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

**RESEARCH ARTICLE**      **Open Access**

CrossMark

# Consequences of paternally inherited effects on the genetic evaluation of maternal effects

Luis Varona[1,2*], Sebastián Munilla[1,3], Joaquim Casellas[4], Carlos Moreno[1,2] and Juan Altarriba[1,2]

## Abstract

**Background:** Mixed models are commonly used for the estimation of variance components and genetic evaluation of livestock populations. Some evaluation models include two types of additive genetic effects, direct and maternal. Estimates of variance components obtained with models that account for maternal effects have been the subject of a long-standing controversy about strong negative estimates of the covariance between direct and maternal effects. Genomic imprinting is known to be in some cases statistically confounded with maternal effects. In this study, we analysed the consequences of ignoring paternally inherited effects on the partitioning of genetic variance.

**Results:** We showed that the existence of paternal parent-of-origin effects can bias the estimation of variance components when maternal effects are included in the evaluation model. Specifically, we demonstrated that adding a constraint on the genetic parameters of a maternal model resulted in correlations between relatives that were the same as those obtained with a model that fits only paternally inherited effects for most pairs of individuals, as in livestock pedigrees. The main consequence is an upward bias in the estimates of the direct and maternal additive genetic variances and a downward bias in the direct-maternal genetic covariance. This was confirmed by a simulation study that investigated five scenarios, with the trait affected by (1) only additive genetic effects, (2) only paternally inherited effects, (3) additive genetic and paternally inherited effects, (4) direct and maternal additive genetic effects and (5) direct and maternal additive genetic plus paternally inherited effects. For each scenario, the existence of a paternally inherited effect not accounted for by the estimation model resulted in a partitioning of the genetic variance according to the predicted pattern. In addition, a model comparison test confirmed that direct and maternal additive models and paternally inherited models provided an equivalent fit.

**Conclusions:** Ignoring paternally inherited effects in the maternal models for genetic evaluation can lead to a specific pattern of bias in variance component estimates, which may account for the unexpectedly strong negative direct-maternal genetic correlations that are typically reported in the literature.

## Background

Genetic evaluation based on mixed models [1] has become the main tool for prediction of breeding values in livestock populations. The most commonly used parameterization accounts for the direct polygenic additive genetic effect inherent to each individual and for several systematic effects (e.g., sex, herd, season, etc.), as well as a residual source of variation. However, some traits may also be affected by maternal genetic effects that refer to the influence from the genome of the dam (e.g., milk production of the dam). These maternal genetic effects are commonly considered in genetic evaluation models for some livestock populations [2, 3]. Nevertheless, variance components estimates that are obtained with models that account for maternal effects have been the subject of a long-standing controversy about the unexpectedly strong negative estimates of the covariance between direct and maternal genetic effects [4–6].

Genomic imprinting [7, 8] is the total or partial silencing of paternal or maternal alleles in the progeny. Several specific genes with imprinted inheritance have been identified in livestock species [9], such as the callipyge mutation in sheep [10, 11] and the *insulin-like growth factor 2* (*IGF2*) gene in pig [12, 13]. Imumorin et al. [14]

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

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 2 of 13

also reported that genomic imprinting is involved in the determinism of quantitative trait loci (QTL) in cattle. More generally, genomic imprinting has been described as a widespread phenomenon in other mammals [15]. Studies on pig growth [16] and bovine fatness traits [17] have also provided statistical evidence for the existence of genetic variation associated with paternal parent-of-origin inheritance that can account for over 10 % of the phenotypic variability in these traits. More recently, Neugenbauer et al. [18, 19] and Tier and Meyer [20] confirmed this result on other cattle and pig traits.

Maternal effects can be statistically confounded with maternal inheritance, as noted by Hager et al. [21]. Furthermore, Meyer and Tier [22] suggested that ignoring parent-of-origin effects may be one of the possible causes for the unexpectedly large negative estimates of the correlation between direct and maternal genetic effects. In this paper, we provide further evidence that supports this latter argument by showing how the existence of a paternal parent-of-origin inheritance caused by the silencing of the alleles in the maternal gametes by imprinting can bias estimates of variance components in models that account for maternal effects. Specifically, we show that for any two individuals that were not related within the paternal lineage, direct and maternal additive genetic effects may generate the same expected covariance between relatives as paternal parent-of-origin or gametic effects. For most livestock pedigrees, the consequence of this is a downward bias in the estimate of the direct-maternal genetic covariance when paternal inheritance occurs and is not properly accounted for. To confirm this postulate, we developed a simulation experiment. In addition, we analysed a dataset of weight records at 210 days for the Pirenaica beef cattle population and we discuss the results in the light of this new evidence.

## Methods

### Quantitative genetic models

Consider a model with direct and maternal additive genetic effects. Hereafter, we use the subscripts "*o*" and "*m*" to differentiate between direct and maternal components, respectively. In addition, capital letters are used to identify different individuals. Under this model, the phenotypic values of individuals $A$ and $B$ are defined as:

$$\begin{aligned} y_A &= \mu + a_{oA} + a_{mA_D} + e_A \\ y_B &= \mu + a_{oB} + a_{mB_D} + e_B \end{aligned},$$

where $\mu$ is the general mean, $a_{oX}$ and $a_{mX}$ are the direct and maternal additive genetic effects, respectively, and $e_X$ is the residual for individual $X = \{A, B, A_D, B_D\}$. According to this notation, $A_D$ and $B_D$ stand for the dams of $A$ and $B$, respectively.

Information for the estimation of variance components estimation originates from covariances between relatives. Under the model with maternal effects described above, and assuming independence of the residuals, the covariance between the phenotypic values of individuals $A$ and $B$ is:

$$\begin{aligned} \mathrm{cov}(y_A, y_b) &= \mathrm{cov}(a_{oA}, a_{oB}) + \mathrm{cov}(a_{oA}, a_{mB_D}) \\ &\quad + \mathrm{cov}(a_{mA_D}, a_{oB}) + \mathrm{cov}(a_{mA_D}, a_{mB_D}) \\ &= r(A, B)\sigma_a^2 + r(A, B)\sigma_{am} + r(A_D, B)\sigma_{am} \\ &\quad + r(A_D, B_D)\sigma_m^2, \end{aligned}$$

where $\sigma_a^2$, $\sigma_m^2$, and $\sigma_{am}$ are the direct additive genetic variance, the maternal additive genetic variance, and the direct-maternal genetic covariance, respectively, and $r(X,Y)$ represents the additive relationship between individuals $X$ and $Y$. In this context, the main sources of information to estimate maternal parameters are the observed correlations between phenotypes of full sibs (FS), maternal half sibs (MHS) and dam-offspring (DO) pairs. Equated to their expectations under the maternal model, the following expressions are obtained:

$$\mathrm{cov(FS)} = \frac{1}{2}\sigma_a^2 + \sigma_{am} + \sigma_m^2,$$

$$\mathrm{cov(MHS)} = \frac{1}{4}\sigma_a^2 + \sigma_{am} + \sigma_m^2,$$

$$\mathrm{cov(DO)} = \frac{1}{2}\sigma_a^2 + \frac{5}{4}\sigma_{am} + \frac{1}{2}\sigma_m^2.$$

Now, let:

$$\sigma_a^2 = 4\sigma_m^2 = -2\sigma_{am}. \tag{1}$$

A little algebra shows that under condition (1) the previous expressions reduce to:

$$\mathrm{cov(FS)} = \frac{1}{4}\sigma_a^2,$$

$$\mathrm{cov(MHS)} = 0,$$
$$\mathrm{cov(DO)} = 0.$$

This is exactly the expectation of the covariance between relatives for these same relationships under a paternal gametic model of inheritance. In the Appendix section, we show that this result can be generalised to any genealogical relationship between a pair of individuals, on the condition that one was not an ancestor of the other within the paternal lineage. Specifically, the covariance between the phenotypic values of any two individuals (say $A$ and $B$) under the condition set by Equation (1) is:

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 3 of 13

$$\text{cov}(y_A, y_B) = \frac{1}{4}r(A_S, B_s)\sigma_a^2,$$

where $A_S$ and $B_S$ are the sires of $A$ and $B$, respectively.

