

Population structure of a colony of dog guides

J. B. Cole*¹, D. E. Franke*², and E. A. Leighton†

*Department of Animal Science, Louisiana State University, Baton Rouge 70803; and

†The Seeing Eye, Inc., Morristown, NJ 07963

ABSTRACT: The objective of this study was to describe changes in genetic diversity in a colony of dog guides since its founding. Two breeds, German Shepherds (GS) and Labrador Retrievers (LR), were evaluated. Data were pedigrees of 4,699 GS and 3,573 LR dogs bred for use as guides by The Seeing Eye, Inc., Morristown, NJ. Rapid increases in average pairwise numerator relationships occurred in both breeds, although the average was approximately one-third higher in the GS population than in the LR population. A similar trend was observed for average inbreeding. The rate of increase in inbreeding has slowed in recent generations. In the current generation, relationship and inbreeding for all animals averaged 25.3 and 26.2% in GS, and 15.5 and 22.0% in LR, respectively. Effective founder numbers initially decreased in GS until Generation 3, and then increased steadily. Effective founder

number in LR constantly increased. There was a constant increase in effective founder number in LR. A similar pattern was noted for effective ancestor number as well. Founder genome equivalents were initially higher in GS but decreased over time in both breeds. Generation intervals averaged 23.7 mo in GS and 23.2 mo in LR. Sires had average service lives of 2.7 and 2.2 generations in GS and LR, respectively. Dams had average service lives of 2.1 generations in both breeds. Litter sizes averaged 5.1 and 7.4 pups per litter for GS and LR, respectively. Effective founder and ancestor numbers have slowly increased over time, and heterozygosity as measured by the number of founder genome equivalents in the population has increased. Limitations on the number of matings permitted for sires and service life of dams has led to a plateau for inbreeding and relationships. The importation of germplasm from other working dog colonies is desirable.

Key Words: Genetic Diversity, Guide Dogs, Population Structure

©2004 American Society of Animal Science. All rights reserved.

J. Anim. Sci. 2004. 82:2906–2912

Introduction

The genetic structure of a population is of interest to geneticists who are tasked with its management. Information about the genetic structure of a population can be recovered from pedigrees when accurate records have been maintained. Coefficients of inbreeding and relationship (Wright, 1922) are the most common measures of diversity calculated from pedigrees. Other measures of diversity derived from pedigrees have been used to assess the health of zoo and wild animal populations for years, but have been applied to domestic animal populations only recently (Sölkner et al., 1998; Roughsedge et al., 1999).

Coefficients of inbreeding are commonly used to quantify the loss of genetic diversity in a population

over time; however, inbreeding trend is only reliable as a measure of diversity when a population is closed to migration, randomly mating, and finite in size (Wright, 1931). Breeding colonies clearly violate these assumptions. Lacy (1989) proposed the idea of the number of effective founders in assessing population diversity. A founder is an ancestor whose parents are unknown. If all founders contribute to the population equally, then effective founders equal actual founders. Unequal founder contributions result in a smaller effective founder number. Boichard et al. (1997) developed the idea of effective ancestor number, which is the minimum number of ancestors necessary to explain the genetic diversity of the current population. Effective ancestors account for bottlenecks and are more accurate in populations undergoing intense selection. Caballero and Toro (2000) discussed the relationships among these and other measures of diversity in small populations, and Toro et al. (2000) demonstrated their use.

The objectives of this study were to estimate and evaluate measures of genetic diversity for German Shepherd (**GS**) and Labrador Retriever (**LR**) dog guides, as well as to identify influential sires and dams in each breed

¹Current address: Animal Improvement Programs Laboratory, USDA-ARS, Beltsville, MD 20705-2350.

²Correspondence: 105 J. B. Francioni Hall (phone: 225-578-3436; fax: 225-578-3279; e-mail: dfranke@agctr.lsu.edu).

Received February 3, 2004.

Accepted June 22, 2004.

Table 1. Summary of founders, parents, and offspring in dog guide populations

Breed ^a	No.	Founders ^b	Sires ^c	Dams ^d	Offspring ^e
GS	4,699	532	58	162	3,947
LR	3,573	19	43	128	3,383

^aGS = German Shepherd, LR = Labrador Retriever.

