

가축 유전체정보 활용 종축 유전능력 평가 연구 - 표지인자 효과 추정 모의실험

조충일 · 이득환*

한경대학교

Study on Genetic Evaluation using Genomic Information in Animal Breeding - Simulation Study for Estimation of Marker Effects

Chungil Cho and Deukhwan Lee*

Department of Animal Life and Environment Science, Hankyong National Univ. Seokjeong-dong Anseong-si Gyeonggi-do, 456-749, Korea

ABSTRACT

This simulation study was performed to investigate the accuracy of the estimated breeding value by using genomic information (GEBV) by way of Bayesian framework. Genomic information by way of single nucleotide polymorphism (SNP) from a chromosome with length of 100cM were simulated with different marker distance (0.1cM, 0.5cM), heritabilities (0.1, 0.5) and half sibs families (20 heads, 4 heads). For generating the simulated population in which animals were inferred to genomic polymorphism, we assumed that the number of quantitative trait loci (QTL) were equal with the number of no effect markers. The positions of markers and QTLs were located with even and scatter distances, respectively. The accuracies of estimated breeding values by way of indicating correlations between true and estimated breeding values were compared on several cases of marker distances, heritabilities and family sizes. The accuracies of breeding values on animals only having genomic information were 0.87 and 0.81 in marker distances of 0.1cM and 0.5cM, respectively. These accuracies were shown to be influenced by heritabilities (0.87 at $h^2=0.10$, 0.94 at $h^2=0.50$). According to half sibs' family size, these accuracies were 0.87 and 0.84 in family size of 20 and 4, respectively. As half sibs family size is high, accuracy of breeding appeared high. Based on the results of this study it is concluded that the amount of marker information, heritability and family size would influence the accuracy of the estimated breeding values in genomic selection methodology for animal breeding.

(Key words : Genomic selection, Animal breeding, Accuracy of selection, QTL, SNP)

서 론

최근 동물에 있어 개체간 다형성이 나타나는 SNP (Single Nucleotide Polymorphisms) 표지인자를 활용하여 개체간의 유사도의 크기를 바탕으로 동물의 유전능력을 추정하는 연구가 미국, 캐나다, 호주 등 축산 선진국을 중심으로 활발히 진행되고 있다 (Calus 등 2008; Gianola와 van kaam 2008; Meuwissen 등 2001; VanRaden 등 2008; Xu 2003).

유전체 선발 (Genomic Selection, GS)은 유전력이 낮은 형질, 도체형질, 질병저항형질 등 선발의 정확도가 낮은 형질에서 유용하며 (Meuwissen 2003), 선발 후 생성되는 형질 (번식형질, 유량)에 대해 혈통정보를 활용하여 평가하는 BLUP 방법과 표지인자정보를 활용하는 MAS (Marker Assisted Selection) 방법간의 유전적 개량량 (Genetic gain)을 비교해 본 결과, 선발 후 1세대 경과되었을

때 MAS 선발이 BLUP 선발에 비해 상대적으로 우수한 유전적 개량량을 얻을 수 있는 연구보고 (Meuwissen와 Goddard 1996)가 있었다. 또한 표현형 데이터와 유전체 정보간의 연관성 규명 (표지인자효과추정) 이후 유전체 정보를 통해 동물의 신생 자축에서 조기선발이 가능해지므로 세대간격을 단축시킬 수 있는 장점이 있다.

유전체 선발의 실용화 단계로 미국 USDA (United States Department of Agriculture)는 2008년부터 젓소 유전능력 평가시 BLUP 및 유전체선발에 대한 유전평가 결과를 각각 제시하고 있다 (VanRaden 등 2008), 최근 우리나라에서도 한우의 유전체 염기서열 분석이 끝나는 등 분자유전학적 기술이 향상되고 있으나 이에 반해 유전체 정보를 활용한 동물의 유전능력 추정에 대한 통계분석 기법 연구는 아직 미흡한 실정이다.

