Dissecting genetic trends to understand breeding practices in livestock: a maternal pig line example

Rostam Abdollahi-Arpanahi, Daniela Lourenco, Andres Legarra, Ignacy Misztal

HAL Id: hal-03481132
https://hal.inrae.fr/hal-03481132

Submitted on 23 Feb 2022

Genetics **S**election **E**volution
GSE

RESEARCH ARTICLE

Dissecting genetic trends to understand breeding practices in livestock: a maternal pig line example

Rostam Abdollahi-Arpanahi[1]*, Daniela Lourenco[1], Andres Legarra[2] and Ignacy Misztal[1]

Abstract

Background: Understanding whether genomic selection has been effective in livestock and when the results of genomic selection became visible are essential questions which we have addressed in this paper. Three criteria were used to identify practices of breeding programs over time: (1) the point of divergence of estimated genetic trends based on pedigree-based best linear unbiased prediction (BLUP) versus single-step genomic BLUP (ssGBLUP), (2) the point of divergence of realized Mendelian sampling (RMS) trends based on BLUP and ssGBLUP, and (3) the partition of genetic trends into that contributed by genotyped and non-genotyped individuals and by males and females.

Methods: We used data on 282,035 animals from a commercial maternal line of pigs, of which 32,856 were genotyped for 36,612 single nucleotide polymorphisms (SNPs) after quality control. Phenotypic data included 228,427, 101,225, and 11,444 records for birth weight, average daily gain in the nursery, and feed intake, respectively. Breeding values were predicted in a multiple-trait framework using BLUP and ssGBLUP.

Results: The points of divergence of the genetic and RMS trends estimated by BLUP and ssGBLUP indicated that genomic selection effectively started in 2019. Partitioning the overall genetic trends into that for genotyped and non-genotyped individuals revealed that the contribution of genotyped animals to the overall genetic trend increased rapidly from ~ 74% in 2016 to 90% in 2019. The contribution of the female pathway to the genetic trend also increased since genomic selection was implemented in this pig population, which reflects the changes in the genotyping strategy in recent years.

Conclusions: Our results show that an assessment of breeding program practices can be done based on the point of divergence of genetic and RMS trends between BLUP and ssGBLUP and based on the partitioning of the genetic trend into contributions from different selection pathways. However, it should be noted that genetic trends can diverge before the onset of genomic selection if superior animals are genotyped retroactively. For the pig population example, the results showed that genomic selection was effective in this population.

Background

Genomic selection has revolutionized the livestock breeding industry and has doubled genetic gain in some species, such as dairy cattle [1, 2]. Other livestock breeding industries, such as for pigs, broilers, beef, etc. have also heavily invested in this technology [3]. Hence, understanding whether implementation of genomic selection has been effective in these species and when the results of genomic selection became visible, are essential questions which need to be addressed.

The genetic progress of a population is a combination of the progress in different selection pathways [4], which are under different selection pressures and that have

*Correspondence: rostam.abdollahi@uga.edu
[1] Department of Animal and Dairy Science, University of Georgia, Athens, GA 30602, USA
Full list of author information is available at the end of the article

different accuracies of selection. The impact of the implementation of genomic selection on different selection pathways can be measured by the decomposition of the overall genetic trend into that contributed by each pathway over the course of selection [5].

When a population is under genomic selection, genetic evaluation based on the conventional pedigree-based method i.e., pedigree-based best linear unbiased prediction (PBLUP), is biased because it neither considers genomic information nor takes genomic preselection into account [6–8]. Genomic preselection occurs when only animals that are selected based on genomic estimated breeding values (GEBV), and therefore on positive Mendelian sampling terms, are phenotyped [9, 10]. In this case, the breeding values predicted by PBLUP for these animals are, on average, underestimated [11], while single-step genomic BLUP (ssGBLUP) is expected to provide unbiased genetic trends [12, 13] because it combines all the information that has been used for selection decisions (pedigree, genomic data, and phenotypic records). Therefore, the point when estimated genetic trends based on PBLUP and ssGBLUP diverge indicates the starting date of effective genomic selection [13]. This is the time-point when young animals are selected based on estimates of their Mendelian sampling terms based on genomics, in addition to parent average and, perhaps, own information. If all genotyped individuals are phenotyped, then the difference between genetic trends by PBLUP and ssGBLUP depends on the corresponding prediction accuracy. In principle, GEBV are more accurate than EBV and result in better genetic progress. However, if multiple-trait selection is practiced, it is important that the genetic trends are estimated using a multiple-trait model.

The average genetic merit of an individual can be decomposed as the sum of the parent average and a Mendelian sampling term that represents the deviation of an animal's breeding value from the average of its parents. Thus, the superiority of the selected candidates over the mean of their parents represents the extra gain compared to the previous generation, by capturing the Mendelian sampling terms. Realized Mendelian sampling (RMS) terms have zero expectation when all animals are genotyped or when genotyping is at random. When animals (e.g., young boars that are kept till they reach sexual maturity and reproduce) are selected based on estimated parent average, their RMS is also 0. However, the average RMS will not be 0 when the genotyped animals are selected based on own phenotype or progeny performance, i.e. with selective genotyping. Moreover, with selective genotyping, the average RMS will be greater with ssGBLUP than with BLUP because the latter does not account for genomic information and preselection. Consequently, a deviation of the average RMS

from 0 would show selective genotyping, and a divergent trend in average estimates of RMS between ssGBLUP and BLUP would display the starting date of genomic selection.

García-Cortés et al. [14] introduced a procedure for partitioning the genetic gain into contributions from parent averages and RMS and for allocating these contributions into pre-defined "paths" (e.g., by country, gender, line, etc.), summarizing path-specific terms to quantify the contributions of different sources to the overall genetic trend. This procedure has been used to quantify the contribution of different countries to the overall genetic trend in Brown Swiss bulls [15] and to explore the impact of national selection and importation in Landrace and Large-White pigs in Croatia [16].

Partitioning the total genetic trend into contributions from genotyped and non-genotyped animals or into contributions from males and females can be used to determine the impact of different selection pathways over time. For instance, if genotyped animals have a greater contribution to genetic gain than non-genotyped individuals, it can indicate selective genotyping (elite animals are genotyped) or that genomic selection is effective and most of the parents are selected from the genotyped candidates. In species such as pigs, the number of progeny is smaller per male and larger per female than in dairy cattle. Therefore, the impact of female paths on genetic progress is potentially higher than that of male paths [17] and it is worth studying this in different species.