^bAnimals with unknown parents that were not used as sires or dams.

^cAll animals used as sires.

^dAll animals used as dams.

^eAll nonfounders not used as parents.

Materials and Methods

Data were provided by The Seeing Eye, Inc. (Morristown, NJ), which breeds and trains dogs for use as guides for blind individuals. The data included pedigrees of 4,699 GS and 3,573 LR dogs bred for use as guides, although not all these animals completed training and worked as guides. The number of founders, parents, and offspring are provided in Table 1. Animals are selected for breeding at 22 to 24 mo of age after partial completion of training. Sires and dams are selected on the basis of a selection index, which weights aptitude to be twice as important as hip quality. Some dogs used as sires or dams subsequently complete training and work as guides for blind people after finishing their time as a breeding animal. With the exception of six GS and seven LR animals imported after foundation, the breeding colony is closed.

Current selection decisions incorporate limits on the number of matings permitted an individual as well as the time span over which animals can be bred. Males are generally permitted to make 10 matings, and bitches are generally kept in the breeding colony until they reach 48 mo of age. The long-term use of such controls on matings may result in an eventual breakdown of the family lines seen in the pedigrees at the moment, and is a desirable step from the perspective of controlling inbreeding and the loss of alleles from the population.

Average coefficients of inbreeding and relationship, effective founder number, effective ancestor number, and founder genome equivalents were estimated for each breed over a 23-yr period (1977 to 2000). Generation intervals were estimated for each sex and breed using parental and litter birth dates. Lengths of sire and dam service lives were also computed.

Coefficients of relationship and inbreeding were calculated using the PyPedal package (Cole and Franke, 2002). Pedigrees were available for all dogs born into the breeding colony since its foundation. As a result, the base population used for the calculation of coefficients of inbreeding and relationship is the actual population of founders. Average coefficients of inbreeding were computed for all animals in each breed. Average numerator relationships among all individuals in the population born in a generation were computed for each breed. Average numerator relationships were computed as the average of all pairwise numerator relationships be-

tween animals in the same breed-generation group. The 10 most influential sires and dams for each breed were identified based on their average pairwise relationship to the breed in Generation 7.5 for GS, and Generation 7 for LR. These influential individuals were included in all analyses.

Founders were assigned generation codes of 0. All other animals were assigned generation codes as follows:

$$GC_o = \frac{(GC_s + GC_d)}{2} + 1$$

where GC_o , GC_s , and GC_d represent offspring, sire, and dam codes, respectively (Pattie, 1965). In this study, generation coefficients were rounded to the nearest half generation. The generation interval is the average age of parents at the time of birth of their selected offspring (Falconer and MacKay, 1996).

Founder animals were defined as individuals with unknown parents. In these data, both parents were known or neither parent was known; there were no half-founders. The effective founder number (f_e) was calculated as follows:

$$f_e = \frac{1}{\sum(p_i^2)}$$

where p_i is the proportion of genes contributed by ancestor i to the current population (Lacy, 1989). If all founders had contributed equally to the population, then f_e would be the same as the actual number of founders. When founders contribute to the population unequally, f_e is smaller than the actual number of founders. The greater the inequity in founder contributions, the smaller the effective founder number.

Lacy (1989) also defined the number of founder genome equivalents (f_g) as a measure of genetic diversity. A founder genome equivalent is the number of founders that would produce a population with the same diversity of founder alleles as the pedigree population assuming all founders contributed equally to each generation of descendants. Founder genome equivalents are calculated as follows:

$$f_g = \frac{1}{\sum\left(\frac{p_i^2}{r_i}\right)}$$

where p_i is the proportion of genes contributed by ancestor i to the current population and r_i is the proportion of founder i genes that are retained in the current population. Like f_e , f_g accounts for unequal founder contributions. Unlike f_e , f_g also accounts for the fraction of founder genomes lost from the pedigree through drift during bottlenecks. Although f_g is the more accurate description of the amount of founder variation present in a population, it can only be calculated directly for simple pedigrees. For large or complex pedigrees, the number of founder genome equivalents must be approximated based on computer simulation of a large number of segregations through the pedigree. This is done by assigning each founder a unique pair of alleles and randomly transmitting those alleles through the pedigree (MacCluer et al., 1986). The number of founder genome equivalents is similar to the effective founder number, but the former has been devalued based on the proportion of its genome that has probably been lost to drift (Lacy, 1989).