본 연구는 유전체 정보를 활용한 동물의 유전능력 평가의 타당성을 알아보기 위하여 수행하였으며 이에 대한 연구의 일환으로 표지

* Corresponding author : Deukhwan Lee, Department of Animal Life and Environment Science, Hankyong National Univ. Seokjeong-dong Anseong-si Gyeonggi-do, 456-749, Korea. Tel: 031-670-5091, Fax: 031-676-5091, E-mail: dhlee@hknu.ac.kr

인자간 거리(Marker distance), 유전력(h^2) 및 반형매 집단크기(Half-sibs family size)에 따라 4개의 모델을 설정하였으며, 각 모델에 대해서 Meuwissen 등(2001)이 제시한 Bayes A 접근방법으로 SNP 표지인자별 효과 및 개체별 육종가를 추정하고 이에 대한 정확도를 알아보았다.

재료 및 방법

1. 모의실험 자료생성

본 연구에 이용된 모의집단은 초기집단으로서 1.0 Morgan (M) 길이의 단일 염색체를 갖는 100두의 집단을 가정하고 개체간 표지인자 및 양적형질 유전자 좌위(QTL)로 변이성이 전혀 없는 개체들로 구성하였으며 표지인자는 분석모델에 따라 0.1cM 및 0.5cM 간격으로 균등하게 위치시켰고, QTL은 표지인자와 동수의 비율로 임의 위치시켰다(총 QTL 및 표지인자 수는 0.1cM와 0.5cM에서 각각 1000개 및 200개 생성). 초기집단으로부터 임의 교배로 얻어진 자손에 대해 수컷과 암컷을 동수의 확률로 세대당 100두씩 생성하여 1000세대까지 집단성장을 발생시켰다. 세대가 성장함에 따라 배우자 형성과정에서 단일염기의 돌연변이를 발생시켜 단일염기 다형(SNP)을 생성하였는데, 여기서 돌연변이 발생율은 표현형 관측치와 연관관계가 없는 표지인자는 2.5×10^{-3} , 양적형질의 표현형과 연관관계가 있는 좌위(QTL)는 2.5×10^{-5} 의 확률로 돌연변이를 발생시켰다. 또한 배우자(gamete) 형성과정에서 동일 염색체 내에 표지인자 또는 QTL 간에 Haldane의 확률(Haldane 1919)로 교차현상을 발생시켰다. QTL에서 돌연변이가 발생할 때, Gamma 분포함수(scale parameter = 1.66, shape parameter = 0.4)를 이용하여 표현형에 영향을 미치는 유전자의 효과를 생성하였으며 유전자 효과의 방향은 동수의 확률로 정의하였다(Table 1). 이 후 1001~1002세대(pheno)는 각 세대당 2,000두의 개체 유전효과 및 표현형 자료를 생성하여 각 표지인자의 유전적 효과를 추정하기 위한 참조집단(Reference Population, RP)을 구성하였다. 여기서 RP의 개체별 표현형 자료생성은 주어진 유전력에 기준하여 단일 염색체 내의 QTL 효과의 총합인 참육종가(TBV)와 임의환경요인으로 가정한 평균이 "0"인 정규분포 집단으로부터 임의의 잔차를 생성하여 이들의 합으로 생성하였다.

1003세대(Juven1)와 1004세대(Juven2)는 오직 유전체자료만 갖는 개체로 구성된 검정집단(Test Population, TP)으로 가정하여 RP에서 추정된 각 표지인자 효과를 바탕으로 TP의 추정 육종가를 계산하였으며, 표지인자간 거리, 유전력, 반형매 집단크기에 따라 4개의 모델을 설정하여 비교, 분석하였다. Model I은 유전력 0.1, 표지인자간 거리 0.1cM, 1001~1004세대의 부계에 대한 평균 반형매 집단크기 20두(종모우 100두, 종빈우 1000두, 종빈우당 자손수 2두로 총 2000두/세대 생성)을 가정한 모형이고, Model II은 유전력 0.1, 표지인자간 거리 0.5cM, 반형매 집단크기 20두, Model III은 유전력 0.5, 표지인자간 거리 0.1cM, 반형매 집단크기

Table 1. The parameters commonly assumed in several simulation genetic models

	Parameters
Number of chromosome	1
Chromosome length	100cM
Effective population size	100
Map scan	Markers-even distance QTL-scatter distance
Mutation rate of Marker loci	2.5×10^{-3}
Mutation rate of QTL	2.5×10^{-5}
Distribution of additive mutational effects	Gamma (Scale = 1.66, Shape = 0.4)
Recombination fraction (Haldane's map function)	$R=(1-\exp(-2d))/2$
Population size/generation (generation 1~1000)	100
Population size/generation (generation 1001~1004)	2000
Mating scheme	Random mating
Total generation	1004
Phenotypic and Genomic record	1001 ~ 1002 gen.
Only Genomic recording	1003 ~ 1004 gen.

