

Estimation of linkage disequilibrium and effective population size using whole genome single nucleotide polymorphisms in Korean native pig and Landrace

Jong-Eun Park^{1,2}, Joon-ho Lee², Ji-Hyun Son² and Deukhwan Lee²

¹Dept. of Animal science & Biotechnology, Seoul National University, Seoul, Korea,

² Graduate school of future convergence technology, Hankyong National University, Anseong, Korea

ABSTRACT: The linkage disequilibrium and effective population size in a closed population of Landrace purebred and Korean native pig crossbred were compared using a whole genomic single nucleotide polymorphism panel. Genotypic data consisted of 661 of Landrace and 1,266 of native pigs born in 2006 to 2012. After controlling SNP quality, linkage disequilibrium measures (r^2) were estimated for over 40,000 SNP. Average distance of SNP pairs was 2.3 Mb for both breeds. From the analysis using SNP, average r^2 between adjacent SNP pairs were 0.34 and 0.48 for native pig and Landrace, respectively. Effective population size estimates based on r^2 with the SNP distance were 27 and 33 for Landrace and native pig, respectively. This result suggests that two breeds in Korea have smaller effective population size than other countries. Crossbreds showed lager genetic diversity than closed populations in Korea.

Keywords: Linkage Disequilibrium ; Effective population size; Single Nucleotide Polymorphism; Pig

Introduction

Genetic diversity of a breed is important for the sustainable use of its genetic resources and its genetic improvement in the future (Falconer and MacKay (1996)), which is why many countries give an effort to control inbreeding through breeding programs (Hill (2000)). Recently, the inbreeding rate is expressed to effective population size (N_e), which is seen as a genetic parameter of the loss of genetic diversity. A useful approach is to estimate past effective population size on the basis of genomic data. Such methods are based on linkage disequilibrium (LD) patterns (Sved (1971); Hayes et al. (2003)). Genome-wide single nucleotide polymorphism (SNP) genotyping became possible in pigs in 2009 (Ramos et al. (2009)). This SNP technology is useful to study the structure of pig populations and allows for a more detailed comparison between populations.

In this study, we compared LD in the two domestic Korean pig breeds and inspected trends in their N_e . The LD-based methods were applied in estimating N_e . Genomic data were also used for pedigree structure analysis. The genomic analysis in this study is based on the genome-wide SNP.

Materials and Methods

Data. Pedigree data were obtained from a swine nucleus herd in Korea, and included 82,830 Korean Landrace and 1,232 Korean native pigs (KNPs). The oldest records were from the early 1980 (1983), but the majority of animals born in but the majority of animals in the data were born after 1998.

DNA samples and genotyping. Among these animals, genotypic data consisted of 661 sows and boars of Landrace and 1,266 (621/645) sows and boars of KNP born in 2006 to 2012. Blood samples were collected from the jugular vein for genomic DNA isolation. The genotyping was performed using PorcineSNP60 BeadChip (Illumina, San Diego, CA) containing 62,163 SNP.

Population structure. The population structure of the two breeds was examined with the multidimensional scaling (MDS) algorithm implemented in PLINK (Purcell et al., 2007). Only SNP that had a missing-ness per individual ≤ 0.1 , a missing-ness per marker ≤ 0.1 , a minor allele frequency ≥ 0.05 and Hardy-Weinberg equilibrium test statistics (χ^2) < 24 (corresponding P-value $\sim 10^{-6}$), and that had been mapped on pig genome build 10 (Sscrofa10, <http://www.ensembl.org>) were used in the calculations. A total of 47,447 SNP in the KNP and 44,525 SNP in the Korean Landrace data met all 4 requirements.

LD and N_e . To calculate the standardized LD value (D') and the squared correlation coefficient between the 2 loci (r^2). Pair-wise LDs were calculated using --r option in PLINK and snpMatrix in R packages (Clayton and Leung (2007)). LD based N_e was calculated as below,

$$r^2 = 1/(4cN_e + 1) \text{ (Sved (1971))}.$$

Results and Discussion

Population structure. The two studied breeds were clearly separated among populations (Figure 1). The KNP showed larger genetic variation than Landrace group in genotype data which implicated crossbred KNP has larger genetic diversity. In actual breeding scheme, the Landrace had been kept separate in the breeding programs for 14 generations without infiltrating other genetic resource from outbreeding. However, KNPs were mated using outbreeding system. Our results reflected this breeding situation well.

