

제주재래흑돼지와 랜드레이스 F₂ 교배축군의 생체중에 대한 유전체와 가계도 기반의 유전력 및 모체효과 추정

박희복^{1,*}, 한상현^{2,*}, 이재봉³, 김상금¹, 강용준¹, 신현숙¹, 신상민¹, 김지향¹, 손준규¹, 백광수¹, 조상래⁴, 조인철¹

¹국립축산과학원 난지축산연구소 ²제주대학교 교육과학연구소 ³전북대학교 인수공통전염병연구소

⁴국립축산과학원 한우연구소

SNP-based and pedigree-based estimation of heritability and maternal effect for body weight traits in an F₂ intercross between Landrace and Jeju native black pigs

Hee-Bok Park^{1,*}, Sang-Hyun Han^{2,*}, Jae-Bong Lee³, Sang-Geum Kim¹, Yong-Jun Kang¹, Hyun-Sook Shin¹, Sang-Min Shin¹, Ji-Hyang Kim¹, Jun-Kyu Son¹, Kwang-Soo Baek¹, Sang-Rae Cho⁴ and In-Cheol Cho^{1,†}

¹Subtropical Livestock Research Institute, National Institute of Animal Science, RDA, 63242 Jeju, Republic of Korea.

²Educational Science Research Institute, Jeju National University, 63243 Jeju, Republic of Korea.

³Korea Zoonosis Research Institute, Chonbuk National University, 54531 Iksan, Republic of Korea.

⁴Hanwoo Research Institute, National Institute of Animal Science, RDA, 25340

ABSTRACT

Growth traits, such as body weight, directly influence productivity and economic efficiency in the swine industry. In this study, we estimate heritability for body weight traits using information from pedigree and genome-wide single nucleotide polymorphism (SNP) chip data. Four body weight phenotypes were measured in 1,105 F₂ progeny from an intercross between Landrace and Jeju native black pigs. All experimental animals were subjected to genotypic analysis using PorcineSNP60K BeadChip platform, and 39,992 autosomal SNP markers filtered by quality control criteria were used to construct genomic relationship matrix for heritability estimation. Restricted maximum likelihood estimates of heritability were obtained using both genomic- and pedigree- relationship matrix in a linear mixed model. The heritability estimates using SNP information were smaller (0.36-0.55) than those which were estimated using pedigree information (0.62-0.97). To investigate effect of common environment, such as maternal effect, on heritability estimation, we included maternal effect as an additional random effect term in the linear mixed model analysis. We detected substantial proportions of phenotypic variance components were explained by maternal effect. And the heritability estimates using both pedigree and SNP information were decreased. Therefore, heritability estimates must be interpreted cautiously when there are obvious common environmental variance components.

(Key word: growth, heritability, genomic relationship matrix, pedigree, maternal effect, pig)

서론

포유동물에서 체중과 같은 성장은 다양한 유전적 요인과 환경적 요인들이 고도로 정밀하게 상호작용하여 발현되는 복잡하고(complex)하고 양적인(quantitative) 생명현상이다(Spencer, 1985). 따라서, 체중형질과 같은 양적인 형질은 전체변이를 나타내는 표현형적 분산성분(phenotypic variance component)은 유전적 분산성분과 환경적 분산성분으로 분할할 수 있으며, 이 두 분산성분을 추정을 통해 유전력(heritability)을 구할 수 있다

(Lee, 2000). 축산업에서 성장형질의 유전적 능력향상은 경제적 효율성과 생산성에 직결되어 있다. 이는 성장형질에 대한 유전력 추정은 축군의 유전적 개선에 필수적인 중요한 단계라 할 수 있다.

전통적으로 선발육종을 위한 유전력의 추정은 기본적으로 가계(pedigree)를 이용하여, 그 가계구성원간의 수리적 관계를 이용하여 구축한 수치기반 상가적(additive) 혈연관계 행렬(numeric relationship matrix, NRM)이 사용되어 왔다. 하지만, 최근에는 단일염기서열(single nucleotide polymorphism, SNP)

* Hee-Bok Park and Sang-Hyun Han contributed equally to this study.