20두, Model IV은 유전력 0.1, 표지인자간 거리 0.1cM, 1001~1004세대의 반형매 집단크기 4두(종모우 500두, 종빈우 1000두, 종빈우당 자손수 2두로 총 2000두/세대 생성) 등을 가정한 모형으로 설정하여 구분하였다(Table 2).

2. 분석방법

(1) Bayesian 추정방법

유전체 선별의 분석방법은 RP에서 표지인자와 표현형 자료를 이용한 표지인자 효과 추정 단계와 추정된 표지인자를 TP에 적용시켜 육종가를 추정하는 단계로 구분하여 분석을 실시하였다.

1) 표지인자효과 분석

4000두의 RP에 표지인자 효과추정을 위한 모형식은 아래와 같이 설정하였다.

$$y_i = 1'_n \mu + \sum_j X_{ij} g_j + e_i$$

여기서 y_i 는 i 번째 개체의 표현형 값, $1'_n$ 는 원소가 1인 n 개의 행 벡터, μ 는 전체평균, X_{ij} 는 i 번째 개체의 j 번째 좌위에서 Major Allele의 수 (0, 1, 2), g_j 는 j 번째 표지인자효과, e_i 는 i 번째 개체의 오차이다.

상기의 분석모형에서 모수에 대한 사전분포는 다음과 같이 가정

Table 2. Information of parameters assumed on each genetic model in simulation study

Model	Population size (1001~1004 generation)				Heritability	Marker & QTL distance (cM)
	No. Sire	No. Dam per Sire	No. Progeny per dam	Half-sibs family size (head)		
Model I	100	1000	2	20	0.1	0.1
Model II	100	1000	2	20	0.1	0.5
Model III	100	1000	2	20	0.5	0.1
Model IV	500	1000	2	4	0.1	0.1

하였다.

$$\mu \propto \text{constanct}$$

$$g_i \sim N(0, \sigma_g^2)$$

$$e_j \sim N(0, \sigma_e^2)$$

상기 모형의 모수를 추정하기 위하여 Meuwissen 등 (2001)이 제시한 Bayes A의 방법으로 아래와 같이 사후분포 확률로부터 표본의 모수를 깁스샘플(Gibbs sampling)하였다(Meuwissen 등 2001; Hayes 2008).

Step 1. 표지인자효과인 g 와 전체평균 μ 에 대해서 임의의 초기값을 설정한다.

여기서 g 는 Random effect, μ 는 Fixed effect

Step 2. Inverse chi-square(X^{-2}) 분포함수로부터 오차분산(σ_e^2)을 샘플한다

$$\sigma_e^2 \sim \chi^{-2}(n-2, e'e)$$

여기서, $e = y - Xg - 1_n' \mu$. X 는 표지인자효과 대한 계수행렬, n 은 기록의 수이다.

Step 3. 아래의 정규분포함수로부터 전체평균 μ 를 샘플한다.

$$\mu \sim N\left(\frac{1}{n}(1_n'y - 1_n'Xg), \frac{\sigma_e^2}{n}\right)$$

Step 4. Inverse chi-square (X^{-2}) 분포함수로부터 각각의 표지인자에 대한 유전분산($\sigma_{g_i}^2$)을 샘플한다.

$$\sigma_{g_i}^2 \sim \chi^{-2}(v+n_j, S+g_j'g_j)$$

여기서, $v=4.012$, $S=0.002$. n_i 는 i 번째 좌위의 대립유전자 수에 하나를 감한 값이다.

