

Research Article

한우 유전자칩(Hanwoo 50K chip)을 이용한 전라북도 한우의 연관불평형 분석

유지석¹, 김도현², 이학교², 오재돈^{2*}¹전북대학교 농축산식품융합학과, ²전북대학교 동물생명공학과

A study on the estimation of Linkage Disequilibrium using SNP(Hanwoo 50K chip) of Hanwoo in Jeollabuk-do

Ji-suk Yu¹, Do-Hyun Kim², Hak-kyo Lee², and Jae-Don Oh^{2*}¹Dept of Agriculture and Livestock Food Convergence, Jeonbuk National University, Jeonju 54896, Korea²Dept of Animal Biotechnology, Jeonbuk National University, Jeonju 54896, Korea***Corresponding author:** Jae-Don Oh, Dept of Animal Biotechnology, Jeonbuk National University, Jeonju 54896, Korea. E-mail : oh5ow@naver.com

ABSTRACT

This study was conducted with the Hanwoo SNP chip (50k_v1) analysis to investigate genotypes of SNPs of Hanwoo (Korean native cattle) in Jeollabuk-do and extent of linkage disequilibrium (LD) among SNP markers. Genomic DNA for genotyping was extracted from the hair root samples of Hanwoo cows (n=1,523). A total of 52,195 SNPs were obtained and 11,056 unnecessary SNPs were removed through the QC process, and a total of 41,139 SNPs were used for LD analysis. The total length of the SNP markers used in the analysis was 2,500 Mb, the mean of minor allele frequency was 0.27 and the average interval distances of adjacent SNP markers by each chromosome was 0.0463 to 0.0725. The results of this study which confirmed by using Hanwoo in Jeollabuk-do, showed a pattern similar to previously reported studies for Hanwoo. The r^2 value was 0.234 for the SNP distance between 0 and 50 Kb and was 0.068 for distance between 150 and 200 Kb. Therefore this study could be utilized as baseline data using genetic improvement for Hanwoo industry.

Keywords: Linkage Disequilibrium, Hanwoo, SNP chip, Jeollabuk-do, Genetic diversity

INTRODUCTION

한우의 개량사업은 후대검정사업을 통해 보증씨수소를 선발하고, 선발된 보증씨수소의 정액을 전국 한우 농가에 공급하면서 한우 산업 전반의 유전적 개량량을 증대시키며 농가 소득향상에 기여하고 있다. 특히 보증씨수소의 선발체계에 유전체를 이용한 유전능력 평가 방법을 활용하면서 평가 정확도가 상승하였고 유전적 개량량 증대에 기여하였다(Park et al. 2019). 농촌진흥청 축산과학원은 한우의 생산능력을 보다 정확하게 예측할 수 있는 한우 유전자칩(Hanwoo 50K SNP BeadChip_v1)을 개발하였다. 이 한우 유전자칩은 기존에 사용되었던 bovine 50K SNP chip에서 한우 품종에서 분석 효율성이 떨어지는 SNP를 제거하고, 한우 품종의 특성을 반영한 SNP와 유전질병에 대한 평가가 가능한 SNP 정보를 추가하여 개발되었다. 그리고 최근에는 도체형질 연관 SNP와 미토콘드리아 SNP 및 유전 질병 연관 SNP를 보강한 한우 유전자칩2(Hanwoo 60K SNP BeadChip_v2)를 개발하여 산업현장에 보급 중에 있다. 현재 보증씨수소의

유전체 선발은 한우 유전자칩을 이용해 이루어지고 있다. 또한 보다 정확한 유전능력 평가를 위한 분석모델의 개발 연구와 다양한 축군에 적용을 통한 산업화 연구도 활발하게 진행되고 있다(Lee et al. 2019; Lopez et al. 2020; Byun et al. 2021; Kim et al. 2022; Lee et al. 2022).

