

Article

Weighted Single-Step Genomic Best Linear Unbiased Prediction Method Application for Assessing Pigs on Meat Productivity and Reproduction Traits

Artem Kabanov ^{1,*} , Ekaterina Melnikova ¹ , Sergey Nikitin ¹ , Maria Somova ¹, Oleg Fomenko ², Valeria Volkova ¹ , Olga Kostyunina ¹ , Tatiana Karpushkina ¹, Elena Martynova ³ and Elena Trebunskikh ⁴

¹ L.K. Ernst Federal Research Center for Animal Husbandry, Dubrovitzky Estate, Podolsk District, 142132 Moscow, Russia; melnikovae_1982@inbox.ru (E.M.); snikitin@vij.ru (S.N.); somova-mm@yandex.ru (M.S.); moonlit_elf@mail.ru (V.V.); kostolan@yandex.ru (O.K.); tati.kriz@gmail.com (T.K.)

² All-Russian Dairy Research Institute, 115093 Moscow, Russia; fomenych@rambler.ru

³ Center of Life Sciences, Skolkovo Institute of Science and Technology, 143026 Moscow, Russia; elenamartynovaster@gmail.com

⁴ LLC "TOPGEN", Verkhnyaya Khava, 396110 Voronezh, Russia; terramio7@mail.ru

* Correspondence: kabanov@mail.ru



Citation: Kabanov, A.; Melnikova, E.; Nikitin, S.; Somova, M.; Fomenko, O.; Volkova, V.; Kostyunina, O.; Karpushkina, T.; Martynova, E.; Trebunskikh, E. Weighted Single-Step Genomic Best Linear Unbiased Prediction Method Application for Assessing Pigs on Meat Productivity and Reproduction Traits. *Animals* **2022**, *12*, 1693. <https://doi.org/10.3390/ani12131693>

Academic Editors: Cristina Sartori and Shogo Tsuruta

Received: 19 March 2022

Accepted: 28 June 2022

Published: 30 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Simple Summary: The effectiveness of genomic selection in pig breeding depends largely on genomic prediction accuracy. The relatively short generation interval and the high selection intensity make the issue of the selection criterion accuracy extremely important. Genotyping animals by SNP markers makes it possible to increase the accuracy of breeding criteria for young replacement animals. However, the applied computational algorithm plays an equally important role, that is, the method of calculating the genomic estimates. The study compared three methods (BLUP AM, ssGBLUP, wssGBLUP) for the assessment of pigs' breeding value according to five main breeding traits.

Abstract: Changes in the accuracy of the genomic estimates obtained by the ssGBLUP and wssGBLUP methods were evaluated using different reference groups. The weighting procedure's reasonableness of application was considered to improve the accuracy of genomic predictions for meat, fattening and reproduction traits in pigs. Six reference groups were formed to assess the genomic data quantity impact on the accuracy of predicted values (groups of genotyped animals). The datasets included 62,927 records of meat and fattening productivity (fat thickness over 6–7 ribs (BF1, mm)), muscle depth (MD, mm) and precocity up to 100 kg (age, days) and 16,070 observations of reproductive qualities (the number of all born piglets (TNB) and the number of live-born piglets (NBA), according to the results of the first farrowing). The wssGBLUP method has an advantage over ssGBLUP in terms of estimation reliability. When using a small reference group, the difference in the accuracy of ssGBLUP over BLUP AM is from -1.9 to $+7.3$ percent points, while for wssGBLUP, the change in accuracy varies from $+18.2$ to $+87.3$ percent points. Furthermore, the superiority of the wssGBLUP is also maintained for the largest group of genotyped animals: from $+4.7$ to $+15.9$ percent points for ssGBLUP and from $+21.1$ to $+90.5$ percent points for wssGBLUP. However, for all analyzed traits, the number of markers explaining 5% of genetic variability varied from 71 to 108, and the number of such SNPs varied depending on the size of the reference group (79–88 for BF1, 72–81 for MD, 71–108 for age). The results of the genetic variation distribution have the greatest similarity between groups of about 1000 and about 1500 individuals. Thus, the size of the reference group of more than 1000 individuals gives more stable results for the estimation based on the wssGBLUP method, while using the reference group of 500 individuals can lead to distorted results of GEBV.

Keywords: genomic evaluation; single-step GBLUP; weighted single-step GBLUP; pigs; SNP effects; estimation reliability; validation of genomic estimates

1. Introduction

The improvement of methods for assessing the farm animals breeding values has two main directions: attracting as much data about the animals and their relatives as possible and developing methods that allow the more efficient use of available information. Thus, genomic evaluation methods have significantly improved the prediction accuracy of the individual genetic value by using marker genotype data [1–3]. The next step in this process was the development of mathematical procedures ensuring the most reasonable application of this information and special algorithms to reduce computational and memory costs, in particular, the determination of the optimal weight of the data involved [4,5]. The density of SNP marker distributions over the genome greatly impacts genomic prediction accuracy [6]. Applying the WGS strategy seems the most preferable for genomic selection because microarray chips cover only a limited number of individual SNPs, while whole-genome sequencing provides information over the entire genome [7,8]. However, the genomic evaluation based on microarray genotyping data is still most widespread due to having the best cost-benefit ratio.

The primary assumption of genomic assessment methods based on linear mixed models is that all markers have the same contribution to the overall genetic variability of a given trait. However, it is difficult to agree that such a situation occurs [9,10]. The ability to account for individual SNP marker weight in genomic estimation has led to the rise of the weighted single-step genomic BLUP (wssGBLUP) methodology, which is able to differentiate the contribution of every single marker effect to the trait variability [11]. A large number of studies are devoted to searching quantitative trait loci and casual genes, the identification of which could significantly reduce the marker panels and improve the individual genetic prediction accuracy. In this regard, many investigations have analyzed the associations between individual SNP markers and/or their positional successions (also known as «windows») and the phenotype variability of key breeding traits in livestock [12]. The associated study results, such as identifying individual genes and large chromosomal segments obtained by different research groups, are often not entirely consistent with each other. This is due to several objective and subjective reasons: trait polygenicity, low marker effect values, the insufficient information content of traditional phenotypic measurements, the significant influence of paratype factors on the phenotype, which cannot always be taken into account, and the utilization of different models and approaches [13–22]. At the same time, other studies have identified the same candidate genes associated with trait variability in different animal species [23]. The idea of preliminary identification and selection of significant SNP markers for inclusion in calculating genomic estimates served as the basis for developing the weighted ssGBLUP procedure [11,24]. In addition, B. Fragomeni et al. (2017) suggested that the proportion of significant markers in the total SNP number can serve as an indirect indicator of the trait polygenicity (“complexity”). They also mention the relatively low efficiency of significant markers identified by any known methods concerning complex quantitative traits [24]. Therefore, the idea of redistributing the “weight” (contribution) of each marker to the individual genetic value assessment while maintaining a standard set of SNP markers on microarray chips (preferably of high density) looks the most promising [25]. Many studies based on real and simulated data compared the effectiveness of using reduced sets of SNP markers through methods such as pre-selection based on GWAS, data from dense and medium-density chips, and whole-genome information. However, improvement in the accuracy of the estimates for the limited marker sets was not always observed (except when the causal genes (QTL) are known), and an increase in the distortion of the results due to genotyping errors is also mentioned by some authors [6,8,26–29]. The fact is that SNPs explain a part of the complex trait heritability, and each of them individually has an infinitely small effect size. Moreover, the main issue is not the number of identified loci influencing the appearance of the trait but rather how the identified loci cumulatively explain the appearance of the trait [30] and whether it can be effectively used for individual prediction, for example, when conducting a genomic evaluation. The assessment of SNPs affects a significant influence on the particular

trait phenotype manifestation variability, and additionally, the values of the SNP markers affect themselves. This depends dramatically on the reference group parameters used for calculation, in particular, the allele frequencies distribution, the linkage disequilibrium extent, the genetic structure, the effective size of the reference population, as well as the polygenicity of the trait itself (the presence of «major» QTLs with significant effects) [7]. Due to this, the competent and reasonable formation of a sufficient size reference group genotyped on high-density chips plays an essential role in forming the individual genomic assessment [6].

