Full Length Research Paper

Parentage verification and identity test of Ghezel sheep using microsatillate markers

Adel Saberivand[1,2]*, Arash Javanmard[3] and Marryam Safdari[4]

[1]Department of Molecular and Cellular Biotechnology, Institute of Biotechnology, Urmia University, Urmia, Iran.
[2]Department of Clinical Sciences, Faculty of Veterinary Medicine, Urmia University, Urmia, Iran.
[3]Department of Gnomics, North and North-west Agriculture Biotechnolo Center, Tabriz, Iran.
[4]Department of Animal Science, Faculty of Agriculture, Tehran University, Karaj, Iran.

The Ghezel sheep is a fat tail high weight Iranian breed which is raised in the North-west of Iran. To design an efficient improvement program and genetic evaluation system for this indigenous breed, accurate estimates of the population genetic parameters is per-required and all pedigrees and relationships should be correctly recorded. Otherwise, it can produce biased evaluations when pedigrees contain errors and procedures utilize information from relatives. The pedigree and genotype data of Ghezel sheep were examined for errors. Parentage control has been performed by amplification of microsatellites. Mean heterozygosities, mean polymorphism index content (PIC) and mean number of alleles per loci were 0.50, 0.43 and 3.71, respectively. Mendelian errors were found following the pedigree corrections. Alleles at the following seven microsatellite loci were identified: BM4307, CSSM004, BM415, RM029, INRA49, BM3205 and OarFCB5. The pedigree was considered incorrect in 6 (12%) out of all the evaluated progeny, as their genotype did not match their parents. The present findings attest to the usefulness of the investigated microsatellites for parentage control in Ghezel sheep.

Key words: Ghezel sheep, microsatellites, genotyping errors, progeny test.

INTRODUCTION

The best linear unbiased predication (BLUP) procedure has been the most widely used in the prediction of the genetic merit in livestock (Mrode, 1996). Selection decisions based on BLUP are more accurate when using dam and ram pedigree information because BLUP takes account of all available information. Animal model evaluations are based on all known genetic relationships between the animals included in the calculation (Henderson1975). The model assumes that all pedigrees, and relationships are correctly recorded. However, BLUP can produce biased evaluations when pedigrees contain errors and procedures utilize information from relatives. The result would have less accurate genetic evaluations and slower genetic progress than it would be possible otherwise. Van Vleck (1970a,b) demonstrated that incorrect identification of sires, in cattle data, can bias estimates of heritabilities, evaluations of sires and estimates of genetic progress due to selection. Bias increased as the fraction of records with errors increased. There is a need to check the correct paternity relationships to be used for predicting the genetic metric of the individuals in the breeding programme through the numerator relationship matrix. It is well known that error in paternity assignment delay genetic progress whose magnitude under certain circumstance reaches that of the amount of paternity errors in the pedigree. Individual identification and parent-age control are essential for consumer protection and efficient management of animal populations study. Therefore correction of parentage for

---

*Corresponding Author. Email: adel64ir@yahoo.com.

breeding stocks is a prerequisite for an efficient breeding program in all farm animal species including sheep. This expectation is not always fulfilled as several studies showed.

There are three main error sources for identification of progenies in our investigated farm: missing identi-fication plastic tags because of fighting of kids, undesired mating of rams to the neighbouring ewes and finally, replacement of new tags by workers without conside-ration of pervious identification of real ID of animals. Therefore, search for some method for parentage test was necessary.

DNA analysis provides a powerful tool for verifying the parentage and identification of individual animals. Micro-satellites or simple sequence repeats (SSRs) discovered in 1981, are tandem repeated motifs of 1-6 nucleotides found in all prokaryotic and eukaryotic genomes (Goldstein et al., 1999). Dinucleotide repeats dominate, followed by mono- and tetra-nucleotide and trinucleotide repeats are least dominant. Repeats of five (penta-) or six (hexa-) nucleotides can also be found. Generally, among dinucleotides, (CA)n repeats are most frequent, followed by (AT)n, (GA)n and (GC)n, the last type of repeat being rare. Microsatellites combine many desirable properties including co-dominance, high variability, rapid and simple assays and uniform genome coverage (Goldstein et al., 1999). These regions are among the most variable in the nuclear DNA molecules. Microsatellites are characterized by high mutation rates, allelic diversity and high heterozygosity, including high variability even among individuals from the same population. Microsatellite data can be used to describe the genetic structuring within and between populations and among subspecies and species. Mutations at micro-satellite loci typically result in length variations, as the number of repeat units either is increased or decreased (Goldstein et al., 1999).

The objective of the present study is to evaluate seven microsatellites for progeny test in Ghezel sheep.