연관불편형(Linkage Disequilibrium: LD)은 염색세상에 두 대립유전자의 비임의적 연관(non-random association)으로 두개의 다른 대립유전자자가 기대치 이상으로 연관되어 나타나는 것을 의미하며, 양적형질 유전자좌위(QTL)의 위치를 탐색하는데 유용하게 활용되고 있다(Meuwissen & Goddard, 2000). 가족의 유전적 형질과 양적형질 유전자좌위(QTL) 등이 집단 내에서 재조합 될 가능성(유전자 이동, 선발, 돌연변이 등)을 추정하고, 연관불편형정도를 이용한 유전적 구조 및 특성의 분석 연구는 경제형질 개량의 효율성 강화를 위해 매우 중요한 자료를 제공하게 된다.

최근 산업축군에서도 한우 유전자칩을 이용한 유전능력 평가 및 선발을 위한 개량사업이 다양한 환경에서 시작되고 있다. 한우 산업축군의 유전체 선발은 시작단계에 있어 다양한 축군에 대한 분석정보가 아직 충분하지 못한 상태이다. 따라서 본연구는 전라 북도의 한우 축군을 대상으로 한우 유전자칩 정보의 연관불편형(LD)을 추정하고 분석하여 지역 한우 축군의 개량을 위한 기초자료로 활용하고자 수행하였다.

MATERIALS AND METHODS

1. 공시재료

본 연구의 공시재료는 전라북도 지역에서 사육되고 있는 1,523두 한우의 꼬리털의 모근에서 채취한 DNA를 이용하였다.

2. 유전자형분석

공시재료에서 추출한 DNA Sample은 농촌진흥청 축산과학원에서 개발된 한우유전자칩(Hanwoo 50K SNP BeadChip_v1)을 이용하여 유전자형 분석을 진행하였다. 유전자형 정보를 GenomeStudio 2.0 software (Illumina Inc, 2016)를 사용하여 Final report 파일을 출력하였고, Perl language를 이용하여 PLINK formatting 과정 후 ped, map file을 작성하였다. 유전자형 분석을 통하여 얻은 총 SNP 수는 52,195이며, Quality control (QC)를 진행하여 조건에 부합하지 않는 SNP를 제거하였다(Table 1). QC 기준은 SNP의 Call rate가 90% 미만인 SNP(1,066개), MAF(Minor Allele Frequency)가 1% 미만인 SNP (9,551개), HWE(Hardy-Weinberg Equilibrium)가 1×10^{-7} 미만인 SNP (439개) 정보를 제거하여 최종적으로 41,605개의 SNP 정보를 사용하였다(Table 2).

Table 1. Test Hanwoo population and Single Nucleotide Polymorphism (SNP)

Materials	No. of marker
Hanwoo	1,523
Total SNP	52,195
Unnecessary SNP	11,056
Used for Analysis SNP	41,139

Table 2. Quality Control(QC) criteria of Single Nucleotide Polymorphism(SNP)

List	Quality control	Thresholds	No. of marker
SNP	Call rates	< 0.90%	1,066
	Minor Allele Frquency	< 0.05%	9,551
	Hardy-Weinberg Equilibrium	< 1×10^{-7}	439

3. 연관불편형 추정

연관불편형의 크기는 서로 다른 두 대립유전자가 서로 연관되어 유전되는 정도를 나타내는 값으로 먼저 D값을 구한 후에 이를 표준화를 진행한 D' 또는 r²값을 이용하여 추정이 가능하다. 통계값 D를 이용한 연관불편형 추정은 집단의 크기가 작거나 대립유전자빈도가 작을 경우 과대추정이 일어날 수 있기 때문에 D 값을 이용하여 두 대립유전자 간의 상관을 보여주는 통계량인 r²값을 추정하였다. 두 개의 마커 간 연관불편형을 계산하기 위하여 Plink1.9 program(Purcell *et al.*, 2007)을 이용하여 분석하였다. freq A_i, freq A_j, freq B_i, freq B_j는 두 마커의 대립유전자의 빈도를 나타낸다. D 값은 다음 같은 수식을 이용하여 계산하였다.

