Full Length Research Paper

Development of polymorphic microsatellite loci for Iranian river buffalo (*Bubalus bubalis*)

Borhan Shokrollahi[1], Cyrus Amirinia[2*], Navid Dinparast Djadid[3], Noor Amirmozaffari[4] and Mohammad Ali Kamali[2]**

[1]Department of Animal Science, Science and Research Branch, Islamic Azad University, Tehran, Iran.
[2]Department of Biotechnology, Animal Science Research Institute of Iran, Karaj, Iran.
[3]Malaria and Vectors Research Group, Biotechnology Research Center, Pasteur Institute of Iran, Tehran, Iran.
[4]Department of Microbiology, Medical University of Iran, Tehran, Iran.

Microsatellite loci were developed using PCR-based isolation of microsatellite arrays (PIMA) for Iranian river buffalo. Blood samples of eighty unrelated individuals from four buffalo populations (Khuzestan, Mazandaran, Guilan and Azarbayejan) were taken and following DNA extraction, isolation of microsatellite loci initiated using enrichment with random amplified polymorphic DNA (RAPD) primers. RAPD-PCR fragments were ligated into PTZ57R TA cloning vector and transformed into DH5α competent cells. Obtained colonies were screened for presence of repetitive elements by repeat-specific and M13 forward and reverse primers. After designing primer pairs for repeat containing fragments, they were tested in all buffalo populations. Two microsatellite loci (RBBSI and RBBSII) were informative and polymorphic. Number of alleles for RBBSI and RBBSII in 80 individuals was 5 and 6, respectively. Expected heterozygosity ranged from 0.65 to 0.81. Significant deviation from Hardy-Weinberg equilibrium expectation occurred for both loci in all populations, but 37.5% of locus/population combination showed the deviation. We postulate that the two newly isolated microsatellite loci during this study could be useful for population genetic studies in *Bubalus bubalis.*

Key words: Microsatellite loci, Iranian river buffalo, *Bubalus bubalis*, PIMA.

INTRODUCTION

Microsatellites also called simple sequence repeats (SSR) are tandemly repeat motifs with 1 - 6 bases length distributed non-randomly in all eukaryotic genome both coding and non-coding (with a high probability) regions (Zane et al., 2002; Li et al., 2002). Their polymorphisms can be affected by location and presence of other linked genes (Slatkin, 1995; Thuillet et al., 2004). These markers can be used in characterization of species popu-lations, genetic diversity (Esmaeilkhanian and Banabazi, 2006) and population studies (Amirinia et al., 2007), as they are hypervariable and widely dispersed through genome. Moreover, they have application in identification of individuals and parentage testing (Seyedabadi et al.,

2006).

Iranian river buffalo is divided to two ecotypes. First ecotype with 400000 heads spreads in North and North West provinces (Mazandaran, Guilan and Azarbayejan). Second ecotype with 100000 heads spreads in south of the country (Khuzestan province), known as Khuze-stanian ecotype. Since the beginning of 20th century, population of Iranian river buffalo has reduced by about 70% (Tavakolian, 1998; Kianzad, 2000). A decrease in the number of buffaloes is associated with three factors: holsteinization, that is, the substitution of low production cows and buffaloes with high production Holstein Friesian cows; mechanization, that is, the substitution of draught animals with tractors and the poor market demand for buffalo products (Borghese and Mazzi, 2005). Iranian buffalo are adapted to special condition of Iran, therefore they are worthwhile stock and a part of genetic resources of Iran. Buffalo play an important role in rural economy in the mentioned provinces.

**Corresponding author. E-mail: ndinparastdjadid@yahoo.com, navid@pasteur.ac.ir. Tel: +98 21 66480780. Fax: +98 21 66465132. *E-mail: amirinya@gmail.com. Tel: +98 261 4439214. Fax: +98 261 4413258

Table 1. List of the RAPD primers, their nucleotide sequence, annealing temperatures and references.

Primer	Sequence (5' to 3')	Annealing temperature (°C)	Reference
Primer 1	GGGACGTCTC	37.5	Joshi et al., (1998)
Primer 2	TTAGCGCCCC	37.5	Joshi et al., (1998)
Primer 3	ATGCCCCTGT	36	Salah et al., (2007)
Primer 4	AAAGCTGCGG	41	Salah et al., (2007)
Primer 5	ACCGCCGAAG	37	Salah et al., (2007)

Researchers applied cattle microsatellite markers for defining the genome make up in buffalo because no systematic studies have been undertaken to develop polymorphic DNA markers specific to this species. Cattle microsatellite markers have many disadvantages like low polymorphism and loss of amplification. For example, Supajit et al. (2008) suggested that cattle microsatellite markers may not be optimize for genetic studies in *Bubalus bubalis*. Navani et al. (2001) reported that 56% of cattle microsatellite markers provided polymorphic band patterns when tested in 25 buffaloes. Mirhoseinie et al. (2005) surveyed the efficiency of 6 cattle microsatellite loci in Iranian endogenous cattle and buffalo populations and suggested that four loci are polymorphic in buffalo populations.