In populations that have undergone a bottleneck the effective number of founders computed using the previous equation is overestimated. Large contributions made by recent ancestors are more important to the population with respect to the loss of genetic diversity than equal contributions made long ago. Boichard et al. (1997) proposed a second measure of diversity to deal with such situations, the effective number of ancestors (f_a), which considers the genetic contribution of all ancestors in the population, not just founders. The effective number of ancestors treats all ancestors in the population the same way, and is computed as:

$$f_a = \frac{1}{\sum(q_i^2)}$$

where q_i is the genetic contribution of the i th ancestor not explained by the previous $i - 1$ ancestors. The ancestors with the greatest contributions are selected iteratively. The number of ancestors with a positive genetic contribution is less than or equal to the actual number of founders.

Mean differences between breeds for length of generation interval, tenure as a breeding animal, number of litters produced, and number of pups produced per litter were tested using two-sample t -tests (Ott, 1993). Breeds were assumed to be independent populations.

Results and Discussion

Demographics

Generations overlap extensively in this population. In GS, Generations 0 and 1 spanned about 20 yr, whereas successive generations spanned shorter periods. In LR, early generations were relatively short, spanning approximately eight birth years per generation, whereas later generations tend to span a larger period of time. Variation in generation length is consid-

erable for both breeds. The majority of the most influential sires and dams in each breed were born in Generations 4 and later. Short generations are indicative of rapid replacement of breeding stock. Patterns were similar for both breeds. Generation 0 consists of all animals with unknown parents, and as long as new breeding stock migrate into the breeding colony that generation will span all birth years in the population.

Sex ratio (West et al., 2002) varies slightly from generation to generation, but over time, it remained approximately 50:50 in GS (48:52) and LR (52:48). The sex ratio is skewed for Generation 0 animals because many founder males were mated to nonfounder females produced within the breeding colony. The number of births seems to decrease in the later generations. This is due to the fact that generations overlap birth years extensively, and animals will be born into the later generations for several more years. Total births will not decrease unless the number of animals required by The Seeing Eye changes.

A downward trend similar to that for births by generation is noted for the number of sires and dams used per generation. As would be expected, more dams than sires were used each generation. It is interesting to note that this represents a much more equitable distribution of selection intensity between the two sexes than is typically the case in livestock breeding. This disparity is usually due to the fact that males can produce large numbers of offspring relative to females. Several dams in each breed had more than 50 offspring, but no dam produced more than 100 offspring. The 10 sires in each breed that produced the most offspring each had more than 100 pups. German Shepherd sires produced an average of 70.8 ± 9.8 offspring, whereas GS dams produced an average of 25.4 ± 1.3 offspring. Labrador Retriever sires produced an average of 81.9 ± 9.6 offspring, whereas LR dams produced an average of 27.5 ± 1.5 offspring. Labrador Retriever and GS sires had tenures, 2.7 ± 1.0 vs. 2.2 ± 1.1 generations, similar to those of LR and GS dams (2.1 ± 0.8 vs. 2.1 ± 0.9 generations). There was no difference in the average number of litters produced by GS and LR sires, 11.2 ± 1.6 vs. 12.2 ± 1.9 . German Shepherd and LR dams produced approximately one-third as many litters on average as sires: 3.9 ± 0.2 and 4.1 ± 0.4 , respectively. Breeds did not differ for the number of litters produced per dam. A total of 58 GS sires and 162 GS dams produced 4,106 offspring in 647 litters (5.1 ± 0.2 pups per litter), whereas 43 LR sires and 128 LR dams produced 3,520 offspring in 527 litters (7.4 ± 0.2 pups per litter; $P < 0.05$). Sires overlapped generations (i.e., sires were able to produce offspring in more than one generation). No LR sire produced litters in more than four distinct half-generation classes, and no GS sire produced litters in more than five distinct half-generation classes. German Shepherd and LR dams were similarly able to produce offspring in more than one distinct half-generation; most dams produced litters in one to three half-generations, whereas only four LR dams produced litters in

