RESEARCH NOTE

Analysis of genetic diversity in a close population of Zandi sheep using genealogical information

FARHAD GHAFOURI-KESBI[*]

*Department of Animal Breeding and Genetics, Animal Science Research Institute of Iran (ASRI),
P. O. Box 1483, Karaj 3146618361, Iran*

Introduction

In a Fisher–Wright idealized population (Fisher 1930; Wright 1931), characterized by Poisson distribution of family size, large number of breeders, random mating, equal sex ratio, absence of selection, mutation, or migration, and discrete generations, the rate of loss of heterozygosity is inversely proportional to population size. In real populations, however, genetic diversity is affected by several ongoing natural processes of mutation, migration, genetic drift and selection. For this reason, Wright (1931) proposed the concept of the effective population size (N_e) for extending population genetics theory developed for ideal populations to real populations. This parameter has been useful in designing strategies for conservation of endangered animal species because of its direct relationship with the level of inbreeding (Caballero and Toro 2000). However, when N_e is estimated from pedigree data, the parameter accuracy may be affected by several factors, as extensively reviewed by Boichard *et al.* (1997).

Consequently, in recent decades, new complementary techniques have been developed for the analysis of genealogies. Lacy (1989) introduced the concept of effective number of founders (f_e) in order to measure the overall founder representation in a managed population accounting for the loss of genetic diversity from unequal founder contributions, and the concept of founder genome equivalents (f_g), also called effective number of founder genomes, to account the effects of bottlenecks, unequal founder contributions, and genetic drift. Boichard *et al.* (1997) developed the concept of effective number of ancestors (f_a) to supplement the effective number of founders. The numbers obtained from

calculation of f_a are essentially same as f_e, but taken into account the decrease in genetic variation in populations that have passed through a bottleneck. Caballero and Toro (2000) proposed the concept of effective number of non-founder genomes (f_{ne}) to describe the relation between effective number of founders and founder genome equivalents. Although, the use of these new tools in analysis of genetic diversity is a recent innovation pioneered by researchers working on endangered wild species, in recent years, application of these approaches has become popular in analysis of domestic animal populations (Boichard *et al.* 1997; Goyache *et al.* 2003; Gutiérrez *et al.* 2003; Hagger 2005).

The aim of the present study was to evaluate genetic diversity in a close population of Zandi sheep using various measures of genetic diversity (Lacy 1989; Boichard *et al.* 1997; Caballero and Toro 2000; Gutiérrez *et al.* 2009).

Materials and methods

Data

Pedigree information used in this study was from the Zandi Sheep Breeding Station, situated in Khojir national park between Tehran and Abali. The location of the experimental flock as well as their natural habitat, description, and husbandry practices were described by Ghafouri-Kesbi *et al.* (2008).

The degree of completeness of the analysed pedigree was assessed by computing: (i) the number of fully traced generations, (ii) the maximum number of generations traced, and (iii) the number of equivalent complete generations. The first is defined as those separating the offspring of the furthest generation where the 2^g ancestors of the individual are known. Ancestors with unknown parent were considered as founders (generation 0). The second is the number of generations separating the individual from its furthest ancestor.

Keywords. genetic diversity; pedigree; sheep; inbreeding.

The equivalent complete generations is computed as the sum over all known ancestors of the terms computed as the sum of $(1/2)^n$ where n is the number of generations separating the individual to each known ancestor.

For the analyses, the reference population was defined as those animals which were alive.

Measures of genetic diversity

Coancestry, f: Coancestry (kinship, Malécot 1948) is defined as the probability that any two alleles, sampled at random (one from each individual), are identical copies of an ancestral allele.

Individual inbreeding coefficient, F: It describes the probability that two alleles at any locus are identical by descent (Wright 1931). F values were computed for each individual following Meuwissen and Luo (1992).