$$D = \text{freq } A_i B_j - \text{freq } A_i \times \text{freq } B_j$$

동일 염색체 상에 존재하는 두 마커 좌위 A와 B의 연관불편형 크기(r²)는 다음의 수식을 이용하여 계산하였다. 아래의 수식에서 freq A_iB_j, freq A_jB_i, freq A_iB_i, freq A_jB_j는 두 마커 사이에 대립유전자로 구성된 haplotype 빈도를 나타낸다. 각 분석으로 얻어진 연관불편형 자료는 R software 패키지(The R Project for Statistical Computing, ver.4.1.2; <http://www.r-project.org>)를 사용하여 수치와 그래프를 나타냈다.

$$r^2 = \frac{D^2}{(\text{freq } A_i \times \text{freq } A_j - \text{freq } B_i \times \text{freq } B_j)}$$

RESULTS AND DISCUSSION

본 연구는 농촌진흥청 축산과학원에서 개발한 한우 유전자칩을 이용해 전라북도에서 사육된 한우 1,523두의 DNA를 공시하여 SNP 유전자형 분석을 실시하였다. 총 52,195개의 SNP의 유전자형이 분석되었고, QC 과정을 통해 11,056개의 SNP가 제거되었으며, 총 41,139개의 SNP가 QC 조건에 부합하여 다음 분석에 이용되었다. 염색체별 SNP표지인자 정보에 대한 요약은 Table 3에 제시하였다. 29개 상염색체의 평균 MAF는 0.27으로 확인되었다. 분석된 결과에서 염색체상 SNP들 간의 거리가 가장 긴 것은 12번 염색체로 확인되었다. SNP간의 평균거리는 0.072Mb으로 확인되었으며 표준편차 역시 가장 큰 것으로 확인되었다(0.12). 가장 짧은 SNP 간의 거리를 가지고 있는 것은 14번 염색체로 확인되었다(0.046 Mb). 14번 염색체의 SNP간의 평균거리의 표준편차는 0.047으로 확인되었다. Lee 등(2021)의 경기 지역에서 사육되어진 한우를 대상으로 연관불편형분석을 수행하여 보고하였으며, 염색체 상에서의 SNP 간의 거리 평균이 0.046-0.073 Mb인 것으로 확인하였다. Hong 등(2017)이 보고한 강원도 집단의 한우에서는 0.06-0.082Mb가 확인되었고, 후보종모우와 후대검정우를 대상으로한 Cho 등(2012)의 보고에서는 0.055-0.074Mb으로 확인되었다. 본 연구결과와 가장 유사한 결과는 경기도 지역의 한우 축군이었으며, 강원도 축군과 가장 큰 차이가 확인되었으며, 후보종모우 및 후대검정우 집단과도 차이가 확인되었다. 이러한 차이는 분석된 축군의 크기가 서로 다른 부분에서 기인한 것도 있지만, 가장 큰 이유는 분석에 이용된 유전자칩의 종류에서 기인한 것임을 확인할 수 있다. 축산과학원에서 개발한 한우 유전자칩은 2018년에 개발되어 보급되었기 때문에 그 이전의 연구에서는 Bovine 50K SNP chip을 이용하여 분석하였다.

분석에 이용된 SNP의 MAF (Minor Allele Frequency)에 따라 분류하여 분포를 확인하였으며, MAF가 0.05 이하인 SNP를 포함한 총 41,139개의 SNP의 분포를 Fig 1에 제시하였다. MAF가 0.05이하의 SNP를 제외하고 나머지 빈도에 분포하는 SNP의 숫자들은 비교적 유사한 것으로 확인되었다. 이는 대체로 SNP 표지인자들의 간격과 대립유전자의 빈도가 일정하게 분포하여 나타난 결과로 확인되어 진다. 최근 유전체 정보를 이용한 유전평가를 위해 특정 형질과 연관된 SNP의 가중치를 조성하여 분석에 활용하는 WGBLUP (weighted genomic best linear unbiased prediction)방법을 통해 평가 정확도를 향상시킨 연구들이 보고되고 있다(Lee *et al.*, 2019; Teissier *et al.*, 2019). 각 SNP 들의 기능적 측면에 대한 고려가 유전평가에 반영이 되면서, 분석의 효율성 강화를 위해 연관불편형 정보의 중요도는 더 높아질 것으로 사료된다.

