

DNA marker traceability in Korean Cattle Jae Chul Kwon¹⁾, Yu-Mi Choi²⁾, Jea-Young Lee³⁾

Abstract

Samples of 33 Hanwoo individuals from Korean elite sire families were used, and five microsatellite markers were selected finally, which were located on chromosomes different chromosomes with the end sequencing of 100 HW-YUBAC that were recorded in the NCBI by Yeungnam University. Ten major microsatellite markers were selected from alleles amplified, their frequencies, H(Heterozygosity) and PIC(Polymorphism information content) with Hardy-Weinberg equilibrium. Next, in order to evaluate the power of the markers selected on the individual animal identification, the match probability(MP) and the relatedness coefficient(R) were computed.

Keywords : Traceability, Microsatellite, Heterozygosity, Match probability, Relatedness, Multiplex PCR

1. 서론

건강에 대한 관심이 늘어나면서 식육에 대한 소비자들의 요구도 양적인 측면에서 질적인 측면으로 변화하고 있으며, 특히 생산지 그리고 생산에서 소비에 이르는 유통과정 상의 신뢰성을 중요시하고 있다. 따라서 둔갑육 판매 방지와 쇠고기 안정성을 위해서 한우의 생산·도축·가공·유통과정의 각 단계별 정보를 기록·관리, 문제 발생 시 이동 경로를 따라 추적 또는 소급하여 신속한 원인규명 및 조치를 가능하게 하는 생산이력체계의 구축이 절실히 필요하게 되었다(Lim et al., 2005). 이에 기존의 개체의 동일성 검사에 사용된 방법으로 혈액형 및 생화학적인 지표의 다형성 분석방법을 사용하였으나 소에 있어서 혈액형검사는 정확한 부계혈통을 결정하지 못하는 단점이 있다.(GLowatzki-Mullis) 이로 인해 DNA다형성에 기초한 다양한 유전적 분석방법들이 개발되었으며 이를 이용하여 개체간의 동일성 검사가 이루어지고 있다.(Peelman, 1998). 이 논문의 목적은 한우를 중심으로 쇠고기 이력체계에서 개체식별에 유리한 DNA marker의 정보(H, PIC)를 알아내고, 그 DNA marker의 정보를 이용하여 여러 가지 DNA marker의 조합들의 검정력(MP, R)을 알아보고, 현실 data에서의 선택된 조합에 적용한 결과분석을 고려하여 최종적으로 개체식별에 유용한 DNA marker의 조합을 찾는 것이다.

1) Graduate, Department of statistics, Yeungnam University, Kyongsan, korea
E-mail : esjtk@yumail.ac.kr

2) Graduate, Department of statistics, Yeungnam University, Kyongsan, korea
E-mail : chldbal84@yumail.ac.kr

3) Professor, Department of statistics, Yeungnam University, Kyongsan, korea
E-mail : jlee@yumail.ac.kr

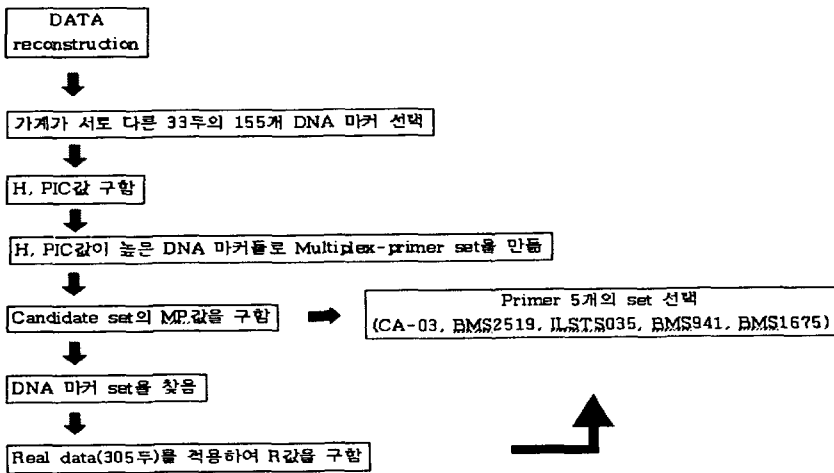
2. 실험 재료 및 방법

2.1 실험재료

본 연구에 공시된 한우는 농협중앙회 가축개량 사업소로부터 한우 집단의 혈연정보가 정확히 기재된 국가 후대검정우 26-29차 36가계 305두와 그 중 조부와 부가 서로 각각 다른 33두 개체를 이용하여 실시하였다.