Individual average relatedness, AR: The AR coefficient for each individual in the pedigree is computed as the average of the coefficients in the row corresponding to the individual in the numerator relationship matrix, A (Gutiérrez and Goyache 2005).

The effective number of population size, N_e: This parameter was obtained following Gutiérrez *et al.* (2009) based on individual increase in inbreeding ($\triangle F_i$). The $\triangle F_i$ coefficients are computed as $\triangle F_i = 1 - \sqrt[t-1]{1 - F_i}$, where F_i is the individual coefficient of inbreeding and t is the equivalent complete generations. This estimate of effective population size ($\overline{N_e}$) can be computed from $\overline{\triangle F}$, that can be computed by averaging the $\triangle F_i$s of the n individuals included in a given reference subpopulation, as $\overline{N_e} = 1/2\overline{\triangle F}$.

The effective number of founders, f_e: Following Lacy (1989), I estimated the effective number of founders as $f_e = 1/\sum_{k=1}^{m} q_k^2$, where q_k is the expected proportional genetic contribution of founder k, calculated by the average relationship of the founder to each animal in the current population, and m is the total number of founders.

The effective number of ancestors, f_a: This parameter complements the information offered by the effective number of founders accounting for the losses of genetic variability produced for the unbalanced use of reproductive individuals producing bottlenecks and computed as $f_a = 1/\sum_{k=1}^{n} p_k^2$, where p_k is the marginal contribution of each ancestor, i.e., the contribution made by an ancestor which is not already explained by a previously chosen ancestor. The sum of marginal contributions of all ancestors is one.

The effective number of founder genomes, f_g: Following Caballero and Toro (2000), parameter f_g was obtained by the inverse of twice the average coancestry of the individuals included in a pre-defined reference population. Since f_g accounts for all causes of gene loss during segregations, f_g always is a smaller number than both f_a and f_e.

The effective number of non-founder genomes, f_{ne}: The 5th type of effective number of animals, the effective number of non-founder genomes, accounts only for the effect of genetic drift in non-founder generations. This effective number is obtained as $f_{ne}^{-1} = f_g^{-1} - f_e^{-1}$ (Caballero and Toro 2000).

Genealogical analyses on pedigree information were carried out using the program ENDOG, version 4.6 (Gutiérrez and Goyache 2005).

Results and discussion

The average values of f, F, AR, number of fully generations, maximum number of generations, number of equivalent generations, and the measures of genetic variability (represented by $\overline{N_e}, f_e, f_a, f_g, and f_{ne}$) for the reference population are shown in table 1. The mean f in the reference population was 1.21% (1.28% in males and 1.21% in females,

Table 1. Results of the pedigree analysis for the reference population.

Item	Total	Male	Female
No. of animals	2566	951	1615
No. of inbred animals	1104	428	676
Average coancestry (%)	1.21	1.28	1.21
Mean inbreeding (%)	1.46	1.38	1.51
Average relatedness (%)	1.91	1.90	1.92
No. of fully generations	1.92	1.95	1.90
Maximum no. of generations	4.78	5.13	4.56
No. of equivalent complete generations	3.15	3.31	3.10
Effective no. of population ($\overline{N_e}$)	71		
Effective no. of founders (f_e)	86		
Effective no. of ancestors (f_a)	74		
No. of ancestors explaining 50%	28		
No. of ancestors explaining 70%	68		
Effective no. of founder genomes (f_g)	41		
Effective no. of non-founder genomes (f_{ne})	79		

respectively). Coancestry describes how much individuals in a breeding programme are related and impacts upon the effectiveness of the selection programme. The opposite of coancestry is diversity. Briefly, the higher the coancestry, the lower will be the genetic diversity. Therefore, for effective selection, we need to manage coancestry. The mean F was 1.46% and the average AR was 1.91%. These values were higher for females than for males, however, these differences were not significant ($P > 0.05$). In addition, in the whole pedigree, these values were higher for males than for females (the mean F was 1.16% in males and 0.96% in females, respectively, and the average AR was, respectively, 1.72% in males and 1.55% in females).

