

DOI: <http://dx.doi.org/10.1590/1678-992X-2017-0351>

ISSN 1678-992X

New insights into genomic selection through population-based non-parametric prediction methods

Leísa Pires Lima¹, Camila Ferreira Azevedo¹*, Marcos Deon Vilela de Resende², Fabyano Fonseca e Silva³, Matheus Massariol Suela⁴, Moysés Nascimento¹, José Marcelo Soriano Viana⁵

¹Universidade Federal de Viçosa – Depto. de Estatística, Av. Peter Henry Rolfs, s/n – 36570-900 – Viçosa, MG – Brasil.

²Embrapa Florestas, Estr. da Ribeira, km 111 – 83411-000 – Colombo, PR – Brasil.

³Universidade Federal de Viçosa – Depto. de Zootecnia.

⁴Universidade Federal de Viçosa – Depto. Fitotecnia.

⁵Universidade Federal de Viçosa – Depto. de Biologia Geral.

*Corresponding author <camila.azevedo@ufv.br>

Edited by: Roberto Fritsche Neto

Received October 14, 2017

Accepted April 07, 2018

ABSTRACT: Genome-wide selection (GWS) is based on a large number of markers widely distributed throughout the genome. Genome-wide selection provides for the estimation of the effect of each molecular marker on the phenotype, thereby allowing for the capture of all genes affecting the quantitative traits of interest. The main statistical tools applied to GWS are based on random regression or dimensionality reduction methods. In this study a new non-parametric method, called Delta-p was proposed, which was then compared to the Genomic Best Linear Unbiased Predictor (G-BLUP) method. Furthermore, a new selection index combining the genetic values obtained by the G-BLUP and Delta-p, named Delta-p/G-BLUP methods, was proposed. The efficiency of the proposed methods was evaluated through both simulation and real studies. The simulated data consisted of eight scenarios comprising a combination of two levels of heritability, two genetic architectures and two dominance status (absence and complete dominance). Each scenario was simulated ten times. All methods were applied to a real dataset of Asian rice (*Oryza sativa*) aiming to increase the efficiency of a current breeding program. The methods were compared as regards accuracy of prediction (simulation data) or predictive ability (real dataset), bias and recovery of the true genomic heritability. The results indicated that the proposed Delta-p/G-BLUP index outperformed the other methods in both prediction accuracy and predictive ability.

Keywords: genomic prediction, selection index, genetic gain, asian rice

Introduction

The evolution of sequencing and genotyping techniques has prompted a great advance in molecular genetics becoming a possibility using DNA information directly for the selection of genetically superior individuals. This framework is called genome-wide selection (GWS), which enables direct early marker-based selection that generally increases the genetic gain per unit of time (Meuwissen et al., 2001). Genome-wide selection is based on the estimation of molecular marker effects on the phenotypes, thus providing accurate additive and dominant genomic value predictions for individuals subject to selection, even for those without observed phenotypic values.

The Genomic Best Linear Unbiased Predictor (G-BLUP) (Goddard, 2009; Habier et al., 2007; Van Raden, 2008) has been extensively applied to GWS and recommended for additive and dominant genomic value predictions (Azevedo et al., 2015; De los Campos et al., 2013; Gianola, 2013; Gianola et al., 2009). However, non-parametric and purely conceptual methods can also be used for this purpose. The Delta-p method was theoretically proposed by Resende (2015) for application to genomic selection, but has not been implemented and evaluated so far. The method is based on the genetic distance between two subpopulations, using the concepts of changes in allele frequency due to selection and the genetic gain theory. It is also possible to combine the genomic values from the G-BLUP with the genomic values

provided by Delta-p under a selection index framework called Delta-p/G-BLUP.

The aim of this study was to propose the Delta-p and Delta-p/G-BLUP index, and to compare it with the traditional G-BLUP method. The efficiency of these methods was evaluated through simulation and real studies. The simulated dataset consisted of eight scenarios comprising the combination of two levels of heritability, two genetic architectures and two dominance status (absence and complete dominance). The empirical dataset consisted of 370 rice accessions phenotyped for seven yield traits, which were genotyped for 44,100 SNP markers.

Materials and Methods

Simulated datasets

The data set was simulated using the Real Breeding software program (Viana, 2011), and the generation of data has been described by Azevedo et al. (2015). A population of 5,000 individuals from 100 families, generated from the random mating of two linkage equilibrium populations, was subjected to five generations of random mating without selection, mutation nor migration. The final population is an advanced generation composite in Hardy-Weinberg equilibrium and linkage disequilibrium (LD).

The prediction of genomic breeding value based on thousands of single nucleotide polymorphisms (SNPs) depends on LD between markers and QTL. Thus, the

LD value (Δ) in a composite population can be calculated from $\Delta_{ab} = \left(\frac{1-2\theta_{ab}}{4}\right)(p_a^1 - p_a^2)(p_b^1 - p_b^2)$ (Kempthorne, 1973; Viana, 2004; Viana et al., 2016), where a and b are two SNPs, or two QTLs, or one SNP and one QTL, θ the frequency of recombinant gametes, and p^1 and p^2 the allele frequencies in the parental populations 1 and 2, respectively.

Azevedo et al. (2015) simulated a total of 2,000 equidistant SNP markers separated by 0.10 centimorgans (cM) across the ten equally sized chromosomes. One hundred of the simulated markers were QTLs randomly distributed in the genome and 95 % of the genetic variation was explained by the markers, according to the expression proposed by Goddard et al. (2011). An amount of 1,000 individuals from 20 full-sib families were genotyped and phenotyped. This simulation study provides a small effective population size ($N_e = 39.22$), which, according to Resende (2015), is the typical value in elite breeding populations of plant species.

The simulated phenotypes considered two types of genetic architecture, one with polygenic inheritance, wherein the 100 QTLs were assumed with limited additive effects on the phenotype, and the other with mixed inheritance, the 5 QTLs with effects accounting for 50 % of the genetic variability and the small additive effects were assigned to the remaining 95 QTLs. The genotypic values for homozygotes were calculated considering $G_{max} = 100(m + a)$ as the upper limit and $G_{min} = 100(m - a)$, as the lower limit where m was the mean of genotypic values and a the homozygote genotypic value. The phenotypic values were computed by adding genotypic values to the true population mean, and error effects sampled from a normal distribution $N(0, \sigma_e^2)$, where the error variance σ_e^2 was computed according to the heritability levels. The additive and dominance effects were assumed to be independents and were normally distributed with zero mean and genetic variance, depending on the assumed genetic architecture and the desired heritability level. In the simulation study, it was also observed that SNPs had minor allele frequency (MAF) greater than 5 %.

Scenarios

Eight different scenarios were used in the analyses: two genetic architectures (polygenic and mixed inheritance); dominance status (absence and complete), two broad-sense heritability levels (around 0.20 and 0.30) for absence of dominance and two broad-sense heritability levels (around 0.30 and 0.50) for complete dominance. These values were chosen because, according to Azevedo et al. (2015), it was expected in these cases, that the genomic selection would be superior to phenotypic selection. The description of the scenarios is presented in Table 1.