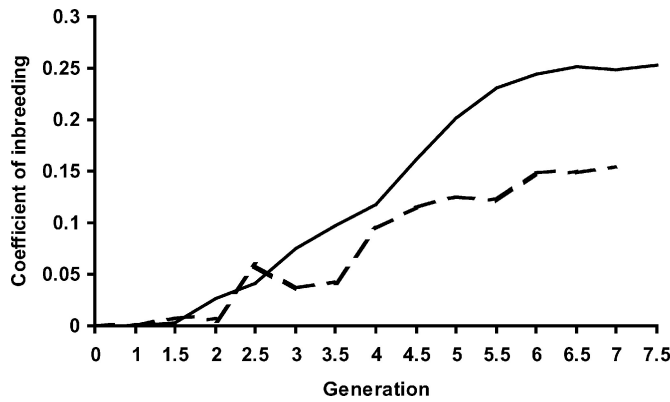


Figure 1. Average coefficients of inbreeding for German Shepherds (—) and Labrador Retrievers (— —).

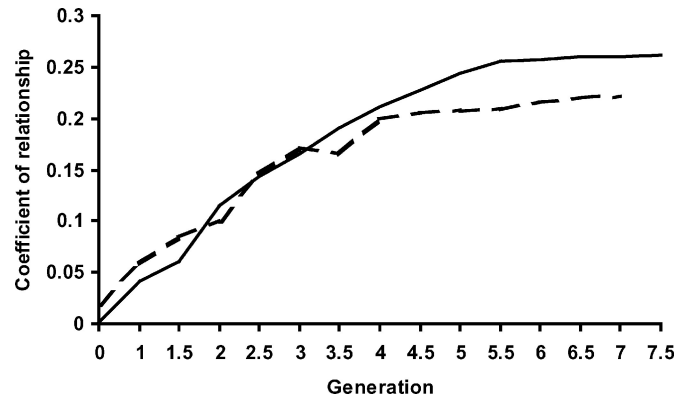


Figure 2. Average pairwise numerator relationships for German Shepherds (—) and Labrador Retrievers (— —).

four half-generations and only one GS dam in five half-generations.

Inbreeding and Relationship

Average coefficients of inbreeding are presented in Figure 1. Generations 7.5 and 7 are the current generations for GS and LR, respectively. The average inbreeding in the GS increased rapidly for several generations, and reached a plateau in Generation 5.5. The breed average inbreeding has remained steady at approximately 25% since Generation 6. Inbreeding decreased in the LR population for the first 2.5 generations. Inbreeding decreased between Generations 2.5 and 4, which may be attributed to the importation of a small number of breeding stock into the population. Inbreeding continually decreased from Generation 4 to 5.5, and then leveled off thereafter. The level of inbreeding in the LR breed for the latest three half-generations has remained steady at approximately 15%. Inbreeding has remained constant in both breeds over the past few generations. This is due to fewer animals born in later generations at present (Table 2), as well as to restrictions placed on use of sires and dams.

Average numerator relationships are shown in Figure 2. Numerator relationships are only equal to coefficients of relationship when parents are not inbred. The rate of increase in relationship was much higher than that for inbreeding, and was similar for both breeds until Generation 3. Relationships in the GS continued

to increase at approximately the same rate, whereas the rate of increase slowed for the LR. The LR reached a plateau starting with Generation 4, whereas the GS did not reach a similar plateau until Generation 5.5. The average pairwise numerator relationships of animals born in Generation 7 were 0.26 ± 0.004 (GS) and 0.22 ± 0.001 (LR), which is similar to paternal half-sibs ($r = 0.25$). The average pairwise numerator relationships between all animals in the pedigree averaged over all generations were 0.16 ± 0.07 (GS) and 0.15 ± 0.06 (LR).

