Accuracy of Genomic Prediction for Meat Quality Traits Using Cow Reference Populations in Hanwoo Cattle

Mohammad Zahangir Alam [†], Shin Dae-Hyun [†], You-Sam Kim, Myung-Hum Park, Yun-Mi Lee [*], Jong-Joo Kim [*]

Posted Date: 19 May 2026

doi: 10.20944/preprints202605.1205.v1

Keywords: accuracy; carcass traits; genomic selection; Hanwoo; reference population

 doi:10.20944/preprints202605.1205.v1

Article

Accuracy of Genomic Prediction for Meat Quality Traits Using Cow Reference Populations in Hanwoo Cattle

Mohammad Zahangir Alam [1,2,†], Shin Dae-Hyun [1,†], You-Sam Kim [3], Myung-Hum Park [3], Yun-Mi Lee [4,*] and Jong-Joo Kim [1,*]

[1] Department of Biotechnology, Yeungnam University, Gyeongsan 38541, Republic of Korea
[2] Department of Genetic Engineering and Biotechnology, Shahjalal University of Science and Technology, Sylhet 3114, Bangladesh
[3] TNT Research Co. Ltd., Jeonju 54810, Republic of Korea
[4] Department of Companion Animal Health, Tongmyong University, Busan, Republic of Korea
* Correspondence: ymlee@tu.ac.kr (Y.M.L.); kimjj@ynu.ac.kr; Tel.: +82-53-810-3027 (J.-J.K.); Tel.: +82-10-2873-1733 (Y.-M.L.)
† These authors have contributed equally to this work.

Simple Summary

Hanwoo has been established as a main beef cattle breed in Korea since its official breeding started in 1979, shifting the focus from labor traits to meat production. In genomic selection, the reference population in estimating breeding value is very important. The accuracy of genetic breeding value based on a large number of cow-based reference population is limited. This is crucial for the Hanwoo industry because it allows the breeding program to dramatically expand the size and genetic diversity of the reference population by including females from many different farms. This study examines the potential of the cow reference population in breeding value estimation. The study concludes that while a steer-based reference population is optimal for maximum accuracy, a cow-based reference population is a practical and valuable alternative. This expansion is the key to improving the accuracy of genomic selection for complex, high-value meat quality traits that are expensive or difficult to measure, helping the industry adapt to changing consumer demands to grow. Incorporating a large number of cows into the reference population is a strategically important step toward building a more robust and powerful genomic selection system for the future.

Abstract

The establishment of a reference population for genomic selection in Korean beef cattle is an ongoing process. There is a high likelihood of sex-specific differences in the composition of the reference and test populations. This study evaluates the accuracy of Genomic Estimated Breeding Values (GEBVs) for carcass traits in Hanwoo cattle, specifically investigating the efficacy of cow-based reference populations. The effectiveness of genomic selection (GS) is heavily dependent on the composition and size of the reference population. Utilizing genotype data from a Hanwoo 50k SNP chip and phenotypic data from 19,168 steers and 6,233 cows, the study estimated GEBV accuracies for carcass weight (CWT), eye muscle area (EMA), backfat thickness (BF), and marbling score (MS) using the GBLUP method. Results demonstrate that steer-based reference populations achieved the highest accuracy (0.64–0.88), averaging 0.78, likely due to standardized management and higher trait heritability (0.39–0.51) compared to cows. In contrast, cow-based reference populations exhibited prediction accuracies (0.55–0.75) in four traits using adjusted residual phenotype, averaging 0.64, but remained highly practical alternatives. While growth traits (CWT and EMA) showed significant bias in cross-sex predictions, fat deposition traits (BF and MS) remained stable across sexes. The study concludes that although steer-based populations provide optimal accuracy, incorporating cows into the reference population is strategically vital in Hanwoo.

 doi:10.20944/preprints202605.1205.v1

Keywords: accuracy; carcass traits; genomic selection; Hanwoo; reference population

1. Introduction

Korean native cattle, Hanwoo (Bos taurus coreanae), were primarily used as agricultural livestock or a source of organic fertilizer in the past. However, Hanwoo became an increasingly important beef breed as agricultural technology progressed and beef consumption increased in Korean society. Now, Hanwoo is recognized as a high-quality beef breed both domestically and internationally for its exceptional meat quality and unique taste. This recognition stems from decades-long Hanwoo breeding programs aimed at producing high-quality meat, focusing on traits such as carcass weight, eye muscle area, backfat thickness, and marbling score. Currently, Hanwoo is a core industry in Korean agriculture, and research is continually conducted to enhance productivity and meat quality through ongoing breeding improvement. However, Korea's Hanwoo industry has also faced a new challenge - shifting consumer tastes. An increasing number of Korean consumers are worried about the high-fat content of Hanwoo beef and its health implications, including obesity and fat-induced diseases. This shift has precipitated a demand for leaner and healthier alternatives, such as Hanwoo heifer meat, which is known to be more expensive than steer or bull meat of the same quality but has seen a downgrade in demand [1]. This market shift emphasizes the importance for the Hanwoo breeding program to adopt a broader perspective that not only prioritizes traditional high-fat content marbling but also focuses on the strategic management of other meat quality and health traits in cows.

Genomic selection (GS) of livestock is gaining popularity due to the rapid progress of genotyping technology, where a reference population of sufficient size is the first requirement, in which all individuals are genotyped with an SNP chip and phenotyped for traits of interest [2]. GS based on genomic estimated breeding value (GEBV) accuracies depends on two criteria, relationships between animals in the reference population and the genetic relatedness between training (reference) and validation populations [3]. In addition, the composition of the reference population acts as a significant factor in the GEBV accuracy [4]. Korean proven bulls (KPNs) are progeny-tested and used for genomic selection in Korea. A core challenge for the Hanwoo breeding program is its unique population structure. However, the accuracy of genomic prediction requires many reference populations with both genotypes and phenotypes, which may not be available for all traits or breeds, and are also expensive to measure. Korean cattle are raised in various breeding environments and management systems, depending on whether they are heifers, cows, steers, or fattening cattle. There is a high possibility that differences might exist in the sex composition of the reference and test populations. These differences might act as potential factors that can reduce the efficiency of the genetic evaluation model. The current reliance on a few sires limits genetic diversity, creating a bottleneck that hinders the full potential of genomic selection. By including cows from diverse farms and lineages, the program can broaden its genetic base, which might further enhance the predictive power of the GS models.

