

Research Article

ssGBLUP을 이용한 한우 도체형질에 대한 유전체육종가 및 신뢰도 추정

김성진^{1†}, 최태정^{2†}, 손지현¹, 이득민³, 이종재⁴, 이정규^{5,6}, 임현태^{5,6}, 구양모^{1*}¹한국종축개량협회 유전정보사업부, ²국립축산과학원 가축개량평가과, ³한경대학교 동물생명환경과학과, ⁴중앙대학교 동물생명공학과, ⁵경상국립대학교 축산학과,⁶경상국립대학교 농업생명과학연구원

Estimation of Genomic Estimated Breeding Value(GEBV) and Reliability for Hanwoo Carcass Traits using ssGBLUP

Sung-Jin Kim^{1†}, Tae-Jeong Choi^{2†}, Ji-Hyun Son¹, Deuk-Min Lee³, Jung-Jae Lee⁴, Jung-Gyu Lee^{5,6}, Hyun-Tae Lim^{5,6}, Yang-Mo Koo^{1*}¹Genetic & IT Solutions Dept. Korea Animal Improvement Association, Seocho, 06668, Korea²Animal Genetics & Breeding Division, NIAS, Cheonan, 31000, Korea³Department of Animal Life & Environment Science, Hankyong National University, Anseong, 17579, Korea⁴Department of Animal Science and Technology, Chung-Ang University, Anseong, 06974, Korea⁵Department of animal Science, GNU, JinJu, 52828, Korea⁶Institute of Agriculture & Life Science, GNU, JinJu, 52828, Korea

*Corresponding author: Yang-Mo Koo, Genetic & IT Solutions Dept. Korea Animal Improvement Association, Seocho, 06668, Korea, Tel : +82-2-588-9301, E-mail : greatman009@gmail.com

ABSTRACT

For this study, the DNA of 8,413 animals collected by various methods from Hanwoo farms nationwide was extracted. The Axiom Bovine 60k version 3(Affymetrix Inc., 2006) SNP panel was used to generate genomic information. Quality Control to increase the accuracy of genomic analysis removed SNPs whose SNPs were on the sex chromosome or whose position on the chromosome was not identified. A total of 64,973 SNP markers were used to remove markers with an SNP call rate of 95% or less, a Minor Allele Frequency(MAF) of 0.01 or less, and a Chi-square value of more than 95% measured for the Hardy-Weinberg Equilibrium(HWE). Among the 8,413 animals with genomic information, duplicate animals, pedigree errors, and animals that did not match dependent variables were removed, and 6,616 references with genomic information were used for analysis. KPN used in the analysis was extracted using semen, and steer performed DNA extraction using tissue samples provided by the Korea Institute for Animal Products Quality Evaluation and the cow collected tail hairs and used them for analysis. As for the pedigree data, five generation pedigree data connected to 6,616 animals were collected from the Korea Animal Improvement Association, and a total of 5,153,168 pedigree data were used for the analysis. The carcass traits data were used for analysis on the remaining 2,376,865 animals after pre-removal of missing and abnormal values among 5,153,168 animals with carcass traits data. Four traits used in the analysis were considered Carcass Weight(CW), Eye Muscle Area(EMA), Backfat Thickness(BF), and Marbling Score(MS). It was estimated using the MiX99(Lidauer, 2015) program to estimate GEBV using ssGBLUP, and reliability estimated was performed using SNP BLUP REL, a self developed reliability estimated Fortran program. The genetic parameters used were the results of Son(2020), and the GEBV of Hanwoo carcass traits was estimated to show the distribution for each trait to determine the normality of GEBV, and to determine whether it was approximated to the normal distribution. EBV estimated by the BLUP method was analyzed as 13.90, 3.78, -0.17, and 0.63, respectively, in CW, EMA, BF, MS in Hanwoo cows. KPN showed 22.51, 6.33, -0.58, and 0.93, which were higher than that of cows, and 16.03, 5.29, -0.83, and 0.84, respectively. GEBV estimated by the single step method was analyzed as 3.32, 1.36, -1.04, and 0.13, respectively, in CW, EMA, BF, and MS in Hanwoo cows. KPN was estimated to be 11.96, 4.14, -1.47, 0.48, Steer was estimated to

be 6.16, 3.20, -1.74, 0.44, and the difference between GEBV and EBV analyzed CW was -10.33, EMA was -2.23, BF was -0.89, and MS was -0.45, and GEBV was relatively lower than EBV. The GEBV of Hanwoo cows, steers and KPN was relatively lower than the estimated breeding value by the BLUP method. The reason for this result was that the proportion of pedigree data is higher than that of genomic data in the process of making H matrix due to the small number of genomic data, so it is estimated that the EBV estimate of the BLUP method using pedigree data is higher. The reliability analysis results of EBV estimated in 6,616 animals with genomic data were 0.48, 0.51, 0.51, and 0.59, respectively, in cows CW, EMA, BF, and MS, and the estimated EBV reliability of KPN was estimated to be 0.92, 0.92, 0.93 and 0.94, which was 40% higher on average. In addition, the GEBV reliability estimated by the single step GBLUP method, the cow CW was 0.66, EMA was 0.68, BF was 0.70 and MS was 0.77, and the KPN was 0.93, 0.93, 0.94, and 0.96. In the case of Hanwoo cows, it was confirmed that the reliability of GEBV increased by 17-19% on average than that of EBV, and in the case of KPN, the reliability of GEBV was slightly higher. It was found that KPNs containing a lot of carcass trait data of progeny had high reliability in EBV and GEBV, and cows with less carcass trait data of progeny than KPN had higher reliability using GEBV than EBV.

Key words: ssGBLUP, RPG, GEBV, Reliability, SNP-BLUP

서론

최근 유전체정보를 이용한 선발 방법인 유전체 선발 (Genomic Selection)을 적용하기 위해 많은 노력들을 하고 있다. 유전체육종가를 추정하기 위한 방법으로는 유전체 정보만을 이용하여 GRM를 구성하는 GBLUP(Genomic BLUP) 방법과 추정된 육종가 (Traditional EBV)와 유전체 정보를 이용하여 추정된 MBV(Molecular Breeding Value)를 혼합 공식 (Blending Formula)을 이용하여 계산하는 Two-step 방법과 기존의 혈연관계 (NRM) 매트릭스를 대신하여 NRM과 GRM (Genomic Relationship Matrix)을 혼합하여 만들어진 H-matrix를 이용하는 single-step GBLUP (Miszta et al., 2009)이 있으며, 최근에는 혈연관계를 이용하여 유전체 정보가 없는 개체의 유전체 정보를 생성하여 계산하는 single-step Super Hybrid 모델 (Fernando et al., 2014)을 이용하여 유전체육종가를 추정하고 있다. 유전체 정보를 활용하는 경우 기존 육종가에 비해 더 높은 신뢰도를 가지며 출생 시 조기 선발에 이용하여도 편익이 발생하지 않는 장점이 있다. 현재 우리나라에서 한우의 개량은 자손의 성적을 이용하여 개체의 능력을 간접 예측하는 후대검정시스템을 적용하고 있으며, 후보KPN의 나이가 약 5년 되었을 때, KPN 당 10두 내외의 후대검정우에 대한 표현형 성적을 얻게 되며 약 70%의 신뢰도로 개체의 유전평가가 실시되고 있다. 한우의 세대 당 유전적 개량량을 향상시키기 위해서는 유전평가의 신뢰도를 향상, 선발강도 향상, 세대간격 단축 등의 노력이 필요하다. 가축에서 유전체 선발의 적용은 기존의 BLUP 추정방법에 비하여 신뢰도 향상 및 세대간 단축을 통해 높은 유전적 개량량을 얻을 수 있으며(Schaeffer, 2006; Hayes et al., 2009), 특히 유전력이 낮은 형질, 개체의 생을 마감한 후에 생성되는 도체형질, 표현형 자료수집에 위험이 따르는 질병저항형질에 대해 더욱 유용하게 활용될 수 있다(Meuwissen et al., 1996; Dekkers et al., 2002). 현재까지 혈통 및 표현형자료를 이용하여 EBV(Estimated Breeding Value)를 추정하여 개체의 선발과 도태를 실시하였지만, 최근 유전체 자료를 기반으로 Single Step Method를 사용하여 표현형, 혈통 및 유전체 마커의 정보를 기반으로 하는 유전체 평가를 많이 사용하고 있다. Single Step Method는 전통적인 방식인 혈통 기반에 유전체 마커 정보를 추가하여 평가할 수 있으며, 이는 혈연관계행렬(Relationship matrix)과 호환되기 위해 혈연관계행렬을 조정하여 평가 방법을 제시하고 있다. 이 방법은 유전자형 자료의 단변량 및 다변량 모델에 대하여 모두 비교가 가능하다. 이를 통해 마커 정보가 있는 Single Step Method는 혈통 기반 방법에 비해 예측 능력을 향상시킬 수 있으며, 혈통을 가진 모든 개체에 대해 향상된 예측값을 제공받을 수 있다(Christensen et al., 2012).