Influential Individuals

A large proportion of the nonzero relationship coefficients in each breed may be attributed to a small number of highly influential sires and dams. Only animals with offspring were considered when identifying influential animals; there are animals with similarly large relationships to their respective breeds that did not produce offspring. As animals with no offspring do not contribute to the genetic structure of the population, they were not of interest for this discussion. Average relationship to the breed was used to identify influential individuals rather than the number of offspring because the former more accurately represents the effect of an individual on the breed at the time of analysis. Contributions to later generations through offspring that become sires and dams of successive generations are also accounted for. The production of a large number of off-

Table 2. Number of animals by breed and generation in dog guide populations

Breed ^a	Generation code														
	0	1	1.5	2	2.5	3	3.5	4	4.5	5	5.5	6	6.5	7	7.5
GS	593	516	539	319	425	282	359	339	312	186	218	242	241	91	37
LR	53	479	229	228	471	373	408	233	234	342	184	208	118	13	0

^aGS = German Shepherd, LR = Labrador Retriever.

spring that do not produce offspring of their own does not mark a sire or dam as influential because they do not affect the population in later generations. It was also noted that several of the animals producing the most offspring were also selected as most influential by this procedure. Regardless of the procedure used to identify influential ancestors, it is clear that this small group of animals is largely responsible for the magnitude of the average coefficients of inbreeding and relationship currently seen in this population.

Most of the 40 influential ancestors identified had pedigrees that extended several generations back to the founders of the breeding colony. The shallowest pedigree was for Kaiser, a late second-generation (2.5) LR sire with 123 offspring and an average relationship to the breed of 0.2. The most influential GS sire was Scotty, an early third-generation (2.6) animal with 277 offspring and an average relationship to the breed of 0.3. The GS dams Xandra and Ophelia are fourth-generation paternal half sibs out of Scotty, and are the youngest of the most influential GS dams. They have a combined total of 27 offspring and an average relationship of 0.3. The LR dams with the shallowest pedigrees are Belle and Pammy, both third-generation offspring. Belle is a Kaiser daughter with 41 offspring and an average relationship of 0.2. Pammy is a Ryan \times Belle daughter with 24 offspring and an average relationship of 0.2; Ryan is not a top-10 LR sire. No founders were identified as most-influential ancestors.

Examination of the pedigrees of the most-influential ancestors shows the extent to which line breeding has been practiced in this breeding population. Historically some males were retained in the population and used as sires for several generations with no appreciable restrictions on their use as breeding animals. These males are the origin of most of the lines seen in the pedigrees today. Hip dysplasia has a sufficiently high heritability ($h^2 = 0.26$; Hamann et al., 2003) that selection for better-quality hips favors line breeding. Research also suggests that various measures of intelligence or aptitude in the dog are also lowly to moderately heritable (Ruefenacht et al., 2002).

Effective Founder and Ancestor Numbers

Estimates of f_e and f_a for each breed by generation are presented in Figures 3 and 4, respectively. The effective ancestor numbers for both breeds were lower than the effective founder numbers, as expected, and exhibited a similar trend over time as f_e . The increase in f_a in GS is probably due to changes in management to increase the diversity of the sire and dam portfolio. A small increase in f_a in GS in 1999 and 2000 may be due to the influence of five breeding animals imported between 1994 and 1997. Those five animals collectively produced 96 offspring, although the average numerator relationship of each to all animals in the breed was less than 1%.

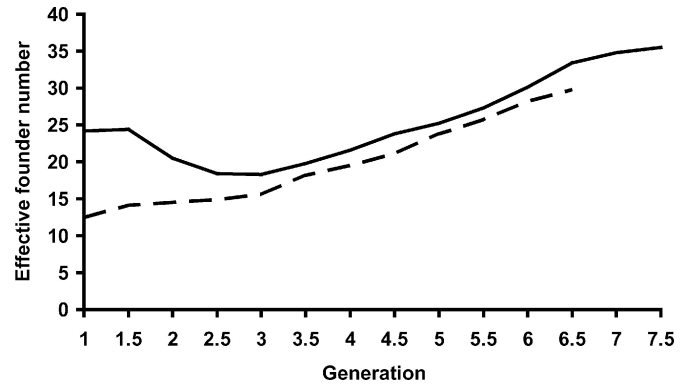


Figure 3. Effective founder number for German Shepherds (—) and Labrador Retrievers (---).