In general, PBLUP and ssGBLUP are expected to estimate similar genetic trends before the starting date of genomic selection. However, when the elite animals are genotyped retroactively or genotyping is done after selection, PBLUP and ssGBLUP can estimate different trends. The pig population that was used here has experienced both situations for some of the traits. Thus, we used three approaches to identify and investigate changes in breeding practices over time, in particular the use of genomic selection, namely: (1) based on differences in genetic trends estimated using PBLUP versus ssGBLUP, (2) based on differences in trends in RMS estimated using PBLUP versus ssGBLUP, and (3) based on partitioning the estimated genetic trends into different selection pathways as in [14]. These approaches were applied to a real dataset from a purebred maternal pig line.

Methods
Data structure
The phenotypes of the Landrace pigs used in this study were collected from 2012 to 2021 and included 228,427 records for birth weight (BW), 101,225 records for average daily gain from birth to the end of the nursery period (ADG) at 11 weeks of age, and 11,444 records for average

daily feed intake during the finishing period (FEED) at 23 weeks of age. FEED was measured on males only. Descriptive statistics for each trait are in Table 1. Pedigree information was available for 282,035 animals, of which 32,856 were genotyped for 36,612 single nucleotide polymorphisms (SNPs) after quality control. The animals in the pedigree were the progeny of 809 sires and 14,674 dams. The number of pigs that had either both or one parent unknown was 2011 and 835, respectively. The ratio of genotyped pigs to all pigs born in each year ranged from 4% in 2015 to 18% in 2019. Of the genotyped pigs, 92% had own records and 32% of those had offspring.

Statistical analysis

Analysis of the data was based on a multi-trait mixed linear model of the three traits considered, with the statistical model for each trait described in the following and with fixed effects denoted in uppercase and random effects in lowercase letters.

For BW, the model was:

$$y_{ijokqn} = S_i + P_j + YHM_o + b(tnb_k) + l_q + a_n + m_n + e_{ijkoqn}, \quad (1)$$

where y_{ijokqn} denotes the BW record of animal n, S_i is the effect of the i-th sex ($i = 1$ or 2), P_j is the effect of the j-th parity ($j = 1, ...,9$), YHM_o is the effect of the o-th herd-year-month ($o = 1,..., 173$), b is the linear regression of BW on total number of piglets born (tnb_k), l_q is the random litter effect of sow q ($q = 1,..., 18,394$), a_n is the random direct genetic effect of animal n ($n = 1, ..., 282,035$), m_n is the maternal genetic effect associated with dam of animal n, and e_{ijkoqn} is the residual effect for the BW record.

For ADG, the model was:

$$y_{ijqn} = S_i + YHW_j + l_q + a_n + e_{ijqn}, \quad (2)$$

where y_{ijqn} denotes the ADG record of animal n, S_i, is the effect of the i-th sex ($i = 1$ or 2), YHW_j is the effect of the j-th year-herd-week ($j = 1,..., 1059$), l_q is the random litter effect of sow q ($q = 1,..., 18,394$), a_n is the random direct genetic effect of animal n ($n = 1, ..., 282,035$), and e_{ijqn} is the residual effect for the ADG record.

For FEED, the model was:

$$y_{ikqon} = B_i + b(Age_k) + l_q + p_o + a_n + e_{ikqon}, \quad (3)$$

where y_{ikqon} denotes the FEED record of animal n, B_i is the fixed effect of barn i ($i = 1,..., 512$), b is the linear regression of FEED on age of weighing (Age_k), l_q is the random litter effect of sow q ($q = 1,..., 18,394$), p_o is the random pen effect ($o = 1,..., 4197$), a_n is the random direct genetic effect of animal n ($n = 1, ..., 282,035$), and e_{ijkoqn} is the residual effect for the FEED record.

In matrix notation, the general model for each trait can be written as:

$$\mathbf{y}_t = \mathbf{X}\mathbf{b}_t + \mathbf{W}_1\mathbf{l}_t + \mathbf{W}_2\mathbf{pe}_t + \mathbf{W}_3\mathbf{a}_t + \mathbf{W}_4\mathbf{m}_t + \mathbf{e}_t, \quad (4)$$

where $\mathbf{y}_t$ is the vector of observations for trait t; t refers to BW, ADG, and FEED; $\mathbf{b}_t$ is the vector of fixed effects; $\mathbf{l}_t$, $\mathbf{pe}_t$, $\mathbf{a}_t$, and $\mathbf{m}_t$ are the vectors of random effects for litter, pen, direct additive genetic, and maternal genetic effects, respectively; $\mathbf{e}_t$ is the vector of residuals; and $\mathbf{X}$, $\mathbf{W}_1$, $\mathbf{W}_2$, $\mathbf{W}_3$, and $\mathbf{W}_4$ are design matrices for the effects in $\mathbf{l}_t$, $\mathbf{pe}_t$, $\mathbf{a}_t$, and $\mathbf{m}_t$, respectively.

The assumed (co)variance structure of random effects for the multiple-trait analysis was as follows:

$$\text{Var}\begin{pmatrix} \mathbf{l}_t \\ \mathbf{pe}_t \\ \mathbf{a}_t \\ \mathbf{m}_t \\ \mathbf{e}_t \end{pmatrix} = \text{Var}\begin{pmatrix} \mathbf{l}_{BW} \\ \mathbf{l}_{ADG} \\ \mathbf{l}_{FEED} \\ \mathbf{pe}_{FEED} \\ \mathbf{a}_{BW} \\ \mathbf{a}_{ADG} \\ \mathbf{a}_{FEED} \\ \mathbf{m}_{BW} \\ \mathbf{e} \end{pmatrix} = \begin{pmatrix} \mathbf{I}\sigma^2_{l_{BW}} & \mathbf{I}\sigma_{l_{BW,ADG}} & \mathbf{I}\sigma_{l_{BW,FEED}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ & \mathbf{I}\sigma^2_{l_{ADG}} & \mathbf{I}\sigma_{l_{ADG,FEED}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ & & \mathbf{I}\sigma^2_{l_{FEED}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ & & & \mathbf{I}\sigma^2_{pe_{FEED}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ & & & & \mathbf{A}\sigma^2_{a_{BW}} & \mathbf{A}\sigma_{a_{BW,ADG}} & \mathbf{A}\sigma_{a_{BW,FEED}} & \mathbf{A}\sigma_{a_{BW,m_{BW}}} \\ & & \text{symmetric} & & & \mathbf{A}\sigma^2_{a_{ADG}} & \mathbf{A}\sigma_{a_{ADG,FEED}} & \mathbf{0} \\ & & & & & & \mathbf{A}\sigma^2_{a_{FEED}} & \mathbf{0} \\ & & & & & & & \mathbf{A}\sigma^2_{m_{BW}} \\ & & & & & & & & \mathbf{R} \otimes \mathbf{I} \end{pmatrix}, \quad (5)$$