This striking result implies that the parameterization of the genetic covariance under the standard maternal animal model is confounded with a model that explains the whole genetic variation as paternally inherited; i.e.

$$\text{cov}(y_A, y_B) = r_g(A^P, B^P)\sigma_s^2,$$

where $r_g(A^P,B^p)$ is the gametic relationship [23] between paternal gametes of individuals $A$ and $B$ ($A^p$ and $B^p$) and $\sigma_s^2$ is the variance caused by paternally inherited gametic effects. Stated the other way around, the existence of paternally inherited variation that is not accounted for by the maternal model will tend to be assigned to the remaining variance components according to Equation (1) in order to accommodate the observed correlations between relatives. As a consequence, variance component estimation will explicitly inflate estimates of the direct and maternal additive genetic variances, with a magnitude close to twice and one half the paternally inherited variance, respectively, and it will reduce the direct-maternal genetic covariance by a value equal to the paternally inherited variance. The following section describes a simulation experiment that we developed in order to test this result.

### Simulation study

Each simulated dataset consisted of a base population of 1000 individuals (500 males and 500 females) and two generations of 5000 phenotyped individuals that comprised 2500 males and 2500 females. Phenotypic records were generated for the individuals of the last two generations based on the following model:

$$\mathbf{y} = \mu + \mathbf{Z_a a_o} + \mathbf{Z_m a_m} + \mathbf{Z_s s} + \mathbf{e},$$

where $\mathbf{y}$ is the vector of phenotypic data, $\mu$ is the general mean, set to 100 units, $\mathbf{a_o}$ and $\mathbf{a_m}$ are vectors of the direct and maternal genetic effects, $\mathbf{s}$ is a vector of paternally inherited gametic effects, and $\mathbf{e}$ is the vector of residuals, while $\mathbf{Z_a}$, $\mathbf{Z_m}$ and $\mathbf{Z_s}$ are incidence matrices. Covariances between genetic ($\mathbf{a_o}$ and $\mathbf{a_m}$) and gametic effects ($\mathbf{s}$) were assumed to be zero and

$$Var(\mathbf{e}) = \mathbf{I}\sigma_{\mathbf{e}}^2,$$

where $\mathbf{I}$ is the identity matrix and $\sigma_e^2$ is the residual

variance. The variances of the direct and maternal additive genetic effects are:

$$Var\begin{pmatrix} \mathbf{a_o} \\ \mathbf{a_m} \end{pmatrix} = \mathbf{A} \otimes \mathbf{T},$$

where $\mathbf{A}$ is the numerator relationship matrix and

$$\mathbf{T} = \begin{pmatrix} \sigma_a^2 & \sigma_{am} \\ \sigma_{am} & \sigma_m^2 \end{pmatrix},$$

For the paternally inherited gametic effects ($\mathbf{s}$):

$$Var(\mathbf{s}) = \mathbf{G}\sigma_s^2,$$

where $\mathbf{G}$ is the gametic relationship matrix

Five simulation scenarios were developed:

Scenario 1 corresponds to a pure direct additive model of inheritance with the following simulated parameters:

$$\sigma_a^2 = 500, \sigma_m^2 = 0, \sigma_{am} = 0, \sigma_s^2 = 0, \sigma_e^2 = 1000.$$

Scenario 2 assumes that the only source of genetic variation is the paternally inherited variance:

$$\sigma_a^2 = 0, \sigma_m^2 = 0, \sigma_{am} = 0, \sigma_s^2 = 250, \sigma_e^2 = 1000.$$

Scenario 3 combines both sources of genetic variation, direct additive and paternally inherited variance. The simulated parameters were:

$$\sigma_a^2 = 500, \sigma_m^2 = 0, \sigma_{am} = 0, \sigma_s^2 = 250, \sigma_e^2 = 1000.$$

Scenario 4 corresponds to the covariance structure that is assumed in a standard maternal animal model. A strong negative covariance between direct and maternal additive genetic effects was also simulated in order to mimic estimates that are frequently achieved in livestock populations for maternally influenced traits:

$$\sigma_a^2 = 500, \sigma_m^2 = 250, \sigma_{am} = -250, \sigma_s^2 = 0, \sigma_e^2 = 1000.$$

Finally, Scenario 5 includes all three types of genetic effects: direct and maternal additive and paternally inherited. The simulated parameters were:

$$\sigma_a^2 = 500, \sigma_m^2 = 250, \sigma_{am} = -250, \sigma_s^2 = 250, \sigma_e^2 = 1000.$$

For each of these simulation scenarios, a total of 10 independent populations were generated.

### Pirenaica beef cattle data

In addition to the simulated scenarios, we analysed data for phenotypic records on the Pirenaica beef cattle breed. The Pirenaica breed is a meat-type beef population from northern Spain with an approximate census of 20 000 individuals

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 4 of 13

that are typically reared under extensive conditions [24]. The dataset consisted of 17 069 records for weight at 210 days (W210), with an average value of 266 kg and a raw standard deviation of 52.9 kg. Phenotypic records were adjusted according to the recommendations of the Beef Improvement Federation [25]. In addition, a pedigree file including 125 974 individual-sire-dam records was used. This information was provided by the National Breeders Confederation (*Confederación Nacional de Asociaciones de Ganado Pirenaico-CONASPI*; http://www.conaspi.es). Ethical approval for animal care and use was not required for this study since all data was field data from the Yield Recording System of the Pirenaica breed; furthermore, data was recorded by the stockbreeders themselves, under standard farm management, with no additional requirements.

An additional simulation study was performed based on the same scenarios as described in the previous section, but using genealogical information of the Pirenaica beef cattle breed and replacing each phenotypic record by a simulated value. The aim was to replicate the same structure of relationships between individuals of the real dataset.

### Statistical models of estimation

Although a realistic implementation of all potential genetic sources of variation in the same hierarchical mixed linear model may result in non-estimable functions, the description of the full model is essential to understand the simplified parameterizations that are used below. The full statistical model is:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Z_a a_o} + \mathbf{Z_m a_m} + \mathbf{Z_s s} + \mathbf{Z_d d} + \mathbf{Z_p p} + \mathbf{Z_h h} + \mathbf{e},$$

where $\mathbf{y}$ is the vector of phenotypic data, $\mathbf{b}$ is the vector of systematic effects, consisting of the general mean for the simulated dataset and the two sexes and 16 age-groups of dam age of parity for the Pirenaica beef cattle data, and $\mathbf{a_o}$, $\mathbf{a_m}$, $\mathbf{s}$ and $\mathbf{d}$ are vectors of direct additive genetic, maternal additive genetic, paternal gametic and maternal gametic effects, respectively. Finally, $\mathbf{p}$ and $\mathbf{h}$ are vectors of random permanent maternal (9224 levels) and herd-year-season (2781 levels) environmental effects (not included in the simulated datasets), $\mathbf{e}$ is a vector of residuals, and $\mathbf{X}$, $\mathbf{Z_a}$, $\mathbf{Z_m}$, $\mathbf{Z_s}$, $\mathbf{Z_d}$, $\mathbf{Z_p}$, and $\mathbf{Z_h}$ are incidence matrices that link the effects with the phenotypic data. Variances and covariances for random sources of variation were defined as follows:

$$\text{var} \begin{bmatrix} \mathbf{a}_o \\ \mathbf{a}_m \\ \mathbf{s} \\ \mathbf{d} \\ \mathbf{p} \\ \mathbf{h} \\ \mathbf{e} \end{bmatrix} = \begin{bmatrix} \mathbf{A}\sigma_a^2 & \mathbf{A}\sigma_{am} & 0 & 0 & 0 & 0 & 0 \\ \mathbf{A}\sigma_{am} & \mathbf{A}\sigma_m^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \mathbf{G}\sigma_s^2 & \mathbf{G}\sigma_{sd} & 0 & 0 & 0 \\ 0 & 0 & \mathbf{G}\sigma_{sd} & \mathbf{G}\sigma_d^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I}\sigma_p^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{I}\sigma_h^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \mathbf{I}\sigma_e^2 \end{bmatrix},$$

where $\sigma_p^2$ and $\sigma_h^2$ represent the variances of the permanent maternal and herd-year-season effects, $\sigma_d^2$ is the

variance of maternal gametic effects, and $\sigma_{sd}$ is the covariance between paternal and maternal gametic effects. All other parameters were defined previously. Note that under a pure direct additive model:

$$\sigma_s^2 = \sigma_d^2 = \sigma_{sd} = \frac{1}{2}\sigma_a^2. \tag{2}$$

A number of models based on reduced parameterizations of this full model were fitted to both the simulated and the real data:

Model A: $\mathbf{y} = \mathbf{Xb} + \mathbf{Z_a a_o} + \mathbf{e},$
Model S: $\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s s} + \mathbf{e},$
Model AS: $\mathbf{y} = \mathbf{Xb} + \mathbf{Z_a a_o} + \mathbf{Z_s s} + \mathbf{e},$
Model SD: $\mathbf{y} = \mathbf{Xb} + \mathbf{Z_s s} + \mathbf{Z_d d} + \mathbf{e},$
Model AM: $\mathbf{y} = \mathbf{Xb} + \mathbf{Z_a a_o} + \mathbf{Z_m a_m} + \mathbf{e},$
Model AMS: $\mathbf{y} = \mathbf{Xb} + \mathbf{Z_a a_o} + \mathbf{Z_m a_m} + \mathbf{Z_s s} + \mathbf{e},$

When the Pirenaica beef data was used, all models also included the terms $\mathbf{z_p p}$ and $\mathbf{z_h h}$.

