Archiv Tierzucht 55 (2012) 4, 375-384, ISSN 0003-9438

Using pedigree information to study genetic diversity and re-evaluating a selection program in an experimental flock of Afshari sheep

Farhad Ghafouri-Kesbi

Islamic Azad University, Karaj Branch, Karaj, Iran

Abstract

The purpose of this paper was to evaluate the inbreeding consequences of a short-term selection experiment which was initiated in 1998 in an Afshari sheep flock. Moreover, the conducted selection experiment was re-evaluated through assessing change in ranking of the first 10 influential ancestors when their genetic contributions were replaced with their breeding values. A total of 1 714 animals were registered in the herdbook with a founder population comprised of 243 animals. The average coancestry (f) and inbreeding (F) in the reference population were 2.1 % and 1.2 %, respectively. Estimated value of the effective population size (N_e) was 50. The effective number of founders (f_e) was estimated to be 40 and the effective number of ancestors (f_a) was 34. Estimates of breeding values revealed that owing to phenotypic selection some ancestors with lower breeding values had greater contribution to the reference genome than those with greater breeding values and, for this reason, the population has been deprived from the maximum genetic improvement that could be achieved if selection was based on breeding values. The effective number of founder genomes (f_g) was computed to be 23 and the effective number of non-founder genomes (f_{ne}) was 55. The index of genetic diversity decreased by almost 2.2 % over the period studied. In general, decrease in genetic variability was low and N_e was not very low for a small-closed population under selection, indicating even in populations which undergo selection, besides achieving genetic gain, the rate of inbreeding is controllable if matings carefully planned.

Keywords: pedigree, sheep, genetic diversity, inbreeding, breeding value

Introduction

Genetic diversity has been defined as the variety of alleles and genotypes present in a population and this is reflected in morphological, physiological and behavioural differences between individuals and populations (Frankham *et al.* 2002). In livestock populations, loss of genetic diversity generally occurs as a direct consequence of improvement programs, especially due to the increased levels of inbreeding and loss of founder alleles through genetic selection and drift (Vozzi *et al.* 2007).

Traditionally genetic diversity in populations of domestic breeds is assessed by estimating effective population size (N_e, Wright 1931). The effective population size is the size of an ideal population that has the same rate of inbreeding as the real population under consideration

with its own complicated pattern of variance in family size, sex ratio, etc. N_e has been considered as a key parameter in conservation genetics because it controls genetic drift and the response to selection. In addition to N_e, in recent years, other criteria have been developed to study genetic diversity. In 1989, Lacy published his key paper on the 'analysis of probabilities of gene origin' in which he introduced the concepts of effective number of founders (f_e) and effective number of founder genomes (f_g) to evaluate genetic diversity. The former only accounts for the unequal contribution of founders, but the later accounts for the unequal contributions of founders, bottlenecks, and random loss of alleles due to genetic drift. Boichard *et al.* (1997) discussed the value of implementing these new approaches to measure genetic variability in domestic animals populations and proposed the idea of effective number of ancestors (f_a) to supplement the effective number of founders. Caballero and Toro (2000) investigated the interrelations among these parameters and introduced a new parameter to assess the loss of genetic diversity by random genetic drift accumulated in non-founder generations, the effective number of non-founder genomes (f_{ne}). Parameters derived from the probabilities of gene origin are of value because they provide a complementary view of the within-population genetic variability. In addition, techniques developed by Boichard *et al.* (1997) are useful in determining which ancestors are the most influential and quantifying their contribution.

In Iran, while many domestic animal populations have undergone selection for years, a few efforts have been made to assess the effects of exerted selection programs on genetic variation in those populations (Ghafouri-Kesbi 2010a, b). The aim of the current study was to evaluate genetic diversity in a close population of Afshari sheep using criteria based on probability of identity-by-descend of genes and criteria based on probabilities of gene origin. In addition, the selection experiment was re-evaluated by estimating the breeding values of the most important ancestors to show if they had both the greater genetic contributions and breeding values or not.

Material and methods

Data

Pedigree information used in this study was from an experimental flock of Afshari sheep maintained at the department of Animal Science at the Zanjan University, Zanjan, Iran, 1 663 m above mean sea level and 35° 35´ S, and 47° 15´ E. The flock was closed in 1998 with 11 fertile rams and 110 breeding ewes. The aim of the experiment was to evaluate the response of Afshari sheep to selection for yearling live weight (selection criterion). The location of the experimental flock, husbandry practices and the results of the selection program were described by Ghafouri-Kesbi *et al.* (2009).

