

Validation of selection accuracy for the total number of piglets born in Landrace pigs using genomic selection

Jae-Don Oh^{1,a}, Chong-Sam Na^{1,a}, and Kyung-Do Park^{1,*}

* **Corresponding Author:** Kyung-Do Park
Tel: +82-63-219-5502, **Fax:** +82-63-219-5505,
E-mail: doobalo@jbnu.ac.kr

¹ Department of Animal Biotechnology, Chonbuk National University, Jeonju 54896, Korea

^a These authors contributed equally to this work.
Submitted May 19, 2016; Revised Jul 22, 2016;
Accepted Aug 10, 2016

Objective: This study was to determine the relationship between estimated breeding value and phenotype information after farrowing when juvenile selection was made in candidate pigs without phenotype information.

Methods: After collecting phenotypic and genomic information for the total number of piglets born by Landrace pigs, selection accuracy between genomic breeding value estimates using genomic information and breeding value estimates of best linear unbiased prediction (BLUP) using conventional pedigree information were compared.

Results: Genetic standard deviation (σ_a) for the total number of piglets born was 0.91. Since the total number of piglets born for candidate pigs was unknown, the accuracy of the breeding value estimated from pedigree information was 0.080. When genomic information was used, the accuracy of the breeding value was 0.216. Assuming that the replacement rate of sows per year is 100% and generation interval is 1 year, genetic gain per year is 0.346 head when genomic information is used. It is 0.128 when BLUP is used.

Conclusion: Genetic gain estimated from single step best linear unbiased prediction (ssBLUP) method is by 2.7 times higher than that the one estimated from BLUP method, i.e., 270% more improvement in efficiency.

Keywords: Genomic Selection, Total Number of Piglets Born, Landrace

INTRODUCTION

As an efficient alternative to accelerate genetic improvement and improve the selection accuracy, genomic selection method using dense single nucleotide polymorphism (SNP) markers covering the whole genome [1,2] has been developed. Genomic selection was first proposed by Meuwissen et al [3]. Through high-density marker mapping, it can predict the genetic ability of individuals. It has been reported that genomic selection is more accurate than conventional breeding value estimation method. Especially when young animals whose phenotypic data could not be obtained for breeding value estimation, highly accurate genomic selection can be used to make juvenile selection possible [2].

To predict genomic information of individuals without genomic information, Gengler et al [4] and VanRaden [5] have proposed algorithms to calculate genomic relationship matrix and the estimation method for obtaining genomic breeding value. Misztal et al [6] have also reported an algorithm to combine conventional pedigree information with genomic information. Recently, Liu et al [7] have developed SNP single step genomic model (SSSmodel) to estimate SNP effect directly. This model have been applied to animals in Germany and Canada for genetic evaluations.

For average daily gain and feed conversion ratio of Duroc pigs in Denmark, accuracy rates of genetic estimation have been compared among best linear unbiased prediction (BLUP)

method (uses pedigree and performance test information for estimation), genomic best linear unbiased prediction (GBLUP) method (uses only genomic information of individual pigs for the estimation), and single step best linear unbiased prediction (ssBLUP, in which all information of individuals with or without genomic information are included in one model for the analysis) [2,8]. It has been reported that the estimation was more accurate when genomic information is used than when only pedigree information is used for animals without genomic information. In addition, ssBLUP was more accurate than BLUP for genetic estimation [2,8].

This experiment was conducted to compare the selection accuracy between genomic breeding value estimation using genomic information and breeding value estimation of BLUP using conventional pedigree information after collecting phenotypic and genomic information for the total number of piglets born in Landrace pigs. In addition, the relationship between estimated breeding value and phenotype information after farrowing was investigated when juvenile selection was made in candidate pigs without phenotype information.

MATERIALS AND METHODS

Single nucleotide polymorphism data and quality control

Using Illumina Porcine SNP60 v1 and v2 beadchip genotype analysis was performed for a total of 1,041 sows and boars with the records of total number of piglets born (TNB) and candidate pigs without records. Genotype information on 62,551 SNP markers were collected, and using 48,245 SNP markers quality control was performed after 12,627 SNPs whose position information was not identified and 1,679 SNPs on sex chromosome were excluded.