For statistical analysis, we applied a Bayesian approach through a Gibbs sampler [26]. The prior distributions for systematic effects and variance components were assumed to be bounded uniform, and the prior distributions for the genetic effects were multivariate Gaussian distributions with mean zero and variance as defined in the previous section. Analyses were run for each simulated or real dataset as a single chain of 525 000 cycles with the first 25 000 iterations being discarded. Convergence was checked by visual inspection of the chains and by applying the test of Raftery and Lewis [27]. All samples were stored to calculate summary statistics.

### Model comparison

Models were compared using the pseudo log-marginal probability of the phenotypic data [28]. If we consider the data vector $\mathbf{y} = (y_i, \mathbf{y}_{-i})$, where $y_i$ is the $i^{\text{th}}$ datum and $\mathbf{y}_{-i}$ is the vector of data with $i^{\text{th}}$ datum deleted, the conditional predictive distribution has a probability density equal to:

$$p(y_i|\mathbf{y}_{-i}) = \int p(y_i|\mathbf{y}_{-i})f(\boldsymbol{\theta}|\mathbf{y}_{-i})d\mathbf{y},$$

where $\boldsymbol{\theta}$ is the vector of parameters. Therefore, $p(y_i|\mathbf{y}_{-i})$ can be interpreted as the probability of each datum given the rest of the data and is known as the conditional predictive ordinate (CPO) for the $i^{\text{th}}$ datum. The pseudo log-marginal probability of the data (LogCPO) is then:

$$\sum_i \ln p(y_i|\mathbf{y}_{-i}).$$

The collection of conditional predictive densities is equivalent to the marginal probability of the data [29]. A

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 5 of 13

Monte Carlo approximation of the LogCPO suggested by Gelfand [28] is:

$$\sum_i \ln \hat{p}(y_i|\mathbf{y}_{-i}),$$

where $\hat{p}(y_i|\mathbf{y}_{-i}) = N\left[\sum_{j=1}^{N} \frac{1}{p(y_i|\theta^j)}\right]^{-1},$

and $N$ is the number of Markov chain Monte Carlo (McMC) draws, and $\theta^j$ is the $j$th draw from the posterior distribution of the corresponding parameter. The interpretation of the results of this test is that the larger is LogCPO, the higher is the marginal probability of data and, thus, the better is the relative fit.

## Results

### Scenario 1: additive model of inheritance

In Scenario 1, all genetic variation was generated as direct additive; there were no maternal genetic or paternally inherited effects. The average posterior mean estimates of variance components and ratios of components of variance are in Table 1. When Model A was used, the estimate of the direct additive genetic variance almost matched the simulated value. Furthermore, when the estimation model only included paternal gametic effects (Model S), a significant proportion of the variation was assigned to these effects and the remaining additive variation was incorporated into the residual variance. Implementation of a more complex parameterization, such as Model AS, attributed a small, although negligible, amount of the simulated variability to the paternally inherited effects. As expected, Model SD absorbed close to half of the additive genetic variance of the paternal gametic effects and another half of the maternal effects. In addition, the correlation between paternal and

maternal effects was close to 1. Finally, the models that included an additive maternal genetic effect (AM and AMS) also correctly assigned most genetic variation to the direct additive effects, and only small amounts to other variance components. The results of the model comparison test showed that Models A and SD presented the best fit, followed closely by Models AM, AS and AMS. Model S had the worst fit.

### Scenario 2: paternal inheritance model

Scenario 2 (Table 2) only included a variance of paternally inherited effects of 250, without considering any other source of genetic variance. The average posterior mean estimate of variance due to paternally inherited effects under Model S was close to the simulated value. However, with Model A, estimates of direct additive genetic variance were on average close to 150, although this variance component was not included in the simulation process. Model AS assigned most of the variance to paternally inherited effects and a small amount to the direct additive effect. Model SD assigned most of the genetic variance to paternal gametic effects, although negligible amounts of variability were also assigned to the gametic maternal variance component. Model AM, which corresponds to the standard maternal animal model, strongly overestimated the direct and maternal additive genetic variances, and resulted in a negative estimate of the covariance between direct and maternal additive genetic effects, which resulted in a negative correlation close to −1. This tendency was reduced when paternally inherited effects were included in the model (Model AMS). Average posterior mean estimates of variances of both direct and maternal additive genetic effects and the covariance between them were smaller than in Model AM. However, the genetic correlation was still negative (−0.90) and a substantial percentage of variability was assigned to paternally inherited effects. It should be noted that the residual

**Table 1** Averages of posterior means (and standard deviations) of (co)variance components for Scenario 1 ($\sigma_a^2 = 500$, $\sigma_e^2 = 1000$)

| Parameter | Model of estimation | | | | | |
|---|---|---|---|---|---|---|
| | A | S | AS | SD | AM | AMS |
| $\sigma_a^2$ | 514.8 (33.2) | - | 498.7 (34.8) | - | 512.5 (43.0) | 469.6 (51.5) |
| $\sigma_m^2$ | - | - | - | - | 28.6 (9.6) | 25.5 (8.9) |
| $\sigma_{am}$ | - | - | - | - | −14.9 (18.65) | 0.88 (19.5) |
| $\sigma_s^2$ | - | 292.4 (27.5) | 27.4 (16.3) | 239.7 (24.2) | - | 30.4 (19.2) |
| $\sigma_d^2$ | - | - | - | 284.7 (25.2) | - | . |
| $\sigma_{sd}$ | - | - | - | 239.8 (21.9) | - | - |
| $\sigma_e^2$ | 1010.8 (26.4) | 1212.8 (26.5) | 1001.8 (26.8) | 1001.3 (29.89) | 1003.7 (29.0) | 1005.8 (29.0) |
| L | −50209.08 | −50517.54 | −50210.99 | −50208.96 | −50213.5 | −50215.8 |

$\sigma_a^2$ and $\sigma_m^2$ are the direct and maternal additive genetic variances and $\sigma_{am}$ is the direct-maternal genetic covariance. $\sigma_s^2$ and $\sigma_d^2$ are the paternal and maternal gametic variances and $\sigma_{sd}$ is the covariance between the paternal and maternal gametic effects. Finally, $\sigma_e^2$ is the residual variance and L is the average logCPO across 10 replicates.

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 6 of 13

**Table 2** Averages of posterior means (and posterior standard deviations) of (co)variance components for Scenario 2 ($\sigma_s^2 = 250$, $\sigma_e^2 = 1000$)

| Parameter | Model of estimation | | | | | |
|---|---|---|---|---|---|---|
| | A | S | AS | SD | AM | AMS |
| $\sigma_a^2$ | 147.5(18.7) | - | 21.3(13.6) | - | 372.2 (43.8) | 111.4 (50.7) |
| $\sigma_m^2$ | - | - | - | - | 105.7 (21.8) | 40.8 (19.0) |
| $\sigma_{am}$ | - | - | - | - | −191.9 (16.6) | −60.9 (29.8) |
| $\sigma_s^2$ | - | 242.0 (31.5) | 208.1(22.9) | 218.6(16.6) | - | 162.0 (30.2) |
| $\sigma_d^2$ | - | - | - | 14.6(9.8) | - | - |
| $\sigma_{sd}$ | - | - | - | 1.7(16.6) | - | - |
| $\sigma_e^2$ | 1101.6(21.9) | 990.3(31.1) | 1019.7(23.1) | 1016.1(23.6) | 960.8 (30.1) | 997.3 (26.9) |
| L | −49730.0 | −49609.5 | −49612.4 | −49614.0 | −49613.2 | −49613.3 |

$\sigma_a^2$ and $\sigma_m^2$ are the direct and maternal additive genetic variances and $\sigma_{am}$ is the direct-maternal genetic covariance. $\sigma_s^2$ and $\sigma_d^2$ are the paternal and maternal gametic variances and $\sigma_{sd}$ is the covariance between the paternal and maternal gametic effects. Finally, $\sigma_e^2$ is the residual variance and L is the average logCPO across 10 replicates.

variance was inflated in Model A and smaller than the simulated value in Model AM. The model with the best fit was Model S, which mimicked the simulation model, and was closely followed by Models AS and SD, which included paternally inherited effects, and by Model AM. Model A had the worst fit.

