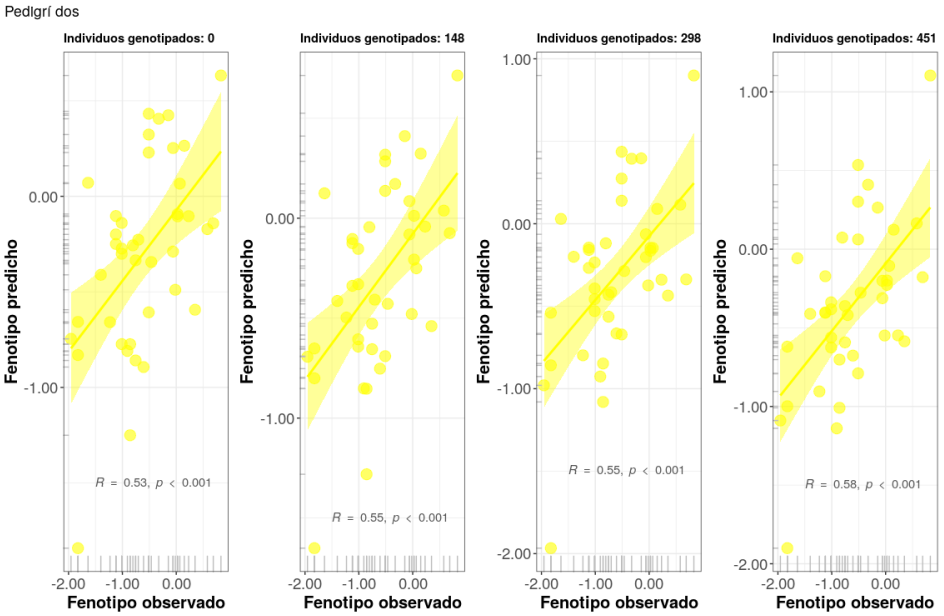
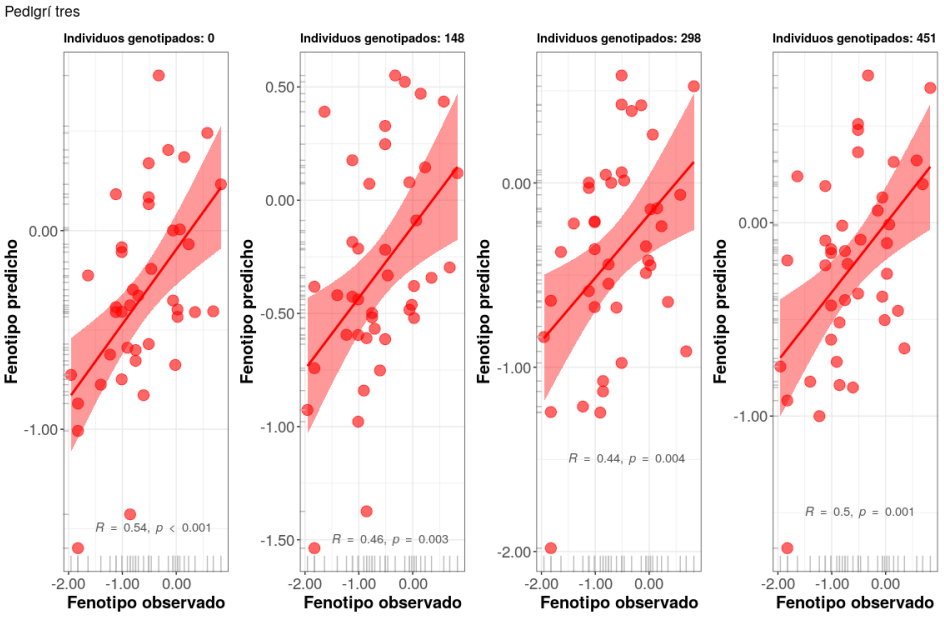

Pedigrí dos
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.53, p < 0.001
R = 0.55, p < 0.001
R = 0.55, p < 0.001
R = 0.58, p < 0.001

Figura A.5:


Pedigrí tres
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.46, p = 0.003
R = 0.44, p = 0.004
R = 0.5, p = 0.001

