

## Research Article

# 국내 Holstein과 Jersey종 젖소집단의 고밀도 유전체 정보 및 연관불평형 비교

서정우<sup>1</sup>, 이학교<sup>2</sup>, 박경도<sup>2</sup>, 이준호<sup>1\*</sup>

<sup>1</sup>주식회사 진앱스, <sup>2</sup>전북대학교 동물생명공학과

## Comparison of SNP variation and extent of linkage disequilibrium in Holstein and Jersey cattle populations in Korea

Jeong-Woo Seo<sup>1</sup>, Hak-Kyo Lee<sup>2</sup>, Kwangdo Park<sup>2</sup>, Joonho Lee<sup>1\*</sup>

<sup>1</sup>GeneApps Co., Ltd., Gangnam, 06105, Korea

<sup>2</sup>Department of animal biotechnology, Jeonbuk National University, Jeonju, Korea

\*Corresponding author: Joon-Ho Lee, GeneApps Co., Ltd., Gangnam, 06105, Korea Tel : +82-2-556-2895, E-mail : [joonho@gene-apps.com](mailto:joonho@gene-apps.com)

### ABSTRACT

Recently, Jersey cattle were introduced and produced in Korea. The number of introduced Jersey cattle was relatively small compared to the Holstein breed, and there are few related studies. This study was conducted to investigate the differences in single nucleotide polymorphism (SNP) variations and to characterize linkage disequilibrium (LD) between Jersey and Holstein populations in Korea. High-density SNP data were collected from 120 Jersey cattle and 921 Holstein cattle. SNP genotypes were identified using Illumina BovineSNP50 version3 BeadChip, which covers 53,218 SNPs located across all chromosomes. For a fair comparison of LD measurement, 29,127 markers which were common after quality control in both breeds were selected. The allele frequencies of all SNP markers were compared in both breeds, especially for the SNPs which showed segregation in only one breed. The measures of the LD ( $r^2$ ), the square of the correlation coefficient between the alleles at two loci that are carried by haplotypes, were calculated. Relatively small genetic diversity was found in Jersey population by the statistics of allele frequencies, heterozygosity,  $r^2$  between adjacent markers, and effective population size

**Key words:** Allele Frequency, Effective Population Size, Holstein, Jersey, Linkage Disequilibrium

### Introduction

우유를 생산하기 위하여 도입된 젖소 품종으로는 잘 알려진 홀스타인종(Holstein) 이외에 저지종(jersey), 건지종(Guernsey) 등이 있으나 국내 우유생산은 홀스타인종에 매우 집중되어 있다(Lim, 2017). 국내 최초의 젖소 도입에 대한 역사를 살펴보면 대한제국시기인 1885년에 저지종 3두를 수입하는 등의 품종 다양화가 가능했을 것으로 보이는 기록이 남아있으나 높은 유량과 내한성에 특화된 홀스타인종이 국내 사육환경에 적합하기 때문에 국내 젖소 집단은 홀스타인종으로 획일화된 경향을 보인다(Lee, 2016; Lim, 2017). 저지종은 홀스타인종에 비해 우유 생산량은 적지만, 사료효율이 높아 사료섭취량 대비 고품분 비율이 높은 우유를 생산할 수 있는 것으로 알려져 있으며(Lim, 2016; Lim, 2020), 특히 내서성(Heat Tolerance)이 강한 품종으로 여름 평균 기온이 상승하는 우리나라의 기후 변화에 대비할 수 있는 젖소 품종 후보군이 될 수 있다(Smith et al., 2013; Lim et al., 2021).

Received December 15, 2022

Revised March 21, 2023

Accepted March 21, 2023

Copyright © 2023 Journal of Animal Breeding and Genomics.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

국내 저지종의 도입은 2010년 젖소 품종의 수입 제한이 개정된 이후에 본격화한 것으로, 수입된 개체수가 적어 집단의 유전적 구성(Gene Pool) 아직까지는 매우 작다고 할 수 있다. 규모가 작은 유전자원의 경우 다형성(Polymorphism) 유실의 비율이 높기 때문에 유전적 다양성 유지를(Genetic Diversity) 위해서는 더 세밀한 관리가 필요하다고 할 수 있다(Stachowicz, 2011). 작은 집단에서는 유전적 부동에 의한 변이 유실이 발생하며, 혈연관계가 존재하는 개체간 교배가 이루어져 집단의 근교계수가 상승할 수 있다(Falconer, 1981).