MATERIALS AND METHODS

Animals

Whole blood sample randomly collected from 50 progeny which belong to pedigree data included 10 rams and 40 ewes (Figure 1). The ewes are routinely bred to rams through controlled natural seasonal mating following heat detection. At the begining of the breeding season, ewes were weighed and randomly assigned to single ram 10 to 20 ewes were. Authorized veterinarian collected blood samples for DNA genotyping from tail vein. Blood was collected on potassium- ethylenediaminetetraacetic acid (K_2EDTA) and stored at -20°C for few weeks or -70°C up to several months.

DNA extraction

DNA was extracted from whole blood according to the previously reported DNA extraction protocol. DNA concentrations were calculated by spectrophotometer by taking the optical density at a wavelength of 260 nm.

Primers and polymerase chain reaction (PCR)

Seven microsatellite markers (BM4307, CSSM04, BM415, RM029, BM3205, INRA049 and OarFCB5) were analyzed for polymorphisms according to Vaiman et al. (1996) (Table 1). The PCR mixture (25 µl) consisted of 1.5 mM $MgCl_2$, 200 µM dNTPs, 0.3 µM primers, 1x PCR buffer, 1 U Taq polymerase (Cinagen Company) and 100 ng of DNA template. The reaction mixture was placed on DNA Thermal cycler (Biometra) and the thermal cycling conditions included touch down Program with the following detail: an initial denaturation step at 94°C for 4 min, 10 cycles of 94°C (60 s), 66°C (decrease)1°C/cycle for 60 s and 72°C (60 s) followed by 25 cycles of 94 (60 s), 56 (120 s) and 72°C (60 s), and a final extension step of 10 min at 72°C. Gels were stained with 0.5 mg/l ethidium bromide for 20 min, destained and photographed.

Proportion of heterozygous genotypes

The equation below was used to calculate the proportion of heterozygous genotypes

$$ph = 1 - \sum\nolimits_{1=1}^{n} p_i^2 \qquad (1)$$

Where: p_i is the frequency of the pth allele among a total of l alleles. The informativeness of a marker can be measured using its heterozygosity or its polymorphism.

Polymorphism information content (PIC)

The probability that one could identify which homologue of a given parent was transmitted to a given offspring is genotyped by the other parent as well. Probability that the parent is heterozygous × probability that the offspring is informative:

$$PIC = 1 - \sum\nolimits_{1=1}^{n} p_i^2 - 2\sum_{i=1}^{l} \sum_{j=1}^{i-1} P_i^2 P_j^2$$

Exclusion power (EPR), the probability of excluding a random individual from the population as a potential parent of an animal based on the genotype of one parent and offspring was calculated from Jamieson (1994, 1997).

Statistical analysis

For estimation of genotype and allele frequencies, we used POPgene software (Yeh et al., 1999). Allelic frequencies and number of alleles per locus were estimated by direct counting from observed genotypes. Heterozygosities, polymorphic information contents and exclusion probabilities were computed using the CERVUS (Ver. 2.0) software (Marshall et al., 1998).

RESULTS AND DICUSSION

The level of information content of a microsatellite is determined, allele number, frequency of frequent allele (FNA), expected Heterozygosities (H_e), observed hetero-zygosities (H_o), Hardy Weinberg test (HWT), F (null), polymorphism informative content (PIC), heterozygosis (Het) and probability of exclusion (PE) values are depen-
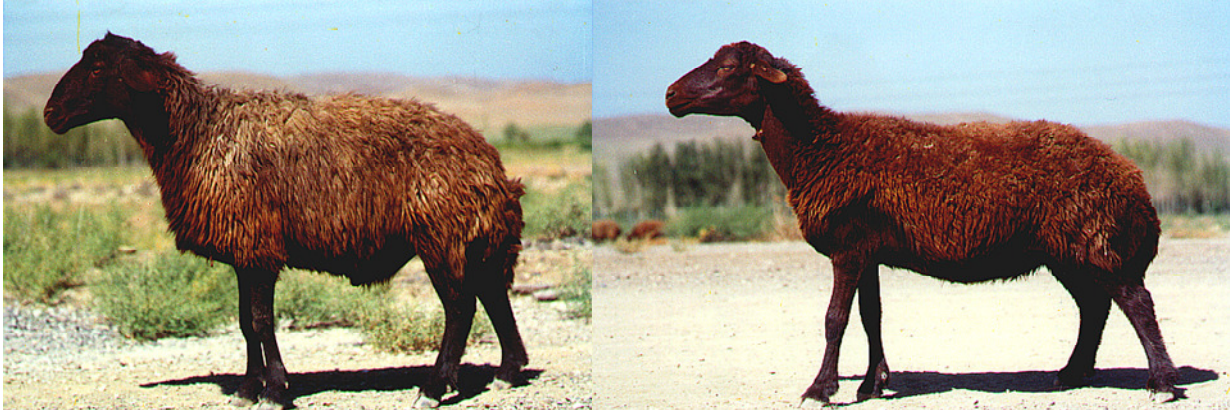
Figure 1. Phenotypic characteristics of fat tail Ghezel sheep.