In the GBLUP model, the realized accuracy, which is an indicator of the model's evaluation accuracy, largely depends on the genetic similarity between the reference and test populations. The reference group serves to provide genetic and phenotypic data for learning the prediction model, and the higher the genetic similarity of this group to the test group, the higher the accuracy of the evaluation. This kind of GS is more practiced and available for dairy cattle. However, Hanwoo is a beef cattle breed, and both sexes are used for meat production. The accuracy of GEBV increases while the genetic relationship (genetic covariance) between the reference and test populations is high [5]. In the previous studies, GEBV accuracies were estimated in Hanwoo steer populations from different regions of Korea [5–8]. Several reports have mentioned the possibility of using the cow as a reference population for dairy traits or any new functional trait [2,9,10]. In a previous study in the Netherlands [11], Holstein heifers from four countries were used to demonstrate that combining cow and bull reference populations increases the accuracy of genomic prediction and genome-wide association

 doi:10.20944/preprints202605.1205.v1

studies. In another study on broiler chickens, the accuracy of estimated breeding values increased in some traits when both sexes were included as the reference population [3]. Therefore, it is necessary to systematically analyse the effect of sex-specific composition of the reference and test populations on genetic evaluation. This study compares the genetic evaluation when cows and steers are used as the reference and test groups, respectively. Furthermore, the results showed significant evidence of the efficiency of the composition of the reference population. The strategic inclusion of this female population might represent the single most important action to expand the reference population and to maximize GEBV accuracy for high-value meat quality traits that are difficult to achieve with a male-centric reference population alone.

2. Materials and Methods

2.1. Animal Phenotypes

Phenotypic data on carcass traits were collected from 6233 Hanwoo cows and 19,168 steers across various regions of Korea. Phenotypic data of carcass weight (CWT), eye muscle area (EMA), backfat thickness (BF), and marbling score (MS) were collected from the one-stop Hanwoo Improvement Information Inquiry Service and the Livestock Products Quality Evaluation Institute. Carcass grading was recorded following the Korean carcass grading procedure established by the National Livestock Cooperatives Federation, as outlined in the Livestock Products Grading Guideline 2011. Briefly, CWT measurements were taken after 24h of post-mortem refrigeration; EMA was measured using a dot–grid technique on a cross-sectional slice between the 13th rib and the 1st lumbar vertebrae perpendicular to the vertebral column, where BF was also measured; MS was graded visually using a categorical system consisting of 9 levels ranging from 1 (no marbling) to 9 (high marbling). The birth years of cows between 2007 and 2022, and steers between 2014 and 2019, were recorded. The slaughter ages of cows ranged from 16 to 194 months, and steers were 24 to 35 months old. The animals were raised under standard feeding and management practices.

2.2. Genotyping and Quality Control

The Hanwoo steer and cow populations were genotyped from the hair root DNA using the Illumina Hanwoo 50k v.1 SNP BeadChip (Illumina Inc., San Diego, CA, USA). Among the total 53,866 SNPs, markers located on sex chromosomes and with unknown or duplicate positions were removed for further quality control (QC) procedures. Finally, 52,116 SNPs were used for the final analysis. Several QC thresholds were set to remove poor-quality SNPs for further analysis. SNPs were discarded from the analysis when the SNP call rate was less than 95%, individuals had a genotyping call rate less than 90%, and the minor allele frequency (MAF) was less than 5% (monomorphic). The genotype frequency significantly deviated ($p < 0.000001$) from the Hardy–Weinberg equilibrium (HWE), and the identity-by-state (IBS) test was performed to determine if there were similar individuals or genotyping errors in the datasets. Pairs of individuals showing a similarity rate greater than 99% were considered either identical animals or indicative of genotyping errors. The entire QC process and IBS test were performed through the PLINK v1.9 toolset [12]. Furthermore, the missing alleles were imputed using Beagle v5.4 software [13]. After conducting the IBS and QC tests, a total of 45,059 SNPs, 19,168 steers, and 6,233 cows remained in the dataset. Since the size of the reference and test groups significantly affects the accuracy of breeding value prediction in GBLUP analysis, steers were removed to ensure an exact match in size with the cow population. Finally, 19,168 steers were adjusted to 6,174, and an equal number of cows (6,174) were used in the analysis; and the distribution before and after adjustment was the same. Based on these individuals, phenotypic, genetic, and residual variances were estimated using the restricted maximum likelihood (REML) method following a mixed linear model based on the genomic relationship matrix (GRM).

 doi:10.20944/preprints202605.1205.v1

2.3. Statistical Analysis

In this study, genetic parameters and genomic estimated breeding values (GEBV) were calculated using the GBLUP method based on the genomic relationship matrix (GRM). The equation used to estimate GEBV was a linear model that can calculate both fixed and random effects, and the equation is as follows.

$$y = Xb + Zu + Z_{farm}f + e \quad (1)$$

Here, y, phenotype vector for each trait (dependent variable); X, Design matrix for fixed effects; b, estimate vector for fixed effects; Z, design matrix for random effects (additive effects); u, estimate vector for additive effects; Z farm, design matrix for random effects (farmer effect); f, estimate vector for farmer effect; e, environmental effect (residual) vector.

