GSE **G**enetics **S**election **E**volution

RESEARCH ARTICLE

Check for updates

A multivariate analysis with direct additive and inbreeding depression load effects

Luis Varona[1*], Juan Altarriba[1], Carlos Moreno[1], María Martínez-Castillero[2] and Joaquim Casellas[3]

Abstract

Background: Inbreeding is caused by mating between related individuals and its most common consequence is inbreeding depression. Several studies have detected heterogeneity in inbreeding depression among founder individuals, and recently a procedure for predicting hidden inbreeding depression loads associated with founders and the Mendelian sampling of non-founders has been developed. The objectives of our study were to expand this model to predict the inbreeding loads for all individuals in the pedigree and to estimate the covariance between the inbreeding loads and the additive genetic effects for the trait of interest. We tested the proposed approach with simulated data and with two datasets of records on weaning weight from the Spanish Pirenaica and Rubia Gallega beef cattle breeds.

Results: The posterior estimates of the variance components with the simulated datasets did not differ significantly from the simulation parameters. In addition, the correlation between the predicted and simulated inbreeding loads were always positive and ranged from 0.27 to 0.82. The beef cattle datasets comprised 35,126 and 75,194 records on weights between 170 and 250 days of age, and pedigrees of 308,836 and 384,434 individual-sire-dam entries for the Pirenaica and Rubia Gallega breeds, respectively. The posterior mean estimates of the variance of inbreeding depression loads were 29,967.8 and 28,222.4 for the Pirenaica and Rubia Gallega breeds, respectively. They were larger than those of the additive variance (695.0 and 439.8 for Pirenaica and Rubia Gallega, respectively), because they should be understood as the variance of the inbreeding depression achieved by a fully inbred (100%) descendant. Therefore, the inbreeding loads have to be rescaled for smaller inbreeding coefficients. In addition, a strong negative correlation (-0.43 ± 0.10) between additive effects and inbreeding loads was detected in the Pirenaica, but not in the Rubia Gallega breed.

Conclusions: The results of the simulation study confirmed the ability of the proposed procedure to predict inbreeding depression loads for all individuals in the populations. Furthermore, the results obtained from the two real datasets confirmed the variability in the inbreeding depression loads in both breeds and suggested a negative correlation of the inbreeding loads with the additive genetic effects in the Pirenaica breed.

Background

Inbreeding is caused by mating between related individuals and is associated with changes in the mean and variance of quantitative traits [1]. Inbreeding depression, which results in a reduction in the phenotypic yield of fitness-related traits [2, 3], has been widely observed in animals, plants, and humans [2, 4] and is the most common consequence of inbreeding. The genetic basis of inbreeding depression stems from the high degree of homozygosity in inbred individuals, which reveals the presence of recessive alleles or losses in the advantage of over-dominance at heterozygous loci [2, 3].

Genetic variation within a population implies that inbreeding depression can vary depending on the

---

*Correspondence: lvarona@unizar.es
[1] Departamento de Anatomía Embriología y Genética Animal, Instituto Agroalimentario de Aragón (IA2), Universidad de Zaragoza, 50013 Saragossa, Spain
Full list of author information is available at the end of the article

Varona *et al. Genet Sel Evol*   (2019) 51:78

Page 2 of 12

genotype of the individuals whose alleles produce identity-by-descent (IBD) states in their progeny. Variability in inbreeding depression has been confirmed within ancestral lineages of Drosophila [5] and in sire families in dairy [6] and beef cattle [7]. Therefore, it is possible to define a specific hidden individual inbreeding depression load [8] (hereafter referred to as inbreeding load), which can be considered a hereditary trait [5, 9] with a phenotype that is only expressed when inbreeding occurs in its offspring.

Lacy et al. [10] proposed a founder decomposition of inbreeding, which assigns a different inbreeding depression to each founder of the population; yet, the common ancestors of inbred individuals are not restricted to the founding generation. Caballero and Toro [11] and García-Cortés et al. [12] proposed a Mendelian decomposition of the inbreeding coefficient that assigns partial inbreeding coefficients to the founders and to the Mendelian sampling of the non-founders. This decomposition was the basis for the development of a mixed-model approach that allows prediction of individual inbreeding loads [8]. However, this model assumes that the founding and non-founding effects on inbreeding depression are distributed independently, and cannot be used to predict the inbreeding load of individuals that do not have inbred progeny. In this study, we propose an alternative parameterization that can predict the inbreeding load for each individual in the pedigree, and can provide estimates of the covariance between the inbreeding loads and the additive genetic effects for the trait of interest. We tested the procedure with a simulation study and two large datasets of records on weaning weight in the Pirenaica and Rubia Gallega beef cattle breeds.

Methods
Theory
Inbreeding load
The recessive alleles that appear in homozygosity caused by IBD from a common ancestor in the paternal and maternal lineages can cause inbreeding depression. Each ancestor might have different sets of recessive alleles; thus, the genotype of each individual determines an inbreeding load that can be interpreted as the effects on the trait performance of its inbred descendants. Inbreeding load can be considered a heritable trait [5, 9] and, from the perspective of the ancestors, it acts additively because the alternative alleles never interact in their inbred progeny.

Thus, the polygenic inbreeding load (i_i) for the ith individual can be decomposed [8] as follows:

$$i_i = i_s + i_d + \varepsilon_i, \tag{1}$$

where i_s and i_d are the inbreeding loads for its sire and dam, respectively, and ε_i is its Mendelian sampling.

Therefore, it follows that the distribution of the inbreeding loads is multivariate Gaussian as:

$$\mathbf{i} \sim N\left(0, \mathbf{A}\sigma_i^2\right),$$

where $\mathbf{A}$ is the numerator relationship matrix [13], and σ_i^2 is the additive genetic variance of the inbreeding loads.
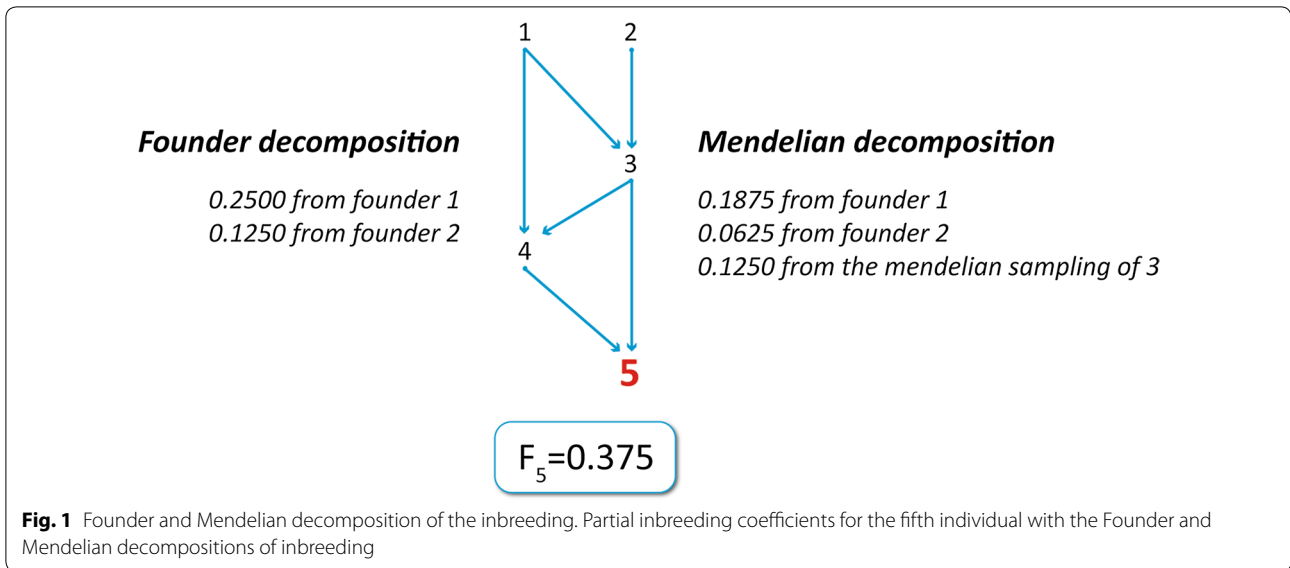
Inbreeding decomposition
Traditionally, inbreeding is split into several units that are attributed to founding individuals [10] by tracing back the origin of the alleles that might be IBD in inbred individuals. Alternatively, inbreeding can be decomposed into the sources of the co-ancestry between the parents of each individual [11, 12], which might include the founders of the population and the Mendelian sampling of the non-founders. The difference between the two approaches is illustrated by a simple pedigree of five individuals (Fig. 1). Individual 5 has an inbreeding of 0.375. The founder decomposition splits it between two founders, and attributes 0.25 and 0.125 to individuals 1 and 2, respectively. In contrast, the Mendelian decomposition attributes 0.1875 and 0.0625 to founders 1 and 2, respectively, and 0.125 to the Mendelian sampling of non-founder 3. Note that, under the founder decomposition, the genealogical inbreeding is the sum of the partial coefficients from the founders, while under the Mendelian decomposition it is the sum of the partial coefficients from the founders and from the Mendelian sampling of the non-founders.