ssGBLUP은 유전체 평가에 선호되는 방법(Aguilar et al., 2010; Christensen et al. 2010)이며, 이는 유전체육종가를 추정 시 유전자 마커 및 혈통 정보를 동시에 포함하여 평가하고 혼합모형식을 풀기 위한 방법으로 유전체와 관련된 G matrix를 역행렬하여 추정한다. 역행렬 계산과 혼합모형식의 반복적 풀이로 G^{-1} 행렬로 인한 고정된 연산의 수는 유전자형을 가진 개체 수에 따라 증가하게 된다. 일

반적으로 유전자형을 가진 개체 수가 30,000개를 초과하면 혼합모형식 분석 시간이 G^{-1} 행렬에 의해 좌우되기 때문이다. 이러한 문제를 완화하기 위해 Legarra et al.(2012)는 몇 가지의 Single Step 혼합모형식을 제시하였으며, 그 중 일부는 G^{-1} 행렬을 없애거나 무효화시키는 방법도 있다(Liu et al., 2014; Fernando et al., 2014).

Misztal et al.(2014) 및 Fragomeni et al.(2015)는 대규모 유전자형 개체군에 대한 추정을 가능하게 하는 검증된 최신 알고리즘(APY)을 발표하였으며, 이는 행렬의 근사된 희소행렬의 역행렬을 사용하도록 제안하였다(Masuda et al., 2016; Strandén et al., 2017). PCG 방법에 의한 반복적인 풀이과정에서 Single Step MME는 일반적인 최소한의 반복 회수를 제공하게 되며(Legarra and Ducrocq, 2012; Strandén and Mäntysaari, 2014), 이는 PCG에서 원래 MME와 동일한 수렴 특징을 가지지만 계산상 간단한 ssGTBLUP이라는 정확한 접근법을 제시하고 있다. 유전체 평가에서 있어 일반적인 평가는 유전체 혈연관계행렬을 가진 BLUP 또는 SNP BLUP 모형을 기초로 한다. Multi-Step 유전체 평가를 하기 위해, 상기 언급한 두 가지의 유전체 모형은 유전체 참조집단에 대해서도 동일한 예측치를 제공하는 것이 입증되었고, 두 가지 모형의 동일한 결과치는 표현형이 없는 어린 유전체 개체에서도 검증이 되었다. 양적형질의 유전 전달을 담당하는 유전자는 SNP 마커의 불완전한 연관불평형과 돌연변이로 인한 SNP 마커들의 유전분산으로 모두 설명할 수는 없다. RPG effect는 일반적으로 불완전한 연관불평형을 설명하기 위한 유전체 모형에 적합하다. 신뢰도 추정에 있어 Multi-Step GBLUP 및 SNP BLUP 모형이 RPG effect를 포함하는 경우 유전체 참조집단과 동일한 예측을 나타낸다는 전제로 이 두 가지 유전체 모형 모두 다양한 유전자를 가지고 있는 형질의 동일성 효과가 표현형이 없는 어린 개체에 대해서도 동일하게 검증이 되었다(Liu, 2017).

본 연구에서는 한우농가에서 직접 사육된 한우의 유전체 자료, 도축성적 및 혈통자료를 수집하여 Single Step 방법을 이용하여 한우의 유전체육종가를 추정하고 RPG effect를 이용하여 신뢰도를 추정하여 농가에서 직접적으로 활용 가능한 유전체 유전능력 추정을 위한 기초자료로 제공하고자 본 연구를 실시하였다.

재료 및 방법

1. 공시자료

본 연구를 위하여 전국의 한우농가에서 다양한 방법으로 채취된 8,413두의 DNA를 추출하였다. 유전체 정보를 생성하기 위해 Axiom Bovine 60k 버전3(Affymetrix Inc., 2006) SNP패널을 이용하였다. 유전체 분석의 정확성을 높이기 위한 Quality Control은 SNP가 염색체에 상에 있거나 염색체상의 위치(Position)가 확인되지 않은 SNP를 제거하였으며, SNP Call rate가 95% 이하, Minor Allele Frequency (MAF)가 0.01 이하, Hardy-Weinberg 불평형(HWE)을 측정한 Chi-square (χ^2) 값이 95%를 초과한 마커를 제거한 총 64,973개의 SNP 마커를 이용하였다. 유전체 정보를 가진 8,413두 중 중복개체, 혈통오류, 종속변수와 매칭이 되지 않는 개체는 제거하여 유전체 정보가 있는 참조 축군 6,616두를 분석에 이용하였다. 분석에 사용된 보증KPN는 정액을 이용하여 DNA추출을 진행하였으며, 거세우는 축산물품질평가원에서 제공받은 조직 샘플을 이용하여 DNA추출을 실시하였고, 암소의 경우, 모근을 채취하여 모근에 있는 DNA를 추출하여 분석에 이용하였다.

혈통자료의 경우 6,616두와 관련된 5계대의 자료를 한국종축개량협회에서 수집하였다. 분석에 이용된 자료는 총 5,153,168두에 대한 혈통자료를 이용하였다. 도체자료는 5,153,168두 중 도체자료를 가지고 있는 개체 중 결측치 및 이상치를 사전제거 한 후 남은 2,376,865두에 대해서 축산물품질평가원에서 자료를 수집하여 분석에 이용하였다. 이용된 형질은 도체중(Carcass Weight, CW), 등심단면적(Eye Muscle Area, EMA), 등지방두께(Backfat Thickness, BF) 및 근내지방도(Marbling Score, MS)의 4개 형질을 고려하였다.

2) 분석방법

① ssGBLUP

ssGBLUP을 이용한 평가 모델은 다음과 같다.

$$y = Xb + Zu + e$$

여기서, y 는 표현형 자료 벡터, b 와 u 는 각각 고정효과와 임의효과에 대한 추정치 벡터, X 와 Z 는 각각 b 와 g 에 대한 계수 행렬, e 는 잔차에 대한 벡터이다. 또한, GRM를 구성을 위해 Van Raden (2008)이 제안한 방법은 다음과 같다.

$$G = \frac{ZZ'}{2 \sum p_i(1-p_i)}$$

여기서, p_i 는 i 번째 SNP 마커의 대립유전자 빈도(allele frequency), Z 는 마커에 대한 근접행렬(incidence matrix)이며, 혼합방정식(MME)은 다음과 같다.

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}Z \\ Z'R^{-1}X & Z'R^{-1}Z + G^{-1} \end{bmatrix} \begin{bmatrix} \hat{\beta} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ Z'R^{-1}y \end{bmatrix}$$

하지만, GBLUP의 평가는 유전체 정보가 있는 개체만 모델에 포함되어 평가되는 단점이 있다. 이를 보완하기 위해 Misztal et al (2009)은 GRM과 NRM을 이용하여 유전체 정보가 있는 개체와 유전체 정보가 없는 개체를 동시에 이용하여 육종가를 한번에 추정하는 single-step GBLUP (ssGBLUP)을 제시하였다. ssGBLUP은 BLUP의 NRM과 GRM 대신 Harmonic Matrix를 이용하는 방식이며, 평가에 이용되어지는 H-matrix 역행렬 (H^{-1}) 추정하는 방법은 다음과 같다.

$$H^{-1} = A^{-1} + \begin{bmatrix} 0 & 0 \\ 0 & wG^{-1} - (1-w)A_{22}^{-1} \end{bmatrix}$$

평가를 하는 개체 수의 증가는 H^{-1} 의 계산 시, 컴퓨터의 성능적인 계산에 어려움이 있지만, Legarra et al.(2012)은 효율적인 H^{-1} 를 만들기 위한 방법을 제안하면서 개체 수가 늘어남에 따라 행렬크기의 증가 문제를 해결할 수 있다. MEM 방법에는 RR-BLUP, SNP-BLUP 및 Bayes 방법이 있다. Bayes 방법은 사전 분포에 적합한 샘플(Random sample)을 추출하여 유전체 정보가 있는 개체보다 SNP 마커의 수가 많은 것과 관계없이 적합한 SNP 효과를 추정하는 것이 가능하다. 상기 모형을 이용하여 GEBV를 추정하는 데 MiX99(Lidauer, 2015) 프로그램을 이용하여 추정하였다.