The dependent variable (y) was set as four traits in total: CWT, EMA, BF, and MS. The fixed effects were set as slaughter month group, birth year, and season group, and the random effects were set as individual effect and farm owner effect. The slaughter age was set as 24 to 35 months in the steer group, formed into a total of 12 groups (Supplementary Table S1). The birth year and season group were set as 1 for March to May, 2 for June to August, 3 for September to November, and 4 for December to February, and combined with the birth year, a total of 22 groups were set from fall 2014 to winter 2019 (Supplementary Table S2). On the other hand, the cow group was classified into 17 slaughter age groups from 13 to 194 months in total (Supplementary Table S3). The birth year and season groups of cows were set in the same way as steers, and were classified into 57 groups from spring 2007 to winter 2022 (Supplementary Table S4). The genetic breeding value was estimated by adjusting for the level of fixed and random effects for each total trait, and $\hat{b}$, $\hat{u}$, $\hat{f}$ were calculated using the following determinants.

$$\begin{bmatrix} \acute{X}X & \acute{X}Z & \acute{X}Z_f \\ \acute{Z}X & \acute{Z}Z + G^{-1}k_1 & \acute{Z}Z_f \\ Z_f'X & \acute{Z}_fZ & \acute{Z}_fZ_f + Ik_2 \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{u} \\ \hat{f} \end{bmatrix} = \begin{bmatrix} \acute{X}y \\ \acute{Z}y \\ Z_f\acute{y} \end{bmatrix} \quad (2)$$

In the above equation, G is the genetic relationship matrix (GRM), and k and I are computed as follows.

$$k_1 = \sigma_e^2/\sigma_a^2 \quad (3)$$

$$k_2 = \sigma_e^2/\sigma_f^2 \quad (4)$$

$$I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (5)$$

HIBLUP [14] software was used for genetic parameter estimation. The heritability for each trait was calculated using the estimated genetic variance and residual variance.

$$Heritability\ (h^2) = \frac{\sigma_a^2}{\sigma_a^2+\sigma_e^2} \quad (6)$$

$$\sigma_a^2: \text{Additive Genetic variance} \quad (7)$$

$$\sigma_e^2: \text{Residual variance (environmental variance)} \quad (8)$$

The genomic relationship matrix (GRM) between all pairs of individuals was calculated using the HIBLUP [14] software. The GRM was calculated based on the SNP marker information according to VanRaden [15], as follows.

$$G = \frac{ZZ^T}{tr\,(ZZ^T)/n} \quad (9)$$

 doi:10.20944/preprints202605.1205.v1

Here, the matrix (Z) is an n × m matrix with the number of **n** individuals and the number of SNPs as dimensions **m,** and *tr* represents the diagonal sum of the matrix. When calculating the genome relationship matrix (G), the additive elements are as follows.

$$G = \frac{zz'}{2\sum_{k=1}^{m} p_k q_k} \quad (10)$$

where the marker matrix (Z) has dimensions of n x m, n is the number of individuals, and m is the number of SNPs used, and each element is composed of (2-2p_k), (1-2p_k), and (-2p_k) while the genotypes are AA, Aa, and aa, respectively. In addition, p_k represents the minor allele frequency (MAF) of the kth SNP, and q_k represents the 1-p_k value.

2.4. Evaluation of Breeding Value Accuracy for Different Structures of Population

The breeding value was estimated using the GBLUP method with the genetic relationship matrix (GRM). The animals were divided into four analysis groups (A, B, C, and D) for the estimation of GEBV and accuracy. The analysis A (Ref_{steer}-$Test_{cow}$) was conducted with steers (n = 6,174) as the reference group, and cows (n = 6,174) as the test group. Analysis B (*Cow cross-validation*) involved a 5-fold cross-validation of cows (n = 6,174). In this study, k-fold cross-validation, a representative cross-validation technique described by Saatchi et al. 2011 [16], was used. Briefly, all populations were randomly divided into five (5) equal groups. For each cross-validation, four (4) out of the five groups were designated as the reference group, and the remaining group was treated as the test population. At this time, the size of each group is nearly identical, and it is crucial to divide the data to prevent duplication between the groups. At each iteration, a new group is designated as a test group, and a total of k evaluations are performed, so that all data participate in the verification process. The average of the test group accuracy of each iteration was used as the final accuracy verification [17]. Analysis C (Ref_{cow}-$Test_{steer}$) used cows (n = 6,174) as the reference group and steers (n = 6,174) as the test group. Analysis D (*Steer cross-validation*) consisted of 5-fold cross-validation of steers. The test groups of analyses A and B were cows, whereas the test groups of analyses C and D were steers. The genetic breeding values, in order, were carcass weight, eye muscle area, backfat thickness, and marbling score. One way to evaluate the genetic breeding value is to calculate the correlation between the actual breeding value of an animal and the estimated genetic breeding value. In addition, maximizing this is the main goal of breeding research [18]. In this study, since it is impossible to obtain the actual breeding value of an animal, the accuracy of the estimated genetic breeding value was evaluated using phenotypic data and the adjusted residual ($\hat{e}$). The adjustment formula is as follows.

$$\hat{e} = y - Xb - Z_{farm}f \quad (11)$$

Here, y : Phenotype vector for each trait (dependent variable); X, Z_{farm} : Design matrix for fixed effects and random effects; b , f : Vector estimates for fixed effects and farmer effects; and $\hat{e}$: Environmental effect (residual) vector.

The realized accuracy of the estimated genetic breeding value is calculated by dividing the Pearson correlation coefficient between the corrected phenotype and the genetic breeding value by the square root of the heritability [19]. The realized accuracy using the total phenotypic data was set to accuracy 1, and the realized accuracy using the adjusted residual ($\hat{e}$) data was set to an accuracy 2.

$$\text{Accuracy 1} = \frac{r(y,GEBV)}{\sqrt{h^2}} \quad (12)$$

$$\text{Accuracy 2} = \frac{r(\hat{e},GEBV)}{\sqrt{h^2}} \quad (13)$$

 doi:10.20944/preprints202605.1205.v1

3. Results

3.1. Summary Statistics of the Phenotypes

The summary statistics for the carcass traits (CWT, EMA, BF, and MS) of the studied population, which consists of 6,233 Hanwoo cows and 19,168 steers, are presented in Table 1. The average CWT, EMA, BF, and MS of the 19,168 steers were 445 kg, 96.7 cm², 14.2 mm, and 5.99, respectively, and those of 6,234 Hanwoo cows were 371 kg, 87.7 cm², 13 mm, and 4.13, respectively.

Table 1. Summary statistics for carcass traits in the Hanwoo population.

Groups	Traits	Mean	Min	Max	SD
Steer (19,168)	CWT (kg)	445	160	692	49.9
	EMA (cm²)	96.7	20.0	160	12.2
	BF (mm)	14.2	1.0	47.0	4.90
	MS (score 1~9)	5.99	1.0	9.00	1.86
Cow (6,233)	CWT (kg)	371	149	602	55.2
	EMA (cm²)	87.7	20.0	153	13.7
	BF (mm)	13.0	2.0	69.0	5.80
	MS (score 1~9)	4.13	1.0	9.0	1.98

SD, standard deviation; Min, minimum; Max, maximum; CWT, carcass weight; EMA, eye muscle area; BF, backfat thickness; MS, marbling score.