† Corresponding author: Dr. In-Cheol Cho

Tel: +82-64-754-5710

E-mail: choic4753@korea.kr

표지를 고밀도로 집적한 chip 기반의 대용량 유전자형 결정법이 일반화 되면서, 집단(population) 구성원 간 유전적 관계를 SNP 표지정보를 이용하여 SNP-기반 유전체 관계 행렬(genomic relationship matrix, GRM)을 구축할 수 있는 방법들이 개발되었고, GRM은 유전력 추정을 위한 혼합모형분석에 있어서 NRM을 대치하여 이용하기도 하였다(VanRaden, 2008; Yang 등 2011b).

최근 국민소득의 증대와 생활수준의 향상은 돈육소비 패턴의 변화를 가져왔다. 즉, 양적인 소비에서 맛으로 대표되는 질적인 소비로 전환되고 있으며 이와 함께 양돈 산업에서 제주 재래흑돼지(Jezu native black pig, JBP)에 대한 관심이 높아지고 있다. JBP는 제주도에서 길러져 오던 재래종으로써 랜드레이스나 요크셔와 같은 외래종에 비하여 체구가 작고 산자수가 적으며, 성장이 느린 단점이 있으나 백색의 단단한 지방과 적색의 육색, 그리고 우수한 상강도(marbling degree)를 보인다(Kim 등, 2009). 1980년대 말 시작된 JBP 복원 및 보존사업으로 순종의 보호와 육성이 진행되어 왔으나 성장형질에 관련된 단점들로 인해 순종 JBP의 산업화는 미흡한 실정이다. 2006년 JBP를 소재로 이용한 본격적인 산업화를 위해 장점인 육질형질을 보존하면서 단점인 성장형질을 분자유전학적 방법을 통해 개량하고자 하는 연구들이 시작되었다. JBP 품종과 랜드레이스 품종 간 교배축군이 조성되어 육질과 성장형질에 연관된 양적형질좌위(quantitative trait loci, QTLs)를 동정해 낼 수가 있었다(Cho 등, 2011, 2015; Yoo 등, 2014).

본 연구는 고 육질형 돼지 품종인 JBP와 고산육형 품종인 랜드레이스 품종간 교배축군(F₂세대)에서 얻어진 체중형질에 모체효과(maternal effect, m^2)에 대하여 SNP 기반과 가계 기반 유전력을 추정하고 이를 비교 하였다.

재료 및 방법

공시 동물

본 연구에 사용된 JBP와 랜드레이스의 상호교차교배(reciprocal intercross) F₂ 집단은 농촌진흥청 국립축산과학원 난지축산연구소에서 생산되고 육성되었다. F₀ 돼지들은 랜드레이스가 17 두, JBP가 19 두가 이용되었으며, 교배를 통하여 F₁ 91 두의 상호교차교배를 통하여 생산된 F₂ 자손 중 총 1,105 두(암 537, 수 568)를 표현형 수집과 유전자형 결정에 이용되었다. 구체적인 공시동물의 생산 선발, 표현형 측정방법은 Cho 등의 보고(2011)에 기술되어 있다.

분석 자료

공시동물에서 측정된 생시체중(body weight at birth, BW_{bir}),

21일령체중(body weight at 21 days after birth, BW₂₁), 70일령체중(body weight at 70 days after birth, BW₇₀), 140일령체중(body weight at 140 days after birth, BW₁₄₀)을 연구에서 이용하였다. SNP 표지의 유전자형 분석은 PorcineSNP60K BeadChip (Illumina Inc., USA)을 이용하여 수행하였다. 결정된 SNP 표지의 유전자형에 대하여 minor allele frequency(MAF)가 0.01 이하, 유전자형 결측 비율이 5% 이상, 하디-바인베르크 평형(Hardy-Weinberg equilibrium)을 극도로 벗어 나는 경우 ($P < 0.000001$)를 제외한 돼지 상염색체 상의 SNP 39,992개의 분석결과를 이용하였다. 이와 같은 quality control 작업은 PLINK 1.07 프로그램(Purcell 등, 2007)을 사용하였다.

