

Trait-specific long-term consequences of genomic selection in beef cattle

Haroldo Henrique de Rezende Neves^{1,2} · Roberto Carneiro¹ · Sandra Aida de Queiroz¹ 

Received: 9 December 2016 / Accepted: 31 October 2017 / Published online: 8 November 2017
© Springer International Publishing AG, part of Springer Nature 2017

Abstract Simulation studies allow addressing consequences of selection schemes, helping to identify effective strategies to enable genetic gain and maintain genetic diversity. The aim of this study was to evaluate the long-term impact of genomic selection (GS) in genetic progress and genetic diversity of beef cattle. Forward-in-time simulation generated a population with pattern of linkage disequilibrium close to that previously reported for real beef cattle populations. Different scenarios of GS and traditional pedigree-based BLUP (PBLUP) selection were simulated for 15 generations, mimicking selection for female reproduction and meat quality. For GS scenarios, an alternative selection criterion was simulated (wGBLUP), intended to enhance long-term gains by attributing more weight to favorable alleles with low frequency. GS allowed genetic progress up to 40% greater than PBLUP, for female reproduction and meat quality. The alternative criterion wGBLUP did not increase long-term response, although allowed reducing inbreeding rates and loss of favorable alleles. The results suggest that GS outperforms PBLUP when the selected trait is under less polygenic background and that attributing

more weight to low-frequency favorable alleles can reduce inbreeding rates and loss of favorable alleles in GS.

Keywords Female reproduction · Low-frequency alleles · Meat quality · Simulation

Introduction

Since pioneer studies highlighted the potential benefits of genomic selection (GS), this strategy has been compared to traditional PBLUP selection mainly to assess the gain in accuracy achievable with GS (e.g. Hayes et al. 2009; Daetwyler et al. 2012). Such focus on accuracy indicates that GS has been mostly evaluated with respect to its potential to enhance short-term response (Bijma 2012), whereas the maintenance of genetic diversity in longer time horizons is essential for the sustainability of breeding schemes (Kristensen and Sørensen 2005).

Long-term genetic gains are invariably related to the maintenance of effective population size (N_e) large enough to cope with this aim, possibly with some sacrifice of short-term response (Bijma 2012), for instance through restriction on inbreeding and optimum contribution selection (OCS). Since the efficacy of such strategies relies on accurate estimation of relatedness as well as on the genetic merit of selection candidates, genomic information could open new opportunities to enhance long-term gains, allowing more precise estimation of the aforementioned parameters (Daetwyler et al. 2007; Ferencakovic et al. 2013).

Alternative selection criteria have also been proposed to increase long-term selection response. Jannink (2010) and Sun and VanRaden (2014), through simulation studies, highlighted the importance of giving larger weight to low-frequency favorable alleles at the beginning of the

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10709-017-9999-1>) contains supplementary material, which is available to authorized users.

✉ Sandra Aida de Queiroz
saquei@fcav.unesp.br

- ¹ Departamento de Zootecnia, School of Agricultural and Veterinarian Sciences (FCAV), São Paulo State University (UNESP), Via de Acesso Prof. Paulo Donato Castellane s/n, Jaboticabal, SP 14884-900, Brazil
- ² GenSys Consultores Associados S/S Ltda., Rua Guilherme Alves, 170. Cj 304, Porto Alegre, RS 90680-000, Brazil

selection process, to balance long-term and short-term responses. Despite the importance of these results, large differences in terms of selection schemes and simulation attributes (e.g. marker density and population size) hinder extrapolating such conclusions to the application of GS in cattle.

While the cost-effectiveness of GS is evident for dairy cattle (e.g. Schaeffer 2006), its adoption in other livestock populations is still limited. In beef cattle, the different structure of the breeding schemes, larger effective population sizes and shorter generation intervals could result in more modest advantage in using GS when compared to traditional schemes, especially for traits routinely recorded before selection decisions takes place.

According to Carvalheiro et al. (2013), application of GS could be envisaged in Brazilian beef cattle, mainly due to the large number of animals routinely recorded by breeding organizations and to the fact that bulls are progeny-tested relatively late in life. In such situations, GS schemes would be advantageous especially for traits in which the progress achieved by conventional selection is currently limited, e.g. for hard-to-measure and late-life traits (Muir 2007; Goddard et al. 2010). Thus, it would be important to evaluate the impact of GS by simulating scenarios mimicking situations of particular importance to beef cattle production, aiming to assess long-term consequences of this process. In the present study, we simulated phenotypes mimicking female reproduction and meat quality traits, so that consequences of GS for traits expressed late in life and hard to measure traits, respectively, could be investigated. The impact of adopting different replacement strategies was also addressed, as more intense use of younger bulls can be envisaged for beef cattle selection schemes (Carvalheiro 2014).

The simulation of repeated cycles of GS considering continuous update of the reference population, as is usual in real applications of GS in cattle (e.g. VanRaden et al. 2013), would contribute to evaluate the feasibility and further consequences of application of GS in beef cattle. This situation is hard to assess by deterministic predictions, especially with respect to inbreeding incidence and genetic diversity in the long-term.

The present study used stochastic simulation to: (1) investigate the long-term impact of GS compared to traditional selection, in scenarios mimicking selection for female reproduction (E1) and meat quality (E2) under different replacement strategies; (2) evaluate the impact of different genomic prediction methods under different scenarios of genetic architecture, in terms of genetic progress, inbreeding incidence and maintenance of genetic diversity; and (3) assess the long-term impact of giving larger weight to low-frequency favorable alleles in genomic predictions.

Materials and methods

Simulation design

The stochastic simulation process was carried out in two stages. In the first stage of the simulations, the QMSim software (Sargolzaei and Schenkel 2009) was employed to simulate a beef cattle population (10 replicates). In a second stage (described later), starting from the base populations generated in the first stage, the recent generations were simulated, through repeated cycles of selection and mating considering different scenarios of prediction method and genetic architecture.

The forward-in-time simulation parameters were defined in QMSim aiming to mimic the pattern of linkage disequilibrium (LD) previously reported for real beef cattle populations, similarly as described in Brito et al. (2011). A first step of simulation of historical generations was carried out, in which 1000 historical generations (H1) were simulated, with N_e kept constant at 1000. After this step, starting from the last generation of H1, 20 historical generations (H2) were created, so that a bottleneck was simulated, through gradual reduction of N_e from 1000 to 200. Such a strategy allowed the initial establishment of mutation-drift equilibrium and the creation of linkage disequilibrium between the simulated loci. For both H1 and H2, the number of animals of each sex was the same. Matings were simulated assuming random union of gametes.

After that, 100 males and 100 females were randomly selected from the last generation of H2 and the population was expanded. In this step, three generations were simulated under random mating and no selection, so that each female produced five offspring per generation and the number of females in reproduction increased exponentially at each generation. From the last generation simulated in the expansion step, 40 males and 1000 females were randomly selected to compose the base population (G0), from which all recent generations (described later) were simulated.

The same simulation framework was employed to produce ten different replicates (base populations), so that the same base populations (G0) were used to simulate the starting point for different selection scenarios, similarly as applied by Bastiaansen et al. (2012). For each scenario and base population, 15 recent generations were simulated, so that different selection strategies were applied to each scenario, which were succeeded by random mating of the animals selected to produce the subsequent generation. At each generation, it was assumed that each female and male in reproduction had 1 and 25 offspring, respectively, totaling 1000 progeny produced.

Simulated genome

The simulated genome was composed by 29 pairs of autosomes (BTA), whose individual length varied from 40 to 146 cM, totaling 2333 cM, similarly as in Brito et al. (2011). In the historical generations, a mutation rate (u) of 10^{-4} was simulated for both markers and QTL, to simulate the target number of segregating loci in the last generation.

Missing genotypes and genotyping errors were not simulated. Mutation events were only simulated in the historical generations, since the recent population involved a small number of generations and thus the effect of mutation could be neglected (Sargolzaei and Schenkel 2009). For both markers and QTL, a recurrent mutation model was adopted. Markers were simulated as bi-allelic, aiming to resemble SNP markers, while the number of QTL alleles at each locus varied from 2 to 4. The markers were evenly spaced along the genome, which resulted in a number of markers per chromosome proportional to length of each BTA. While the number of QTL was also proportional to the length of each BTA, QTL positions within each chromosome were set randomly.

When a large number of historical generations are simulated, aiming to achieve mutation-drift equilibrium, a large proportion of non-segregating loci can be obtained in the last generation. Thus, in the last historical generation, 40,000 segregating markers were randomly selected among those with $MAF > 0.01$, in order to simulate the information available in marker panels in the recent generations. In addition, 1000 QTL loci with $MAF > 0.01$ were randomly selected to simulate phenotypes in the recent generations. Such procedures allowed to simulate marker density similar to that obtained when using commercial SNP panels designed for cattle, after the quality control of genotypes is carried out (e.g. Sargolzaei et al. 2008; Hayes et al. 2009).