Table 1 Descriptive statistics of the Landrace pig breed dataset

Trait (unit)	Number of records	Mean	SD	h²	Minimum	Maximum
BW (g)	228,427	1281.6	327.8	0.05	45	3992
ADG (g)	101,255	343.5	61.5	0.22	200	600
FEED (g)	11,444	1917.8	301.9	0.26	827	3872

BW Birth weight, *ADG* average daily gain through the end of the nursery, *FEED* feed intake

where σ_i^2 is the variance of the i-th random effect; σ_{i_j} denotes the covariance components of the i-th effect for the j-th combination of traits, and $\mathbf{R} = 3 \times 3$ matrix with (co)variance between traits; $\mathbf{A}$ is the numerator relationship matrix constructed based on pedigree information for PBLUP, and $\mathbf{I}$ is an identity matrix. For the ssGBLUP analysis, $\mathbf{A}$ was replaced by $\mathbf{H}$, with $\mathbf{H}^{-1}$ computed as in Aguilar et al. [8]:

$$\mathbf{H}^{-1} = \mathbf{A}^{-1} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{G}^{-1} - \mathbf{A}_{22}^{-1} \end{bmatrix},$$

where $\mathbf{G}^{-1}$ is the inverse of the genomic relationship matrix and $\mathbf{A}_{22}^{-1}$ is the inverse of the pedigree relationship matrix for genotyped individuals. The genomic relationship matrix ($\mathbf{G}$) was constructed using the first method of VanRaden [18]:

$$\mathbf{G} = \frac{\mathbf{Z}\mathbf{Z}'}{2\sum p_i(1 - p_i)},$$

where $\mathbf{Z}$ is a matrix of genotypes coded as 0, 1, and 2 for AA, AB, and BB, respectively, and then centered by subtracting twice the frequency of the major allele of SNP i (p_i) ($i = 1, ..., 36{,}612$). To avoid singularity problems, $\mathbf{G}$ was blended with $\mathbf{A}_{22}$ as $\mathbf{G} = 0.95\,\mathbf{G} + 0.05\,\mathbf{A}_{22}$.

Solutions for the multi-trait PBLUP and ssGBLUP were obtained using the preconditioned conjugate gradient algorithm with iteration on data, as implemented in the BLUP90IOD2 program [19]. The (co)variance components were the most recent estimates derived using PBLUP. The GEBV from ssGBLUP were set to the same base (i.e., year 2015) as the mean EBV from PBLUP. To facilitate comparisons between traits, G(EBV) were divided by the square root of the additive genetic variance.

Criteria to dissect genetic trends
Divergence of genetic trends
The point of divergence of the genetic trends obtained by ssGBLUP and PBLUP was used to identify the onset of genomic selection. Details on the theory of predicting breeding values by PBLUP and ssGBLUP, are in Abdollahi-Arpanahi et al. [13]. To estimate the genetic trends using PBLUP and ssGBLUP, the (G)EBV for a given trait were averaged by year of birth for animals with both phenotypes and genotypes. The reason for using only animals with genotypes and phenotypes to estimate genetic trends is that young animals without genotypes and phenotypes (own and progeny) do not contribute information to the evaluation and their average EBV is equal to the parent average.

Realized Mendelian sampling terms
The RMS for individual i for a given trait was estimated as:

$$RMS_i = (G)EBV_i - PA_i, \tag{6}$$

where PA is the parent average (average (G)EBV of the parents) and $(G)EBV_i$ denotes the (genomic) estimated breeding value of individual i. More theoretical details about the RMS can be found in Abdollahi-Arpanahi et al. [13].

When animals are randomly sampled for genotyping at a young age before any source of information (not even genomics) is available, RMS is 0 on average. However, if the "best" animals based on progeny testing or own performance are genotyped (i.e. selective genotyping), then the RMS of the genotyped animals will be nonzero, (i.e. positive if they selected for a higher value and negative if they are selected for a lower value). Since genomic preselection has taken place in most of the livestock populations, the divergence in RMS trends obtained based on EBV and GEBV of genotyped animals can also indicate the starting point of genomic selection. The same animals as used to estimate genetic trends were also used to estimate RMS trends.

Partitioning of genetic trends
Predictions of breeding values can be partitioned to quantify the contribution of genotyped versus non-genotyped or males versus females as follows:

$$\widehat{a}_i = \frac{1}{2}\widehat{a}_{s(i)} + \frac{1}{2}\widehat{a}_{d(i)} + \widehat{m}_i, \tag{7}$$

where $\widehat{m}_i$ is the estimate of the RMS and $\widehat{a}$ is the (G)EBV; subscripts s and d refer to the sire and dam of animal i, respectively. for founder animals $\widehat{a}_i = \widehat{m}_i$. For the whole population Eq. (7) can be written as:

$$\widehat{\mathbf{a}} = \mathbf{T}\widehat{\mathbf{m}}, \tag{8}$$

where $\mathbf{T}$ is a triangular matrix that relates each animal to its parents [20]. Following Eq. (8) and considering that $\widehat{\mathbf{m}} = \mathbf{T}^{-1}\widehat{\mathbf{a}}$, the vector of (G)EBV for the entire population (Eq. (8)) can be partitioned into contributions of defined selection pathways [14] as:

$$\widehat{\mathbf{a}} = \mathbf{T}\mathbf{P}_1\mathbf{T}^{-1}\widehat{\mathbf{a}} + \mathbf{T}\mathbf{P}_2\mathbf{T}^{-1}\widehat{\mathbf{a}} = \widehat{\mathbf{a}}_1 + \widehat{\mathbf{a}}_2, \tag{9}$$

where $\mathbf{P}_i$ is a diagonal matrix of 1s and 0s to select the corresponding columns of $\mathbf{T}$ and is used to allocate the RMS of males versus females or of genotyped versus non-genotyped individuals to the i-th partition of $\widehat{\mathbf{a}}$.