일배체형 유전체 정보를 이용한 연관불평형(Linkage Disequilibrium; LD)의 추정치는 집단 내 유전적 다양성의 정도를 나타낼 수 있으며, 동일한 조건에서 계산된 연관불평형은 집단간 유전적 다양성의 상대적 비교에 이용될 수 있다. 연관불평형의 추정은 SNP 또는 MS (Microsatellite) 표지인자를 통해 다양하게 측정 가능하며, 집단의 유전적 다양성이 클 경우 일배체형 구성을 이루는 두 마커간 교차 현상이 발생한 개체들이 포함될 확률이 높아지므로 연관불평형은 반대로 낮게 추정된다(Qanbari et al., 2010). 연관불평형 추정은 유전체 선발의 타당성을 간접적으로 증명하거나 양적형질좌위(QTL) 발굴의 도구로 사용되기도 한다(de Roos et al., 2008; Frances et al., 2020). 또한 일배체형 구성을 이루는 두 마커간의 물리적 거리 별 연관불평형을 통하여 유효집단 크기를 추정할 수 있으며, 이는 현 집단 상황의 유전체 정보에서 드러나는 유전적 집단 규모를 가늠할 수 있게 한다(Cho et al., 2012; Do et al., 2014).

최근에 소규모로 도입되어 생산이 활성화된 저지종 집단이기 때문에 홀스타인 품종에 비하여 상대적으로 연구가 잘 이루어지지 않았으며, 특히 유전 육종에 관련된 국내 연구는 많이 부족한 실정이다. 본 연구의 목적은 소규모 신품종인 저지종 유전자원 관리에 유전체 정보를 이용함으로써 더 정밀하고 직접적인 관리를 가능하게 하는 것으로, 기존의 등록정보 및 혈통관리 뿐만 아니라 고밀도 유전체정보를 통하여 집단의 유전적 구성과 현황을 더 명확하게 판단하기 위함이다. 이를 위하여 현재 국내 우점 품종인 홀스타인종과 신품종인 저지종 집단의 유전체 정보를 수집하여 유전적 다양성과 유효집단 크기를 직접 비교하였고 이는 유전자원관리를 위한 중장기적 계획 수립에 도움이 되는 기초 자료로 활용될 수 있을 것으로 예상된다.

## Materials and Methods

### 유전체자료

본 연구에 이용된 고밀도 유전체 정보는 홀스타인종 집단과 저지종 집단 모두에서 같은 종류의 SNP chip을(Illumina BovineSNP50 version3) 이용하여 분석된 유전자형 자료로 총 SNP 마커 수는 53,218개이고 저지종은 120두, 홀스타인종은 921두가 분석에 이용되었다. 유전자형 정보를 finalReport 형식의 텍스트 파일에서 추출하여 행렬의 형태로 전환하였으며, 각 집단에서 유전자형 빈도 및 유전자 빈도를 계산하여 비교하였다. 두 집단의 유전체 정보를 비교하기 위하여 한 쪽 품종에서만 유전자형이 고정된 경우를 탐색했으며, 이후 유전체 전장의 품종집단 별 유전자 빈도 편차를 계산하였다.

### 연관불평형 분석

집단의 다양성을 추가적으로 분석하기 위하여 일배체형으로 구성되어지는 두 좌위의 대립유전자간 빈도 상관계수의 제공으로 나타내어지는 연관불평형 분석을 실시하였다. 연관불평형 분석을 위해서 두 품종 모두에서 MAF (minor allele frequency)가 0.05 이상, 결측률이 10% 이하, 상염색체에 존재하면서 물리적 위치정보(SNP position)가 존재하는 마커들만 추출하여 이용하였다. 이 과정에서 두 품종에서 선택되어지는 SNP 마커가 서로 다르게 나타나기 때문에 두 품종에서 중복되는 마커들을 모은 결과, 29,127개 마커가 선택되었다.

