

제주마 집단의 혈연 정보량과 정보 오류가 유전 모수 추정치에 미치는 영향

김남영¹ · 이성수¹ · 양영훈^{2*}

¹농촌진흥청 국립축산과학원, ²제주대학교

Influence of Amount of Pedigree Information and Parental Misidentification of Progeny on Estimates of Genetic Parameters in Jeju Race Horses

Nam-Young Kim¹, Sung-Soo Lee¹ and Young-Hoon Yang^{2*}

¹National Institute of Animal Science, RDA, Jeju 690-150, Korea

²Faculty of Animal Biotechnology, Jeju National University, Jeju 690-756, Korea

ABSTRACT

The pedigree information and race records of 1,000 m finishing time of Jeju race horses at KRA were used to study the effect of amount of pedigree information and parental misidentification on the estimates of genetic parameters. The modified data sets were made at the range of 2.5 to 25% parental misidentifications or loss of parental information of individuals with an increment of 2.5 percent. For each incremental level, 20 randomly replicated data sets were obtained and analyzed by single-trait analysis with a DF-REML(AI) algorithm. As the rate of misidentification increased or the amount of pedigree information decreased, the estimates of fraction of additive genetics variance component gradually decreased almost linearly ($p < 0.05$), while the estimated fractions of error variance and permanent environmental variance components gradually increased for the finishing time. Regression coefficients of the percentage amount of both parents' information loss and incorrect pedigree information on additive genetic variances were -0.079 and -0.114 , respectively ($p < 0.01$). The estimate of heritability decreased by 0.92% for one percent loss of both parents' information and 1.39% for one percent increase of both parental misidentifications of progeny ($p < 0.01$). For the consideration of probable incorrect and missing parent information of progeny in this early population of Jeju horses, the estimates of additive genetic parameters would be biased downward about ten percent. This results indicate that the amount of pedigree information loss and misidentification of progeny would severely affect estimates of genetic parameters and would reduce genetic gains for selection in Jeju horse population.

(Key words : misidentification, pedigree information, genetic parameter estimates, Jeju race horses)

서 론

집단을 구성하는 개체들의 혈연관계에 대한 정보는 가축의 유전 능력 평가에 필수적 기초 자료로 이용되고 있다. 개체 육종가(B.V.)의 추정은 상가적 혈연 계수의 행렬식(additive relationship matrix)을 필요로 하게 되는데(Henderson, 1984), 행렬식을 구성하는 혈통 정보량과 정보의 정확성은 개체의 유전적 능력 평가에 그만큼 중요한 요인으로 작용하게 된다.

오랜 기간 폐쇄된 소규모 가축 집단에서는 부모 및 조상의 혈통 정보에 대한 기록이 없을지라도, 집단을 구성하고 있는 개체들 사이에는 다양한 정도의 혈연관계가 있을 것이라는 추측은 당연하다. 최근에는 보다 정확한 유전 능력 평가를 위해

서 유실된 혈연 정보를 최대한 수집하고, 정보의 오류를 교정하기 위하여 분자유전학적인 기법이 사용되기도 한다. 또한, 정확한 유전 능력 평가를 위한 노력의 하나로 혈연관계를 나타내는 상가적 행렬 계수의 구성을 DNA 표지나 SNP 좌위를 이용한 분자유전학적인 방법으로 접근하는 방법이 보고된 바도 있다(Wilson 등, 2003; Hay와 Goddard, 2008).

Senneke 등(2004)과 Sanders 등(2006)은 헤어포드 품종 집단과 Angeln 젓소 집단에서 아비에 대한 정보 오류가 생식 체중과 이유시 체중에 대한 유전 모수 추정치에 영향을 보고한바 있다. 특히 Sander 등(2006)은 종모우당 자손수가 적거나, 유전력이 낮은 형질의 경우에는 아비에 대한 정보의 유실 또는 정보의 오류는 선발에 따른 유전적 개량량에도 심각한 영향을

* 본 논문은 농촌진흥청 연구사업(과제번호 : PJ01022201)의 지원에 의해 이루어진 것임.

† Correspondence : yhyang@jeju.ac.kr

미친다고 하였다. Van Vleck(1970)은 부가계로 구성된 암소 집단에서 개체의 아버지에 대한 정보 오류는 형질의 유전력을 감소시키고, 정확한 유효 정보의 비율(p)을 기준으로 한다고 하면, 유전력 추정치의 감소율은 대략 $(1-p^2)$ 에 이른다고 요약한 바 있다. Woolliams(2006)은 개체에 대한 아버지의 오류와 유실은 육종가 추정치에 대한 신뢰도와 개량량을 감소시킨다는 이론을 논리적으로 전개한 바 있다.

일반적으로 가축의 유전 능력 평가와 선발 반응 예측에 이용되는 유전 모수 추정치는 혈통 정보를 기초로 계산된다. 혈액학적 및 분자유전학적 다양한 혈통 확인 방법이 개발되기 이전부터 존재한 혈통 기록에서는 그 정보의 오류가 완벽하게 배제되었다고 확신하기 어려운 것도 사실이다. 친자 확인 기법에 의하면 일반적으로 혈통에 대한 기록의 오류는 육우, 젓소, 염소 등의 집단에서 7~12% 내외로 존재하는 것으로 보고되고 있다(Visscher 등, 2002; Sanders 등, 2006; Saberivand 등, 2011).