Real datasets

Asian rice (*Oryza sativa*) is one of the most widely consumed foods in the world (Subudhi et al., 2006).

Table 1 – Scenarios with the respective true additive heritabilities (h_a^2), due to dominance (h_d^2) and total (h_g^2), genetic architectures (traits controlled by genes of small effect - polygenic inheritance, and traits controlled by genes of small and greater effect - mixed inheritance) and dominance status (absence and complete).

Scenario	Genetic Architectures	Dominance status	h_a^2	h_d^2	h_g^2
Scenario 1	Polygenic inheritance	Absence	0.22	-	0.22
Scenario 2	Polygenic inheritance	Absence	0.33	-	0.33
Scenario 3	Mixed inheritance	Absence	0.20	-	0.20
Scenario 4	Mixed inheritance	Absence	0.35	-	0.35
Scenario 5	Polygenic inheritance	Complete	0.21	0.10	0.31
Scenario 6	Polygenic inheritance	Complete	0.35	0.17	0.52
Scenario 7	Mixed inheritance	Complete	0.20	0.13	0.33
Scenario 8	Mixed inheritance	Complete	0.33	0.21	0.54

Thus, the high population growth rate has justified the interest of researchers in developing high-yielding rice varieties. The real/empirical dataset consisted of 370 rice accessions phenotyped for seven yield traits, which were genotyped for 44,100 SNPs markers. The dataset is publicly available and is part of two projects, the OryzaSNP Project and the OMAP Project (Ammiraju et al., 2006; Zhao et al., 2011), and is available at <http://ricediversity.org/>.

Plantations were supervised throughout the access phase, from May to Oct of the years 2006 and 2007. A complete block design with two replications was used, in which the planting lines had a length equal to 5 m. The plants were spaced 25 cm apart and 0.50 m between rows. Rice accessions were evaluated in the field at Stuttgart, Arkansas coordinates, Lat: 034°49.5805' N; Long -091°54.8357' W, 100 m. Further details can be found in Zhao et al. (2011). The quality control procedures were made considering a call rate of 70 % and minor allele frequency (MAF) less than 1 %. After the quality control of the genomic database the total was 36,901 SNPs markers.

The seven traits evaluated were: (i) Panicle number per plant, (ii) Plant height, (iii) Panicle length, (iv) Primary panicle branch number, (v) Seed number per panicle, (vi) Florets per panicle, (vii) Panicle fertility.

G-BLUP method

The Genomic Best Linear Unbiased Predictor (G-BLUP) method is based on the following linear mixed model:

$$y = 1\mu + Za + Zd + e,$$

where, y is a vector of phenotypes ($N \times 1$, being N the number of genotyped and phenotyped individuals); μ the general mean and 1 the vector with dimension ($N \times 1$); a the vector of additive genomic values ($N \times 1$) with incidence matrix Z ($N \times N$), whose assumed distribution is $a \sim N(0, G_a \sigma_a^2)$, where σ_a^2 is the genetic additive variance and G_a ($N \times N$) the additive genomic relationship

matrix; d the vector of dominance genomic values ($N \times 1$) with incidence matrix Z ($N \times N$), whose assumed distribution is $d \sim N(0, G_d \sigma_d^2)$, where σ_d^2 is the dominance variance and G_d ($N \times N$) the dominance genomic relationship matrix; and e the random residual vector, assumed as $e \sim N(0, I \sigma_e^2)$, where σ_e^2 is the residual variance and I an identity matrix.

According to Vitezica et al. (2013), the genomic relationship matrices for the additive and dominance effects, G_a and G_d , are given, respectively, by:

$$G_a = \frac{WW'}{\sum_{i=1}^n (2p_i q_i)} \text{ and } G_d = \frac{SS'}{\sum_{i=1}^n (2p_i q_i)^2};$$

where p_i and q_i are the allele frequencies of the locus i . The elements of W and S are given by:

$$W = \begin{cases} \text{if } MM, \text{ then } 2 \rightarrow 2q \\ \text{if } Mm, \text{ then } 1 \rightarrow q - p \\ \text{if } mm, \text{ then } 0 \rightarrow 2p \end{cases} \text{ and } S = \begin{cases} \text{if } MM, \text{ then } 0 \rightarrow 2q^2 \\ \text{if } Mm, \text{ then } 1 \rightarrow 2pq \\ \text{if } mm, \text{ then } 0 \rightarrow -2p^2 \end{cases}$$

Delta-p method

The Delta-p method uses the concepts of changes in allele frequency due to selection and genetic gain theory (contrast between the averages of two subpopulations). The training population is initially divided into two subpopulations: one with individuals above the general mean (subpopulation one, with a higher average genetic value u_1) and the other with individuals below the general mean (subpopulation two, with lower average genetic value u_2). The difference between the average genetic values of the two subpopulations ($u_1 - u_2$) is due to the frequency of favorable alleles (p) among the two populations. Thus, $(u_1 - u_2)$ is explained by $\Delta p = p_1 - p_2$, where Δp is the difference of allelic frequencies between the two subpopulations. The Δp values are calculated for each marker i , so that the positive signals are allocated as favorable. Thus, the allelic substitution effects (α_i) of these markers are taken as positive, whereas for those with negative signs of Δp their α_i 's are assigned as negative.

In this context, the incidence matrix W for the vector of additive effects of markers is reparametrized, making the markers genotypes as 0 (bb), 1 (Bb) and 2 (BB). Here, the B allele is the favorable allele, and the BB or bb allocation is dictated by the α_i signal. Thus, the correct allocation in BB or bb is probabilistic. In general, it is expected that the highest number of allocation errors will occur in loci with very minor effects.

The Delta-p method requires the estimation of the α_i effects, and later the prediction of the genomic values (a). Taking Δp_i as an indicator of the relative magnitude of α_i (the higher the Δp_i , the higher the chance of the α_i being higher), it is possible to define the proportion $\Delta p_i / \Delta p_m$ (with their respective positive or negative signals); where Δp_m is the average delta-p computed using the absolute value of Δp_i multiplied by the average α (α_m), which results in the α_i effects. Thus, $\alpha_i = (\Delta p_i / \Delta p_m) \alpha_m$.