### Scenario 3: additive and paternal inheritance model

Scenario 3 included an additive genetic variance component of 500 and a paternally inherited component with variance equal to 250. The results are in Table 3. The first and second models of estimation (Models A and S) clearly overestimated the additive genetic variance and the paternal gametic variance, respectively. This overestimation disappeared with Model AS, which replicated the model of simulation by including a paternal gametic effect, and the estimates were close to the simulated values. Models AM and AMS overestimated the additive genetic variance

and assigned some of the variance to the maternal genetic effects. Furthermore, these models generated a negative covariance between additive and maternal genetic effects. This overestimation of variances was reduced with Model AMS, for which a substantial amount of the variability was assigned to the paternally inherited effects, although less than the simulated value. The model comparison test indicated that Models AS, SD, AM and AMS had the best fit, with similar average LogCPO values, whereas Models A and S had a substantially poorer fit.

### Scenario 4: maternal animal model

Scenario 4 corresponds to the standard maternal animal model. Direct and maternal additive genetic variances of 500 and 250 were generated, but with a strong negative correlation between these two effects (−0.707). Average posterior mean estimates are in Table 4. As expected, the average posterior mean estimates from Model AM

**Table 3** Averages of posterior means (and posterior standard deviations) of (co)variance components for Scenario 3 ($\sigma_a^2 = 500$, $\sigma_s^2 = 250$, $\sigma_e^2 = 1000$)

| Parameter | Model of estimation | | | | | |
|---|---|---|---|---|---|---|
| | A | S | AS | SD | AM | AMS |
| $\sigma_a^2$ | 672.1 (37.6) | - | 536.7 (39.8) | - | 872.0 (60.0) | 692.4 (97.8) |
| $\sigma_m^2$ | - | - | - | - | 96.1 (22.8) | 57.7 (21.1) |
| $\sigma_{am}$ | - | - | - | - | −181.1 (35.7) | −94.5 (44.6) |
| $\sigma_s^2$ | - | 566.3 (40.5) | 201.4 (38.0) | 475.5 (35.7) | - | 116.8 (60.4) |
| $\sigma_d^2$ | - | - | - | 256.0 (26.2) | - | - |
| $\sigma_{sd}$ | - | - | - | 281.3 (30.8) | - | - |
| $\sigma_e^2$ | 1091.0 (28.8) | 1185.7 (37.0) | 1025.8 (32.6) | 1034.5 (35.1) | 972.9 (37.0) | 992.4 (35.4) |
| L | −50857.1 | −51040.3 | −50811.7 | −50810.5 | −50813.5 | −50812.9 |

$\sigma_a^2$ and $\sigma_m^2$ are the direct and maternal additive genetic variances and $\sigma_{am}$ is the direct-maternal genetic covariance. $\sigma_s^2$ and $\sigma_d^2$ are the paternal and maternal gametic variances and $\sigma_{sd}$ is the covariance between the paternal and maternal gametic effects. Finally, $\sigma_e^2$ is the residual variance and L is the average logCPO across10 replicates.

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 7 of 13

**Table 4** Averages of posterior means (and posterior standard deviations) of (co)variance components for Scenario 4 ($\sigma_a^2 = 500$, $\sigma_m^2 = 250$, $\sigma_{am} = -250$, $\sigma_e^2 = 1000$)

| Parameter | Model of estimation | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | A | S | AS | SD | AM | AMS |
| $\sigma_a^2$ | 348.0 (28.5) | - | 264.0 (24.2) | - | 523.4 (47.0) | 367.5 (76.5) |
| $\sigma_m^2$ | - | - | - | - | 268.8 (30.2) | 228.8 (38.2) |
| $\sigma_{am}$ | - | - | - | - | −266.3 (33.1) | −187.1 (49.4) |
| $\sigma_s^2$ | - | 282.6 (25.7) | 173.3 (29.2) | 284.5 (29.8) | - | 92.9 (41.3) |
| $\sigma_d^2$ | - | - | - | 237.4 (25.3) | - | - |
| $\sigma_{sd}$ | - | - | - | 0.3 (25.3) | - | - |
| $\sigma_e^2$ | 1159.0 (21.1) | 1224.6 (24.4) | 1077.7 (25.3) | 988.6 (24.9) | 986.1 (24.6) | 1012.0 (23.4) |
| L | −50427.0 | −50375.2 | −50329.7 | −50318.3 | −50318.7 | −50494.3 |

$\sigma_a^2$ and $\sigma_m^2$ are the direct and maternal additive genetic variances and $\sigma_{am}$ is the direct-maternal genetic covariance. $\sigma_s^2$ and $\sigma_d^2$ are the paternal and maternal gametic variances and $\sigma_{sd}$ is the covariance between the paternal and maternal gametic effects. Finally, $\sigma_e^2$ is the residual variance and L is the average logCPO across 10 replicates.

almost matched the simulated values for the direct and maternal additive genetic effects, and the covariance between them. Moreover, all models except AM gave biased estimates of variance components. In particular, Model A resulted in an average posterior mean estimate of the direct additive genetic variance that was smaller than the simulated value. Model S assigned a significant proportion of the variation to paternally inherited effects, although they were not included in the simulation model. This was confirmed with Model AS, which assigned a significant proportion of the variation to the paternally inherited effects. Model SD assigned a significant proportion of the variance to both paternal and maternal gametic effects, but the covariance between them was almost null. Finally, with Model AMS, the presence of a significant amount of variation in the paternal gametic effects reduced the magnitude of the direct and maternal additive genetic variance components compared to the results obtained with Model AM. For this scenario, the models with the best fit were SD and AM, followed by Models AS and S. The models that fitted the simulated data worst were A and AMS.

### Scenario 5: direct, maternal and paternal inheritance model

Scenario 5 included direct and maternal additive genetic, and paternal gametic effects. Direct and maternal additive genetic effects with variances of 500 and 250 were generated, with a strong negative correlation between them (−0.707). The paternally inherited variance was 250. The average posterior mean estimates are in Table 5. The average results of Model AMS were very close to the simulated values. However, the standard deviation of the estimates was substantially higher than for the other models and also than for the previous scenarios. The results of Model A were very close to the simulated value

for the direct additive genetic effect (510.8 vs 500), although residual variance was significantly overestimated. Model S, AS and SD overestimated the paternal gametic variance, which captured the sources of variation that were ignored in the estimation model. It is remarkable that Model SD provided almost a null covariance between sire and dam gametic effects. Finally, the results of Model AM reflected a redistribution of the ignored paternally inherited variance by increasing the direct and maternal additive genetic variances (to 924.6 and 371.0, respectively), and by generating a strong negative direct-maternal genetic covariance (−473.5). The results of the LogCPO test indicated that Model AMS, which replicated the simulation model, had the best fit. It was closely followed by Models AM and SD, whereas Models AS, S and A fitted the simulated data worst.