Counter to expectations, diversity increased in both breeds as measured by effective founder and ancestor numbers. The use of sires such as Scotty, who was used as a sire for 6 yr and who produced 277 GS pups, and Scotty's grandsire, Quasar, who served as a sire for 7 yr and produced 342 pups, was expected to result in decreases in f_e and f_a . In comparison, no LR sire produced more than 245 pups, although five sires did produce 200 or more offspring, and increases in f_e and f_a in the LR trailed those in the GS. The observed increase in variability must be due to the continued immigration of parents into both breeds over time. The GS population has always been larger than the LR population, allowing for the use of more sires and dams in that breed, and resulting in an increase in both of these measures of diversity over time.

Founder Genome Equivalents

Estimates of f_g for each breed by generation are presented in Figure 5. The effective founder numbers presented above are not adjusted for loss of genetic vari-

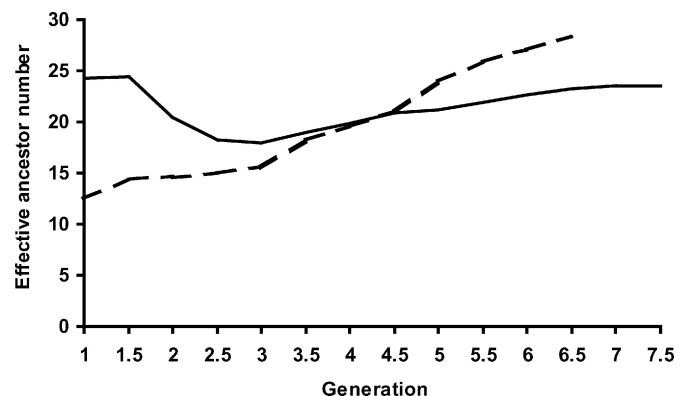


Figure 4. Effective ancestor number for German Shepherds (—) and Labrador Retrievers (---).

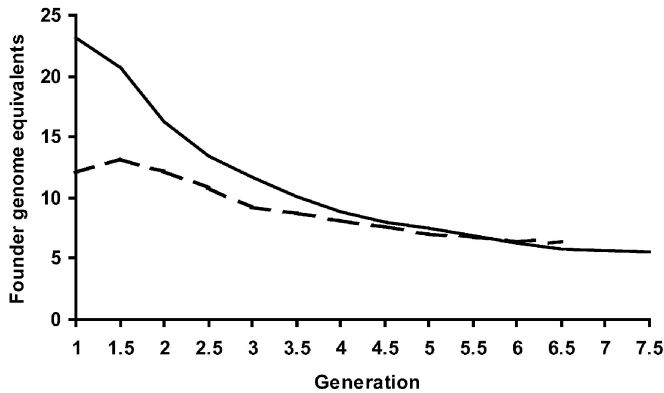


Figure 5. Founder genome equivalents for German Shepherds (—) and Labrador Retrievers (---).

ability due to genetic drift over time, and may be biased upwards. In populations that have undergone bottlenecks, genetic drift is an important factor contributing to the loss of diversity. Lacy (1989) showed that $f_g \approx n_e/g$, where n_e is the effective population size and g is the generation since foundation. The number of founder genome equivalents should decrease over time, which is consistent with the results shown in Figure 5. Almost twice as many founders were in the LR in Generation 0, but a greater loss of diversity occurred in the GS since foundation. Founder genes were lost as a quadratic function of time in the GS and as a linear function of time in the LR. The rate of allele loss has slowed in both breeds in recent years.

Generation Interval

Overall generation intervals were similar between the two breeds at 26.6 ± 0.8 mo for GS, and 25.1 ± 0.5 mo for LR. Separate generation intervals were calculated for each sex-breed combination due to the fact that fewer sires than dams were used in the breeding colony. German Shepherd sires had an average generation interval (28.1 ± 1.9 mo), which was similar to dams (25.9 ± 0.8 mo). Generation intervals for LR sires (25.5 ± 1.8 mo) and dams (24.8 ± 0.4 mo) also were similar.

Assuming onset of sexual maturity in the dog at 6 mo of age, and a 60-d gestation length, the shortest theoretical generation interval in the dog is 8 mo, although 12 to 14 mo is more probable. The average generation interval in both breeds and sexes is approximately 24 mo and reflects the health screening and training that dog guides must receive. The use of relatively old dogs as sires and dams is due to the use of animals early in the breeding program for many litters over several half-generations.