Model
Based on the Mendelian decomposition of inbreeding, Casellas [8] proposed a linear model that explains the phenotypic data ($\mathbf{y}$) with two vectors of random effects: the additive genetic ($\mathbf{a}$), and one extra effect ($\boldsymbol{\varepsilon_i}$) that is attributed to the inbreeding depression generated by the inbreeding loads of their ancestors (the founders and the Mendelian sampling of the non-founders). Thus,

$$\mathbf{y} = \mathbf{1}'\mu + \mathbf{Za} + \boldsymbol{\varepsilon_i} + \mathbf{e}, \tag{2}$$

where μ is the general mean, $\mathbf{e}$ is the vector of the residuals, 1 is a vector of ones with the same length of $\mathbf{y}$, $\mathbf{Z}$ is the incidence matrix and $\mathbf{T}$ is the matrix that contains the partial inbreeding coefficients from the Mendelian decomposition and connects the phenotypic data of inbred individuals and the inbreeding loads of their common ancestors from the paternal and maternal lineages. Furthermore, $\mathbf{a} \sim N\left(0, \mathbf{A}\sigma_a^2\right)$ and $\mathbf{e} \sim N\left(0, \mathbf{I}\sigma_e^2\right)$, where $\mathbf{A}$ is the numerator relationship matrix and σ_a^2 and σ_e^2 are their associated variance components. Casellas [8] proposed that the distribution of $\boldsymbol{\varepsilon_i}$ is $\boldsymbol{\varepsilon_i} \sim N\left(0, \mathbf{I}\sigma_i^2\right)$ with σ_i^2 being the variance component of the inbreeding loads.

Varona *et al. Genet Sel Evol*   (2019) 51:78

Page 3 of 12

Fig. 1 Founder and Mendelian decomposition of the inbreeding. Partial inbreeding coefficients for the fifth individual with the Founder and Mendelian decompositions of inbreeding

However, this assumption is incorrect because the variance that is attributed to the inbreeding loads of the founders must be larger than the variance that is attributed to the Mendelian sampling of the non-founders. Thus, the appropriate distribution of ε_i is $\varepsilon_i \sim N\left(0, \mathbf{Q}\sigma_i^2\right)$, where $\mathbf{Q}$ is a diagonal matrix with a value 1 for the elements corresponding to founder individuals, and a value of $\frac{1}{2}\left(1 - \frac{F_S + F_D}{2}\right)$ for the elements corresponding to the non-founder individuals [11]. F_S and F_D are the genealogical inbreeding coefficient of the sire and dam of the individual, respectively.

In this study, we reparametrize the model in terms of individual inbreeding loads ($\mathbf{i}$) that, given their additive nature, can be represented as follows:

$$\mathbf{i} = \mathbf{Pi} + \boldsymbol{\varepsilon}_\mathbf{i},$$

where $\mathbf{P}$ is a matrix with a diagonal of 0 s and 0.5 in the elements that link an individual with its sire and dam, such that it establishes a recurrent relationship between the inbreeding loads of each individual with their parents, as in Eq. (1). As a result, $\boldsymbol{\varepsilon}_\mathbf{i} = (\mathbf{I} - \mathbf{P})\mathbf{i}$ and Model (2) can be rewritten as follows:

$$\mathbf{y} = \mathbf{1}'\mu + \mathbf{Za} + \mathbf{Ki} + \mathbf{e}, \tag{3}$$

where $\mathbf{K} = \mathbf{T}(\mathbf{I} - \mathbf{P})$ and $\mathbf{i} \sim N\left(0, \mathbf{A}\sigma_i^2\right)$. As a result, a multiple trait model that includes the genetic covariance between additive effects and inbreeding loads can then be defined as:

$$\begin{pmatrix} \mathbf{a} \\ \mathbf{i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{G} \otimes \mathbf{A}\right), \tag{4}$$

where $\mathbf{G} = \begin{pmatrix} \sigma_a^2 & \sigma_{ai} \\ \sigma_{ai} & \sigma_i^2 \end{pmatrix}$, and σ_{ai} is the covariance between additive effects and inbreeding loads.

Example

Given the pedigree of five individuals (Fig. 1), we assume that individuals 3, 4, and 5 have phenotypes 113, 87, and 96, respectively. Applying the method proposed by García-Cortés et al. [12] a partial inbreeding coefficient of 0.250 is obtained for the animal 4 generated by individual 1, and three partial inbreeding coefficients (0.1875, 0.0625, and 0.125) for animal 5, generated by individuals 1, 2, and 3, respectively. Therefore, the mixed-model equations for the implementation of the model defined by Eq. (3) requires the following vector and matrices:

$$\mathbf{y} = \begin{bmatrix} 113 \\ 87 \\ 96 \end{bmatrix},$$

$$\mathbf{Z} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{K} = \mathbf{T}(\mathbf{I} - \mathbf{P}) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0 & 0 & 0 & 0 \\ 0.125 & 0 & 0.125 & 0 & 0 \end{bmatrix},$$

where $\mathbf{T} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0 & 0 & 0 & 0 \\ 0.1875 & 0.0625 & 0.125 & 0 & 0 \end{bmatrix},$

Varona *et al. Genet Sel Evol*   (2019) 51:78

Page 4 of 12

and $\mathbf{I} - \mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ -0.5 & -0.5 & 1 & 0 & 0 \\ -0.5 & 0 & -0.5 & 1 & 0 \\ 0 & 0 & -0.5 & -0.5 & 1 \end{bmatrix}$.