Table 3. ber of markers, map distances and adjacent linkage disequilibrium on 29 autosomes

Chromosome	No of SNP	Mean r^2	SD	Total length (Mb)	Average Interval Size	Interval Standard Deviation
1	2,591	0.268	46.048	158.09	0.061	0.055
2	2,142	0.267	41.443	136.66	0.064	0.069
3	2,047	0.267	34.904	121.14	0.059	0.062
4	1,912	0.271	34.918	120.01	0.063	0.055
5	1,678	0.265	36.839	121.08	0.072	0.071
6	2,551	0.269	31.076	118.98	0.047	0.056
7	2,000	0.275	30.376	112.38	0.056	0.064
8	1,750	0.267	32.317	113.01	0.065	0.055
9	1,613	0.263	31.000	105.46	0.065	0.064
10	1,854	0.265	32.705	104.17	0.056	0.093
11	1,696	0.268	32.426	107.18	0.063	0.060
12	1,254	0.267	26.255	90.83	0.072	0.120
13	1,308	0.273	24.455	83.84	0.064	0.056
14	1,796	0.258	23.167	83.15	0.046	0.047
15	1,317	0.269	23.576	84.73	0.064	0.065
16	1,226	0.270	23.495	81.41	0.066	0.071
17	1,205	0.261	22.673	74.85	0.062	0.067
18	1,025	0.268	19.309	65.16	0.064	0.063
19	1,112	0.273	18.021	63.54	0.057	0.053
20	1,223	0.263	20.993	71.60	0.059	0.052
21	1,124	0.278	19.926	71.10	0.063	0.069
22	962	0.266	18.297	61.22	0.064	0.053
23	919	0.277	14.859	52.10	0.057	0.056
24	993	0.268	18.443	62.10	0.063	0.054
25	781	0.281	12.644	42.65	0.055	0.044
26	815	0.273	14.438	50.95	0.063	0.049
27	740	0.276	12.688	45.33	0.061	0.060
28	707	0.267	13.650	46.18	0.065	0.059
29	798	0.273	14.611	51.10	0.064	0.067
Mean/Total	1,419	0.269	25.019	2,500	0.061	0.062

*SD: Standard deviation

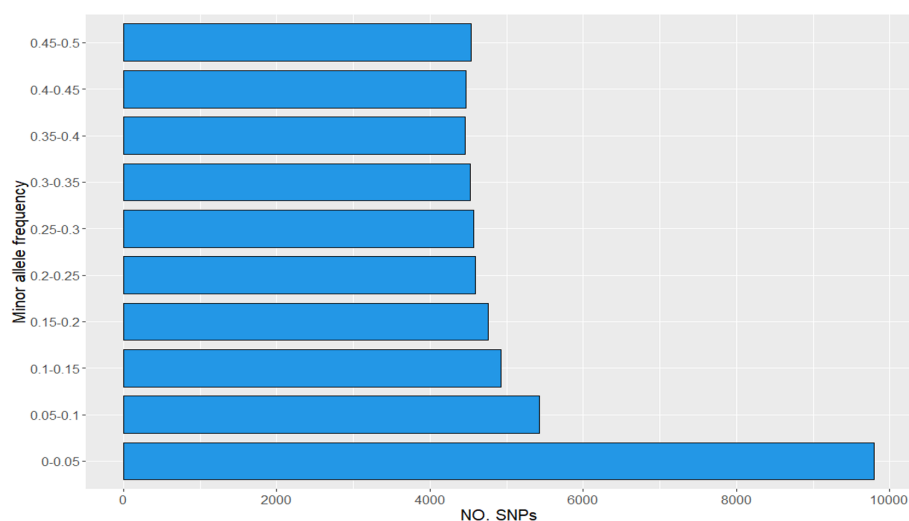
**Figure 1.** Minor Allele Frequency of SNPs