2.2 통계적 분석 과정

이 연구에 대한 통계적 분석과정은 다음 <그림1>의 과정으로 이루어졌다.



<그림1> DNA 마커 자료의 통계 분석 과정

3. Informativeness(DNA marker 정보력의 측도)

3.1 H(Heterozygosity) & PIC(Polymorphism Information Content)

DNA marker의 정보력이란 개체식별의 측면에서 보았을 때 다양성의 측도라고 할 수 있으며 이 정보력의 측도로 H, PIC를 사용한다.

H(Heterozygosity)는 Locus의 이형접합체는 모집단의 데이터 셋에 있는 이형접합 개체의 비율을 나타내며, 다음과 같이 정의된다.

$$H = (1 - \sum_{i=1}^n p_i^2) \quad (2)$$

단, n = no. of alleles of locus

p_i = estimated allelic frequency of allele i of locus k .

단, 유전자 locus 안에서 대립유전자들 간에 하디-와인버그 평형을 유지해야 한다.

PIC(Polymorphism Information Content)는 부모로부터 자손에 전달되는 대립유전자를 구별해 낼 수 있는 확률이며 다음 (3)식과 같다.

$$PIC = (1 - \sum_{i=1}^n p_i^2 - \sum_{i=1}^{n-1} \sum_{j=i+1}^n 2p_i p_j^2) \quad (3)$$

즉, n = no. of alleles of locus

$\hat{p}_{i(j)}$ = estimated allelic frequency of allele $i(j)$ of locus

이때, allele의 수가 크면, 다양성 정보량(PIC)값도 이형접합체(H)값과 점점 비슷해진다.

<표2> DNA marker의 Allele frequencies

Marker	Alleles amplified										H	PIC
	1	2	3	4	5	6	7	8	9			
BMS2519	μ	0.226	0.016	0.081	0.177	0.242	0.161	0.081	0.016		0.8195	0.7947
	σ	0.053	0.016	0.034	0.048	0.054	0.047	0.034	0.016			
TGLA44	μ	0.097	0.129	0.210	0.210	0.161	0.177	0.016			0.8282	0.8047
	σ	0.037	0.042	0.052	0.052	0.047	0.048	0.016				
ILSTS035	μ	0.059	0.324	0.147	0.088	0.088	0.206	0.088			0.8043	0.7824
	σ	0.0410	0.081	0.0616	0.049	0.049	0.070	0.049				
BMS1242	μ	0.206	0.029	0.206	0.235	0.118	0.206				0.8027	0.7856
	σ	0.0704	0.029	0.070	0.073	0.056	0.070					
HUJ223	μ	0.076	0.106	0.258	0.015	0.106	0.227	0.076	0.121	0.015	0.8328	0.8125
	σ	0.0339	0.039	0.0560	0.015	0.039	0.053	0.033	0.041	0.015		
BMS941	μ	0.016	0.032	0.177	0.29	0.113	0.065	0.161	0.081	0.065	0.8296	0.8095
	σ	0.016	0.022	0.048	0.058	0.040	0.031	0.047	0.034	0.031		
BMS1675	μ	0.061	0.039	0.212	0.121	0.076	0.076	0.015	0.045		0.7677	0.7401
	σ	0.029	0.061	0.051	0.041	0.032	0.032	0.015	0.025			
BM3507	μ	0.081	0.097	0.339	0.048	0.242	0.194				0.7706	0.7506
	σ	0.034	0.037	0.060	0.027	0.054	0.050					
HW-YU2-CA-95	μ	0.207	0.138	0.069	0.293	0.293					0.7615	0.7480
	σ	0.053	0.045	0.033	0.060	0.060						
HW-YU2-CA03	μ	0.183	0.183	0.017	0.333	0.233	0.050				0.7651	0.7285
	σ	0.050	0.050	0.016	0.061	0.055	0.028					

4. Evaluation

개체식별에 유용하다고 선택된 DNA marker들의 검정력을 평가하기 위해서는 MP, R을 사용한다.