For each replicate, 3064 evenly spaced marker loci (IBD loci) were used to monitor identity-by-descent: in each chromosome, about 100 IBD markers were positioned evenly spaced and monitored through the attribution of specific tags to each founder allele available in G_0 , similarly as in Sonesson et al. (2012). Markers with $MAF < 5\%$ were also monitored, aiming to simulate the evolution of loci potentially associated to deleterious mutations (homozygous mutation loci, HML), similarly as in Keller et al. (2011). Both IBD and HML loci were not included in the set of markers considered in genomic predictions, so that they were not directly involved in selection.

A summary of the simulation parameters is available in Online Resource 1, whereas a schematic representation of the simulation steps is provided in Online Resource 2.

Simulated scenarios

Two selection schemes were simulated, aiming to replicate selection for female reproductive traits and meat quality (schemes E1 and E2, respectively). A schematic overview of such selection schemes is provided in Online Resource 3. Within a given scheme, each scenario was defined by the application of a particular selection criterion, as a way to compare genomic prediction under different methods to traditional PBLUP selection.

Since the version of QMSim employed in this study did not allow simulating repeated cycles of selection and mating using different genomic prediction methods, custom routines were coded in R language (R Development Core Team 2013) to allow simulation of such scenarios. In order to simulate gametogenesis, the recombination was modeled by assuming that the number of crossovers follows a Poisson distribution with average equal to the chromosome length (in Morgans). Crossover positions were randomly assigned along each chromosome. For each scenario and replicate, the genotypes of animals of a given generation were simulated assuming random union of gametes of the selected parents (1000 dams and 40 sires).

The true breeding value (TBV) of each animal was calculated as: $TBV = Zv$, where Z and v are the incidence matrix relating individuals to QTL alleles and the matrix of additive effects of QTL alleles, respectively. In all scenarios, a continuous phenotype Y was simulated by adding an overall mean ($\mu = 10$) and a random residual (ϵ) to the TBV of each animal. The residuals were simulated so that $\epsilon \sim N(0, \sigma_\epsilon^2)$, in which σ_ϵ^2 was defined in each combination of scenario and replicate to meet a pre-specified heritability.

The following selection schemes were simulated:

Female reproduction (E1) A continuous trait of heritability equal to 0.15, influenced by 1000 QTL and only expressed in females was simulated. QTL effects were drawn from a gamma distribution (shape parameter = 0.40). The first generation was obtained after the random mating of 1000 dams and 40 sires from the base population (G_0), producing 1000 individuals (G_1), about half of each sex. After G_1 was generated, the following selection criteria were employed to select the parents of the next generations (G_2 – G_{15}):

1. PBLUP: estimated breeding value (EBV). This criterion intended to evaluate the consequences of traditional selection, based on predictions obtained considering phenotypic records and pedigree information.
2. GBLUP: genomic estimated breeding value (GEBV), where $GEBV = \sum x_j \hat{\beta}_j$, x_j is the genotype for the j th marker and $\hat{\beta}_j$ is the respective estimated allele substitution effect, obtained using ridge regression (Hoerl and Kennard 1970), assuming that the effects of all loci

follow a normal distribution. This criterion intended to evaluate the consequences of GS schemes.

3. wGBLUP: weighted GEBV (wGEBV), where $wGEBV = x_j \hat{\beta}_j p_j^{-0.5}$, in which x_j and $\hat{\beta}_j$ are the genotype and estimated allele substitution effect, respectively, similarly as described for GBLUP, and p_j is the respective frequency of the favorable allele for the j -th marker. This criterion is identical to that applied by Jannink (2010) aiming to reduce the loss of favorable alleles of lower frequency, by means of attributing more weight to them when computing genomic predictions.
4. TBV: true breeding value, as described previously. The simulation of this criterion intended to investigate the consequences of very high prediction accuracies on genetic progress and genetic diversity.

For all scenarios, it was assumed that the phenotype of each female would be available only after the selection decisions. Thus, when selecting the parents of generation $\mathbf{i} + 1$, genetic predictions for females from generation \mathbf{i} were based only on the phenotypic information accumulated until generation $(\mathbf{i} - 1)$. After this step, among the females of generation \mathbf{i} , only the phenotypes of those selected for reproduction were available and considered in the subsequent genetic evaluation. For the sake of simplicity, it was assumed that each female expressed the phenotype only once during its life.

At each generation, the BLUPF90 software (Misztal 2014) was employed to predict EBVs of all animals. All pedigree and phenotypic information accumulated and available until each generation was considered in BLUP equations, as well as the variance ratios calculated with the simulated values for the variance components.

In the scenarios under GS, the 1000 females from G0 were assumed to be genotyped with a 50 k panel, composing the initial reference population employed to estimate marker effects. Custom routines were employed to obtain marker effect estimates using R software, similarly as described in Neves et al. (2012).

Before each new cycle of selection and reproduction, both PBLUP and genomic predictions were updated with phenotypes and genotypes of the females selected in the previous generation. In order to select the parents of generation $\mathbf{i} + 1$, it was assumed that animals from generation \mathbf{i} and their parents were available for reproduction (selection candidates). In addition to the females with phenotypes available, all selection candidates were assumed to be genotyped. For the sake of simplicity, all genotyped animals have genotypes available for all simulated markers.

For each selection criterion, two replacement strategies were simulated: (1) culling of the worse 20% animals (8 males and 200 females) among the parents of generation

\mathbf{i} (v1, “fixed replacement rate”) and selection of the top 8 males and top 200 females from generation \mathbf{i} to replace them; (2) selection of the top 1000 females and top 40 males for that criterion, considering all animals available for reproduction, regardless of their generation coefficient (v2, “variable replacement rate”). Thus, for E1, eight different scenarios were simulated, as the combination between selection criterion (PBLUP, GBLUP, wGBLUP or TBV) and replacement strategy (v1 or v2), for ten replicates.

Meat quality (E2) Three different traits (A, B and C), expressed in both sexes, were simulated with heritability equal to 0.35, according to the following scenarios of genetic architecture:

- A. 1000 QTL, whose effects were drawn from a gamma distribution (shape = 0.40);
- B. 100 QTL, whose effects were drawn from a gamma distribution (shape = 0.40);
- C. 1000 QTL, five of which had larger effect (“major QTL”), each one explaining an arbitrarily large proportion of the additive variance (7%) and with position randomly sampled among the QTL loci. The remainder QTL (995) had effects drawn from a normal distribution, scaled so that together they explained the remainder 65% of the additive variance. Thus, these traits mimicked different scenarios of genetic architecture: more polygenic control (A), trait influenced by smaller number of loci (B) and trait influenced by large number of loci, few of them of large effect (C).

Similarly as in scheme E1, the first generation was obtained after the mating between the 1000 females and 40 males from G0, producing 1000 offspring, about a half of each sex. After G1 was obtained the following selection criteria were employed to select the parents of the next generation (G2), as well as those of the following generations (up to G15):

(1–4) TBV, PBLUP, GBLUP and wGBLUP (respectively, similarly as previously described for E1).

(5) EN: GEBV calculated with estimated marker effects obtained using the variable selection method elastic net (EN) (Zou and Hastie 2005). This method can be understood as an extension of the LASSO method (Tibshirani 1996), in which solutions are obtained under constraints imposed by a mixed penalty parameter, resultant from the combination between the LASSO penalty and the ridge regression penalty. According to Waldmann et al. (2013), setting a weight (α) close to 1 for the LASSO penalty makes the EN to behave similarly to the LASSO (i.e. strong variable selection), while improving robustness to strongly correlated predictor variables, as is the case of SNP markers. In the present study, α was set to 0.95. Estimation of marker effects was carried out using the glmnet R package.

(6) wEN $wGEBV$ estimated similarly as previously described for $wGBLUP$ (larger weight for rare alleles), except by the fact that marker effect estimates were obtained via EN.

It was assumed that, starting from G1, phenotypes would be measured on 250 animals randomly chosen at each generation, being that phenotyping would require the slaughtering of these animals, as is the case of most carcass and meat quality traits. In order to select the parents of generation $i + 1$, it was assumed that the parents of generation i were available for reproduction, as well as the animals from generation i that were not slaughtered. For selection candidates from generation i or older, each selection criterion was computed considering all information accumulated until that generation (including phenotypes on half-sibs from generation i).

For scenarios under GS, the 250 phenotypes from G1 animals comprised the initial reference population, employed to estimate marker effects under the different methods previously described. At each generation, marker effect estimates were updated, including the new phenotypes and genotypes in the reference set. In addition to the animals slaughtered, it was assumed that all animals available for reproduction (selection candidates) were genotyped, similarly as in E1.

Aiming to investigate the effect of replacement strategies in this scheme, the same strategies previously described as fixed (v1) and variable replacement rate (v2) were applied to the scenarios with selection for trait A, while only the strategy of fixed replacement rate was applied to the traits B and C. Thus, for each selection criterion ($n = 6$), four trait-replacement strategy combinations were simulated (A_v1, A_v2, B_v1 and C_v1), totaling 24 scenarios, for ten replicates.