통계 분석

측정 형질들에 대한 유전력을 추정하기 선형혼합모형은 다음과 같다:

$$y = Xb + Zu + e \quad (\text{Model 1})$$

여기서,

y 는 표현형 관측치에 대한 벡터이고 b 는 전체평균, 성별, 그리고 배치의 고정효과 벡터이며, X 는 관측치에 각 고정효과를 연관시키는 빈도행렬(incidence matrix)이며, u 는 개체의 상가적 임의 유전효과 벡터이며, Z 는 u 에 연관이 된 빈도행렬이며, e 는 임의 잔차효과 벡터이다. e 에 대한 평균과 분산은 $e \sim N(0, I\alpha^2)$ 로 정의하였으며, u 에 대한 평균과 분산은 $u \sim N(0, A\alpha^2)$ 또는 $u \sim N(0, G\alpha^2)$ 과 같이 정의하였다. 여기서 I 는 항등행렬(identity matrix)이고, A 는 혈연관계 행렬(NRM)이며(Henderson, 1984), G 는 유전체 관계행렬(GRM)을 나타낸다.

G 를 구축하기 위해 SNP 표지의 유전자형과 대립 유전자의 빈도를 이용하는 아래와 같은 행렬 추정식이 이용되었다(VanRaden, 2008; Yang 등, 2011a):

$$G = \frac{(M - 2P)(M - 2P)'}{2 \sum_i^m P_i(1 - P_i)}$$

여기서,

M 은 행렬로서 사용된 SNP 표지의 수를 m 개라하고, 개체수를 n 이라 할 때, 그 크기는 $n \times m$ 으로 설정된다. M 행렬의 각 원소는 각 SNP 표지에 대해서 다수대립유전자(major allele)의 동형접합체는 1, 소수대립유전자(minor allele)의 동형접합체는 -1, 그리고 이형접합체는 0으로 하여 SNP 유전자형을 지정하였다. 따라서 MM' 행렬의 대각선(diagonal)의 경우 각 원소들은 각 개체의 동형접합체의 개수이며, 비대각선(off-diagonal)의 경우는 집단 내 두 개체 간에 공유하는 대립유전자의 수이다.

P행렬의 각 원소는 *i*번째 SNP 표지의 소수대립유전자의 유전자빈도와 기준빈도 0.5의차에 2를 곱해 준 것이다. P_i 는 *i*번째 SNP 표지의 소수대립인자의 유전자빈도이다. GRM 구축에는 Yang등(2011a)이 개발한 GCTA (Genome-wide Complex Trait Analysis) 프로그램을 이용 하였다

체중 형질들의 유전력을 추정에 있어서 모돈효과를 추정하기 위한 선형혼합모형은 다음과 같다

$$y = Xb + Z_1u + Z_2m + e \quad (\text{Model 2})$$

여기서,

y 는 표현형 관측치에 대한 벡터이고 b 는 전체평균, 성별, 그리고 배치의 고정효과벡터이며, X 는 관측치에 각 고정 효과를 연관시키는 빈도 행렬이며 u 는 개체의 상가적 임의유전효과벡터이며, Z_1 는 u 에 연관된 빈도행렬이며 m 은 F₁ 모돈의 임의효과 벡터이며, Z_2 는 m 에 연관된 빈도행렬이다. e 는 임의 산차효과 벡터이다. e 에 대한 평균과 분산은 $e \sim N(0, I\alpha_e^2)$ 로 정의 하였으며, u 에 대한 평균과 분산은 $u \sim N(0, A\alpha_u^2)$ 또는 $u \sim N(0, G\alpha_u^2)$ 과 같이 정의 하였다. m 에 대한 평균과분산은 $m \sim N(0, I\alpha_m^2)$ 로 정의 하였으며, I 는 모체 임의효과에 대한 행렬을 의미한다.