For quality control, markers with more than 10% of missing rate, markers without polymorphism (all homo or hetero genotype markers), markers with less than 1% of minor allele frequency, and markers with more than 23.93 ($p < 10^{-6}$) of Hardy Weinberg disequilibrium chi square value, and genomic information of animals with more than 20% of SNP missing rate were excluded. A total of 1,038 animals passed quality control and a total of 36,392 SNP markers were used in this experiment (Table 1).

Phenotypic data for validation of selection accuracy

Of the 1,038 pigs with SNP information obtained from Sunjin GGP farm, a total of 916 pigs (836 sows and 80 boars) were used for the selection accuracy test. Data from 122 candidate pigs were not analyzed. Of the 836 sows with SNP information, 673 sows had records of total number of piglets born. The remaining 163 pigs were gilts without records of total number of piglets born (Table 2). Using the records from 916 pigs, the estimated breeding values for total number of piglets born were compared. To determine the relationship of total number of piglets born with

Table 1. Frequency of variables after the quality control of raw single nucleotide polymorphisms dataset

Variable	Frequency
Selected animals	1,038
Markers on autosomal chromosomes	48,245
Selected (useful) markers	36,392
Outlier markers ¹⁾	11,853
All missing markers	1,551
All homo genotype markers	6,073
All hetero genotype markers	1
Markers with missing over 10%	3,316
Markers with minor allele frequency less than 0.01	10,398
Markers with HW chisquare over 23.93	817

¹⁾Total number of outlier markers across all criteria (not column sum).

phenotype by selection ratio, records for the total number of piglets born from the 163 gilts were collected from 2014 to 2015. Total number of animals in the A matrix and generations were 2,864 heads and 12, respectively.

Statistical models

Linear model for the estimation of genetic parameters and breeding values was as follows:

$$y_{ijkl} = \mu + P_i + YS_j + a_k + pe_k + e_{ijkl}$$

Where, y_{ijkl} = observed value of total number of piglets born, overall average, P_i = i th fixed effect of parity, YS_j = j th fixed effect of farrowing year-season, a_k = k th additive genetic effect (breeding value), pe_k = k th permanent environmental effect of animals, e_{ijkl} = residual effects.

Genetic parameters: Variance components for total number of piglets born were estimated using VCE ver. 6.0 (Institute of Farm Animal Genetics, Friedrich Loeffler Institute, Neustadt, Germany) and heritability (h^2) and repeatability (r) calculation formulars were as follows:

$$h^2 = \frac{\sigma_a^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2}, \quad r = \frac{\sigma_a^2 + \sigma_{pe}^2}{\sigma_a^2 + \sigma_{pe}^2 + \sigma_e^2}$$

Table 2. Distributional properties of data for the total number of piglets born (TNB)

Farrowing year	No. of records	TNB (head) Mean±SD	No. of sows		
			Total	With SNPs	Without SNPs
2007	561	11.3±3.0	284	16	268
2008	536	11.7±3.0	293	54	239
2009	543	11.6±3.0	296	144	152
2010	509	11.3±2.9	291	261	30
2011	552	11.9±3.0	321	304	17
2012	546	12.2±2.9	309	296	13
2013	553	12.0±3.2	316	293	23
Overall	3,800	11.7±3.0	1,282 ¹⁾	673 (163) ²⁾	446

SD, standard deviation; SNPs, single nucleotide polymorphisms.

¹⁾Total number of sows across all farrowing year (not column sum).

²⁾Without TNB records.

where, σ_a^2 = additive genetic variance, σ_{pe}^2 = permanent environmental variance and σ_e^2 = error variance.

Estimation of genomic breeding values: For the fixed effect, parity and farrowing year-month were included and the analysis model was as follows:

$$y = Xb + Za + Wp + e$$

Where, y = $n \times 1$ vector of observation, b = $p \times 1$ vector of fixed effect, a = $q \times 1$ vector of additive genetic random effect, p = $q \times 1$ vector of permanent environmental random effect, $X(n \times p)$, $Z(n \times q)$, and $W(n \times q)$ known incidence matrix corresponding to b , a , and p , $e = n \times 1$ vector of residual effect. Mixed model equation (MME) was as follows:

$$\begin{bmatrix} X'X & X'Z & X'W \\ Z'X & Z'Z + \alpha_1 H^{-1} & Z'W \\ W'X & W'Z & W'W + \alpha_2 I \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \\ W'y \end{bmatrix}$$