On the other hand, there were few studies on genetic structure and population variation, with no published data on isolation and characterisation of microsatellite markers in Iranian buffalo. Therefore, it seems necessary to isolate and develop these kinds of genetic markers in *B. bubalis*. Development of particular microsatellite markers for Iranian river buffalo could be helpful for the study of genetic diversity within and among its populations and defining the functional composition of their genome. Microsatellites have been traditionally isolated from partial genomic libraries selected for small insert size by screening a large number of clones through colony hybridization with repeat containing probes (Rassmann et al., 1991). Generally, this method is quite tedious and inefficient. Therefore, in order to reduce the time invested in microsatellite isolation and to significantly increase yield, several alternative strategies have been attempted. One strategy is RAPD-based methods. Studies showed that RAPD fragments contain microsatellite repeats more frequently than random genomic clones (Cifarelli et al., 1995; D'Amato et al., 1999). PIMA (PCR-based isolation of microsatellite array) has been proposed by Lunt et al. (1999) and is more efficient (the number of positive colonies) and accurate (the number of positive colonies containing a microsatellite) than traditional methods. Not only it alleviates hazardous waste from hybridization using radioactive elements, but also it increases the rate of producing sufficient flanking regions.

In this study we tried to develop polymorphic markers for Iranian river buffalo, which can be applied to various population genetic researches.

MATERIAL AND METHODS

Sampling

Eighty unrelated Iranian river buffalo was selected from Khuzestan, Mazandaran, Guilan and Azarbayejan provinces.

DNA extraction and PCR amplification

Blood samples were taken from the jugular vein and DNA was extracted by Salting out procedure (Miller et al., 1988). PCR was optimized with RAPD primers. Five RAPD primers were used to amplify fragments from the genome in separate reactions (Table 1). PCR amplifications were performed in a 25 µl reaction volume containing 1x PCR buffer, 200 µM of each dNTP, 1.5 mM MgCl$_2$, 20 - 100 ng of the template DNA, 10 pM of each primer and 1 unit of Taq DNA polymerase (Metabion co, Germany). Amplifications were performed in a thermal cycler (Eppendorf gradient Master cycler) under the following conditions: 5 min at 94°C; 35 cycles of 30 s at 94°C, 50 s at specific optimized annealing temperature for each primer (Table 1), 1 min at 72°C and a final extension of 10 min at 72°C.

Cloning of PCR products

PCR product (RAPD fragments) was purified using clean up kit (Fermentas) and ligated into the PTZ57R TA cloning vector (Fermentas) with T4 DNA ligase at 4°C overnight. Heat shock transformation with *Escherichia coli* DH5α competent cells was conducted at 42°C for 90 s and the competent cells were plated on LB-agar plates, which contained IPTG, X-Gal and ampicillin to allow blue and white colony selection. The colonies were identified by colony PCR with M13 forward and reverse primers that may technically show greater than 181 bp bands on agarose gel electrophoresis if the inserted fragments exist in the colony, then screening of microsatellite containing clones (positive clones) was conducted by colony PCR using M13 forward and reverse primers plus a repeat-specific primer ((CA$_8$) A and (AGC$_8$) A). The A in the end of repeat specific primers allow the annealing and extension of the array with its 3' A in the flanking region and not within internal sites in a repeat array. This ensures a single product for each colony that contained a repeat element (Lunt, 2005).