4.1 MP(Match probability) & R (Relatedness coefficient)

MP(Match Probability)는 독립인 두 동물 사이에 동일한 allele을 가질 확률 값으로, 유전자형의 도수들은 독립성이 보장되어야 함을 전제조건으로 하며, 다음 (4)식과 같이 정의 된다.

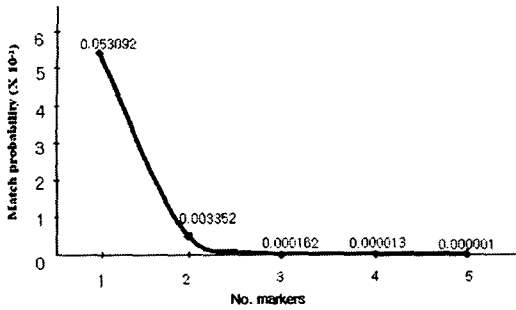
$$MP = \prod_{k=1}^m \left(\sum_{i=1}^{n_k} (p_{k_i}^2) + \sum_{i=1}^{n_k} \sum_{j=i+1}^{n_k} (2p_{k_i} p_{k_j})^2 \right) \tag{4}$$

단, m = loci의 수

n_k = locus k의 allele의 수

p_{k_i} = locus k의 i allele의 추정된 도수 비율

p_{k_j} = locus k의 j allele의 추정된 도수 비율



사용된 마커 수	DNA marker 조합	동물 오분류 확률(%)
1	CA-03	13.7446
2	CA-03 BMS2519	4.38093
3	CA-03 BMS2519 ILSTS035	3.71657
4	CA-03 BMS2519 ILSTS035 BMS941	0.11109
5	CA-03 BMS2519 ILSTS035 BMS941 BMS1675	0.022649

<그림 2> 독립인 두 동물 사이에 동일한 allele을 가질 확률값(MP) (확률 값 X10⁻²)과 마커조합

R 계수는 관련 동물들 사이의 닮음의 정도를 측정하는 도구로 사용되어져 왔으며, 개별 혹은 자연적인 동물들 사이에 관련된 경험적인 확률(empirical probability)을 나타낸다(Queller and Goodnigh, 1989).

$$R(i, j) = \frac{v(i, j)}{w(i)}$$

$$v(i, j) = \sum_{k=1}^2 \sum_{p=1}^2 \sum_{l=1}^r \sum_{m=1}^{s_r} \begin{cases} 0 & \text{if } x_k \neq y_p \neq m \\ [1-f(m, l)] & \text{if } x_k = y_p = m \end{cases}$$

$$w(i) = \sum_{k=1}^2 \sum_{l=1}^r \sum_{m=1}^{s_r} \begin{cases} 0 & \text{if } x_k \neq m \\ [1-f(m, l)] & \text{if } x_k = m \end{cases} \quad (5)$$

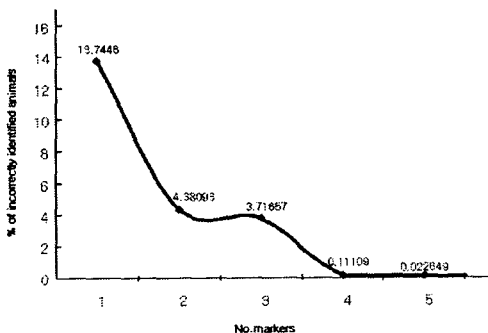
- 즉, r = 마커의 수
- s_r = 마커 r 의 alleles의 수
- m = 마커 r 의 각각의 alleles
- x_k = i 번째 동물에서 나타나는 allele, k 번째 위치, l 번째 locus
- y_p = j 번째 동물에서 나타나는 allele, p 번째 위치, l 번째 locus
- $f(m, l)$ = l 번째 locus에 마커 m 의 allele 도수 비율

4.2 Percentage of animal incorrectly identified

R값이 0.85보다 크면 더 많은 DNA 마커의 사용이 고려되어야 한다. (Cunningham et al.(1999)) 이에 Percentage of animal incorrectly identified는 다음과 (6)식과 같은 공식으로 구할 수 있다.