Table 4에서는 SNP 표지인자들을 50Kb 간격으로 나누어 r^2 값에 대한 평균 및 표준편차 그리고 거리별 SNP의 개수를 확인하여 제시하였다. 총 750,800개의 SNP 쌍이 고르게 분포되어 있으며, SNP 사이의 거리가 가까울수록 r^2 값이 높은 것이 확인되었다. 거리별 SNP의 개수는 거리가 850-900 Kb일 때 35,656쌍으로 가장 적었고, 50~100 Kb일 때 43,699으로 가장 많았다. r^2 값은 SNP 거리가 0~50 Kb인 경우 0.234, SNP 거리가 150~200 Kb인 경우 0.068로 연관불편형이 급격히 감소하는 현상을 확인하였으며, Fig 2에 그래프로 제시하였다. 서로 다른 한우 집단에서 연관불편형 분석을 보고한 Cho 등(2012), Hong 등(2017) 그리고 Lee 등(2021)의 보고에서도 200 Kb 미만에서 연관불편형이 급격히 감소하는 현상이 나타나고, r^2 의 값이 200Kb 이후에는 감소하는 것으로 확인되었다.

Table 4. Pairwise Linkage disequilibrium by distance

Fragment size (Kb)	No.	Mean (r^2)	Median	SD	$R^2 > 0.2(\%)$	$R^2 > 0.3(\%)$
0-50	36,887	0.234	0.10	0.29	35.1	27.0
50-100	43,699	0.125	0.05	0.19	18.7	12.3
100-150	42,048	0.087	0.03	0.15	11.4	6.9
150-200	40,663	0.068	0.02	0.13	8.1	4.6
200-250	39,646	0.057	0.02	0.11	5.8	3.2
250-300	38,723	0.051	0.02	0.10	4.6	2.5
300-350	37,877	0.047	0.02	0.10	4.1	2.1
350-400	37,900	0.043	0.02	0.09	3.4	1.9
400-450	37,258	0.041	0.02	0.09	3.1	1.7
450-500	36,928	0.040	0.02	0.09	3.0	1.6
500-550	36,326	0.040	0.01	0.09	2.9	1.7
550-600	36,350	0.038	0.01	0.09	2.7	1.6
600-650	36,084	0.037	0.01	0.08	2.6	1.5
650-700	36,149	0.037	0.01	0.08	2.5	1.5
700-750	35,990	0.037	0.01	0.08	2.4	1.5
750-800	35,844	0.036	0.01	0.08	2.4	1.5
800-850	35,713	0.036	0.01	0.08	2.4	1.4
850-900	35,656	0.036	0.01	0.08	2.4	1.4
900-950	35,689	0.035	0.01	0.08	2.4	1.4
950-1,000	35,370	0.035	0.01	0.08	2.3	1.3

*SD: Standard deviation (r^2)

최근 유전체 정보를 이용한 유전능력 평가 정확도의 상승은 한우산업 전반의 개량량을 증대시키고 개량의 효율성을 강화하는데 큰 역할을 할 것으로 기대되고 있다. 지금까지 한우의 개량사업은 보증씨수소의 선발과 인공수정에 집중된 국가단위의 개량이 중심이 되어왔다. 이는 사육중인 암소의 유전능력을 정확히 평가할 수 있는 방법이 미흡했기 때문이며, 유전체 정보를 이용한 유전능력 평가 기술은 사육중인 암소들의 유전능력을 보다 정확히 평가하고 지역단위, 축군단위의 개량사업 추진을 지원하는 효과적인 기술이 될 것으로 기대된다. 효율적이고 다양한 개량사업의 추진을 위해서 다양한 축군과 환경에서 분석된 유전체 정보가 서로 공유되고 참고된다면, 보다 효율적인 개량 사업의 전략 수립에 큰 도움이 될 것으로 기대된다. 따라서 본 연구는 전라북도 지역의 한우 축군에 대한 유전적인 기초 자료를 제공하고 공유함으로써 한우산업의 효율적인 개량 사업 추진에 도움이 되고자 수행하였다.