This procedure was implemented using the R package AlphaPart 0.8.1. [21], using the GEBV obtained using ssGBLUP for all animals (i.e., genotyped and non-genotyped) as the input. The contribution of each pathway for each birth year was expressed as a percentage by dividing the average GEBV of a partition by the average GEBV of the whole population.

Results
Genetic trends
Estimated genetic trends for genotyped individuals, in genetic standard deviation units, based on PBLUP and ssGBLUP are presented in Fig. 1. The genetic trends were favorable for all traits, with a faster improvement in recent years. The changes in average EBV from 2015 to 2020 for BW, ADG, and FEED were 0.66, 0.72, and 0.20, respectively based on PBLUP and 0.65, 1.03, and
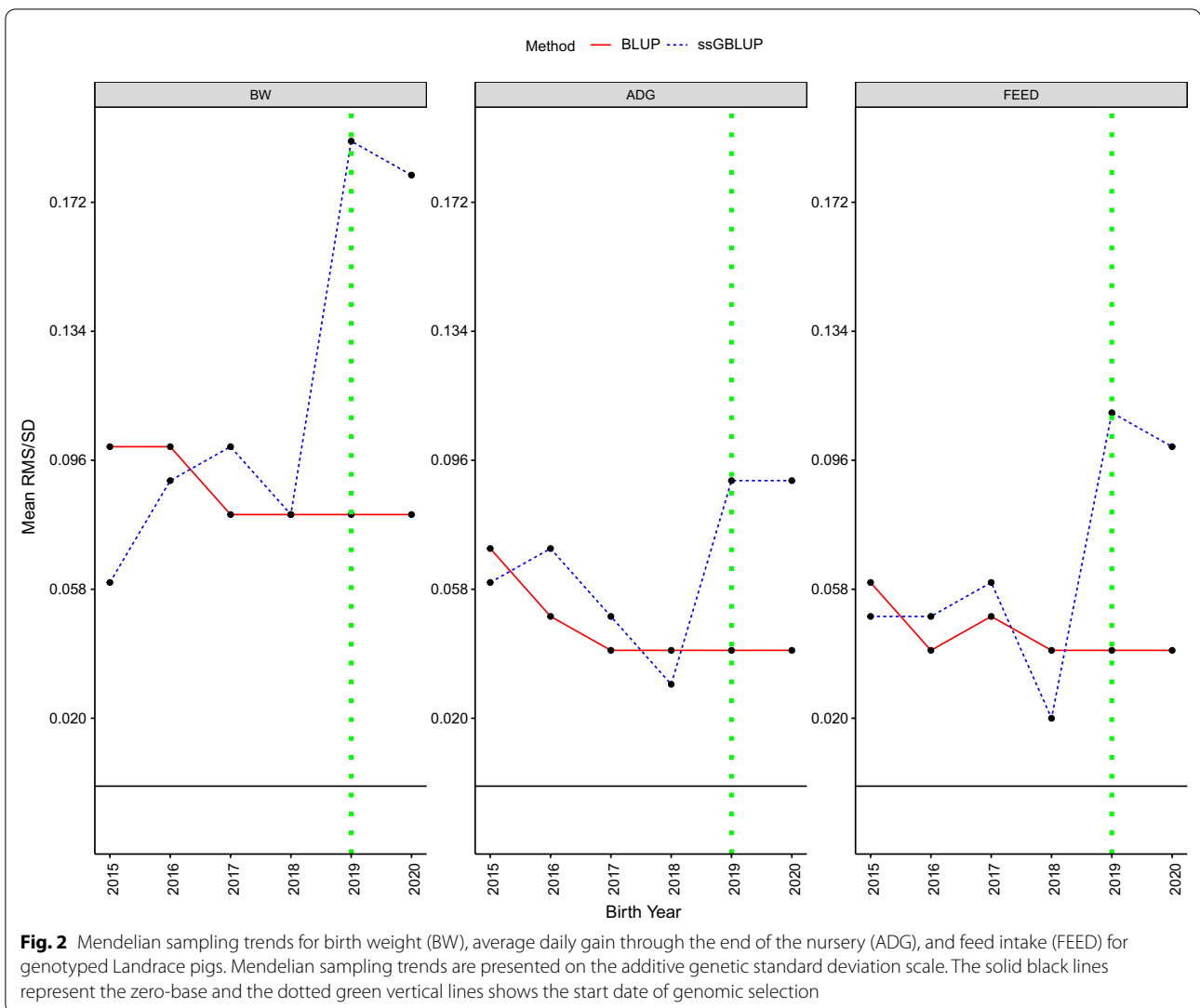
0.31 based on ssGBLUP. For ADG and FEED, the genetic trends estimated using PBLUP and ssGBLUP started to diverge in 2018, but for BW, which is not under direct selection, there was no evidence of divergence. In the last year of data (i.e., 2020), the difference between average EBV based on ssGBLUP and PBLUP were − 0.01, 0.31, and 0.11 SD for BW, ADG, and FEED, respectively. The positive genetic trend for BW is due to its indirect response to selection for increasing the maternal effect on BW and for other correlated traits in the selection index.