동일 염색체 내 연관불평형을 추정하기 위한  $r^2$ 는 아래와 같이 계산하였다.

$$r^2 = \frac{D^2}{p(A_1) \times p(A_2) \times p(B_1) \times p(B_2)}$$

여기서,  $p(A_1), p(A_2), p(B_1)$  및  $p(A_2)$ 는 집단 내 A 또는 B 두 좌위의 각 두 유전자 (biallelic) 빈도 추정 자료이며,  $D$ 는 아래와 같다.

$$D = p(A_1B_1) \times p(A_2B_2) - p(A_1B_2) \times p(A_2B_1)$$

여기서,  $p(A_1B_1), p(A_2B_2), p(A_1B_2)$  및  $p(A_2B_1)$ 는 집단 내 A 또는 B 두 좌위의 대립유전자로 구성된 일배체형의 빈도이다. 두 좌위가 모두 이형일 경우, 일배체형이  $A_1B_1/A_2B_2$  인지  $A_1B_2/A_2B_1$  인지 구분할 수 없기 때문에 EM algorithm (Expectation Maximization Algorithm)를 이용하여 조건부 확률을 계산하여 변화량이 일정량 이하로( $10^{-5}$ ) 감소할 때까지 반복 연산하여 일배체형의 빈도를 추정하였다.

### 연관불평형을 이용한 집단의 유효집단크기 추정

Sved (1971)는 돌연변이가 발생하지 않는다는 가정에서  $r^2$ 의 기대치와 유효집단크기( $N_e$ )에 대한 관계식을 아래와 같이 발표하였다.

$$E(r^2) = \frac{1}{1 + 4N_e c}$$

여기서,  $c$ 는 Morgan 단위의 마커간 거리이며, 세대( $t$ ) 별  $N_e$  추정은  $t = \frac{1}{2c}$ 에 1cM(centimorgan)을 1Mb (mega base pair) 적용하여 계산하였다.

## Results and Discussion

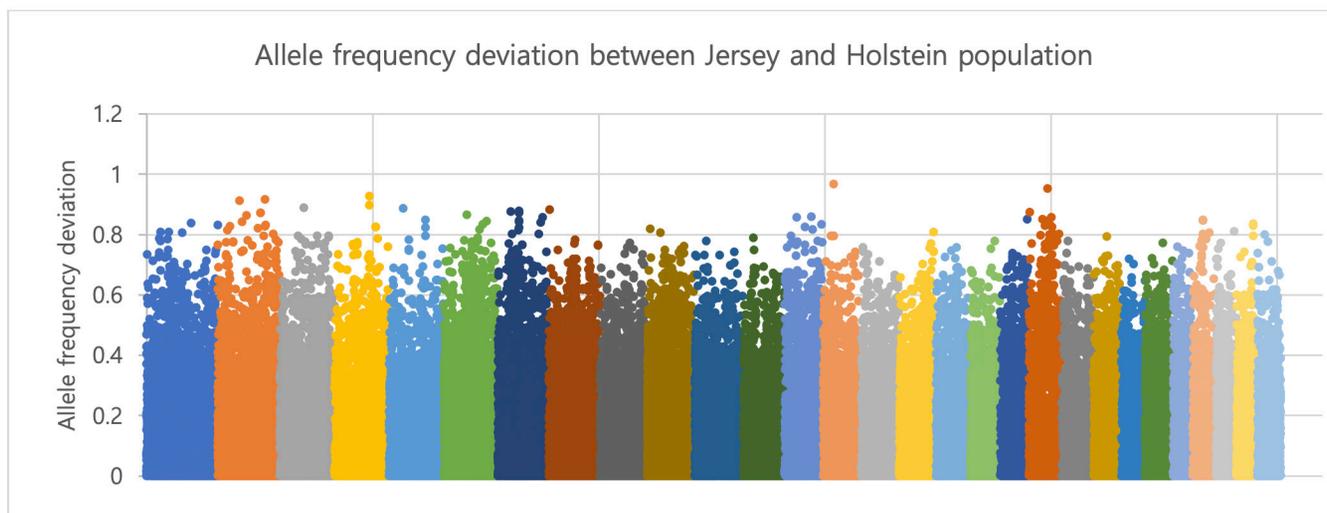
### 품종 특이적 유전적 다양성 차이 비교

두 집단의 유전체 정보를 비교하기 위하여 한 쪽 품종에서만 유전자형이 고정된 경우를 탐색했으며 유전체 전장의 품종집단 별 유전자 빈도를 비교하였다. Table 1에서와 같이 저지종 집단과 홀스타인 집단 모두에서 다형성이 존재하는 많은 마커는 상염색체에 2,726개 발견되었으며, 저지종 집단에서는 다형성이 존재하지 않으나 홀스타인 집단에서는 변이가 존재하는 마커는 10,111개, 반대로 저지종 집단에서만 다형성이 존재하는 마커는 41개만 발견되어 홀스타인 집단에서 유전적 다양성이 훨씬 더 크게 나타나는 것을 확인할 수 있었다.