3.2. Estimation of Variance Components and Heritability

In the studied population, which contains a total of 12,348 heads, including 6,174 steers and 6,174 cows. The phenotypic variance (σ^2_P), genetic variance (σ^2_a), residual variance (σ^2_e), and heritability (h^2) were estimated (Table 2) using the populations. Among the three groups, the group with the largest phenotypic variance (σ^2_P) for all traits consisted of only cows. The heritability (h^2) of cows was 0.28, 0.22, 0.32, and 0.36 for CW, EMA, BF, and MS traits, respectively. In comparison, it showed the lowest heritability for all traits. The population with the lowest phenotypic variance (σ^2_P) consisted of steers, and conversely, it showed the highest heritability. The heritability of the four traits of steers, CW, EMA, BF, and MS, was 0.46, 0.39, 0.39, and 0.51, respectively, and the trait with the lowest heritability was EMA, while the marbling score has the highest heritability. The population with an intermediate size of phenotypic variance (σ^2_P) and heritability (h^2) was a mixture of steers and cows, with h^2 of CW, EMA, BF, and MS being 0.33, 0.28, 0.35, and 0.41, respectively, and the trait with the lowest h^2 was EMA, and the highest h^2 was MS, as observed in the other two populations.

Table 2. Estimates of genetic parameters of the four carcass traits in Hanwoo.

Population (n)	Traits	σ^2_p	σ^2_a	σ^2_a (SD)	σ^2_e	σ^2_e(SD)	h^2
Steer + cow (12,348)	CWT	2182.34	718.97	39.24	1463.37	27.98	0.33
	EMA	144.48	40.92	2.50	103.56	1.93	0.28
	BF	26.92	9.45	0.51	17.47	0.34	0.35
	MS	3.19	1.31	0.06	1.88	0.04	0.41
Steer (6,174).	CWT	2093.86	968.61	63.11	1125.25	41.92	0.46
	EMA	131.52	51.31	3.78	80.22	2.76	0.39
	BF	23.17	9.13	0.68	14.04	0.49	0.39
	MS	3.06	1.57	0.10	1.50	0.06	0.51
Cow (6,174)	CWT	2277.19	641.28	57.14	1635.91	45.82	0.28
	EMA	154.25	34.40	3.59	119.85	3.20	0.22
	BF	30.19	9.70	0.82	20.49	0.61	0.32
	MS	3.25	1.17	0.09	2.09	0.06	0.36

 doi:10.20944/preprints202605.1205.v1

$\sigma_p^2, Phenotypic\ variance$; $\sigma_a^2, Additive\ genetic\ variance$; $\sigma_e^2, Residual\ variance$; $h^2, Heritability$; SD , standard deviation; CWT, carcass weight; EMA, eye muscle area; BF, backfat thickness; MS, marbling score.

3.3. Estimation of genetic breeding value

The average genetic breeding values of CWT, EMA, BF and MS were 2.81, 0.50, -0.01, and 0.09, respectively, for analysis A and 0.21, 0.03, 0.01, and 0.00, respectively, for analysis B. Additionally, the GEBVs of CWT, EMA, BF, and MS were –2.01, -0.35, 0.00, and -0.05, respectively, for analysis C and 0.11, 0.01, 0.00, and 0.00 for analysis D. The results (Table 3) reveal that the reliability of GEBVs is highly sensitive to sex-specific differences in growth traits compared to carcass quality traits. When applying a reference population of one sex to a testing population of the other (Analyses A and C), Carcass Weight and Eye Muscle Area exhibit significant systematic bias, with CWT showing an overestimation of 2.81 in cows and an underestimation of -2.01 in steers. Conversely, BF and MS remain consistently stable with average GEBVs near 0.00, indicating that the genetic markers for fat deposition are expressed similarly across both sexes.

Table 3. Genomic estimated breeding values of four carcass traits in Hanwoo using different sex combinations.

Analysis	Reference Population (6174)	Testing Population (6174)	Trait	GEBV			
				(Min)	(Max)	(Average)	(SD)
A	Steer	Cow	CWT	-71.68	95.26	2.81	24.27
			EMA	-17.51	18.29	0.50	5.21
			BF	-5.99	8.01	-0.01	1.77
			MS	-2.86	3.05	0.09	0.87
B	Cow cross-validation		CWT	-47.72	70.08	0.21	16.60
			EMA	-11.93	13.44	0.03	3.39
			BF	-6.18	6.49	0.01	1.62
			MS	-2.81	2.75	0.00	0.71
C	Cow	Steer	CWT	-51.57	71.91	-2.01	16.60
			EMA	-12.44	12.47	-0.35	3.64
			BF	-6.72	6.41	0.00	1.71
			MS	-2.22	2.37	-0.05	0.63
D	Steer cross-validation		CWT	-75.64	86.53	0.11	23.25
			EMA	-17.10	17.71	0.01	5.04
			BF	-6.73	8.01	0.00	1.80
			MS	-2.54	2.96	0.00	0.78

3.4. Evaluation of Breeding Value Accuracy for Different Structures of Population

3.4.1. Accuracy of Cows

Two metrics were used to evaluate prediction accuracy: Accuracy 1, defined as the Pearson correlation between the phenotype and genetic breeding value divided by the heritability (h^2), and Accuracy 2, which utilized the adjusted residual phenotype ($\hat{e}$) in the same calculation. In Analysis A, Accuracy 1 values for CWT, EMA, BF, and MS were 0.58, 0.51, 0.43, and 0.62, respectively, while Analysis B yielded slightly higher results of 0.59, 0.52, 0.45, and 0.62 (Table 4). Analysis B generally outperformed Analysis A across most traits during cow cross-validation. For Accuracy 2, Analysis A recorded values of 0.65, 0.57, 0.48, and 0.63, compared to Analysis B's 0.61, 0.53, 0.48, and 0.62. Although both analyses demonstrated higher accuracy when using adjusted residuals, the improvement was notably more significant in Analysis A.

 doi:10.20944/preprints202605.1205.v1

Table 4. The accuracy of GEBV of the four carcass traits using phenotypes and adjusted residuals in different combinations of cows and steers as reference and testing population.