Table 1: Average linkage disequilibrium (LD) in different autosomal chromosomes for SNP 30 kb and 3 Mb apart

Item	30kb between SNP						3 Mb between SNP					
	Korean native			Korean Landrace			Korean native			Korean Landrace		
	n	r ²	D'	n	r ²	D'	n	r ²	D'	n	r ²	D'
Chr												
1	2,418	0.4	0.87	2,272	0.52	0.92	278,330	0.21	0.71	244,760	0.24	0.73
2	1,605	0.34	0.84	1,508	0.46	0.91	159,578	0.16	0.65	147,233	0.19	0.66
3	1,339	0.34	0.82	1,180	0.49	0.92	131,488	0.15	0.61	108,708	0.18	0.65
4	1,492	0.35	0.85	1,439	0.46	0.9	180,215	0.16	0.65	165,121	0.18	0.65
5	924	0.35	0.82	829	0.44	0.88	102,692	0.15	0.61	80,053	0.16	0.61
6	1,681	0.34	0.83	1,712	0.48	0.9	146,994	0.15	0.63	142,973	0.19	0.63
7	1,282	0.36	0.86	1,251	0.47	0.89	156,915	0.17	0.66	149,418	0.19	0.63
8	1,300	0.34	0.83	1,170	0.44	0.88	129,676	0.16	0.63	113,984	0.16	0.62
9	1,529	0.31	0.81	1,526	0.46	0.88	146,187	0.14	0.61	143,196	0.19	0.62
10	874	0.34	0.83	789	0.44	0.88	88,915	0.16	0.66	76,369	0.14	0.56
11	753	0.33	0.83	711	0.44	0.88	85,737	0.14	0.6	80,480	0.16	0.61
12	819	0.37	0.84	817	0.46	0.88	70,932	0.16	0.63	73,384	0.15	0.56
13	1,511	0.4	0.88	1,541	0.48	0.9	168,443	0.23	0.72	163,250	0.21	0.7
14	1,550	0.43	0.88	1,449	0.53	0.9	185,480	0.2	0.69	166,088	0.3	0.72
15	1,277	0.3	0.81	1,078	0.51	0.91	132,828	0.13	0.59	105,536	0.19	0.65
16	834	0.35	0.85	708	0.48	0.9	89,509	0.16	0.64	73,153	0.18	0.6
17	789	0.14	0.42	667	0.43	0.9	89,284	0.06	0.35	72,756	0.17	0.64
18	637	0.01	0.2	575	0.51	0.91	64,580	0.01	0.2	56,606	0.17	0.63
Total	22,614	0.34	0.81	21,222	0.48	0.9	2,407,783	0.16	0.63	2,163,068	0.2	0.65

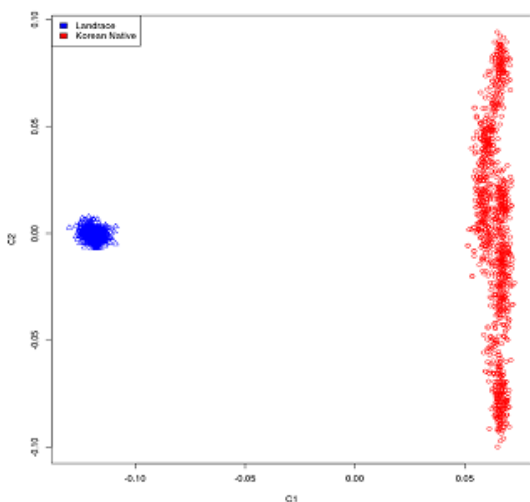


Figure 1: Population structure of the genotyped data. The two most informative axes of the multidimensional scaling are presented for Korean Landrace (blue) and Korean native pig (red).

Average LD. After controlling SNP quality, linkage disequilibrium measures (r^2) were estimated for over 40,000 (44,525 and 47,447 for Landrace and KNP, respectively) of SNP. Table 1 showed the average LD for each chromosome for distances of 30 kb (mean distance between markers in the SNP panel) and 3 Mb between SNP. Average distance of SNP pairs was 2.3 Mb for both breeds. From the LD analysis of 30kb between SNP, average LD (r^2) between adjacent SNP pairs was 0.34 and 0.48 (50 and 40% of the SNP pairs had $r^2 > 0.2$) for KNP and Landrace,

respectively. Similar trend was observed in 3 Mb between SNP. We also found non-linear relationship between LD and number of SNP in both populations. The average LD across all autosomal chromosomes showed that overall LD was stronger in Korean Landrace compared with KNP, their average $r^2 > 0.2$ extending to 1.0 and 0.7 Mb, respectively (Figure 2). Observed larger genetic variation of KNP in MDS analysis agreed with smaller LD results, which gave a larger N_e .