② Reliability with RPG effect

다음의 Mixed linear model(혼합 선형 모형)은 유전체 참조집단의 표현형 데이터를 분석하는데 이용된다(Liu et al., 2014).

$$y = Xb + W_2\mu_2 + e$$

여기서

- y 는 후대성적이 있는 KPN의 역회귀육종가, 정확한 표현형 자료들의 $n_T \times 1$ 번째 벡터(Jairath et al., 1998);
- n_T 는 표현형 자료들의 수; b 는 모든 fixed effects의 $n_F \times 1$ 번째 벡터;
- n_F 는 모든 fixed effects의 수; μ_2 는 additive genetic effects의 $n \times 1$ 번째 벡터;
- n 는 reference animals(참조집단)의 수; e 는 residual effects의 $n_T \times 1$ 의 벡터;
- X 는 $n_T \times n_F$ 번째 값,
- W_2 는 $n_T \times n$ 의 b 와 μ_2 효과들의 incidence matrices 값

상기 유전체 모형은 다형질 유전체 평가로 확장할 수 있다. 표현형 자료가 있는 참조집단의 개체 중에서 몇 가지 형질에 대해 결측치가 있어도 참조모집단에 포함시켜서 평가할 수 있다. 다형질 참조 모집단을 한 번에 평가하는 것이 각 형질별로 분리하여 평가하는 것보다 컴퓨터 연산방법이 효율적이다. 이는 모형에 포함되는 효과들을 병렬구조로 평가할 수 있기 때문에 가능하다. 유전체 평가를 위해 선별된 m SNP 마커는 불완전한 연관불평형 때문에 모든 상가적 유전분산을 설명할 수 없다고 가정한다. 여기서 k 는 모든 m SNP 마커에 의해 설명되지 않는 상가적 유전분산의 비율이다. 따라서, 상가적 유전효과는 다음과 같이 나눌 수 있다.

$$u_2 = Z_g + a_2$$

여기서

g 는 m SNP 마커의 상가적 유전효과의 $m \times 1$ 번째 벡터

a_2 는 참조집단의 RPG 효과의 $n \times 1$ 번째 벡터

Z 는 참조집단의 회귀계수가 포함된 m SNP 마커의 $n \times m$ 번째 유전자형 design matrix: $2 - 2p_j, 1 - 2p_j, -2p_j, j$ 번째 마커의 AA, AB, BB의 유전자타입(VanRaden, 2008)

여기서

p_j 는 j 번째 SNP 마커의 allele 빈도이며 DGV를 다음과 같이 쓸 수 있다.

$$u_{DGV} = Zg$$

모형식의 u_2 를 방정식으로 대체하면 SNP BLUP 모형이 된다.

$$y = Xb + W_2 Z_g + W_2 a_2 + e$$

$\text{var}(e) = R$ 로 두면, 참조집단의 RPG 효과는 정규분포를 갖는다.

$$a_2 \sim N(0, k \sigma_u^2 A_{22})$$

여기서, A_{22} 는 참조집단 개체들의 혈연계수행렬을 나타낸다.

잔차 유전체 분산성분인(Residual polygenic variance parameter) k 는 0과 1 사이의 값을 취할 수 있지만 2개의 경계 값인 $k = 0$, $k = 1$ 은 제외한다. 또한 SNP 마커의 상가적 유전효과는 정규분포를 따른다고 가정한다.

$$g \sim N(0, (1 - k) \sigma_u^2 B)$$

모든 m SNP 마커가 SNP BLUP 모형과 같은 상가적 유전분산을 설명하는 경우, 연관성이 없는 SNP효과의 가정하에서 B 는 $m \times m$ 의 대각선 행렬이다(Meuwissen et al., 2001; VanRaden, 2008; Liu, 2011). 그러면 다음과 같이 나타낼 수 있다.

$$B = \left[\sum_{j=1}^m 2p_j(1 - p_j) \right]^{-1} I$$

g 와 a_2 사이의 공분산을 0이라고 가정할 때:

$$\text{Cov}(g, a_2) = 0$$

관측된 유전체 관계행렬은 다음과 같이 표시한다.

$$G_{DGV} = ZBZ'$$

Liu et al.(2014)에서 보고한 바와 같이, 다음과 같이 나타낼 수 있다.

$$\text{var}(u_2) = \text{var}(Zg + a_2) = [(1 - k)ZBZ' + kA_{22}] \sigma_u^2 = [(1 - k)G_{DGV} + kA_{22}] \sigma_u^2 = \sigma_u^2 G_{22}$$

상기 공식에서 유전체 관계행렬 G_{22} 을 볼 수 있는데 이는 VanRaden(2008)이 처음으로 제시하였고, SNP BLUP 모델에서 가중치 $(1 - k)$ 를 가지는 유전체 혈연계수행렬과 가중치 k 를 가지는 추정 혈통계수행렬 A_{22} 의 선형합수이다.

RPG 효과 a_2 와 상가적 유전효과 u_2 사이의 공분산은 다음과 같다.

$$\text{cov}(a_2, u_2) = \text{cov}(a_2, Zg + a_2) = \text{var}(a_2) = k\sigma_u^2 A_{22}$$

상기 식의 공분산 식을 이용하여, RPG 효과 a_2 는 GEBV $\hat{u}_2$ 에서 예측할 수 있다.

$$\hat{a}_2 = \text{cov}(a_2, u_2)[\text{var}(u_2)]^{-1}\hat{u}_2 = k\sigma_u^2 A_{22}\sigma_u^{-1} = kA_{22}G_{22}^{-1}\hat{u}_2$$

상기 식을 기반으로 계산을 하면, 다음의 식을 얻을 수 있다.

$$A_{22}^{-1}\hat{a}_2 = kG_{22}^{-1}\hat{u}_2$$

유사한 방법으로 DGV와 u_2 의 공분산을 얻을 수 있다.

$$\text{cov}(Zg, u_2) = \text{cov}(Zg, Zg + a_2) = \text{var}(Zg) = (1-k)\sigma_u^2 ZBZ' = (1-k)\sigma_u^2 G_{\text{DGV}}$$

$\hat{u}_{\text{DGV}}$ 참조집단의 DGV는 $\hat{u}_2$ 에서 다음과 같이 예측할 수 있다.

$$\hat{u}_{\text{DGV}} = Z\hat{g} = \text{cov}(Zg, u_2)[\text{var}(u_2)]^{-1}\hat{u}_2 = (1-k)\sigma_u^2 G_{\text{DGV}}\sigma_u^{-2}G_{22}^{-1}\hat{u}_2 = (1-k)G_{\text{DGV}}G_{22}^{-1}\hat{u}_2$$

$(G_{\text{DGV}})^{-1}Z'B$ 에 의한 사전 다중 방정식은 SNP효과 g 에 대한 방정식으로 이어진다.

$$\hat{g} = (1-k)BZ'G_{22}^{-1}\hat{u}_2$$

상기 방정식은 VanRaden (2008)이 최초로 제시하였고, 동등한 예측치를 나타내는 공식은 Stranden and Garrick (2009)이 제시하였다.

GBLUP 모형의 적합한 마커 개수보다 적은 수의 참조집단이 존재하더라도 역행렬 $G_{\text{DGV}}^{-1} = (ZBZ')^{-1}$ 의 존재는 보장되지 않는다(계산이 되지 않고 바로 다음 단계를 수행할 수 있다.). 이는 마커의 유전자형이 연관불평형 때문에 선형적으로 독립적이지 않거나 동일한 유전자가 나타나는 경우 등이 해당된다. 참조집단의 개체수가 마커수를 초과할 경우, 유전체 행렬은 비변환적(불가역, 비가역)일 가능성이 더 높다. 단, residual polygenic effect를 포함하면 수치의 안정성이 향상될 수 있으며, 다음과 같은 방정식을 도출할 수 있다.

$$B_{\hat{g}}^{-1} = (1-k)Z'G_{22}^{-1}\hat{u}_2$$

상기 방정식을 사용하여 유전체육종가를 참조집단에 대해 각각 RPG와 DGV로 변환할 수 있다. SNP BLUP 모형을 이용하여 참조집단의 RPG효과를 추정하기 위해 첨자 0으로 표시된 유전체 자료가 없는 선조와 참조집단의 혈연계수행렬은 다음과 같이 정의될 수 있다.

