Estimation of Effective Population Size in the Sapsaree:
A Korean Native Dog (*Canis familiaris*)

M. Alam, K. I. Han[1], D. H. Lee[2], J. H. Ha[3] and J. J. Kim*

School of Biotechnology, Yeungnam University, Gyeongsan, Gyeongbuk, Korea

ABSTRACT: Effective population size (N_e) is an important measure to understand population structure and genetic variability in animal species. The objective of this study was to estimate N_e in Sapsaree dogs using the information of rate of inbreeding and genomic data that were obtained from pedigree and the Illumina CanineSNP20 (20K) and CanineHD (170K) beadchips, respectively. Three SNP panels, *i.e.* Sap134 (20K), Sap60 (170K), and Sap183 (the combined panel from the 20K and 170K), were used to genotype 134, 60, and 183 animal samples, respectively. The N_e estimates based on inbreeding rate ranged from 16 to 51 about five to 13 generations ago. With the use of SNP genotypes, two methods were applied for N_e estimation, *i.e.* pair-wise r^2 values using a simple expectation of distance and r^2 values under a non-linear regression with respective distances assuming a finite population size. The average pair-wise N_e estimates across generations using the pairs of SNPs that were located within 5 Mb in the Sap134, Sap60, and Sap183 panels, were 1,486, 1,025 and 1,293, respectively. Under the non-linear regression method, the average N_e estimates were 1,601, 528, and 1,129 for the respective panels. Also, the point estimates of past N_e at 5, 20, and 50 generations ago ranged between 64 to 75, 245 to 286, and 573 to 646, respectively, indicating a significant N_e reduction in the last several generations. These results suggest a strong necessity for minimizing inbreeding through the application of genomic selection or other breeding strategies to increase N_e, so as to maintain genetic variation and to avoid future bottlenecks in the Sapsaree population. (**Key Words:** Effective Population Size, Inbreeding Rate, Linkage Disequilibrium, SNP, Sapsaree Dog)

INTRODUCTION

Many dog breeds are originated and established by extensive selection from a small gene pool. During the selection process, some breeds are developed through population bottleneck causing loss in genetic diversity. The Sapsaree, an aboriginal Korean dog breed, also faced a severe bottleneck a few decades ago during Japanese colonization, while conservation and systematic breeding programs have been implemented since 1980.

Effective population size (N_e) is an useful criterion in classifying a breed population in view of degree of endangerment (FAO, 1998; Schwartz et al., 1998; Duchev et al., 2006), and a key parameter in conservation and population genetics (Gutierrez et al., 2008), because it is related to inbreeding, fitness, and loss of genetic variation through random genetic drift (Crow and Kimura, 1970; Falconer and Mackay, 1996).

Recently the use of inbreeding rate for N_e estimation has been replaced by molecular markers (Schwartz et al., 1998; Beaumont, 2003; Leberg, 2005; Wang, 2005), especially by high-throughput SNP panels. These genome-wide SNPs are expected to remarkably improve the reliability in estimating N_e (Nomura, 2009). Linkage disequilibrium (LD), *i.e.* non-random association of alleles between two SNP markers, is caused by the finite population size, mutation, migration, and selection (Lander and Schork, 1994) and is widely used in N_e estimation (Sved, 1971; Hill, 1981; Hayes et al., 2003). Also, LD between densely spaced SNP makers contains information about the historical population size (Hayes et al., 2003).

Herein, we investigated the population structure, extent of LD and effective population size (N_e) of the Sapsaree breed in the first such study of Korean aboriginal dog breeds.

* Corresponding Author: Jong-Joo Kim. Tel: +82-53-810-3027, Fax: +82-53-801-3027, E-mail: kimjj@ynu.ac.kr
[1] Sapsaree Breeding Research Institute, Hayang, Gyeongsan, Gyeongbuk, Korea.
[2] The Korean Sapsaree Foundation, Kyungpook National University, Taegu, Korea.
[3] Department of Biotechnology, Kyungpook National University, Taegu, Korea.

MATERIALS AND METHODS

Animals, pedigree and genotypes

A total of 1,082 Sapsaree dogs with a pedigree size of 8,264 were generated (13 generations) between 1989 and 2009 at the Sapsaree Breeding Research Institute, Hayang, Gyeonsan, Korea. Among the individuals, 183 dogs were selected that were least genetically related to each other. The chosen Sapsarees were genotyped in two phases; a set of 134 Sapsarees with the Illumina CanineSNP20 (>22,000 SNPs) and a second set of 60 individuals, including the eleven dogs that were also genotyped using the CanineSNP20s, with the Illumina CanineHD BeadChip (>170,000 SNPs). The SNPs were genotyped according to the Illumina's Infinium and Infinium HD Assay Protocols, respectively (Illumina Inc., USA). Both the chips contained evenly spaced and validated SNPs that were derived from CanFam2.0 assembly.

Inbreeding effective population size

Three reference animal sets with all pedigree, phenotypes or genotypes were used for analysis. The increase in inbreeding (ΔF) for each animal was calculated using the formulae $\Delta F_i = 1 - {}^{t_i-1}\!\sqrt{1-F_i}$ (Gutierrez et al., 2009), where individual inbreeding coefficients (F_i) were derived from pedigree, and t_i referred to the number of complete generations for the F_i estimation (Maignel et al., 1996). An average increase in inbreeding ($\overline{\Delta F}$) across the individuals was obtained, and an inbreeding N_e and its standard error ($\sigma_{\overline{N_e}}$) were estimated by the equations,

$$N_e = \frac{1}{2\overline{\Delta F}} \quad \text{and} \quad \sigma_{\overline{N_e}} = \frac{2}{\sqrt{N}} \overline{N_e}^2 \sigma_{\Delta F},$$ in which N represents

the number of animals in each sample, and $\sigma_{\Delta F}$ is the standard deviation of ΔF estimates (Gutierrez et al., 2008). Accuracy in inbreeding coefficient was measured with the pedigree completeness index (PEC value; MacCluer et al., 1983) for each animal in the process. Primarily a PEC value of 0.7 or greater value was accounted in N_e estimation. To validate the robustness of N_e estimates, the three types of samples and a lower limit for PEC (≥ 0.4, ≥ 0.5, ≥ 0.6) were also used. Several levels of ancestral generations were used in computing the PEC index, *e.g.* 5, 8, 10, and 13. Data formulation, inbreeding rate and PEC indexes were derived using the software packages, CFC1.0 (Sargolzaei et al., 2006) and EVA interface 1.3 (Berg, 2009).