Mendelian sampling trends
The RMS trends for genotyped individuals, in genetic standard deviation units, estimated using PBLUP and ssGBLUP are shown in Fig. 2. The pattern of changes in RMS trends was the same for the three traits and started
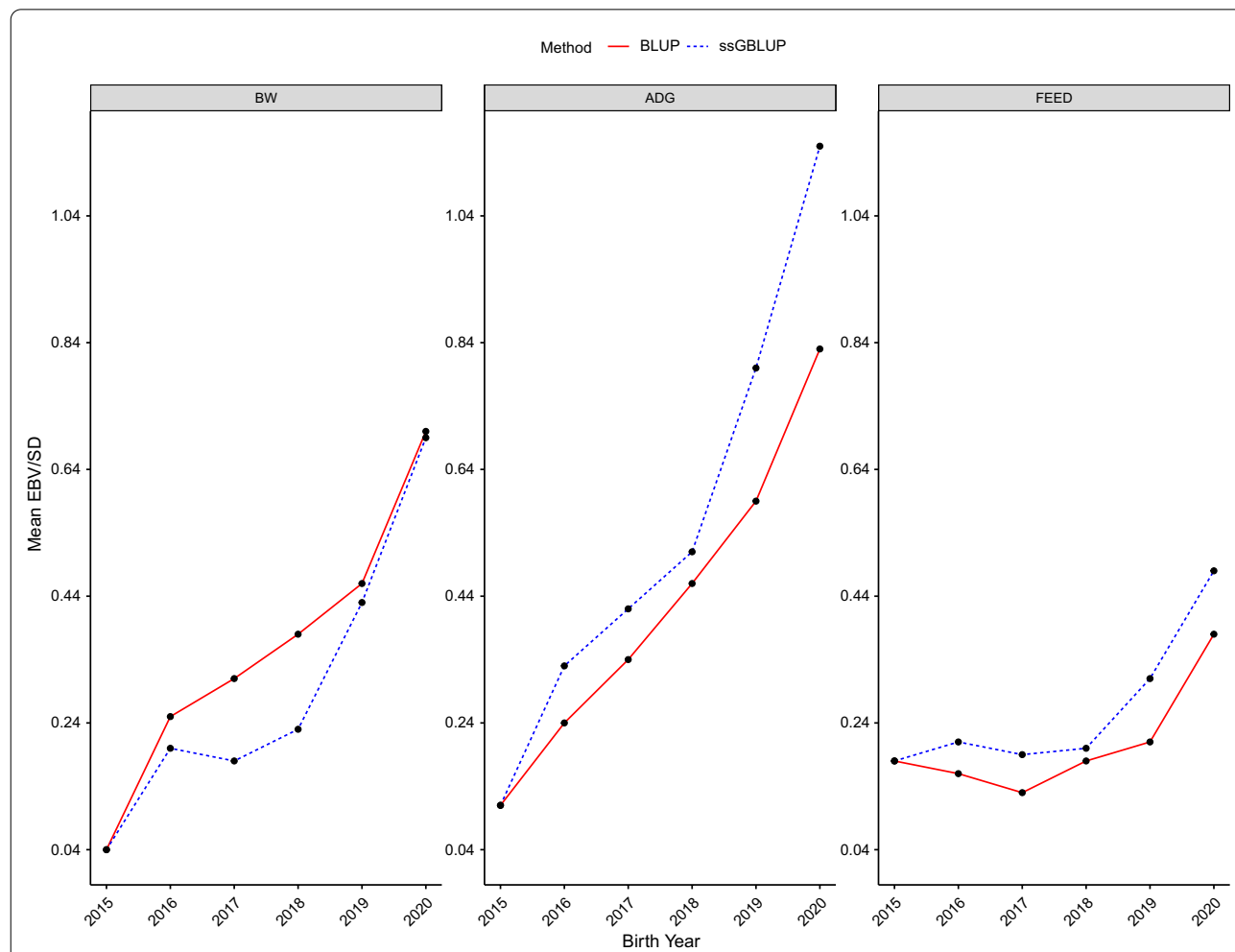
Fig. 1 Genetic trends for birth weight (BW), average daily gain through the end of the nursery (ADG), and feed intake (FEED) for genotyped Landrace pigs. Genetic trends are presented on the additive genetic standard deviation scale and the genetic base was adjusted to the 2015 birth year

Abdollahi-Arpanahi *et al. Genetics Selection Evolution*   (2021) 53:89

Page 6 of 10

Fig. 2 Mendelian sampling trends for birth weight (BW), average daily gain through the end of the nursery (ADG), and feed intake (FEED) for genotyped Landrace pigs. Mendelian sampling trends are presented on the additive genetic standard deviation scale. The solid black lines represent the zero-base and the dotted green vertical lines shows the start date of genomic selection

to diverge in 2019. In the last year of data (i.e., 2020), the difference between average estimates of RMS based on PBLUP and ssGBLUP was 0.10, 0.05, and 0.06 SD for BW, ADG, and FEED, respectively. The positive estimated trend in RMS and the considerable difference of RMS from 0 from 2015 to 2017 are due to the genotyping in 2018 of elite culled or active boars that were born before 2018 and were retrieved from stored tissue samples.
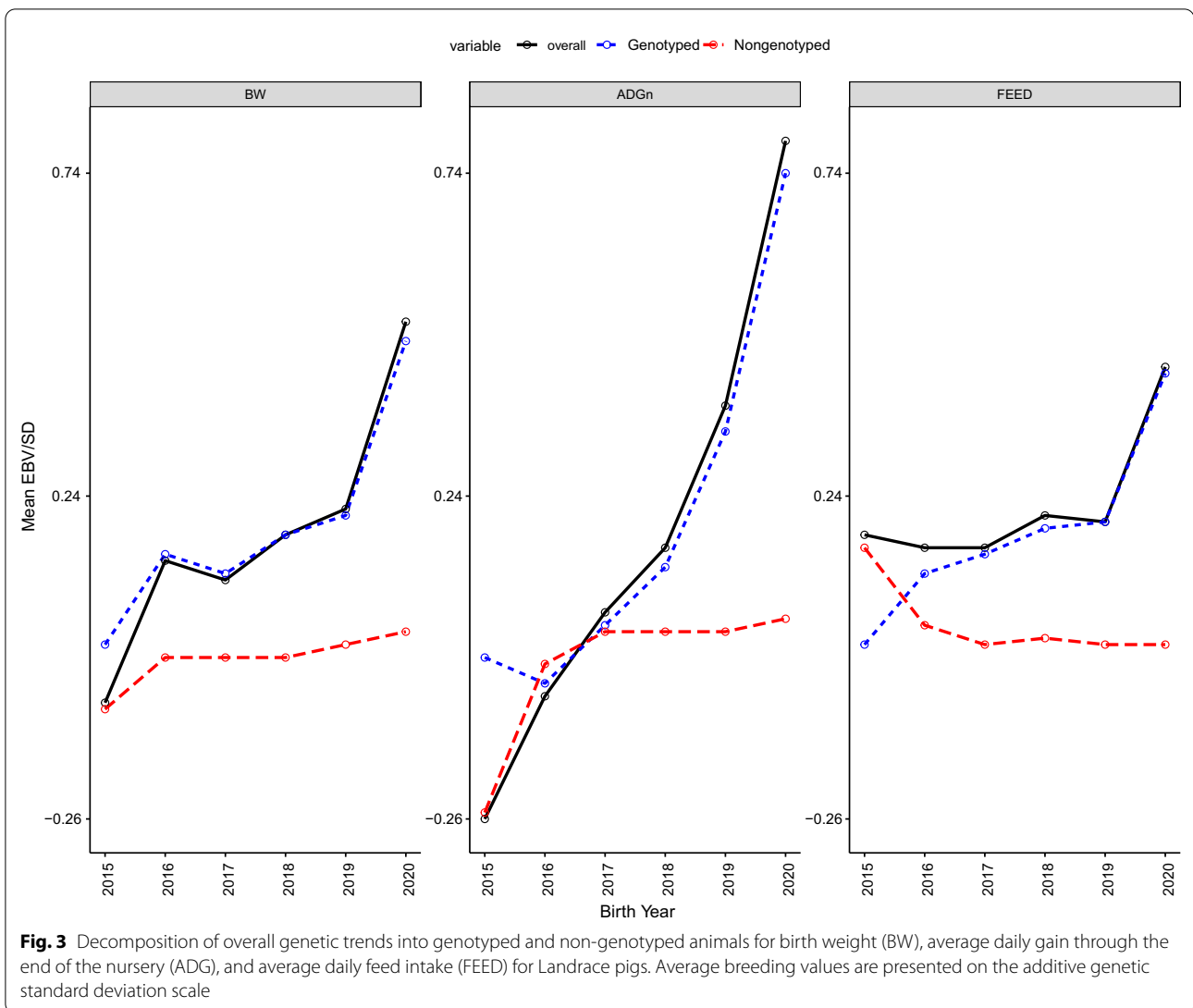
Contributions of genotyped versus non-genotyped individuals to genetic trends