$$\begin{bmatrix} A_{00} & A_{02} \\ A_{20} & A_{22} \end{bmatrix},$$

그리고 상기 혈연계수행렬을 역행렬 하면,

$$\begin{bmatrix} A_{00} & A_{02} \\ A_{20} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A^{00} & A^{02} \\ A^{20} & A^{22} \end{bmatrix}$$

분할(partitioned) 행렬의 역행렬 규칙에 따라(Harville, 1997) 다음과 같이 표시된다.

$$A_{22}^{-1} = A^{22} - A^{20}(A^{00})^{-1}A^{02}$$

참조집단의 유전체자료가 없는 선조의 RPG에 대한 방정식을 추가함으로써 SNP BLUP 모델에 대한 Henderson의 혼합 모형 방정식은 방정식과 같다. 여기서 $\hat{a}_0$ 는 참조집단의 유전체자료가 없는 조상의 RPG를 나타낸다.

$$\begin{bmatrix} X'R^{-1}X & X'R^{-1}W_2 & 0 & X'R^{-1}W_2Z \\ W_2'R^{-1}W_2 + \frac{1}{k}\sigma_u^{-2}A^{22} & \frac{1}{k}\sigma_u^{-2}A^{20} & 0 & W_2'R^{-1}W_2Z \\ \frac{1}{k}\sigma_u^{-2}A^{00} & 0 & 0 & 0 \\ \text{symm.} & Z'W_2'R^{-1}W_2Z + \frac{1}{1-k}\sigma_u^{-2}B^{-1} & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a}_2 \\ \hat{a}_0 \\ \hat{g} \end{bmatrix} = \begin{bmatrix} X'R^{-1}y \\ W_2'R^{-1}y \\ 0 \\ Z'W_2'R^{-1}y \end{bmatrix}$$

Single step SNP model은 Interbull 등에서 사용되고 있으며(Liu et al, 2014) Single step SNP model을 사용하기 위해서는 먼저 유전체 혈연계수행렬의 방법을 고려해야 한다. 현재 Interbull의 유전체 신뢰도 추정 방법은 유전체 개체 수에 대한 제한이 없으며, 모든 참조집단을 포함하는 SNP 관계 매트릭스의 역행렬을 구할 수 있다. 다른 대안들은 역행렬 시 너무 커지게 되는 유전체 혈연계수행렬과 함께 계산되기에 효율성이 떨어진다. Interbull 유전체 신뢰도 추정 방법의 또 다른 특징은 EDC를 얻기 위해 역행렬을 수행할 필요가 없다는 것이다. Interbull 유전체 신뢰도 추정 방법은 Interbull 표준 절차에 따라 계산된 개체들의 EDC를 요구하며, Interbull 유전체 신뢰도 추정 방법은 유전체 검증을 유전체 신뢰도 추정 절차의 일부로 통합함으로써 실현되며, 유전체의 신뢰성에 근사치를 두고 계산하게 된다.

본 연구에서는 상기 방법을 기반으로 자체 개발된 Fortran 프로그램은 SNP BLUP REL을 사용하였다. 유전체 신뢰도 추정에 대한 논문에서는 genomic 평가에 사용되는 것과 동일한 allele frequencies를 SNP BLUP 신뢰도 추정 소프트웨어로 권장하고 있다.

③ 유전모수

본 연구에서 사용된 유전모수 값을 Table 1에 제시하였다. 사용된 유전모수 값은 손(2020)의 결과를 이용하였다. 도체중, 등심단면적, 등지방두께 및 근내지방도의 유전분산 값은 각각 500.8 ± 47.01 , 29.92 ± 3.06 , 8.738 ± 0.85 및 1.545 ± 0.26 , 잔차분산 값은 각각 $1,298.5 \pm 31.01$, 77.28 ± 2.82 , 15.973 ± 0.50 및 1.662 ± 0.18 , 유전력은 각각 0.28 ± 0.02 , 0.28 ± 0.03 , 0.35 ± 0.03 및 0.48 ± 0.07 로 제시되었다.

Table 1. Estimated genetic(σ_a^2) and residual(σ_e^2) variance components, heritabilities(h^2), phenotype genetic correlations among traits of Hanwoo steer

Trait	h^2	Component of Variance		Correlations			
		σ_a^2	σ_e^2	CW	EMA	BF	MS
CW	0.28 ± 0.02	500.8 ± 47.01	$1,298.5 \pm 31.01$		0.43	0.20	0.14
EMA	0.28 ± 0.03	29.92 ± 3.06	77.28 ± 2.82	0.50		-0.29	0.58
BF	0.35 ± 0.03	8.738 ± 0.85	15.973 ± 0.50	0.34	0.11		-0.13
MS	0.48 ± 0.07	1.545 ± 0.26	1.662 ± 0.18	0.18	0.33	0.14	

④ 정규성 검정

한우 도체형질의 유전체육종가를 추정하여 추정된 유전체육종가의 정규성을 파악하기 위하여 각 형질별로 분포를 나타내어 정규분포에 근사되는지를 파악하였다. 정규분석은 SAS UNIVARIATE Procedure의 Checking Variables for Normality를 이용하여 Kolmogorov-Smirnov 통계량(D)으로 검정하였다. 이러한 적합도 검정으로 지정 분포로부터 관측값이 제대로 작성되었는지를 확인할 수 있는 비모수적 검정이다. Kolmogorov-Smirnov 통계량(D) 추정을 위한 공식은 아래와 같다.

$$D = \sup_x |F_n(x) - F(x)|$$

결과 및 고찰

한우농가에서 사육중인 암소를 평가하기 위한 유전체 분석 집단의 구성은 살아있는 암소인 대상개체의 모근 및 도축된 후대의 조직샘플, 후대의 아버지 개체의 정액을 한 쌍으로 묶어 수집 후 유전체자료를 생성하였으며, 분석된 내역은 Table 2에 나타내었다. 분석된 6,616두 중 도축된 개체는 3,091두로 약 47%의 비중을 차지하였다. 이 중, 암소와 거세우는 각각 2,449두 및 642두 있었으며, 분석에 이용된 보증KPN와 살아있는 암소는 각각 275두 및 3,053두였다. 거세우 197두는 아직 도축이 되지 않아 도축성적을 가지고 있지 않은 것으로 나타났다. 한우 암소는 도축월령이 54개월령 내외로 초산일령이 24.4개월령, 임신기간이 280일 전후임을 감안했을 때 평균적으로 2산에서 3산 정도 후대를 생산한 후 도축된다(손, 2020). Table 2를 보면 유전체자료를 가지고 있는 암소 평균 도축월령이 64.6개월령으로 유전체자료를 가지고 있지 않는 암소 평균 도축월령인 47.4개월령으로 다소 높게 나타났는데 이는 유전체 분석을 위한 암소 집단의 조건으로 혈통의 연결이 잘 되어 있으며, 후대를 많이 생산한 개체를 우선 대상으로 지정하였기 때문이다. 도축성적 간 비교를 해보면 유전체 자료가 있는 암소개체의 도체성적은 평균 도체중, 등심단면적, 등지방두께 및 근내지방도가 각각 357.29 kg, 85.18 cm², 14.64 mm 및 4.21점으로 나타났으며, 유전체 자료가 없는 암소개체의 도체성적은 각각 352.92 kg, 83.73 cm², 13.32 mm 및 4.36점으로 나타났다. 거세우의 도축 시기는 평균적으로 29개월에서 31개월령에 도축 축하하게 되는데, 유전체자료를 가지고 있는 거세우 2,499두의 평균 도축월령은 29.96개월령이며, 평균 도체중, 등심단면적, 등지방두께 및 근내지방도는 각각 458.94 kg, 96.44 cm², 13.80 mm 및 6.39점으로 나타났다. 이는 유전체자료가 없는 거세우 1,431,299두에서 도축월령 평균 30.07개월령과는 큰 차이가 없었으나, 도체중, 등심단면적, 등지방두께 및 근내지방도의 평균성적이 각각 443.51 kg, 93.72 cm², 13.62 mm 및 5.94점 보다 다소 높게 나타났다. 자신의 유전체 자료와 도축자료를 가지고 있는 개체는 유전체 자료를 가지고 있는 개체들 6,616두 중에 암소 642두, 거세우 2,449두였다. 유전체 자료를 가지고 있는 개체와 혈통으로 연결 되었지만 자신의 유전체 자료를 가지고 있지 않은 개체는 암소에서는 945,566두, 거세우에서는 약 140만두 가량이 되었다. 암소에서는 유전체 자료를 가지고 있는 개체는 그렇지 않은 개체들보다 약 17개월 가량 더 늦게 도축이 되었지만 그 밖의 도체형질에 대한 평균은 큰 차이를 보이지 않았다. 거세우의 경우 출하시기가 수익에 많은 영향을 끼치므로 도축시기는 큰 차이가 없었으며, 암소와 마찬가지로 도체성적에 대한 값도 큰 차이를 보이지 않았다.