The α_m value can be obtained using the theoretical expression of genetic gain. For a given locus i in subpopulation one, $u_{1i} = (p_{1i} - q_{1i}) \alpha_i = (2p_{1i} - 1) \alpha_i$, where $i = 1, 2, \dots, n$; and in subpopulation two, $u_{2i} = (p_{2i} - q_{2i}) \alpha_i = (2p_{2i} - 1) \alpha_i$. Thus, $(u_{1i} - u_{2i}) = (2p_{1i} - 2p_{2i}) \alpha_i = 2\Delta p_i \alpha_i$. Summing all loci, $(u_1 - u_2) = 2\sum(\Delta p_i \alpha_i)$. The quantity Δp_m is defined by Falconer (1989) as: $\Delta p_m = sq(1 - q)$, where $s = \alpha k / \sigma_p$, where s is the selective coefficient of the locus, assuming an additive-dominance model. Therefore, $\Delta p_m = \alpha k p q / \sqrt{\sigma_p^2}$ where k is the standardized selection intensity and σ_p^2 the phenotypic variance. The genetic gain is defined as $G_s = k \sigma_a^2 / \sqrt{\sigma_p^2}$. Thus, $k / \sqrt{\sigma_p^2} = G_s / \sigma_a^2$. Replacing Δp_m in the expression it is possible to obtain $\Delta p_m = \alpha p q G_s / \sigma_a^2$.

The genetic gain associated with the first differential selection is given by $G_{s1} = h_a^2 (u_1 - u_0)$, whereas the second one is $G_{s2} = h_a^2 (u_0 - u_2)$. These gains are symmetrical and must be summed to obtain the total gain $2G_s$, given by: $2G_s = G_{s1} + G_{s2} = h_a^2 (u_1 - u_2)$. Thus, $G_s = (1/2) h_a^2 (u_1 - u_2)$. Replacing G_s into Δp_m , $\Delta p_m = \alpha p q G_s / \sigma_a^2$ and $\Delta p_m = \alpha p q (1/2) h_a^2 (u_1 - u_2) / \sigma_a^2$; where $\sigma_a^2 = 2pq\alpha^2$. Replacing and dividing by Δp_m , we have:

$$1 = [\alpha p q (1/2) h_a^2 (u_1 - u_2)] / (2pq\alpha^2 \Delta p_m),$$

thus, $1 = (1/2) h_a^2 (u_1 - u_2) / (2\alpha \Delta p_m)$. Isolating α from the last expression, $\alpha = (1/2) h_a^2 (u_1 - u_2) / (2\Delta p_m)$ for n locus and the average α per locus is given by:

$$\alpha_m = 0.5 h_a^2 (u_1 - u_2) / [n_{\text{markers}} 2\Delta p_m].$$

Therefore, the additive markers effects (α) are non-parametrically estimated in the training population (individuals with known phenotypes and genotypes). Based on these effects, the genetic values (a) are estimated by the following expression: $\hat{a} = W\hat{\alpha}$, being W the incidence matrix of the additive genetic effects of the markers. This approach does not demand an iterative computational method and uses only genetic distance concepts (magnitude of Δp_i) and the genetic gains of two subpopulations. Additionally, it also does not use differential shrinkage based on allele frequencies, which benefits the loci with higher MAFs. In summary, the Delta-p method can be implemented by taking the followings steps: i) subdivision of the training population into two, according to the phenotype adjusted for systematic effects; ii) calculation of Δp_i and Δp_m ; iii) calculation of α_m ; and iv) calculation of $\alpha_i = (\Delta p_i / \Delta p_m) \alpha_m$.

For the dominance effects, the quantity $\Delta(2pq)$ (difference between frequencies of heterozygous genotypes of subpopulations one and two) is used instead of Δp . The better the phenotype and the higher the frequency of heterozygous genotypes in the population, the more favorable the value of the genotype corrected for the additive genetic effects (i.e., higher the of dominance deviations). In order to infer the effects of dominance, the interest is to find this association (frequency of heterozygous and values of phenotypes). For this, the

phenotype must be returned as the frequency of heterozygous, corrected for additive effects. The incidence matrix S is inherent to the frequency of heterozygous among the three genotype classes. Then, g_m is defined as the genotype value of the heterozygous, which is corrected for the additive effect (α_i). It gives the dominance deviation of the locus i (δ_i), and the genomic dominance value (vector d) is defined as $\hat{d} = S\hat{\delta}$, where S is the previously mentioned incidence matrix.

Delta-p/G-BLUP Index method

Based on the principle of combined selection (Lush, 1947; Falconer, 1989), the G-BLUP can be jointly used with the Delta-p method. A selection index of the form $I = b_1\hat{a}_1 + b_2\hat{a}_2$ is defined based on information from predicted additive genomic values from G-BLUP (\hat{a}_1) and Delta-p (\hat{a}_2), weighted by b_1 and b_2 coefficients, respectively. The index can be also represented by $I = b_1f(y, W) + b_2f((u_1 - u_2), \Delta p)$; where y is the vector of individual phenotypic data, W the centralized allelic dosage matrix, $(u_1 - u_2)$ the contrast of averages from subpopulations and Δp the differential frequency between subpopulations. The weights of the index are given by $b = P^{-1}C$ as follows:

$$C = \begin{bmatrix} \sigma_{a_1}^2 r_{\hat{a}_1 a}^2 & \\ \sigma_{a_2}^2 r_{\hat{a}_2 a}^2 & \end{bmatrix}, P = \begin{bmatrix} \sigma_{a_1}^2 r_{\hat{a}_1 a}^2 & cov(a_1, a_2) r_{\hat{a}_1 a}^2 r_{\hat{a}_2 a}^2 \\ cov(a_1, a_2) r_{\hat{a}_1 a}^2 r_{\hat{a}_2 a}^2 & \sigma_{a_2}^2 r_{\hat{a}_2 a}^2 \end{bmatrix} = \begin{bmatrix} \sigma_{a_1}^2 r_{\hat{a}_1 a}^2 & \sigma_{a_1 a_2}^2 r_{\hat{a}_1 a}^2 r_{\hat{a}_2 a}^2 \\ \sigma_{a_2}^2 r_{\hat{a}_2 a}^2 & \sigma_{a_2}^2 r_{\hat{a}_2 a}^2 \end{bmatrix}$$

where C is a matrix of variance and reliability (r^2) of estimated genomic values through Delta-p and G-BLUP, P a covariance matrix between the estimated genomic values through Delta-p and G-BLUP and a the true additive genetic value of individuals.

Dividing P and C by $\sigma_{a_1}^2$, and considering $\Delta^2 = \sigma_{a_2}^2 / \sigma_{a_1}^2$, it is possible to obtain:

$$C = \begin{bmatrix} r_{\hat{a}_1 a}^2 & \\ \Delta^2 r_{\hat{a}_2 a}^2 & \end{bmatrix}; P = \begin{bmatrix} r_{\hat{a}_1 a}^2 & \Delta^2 r_{\hat{a}_1 a}^2 r_{\hat{a}_2 a}^2 \\ \Delta^2 r_{\hat{a}_1 a}^2 r_{\hat{a}_2 a}^2 & \Delta^2 r_{\hat{a}_2 a}^2 \end{bmatrix}$$