In addition, the numerator relationship matrix ($\mathbf{A}$) and its inverse ($\mathbf{A}^{-1}$) are as follows:

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0.500 & 0.750 & 0.625 \\ 0 & 1 & 0.500 & 0.250 & 0.375 \\ 0.500 & 0.500 & 1 & 0.750 & 0.875 \\ 0.750 & 0.250 & 0.750 & 1.250 & 1 \\ 0.625 & 0.375 & 0.875 & 1 & 1.375 \end{bmatrix},$$

and

$$\mathbf{A}^{-1} = \begin{bmatrix} 2 & 0.5 & -0.5 & -1 & 0 \\ 0.5 & 1.5 & -1 & 0 & 0 \\ -0.5 & -1 & 3.071 & -0.429 & -1.143 \\ -1 & 0 & -0.429 & 2.571 & -1.143 \\ 0 & 0 & -1.143 & -1.143 & 2.286 \end{bmatrix},$$

which leads to the following mixed-model equations:

$$\begin{bmatrix} \mathbf{1}'\mathbf{1} & \mathbf{1}'\mathbf{Z} & \mathbf{1}'\mathbf{K} \\ \mathbf{Z}'\mathbf{1} & \mathbf{Z}'\mathbf{Z} + \mathbf{A}^{-1}g^{11}\sigma_e^2 & \mathbf{Z}'\mathbf{K} + \mathbf{A}^{-1}g^{12}\sigma_e^2 \\ \mathbf{K}'\mathbf{1} & \mathbf{K}'\mathbf{Z} + \mathbf{A}^{-1}g^{21}\sigma_e^2 & \mathbf{K}'\mathbf{K} + \mathbf{A}^{-1}g^{22}\sigma_e^2 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{a}} \\ \hat{\mathbf{i}} \end{bmatrix}$$
$$= \begin{bmatrix} \mathbf{1}'\mathbf{y} \\ \mathbf{Z}'\mathbf{y} \\ \mathbf{K}'\mathbf{y} \end{bmatrix},$$

where g^{11}, g^{12}, g^{22} are the elements of the inverse ($\mathbf{G}^{-1}$) of the covariance matrix between the additive and the inbreeding load effects. Assuming $\sigma_e^2 = 10$, $\sigma_a^2 = 2$, $\sigma_i^2 = 1$, and $\sigma_{ai} = -0.25$, then

$$\mathbf{G}^{-1} = \begin{bmatrix} g^{11} & g^{12} \\ g^{21} & g^{22} \end{bmatrix} = \begin{bmatrix} 2 & -0.25 \\ -0.25 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 0.516 & 0.129 \\ 0.129 & 1.032 \end{bmatrix},$$

and the equations become the following:

with solutions:

$$\begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{a}}_1 \\ \hat{\mathbf{a}}_2 \\ \hat{\mathbf{a}}_3 \\ \hat{\mathbf{a}}_4 \\ \hat{\mathbf{a}}_5 \\ \hat{\mathbf{i}}_1 \\ \hat{\mathbf{i}}_2 \\ \hat{\mathbf{i}}_3 \\ \hat{\mathbf{i}}_4 \\ \hat{\mathbf{i}}_5 \end{bmatrix} = \begin{bmatrix} 104.88 \\ -2.64 \\ 4.37 \\ 1.96 \\ -9.81 \\ -6.24 \\ -4.50 \\ -13.49 \\ -9.37 \\ -5.41 \\ -7.07 \end{bmatrix}.$$

Simulation

A pedigree of 30,000 individuals arranged in six discrete generations of 5000 individuals (2500 sires and 2500 dams) was simulated. Each generation was obtained from the random mating of 20 sires and 500 dams that were chosen randomly from the previous generation. Thus, neither selection nor purging were simulated. Once the pedigree was obtained, the direct additive genetic effects ($\mathbf{a}$) and the inbreeding loads ($\mathbf{i}$) were simulated based on the multivariate distribution from Eq. (4). The vector of the phenotypic records ($\mathbf{y}$) was generated by Eq. (3). The partial inbreeding coefficients required for the $\mathbf{T}$ matrix were derived from the procedure of García-Cortés et al. [12].

Two cases (i) and (ii) were simulated based on the following parameters:

i. $\mu = 100$, $\sigma_a^2 = 100$, $\sigma_i^2 = 10{,}000$, $\sigma_{ai} = -500$, r($\mathbf{a}$, $\mathbf{i}$) $= -0.5$ and $\sigma_e^2 = 100$,

ii. $\mu = 100$, $\sigma_a^2 = 100$, $\sigma_i^2 = 1000$, $\sigma_{ai} = 0$, r($\mathbf{a}$, $\mathbf{i}$) $= 0.0$ and $\sigma_e^2 = 100$,

$$\text{where r}(\mathbf{a}, \mathbf{i}) = \frac{\sigma_{ai}}{\sigma_a \sigma_i}.$$

$$\begin{bmatrix} 3 & 0 & 0 & 1 & 1 & 1 & 0.375 & 0 & 0.125 & 0 & 0 \\ 0 & 1.032 & 0.258 & -0.258 & -0.516 & 0 & 0.258 & 0.0645 & -0.0645 & -0.129 & 0 \\ 0 & 0.258 & 0.774 & -0.516 & 0 & 0 & 0.0645 & 0.1935 & -0.129 & 0 & 0 \\ 1 & -0.258 & -0.516 & 2.5849 & -0.2211 & -0.5897 & -0.0645 & -0.129 & 0.3962 & -0.0553 & -0.1474 \\ 1 & -0.516 & 0 & -0.2211 & 2.327 & -0.5897 & 0.121 & 0 & -0.0553 & 0.332 & -0.1474 \\ 1 & 0 & 0 & -0.5897 & -0.5897 & 2.1794 & 0.125 & 0 & -0.022 & -0.1474 & 0.2949 \\ 0.375 & 0.258 & 0.0645 & -0.0645 & 0.121 & 0.125 & 2.142 & 0.516 & -0.500 & -1.032 & 0 \\ 0 & 0.0645 & 0.1935 & -0.129 & 0 & 0 & -0.500 & 1.548 & -1.032 & 0 & 0 \\ 0.125 & -0.0645 & -0.129 & 0.3962 & -0.0552 & -0.022 & 1.0465 & -1.032 & 3.185 & -0.4423 & -1.1794 \\ 0 & -0.129 & 0 & -0.0553 & 0.3317 & -0.1474 & -1.032 & 0 & -0.4423 & 2.6537 & -1.1794 \\ 0 & 0 & 0 & -0.1474 & -0.1474 & 0.2948 & 0 & 0 & -1.1794 & -1.1794 & 2.3589 \end{bmatrix} \begin{bmatrix} \hat{\mu} \\ \hat{\mathbf{a}}_1 \\ \hat{\mathbf{a}}_2 \\ \hat{\mathbf{a}}_3 \\ \hat{\mathbf{a}}_4 \\ \hat{\mathbf{a}}_5 \\ \hat{\mathbf{i}}_1 \\ \hat{\mathbf{i}}_2 \\ \hat{\mathbf{i}}_3 \\ \hat{\mathbf{i}}_4 \\ \hat{\mathbf{i}}_5 \end{bmatrix} = \begin{bmatrix} 296 \\ 0 \\ 0 \\ 113 \\ 87 \\ 96 \\ 33.75 \\ 0 \\ 12 \\ 0 \\ 0 \end{bmatrix},$$

Varona *et al. Genet Sel Evol* (2019) 51:78

Page 5 of 12

The simulated variances of inbreeding loads (σ_i^2) were greater than the additive variances (σ_a^2) as they reflect the inbreeding depression achieved by a completely inbred (100%) descendent. The percentage of phenotypic variance explained by the inbreeding depression loads depends on the magnitude of the partial inbreeding coefficients. Therefore, in the first simulation case (i), the variance of inbreeding load for a partial inbreeding coefficient of 0.10 will be equal to the additive variance (100), whereas in the second (ii), it will be a tenth part (10).