Table 2. Information of Bull and cow records on genotype data

	No. of Bull	No. Cow	No. Steer	Sum.
Slaughter	-	642	2,449	3,091
Alive	275	3,053	197	3,525
Amount	275	3,695	2,646	6,616

Table 3. Basic statistics on carcass traits in Hanwoo using genomic data linked to pedigree

Traits (unit5)	Included Genotype data			Without Genotype data		
	Obs.	Mean	STD	Obs.	Mean	STD
Cows						
Age at slaughter	642	64.63	27.71	945,566	47.38	15.27
CW (kg)	642	357.29	46.79	945,566	352.92	46.77
EMA (cm ²)	642	85.18	11.87	945,566	83.73	11.49
BF (mm)	642	14.64	6.05	945,566	13.32	5.37
MS (score)	642	4.21	1.94	945,566	4.36	1.91
Steers						
Age at slaughter	2,449	29.96	1.70	1,431,299	30.07	1.89
CW (kg)	2,449	458.94	49.66	1,431,299	443.51	49.80
EMA (cm ²)	2,449	96.44	11.92	1,431,299	93.72	11.53
BF (mm)	2,449	13.80	5.00	1,431,299	13.62	5.05
MS (score)	2,449	6.39	1.66	1,431,299	5.94	1.84

CW: Carcass weight (kg), EMA: Eye muscle area (cm²), BF: Backfat thickness (mm), MS: Marbling score.

한우 암소, 거세우, KPN의 육종가를 BLUP방법과 ssGBLUP방법으로 추정하였을 때 기대되는 육종가를 Table 4.에 나타내었다. BLUP방법으로 추정된 육종가는 한우 암소에서 도체중, 등심단면적, 등지방두께 및 근내지방도에서 각각 13.90, 3.78, -0.17, 0.63으로 분석되었다. KPN에서는 그보다 높은 22.51, 6.33, -0.58, 0.93을 나타냈으며, 거세우에서는 16.03, 5.29, -0.83, 0.84를 나타냈다. Single Step 방법으로 추정된 유전체육종가는 한우 암소에서 도체중, 등심단면적, 등지방두께 및 근내지방도에서 각각 3.32, 1.36, -1.04, 0.13으로 분석되었다. KPN에 대해서는 11.96, 4.14, -1.47, 0.48의 값으로 나타났고, 거세우에서는 6.16, 3.20, -1.74, 0.44로 분석되어 유전체육종가와 육종가 간의 차이는 도체중에서 -10.33, 등심단면적에서 -2.23, 등지방두께에서 -0.89, 근내지방도에서 -0.45로 나타났으며, 상대적으로 유전체육종가 값이 육종가 값보다 낮게 나타났다. 한우 암소, 거세우, KPN의 유전체육종가는 상대적으로 BLUP방법으로 추정된 육종가보다 낮게 나타났다. 이러한 결과가 나오는 이유는 유전체육종가의 자료수가 작아 H matrix를 만드는 과정에서 혈연계수행렬에 유전체 자료의 비중보다 혈통자료에 대한 비중이 높게 차지하여 혈통자료를 이용하는 BLUP방법의 EBV 추정치가 더 높은 수치로 추정된다고 사료된다.

Table 4. Means(SD) for EBVs of CW, EMA, BF, and MS on genotyped animals in Hanwoo using BLUP and ssGBLUP

Traits	BLUP	ssGBLUP	Diff.
Cow (N=3,695)			
CW	13.90(15.79)	3.32(18.46)	-10.58
EMA	3.78(4.01)	1.36(4.57)	-2.42
BF	-0.17(1.94)	-1.04(2.12)	-0.87
MS	0.63(0.91)	0.13(0.98)	-0.50
Bull (N=275)			
CW	22.51(22.16)	11.96(22.83)	-10.55
EMA	6.33(5.20)	4.14(5.35)	-2.19
BF	-0.58(2.33)	-1.47(2.36)	-0.89
MS	0.93(0.89)	0.48(0.90)	-0.45
Steer (N=2,646)			
CW	16.03(18.47)	6.16(21.13)	-9.87
EMA	5.29(4.20)	3.20(4.66)	-2.09
BF	-0.83(1.89)	-1.74(2.13)	-0.91
MS	0.84(0.88)	0.44(0.95)	-0.40

CW: Carcass weight (kg), EMA: Eye muscle area (cm²), BF: Backfat thickness (mm), MS: Marbling score.

Table 5. Comparison of estimated breeding value(EBV) and genomic estimated breeding value(GEBV) reliability(REL)

Traits	EBV_REL	GEBV_REL	Diff.
Cow (n = 3,695)			
CW	0.48	0.66	0.18
EMA	0.51	0.68	0.17
BF	0.51	0.7	0.19
MS	0.59	0.77	0.18
Bull (n = 275)			
CW	0.92	0.93	0.01
EMA	0.92	0.93	0.01
BF	0.93	0.94	0.01
MS	0.94	0.96	0.02

Diff : Difference, CW: Carcass weight (kg), EMA: Eye muscle area (cm²), BF: Backfat thickness (mm), MS: Marbling score.

한우 240만두의 개체의 유전체육종가 값이 정규분포를 따르는지 확인하기 위하여 콜모고로프-스미노프(Kolmogorov-Smirnov)검정으로 정규성 검정을 진행하여 Figure 1~4에 제시하였다. 육종가와 유전체육종가에 대하여 정규성을 나타내는지 검정하고자 정규분포함수에 적합시킨 히스토그램 및 QQ-PLOT을 도체형질별로 나타내었다. Figure. 1은 한우 암소의 육종가에 대한 정규성 검정이며 Figure. 2는 한우 암소의 유전체육종가에 대한 정규성 검정 결과이다. Figure. 3은 한우 거세우와 KPN를 전부 수소로 묶어 육종가에 대한 정규성분포를 그린 것이며, Figure. 4도 한우 수소에 대하여 유전체육종가에 대한 정규성 분포를 그린 것이다. Kolmogorov Smirnov검정한 결과 네 가지 도체형질에 대한 육종가 및 유전체육종가 값은 정규분포를 따르는 것으로 나타났다($P \text{ value} \leq 0.05$).

유전체 자료를 가지고 있는 한우 6,616두에 대한 도축자료로부터 추정된 육종가의 신뢰도는 암소에서 도체중, 등심단면적, 등지방두께 및 근내지방도 각각 0.48, 0.51, 0.51 및 0.59로 나타났으며 한우 KPN의 추정된 육종가 신뢰도는 0.92, 0.92, 0.93 및 0.94로 평균 40%정도 높게 추정되었다. 반면 Single Step GBLUP방법으로 추정된 유전체육종가 신뢰도는 암소에서 도체중 0.66, 등심단면적 0.68, 등지방두께 0.70, 근내지방도 0.77을 나타냈으며, 한우 KPN에서는 차례대로 0.93, 0.93, 0.94, 0.96으로 나타났다. 한우 암소의 경우 유전체육종가의 신뢰도가 육종가의 신뢰도 보다 평균 17~19% 상승되는 것을 확인하였으며, KPN의 경우 미약하지만 유전체육종가의 신뢰도가 약간 높게 나타났다. 이는 후대의 도체성적이 많이 포함된 KPN의 경우는 육종가 및 유전체육종가에서 높은 신뢰도를 가지고 있으며, 후대의 도체성적이 KPN보다 적은 암소의 경우에는 육종가보다 유전체육종가를 이용한 신뢰도가 높다는 것을 알 수 있었다.