유전 능력 평가를 위하여 초기 가축 집단에서 만족할 만큼 충분한 혈통 정보를 확보하는 것은 쉽지 않다. 또한 혈통 정보가 존재할지라도 혈통 정보 자체의 오류를 완벽하게 배제할 수 있는 조치도 쉽지 않다(Visscher 등, 2002; Sanders 등, 2006). 소규모 가축 집단을 평가함에 있어서 개체들 사이에 어느 정도의 혈연 정보량을 확보해야 신뢰성 있는 평가가 되는지, 또는 정보의 유실량이 모수 추정치에 미치는 영향은 유전적 능력 평가에 항상 의문이 아닐 수 없다.

따라서 본 연구는 등록 관리되고 있는 제주마 집단에서 제주마의 유전적 평가를 위한 기초 연구로 혈통 정보의 유실 정도와 오류량이 유전 분산 등 유전 모수의 추정에 미치는 영향을 분석하였다.

재료 및 방법

제주특별자치도에서 혈통 관리되는 제주마 1,763두(수컷 773두, 암컷 990두)의 개체의 혈통 기록과 제주경마공원의 1,000 m 경주 기록을 수집하여 분석에 이용하였다. 수집된 기록 중 경주로 상태는 함수율에 따라 건조(1~5%), 양호(6~9%), 다습(10~14%), 포화(15~19%), 불량(20% 이상)으로 구분하였으며, 부담 중량은 경주 조건, 과거 성적, 최근 성적 흐름, 출전 등록마 간 능력 차이 등을 고려하여 한국마사회 핸디캡 위원에서 결정한 중량을 분석에 이용하였다. 제주마 1,763두 개체들 가운데 부모에 대한 정보를 소유한 개체는 1,565두(정보 보유율 88.8%)였다. 경주 기록 보유 개체는 224두(12.7%)였으며, 1,000 m 경주 기록 수는 1,865개로 제주마 개체 당 평균 8.3개의 기록을 보유하고 있었다.

혈통 정보의 유실량이 모수 추정치에 영향을 관찰하기 위

해서 혈통원부의 자료를 이용하여 다양한 유형의 모의 자료를 구성하였다. 자료의 구성은 Fortran 램덤 함수를 이용하여 혈통 자료에서 0~25%까지 2.5%의 등간 간격(0, 2.5, 5.0, 7.5, 10.0, 12.5, 15.0, 17.5, 20.0, 22.5, 25%)으로 혈통 정보를 삭제하면서 33개 유형의 모의 혈통 자료(부모 정보 동시 유실 유형 11개, 아버 정보 유실 유형 11개, 어미 정보 유실 유형 11개)를 구성하였다. 혈통 정보의 오류량이 모수 추정치에 영향에 대한 분석도 마찬가지로 0~25%까지 2.5%의 등간 간격으로 모의 혈통 자료들을 구성하였는데, 개체의 생년월일을 기준으로 하여 동일년도에 태어나 이웃하고 있는 개체들 사이에 서로 램덤으로 교환되는 방법으로 33개의 유형(부모 혈통이 동시 뒤바뀐 유형 11개, 아버 정보 바뀐 유형 11개, 어미 정보 바뀐 유형 11개)의 모의 자료를 구성하였다.

정보유실 또는 오류에 대한 영향은 각각의 유형에서 다시 20회 반복으로 램덤 자료를 재구성하면서 모수를 추정한 후, 20회 반복의 모수 추정치들의 평균을 중심점으로 하여 혈통 정보의 변동에 따른 유전 모수 추정치의 변화량에 대한 회귀 분석을 수행하였다.

분산 성분과 유전 모수의 추정은 1,000 m 경주의 주과속도에 영향을 미치는 연도, 나이, 성별, 계절, 경주로의 상태를 고정 효과로 하였고, 부담 중량의 공변량과 임의 효과인 영구 환경 효과와 상가적 유전 효과를 모형에 포함하여 WOMBAT (Meyer K., 2010) 프로그램을 이용하여 DF-REML-AI 알고리즘으로 추정하였다. 통계적 모형은 아래와 같다.

$$Y = Xb + Z_1a + Z_2pe + e$$

여기서,

Y는 1,000 m 경주 기록인 주과 시간에 대한 벡터,

X, Z₁, Z₂는 고정 효과와 임의 효과에 대한 각각의 빈도 행렬 및 벡터,

b는 연도, 나이, 계절, 성별의 고정 효과,

a는 개체의 상가적 유전 효과,

pe는 영구 환경 효과,

a는 N(0, σ_A^2R)의 분포로 R는 혈연 정보를 이용한 혈연 계수의 행렬식,

pe는 N(0, σ_{PE}^2I)인 분포, e는 N(0, σ_E^2I)의 분포로 가정하였다.