Criteria for analysis of results

For each scenario and replicate, the following statistics were computed at each generation:

- ΔG : accumulated genetic gain, computed as the difference between the average TBV at generation i and the average TBV in the respective base population (G0), standardized by the SD of TBV in G0.
- SD.TBV: standard deviation of the true breeding values (TBV) at generation i .
- aveL: average generation coefficient of the parents of generation i , computed as a proxy for generation interval in each scenario.
- IBD: Percentage of genome that was homozygous due to identity-by-descent, calculated as in Sonesson et al. (2012). For each locus designated to monitor IBD, population homozygosity was computed as $f_j = \sum f_{kj}^2$, where f_{kj}

is the frequency of the founder allele k at locus j in the generation i . The IBD coefficient was calculated as the average of f_j over all monitored loci.

- Fped: average of pedigree-based inbreeding coefficient, over all animals from generation i .
- HML: Homozygosity at loci potentially associated to deleterious mutations. Calculated as the average of the homozygosity for the lower frequency variants, i.e. for all “homozygous mutation loci” (Keller et al. 2011).
- Number of favorable alleles lost (Nlost): the accumulated number of favorable alleles lost along the selection process.

At each generation, all statistics were computed separately for each replicate (population). Aiming to enhance objective comparisons among scenarios, the results of each statistic were analyzed by fitting the following linear model, considering the data regarding to selection for a same trait:

$$s_{ij} = c_i + p_j + e_{ij},$$

where s_{ij} is the j th observation for statistic s in the i th scenario, c_i is the effect of the i th scenario on s , p_j is the effect of the j th base population and $e_{ij} \sim N(0, \sigma_e^2)$. The assumptions of normally distributed and homoscedastic residuals were checked using Shapiro–Wilk and Breusch–Pagan tests, respectively. For each statistic, generation and trait, least squares means of each scenario were contrasted using t tests, adjusted for false discovery rate ($\alpha = 5\%$). In the case of IBD, Fped and HML, each statistic was log transformed in order to meet the assumptions of the linear model.

Results

Linkage disequilibrium and summary statistics

In the base generation (G0), the average (SD) linkage disequilibrium (LD) between adjacent markers over the 29 simulated chromosomes, measured by the r^2 statistic (Hill and Robertson 1968), was 0.23 (0.24), while the average (SD) of homozygosity was around 0.65 (0.01). The pattern of LD decay with genomic distance (data not shown) was also consistent with that previously reported for beef cattle breeds (e.g. Lu et al. 2012).

Selection for female reproduction (E1)

Genetic progress

For a same selection criterion, greater genetic progress was achieved under the scenarios with variable replacement rate, being that the superiority over scenarios with fixed replacement rate reached 20% in the last generation (Fig. 1;

Fig. 1 Trends for genetic progress and genetic diversity* under different scenarios** of selection for female reproduction trait. **deltaG* accumulated genetic gain (in units of additive SD of the base population), *aveL* average generation interval, *SD.TBV* $10 \times$ SD of true breeding values (TBV), *IBD* proportion of genome that is homozygous due to identity-by-descent, in %, *Fped* average of individual pedigree-based inbreeding coefficients, in %, *HML* homozygosity for alleles potentially associated to deleterious mutations, in %, *Nlost* accumulated number of favorable alleles lost. Averages of 10 replicates are plotted for each statistic and generation. **Each scenario is defined by combination of: replacement strategy (replacement rate fixed at 20% or variable rate, v1 and v2, respectively) and selection criterion: TBV, estimated breeding value (PBLUP), genomic estimated breeding value using GBLUP (GBLUP) and GBLUP giving more weight to low-frequency favorable alleles (wGBLUP)

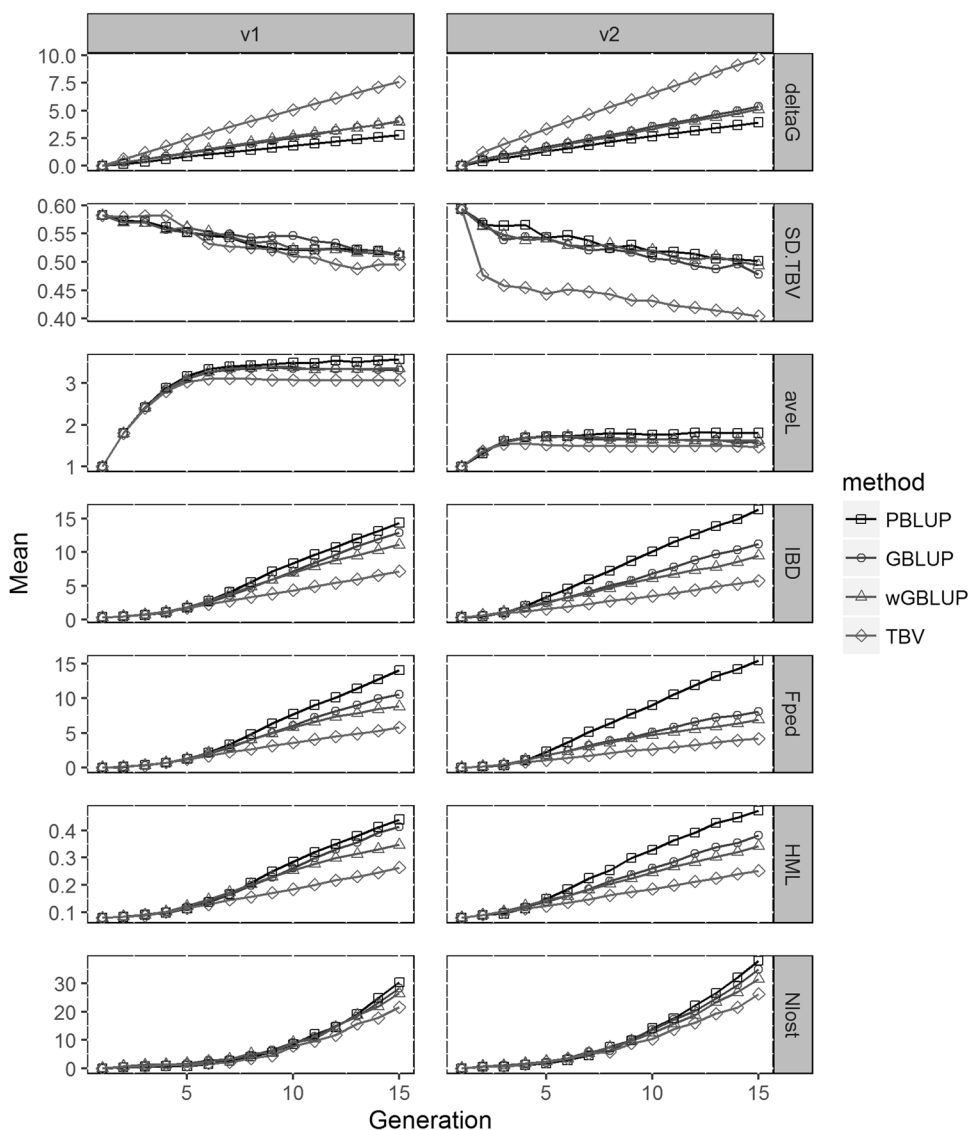


Table 1). The greater genetic progress achieved with variable replacement rate is consistent with the smaller generation interval (*aveL*) in the scenarios with this replacement strategy (Fig. 1; Table 1). Conversely to what was observed for PBLUP (for which *aveL* remained nearly constant since generation 7), there was a slight trend of reduction in *aveL* until the last generations of GS scenarios.

Genetic diversity and inbreeding

Starting from generation 7, the increase in the average inbreeding coefficients based on IBD loci was larger in the scenarios under PBLUP selection, especially when this method was applied without constraint on the replacement rate (v2), reaching values close to 15% in generation 15 (Table 1). The smallest inbreeding levels (according to IBD) were achieved under selection based on TBV (Fig. 1).

Trends for homozygosity at HML loci and average of *Fped* coefficients were very similar to the trends for IBD, being that the largest values were equal to 0.45 and 15%, respectively, for PBLUP at generation 15 (Table 1). The number of favorable alleles lost exhibited a quadratic trend, with more pronounced losses after generation 10 (Fig. 1).

Contrasts at generation 15

For the sake of brevity, and since the focus of this study is on long-term consequences of the different strategies under investigation, detailed results for contrasts between scenarios at generation 15 are presented in this section.