	Analysis	Reference Population (6174)	Testing Population (6174)	Parameters	CWT	EMA	BF	MS	*Average accuracy*
				h^2	**0.33**	**0.28**	**0.35**	**0.41**	
Total Phenotype	A	Steer	Cow	r (y, GEBV)	0.33	0.27	0.26	0.40	***0.54***
				Accuracy 1	0.58	0.51	0.43	0.62	
	B	Cow cross-validation		r (y, GEBV)	0.34	0.28	0.27	0.40	***0.55***
				Accuracy 1	0.59	0.52	0.45	0.62	
	C	Cow	Steer	r (y, GEBV)	0.38	0.32	0.30	0.35	***0.58***
				Accuracy 1	0.67	0.59	0.51	0.54	
	D	Steer cross-validation		r (y, GEBV)	0.44	0.41	0.35	0.41	***0.69***
				Accuracy 1	0.77	0.77	0.60	0.63	
Adjusted Residuals	A	Steer	Cow	r (ê, GEBV)	0.37	0.30	0.28	0.40	***0.58***
				Accuracy 2	0.65	0.57	0.48	0.63	
	B	Cow cross-validation		r (ê, GEBV)	0.35	0.28	0.28	0.40	***0.56***
				Accuracy 2	0.61	0.53	0.48	0.62	
	C	Cow	Steer	r (ê, GEBV)	0.43	0.35	0.33	0.39	***0.64***
				Accuracy 2	0.75	0.66	0.55	0.61	
	D	Steer cross-validation		r (ê, GEBV)	0.50	0.46	0.38	0.46	***0.78***
				Accuracy 2	0.88	0.86	0.64	0.72	

Accuracy 1 is the Pearson correlation coefficient between the observed total phenotype and predicted genetic breeding value, divided by the square root of the heritability; Accuracy 2 uses the adjusted residual phenotype.

3.4.2. Accuracy of Steers

In Analysis C, which utilized cows as the reference population, Accuracy 1 values for CWT, EMA, BF, and MS were 0.67, 0.59, 0.51, and 0.54, respectively. These metrics improved significantly in Analysis D (the steer cross-validation group) to 0.77, 0.77, 0.60, and 0.63 (Table 4). A similar trend was observed for Accuracy 2, which rose from 0.75, 0.66, 0.55, and 0.61 in Analysis C to 0.88, 0.86, 0.64, and 0.72 in Analysis D (Figure 1). Across all traits and evaluation metrics, Analysis D consistently demonstrated superior performance, indicating that employing steers as a reference population yields higher genomic prediction accuracy than using cows.

 doi:10.20944/preprints202605.1205.v1

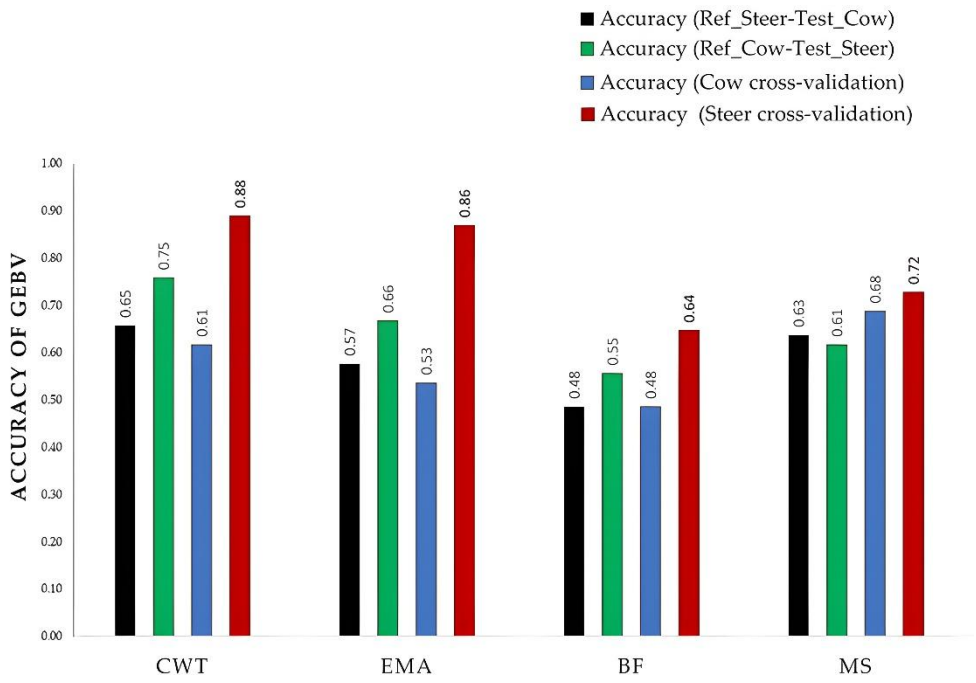

Figure 1. Accuracy scenario of GEBV of the carcass traits using adjusted residual phenotype.