결 과

개체의 부모 정보를 램덤 함수에 의해서 의도적으로 소실시킨 자료를 이용한 분석에서는 부모 정보의 소실량의 증가됨(0~25%)에 따라, 전체 분산 성분은 변화가 거의 없었으나, 오차 분산 성분과 영구 환경 분산 성분은 증가하는 경향을 보여

주고 있었고($p < 0.01$), 상가적 유전 분산 성분은 뚜렷하게 감소($p < 0.01$)하는 경향으로 나타났다(Table 1). 이에 따라 추정된 분산 성분에 의한 유전력 추정치도 혈통 정보 소실 비율이 증가함에 따라 선형적($b = -0.006 \pm 0.000, p < 0.01$)으로 낮아지는 경향을 보였다. 집단에 소속된 개체들의 부모에 대한 정보가 1% 추가로 유실될 때마다 유전력 추정치는 0.92%씩 감소하는 반응을 보이고 있다.

개체의 부모 정보 가운데 아비에 대한 정보만 점증적으로 소실시킨 경우에는 오차 분산 성분은 증가하는 경향을 보였지만, 영구 환경 분산 성분은 감소하는 경향을 보였고($p < 0.05$), 상가적 유전 분산 성분의 양은 감소하는 경향을 보였으나($p < 0.01$), 유전력 추정치에 대한 선형적 감소 경향($p > 0.05$)은 나타나지 않았다(Table 2). 개체의 어미에 대한 정보를 점증적으로 유실시킨 자료로부터 얻은 모수 추정치들은 부모 정보가 동시에 소실된 자료에서 얻은 결과와 유사한 경향으로 상가적 유전 분산과 유전력 추정치의 선형적 감소를 확인할 수 있었다(Table 3). 어미 정보 유실량 증가에 따른 오차 분산 성분의 선형적 변화에 대한 유의성은 없었다($p > 0.05$).

이웃하는 개체들 사이에 무작위로 부모에 대한 정보가 동시에 교환된 자료에서는 정보 오류가 점증적으로 증가(0~25%)됨에 따라 오차 분산 성분과 영구 환경 분산 성분은 증가를 보였으며($p < 0.01$), 전체 분산 성분과 상가적 유전 분산 성분($p < 0.01$)은 뚜렷하게 감소하는 경향을 보였다($p < 0.01$). 개체의 부모에 대한 정보 오류량이 증가함에 따라 유전력 추정치도 선형적으로($b = -0.008 \pm 0.001, p < 0.01$) 뚜렷하게 감소되는 것으로 나타났다(Table 4). 개체의 부모에 대한 혈통 정보의 오류가 1% 증가됨에 따라 유전력 추정치는 1.39%씩 낮아지는 반응을 보이고 있었다($p < 0.01$).

개체의 아비에 대한 정보가 바뀌는 경우에는 상가적 유전 분산과 유전력의 감소로 이어지고 있으나(Table 5), 정보의 오류량이 증가에 따라 나타나는 선형적 감소 현상을 관찰되지 않았다($p > 0.05$). 개체의 어미에 대한 정보가 바뀌는 경우에는 전체 분산 성분과 상가적 유전 분산 성분의 감소로 이어지고 있었다($p < 0.01$). 어미에 대한 정보의 오류량이 점진적으로 증가함에 따라 유전력 추정치도 선형적으로 감소($b = -0.006 \pm 0.000, p < 0.01$)하는 현상을 뚜렷하게 보여주고 있었다($p < 0.01$).

Table 1. Parameter estimates for the finish time at 1000-meter race and regression coefficients(b) of losing percent rates of estimates with respect to losing percent of pedigree information using the pedigree data with partially deleted sire and dam information

Item	Percent losing rate of sire and dam information											b	SE(b)	t-value
	0.0	2.5	5.0	7.5	10.0	12.5	15.0	17.5	20.0	22.5	25.0			
Var(e)	2.621	2.621	2.621	2.621	2.622	2.622	2.622	2.622	2.622	2.622	2.623	0.000	0.000	11.692**
Var(a)	8.063	7.877	7.700	7.464	6.876	6.292	6.283	5.853	4.929	4.752	4.318	-0.079	0.004	-18.497**
Var(pe)	2.714	2.845	3.009	3.174	3.652	4.147	4.120	4.445	5.314	5.483	5.815	0.065	0.004	16.066**
Var(ph)	13.398	19.493	13.331	13.259	13.150	13.061	13.025	12.793	12.866	12.857	12.756	-0.059	0.034	-1.717 ^{NS}
Repeatability	0.804	0.803	0.803	0.802	0.801	0.799	0.799	0.797	0.796	0.796	0.794	0.000	0.000	-24.370**
Heritability	0.602	0.590	0.577	0.563	0.523	0.481	0.482	0.453	0.382	0.368	0.338	-0.006	0.000	-16.502**
e/a	0.325	0.333	0.340	0.351	0.381	0.417	0.417	0.448	0.532	0.552	0.607	0.006	0.001	10.231**
e/pe	0.966	0.922	0.871	0.826	0.718	0.632	0.636	0.590	0.494	0.478	0.451	-0.011	0.001	-19.813**
	Percent change against the initial value (intact)													
Var(e)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.001	0.000	11.692
Var(a)	0.0	-2.3	-4.5	-7.4	-14.7	-22.0	-22.1	-27.4	-38.9	-41.1	-46.4	-0.978	0.053	-18.497
Var(pe)	0.0	4.8	10.9	16.9	34.5	52.8	51.8	63.8	95.8	102.0	114.3	2.411	0.150	16.066
Var(ph)	0.0	45.5	-0.5	-1.0	-1.9	-2.5	-2.8	-4.5	-4.0	-4.0	-4.8	-0.437	0.255	-1.717
Repeatability	0.0	-0.1	-0.1	-0.3	-0.5	-0.6	-0.7	-0.9	-1.0	-1.0	-1.2	-0.026	0.001	-24.370
Heritability	0.0	-1.9	-4.0	-6.5	-13.1	-20.1	-19.9	-24.8	-36.5	-38.8	-43.9	-0.922	0.056	-16.502