Step 5. 아래의 정규분포함수로부터 각각의 표지인자효과(g_i)를 샘플한다.

$$g_i \sim N\left(\frac{X_j'y - X_j'Xg_{(-j)} - X_j'1_n\mu}{X_j'X + \frac{\sigma_e^2}{\sigma_{g_i}^2}}, \frac{\sigma_e^2}{X_j'X + \frac{\sigma_e^2}{\sigma_{g_i}^2}}\right)$$

여기서, X_j' 은 X 행렬의 전치행렬에서 j 번째 행벡터이고, $g_{(-j)}$ 는 j 번째 표지인자를 제외한 나머지 표지인자에 대해 샘플한 값의 열벡터이다.

Step 2에서 Step 5의 과정을 10,000회 반복하여 표본에서 모수를 샘플한 후, 초기 5,000회는 burn-in 구간으로 간주하여 제거하

였으며, 나머지 표본추출 값에 대한 평균값을 각 모수($\mu, g_i, \sigma_{g_i}^2, \sigma_e^2$)의 추정치로 간주하였다.

2) 육종가 추정

RP에서 추정된 표지인자효과(\hat{g})를 이용하여 조기선발 집단인 TP에서 i 번째 개체의 육종가를 다음과 같이 추정하였다.

$$BV_i = X_{ij} \hat{g}_j$$

여기서, BV_i 는 i 번째 개체의 추정 육종가, X_{ij} 는 i 번째 개체의 j 번째 좌위에 SNP 유전자형 표지인자효과(0, 1, 2)의 계수행렬 (design matrix); \hat{g}_j 는 j 번째 좌위에서 추정된 표지인자효과이다.

TP에서 추정된 육종가(EBV)와 참육종가(TBV)간의 상관분석을 통하여 유전평가분석의 정확도($r(TBV,EBV)$)를 계산하였으며, 본 모의실험은 각 모형별 모의실험자료 및 분석을 20회 실시하여 기초통계량을 계산하였다(Table 3).

(2) BLUP 추정방법

유전력이 낮은 형질에서 유전체를 활용한 Bayes A 방법과 혈통기반의 BLUP 추정간 유전분석의 정확도를 비교하고자 아래와 같이 개체 모형설정 후 BLUP 추정하였다.

$$y_i = 1_n' \mu + a_i + e_i$$

여기서 y_i 는 i 번째 개체의 표현형 값, $1_n'$ 는 원소가 1인 n 개의 벡터, μ 는 전체평균, a_i 는 i 번째 개체효과, e_i 는 i 번째 개체의 오차이다.

상기의 개체모형을 바탕으로 저도의 유전력($h^2=0.1$)을 갖는 Model I의 모의실험 자료를 이용하여 BLUP 추정방법에 의한 개체별 육종가 추정치의 정확도를 제시하였다(Table 3).

결과 및 고찰

표현형자료를 갖는 Pheno 세대는 참조집단(RP)으로 표현형자료와 유전체자료를 통해 각 표지인자별 효과를 추정하였으며, 양적형질 유전자좌위(QTL)의 참 효과값과의 상관관계를 통하여 정확도를 제시하였다. Juven1 및 Juven 2 세대는 도축되지 않고 생존해

Table 3. Correlation coefficients between the estimated and true breeding values on generation of Pheno, Juven 1 and Juven 2 in the simulated population

Population	Bayes A				4) BLUP in Model I
	Model I	Model II	Model III	Model IV	
1) Pheno	0.88 ± 0.05	0.85 ± 0.09	0.95 ± 0.02	0.86 ± 0.04	0.46 ± 0.12
2) Juven 1	0.87 ± 0.05	0.81 ± 0.11	0.94 ± 0.03	0.84 ± 0.05	0.31 ± 0.09
3) Juven 2	0.86 ± 0.04	0.80 ± 0.09	0.93 ± 0.03	0.83 ± 0.05	0.25 ± 0.09

- 1) Pheno: Parental animals have phenotype data and genotype data.
- 2) Juven 1: Juvenile animals have only genotype data.
- 3) Juven 2: Progenies of Juven1 have only genotype data.
- 4) BLUP: BLUP was based on Animal model considering phenotype and pedigree.

있는 개체로 유전체자료만 갖는 검정집단(TP)이며, 이들의 각 표지인자별 효과에 대한 추정치들의 합으로 표현된 추정 육종가는 RP에서 추정된 표지인자 효과값의 벡터와 SNP 유전자형 계수행렬 (design matrix)을 통해 계산하였다.