Table 1 Genetic progress and genetic diversity¹ under different scenarios² of selection for female reproduction trait at generation 15*

Repl.	Criterion	deltaG	SD.TBV	aveL	IBD (%)	HML (%)	Fped (%)	Nlost
v1	PBLUP	2.8 ^{(0.15)e}	0.51 ^{(0.02)a}	3.6 ^{(0.03)a}	14.1 ^{(0.07)ab}	0.43 ^{(0.06)ab}	13.8 ^{(0.07)a}	30.3 ^{(1.69)cd}
v1	GBLUP	4.0 ^{(0.12)d}	0.51 ^{(0.01)a}	3.3 ^{(0.02)b}	12.5 ^{(0.09)bc}	0.41 ^{(0.05)ab}	10.2 ^{(0.08)b}	28.3 ^{(1.83)cde}
v1	TBV	7.6 ^{(0.08)b}	0.50 ^{(0.01)ab}	3.1 ^{(0.00)c}	7.1 ^{(0.04)e}	0.26 ^{(0.03)d}	5.8 ^{(0.03)e}	21.6 ^{(0.95)f}
v1	wGBLUP	4.0 ^{(0.17)d}	0.51 ^{(0.01)a}	3.4 ^{(0.03)b}	10.9 ^{(0.06)cd}	0.34 ^{(0.06)c}	8.7 ^{(0.06)c}	26.5 ^{(1.40)de}
v2	PBLUP	3.9 ^{(0.15)d}	0.50 ^{(0.01)a}	1.8 ^{(0.03)d}	16.0 ^{(0.08)a}	0.46 ^{(0.07)a}	15.0 ^{(0.09)a}	37.9 ^{(1.32)a}
v2	GBLUP	5.4 ^{(0.10)c}	0.48 ^{(0.01)b}	1.6 ^{(0.02)e}	11.1 ^{(0.04)c}	0.38 ^{(0.03)bc}	8.0 ^{(0.04)cd}	35.0 ^{(2.02)ab}
v2	TBV	9.7 ^{(0.12)a}	0.40 ^{(0.01)c}	1.5 ^{(0.01)f}	5.7 ^{(0.03)f}	0.25 ^{(0.04)d}	4.2 ^{(0.03)f}	26.2 ^{(2.08)e}
v2	wGBLUP	5.1 ^{(0.08)c}	0.49 ^{(0.01)ab}	1.6 ^{(0.01)e}	9.4 ^{(0.04)d}	0.34 ^{(0.03)c}	6.9 ^{(0.04)d}	31.5 ^{(1.34)bc}

¹deltaG accumulated genetic gain (in units of additive SD of the base population), SD.TBV standard deviation of true breeding values, aveL average generation interval, IBD proportion of genome that is homozygous due to identity-by-descent, in %, Fped average of individual pedigree-based inbreeding coefficients, in %, HML homozygosity for alleles potentially associated to deleterious mutations, in %. Nlost accumulated number of favorable alleles lost

²Each scenario is defined by combination of: (a) replacement strategy (Repl.): fixed rate (20%, v1) or variable rate (v2) and (b) selection criterion: TBV, estimated breeding value (PBLUP), genomic estimated breeding value using GBLUP (GBLUP) and GBLUP giving more weight to low-frequency favorable alleles (wGBLUP)

*Averages of ten replicates (standard error, SE, within brackets). For a same statistic (column), averages with the same superscript letter do not differ (t-test, adjusted p-value > 0.05)

Replacement strategy

Variable replacement rate (v2) resulted in larger genetic gain than that achieved with replacement rate fixed at 20% (v1), irrespective of selection criterion (Table 1). The genetic gain accumulated until generation 15 was about 40% greater for v2 under selection based on PBLUP, while with GS this difference was 30%. The variable replacement rate (v2) resulted in smaller genetic variability for the scenarios under selection based on TBV and GBLUP (Table 1). Under PBLUP selection, there was no significant difference in inbreeding incidence between replacement strategies, whereas under GS, smaller inbreeding levels were obtained with variable replacement rate (Table 1), notably in the case of Fped.

PBLUP vs GS

Genomic selection resulted in genetic progress significantly greater than selection based on PBLUP (Table 1), so that the genetic gain accumulated up to generation 15 was between 25 and 40% greater for GS. GS methods resulted in pedigree-based inbreeding coefficients significantly smaller than PBLUP (between 25 and 54% smaller), regardless of the replacement strategy. However, at generation 15, under fixed replacement rate, only wGBLUP resulted in estimates significantly lower than PBLUP for estimates of IBD, HML and number of alleles lost (Nlost) (Table 1).

Weighting on low frequency favorable alleles (wGBLUP vs GBLUP)

There was no significant difference between GBLUP and wGBLUP in terms of genetic progress (Table 1). Regarding to inbreeding incidence, wGBLUP resulted in smaller averages of Fped and HML (both about – 15%), under fixed replacement rate. Under variable replacement rate, wGBLUP resulted in average IBD coefficients 15.5% smaller than GBLUP (Table 1).

Selection for meat quality (E2)

Genetic progress

Because the results pertaining to the comparison between replacement rates in scheme E2 were quite similar to those previously reported for female reproduction (scheme E1, Fig. 1; Table 1) only the results of genetic progress and genetic diversity obtained with one of the replacement strategies (v1) will be detailed (Figs. 2, 3).

Regardless the simulated genetic architecture, GS outperformed PBLUP in terms of genetic progress (Fig. 2; Table 2). For trait B (100 QTL), the similarity between different GS methods in terms of genetic progress was evident, while a slightly difference between them was observed for traits A and C (Fig. 2), for which, in terms of deltaG, GBLUP outperformed EN (Fig. 2). When GS was applied, a trend of reduction in the generation interval (aveL) was observed for all traits, especially after generation 10 (Fig. 2).

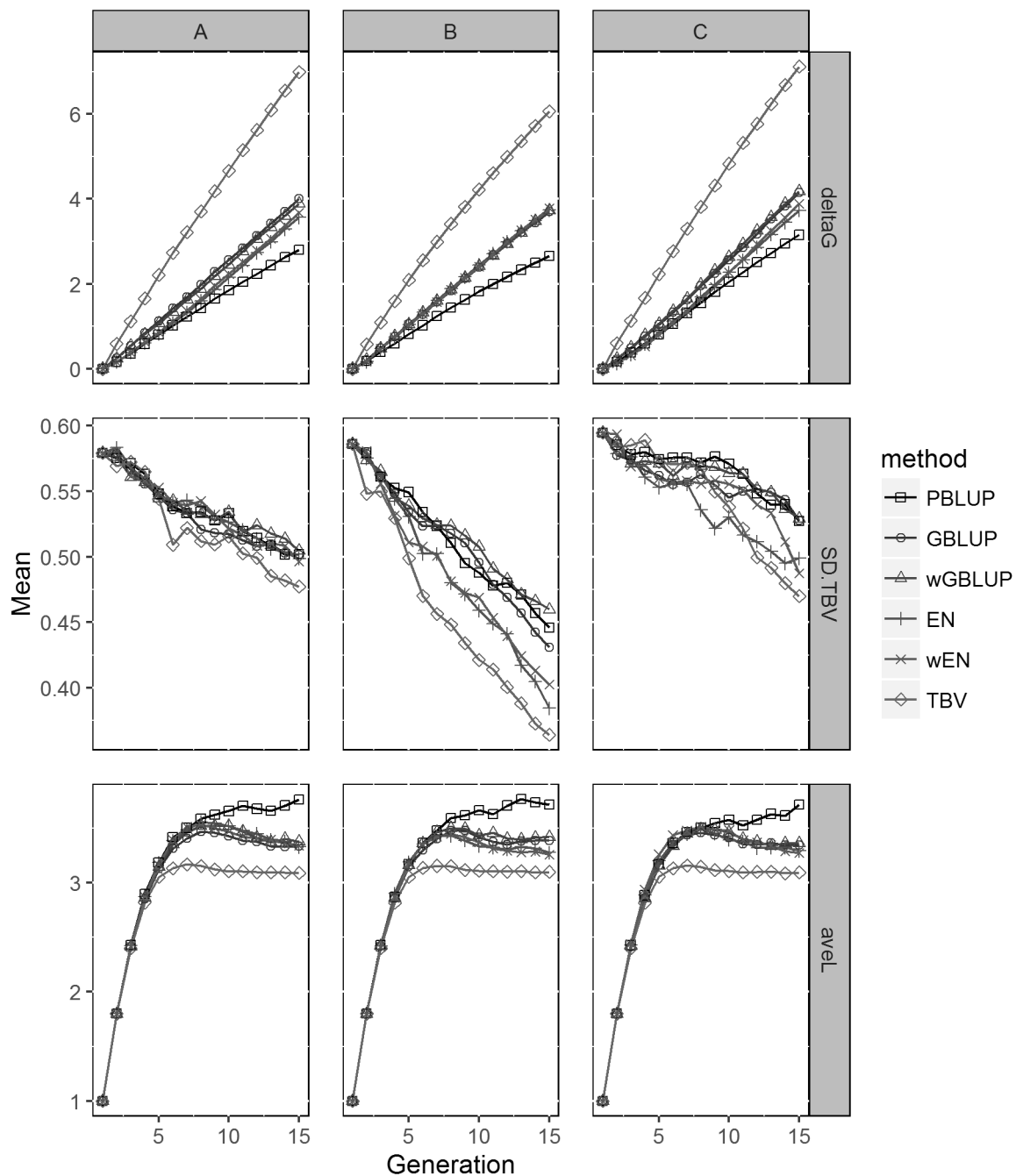


Fig. 2 Trends of genetic progress and genetic variability* for different scenarios** of selection for meat quality traits (a, b, c)***. **deltaG* accumulated genetic gain (in units of additive SD of the base population), *aveL* average generation interval, *SD.TBV* SD of true breeding values (TBV), Averages of 10 replicates plotted for each statistic and generation. **Each scenario is defined by a selection crite-

riion: TBV, EBV (PBLUP), GS using GBLUP or EN (GBLUP or EN) and GBLUP or EN giving more weight to low-frequency favorable alleles (wGBLUP or wEN). ***Trait A (1000 QTL), trait B (100 QTL), trait C (1000 QTL, five of which had larger effect and each explained 7% of the additive variance). A fixed replacement rate of 20% was adopted (v1)

Genetic diversity and inbreeding

For all simulated traits, PBLUP selection resulted in more pronounced increase in inbreeding coefficients (Fig. 3; Table 3), while the selection for TBV resulted in the smallest

inbreeding incidence. For trait A, EN and wEN resulted in more inbreeding than the other GS methods. For traits B and C, lower inbreeding incidence was obtained with wGBLUP and wEN, when compared to GBLUP and EN, respectively (Fig. 3; Table 3).