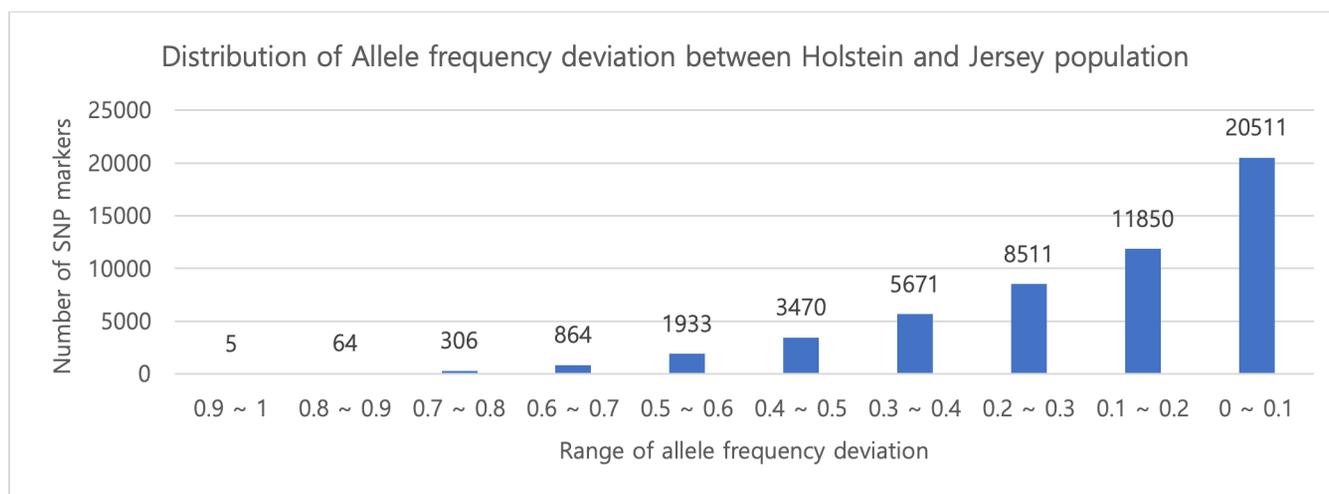
**Table 1.** Number of SNP markers which were only polymorphic in specific breeds

Condition	all SNP markers	markers on autosomal chromosome
no segregation in Jersey	13,275	12,621
no segregation in Jersey / segregation in Holstein	10,502	10,111
no segregation both in Jersey and Holstein	2,773	2,726
no segregation in Holstein / segregation in Jersey	46	41
no segregation in Holstein	2,819	2,551

유전체 전장에서의 품종 별 유전자 빈도 차이를 Fig 1에 표시하였다. Fig 2에서와 같이 대부분의 마커들은 우전자형 빈도 차이가 0에서 0.1로 나타났으나 품종 별 유전자 빈도 차이가 0.8 이상의 큰 편차를 보이는 마커가 64개, 0.9 이상을 나타내는 마커도 5개가 발견되어(Table 2) 이러한 마커를 조합할 경우 품종 특성 구분에 활용될 수 있을 것으로 사료된다. 모든 마커의 평균 이형접합자 빈도 평균도 홀스타인 집단에서는 0.30, 저지 집단에서는 0.26으로 나타나 홀스타인 집단의 이형접합자 빈도가 더 높게 나타났다.



