Prediction of gametic variance and its use in bovine breeding programs

C Hozé, A Baur, S Fritz, D Boichard

▶ **To cite this version:**

C Hozé, A Baur, S Fritz, D Boichard. Prediction of gametic variance and its use in bovine breeding programs. 12th World Congress on Genetics Applied to Livestock Production, Jul 2022, Rotterdam, Netherlands. hal-03731280

HAL Id: hal-03731280
https://hal.inrae.fr/hal-03731280

Submitted on 20 Jul 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Prediction of gametic variance and its use in bovine breeding programs

C. Hozé[1, 2], A. Baur[1,2], S. Fritz[1, 2], D. Boichard[2]

[1]Allice, 75595 Paris, France, [2]Université Paris Saclay, INRAE, AgroParisTech, GABI, 78350 Jouy-en-Josas, France; chris.hoze@allice.fr

Abstract

With genomic selection, gametic variance has been of increasing interest for breeding top animals. Two gametic variance estimates based on SNP effects estimated in 2018, were obtained deterministically and by simulating progeny. They were compared with each other, with the variance of the genomic breeding values of real progeny in fall 2020, and with the variability of the phenotypes of the female offspring. As expected, both predictors of gametic variances were similar, provided that the number of simulated progeny is large enough (500). The correlation between those predictors and gametic variances observed on real progeny was rather high, although limited by the number of offspring available. The correlation between the gametic variance of the sires and the phenotypic variance ranged from 0.14 to 0.85 (mean=0.33) according to the heritability of the trait. Gametic variance is useful to select parents of future bulls, but its potential to control phenotypic variability is more limited.

Introduction

Since its implementation in France in 2009, genomic selection has had a tremendous development, all bulls are genomically evaluated and more and more females are genotyped annually. Young bulls are selected on the basis of their genomic values among a large number of male progeny of bull sires and bull dams. Under the polygenic model, gametic variances of each parent or, equivalently, the Mendelian sampling (MS) variance of the progeny depend only on inbreeding F_s and F_d of the parents [$Var(MS) = 0.5\ \sigma_g^2\ (1 - 0.5\ F_s - 0.5F_d)$] and is constant without inbreeding . In genomic selection, not only the MS term of the progeny can be predicted but also the gametic variance of each parent, provided that SNP effects and genotype phases of the parents are known. According to the large investments necessary to select the very best young bulls (matings, genotyping, embryo production and transfer), it is essential to maximize the probability to generate and find them. This probability depends not only on the genetic merit of the parents but also on their gametic variance. As already shown in the literature, gametic variance is of major importance when breeding top animals (Bijma et al., 2020; Segelke et al., 2014). However, in spite of its increasing interest, little investigation was done on its practical prediction and use. In this study, we (1) compared two approaches to predict the gametic variance of the genotyped parents; (2) investigated the impact of gametic variance in the selection success of candidates; (3) compared the predicted gametic variances with the variance of genomic breeding values of real genotyped progeny, and (4) with the phenotypic variability of the daughters. In this validation procedure, SNP effects were estimated in French Holstein population with the information available in June 2018 and genomic breeding values and phenotypes of progeny were those available in fall 2021.

Materials & Methods

Population studied.

We used genotypes and breeding values for 426 widely used bulls in Holstein breeds born between 1991 and 2016. Bulls were selected to have more than 200 daughters with morphology phenotypes, 1000 daughters with lactation phenotypes and 500 genotyped progeny. Genotypes of sires were obtained with different Illumina chip platforms with at least 50K SNPs, and imputed and phased with Fimpute3 software (Sargolzaei et al., 2014).

SNP effects were obtained from the French national genomic evaluation (Boichard et al, 2012) computed in June 2018. These effects were used as input for simulation software and gametic variability prediction. Five production and 10 morphological traits were studied.

Simulation procedure and prediction of gametic variance.

Two different approaches were used to predict the gametic variance of each bull. The first one was based on simulation (Segelke et al, 2014). For each bull mated to one simulated fully homozygous dam, 500 progeny were simulated. For each progeny, segregations were randomly sampled and recombinations were simulated assuming 1cM corresponds to 1 Mbase. Progeny direct genomic value was obtained by combining the SNP effects with its genotypes. The MS variance was estimated by the variance of the DGV (direct genomic value) of the 500 progeny. Because SNP effects are estimated with an accuracy lower than 1, the MS variance was underestimated.

The second approach is the one described by Santos et al (2019). The gametic variance (V_{theo}) of each sire was predicted from the contrasts within loci with the following formula:

$$V_{theo} = 0.25 \left[\Sigma_i \alpha_i^2 + 2 \Sigma_i \Sigma_{j>i} \alpha_i \alpha_j (1-2r_{ij}) \right] \quad (1)$$

where : α_ι is the signed contrast between effects of the paternal and maternal alleles for marker i, and r_{ij} is the recombination rate between markers i and j. Recombination rate was assumed equal to 0.5 across chromosomes.