The simulated datasets were analyzed under Model (3). We used a Bayesian approach which included uniform priors for the mean (μ) and for the variance components (**G**) and σ_e^2, and a multivariate Gaussian distribution for **a** and **i**. The analysis was implemented through a Gibbs Sampler [14] with two chains of 550,000 iterations, after discarding the first 50,000 iterations.

Beef cattle data

The dataset included 35,126 (Pirenaica) and 75,194 (Rubia Gallega) records on live weights (one per individual) between 170 and 250 days of age, which were provided by the two breeding associations and had been recorded since 1989 and 1991, respectively, until 2018. The average ($\pm$ standard deviation) weights in the Pirenaica and the Rubia Gallega breeds were 258.4 kg ($\pm$ 58.0 kg) and 281.4 kg ($\pm$ 46.7 kg), respectively. The genealogical information dated back to 1920 (Pirenaica) and 1955 (Rubia Gallega), and comprised 308,836 and 384,434 individual-sire-dam records, respectively. The average depth of the pedigree of the individuals for which there was phenotypic information was 5.9 and 3.9 generations for the Pirenaica and the Rubia Gallega breeds, respectively. In the Pirenaica breed, 248,471 (80.5%) individuals were inbred and had an average inbreeding ($\pm$ standard deviation) coefficient of 0.030 ($\pm$ 0.056). In the Rubia Gallega breed, inbreeding affected 176,048 (45.8%) individuals, and the average inbreeding coefficient ($\pm$ standard deviation) was 0.020 ($\pm$ 0.048). The evolutions of the average inbreeding and the percentage of inbred individuals from 1975 to 2017 are shown in Additional file 1: Figures S1 and Additional file 2: Figure S2 for Pirenaica and Rubia Gallega, respectively.

The partial inbreeding coefficients from the Mendelian decomposition of inbreeding were derived from the procedure of García-Cortés et al. [12], which generated 16,099,374 and 5,080,457 coefficients from 8721 and 3601 ancestors in the Pirenaica and Rubia Gallega breeds, respectively. The average partial inbreeding coefficients ($\pm$ standard deviation) were 5.7×10^{-4} ($\pm 4.2 \times 10^{-3}$) in the Pirenaica and 1.5×10^{-3} ($\pm 6.9 \times 10^{-3}$) in the Rubia Gallega breed. Most of the partial inbreeding coefficients

Table 1 Distribution of partial inbreeding coefficients (number and percentage) based on their magnitude in the Pirenaica and Rubia Gallega breeds

Partial inbreeding	Pirenaica	Rubia Gallega
< 10-4	10,996,924 (68.30%)	1,564,740 (30.80%)
10-4-10-3	3,967,842 (24.65%)	2,559,733 (50.38%)
10-3-10-2	1,028,578 (6.39%)	823,410 (16.21%)
10-1-10-2	126,238 (0.78%)	124,155 (2.44%)
> 10-1	9792 (0.06%)	8.419 (0.17%)

were very low (Table 1); 99.3 and 93.0%, and 97.4 and 81.2% were lower than 0.01 and 0.001 in the Pirenaica and Rubia Gallega breeds, respectively.

Once the partial inbreeding coefficients from the Mendelian decomposition of inbreeding were available, the phenotypic data were analyzed with the following model:

$$\mathbf{y} = \mathbf{t}c + \mathbf{Xb} + \mathbf{Wp} + \mathbf{Za} + \mathbf{Ki} + \mathbf{e},$$

where **t** is the vector of age at recording, **b** is the vector of systematic effects, which included sex and age of the dam (15 levels), **p** is the vector of herd-year-season effects (Pirenaica, 6503 levels and Rubia Gallega, 5251 levels), and **a**, **i**, and **e** are the vectors of direct genetic, inbreeding load, and residual effects, respectively. Furthermore, c is the covariate with age at recording, and **X**, **W**, **Z**, and **K** are the appropriate incidence matrices.

Under a hierarchical Bayesian analysis, it was assumed that prior distributions for the permanent environmental, additive, and inbreeding load effects followed a multivariate Gaussian distribution:

$$\mathbf{p} \sim \mathrm{N}\left(0, \mathbf{I}\sigma_p^2\right),$$

where σ_p^2 is the variance of the permanent environmental effects and

$$\begin{pmatrix} \mathbf{u} \\ \mathbf{i} \end{pmatrix} \sim N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \mathbf{G} \otimes \mathbf{A} \right),$$

as in Model (4).

The prior distribution for systematic effects were uniform between $-\mathrm{M}$ and M, being M a very large value. The prior distributions for σ_p^2 was an inverted Chi-square $\left(\chi^{-2}(s, v)\right)$ distribution with parameters $s = 0$ and $v = -2$, and the prior distribution for **G** were inverted Chi-square and inverted Wishart ($IW(\mathbf{S}, v)$) distribution with:

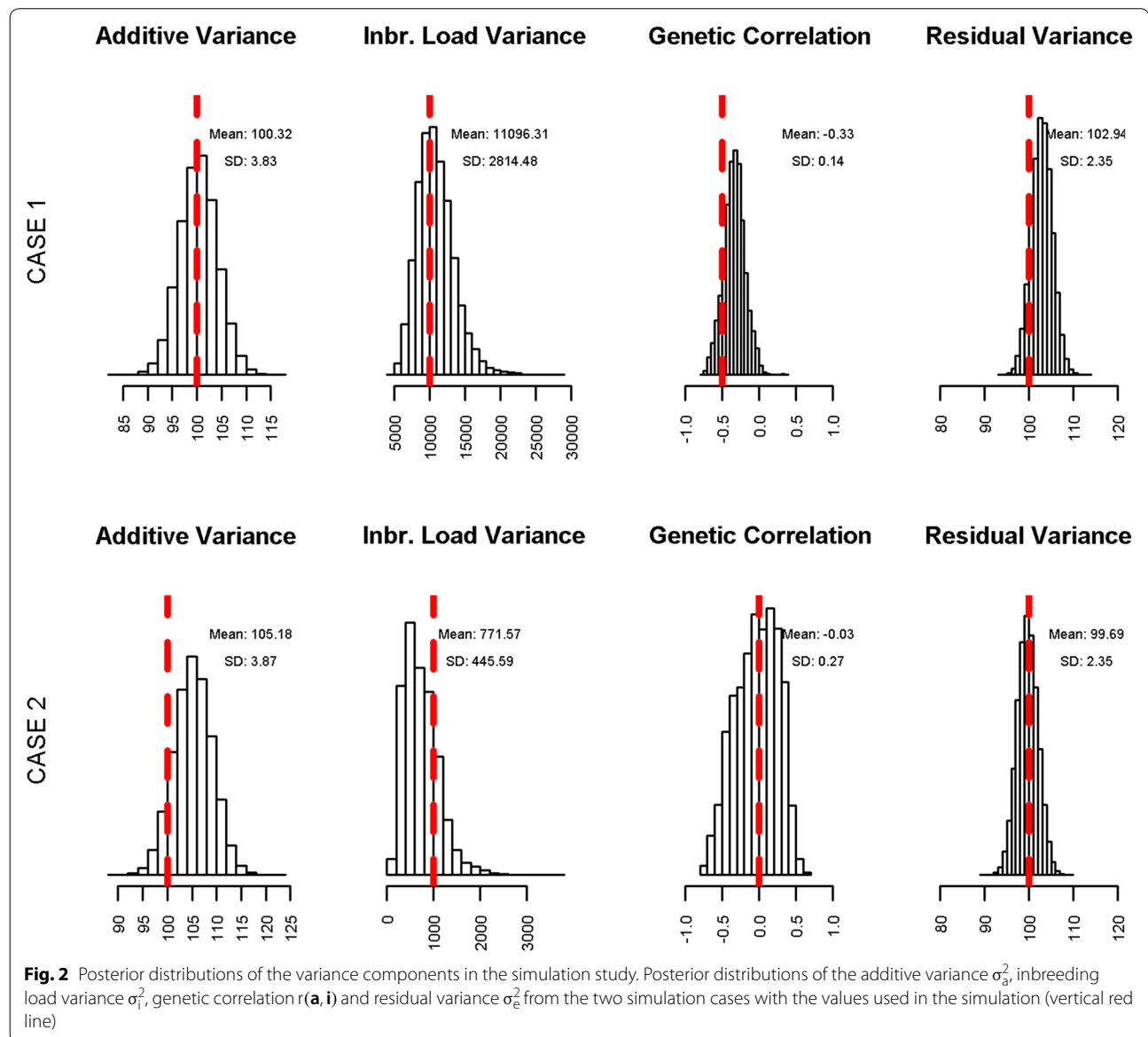$$\mathbf{S} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix},$$