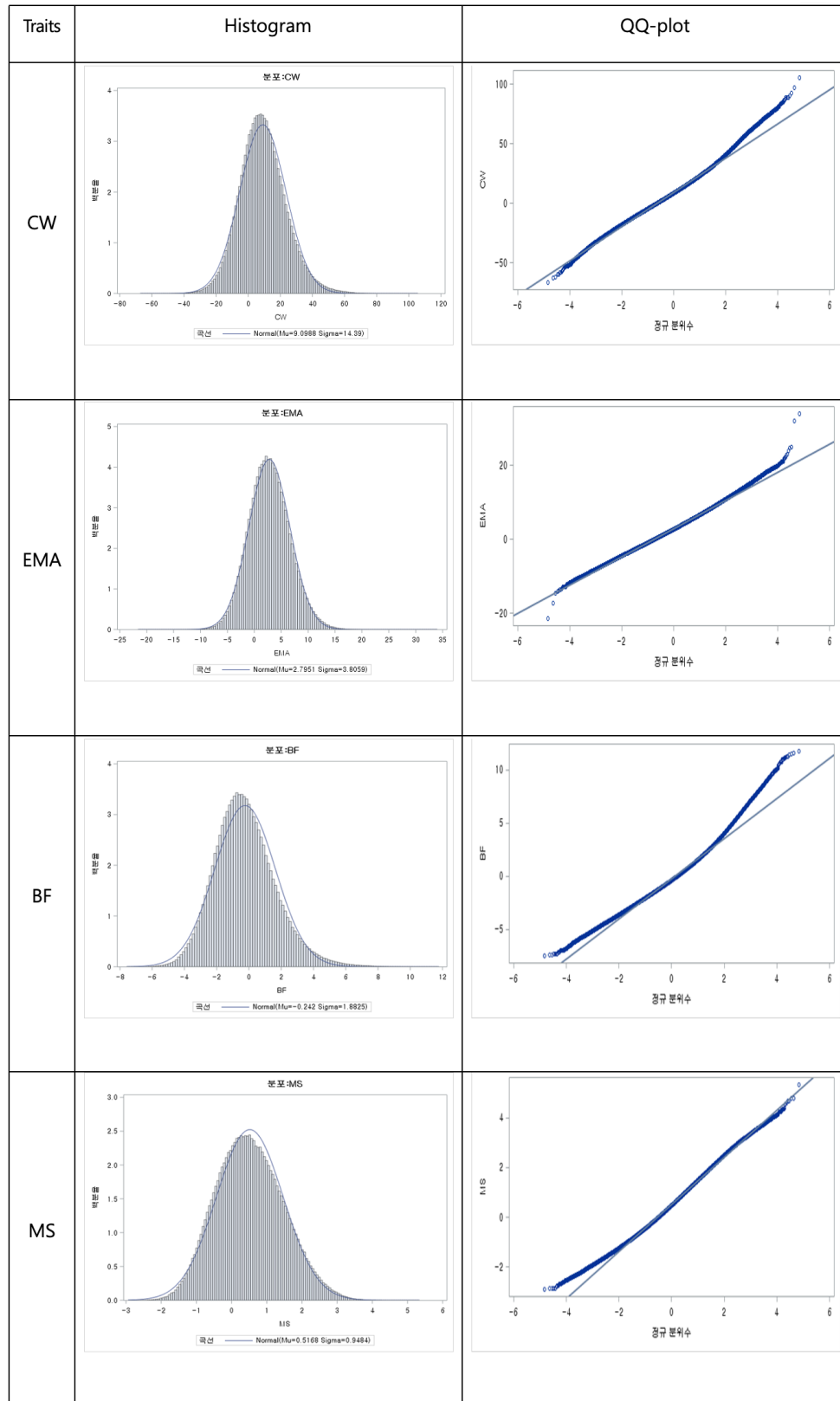


Figure 1. The histogram (left) and QQ-plot (right) for normality test using the SNP effects for CW, EMA, BF, MS in Hanwoo Cow

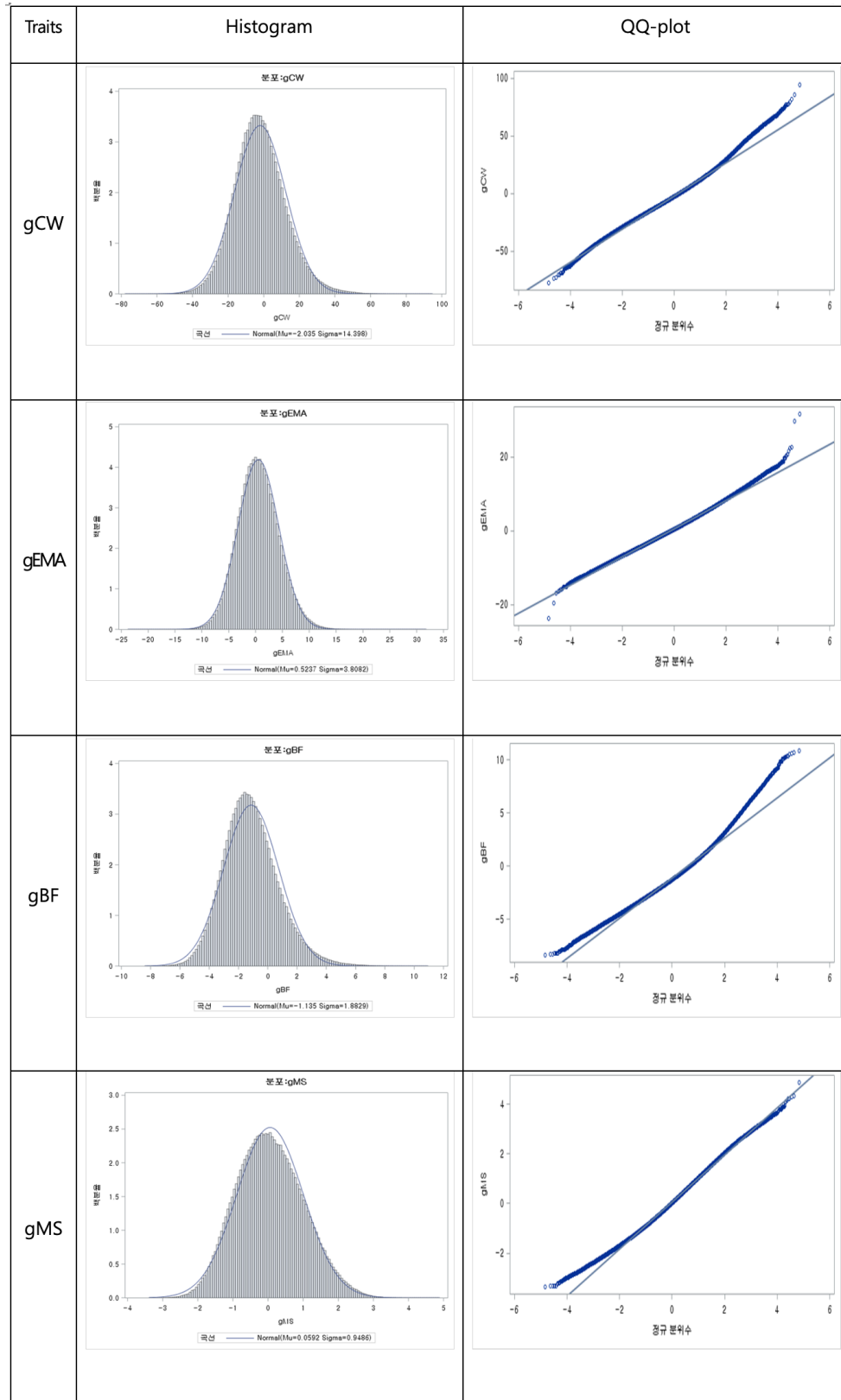


Figure 2. The histogram (left) and QQ-plot (right) for normality test using the SNP effects for gCW, gEMA, gBF, gMS in Hanwoo Cow