α_u^2 (유전분산성분)과 α_m^2 (모돈분산성분), 그리고 α_e^2 (환경분산성분)의 추정에는 restricted maximum likelihood(REML) 방법에 기반 한 ASReml 프로그램(VSN international, UK)을 이용 하였다. 모형 1을 이용해서 얻은 분산성분추정을 위한 결과를 이용해서는 아래의 식을 이용하여 유전력을 계산하였다

$$h_{\text{model1}}^2 = \sigma_a^2 / \sigma_a^2 + \sigma_e^2$$

모형 2을 이용해서 얻은 분산성분추정 결과를 이용해서 아래의 식을 이용하여 유전력과 모체효과(m^2)를 계산하였다

$$h_{\text{model2}}^2 = \sigma_a^2 / \sigma_a^2 + \sigma_m^2 + \sigma_e^2$$

$$m^2 = \sigma_m^2 / \sigma_a^2 + \sigma_m^2 + \sigma_e^2$$

h_{model1}^2 , h_{model2}^2 그리고 m^2 의 표준오차(standard error)는 ASReml프로그램의 pin파일이 제공하는 테일러전개식의 1차 근사(first order Taylor series approximation)를 이용한 델타방법(delta method)을 이용 하여 구하였다

결과 및 고찰

측정형질의 기초 통계치들은 Yoo등의 보고(2014)에 제시되어있으며, 생시체중(BW_{bir}), 21일령체중(BW21), 70일령체중(BW70), 그리고 140일령체중(BW140)의 평균은 각각 1.5 kg, 6.3 kg, 24.9 kg, 그리고 71.8 kg이었다. Porcine SNP 60K BeadChip을 이용하여 유전자형이 분석된 총62,163개의 SNP 표지들 중에서 QC 작업을 거친 전체 39,992개 SNP의 상염색체상의 위치에 분포에 대한 정보는 Table1과 같다.

Table2는 공식 육군에서 BW_{bir}, BW21, BW70, BW140에 대한 유전분산성분과 환경분산성분 유전력을 모형 1을 기반으로 NRM과 GRM을 이용하여 REML 방법으로 추정하였다. 모형 1에서 NRM 유전력의 추정치는 0.62에서 0.97이였으며, GRM 유전력의 추정치는 0.36에서 0.55이었다. 같은 통계모형과 동일한 공식동물에서 비롯된 표현형 자료가 사용되었음에도 불구하고, 이상과 같은 상당한 차이를 보이는 이유로 SNP표지를 이용한 전장 유전체연구에서 제안된 “사라진 유전력(missing heritability)”에 기인한 것으로 추정된다(Manolio 등, 2009; Witt 등, 2014). 전장 유전체 유전분석에 쓰이고 있는 DNA chip상의 유의한 모든 SNP 표지들을 유전력 추정에 사용 하더라도, 가계를 이용한 상가적 유전효과(additive genetic effect)를 반영하는 전통적인 유전력 추정치보다 일반적으로 낮게 나타나는데 이는 현재 DNA chip 상에 집적된 SNP들이 전체 상가적 유전효과에 영향을 줄 수 있는 희귀 변이(rare variant)나 복제수 변이(copy-number variation, CNV)와같은 구조적 변이(structural variant)등을 반영하기 어렵기 때문이다. 인류 집단 연구에서도 체질량 계수(body mass index, BMI)와 신장에 대한 NRM 유전력이 GRM 유전력 보다 높은 수준을 보였다(Magnusson과Rasmussen, 2002; Silventoinen등, 2003; Yang 등, 2011b).

하지만, 유전력 추정에 있어서 가계를 이용할 시에는 모체효과 또는 가정효과(household effect)와 같이 환경을 공유하면서 발생할 수 있는 추가적인 분산성분이 있는 것에 주의할 필요가 있다. 이들을 유전력 추정을 위한 혼합 모형에 포함시키지 않을 경우, 유전력의 추정치가 실제보다 과장되게 계산되어 나올 수도 있기 때문이다(Yang 등, 2011b). 본 연구 결과에 있어서도 BW_{bir}, BW21, BW70의 경우 추정치가 0.9를 넘어(Table2), 과대하게 추정 되었을 가능성을 배제하지 못한다. 따라서 본 연구에서는 F₂육군 NRM 유전력 추정모형에 모체효과를 포함시켜 분산성분을 추정함으로써 보다 더 정확한 유전력 추정 하고자 하였다. Table3은 공식육군에서 측정형질에 대한 유전분산성분과 환경분산성분 모체효과분산성분 그리고 유전력과 모체효과를 모형2를 기반으로 NRM과 GRM을 이용하여 REML 방법으로 추정하였다. 모형 2에서