The effective population size ($\overline{N_e}$) was estimated to be 71. The FAO (1998) set an effective population size of 50 animals as a critical level. Current estimate of N_e is higher than this critical value; however, it should be noted that estimates of N_e are not constant and change with time following change in inbreeding level of the population, i.e., current estimate of N_e can fall below this critical value, if the rate of inbreeding increases.

The number of founders contributed in the reference population was 615. In contrast, the effective number of founders (f_e) was 86. This shows the excessive use of some animals as parents. Goyache *et al.* (2003) analysed the pedigree information of the endangered Xalda sheep breed of Asturias which included 805 animals (562 alive animals) and 329 founders, and estimated f_e for the population as 80. The small effective number of founders suggests a smaller number of founder animals or a higher disequilibrium between founder contributions. The ratio of f_e to the number of founders was 0.14 which was smaller than that observed by Goyache *et al.* (2003) in Xalda sheep as 0.24, indicating higher disequilibrium between founder contributions in the Zandi population.

The total number of ancestors contributed in the reference population was 598. This relatively high number of ancestors is typical of populations lacking long historical pedigrees showing a short time period between the youngest generations and the founders (Gutiérrez *et al.* 2003). The effective number of ancestors (f_a) was computed to be 74. In the Xalda sheep population which included 236 ancestors, Goyache *et al.* (2003) estimated f_a as 40. Increasing generations will also increase the chance of bottleneck; therefore, in populations with a long historical pedigree knowledge, smaller estimates of f_a would be expected. Gutiérrez *et al.* (2003) reported that estimates of f_e and f_a are higher in the larger populations, especially when the size of their founder population was initially high, and are not directly dependent upon the size of their populations of reference. The calculation of f_a involves looking for marginal contribution of each ancestor to the gene pool of the reference population, starting from the total contribution of the most influential ancestor. Table 2 gives the proportion of genes contributed to the reference population by the most influential ancestors. The most

influential ancestor (a sire born in 1992 with 60 offspring) explained 4.66% of the genetic diversity of the reference population. This was 4.14%, 3.75%, 3.28%, and 3.17% for 2nd, 3rd, 4th and 5th most influential ancestors, respectively. Of 100% of the genetic diversity presented in the reference population, 50% and 70%, respectively, would be explained by 28 and 68 ancestors (figure 1) which were higher than those reported by Goyache *et al.* (2003) in Xalda sheep where the number of ancestors needed for explaining 50% and 70% of the genetic variability was 13 and 36 ancestors, respectively. The information yielded by estimating f_a is of value for assessing the applied breeding programme. Influential ancestors, selected by the estimation of f_a, specify animals that have a high contribution to the current population. These animals can then be re-examined in order to determine if they were better than average in their contemporary group, indicating they possess beneficial alleles for the selected trait.

The effective number of founder genomes, accounted for all losses of diversity, was 41 and, as expected, was smaller than both f_a and f_e. The effective number of founder genomes measures how many founders' genes have been maintained in a population for a given locus and directly related to genetic diversity. The lower f_g indicates the presence of a lower proportion of founder genes in the reference population.

The effective number of non-founder genomes was estimated to be 79. This parameter measures precisely the amount of genetic drift that has occurred during the history of the population since its foundation.

In general, deviation of effective number of founders and ancestors from the number of founders and ancestors provide evidence that the population has lost a part of its genetic diversity and highlights the necessity for planning a suitable conservation programme. In this population of Zandi sheep, breeders have tried to avoid mating of related animals and animals are generally selected based on phenotypic values, however, selection due to BLUP-breeding values is in progress. BLUP selection tends to select members of the same families, thereby increasing relatedness among the population. In addition, in recent years, artificial insemination has been considered as the major tool in the mating process. It has been shown that artificial insemination has a strong influence on the genetic diversity of domestic animal populations. Hagger (2005) reported that due to the heavy use of artificial insemination, the genetic diversity in most dairy cattle breeds is relatively small. For example, in Holstein cattle, it was discovered that only three Y chromosomes of 2500 widely-used bulls, with only 46 maternal grandsires contribute in the genome of the whole population (McDaniel 2001). In addition, Boichard *et al.* (1997) who worked on three large French bovine populations (Abondance, Normande and Limousine), found greater genetic diversity in Limousine because of the dominance of natural mating.