Var(e) : residual variance, Var(a) : additive variance, Var(pe) : permanent environmental variance, Var(ph) : phenotypic variance. All the parameter estimates are of the means of 20 estimets from randomly replicated data sets.

** $p < 0.01$, ^{NS} not significant.

Table 2. Parameter estimates for the finish time at 1000-meter race and regression coefficients(b) of losing percent rate of estimates with respect to losing percent of pedigree information using the pedigree data with partially deleted sire information

Item	Percent losing rate of sire information											b	SE(b)	t-value
	0.0	2.5	5.0	7.5	10.0	12.5	15.0	17.5	20.0	22.5	25.0			
Var(e)	2.621	2.621	2.621	2.622	2.622	2.622	2.622	2.623	2.623	2.623	2.623	0.000	0.000	10.562**
Var(a)	8.063	8.017	8.017	7.969	7.989	7.719	7.708	7.625	7.620	7.918	7.677	-0.008	0.002	-3.603**
Var(pe)	2.714	2.732	2.745	2.677	2.538	2.711	2.561	2.722	2.615	2.210	2.448	-0.007	0.002	-2.822*
Var(ph)	13.398	13.371	13.383	13.268	13.149	13.052	13.021	12.969	12.858	12.751	12.747	-0.015	0.001	-18.757**
Repeatability	0.804	0.804	0.804	0.802	0.801	0.799	0.799	0.798	0.796	0.794	0.794	0.000	0.000	-18.648**
Heritability	0.602	0.600	0.599	0.601	0.607	0.591	0.592	0.588	0.593	0.621	0.603	0.000	0.000	0.321 ^{NS}
e/a	0.325	0.327	0.327	0.329	0.328	0.340	0.340	0.344	0.344	0.331	0.342	0.000	0.000	3.604**
e/pe	0.966	0.959	0.955	0.980	1.033	0.967	1.024	0.964	1.003	1.187	1.072	0.003	0.001	2.699*
Percent change against the initial value (intact)														
Var(e)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.1	0.1	0.1	0.1	0.001	0.000	
Var(a)	0.0	-0.6	-0.6	-1.2	-0.9	-4.3	-4.4	-5.4	-5.5	-1.8	-4.8	-0.101	0.028	
Var(pe)	0.0	0.7	1.1	-1.4	-6.5	-0.1	-5.6	0.3	-3.7	-18.6	-9.8	-0.248	0.088	
Var(ph)	0.0	-0.2	-0.1	-1.0	-1.9	-2.6	-2.8	-3.2	-4.0	-4.8	-4.9	-0.109	0.006	
Repeatability	0.0	-0.1	0.0	-0.2	-0.5	-0.7	-0.7	-0.8	-1.0	-1.3	-1.3	-0.028	0.002	
Heritability	0.0	-0.4	-0.4	-0.2	0.9	-1.7	-1.6	-2.3	-1.5	3.3	0.2	0.010	0.031	

Var(e) : residual variance, Var(a) : additive variance, Var(pe) : permanent environmental variance, Var(ph) : phenotypic variance. All the parameter estimates are of the means of 20 estimates from randomly replicated data sets.

* $p < 0.05$, ** $p < 0.01$, ^{NS} not significant.

Table 3. Parameter estimates for the finish time at 1000-meter race and regression coefficients(b) of losing percent rate of estimates with respect to losing percent of pedigree information using the pedigree data with partially deleted dam information.