Table 3은 각 모형별 20번 반복 모의실험한 육종가 추정치들에 대하여 정확도 및 몬테카를로 표준편차(Monte Carlo Standard deviations, MS)를 나타냈다. 여기서 Model I, III, IV는 반복 모의실험간 편차(MS)가 비교적 적게 나타났음에 반해 Model II는 반복 모의실험간 편차가 높은 것으로 나타났다. 이는 표지인자간 거리가 다른 모델에 비해 조밀하지 못하여 발생한 편차로 사료되며 보다 정확한 분석을 위해 모의실험간 반복수를 높여 분석해야 될 것으로 사료되었다.

1. 표지인자간 거리에 따른 육종가 정확도 비교

Model I 과 Model II간의 비교로 유전력 0.1, 반형매 집단크기 20두 일 때 표지인자간 거리(0.1cM, 0.5cM)에 따른 육종가 정확도에 대해 Fig. 1에 제시하였다.

표지인자간 거리가 고밀도인(0.1cM) 경우 Pheno, Juven1 및 Juven 2세대의 육종가의 정확도는 각각 0.88, 0.87, 0.86였고, 반면 표지인자간 거리가 저밀도인(0.5cM) 경우 육종가의 정확도는 각각 0.85, 0.81, 0.80로 표지인자간 거리가 조밀할수록 육종가의 정확도가 높아짐을 알 수 있었다.

Meuwissen 등(2001)은 동일크기의 염색체 10개, 각 염색체의 길이는 100cM, 유전력은 0.5, 표지인자간 거리는 1cM 간격이며 유효집단의 크기가 100두로 설정한 모의실험에서 참육종가(QTL)와 추정 육종가(표지인자효과) 간의 정확도를 분석한 결과, 표현형 자료가 없는 Juven1 세대에서 Bayes A 방법으로 분석하였을 때 육종가의 정확도는 0.80로 나타났다. 또한 Solberg 등(2008)은 동일크기의 염색체 10개, 전장의 유전체길이는 10Morgan, 유전력이 0.5로 설정한 모의실험에서 참육종가와 추정 육종가간의 정확도를 알아보기 위하여 연관불평형(Linkage Disequilibrium)의 크기(r^2)가 유효집단의 크기(Ne)와 표지인자간 재조합률(c)의 함수로 표현

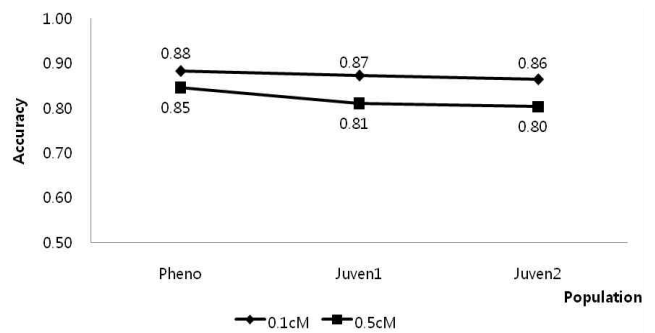


Fig. 1. Accuracy of breeding value for a heritability of 0.1, half-sibs family size of 20 and with different marker distance (0.1cM, 0.5cM).

($E(r^2) = 1/(1 + 4N_e c)$; Sved, 1971) 될 수 있기 때문에 유효집단을 고려한 표지인자간 거리를 N_e/M (Morgan)으로 표현할 때, 1 N_e/M 에서 0.69, 2 N_e/M 일 때 0.79, 4 N_e/M 일 때 0.84 및 8 N_e/M 일 때 0.86로 고정된 유효집단에서 고밀도 표지인자일 수록 정확도가 향상되는 것으로 보고하였다. 본 연구의 0.5cM 표지인자 거리에서 정확도가 0.81로 Solberg 등(2008)의 0.79 보다 조금 높게 추정되었는데, 이는 염색체수가 늘어남에 따라 오차율이 증가해서 나타난 결과로 해석된다.

2. 유전력별 육종가 정확도 비교

Model I 과 Model III 간의 비교로 표지인자간 거리 0.1cM, 반형매 집단크기 20두 일 때 유전력(0.1, 0.5)에 따른 육종가 정확도에 대해 Fig. 2에 제시하였다.