F_{IT}, F_{ST} and F_{IS} were also obtained for each group of individuals with the same birth year. F_{IT} is an average inbreeding measure of the group, F_{ST} is an expected value of inbreeding under random mating, and F_{IS} measures the deviation from randomness in the actual breeding.

When $F_{IS} > 0$, the actual inbreeding (F_{IT}) exceeds the level expected under random mating (F_{ST}), implying that mating among genetically related parents happens more often, or that the population is partitioned into subpopulations and mating is restricted within each subpopulation. In contrast, in the population with $F_{IS} < 0$, avoidance of inbreeding or mating between subpopulations is carried out predominantly. The statistic $F_{IS} = (F_{IT} - F_{ST})/(1 - F_{ST})$ was described in Wright (1969).

Effective population size from SNP genotypes

Three sets of Sapsarees were genotyped with the SNP panels of 20K (Sap134), 170K (Sap60), and the combined SNP set of 20K and 170K (Sap183) using 134, 60, and 183 individuals, respectively. Among the SNPs in the SNP panels that were located on autosomal chromosomes, those with minor allele frequencies less than 0.05 and Hardy-Weinberg equilibrium test statistics, $x^2 > 10.83$ (corresponding p<0.001) were deleted (Table 1). From the pairwise haplotype frequencies between each pair of markers, *e.g.* A and B, each with two alleles, *e.g.* 1 and 2, (Falconer and MacKay, 1996), the D value (one measure of linkage disequilibrium, LD) was calculated from $D_{AB} = P_{A_1B_1} \times P_{A_2B_2} - P_{A_1B_2} \times P_{A_2B_1}$ (Lewontin, 1964; Hill, 1981), where p is the frequency of haplotypes for either homozygotes or heterozygotes in the population. A squared correlation coefficient (r^2) was calculated for all pair of SNPs on a given chromosome, as $r^2 = \dfrac{D_{AB}^2}{P(A_1)P(A_2)P(B_1)P(B_2)}$, where, P_{A1}, P_{A2}, P_{B1}, P_{B2} are frequencies of A_1 A_2, B_1, B_2 alleles in the population, respectively (Hill and Robertson, 1968; Hill, 1981). The expectation of r^2 was used to derive N_e from $E(r^2) = \dfrac{1}{(1 + 4N_e c)}$ (Sved, 1971), where N_e is the finite effective population size, and c is the length of the chromosome segment in Morgan units. The pairs of SNPs that were located within 5 megabase (Mb) were used to estimate the average pair-wise N_e. A unit physical distance

Table 1. Number of available SNPs and average interval distances between adjacent SNPs across all 38 autosomes in different canine SNP panels

	Sap134[b]	Sap60[b]	Sap183[b]
Total length (Mb)	2,181	2,198	2,178
Total number of SNP	17,441	89,788	14,055
Average interval (Kb)	125	24	125
SD[a] (Kb)	135	27	135

[a] Standard deviation of average intervals.

[b] Three sets of SNP panels were used with genotyped with the Illumina CanineSNP20K chips (Sap134), CanineHD 170K BeadChip (Sap60), and the combined SNP set of 20K and 170K (Sap183) using 134, 60, and 183 individuals, respectively.

(Mb) between SNPs was simply assumed as a unit genetic distance (centimorgan), *i.e.* 1 cM ≈1 Mb. Another measure of N_e was applied using a non-linear regression, in which a corrected r^2 for each pair of SNPs was regressed in respect to the physical distances. The non-linear regression model was $y_i = \dfrac{1}{(\alpha + \beta c_i)} + e_i$, where $y_i = \left(r_i^2 - \dfrac{1}{n} \right)$ and y_i is the adjusted r^2 for chromosome sample size (n) for SNP pair i, α is 1 in absence of mutation, β is equivalent of $4N_e$ and c_i is the recombination distance (in Morgan). Chromosome sample size (n) was adjusted according to Weir and Hill (1980), as it induced chance disequilibrium of variance(r) from experimental sampling. Parameter β was estimated iteratively using a least squares equation.

Past effective population size from SNP genotype

The past effective population sizes were observed from LD values assuming that the population had a linear growth, and using $t = \dfrac{1}{2c}$, where recombination distance, c, reflected the ancestral effective population size at t generation ago. The final point estimate of past N_e was an average of the pair-wise N_e at a certain distance in context to a particular t value. The only r^2 values ranging between 0.01 and 0.99 were used, because of the limited parameter space of the r^2 statistic where the point estimates of N_e become infinite at $r^2 = 0$, and zero when $r^2 = 1$ (Uimari and Tapio, 2011). Also, another normalized measure of LD, D' was applied, which enabled the dependence on allele frequency to be avoided (Lewontin, 1964).

$$D' = \frac{|D_{AB}|}{D_{AB}^{\max}}$$

Where,

$$D_{AB}^{\max} = \begin{cases} \min(P_{A_1} \times P_{B_1}, P_{A_2} \times P_{B_2}), & if\ D < 0 \\ \min(P_{A_1} \times P_{B_2}, P_{A_2} \times P_{B_1}), & if\ D > 0 \end{cases}$$

RESULTS

Pedigree completeness (PEC), rate of inbreeding and effective population size from pedigree

Three reference animal sets, *i.e.* with all pedigree, phenotypes, and genotypes, revealed average PEC index values of 0.80, 0.65 and 0.73, respectively, assuming that five ancestral generations contributed to pedigree completeness. An average of five equivalent complete generations was traced in the pedigree with a maximum of 8 generations by 2009. The average inbreeding coefficient was 0.10 with an increasing rate of inbreeding per generation (Figure 1), which stabilized in the last few generations. Actual inbreeding (F_{IT}) surpassed the expectation under random mating (F_{ST}) in several generations.