The quality of pedigree used was high. Out of 1 496 lambs that were born during the experiment, the fathers of only five lambs were not known. A total of 1 714 animals were registered in the herdbook from 1998 through 2005. For the analyses a reference population had to be defined. In the current study, the reference population was defined as 281 animals (136 males and 145 females) which were born in 2005.

The degree of completeness of the analysed pedigree was assessed by computing the number of equivalent complete generations. This was 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.

The generation interval was computed as the average age of parents at the birth of their progeny kept for reproduction. This was computed for four genetic pathways, sire to son (L_{ss}), sire to daughter (L_{sd}), dam to son (L_{ds}), and dam to daughter (L_{dd}). The average generation interval (L) was computed from:

$$L = \frac{L_{ss}+L_{sd}+L_{ds}+L_{dd}}{4} \quad (1)$$

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 2 alleles at any locus are identical-by-descent (Wright 1931).

The effective number of population size, N_e: This parameter was obtained following Cervantes *et al.* (2011) based on increase in coancestry for all pairs of individuals *j* and *k* (Δc_{jk}) in a reference subpopulation. This parameter is computed as:

$$\Delta c_{jk} = 1-\sqrt[\left(\frac{g_j+g_k}{2}\right)]{1-c_{jk}} \quad (2)$$

where c_{jk} is the inbreeding value corresponding to an offspring from *j* and *k*, and g_j and g_k are the discrete equivalent generation of individuals *j* and *k*. Averaging the increase in coancestry for all pairs of individuals in a reference subpopulation, we can estimate a realised effective population size based on coancestries as:

$$\overline{N_{ec}} = \frac{1}{2\overline{\Delta c}} \quad (3)$$

The effective number of founders, f_e: The parameter f_e indicates the number of equally contributing founders that would produce the same level of genetic diversity as that observed in the current population. Lacy (1989) estimated the effective number of founders as:

$$f_e = \frac{1}{\sum_{k=1}^{m} q_k^2} \quad (4)$$

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 number, as outlined by Boichard *et al.* (1997) is the minimum number of ancestors, not necessary founders, explaining the complete genetic diversity of the current population. 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:

$$f_a = \frac{1}{\sum_{k=1}^{n} p_k^2} \quad (5)$$

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.

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. To do such an evaluation, I used breeding values of the first 10 influential ancestors for the selection criteria (yearling weight) to evaluate change in their ranking when their genetic contributions were replaced with their breeding values. Breeding values of individual animals were obtained with the best linear unbiased prediction (BLUP) method (Ghafouri-Kesbi *et al.* 2009).

The effective number of founder genomes, f_g: The effective number of founder genomes show how many founders would be needed to produce the same genetic diversity as found in the population if all founders contributed equally and no founder alleles were lost through drift under random mating. As f_g accounts for all causes of gene loss during segregations (Lacy 1989), f_g always is a smaller number than both the f_a and f_e. Following Caballero & 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. Estimates of f_g could be used to evaluate genetic diversity in the reference subpopulations relative to a base line. The amount of genetic diversity (GD) in the reference population relative to the base population is approximated as below (Lacy 1989), when genetic diversity is expressed with the »expected hetrozygosity« (Nei 1973):

$$GD = 1 - \frac{1}{2fg} \tag{6}$$

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. GD at a given generation t is simply $1-f$ and $f_g=1/2f$ (being f the average coancestory). An estimate of GD in the base population can be computed as:

$$GD^* = 1 - \frac{1}{\frac{1}{2fe}} \tag{7}$$

The difference between GD and GD^* is an estimate of the random drift (RD) that can be expressed using f_{ne} as:

$$f_{ne} = \frac{1}{2RD} \tag{8}$$

Using estimates of f_e and f_g, one could estimate f_{ne} as (Caballero & Toro 2000):

$$f_{ne}^{-1} = f_g^{-1} - f_e^{-1} \tag{9}$$

Genealogical analyses on pedigree information were carried out using the program ENDOG ver. 4.8 (Gutiérrez & Goyache 2005).