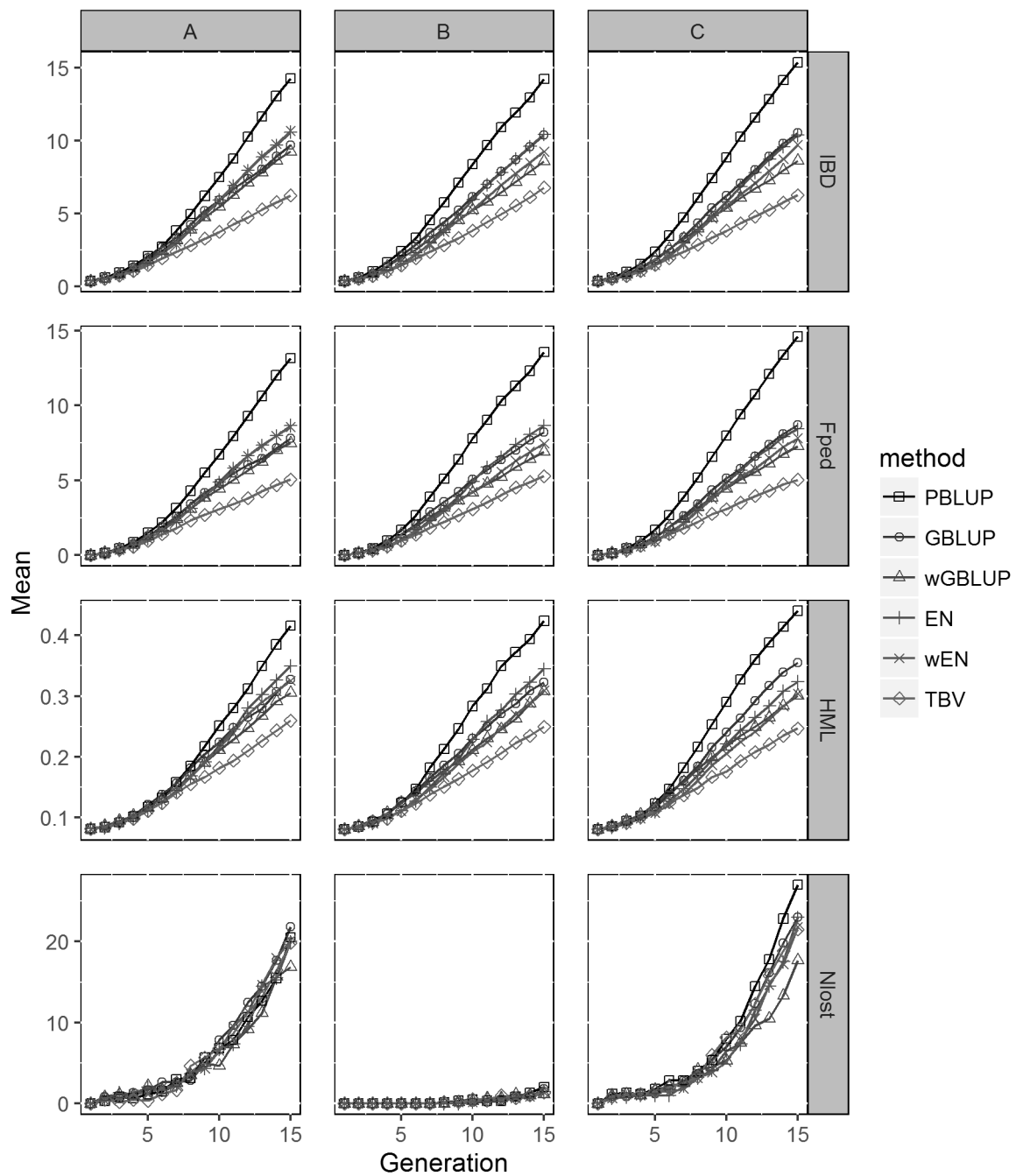


Fig. 3 Trends for genetic diversity* under different scenarios** of selection for meat quality traits (a, b, c)***. *IBD proportion of genome that is homozygous due to identity-by-descent, in %, Fped average of individual pedigree-based inbreeding coefficients, in %, HML homozygosity for alleles potentially associated to deleterious mutations, in %. Nlost accumulated number of favorable alleles lost. Averages of 10 replicates are plotted for each statistic and generation.

Selection criteria: true breeding value (TBV), estimated breeding value (PBLUP), genomic estimated breeding value using GBLUP or EN (GBLUP or EN) and GBLUP or EN giving more weight to low-frequency favorable alleles (wGBLUP or wEN). *Trait A (1000 QTL), trait B (100 QTL), trait C (1000 QTL, five of which had larger effect and explained 7% of the additive variance). A fixed replacement rate of 20% was adopted (v1)

Table 2 Genetic progress and genetic diversity¹ under different scenarios² of selection for meat quality trait at generation 15*, for different replacement rates

Repl.	Criterion	deltaG	SD.TBV	aveL	IBD (%)	HML (%)	Fped (%)	Nlost
v1	PBLUP	2.8 ^(0.13) i	0.50 ^(0.01) ab	3.8 ^(0.07) a	14.0 ^(0.06) a	0.41 ^(0.06) b	12.9 ^(0.07) a	21.6 ^(1.72) de
v1	GBLUP	4.0 ^(0.07) efg	0.50 ^(0.01) ab	3.3 ^(0.02) b	9.6 ^(0.03) c	0.33 ^(0.04) def	7.8 ^(0.03) bc	21.6 ^(1.16) de
v1	EN	3.6 ^(0.12) h	0.50 ^(0.01) ab	3.3 ^(0.02) b	10.4 ^(0.07) bc	0.34 ^(0.06) cdef	8.5 ^(0.06) bc	18.7 ^(1.13) ef
v1	wGBLUP	3.9 ^(0.14) fgh	0.50 ^(0.01) a	3.4 ^(0.03) b	6.2 ^(0.04) d	0.26 ^(0.04) g	5.0 ^(0.04) d	19.1 ^(1.20) ef
v1	wEN	3.7 ^(0.09) gh	0.50 ^(0.01) abc	3.3 ^(0.03) b	9.0 ^(0.07) c	0.30 ^(0.04) f	7.4 ^(0.06) c	15.9 ^(1.40) f
v1	TBV	7.0 ^(0.08) b	0.48 ^(0.01) c	3.1 ^(0.01) c	10.3 ^(0.09) bc	0.32 ^(0.06) ef	8.3 ^(0.08) bc	19.3 ^(2.30) ef
v2	PBLUP	4.1 ^(0.20) ef	0.50 ^(0.01) ab	2.2 ^(0.04) d	16.3 ^(0.06) a	0.49 ^(0.06) a	15.0 ^(0.06) a	36.4 ^(1.54) a
v2	GBLUP	5.1 ^(0.10) c	0.48 ^(0.01) bc	1.9 ^(0.01) e	10.3 ^(0.07) bc	0.36 ^(0.07) bcd	7.9 ^(0.07) bc	29.0 ^(2.59) bc
v2	EN	4.7 ^(0.18) d	0.49 ^(0.01) abc	1.9 ^(0.04) e	11.8 ^(0.09) b	0.41 ^(0.08) b	8.9 ^(0.09) b	30.1 ^(2.95) b
v2	TBV	8.6 ^(0.08) a	0.42 ^(0.01) d	1.7 ^(0.01) f	4.9 ^(0.03) e	0.22 ^(0.02) h	3.6 ^(0.03) e	16.4 ^(1.11) f
v2	wGBLUP	5.3 ^(0.11) c	0.50 ^(0.01) ab	1.9 ^(0.02) e	10.2 ^(0.06) bc	0.37 ^(0.04) bc	7.6 ^(0.07) bc	24.6 ^(1.78) d
v2	wEN	4.3 ^(0.19) e	0.49 ^(0.01) abc	1.8 ^(0.01) e	11.7 ^(0.11) b	0.35 ^(0.07) cde	8.7 ^(0.09) b	25.6 ^(1.80) cd

¹deltaG accumulated genetic gain (in units of additive SD of the base population), SD.TBV standard deviation of true breeding values, aveL average generation interval, IBD proportion of genome that is homozygous due to identity-by-descent, in %, HML homozygosity for alleles potentially associated to deleterious mutations, in %, Fped average of individual pedigree-based inbreeding coefficients, in %, Nlost accumulated number of favorable alleles lost

²Each scenario is defined by combination of: (a) Repl.=replacement strategy (fixed rate=20%, v1) or variable rate (v2) and (b) selection criterion: TBV, estimated breeding value (PBLUP), genomic estimated breeding value using GBLUP or EN (GBLUP) and GBLUP or EN giving more weight to low-frequency favorable alleles (wGBLUP or wEN). The selection for a trait influenced by 1000 QTL was simulated (trait A)*

*Averages of 10 replicates (standard error, SE, within brackets). For a same statistic (column), averages with the same superscript letter do not differ (t-test, adjusted p-value > 0.05)

Contrasts at generation 15

Replacement strategy

For a given selection criterion, significantly larger genetic progress was obtained with variable replacement rate (Table 2), confirming the same trend verified in the scheme mimicking selection for female reproduction (E1). Inbreeding incidence (Fped, IBD and HML) was also greater under variable replacement rate than with fixed replacement (Table 2).