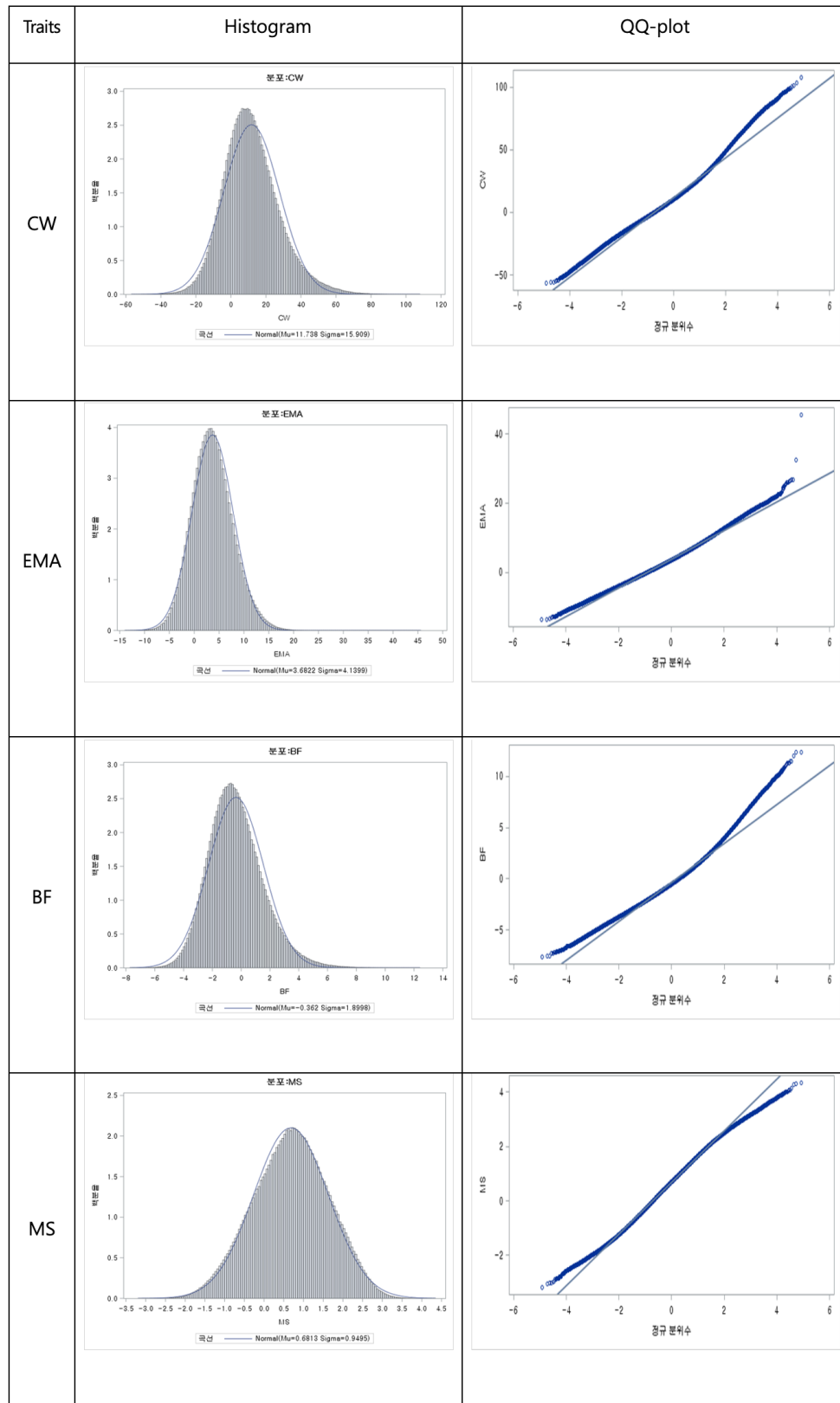


Figure 3. The histogram (left) and QQ-plot (right) for normality test using the SNP effects for CW, EMA, BF, MS in Hanwoo Bull and Steer

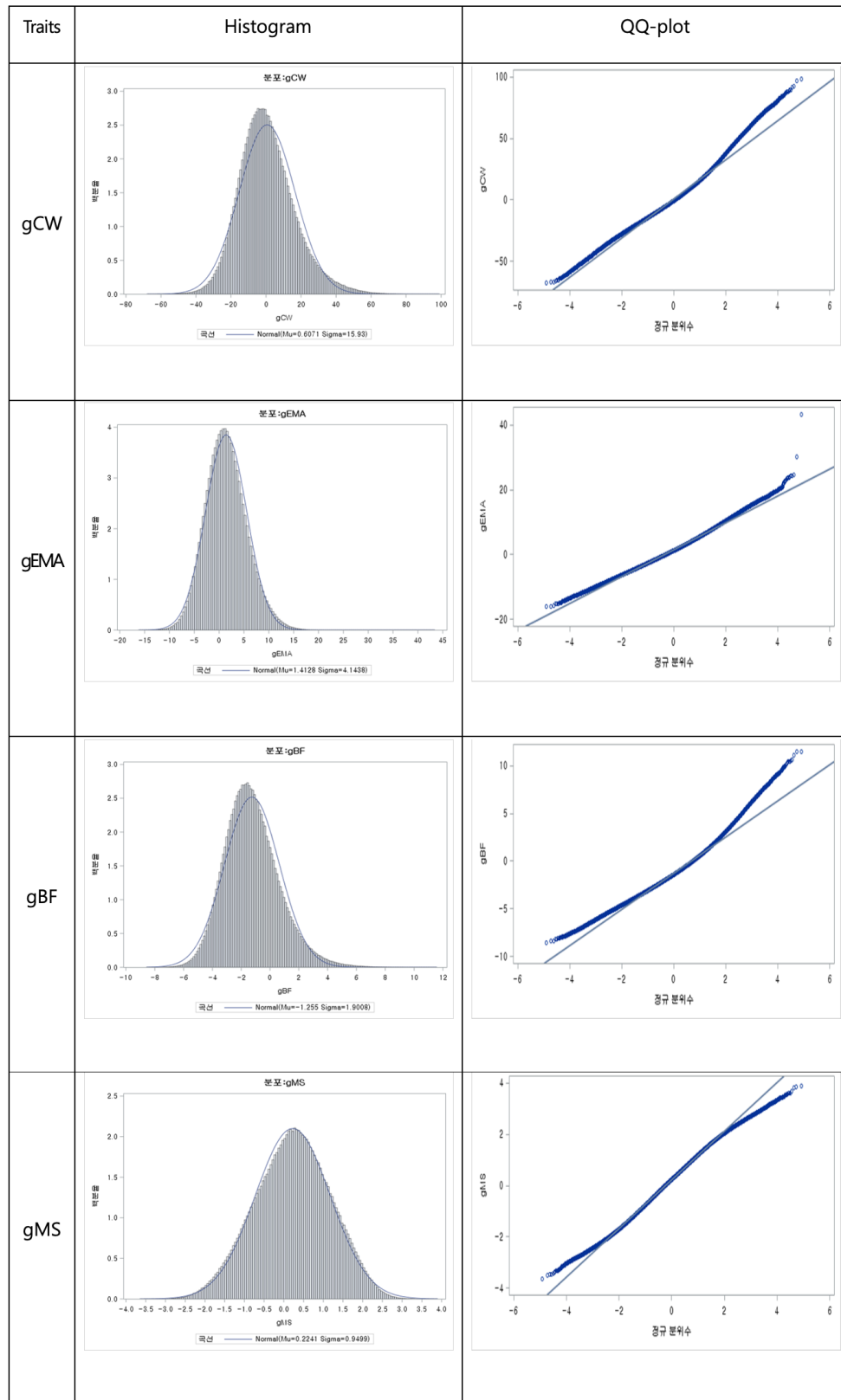


Figure 4. The histogram (left) and QQ-plot (right) for normality test using the SNP effects for gCW, gEMA, gBF, gMS in Hanwoo Bull and Steer

요약

본 연구를 위하여 전국의 한우농가에서 다양한 방법으로 채취된 8,413두의 DNA를 추출하였다. 유전체 정보를 생성하기 위해 Axiom Bovine 60k ver3(Affymetrix Inc., 2006) SNP패널을 이용하였다. 유전체 분석의 정확성을 높이기 위한 Quality Control은 SNP가 성염색체에 상에 있거나 염색체상의 위치(Position)가 확인되지 않은 SNP를 제거하였으며, SNP Call rate가 95% 이하, Minor Allele Frequency(MAF)가 0.01 이하, Hardy-Weinberg 불평형(HWE)을 측정한 Chi-square (χ^2) 값이 95%를 초과한 마커를 제거한 총 64,973개의 SNP 마커를 이용하였다. 유전체 정보를 가진 8,413두 중 중복개체, 혈통오류, 종속변수와 매칭이 되지 않는 개체는 제거하고 유전체 정보가 있는 참조축군 6,616두를 분석에 이용하였다. 분석에 사용된 보증KPN는 정액을 이용하여 DNA 추출하였으며, 거세우는 축산물품질평가원에서 제공받은 조직 샘플을 이용하여 DNA 추출하였고, 암소의 경우, 모근을 채취하여 DNA를 추출하여 분석에 이용하였다. 혈통자료는 6,616두와 연관된 5계대의 자료를 한국축종개량협회에서 수집하여 총 5,153,168두에 대한 혈통자료를 이용하였다. 도체성적을 보유한 개체 5,153,168두 중 결측치 및 이상치를 사전제거 한 후 남은 2,376,865두에 대해서 축산물품질평가원에서 자료를 수집하여 분석에 이용하였다. 이용된 형질은 도체중(Carcass Weight, CW), 등심단면적(Eye Muscle Area, EMA), 등지방두께(Backfat Thickness, BF) 및 근내지방도(Marbling Score, MS)의 4개 형질을 고려하였다. ssGBLUP을 이용하여 GEBV를 추정하는데 MiX99(Lidauer, 2015) 프로그램을 이용하여 추정하였고 신뢰도 추정은 자체 개발된 신뢰도 추정 Fortran 프로그램인 SNP BLUP REL을 사용하였다. 사용된 유전모수 값은 손(2020)의 결과를 이용하였으며, 한우 도체형질의 유전체육종가를 추정하여 추정된 유전체육종가의 정규성을 파악하기 위하여 각 형질별로 분포를 나타내어 정규분포에 근사되는지를 파악하였다.