The decomposition of genetic trends into Mendelian sampling contributions from genotyped and non-genotyped individuals is shown in Fig. 3. The percentage of individuals born from 2015 to 2020 that were genotyped ranged from 5 to 21%. All genetic gain in BW was due

to non-genotyped individuals before 2016, but from 2016 onwards, genotyped individuals were responsible for the genetic gain. For ADG, non-genotyped individuals had a greater contribution to the genetic trend than genotyped animals until 2016, but from 2017 onwards, the contribution of genotyped individuals to genetic gain increased from 74% in 2016 to 94% in 2020. For feed intake, in 2015, all the genetic gain was driven by non-genotyped pigs, but the contribution of genotyped pigs increased rapidly after that, from 76% in 2016 to 97% in 2020.

Contributions of females versus males to genetic trends

Figure 4 shows the decomposition of genetic trends into contributions from males and females for all traits. For BW, most of the genetic gain was driven by females from 2019 onwards. For ADG, males and females had similar

Fig. 3 Decomposition of overall genetic trends into genotyped and non-genotyped animals for birth weight (BW), average daily gain through the end of the nursery (ADG), and average daily feed intake (FEED) for Landrace pigs. Average breeding values are presented on the additive genetic standard deviation scale

contributions to the genetic trend up to 2017, but from 2018 onwards, the contribution of females was greater than that of males. The pattern of genetic trend for feed intake was similar to that of males, while the contribution of females to the genetic trend for feed intake was in the undesirable direction. For BW and ADG, the impact of the female pathway increased after implementation of genomic selection.

Discussion
Genetic trends
We used the divergence of genetic and RMS trends obtained by PBLUP versus ssGBLUP to determine the effectiveness of genomic selection. We also partitioned genetic trends into contributions from genotyped versus non-genotyped animals and from males

versus females. Quantifying the contribution of genotyped individuals to the genetic trends revealed the presence of selective genotyping (only elite animals were genotyped) and the impact of genomic selection (faster genetic improvement) in this breeding program. Decomposition of genetic trends into contributions from males versus females allowed the monitoring of the contribution of males versus females to genetic trends after the implementation of genomic selection.
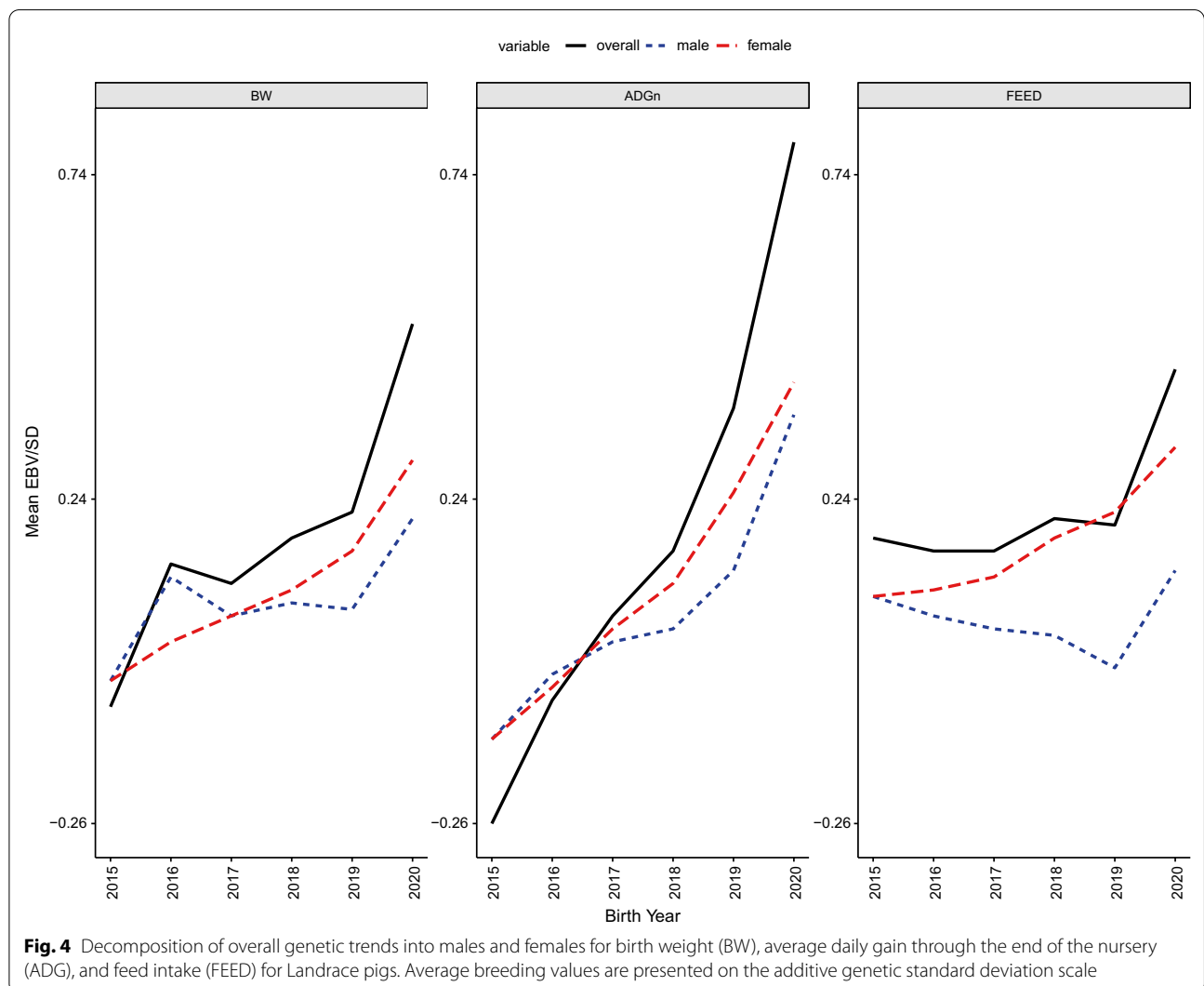
The effective starting point of genomic selection was found to be in 2019 in the offspring of the first animals selected based on GEBV in 2018. This finding agrees with the history of genomic selection in this pig population. Genotyping started in late 2016, for pigs born between 2015 and 2017 that were active boars or culled boars with stored tissue samples. Consequently, the higher genetic