and $v = -3$, that reduces to a uniform distribution. The Gibbs sampler [14] was implemented with two chains of 550,000 iterations, after discarding the first 50,000 iterations. Convergence was confirmed by visual inspection of the chains (see Additional file 3: Figures S3 and Additional file 4: Figure S4) and by the use of the CODA package in the R software [15].
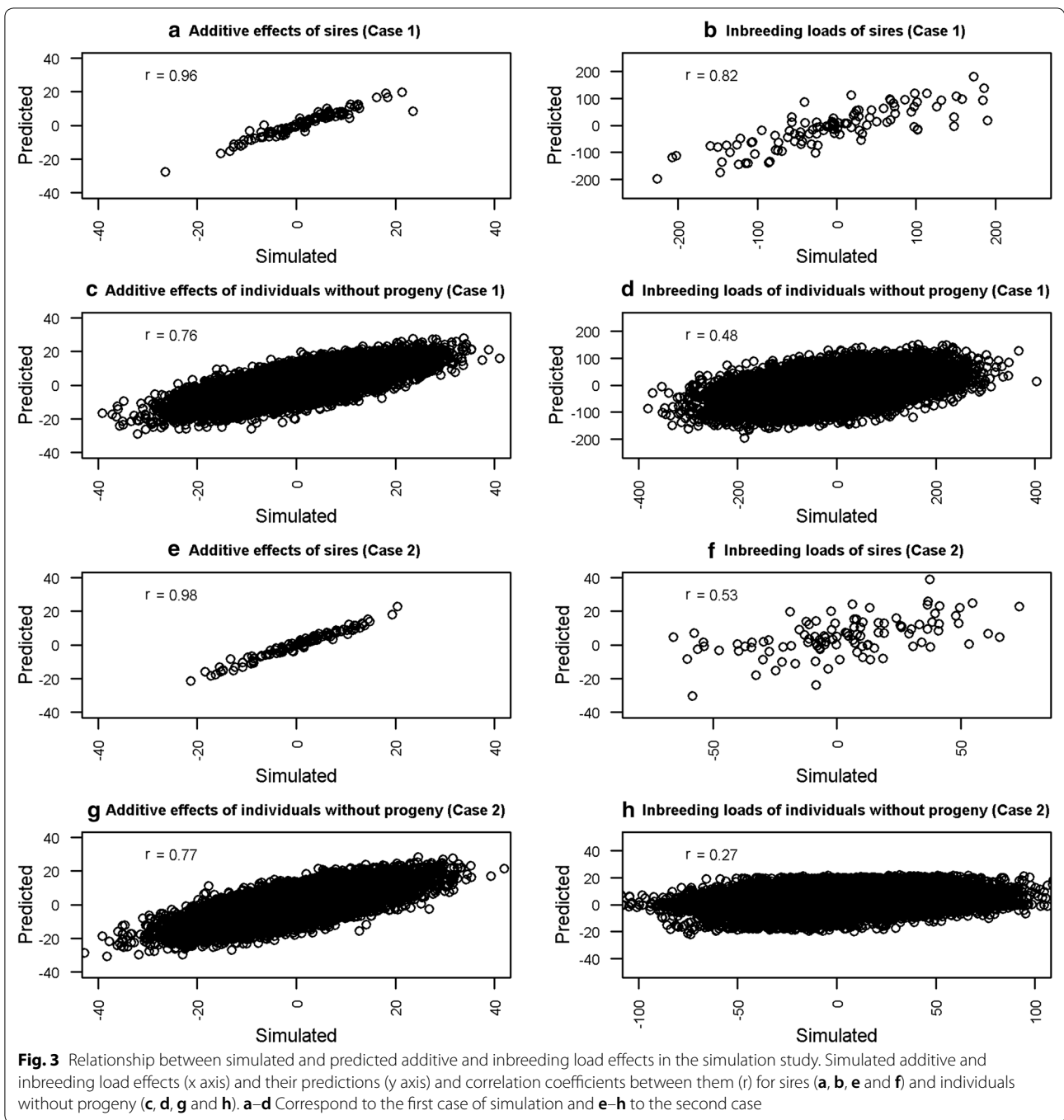
Results

Simulation study

The posterior distributions of σ_a^2, σ_i^2, $r(\mathbf{a}, \mathbf{i})$ and σ_e^2 obtained from the simulation study are shown in Fig. 2. The variance components used in the simulation were within the posterior mean ± 2 times the posterior standard deviation. Figure 3 shows the plots of the simulated direct genetic and inbreeding load effects and their predictions (posterior mean estimates) for sires and for individuals without progeny. In addition, the correlations between the simulated and predicted effects for sires and individuals without progeny are shown. The correlations between the direct genetic effects and their predictions ranged from 0.76 (individuals without progeny in case (i)) to 0.98 (sires in case (ii)), and were always higher than their counterparts for the inbreeding loads. Nevertheless, the correlations between simulated and predicted inbreeding loads were always positive, and ranged from 0.27 (individuals without progeny in case (ii)) to 0.82 (sires in case (i)).

Fig. 2 Posterior distributions of the variance components in the simulation study. Posterior distributions of the additive variance σ_a^2, inbreeding load variance σ_i^2, genetic correlation $r(\mathbf{a}, \mathbf{i})$ and residual variance σ_e^2 from the two simulation cases with the values used in the simulation (vertical red line)

Fig. 3 Relationship between simulated and predicted additive and inbreeding load effects in the simulation study. Simulated additive and inbreeding load effects (x axis) and their predictions (y axis) and correlation coefficients between them (r) for sires (**a**, **b**, **e** and **f**) and individuals without progeny (**c**, **d**, **g** and **h**). **a**–**d** Correspond to the first case of simulation and **e**–**h** to the second case

Beef cattle datasets

The posterior mean estimates of the variance components in the Pirenaica and Rubia Gallega breeds are in Table 2. The posterior mean estimates of the additive genetic variance and inbreeding load variance components were equal to 695.0 kg^2 and 29,966.8 kg^2 in Pirenaica, and 439.8 kg^2 and 28,222.4 kg^2 in Rubia Gallega, respectively. The posterior mean estimates of the genetic correlation between additive genetic and inbreeding load effects were − 0.43 in Pirenaica and − 0.04 in Rubia Gallega. The posterior mean estimates of the permanent environmental and residual variances were 1035.2 kg^2 and 483.1 kg^2 in Pirenaica, and 320.0 kg^2 and 1018.2 kg^2 in Rubia Gallega.