혈통 정보를 이용한 NRM 유전력의 범위는 0.41에서 0.59이었으나, GRM 유전력은 0.04에서 0.46이었다. 전반적으로 모형 2를 이용한 유전력 추정치들이 모형 1의결과보다 낮았는데, 이는 모체효과의 분산성분 때문으로 가계의 유전력추정에 있어 모체효과와 같은 공통된 환경효과가 존재할 때에는 그 해석에 있어서 신중을 기할 필요가 있을 것으로 생각된다

집단에서 가계 정보 없이 GRM 유전력을 추정할 경우에 있어서도 주의가 필요하다고 하겠다. Yang 등(2011b)은 이러한 경우 공동 환경효과가 의심되는 부모 형제지간 같이 가까운 혈연관계인 개체들을 제외할 것을 권고하고 있다

본 연구에서는 제주재래흑돼지와 랜드레이스의 교배 축군으로 부터 성장형질의 유전력 추정을 위하여 기존의 가계정

Table 1. The information of assigned SNP markers in each porcine chromosome

Chromosome no.	Length of chromosome (Mb)	Number of SNP markers
1	314.9	4,937
2	162.1	2,575
3	144.7	2,196
4	143.4	2,678
5	111.2	1,832
6	157.6	2,655
7	14.6	2,509
8	147.7	2,198
9	153.6	2,487
10	77.8	1,388
11	87.6	1,416
12	63.4	1,177
13	218.1	3,034
14	153.6	2,898
15	157.4	2,205
16	86.6	1,461
17	68.9	1,295
18	60.5	1,051
Total		39,992

Table 2. Estimates of heritability(h^2) using Model 1

Trait	NRM			GRM		
	α_a^2	α_e^2	h^2	α_a^2	α_e^2	h^2
BW_bir	8.82E-02	7.62E-03	0.92 (0.11) ¹	2.78E-02	4.98E-02	0.36 (0.05)
BW21	2.01	5.22E-02	0.97 (0.10)	0.57	0.93	0.38 (0.05)
BW70	19.03	2.10	0.90 (0.11)	7.79	9.35	0.45 (0.05)
BW140	60.96	37.23	0.62 (0.10)	50.14	40.83	0.55 (0.04)

¹ Numbers in parenthesis represent SE of h^2

Table 3. Estimates of heritability (h^2) and maternal effect (m^2) using model 2

Trait	NRM					GRM				
	α_a^2	α_m^2	α_e^2	h^2	m^2	α_a^2	α_m^2	α_e^2	h^2	m^2
BW_bir	0.04	0.01	0.03	0.50 (0.28)	0.17 (0.11)	0.003	0.03	0.05	0.036 (0.03)	0.38 (0.06)
BW21	0.72	0.39	0.64	0.41 (0.26)	0.22 (0.11)	0.17	0.60	0.91	0.099 (0.04)	0.36 (0.06)
BW70	8.63	3.50	7.26	0.44 (0.27)	0.18 (0.12)	4.33	5.28	9.05	0.23 (0.05)	0.28 (0.05)
BW140	57.80	1.10	38.80	0.59 (0.25)	0.01 (0.09)	42.08	8.20	41.14	0.46 (0.05)	0.09 (0.04)

¹ Numbers in parenthesis represent SE of h^2 and m^2

보를 이용한 방법과 SNP chip 분석에서 산출된 유전자형 정보를 활용한 방법을 비교한 결과 그 차이를 “사라진 유전력”을 이용하여 고찰하였다. 그리고, 가계 및 집단의 유전력 추정에 있어서도 그 해석에 주의가 필요함을 모체효과의 분산 성분 추정을 통하여 보여주었다. 본 연구의 결과는 제주재래돼지를 기반으로 계통집단구축에 있어 성장형질 개량 연구에 기초자료로 활용될 것으로 기대된다.