**Figure 1.** Absolute allele frequency deviation between Jersey and Holstein by chromosome (1 to 29)



**Figure 2.** Distribution of allele frequency deviation between Holstein and Jersey breed

**Table 2.** SNP markers which were breed specific (allele frequency deviation over 0.9)

SNP name	Chr.	Position	Heterozygosity deviation	Absolute allele Frequency deviation
BTB-00557585	14	24,607,527	0.007	0.967
Hapmap51736-BTA-50602	20	41,777,888	0.008	0.951
ARS-BFGL-NGS-116590	4	77,635,835	-0.032	0.926
ARS-BFGL-BAC-5705	2	103,707,675	-0.073	0.916
Hapmap48352-BTA-47529	2	47,873,555	0.115	0.911

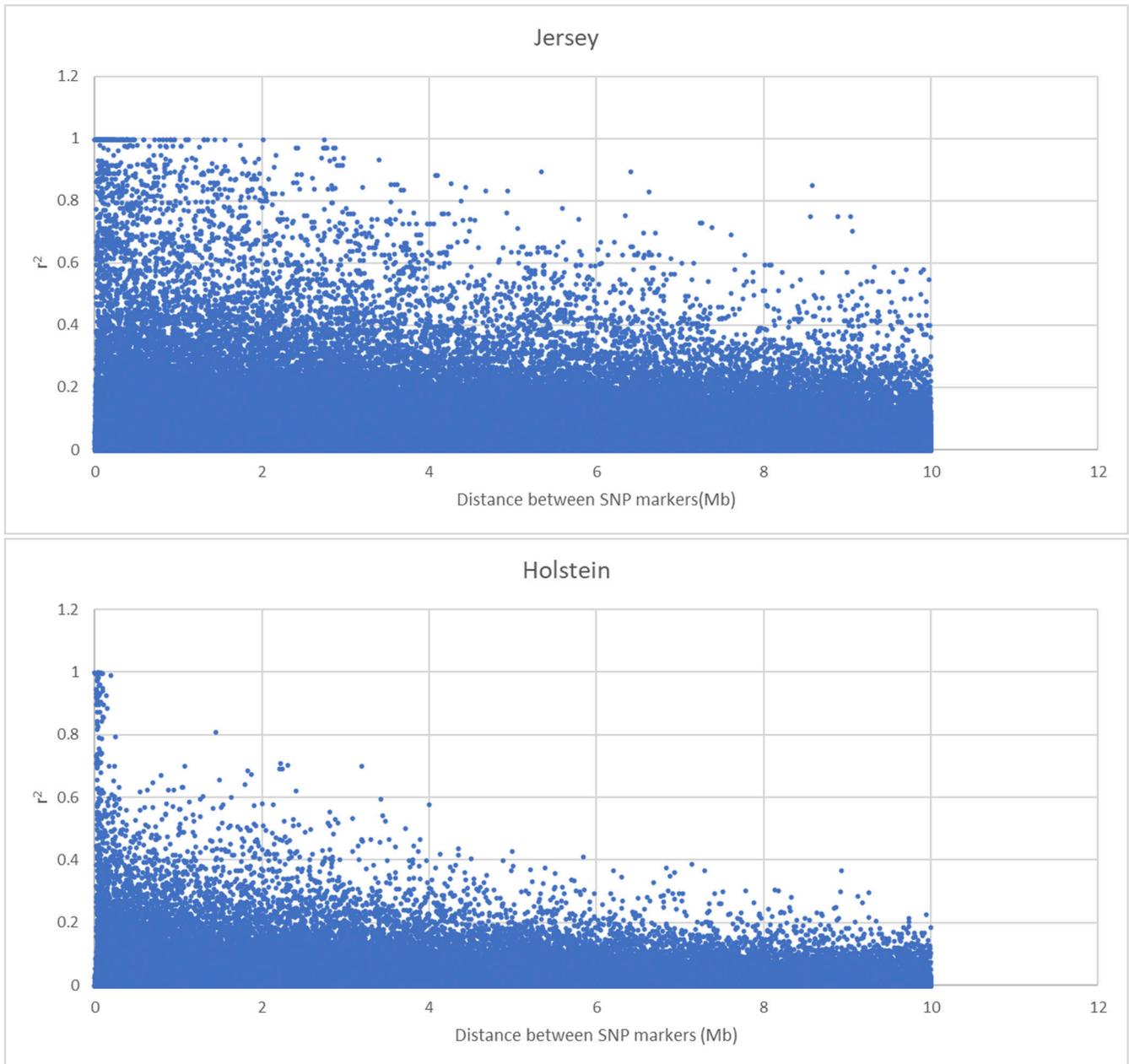
## 품종별 연관불평형 비교

유전체 정보에서 드러난 집단 특이적 특징을 더욱 명확하게 구명하기 위하여 같은 SNP 마커쌍의 조합으로 구성된 일배체형을 이용하여 LD( $r^2$ )를 계산하였다. Table 3에 각 염색체별 일배체형 조합 수,  $r^2$  평균 및 인접 마커간  $r^2$  평균을 품종별로 나타내었다. 저지종 집단의 LD가 모든 마커쌍(SNP marker pair)으로 구성된 일배체형 평균이나 인접 마커쌍으로 구성된 일배체형 평균에서 모두 홀스타인 집단에 비하여 높게 나타났으며, 이는 저지종 집단의 유전적 다양성이 부족하여 마커간 연관이 강하게 나타난 것으로 예상된다. 이러한 현상은 물리적 거리별 LD (LD Decay) 감소를 품종별로 나타낸 Fig 3에서도(29번 염색체 예, 10 Mb까지 표시) 명확하게 드러난다.

연관불평형의 정도를 나타내는  $r^2$ 는 동일 염색체에 존재하는(Syntenic) 마커간 물리적 거리가 멀수록 연관정도가 낮아지는 경향을 보이는데, Fig 3의 홀스타인종에서는 이러한 현상이 명확하게 나타나지만 저지종의 경우 물리적 거리가 먼 마커쌍 사이에서도 일부 연관이 강하게 나타나는 현상이 보인다. 이것은 마커쌍 사이에 교차현상이 일어난 개체수가 집단 내 상대적으로 적은 것으로 그만큼 유전적 다양성이 적은 것을 의미한다.