Generation intervals average approximately 24 mo in both breeds, and although it would be desirable to select sires at a younger age from a theoretical perspective, it is not practical in this case due to the intense selection for health and aptitude practiced in this par-

ticular colony. However, the relatively short generation intervals in concert with good screening and strong selection has probably contributed to the excellent hip quality seen in these dogs (Leighton, 1997). The number of sires and dams used each generation, as well as the number of puppies whelped, has remained fairly constant over time. This reflects demand for dog guides, resources available for raising and training puppies, and resources available for matching and training dogs and handlers (The Seeing Eye, Inc., 2002). The high, nonrecoverable cost of producing dog guides also contributes to this. Sire and dam service lives are reasonably short and reflect the availability of new breeding stock. Current management practices should be maintained, notably the periodic immigration of breeding stock.

Implications

The results with respect to genetic diversity in populations of dog guides are more promising than the levels of inbreeding and relationship in the population might suggest. Effective founder and ancestor numbers have slowly increased over time, and heterozygosity as measured by the number of founder genome equivalents in the population is increasing. Steps taken in the past few years to limit the number of matings permitted for sires and dams has led to a plateau for inbreeding and relationship levels, and will no doubt result in a decrease in those levels over time.

Literature Cited

- Boichard, D., L. Maignel, and E. Verrier. 1997. The value of using probabilities of gene origin to measure genetic variability in a population. *Genet. Sel. Evol.* 29:5–23.
- Caballero, A., and M. A. Toro. 2000. Interrelations between effective population size and pedigree tools for the management of conserved populations. *Genet. Res.* 75:331–343.
- Cole, J. B., and D. E. Franke. 2002. PyPedal: A package for pedigree analysis using the Python programming language. *J. Dairy Sci.* 85(Suppl.1):323. (Abstr.)
- Falconer, D. S., and T. C. F. MacKay. 1996. *Introduction to Quantitative Genetics*. 4th ed. Longman Group Ltd., Essex, U.K.
- Hamann, H., T. Kirchhoff, and O. Distl. 2003. Bayesian analysis of heritability of canine hip dysplasia in German Shepherd Dogs. *J. Anim. Breed. Genet.* 120:258–268.
- Lacy, R. C. 1989. Analysis of founder representation in pedigrees: Founder equivalents and founder genome equivalents. *Zoo Biol.* 8:111–123.
- Leighton, E. A. 1997. Genetics of canine hip dysplasia. *J. Am. Vet. Med. Assoc.* 210:1474–1479.
- Ott, R. L. 1993. *An Introduction to Statistical Methods and Data Analysis*. 4th ed. Duxbury Press, Belmont, CA.
- Pattie, W. 1965. Selection for weaning weight in Merino sheep. *Aust. J. Agric. Exp. Anim. Husb.* 5:353–360.
- Roughsedge, T., S. Brotherstone, and P. M. Visscher. 1999. Quantifying genetic contributions to a dairy cattle population using pedigree analysis. *Livest. Prod. Sci.* 60:359–369.
- Ruefenacht, S., S. Gebhardt-Henrich, T. Miyake, and C. Gaillard. 2002. A behaviour test on German Shepherd dogs: Heritability of seven different traits. *Appl. Anim. Behav. Sci.* 79:113–132.
- Sölkner, J., L. Filipic, and N. Hampshire. 1998. Genetic variability of populations and similarity of subpopulations in Austrian cattle breeds determined by analysis of pedigrees. *Anim. Sci.* 67:249–256.

- The Seeing Eye, Inc. 2002. The Seeing Eye 2002 Annual Report. Available: <http://www.seeingeye.org/docs/AnnualReport2002.pdf>. Accessed May 14, 2004.
- Toro, M. F., J. Rodríguez, L. Silio, and C. Rodríguez. 2000. Genealogical analysis of a closed herd of Black Hairless Iberian Pigs. *Conserv. Biol.* 14:1843–1851.
- West, S. A., S. E. Reece, and B. C. Sheldon. 2002. Sex ratios. *Heredity* 88:117–124.
- Wright, S. 1922. Coefficients of inbreeding and relationship. *Am. Nat.* 56:330–338.
- Wright, S. 1931. Evolution in Mendelian populations. *Genetics* 16:97–159.