Overall, the average inbreeding N_e ranged from 16 to 19, 34 to 39, and 36 to 51 using PEC index at 5, 10 and 13 generations, respectively (Table 2). In general, the N_e estimates were consistent between different PEC limits, and greater N_e values were estimated as the PEC generation index increased, partly due to the inclusion of earlier generations in which the individuals including founders had a lower inbreeding rate (Figure 2).

Linkage disequilibrium, recent and past effective population sizes from SNP

The average r^2 estimates were 0.12 and 0.07 when the pairs of SNPs were located within 100 Kb and 5 Mb, respectively, with the Sap134 panels in which 22K SNPs were embedded. When the Sap60 panels with 170K SNPs were applied, the average r^2 estimates were 0.16 and 0.10, within the respective distances (Table 3). Because map
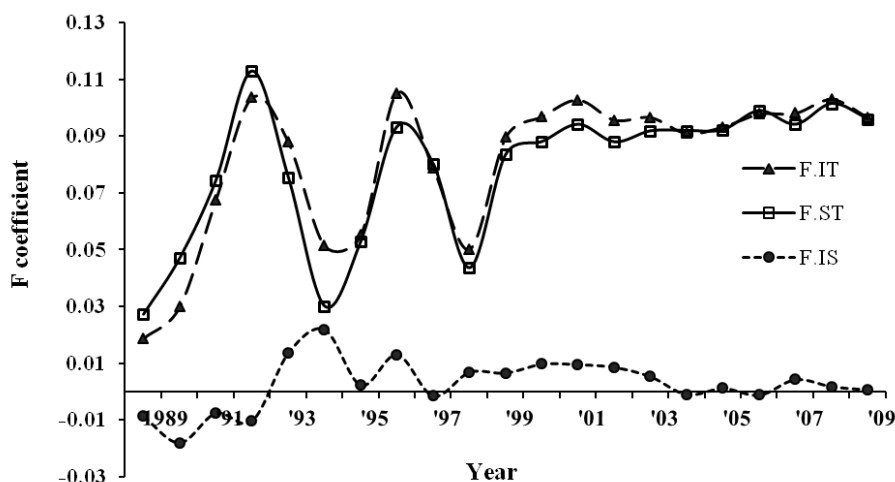
Figure 1. Average inbreeding coefficients (F_{IT}, F_{ST}, F_{IS}) of the Sapsaree population along birth year.

Table 2. Average effective population size (N_e) of Sapsaree based on reference animals with different pedigree completeness (PEC) generations and limits

| Type of data	PEC Gen. Ind.[a]	PEC limit							
		>0.7		>0.6		>0.5		>0.4	
		N_e	SE	N_e	SE	N_e	SE	N_e	SE
Whole pedigree	5	17.7	0.1	18.0	0.1	18.5	0.1	18.6	0.1
	8	29.2	0.2	30.0	0.2	30.6	0.2	31.4	0.2
	10	36.7	0.7	38.1	0.3	38.4	0.3	39.4	0.3
	13	-	-	35.6	4.5	49.0	0.5	51.1	0.4
Phenotyped	5	17.0	0.3	16.3	0.3	15.2	0.4	17.0	0.3
	8	26.5	0.7	26.7	0.6	27.3	0.5	28.3	0.5
	10	-	-	35.5	1.5	33.8	0.8	35.2	0.7
	13	-	-	-	-	44.8	2.3	45.1	1.1
Genotyped	5	17.3	0.6	18.1	0.7	15.6	0.6	18.7	0.7
	8	26.1	1.2	27.5	0.9	29.9	1.1	31.6	1.2
	10	-	-	36.1	2.3	34.7	1.2	38.7	1.4
	13	-	-	-	-	45.8	3.3	46.0	1.6

[a] PEC Gen. Ind. = PEC generation index.

density between flanking SNPs in the latter panels was higher (Table 1), the LDs values were greater than when the Sap134 panels were used. The greatest r^2 values were observed on canine chromosome (CFA) 37, *i.e.* 0.15 and 0.21 for the pairs of SNPs within 100 Kb in the Sap134 and Sap60 panels, respectively. The distribution of the D' estimates had similar tendency as for the r^2 values (Table 3). Figure 2 displayed LD profiles along the distances between the pairs of SNPs, showing that LD was greatest when the pairs of SNPs were very closely located, and decreased with the increasing distances between the pairs of SNPs.

The N_e estimates were 1,486, 1,025, and 1,293 from the pair-wise method, and 1,601, 528, and 1,129 from the non-linear method using Sap134, Sap60 and Sap183 panel, respectively, within a 5 Mb distance (Table 4). Strong positive correlations on the N_e estimates from the non-linear approach were observed between the data sets (Sap134, Sap60 and Sap183 panels), *i.e.* the correlation coefficients ranged between 0.41 and 0.65 (Figure 3).