Results and discussion

Mean of coancestry (*f*) and inbreeding (*F*) by year of birth are shown in Figure 1. The mean coancestry in the reference population was 2.1 %. This parameter highlights how much animals in a breeding population are genetically linked and limits the effectiveness of selection programs. High coancestry means that most individuals carry similar alleles, i.e., allelic variation is low in populations with high coancestry. *F* was 0 from the foundation of the herdbook until 2001; but accumulated in three last years and reached to 1.2 % in 2005. Of 1 733 animals registered in the herdbook of Afshari sheep, only 5 % were inbred, however, their mean inbreeding was relatively high (≈10 %).

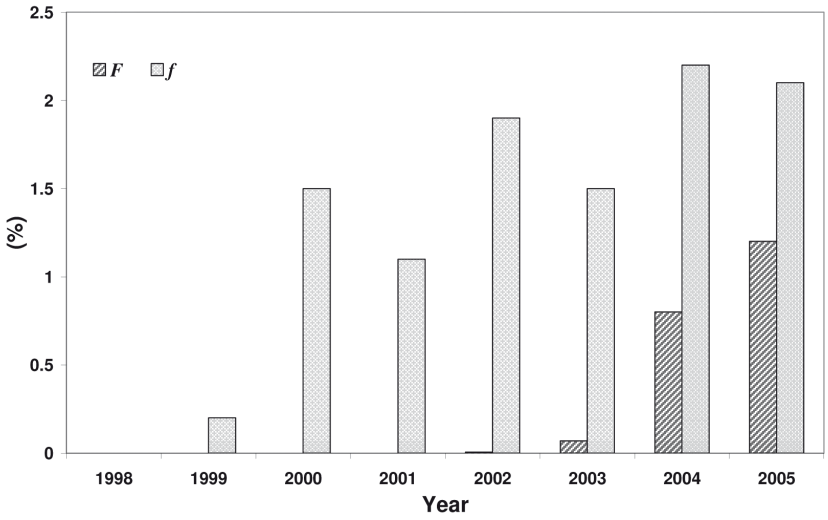

Figure 1
Means of coancestry (*f*) and inbreeding (*F*) by year of birth.

The estimates of $\overline{N_{ec}}$, f_e, f_a, f_g, f_{ne}, number of known generations and number of equivalent generations are shown in Table 1. Estimated value of N_e was 50 (Table 1), which was in the range cited for breeds of domestic animals (Bozzi *et al.* 2006, Vozzi *et al.* 2007, Ghafouri-Kesbi 2010a, b). This concept, developed by Wright (1931), gives a good intuitive understanding of the risks of selection and has a large influence on the overall level of genetic diversity in populations and on the fate of alleles under selection. Selection for advantageous alleles is more efficient in populations with a large N_e (Criscione *et al.* 2005). Evolutionary biologists have recommended that an effective population size in the range of 500-5 000 is necessary to secure evolutionary potential of natural populations (Frankham *et al.* 2002). In the field of animal breeding, the figure ranges between 50 to 100 (Bijma 2000). The reduction in effective population size, as a direct consequence of reduction in genetic diversity, associates with various unfavorable phenomena such as inbreeding depression in fitness-related traits and an increased fluctuation in selection response (Falconer & MacKay 1996). Therefore, active management of the rate of inbreeding is one of the main tasks of breeders to achieve an optimum N_e to allow the population to face future economic and environment changes and to assure a possible long-term response to selection.

Table 1
Results of the pedigree analysis for the reference population

Average co-ancestry, f	2.1 %
Average inbreeding, F	1.2 %
No. of founders	243
No. of ancestors	148
Effective no. of population, $\overline{N_{ec}}$	50
Effective no. of founders, f_e	40
Effective no. of ancestors, f_a	34
Effective no. of founder genomes, f_g	23
Effective no. of non-founder genomes, f_{ne}	55
No. of known generations	4
No. of equivalent generations for whole pedigree	1.42
No. of equivalent generations for inbred animals	2.45

The total number of founders contributed in the reference population was 243 and the effective number of founders was 40. Current estimate of f_e is lower than that reported by Goyache *et al.* (2003) for the Xalda breed of sheep including 329 founders as 80 and that reported by Ghafouri-Kesbi (2010a) for a close population of Zandi sheep including 615 founders as 86. The small ratio of the effective number of founders to the number of founders (0.16) shows a sizeable disequilibrium between founder contributions to the gene pool of the reference population. Whatever the founder contribution is unbalanced, the f_e is smaller.