**Table 3.** Number of Haplotypes, average  $r^2$  (all and adjacent) by chromosome and breed

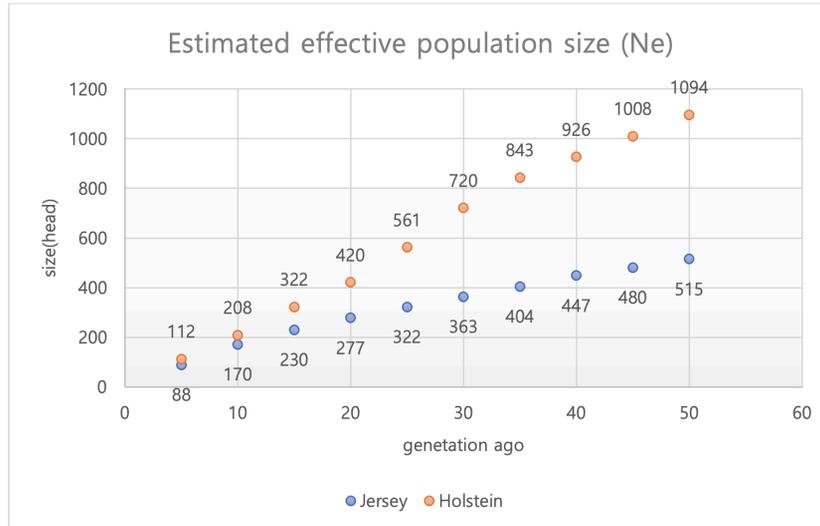
Chromosome	Number of Haplotypes	Jersey		Holstein	
		All	adjacent	All	adjacent
1	1,648,020	0.043	0.374	0.013	0.267
2	1,014,600	0.043	0.387	0.014	0.257
3	937,765	0.044	0.363	0.014	0.246
4	925,480	0.043	0.339	0.013	0.230
5	753,378	0.042	0.325	0.015	0.253
6	1,437,360	0.055	0.437	0.016	0.320
7	1,076,778	0.046	0.398	0.022	0.322
8	808,356	0.045	0.327	0.015	0.251
9	642,411	0.046	0.352	0.016	0.257
10	889,111	0.042	0.346	0.019	0.311
11	816,003	0.040	0.312	0.014	0.233
12	444,153	0.043	0.296	0.016	0.227
13	352,380	0.046	0.322	0.018	0.232
14	539,241	0.046	0.323	0.018	0.274
15	450,775	0.044	0.318	0.015	0.211
16	491,536	0.052	0.326	0.020	0.294
17	409,060	0.049	0.292	0.017	0.238
18	297,606	0.045	0.335	0.015	0.216
19	289,180	0.049	0.315	0.018	0.231
20	389,403	0.076	0.452	0.024	0.306
21	339,076	0.051	0.339	0.018	0.256
22	238,395	0.048	0.299	0.018	0.232
23	218,791	0.051	0.305	0.022	0.251
24	255,255	0.051	0.312	0.019	0.235
25	141,778	0.057	0.300	0.021	0.209
26	167,331	0.053	0.308	0.023	0.224
27	154,290	0.053	0.303	0.019	0.192
28	122,760	0.053	0.265	0.019	0.184
29	182,106	0.059	0.299	0.021	0.200
Total		0.047	0.344	0.017	0.256