The weights (b_1 and b_2) and accuracy (r_{Ia}) between the index and the true additive genetic value a are given respectively by:

$$b_1 = \frac{1 - r_{\hat{a}_2 a}^2 \Delta^2}{1 - r_{\hat{a}_1 a}^2 r_{\hat{a}_2 a}^2 \Delta^2}; b_2 = \frac{1 - r_{\hat{a}_1 a}^2}{1 - r_{\hat{a}_1 a}^2 r_{\hat{a}_2 a}^2 \Delta^2};$$

$$r_{Ia} = \left[\frac{r_{\hat{a}_1 a}^2 + r_{\hat{a}_2 a}^2 \Delta^2 - 2r_{\hat{a}_1 a}^2 r_{\hat{a}_2 a}^2 \Delta^2}{1 - r_{\hat{a}_1 a}^2 r_{\hat{a}_2 a}^2 \Delta^2} \right]^{1/2} = \left[1 - \frac{(1 - r_{\hat{a}_1 a}^2)(1 - r_{\hat{a}_2 a}^2 \Delta^2)}{1 - r_{\hat{a}_1 a}^2 r_{\hat{a}_2 a}^2 \Delta^2} \right]^{1/2}$$

Under this approach, the reliabilities of \hat{a}_1 ($r_{\hat{a}_1 a}^2$, obtained by G-BLUP) and \hat{a}_2 ($r_{\hat{a}_2 a}^2$, obtained by Delta-p) and

the proportion of variances of \hat{a}_2 and \hat{a}_1 are required. If $\Delta^2 = \sigma_{a_2}^2 / \sigma_{a_1}^2$ provides a value higher than one, then Δ^2 must be at least 1. Another option is to calculate $cov(a_1, a_2)$ directly through $cov(\hat{a}_1, \hat{a}_2)$. Considering that the dominance effects of the methods are analogous to that described above only the genomic value due to the dominance should be used (\hat{d}) instead of the additive genomic value (\hat{a}) in derived equations.

Computer resources

All computational implementations of the proposed methods were performed using the R software program (version 3.3.1). The G-BLUP was performed in the *rrBLUP* package through the *mixed.solve* function. The algorithm used for the development of the Delta-p method is available in http://www.novoscurtos.ufv.br/departamentos/ufv/det/www/?page_id=1374.

Comparisons of the proposed methods

Simulated dataset

Each type of population was simulated 10 times under the same parameter settings, which preserved the same features and provided samples that were effectively from the same conceptual population. Nine replicates were used as training populations, and one replicate was used as a validation population. The estimations based on each of the nine replicates were validated by obtaining estimates of the efficiency measures. The efficiency measures of genomic prediction were given by the accuracies (r_{aa} and r_{da}), bias ($b_{y\hat{a}}$ e $b_{y\hat{d}}$), recovered additive and dominance genomic heritabilities (h_a^2 e h_d^2) and relative efficiency (RE).

The accuracy is defined as the correlation between the predicted values and the parametric genetic values; and the prediction bias is defined as the regression coefficient associated with this correlation. For regression coefficients below one (< 1) it is understood that predicted values had been overestimated, whereas for coefficients above one (> 1), it was concluded that predicted values had been underestimated. The recovered additive genomic heritability is given by

$$h_a^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_d^2 + \sigma_e^2},$$

where $\sigma_a^2 = \sum_{i=1}^n 2p_i q_i m_i^2$ is the additive genomic variance, m_i^2 the square of the i th marker effect with allelic frequencies equal to p_i and q_i . The genomic heritability due to dominance is given by

$$h_d^2 = \frac{\sigma_d^2}{\sigma_a^2 + \sigma_d^2 + \sigma_e^2},$$

where $\sigma_d^2 = \sum_{i=1}^n (2p_i q_i d_i)^2$ and d_i is the genotypic value of the heterozygous. The relative efficiency is given by the ratio between the accuracies from the compared methods. All these measures were obtained for each replicate in each scenario and the general results were reported as average values.

Real dataset

We obtained the measures of efficiency using five-fold cross validation ($k = 5$), in which in each fold the training population consisted of 296 individuals (80 % of total) and the validation population consisted of 74 individuals (20 % of total). The efficiency measures used were the predictive ability ($r_{\hat{a}_y}$), which consisted of the correlation between the estimated additive genomic values and the phenotypic values of the validation population, and prediction bias.

Results and Discussion

The average and the respective estimated standard deviations for the accuracy, bias, heritabilities and relative efficiency were obtained by both Delta-p and G-BLUP and are shown in Table 2. We note an increase in accuracy with the use of the G-BLUP. The correlation between the genomic values predicted by the Delta-p and the G-BLUP method was high ($r_{\hat{a}_p\hat{a}_g}$), ranging from 0.80 to 0.88. Although the G-BLUP is widely applied to GWS and recommended for the prediction of additive and dominant genomic values (Azevedo et al., 2015; De los Campos et al., 2013; Gianola, 2013; Gianola et al., 2009), this study proposed combining the values derived from the Delta-p and G-BLUP methods to construct an index giving different weights to each of them which can be seen as an improvement in the G-BLUP additive and dominant genomic predictions. For the additive model, the increase in accuracy using the Delta-p/G-BLUP index (0.80 to 0.87) was 0.03 units on average compared to the accuracy of the G-BLUP method (0.77 to 0.85). For the additive-dominant model, this increase was more evident, being 0.08 units. The Delta-p/G-BLUP index, considering the additive model, provided a relative average efficiency of 103 % to 105 %, while in the additive-dominant model this efficiency increased to 109 % to 116 %. It is noted that the Delta-p/G-BLUP index improved the prediction of additive and dominant genomic values through G-BLUP, since the index provided greater accuracy. It is important to report that this advantage in terms of extra percentage points has no additional computational costs.

As already mentioned, the Delta-p/G-BLUP index method is based on the combined selection index. According to Lush (1947) and Falconer (1989), this process outperformed other selection methods, for example, individual or among or within family selection. Similarly, in the present study, the Delta-p/G-BLUP index method, which is based on the combined selection index, outperformed the G-BLUP and Delta-p methods. Gains of 5 % in accuracy are already significant in genetic improvement, often equivalent to the gain that is obtained in a complete cycle of improvement (Resende et al., 2015).

The differences between allele frequencies were effective in estimating the marker effects. This result corroborates the principle of the Delta-p method. Weller et al. (2014) and Weller (2016) reported that marker se-

Table 2 – Average and standard deviation of accuracy (r_a), bias (b_y), additive heritability (h_a^2), correlation between additive genomic values estimated by Delta-p and Genomic Best Linear Unbiased Predictor - G-BLUP ($r_{\hat{a}_p\hat{a}_g}$), proportion between the additive variances estimated (Δ^2), weights (b_1 and b_2), accuracy (r_a) and bias (b_y) of the Delta-p/G-BLUP index method, relative efficiency (RE) between the index and G-BLUP considering simulated dataset.