사 사

본 연구는 국립축산과학원 경상과제 “제주재래돼지 기반 합성 종개량을 위한 계통조성돈 유지 및 개발과제번호: PJ01010502”의 지원과 2016년도 농촌진흥청 국립축산과학원 박사후 연수과정 지원사업에 의해 이루어진 것임

REFERENCES

- Cho IC, Park HB, Yoo CK, Lee GJ, Lim HT, Lee JB, Jung EJ, Ko MS, Lee JH and Jeon JT. 2011. QTL analysis of white blood cell, platelet and red blood cell-related traits in an F₂ intercross between Landrace and Korean native pigs. *Anim. Genet.* 42:621-626.
- Cho IC, Yoo CK, Lee JB, Jung EJ, Han SH, Lee SS, Ko MS, Lim HT and Park HB. 2015. Genome-wide QTL analysis of meat quality-related traits in a large F₂ intercross between Landrace and Korean native pigs. *Genet. Sel. Evol.* 47:7.
- Henderson C. 1984. Applications of linear models in animal breeding. University of Guelph, Guelph. Canada.
- Kim DH, Seong PN, Cho SH, Kim JH, Lee JM, Jo C and Lim DG. 2009. Fatty acid composition and meat quality traits of organically reared Korean native black pigs. *Livest. Sci.* 120:96-102.
- Lee C. 2000. Methods and Techniques for Variance Component Estimation in Animal Breeding. *Asian-Australas. J. Anim. Sci.* 13:413-422.
- Magnusson PK and Rasmussen F. 2002. Familial resemblance of body mass index: A study of young men in Sweden. *Int. J. Obes. Relat. Metab. Disord.* 26:1225-1231.
- Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorf LA, Hunter DJ, McCarthy MI, Ramos EM, Cardon LR, Chakravarti A, Cho JH, Guttmacher AE, Kong A, Kruglyak L, Mardis E, Rotimi CN, Slatkin M, Valle D, Whittemore AS, Boehnke M, Clark AG, Eichler EE, Gibson G, Haines JL, Mackay TF, McCarroll SA and Visscher PM. 2009. Finding the missing heritability of complex diseases. *Nature.* 461:747-753.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly MJ and Sham PC. 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81:559-575.
- Silventoinen K, Sammalisto S, Perola M, Boomsma DI, Cornes BK, Davis C, Dunkel L, De Lange M, Harris JR, Hjelmborg JV, Luciano M, Martin NG, Mortensen J, Nisticò L, Pedersen NL, Skytthe A, Spector TD, Stazi MA, Willemsen G and Kaprio J. 2003. Heritability of adult body height: a comparative study of twin cohorts in eight countries. *Twin Res.* 6:399-408.
- Spencer GSG. 1985. Hormonal systems regulating growth. *Livest. Prod. Sci.* 12:31-46.
- VanRaden PM. 2008. Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91:4414-4423.
- Witte JS, Visscher PM and Wray NR. 2014. The contribution of genetic variants to disease depends on the ruler. *Nat. Rev. Genet.* 15:765-776.
- Yang J, Lee SH, Goddard ME and Visscher PM. 2011a. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* 88:76-82.
- Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, Cunningham JM, de Andrade M, Feenstra B, Feingold E, Hayes MG, Hill WG, Landi MT, Alonso A, Lettre G, Lin P, Ling H, Lowe W, Mathias RA, Melbye M, Pugh E, Cornelis MC, Weir BS, Goddard ME and Visscher PM. 2011b. Genome partitioning of genetic variation for complex traits using common SNPs. *Nat. Genet.* 43:519-525.
- Yoo CK, Park HB, Lee JB, Jung EJ, Kim BM, Kim HI, Ahn SJ, Ko MS, Cho IC and Lim HT. 2014. QTL analysis of body weight and carcass body length traits in an F₂ intercross between Landrace and Korean native pigs. *Anim. Genet.* 45:589-592.