Therefore, our study aimed to assess the impact of the SNP genotypes' data amount on individual genomic estimate accuracy based on the ssGBLUP and wssGBLUP methods. The validity of applying the weighting procedure to improve the accuracy of genomic predictions for pigs' meat, fattening, and reproduction traits were considered.

2. Materials and Methods

2.1. Phenotypes

The phenotypic features of meat, fattening and reproductive traits of Large White pigs were used. Animals born between 2015 and 2020 were included in the dataset. The following characteristics of meat and fattening productivity were evaluated: fat thickness over 6–7 ribs (BF1, mm), muscle depth (MD, mm) and precocity up to 100 kg (age, days). In addition, according to the results of the first farrowing, the number of all-born piglets (TNB) and the number of live-born piglets (NBA) were taken into account. Descriptive statistics of the complete data set are given in Table 1.

Table 1. The descriptive statistics of the whole dataset.

Trait	Mean	SD	Range	h^2	N
BF1	16.13	3.63	7.0 to 34.0	0.428	62,927
MD	59.52	7.18	40.0 to 99.0	0.195	
Age	154.4	9.17	109 to 205	0.323	
TNB	15.4	4.14	1 to 29	0.119	16,070
NBA	14.2	3.95	0 to 28	0.112	

To assess the genomic data amount effect on the accuracy of predictive values for several traits, we formed six reference groups (groups of genotyped animals) and calculated them according to the scenarios presented in Table 2.

Table 2. Calculation scenarios.

Scenario	Methods	Phenotypes and Pedigree Data	Genomic Data	Results
Scenario 1	ssGBLUP wssGBLUP	WD	RG1 (RG4)	$GEBV_{ss1/4}$ $GEBV_{wss1/4}$
Scenario 2	ssGBLUP wssGBLUP	WD	RG2 (RG5)	$GEBV_{ss2/5}$ $GEBV_{wss2/5}$
Scenario 3	ssGBLUP wssGBLUP	WD	RG3 (RG6)	$GEBV_{ss3/6}$ $GEBV_{wss3/6}$
Scenario 4	BLUP AM	WD	-	EBV_{WD}
Scenario 5 (validation)	BLUP AM ssGBLUP wssGBLUP	PD	- RG3 (RG6) RG3 (RG6)	PA_{PD} $GEBV_{ssPD}$ $GEBV_{wssPD}$

The groups were selected in such a way as to simulate a situation where genotyping was carried out on the part of the livestock in the estimated population at three stages, with a cumulative effect. An analysis of meat and fattening qualities was carried out on a dataset that included 62,927 records of in vivo measurements of backfat thickness,

muscle depth and precocity in gilts and boars when they reached 100 kg live weight. Studies of reproductive traits were carried out on a dataset, including results of the first farrowing (16,070 observations). Accordingly, the reference groups of genotyped individuals according to the meat and reproduction traits were different but formed according to a single algorithm. Thus, the first group for meat and fattening qualities included 530 pigs (boars, sows, piglets) from the 2010–2015 year of birth, and the second group included 1178 individuals, including the first group (boars, sows, piglets) born in 2010–2017. The third group consisted of animals from the second group, to which 315 more heads were added, for a total of 1493 individuals born from 2010–2021. According to the reproduction traits, only genomic data of sows were used in the analysis. Therefore, the following values characterized the groups: the fourth included 396 individuals (sows born from 2010–2015), the fifth included 870 animals (sows born from 2010–2017), and the sixth included 1228 sows born from 2010–2020 (Table 3). Based on the groups' genomic data, the values of the genomic scores of all individuals in the total analyzed data sets were determined.

Table 3. Reference groups description.

Parameter	Meat and Fattening Traits		
	Group 1	Group 2	Group 3
N	530	1178	1493
Ne	85	90	96
Reproduction traits			
	Group 4	Group 5	Group 6
N	396	870	1228
Ne	96	107	107

N: number of animals in groups; Ne: effective population size.

The most complete (numerous) reference groups of individuals 3 and 6 were characterized by the following average indicators and standard deviations of phenotypic values (mean (SD)): group 3, BF1—14.61 (3.13), MD—57.95 (5.96), age—156.1 (9.78); group 6, TNB—13.48 (3.87), NBA—12.76 (3.75). The phenotype's limiting values (ranges) of the reference groups did not differ from those of the general population. A set of partial data was formed for validation by deleting records of validated animal phenotypes. The relationships between the validated animals and their descendants were reset in the pedigree. The group of validated animals for meat and fattening traits included 200 individuals: 79 boars (mean rel = 0.954) and 121 sows (mean rel = 0.849). The validated animals for reproduction traits group included 165 sows (mean rel = 0.572).

2.2. Genotypes

We utilized the GGP Porcine HD hybridization chips (Illumina/Neogen, Lincoln, NE, USA) to genotype boars and sows used in this study to form reference groups. The GGP Porcine HD hybridization chip contains 70,000 SNPs with a mean intermarker interval of approximately 42 kb and bears 20 major causative mutations. These SNPs are identified within and between all major pig breeds. Genome-wide genotyping data were presented as 1483 individual genotypes exported from the GenomeStudio 2 program (Illumina Inc., San Diego, CA, USA).

We performed data quality control (QC) by using Plink 1.9 [31] to ensure the accurate identification of SNPs and for the correct interpretation of genetic association results. Individuals and SNPs that did not pass the following thresholds were removed: SNPs with accuracy (GC Score, GT Score) less than 0.2, animal records with more than 10% genotype absence, SNP markers absent in more than 10% of genotyped animals, SNP markers with a minor allele frequency of less than 5%; SNP markers deviating from the Hardy–Weinberg equilibrium with a threshold value of 10^{-6} ; and animal records with Mendelian errors of allele inheritance (the presence of such errors was regarded as errors

in recording animal pedigrees). Afterward, the QC genomic data included 50,740 SNPs of 1483 genotyped animals.

2.3. Statistical Analysis and Assessment

Descriptive statistics calculations were performed in the program STATISTICA 10, and animal pedigrees were analyzed by the CFC software package [32]. Primary processing and transformation of genomic data were carried out using the local programs using integrated development environment (IDE) RStudio for the R language. SNeP v1.1 was used to compute effective population sizes [33], and the BLUPF90 software package

Reference [34] was used to obtain individual breeding value estimates (EBV, GEBV_{ss}, GEBV_{wss}). The genetic population parameters (heritability, genetic variance and covariance) were estimated based on the restricted maximum likelihood method in the REMLF90 program [34,35].

The mixed linear model on the BLUP AM method is:

$$y = Xb + Za + e,$$

where y represents the vector of phenotype observations on the trait (TNB, NBA, BF1, MD and Age); b represents vector of fixed effects («farm-year—contemporary group»; «sex of the animal»; «weight of the animal»); a represents vector of random animal effect; e represents the vector of the random residual effect; and X and Z represent the incidence matrices.

The significance of the fixed effects was assessed based on the exact Fisher criterion. To optimize the set of random effects, the Akaike criterion was used [36]. The genomic relationship matrix G was obtained using the linear method proposed by Van Raden [37]:

$$G = \frac{M'M}{2 \sum_{i=1}^m p_i(1-p_i)}$$

where m is the number of SNPs, p_i is the frequency of the minor allele of SNP i , and M is the centered genotype matrix.

Fixed and random effects in the ssGBLUP and wssGBLUP models are the same as in the BLUP AM model. The relationship matrix (A^{-1}) is replaced by (H^{-1}).

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & \tau(\alpha G + \beta A_{22})^{-1} - \omega A_{22} \end{bmatrix}$$

where A is the relationship matrix, A_{22}^{-1} is the inverse relationship matrix for genotyped animals, G is the genomic relationship matrix, and the coefficients $\alpha = 0.95$, $\beta = 0.05$, $\tau = \omega = 1$ (by default) [38].