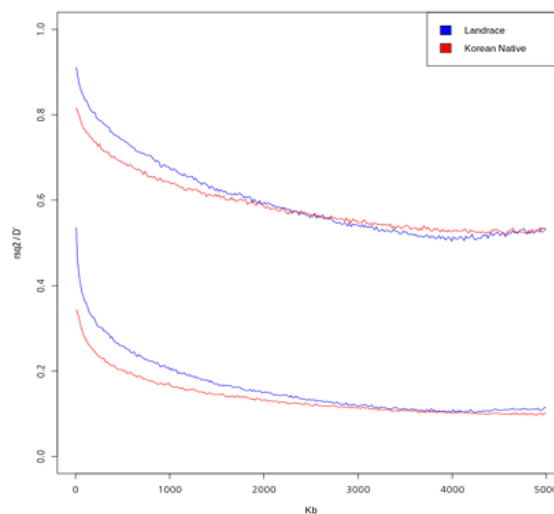


Figure 2: Average linkage disequilibrium (LD) over all chromosomes for Korean Landrace (blue) and for Korean native pig (red). The 2 upper lines correspond to D' (standardized LD value) and the 2 lower lines to r^2 (squared correlation coefficient between SNP).

Effective population size. Effective population size estimates based on r^2 with the SNP distance were different to those based on the pedigree data as 27 and 33 for Landrace and KNP, respectively (Figure 3). The difference in N_e showed weak level of significance (Student's t test, $p = 0.06 < 0.1$). Thus, the results indicate that the effective population size of current both breed has a clear effect on the extent of LD. Even effective population size of KNP is a bit larger than Landrace, both breeds is smaller than recommended minimum of 50 and may get smaller than that in the near future. Therefore, breeding action should be taken to balance the inbreeding rate and selection response. Because a moderate level of LD extends over half of SNP pairs, we suggested that some outbreeding system should be implemented from importing new genetic resources and selection based on whole genome SNP markers (genomic selection) would be also efficient for both breeds.

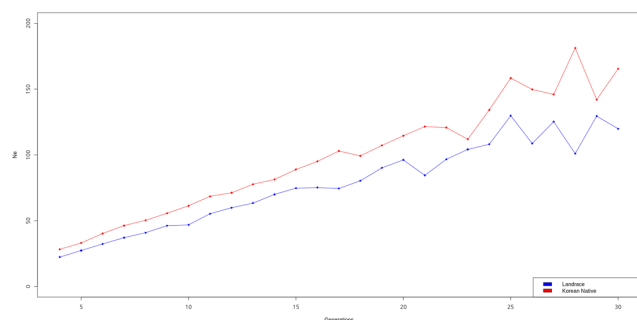


Figure 3: Estimates of effective population size (N_e) for Korean Landrace (blue) and for Korean native (red) t ancestral generations ago based on the squared correlation coefficient (r^2) between SNPs

Conclusion

The two studied breeds were clearly separated among populations. KNP crossbred showed larger genetic variation than Landrace group in genotype data. Average LD of (r^2) between adjacent SNP pairs (<30kb) were 0.34 and 0.48 for KNP and Landrace, respectively. These results reflect the different breeding situation between two breeds well. LD based effective population size was much smaller than pedigree based effective population size and may get smaller than that in the near future. Therefore, breeding action should be taken to balance the inbreeding rate and selection response.

Acknowledgments

This work was supported by a grant from the IPET Program (No. 20093068(309022-5) and 20100139), Ministry of agriculture, food and rural affairs, Republic of Korea.

Literature Cited

- Clayton, D., and Leung, H. T. (2007). *Human heredity*. 64(1), 45-51.
- Falconer, D. S., and T. F. C. MacKay. (1996). Longmans Green, Harlow, Essex, UK.
- Hayes B. J., Visscher P. M., McPartlan H. C., et al. (2003). *Genome Res.* 13:635–643.
- Hill W. G. (2000). *Livest. Prod. Sci.* 63:99–109.
- Purcell, S., Neale, B., Todd-Brown, K., et al. (2007). *The American Journal of Human Genetics*, 81(3), 559-575.
- Ramos A. M., Crooijmans R. P., Affara N. A., et al. (2009). *PLoS ONE* 4:e6524.
- Sved J. A. (1971). *Theor. Popul. Biol.* 2:125–141.
- Uimari, P., and Tapio, M. (2011). *Journal of animal science*, 89(3), 609-614.