### Pirenaica beef cattle breed data

The results of the analysis using real data on the Pirenaica beef cattle breed are in Table 6. With Model A, the posterior mean estimate of the direct heritability ($h_d^2$) was equal to 0.367, but with Model AS, it was only equal to 0.174, whereas the posterior mean estimate of paternal gametic heritability was 0.379, which increased to 0.505 with Model S that did not take direct additive genetic effects into account. Model SD returned the highest posterior mean estimate of the paternal heritability (0.463), a lower value for maternal heritability (0.130) and a very low estimate of the direct-maternal genetic correlation (0.059). The standard maternal animal model (AM) produced a very large posterior estimate of direct additive heritability (0.471), along with a smaller posterior estimate of maternal heritability (0.157), and the correlation between direct and maternal effects was strong and negative (−0.734). Finally, with Model AMS, both heritabilities fell to very low values

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 8 of 13

**Table 5** Averages of posterior means (and posterior standard deviations) of (co)variance components for Scenario 5 ($\sigma_a^2 = 500$, $\sigma_m^2 = 250$, $\sigma_{am} = -250$, $\sigma_s^2 = 250$, $\sigma_e^2 = 1000$)

| Parameter | Model of estimation | | | | | |
|---|---|---|---|---|---|---|
| | A | S | AS | SD | AM | AMS |
| $\sigma_a^2$ | 510.82 (30.39) | - | 290.21 (24.10) | - | 924.60 (50.17) | 569.54 (121.68) |
| $\sigma_m^2$ | - | - | - | - | 370.97 (26.72) | 279.33 (34.50) |
| $\sigma_{am}$ | - | - | - | - | −473.53 (29.09) | −292.76 (58.38) |
| $\sigma_s^2$ | - | 551.14 (37.39) | 414.96 (36.31) | 544.03 (35.98) | - | 223.57 (77.09) |
| $\sigma_d^2$ | - | - | - | 236.24 (19.56) | - | - |
| $\sigma_{sd}$ | - | - | - | −5.05 (28.14) | - | - |
| $\sigma_e^2$ | 1236.37 (24.73) | 1207.52 (29.87) | 1056.10 (31.02) | 973.73 (33.36) | 923.52 (32.99) | 973.03 (32.86) |
| L | −51048.4 | −51010.0 | −50902.3 | −50868.6 | −50860.6 | −50858.1 |

$\sigma_a^2$ and $\sigma_m^2$ are the direct and maternal additive genetic variances and $\sigma_{am}$ is the direct-maternal genetic covariance. $\sigma_s^2$ and $\sigma_d^2$ are the paternal and maternal gametic variances and $\sigma_{sd}$ is the covariance between the paternal and maternal gametic effects. Finally, $\sigma_e^2$ is the residual variance and L is the average logCPO across 10 replicates

(0.045 and 0.056) and a significant proportion of the genetic variation was assigned to the paternal gametic effects (0.436). The results of the comparison of models based on LogCPO showed that Model SD had the best predictive ability, followed by Models AM and ASs. Model AMS was next, and Models S and A showed the worst fit.

## Discussion

We present a theoretical development that shows that the structure of the (co) variance components assumed by the standard maternal animal model is partially confounded with a model that considers a paternal gametic effect. Meyer and Tier [22] postulated that this may be one

**Table 6** Posterior means (and posterior standard deviations) of (co)variance components for weight at 210 days in the Pirenaica beef cattle population

| Parameter | Model of estimation | | | | | |
|---|---|---|---|---|---|---|
| | A | S | AS | SD | AM | AMS |
| $\sigma_a^2$ | 798.3 (52.0) | - | 418.8 (60.7) | - | 1317.0 (93.2) | 94.0 (35.9) |
| $\sigma_m^2$ | - | - | - | - | 426.7 (54.1) | 139.3 (30.1) |
| $\sigma_{am}$ | - | - | - | - | −553.9 (64.3) | 31.2 (26.5) |
| $\sigma_s^2$ | - | 1236.2 (93.4) | 891.3 (100.5) | 1100.0 (96.1) | - | 1029.8 (90.0) |
| $\sigma_d^2$ | - | - | - | 326.3 (43.2) | - | . |
| $\sigma_{sd}$ | - | - | - | 30.4 (53.4) | - | - |
| $\sigma_h^2$ | 534.4 (24.6) | 477.6 (23.0) | 454.4 (22.1) | 446.5 (22.2) | 485.7 (23.4) | 443.2 (22.0) |
| $\sigma_p^2$ | 94.5 (17.2) | 191.2 (17.2) | 109.2 (18.3) | 69.1 (20.1) | 52.8 (22.9) | 47.9 (22.9) |
| $\sigma_e^2$ | 747.7 (32.4) | 539.4 (48.4) | 501.9 (43.9) | 445.5 (53.9) | 492.3 (46.9) | 591.1 (47.5) |
| $h_a^2$ | 0.367 (0.021) | - | 0.176 (0.026) | - | 0.474 (0.023) | 0.040 (0.015) |
| $h_m^2$ | - | - | - | - | 0.153 (0.016) | 0.059 (0.012) |
| $r_{am}$ | - | - | - | - | −0.739 (0.016) | 0.311 (0.187) |
| $h_s^2$ | - | 0.505 (0.030) | 0.375 (0.037) | 0.460 (0.032) | - | 0.439 (0.032) |
| $h_d^2$ | - | - | - | 0.137 (0.018) | - | - |
| $r_{sd}$ | - | - | - | 0.053 (0.091) | - | - |
| $c_h^2$ | 0.246 (0.010) | 0.195 (0.010) | 0.191 (0.009) | 0.187 (0.010) | 0.175 (0.009) | 0.189 (0.010) |
| $c_p^2$ | 0.043 (0.008) | 0.078 (0.007) | 0.046 (0.008) | 0.029 (0.008) | 0.019 (0.008) | 0.020 (0.010) |
| LogCPO | −86767 | −86703 | −86513 | −86178 | −86378 | −86647 |

$\sigma_a^2$ and $\sigma_m^2$ are the direct and maternal additive genetic variances and $\sigma_{am}$ is the direct-maternal genetic covariance. $\sigma_s^2$ and $\sigma_d^2$ are the paternal and maternal gametic variances and $\sigma_{sd}$ is the covariance between the paternal and maternal gametic effects. $\sigma_h^2$ and $\sigma_p^2$ are the herd and maternal permanent environmental variances. $\sigma_e^2$ is the residual variance. $h_a^2$, $h_m^2$, $h_s^2$ and $h_d^2$ are the ratios of the direct additive genetic, maternal additive genetic, paternal gametic and maternal gametic variances to the phenotypic variance, respectively. $r_{am}$ is the genetic correlation between direct and maternal additive effects and $r_{sd}$ is the correlation between paternal and maternal gametic effects. $c_h^2$ and $c_p^2$ are the ratios of herd and permanent maternal environmental effects to the phenotypic variance.

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 9 of 13

possible cause for the strong negative estimates of the genetic correlation between direct and maternal genetic effects that are typically obtained. We derived results that show that, for most of the genetic relationships in livestock pedigrees, the covariance between relatives generated by a paternally inherited effect is mimicked by the maternal animal model when variance components follow the pattern defined by Equation (1). Thus, when paternally inherited effects are ignored in the estimation model, its variance is assigned to the remaining variance components in order to accommodate the observed correlations between relatives. Under the maternal animal model, its consequences are an inflation of the direct and maternal additive genetic variances, and a downward bias of the direct-maternal genetic correlation.

These results were confirmed by the simulation study. When simulation and estimation models were analogous, the Bayesian analysis was able to provide suitable estimates of the variance components. Furthermore, when the true simulation structure of the (co)variance components was included in the estimation model together with other variance components, the procedure correctly estimated the true parameters, whereas estimates of the non-simulated variance components were small. However, the presence of paternal gametic effects (Scenarios 2 and 3) and their absence in the estimation model resulted in an overestimation of the direct additive genetic variance, as previously described by other authors [17, 22]. Furthermore, when paternally inherited effects that play a significant role in the genetic variation were ignored in a model that included correlated direct and maternal additive genetic effects, the direct additive genetic variance component was increased by almost twice the paternal gametic variance, whereas the genetic covariance between direct and maternal effects became negative and the estimate of the maternal additive genetic variance increased by almost one half of the simulated paternal gametic variance, as was observed in Scenarios 2, 3 and 5. As described above, this redistribution of the gametic paternal variance is linked to the constraint imposed in Equation (1). These results were also consistent with the presence of a significant percentage of paternal gametic variance for Models S, AS, SD and AMS in Scenario 4, where only direct and maternal additive effects were simulated, and with the increase of paternal gametic variance in Models S, AS and SD in Scenario 5. Finally, the decrease in covariance between paternal and maternal effects under Model SD in Scenarios 4 and 5 may also be attributed to this phenomenon.

More generally, when any relevant variance component in the simulation model was not considered in the estimation, substantial biases on the variance component estimation were noted. For example, ignoring a true direct additive genetic effect while assuming paternal and maternal gametic effects in the estimation model, led to the absorption of close to half of the additive genetic variance by the paternal gametic variance and half by the maternal gametic variance. In addition, the increase in the estimate of the covariance between paternal and maternal gametic effects was also of this magnitude. This distribution of the direct additive variance followed the pattern described in Equation (2) and was exploited in the model proposed by Neugenbauer et al. [18, 19] in order to detect paternal and maternal inherited effects.

The model comparison test that we performed provided support to our main argument. To illustrate this point, in Scenarios 2 and 3, where paternally inherited effects were included in the simulation model, the goodness of fit of Model AM was very close to that obtained with Models S and AS, which mimicked the respective simulation models. Moreover, under Scenarios 4 and 5, Model SD performed almost as well as Models AM or AMS, which replicated the true simulation models. Clearly, the constraint imposed by Equation (1) explains the similarity observed between the maternal animal model and the paternally inherited model in those cases.