**Figure 3.** LD decay plotting (BTA(Chromosome) 29) of Holstein breed (upper) and Jersey breed (lower).

#### 유효집단 크기 차이

연관불평형을 이용하여 품종 별 유전체 유효집단 크기를 추정된 결과 Fig 4에서와 같이 저지종의 유전체유효집단 크기가 훨씬 작은 것으로 나타났다. 이것은 유전적 다양성이 낮게 나타난 이전 결과자료와 일치하는 결과이지만 Rincon 등(2011)의 연구결과에서도 저지종의 유전적 다양성이 홀스타인종 보다 낮게 나타나, 저지종의 유전적 다양성이 적은 것이 저지종 특이적 특성일 수 있다. 그러나 종 특이적 특성은 많은 세대 이전의 유전체 유효집단 크기에서 반영되는 결과이며, 최근 세대의 경향을 나타내는 1~5세대 유효집단 크기도 홀스타인종에서 112두이지만 저지종에서 88두로 나타나 국내 저지종 집단의 현 유효집단크기도 작은 것으로 보인다.



**Figure 4.** Historical effective population size (Ne) in Jersey and Holstein breed

## Conclusion

유전체 정보를 이용한 여러가지 집단유전학적 지표에서 국내 저지종 유전자원의 다양성(Genetic Diversity)이 부족하다는 결과를 보여주고 있다. 현재 도입된 저지종의 개체수도 작을 뿐 아니라 연구에 이용된 표본 수도 120두로 정밀한 유전체 자료분석에는 다소 부족한 면이 존재하나 집단의 개량 측면이나 근친도 관리 측면에서 유전적 다양성 유지는 매우 중요하며, 특히 초기 집단에서의 유전적 다양성 유실은 후대 개체들의 유전적 구성에 큰 영향을 미칠 것이기 때문에 초기 집단의 유전적 다양성을 유전체 정보로 확인하는 것이 중요하다고 할 수 있다. 우리나라의 사육환경이 점차 변화되고, 이에 대응하기 위하여 새로운 품종을 도입하려는 시도는 유전자원의 다양화 측면에서 바람직하다고 볼 수 있다. 새로운 품종의 도입에는 검정, 등록 및 혈통기록 체계 개발이 우선적으로 선행되어야 하며, 더불어 유전집단 유지 및 개량을 위해서는 집단유전학적 분석을 수행해야 하는데 여기에 유전체 정보를 활용하는 것이 명확하고 효율적인 관리를 가능하게 할 것이다. 현재 저지종 집단의 유전적 다양성 유지를 위해서는 더 많은 유전자원이 외부로부터 도입되어야 할 것으로 사료되며, 집단이 구성되는 초기 단계이므로 집단 내 교배를 실시할 경우 유전적 거리가 먼 개체들로 구성된 교배조합을 구성해야 한다. 유전체 자료를 이용하여 이형접합자 비율을 최대화하는 상보적 교배조합 개발이 집단의 효율적인 유전적 다양성 유지를 위하여 필요하며, 더 많은 개체의 유전체 정보 확보가 필요할 것으로 보인다.

## Acknowledgements

본 논문은 농촌진흥청 연구사업 (세부과제번호: PJ01620003)의 지원에 의해 이루어진 것임.