While the number of ancestors contributing to the reference population was 148, the effective number of ancestors was 34, or less than a fourth. Current estimate of f_a is lower than other reports: 40 by Goyache *et al.* (2003) for the Xalda breed of sheep included 236 ancestors and of 86 by Ghafouri-Kesbi (2010a) for a close population of Zandi sheep included 598 ancestors. According to Gutiérrez *et al.* (2003) estimates of f_e and f_a are higher in the larger populations, especially when the size of their founder population was initially high.

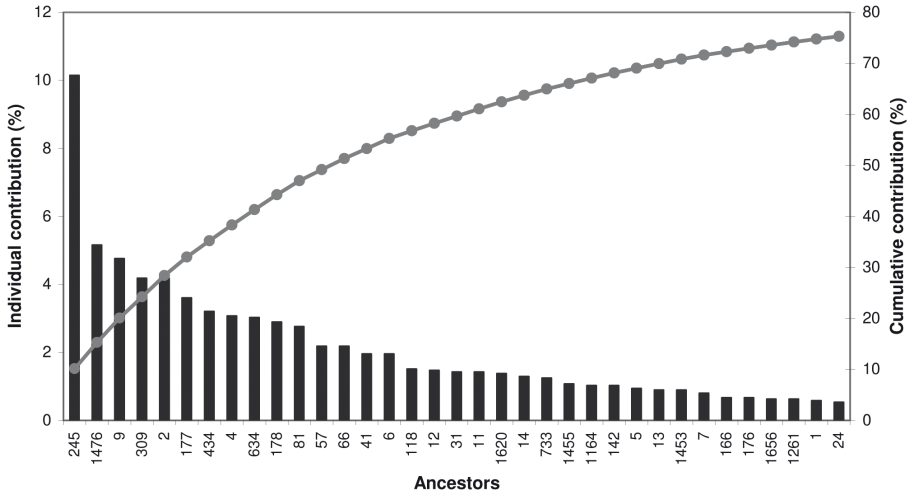

Figure 2
Individual and cumulative marginal contribution of the first 35 most influential ancestors presented 75 % of the total diversity.

As procedure of estimating f_a takes in account the possible bottlenecks in the pedigree, this resulted in a smaller estimated value than f_e, concordant with other reports (Goyache *et al.* 2003, Bozzi *et al.* 2006, Vozzi *et al.* 2007, Ghafouri-Kesbi 2010a). The comparison between the effective number of ancestors and the effective number of founders allows to reveal the decrease in genetic variation in populations that have passed through a bottleneck (Boichard *et al.* 1997). In the current study, the f_a/f_e ratio was 0.85, illustrating that a narrow bottleneck have occurred in development of this population of Afshari sheep. Figure 2 shows the marginal contribution of first 35 most influential ancestors explaining 75 % of the genetic diversity presented in the gene pool of the reference population. As shown, only 13 ancestors were necessary to explain 50 % of the genetic variability presented in the reference genome.

Table 2, gives the proportion of genes contributed to the reference population by the 10 most influential ancestors together with their breeding values for yearling weight. The first most influential ancestor was a sire born in 2000 with 104 offspring and contributed 10.1 % in the genes of the reference population. The 2nd and 3rd most influential ancestors were also sire with 78 and 49 offspring, respectively, and contributed in the genome of the reference population as 5.2 % and 4.7 %, respectively.

Table 2
Description of the most important ancestors. Their genetic contribution and their breeding values.

Ancestor	No. of equivalent generations	Contribution, %	Breeding value, kg
1st ancestor	1	10.10	2.344
2nd ancestor	1	5.16	0.965
3rd ancestor	1	4.75	4.108
4th ancestor	1	4.18	3.205
5th ancestor	0	4.17	3.679
6th ancestor	0	3.60	1.684
7th ancestor	1	3.20	4.894
8th ancestor	0	3.07	1.845
9th ancestor	1.5	3.02	1.907
10th ancestor	0	2.89	2.321