At generation 15, there was no difference between replacement strategies in terms of IBD and Fped, while the variable replacement rate resulted in significantly greater loss of favorable alleles (v2, Table 2).

PBLUP vs GS

Most GS scenarios resulted in genetic progress significantly greater than achieved in correspondent scenarios under PBLUP selection, being that the advantage of GS over PBLUP in terms of genetic progress reached +40% in generation 15 (Table 3). The smallest advantage of GS over PBLUP in terms of genetics progress was verified for trait C, for the scenario in which breeding values were predicted using EN (15.6%).

As a general rule, for all scenarios of genetic architecture, lower levels of inbreeding were obtained under GS compared to PBLUP. In generation 15, the averages of the pedigree-based inbreeding coefficients were between 34% (EN) and 42% (wGBLUP) lower under GS scenarios, when compared to PBLUP, in the scenario in which more polygenic background was simulated (trait A, Table 3). For all measures associated to inbreeding (Fped, HML and IBD), the advantage of GS over PBLUP was even greater when a smaller number of QTL (trait B) or major QTL were simulated (trait C) (Table 3).

Variable selection methods vs GBLUP

For the more polygenic trait (trait A), the selection based on EN prediction resulted in genetic progress between 16 and 11% lower than GBLUP, being these differences significant (Table 3). At generation 15, EN also resulted in inbreeding coefficients larger than GBLUP, although such differences were not significant (Table 3).

When 100 QTL were simulated (trait B), there was no significant difference between EN and GBLUP in terms of genetic progress (Table 3). It is worth to emphasize that in the first simulated generations, deltaG was slightly lower for EN (data not shown). As more animals were included in the reference population, the genetic progress achieved with EN was comparable to GBLUP (Table 3). For this same

Table 3 Genetic progress and genetic diversity¹ under different scenarios² of selection for meat quality traits³ at generation 15*

Trait	Criterion	deltaG	SD.TBV	aveL	IBD (%)	HML (%)	Fped (%)	Nlost
A	PBLUP	2.8 ^{(0.13)e}	0.50 ^{(0.01)a}	3.8 ^{(0.07)a}	14.0 ^{(0.06)a}	0.41 ^{(0.06)a}	12.9 ^{(0.07)a}	21.6 ^{(1.72)a}
	GBLUP	4.0 ^{(0.07)b}	0.50 ^{(0.01)a}	3.3 ^{(0.02)b}	9.6 ^{(0.03)b}	0.33 ^{(0.04)bc}	7.8 ^{(0.03)bc}	21.6 ^{(1.16)a}
	EN	3.6 ^{(0.12)d}	0.50 ^{(0.01)a}	3.3 ^{(0.02)b}	10.4 ^{(0.07)b}	0.34 ^{(0.06)b}	8.5 ^{(0.06)b}	18.7 ^{(1.13)ab}
	TBV	7.0 ^{(0.08)a}	0.48 ^{(0.01)b}	3.1 ^{(0.01)c}	6.2 ^{(0.04)c}	0.26 ^{(0.04)d}	5.0 ^{(0.04)d}	19.1 ^{(1.20)ab}
	wGBLUP	3.9 ^{(0.14)bc}	0.50 ^{(0.01)a}	3.4 ^{(0.03)b}	9.0 ^{(0.07)b}	0.30 ^{(0.04)c}	7.4 ^{(0.06)c}	15.9 ^{(1.40)b}
	wEN	3.7 ^{(0.09)cd}	0.50 ^{(0.01)ab}	3.3 ^{(0.03)b}	10.3 ^{(0.09)b}	0.32 ^{(0.06)bc}	8.3 ^{(0.08)bc}	19.3 ^{(2.30)ab}
B	PBLUP	2.7 ^{(0.11)c}	0.45 ^{(0.03)ab}	3.7 ^{(0.06)a}	13.8 ^{(0.08)a}	0.41 ^{(0.08)a}	13.1 ^{(0.09)a}	3.0 ^{(0.76)a}
	GBLUP	3.7 ^{(0.21)b}	0.43 ^{(0.03)b}	3.4 ^{(0.04)b}	10.2 ^{(0.06)b}	0.32 ^{(0.05)b}	8.1 ^{(0.06)bc}	2.4 ^{(0.48)ab}
	EN	3.7 ^{(0.21)b}	0.38 ^{(0.02)cd}	3.3 ^{(0.03)c}	10.1 ^{(0.09)bc}	0.34 ^{(0.06)b}	8.4 ^{(0.08)b}	2.1 ^{(0.59)ab}
	TBV	6.1 ^{(0.28)a}	0.36 ^{(0.02)d}	3.1 ^{(0.01)d}	6.7 ^{(0.04)d}	0.25 ^{(0.03)c}	5.3 ^{(0.04)d}	1.4 ^{(0.48)b}
	wGBLUP	3.7 ^{(0.17)b}	0.46 ^{(0.02)a}	3.4 ^{(0.03)b}	8.5 ^{(0.04)c}	0.31 ^{(0.04)b}	6.9 ^{(0.03)c}	1.5 ^{(0.34)b}
	wEN	3.8 ^{(0.23)b}	0.40 ^{(0.02)c}	3.3 ^{(0.02)c}	9.0 ^{(0.08)bc}	0.30 ^{(0.07)b}	7.3 ^{(0.07)bc}	1.9 ^{(0.59)ab}
C	PBLUP	3.2 ^{(0.13)d}	0.53 ^{(0.01)a}	3.7 ^{(0.07)a}	15.2 ^{(0.06)a}	0.44 ^{(0.06)a}	14.5 ^{(0.07)a}	25.7 ^{(1.55)a}
	GBLUP	4.2 ^{(0.07)b}	0.53 ^{(0.01)a}	3.4 ^{(0.02)bc}	10.4 ^{(0.03)b}	0.35 ^{(0.04)b}	8.6 ^{(0.03)b}	22.2 ^{(1.61)ab}
	EN	3.7 ^{(0.12)c}	0.50 ^{(0.01)ab}	3.3 ^{(0.02)cd}	10.2 ^{(0.07)b}	0.32 ^{(0.06)bc}	8.3 ^{(0.06)bc}	22.1 ^{(1.20)abc}
	TBV	7.1 ^{(0.08)a}	0.47 ^{(0.01)b}	3.1 ^{(0.01)e}	6.2 ^{(0.04)d}	0.25 ^{(0.04)d}	5.0 ^{(0.04)d}	20.9 ^{(1.51)bc}
	wGBLUP	4.2 ^{(0.14)b}	0.53 ^{(0.01)a}	3.4 ^{(0.03)b}	8.5 ^{(0.07)c}	0.30 ^{(0.04)c}	7.1 ^{(0.06)c}	18.4 ^{(1.31)c}
	wEN	3.9 ^{(0.09)bc}	0.49 ^{(0.01)b}	3.3 ^{(0.03)d}	9.4 ^{(0.09)bc}	0.30 ^{(0.06)c}	7.5 ^{(0.08)bc}	19.8 ^{(2.00)bc}

¹deltaG accumulated genetic gain (in units of additive SD of the base population), SD.TBV SD of true breeding values (TBV), IBD proportion of genome that is homozygous due to identity-by-descent, in %, HML homozygosity for alleles potentially associated to deleterious mutations, in %, Fped average of individual pedigree-based inbreeding coefficients, in %, Nlost accumulated number of favorable alleles lost, aveL average generation interval, \bar{y} average of 10 replicates, SE standard error

²Each scenario is defined by a selection criterion: TBV, estimated breeding value (PBLUP), genomic estimated breeding value using GBLUP or EN (GBLUP or EN) and GBLUP or EN giving more weight to low-frequency favorable alleles (wGBLUP or wEN)

³trait A (1000 QTL), trait B (100 QTL), trait C (1000 QTL, five of which had larger effect and explained 7% of the additive variance). A fixed replacement rate of 20% was adopted

*Averages of 10 replicates (standard error, SE, within brackets). For a same statistic (column), averages with the same superscript letter do not differ (t-test, adjusted p-value > 0.05)

trait, there was also no significant difference between EN and GBLUP in terms of inbreeding incidence (Table 3).

When five major QTL were simulated in a scenario with 1000 QTL (trait C), EN resulted in genetic gain significantly lower than GBLUP (difference about 10%, Table 3). In this situation, while inbreeding incidence was significantly lower for EN in the first generations (data not shown), such differences tended to diminish along the selection process, so that there was no difference between EN and GBLUP in terms of inbreeding incidence at generation 15 (Table 3).