Abdollahi-Arpanahi *et al. Genetics Selection Evolution* (2021) 53:89

Page 8 of 10

trends estimated from 2015 to 2018 for ADG using ssGBLUP than using PBLUP can be the result of selective genotyping. The decline in genetic trend for ADG estimated by PBLUP after 2018 is due to genomic preselection bias. When only genotyped animals receive phenotype records, PBLUP do not account for the positive RMS of those young animals. Similar results are reported in the literature [12, 13, 22]. Masuda et al. [12] compared genetic trends estimated using PBLUP and ssGBLUP for milk production traits in US Holstein cattle and found that after the implementation of genomic selection, the genetic trend based on PBLUP was underestimated because of genomic preselection. In a simulation study, Jibrila et al. [22] showed that genomic preselection caused bias in estimates of genetic gain based on PBLUB, while the bias was smaller when based on ssGBLUP. According to Abdollahi-Arpanahi et al. [13], after implementing genomic selection in pig, broiler, and beef

cattle populations, the genetic trends obtained by ssGBLUP accelerated and those estimated using PBLUP decelerated.

In the pig population under study, genotyping was done retroactively, which means that for the genotyped animals born from 2015 to 2017 the selection decisions were practiced by another method such as PBLUP, thus even if ssGBLUP during this period results in higher accuracy than PBLUP, we do not expect a higher genetic trend for ssGBLUP. In fact, the accuracy of ssGBLUP will be greater than the accuracy of PBLUP at any point when a sufficient number of animals is genotyped. However, the prediction accuracy at the time of selection is what is reflected in the genetic trend. The reason is that if the company/breeder invested in genotyping but has not used the genomic information in selection decisions, the higher accuracy of evaluation by ssGBLUP compared to PBLUP using accumulated data does not necessarily

Fig. 4 Decomposition of overall genetic trends into males and females for birth weight (BW), average daily gain through the end of the nursery (ADG), and feed intake (FEED) for Landrace pigs. Average breeding values are presented on the additive genetic standard deviation scale

Abdollahi-Arpanahi *et al. Genetics Selection Evolution*   (2021) 53:89

Page 9 of 10

translate into the genetic trend. Overall, the changes of prediction accuracies over time may not follow the genetic trends estimated by PBLUP or ssGBLUP.

Although investigating the fluctuations in the genetic trend across time for each trait is beyond the scope of this study, the changes in genetic trends observed are consistent with the breeding practices and the periodic modifications of weights in the selection index and the genetic correlations between traits under selection. One example is the estimated increase of 0.6 and 0.7 SD in BW from 2015 to 2020 based on ssGBLUP and PBLUP, respectively. Although the direct genetic value of BW has not been selected for in this population, the maternal genetic value for BW has. The genetic correlation between direct and maternal genetic values is about 0.1. However, the genetic correlations of direct BW with growth rate in the nursery, finisher average daily gain, and finisher average daily feed intake are positive and high, e.g., 0.42, 0.31 and 0.25, respectively. Therefore, we believe that the observed genetic trend for BW is due to the correlated responses to selection.

Mendelian sampling trends

The RMS trends revealed signatures of selective genotyping for the three traits. All males born from 2017 to 2020 were genotyped, but only a subset of females selected based on phenotypes were genotyped during this period. Hence, the deviations of RMS from 0 for birth years after 2018 are due to the strong selective genotyping of females rather than of males. Moreover, inferior males, e.g., those with a small BW, may be removed from the tested population before genotyping, which can result in positive RMS. As genotyping becomes less expensive, genotyping more young animals becomes economically justified and we expect a convergence of the RMS trends estimated using PBLUP versus ssGBLUP if phenotypic records are available for all animals.

The advantages of ssGBLUP in reducing prediction bias increase when animals have been preselected based on GEBV. Genetic evaluation using PBLUP assumes that RMS average 0, but when genotyped animals with positive or negative RMS receive phenotypes or progeny, the average RMS is no longer 0 [7, 9]. In this regard, a simulation study showed that the RMS for bulls clearly deviated from 0 after genomic preselection was implemented in a dairy cattle population [10].

Decomposition of genetic trends

To quantify the contribution of genotyped individuals to genetic trends, we partitioned the genetic trends into the genetic gain derived by genotyped individuals and that achieved by non-genotyped individuals. Regardless of the trait, in recent years, genotyped individuals had a greater contribution to genetic gain than non-genotyped individuals. The greater contribution of genotyped individuals to genetic trends does not necessarily depict the effectiveness of genomic selection. For instance, if genotyped animals are preselected based on PBLUP EBV, we expect genetic trends to be higher for genotyped than for non-genotyped animals, which is the case for the period from 2015 to 2018 before genomic selection started.

A greater impact of females on genetic trends would be because selection decisions in a maternal line are placed more on females than males, and in pigs, each selected female has a larger contribution because it produces more progeny. However, the pig breed analyzed here is a maternal line and 40% of the traits under selection are only measured in females, which results in the females being the main drivers of changes in these traits. Thus, it is expected that females contribute more to the next generation than males in a pig breeding program, although the selection intensity for males is higher than for females. We found that the contribution of females to the genetic trends for BW and ADG was greater than that of males after the implementation of genomic selection.