Table 1. Microsatillate markers in different chromosome used to reveal error in identification (Vaiman et al 1996).

Loci	Primer Sequence	PCR size (bp)	Type of SSR
BM4307	ATAACACAAAAAGTGGAAAACACTC ATTTTATCTCAGGTCCCTTTTTATC	187 - 203	$(CA)_{12}(TA)_6$
BM3205	TCTTGCTTCCTTCCAAATCTC TGCCCTTATTTTAACAGTCTGC	204 - 219	$(TA)_9(TG)_{11}$
CSSM004	ATGCGTCCTAGAAACTTGAGATTG GAAATCATCTGGTCATTATCAGTG	183 - 220	$(GT)_{10}(TA)_5$
INRA049	TGTATTAGTTTGTGTTCTTTGGC TTGGCTTCCACAATCACACA	56 - 172	$(TA)_{10}(TG)_{11}$
OarFCB5	AAGTTAATTTTCTGGCTGGAAAACCC GACCTGACCCTTACTCTCTTCACTC	79 - 130	$(GT)_3A(TG)_6TA(TG)_4$
RM029	ATATGTCTCTGCATATCTGTTTAT CTAATCCCATAGTGAGCAGACC	80 - 102	$(CA)_{12}$
BM415	CAAATTGACTTATCCTTGGCTG TGTAACATCTGGGCTGCATC	137 - 157	$(TG)_{18}$

dent on the number of alleles and on the frequency distribution of these alleles on the population which is shown in Table 2. The results of the microsatellite markers potential use in paternity tests and the control of individual identification on the studied population are shown in Table 2.

Pedigree errors were shown to occur in 12.0% of cases tested in this study. This is within the ranges of pedigree errors revealed in other animal studies (8.7 to 15.5% in sheep: Barnett et al., 1999; 2 to 22% in cattle: Visscher et al., 2002) using microsatellite markers. The 7 micro-satellites evaluated in this study provide sufficient power (75.3%) to be useful in a parentage determination kit to confirm parentage of offspring from that going to be selected as future breeding animals. The possibility of multiple paternities from sheep may require a greater power of exclusion, which would be provided by using more loci.

The average exclusion probability is a measure of effi-

ciency in paternity testing; it refers to the prior ability of tests to detect paternity inconsistencies. This parameter measures the capacity of the system to detect a false accusation of paternity. Traditionally, this average exclusion probability has been estimated as the probability of excluding a male who is not the father by an inconsistency in at least one of the studied loci.

In a classical analysis of genetic relationship, one of the useful parameters is the power of exclusion (PE), the power of a genetic marker in excluding a non-related individual chosen by chance in a specific population, as an alleged father in a paternity investigation. The paternity PE is the expected average probability that a polymorphic locus shows the exclusion of a man without kinship with the biological father. This index depends on the informative content of a locus, which depends on its number of alleles and its respective frequencies. From the probabilities of exclusion of several loci, it is possible to calculate the combined PE (PEC), by simple multiplication

Table 2. Allele number, frequency of frequent alleles, expected heterozygosities, observed heterozygosities, Hardy Weinberg test, F (null), PE1, heterozygosis and probability of exclusion of 7 microsatellite markers in Ghezel breed.

Loci	Allele number	FNA[a]	H(e)	H(o)	PIC	HWT	F (Null)	PE1
BM4307	2	0.66	0.44	0.33	0.34	ns	+0.142	0.90
CSSM004	5	0.65	0.55	0.68	0.51	**	-0.154	0.83
BM415	3	0.45	0.65	0.68	0.57	ns	-0.033	0.79
RM029	7	0.58	0.63	0.54	0.59	ns	+0.082	0.76
BM3205	3	0.73	0.42	0.22	0.37	ns	+0.279	0.91
INRA049	4	0.79	0.35	0.35	0.31	ns	-0.072	0.93
OarfCB5	2	0.68	0.44	0.47	0.34	ns	-0.045	0.90

He, Expected heterozygosities; Ho, observed heterozygosities; HWT, Hardy Weinberg test;; Het, heterozygosis; PE, probability of exclusion; Ns: non significant; **, significant level at p < 0.01; [a], frequency of most frequent allele (FNA).

of the values for each locus. The value of PEC is a function of the examined loci number, as well as of the informative content of each locus. The knowledge of the PE and the PEC can define the loci to be used in an analysis of genetic relationship. Genetic paternity testing can provide sire identity data for offspring when females have been exposed to multiple males. However, correct paternity assignment can be influenced by factors determined in the laboratory and by size and genetic composition of breeding groups. When parents are assigned on the basis of LOD scores, the most likely candidate parent is the candidate parent with highest LOD score. The likelihood ratio is the likelihood that the candidate parent is the true parent divided by the likelihood that the candidate parent is not the true parent. Negative LOD score shows that the candidate parent is less likely to be the true parent than an arbitrary randomly chosen individual. A LOD score of zero candidate parents is equally likely to be the true parent as an arbitrary randomly chosen individual. Positive LOD score candidate parent is most likely to be the true parent than arbitrary randomly chosen individual.