It should be emphasized that the constraint imposed by Equation (1) is not a pure parametric equivalence, since it is limited by the condition that available correlations between relatives do not involve genealogical relationships where one individual is an ancestor of the other within the paternal lineage. Therefore, the ability to separate direct and maternal additive genetic effects and sire imprinted effects can only be achieved if relevant records from sires and their offspring are available or, more generally, if there exist records from individuals that are genetically linked through the paternal pathway. However, in beef cattle datasets, phenotypes on sire-offspring pairs are usually rare. In the Pirenaica beef cattle dataset analyzed here, there were 527 463 paternal half-sibs, 15 128 maternal half-sibs, 2707 full-sibs, 2411 dam-progeny and only 4566 sire-progeny pairs with phenotypic information on both individuals. Thus, most of the information available for estimation of variance components was provided by half-sib relatives, whose covariance generated by paternally inherited effects can be also achieved by the maternal animal model under the constraint imposed by Equation (1).

This latter explanation helps to understand the results of the analysis of the Pirenaica beef cattle dataset. Estimates of variance components obtained from Model A were concordant with previous results obtained in the same [30] and other beef cattle populations [31, 32]. In turn, alternative estimation models, such as S, AS and SD, suggested that paternal gametic effects may play an important role in the variation of weight at 210 days in this breed. In contrast, under Model AM, maternal additive effects were also very relevant. Furthermore, when

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 10 of 13

both effects were fitted in the estimation model (AMS), the maternal effects contributed less to the total variation than the paternal gametic effects. A similar reduction in the maternal contribution on weaning weight was also found in the Bruna dels Pirineus beef cattle breed when paternal gametic effects were considered in the estimation model [33]. Indeed, more evidence for the presence of parent-of-origin variation is derived from the fact that when Model SD was used, a very low correlation between paternal and maternal gametic effects was obtained. This is in contrast to the results reported by Neugenbauer *et al.* [18, 19], which provided a high correlation between them, as in the outcome of the first simulation scenario, where a pure additive model was simulated. However, as observed with Scenarios 4 and 5, independence between paternal and maternal gametic effects, together with the presence of a significant paternally inherited variance component can also be due to the existence of a strong and negative genetic correlation between direct and maternal genetic effects. Given the structure of the information in the Pirenaica dataset, the genetic variance is distributed between paternal gametic and direct and maternal effects depending on the model of estimation and, thus, it is very difficult to clearly identify the true sources of variation. However, at this point, it is worth mentioning that the results of the simulation study using the Pirenaica's breed genealogical structure were similar to those obtained with the above described simulation study (results in Tables S1, S2, S3, S4 and S5 [See Additional file 1: Tables S1-S5]). This allowed us to rule out the possibility of a statistical artefact caused by the structure of the information. Nevertheless, the model comparison using logCPO indicated that Model SD was most suited to the Pirenaica beef cattle data. It should be noted that this model performed well in all simulation scenarios (Tables 1, 2, 3, 4 and 5), since it can accommodate a significant amount of paternal inheritance and also a strong negative correlation between direct or maternal effects, as in Scenario 4. In fact, Model AM had the second best predictive ability based on the LogCPO approach and showed a better fit than Model AS. However, the very large difference in LogCPO between Models SD and AM (−86178 vs −86378) and the results of Model AMS suggest that the existence of significant paternal gametic effects is plausible in the Pirenaica population.

Many studies have reported negative estimates of the genetic covariance between direct and maternal additive effects for beef cattle populations [4–6]. Several explanations have been proposed, including environmental correlations between progeny performance and maternal environmental effects [34, 35] and sire × herd interactions [36, 37]. An additional explanation that should be considered is the presence of paternally inherited effects,

as previously suggested by Meyer and Tier [22]. Our evidence is consistent with the effect associated with imprinted genes, such as the *insulin-like growth factor 2* (*IGF2*) gene involved in fetal growth [38, 39] or postnatal growth and fat deposition in pigs [11, 12, 40] and mice [41]. Genomic imprinting of the *IGF2* gene has also been shown to affect meat quality traits in beef cattle [42, 43] and some mutations in *IGF2* have been associated with carcass traits [44] and body weight [45].

Furthermore, a wealth of evidence on the role of parental inheritance in fetal and early growth in mammals supports the imprinting hypothesis [46, 47]. This is related with parent-offspring conflict theories about the evolutionary origin of paternal effects [48], which are based on the hypothesis that the self-interest of the paternal genome is to increase the growth of the fetus or young individuals, whereas the aim of the maternal genome is to sufficiently limit growth to ensure the success of future offspring without jeopardizing the mother's survival. According to this hypothesis, genes that enhance growth tend to be more expressed if they are paternally inherited than when they are maternally inherited. This phenomenon should be even more marked in species with a low rate or absence of full siblings and a high frequency of multiple paternity [49], as is the case for cattle.

Finally, other causes that can explain overestimation of paternal gametic variance are a large difference in allele frequencies between sires and dams or the presence of significant dominance variation. However, these are not relevant for close populations and that include few full-sib relationships, respectively. In future studies, a very promising alternative to distinguish between paternally and maternally inherited effects arises from the use of molecular information that may contribute to better discriminate between paternal and maternal alleles.

## Conclusions
Ignoring paternally inherited effects in genetic evaluations with maternal animal models can lead to a specific pattern of bias in the estimates of variance components. This phenomenon may account for the unexpectedly strong negative direct-maternal genetic correlations that are reported in the literature. In addition, we showed that the existence of paternal effects may play a role in the explanation of the partitioning of genetic variation in growth traits and possibly in some other relevant traits in livestock production. The main implication of this study is that the presence of relevant paternally inherited variation should be investigated when the maternal animal model provided negative correlation estimates between direct and maternal effects. In addition, in those cases, applying maternal animal models should be done with caution.

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 11 of 13

## Appendix

Consider a model that fits direct and maternal additive genetic effects. As pointed out in the Methods section, the phenotypic values of any two individuals, say $A$ and $B$, are modelled as:

$$y_A = \mu + a_{oA} + a_{mA_D} + e_A$$
$$y_B = \mu + a_{oB} + a_{mB_D} + e_B$$

Assuming independence of the residuals, the covariance between both phenotypic values can be described as a function of the following four pairwise covariances:

$$\text{cov}(y_A, y_B) = \text{cov}(a_{oA}, a_{oB}) + \text{cov}(a_{mA_D}, a_{mB_D}) + \text{cov}(a_{mA_D}, a_{oB}) + \text{cov}(a_{mA_D}, a_{mB_D}).$$

Now, we will use the standard regression of an individual's breeding value on half the breeding values of its parents plus a residual due to Mendelian segregation [49], i.e.

$$a_{oX} = \frac{1}{2}a_{oX_S} + \frac{1}{2}a_{oX_D} + \phi_{oX},$$

to expand all the covariance terms to the parental generation. Replacing first the direct breeding value of individual $B$ and working out the resulting expression, the first term in the right hand side can be expressed as:

$$\text{cov}(a_{oA}, a_{oB}) = \frac{1}{2}\text{cov}(a_{oA}, a_{oB_S}) + \frac{1}{2}\text{cov}(a_{oA}, a_{oB_D}) + \text{cov}(a_{oA}, \phi_{oB}).$$

A further step by replacing the breeding value of individual $A$, results in:

$$\text{cov}(a_{oA}, a_{oB}) = \frac{1}{4}\text{cov}(a_{oA_S}, a_{oB_S}) + \frac{1}{4}\text{cov}(a_{oA_D}, a_{oB_S})$$
$$+ \frac{1}{2}\text{cov}(\phi_{oA}, a_{oB_S}) + \frac{1}{4}\text{cov}(a_{oA_S}, a_{oB_D})$$
$$+ \frac{1}{4}\text{cov}(a_{oA_D}, a_{oB_D}) + \frac{1}{2}\text{cov}(\phi_{oA}, a_{oB_D})$$
$$+ \frac{1}{2}\text{cov}(a_{oA_S}, \phi_{oB}) + \frac{1}{2}\text{cov}(a_{oA_D}, \phi_{oB}) + \text{cov}(\phi_{oA}, \phi_B)$$