유전력이 0.1일 때 Pheno, Juven 1 및 Juven 2 세대의 육종가의 정확도는 각각 0.88, 0.87, 0.86 였고, 유전력이 0.5인 경우 육종가의 정확도는 각각 0.95, 0.94, 0.93로 유전력이 높을수록 육종가의 정확도는 높게 나타났다.

Nielsen 등(2009)은 Meuwissen 등(2001)이 제시한 GBLUP 방법으로 표지인자간 거리가 $0.5N_e / \text{Morgan}$ 이며, 표현형데이터가

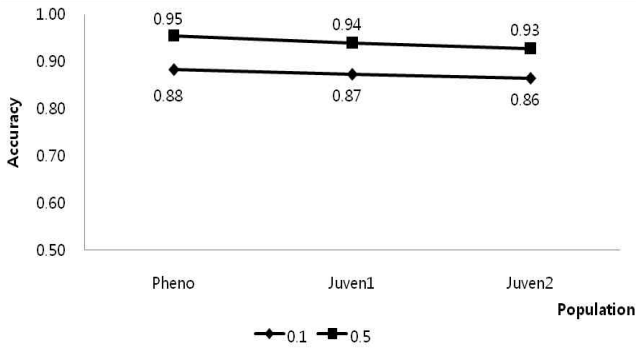


Fig. 2. Accuracy of breeding value for a marker distance of 0.1cM, half-sibs family size of 20 and with different heritability (0.1, 0.5).

없는 G2 세대에서 유전력 0.2와 0.4일 때 정확도는 각각 0.71과 0.75로 보고하여 유전평가의 정확도가 본 연구 보다 다소 낮게 나타났는데 이는 추정방법간 차이에서 기인한 부분으로 사료되며 (Meuwissen 등 2001) Nielsen 등 (2009)의 연구도 본 연구와 동일하게 유전력이 높을 때 정확도가 높게 나타났다. 또한 본 연구의 유전분석 결과, 낮은 유전력 ($h^2=0.1$)에서도 0.86 이상의 높은 정확도를 보였는데, 이러한 결과로 살펴볼 때 저도의 유전력을 갖는 형질에서 유전체 선발을 실시할 경우 개량의 효율이 향상될 것으로 사료되었다 (Meuwissen, 2003).

3. 분석방법간 육종가 정확도 비교

Model I에서 생성된 모의실험데이터를 활용하여 Bayes A와 BLUP 방법간의 정확도를 비교하였다 (Table 3). Bayes A 추정시 Pheno, Juven 1 및 Juven 2 세대의 육종가의 정확도는 각각 0.88, 0.87, 0.86 인데 반해 BLUP 추정시 0.46, 0.31 및 0.25로 매우 낮은 정확도를 보였다. 이는 유전체 분석의 우수성을 나타내는 부분이며, 특히 BLUP 추정의 경우 표현형기록이 없는 Juven 세대 에서 유전분석의 정확도가 급격히 낮아지는 반면 유전체정보를 활용한 Bayes A의 정확도는 0.86 이상으로 높게 나타나 육성기 및 수 정란 단계에서의 조기선발이 가능함을 나타낸다고 할 수 있다.

4. 반형매 집단크기별 육종가 정확도 비교

Model I 과 Model IV간의 비교로 표지인자간 거리 0.1cM, 유전력 0.1일 때 부계에 대한 평균 반형매 집단크기(20두, 4두)에 따른 육종가 정확도에 대해 Fig. 3에 제시하였다.

1001~1004세대의 반형매 집단크기가 20두일 때 Pheno, Juven 1 및 Juven 2 세대의 육종가의 정확도는 각각 0.88, 0.87, 0.86 였고, 반형매 집단크기가 4두인 경우 육종가의 정확도는 각각 0.86, 0.84, 0.83로 반형매 집단크기가 클수록 육종가의 정확도는 높게 나타났다. 이는 개체간의 유사도에 의해 유전체분석의 정확도가 변화되었음을 의미하며, 유전체분석을 통한 현장적용시 특정집단에서