4. Discussion

The primary objective of this study was to evaluate the accuracy of Genomic Estimated Breeding Values (GEBVs) for carcass traits in Hanwoo cattle, specifically focusing on the performance of a cow-based reference population. The precision of genomic selection is multifaceted, governed by trait heritability [18], reference population size [20], marker density [21], prediction methodologies [22], and the genomic relationship between the reference and test populations [23]. To isolate the impact of the reference population's sex composition, we standardized these external variables across our analyses. Our descriptive statistics for carcass weight (CWT), eye muscle area (EMA), backfat thickness (BF), and marbling score (MS) largely aligned with the Livestock Product Grading Statistical Yearbook 2023, which reported averages of 431.7 kg, 91.1 cm², 13.2 mm, and 5.7, respectively. Our phenotypic data also showed strong consistency with previous large-scale studies. For instance, Lee et al. [5] and Lopez [6] reported CWT values ranging from 441.2 to 442.03 kg, EMA 95.82 to 96.04 cm², BF values from 14.2 to 14.23 mm and MS showed exactly similar value of 5.95 while Lee [24] observed 442.33 kg, 96.53 cm², 14.26 mm, and 5.96 for CWT, EMA, BF, and MS, respectively across a national sample of 18,499 heads. The close alignment of our basic statistics with these contemporary studies validates the representativeness of our steer-based reference data. Previous literature indicates that higher heritability and larger reference populations are the primary drivers of GEBV accuracy [25]. In Hanwoo, carcass trait heritability typically ranges from 0.26 to 0.61, categorized as moderate to high [26–30]. In our study, heritability estimates for all four carcass traits were higher in steers than in cows. This discrepancy likely stems from the differing production lifecycles; steers were slaughtered at approximately 30 months, whereas cows were slaughtered at 50 months. In Japanese black cattle, using a reference population of more than 5000 animals with a heritability of 0.1~ 0.5 could be practical for the expected accuracy of GEBV for a polygenic trait [31]. Another review [32] examined the studies over 42 years of publications in scientific literature, suggesting that heritability estimates for most carcass traits varied greatly, which could be due to differences in breed composition, methods of estimation, effects in the model, number of observations, measurement errors, sex, and management. The extended lifespan of cows (an additional 20 months) increases exposure to diverse environmental variables, likely inflating the environmental variance relative to genetic variance. Consequently, the lower heritability observed in cows suggests that environmental noise may mask genetic potential more significantly than in steers.

 doi:10.20944/preprints202605.1205.v1

Our study addressed this by utilizing the average heritability of the combined sex population to predict GEBV accuracy. Furthermore, by utilizing a 50K SNP chip, we adhered to the marker density recommendations established for commercial Hanwoo populations [6].

We utilized the GBLUP method to estimate breeding values, comparing cross-validation within sexes against cross-sex predictions (i.e., using cows to predict steers and vice versa). Realized accuracy was determined by the Pearson correlation between the genetic breeding values and adjusted residual phenotypes, normalized by the square root of the heritability. The superior performance of steer-based models is likely due to the more uniform environmental conditions and standardized slaughter ages associated with steer production. Furthermore, because the genomic relationship between the steers and cows in this study was relatively low, the influence of pedigree relatedness on breeding value accuracy was likely minimal, highlighting the importance of reference population composition over simple kinship. Bedhane et al. [27] reported accuracies between 0.29 and 0.46 for meat quality in Hanwoo, while Nellore cattle studies have shown lower accuracies ranging from 0.21 to 0.46 [33]. Notably, Lee et al. [5] achieved GBLUP accuracies of 0.62–0.66 for carcass traits using a massive steer reference population to predict a small cow test group, a pattern mirrored in our findings. In light of their close genetic proximity to Hanwoo, Korean Brindle cattle achieved genomic estimated breeding value (GEBV) accuracies of 0.44 for carcass weight (CWT), 0.43 for eye muscle area (EMA), 0.42 for backfat thickness (BF), and 0.44 for marbling score (MS) when utilizing a Hanwoo steer reference population [34]. While these results underscore a significant hereditary link, the observed reduction in accuracy compared to intra-breed predictions is likely attributable to inherent breed-specific divergence and the limited size of the test population. Consequently, these findings suggest that while cross-population reference sets offer a viable starting point, refining GEBV precision for Brindle cattle will necessitate larger, breed-specific datasets to account for subtle genomic variations. While some studies, such as Kim [35], reported exceptionally high accuracies (0.78–0.81), these are often attributed to significantly smaller test populations (e.g., n=46). In contrast, Lee G.H. et al. [24] reported more moderate realized accuracies (0.40–0.54) using a large, mixed-sex population of 18,499 heads. Our findings confirm that while a cow-based reference population provides acceptable GEBV accuracy, the inherent environmental variance in older female cattle remains a challenge. Future research should focus on refining models to better account for the environmental factors unique to cow populations to bridge the accuracy gap between the sexes.

5. Conclusions

In conclusion, this study evaluated the accuracy of genomic prediction using the GBLUP method across various sex-specific reference population configurations. Our findings reveal that utilizing cows as a reference population resulted in an approximate 10% increase in average GEBV accuracy for carcass traits in steers, with the notable exceptions of Backfat Thickness (BF) and Marbling Score (MS). This discrepancy likely stems from the physiological divergence in marbling characteristics between cows and steers. Ultimately, these results underscore that the genetic connectivity between reference and test populations, regardless of sex, is a primary driver of prediction accuracy. To fully establish Hanwoo cows as a viable reference for commercial herd evaluation, future efforts must prioritize the continuous expansion of reference datasets and the rigorous validation of genomic models.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org, Table S1: Classification of slaughter age in the Hanwoo steer population; Table S2: Classification of birth and season in Hanwoo steer population; Table S3: Classification of slaughter age in the Hanwoo cow population; Table S4: Classification of birth year season in Hanwoo cow population.

Author Contributions: Conceptualization, M.Z.A, S.D.H, and J.-J.K; methodology, M.Z.A., S.D.H., and J.-J.K; software, S.D.H., Y.-S.K, M.H.P, and Y.-M.L.; validation, S.D.H., Y.-M.L., J.-J.K.; formal analysis, S.D.H., Y.-M.L.; investigation, S.D.H., Y.-M.L.; data curation, M.Z.A., S.D.H., Y.-M.L.; writing - original draft, M.Z.A.; writing -

 doi:10.20944/preprints202605.1205.v1

review & editing, M.Z.A., S.D.H., Y.-M.L., and J.-J.K.; supervision, J.-J.K.; project administration, Y.-M.L.; funding acquisition, J.-J.K., Y.-M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through the Smart Farm Innovation Technology Development Program funded by the Ministry of Agriculture, Food and Rural Affairs (RS-2025-02307074).

Institutional Review Board Statement: No ethical approval was required to collect animal data, as it was a routine process of a beef cattle slaughterhouse. However, the care and slaughter of all animals used in this study were approved by the Animal Care and Use Committee of the National Institute of Animal Science (NIAS), Rural Development Administration (RDA), South Korea. All animal health and welfare practices followed the approved guidelines. The genotype and phenotype data were collected by the Hanwoo Improvement Center (HIC) of the National Agricultural Cooperative Federation.

Informed Consent Statement: Not applicable.

Data Availability Statement: None of the data was deposited in an official repository due to the privacy of the funders' intellectual property. However, the dataset is available from the corresponding author upon reasonable request.