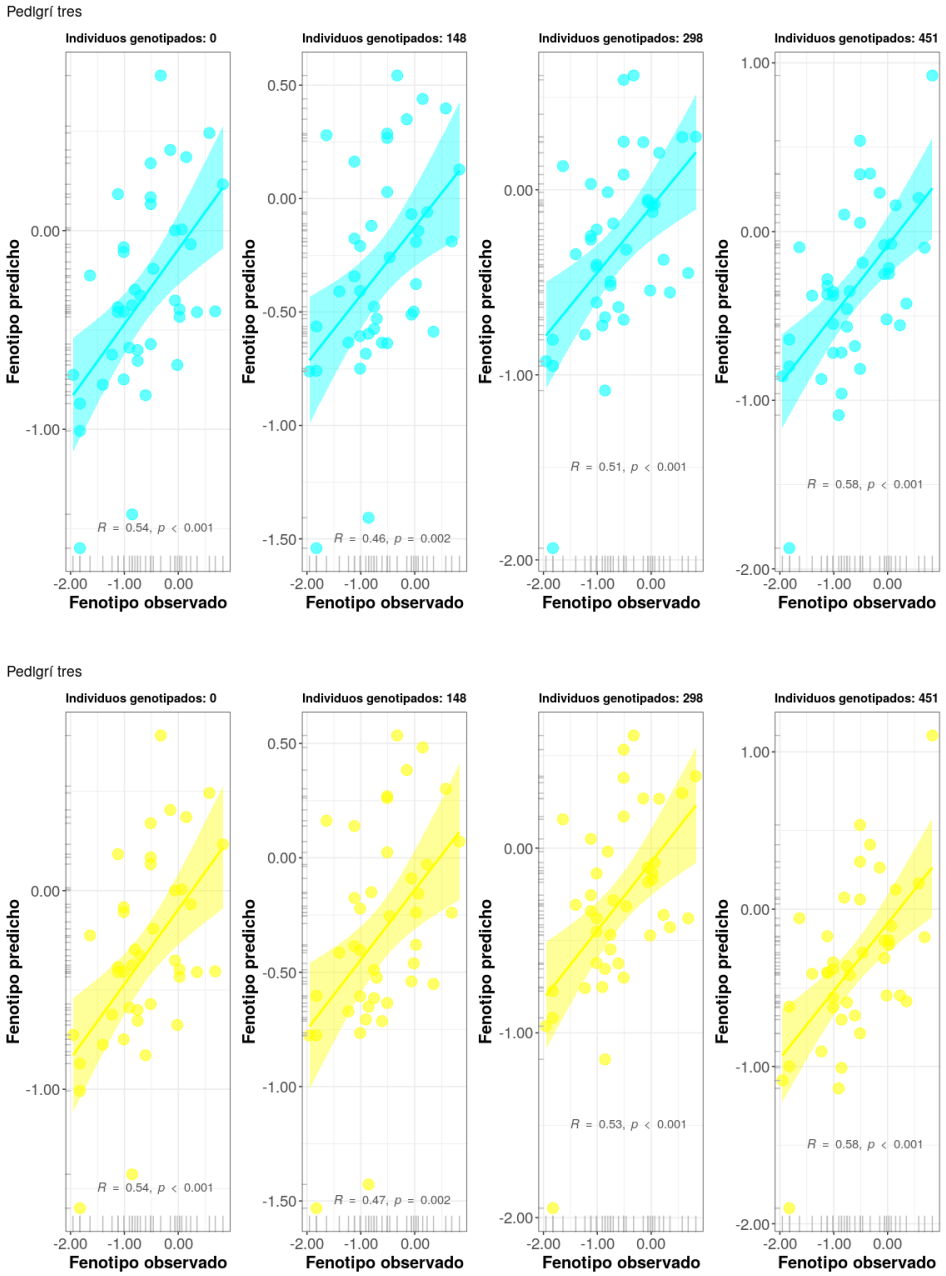

Pedlgrí tres
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.46, p = 0.002
R = 0.51, p < 0.001
R = 0.58, p < 0.001
Pedlgrí tres
Individuos genotipados: 0
Individuos genotipados: 148
Individuos genotipados: 298
Individuos genotipados: 451
Fenotipo predicho
Fenotipo observado
R = 0.54, p < 0.001
R = 0.47, p = 0.002
R = 0.53, p < 0.001
R = 0.58, p < 0.001

Figura A.6:

Bibliografía

Blasco, A. 2021. *Mejora genética animal*. 1st edition. EDITORIAL SÍNTESIS, S. A.

Blasco, A., y M. A. Toro. 2014. «A short critical history of the application of genomics to animal breeding». *Livestock Science* 166: 4-9.

Desta, Z. A., y R. Ortiz. 2014. «Genomic selection: genome-wide prediction in plant improvement». *Trends in Plant Science* 19 (9): 592-601.

Fernández, J., y M. Toro. 2006. «A new method to estimate relatedness from molecular markers». *Molecular Ecology* 15: 1657-67.