The EBV reliability (GEBV) was calculated as follows:

$$Rel = 1 - \frac{PEV}{varA}$$

where PEV is the estimated error variance, and $varA$ is the additive genetic variance.

In the wssGBLUP approach (weighted ssGBLUP) genomic matrix (G) is replaced by a weighted G^* matrix, which is constructed as follows:

$$G^* = \frac{M'DM}{2 \sum_{i=1}^m p_i(1-p_i)}$$

In the wssGBLUP analysis, the D matrix is a diagonal matrix of the SNP weights [11]. SNP weights were obtained by solving the inverse GEBV equation using POSTGSF90 software [36] based on the non-linear method for calculating SNP weights (by Van Raden [37]):

$$d_{ii} = 1.125 \frac{|\hat{a}_i|}{sd(\hat{a})} - 2$$

The wssGBLUP approach is based on an iterative algorithm with the following steps: (1) perform ssGBLUP with matrix G^* (at iteration 1, all SNP weights in matrix D are equal

to 1 and are equivalent to ssGBLUP), (2) estimate the effects of SNP values of genomic scores in the previous step, (3) evaluate the effect of each SNP, (4) normalize the SNP effect vector to obtain SNP weights (the normalization process ensures that the sum of the SNP variances remains constant), (5) use the SNP weights to construct a new D-matrix, and (6) move to step (1).

3. Results

The pig breeding value estimates and their reliability were calculated according to three meat and fattening traits and two reproduction traits. Genomic scores (GEBV) were obtained using all three compared reference groups (Scenarios 1–3) for genotyped animals of the smallest groups: the first and fourth. A comparison of the average reliability values of the GEBV estimates indicates that accuracy grows for each analyzed method with increasing the reference group size from +4.0 to +9.9% points and from +0.0 to +2.4% points for the ssGBLUP and wssGBLUP methods, respectively. Thus, the accumulation of genomic data for the reference group gives the most significant increase in accuracy when using the ssGBLUP method. In contrast, the reference group size enlargement leads to a relatively smaller increase in the prediction accuracy for the wssGBLUP method.

In the previous article, we discussed the increase in the accuracy of breeding value estimates by including genomic data on the characteristics of meat and fattening productivity and reproduction traits in pigs [39]. However, in this study, we decided to evaluate the validity of the procedure application for individual SNP marker effect weighting and compared the obtained estimates' reliability with traditional BLUP AM (Scenario 4) and ssGBLUP (Table 4).

Table 4. Average animal estimate reliability of the first and fourth reference groups (530 and 396 animals), depending on the method and amount of genomic data.

Trait	BLUP AM		ssGBLUP		wssGBLUP		
	Scenario 4	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3
Meat and fattening traits							
BF1	0.723	0.716	0.757	0.767	0.919	0.931	0.933
MD	0.597	0.606	0.654	0.666	0.931	0.942	0.944
Age	0.679	0.675	0.718	0.728	0.983	0.986	0.986
Reproduction traits							
NBA	0.438	0.465	0.484	0.499	0.803	0.803	0.813
TNB	0.441	0.473	0.492	0.511	0.826	0.828	0.840

BF1: backfat thickness over 6–7 ribs, MD: muscle depth, Age: days to 100 kg, TNB: number of all piglets born at the first farrowing, NBA: number of piglets born alive at the first farrowing.

We note that the wssGBLUP method has an advantage over ssGBLUP in estimated reliability. For example, when using a small reference group (Scenario 1), the difference in the accuracy of ssGBLUP over BLUP AM is from -1.9 to $+7.3$ percent points, while for wssGBLUP, the change in accuracy varies from $+18.2$ to $+87.3$ percent points. Furthermore, the superiority of the second method is also maintained for the largest group of genotyped animals (Scenario 3): from $+4.7$ to $+15.9$ percent points for ssGBLUP and from $+21.1$ to $+90.5$ percent points for wssGBLUP. The application of the wssGBLUP method using the genomic data for a reference group of 500 animals (Scenario 1) at first glance already makes it possible to obtain highly reliable predictive values of the pig breeding values in terms of meat fattening productivity and reproduction indicators.

The correlation coefficients between animal estimates of the first group (530 animals) GEBV_{wss} and GEBV_{ss} based on meat and fattening productivity, obtained using three reference groups, were close to 1 (from 0.950 to 0.994 for GEBV_{ss} and from 0.970 to 0.995 for GEBV_{wss}). In our opinion, this indicates the high reliability of both methods concerning highly inherited traits. On the other hand, for reproduction traits, the correlation coefficients

between estimates were characterized by lower Pearson correlation values: from +0.823 to +0.980 for $GEBV_{ss}$ and from +0.878 to +0.981 $GEBV_{wss}$ (results not shown in the table).

The validation of genomic estimates using ssGBLUP and wssGBLUP methods was carried out by comparing genomic predictions based on a partial data set (PD) (Scenario 5) with the most accurate genomic estimates based on complete information about animals, whole data set (WD) (Scenario 3) (Table 5).

Table 5. Correlation of the weighted final genomic estimates ($GEBV_{wssWD}$) of the validated group (200 animals) with their parent's average predictions (PA), genomic predictions based on pedigree and genome data using the ssGBLUP ($GEBV_{ssPD}$) and wssGBLUP ($GEBV_{wssPD}$) methods (Scenario 5) *.

Correlation between		Trait				
		BF1	MD	Age	TNB	NBA
EBV_{WD}	PA_{PD}	0.358	0.336	0.387	0.787	0.767
	$GEBV_{ssPD}$	0.509	0.500	0.452	0.697	0.690
	$GEBV_{wssPD}$	0.505	0.436	0.426	0.618	0.594
$GEBV_{ssWD}$	PA_{PD}	0.624	0.506	0.699	0.751	0.721
	$GEBV_{ssPD}$	0.993	0.942	0.969	0.875	0.878
	$GEBV_{wssPD}$	0.569	0.520	0.488	0.799	0.791
$GEBV_{wssWD}$	PA_{PD}	0.356	0.278	0.373	0.615	0.574
	$GEBV_{ssPD}$	0.565	0.498	0.479	0.686	0.693
	$GEBV_{wssPD}$	0.562	0.457	0.466	0.713	0.716

* All correlation coefficients are significant with p -values < 0.001. BF1: backfat thickness over 6–7 ribs, MD: muscle depth, Age: days to 100 kg, TNB: number of all piglets born at the first farrowing, NBA: number of piglets born alive at the first farrowing.

A preliminary evaluation of young animals was performed based on the pedigree and was a prediction based on the parent average (PA). In our study, the predictive indicators of a validated group of individuals on a partial data set (PA_{PD} , $GEBV_{ssPD}$, $GEBV_{wssPD}$) were compared with their most accurate breeding value estimates ($GEBV_{wssWD}$) on the whole data set.

A genomic estimation method validation was carried out on genomic predictions correlated with highly reliable estimates for each method (Scenario 5). Of note is the fact that for traits with relatively higher heritability (meat and fattening qualities), the relationship of the predicted values with the final estimates was the lowest for PA_{PD} , with the exception of the correlation $r(PA_{PD}, GEBV_{ssWD})$. Estimates based on weighted values are either close to or slightly inferior to the $GEBV_{ssPD}$ estimates. The final estimates (EBV_{WD} , $GEBV_{ssWD}$, $GEBV_{wssWD}$) have the highest correlation with predictions calculated using the same method. However, highly reliable estimates based on BLUP are more closely interrelated with genomic predictions. The final weighted estimates correlate almost at the same level with the predictions of $GEBV_{ssPD}$ and $GEBV_{wssPD}$. For traits with low heritability (TNB, NBA), $GEBV_{wssWD}$ scores are also inferior to $GEBV_{ssWD}$. In the case of traditional BLUP, the greatest relationship is characterized by PA_{PD} . It follows from the results that genomic methods are most justified in predicting the true breeding value of individuals for traits with relatively high heritability. For low-inherited traits, the ssGBLUP method has an advantage over wssGBLUP.