Similarly,

$$\text{cov}(a_{oA}, a_{mB_D}) = \frac{1}{2}\text{cov}(a_{oA_S}, a_{mB_D}) + \frac{1}{2}\text{cov}(a_{oA_D}, a_{mB_D}) + \text{cov}(\phi_{oA_S}, a_{mB_D}),$$

and

$$\text{cov}(a_{mA_D}, a_{oB}) = \frac{1}{2}\text{cov}(a_{mA_D}, a_{oB_S}) + \frac{1}{2}\text{cov}(a_{mA_D}, a_{oB_D}) + \text{cov}(a_{mA_D}, \phi_{oB}).$$

Collecting all these results,

$$\text{cov}(y_A, y_B) = \frac{1}{4}\text{cov}(a_{oA_S}, a_{oB_S}) + \frac{1}{4}\text{cov}(a_{oA_D}, a_{oB_S})$$
$$+ \frac{1}{2}\text{cov}(\phi_{oA}, a_{oB_S}) + \frac{1}{4}\text{cov}(a_{oA_S}, a_{oB_D})$$
$$+ \frac{1}{4}\text{cov}(a_{oA_D}, a_{oB_D}) + \frac{1}{2}\text{cov}(\phi_{oA}, a_{oB_D})$$
$$+ \frac{1}{2}\text{cov}(a_{oA_S}, \phi_{oB}) + \frac{1}{2}\text{cov}(a_{oA_D}, \phi_{oB})$$
$$+ \text{cov}(\phi_{oA}, \phi_{oB}) + \frac{1}{2}\text{cov}(a_{oA_S}, a_{mB_D})$$
$$+ \frac{1}{2}\text{cov}(a_{oA_D}, a_{mB_D}) + \text{cov}(\phi_{oA}, a_{mB_D})$$
$$+ \frac{1}{2}\text{cov}(a_{mA_D}, a_{oB_S}) + \frac{1}{2}\text{cov}(a_{mA_D}, a_{oB_D})$$
$$+ \text{cov}(a_{mA_D}, \phi_{oB}) + \text{cov}(a_{mA_D}, a_{mB_D})$$

Now, recalling that:

$$\text{cov}(a_{oX}, a_{oY}) = r(X, Y)\sigma_a^2,$$
$$\text{cov}(a_{oX}, a_{mY}) = r(X, Y_D)\sigma_{dm},$$
$$\text{cov}(a_{mX}, a_{mY}) = r(X_D, Y_D)\sigma_m^2,$$

where $r(X,Y)$ stands for the additive relationship between any two individuals $X$ and $Y$, the previous expression can be reformulated as:

$$\text{cov}(y_A, y_B) = \frac{1}{4}r(A_S, B_S)\sigma_d^2 + \frac{1}{4}r(A_D, B_S)\sigma_d^2$$
$$+ \frac{1}{2}\text{cov}(\phi_{oA}, a_{oB_s}) + \frac{1}{4}r(A_S, B_D)\sigma_d^2$$
$$+ \frac{1}{4}r(A_D, B_D)\sigma_d^2 + \frac{1}{2}\text{cov}(\phi_{oA}, a_{oB_D})$$
$$+ \frac{1}{2}\text{cov}(a_{oA_S}, \phi_{oB}) + \frac{1}{2}\text{cov}(a_{oA_D}, , \phi_{oB})$$
$$+ \text{cov}(\phi_{oA}, \phi_{oB}) + \frac{1}{2}r(A_S, B_D)\sigma_{dm}$$
$$+ \frac{1}{2}r(A_D, B_D)\sigma_{dm} + \text{cov}(\phi_{oA}, a_{mB_D})$$
$$+ \frac{1}{2}r(A_D, B_S)\sigma_{dm} + \frac{1}{2}r(A_D, B_D)\sigma_{dm}$$
$$+ \text{cov}(a_{mA_D}, \phi_{oB}) + \frac{1}{2}r(A_D, B_D)\sigma_m^2$$

At this point, we introduce the condition set by Equation (1). Let:

$$\sigma_d^2 = -2\sigma_{dm} = 4\sigma_m^2.$$

Then, after proper cancelations and making use of this identity:

$$\frac{\text{cov}(\phi_{oA}, a_{oB_D})}{\text{cov}(\phi_{oA}, a_{mB_D})} = \frac{\sigma_a^2}{\sigma_{am}},$$

the expression of the covariance between the two phenotypic values is reduced to the following expression:

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 12 of 13

$$\mathrm{cov}(y_A, y_B) = \frac{1}{4}r(A_s, B_s)\sigma_d^2 + \frac{1}{2}\mathrm{cov}(\phi_{oA}, a_{oB_s})$$
$$+ \frac{1}{2}\mathrm{cov}(a_{oA_s}, \phi_{oB}) + \mathrm{cov}(\phi_{oA}, \phi_{oB}).$$

Notice that the $\mathrm{cov}(\phi_{oX}, a_{oY_S}) > 0$ only if $X$ is an ancestor of $Y$ through the paternal line. In turn, the $\mathrm{cov}(\phi_{oX}, \phi_{oY}) > 0$ only if $X = Y$. Thus, for any other two individuals in the population:

$$\mathrm{cov}(y_A, y_B) = \frac{1}{4}r(A_S, B_S)\sigma_d^2.$$

## Additional file

**Additional file 1: Table S1.** (Co) variance components posterior means (and posterior standard deviations) for the first case of simulation ($\sigma_a^2 = 500$, $\sigma_e^2 = 1000$) using the genealogical and phenotypic information of the Pirenaica dataset. $\sigma_a^2$ and $\sigma_m^2$ are the direct and maternal additive genetic variances and $\sigma_{am}$ is the direct-maternal genetic covariance. $\sigma_s^2$ and $\sigma_d^2$ are the paternal and maternal gametic variances and $\sigma_{sd}$ is the covariance between the paternal and maternal gametic effects. $\sigma_h^2$ and $\sigma_p^2$ are the herd and maternal permanent environmental variances. $\sigma_e^2$ is the residual variance. **Table S2.** (Co) variance components posterior means (and posterior standard deviations) for the first case of simulation ($\sigma_s^2 = 250$, $\sigma_e^2 = 1000$) using the genealogical and data structure of the Pirenaica dataset. Description: same as for Table S1. **Table S3.** (Co) variance components posterior means (and posterior standard deviations) for the first case of simulation ($\sigma_a^2 = 500$, $\sigma_s^2 = 250$, $\sigma_e^2 = 1000$) using the genealogical and data structure of the Pirenaica dataset. Description: same as for Table S1. **Table S4.** (Co) variance components posterior means (and posterior standard deviations) for the first case of simulation ($\sigma_a^2 = 500$, $\sigma_m^2 = 250$, $= -250$, $\sigma_e^2 = 1000$) using the genealogical and data structure of the Pirenaica dataset. Description: same as for Table S1. **Table S5.** (Co) variance components posterior means (and posterior standard deviations) for the first case of simulation ($\sigma_a^2 = 500$, $\sigma_m^2 = 250$, $\sigma_{am} = -250$, $\sigma_s^2 = 250$, $\sigma_e^2 = 1000$) using the genealogical and data structure of the Pirenaica dataset. Description: same as for Table S1. (DOCX 168 kb)

## Competing interests
The authors declare that they have no competing interest.

## Authors' contributions
LV derived the mathematical results, conceived the simulation study, developed the software code and drafted the initial manuscript. JA prepared the genealogical and phenotypic datasets of the Pirenaica beef cattle population. SM and JC aided in the preparation of the manuscript. SM, JC, CM and JA helped in the interpretation of the results. All authors read and approved the final manuscript.

## Author details
[1]Unidad de Genética Cuantitativa y Mejora Animal, Universidad de Zaragoza, 50013 Zaragoza, Spain. [2]Instituto Agroalimentario de Aragón (IA2), Universidad de Zaragoza, 50013 Zaragoza, Spain. [3]Departamento de Producción Animal, Facultad de Agronomía, Universidad de Buenos Aires, 1417 Ciudad Autónoma de Buenos Aires, Argentina. [4]Grup de Recerca en Remugants, Departament de Ciència Animal i dels Aliments, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain.