BLUP방법으로 추정된 육종가는 한우 암소에서 도체중, 등심단면적, 등지방두께 및 근내지방도에서 각각 13.90, 3.78, -0.17, 0.63으로 분석되었다. KPN에서는 그보다 높은 22.51, 6.33, -0.58, 0.93을 나타냈으며, 거세우에서는 16.03, 5.29, -0.83, 0.84를 나타냈다. Single Step 방법으로 추정된 유전체육종가는 한우 암소에서 도체중, 등심단면적, 등지방두께 및 근내지방도에서 각각 3.32, 1.36, -1.04, 0.13으로 분석되었다. KPN에 대해서는 11.96, 4.14, -1.47, 0.48의 값으로 나타났고, 거세우에서는 6.16, 3.20, -1.74, 0.44로 분석되어 유전체육종가와 육종가 간의 차이는 도체중에서 -10.33, 등심단면적에서 -2.23, 등지방두께에서 -0.89, 근내지방도에서 -0.45로 나타났으며, 상대적으로 유전체육종가 값이 육종가 값보다 낮게 나타났다. 한우 암소, 거세우, KPN의 유전체육종가는 상대적으로 BLUP방법으로 추정된 육종가보다 낮게 나타났다. 이러한 결과가 나오는 이유는 유전체육종가의 자료수가 작아 H matrix를 만드는 과정에서 혈연계수행렬에 유전체 자료의 비중보다 혈통자료에 대한 비중이 높게 차지하여 혈통자료를 이용하는 BLUP방법의 EBV 추정치가 더 높은 수치로 추정된다고 사료된다. 유전체 자료를 가지고 있는 한우 6,616두에 대한 도축자료로부터 추정된 육종가의 신뢰도는 암소에서 도체중, 등심단면적, 등지방두께 및 근내지방도 각각 0.48, 0.51, 0.51 및 0.59로 나타났으며 한우 KPN의 추정된 육종가 신뢰도는 0.92, 0.92, 0.93 및 0.94로 평균 40%정도 높게 추정되었다. 반면 Single Step GBLUP방법으로 추정된 유전체육종가 신뢰도는 암소에서 도체중 0.66, 등심단면적 0.68, 등지방두께 0.70, 근내지방도 0.77을 나타냈으며, 한우 KPN에서는 차례대로 0.93, 0.93, 0.94, 0.96으로 나타났다. 한우 암소의 경우 유전체육종가의 신뢰도가 육종가의 신뢰도 보다 평균 17~19% 상승되는 것을 확인하였으며, KPN의 경우 미약하지만 유전체육종가의 신뢰도가 약간 높게 나타났다. 이는 후대의 도체성적이 많이 포함된 KPN의 경우는 육종가 및 유전체육종가에서 높은 신뢰도를 가지고 있으며, 후대의 도체성적이 KPN보다 적은 암소의 경우에는 육종가보다 유전체육종가를 이용한 신뢰도가 높다는 것을 알 수 있었다.

검색어 : ssGBLUP, RPG, GEBV, Reliability, SNP-BLUP

REFERENCES

- Aguilar I., Misztal I., Johnson D.L., Legarra A., Tsuruta S., Lawlor T.J. 2010. Hot topic: a unified approach to utilize phenotypic, full pedigree, and genomic information for genetic evaluation of Holstein final score. *J. Dairy Sci.* 93, 743-752.
- Christensen O.F., Lund M.S. 2010. Genomic prediction when some animals are not genotyped. *Gen. Sel. Evol.* 42, 2.
- Christensen O.F., Madsen P., Nielsen B., Ostersen T., Su G. 2012. Single-step methods for genomic evaluation in pigs. *Animal*. Vol 6, Issue 10, p. 1565-1571.
- Dekkers J., Hospital F. 2002. The use of molecular genetics in the improvement of agricultural populations. *Nature Reviews Genetics* 3, 22-32.
- Fernando R.L., Dekkers J.C., Garrick D.J. 2014. A class of Bayesian methods to combine large numbers of genotyped and non-genotyped animals for whole-genome analyses. *Genetics selection evolution*, 46(1), p. 50.
- Fragomeni B.O., Lourenco D.A.L., Tsuruta S., Masuda Y., Aguilar I., Legarra A., Lawlor T.J., and Misztal I. 2015. Hot topic: Use of genomic single-step genomic BLUP with a large number of genotypes. *J. Dairy Sci.* (in press). doi:10.3168/jds.2014-9125.

- Harville D.A. 1997. Matrix Analysis from a Statistician's Perspective. Springer-Verlag, Berlin.
- Hayes B.J., Bowman Phillip J, Chamberlain AC, Kara Verbyla and Mike E. Goddard. 2009. Accuracy of genomic breeding values in multi-breed dairy cattle populations. Genetics selection evolution, 41(1), p. 51.
- Jairath L., Dekkers J.C.M., Schaeffer L.R., Liu Z., Burnside E.B., Kolstad B. 1998. Genetic evaluation for herd life in Canada. J. Dairy Sci. 81:550-562.
- Legarra A., Ducrocq V. 2012. Computational strategies for national integration of phenotypic, genomic, and pedigree data in a single-step best linear unbiased prediction. J. Dairy Sci. 95:4629-4645.
- Lindauer M., Hoos H., Hutter F., and Schaub T. 2015. Autofolio: An automatically configured algorithm selector. JAIR. 53:745-778
- Liu Z. 2011. Use of MACE results as input for genomic models. Interbull Bull. 43:1-4.
- Liu Z., Goddard M.E., Reinhardt F., and Reents R. 2014. A single-step genomic model with direct estimation of marker effects. J. Dairy Sci. 97:5833-5850.
- Liu Z., Goddard M.E., Hayes B.J., Reinhardt F., and Reents R. 2016. Technical note: Equivalent genomic models with a residual polygenic effect. J. Dairy Sci. 99:2016-2025.
- Masuda Y., Misztal I., Tsuruta S., Legarra A., Aguilar I., Lourenco D.A.L., Fragomeni B.O., Lawlor T.J. 2016. Implementation of genomic recursions in single-step genomic best linear unbiased predictor for US Holsteins with a large number of genotyped animals. J. Dairy Sci. 99:1968-1974.
- Meuwissen T.H. and Goddard M.E. 1996. The use of marker haplotypes in animal breeding schemes. Genetics, Selection, Evolution 28. 161-176.
- Meuwissen T.H., Hayes B.J., and Goddard M.E. 2001, Prediction of total genetic value using genome-wide dense marker maps. Genetics 157: 1819-1829.
- Misztal I., Legarra A., Aguilar I. 2009. Computing procedures for genetic evaluation including phenotypic, full pedigree, and genomic information. J Dairy Sci. 92(9):4648-55.
- Misztal I., Legarra A., and Aguilar I. 2014. Using recursion to compute the inverse of the genomic relationship matrix. J. Dairy Sci. 97:3943-3952.
- Schaeffer L.R. 2006. Strategy for applying genome-wide selection in dairy cattle. Journal of Animal Breeding and Genetics. 123(4), 218-223.
- Strandén I., and Garrick D.J. 2009. Derivation of equivalent computing algorithms for genomic predictions and reliabilities of animal merit. J. Dairy Sci. 92(6):2971-2975.
- Strandén I., Mäntysaari E.A. 2014. Comparison of some equivalent equations to solve single-step GBLUP. Proc. 10th World Congr. Genet. Appl. Livest. Prod., Vancouver, Canada: p. 069.
- Strandén I., Matilainen K., Aamand G.P., Mäntysaari E.A. 2017. Solving efficiently large single-step genomic best linear unbiased prediction models. J. Anim. Breed. Genet. 134:264-274.
- VanRaden P.M. 2008. Efficient methods to compute genomic predictions. J. Dairy Sci. 91:4414-4423.
- 손지현. 2020. 한우 유전능력 평가체계 및 육종가 정확도 개선 방안에 대한 연구. 한경대학교 박사학위 논문.