Model	Scenarios	Method	$r_{\hat{a}_a}$	$b_{y\hat{a}}$	h_a^2	$r_{\hat{a}_p\hat{a}_g}$	Δ^2	b_1	b_2	r_a	$b_{y\hat{a}}$	RE	
Additive	1	Delta-p	0.67 ± 0.03	1.13 ± 0.10	0.11 ± 0.01	0.86 ± 0.03	0.62 ± 0.09	0.87 ± 0.02	0.49 ± 0.02	0.80 ± 0.01	0.83 ± 0.04	1.05 ± 0.01	
		G-BLUP	0.77 ± 0.01	1.02 ± 0.06	0.17 ± 0.02	0.82 ± 0.03	0.71 ± 0.12	0.85 ± 0.03	0.42 ± 0.02	0.85 ± 0.01	0.86 ± 0.06	1.03 ± 0.01	
	2	Delta-p	0.71 ± 0.01	1.00 ± 0.08	0.25 ± 0.03	0.88 ± 0.01	0.67 ± 0.17	0.85 ± 0.04	0.47 ± 0.03	0.38 ± 0.03	0.87 ± 0.01	0.87 ± 0.06	1.04 ± 0.01
		G-BLUP	0.83 ± 0.01	1.02 ± 0.09	0.35 ± 0.04	0.86 ± 0.02	0.69 ± 0.09	0.86 ± 0.03	0.38 ± 0.03	0.69 ± 0.04	0.87 ± 0.01	0.87 ± 0.04	1.03 ± 0.01
	3	Delta-p	0.69 ± 0.02	1.15 ± 0.07	0.11 ± 0.01	0.84 ± 0.04	1.09 ± 0.32	0.78 ± 0.05	0.69 ± 0.04	0.73 ± 0.02	0.73 ± 0.02	0.66 ± 0.03	1.15 ± 0.04
		G-BLUP	0.79 ± 0.01	1.05 ± 0.11	0.18 ± 0.03	0.80 ± 0.03	1.32 ± 0.35	0.68 ± 0.12	0.67 ± 0.06	0.80 ± 0.04	0.80 ± 0.04	0.71 ± 0.04	1.16 ± 0.06
4	Delta-p	0.72 ± 0.01	1.05 ± 0.06	0.24 ± 0.02	0.85 ± 0.04	0.60 ± 0.17	0.88 ± 0.03	0.68 ± 0.05	0.68 ± 0.05	0.66 ± 0.04	0.59 ± 0.05	1.10 ± 0.03	
	G-BLUP	0.85 ± 0.01	1.00 ± 0.05	0.36 ± 0.03	0.81 ± 0.02	0.68 ± 0.27	0.85 ± 0.08	0.64 ± 0.04	0.71 ± 0.05	0.71 ± 0.05	0.54 ± 0.03	1.09 ± 0.04	
Additive-Dominant	5	Delta-p	0.55 ± 0.06	0.73 ± 0.10	0.16 ± 0.03	0.80 ± 0.03	0.80 ± 0.03	0.80 ± 0.03	0.80 ± 0.03	0.80 ± 0.03	0.80 ± 0.03	0.80 ± 0.03	1.16 ± 0.06
		G-BLUP	0.63 ± 0.02	1.00 ± 0.08	0.13 ± 0.02	0.85 ± 0.04	1.09 ± 0.32	0.78 ± 0.05	0.69 ± 0.04	0.73 ± 0.02	0.73 ± 0.02	0.66 ± 0.03	1.15 ± 0.04
	6	Delta-p	0.59 ± 0.03	0.70 ± 0.08	0.33 ± 0.07	0.80 ± 0.03	1.32 ± 0.35	0.68 ± 0.12	0.67 ± 0.06	0.80 ± 0.04	0.80 ± 0.04	0.71 ± 0.04	1.16 ± 0.06
		G-BLUP	0.69 ± 0.02	1.02 ± 0.06	0.26 ± 0.03	0.85 ± 0.04	0.60 ± 0.17	0.88 ± 0.03	0.68 ± 0.05	0.68 ± 0.05	0.66 ± 0.04	0.59 ± 0.05	1.10 ± 0.03
	7	Delta-p	0.55 ± 0.07	1.10 ± 0.17	0.07 ± 0.01	0.81 ± 0.02	0.68 ± 0.27	0.85 ± 0.08	0.64 ± 0.04	0.71 ± 0.05	0.71 ± 0.05	0.54 ± 0.03	1.09 ± 0.04
		G-BLUP	0.60 ± 0.04	1.02 ± 0.25	0.12 ± 0.01	0.81 ± 0.02	0.68 ± 0.27	0.85 ± 0.08	0.64 ± 0.04	0.71 ± 0.05	0.71 ± 0.05	0.54 ± 0.03	1.09 ± 0.04
8	Delta-p	0.57 ± 0.05	0.89 ± 0.11	0.15 ± 0.04	0.81 ± 0.02	0.68 ± 0.27	0.85 ± 0.08	0.64 ± 0.04	0.71 ± 0.05	0.71 ± 0.05	0.54 ± 0.03	1.09 ± 0.04	
	G-BLUP	0.65 ± 0.03	0.89 ± 0.07	0.23 ± 0.04	0.81 ± 0.02	0.68 ± 0.27	0.85 ± 0.08	0.64 ± 0.04	0.71 ± 0.05	0.71 ± 0.05	0.54 ± 0.03	1.09 ± 0.04	

Scenarios with traits controlled by small effects genes - Scenario 1 ($h_a^2 = 0.22$), Scenario 2 ($h_a^2 = 0.33$), Scenario 5 ($h_a^2 = 0.21$ and $h_d^2 = 0.10$), Scenario 6 ($h_a^2 = 0.35$ and $h_d^2 = 0.17$). Scenarios with traits controlled by genes of small and large effects - Scenario 3 ($h_a^2 = 0.20$), Scenario 4 ($h_a^2 = 0.35$), Scenario 7 ($h_a^2 = 0.20$ and $h_d^2 = 0.13$), Scenario 8 ($h_a^2 = 0.33$ and $h_d^2 = 0.21$).

lection based on differences between allelic frequencies between groups of young and older bulls, i.e. between two contrasting subpopulations, can be effective in genomic selection in dairy cattle.

Considering the additive model, it was observed that for all scenarios, the G-BLUP method was the one that presented an estimated additive heritability close to the parametric value (Table 2). Considering the additive-dominant model, it was observed that for the scenarios whose traits are controlled by genes of small effects (scenarios 5 and 6), the Delta-p method presented the additive heritability estimates closest to the parametric values. On the other hand, for the scenarios whose traits are controlled by genes of both small and sizeable effects (scenarios 7 and 8), the G-BLUP method presented better performance.

Considering the additive model, the G-BLUP and Delta-p methods had non-biased predicted additive genomic values (Table 2), except for the Delta-p method in scenarios 1 and 3, where the method underestimated these values ($b_{yda} > 1$). Considering the additive-dominant model, the G-BLUP method provided non-biased additive genomic values, except for scenario 8, where the method overestimated these values ($b_{yda} < 1$). In contrast, the Delta-p method presented predicted additive genomic values which were overestimated for scenarios 5, 6 and 8; and underestimated for scenario 7. Already, the Delta-p/G-BLUP index overestimated the predicted additive genomic values for all scenarios considered ($b_{yda} < 1$). However, the index had biases considerably closer to one in the additive model than in the additive-dominant model (Table 2).