The PE is a parameter to solve problems of some genetic markers in a population and is most commonly used as molecular markers in pedigree verification (Luikart et al., 1999). Jakabova et al. (2002) have also shown that at least, five microsatellites with the highest individual PE values that have a 97% total exclusion probability should be used to obtain a high degree of excluding incorrect parentage. Usha et al. (1994) also reported a total PE of 0.88 for two microsatellite loci used in cattle parentage control. Marklund et al. (1994) analyzed eight microsatellite loci in paternity testing to reach a total exclusion probability of 0.96 to 0.99 in different horse breeds.

Comparison of our results with these various results clearly shows that our selected microsatellites have greater power of exclusion given the fact that we could reach a high level of exclusion with only seven loci (PE =

0.97). Analyses of more loci will allow the increase of the combination efficiency.

Simulation based on allele frequencies of parents suggested that six loci would allow assignment of progeny to their correct dam in 92% cases. In reality, however, assignment success was half of this. Discrepancies between the simulations and real data sets were considered to be largely due to the presence of null alleles at loci that were not accounted for the simulations. The validity of this kind of research would need more genotyping of loci and screening for large numbers of offspring is required.

ACKNOWLEDGEMENT

REFERENCES

Barnett N, Purvis I, Van Hest B, Franklin I (1999). The accuracy of current dam pedigree recording strategies employed by stud Merino breeders. In: Proceedings of the Association for the Advancement of Animal Breeding and Genetics. Mandurah, Western Australia, 13: 373-376.
Goldstein DB, Linares AR, Cavallil-Sforza LL, Feldman MV (1995). Genetic absolute dating based on microsatellites and the origin of mordern humans. Proc. Natl. Acad. Sci. USA. 92: 11549-11552.
Henderson C (1975). Best linear unbiased estimation and prediction under a selection model. Biometrics, 31: 423-447.
Jakabova D, Trandzik J, Chrastina J (2002). Effectiveness of six highly polymorphic microsatellite markers in resolving paternity cases in Thoroughbred horses in Slovakia. Czech. J. Anim. Sci. 47: 497-501.
Jamieson A (1994). The effectiveness of using co-dominant polymorphic allelic series for (1) checking pedigrees and (2) distinguishing full-sib pair members. In: Festschrift in Honour of Dr Clyde J. Stormont. Anim. Genet. 25(1): 37-44.
Jamieson A, Taylor St CS (1997). Comparison of three probability formulae for parentage exclusion. Anim. Genet. 28: 397-400.
Luikart G, Biju-Duval MP, Ertugrul Y, Zagdsuren C, Maudet C, Taberlet P (1999). Power of 22 microsatellite markers in fluorescent multi-

plexes for parentage testing in goats (*Capra hircus*). Anim. Genet. 30: 431-438.

Marklund S, Ellegren H, Eriksson S, Sandberg K, Andersson L (1994). Parentage testing and linkage analysis in the horse using a set of highly polymorphic microsatellites. Anim. Genet. 25: 19-23.

Marshall TC, Slate J, Kruuk L, Pemberton JM (1998). Statistical confidence for likelihood-based paternity inference in natural populations. Mol. Ecol. 7: 639-655.

Mrode RA (1996) Linear Models for the Prediction of Animal Breeding Values. Biddles, Guildford, p. 184.

Usha AP, Simpson SP, Williams JL (1994). Evaluation of microsatellite markers for parentage verification. In: Proceedings of the 24th ISAG Conference. Anim. Genet. 25: p. 41.

Vaiman D, Schibler L, Bourgeois F, Oustry A, Amigues Y, Cribiu E (1996). A genetic linkage map of the male goat genome. Genetics, 144: 279-305.

Van Vleck LD (1970a). Misidentification in estimating the paternal sib correlation. J. Dairy Sci. p. 31469.

Van Vleck LD (1970b). Misidentification and sire evaluation. J. Dairy Sci. pp. 531-697.

Visscher PM, Woolliams JA, Smith D, Williams JL, (2002). Estimation of pedigree errors in the UK dairy population using microsatellite markers and the impact on selection. J. Dairy Sci. 85: 2368-2375.

Yeh, Francis C, Rong-cai Y, Tim B (1999). POPGENE Ver. 3.31. Microsoft Window-based freeware for population genetic analysis. p. 208.