Where, $\alpha_1 = \sigma_e^2 / \sigma_a^2$, $\alpha_2 = \frac{\sigma_e^2}{\sigma_{pe}^2}$, and $H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & G^{-1} - A_{22}^{-1} \end{bmatrix}$. A^{-1} = inverse matrix of numerator relationship matrix, G^{-1} = inverse matrix of genomic relationship matrix, A_{22}^{-1} = inverse matrix of numerator relationship matrix of pigs with genomic information. Reliability (r^2) of breeding value was calculated using Prediction error variance (PEV) value by the following formular.

$$r^2 = 1 - \left(\frac{PEV}{\sigma_a^2} \right)$$

RESULTS AND DISCUSSION

Genetic parameters

In this experiment, the estimated heritability of total number of piglets born was 0.096, which was higher than that (0.08) of Su et al [9] and Chen et al [10]. It has been reported that the heritability estimates of total number of piglets born in Yorkshire, Duroc, and Berkshire pigs were 0.10, 0.09, and 0.07, respectively [10,11]. The heritability for total number of piglets born could vary due to different genetic characteristics of the population. Generally, it is around 0.1 or lower [12]. The estimated repeatability for total number of piglets born was 0.195. Variance components and genetic parameters are shown in Table 3.

Table 3. Additive genetic (σ_a^2), permanent environmental (σ_{pe}^2), residual (σ_e^2) variance components, heritabilities (h^2) and standard errors (SE), and repeatabilities (r) for the total number of piglets born (TNB).

Trait	Variance components			$h^2 \pm SE$	r
	σ_a^2	σ_{pe}^2	σ_e^2		
TNB	0.8284	0.8616	6.9664	0.096 ± 0.001	0.195

Table 4. Means and standard deviations of estimated breeding value (EBV) and genomic estimated breeding value (GEBV) accuracies in animals with or without records for the total number of piglets born (TNB) by the estimation methods

Method	Overall (836 head)	Sows with records (673 head)	Sows without records (163 head)
BLUP	0.262 ± 0.115	0.306 ± 0.080	0.080 ± 0.024
ssBLUP	0.356 ± 0.094	0.390 ± 0.065	0.216 ± 0.055

BLUP, best linear unbiased prediction; ssBLUP, single step best linear unbiased prediction.

Accuracy of estimated breeding value and genomic estimated breeding value

Since the accuracies of estimated breeding value can vary with different numbers of test records, pedigree information, and heritability, after classifying sows with or without records for the total number of piglets born, the accuracies of breeding values obtained from BLUP and ssGBLUP were estimated. Results are shown in Table 4. When ssBLUP method was used, the accuracy was higher than that when conventional BLUP method was used. The difference in accuracy was bigger in pigs without phenotypic records. This results coincided with those of Forni et al [13]. As a result, there was relatively high accuracy difference in young animals without phenotypic data. When genomic information was used, the accuracy was increased. Such improved accuracy might be due to the fact that numerator relationship matrix used for BLUP was substituted by the genomic relationship matrix derived from the similarity of genotype of all markers with Mendelian sampling.

Correlation analysis was performed between genomic estimated breeding value (GEBV) estimated from ssBLUP and breeding value estimated from conventional BLUP using only pedigree information (Figure 1). The estimated correlation coefficient of between GEBV and estimated breeding value (EBV) was 0.83. The estimated regression equation was $y = 0.12 + 0.94x$ ($r^2 = 0.68$).

When there was no phenotypic data in the progeny of pigs with the estimated genomic breeding value, breeding value

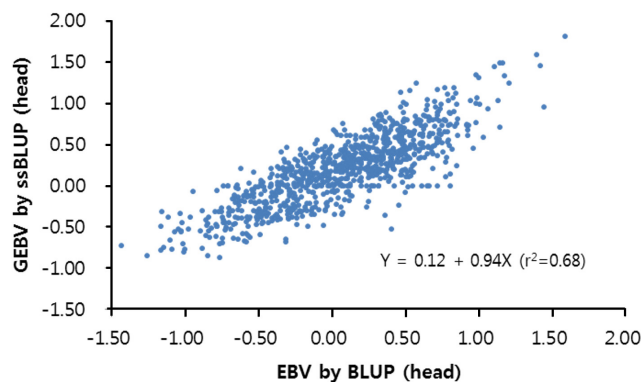


Figure 1. Correlation between estimated breeding value (EBV) and genomic estimated breeding value (GEBV) for the total number of piglets born (TNB).