The past N_e estimates ranged between 5,381 and 5,699 about five hundred generations ago, reducing to 1,126 to 1,265 about 100 generations ago. Since then, N_e continued to decrease to 64 to 75 about five generations ago, indicating a faster drop during the last ten generations (Figure 4, Table 5). The Sap60 panel showed a pattern of constant reduction in the past N_e estimates, probably due to
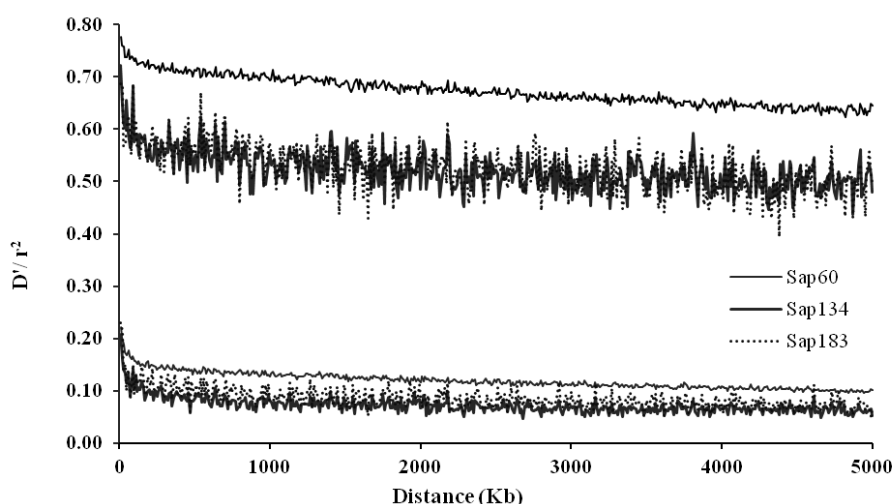
Figure 2. Average linkage disequilibrium (LD) over all autosomal chromosomes within 5 Mb distances between the pairs of SNPs for Sapsaree dogs. The upper and lower 3 lines correspond to D' and r^2 measures for the three samples (Sap134, Sap60 and Sap183), in which the SNP panels of 20K, 170K, and the combined SNP set of 20K and 170K were used using 134, 60, and 183 individuals, respectively.

Table 3. Linkage disequilibrium (LD) measures in autosomal chromosomes using pairs of SNPs that are distanced within 100 Kb and 5 Mb

| CFA[a] | 100 Kb between SNPs | | | | | | 5 Mb distance between SNPs | | | | | |
| | Sap134[b] | | | Sap60[b] | | | Sap134[b] | | | Sap60[b] | | |
	n	r^2	D'	n	r^2	D'	n	r^2	D'	n	r^2	D'
1	570	0.12	0.59	14,035	0.15	0.70	5,405	0.07	0.51	135,695	0.11	0.63
2	288	0.11	0.60	8,575	0.15	0.70	2,616	0.07	0.50	77,137	0.10	0.62
3	462	0.13	0.62	12,776	0.17	0.70	4,494	0.07	0.49	124,260	0.11	0.61
4	373	0.13	0.60	9,890	0.16	0.72	3,350	0.07	0.50	95,411	0.12	0.65
5	388	0.12	0.57	10,192	0.16	0.70	3,653	0.07	0.49	93,254	0.11	0.62
6	305	0.11	0.59	9,212	0.16	0.70	2,905	0.07	0.49	81,648	0.11	0.61
7	379	0.11	0.61	10,065	0.16	0.70	3,690	0.07	0.51	94,366	0.11	0.63
8	272	0.12	0.61	7,203	0.17	0.70	2,672	0.07	0.50	68,456	0.11	0.62
9	251	0.12	0.63	5,782	0.15	0.70	1,968	0.07	0.52	53,236	0.11	0.63
10	308	0.12	0.57	8,013	0.15	0.69	2,944	0.06	0.47	75,968	0.09	0.59
11	315	0.11	0.56	6,719	0.15	0.70	2,731	0.07	0.49	61,190	0.10	0.61
12	328	0.12	0.60	9,331	0.14	0.69	2,632	0.07	0.51	86,870	0.10	0.61
13	366	0.11	0.59	8,147	0.16	0.70	3,034	0.07	0.50	77,424	0.12	0.63
14	232	0.12	0.63	6,908	0.15	0.68	2,554	0.07	0.50	64,981	0.11	0.62
15	287	0.11	0.59	7,968	0.16	0.69	2,495	0.07	0.48	74,810	0.11	0.61
16	264	0.11	0.59	7,740	0.15	0.67	2,641	0.07	0.48	72,492	0.09	0.58
17	354	0.11	0.58	8,222	0.16	0.69	3,396	0.07	0.49	77,955	0.10	0.60
18	178	0.10	0.57	6,490	0.16	0.69	1,670	0.07	0.50	57,179	0.10	0.60
19	267	0.12	0.62	6,579	0.15	0.69	2,398	0.07	0.49	63,658	0.10	0.59
20	246	0.15	0.61	5,973	0.17	0.71	2,080	0.07	0.53	51,562	0.12	0.62
21	301	0.10	0.55	7,318	0.14	0.69	2,397	0.07	0.48	66,089	0.09	0.60
22	252	0.11	0.58	6,401	0.16	0.72	2,040	0.07	0.50	58,410	0.12	0.64
23	388	0.12	0.61	7,257	0.15	0.69	3,422	0.06	0.50	70,489	0.10	0.60
23	304	0.13	0.61	6,300	0.16	0.70	2,719	0.07	0.50	60,028	0.11	0.61
25	288	0.12	0.57	7,458	0.16	0.67	2,883	0.07	0.50	72,548	0.10	0.58
26	196	0.14	0.60	5,257	0.17	0.71	1,615	0.07	0.50	42,076	0.10	0.59
27	178	0.11	0.60	5,628	0.14	0.68	1,787	0.07	0.51	47,172	0.10	0.62
28	257	0.12	0.60	4,536	0.17	0.71	2,042	0.07	0.48	42,146	0.11	0.60
29	208	0.09	0.59	4,651	0.15	0.70	1,821	0.07	0.49	43,795	0.10	0.61
30	188	0.13	0.61	5,346	0.15	0.67	1,937	0.07	0.49	47,916	0.09	0.56
31	166	0.10	0.58	5,381	0.15	0.69	1,673	0.06	0.48	45,250	0.09	0.59
32	165	0.12	0.61	5,232	0.15	0.69	1,636	0.07	0.49	47,742	0.10	0.59
33	161	0.12	0.62	3,683	0.17	0.71	1,522	0.06	0.48	34,370	0.09	0.57
34	218	0.12	0.58	6,054	0.16	0.70	2,050	0.06	0.48	53,411	0.09	0.59
35	164	0.11	0.58	4,246	0.14	0.68	1,467	0.06	0.45	37,795	0.08	0.53
36	116	0.10	0.54	4,306	0.13	0.66	1,009	0.06	0.46	38,095	0.08	0.56
37	129	0.15	0.62	2,913	0.21	0.73	1,135	0.07	0.49	25,834	0.10	0.60
38	138	0.10	0.64	3,484	0.14	0.68	1,259	0.05	0.47	30,795	0.08	0.55
All[c]	10,250	0.12	0.59	265,271	0.16	0.69	93,742	0.07	0.49	2,451,513	0.10	0.60

[a] CFA = Canine chromosome.