Item	Percent losing rate of dam information											b	SE(b)	t-value
	0.0	2.5	5.0	7.5	10.0	12.5	15.0	17.5	20.0	22.5	25.0			
Var(e)	2.621	2.621	2.621	2.621	2.620	2.621	2.621	2.621	2.621	2.621	2.621	0.000	0.000	0.380 ^{NS}
Var(a)	8.063	7.664	7.756	7.462	6.840	6.912	6.718	5.962	5.360	5.203	4.934	-0.065	0.005	-13.993**
Var(pe)	2.714	3.093	3.034	3.325	3.894	3.828	3.982	4.660	5.191	5.332	5.569	0.059	0.004	15.018**
Var(ph)	13.398	13.378	13.411	13.409	13.356	13.362	13.322	13.243	13.172	13.155	13.124	-0.006	0.001	-7.333**
Repeatability	0.804	0.804	0.805	0.804	0.804	0.804	0.803	0.802	0.801	0.801	0.800	0.000	0.000	-7.300**
Heritability	0.602	0.573	0.578	0.556	0.512	0.517	0.504	0.454	0.406	0.395	0.375	-0.005	0.000	-14.184**
e/a	0.325	0.342	0.338	0.351	0.383	0.379	0.390	0.440	0.489	0.504	0.531	0.004	0.000	10.081**
e/pe	0.966	0.847	0.864	0.788	0.673	0.685	0.658	0.562	0.505	0.492	0.471	-0.010	0.001	-15.367**
Percent change against the initial value (intact)														
Var(e)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.000	0.000	
Var(a)	0.0	-4.9	-3.8	-7.4	-15.2	-14.3	-16.7	-26.1	-33.5	-35.5	-38.8	-0.807	0.058	
Var(pe)	0.0	14.0	11.8	22.5	43.5	41.1	46.7	71.7	91.3	96.4	105.2	2.174	0.145	
Var(ph)	0.0	-0.1	0.1	0.1	-0.3	-0.3	-0.6	-1.2	-1.7	-1.8	-2.0	-0.045	0.006	
Repeatability	0.0	0.0	0.0	0.0	-0.1	-0.1	-0.1	-0.3	-0.4	-0.4	-0.5	-0.011	0.002	
Heritability	0.0	-4.8	-3.9	-7.5	-14.9	-14.1	-16.2	-24.6	-32.5	-34.4	-37.6	-0.778	0.055	

Var(e) : residual variance, Var(a) : additive variance, Var(pe) : permanent environmental variance, Var(ph) : phenotypic variance. All the parameter estimates are of the means of 20 estimates from randomly replicated data sets.

* $p < 0.05$, ** $p < 0.01$, ^{NS} not significant.

Table 4. Parameter estimates for the finish time at 1000-meter race and regression coefficients(b) of losing percent rate of estimates with respect to losing percent of pedigree information using the pedigree data with partially exchanged sire and dam information between contemporary individuals

Item	Percent exchange rate of sire and dam information											b	SE(b)	t-value
	0.0	2.5	5.0	7.5	10.0	12.5	15.0	17.5	20.0	22.5	25.0			
Var(e)	2.621	2.621	2.621	2.622	2.622	2.622	2.622	2.622	2.622	2.622	2.622	0.000	0.000	9.454**
Var(a)	8.063	7.563	6.850	6.144	5.090	4.929	3.755	4.398	3.224	2.771	2.541	-0.114	0.007	-15.829**
Var(pe)	2.714	3.198	3.872	4.551	5.460	5.608	6.694	6.057	7.127	7.520	7.706	0.102	0.007	14.800**
Var(ph)	13.398	13.385	13.343	13.316	13.172	13.159	13.070	13.077	12.973	12.912	12.870	-0.011	0.001	-19.760**
Repeatability	0.804	0.804	0.804	0.803	0.801	0.801	0.799	0.799	0.798	0.797	0.796	0.000	0.000	-18.962**
Heritability	0.602	0.565	0.513	0.461	0.384	0.374	0.284	0.335	0.246	0.214	0.196	-0.008	0.001	-15.582**
e/a	0.325	0.347	0.383	0.427	0.515	0.532	0.698	0.596	0.813	0.946	1.032	0.014	0.001	10.453**
e/pe	0.966	0.820	0.677	0.576	0.480	0.467	0.392	0.433	0.368	0.349	0.340	-0.011	0.002	-7.076**
Percent change against the initial value (intact)														
Var(e)	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.001	0.000	
Var(a)	0.0	-6.2	-15.0	-23.8	-36.9	-38.9	-53.4	-45.4	-60.0	-65.6	-68.5	-1.409	0.089	
Var(pe)	0.0	17.8	42.6	67.7	101.2	106.6	146.6	123.2	162.6	177.1	183.9	3.769	0.255	
Var(ph)	0.0	-0.1	-0.4	-0.6	-1.7	-1.8	-2.4	-2.4	-3.2	-3.6	-3.9	-0.084	0.004	
Repeatability	0.0	0.0	-0.1	-0.2	-0.4	-0.4	-0.6	-0.6	-0.8	-0.9	-1.0	-0.022	0.001	
Heritability	0.0	-6.1	-14.7	-23.4	-36.1	-37.9	-52.7	-44.3	-59.0	-64.5	-67.4	-1.386	0.089	

Var(e) : residual variance, Var(a) : additive variance, Var(pe) : permanent environmental variance; Var(ph), phenotypic variance. All the parameter estimates are of the means of 20 estimates from randomly replicated data sets.