Table 5. Distributional properties of validation data of the phenotypic total number of piglets born (TNB) by best linear unbiased prediction (BLUP) and single step best linear unbiased prediction (ssBLUP) methods

Farrow year	No. of records	TNB (head)		No. of sows	
		Mean±SD	Total	With SNPs	Without SNPs
2007	561	11.3±3.0	284	16	268
2008	536	11.7±3.0	293	54	239
2009	543	11.6±3.0	296	144	152
2010	509	11.3±2.9	291	261	30
2011	552	11.9±3.0	321	304	17
2012	546	12.2±2.9	309	296	13
2013	553	12.0±3.2	316	293	23
2014	576	12.4±3.3	348	311 (156) ²⁾	37
2015	398	12.9±3.1	262	157 (7) ²⁾	105
Overall	4,774	11.9±3.1	1,394 ¹⁾	836	558

SD, standard deviation; SNPs, single nucleotide polymorphisms.

¹⁾ Total number of sows across all farrowing year (not column sum).

²⁾ Number of candidate sows that had the first TNB records with SNPs information.

estimated using BLUP method was the average breeding value of parents. Therefore, pigs with same parents had the same breeding values. On the other hand, since genomic breeding value was estimated using individual genomic information, accurate breeding value with Mendelian sampling could be calculated. Therefore, genomic breeding value seems to be more efficient than breeding value estimated using BLUP for selecting candidate pigs.

Relationship between selection ratio and phenotype

From late 2013, genomic information was collected from 63 candidate pigs. They began to farrow from the year of 2014. Records for total number of piglets born were collected. Of 348 sows farrowed in 2014, genomic information was obtained from 311 sows. The remaining 37 sows had no genomic information. Of the 311 sows with genomic information, 156 were candidate sows with first parity. By 2015, the overall average for total number of piglets born in GGP farms was 11.9. The distributional properties of validation data by year are shown in Table 5.

To determine the relationship between selection ratio and phenotype, 163 candidate pigs were divided into 5 groups by selection ratio (from top 10% to middle 50% candidate pigs). The mean values of total number of piglets born were compared (Table 6).

When candidate pigs for the next generation were selected by breeding values estimated from the conventional BLUP method, the total numbers of piglets born were severely fluctuated in pigs from Top 10% to 50%. In pigs of top 20% to 30%, the total number of piglets born was lower than the population average, suggesting that the accuracy of the estimated breeding value was very low. On the other hand, when they were selected by GEBV of ssBLUP using genomic information, it was always higher than the population average (Figure 2). These results

Table 6. Means and standard deviations for the total number of piglets born (TNB) by selection ratio and the estimation method

Top %	No. of sows	Total number of piglets born (TNB)		
		No. of records	BLUP	ssBLUP
10	16	42	11.90±4.24	12.50±3.71
20	16	41	12.54±3.33	12.30±3.93
30	16	37	11.70±3.45	12.07±3.48
40	17	47	12.87±2.97	12.50±3.42
50	16	40	12.05±3.03	12.20±3.64
Overall	163	386	12.12±3.89	12.12±3.89

BLUP, best linear unbiased prediction; ssBLUP, single step best linear unbiased prediction.

proved that the accuracy was higher when genomic information was used for selection than that when pedigree information was used for selection.

Genetic gain

To calculate the genetic gain using the formula of $(\Delta G) = r_{GP} \times i \times \sigma_a / L$, selection of sows from the tested sows for the total number of piglets born was based on pedigree information. Of 5,000 tested sows, 550 were selected (recently replacement rate of sows is 100%). Of 500 tested boars, 45 were selected. Therefore, the selection intensity (*i*) was 1.76. Genetic standard deviation (σ_a) for the total number of piglets born was 0.91. Since there was no data on the total number of piglets born for candidate pigs, the accuracy of the breeding value estimated from pedigree information was 0.080. When genomic information was used, the accuracy was 0.216 (Table 7). Assuming the replacement rate of sows per year at 100% and generation interval of 1 year, genetic gain per year would be 0.346 heads when genomic information was used and 0.128 when BLUP was used. Therefore, genetic gain estimated using ssBLUP method was 2.7 times higher than that estimated using BLUP method, i.e., 270% more efficient in improvement efficiency.