This is also confirmed by the analysis of the inflation of the final estimates for preliminary predictions (EBV_{WD} , $GEBV_{ssWD}$, $GEBV_{wssWD}$ on PA_{PD} , $GEBV_{ssPD}$, $GEBV_{wssPD}$). The level of the regression coefficient (b), when predictions and final estimates were based on the same method, varied. For the BLUP method, it ranged from 0.64 to 0.72 (R^2 from 0.11 to 0.15) and 1.06 (R^2 from 0.59 to 0.62); for the ssGBLUP method, it ranged from 0.78 to 0.87 (R^2 from 0.26 to 0.34) and from 0.99 to 1.01 (R^2 from 0.77 to 0.77); and for the wssGBLUP method, it ranged from 0.57 to 0.78 (R^2 from 0.21 to 0.32) and from 0.76 to 0.77 ($R^2 = 0.51$) for meat and fattening traits and reproduction traits, respectively. In the case when different methods obtained the prediction and final estimates, the best values of inflation (close to

one) and the coefficient of determination were identified for $GEBV_{ssPD}$ (from 0.70 to 0.88) (R^2 from 0.20 to 0.32) for meat and growth traits, from 0.69 to 1.27 (R^2 from 0.47 to 0.48) for reproduction traits), the worst values of b for meat and fattening traits had PA_{PD} , and for reproduction traits— $GEBV_{wssPD}$ (the data is given in the Supplementary Table S1).

In our opinion, such a significant increase in the reliability of $GEBV_{wss}$ estimates compared to $GEBV_{ss}$ is associated with a decrease in the residual variance due to the use of more optimal contributions of SNP markers when using the weighting procedure for the effect of each SNP marker (with the constant genetic variance). However, the validation of genomic methods (ssGBLUP and wssGBLUP), a test of the accuracy of genomic estimates [10], indicates a relatively higher predictive ability of the analyzed methods of the genomic assessment of complex polygenic traits in comparison with the assessment by pedigree, it reveals a slight improvement in the prognosis due to the use of the weighting procedure. This is due to several factors: the absolute effect values of individual SNP markers are very small, and there are no main effects for the analyzed traits. Thus, with a given limited amount of genomic data (approximately 1.5 thousand samples), a change in the set of phenotypic data (the phenotypes of the validated group were excluded from the calculation) leads to a significant change in these effect estimates and, accordingly, to a redistribution of each marker weight, which is observed during validation. For a given amount of genomic data, the wssGBLUP method is associated with the risk of obtaining biased results if there is a significant change in the reference group based on which the marker weights were estimated (which will be shown below).

We analyzed the levels of the EBV relationship for BLUP AM (Scenario 4) in individuals of the first and fourth groups, obtained for three methods using genomic data from different reference groups (Scenarios 1–3) (Table 6).

Table 6. Correlation of EBV/GEBV of individuals of the first and fourth reference groups (530 and 396 heads) obtained by compared methods *.

EBV by Trait		GEBV _{ss}			GEBV _{wss}		
Meat and fattening traits							
Scenario 4	Scenario 1	Scenario 2	Scenario 3	Scenario 1	Scenario 2	Scenario 3	
EBV(BF1)	0.974	0.950	0.945	0.977	0.956	0.951	
EBV(MD)	0.957	0.927	0.910	0.926	0.918	0.911	
EBV(Age)	0.961	0.937	0.931	0.965	0.948	0.944	
Reproduction traits							
EBV(NBA)	0.944	0.781	0.823	0.792	0.843	0.841	
EBV(TNB)	0.944	0.769	0.808	0.775	0.840	0.846	

* All correlation coefficients are significant with p -values < 0.001 . EBV(BF1): estimated breeding value on BLUP AM by backfat thickness over 6–7 ribs, EBV(MD): estimated breeding value on BLUP AM by muscle depth, EBV(Age): estimated breeding value on BLUP AM by days to 100 kg, EBV(TNB): estimated breeding value on BLUP AM by the number of all piglets born at the first farrowing, EBV(NBA): estimated breeding value on BLUP AM by the number of piglets born alive at the first farrowing.

As can be seen from the results, an increase in the size of the reference group for both methods reduces the relationship level between EBV and GEBV scores, which is most pronounced for the $GEBV_{wss}$ estimates, apparently due to the adjustment of marker effects.

It should be noted that the level of correlation between the genomic estimates themselves, obtained by the two compared methods, is very high (for meat and fattening traits $r_{(GEBV_{ss}, GEBV_{wss})}$ varies from +0.887 to +0.999 depending on the group; for reproduction traits it varies from +0.737 to +0.868), and from the point of view of pigs practical selection, these differences in results are not so significant. However, the change in the correlation coefficients indicates a stable decrease in the relationship between genomic scores and EBV for meat and fattening traits with the accumulation of genomic information, while changes in reproductive qualities are less predictable.

We analyzed the values of the SNP marker weights based on the contribution of each marker effect to the final breeding value assessment, as well as the part of the total genetic

variance described by each marker and the p -values characterizing the significance of the marker effect contribution into the trait variability (Table 7).

Table 7. Correlation of parameters of compared reference groups.

Trait/Group		Part of Variance Explained by Each SNP Marker		p -Value *		SNP Marker «Weight»	
Meat and fattening traits							
BF1	Scenario 2	0.429	-	0.382	-	0.408	-
	Scenario 3	0.352	0.850	0.314	0.834	0.352	0.839
MD	Scenario 2	0.511	-	0.454	-	0.479	-
	Scenario 3	0.435	0.816	0.353	0.790	0.390	0.796
Age	Scenario 2	0.394	-	0.366	-	0.404	-
	Scenario 3	0.335	0.822	0.291	0.809	0.330	0.811
Reproduction traits							
NBA	Scenario 2	0.424	-	0.395	-	0.436	-
	Scenario 3	0.318	0.720	0.277	0.714	0.331	0.732
TNB	Scenario 2	0.420	-	0.384	-	0.423	-
	Scenario 3	0.294	0.725	0.263	0.716	0.313	0.733

* All correlation coefficients are significant with p -values < 0.001. BF1: backfat thickness over 6–7 ribs, MD: muscle depth, Age: days to 100 kg, TNB: number of all piglets born at the first farrowing, NBA: number of piglets born alive at the first farrowing.

The correlation coefficient analysis of the variance proportions explained by SNP markers indicates a redistribution of these values for groups that differ significantly in size (1 and 3, 4 and 6), which was also confirmed by the relatively low consistency of the SNP markers' weight values. Each marker's contribution varies depending on the reference group examined (Scenarios 1–3). A similar trend was observed for the p -values, confirming the more remarkable similarity of the significant marker sets identified in groups 2 and 3. The SNP marker allele frequencies analysis revealed a slight change in the compared reference groups. The correlation of allele frequencies between different groups was as follows: $r_{(gr1,gr2)} = +0.929$, $r_{(gr1,gr3)} = +0.901$, and $r_{(gr2,gr3)} = +0.959$ for meat and fattening traits, and $r_{(gr1,gr2)} = +0.936$, $r_{(gr1,gr3)} = +0.904$, and $r_{(gr2,gr3)} = +0.959$ for reproduction traits.

The power of the method to take into account the individual SNP marker weight as a predictor of breeding value becomes significantly lower in situations when the true QTLs of the greatest influence on the trait variability are unknown and the traits themselves are polygenic. The traits analyzed in the study are characterized by very low proportions of variance described by individual SNP markers (Figure 1). Thus, about 5% of the genetic variance of the BF1 trait is explained by 88 SNP markers when Scenario 1 using. For all analyzed traits, the number of markers explaining 5% of genetic variability varied from 71 to 108, and the number of such SNPs varied depending on the reference group size (79–88 for BF1, 72–81 for MD, 71–108 for age) (Figure 2).