Table 3 shows the average results obtained through a simulated dataset, the methods for the genomic values due to dominance considering the additive-dominant model and also the weights, accuracy and efficiency obtained through the Delta-p/G-BLUP index. In this case, the improvement in the prediction of genomic values with the use of the index (Table 3) was less evident. Only 0.01 units improved in prediction accuracy compared to the accuracy of G-BLUP. Bennewitz and Meuwissen (2010), Hill et al. (2008), and Wellmann and Bennewitz (2012) discuss the relevance of the inclusion of dominance in Quantitative Genomics. However, according to the results reported by Azevedo et al. (2015), genomic values due to dominance are difficult to predict and are consequently associated with low accuracy values and higher bias. This corroborates the low accuracy level recorded for the Delta-p/G-BLUP index and G-BLUP when considering only dominance genomic values.

Although the dominance genomic values, considering the additive-dominant model, were inaccurately estimated, it can be seen that the index and the G-BLUP accuracy were similar. This result agrees with Azevedo et al. (2015), which used several methods applied to GWS and obtained results of similar dominance accuracy. Thus, it should be noted that the main criteria for the comparison of methods in these scenarios is the recovered heritabilities, due to dominance and prediction bias.

Table 3 – Average and standard deviation of accuracy (r_{da}), bias (b_{yda}), additive heritability (h_a^2), correlation between dominant genomic values estimated by Delta-p and Genomic Best Linear Unbiased Predictor - G-BLUP ($r_{da\phi_2}$), proportion between the dominant variances estimated (Δ^2), weights (b_1 and b_2), accuracy (r_d) and bias (b_{ydf}) of the Delta-p/G-BLUP index method, relative efficiency (RE) between the index and G-BLUP considering simulated dataset.

Model	Scenarios	Method	r_{da}	b_{yda}	h_a^2	$r_{da\phi_2}$	Δ^2	b_1	b_2	r_d	b_{ydf}	RE
5		Delta-p	0.13 ± 0.04	1.28 ± 0.31	0.01 ± 0.00	0.67 ± 0.11	0.20 ± 0.08	1.00 ± 0.00	0.96 ± 0.01	0.21 ± 0.02	0.62 ± 0.11	1.04 ± 0.04
		GBLUP	0.20 ± 0.02	0.84 ± 0.13	0.08 ± 0.01							
6		Delta-p	0.16 ± 0.02	1.18 ± 0.24	0.03 ± 0.01	0.66 ± 0.06	0.18 ± 0.09	1.00 ± 0.00	0.94 ± 0.01	0.25 ± 0.01	0.49 ± 0.05	1.04 ± 0.03
		GBLUP	0.24 ± 0.01	0.62 ± 0.11	0.17 ± 0.03							
7		Delta-p	0.02 ± 0.08	0.95 ± 0.18	0.03 ± 0.01	0.60 ± 0.12	0.39 ± 0.12	1.00 ± 0.01	1.00 ± 0.00	0.07 ± 0.06	0.52 ± 0.08	3.62 ± 8.84
		GBLUP	0.05 ± 0.07	0.71 ± 0.42	0.15 ± 0.21							
8		Delta-p	0.01 ± 0.05	0.89 ± 0.16	0.06 ± 0.02	0.48 ± 0.12	0.35 ± 0.14	1.00 ± 0.00	1.00 ± 0.00	0.07 ± 0.02	0.44 ± 0.07	1.48 ± 1.27
		GBLUP	0.06 ± 0.03	0.50 ± 0.11	0.19 ± 0.04							

Scenarios with traits controlled by small effects genes - Scenario 5 ($h_a^2 = 0.21$ and $h_d^2 = 0.10$), Scenario 6 ($h_a^2 = 0.35$ and $h_d^2 = 0.17$). Scenarios with traits controlled by genes of small and large effects - Scenario 7 ($h_a^2 = 0.20$ and $h_d^2 = 0.13$), Scenario 8 ($h_a^2 = 0.33$ and $h_d^2 = 0.21$).

When analyzing the heritabilities due to dominance, it was observed that for the four scenarios, the closest estimates were provided by the G-BLUP method. However, in scenarios 5 and 8, the heritabilities, due to dominance, were slightly underestimated and in scenario 7 the heritabilities, due to dominance, were slightly overestimated. The Delta-p method underestimated the genomic values due to dominance in the scenarios whose traits were controlled by genes with small effects (scenarios 5 and 6) and overestimated the scenarios whose traits were controlled by genes with both small and sizeable effects. The G-BLUP overestimated the values in all scenarios, with bias being considerably lower in relation to the Delta-p method. In addition, the Delta-p/G-BLUP index also overestimated the genomic values due to dominance for all scenarios considered ($b_{y_{la}} < 1$) (Table 3).

Although we have previously analyzed the additive genomic values and because of the dominance, considering the additive-dominant model, it is also important to present the results referring to the total genomic value ($a + d$), since this is related to individual performance (Table 4). Analogous to Tables 2 and 3, the G-BLUP presented higher accuracy in the prediction of total genomic values, when compared to the Delta-p method. The increase in accuracy using the Delta-p/G-BLUP index (0.65 to 0.71) was 0.07 units, on average, compared to the accuracy of the G-BLUP method (0.56 to 0.65), showing again the efficiency of the index. The Delta-p/G-BLUP index provided a relevant relative average efficiency that was 108 % to 114 % (Table 4). It was observed that, under the four scenarios considered, the G-BLUP had a better performance in recovering the total genomic heritability than the Delta-p. However, these were still underestimated (Table 4). In addition, the three methods produced predicted total genomic values that were overestimated ($b_{y_{\hat{g}}} < 1$). Although the Delta-p/G-BLUP index presented biased total genomic values, this method has the great advantage of obtaining greater accuracy when compared to the G-BLUP method.

To elucidate the importance of the Delta-p/G-BLUP index for predicting genomic selection, the results for an applied breeding program of Asian rice are presented in Table 5. Analogous to previous results considering additive model (Table 2), G-BLUP showed more accurate genomic values than Delta-p. The Delta-p method presented values showing overestimated additive genomic values ($b_{y_{\hat{a}}} < 1$), except for the Panicle number per plant, while the G-BLUP presented values that were less biased. The Delta-p/G-BLUP index, considering the real dataset, provided a relative average efficiency of 101 % to 116 %, being on average 8 % more efficient than the G-BLUP. Eight % is highly expressive and under genomic selection performed annually, such as in rice cultivation, these gains are cumulative and grow rapidly (Resende et al., 2015). Thus, the index proposed in this study can be satisfactorily implemented in rice breeding programs.