## References
1. Henderson CR. Application of linear models in animal breeding. 3rd ed. Guelph: University of Guelph; 1984.
2. Willham RL. The role of maternal effects in animal breeding. 3. Biometrical aspects of maternal effects in animals. J Anim Sci. 1972;35:1288–93.
3. Meyer K. Variance-components due to direct and maternal effects for growth traits of Australian beef-cattle. Livest Prod Sci. 1992;31:179–204.
4. Koch RM. The role of maternal effects in animal breeding. VI. Maternal effects in beef cattle. J Anim Sci. 1972;35:1316–23.
5. Robinson DL. Estimation and interpretation of direct and maternal genetic parameters for weights of Australian Angus cattle. Livest Prod Sci. 1996;45:1–11.
6. Meyer K. Estimates of genetic parameters for weaning weight of beef cattle accounting for direct-maternal environmental covariances. Livest Prod Sci. 1997;52:187–99.
7. Barlow DP. Competition - A common motif for the imprinting mechanism? EMBO J. 1997;16:6899–905.
8. Reik W, Walker J. Imprinting mechanisms in mammals. Curr Opin Genet Dev. 1998;8:154–64.
9. Magee DA, Spillane C, Berkowicz EW, Sikora KM, MacHugh DE. Imprinted loci in domestic livestock species as epigenomic targets for artificial selection of complex traits. Anim Genet. 2014;45:25–39.
10. Cockett NE, Jackson SP, Shay TL, Farnir S, Berghmans S, Snowder GD, et al. Polar overdominance at the ovine *callipyge* locus. Science. 1996;273:236–8.
11. Georges M, Charlier C, Cockett N. The *callipyge* locus: evidence for the trans interaction of reciprocally imprinted genes. Trends Genet. 2003;19:248–52.
12. Jeon JT, Carlborg O, Törnsten A, Giuffra E, Amarger V, Chardon P, et al. A paternally expressed QTL affecting skeletal and cardiac muscle mass in pigs maps to the *IGF2* locus. Nat Genet. 1999;21:157–8.
13. Nezer C, Moreau L, Brouwers B, Coppieters W, Detilleux J, Hanset R, et al. An imprinted QTL with major effect on muscle mass and fat deposition maps to the IGF2 locus in pigs. Nat Genet. 1999;21:155–6.
14. Imumorin IG, Kim EH, Lee YM, De Koning DJ, van Arendonk J, De Donato M, et al. Genome scan for parent-of-origin QTL effects on bovine growth and carcass traits. Front Genet. 2011;2:44.
15. Morison IM, Ramsay JP, Spencer HG. A census of mammalian imprinting. Trends Genet. 2005;21:457–65.
16. De Vries AG, Kerr R, Tier B, Long T, Meuwissen THE. Gametic imprinting effects on rate and composition of pig growth. Theor Appl Genet. 1994;88:1037–42.
17. Engellandt T, Tier B. Genetic variances due to imprinted genes in cattle. J Anim Breed Genet. 2002;119:154–65.
18. Neugenbauer N, Luther H, Reinsch N. Parent-of-origin effects cause genetic variation in pig performance traits. Animal. 2010;4:672–81.
19. Neugenbauer N, Räder I, Schild HJ, Zimmer D, Reinsch N. Evidence of parent-of-origin effects on genetic variability of beef traits. J Anim Sci. 2010;88:523–32.
20. Tier B, Meyer K. Analysing quantitative parent-of-origin effects with examples from ultrasonic measures of body composition in Australian beef cattle. J Anim Breed Genet. 2012;129:359–68.
21. Hager R, Cheverud JM, Wolf JB. Maternal effects as the cause of parent-of-origin effects that mimic genomic imprinting. Genetics. 2008;178:1755–62.
22. Meyer K, Tier B. Estimates of variances due to parent of origin effects for weights of Australian beef cattle. Anim Prod Sci. 2012;52:215–24.
23. Schaeffer LR, Kennedy BW, Gibson JP. The inverse of the gametic relationship matrix. J Dairy Sci. 1989;72:1266–72.
24. Sánchez A, Ambrona J, Sánchez L. Razas ganaderas españolas bovinas. Madrid: MAPA-FEAGAS; 2002.
25. Beef Improvement Federation. Guidelines for uniform beef improvement programs. Raleigh: North Carolina University Press; 2010.
26. Gelfand AE, Smith AFM. Sampling-based approaches to calculating marginal densities. J Am Statist Assoc. 1990;85:398–409.
27. Raftery AE, Lewis SM. How many iterations in the Gibbs Sampler. In: Bayesian Statistics IV, editor. Edited by Bernardo JM, Berger JO, Dawid AP, Smith AFM. New York: Oxford University Press; 1992. p. 763–74.
28. Gelfand AE. Model determination using sampling-based methods. In: Gilks WR, Richardson S, Spiegelhalter DJ, editors. Markov chain Monte Carlo in practice. London: Chapman Hall; 1996. p. 145–61.

Varona *et al. Genetics Selection Evolution* (2015) 47:63

Page 13 of 13

29. Besag J. Spatial interaction and the statistical analysis of lattice systems (with discussion). J Roy Stat Soc Lond B. 1974;36:192–225.

30. Altarriba J, Yagüe G, Moreno C, Varona L. Exploring the possibilities of genetic improvement from traceability data an example in the Pirenaica beef cattle. Livest Sci. 2009;125:115–20.

31. Utrera AR, Van Vleck LD. Heritability estimates for carcass traits of cattle: a review. Genet Mol Res. 2004;3:380–94.

32. Bouquet A, Fouilloux MN, Renand G, Phocas F. Genetic parameters for growth, muscularity, feed efficiency and carcass traits of young beef bulls. Livest Sci. 2010;129:38–48.

33. Fina M, Varona L, Piedrafita J, Casellas J. Sources of sire-specific genetic variance for birth and weaning weight in Bruna dels Pirineus beef calves. Animal. 2012;6:1931–8.

34. Quintanilla R, Varona L, Pujol MR, Piedrafita J. Maternal animal model with correlation between maternal environmental effects of related dams. J Anim Sci. 1999;77:2904–17.

35. Bijma P. Estimating maternal genetic effects in livestock. J Anim Sci. 2006;84:800–6.

36. Dodenhoff J, Van Vleck LD, Wilson DE. Comparison of models to estimate genetic effects for weaning weight of Angus cattle. J Anim Sci. 1999;77:3176–84.

37. Gutiérrez JP, Fernández I, Álvarez I, Royo LJ, Goyache F. Sire x contemporary group interactions for birth weight and preweaning growth traits in the Asturiana de los Valles beef cattle breed. Livest Sci. 2006;99:61–8.

38. DeChiara TM, Efstratiadis A, Robertson EJ. A growth-deficiency phenotype in heterozygous mice carrying an insulin-like growth factor II gene disrupted by targeting. Nature. 1990;345:78–80.

39. Rotwein P, Hall LJ. Evolution of insulin-like growth factor 2: characterization of the mouse IGF2 gene and identification of two pseudo-exons. DNA Cell Biol. 1990;9:725–35.

40. Van Laere AS, Nguyen M, Braunschweig M, Nezer C, Collette C, Moreau L, et al. A regulatory mutation in *IGF2* causes a major QTL effect on muscle growth in the pig. Nature. 2003;42:832–6.

41. Rogler CE, Yang D, Rossetti L, Donohoe J, Alt E, Chang CJ, et al. Altered body composition and increased frequency of diverse malignancies in insulin-like growth factor-II transgenic mice. J Biol Chem. 1994;269:13779–84.

42. Dindot SV, Kent KC, Evers B, Loskutoff N, Womack J, Piedrahita JA. Conservation of genomic imprinting at the XIST, IGF2 and GTL2 loci in the bovine. Mamm Genome. 2004;15:966–74.

43. Curchoe C, Zhang S, Bin Y, Zhang X, Yang L, Feng D, et al. Promoter-specific expression of the imprinted *IGF2* gene in cattle (*Bos taurus*). Biol Reprod. 2005;73:1275–81.

44. Goodall JJ, Schmutz SM. *IGF2* gene characterization and association with rib eye area in beef cattle. Anim Genet. 2007;38:154–61.

45. Zwierzchowski L, Siadkowska E, Oprzadek J, Flisikowski K, Dymnicki E. An association of C/T polymorphism in exon 2 of the bovine *insulin-like growth factor 2* gene with meat production traits in Polish Holstein-Friesian cattle. Czech J Anim Sci. 2010;55:227–33.

46. Bartolomei MS, Tilghman SM. Genomic imprinting in mammals. Annu Rev Genet. 1997;31:493–525.

47. Tilghman SM. The sins of the fathers and mothers: genomic imprinting in mammalian development. Cell. 1999;96:185–93.

48. Haig D. Genomic imprinting and the theory of parent-offspring conflict. Semin Cell Dev Biol. 1992;3:153–60.

49. Haig D. Multiple paternity and genomic imprinting. Genetics. 1999;151:1229–31.