It is noteworthy that the results obtained for the genetic variance distribution over the entire set of SNP markers for the analyzed traits have the most remarkable similarity between the groups Scenario 2 and 3. A reference group size of about 1000 (Scenario 2) and about 1500 heads (Scenario 3) gives more consistent results for evaluation based on the wssGBLUP method. Thus, applying this procedure using a reference group of 500 heads can lead to somewhat distorted GEBV results.

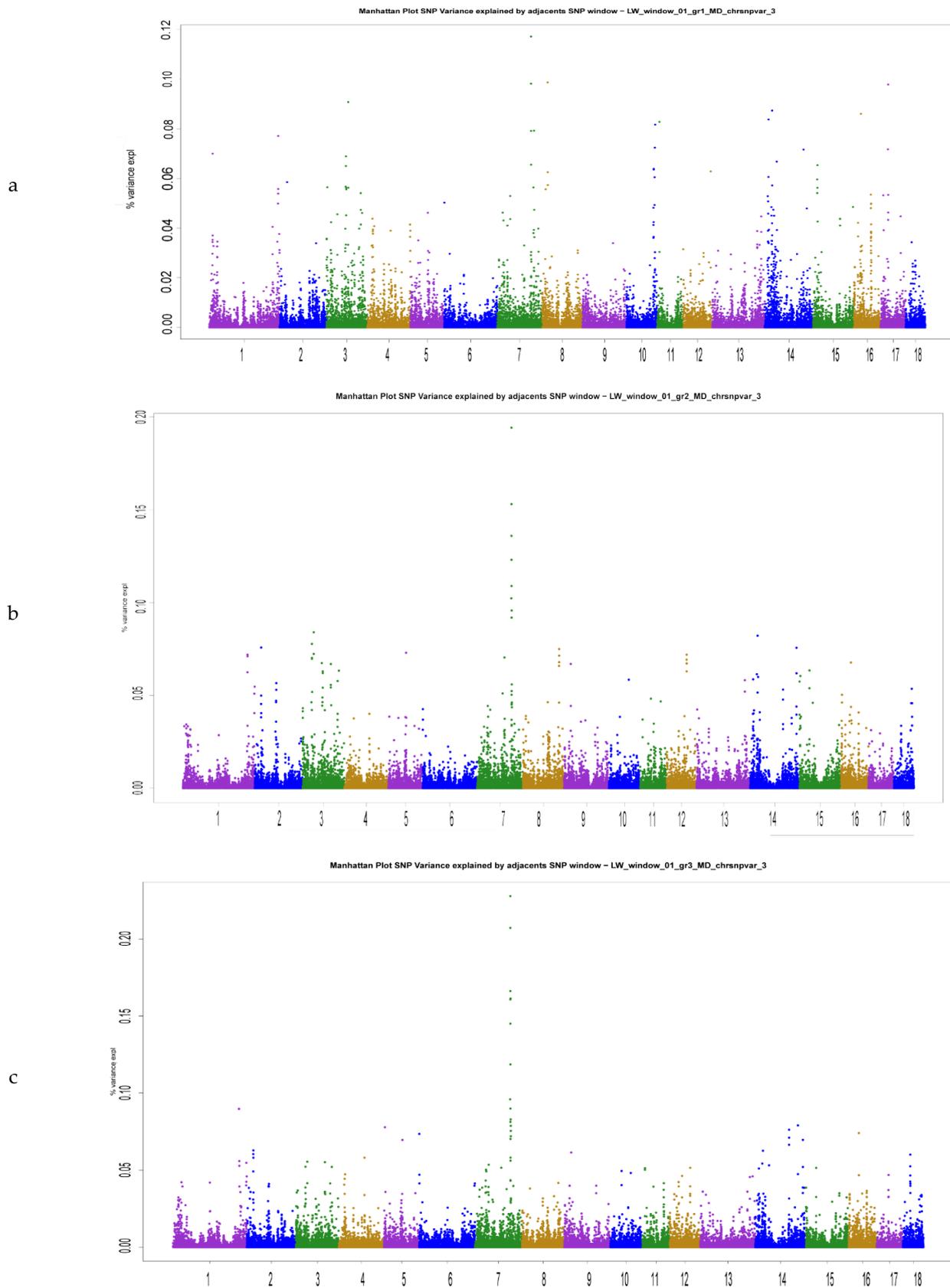


Figure 1. Manhattan plots of muscle depth (MD) variance explained by SNP markers based on Scenarios 1–3 (a—Scenario 1, b—Scenario 2, c—Scenario 3).

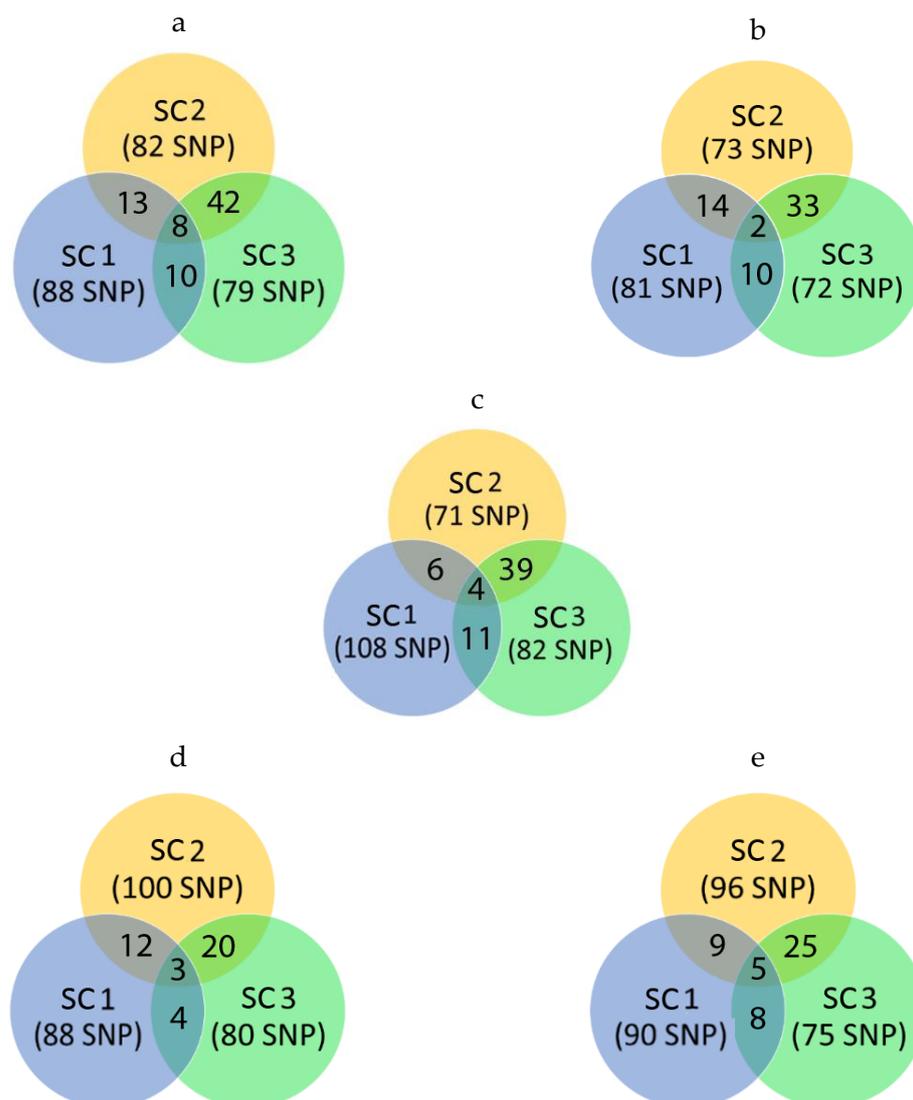


Figure 2. Trait/Venn diagrams showing common SNP in different groups. Values in brackets indicate the total number of SNP markers which explain 5% of a trait variability based on the respective reference group. Set intersections indicate the number of SNP markers that explain 5% of the trait variability common for calculations based on Scenarios 1–3 (SC1—Scenario 1, SC2—Scenario 2, SC—Scenario, 3) **a**—for backfat thickness over 6–7 ribs (BF1), **b**—for muscle depth (MD), **c**—for days to 100 kg (Age), **d**—for number of all piglets born at the first farrowing (TNB), **e**—for number of piglets born alive at the first farrowing (NBA).