* $p < 0.05$, ** $p < 0.01$, ^{NS} not significant.

Table 5. Parameter estimates for the finish time at 1000-meter race and regression coefficients(b) of losing percent rate of estimates with respect to losing percent of pedigree information using the pedigree data with partially exchanged sire information between contemporary individuals

Item	Percent exchange rate of sire information											b	SE(b)	t-value
	0.0	2.5	5.0	7.5	10.0	12.5	15.0	17.5	20.0	22.5	25.0			
Var(e)	2.621	2.621	2.621	2.621	2.621	2.621	2.622	2.621	2.621	2.621	2.622	0.000	0.000	1.249 ^{NS}
Var(a)	8.063	7.852	7.967	7.881	8.057	7.876	7.498	8.125	7.914	8.008	7.847	-0.001	0.003	-0.370 ^{NS}
Var(pe)	2.714	2.819	2.829	2.935	2.858	3.028	2.997	2.824	3.027	2.969	3.081	0.005	0.001	3.681**
Var(ph)	13.398	13.392	13.428	13.437	13.536	13.526	13.534	13.571	13.562	13.597	13.549	0.004	0.001	6.711**
Repeatability	0.804	0.804	0.805	0.805	0.806	0.806	0.806	0.807	0.807	0.807	0.806	0.000	0.000	6.661**
Heritability	0.602	0.586	0.594	0.586	0.595	0.582	0.585	0.599	0.584	0.589	0.579	0.000	0.000	-1.768 ^{NS}
e/a	0.325	0.334	0.329	0.333	0.325	0.333	0.350	0.323	0.331	0.327	0.334	0.000	0.000	0.375 ^{NS}
e/pe	0.966	0.930	0.926	0.893	0.917	0.866	0.875	0.928	0.866	0.883	0.851	-0.002	0.000	-3.677**
Percent change against the initial value (intact)														
Var(e)	0.000	0.010	0.005	0.006	0.003	0.008	0.019	-0.001	0.011	0.000	0.025	0.000	0.000	
Var(a)	0.000	-2.618	-1.197	-2.256	-0.080	-2.321	-7.006	0.773	-1.854	-0.676	-2.677	-0.015	0.042	
Var(pe)	0.000	3.873	4.244	8.140	5.308	11.583	10.422	4.070	11.518	9.413	13.507	0.197	0.054	
Var(ph)	0.000	-0.042	0.224	0.294	1.031	0.952	1.017	1.289	1.220	1.486	1.130	0.030	0.005	
Repeatability	0.000	0.031	0.098	0.110	0.290	0.271	0.289	0.354	0.335	0.404	0.308	0.008	0.001	
Heritability	0.000	-2.629	-1.327	-2.591	-1.133	-3.263	-2.852	-0.561	-3.066	-2.184	-3.888	-0.037	0.021	

Var(e) : residual variance, Var(a) : additive variance, Var(pe) : permanent environmental variance, Var(ph) : phenotypic variance. All the parameter estimates are of the means of 20 estimates from randomly replicated data sets.

* $p < 0.05$, ** $p < 0.01$, ^{NS} not significant.

추정된 유전력은 어미의 정보 오류량이 1%가 증가됨에 따라 1.04% 감소하는 반응을 보였다(Table 6).

고찰

본 연구에서 개체의 부모에 대한 혈통 정보의 유실량과 오류량은 모두 상당한 정도의 상가적 유전 분산량을 감소시키는 것을 확인해 주고 있었다. 특히 자손의 부모에 대한 정보 유실보다는 정보의 오류가 상가적 유전분산량과 유전력 추정치를 보다 낮게 왜곡하는 것으로 나타나서, 유전 모수 추정을 위한 분석에 앞서 미기록 혈통 정보의 추가적 확보보다는 기록된 정보의 오류를 여러 가지 방법으로 재확인하는 것이 바람직하다고 생각되었다(Table1, Table4).

제주마의 혈통은 2000년 이후부터 분자유전학적인 기법의 도움으로 혈통 관리가 되고 있다. 등록 초창기에는 조상이 이미 폐사하여 혈통 기록에 대한 진위가 확인되지 못한 혈통에 대한 기록도 포함되었고, 부모에 대한 정보가 없는 일부의 개

체들도 포함되어 제주마 등록 집단을 형성하고 있는 상태다. 등록 초기에 부모에 대한 정보가 확인되지 못한 개체들은 기초 등록마(11.2%)로 분류가 되었다. 사실 이들 개체들은 집단을 구성하고 있는 구성원들 사이에 이미 혈연관계가 있는 것으로 추정되고 있지만, 검증할만한 표본이 없어서 혈연관계가 없는 것으로 가정하여 관리되고 있는 셈이다. 이 같은 자료의 구성에서는 유전 모수의 추정과 육종가와 선발 반응에 왜곡 현상을 배제할 수 없다고 생각된다.