Sequencing and data analysis

The positive colonies were used to extract plasmid DNA for sequencing. Sequencing was performed using M13 vector primers. After sequencing of plasmids, sequences were verified by alignment using the BioEdit software (version 7.0.9.0) to remove any redundant plasmids. The sequences were then applied to microsatellite repeats finder program (www.biophp.org/minitools/
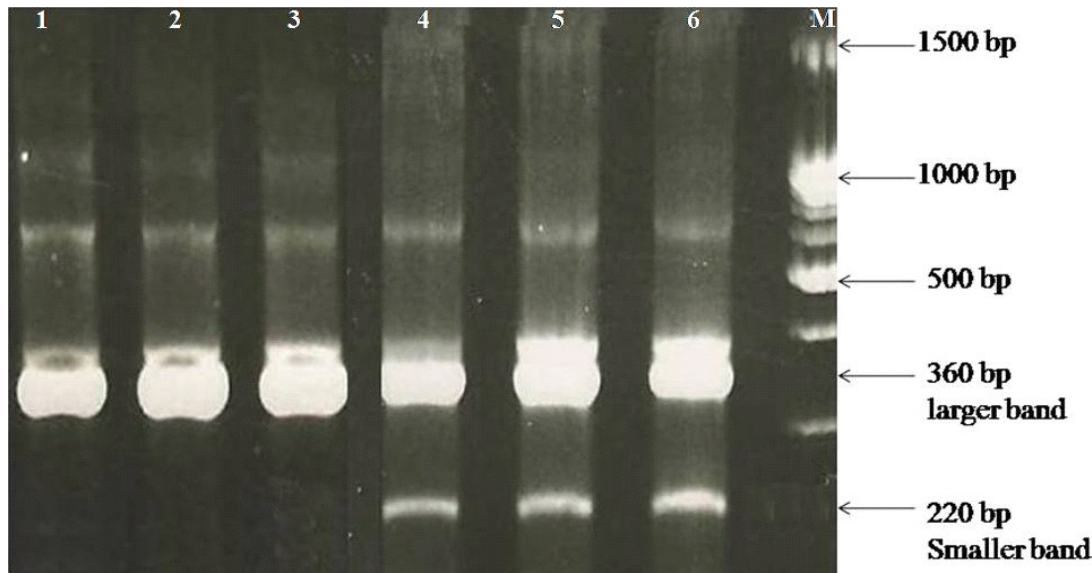
Figure 1. Selection of microsatellite containing clones by PCR amplification. Smaller bands indicate possible microsatellite regions, whereas larger bands indicate the whole cloned fragment. In all lanes (1 - 6) PCR amplification (360 bp) occurred using M13 forward and reverse primers, and in some lanes (4 - 6) PCR amplification (360 bp and 220 bp) occurred using M13 forward and reverse primers and a repeat specific primer. The products of amplification and molecular size marker (M) were separated in a 1% agarose gel containing ethidium bromide.

Microsatellite_repeats_finder/demo), which is a Web-based program for searching repetitive sequences. The repeats then analysed by BLAST program, followed by designing primers for the flanking regions of repetitive sequences using Oligo 6 software. Two primer sets were designed for an annealing temperature of 54°C. Each primer pair was tested on genomic DNA using optimized condition (94°C for 30 s, 54°C for 50 s and 72°C for 55 s) for Iranian river buffalo. 8 µl PCR product was diluted with an equal volume of loading dye solution (0.05% xylene cyanol FF, 0.05% bromophenol blue and 95% of formamide) and then were resolved on 8% denaturing acrylamide gels containing 7 M urea. The gels were run at a constant 80 V at overnight. After electrophoresis, the gels were visualized with silver nitrate staining and microsatellite length and the allele frequencies were estimated using GELPRO analyser software (version 3.1). The POPGENE computer package (Yeh et al., 1999) was employed to calculate allele frequencies, Hardy–Weinberg equilibrium test and construction of dendrogram. Observed heterozygosity (H$_O$) was calculated as the proportion of total heterozygous animals and expected heterozygosity (H$_E$) was estimated according to the following equation (Hedrick, 1999):

$$H_E = 1 - \sum_{i=1}^{n} P_i^2$$

Where Pi stands for the ith allele frequency.

Expected heterozygosity and observed heterozygosity in each population were estimated to determine the genetic variation within and between populations. Genetic distance was calculated according to Nei (1978) standard method.

RESULT

Colonies of RAPD fragments were analysed to verify existence of repetitive sequences. The result of a representative colony-PCR is shown in Figure 1; the size of larger and smaller amplified product fragments were 360 and 220 bp, respectively. The latter product is the result of amplification with a repeat specific primer and M13 forward or reverses primer.

Sequencing of verified colonies with M13 primers revealed that approximately 40% of clones contained repetitive sequences. 12 clones were selected that contained repetitive sequences of at least 7 repeats in length. Finally 2 clones were submitted to GenBank with accession numbers as shown in Table 2. It needs to be noted that duplicate clones and those which matched with known microsatellite sequences in GenBank were not analysed further.