[b] Two sets of SNP panels with the Illumina CanineSNP20K chips (Sap134), CanineHD 170K BeadChip (Sap60), using 134 and 60 individuals, respectively. n = Total number of SNP pairs. D' = Standardized LD; r^2 = Squared correlation coefficient between a paire of SNPs.

[c] Average values across chromosomes and standard error of the estimates were lower than 0.0001 (data not shown).

Table 4. Estimates of effective population size (N_e) using different sets of SNP panels that were estimated using r^2 with the pairs of SNPs that are distanced within 5 Mb

CFA[a]	Population					
	Sap134[b]		Sap60[b]		Sap183[b]	
	N_{e1}[c]	N_{e2}[c]	N_{e1}[c]	N_{e2}[c]	N_{e1}[c]	N_{e2}[c]
1	923	1,412	1,055	500	940	1,188
2	1,155	1,864	999	563	1,220	1,184
3	850	1,176	1,008	401	909	822
4	971	1,393	883	379	981	914
5	862	1,486	955	503	818	972
6	801	1,435	1,053	500	760	1,086
7	825	1,458	905	429	971	1,106
8	1,120	1,374	1,021	384	1,660	750
9	886	1,470	1,187	519	837	894
10	6,330	1,624	1,215	594	9,185	1,378
11	2,677	1,990	1,092	603	881	1,561
12	861	1,597	986	647	808	1,148
13	4,693	1,403	1,044	363	827	1,044
14	1,250	1,149	941	520	1,740	833
15	885	1,532	976	420	810	952
16	755	1,745	1,016	764	731	1,289
17	2,229	1,656	1,013	452	768	1,109
18	884	1,587	924	506	823	1,108
19	781	1,514	942	483	821	1,100
20	892	1,218	887	389	872	618
21	839	1,637	1,044	781	833	1,369
22	988	1,667	1,174	387	1,178	1,154
23	801	1,605	1,040	630	794	1,171
24	2,505	1,285	891	425	4,461	806
25	967	1,706	1,208	466	1,311	1,461
26	842	1,285	1,061	460	1,027	930
27	1,666	1,598	951	691	856	1,375
28	859	1,770	883	377	1,585	1,102
29	883	1,645	1,009	471	1,122	1,087
30	1,092	1,473	952	633	1,388	1,201
31	1,160	2,503	946	675	776	1,628
32	7,927	1,736	845	559	1,093	1,275
33	1,022	1,921	1,057	424	929	979
34	773	1,558	1,044	484	955	1,081
35	759	1,551	1,395	738	820	1,302
36	912	2,570	1,377	1,002	775	1,594
37	862	1,170	872	227	908	753
38	975	2,090	1,106	722	976	1,594
Mean	1,486	1,601	1,025	528	1,293	1,129
SD[d]	1,549	312	127	150	1,453	250

[a] CFA = Canine chromosome.

[b] Three sets of SNP panels were used with genotyped with the Illumina CanineSNP20K chips (Sap134), CanineHD 170K BeadChip (Sap60), and the combined SNP set of 20K and 170K (Sap183) using 134, 60, and 183 individuals, respectively.

[c] Effective population size estimated by a pair-wise method (N_{e1}) and a non-linear regression method (N_{e2}).

[d] SD = Standard deviation of the N_e estimates across all CFAs.
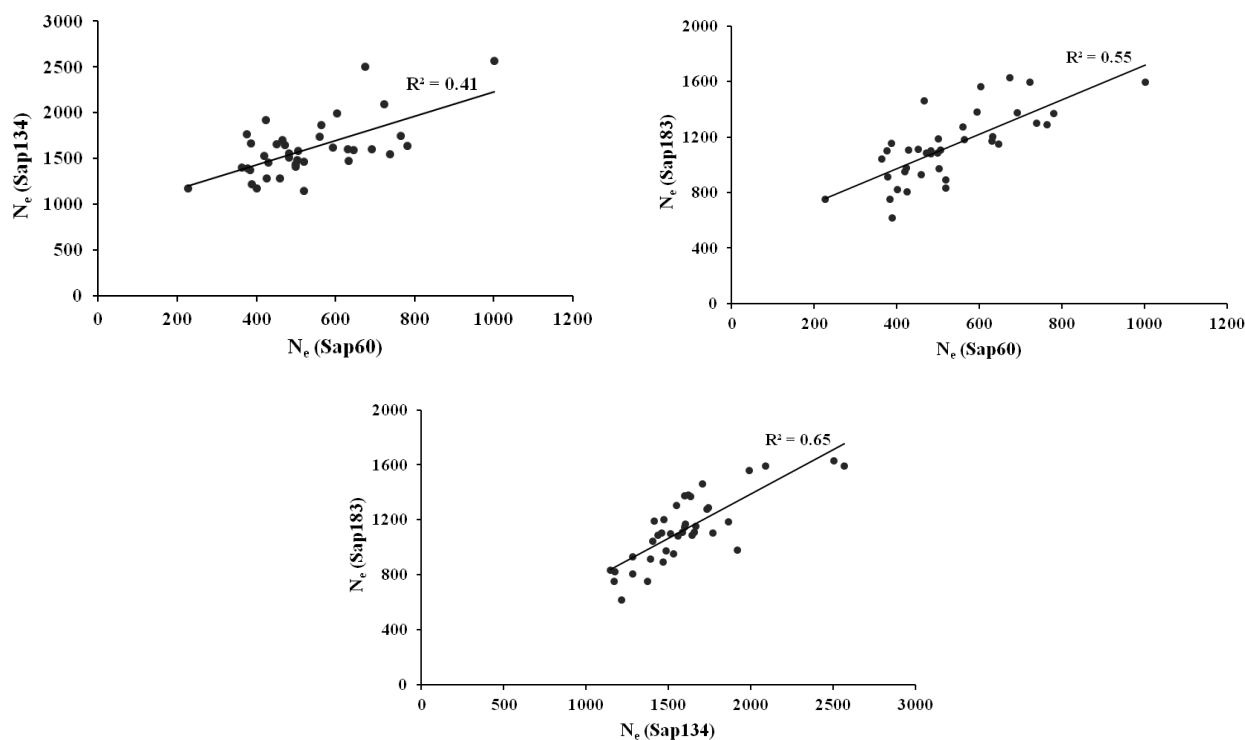
Figure 3. Correlations between the non-linear N_e estimates from the Sap134, Sap60 and Sap183 panels for the pairs of SNPs that were distanced within 5Mb.