Figure 4 presents, for each breed, a bivariate plot of the predictions (posterior mean estimates) of the direct

Varona *et al. Genet Sel Evol* (2019) 51:78

Page 8 of 12

Table 2 Posterior mean estimates (and posterior standard deviation) of variance components in the Pirenaica and Rubia Gallega breeds

	Population	
	Pirenaica	Rubia Gallega
σ_a^2	695.016 (25.688)	439.803 (18.121)
σ_i^2	29,966.800 (5868.275)	28,222.360 (5454.273)
$r(\mathbf{a}, \mathbf{i})$	-0.429 (0.102)	-0.043 (0.087)
σ_p^2	1035.209 (27.714)	320.023 (11.273)
σ_e^2	483.060 (15.010)	1018.207 (12.612)

σ_a^2 is the additive genetic variance, σ_i^2 is the variance of individual inbreeding loads, $r(\mathbf{a}, \mathbf{i})$ is the correlation between additive genetic and individual inbreeding loads, σ_p^2 is the permanent environmental variance and σ_e^2 is the residual variance

additive and inbreeding load effects and a histogram that represents the distribution of the predicted inbreeding loads. In the Pirenaica breed, the bivariate plot indicated a clear negative relationship between direct genetic and inbreeding load effects, which was not apparent in the Rubia Gallega breed. The average (and standard deviation) of the predicted inbreeding loads were -69.3 (58.6) kg and -32.3 (41.7) kg in Pirenaica and Rubia Gallega breeds, respectively, and ranged from -369.1 to 234.7 in Pirenaica and from -374.1 to 266.1 in Rubia Gallega. The proportion of individuals in the Pirenaica and Rubia Gallega breeds that had a positive predicted inbreeding load was equal to 10.5% and 22.5%, respectively.
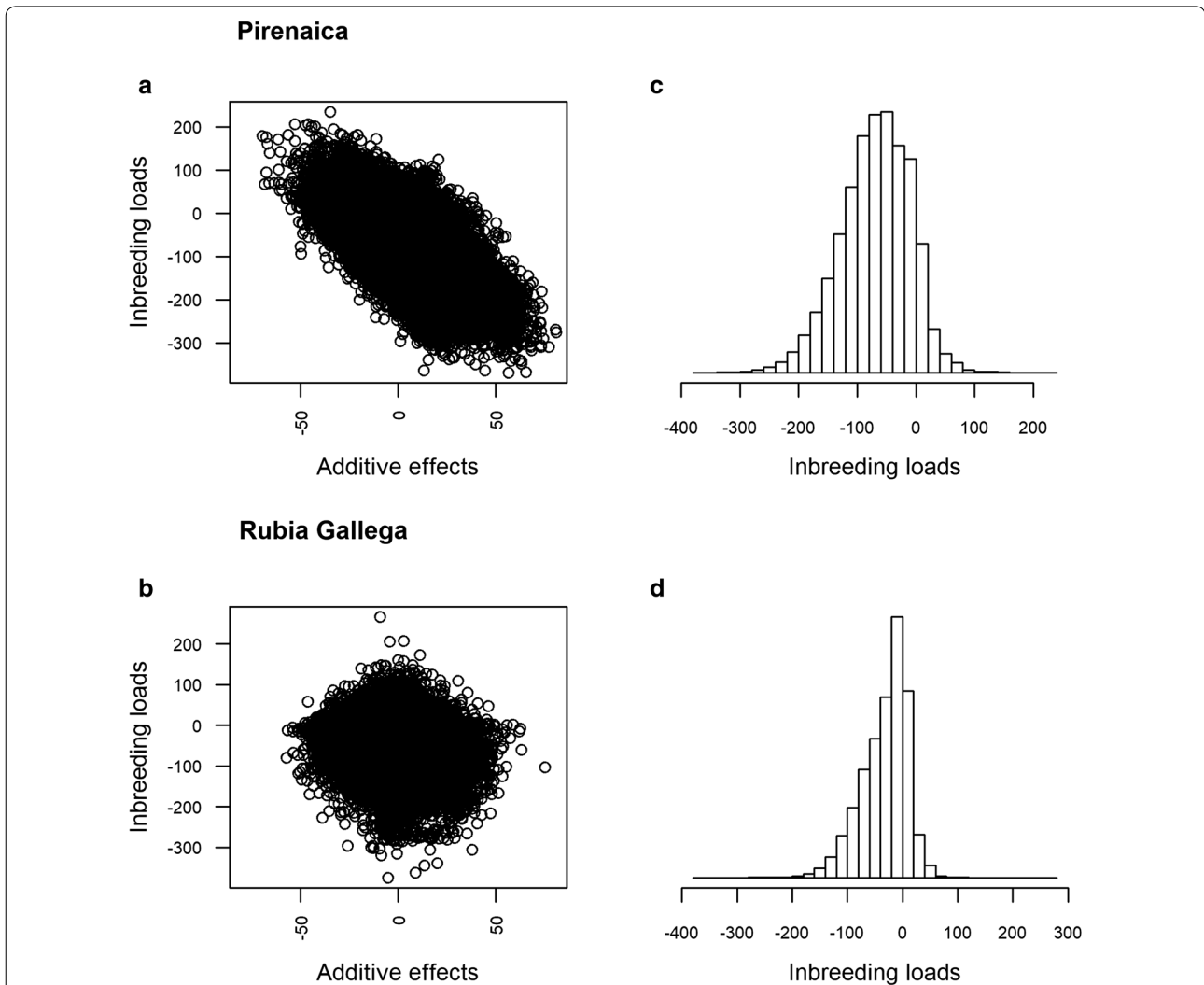

Fig. 4 Relationship between predicted additive and inbreeding load effects and distribution of inbreeding load effects for weaning weight in Pirenaica and Rubia Gallega. Predictions of the additive (x axis) and inbreeding load (y axis) effects for weaning weight in Pirenaica (**a**) and Rubia Gallega (**b**) and histograms of predictions of inbreeding load effects for weaning weight in Pirenaica (**c**) and Rubia Gallega (**d**)

Discussion

Inbreeding load and inbreeding depression are heterogeneous among populations [5] or families [6, 7], as inbreeding depression is affected by the variants of recessive alleles in each population, or even in each ancestor that generates inbreeding. Variability in inbreeding depression has been confirmed in several livestock populations [16–21] with the approach proposed by Lacy et al. [10], which is consistent with the results of previous studies [22–24] that revealed heterogeneity in the contribution of genomic regions to inbreeding depression in several traits and species.

Management of inbreeding to prevent the undesirable consequences is an important aspect of conservation genetics [25] and in livestock and plant breeding programs [26, 27]. Purging is a natural process of selection against recessive alleles that are expressed in the homozygous state [28]. Purging during past generations can be inferred in pedigreed populations [29–31], and evidence of purging has been detected in natural [32–34] and livestock [29, 35] populations.

Genetic variability of the inbreeding loads among individuals within a population can be taken into account in 'artificial' purging strategies, to avoid the use of those individuals with poor inbreeding loads for reproductive purposes. However, it requires prediction of the inbreeding loads of young individuals without inbred progeny. The approach proposed by Lacy et al. [10] is useful for describing the heterogeneity of inbreeding depression among founders, yet in most livestock populations, the founders live several generations before the current candidates for selection and, therefore, the prediction of the inbreeding loads based on Lacy's approach is not useful for practical 'artificial' purging strategies. Casellas [8] proposed an approach that uses the Mendelian decomposition of inbreeding and provides prediction of the inbreeding loads from both founders and the Mendelian sampling of the non-founders. However, it can only provide prediction of the inbreeding loads of the individuals that contribute to the inbreeding of individuals with known phenotypes. Moreover, its estimates of the variance of the inbreeding loads are biased, since it ignores the reduction in variance due to the Mendelian segregation of the progeny.