4. Discussion

Causal QTN information can be included in the ssGBLUP procedure using a weighted genomic relatedness matrix. For example, a study by [24] performed on simulated data showed that the inclusion of realistic weights for causal QTNs allows for genomic estimates of higher accuracy. Moreover, it was assumed that the pre-selection of SNP markers (as a tool for optimizing the model prediction) should increase the accuracy of genomic estimates. However, studies [29] for the Japanese flounder (*Paralichthys olivaceus*) based on disease resistance showed that a significant increase in the prediction reliability was not observed with the preliminary selection of SNPs based on both SNP-GWAS (by p -value) and ssGWAS (by a fraction of the marker explained variance) [29]. In this connection, the authors conclude that the preliminary marker selection does not significantly increase the accuracy of the estimates while analyzing low heritable traits. In our study, the results also indicate a rather significant redistribution of the genetic variance proportions described by

individual SNP markers of a trait, which does not allow for a preliminary marker selection for genomic prediction based on the estimated traits in pigs.

A comparison of the genomic estimate accuracy obtained using wssGBLUP with traditional BLUP estimates was carried out by [40] for Holstein cows based on calving interval. The gain in GEBV accuracy with wssGBLUP was +5.4 to +5.7 (first-calf heifers) and +9.4 to +9.7 (multi-calving cows) percent points compared to pedigree-based BLUP accuracy [40]. Lourenco et al. [25] also compared the genomic prediction accuracy for dairy productivity traits in cows obtained using several methods (Bayesian regression, genomic BLUP, single-step GBLUP and weighted single-step GBLUP). The authors found the superiority of the ssGBLUP method for populations with relatively few genotyped animals. However, it was also noted that weighted ssGBLUP has the potential to improve the accuracy of the estimates [25]. Lourenco et al. showed on simulated data that wssGBLUP is more efficient when the genotyping population is small and a small number of QTLs affect the trait. When using medium-size genotyped populations, ssGBLUP can be applied without losing accuracy compared to wssGBLUP [41].

The most accurate marker weights in the study by Atashi et al. [40] were obtained at the second iteration of the wssGBLUP procedure. Similar results for SNP marker weight accuracy assessment at the second iteration were obtained in the study by Mehrban et al. (2021) using a non-linear calculation method for meat traits and carcass quality in beef cattle. Hanwoo [42], as well as Teissier M et al. [43,44], assessed for various traits in dairy goats using the quadratic SNP weighting procedure and windows of 40 SNPs. Our study found the highest estimate accuracy at the third iteration for all analyzed traits, similar to the study by Zhang et al. (2016) [45] on simulated data. The effectiveness of the weighted SNP effects was compared on dairy goats [43] for milk protein content. As a result, the authors show that the weighted ssGBLUP is more accurate than ssGBLUP (+8 to +12 percent points). The same authors [44] found that the wssGBLUP method gives a slight advantage over ssGBLUP and only for traits with identified QTLs when studying milk productivity traits, udder quality, and the number of somatic cells in goat milk [43,44]. Our research results confirm the increase in genomic estimate reliability when using wssGBLUP compared to ssGBLUP and BLUP AM. Moreover, the increase in the reliability of $GEBV_{wss}$ relative to EBV ranges from +18.2 to +90.5 percent points. Mehrban et al. (2021) showed a non-uniform improvement in the accuracy of beef trait weighted scores in beef cattle compared to ssGEBV [42]. Although accuracy gains ranged from 5 to 13 percentage points (when using individual SNPs), for BF (fat thickness) and MS (marbling score), significant accuracy gains between ssGBLUP and wssGBLUP were not observed.

Embedding genomic assessment into the overall livestock breeding value assessment system implies the presence of a reference population (group) of genotyped animals of sufficient size with optimal parameters to obtain the most reliable genomic prediction results. However, the issue of determining these parameters and volume still cannot be resolved unambiguously. Alvarenga et al. (2020) [46] compared the ssGBLUP and wssGBLUP methods based on different training populations (only purebred individuals and the combined population—purebred plus crossbreeds). As a result, it was found that the ssGBLUP method gave the highest accuracy and the smallest GEBV bias. However, the authors noted the stratification of the genotyped population (purebred and crossbreeds) [46]. In our study, there was no stratification in the compared groups; therefore, the advantage of the wssGBLUP method was observed.

Ardestani et al. (2021) [3] proved the superiority of genomic methods for predicting breeding value over the traditional BLUP procedure in Duroc, Landrace and Yorkshire pigs for meat traits. The greatest estimate reliability for fat thickness, muscle depth and average daily gain was achieved by ssGBLUP and BayesCp. However, the increase in reliability was different for each breed studied. According to the authors, this occurrence can be caused by differences in the reference group's volume, the amount of data available for evaluation, and the differences in the breeding selection goals [3].

The accuracy of genomic estimates in the study by Grossi et al. (2018) [27] on pigs of three breeds (Duroc, Landrace and Yorkshire), calculated as the Pearson's correlation of deregressed breeding value estimates with genomic predictions for fat thickness, muscle depth, and number of piglets born ranged from 0.39 to 0.55 (by BF1), from 0.40 to 0.68 (by MD), and from 0.47 to 0.58 (by TNB) with a decrease in these indicators with a decrease in the number of included SNP markers (by the development of low-density panels), and slightly exceeded the pedigree-based prediction (for BF1 accuracy, this value ranged from 0.40 to 0.54; for MD, it ranged from 0.52 to 0.61; for TNB, it ranged from 0.26 to 0.64) [27]. In our study, validation revealed a similar level of correlation between wssGBLUP and ssGBLUP genomic predictions (Scenario 5) with highly accurate final estimates of $GEBV_{wssWD}$ (Scenario 3): for BF1, values ranged from 0.51 to 0.56; for MD, values ranged from 0.45 to 0.50; for age, values ranged from 0.43 to 0.48; for TNB, values ranged from 0.69 to 0.71; and for NBA, values ranged from 0.69 to 0.72, which exceeds the accuracy of predictions based on the pedigree using the BLUP AM method (the correlation of PA with $GEBV_{wssWD}$ was 0.36 for BF1, 0.34 for MD, 0.39 for age, 0.62 for TNB and 0.57 for NBA).

It should be noted that the genomic assessment procedure (based on the ssGBLUP and wssGBLUP methods) can significantly improve the preliminary prediction accuracy of the individual genetic values within a particular population. An increase in the reference group size provides more stable individual marker weight values, and the proportion of genetic variance explained by each of them, especially when the reference group is closely related to the assessed population [23]. Since the strength of the LD between SNP markers and QTL may decay over time due to the impact of genetic forces, genomic predictions calculated based on the reference group may lose accuracy. In this regard, constant retraining of the prediction model, reassessment of SNP effects and constant replenishment of the phenotypic data set are required [7]. An additional increase in the genomic prediction reliability can be achieved by using SNP marker weights, which should be calculated separately for each trait [25].

5. Conclusions

Using the ssGBLUP and wssGBLUP methods makes it possible to assess pigs' genomic breeding values based on the main breeding traits (meat and fattening qualities and reproduction indicators of the first farrowing) with high precision. Using the weighting procedure to evaluate SNP markers effects significantly improves the accuracy of the genomic predictions. The wssGBLUP implementation is reasoned for genotyped reference groups of as few as 1000 individuals (meat and fattening qualities) or 1500 animals (reproduction indicators). Further replenishment of the reference group involves the mandatory recalculation of the SNP markers weights for genomic assessment of polygenic traits influenced by a large number of QTLs, such as meat, fattening and reproduction qualities in pigs.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ani12131693/s1>, Table S1: Regression coefficients (b) of final estimates (EBV_{WD} , $GEBV_{ssWD}$, $GEBV_{wssWD}$) on predictions (PA_{PD} , $GEBV_{ssPD}$, $GEBV_{wssPD}$) of validated animals.