DGV and yield deviations (YD) obtained from the French National evaluation run computed in November 2021 were used to validate gametic variance predictions on real data. YD were performances of the female offspring adjusted for fixed effects of the evaluation model and dam genetic merit. For each sire, we computed the variance (V_{obs}) of DGV of progeny genotyped after 2018. The phenotypic variance (V_{yd}) was obtained by computing YD variance. Due to the increase in reference population size and change in French evaluation model (especially change in imputation results due to switch from UMD3 to ARS-UCD1 assembly), markers and haplotypes effects might differ between June 2018 and November 2021, allowing us to test the robustness of our predictors.

Results

Comparison of gametic variance predictions.

The predicted and simulation based gametic variances were very similar, with the same mean, very close standard deviations, and correlations in the [0.82 – 0.94] interval. A larger number of simulated progeny would probably have led to higher correlations. Therefore in practice, the deterministic approach is preferable because it is more accurate and faster to compute. But simulations can be a more flexible tool especially to study covariations between traits or more complex genetic determinisms with dominance or epistasis.

The same indicators computed with 2021 SNP estimates were consistent with those of 2018, showing a good conservation of SNP effect: the correlations between V_{theo} in 2018 and 2021 ranged from 0.87 to 0.99 according to traits. Correlations are presented on Table 2.

Heterogeneity of gametic variance between bulls.

Gametic variances observed on real progeny DGVs (V_{obs}) varied widely between sires and ranged from simple to triple for most traits. This is much more than expected under the polygenic model. Figure 1 shows the distribution of V_{obs} for fat content (left) and stature (right). Differences were particularly large for fat content due to the effect of DGAT1 major gene, and lower but still important for stature with most values ranging from 0.3 to 0.9. This large variability of gametic variance observed across animals highlights its interest when breeding

top animals. Assuming these predictions are accurate and reflect true gametic variance (this depends on the accuracy of SNP effects), very good predictions of extreme deciles can be obtained, as well as the probability to obtain an elite progeny above a given threshold. For instance, for milk yield, the prediction of the 90th percentile was: P90% = 0.999 Genomic Pedigree Index + 1.29 σ_{MS} (R^2 = 0.99), with σ_{MS} the MS standard deviation for a given couple of parents.

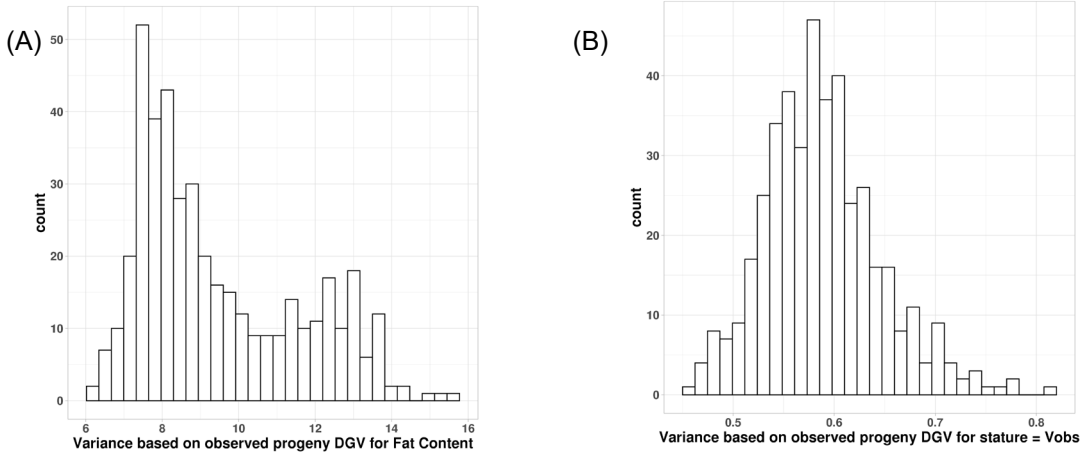

Figure 1. Distribution of progeny DGV variances (V_{obs}) for fat content (A) and stature (B)

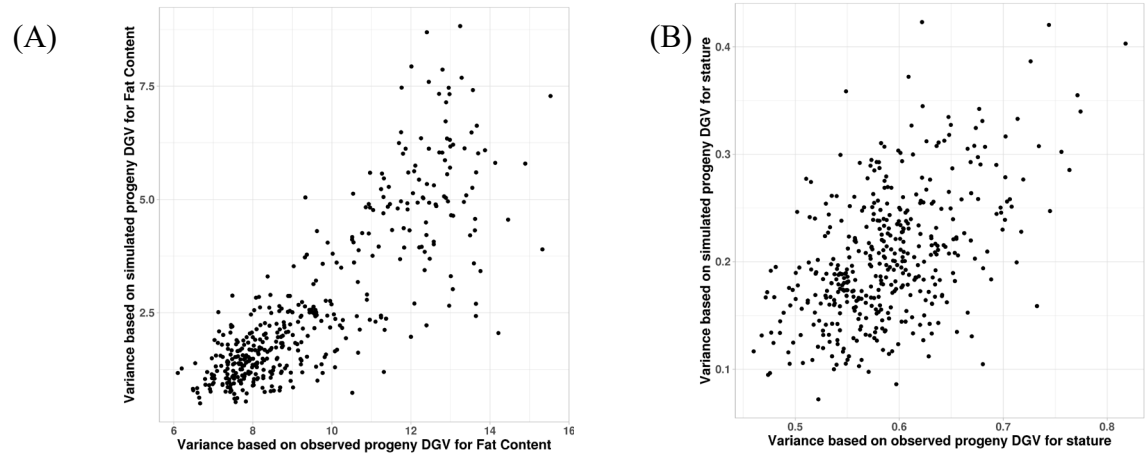

Figure 2. Relationship between DGV within sire variances for simulated and observed progeny, for fat content (A) and stature (B)