It is expected that the breeding values of the most influential ancestors, those which have been used for a longer time, be greater than ancestors with lower contribution, those which have been used for a shorter time. But it is not the case here. As shown in the Table 3, the estimated breeding value for the most influential ancestor (2.344kg) is lower than some other ancestors with lower contributions. The 2nd, 3rd and 7th influential ancestors belong to a contemporary group (they were born in the same year and selected at the same age). Surprisingly when their breeding values were taken in to consideration their ranking became vice versa, i.e., the 7th influential ancestor was ranked as first and 3rd and 2nd ancestors were ranked as the second and third, respectively. Probably it is because in this population phenotypic values were used to select superior animals instead of breeding values. If selection was based on estimates of breeding values, the 7th influential ancestor should have contributed in the population more than 2nd and 3rd ancestors because it has a greater breeding value for the yearling weight. In the other words, if animals had chosen based on their breeding values, influential ancestors would have the same ranking either based on their genetic contribution or their breeding values. Of course, other factors such

as different ability for fitness or reproductive success, breeders' decisions and pedigree known for each identified ancestor, individual nearer or belonging to the base population would tend to have lower breeding values, may change the genetic representation of an ancestor. However, in the current study, some of the most influential ancestors belong to a contemporary group (2nd, 3rd and 7th ancestors) which makes comparsion to be possible. For this reason, one could advocate the idea that genetic improvement that has achieved in this population through performed selection program should have been greater than that reported by Ghafouri-Kebi *et al.* (2009) if selection procedure was based on breeding values.

Table 3
Generation intervals (in years) for the four genetic pathways and the average generation interval.

Pathway	N	Generation interval ± SE
Sire-son	8	3.50±0.53
Sire-daughter	71	3.16±0.42
Dam-son	8	4.50±0.54
Dam-daughter	71	3.85±0.52
Total	158	3.55±0.11

Effective number of founder genomes (f_g) which accounted for all losses of genetic diversity and as pointed out by Lacy (1989) is directly related to loss of gene diversity was 23. f_g is always less than N_e, and also less than both f_e and f_a. Similar to f_e, f_g accounts for unequal founder contributions, but f_g also accounts for the fraction of founder genomes lost from the pedigree through drift during bottlenecks. The corresponding genetic diversity index, which takes all the causes of reduction in genetic diversity, decreased by 2.2 % over the period studied. In the other word, out of 100 % of the heterozygosity presented in the initial population, 2.2 % has disappeared from the population mainly due to unbalanced contributions of reproductive animals. However, the loss of genetic diversity is low for a population under selection. This is because in the population the birth information of all reproductive animals were available, made it easy to manage matings aimed at avoiding mating between close-relative animals.

Effective number of non-founder genomes (f_{ne}) which measures precisely the amount of genetic drift that has occurred during the history of the population since its foundation was 55 and was larger than f_e, suggesting that for the reduction of genetic diversity, unequal contribution of founders is far important than the random genetic drift.

The average generation interval of the reference population was 3.55 years (Table 3). Because sires were replaced earlier, the generation interval was shorter on the sire side (3.33 years) compared with the dam side (4.17 years). A shorter generation interval in males than in females has been previously reported by Goyache *et al.* (2003) in Xalda breed of sheep. From a conservation genetics perspective, a long generation interval has an important benefit, 'maintaining genetic diversity'. Briefly, the longer the generation interval, the lower will be the annual loss of genetic diversity. Therefore, prolongation of the generation interval can be used to preserve genetic diversity. In captive-small populations, to avoid inbreeding it is therefore advantageous if the generation interval be as long as possible. Breeding animals should be used in breeding for as long as possible and new breeding animals should be chosen from the last litter of the old parents. However, this strategy would not be practical

in domestic animal populations because generation interval is inversely proportional to the genetic gain. For this reason, application of this strategy is limited to captive population of wild species. In populations under selection to achieve response to selection a reasonable choice of generation interval must be achieved. It would seem that a generation interval of around 3 years can be considered as a goal to be achieved in sheep breeding aimed at having both a most favorable selection response and an acceptable loss in genetic diversity. It is easy to achieve by planning a mating program in which ewes are on average 2 years old at first lambing and lamb five times at intervals of 1 year, and sires start breeding at 2 years of age and are used consecutively for the next 2 years.

Inbreeding has been shown to adversely affect many mammalian populations. In sheep, for example, Lamberson & Thomas (1984) with a literature survey reported that in sheep most important economical traits are affected adversely by inbreeding depression. Hence, to maintain a selected group of animals with constant or increasing production and fitness, inbreeding must be controlled. Minimization of the average coancestry has been theoretically proven to be the most efficient method to preserve genetic diversity when only pedigree information is available (Fernández & Caballero 2001). Breeding the animals with lowest coancestory will necessarily maximise genetic diversity in the next generation, as it ensures that the founder alleles with lowest frequency are preferentially propagated. This method would be used in the population to decline the rate of inbreeding by breeding animals those with the lowest mean coancestry with the others.