Author Contributions: Conceptualization, E.M. (Ekaterina Melnikova); data curation, A.K., S.N., O.K., T.K., E.M. (Elena Martynova) and E.T.; funding acquisition, V.V.; investigation, A.K. and E.M. (Ekaterina Melnikova); methodology, A.K. and E.M. (Ekaterina Melnikova); resources, A.K., S.N., O.K., T.K., E.M. (Elena Martynova) and E.T.; software, A.K., S.N. and O.F.; validation, E.M. (Ekaterina Melnikova), M.S. and O.F.; visualization, S.N.; writing—original draft, A.K. and E.M. (Ekaterina Melnikova); writing—review & editing, E.M. (Ekaterina Melnikova), M.S. and O.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Russian Science Foundation, grant number 21-76-10038.

Institutional Review Board Statement: The study was performed in accordance with the ethical guidelines of the L.K. Ernst Federal Research Center for Animal Husbandry. Protocol No 12/1 was approved by the Commission on the Ethics of Animal Experiments of the L.K. Ernst Federal Research Center for Animal Husbandry on 4 December 2019.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset (phenotype and genotype information) of the Large White pigs used and analysed during the current study are owned by LLC “TOPGEN” and were provided via a signed data access agreement, which does not allow for data sharing.

Acknowledgments: The authors are grateful for the information provided for the research to Permyakov, A.G., Kazmina, N.A., LLC “TOPGEN” (Voronezh region, Verkhnyaya Khava village). The authors express their sincere gratitude to the three reviewers for their helpful comments and suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* **2001**, *157*, 1819–1829. [CrossRef] [PubMed]
2. Schaeffer, L.R. Strategy for applying genome-wide selection in dairy cattle. *J. Anim. Breed. Genet.* **2006**, *123*, 218–223. [CrossRef] [PubMed]
3. Ardestani, S.S.; Jafarikia, M.; Sargolzaei, M.; Sullivan, B.; Miar, Y. Genomic Prediction of Average Daily Gain, Back-Fat Thickness, and Loin Muscle Depth Using Different Genomic Tools in Canadian Swine Populations. *Front. Genet.* **2021**, *12*, 735. [CrossRef]
4. Misztal, I. Inexpensive Computation of the Inverse of the Genomic Relationship Matrix in Populations with Small Effective Population Size. *Genetics* **2015**, *202*, 401–409. [CrossRef]
5. Pocrnic, I.; Lourenco, D.A.L.; Masuda, Y.; Legarra, A.; Misztal, I. The Dimensionality of Genomic Information and Its Effect on Genomic Prediction. *Genetics* **2016**, *203*, 573–581. [CrossRef] [PubMed]
6. Badke, Y.M.; Bates, R.O.; Ernst, C.W.; Fix, J.; Steibel, J.P. Accuracy of Estimation of Genomic Breeding Values in Pigs Using Low-Density Genotypes and Imputation. *G3-Genes Genomes Genet.* **2014**, *4*, 623–631. [CrossRef]
7. MacLeod, I.M.; Hayes, B.; Goddard, M. The Effects of Demography and Long-Term Selection on the Accuracy of Genomic Prediction with Sequence Data. *Genetics* **2014**, *198*, 1671–1684. [CrossRef]
8. Pérez-Enciso, M.; Rincón, J.C.; Legarra, A. Sequence- vs. chip-assisted genomic selection: Accurate biological information is advised. *Genet. Sel. Evol.* **2015**, *47*, 43. [CrossRef]
9. Mrode, R.A. *Linear Models for the Prediction of Animal Breeding Values*, 3rd ed.; CABI: Boston, MA, USA, 2013.
10. Lourenco, D.A.L. Introduction to Genomic Selection. University of Chicago: Piracicaba, Brazil, 2019. Available online: http://nce.ads.uga.edu/wiki/lib/exe/fetch.php?media=gs_lourenco_2019a.pdf (accessed on 29 June 2022).
11. Wang, H.; Misztal, I.; Aguilar, I.; Legarra, A.; Muir, W.M. Genome-wide association mapping including phenotypes from relatives without genotypes. *Genet. Res.* **2012**, *94*, 73–83. [CrossRef]
12. Braz, C.U.; Taylor, J.F.; Bresolin, T.; Espigolan, R.; Feitosa, F.L.B.; Carvalheiro, R.; Baldi, F.; De Albuquerque, L.G.; De Oliveira, H.N. Sliding window haplotype approaches overcome single SNP analysis limitations in identifying genes for meat tenderness in Nelore cattle. *BMC Genet.* **2019**, *20*, 8. [CrossRef]
13. Cherel, P.; Pires, J.; Glénisson, J.; Milan, D.; Iannuccelli, N.; Héroult, F.; Damon, M.; Le Roy, P. Joint analysis of quantitative trait loci and major-effect causative mutations affecting meat quality and carcass composition traits in pigs. *BMC Genet.* **2011**, *12*, 76. [CrossRef] [PubMed]
14. Fontanesi, L.; Schiavo, G.; Galimberti, G.; Calò, D.G.; Scotti, E.; Martelli, P.L.; Buttazzoni, L.; Casadio, R.; Russo, V. A genome wide association study for backfat thickness in Italian Large White pigs highlights new regions affecting fat deposition including neuronal genes. *BMC Genom.* **2012**, *13*, 583. [CrossRef] [PubMed]
15. Jung, E.J.; Park, H.B.; Lee, J.B.; Yoo, C.K.; Kim, B.M.; Kim, H.I.; Kim, B.W.; Lim, H.T. Genome-wide association analysis identifies quantitative trait loci for growth in a Landrace purebred population. *Anim. Genet.* **2014**, *45*, 442–444. [CrossRef] [PubMed]
16. Sanchez, M.-P.; Tribout, T.; Iannuccelli, N.; Bouffaud, M.; Servin, B.; Tenghe, A.; Dehais, P.; Muller, N.; Del Schneider, M.P.; Mercat, M.-J.; et al. A genome-wide association study of production traits in a commercial population of Large White pigs: Evidence of haplotypes affecting meat quality. *Genet. Sel. Evol.* **2014**, *46*, 12. [CrossRef]
17. Strucken, E.M.; Schmitt, A.O.; Bergfeld, U.; Jurke, I.; Reissmann, M.; Brockmann, G.A. Genome wide study and validation of markers associated with production traits in German Landrace boars. *J. Anim. Sci.* **2014**, *92*, 1939–1944. [CrossRef]
18. Iqbal, A.; Kim, Y.-S.; Kang, J.-M.; Lee, Y.-M.; Rai, R.; Jung, J.-H.; Oh, D.-Y.; Nam, K.-C.; Lee, H.-K.; Kim, J.-J. Genome-wide Association Study to Identify Quantitative Trait Loci for Meat and Carcass Quality Traits in Berkshire. *Asian-Australas. J. Anim. Sci.* **2015**, *28*, 1537–1544. [CrossRef]
19. Meuwissen, T.; Hayes, B.; Goddard, M. Genomic selection: A paradigm shift in animal breeding. *Anim. Front.* **2016**, *6*, 6–14. [CrossRef]