Acknowledgments: You-Sam Kim and Myung-Hum Park provided SNP chip data in this research work. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

Conflicts of Interest: The authors declare no conflicts of interest, and the funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

IBS	Identity By State
GBLUP	Genomic Best Linear Unbiased Prediction
REML	Restricted Maximum Likelihood
GRM	Genetic Relationship Matrix
GEBV	Genomic Estimated Breeding Value
CWT	Carcass Weight
EMA	Eye Muscle Area
BF	Back Fat Thickness
MS	Marbling Score
GS	Genomic Selection
KPN	Korean Proven Number/Korean Proven Bull's No.

References

1. Hossain, M.A.; Aung, S.H.; Park, J.Y.; Kim, S.H.; Lee, S.S.; Nam, K.C. Effects of Gender and Slaughter Age on Physicochemical and Quality Traits of Korean Hanwoo Striploin. *Journal of Animal Science and Technology* **2024**, *66*, 614–629, doi:10.5187/jast.2023.e127.
2. Ding, X.; Zhang, Z.; Li, X.; Wang, S.; Wu, X.; Sun, D.; Yu, Y.; Liu, J.; Wang, Y.; Zhang, Y.; et al. Accuracy of Genomic Prediction for Milk Production Traits in the Chinese Holstein Population Using a Reference Population Consisting of Cows. *Journal of Dairy Science* **2013**, *96*, 5315–5323, doi:10.3168/jds.2012-6194.
3. Lourenco, D.A.L.; Fragomeni, B.O.; Tsuruta, S.; Aguilar, I.; Zumbach, B.; Hawken, R.J.; Legarra, A.; Misztal, I. Accuracy of Estimated Breeding Values with Genomic Information on Males, Females, or Both: An Example on Broiler Chicken. *Genetics Selection Evolution* **2015**, *47*, 1–9, doi:10.1186/s12711-015-0137-1.
4. van den Berg, I.; Meuwissen, T.H.E.; MacLeod, I.M.; Goddard, M.E. Predicting the Effect of Reference Population on the Accuracy of within, across, and Multibreed Genomic Prediction. *Journal of Dairy Science* **2019**, *102*, 3155–3174, doi:10.3168/jds.2018-15231.

 doi:10.20944/preprints202605.1205.v1

5. Lee, H.S.; Kim, Y.; Lee, D.H.; Seo, D.; Lee, D.J.; Do, C.H.; Dinh, P.T.N.; Ekanayake, W.; Lee, K.H.; Yoon, D.; et al. Comparison of Accuracy of Breeding Value for Cow from Three Methods in Hanwoo (Korean Cattle) Population. *Journal of Animal Science and Technology* **2023**, *65*, 720–734, doi:10.5187/JAST.2023.E5.
6. Lopez, B.I.; Lee, S.H.; Shin, D.H.; Oh, J.D.; Chai, H.H.; Park, W.; Park, J.E.; Lim, D. Accuracy of Genomic Evaluation Using Imputed High-Density Genotypes for Carcass Traits in Commercial Hanwoo Population. *Livestock Science* **2020**, *241*, doi:10.1016/j.livsci.2020.104256.
7. Kim, E.H.; Kang, H.C.; Myung, C.H.; Kim, J.Y.; Sun, D.W.; Lee, D.H.; Lee, S.H.; Lim, H.T. Comparison on Genomic Prediction Using Pedigree BLUP and Single Step GBLUP through the Hanwoo Full-Sib Family. *Animal Bioscience* **2023**, *36*, 1327–1335, doi:10.5713/ab.22.0327.
8. Park, M.N.; Alam, M.; Kim, S.; Park, B.; Lee, S.H.; Lee, S.S. Genomic Selection through Single-Step Genomic Best Linear Unbiased Prediction Improves the Accuracy of Evaluation in Hanwoo Cattle. *Asian-Australasian Journal of Animal Sciences* **2020**, *33*, 1544–1557, doi:10.5713/ajas.18.0936.
9. Buch, L.H.; Kargo, M.; Berg, P.; Lassen, J.; Sørensen, A.C. The Value of Cows in Reference Populations for Genomic Selection of New Functional Traits. *Animal* **2012**, *6*, 880–886, doi:10.1017/S1751731111002205.
10. Nayee, N.; Gajjar, S.; Sudhakar, A.; Saha, S.; Trivedi, K.; Vataliya, P. Genomic Selection in Gir Cattle Using Female Reference Population. *Indian Journal of Animal Sciences* **2021**, *90*, 1612–1616, doi:10.56093/ijans.v90i12.113193.
11. Calus, M.P.L.; de Haas, Y.; Veerkamp, R.F. Combining Cow and Bull Reference Populations to Increase Accuracy of Genomic Prediction and Genome-Wide Association Studies. *Journal of Dairy Science* **2013**, *96*, 6703–6715, doi:10.3168/jds.2012-6013.
12. 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. *American journal of human genetics* **2007**, *81*, 559–575, doi:10.1086/519795.
13. Browning, B.L.; Zhou, Y.; Browning, S.R. A One-Penny Imputed Genome from Next-Generation Reference Panels. *The American Journal of Human Genetics* **2018**, *103*, 338–348, doi:10.1016/J.AJHG.2018.07.015.
14. Yin, L.; Zhang, H.; Tang, Z.; Yin, D.; Fu, Y.; Yuan, X.; Li, X.; Liu, X.; Zhao, S. HIBLUP: An Integration of Statistical Models on the BLUP Framework for Efficient Genetic Evaluation Using Big Genomic Data. *Nucleic acids research* **2023**, *51*, 3501–3512, doi:10.1093/nar/gkad074.
15. VanRaden, P.M. Efficient Methods to Compute Genomic Predictions. *Journal of Dairy Science* **2008**, *91*, 4414–4423, doi:10.3168/jds.2007-0980.
16. Saatchi, M.; McClure, M.C.; McKay, S.D.; Rolf, M.M.; Kim, J.; Decker, J.E.; Taxis, T.M.; Chapple, R.H.; Ramey, H.R.; Northcutt, S.L.; et al. Accuracies of Genomic Breeding Values in American Angus Beef Cattle Using K-Means Clustering for Cross-Validation. *Genetics Selection Evolution* **2011**, *43*, 40, doi:10.1186/1297-9686-43-40.
17. Berrar, D. Cross-Validation. In *Reference Module in Life Sciences*; Elsevier, 2024.
18. Goddard, M. Genomic Selection: Prediction of Accuracy and Maximisation of Long Term Response. *Genetica* **2009**, *136*, 245–257, doi:10.1007/s10709-008-9308-0.
19. Hayes, B.J.; Bowman, P.J.; Chamberlain, A.J.; Goddard, M.E. Invited Review: Genomic Selection in Dairy Cattle: Progress and Challenges. *Journal of Dairy Science* **2009**, *92*, 433–443, doi:10.3168/jds.2008-1646.
20. Daetwyler, H.D.; Villanueva, B.; Woolliams, J.A. Accuracy of Predicting the Genetic Risk of Disease Using a Genome-Wide Approach. *PLoS ONE* **2008**, *3*, doi:10.1371/journal.pone.0003395.
21. Clark, S.A.; Hickey, J.M.; Van Der Werf, J.H.J. Different Models of Genetic Variation and Their Effect on Genomic Evaluation. *Genetics Selection Evolution* **2011**, *43*, 1–9, doi:10.1186/1297-9686-43-18.
22. Zhou, L.; Mrode, R.; Zhang, S.; Zhang, Q.; Li, B.; Liu, J.F. Factors Affecting GEBV Accuracy with Single-Step Bayesian Models. *Heredity* **2018**, *120*, 100–109, doi:10.1038/s41437-017-0010-9.
23. Kim, B.; Ekanayake, W.; Shin, J.; Kim, Y.; Lee, D.; Kim, Y.; Lee, S.H. Effect of Population Structure on the Accuracy of Genomic Prediction for Carcass Traits in Hanwoo Cattle: A Simulation Study. *Journal of Animal Breeding and Genomics* **2023**, *7*, 157–166, doi:10.12972/jabng.20230017.
24. Lee, G.H.; Lee, Y.H.; Kong, H.S. Genetic Evaluation and Accuracy Analysis of Commercial Hanwoo Population Using Genomic Data. *Journal of Animal Reproduction and Biotechnology* **2023**, *38*, 32–37, doi:10.12750/jarb.38.1.32.