In conclusion, the estimates of f_e, f_a, f_g and f_{ne} indicated unequal contribution of founders, presence of bottleneck and genetic drift in formation of this population of Afshari sheep. In addition, estimated value of genetic diversity index highlights that approximately 2.2 % of the total heterozygosity has disappeared from the population. While some ancestors with lower breeding values had greater contribution to the population than ancestors with greater breeding values, breeding values of the most used ancestors are positive, genetic variability reduction is low, and N_e is not very low for a small population under selection, indicating that even in populations which undergo selection, besides achieving genetic gain, the rate of inbreeding is controllable if matings carefully planned.

Acknowledgements

Thanks to Dr. J. P. Gutiérrez of Madrid University for assistance in using the ENDOG program, to three anonymous reviewers for valuable comments that helped to improve this manuscript considerably and to people who collected data over the experiment, in particular H. S. Mohammadi.

References

Bijma P (2000) Long-term genetic contributions. Prediction of rates of inbreeding and genetic gain in selected populations. Ph.D. Thesis, Wageningen University, The Netherlands

Boichard D, Maignel L, Verrier E (1997) The value of using probabilities of gene origin to measure genetic variability in a population. Gen Sel Evol 29, 5-23

Bozzi R, Franci O, Forabosco F, Pugliese C, Crovetti A, Francesco F (2005) Genetic variability in three Italian beef cattle breeds derived from pedigree information. Ital J Anim Sci 5, 129-137

Caballero A, Toro MA (2000) Interrelations between effective population size and other pedigree tools for the management of conserved populations. Gen Res 75, 331-343

Cervantes I, Goyache F, Molina A, Valera M, Gutiérrez JP (2011) Estimation of effective population size from the rate of coancestry in pedigreed populations. J Anim Breed Genet 126, 56-63

Criscione CD, Blouin MS (2005) Effective sizes of macroparasite populations: a conceptual model. Trend Parasit 21, 212-217

Falconer DS, MacKay TFC (1996) Introduction to quantitative genetics. 4th ed. Longman Group press, Harlow, UK

Fernández J, Caballero A (2001) A comparison of management strategies for conservation with regard to population fitness. Conservation Genetics 2,121-131

Frankham R, Ballou JD, Briscoe DA (2002) Introduction to Conservation Genetics. Cambridge University Press, Cambridge, UK

Ghafouri-Kesbi F, Eskandarinasab MP, Hassanabadi A (2009) Short-term selection for yearling weight in a small-experimental Iranian Afshari sheep flock. Can J Anim Sci 89, 301-307

Ghafouri-Kesbi F (2010a) Analysis of genetic diversity in a close population of Zandi sheep using genealogical information. J Genet 89, 479-483

Ghafouri-Kesbi F (2010b) Change in genetic size of small-closed populations: lessons from a domestic mammal population. Genet Mol Biol 33, 657-662

Goyache F, Gutiérrez JP, Fernández I, Gómez E, Álvarez I, Díez J, Royo LJ (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 JP, Goyache F (2005) A note on ENDOG: a computer program for analysing pedigree information. J Anim Breed Genet, 122, 357-360

Gutiérrez JP, Altarriba J, Díaz C, Quintanilla R, Cañón J, Piedrafita J (2003) Pedigree analysis of eight Spanish beef cattle breeds. Gen Sel Evol 35, 43-64

Lacy RC (1989) Analysis of founder representation in pedigrees: founder equivalents and founder genome equivalents. Zoo Biol 8,111-123

Lamberson WR, Thomas DL (1984) Effects of inbreeding in sheep: a review. Animl Breed Abst 52, 287-297

Malécot G (1948) [The Mathematics of Heredity] Masson et Cie press, Paris, France [in French]

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Nat Acad Sci 70, 3321-3323

Vozzi PA, Marcondes CR, Bezerra LAF, Lobo RB (2007) Pedigree analyses in the Breeding Program for Nellore Cattle. Gen Mol Res 29, 482-485

Wright S (1931) Evolution in Mendelian populations. Genetics 16, 97-159

Received 16 December 2011, accepted 30 January 2012.

Corresponding author:

Farhad Ghafouri-Kesbi
email: farhad_ghy@yahoo.com

Islamic Azad University-Karaj Branch, P. O. Box: 31485-313, Karaj, Iran