Two primer sets designed for flanking the newly identified microsatellite loci were tested for their variability in 80 individuals from four populations. Expected size of alleles, observed hetrozigosity (H$_O$) and expected hetrozygosity (H$_E$) are shown in Table 2. Numbers of alleles of RBBSI locus in Khuzestan and Guilan populations were five whereas in Mazandaran and Azarbayejan populations were four and number of alleles of RBBSII in Khuzestan and Mazandaran populations were five whereas those of Guilan and Azarbayejan populations were six. The estimated mean expected hetrozygosity per locus in all the populations was 0.697 and 0.788 (Table 3). The level of within populations hetrozygosity were low in Guilan (H$_E$ = 0.69) and high in Khuzestan (H$_E$ = 0.77). The chi-square test for Hardy-Weinberg equilibrium is shown in Table 4. Both loci in all population showed significant

Table 2. Two microsatellite loci from four Iranian river buffalo populations.

Locus	Primer sequences (5' to 3')	Repeat motif	Size (bp)	No. of alleles	H_O	H_E	GenBank Accession no.
RBBSI	Forward: TGTCTATAGTCCATGGAGTC Reverse: GGCTAATACAGACCTTCTG	(CA) 7-21	290 - 318	5	0.725	0.715	FJ744166
RBBSII	Forward: CTCCAGTACTTTTGCCTAG Reverse: CCTCTCTGATTTGGGTAAG	(AGC) 7-19	297 - 334	6	0.712	0.793	FJ744167

H_O: Observed hetrozygosity, H_E: expected hetrozygosity.

Table 3. Heterozygosity and number of alleles of two microsatellite loci in Iranian river buffalo populations.

Locus	Khuzestan			Mazandaran			Guilan			Azarbayejan			Mean	
	N_A	N_E	H_E	N_A	N_E	H_E	N_A	N_E	H_E	N_A	N_E	H_E	N_E	H_E
RBBSI	5	3.72	0.75	4	3.29	0.714	5	2.7	0.646	4	2.97	0.68	3.17	0.697
RBBSII	5	4.51	0.798	5	4.44	0.794	6	3.75	0.752	6	4.84	0.81	4.38	0.788

N_A: Number of alleles, N_E: number of effective alleles, H_E: expected hetrozygosity.

Table 4. Probability test for Hardy-Weinberg Equilibrium at two loci for four populations of Iranian river buffalo.

Locus	HW for populations				HW for all	
	Khuzestan	Mazandaran	Guilan	Azarbayejan	χ^2	P
RBBSI	0.004	0.062	0.86	0.220	34.6	0.000
RBBSII	0.066	0.008	0.40	0.003	56.3	0.000

deviation from Hardy-Weinberg but approximately 37.5% of locus/population combinations showed significant departure from Hardy-Weinberg at $p < 0.05$.

The genetic distances among populations were shown in Table 5. The highest distance was observed between Khuzestan and Guilan populations and the lowest distance was between Guilan and Mazandaran populations. Cluster analysis was performed according to UPGMA. Result showed that Guilan and Mazandaran populations were more related to each other and Khuzestan and Azarbayejan populations had the highest genetic distance among Iranian river buffalo populations (Figure 2).

DISCUSSION

In current study we characterized two new polymorphic loci for Iranian river buffalo using PIMA (PCR-based isolation of microsatellite arrays). There are several procedures for isolation of microsatellite loci, among those; PIMA seems to have some advantages over other techniques that make it more suitable for isolation of microsatellite loci in this kind of studies. First, the PIMA is significantly cheaper in a laboratory which is not regularly performing hybridization techniques. Second, it has the ability to isolate both flanking regions simultaneously and third, PIMA avoids handling radioactivity and utilizes the reagents, experience and equipment which need for PCR analysis. According to the results of this study, PIMA seems to have similar efficiency as reported in other species such as fishes (Hsu et al., 2004; Sanches and JR, 2006; Chung-Jian et al., 2007; Lin et al., 2008). In present study we used di-nucleotide $(CA)_8$ and tri-nucleotide $(AGC)_8$ repeat specific primers for isolation of microsatellite loci, because these repeats are highly prevalent in the genome of mammals (Zane et al., 2002).

Both loci (RBBSI and RBBSII) had either 5 or 6 alleles and were polymorphic, other loci were either monomorphic or failed to amplify and numbers of repeats were too low. In general, observed heterozygosity was high at both loci. These data are in agreement with the study of Aminafshar (2008), who concluded that there were high mean percentage observed heterozygosity ($H_O = 0.9$) in 3 population of Iranian buffalo using 15 cattle microsatellite. Takezaki and Nei (1996) suggested that average heterozygosity must be between 0.3 and 0.8 in a population, in order to be a useful marker for measuring genetic variation. Our results for mean heterozygosity were within that range; therefore the identified markers in this study are suitable for measuring genetic variation. Number of alleles in two loci varied from 4 to 6 in all populations.