ACKNOWLEDGEMENTS

본 연구는 농촌진흥청(PJ01621803)의 지원으로 수행되었다.

REFERENCES

- Byun, S.-K., Kim, D.-H., Oh, J.-D., and Lee, H.-K., (2021). The Analysis on Evaluation of Genomic breeding value in Brindle Cattle using reference population of Hanwoo. *Journal of Animal Breeding and Genomics*, 5(3), 91-98.
- Choi, T. (2009). Establishment of phylogenomic characteristics for Korean traditional cattle breeds (Hanwoo, Korean brindle and black). *Doctors degree thesis. Chonbuk National University*, 1-180.
- Hong, M.-W., Choi, S.-Y., Kim, H., Yang, S.-Y., Kwak, K., Kim, J.-B., and Lee, S.-J., (2017). Linkage Disequilibrium and Association Analysis of Hanwoo Steer in Gangwon Region using Bovine SNP50K BeadChip. *Annals of Animal Resource Sciences*, 28(2), 46-55.
- Kim, H., Cho, Y. M., Ko, Y.-G., Kim, S. W., & Seong, H.-H. (2014). Analysis of Hematologic Characteristics of Korean Native Stripped Cattle Chikso according to the Ages. *Journal of Embryo Transfer*, 29(3), 313-319.
- Kim, N., Lee, J., Dang, C., Jung, Y., & Cheon, H. (2014). Prospect and retrospect on improvement of Hanwoo (Korean cattle) for 50 Years. *Bulletin of the Animal Biotechnology*, 6, 1-13.
- Kim, J.-W., (2018), Study on Regional Genetic Evaluation System in Hanwoo Cow, Master's thesis, *Jeonbuk National University*, Korea
- Kim, D.-H., Oh, J.-D., and Lee, H.-K., (2022). A study on analysis efficiency through the breeding value evaluation system according to the pedigree depth of Hanwoo in Jeonbuk region. *Journal of Animal Breeding and Genomics*, 6(1), 11-18.
- Lee, G. H., & Kong, H. S. (2021). Linkage Disequilibrium Analysis of Hanwoo in Gyeonggi Region using Hanwoo SNP Chip. *Journal of Animal Breeding and Genomics Vol*, 5(4).
- Lee, G. H., Lee, Y. S., Moon, S. J., & Kong, H. S. (2022). The Accuracy of Genomic Estimated Breeding Value Using a Hanwoo SNP Chip and the Pedigree Data of Hanwoo Cows in Gyeonggi Province. *Journal of Life Science*, 32(4), 279-284.
- Lopez, B. I., Lee, S.-H., Park, J.-E., Shin, D.-H., Oh, J.-D., de Las Heras-Saldana, S., . . . Lim, D. (2019). Weighted genomic best linear unbiased prediction for carcass traits in Hanwoo cattle. *Genes*, 10(12), 1019.
- Lopez, B. I., Lee, S.-H., Shin, D.-H., Oh, J.-D., Chai, H.-H., Park, W., . . . Lim, D. (2020). Accuracy of genomic evaluation using imputed high-density genotypes for carcass traits in commercial Hanwoo population. *Livestock Science*, 241, 104256.
- Meuwissen, T., & Goddard, M. E. (2000). Fine mapping of quantitative trait loci using linkage disequilibria with closely linked marker loci. *genetics*, 155(1), 421-430.
- Park, M. N., Alam, M., Kim, S., Park, B., Lee, S. H., & Lee, S. S. (2020). Genomic selection through single-step genomic best linear unbiased prediction improves the accuracy of evaluation in Hanwoo cattle. *Asian-Australasian Journal of Animal Sciences*, 33(10), 1544.
- Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, Maller J, Sklar P, de Bakker PI, Daly M J. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet*, 81: 559-575.
- Teissier, M., Larroque, H., & Robert-Granié, C. (2019). Accuracy of genomic evaluation with weighted single-step genomic best linear unbiased prediction for milk production traits, udder type traits, and somatic cell scores in French dairy goats. *Journal of dairy science*, 102(4), 3142-3154.