 doi:10.20944/preprints202605.1205.v1

25. Karaman, E.; Su, G.; Croue, I.; Lund, M.S. Genomic Prediction Using a Reference Population of Multiple Pure Breeds and Admixed Individuals. *Genetics Selection Evolution* **2021**, *53*, doi:10.1186/s12711-021-00637-y.
26. Oh, J.D.; Lee, G.H.; Kong, H.S. Estimation of Heritability and Genetic Parameters for Carcass Traits and Primal Cut Production Traits in Hanwoo. *Journal of Animal Reproduction and Biotechnology* **2024**, *39*, 114–120, doi:10.12750/jarb.39.2.114.
27. Bedhane, M.; van der Werf, J.; de las Heras-Saldana, S.; Ackerson, L.; Lim, D.; Park, B.; Park, M.N.; Roh, S.; Clark, S. Parameter Estimation and Assessment of Bias in Genetic Evaluation of Carcass Traits in Hanwoo Cattle Using Real and Simulated Data. *Journal of Animal Science and Technology* **2023**, *65*, 1180–1193, doi:10.5187/JAST.2023.E36.
28. Alam, M.; Lee, S.H.; Lee, D.H.; Cho, C.; Park, M.N. Genetic Analysis of Major Carcass Traits of Korean Hanwoo Males Raised for Thirty Months. *Animals* **2021**, *11*, doi:10.3390/ani11061792.
29. Naserkheil, M.; Lee, D.H.; Mehrban, H. Improving the Accuracy of Genomic Evaluation for Linear Body Measurement Traits Using Single-Step Genomic Best Linear Unbiased Prediction in Hanwoo Beef Cattle. *BMC Genetics* **2020**, *21*, doi:10.1186/s12863-020-00928-1.
30. Mehrban, H.; Lee, D.H.; Naserkheil, M.; Moradi, M.H.; Ibáñez-Escriche, N. Comparison of Conventional BLUP and Singlestep Genomic BLUP Evaluations for Yearling Weight and Carcass Traits in Hanwoo Beef Cattle Using Single Trait and Multi-Trait Models. *PLoS ONE* **2019**, *14*, doi:10.1371/journal.pone.0223352.
31. Takeda, M.; Inoue, K.; Oyama, H.; Uchiyama, K.; Yoshinari, K.; Sasago, N.; Kojima, T.; Kashima, M.; Suzuki, H.; Kamata, T.; et al. Exploring the Size of Reference Population for Expected Accuracy of Genomic Prediction Using Simulated and Real Data in Japanese Black Cattle. *BMC Genomics* **2021**, *22*, doi:10.1186/s12864-021-08121-z.
32. Utrera, A.R.; Dale, L.; Vleck, V. *Heritability Estimates for Carcass Traits of Cattle: A Review*; 2004;
33. Fernandes Júnior, G.A.; Rosa, G.J.M.; Valente, B.D.; Carvalheiro, R.; Baldi, F.; Garcia, D.A.; Gordo, D.G.M.; Espigolan, R.; Takada, L.; Tonussi, R.L.; et al. Genomic Prediction of Breeding Values for Carcass Traits in Nellore Cattle. *Genetics Selection Evolution* **2016**, *48*, doi:10.1186/s12711-016-0188-y.
34. Byun, S.K.; Kim, D.H.; Oh, J.D.; Lee, H.K. The Analysis on Evaluation of Genomic Breeding Value in Brindle Cattle Using Reference Population of Hanwoo. *Journal of Animal Breeding and Genomics* **2021**, *5*, doi:10.12972/jabng.20210009.
35. Kim, E.H.; Sun, D.W.; Kang, H.C.; Kim, J.Y.; Myung, C.H.; Lee, D.H.; Lee, S.H.; Lim, H.T. A Study of the Genomic Estimated Breeding Value and Accuracy Using Genotypes in Hanwoo Steer (Korean Cattle). *Korean Journal of Agricultural Science* **2021**, *48*, 681–691, doi:10.7744/kjoas.20210057.