Table 2. Description of the most influential ancestors.

Proportion of genes contributed	Sex	Founder	No. of offspring	Contribution (%)
1st ancestor	Male	No	60	4.66
2nd ancestor	Male	No	55	4.14
3rd ancestor	Male	Yes	19	3.75
4th ancestor	Male	Yes	45	3.28
5th ancestor	Male	No	125	3.17
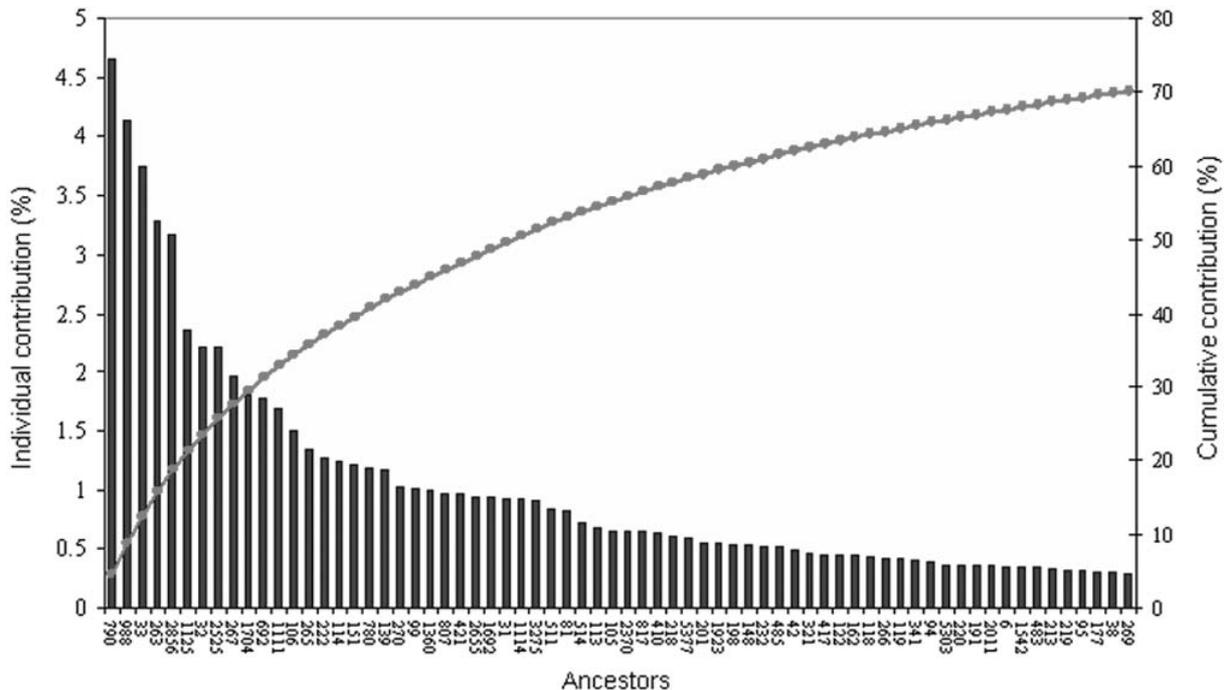
Figure 1. Individual and cumulative marginal contribution of the first 68 most influential ancestors presented 70% of the total diversity.