## 요약

최근 국내에 저지 젖소 품종이 도입되어 사육되고 있으나 도입된 저지소의 개체수는 홀스타인 품종에 비해 상대적으로 적고 관련 연구도 부족한 상황이다. 본 연구는 국내 저지종 집단과 홀스타인종 집단에서 구성한 고밀도 유전체 정보의 유전적 다양성 및 연관불평형의 차이를 비교를 통하여 집단유전학적 특성을 비교하기 위하여 수행하였다. 고밀도 유전체 정보는 홀스타인종 집단과 저지종 집단 모두에서 같은 종류의 SNP chip을(Illumina BovineSNP50 version3) 이용하여 분석된 유전자형 자료로 총 SNP 마커 수는 53,218개이고 저지종은 120두, 홀스타인종은 921두가 분석에 이용되었다. 연관불평형에 대한 공평한 비교를 위하여 품질평가를 완료한 뒤 두 품종에서 공통된 29,127개의마커를 선택하였다. 유전체 전장의 모든 SNP 마커의 유전자 빈도를 비교하였으며, 특히 한 품종에서만 다형성이 존재하는 마커들을 탐색하였다. 일배체형으로 구성되어지는 두 좌위의 대립유전자간 빈도 상관계수의 제공으로 나타내어지는  $r^2$ 의 계산을 통하여 품종별 연관불평형을 추정하였다. 대립유전자빈도, 이형접합성, 일배체형의  $r^2$  및 유효집단 크기 비교에서 저지종의 유전적 다양성이 작은 것으로 나타났다.

**검색어:** 저지, 홀스타인, 대립유전자 빈도, 연관불평형, 유효집단크기

## Reference

- Lim DH. 2017. Milk composition and characteristics of Jersey cows. *Food Sci. Anim. Resour.* 6(1): 5-11.
- Lee KJ. 2016. Production and Supply of Milk in Joseon during Japanese Colonial Period (1910~1945). *J. Korean Soc. Food Cult.* 31(5): 400-410
- Lim DH, Mayakrishnan V, Lee HJ, Ki KS, Kim TI and Kim YH. 2020. comparative study on milk composition of Jersey and Holstein dairy cows during the early lactation. *J. Anim. Sci. Tech.* 62(4):565-576. doi: 10.5187/jast.2020.62.4.565.
- Smith DL, Rude BJ and Ward SH. 2013 Short communication: Comparison of the effects of heat stress on milk and component yields and somatic cell score in Holstein and Jersey cows. *J. Dairy Sci.* 96(5): 3028-3033. doi:/10.3168/jds.2012-5737
- Lim DH, Kim TI, Park SM, Ki KS and Kim YH. 2021. Evaluation of heat stress responses in Holstein and Jersey cows by analyzing physiological characteristics and milk production in Korea. *J. Anim. Sci. Tech.*63(4):872-883. doi: 10.5187/jast.2021.e62
- Stachowicz K, Sargolzaei M, Miglior F and Schenke FS. 2011. Rates of inbreeding and genetic diversity in Canadian Holstein and Jersey cattle. *J. Dairy Sci.* 94 :5160–5175 doi: 10.3168/jds.2010-3308
- Falconer DS. 1981. Introduction to quantitative genetics, Ed. 2. Longmans Green, London/New York.
- Qanbari S, Pimentel ECG, Tetens J, Thaller G, Lichtner P, Sharifi AR and Simianer H. 2010. The pattern of linkage disequilibrium in German Holstein cattle. *Anim. Genet.* 41:346-356 doi: 10.1111/j.1365-2052.2009.02011.x
- de Roos APW, Hayes BJ, Spelman RJ and Goddard ME. Linkage disequilibrium and persistence of phase in Holstein–Friesian, Jersey and Angus cattle. 2008. *Genetics.* 179(3): 1503–1512. doi: 10.1534/genetics.107.084301
- Thistlethwaite FR, El-Dien OG, Ratcliffe B, Klápště J, Porth I, Chen C, Stoehr MU, Ingvarsson PK and El-Kassaby YA. 2020. Linkage disequilibrium vs. pedigree: Genomic selection prediction accuracy in conifer species. *Plos One.* Doi: 10.1371/journal.pone.0232201
- Cho CI, Lee JH and Lee DH. 2012. Estimation of linkage disequilibrium and effective population size using whole genome single nucleotide polymorphisms in Hanwoo. *J. Life Sci.* 22(3):366-372
- Do KT, Lee JH, Lee HK, Kim J, Park KD. 2014. Estimation of effective population size using single-nucleotide polymorphism (SNP) data in Jeju horse. *J. Anim. Sci. Tech.* 56(8):1-6. doi: 10.1186/2055-0391-56-28
- Sved JA. 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite population. *Theor. Popul. Biol.* 2:125-141
- Rincon G, Weber KL, Van Eenennaam AL, Golden BL, and Medrano JF. 2011. Hot topic: Performance of bovine high-density genotyping platforms in Holsteins and Jerseys. *J. Dairy Sci.* 94:6116–6121. doi: 10.3168/jds.2011-476