Fisher, R. A. 1918. «The correlaction between relatives under the supposition of Mendelian inheritance». *Transactions of the Royal Society of Edinburgh* 52: 399-433.

Hayes, B. J., P. M. Visscher, y M. E. Goddard. 2009. «Increased accuracy of artificial selection by using the realized relationship matrix». *Genetics Research* 91: 47-60.

Imai, A., T. Kuniga, T. Yoshioka, K. Nonaka, N. Mitani, H. Fukamachi, N. Hiehata, M. Yamamoto, y T. Hayashi. 2019. «Single-step genomic prediction of fruit-quality traits using phenotypic records of non-genotyped relatives in citrus». *PLoS ONE* 14 (8). https://doi.org/https://doi.org/10.1371/journal.pone.0221880.

Jurcic, E. J., P. V. Villalba, P. S. Pathauer, D. A. Palazzini, G. P. J. Oberschelp, L. Harrand, M. N. Garcia, et al. 2021. «Genomic selection: genome-wide prediction in plant improvement». *Trends in Plant Science* 127: 176-89.

Legarra, A., I. Aguilar, y I. Misztal. 2009. «A relationship matrix including full pedigree and genomic information». *Journal of Dairy Science* 92: 4656-63. https://doi.org/10.3168/jds.2009-2061.

Legarra, A., O. F. Christensen, I. Aguilar, y I. Misztal. 2014. «Single Step, a general approach for genomic selection». *Livestock Science*. https://doi.org/http://dx.doi.org/10.1016/j.livsci.2014.04.029.

Lourenco, D., A. Legarra, S. Tsuruta, Y. Masuda, I. Aguilar, y I. Misztal. 2020. «Single-Step Genomic Evaluations from Theory to Practice: Using SNP Chips and Sequence Data in BLUPF90». *Genes* 11: 790. https://doi.org/doi:10.3390/genes11070790.

Misztal, I., S. E. Aggrrey, y W. M. Muir. 2012. «Experiences with a single-step genome evaluation». *Poultry Science* 92: 2530-4.

Misztal, I., A. Legarra, y I. Aguilar. 2009. «Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information». *Journal of Dairy Science* 92: 4648-55. https://doi.org/10.3168/jds.2009-2064.

Misztal, I., D. Lourenco, y A. Legarra. 2020. «Current status of genomic evaluation». *Journal of Animal Science* 98 (4): 1-14. https://doi.org/10.1093/jas/skaa101.

Nakaya, A., y S. N. Isobe. 2012. «Will genomic selection be a practical method for plant breeding?» *Annals of Botany* 110: 1303-16.

Nelson, R. M., M. E. Pettersson, y Ö. Carlborg. 2012. «A century after Fisher: time for a new paradigm in quantitative genetics». *Trends in Genetics* 29 (9): 669-76.

Pérez-Enciso, M. 2017. «Animal breeding learning from machine learning». *Journal of Animal Breeding and Genetics* 134: 85-86.

Pérez-Rodríguez, P., J. Crossa, J. Rutkoski, J. Poland, R. Singh, A. Legarra, E. Autrique, J. Burgueño G. de los Campos, y S. Dreisigacker. 2017. «Single-step genomic and pedigree genotype x environment interaction models for predicting wheat lines in international environments». *Plant Genome* 10 (2). https://doi.org/10.3835/plantgenome2016.09.0089.

Tong, H., y Z. Nikoloski. 2021. «Machine learning approaches for crop improvement: leveraging phenotypic and genotypic big data». *Journal of Plant Physiology* 257: 153354. https://doi.org/10.1016/j.jplph.2020.153354.

Turelli, M. 2017. «Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps». *Theoretical Population Biology* 118: 46-49.

Villemereuil, P. de, H. Schielzeth, S. Nakagawa, y M. Morrissey. 2016. «General methods for evolutionary quantitative genetic inference from generalized mixed models». *Genetics* 204: 1281-94.

Vourlaki, I., R. Castanera, S. Ramos-Onsins, J. Casacuberta, y M. Pérez-Enciso. s. f. «Transposable element polymorphisms improve prediction of complex agronomic traits in rice». *Frontiers in Plant Science*.

Wright, S. 1922. «Coefficients of inbreeding and relationship». *The American Naturalist* 56: 330-38.

Xua, S., D. Zhub, y Q. Zhang. 2014. «Predicting hybrid performance in rice using genomic best linear unbiased prediction». *Proceedings of the National Academy of Sciences of the United States of America* 111 (34): 12456-61. https://doi.org/10.1073/pnas.1413750111.

Agradecimientos