20. Guo, Y.; Huang, Y.; Hou, L.; Ma, J.; Chen, C.; Ai, H.; Huang, L.; Ren, J. Genome-wide detection of genetic markers associated with growth and fatness in four pig populations using four approaches. *Genet. Sel. Evol.* **2017**, *49*, 21. [[CrossRef](#)]
21. Zhu, D.; Liu, X.; Max, R.; Zhang, Z.; Zhao, S.; Fan, B. Genome-wide association study of the backfat thickness trait in two pig populations. *Front. Agric. Sci. Eng.* **2014**, *1*, 91–95. [[CrossRef](#)]
22. Fabbri, M.C.; Zappaterra, M.; Davoli, R.; Zambonelli, P. Genome-wide association study identifies markers associated with carcass and meat quality traits in Italian Large White pigs. *Anim. Genet.* **2020**, *51*, 950–952. [[CrossRef](#)]
23. Liao, Y.; Wang, Z.; Glória, L.S.; Zhang, K.; Zhang, C.; Yang, R.; Luo, X.; Jia, X.; Lai, S.-J.; Chen, S.-Y. Genome-Wide Association Studies for Growth Curves in Meat Rabbits Through the Single-Step Nonlinear Mixed Model. *Front. Genet.* **2021**, *12*, 750939. [[CrossRef](#)] [[PubMed](#)]
24. Fragomeni, B.O.; Lourenco, D.A.L.; Masuda, Y.; Legarra, A.; Misztal, I. Incorporation of causative quantitative trait nucleotides in single-step GBLUP. *Genet. Sel. Evol.* **2017**, *49*, 59. [[CrossRef](#)] [[PubMed](#)]
25. Lourenco, D.; Misztal, I.; Tsuruta, S.; Aguilar, I.; Ezra, E.; Ron, M.; Shirak, A.; Weller, J. Methods for genomic evaluation of a relatively small genotyped dairy population and effect of genotyped cow information in multiparity analyses. *J. Dairy Sci.* **2014**, *97*, 1742–1752. [[CrossRef](#)]
26. Veerkamp, R.F.; Bouwman, A.C.; Schrooten, C.; Calus, M.P.L. Genomic prediction using preselected DNA variants from a GWAS with whole-genome sequence data in Holstein–Friesian cattle. *Genet. Sel. Evol.* **2016**, *48*, 95. [[CrossRef](#)] [[PubMed](#)]
27. Grossi, D.A.; Brito, L.F.; Jafarikia, M.; Schenkel, F.S.; Feng, Z. Genotype imputation from various low-density SNP panels and its impact on accuracy of genomic breeding values in pigs. *Animal* **2018**, *12*, 2235–2245. [[CrossRef](#)] [[PubMed](#)]
28. Moghaddar, N.; Khansefid, M.; Van Der Werf, J.H.J.; Bolormaa, S.; Duijvesteijn, N.; Clark, S.A.; Swan, A.A.; Daetwyler, H.D.; MacLeod, I.M. Genomic prediction based on selected variants from imputed whole-genome sequence data in Australian sheep populations. *Genet. Sel. Evol.* **2019**, *51*, 72. [[CrossRef](#)]
29. Lu, S.; Liu, Y.; Yu, X.; Li, Y.; Yang, Y.; Wei, M.; Zhou, Q.; Wang, J.; Zhang, Y.; Zheng, W.; et al. Prediction of genomic breeding values based on pre-selected SNPs using ssGBLUP, WssGBLUP and BayesB for Edwardsiellosis resistance in Japanese flounder. *Genet. Sel. Evol.* **2020**, *52*, 49. [[CrossRef](#)]
30. Tam, V.; Patel, N.; Turcotte, M.; Bossé, Y.; Paré, G.; Meyre, D. Benefits and limitations of genome-wide association studies. *Nat. Rev. Genet.* **2019**, *20*, 467–484. [[CrossRef](#)]
31. Purcell, S.; Neale, B.; Todd-Brown, K.; Thomas, L.; Ferreira, M.A.R.; Bender, D.; Maller, J.; Sklar, P.; de Bakker, P.I.W.; Daly, M.J.; et al. Plink: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *Am. J. Hum. Genet.* **2007**, *81*, 559–575. [[CrossRef](#)]
32. Sargolzaei, M.; Iwaisaki, H.; Colleau, J.-J. A fast algorithm for computing inbreeding coefficients in large populations. *J. Anim. Breed. Genet.* **2005**, *122*, 325–331. [[CrossRef](#)]
33. Barbato, M.; Orozco-Terwengel, P.; Tapio, M.; Bruford, M.W. SNEP: A tool to estimate trends in recent effective population size trajectories using genome-wide SNP data. *Front. Genet.* **2015**, *6*, 109. [[CrossRef](#)] [[PubMed](#)]
34. Misztal, I.; Tsuruta, S.; Lourenco, D.; Masuda, Y.; Aguilar, I.; Legarra, A.; Vitezica, Z. BLUPF90 Family of Programs. University of Georgia: Athens, GA, USA, 2018. Available online: <http://nce.ads.uga.edu/wiki/doku.php> (accessed on 27 April 2022).
35. Aguilar, I.; Misztal, I.; Johnson, D.L.; Legarra, A.; Tsuruta, S.; Lawlor, T.J. 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*, 743–752. [[CrossRef](#)] [[PubMed](#)]
36. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **1974**, *19*, 716–723. [[CrossRef](#)]
37. VanRaden, P. Efficient Methods to Compute Genomic Predictions. *J. Dairy Sci.* **2008**, *91*, 4414–4423. [[CrossRef](#)]
38. Lourenco, D.; Legarra, A.; Tsuruta, S.; Masuda, Y.; Aguilar, I.; Misztal, I. Single-Step Genomic Evaluations from Theory to Practice: Using SNP Chips and Sequence Data in BLUPF90. *Genes* **2020**, *11*, 790. [[CrossRef](#)] [[PubMed](#)]
39. Melnikova, E.; Kabanov, A.; Nikitin, S.; Somova, M.; Kharitonov, S.; Otradnov, P.; Kostyunina, O.; Karpushkina, T.; Martynova, E.; Sermyagin, A.; et al. Application of Genomic Data for Reliability Improvement of Pig Breeding Value Estimates. *Animals* **2021**, *11*, 1557. [[CrossRef](#)]
40. Atashi, H.; Salavati, M.; De Koster, J.; Crowe, M.A.; Opsomer, G.; Hostens, M. The GplusE Consortium A Genome-Wide Association Study for Calving Interval in Holstein Dairy Cows Using Weighted Single-Step Genomic BLUP Approach. *Animals* **2020**, *10*, 500. [[CrossRef](#)]
41. Lourenco, D.A.L.; Fragomeni, B.O.; Bradford, H.L.; Menezes, I.R.; Ferraz, J.B.S.; Aguilar, I.; Tsuruta, S.; Misztal, I. Implications of SNP weighting on single-step genomic predictions for different reference population sizes. *J. Anim. Breed. Genet.* **2017**, *134*, 463–471. [[CrossRef](#)]
42. Mehrban, H.; Naserkheil, M.; Lee, D.; Cho, C.; Choi, T.; Park, M.; Ibáñez-Escriche, N. Genomic Prediction Using Alternative Strategies of Weighted Single-Step Genomic BLUP for Yearling Weight and Carcass Traits in Hanwoo Beef Cattle. *Genes* **2021**, *12*, 266. [[CrossRef](#)]
43. Teissier, M.; Larroque, H.; Robert-Granie, C. Accuracy of genomic evaluation with weighted single-step genomic best linear unbiased prediction for milk production traits, udder type traits, and somatic cell scores in French dairy goats. *J. Dairy Sci.* **2019**, *102*, 3142–3154. [[CrossRef](#)]

44. Teissier, M.; Larroque, H.; Robert-Granié, C. Weighted single-step genomic BLUP improves accuracy of genomic breeding values for protein content in French dairy goats: A quantitative trait influenced by a major gene. *Genet. Sel. Evol.* **2018**, *50*, 31. [[CrossRef](#)] [[PubMed](#)]
45. Zhang, X.; Lourenco, D.; Aguilar, I.; Legarra, A.; Misztal, I. Weighting Strategies for Single-Step Genomic BLUP: An Iterative Approach for Accurate Calculation of GEBV and GWAS. *Front. Genet.* **2016**, *7*, 151. [[CrossRef](#)] [[PubMed](#)]
46. Alvarenga, A.; Veroneze, R.; De Oliveira, H.R.; Marques, D.B.D.; Lopes, P.S.; Silva, F.F.; Brito, L.F. Comparing Alternative Single-Step GBLUP Approaches and Training Population Designs for Genomic Evaluation of Crossbred Animals. *Front. Genet.* **2020**, *11*, 263. [[CrossRef](#)] [[PubMed](#)]