본 연구에서 개체의 아버지 또는 어미에 대한 정보의 유실량과 오류량 정도가 유전 모수 추정치에 대한 영향력은 앞서 보고한 연구자들(Van Vleck, 1970; Geldermann 등, 1986; Israel과 Weller, 2000; Senneke 등, 2004; Sanders 등, 2006)의 결과와 매우 유사한 형태로 오차 분산 성분과 영구 환경 분산 성분의 증가를 유발시키고, 상가적 유전 분산 성분의 감소를 가져오는 것을 확인해 주고 있었다. 원본의 혈통 자료에서 얻은 모수를 기준으로 했을 때, 개체의 부모 정보 동시 유실량과 오류량의 증가는 유전력 추정치를 선형적 부(유실량 : $b = -$

Table 6. Parameter estimates for the finish time at 1000-meter race and regression coefficients(b) of losing percent rate of estimates with respect to losing percent of pedigree information using the pedigree data with partially exchanged dam information between contemporary individuals

Item	Percent exchange rate of dam information											b	SE(b)	t-value
	0.0	2.5	5.0	7.5	10.0	12.5	15.0	17.5	20.0	22.5	25.0			
Var(e)	2.621	2.621	2.621	2.604	2.622	2.622	2.622	2.622	2.622	2.622	2.622	0.000	0.000	0.741 ^{NS}
Var(a)	8.063	7.565	7.285	6.647	6.006	5.737	5.317	4.793	5.088	3.710	3.740	-0.087	0.005	-17.677**
Var(pe)	2.714	3.219	3.550	4.078	4.633	4.895	5.243	5.708	5.446	6.638	6.528	0.077	0.005	17.078**
Var(ph)	13.398	13.405	13.457	13.347	13.261	13.254	13.181	13.123	13.156	12.970	12.975	-0.010	0.001	-9.861**
Repeatability	0.804	0.804	0.805	0.804	0.802	0.802	0.801	0.800	0.801	0.798	0.798	0.000	0.000	-9.950**
Heritability	0.602	0.564	0.541	0.497	0.451	0.431	0.402	0.364	0.385	0.285	0.294	-0.006	0.000	-17.226**
e/a	0.325	0.347	0.360	0.392	0.436	0.457	0.493	0.547	0.515	0.707	0.701	0.008	0.001	9.465**
e/pe	0.966	0.814	0.738	0.638	0.566	0.536	0.500	0.459	0.481	0.395	0.402	-0.010	0.001	-8.672**
Percent change against the initial value (intact)														
Var(e)	0.0	0.0	0.0	-0.7	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.003	0.004	
Var(a)	0.0	-6.2	-9.6	-17.6	-25.5	-28.8	-34.1	-40.6	-36.9	-54.0	-53.6	-1.083	0.061	
Var(pe)	0.0	18.6	30.8	50.3	70.7	80.4	93.2	110.3	100.7	144.6	140.5	2.834	0.166	
Var(ph)	0.0	0.1	0.4	-0.4	-1.0	-1.1	-1.6	-2.1	-1.8	-3.2	-3.2	-0.072	0.007	
Repeatability	0.0	0.0	0.1	-0.1	-0.3	-0.3	-0.4	-0.5	-0.5	-0.8	-0.8	-0.018	0.002	
Heritability	0.0	-6.2	-10.1	-17.5	-25.0	-28.4	-33.2	-39.5	-36.1	-52.6	-51.1	-1.039	0.060	

Var(e) : residual variance, Var(a) : additive variance, Var(pe) : permanent environmental variance, Var(ph) : phenotypic variance. All the parameter estimates are of the means of 20 estimets from randomly replicated data sets.
 * $p < 0.05$, ** $p < 0.01$, ^{NS} not significant.

0.92, 오류량 : $b = -1.38$)의 관계로 뚜렷하게 감소시키며($p < 0.01$), 정보의 오류량은 정보의 유실량보다 더 큰 영향을 미친다고 하겠다.

일반적으로 발생하는 정보의 오류량과 초기집단에 존재하는 혈통 정보의 누락 등에 의하면, 약 10% 내외의 정보 소실과 오류가 존재할 수 있는 것으로 추정할 수 있겠는데, 이와 같은 상황이라면 제주마의 1,000m 주파 속도에 대한 유전력은 0.60으로 추정되었지만, 유전력 추정치는 실제보다도 10% 정도는 낮게 추정되고 있다고 사료된다. 결과적으로 이처럼 낮게 추정된 유전 모수들은 선발 반응에도 당연히 영향을 미치게 될 것이다. Visscher 등(2002)도 아버지에 대한 정보 오류가 10%로 확인된 젓소 집단에서 선발 반응이 약 2~3% 정도 감소한 것으로 추정한 바 있다. 이는 후대 검정 효율과 육종가의 예측에도 실질적으로 영향하고 있음을 말해주고 있다. 제주마 집단에서 어미의 정보가 아버지의 정보보다 유전 모수 추정에 더 많은 영향을 주는 것으로 분석되었는데, 이는 번식계절에 극소수의 종모마를 이용하여 번식시키는 제주마 초기 집단의 특이적인 번식 형태가 영향을 미친 것으로 생각된다(종마 두당 평균 34.1두의 자마 생산). 또한 어미에 대한 정보 오류는 자마의 자연 포유 관계의 형성으로 현실적으로 그리 크게 발생하지는 않을 것으로 생각된다.