Although the approach proposed by Casellas [8] was the basis of the approach proposed in our study, rather than using directly the partial inbreeding coefficients derived from García-Cortes et al. [12], the coefficients were transformed by the I–P recursive transformation in order to predict the total individual inbreeding load ($\mathbf{i}$). The parametrization is crucial because it allows the assumption that individual inbreeding loads follow a multivariate Gaussian distribution and permits the use of the numerator relationship matrix (**A**) to predict the inbreeding loads of individuals that have yet to contribute to inbreeding but have relatives that have done so. The approach is illustrated with the results obtained from a simulation study, which used the same model of analysis to generate the simulated datasets. In spite of the limited amount of data, the posterior distribution of the variance components did not differ significantly from the simulated ones (Fig. 2), which suggests that there is enough information to recover the variance of inbreeding loads and their covariance with the direct additive effects. The most outstanding result of the simulation study is that it demonstrates the ability of prediction of the inbreeding loads of individuals without progeny, since the correlations between the predicted and simulated inbreeding loads were 0.27 and 0.48 in the first (i) and second (ii) case of simulation, respectively (Fig. 3). The higher correlation in the second case was caused by the larger σ_i^2 used in the simulation. The correlations obtained from the individuals without progeny were lower than those for the sires, which had many inbred progeny, and they were also lower than the correlations between simulated and predicted additive effects. Nonetheless, these results might provide a basis for the development of "artificial" purging strategies in order to avoid the undesirable consequences of inbreeding depression. One approach might be to develop a selection index by weighting the breeding values and inbreeding loads in an appropriate way.

Implementation of our model involved the construction of the mixed-model equations, which followed the same rationale as that of the maternal animal model [36]. The model attributes two genetic effects to each individual; one, i.e. **a** that it is expressed directly in the individual phenotype, and another, i.e. **i** that it is expressed only in the phenotype of its inbred descendants. In spite of the complexity of the model, the results obtained from the simulation study (Fig. 2) demonstrated its ability to recover the simulated parameters from the genealogical and phenotypic information. Computationally, the additional complexities beyond the standard mixed-model are the calculation of the partial inbreeding coefficients from the Mendelian decomposition of inbreeding and the calculation and storage of the $\mathbf{K'K}$ block. First, the tabular method that was implemented for the inbreeding decomposition described in [12] involves the construction of a square matrix with the dimension of the pedigree. Thus, for computational feasibility, the partial inbreeding coefficients of each individual were calculated sequentially from a reduced pedigree that only included its ancestors. Second, the $\mathbf{K'K}$ matrix is more dense than is the $\mathbf{Z'Z}$ block and it requires more memory storage and computational time. Nevertheless, the contributions of the partial inbreeding coefficients to the $\mathbf{K'K}$ matrix were squared and, thus, the very small partial inbreeding coefficients generated almost null values, which could be discarded.

Varona *et al. Genet Sel Evol*   (2019) 51:78

Page 10 of 12

The posterior distributions of the inbreeding load variances for weights at 210 days in the Pirenaica and Rubia Gallega breeds indicated high variability among them, which is consistent with the variability of inbreeding depression among sire families reported by Carolino and Gama [7]. The posterior mean estimates were much larger (29,966.8 and 28,222.4 in Pirenaica and Rubia Gallega breeds, respectively) than those of the other variance components in the model, which was confirmed by the broad range of the prediction of the inbreeding loads. However, it should be noted that the model provides estimates of the inbreeding loads that must be understood as the inbreeding depression achieved by a fully inbred (100%) descendent. Logically, the predictions are not realistic and have to be rescaled. Thus, the expected inbreeding depression from an ancestor that generates a partial inbreeding coefficient of 0.10 and has a predicted inbreeding load of -100 kg is -10 kg.

The analysis of the beef cattle datasets indicated that, for most of the individuals, the prediction of their inbreeding load was negative, which is consistent with several studies that have analyzed the inbreeding depression in body weights of beef cattle [7, 37–39]. However, a significant proportion of the individuals (10.52% in Pirenaica, and 22.54% in Rubia Gallega) had a positive prediction of their inbreeding loads. In spite of the prediction error variance of the inbreeding loads, these figures may indicate that a relevant proportion of the individuals might have had an incremental effect on the trait of interest. Therefore, the potential inbreeding caused by mating between the descendants might not need to be avoided or could even be favored, at least for the analyzed trait. It should be noted that with a similar posterior mean estimate of the variance component (29,966.8 and 28,222.4), the prediction of the inbreeding load effects was more variable in the Pirenaica (SD: 58.6 kg) than in the Rubia Gallega (SD: 41.7) breed. This was due to the different amounts of pedigree information available for the estimation of the inbreeding loads (5.9 vs. 3.9 generations in Pirenaica and Rubia Gallega breeds, respectively). A deeper pedigree generates more partial inbreeding coefficients (Table 1) and provides more reliable information for the prediction of the inbreeding loads, which is reflected in larger variances for the predictors.

The additive nature of the inbreeding loads implies that it can be genetically correlated with other traits such as the direct additive effect on the analyzed traits. The mean posterior estimate of the genetic correlation between the inbreeding loads and the direct additive effect in the Pirenaica breed was clearly negative (-0.43) and the posterior probability that it was negative was higher than 0.99. In contrast, the estimate of the genetic correlation in the Rubia Gallega breed was near zero. To our knowledge,

to date there are no published studies that provide estimates of the correlations between additive genetic and inbreeding load effects in livestock breeds, although Carolino and Gama [7] did not find a significant correlation between the breeding values of the sires and the inbreeding depression associated with them. The only study available which provided an estimate of the genetic correlation [40], although retracted, suggested a very small, negative genetic correlation [41]. A possible explanation of this negative correlation is that older individuals may have a smaller additive genetic effect due to selection and a higher inbreeding load as they are more exposed to purging than recent ancestors. This hypothesis may also explain the differences in the estimates of the genetic correlation between the Pirenaica and Rubia Gallega breeds, since the depth of the pedigree in the Pirenaica was larger, and, therefore, older individuals with more opportunities of purging were included. Regardless of its cause, if the negative correlation is confirmed, it will mean that individuals with high breeding values tend to cause worse inbreeding depression if their descendants are inbred.

Finally, it is worth noting that the assumed model uses a prior distribution of the inbreeding loads centered at zero. Further research is necessary to define alternative prior distributions that allow the mean of the inbreeding loads to differ from zero.

Conclusions

Our study proposes a mixed-model approach that includes individual inbreeding depression loads, direct additive effects and the covariance between them. The approach was applied to simulated data and two datasets with records on weaning weight from beef cattle breeds (Pirenaica and Rubia Gallega) in Spain. The results demonstrated the ability of the model to recover the simulated parameters and to provide prediction of the individual inbreeding loads of candidates for selection without inbred progeny.

Supplementary information

Additional file 1: Figure S1. Plot of the Gibbs Sampler chains for the inbreeding load variance in the Pirenaica Breed.

Additional file 2: Figure S2. Plot of the Gibbs Sampler chains for the inbreeding load variance in the Rubia Gallega Breed.

Additional file 3: Figure S3. Evolution of the number of inbred and non-inbred individuals and the average percentage of inbreeding from 1975 to 2017 in the Pirenaica Breed.

Additional file 4: Figure S4. Evolution of the number of inbred and non-inbred individuals and the average percentage of inbreeding from 1975 to 2017 in the Rubia Gallega Breed.