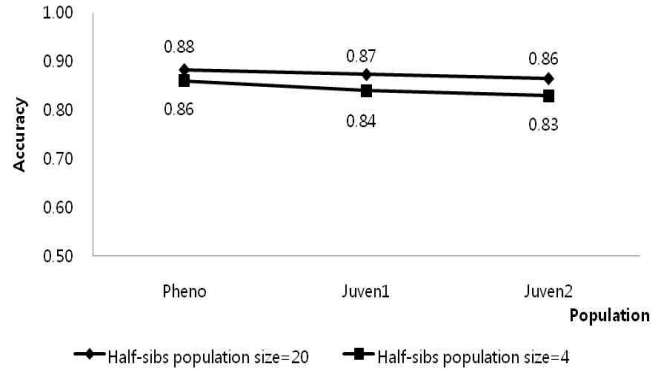


Fig. 3. Accuracy of breeding value for a marker distance of 0.1cM, heritability of 0.1 and with different half-sibs family size (20 heads, 4 heads).

분석한 표지인자효과를 다른 집단에 적용할 경우 육종가의 정확도가 떨어질 수 있을 것으로 사료되었다.

Hayes 등 (2008)은 유전체 선발기술은 대가축의 개량에 있어서 혁명적인 신기술이라고 표현하고 있다. 하지만 현재 유전체 분석을 위한 고가의 Chip 비용으로 산업화 하는데 많은 비용이 소요되지만 금후 분석비용의 가격하락이 예상되고 종축의 가치를 고려할 때 산업적 가치가 매우 클 것으로 예상된다. 유전체 선발을 한우 종축 개량체계에 실용화하기 위해서는 Reference Population으로 약 2,000두 이상의 유전체 및 표현형자료가 필요하며 (Goddard, 2008). 유전력을 0.5로 가정하였을 때 유전평가의 정확도는 0.5~0.8로 추정된다 (Goddard, 2008). 유전체 정보를 이용한 종축 선발의 방법에서 고려해야 할 것은 세대가 지남에 따라 집단의 유전변이가 변화하기 때문에 유전체와 표현형과의 관련성을 비교할 수 있는 능력검정 자료를 지속적으로 생성해야 할 것이며, 또한 유전체선발로 인한 특정유전자 선호로 유전적 부동 (Genetic Drift) 현상이 발생할 수 있으므로 추가적인 SNP 발굴이 필요할 것 (Meuwissen 2003)으로 사료되었다.

요 약

본 연구는 유전체분석에 대해 모의실험한 연구로서 Reference Population (RP)이 구성되었을 때, 표현형 자료가 없고 유전체자료만 있는 Juven 1 또는 Juven 2 세대에 대해 유전평가의 정확도에 대해 알아보고자 연구를 실시하였다. 모의실험의 가정으로 염색체는 1개이며 염색체길이는 100cM로 가정하였다. 초기의 유효집단의 수는 100두의 다형성이 없는 초기집단에서 유전자 효과가 없는 표지인자 (Marker)를 0.1cM 및 0.5cM 간격으로 균등하게 단일 염기 돌연변이에 의한 다형성을 발생시켰고 유전자 효과가 있는 QTL 좌위는 Marker와 동수의 비율로 임의위치를 지정하여 돌연변이에 의한 변이성을 생성하였으며 이때 유전자 효과는 Gamma 분포함수 (scale=1.66, shape=0.4)에서 생성하였다. 배우자 (gamete) 형성과정에서 Haldane의 가정에 유전자 재조합을 생성하였으며 돌연변이 발생율은 Marker 및 QTL 좌위에서 2.5×10^{-3} 및

2.5×10^{-5} 의 확률로 발생시켜 1000세대까지 세대번식을 유지하였다. 이 후 1001세대부터 1004세대까지 세대당 2000두의 자손을 생성하였으며 이 때 유전력을 0.1 및 0.5의 가정하에 1001~1002세대에서 표현형 자료를 생성하였고, 1003~1004세대는 오직 유전체자료만 생성하였다. Bayesian 방법을 이용하여 개체별 육종가를 추정하였으며 표지인자간 거리(0.1cM, 0.5cM), 유전력(0.1, 0.5) 및 반형매 집단크기(20두, 4두)에 따라 참육종가와 추정 육종가간의 상관으로 표현되는 육종가 정확도에 대해 비교한 결과 1003세대에서 표지인자간 거리가 0.1cM 및 0.5cM일 때 육종가의 정확도는 각각 0.87, 0.81였고, 유전력이 0.1 및 0.5 일 때 각각 0.87, 0.94로 추정되었으며, 반형매 집단의 크기가 20두 일 때 0.87, 4두 일 때 0.84로 추정되었다.