Table 4 – Average and standard deviation of accuracy ($r_{\hat{g}g}$), bias ($b_{y_{\hat{g}}}$), additive heritability ($h_{\hat{g}}^2$), correlation between total genomic values estimated by Delta-p and Genomic Best Linear Unbiased Predictor - GBLUP ($r_{\hat{g}\hat{g}_2}$), proportion between the total variances estimated (Δ^2), weights (b_1 and b_2), accuracy ($r_{\hat{g}}$) and bias ($b_{y_{\hat{g}}}$) of the Delta-p/G-BLUP index method, relative efficiency (RE) between the index and GBLUP considering real dataset.

Model	Scenarios	Method	$r_{\hat{g}g}$	$b_{y_{\hat{g}}}$	$h_{\hat{g}}^2$	$r_{\hat{g}\hat{g}_2}$	Δ^2	b_1	b_2	$r_{\hat{g}}$	$b_{y_{\hat{g}}}$	RE
5		Delta-p	0.51 ± 0.03	0.66 ± 0.07	0.17 ± 0.02	0.78 ± 0.04	0.62 ± 0.15	0.89 ± 0.02	0.68 ± 0.03	0.65 ± 0.03	0.44 ± 0.04	1.10 ± 0.02
		GBLUP	0.60 ± 0.03	0.60 ± 0.06	0.21 ± 0.02							
6		Delta-p	0.56 ± 0.02	0.64 ± 0.05	0.36 ± 0.04	0.71 ± 0.05	0.63 ± 0.16	0.88 ± 0.03	0.62 ± 0.03	0.71 ± 0.02	0.47 ± 0.03	1.08 ± 0.02
		GBLUP	0.65 ± 0.02	0.59 ± 0.06	0.43 ± 0.10							
7		Delta-p	0.51 ± 0.04	0.58 ± 0.07	0.10 ± 0.01	0.81 ± 0.05	0.71 ± 0.12	0.86 ± 0.02	0.73 ± 0.02	0.64 ± 0.02	0.38 ± 0.04	1.14 ± 0.02
		GBLUP	0.56 ± 0.02	0.54 ± 0.06	0.27 ± 0.04							
8		Delta-p	0.52 ± 0.02	0.46 ± 0.05	0.21 ± 0.06	0.75 ± 0.03	0.85 ± 0.25	0.84 ± 0.04	0.69 ± 0.04	0.69 ± 0.02	0.34 ± 0.02	1.13 ± 0.05
		GBLUP	0.61 ± 0.03	0.47 ± 0.04	0.42 ± 0.08							

Additive-Dominant

Scenarios with traits controlled by small effects genes - Scenario 5 ($h_a^2 = 0.21$ and $h_d^2 = 0.10$), Scenario 6 ($h_a^2 = 0.35$ and $h_d^2 = 0.17$). Scenarios with traits controlled by genes of small and large effects - Scenario 7 ($h_a^2 = 0.20$ and $h_d^2 = 0.13$), Scenario 8 ($h_a^2 = 0.33$ and $h_d^2 = 0.21$).

Table 5 – Predictive ability ($r_{\hat{y}}$), bias ($b_{\hat{y}}$), additive heritability (h_a^2), correlation between additive genomic values estimated by Deltap and Genomic Best Linear Unbiased Predictor - GBLUP ($r_{\hat{y}, \hat{y}_G}$), proportion between the additive variances estimated (Δ^2), weights (b_1 and b_2), predictive ability ($r_{\hat{y}_G}$) and bias ($b_{\hat{y}_G}$) of the Deltap/G-GLUP index method, relative efficiency (RE) between the index and G-GLUP considering real dataset.

Yield Trait	Method	$r_{\hat{y}}$	$b_{\hat{y}}$	h_a^2	$r_{\hat{y}, \hat{y}_G}$	Δ^2	b_1	b_2	$r_{\hat{y}_G}$	$b_{\hat{y}_G}$	RE
Panicle Number Per Plant	Deltap	0.69 ± 0.11	1.74 ± 0.31	0.16 ± 0.11	0.85 ± 0.04	0.25 ± 0.05	0.96 ± 0.01	0.35 ± 0.11	0.83 ± 0.07	0.93 ± 0.13	1.01 ± 0.01
	GBLUP	0.82 ± 0.07	1.03 ± 0.12	0.65 ± 0.02							
Plant Height	Deltap	0.47 ± 0.16	0.84 ± 0.27	0.31 ± 0.12	0.63 ± 0.10	0.56 ± 0.12	0.91 ± 0.03	0.41 ± 0.11	0.81 ± 0.06	0.93 ± 0.11	1.01 ± 0.01
	GBLUP	0.80 ± 0.07	1.09 ± 0.10	0.55 ± 0.07							
Panicle Length	Deltap	0.40 ± 0.15	0.66 ± 0.23	0.36 ± 0.13	0.60 ± 0.09	0.78 ± 0.17	0.89 ± 0.03	0.55 ± 0.13	0.73 ± 0.09	0.80 ± 0.21	1.04 ± 0.04
	GBLUP	0.71 ± 0.11	1.06 ± 0.10	0.47 ± 0.11							
Primary Panicle Branch Number	Deltap	0.34 ± 0.19	0.43 ± 0.26	0.65 ± 0.07	0.57 ± 0.15	1.00 ± 0.00	0.67 ± 0.41	0.71 ± 0.13	0.71 ± 0.15	0.62 ± 0.24	1.11 ± 0.10
	GBLUP	0.65 ± 0.13	1.06 ± 0.23	0.38 ± 0.19							
Seed Number Per Panicle	Deltap	0.30 ± 0.24	0.39 ± 0.33	0.81 ± 0.06	0.55 ± 0.16	1.00 ± 0.00	0.97 ± 0.02	0.66 ± 0.15	0.68 ± 0.18	0.53 ± 0.28	1.16 ± 0.20
	GBLUP	0.58 ± 0.13	1.12 ± 0.30	0.28 ± 0.37							
Florets Per Panicle	Deltap	0.41 ± 0.18	0.59 ± 0.26	0.51 ± 0.07	0.66 ± 0.06	1.00 ± 0.00	0.92 ± 0.02	0.55 ± 0.15	0.74 ± 0.12	0.72 ± 0.21	1.08 ± 0.09
	GBLUP	0.69 ± 0.11	1.09 ± 0.18	0.40 ± 0.11							
Panicle Fertility	Deltap	0.27 ± 0.10	0.39 ± 0.16	0.50 ± 0.13	0.49 ± 0.13	1.00 ± 0.00	0.99 ± 0.01	0.71 ± 0.16	0.59 ± 0.14	0.55 ± 0.19	1.16 ± 0.15
	GBLUP	0.52 ± 0.16	0.99 ± 0.32	0.31 ± 0.18							

Thus, the Delta-p/G-BLUP index is defined based on information from predicted genomic values from G-BLUP and Delta-p. This approach is rarely applied in the context of GWS. This method, as much for the simulated data as for the rice data, showed itself to be superior to the G-BLUP method. In addition, it led to genomic, additive and dominance values that are more accurate than the traditional G-BLUP and required reduced computational time. It can also be easily implemented in the context of GWS.