Validation of predicted gametic variance

Depending on the traits, the correlations between predicted variance V_{theo} in 2018 and observed variance V_{obs} in 2021 ranged from 0.47 to 0.92 and reached 0.62 on average. Figure 2 presents these results for fat content and stature. These moderate values probably reflect both the limited number of progeny and the variability due to dams.

The correlation between V_{theo} and the within bull variance of phenotypes was quite low with an average of 0.25 across traits. On the youngest animal (58 bulls born after 2015), correlations were a bit higher with an average of 0.33. Except for a high correlation for fat content (0.85), correlations were moderate for heritable traits (0.44 for protein yield or 0.38 for body depth) and as low as 0.14 for low heritability traits such as rear udder height. All correlations are presented on table 2.

Discussion

This study in Holstein cattle breed investigates the opportunity to predict progeny variability from genotypes and SNP effects.

We observed large variation in progeny variability among bulls supporting the growing interest for an indicator of gametic variability, especially in the choice of bull sires and dams to maximize the probability to conceive top ranking bulls. Validation on real data showed that the proposed theoretical indicator performed quite well to predict the genetic variability of progeny but was not accurate enough to predict the phenotypic performance variability, highly dependent on environmental factors. Computation of the theoretical predictor with formula (1) is fast, meaning that it can be obtained routinely and used to select animals for the breeding scheme. In this perspective, development of specific selection criteria such as those developed by Bijma et al. (2020) is expected on a routine basis. In spite of its cost of computation, the estimation of gametic variance by simulation of progeny may also be interesting especially in complex situations, mainly in case of non-additive determinism.

Table 2. Correlations between theoretical gametic variance computed in June 2018 (V_{theo}_June18), others indicators of gametic variance and observed progeny variance.

Traits	V_{sim}_June18	V_{sim}_Nov21	V_{theo}_Nov21	V_{obs}_Nov21	V_{yd}_Nov21
Protein Yield	0.91	0.76	0.87	0.42	0.44
Fat Yield	0.94	0.94	0.98	0.77	0.16
Protein content	0.88	0.89	0.99	0.84	0.46
Fat Content	0.95	0.95	1.00	0.93	0.85
Milk	0.90	0.84	0.97	0.73	0.15
Udder support	0.87	0.77	0.88	0.59	0.17
Udder Depth	0.84	0.78	0.93	0.37	0.35
Fore Udder Attachment	0.94	0.88	0.95	0.63	0.23
Rear Udder Height	0.91	0.88	0.92	0.55	0.14
Fore Teat Distance	0.90	0.89	0.95	0.64	0.32
Stature	0.80	0.83	0.94	0.64	0.47
Body depth	0.90	0.85	0.94	0.71	0.38
Width at ischium	0.96	0.92	0.96	0.66	0.27
Locomotion	0.87	0.83	0.89	0.51	0.24
Body Condition Score	0.93	0.83	0.94	0.72	0.35
Average	0.90	0.86	0.94	0.65	0.33

[1] Theoretical gametic variance computed either in November 2021 (**Vtheo_Nov21**), Simulated gametic variance computed either in November 2021 (**Vsim_Nov21**) or June 2018 (**Vsim_June18)**, Observed gametic variance computed in November 2021 on DGV (**Vobs_Nov21)** or on YD (**Vyd_Nov21**)

Acknowledgments:
This research has received funding from European Union's Horizon 2020 research and innovation program - GenTORE - under grant agreement N° 727213.

References

Bijma P., Wientjes Y.C.J., and Calus M.P.L (2020) *Genetics*. 214 (1). 91-107. https://doi.org/10.1534/genetics.119.302643

Sargolzaei M., Chesnais J., and Schenkel F. (2014) *BMC Genomics* 15(1):478. http://doi.org/10.1186/1471-2164-15-478

Boichard D., Guillaume F., Baur A., Croiseau P., Rossignol M. N. et al. (2012) *Anim Prod Sci* 52. 115-120. https://doi.org/10.1071/AN11119

Segelke D., Reinhardt F., Liu Z. and Thaller G. (2014) *Genet Sel Evol* 46. 42. https://doi.org/10.1186/1297-9686-46-42

Santos D.J.A., Cole J.B., Lawlor T.J.Jr., VanRaden P.M., Tonhati H., et al (2019). J Dairy Sci 102(6) 5279-5294. https://doi.org/10.3168/jds.2018-15971