CONCLUSION

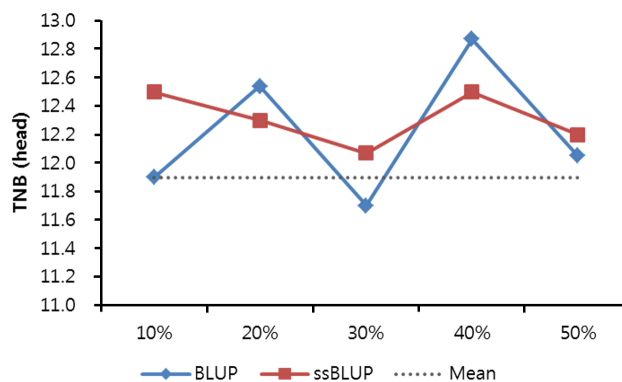


Figure 2. Trends on the total number of piglets born (TNB) by selection ratio and the estimation method.

Table 7. Genetic gains of the total number of piglets born (TNB) per year by the selection method of candidate sows

Method	σ_a	i	r_{GP}	L	ΔG
BLUP	0.91	1.76	0.080	1	0.128
ssBLUP	0.91	1.76	0.216	1	0.346

BLUP, best linear unbiased prediction; ssBLUP single step best linear unbiased prediction.

σ_a , genetic standard deviation; i , selection intensity, r_{GP} , accuracy; L , generation interval.

The selection accuracy was higher when sows were selected through the GEBV method using many markers than that when sows were selected through the conventional BLUP method using pedigree information. The selection accuracy was especially high for pigs with only genomic information (no phenotypic data). Accurate juvenile selection of them is possible using the GEBV method.

CONFLICT OF INTEREST

We certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

ACKNOWLEDGMENTS

This work was supported by a grant (PJ01111502) of the Next Generation BioGreen 21 Program funded by Rural Development Administration, Republic of Korea.

REFERENCES

1. VanRaden PM, Van Tassell CP, Wiggans GR, et al. Invited review: Reliability of genomic predictions for North American Holstein bulls. *J Dairy Sci* 2009;92:16-24.
2. Christensen OF, Madsen P, Nielsen B, Ostensen T, Su G. Single-step methods for genomic evaluation in pigs. *Animal* 2012;6:1565-71.
3. Meuwissen THE, Hayes BJ, Goddard ME. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 2001;157:1819-29.
4. Gengler, N, Abras S, Verkenne C, et al. Accuracy of prediction of gene content in large animal populations and its use for candidate gene detection and genetic evaluation. *J Dairy Sci* 2008;91:1652-59.
5. VanRaden PM. Efficient methods to compute genomic predictions. *J Dairy Sci* 2008;91:4414-23.
6. Misztal I, Legarra A, Aguilar I. A relationship matrix including full pedigree and genomic information. *J Dairy Sci* 2009;92:4656-63.
7. Liu M, Goddard ME, Reinhardt F, Reent R. A single-step genomic model with direct estimation of marker effects. *J Dairy Sci* 2014; 97:5833-50.
8. Christensen OF, Lund MS. Genomic prediction when some animals are not genotyped. *Genet Sel Evol* 2010;42:1.
9. Su G, Lund MS, Sorensen D. Selection for litter size at day five to improve litter size at weaning and piglet survival rate. *J Anim Sci* 2007;85:1385-92.
10. Chen P, Baas TJ, Mabry JW, Koehler KJ, Dekkers JC. Genetic parameters and trends for litter traits in U.S. Yorkshire, Duroc, Hampshire, and Landrace pigs. *J Anim Sci* 2003;81:46-53.
11. Lee JH, Song KD, Lee HK, et al. Genetic parameters of reproductive and meat quality traits in Korean Berkshire pigs. *Asian-Australas J Anim Sci* 2015;28:1388-93.
12. Arango J, Misztal I, Tsuruta S, Culbertson M, Herring W. Threshold-linear estimation of genetic parameters for farrowing mortality, litter size, and test performance of Large White sows. *J Anim Sci* 2005;83:499-506.
13. Forni S, Aguilar I, Misztal I. Different genomic relationship matrices for single-step analysis using phenotypic, pedigree and genomic information. *Genet Sel Evol* 2011;43:1.