Varona *et al. Genet Sel Evol*   (2019) 51:78

Page 11 of 12

Author details
[1] Departamento de Anatomía Embriología y Genética Animal, Instituto Agroalimentario de Aragón (IA2), Universidad de Zaragoza, 50013 Saragossa, Spain. [2] Dipartimento di Agronomia Animali, Alimenti Risorce Naturali e Ambiente, Università degli Studi di Padova, 35122 Padua, Italy. [3] Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain.

References
1. Falconer DS, Mackay TFC. Introduction to quantitative genetics. Harlow: Longman Group; 1996.
2. Charlesworth D, Willis JH. The genetics of inbreeding depression. Nat Rev Genet. 2009;10:783–96.
3. Hedrick PW, Garcia-Dorado A. Understanding inbreeding depression, purging, and genetic rescue. Trends Ecol Evol. 2016;31:940–52.
4. Leroy G. Inbreeding depression in livestock species: review and meta-analysis. Anim Genet. 2014;45:618–28.
5. Fowler K, Whitlock MC. The variance in inbreeding depression and the recovery of fitness in bottlenecked populations. Proc Biol Sci. 1999;266:2061–6.
6. Miglior F, Burnside EB, Hohenboken WD. Heterogeneity among families of Holstein cattle. In: Proceedings of the 5th world congress on genetics applied to livestock production: 7–12 August 2014. Guelph; 1994. p. 479–82.
7. Carolino N, Gama LT. Inbreeding depression on beef cattle traits: estimates, linearity of effects and heterogeneity among sire-families. Genet Sel Evol. 2009;40:511–27.
8. Casellas J. On individual-specific prediction of hidden inbreeding depression load. J Anim Breed Genet. 2018;135:37–44.
9. Moorad JA, Wade MJ. A genetic interpretation of the variation in inbreeding depression. Genetics. 2005;170:1373–84.
10. Lacy RC, Aleks G, Walsh A. Hierachical analysis of inbreeding depression in *Peromyscus polionotus*. Evolution. 1996;50:2187–200.
11. Caballero A, Toro MA. Interrelations between effective population size and other pedigree tools for the management of conserved populations. Genet Res. 2000;75:331–43.
12. García-Cortés LA, Martínez-Ávila JC, Toro MA. Fine decomposition of the inbreeding and the coancestry coefficients by using the tabular method. Conserv Genet. 2010;11:1945–52.
13. Wright S. Coefficients of inbreeding and relationship. Am Nat. 1922;56:330–8.
14. Gelfand AE, Smith AFM. Sampling-based approaches to calculating marginal densities. J Am Stat Assoc. 1990;85:398–409.
15. Plummer M, Best N, Cowles K, Vines K. CODA: convergence diagnosis and output analysis for MCMC. R News. 2006;6:7–11.
16. Rodrigáñez J, Toro MA, Rodriguez MC, Silió L. Effect of founder allele survival and inbreeding depression on litter size in a closed line of Large White pigs. Anim Sci. 1998;67:573–82.
17. Casellas J, Piedrafita J, Caja G, Varona L. Analysis of founder-specific inbreeding depression on birth weight in Ripollesa lambs. J Anim Sci. 2009;87:72–9.
18. Casellas J, Vidal-Roqueta D, Flores E, Casellas-Vidal D, Llach-Vila M, Salgas-Fina R, et al. Epistasis for founder-specific inbreeding depression in rabbits. J Hered. 2011;102:157–64.
19. Casellas J, Varona L, Ibáñez-Escriche N, Quintanilla R, Noguera JL. Skew distribution of founder-specific inbreeding depression effects on the longevity of Landrace sows. Genet Res. 2008;90:499–508.
20. Gulisija D, Gianola D, Weigel KA, Toro MA. Between-founder heterogeneity in inbreeding depression for production in Jersey cows. Livest Sci. 2006;104:244–53.
21. Todd ET, Ho SYW, Thomson PC, Ang RA, Velie BD, Hamilton NA. Founder-specific inbreeding depression affects racing performance in Thoroughbred horses. Sci Rep. 2018;8:6167.
22. Pryce JE, Haile-Mariam M, Goddard ME, Hayes BJ. Identification of genomic regions associated with inbreeding depression in Holstein and Jersey dairy cattle. Genet Sel Evol. 2014;46:71.
23. Saura M, Fernández A, Varona L, Fernández AI, De Cara MÁR, Barragán C, et al. Detecting inbreeding depression for reproductive traits in Iberian pigs using genome-wide data. Genet Sel Evol. 2015;47:1.
24. Ferenčaković M, Sölkner J, Kapš M, Curik I. Genome-wide mapping and estimation of inbreeding depression of semen quality traits in a cattle population. J Dairy Sci. 2017;100:4721–30.
25. Hedrick PW, Kalinowski ST. Inbreeding depression in conservation biology. Annu Rev Ecol Syst. 2000;31:139–62.
26. Meuwissen TH. Maximizing the response of selection with a predefined rate of inbreeding. J Anim Sci. 1997;75:934–40.
27. Woolliams JA, Berg P, Dagnachew BS, Meuwissen THE. Genetic contributions and their optimization. J Anim Breed Genet. 2015;132:89–99.
28. Crnokrak P, Barrett SCH. Perspective: purging the genetic load: a review of the experimental evidence. Evolution. 2002;56:2347–58.
29. Gulisija D, Crow JF. Inferring purging from pedigree data. Evolution. 2007;61:1043–51.
30. García-Dorado A, Wang J, López-Cortegano E. Predictive model and software for inbreeding-purging analysis of pedigreed populations. G3 (Bethesda). 2016;6:3593–601.
31. López-Cortegano E, Bersabé D, Wang J, García-Dorado A. Detection of genetic purging and predictive value of purging parameters estimated in pedigreed populations. Heredity. 2018;121:38–51.
32. Boakes EH, Wang J, Amos W. An investigation of inbreeding depression and purging in captive pedigreed populations. Heredity. 2007;98:172–82.
33. Facon B, Hufbauer RA, Tayeh A, Loiseau A, Lombaert E, Vitalis R, et al. Inbreeding depression is purged in the invasive insect *Harmonia axyridis*. Curr Biol. 2011;21:424–7.
34. Moreno E, Pérez-González J, Carranza J, Moya-Laraño J. Better fitness in captive Cuvier's Gazelle despite inbreeding increase: evidence of purging? PLoS One. 2015;10:e0145111.
35. Mc Parland S, Kearney F, Berry DP. Purging of inbreeding depression within the Irish Holstein-Friesian population. Genet Sel Evol. 2009;41:16.
36. Henderson CR. Applications of linear models in animal breeding. Guelph: University of Guelph; 1984.

Varona *et al. Genet Sel Evol*   (2019) 51:78

Page 12 of 12

37. Burrow HM. The effects of inbreeding in beef cattle. Anim Breed Abstr. 1993;61:737–50.
38. Pereira RJ, Santana ML Jr, Ayres DR, Bignardi AB, Menezes GR, Silva LO, et al. Inbreeding depression in Zebu cattle traits. J Anim Breed Genet. 2016;133:523–33.
39. Reverter A, Porto-Neto LR, Fortes MRS, Kasarapu P, de Cara MAR, Burrow HM, et al. Genomic inbreeding depression for climatic adaptation of tropical beef cattle. J Anim Sci. 2017;95:3809–21.
40. Tomkins JL, Penrose MA, Greeff J, LeBas NR. Additive genetic breeding values correlate with the load of partially deleterious mutations. Science. 2010;328:892–4.
41. Tomkins JL, Penrose MA, Greeff J, Lebas NR. Retraction. Science. 2011;333:1220.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.