Percentage of animal incorrectly identified = $\frac{I_{A(R)}}{\text{Total number of compares}}$ (6)

$$I_A = \begin{cases} 1 & R \in A \\ 0 & R \notin A \end{cases} \quad A = \{R | R > 0.85\}$$



사용된 마커 수	DNA marker 조합	MP
1	TGLA44	0.053092
2	TGLA44 ILSTS035	0.003352
3	TGLA44 ILSTS035 HUJ223	0.000162
4	TGLA44 ILSTS035 HUJ223 BMS1675	0.000013
5	TGLA44 ILSTS035 HUJ223 BMS1675 CA-03	0.000001

<그림 3> DNA 마커조합에 의한 R 계수의 Percentage of animal incorrectly identified값(확률값 X10⁻²)과 마커조합

5. 결론

국가 후대 검정우 26-29차 36가계 305두의 Chromosomes에서 155개의 DNA 마커를 사용하여 개체식별에 유용한 DNA 마커(BMS2519, TGLA44, ILSTS035, BMS1242, HUI223, BMS941, BMS1675, BM3507 HW-YU2-CA-95, HW-YU2-CA03)를 찾았고, 그 조합들 중 MP값과 R값을 고려했을 때 개체식별에 가장 유리한 조합은 BMS2519, ILSTS035, BMS941, BMS1675, HW-YU2-CA-03으로 밝혀졌다. 이 조합을 305두에 적용한 동물 오분류 확률은 대략적으로 10,000두 중 2두 정도로 나타났다. 따라서 이 5개의 마커로 당장 현실 개체식별에 적용하기에는 다소 불충분하다고 볼 수 있다. 그러므로 이 실험의 결과를 현실 자료에 적용하기 위해서는 마커 수를 약간 더 늘려 분석해야 한다는 결론을 내릴 수가 있다.

참고자료 및 문헌

1. Arana, A., Soret, B., Lasa, I and Alfonso L. (2002) Meat traceability using DNA markers : application to the beef industry. *Meat Science*. 61, 367-373.
2. Cunningham, E., Meghen, C., Scott, C., Bradley, D., MacHugh, D., Loftus, R. (1999) DNA Traceability techniques for the meat industry. In F. Toldra and Troy D.(Eds.) New developments in guaranteeing the optimal sensory quality of meat. 1-7 *Valencia*(Spain).
3. GLowatzki-Mullis, M., Gaillard, C., Wigger, G. and Fries, R. (1995) Microsatellite-based parentage control in cattle. *Anim. Genorn.* 26, 7-12.
4. Lim, H., Min, H., Moon W., Lee, J., Kim, J. Cho, I., Lee, H., Lee, Y., Lee, J., and Jeon, J. (2005) Analysis and selection of microsatellites markers for individual traceability system in Hanwoo. *J. Anim. Sci & Technol.* 47-4, 491-500.
5. Yoon, D., Kong, H., Oh, J., Lee, J., Cho, B., Kim, J., Jeon, K., Jo, C., Jeon, G. and Lee, H. (2005) Establishment of an individual identification system based on microsatellite polymorphisms in korean cattle(Hanwoo). *Asian-Aust. J. Anim. Sci.* 18-6, 762-766.
6. Ott. Jurg. (1999) *Analysis of human genetic linkage, 3rd ed.* The Johns Hopkins university press.
7. Peelman, L., Mortiaux, F., Zeveren, A., Dansercoer, A., Mommens, G., Coopman, F., Bouquet, Y., Burny, A., Renaville, R and portetelle, D. (1998) Evaluation of the genetic variability of 23 bovine microsatellite markers in Four Belgian cattle breeds. *Anim. Genet.* 29, 161-167.
8. Queller, D. C., & Goodnigh, K. F. (1989) Estimating relatedness using genetic markers. *Evolution* 43-2, 258-275
9. Weir, B. (1996) *Genetic data analysis II. Methods for discrete population genetic data*. Sunderland, Massachusetts: Sinauer Associates, Inc. Publishers.
10. Weiss, K. (1993) *Genetic variation and human disease. principles and evolutionary approaches*. Cambridge, Great Britain: Cambridge University Press.