the contribution of more SNPs at closer distances than the lower density panels (Sap134 and Sap183), which produced a weighted average at certain generations despite a lowest chromosome sample size.

DISCUSSION

In this study, we applied various density panels of SNP to determine overall, current and past effective population sizes in order to investigate rates of genetic erosion, fixation of deleterious alleles and inbreeding (Wright, 1969). The past N_e estimates, *e.g.* 64 to 75 about five generations ago, were in accordance with Calboli et al. (2008), in which the N_e was estimated between 17 and 76 in multiple breeds such as Greyhound, Rough Collie, Akita Inu, Boxer, English bulldog, Chow chow, Golden retriever, English springer spaniel and German shepherds. However, these breeds had lower inbreeding rates (0.024 to 0.073) than in the Sapsaree breed (Figure 1). Likewise, the wolves as dog's early ancestors, had lower effective population sizes, *e.g.* the Italian wolves with N_es ranging 30 to 50 since last century (Randi et al., 2000) or the Finnish wolf with the current N_e of 40 (Aspi et al., 2006).

It is known that the extent of LD in dogs was greater than in cattle and humans (Lindblad-Toh et al., 2005; Tenesa et al., 2007). Sutter et al. (2004) reported that LDs in dogs extended even up to 100 times higher than in humans, *e.g.* in Akita, Bernese mountain dog, Golden retriever,

Labrador retriever, and Pekingese. Lou et al. (2003) observed extensive LDs within certain chromosomal distances (5 to 10 cM), and we also found strong LDs in Sapsaree (Figure 2).

The noticeable expansion of LD in dogs may reflect a narrow bottleneck in their domestication history (Ostrander and Kruglyak, 2000), and the LD at long distances reflects a more recent population history and vice versa (Hayes et al., 2003). The tight population bottleneck history in Akita and Bernese mountain dogs was followed by a large reduction in N_e (Rogers and Brace, 1995; Wilcox and Walkowicz, 1995; AKC, 1998) and reflects a long extented LD (Sutter et al., 2004). This result is consistent with the Sapsaree population history, in which significant reduction of the population size occurred during Japanese colonization of Korea more than half a century ago, wherein the breed reached extinction and was re-established about 30 years ago (Han et al., 2010). Moreover, the past N_e with significant reduction of the size in recent generations may cause an increasing inbreeding rate in the Sapsaree individuals of the current generation (Figure 1).

In the non-linear analyses, sampling errors were corrected to get unbiased r^2 estimates with the consideration of a chromosome sample size (Weir and Hill, 1980). Perhaps, putting restrictions over r^2 measures to avoid complexity in computation might lead some biases in estimates and, especially for the different platforms (*i.e.* dog breeds) in the high-throughput SNP panels, in which wrong
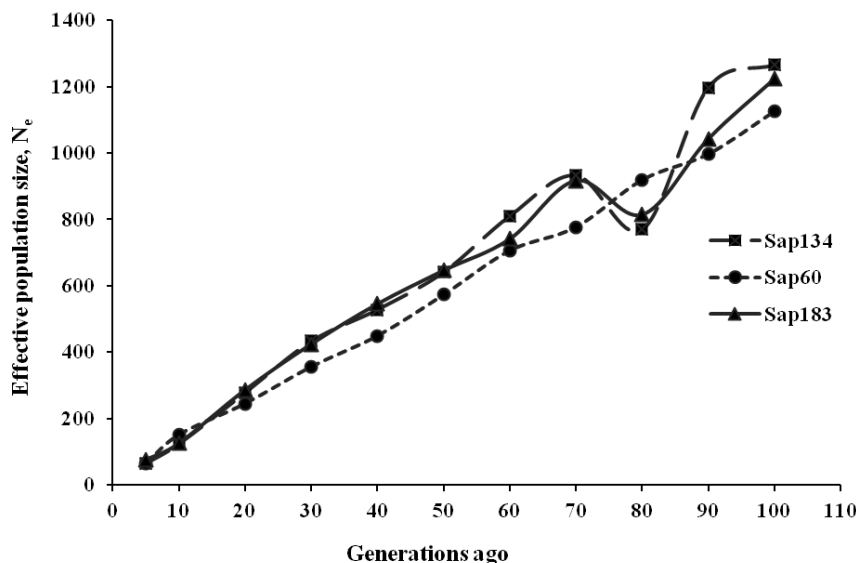
Figure 4. Estimates of effective population size (N_e) for Sapsaree at t ancestral generations ago based on the squared correlation coefficient (r^2) between SNPs. Three SNP panels (Sap134, Sap60 and Sap183) were used, in which 20K, 170K, and the combined set of 20K and 170K SNPs were genotyped with 134, 60, and 183 individuals, respectively.

marker order or the relative distances between the markers are a primary concern, even though the biases may be diluted by the large numbers of SNP pairs. Another factor of biased r^2 estimates is recombination hot-spots on the test chromosomes, causing overestimated N_e if the genetic distance was estimated to be smaller or greater. Thus, more reliability in estimating N_e requires least biased estimates of genetic distances between SNPs. However, the small sample size in the study may be sufficient to obtain precise N_e estimates. According to Bartley et al. (1992), a sample size over 90 yielded genetic estimates with good precision. In this study, three SNP panels with different map densities and sample sizes were applied (Table 1), and the past N_e estimates were in general consistent across the three SNP panels (Table 5).