Weighting on low frequency favorable alleles (wGBLUP and wEN)

In all scenarios of genetic architecture, there was no significant difference between GBLUP and wGBLUP in terms of genetic gain, what also occurred when EN and wEN were contrasted (Table 3). For the more polygenic trait, although the averages of Fped and IBD were slightly lower for wGBLUP and wEN (compared to GBLUP and EN, respectively), such differences were not significant (trait A, Table 3), being

that the only benefit of weighting on low-frequency favorable alleles was obtained for wGBLUP with respect to the number of favorable alleles lost (Nlost 26% lower at generation 15, compared to GBLUP).

For the less polygenic traits (B and C), inbreeding incidence was significantly lower under wGBLUP in the last generations (Table 3). No significant difference was found between wEN and EN for all statistics related to inbreeding incidence (Table 3).

Discussion

Long-term consequences of GS

For schemes simulating selection for female reproduction (E1) and meat quality (E2), GS allowed considerable gains in terms of genetic progress, regardless of the time horizon considered. In these situations, the genetic gain accumulated until the last generation (15th) was up to 40% greater for GS when compared to selection based on PBLUP.

As a general rule, for scenarios in which greater genetic progress was achieved using GS, smaller pedigree-based inbreeding coefficients were also estimated, which were relatively consistent with the smaller levels of homozygosity due to identity-by-descent tracked by the simulation of IBD loci. The relative benefit of GS in terms of inbreeding incidence increased along the generations, so that inbreeding levels estimated using pedigree or IBD information were at least 25% smaller under GS in the last generation, when compared to similar figures estimated under PBLUP selection. Similarly, GS also allowed significant benefits in terms of diminishing the loss of favorable alleles as well as reducing homozygosity for alleles potentially associated to deleterious mutations.

Such results partially confirm the theoretical expectations found in Daetwyler et al. (2007) and Dekkers (2007), who suggested that lower levels of inbreeding could be achieved through the application of GS, as a result of the lower emphasis on family information under this strategy, due to the possibility of obtaining better estimates of the Mendelian sampling (MS) term.

The comparison between different scenarios of genetic architecture (scheme E2) suggests that larger advantage of GS over PBLUP in terms of inbreeding incidence can be expected under less polygenic background. The largest benefit for GS over PBLUP in terms of inbreeding levels was observed for the scenarios where a smaller number of QTL or QTL of larger effect were simulated. Nevertheless, while Bastiaansen et al. (2012) also reported lower incidence of inbreeding for GS when compared to PBLUP, these authors did not find any evidence of association between genetic architecture and inbreeding incidence.

When different replacement strategies were compared, larger genetic progress was achieved in scenarios without restriction on the replacement rate, what, in practice, meant that a larger proportion of younger animals were selected. For both schemes in which replacement strategies were compared (E1 and E2), a larger replacement rate resulted in larger benefit for PBLUP scenarios, where the accumulated genetic gain was about 40% larger compared to a situation when a fixed replacement rate of 20% was applied (v2 vs v1). Under GS scenarios, the advantage in terms of accumulated genetic gain attributable to the allowance of variable replacement rate was about 30%.

Conversely, the inbreeding coefficients were larger for most of the scenarios in which a larger replacement rate was applied, and noticeably larger for selection based on PBLUP. The larger reduction in generation intervals under the scenarios with no constraint on the replacement rate would imply in a larger increase in inbreeding levels. Generation intervals were smaller for GS scenarios than for PBLUP in most situations and regardless of the replacement strategy.

According to Pryce et al. (2012), selection schemes in which the annual increase in inbreeding is high would require strategies such as OCS, which can be more effective to manage genetic diversity. In the present study, although a considerable advantage was associated to a larger replacement rate in terms of genetic progress, the sharp reduction in the generation intervals under this strategy and the consequent increase in inbreeding rates draw attention for the need to design alternative GS schemes aiming to prevent severe increases in inbreeding levels, possibly considering OCS.

Genomic information was assumed to be available for all selection candidates in the present study. Although the strategic use of imputation, with genotyping of all selection candidates from an extremely-low density has been suggested as a cost-effective strategy in swine breeding (Huang et al. 2012), this can be judged as a scenario still hard to be met in some real beef cattle populations, given the costs involved in genotyping and considering that extremely-low density panels may not be feasible to apply in some populations, because of large effective population size and smaller extent of linkage disequilibrium. Previous studies on long-term consequences of GS also assumed all selection candidates being genotyped (e.g. Sonesson et al. 2012; Bastiaansen et al. 2012). Such assumption was judged to be suitable to allow testing the hypotheses investigated in the present study and making it possible to compare its results to those from studies carried out under similar conditions.

Performance of GS methods under different scenarios of genetic architecture

For the meat quality selection scheme (E2), where different scenarios of genetic architecture were simulated, no significant benefit was found for variable selection methods over GBLUP in all time horizons considered. Especially for the scenario where 100 QTL were simulated, we expected larger accuracy for the predictions obtained with EN (Daetwyler et al. 2010; Clark et al. 2011).

The EN method, such as implemented in this study (i.e. weight for LASSO penalty close to 1), was possibly more sensitive to the small size of the reference population size in the first generations simulated. It is known that LASSO method selects at most as many variables as there are observations in the reference population (Zou and Hastie 2005; Ogutu et al. 2012). It can be expected that, in the first generations, due to the small size of the reference population, the number of markers selected by EN may have been smaller than needed to cover all QTL regions, resulting in smaller accuracy for EN prediction compared to PBLUP. The consideration of a small weight to the ridge penalty in the EN predictions in the present study aimed to alleviate this problem, but it did not seem to be effective. Such hypothesis could be corroborated by the fact that, as the reference

population increased, the genetic gain accumulated under EN prediction was slightly larger than that under GBLUP, when 100 QTL were simulated. Even though the difference was not significant, this result would suggest some benefit of EN prediction in this scenario, also drawing attention for the limitations of attributing large weight for the LASSO penalty when reference populations are small.

No significant differences between EN and GBLUP were found in terms of inbreeding incidence, although for the trait under more polygenic control (E2, trait A) the inbreeding levels were slightly larger for EN. Bastiaansen et al. (2012) also reported that larger pedigree-based inbreeding coefficients were obtained under two variable selection methods (Bayesian regression and partial least squares), when compared to GBLUP (coefficients about 10% larger), although differences between methods in terms of genomic inbreeding were not reported.

Sonesson et al. (2012) simulated a trait influenced by 1000 QTL, which was measured on full-sibs of the selection candidates. In this situation, using a variable selection method (Bayes B) to compute genomic predictions did not result in larger genetic progress when compared to GBLUP, under truncation selection, while it resulted in larger IBD coefficient (+ 12% in the last generation simulated). According to these authors, this result was associated to the fact that variable selection methods apply larger selective pressure nearby QTL of larger effect, increasing the levels of homozygosity in these regions (phenomenon known as “hitchhiking effect”, Smith and Haigh 1947) and thus increasing IBD coefficients. In the present study, the larger estimates of IBD obtained for EN in the last generations of a scenario with comparable genetic architecture (E2, trait A) would corroborate such hypothesis.

Impact of weighting on low-frequency favorable alleles

There was no conclusive evidence in favor of the hypothesis that attributing more weight to favorable alleles of lower frequency (as applied in the definition of wGBLUP and wEN) would enhance long-term genetic gain, although in most of the scenarios using this weighting strategy tended to produce lower levels of inbreeding and reduced the loss of favorable alleles. When traits under less polygenic control were simulated, slightly larger genetic gain was achieved with wGBLUP, when compared to GBLUP at generation 15, but such differences were not significant. Under a more polygenic scenario (3000 QTL), simulated by Sun and VanRaden (2014), the weighting strategy employed in wGBLUP lead to smaller accumulated genetic gain compared to an unweighted scenario, although these authors simulated a more simplistic situation, where true QTL effects (and not estimated marker effects) were considered in predictions.

Conversely, Jannink (2010) reported that greater long-term genetic gain was achieved under adoption of a selection criterion similar to wGBLUP. Some of the explanations for such divergences could be related to the genetic architecture of the simulated traits in each case. This author simulated 100 QTL all explaining an equal proportion of the additive variance, in a way that QTL of lower MAF had larger allele substitution effects. This constraint was not applied in the present study, what could have reduced the relative importance of losing favorable QTL alleles that did not contribute much to the additive variance. Moreover, Jannink (2010) simulated bi-allelic QTL with same properties of the markers, suggesting that this approach may have resulted in higher LD between QTL and markers than in a situation in which they had different properties, as is the case of the present study.

In addition, due to the quadratic trend observed for N_{lost} , it could be expected that more pronounced differences between methods (e.g. GBLUP vs wGBLUP) in terms of loss of favorable alleles could be observed if more generations were simulated. Another point that is worth to mention is that the function employed to derive wGBLUP, attributing more weight to low-frequency favorable alleles, could have been ineffective to ensure optimal weighting in the simulated population. Sun and VanRaden (2014) suggested new functions that would allow better balance between short- and long-term gains. Further studies are needed to investigate alternative weighting strategies, possibly exploring the alternatives proposed by Sun and VanRaden (2014), in scenarios mimicking beef cattle populations.