Conclusions

The proposed method, the Delta-p/G-BLUP index, is easy to implement in the context of genomic selection. It demands reduced computational cost and leads to genomic, additive, dominance and total values that are more accurate than the G-BLUP method. On the other hand, the Delta-p method provided lower values of accuracy and dominance bias estimates than the G-BLUP method. However, for additive effects, under almost all scenarios evaluated, the estimates were more biased for the Delta-p model than for the G-BLUP. For the traits evaluated in rice it was concluded that the Delta-p/G-BLUP index was superior to the G-BLUP method, and increased predictive ability.

Acknowledgments

The authors thank the following Brazilian funding organizations: the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Fundação Arthur Bernardes (FUNARBE).

Authors' Contributions

Conceptualization: Lima, L.P., Azevedo, C.F., Resende, M.D.V., Silva, F.F. Data acquisition: Lima, L.P., Azevedo, C.F., Suela, M.M., Viana, J.M.S. Data Analysis: Lima, L.P., Azevedo, C.F., Resende, M.D.V., Suela, M.M. Design of methodology: Lima, L.P., Azevedo, C.F., Resende, M.D.V., Silva, F.F. Software development: Lima, L.P., Azevedo, C.F., Resende, M.D.V., Suela, M.M., Viana, J.M.S. Writing and Editing: Lima, L.P., Azevedo, C.F., Resende, M.D.V., Silva, F.F., Suela, M.M., Nascimento, M.

References

- Ammiraju, J.S.S.; Luo, M.; Goicochea, J.L.; Wang, W.; Kudrna, D.; Mueller, C.; Talag, J.; Kim, H.; Sisneros N.B.; Blackmon, B.; Fang, E.; Tomkins, J.B.; Brar, D.; Mackill, D.; MacCouch, S.; Kurata, N.; Lambert, G.; Galbraith, D.W.; Arumuganathan, K.; Rao, K.; Walling, J.G.; Gill, N.; Yu, Y.; Sanmiguel, P.; Soderlund, C.; Jackson, S.; Wing, R.A. 2006. The Oryza bacterial artificial chromosome library resource: construction and analysis of 12 deep-coverage large-insert BAC libraries that represent the 10 genome types of the genus *Oryza*. *Genome Research* 16: 140-147.

- Azevedo, C.F.; Resende, M.D.V.; Silva, F.F.; Viana, J.M.S.; Valente, M.S.F.; Resende Junior, M.F.R.; Muñoz, P. 2015. Ridge, Lasso and Bayesian additive-dominance genomic models. *BMC Genetics* 16: 105.
- Bennewitz, J.; Meuwissen, T.H.E. 2010. The distribution of QTL additive and dominance effects in porcine F2 crosses. *Journal of Animal Breeding and Genetics* 127: 171-179.
- De los Campos, G.; Hickey, J.M.; Pong-Wong, R.; Daetwyler, H.D.; Callus, M.P.L. 2013. Whole genome regression and prediction methods applied to plant and animal breeding. *Genetics* 193: 327-345.
- Falconer, D.S. 1989. *Introduction to Quantitative Genetics*. Longman Scientific and Technical, New York, NY, USA.
- Gianola, D. 2013. Priors in whole-genome regression: the Bayesian alphabet returns. *Genetics* 194: 573-596.
- Gianola, D.; De los Campos, G.; Hill, W.G.; Manfredi, E.; Fernando, R. 2009. Additive genetic variability and the Bayesian alphabet. *Genetics* 183: 347-363.
- Goddard, M.E. 2009. Genomic selection: prediction of accuracy and maximization of long term response. *Genetics* 136: 345-357.
- Goddard, M.E.; Hayes, B.J.; Meuwissen, T.H.E. 2011. Using the genomic relationship matrix to predict the accuracy of genomic selection. *Journal of Animal Breeding and Genetics* 128: 409-421.
- Habier, D.; Fernando, R.L.; Dekkers, J.C.M. 2007. The impact of genetic relationship on genome-assisted breeding values. *Genetics* 117: 2389-2397.
- Hill, W.G.; Goddard, M.E.; Visscher, P.M. 2008. Data and theory point to mainly additive genetic variance for complex traits. *PLoS Genetics* 4: e1000008.
- Kempthorne, O. 1973. *An Introduction to Genetic Statistics*. Iowa State University Press, Ames, IA, USA.
- Lush, J.L. 1947. Family merit and individual merit as basis for selection. *American Naturalist* 81: 241-261.
- Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157: 1819-1829.
- Resende, M.D.V. 2015. *Quantitative and Population Genetics = Genética Quantitativa e de Populações*. Editora Suprema, Visconde do Rio Branco, MG, Brazil (in Portuguese).
- Resende, M.D.V.; Ramalho, M.A.P.; Guilherme, S.R.; Abreu, A.F.B. 2015. Multi generation index in the within progenies bulk method for breeding of self-pollinated plants. *Crop Science* 55: 1202-1211.
- Subudhi, P.K.; Sasaki, T.; Khush, G.S. 2006. Rice. v.1, p. 1-9. In: Kole, C., ed. *Genome mapping and molecular breeding in plants*. Springer, Berlin, Germany.
- Van Raden, P.M. 2008. Efficient methods to compute genomic predictions. *Journal of Dairy Science* 91: 4414-4423.
- Viana, J.M.S. 2004. *Quantitative genetics theory for non-inbred populations in linkage disequilibrium*. *Genetics and Molecular Biology* 27: 594-601.
- Viana, J.M.S. 2011. Program for molecular and quantitative data analysis Real Breeding = Programa para análises de dados moleculares e quantitativos Real Breeding. Universidade Federal de Viçosa, Viçosa, MG, Brazil (in Portuguese).
- Viana, J.M.S.; Piepho, H.P.; Silva, F.F. 2016. Quantitative genetics theory for genomic selection and efficiency of breeding value prediction in open-pollinated populations. *Scientia Agricola* 73: 243-251.
- Vitezica, Z.G.; Varona, L.; Legarra, A. 2013. On the additive and dominance variance and covariance of individuals within the genomic selection scope. *Genetics* 195: 1223-1230.
- Weller, J.I. 2016. *Genomic Selection in Animals*. John Wiley, Hoboken, NJ, USA.
- Weller, J.I.; Glick, G.; Ezra, E.; Seroussi, E.; Shemesh, M.; Zeron, Y.; Ron, M. 2014. Predictive ability of selected subsets of single nucleotide polymorphisms (SNPs) in a moderately sized dairy cattle population. *Animal* 8: 208-216.
- Wellmann, R.; Bennewitz, J. 2012. Bayesian models with dominance effects for genomic evaluation of quantitative traits. *Genetics Research* 94: 21-37.
- Zhao, K.; Tung, C.W.; Eizenga, G.C.; Wright, M.H.; Ali, M.L.; Price, A.H.; Norton, J. G.; Islam, A.R.; Reynolds, A.; Mezey, J.; McClung, A.M.; Bustamante, C.D.; McClung, A.M. 2011. Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nature Communications* 2: 467-477.