A reduction in the effective population size happens during domestication, breed formation, or artificial breeding. For example, the N_e was estimated as 20 in Japanese black cattle in 1990s (Nomura et al., 2001), 100 in Holstein cattle (Young and Seykora, 1996), 55 in Finnish Yorkshire pig, and 80 in Finnish Landrace (Uimari and Tapio, 2011). The lower genetic diversity caused by the small N_e is a great concern for animal breeders regarding species viability and genetic disorders. A minimum N_e size of 50 or 100 needs to be maintained in a population, breed, or species (FAO, 2000; Meuwissen, 2009).

In Sapsaree, the increasing inbreeding rate due to a small number of founders at the re-establishing stage may have caused a low N_e despite the rapid increase of the population in the last few decades. Although the breeding programs focused on conforming to a breed standard, controlled inbreeding is a priority in maintaining the

population, especially in the Sapsaree breed, in which a few founders were used to form the breed population. A long-term genetic variability is also needed for future genetic gain in the breeding plans. Therefore, appropriate selection methods have to be chosen to maximize selection responses while fixing or minimizing rate of inbreeding (Meuwissen, 1997; Colleau and Tribout, 2008) or to optimize the use of the genetic resources of parental generation (Sanchez et al., 2003). Genomic selection (Meuwissen et al., 2001) can be an alternative method in the Sapsaree breeding plans to estimate genomic breeding values (GEBV), which enables rate of inbreeding to decrease, compared with current BLUP-based BV estimation (Daetwyler et al., 2007).

In conclusion, our results revealed small effective

Table 5. Effective population size (N_e) of Sapsaree in the past generations

Generations ago	SNP panels[a]		
	Sap134	Sap60	Sap183
5	65	64	75
10	124	151	125
20	245	278	286
50	573	641	646
80	919	769	814
100	1,126	1,265	1,224
200	2,205	2,952	2,240
300	3,464	3,627	4,195
400	4,685	4,665	3,344
500	5,614	5,699	5,381

[a] Three sets of SNP panels were used with genotyped with the Illumina CanineSNP20K chips (Sap134), CanineHD 170K BeadChip (Sap60), and the combined SNP set of 20K and 170K (Sap183) using 134, 60, and 183 individuals, respectively.

population sizes of the Sapsaree breed in the last few generations. Therefore, while the breeding programs are implemented for body shape, morphology, and behavior toward the breed standards genetic variability which was ignored during the last several decades due to the fast process of the breed re-establishment must now be considered to further avoid inbreeding. For efficient implementation of the breeding program in the Sapsaree population, genomic selections can be an effective tool to increase genetic gains for the target traits, while selecting parents with least genetic relatedness to produce progeny in the successive generations.

ACKNOWLEDGEMENTS

REFERENCES

AKC (The American Kennel Club). 1998. The complete dog book. Howell Book House, New York, USA.

Aspi, J., E. Roininen, M. Ruokonen, I. Kojola and C. Vilà. 2006. Genetic diversity, population structure, effective population size and demographic history of the Finnish wolf population. Mol. Ecol. 15:1561-1576.

Bartley, D., M. Bagley, G. Gall and B. Bentley. 1992. Use of linkage disequilibrium data to estimate effective size of hatchery and natural fish populations. Conserv. Biol. 6:365-375.

Beaumont, M. A. 2003. Conservation genetics. pp 779-812 in Handbook of Statistical Genetics (Ed. D. J. Balding, M. Bishop and C. Cannings). Wiley, London.

Berg, P. 2009. Original Eva User's Guide by Peer Berg. Danish Faculty of Agricultural Sciences, Department of Animal, Breeding and Genetics, P.O. Box 50, DK-8830 Tjele, Denmark, E-mail: Peer.Berg@agrsci.dk.

Calboli, F. C. F., J. Sampson, N. Fretwell and D. J. Balding. 2008. Population structure and inbreeding from pedigree analysis of pure bred dogs. Genetics 179:593-601.

Colleau, J. J. and T. Tribout. 2008. Optimized management of genetic variability in selected pig populations. J. Anim. Breed. Genet. 125:291-300.

Crow, J. F. and M. Kimura. 1970. An introduction to population genetics theory. Burgess Publishing, Minneapolis.

Daetwyler, H. D., B. Villanueva, P. Bijma and J. A. Woolliams. 2007. Inbreeding in genome-wide selection. J. Anim. Breed. Genet. 124:369-376.

Duchev, Z., O. Distl and E. Groeneveld. 2006. Early warning system for loss of diversity in European livestock breeds. Archiv. Anim. Breed. 49:521-531.

Falconer, D. S. and F. C. Mackay. 1996. Introduction to Quantitative Genetics, 4th ed., Longman Group Ltd, England.

FAO. 1998. Secondary guidelines for development of national farm animal genetic resources management plans: Management of small populations at risk. UN Food and Agric. Org. http://www.fao.org/AG/AGAInfo/resources/en/pubs_gen.html.

FAO. 2000. Secondary guidelines for development of national farm animal genetic resources management plans: Management of small populations at risk. UN Food and Agric. Org. http://www.fao.org/AG/AGAInfo/resources/en/pubs_gen.html.

Gutierrez, J. P., I. Cervantes, A. Molina, M. Valera and F. Goyache. 2008. Individual increase in inbreeding allows estimating effective sizes from pedigrees. Genet. Sel. Evol. 40:359-378.