위의 결과로 미루어 보아 다량의 SNP 표지정보 및 반형매 집단의 크기가 클수록 즉, 혈연계수가 높은 집단일 때 육종가의 정확도는 높게 나타났다. 유전체선발의 활용시 비교적 높은 정확도로써 조기선발이 가능하며 이로 인한 세대간격을 단축시킬 수 있어 개량의 효율을 높일 수 있을 것으로 사료된다. 반면에 유전체선발은 분석비용이 비싸며, 지속적인 유전체 선발시 특정유전자 선호로 인한 유전적 부동(Genetic Drift) 현상이 발생할 수 있기 때문에 지속적인 SNP 발굴에 대한 노력이 필요한(Meuwissen 2003) 단점이 있으나 한우 또는 젖소와 같은 대가축과 같이 세대간격이 긴 가축에서 유전체선발 할 경우 조기선발로 인한 세대간격 단축과 유전평가의 높은 정확도(0.8 이상)로 인해 개량의 효율을 극대화 할 수 있을 것으로 사료된다.

(주제어: 유전체 선발, 동물육종, 정확도, QTL, SNP)

사 사

본 연구는 2010년 농림기술행과제의 일환으로 수행하였음.

인 용 문 헌

Calus, M. P. L., Meuwissen, T. H. E., de Roos, A. P. W. and Veerkamp, R. F. 2008. Accuracy of genomic selection using different methods to define haplotypes. *Genetics* 178:553-561.

Gianola, D. and van Kaam, J. B. C. H. M. 2008. Reproducing kernel hilbert Spaces regression methods for genomic assisted prediction of quantitative traits. *Genetics* 178:2289-2303.

Goddard, M. E. 2008. Genomic selection: Prediction of accuracy and maximization of long term response. *Genetica*. Doi:10.1007/s10709-008-9308-0.

Hayes, B. J. 2008. Course Notes 'QTL Mapping, Mas, and Genomic Selection'.

Hayes, B. J. Bowman, P. J. Chamberlain, A. J. and Goddard, M. E. 2009. Invited review: Genomic selection in dairy cattle: Progress and challenges. *J Dairy Sci*, 92(2):433-443.

Haldane, J. B. S. 1919. The combination of linkage values and the calculation of distances between the loci of linked factors. *J Genet* 8:299-309.

Meuwissen, T. H. E. and Goddard, M. E. 1996. The use of marker haplotypes in animal breeding schemes. *Genetics Selection Evolution* 28:161-176.

Meuwissen, T. H. E., Hayes, B. J. and Goddard, M. E. 2001. Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157:1819-1829.

Meuwissen, T. H. E. 2003. 'Genomic selection: The future of marker assisted selection and animal breeding'.

Nielsen, H. M. 2009. Comparison of accuracy of genome-wide and BLUP breeding value estimates in sib based aquaculture breeding schemes. *Aquaculture* 289(2009) 259-264.

Sved, J. A., 1971 Linkage disequilibrium and homozygosity of chromosome segments in finite population. *Theor. Popul. Biol.* 2:125-141.

Solberg, T. R., Sonesson, A. K., Wooliams, J. A. and Meuwissen, T. H. E. 2008. Genomic selection using different marker types and densities. *J. Anim. Sci.* 86:2447-2454 .

VanRaden, P. M., Van Tassell, C. P., Wiggans, G. R., Sonstegard, T. S., Schnabel, R. D. and Taylor, J. F. 2008. Invited review: Reliability of genomic predictions for North American Holstein bulls. *J. Dairy. Sci.* 92:16-24.

VanRaden, P. M. 2008. Efficient methods to compute genomic predictions. *J. Dairy. Sci.* 91:4414-4423.

Xu, S. Z., 2003 Estimating polygenic effects using markers of the entire genome. *Genetics* 163:789-801.

(Received Jun. 9, 2010; Revised Oct. 13, 2010; Accepted Oct. 20, 2010)