In the present study, effects of inbreeding depression were not simulated, what means that larger levels of inbreeding in a given scenario did not reflect in worsen phenotypic performance. Because inbreeding coefficients under wGBLUP (and wEN) tended to be smaller, the benefit of weighting on low-frequency favorable alleles could be larger for traits affected by inbreeding depression. Thus, future studies could investigate the consequences of selection based on criteria similar to wGBLUP also considering some form of discount rate for losses due to inbreeding incidence, e.g. similarly as in Pryce et al. (2012).

Conclusions

GS allowed considerable long-term benefit to female reproduction and meat quality traits, when compared to PBLUP.

In terms of controlling inbreeding rates, larger advantage for GS over PBLUP was obtained for less polygenic traits.

More weight for favorable alleles of lower frequency tended to produce lower levels of inbreeding and to reduce the loss of favorable alleles in GS.

Acknowledgements The first author was supported by grants from State of São Paulo Foundation for Research Supporting (FAPESP), Process 2010/06185-4, within the postgraduate program on Genetics and Animal Breeding at FCAV/UNESP, Brazil. The last two authors receive a research scholarship from National Council for Scientific and Technological Development (CNPq), Processes: 308636/2014-7 and 307440/2013-3, respectively.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Bastiaansen JWM, Coster A, Calus MPL, Van Arendonk JAM, Bovenhuis H (2012) Long-term response to genomic selection: effects of estimation method and reference population structure for different genetic architectures. *Genet Sel Evol* 44:3
- Bijma P (2012) Long-term genomic improvement—new challenges for population genetics. *J Anim Breed Genet* 129:1–2. <https://doi.org/10.1111/j.1439-0388.2011.00985.x>
- Brito FV, Braccini Neto J, Sargolzaei M, Cobuci JA, Schenkel FS (2011) Accuracy of genomic selection in simulated populations mimicking the extent of linkage disequilibrium in beef cattle. *BMC Genet* 12:80. <https://doi.org/10.1186/1471-2156-12-80>
- Carvalho R (2014) Genomic selection in Nelore cattle in Brazil. In: Proceedings of the 10th WCGALP, Vancouver, Canada
- Carvalho R, McEwan JC, Neves HHR, Utsunomiya YT, O'Brien AMP, Boison SA, Sölkner J, Schenkel FS, Van Tassell CP, Sonstegard TS, Garcia JF (2013) Accuracy of genomic predictions in Nelore beef cattle. In: Proceedings of the 20th AAABG Conference, Napier, New Zealand
- Clark SA, Hickey JM, Van der Werf JHJ (2011) Different models of genetic variation and their effect on genomic evaluation. *Genet Sel Evol* 43:18
- Daetwyler HD, Villanueva B, Bijma P, Wolliams JA (2007) Inbreeding in genome-wide selection. *J Anim Breed Genet* 124:369–376. <https://doi.org/10.1111/j.1439-0388.2007.00693.x>
- Daetwyler HD, Pong-Wong R, Villanueva B, Wolliams JA (2010) The impact of genetic architecture on genome-wide evaluation methods. *Genetics* 185:1021–1031. <https://doi.org/10.1534/genetics.110.116855>
- Daetwyler HD, Swan AA, Van der Werf JH, Hayes BJ (2012) Accuracy of pedigree and genomic predictions of carcass and novel meat quality traits in multi-breed sheep data assessed by cross-validation. *Genet Sel Evol* 44:33. <https://doi.org/10.1186/1297-9686-44-33>
- Dekkers JCM (2007) Prediction of response to marker-assisted and genomic selection using selection index theory. *J Anim Breed Genet* 124:331–341. <https://doi.org/10.1111/j.1439-0388.2007.00701.x>
- Ferencakovic M, Hamzic E, Gredler B, Solberg TR, Klemetsdal G, Curik I, Sölkner J (2013) Estimates of autozygosity derived from runs of homozygosity: empirical evidence from selected cattle populations. *J Anim Breed Genet* 130:286–293. <https://doi.org/10.1111/jbg.12012>
- Goddard ME, Hayes BJ, Meuwissen THE (2010) Genomic selection in livestock populations. *Genet Res Camb* 92:413–421. <https://doi.org/10.1017/S0016672310000613>
- Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME (2009) Invited review: genomic selection in dairy cattle: Progress and challenges. *J Dairy Sci* 92:433–443. <https://doi.org/10.3168/jds.2008-1646>
- Hill WG, Robertson A (1968) Linkage disequilibrium in finite populations. *Theor Appl Genet* 38:226–231. <https://doi.org/10.1007/bf01245622>
- Hoerl AE, Kennard RW (1970) Ridge regression: biased estimation for nonorthogonal problems. *Technometrics* 12:55–67. <https://doi.org/10.2307/1267351>
- Huang Y, Hickey JM, Cleveland MA (2012) Assessment of alternative genotyping strategies to maximize imputation accuracy at minimal cost. *Genet Sel Evol* 44:25
- Jannink JL (2010) Dynamics of long-term genomic selection. *Genet Sel Evol* 42:35. <https://doi.org/10.1186/1297-9686-42-35>
- Keller MC, Visscher PM, Goddard ME (2011) Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data. *Genetics* 89:237–249. <https://doi.org/10.1534/genetics.111.130922>
- Kristensen T, Sørensen A (2005) Inbreeding—lessons from animal breeding, evolutionary biology and conservation genetics. *Anim Sci* 80(2):121–133. <https://doi.org/10.1079/ASC41960121>
- Lu D, Sargolzaei M, Kelly M, Li C, Van der Voort G, Wang Z, Plastow G, Moore S, Miller SP (2012) Linkage disequilibrium in Angus, Charolais, and Crossbred beef cattle. *Front Genet* 3:152
- Misztal I (2014) BLUPF90 - a flexible mixed model program in Fortran 90. <http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=blupf90.pdf>. Accessed 19 July 2014
- Muir WM (2007) Comparison of genomic and traditional BLUP estimated breeding value accuracy and selection response under alternative trait and genomic parameters. *J Anim Breed Genet* 124:342–355. <https://doi.org/10.1111/j.1439-0388.2007.00700.x>
- Neves HHR, Carvalho R, Queiroz SA (2012) A comparison of statistical methods for genomic selection in a mice population. *BMC Genet* 13:100. <https://doi.org/10.1186/1471-2156-13-100>
- Ogutu JO, Schulz-Streeck T, Piepho HP (2012) Genomic selection using regularized linear regression models: ridge regression, lasso, elastic net and their extensions. *BMC Proceed* 6:S10. <https://doi.org/10.1186/1753-6561-6-S2-S10>
- Pryce JE, Hayes BJ, Goddard ME (2012) Novel strategies to minimize progeny inbreeding while maximizing genetic gain using genomic information. *J Dairy Sci* 95:377–388. <https://doi.org/10.3168/jds.2011-4254>
- R Development Core Team (2013) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013. ISBN 3-900051-07-0. <http://www.R-project.org>. Accessed 10 Jan 2014
- Sargolzaei M, Schenkel FS (2009) QMSim: a large-scale genome simulator for livestock. *Bioinformatics* 25:680–681. <https://doi.org/10.1093/bioinformatics/btp045>
- Sargolzaei M, Schenkel FS, Jansen GB, Schaeffer LR (2008) Extent of linkage disequilibrium in Holstein Cattle in North America. *J Dairy Sci* 91:2106–2117. <https://doi.org/10.3168/jds.2007-0553>
- Schaeffer LR (2006) Strategy for applying genome wide selection in dairy cattle. *J Anim Breed Genet* 123:218–223. <https://doi.org/10.1111/j.1439-0388.2006.00595.x>
- Smith MJ, Haigh J (1947) The hitch-hiking effect of a favorable gene. *Genet Res* 23:23–35
- Sonesson AK, Wolliams JA, Meuwissen THE (2012) Genomic selection requires genomic control of inbreeding. *Genet Sel Evol* 44:27. <https://doi.org/10.1186/1297-9686-44-27>
- Sun C, VanRaden PM (2014) Increasing long-term response by selecting for favorable minor alleles. *Plos One* 9:e88510. <https://doi.org/10.1371/journal.pone.0088510>
- Tibshirani R (1996) Regression shrinkage and selection via the lasso. *J R Stat Soc B* 58:267–288

- VanRaden PM, Null DJ, Sargolzaei M, Wiggans GR, Tooker ME, Cole JB, Sonstegard TS, Connor EE, Winters M, Van Kaam J, Van Doormaal BJ, Faust MA, Doak GA (2013) Genomic imputation and evaluation using high density Holstein genotypes. *J Dairy Sci* 96:668–678. <https://doi.org/10.3168/jds.2012-5702>
- Waldmann P, Mészáros G, Gredler B, Fuerst C, Sölkner J (2013) Evaluation of the lasso and the elastic net in genome-wide association studies. *Front Genet* 4:270. <https://doi.org/10.3389/fgene.2013.00270>
- Zou H, Hastie T (2005) Regularization and variable selection via the elastic net. *J Roy Stat Soc B* 67:301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>