Gutierrez, J. P., I. Cervantes and F. Goyache. 2009. Improving the estimation of realized effective population sizes in farm animals. J. Anim. Breed. Genet. 126:327-332.

Han, K. I., M. Alam, Y. M. Lee, D. H. Lee, J. H. Ha and J. J. Kim. 2010. A study on the morphology and Behavior of the Sapsaree: A Korean native dog (*Canis familiaris*). J. Anim. Sci. Technol. 52:481-490.

Hayes, B. J., P. M. Visscher, H. C. McPartlan and M. E. Goddard. 2003. Novel multi locus measure of linkage disequilibrium to estimate past effective population size. Genome Res. 13:635-643.

Hill, W. G. 1981. Estimation of effective population size from data on linkage disequilibrium. Genet. Res. 38:209-216.

Hill, W. G. and A. Robertson. 1968. Linkage disequilibrium in finite populations. Theor. Appl. Genet. 38:226-231.

Lander, E. S. and N. J. Schork. 1994. Genetic dissection of complex traits. Science 265:2037-2048.

Leberg, P. 2005. Genetic approaches for estimating the effective size of populations. J. Wildl. Manage. 69:1385-1399.

Lewontin, R. C. 1964. The interaction of selection and linkage. I. General considerations: heterotic models. Genetics 49:49-67.

Lindblad-Toh, K., C. M. Wade, T. S. Mikkelsen, E. K. Karlsson, D. B. Jaffe, M. Kamal, M. Clamp, J. L. Chang, E. J. Kulbokas, M. C. Zody et al., 2005. Genome sequence, comparative analysis and haplotype structure of the domestic dog. Nature 438:803-819.

Lou, X.Y., R. J. Todhunter, M. Lin, Q. Lu, T. Liu, Z. Wang, S. P. Bliss, G. Casella, G. M. Acland, G. Lust and R. Wu. 2003. The extent and distribution of linkage disequilibrium in a multi-hierarchic outbred canine pedigree. Mamm. Genome 14:555-564.

MacCluer, J. W., A. J. Boyce, B. Dyke, L. R. Weitkamp, D. W. Pfenning and C. J. Parsons. 1983. Inbreeding and pedigree structure in Standardbred horses. J. Hered. 74:394-399.

Maignel, L., D. Boichard and E. Verrier. 1996. Genetic variability of French dairy breeds estimated from pedigree information. Bulletin – International Bull Evaluation Service 14:49-54.

Meuwissen, T. H. E. 1997. Maximising the response of selection with a predefined rate of inbreeding. J. Anim. Sci. 75:934-940.

Meuwissen, T. H. E. 2009. Genetic management of small populations: A review. Acta Agric. Scand. Section A. Anim. Sci. 59:71-79.

Meuwissen, T. H. E., B. Hayes and M. E. Goddard. 2001. Prediction of total genetic value using genome-wide dense marker maps. Genetics 157:1819-1829.

Nomura, T. 2009. Interval estimation of the effective population size from heterozygote-excess in SNP markers. Biometrical Journal 51:996-1016 DOI:10.1002/bimj.200900097.

Nomura, T., T. Honda and F. Mukai. 2001. Inbreeding and effective population size of Japanese Black cattle. J. Anim. Sci. 79:366-370.

Ostrander, E. A. and L. Kruglyak. 2000. Unleashing the canine

genome. Genome Res. 10:1271-1274.

Randi, E., V. Lucchini, M. F. Christensen, N. Mucci, S. M. Funk, G. Dolf and V. Loeschke. 2000. Mitochondrial DNA variability in Italian and east European wolves: detecting the consequences of small population size and hybridization. Conserv. Biol. 14:464-473.

Rogers, C. A. and A. H. Brace. 1995. The international encyclopedia of dogs. Howell Book House, New York, USA.

Sanchez, L., P. Bijma and J. A. Woolliams. 2003. Minimizing inbreeding by managing genetic contributions across generations. Genetics 164:1589-1595.

Sargolzaei, M., H. Iwaisaki and J. J. Colleau. 2006. CFC (Contribution, Inbreeding and Coancestry) Release 1.0. A software package for pedigree analysis and monitoring genetic diversity. Graduate School of Science and Technology, Niigata University, Niigata, 950-2181, Japan.

Schwartz, M. K., D. A. Tallmon and G. Luikart. 1998. Review of DNA-based census and effective population size estimators. Anim. Conserv. 1:293-299.

Sutter, N. B., M. A. Eberle, H. G. Parker, B. J. Pullar, E. F. Kirkness, L. Kruglyak and E. A. Ostrander. 2004. Extensive and breed-specific linkage disequilibrium in Canis familiaris. Genome Res. 14:2388-2396.

Sved, J. A. 1971. Linkage disequilibrium and homozygosity of chromosome segments in finite populations. Theor. Popul. Biol. 2:125-141.

Tenesa, A., P. Navarro, B. J. Hayes, D. L. Duffy, G. M. Clarke, M. E. Goddard and P. M. Visscher. 2007. Recent human effective population size estimated from linkage disequilibrium. Genome Res. 17:520-526.

Uimari, P. and M. Tapio. 2011. Extent of linkage disequilibrium and effective population size in Finnish Landrace and Finnish Yorkshire pig breeds. J. Anim. Sci. 89:609-614.

Wang, J. 2005. Estimation of effective population sizes from data on genetic markers. Philos. Trans. R. Soc. B. 360:1395-1409.

Weir, B. S. and W. G. Hill. 1980. Effect of mating structure on variation in linkage disequilibrium. Genetics 95:477-488.

Wilcox, B. and C. Walkowicz. 1995. Atlas of Dog Breeds of the World. T. F. H. Publications, Neptune City, NJ, USA.

Wright, S. 1969. Evolution and the genetics of populations: The Theory of Gene Frequencies, Vol. II, University of Chicago Press, Chicago, USA.

Young, C. W. and A. J. Seykora. 1996. Estimates of inbreeding and relationship among registered Holstein females in the Unites States. J. Dairy Sci. 79:502-505.