끝으로 유전 능력 평가에 혈통 정보의 확보량과 오류량에 따른 왜곡현상의 배제를 위해서는 혈통에 대한 최대의 정보 수집과 다양한 방법으로 혈통 진위에 대한 확인 점검이 필수적이라 생각되며, 참값에 보다 근접할 수 있는 유전 모수 추정을 위해서는 제주마 집단뿐만 아니라, 유전 모수를 추정하려는 모든 소규모 가축 집단에서는 추가적인 혈통 정보의 확보와 오류 확인에 대한 노력이 필요하다고 사료된다.

적 요

제주마 경주 능력 개량을 위한 기초 자료 확보를 위하여 혈통 정보량과 오류량이 유전 모수 추정치에 미치는 영향을 분석하였다. 혈통 자료는 램덤으로 혈통 정보의 삭제(0~25%) 또는 오류 유발(0~25%)과 각 수준별 20회 반복으로 모의자료를 형성하였다. 분석에 이용된 형질은 제주마 1,000m 경주의 주파 시간을 이용하였고, 상가적 유전 분산의 추정에는 DF-REML-AI 알고리즘을 이용하였다.

개체의 양부모에 대한 정보 유실량과 오류량의 증가는 선형적인 회귀관계로 상가적 유전 분산 성분(유실 $b = -0.079$, 오류 $b = -0.114$)과 유전력 추정치(유실, $b = -0.006$; 오류, $b = 0.008$)의 감소를 초래했다($p < 0.01$). 특히 개체의 양부모에 대한 정보유실보다는 정보 오류가 상가적 유전 분산량과 유전력 추정치를 보다 크게 왜곡하는 것으로 나타났다. 부 또는 모에

대한 정보의 유실과 오류의 증가도 대부분 오차 분산 성분과 영구 환경 분산 성분의 증가를 유발시키고, 상가적 유전 분산 성분의 감소를 가져왔다.

원자료에서 얻은 모수를 기준으로 했을 때 양부모에 대한 정보의 유실량과 오류량이 1% 증가함에 따라 유전력은 각각 0.92% 및 1.39%씩 낮아지는 관계를 보였다 ($p < 0.01$).

제주마의 1,000m 주파 속도에 대한 유전력은 0.60으로 추정되었지만, 초기 집단에 존재하는 혈통 정보의 누락 및 오류를 고려한다면 실제보다도 10% 정도는 낮게 추정되고 있다고 사료되었다. 이와 같은 결과는 상가적 유전 분산량의 감소는 물론, 선발에 따른 제주마의 경주 능력에 대한 유전적 개량량에도 영향을 줄 것으로 판단된다.

참 고 문 헌

- Geldermann H, Pieper U and Weber WE. 1986. Effect of misidentification on the estimation of breeding value and heritability in cattle. *J. Anim. Sci.* 63:1759-1768.
- Henderson CR. 1984. Applications of linear models in animal breeding. Canadian Cataloguing in Publication Data. Univ. Guelph, Ontario, Canada.
- Hayes BJ and Goddard ME. 2008. Technical note: Prediction of breeding values using marker-derived relationship matrices. *J. Anim. Sci.* 86:2089-2091.
- Isreal C and Weller JI. 2000. Effect of misidentification on genetic gain and estimation of breeding value in dairy cattle population. *J. Dairy Sci.* 83:181-187.
- Meyer K. 2010. WOMBAT, Animal Genetics and Breeding Unit, University of New England, Armidale, NSW 2351, Australia.
- Saberivand A, Javanmard A and Safdari M. 2011. Parentage verification and identity test of Ghezel sheep using microsatellite markers. *African Journal of Biotechnology* 10 (31): 5815-5819.
- Sanders K, Bennewitz J and Kalm E. 2006. Wrong and missing sire information affects genetic gain in the Angeln dairy cattle population. *J. Dairy Sci.* 89:315-321.
- Senneke SL, MacNeil MD and Van Vleck LD. 2004. Effect of sire misidentification on estimates of genetic parameters for birth and weaning weights in Hereford cattle. *J. Anim. Sci.* 2004. 80:2307-2312.
- Visscher PM, Woolliams JA, Smith D and Williams JL. 2002. Estimation of pedigree errors in the UK dairy population using microsatellite markers and the impact on selection. *J.*

- Dairy Sci. 85:2368-2375.
- Wilson AJ, McDonald G, Moghadam HK, Herbinger CM and Ferguson MM. 2003. Marker-assisted estimation of quantitative genetic parameters in rainbow trout, *Oncorhynchus mykiss*. Genet. Res. 81:145-156.
- Woolliams JA. 2006. A note on the differential impact of wrong and missing sire information on reliability and gain. J. Dairy Sci. 89:4901-4902.
-
- Received April 21, 2014, Revised May 6, 2014, Accepted September 22, 2014