Different methods have been proposed to minimize or delay the loss of genetic diversity or the appearance of inbreeding (Meuwissen 1997; Caballero and Toro 2000). With respect to close populations, like this population of Zandi sheep, the best strategy in the long term is by preferentially choosing as breeding animals those with the lowest mean coancestry with the others. This approach gives breeding priority to animals with the highest probability of carrying unique alleles, or to animals with the lowest representation in the descent population (Caballero and Toro 2000). This strategy can be implemented in the population to control contributions of each parent to the next generation. It could also be possible to split the population into several families of related animals and to organize mating among animals from different families. Recently, optimum genetic contribution theory (Woolliams and Thompson 1994), which varies contributions of each animal to the next generation and selects the least related individuals, has suggested aimed at maximizing gain with a restriction on the weighted coancestry of the breeders. In comparison to BLUP truncation selection, selection using optimum genetic contribution theory helps us to have either higher response for the same level of inbreeding or lower inbreeding for the same level of response. This approach can be implemented to consider the inbreeding rate in selection strategy of small populations using computer programs such as the GENCONT package (Meuwissen 2002).

References

Boichard D., Maignel L. and Verrier E. 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 Toro M. A. 2000 Interrelations between effective population size and other pedigree tools for the management of conserved populations. *Genet. Res.* **75**, 331–343.

FAO 1998 *Secondary guidelines for development of national farm animal genetic resources management plans: management of small populations at risk.* Food and Agriculture Organization, Rome, Italy.

Fisher R. A. 1930 *The genetical theory of natural selection*, 2nd edition. Dover, New York, USA.

Ghafouri-Kesbi F., Eskandarinasab M. P. and Hassanabadi A. 2008 Investigation of genetic variability and inbreeding characteristics in a population of Zandi sheep. *Can. J. Anim. Sci.* **88**, 409–417.

Goyache F., Gutiérrez J. P., Fernández I., Gómez E., Álvarez I., Díez J. and Royo L. J. 2003 Using pedigree information to monitor genetic variability of endangered populations: the Xalda sheep breed of Asturias as an example. *J. Anim. Breed. Genet.* **120**, 95–103.

Gutiérrez J. P. and Goyache F. 2005 A note on ENDOG: a computer program for analysing pedigree information. *J. Anim. Breed. Genet.* **122**, 357–360.

Gutiérrez J. P., Altarriba J., Díaz C., Quintanilla R., Cañón J. and Piedrafita J. 2003 Pedigree analysis of eight Spanish beef cattle breeds. *Genet. Sel. Evol.* **35**, 43–64.

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

Hagger C. 2005 Estimation of genetic diversity in the brown cattle population of Switzerland obtained from pedigree information. *J. Anim. Breed. Genet.* **122**, 405–413.

Lacy R. C. 1989 Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. *Zoo. Biol.* **8**, 111–123.

Malécot G. 1948 *Les Mathématiques de Ì Hérédité*. Masson et Cie. Paris, France.

McDaniel B. T. 2001 Uncontrolled inbreeding. *J. Dairy Sci.* **84**, (E. suppl.), E185–E186.

Meuwissen T. H. E. 1997 Maximizing the response of selection with a predefined rate of inbreeding. *J. Anim. Sci.* **75**, 934–940.

Meuwissen T. H. E. 2002 Gencont: an operational tool for controlling inbreeding in selection and conservation schemes. In *Proceedings of the 7th Word Congress on Genetics Applied to Livestock Production*. CD-ROM communication no. 28-20. montpellier, France.

Meuwissen T. I. and Luo Z. 1992 Computing inbreeding coefficients in large populations. *Genet. Sel. Evol.* **24**, 305–313.

Wright S. 1931 Evolution in Mendelian populations. *Genetics* **16**, 97–159.

Woolliams J. A. and Thompson R. 1994 A theory of genetic contributions. In *Proceedings of the 5th World Congres on Genetics Applied to Livestock Production*. University of Guelph, Ontario, Canada. **19**, 127–134.