For FEED, while a flat to slightly positive genetic trend was observed for males, the trend for females was positive and unfavorable. Feed intake has a negative economic value, while ADG has a positive value in the index. Therefore, the breeding objective is to achieve a positive response in growth rate and a flat or slightly positive response in FEED, which, in turn, improves feed efficiency. Few studies investigated the contribution of different selection paths to genetic trends. For example, García-Ruiz et al. [2] demonstrated that 73 to 90% of the selection differential for milk production traits in US Holstein cattle is due to the sire of the bull and sire of the cow pathways.

Conclusions

Divergence of genetic trends for genotyped animals estimated using PBLUP versus ssGBLUP indicates the presence of genomic selection. This divergence may occur before the onset of genomic selection if superior animals are genotyped retroactively. Presence of nonzero average RMS by ssGBLUP or PBLUP indicates selective genotyping. Selective genotyping can be deliberate, e.g., genotyping of animals with superior genotypes, or incidental due to removal of weak/sick/dead animals. Under genomic selection, trends for RMS are higher when estimated using ssGBLUP than using PBLUP, with the point of divergence indicating the effective onset of genomic selection. Partitioning of genetic trends into contributions by various classes of animals such as genotyped versus ungenotyped or males versus females allows the

Abdollahi-Arpanahi *et al. Genetics Selection Evolution*   (2021) 53:89

Page 10 of 10

determination of the relative impact of genotyping for different groups of animals. In particular, the observation that nearly all the genetic progress is contributed by genotyped animals confirms the increasing interest in genotyping animals, and the observation that a large fraction of the genetic progress is contributed by females validates the importance of the females in the genetic progress of a pig population. In summary, post-processing of EBV and GEBV can help to investigate the effectiveness of genomic selection and assess breeding program practices.

Availability of data and materials
Provider of data does not intend to disclose its identity.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

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

Author details
[1]Department of Animal and Dairy Science, University of Georgia, Athens, GA 30602, USA. [2]UMR GenPhySE, INRA Toulouse, Castanet Tolosan, France.

References
1. Georges M, Charlier C, Hayes B. Harnessing genomic information for livestock improvement. Nat Rev Genet. 2019;20:135–56.
2. García-Ruiz A, Cole JB, VanRaden PM, Wiggans GR, Ruiz-López FJ, Van Tassell CP. Changes in genetic selection differentials and generation intervals in US Holstein dairy cattle as a result of genomic selection. Proc Natl Acad Sci USA. 2016;113:E3995-4004.
3. Misztal I, Lourenco D, Legarra A. Current status of genomic evaluation. J Anim Sci. 2020. https://doi.org/10.1093/jas/skaa101.
4. Rendel JM, Robertson A. Estimation of genetic gain in milk yield by selection in a closed herd of dairy cattle. J Genet. 1950;50:1–8.
5. Woolliams JA, Bijma P, Villanueva B. Expected genetic contributions and their impact on gene flow and genetic gain. Genetics. 1999;153:1009–20.
6. Party C, Ducrocq V. Bias due to genomic selection. Interbull Bull. 2009;39:77–82.
7. Patry C, Ducrocq V. Evidence of biases in genetic evaluations due to genomic preselection in dairy cattle. J Dairy Sci. 2011;94:1011–20.
8. Aguilar I, Misztal I, Johnson DL, Legarra A, Tsuruta S, Lawlor TJ. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. J Dairy Sci. 2010;93:743–52.
9. Patry C, Ducrocq V. Accounting for genomic pre-selection in national BLUP evaluations in dairy cattle. Genet Sel Evol. 2011;43:30.
10. Tyrisevä AM, Mäntysaari EA, Jakobsen J, Aamand GP, Dürr J, Fikse WF, et al. Detection of evaluation bias caused by genomic preselection. J Dairy Sci. 2018;101:3155–63.
11. Koivula M, Stranden I, Aamand GP, Mantysaari EA. Reducing bias in the dairy cattle single-step genomic evaluation by ignoring bulls without progeny. J Anim Breed Genet. 2018;135:107–15.
12. Masuda Y, VanRaden PM, Misztal I, Lawlor TJ. Differing genetic trend estimates from traditional and genomic evaluations of genotyped animals as evidence of preselection bias in US Holsteins. J Dairy Sci. 2018;101:5194–206.
13. Abdollahi-Arpanahi R, Lourenco D, Misztal I. Detecting effective date of genomic selection by divergent trends from BLUP and ssGBLUP in pigs, beef cattle, and broilers. BioRXiv. 2021;101:5166.
14. García-Cortés LA, Martínez-Ávila JC, Toro MA. Partition of the genetic trend to validate multiple selection decisions. Animal. 2008;2:821–4.
15. Gorjanc G, Potočnik K, García-Cortés LA, Jakobsen J, Dürr J. Partitioning of international genetic trends by origin in Brown Swiss bulls. Interbull Bull. 2011;44:81–6.
16. Škorput D, Gorjanc G, Kasap A, Luković Z. Partition of genetic trends by origin in Landrace and Large-White pigs. Animal. 2015;9:1605–9.
17. Lourenco DAL, Fragomeni BO, Tsuruta S, Aguilar I, Zumbach B, Hawken RJ, et al. Accuracy of estimated breeding values with genomic information on males, females, or both: an example on broiler chicken. Genet Sel Evol. 2015;47:56.
18. VanRaden PM. Efficient methods to compute genomic predictions. J Dairy Sci. 2008;91:4414–23.
19. Misztal I, Tsuruta S, Lourenco DAL, Masuda Y, Aguilar I, Legarra A, et al. Manual for BLUPF90 family of programs. Athens: University of Georgia; 2014.
20. Quaas RL. Computing the diagonal elements and inverse of a large numerator relationship matrix. Biometrics. 1976;32:949–53.
21. Obšteter J, Holl J, Hickey JM, Gorjanc G. AlphaPart-R implementation of the method for partitioning genetic trends. Genet Sel Evol. 2021;53:30.
22. Jibrila I, ten Napel J, Vandenplas J, Veerkamp RF, Calus MPL. Investigating the impact of preselection on subsequent single-step genomic BLUP evaluation of preselected animals. Genet Sel Evol. 2020;52:42.

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