Table 5. Genetic distance between four Iranian river buffalo populations.

	Khuzestan	Mazandaran	Guilan
Mazandaran	0.1453		
Guilan	0.1863	0.0340	
Azarbayejan	0.1738	0.0666	0.0441

Figure 2. The relationship among the Iranian river buffalo populations shown by dendrogram.

Barker (1994) suggested that microsatellite loci in genetic distance studies should have not less than four alleles because it increases the standard error of distance estimates. Considering number of alleles in all popula-tions, these loci can be suitable candidate for analysis of genetic diversity for *B. bubalis.* Genotype frequencies and exclusion probabilities using allele frequency in populations should be consistent with Hardy–Weinberg expectation (Schnabel et al., 2000). As shown as in Table 4, significant deviations from Hardy-Weinberg were observed for both loci using probably test (p < 0.05); these deviations might be caused by small effective population sizes, the difficulty in collecting enough unrelated individual samples.

The genetic distance between Guilan and Mazandaran populations was the lowest. Both populations are located in the same geographical and climatic conditions and with Azarbayejan population belonging to North ecotype of Iranian river buffalo. These populations were located in one cluster and Khuzestan river buffalo population at another (Figure 2). Results of this study are in agreement with Aminafshar (2008) study who investigated phylo-genetic relationship of the Iranian river buffalo using 15 cattle microsatellite markers and 360 DNA samples in three buffalo populations (Mazandaran, Azarbayejan and Khuzestan) and disagreed with Mirhoseinie et al. (2005) who used 4 cattle polymorphic markers and 90 samples in the same buffalo populations. The latter study showed that Khuzestan and Azarbayejan had the lowest genetic distance and they had high genetic distance with Mazandaran population. It seems that the first study was more accurate and reliable than the latter one, mainly based on the higher number of microsatellite genetic markers and DNA samples used by Aminafshar (2008).

The results of the present study are promising, highlighting the extreme importance of developing of microsatellite loci in Iranian river buffalo. The PIMA methodology can be effectively utilized to isolate micro-satellite loci and being a fairly simple technique can be applied in laboratories with a minimum of available resources, dispensing with the need for radioactive materials and the construction of genomic libraries of the organism of interest. Moreover, the description of these first two polymorphic microsatellite loci in Iranian river buffalo *(B. bubalis)* may contribute effectively to studies on the population genetics and genome mapping of *B. bubalis.*
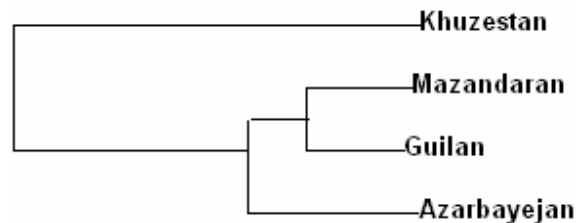
REFERENCES

Aminafshar M (2008). Phylogenetic study of the Iranian buffalo by microsatellite markers. In: Department of animal science, PhD Thesis, p. 200. Tehran: Islamic Azad university-Science and Research branch.

Amirinia C, Seyedabadi H, Banabazi MH, Kamali MA (2007). Bottleneck study and genetic structure of Iranian caspian horse population using microsatellites. Pak. J. biol. Sci. 9: 1540-1543

Barker JSF (1994). A global protocol for determining genetic distances among domestic livestock breeds. In: 5th World Congress on Genetics Applied to Livestock Production, pp. 501-508.

Borghese A, Mazzi M (2005). Buffalo population and strategies in the world. pp. 1-39. In Rome: FAO

Chung-Jian LJP, Wang JP, Lin HD, Chiang TY (2007). Isolation and characterization of microsatellite loci in *Hemibarbus labeo* (Cyprinidae) using PCR-based isolation of microsatellite arrays (PIMA). Mol. Ecol. 7: 788-790

Cifarelli RA, Gallitelli M, Cellini F (1995). Random amplified hybridization microsatellites (RAHM): isolation of a new class of microsatellite-containing DNA clones. Nucleic Acids Res. 23: 3802-3803.

D'Amato ME, Lunt DH, Carvalho GR (1999). Microsatellite markers for the hake *Macruronus magellanicus* amplify other gadoid fish. Mol. Ecol. 8: 1086-1087.

Esmaeilkhanian S, Banabazi MH (2006). Genetic Variation within and between five Iranian sheep populations using microsatellite markers. Pak. J. biol. Sci. 13: 2488-2492

Hedrick PW (1999). Genetics of populations (2nd Ed). Jones and Bartlett Publishers. Sudbury, MA, USA.

Hsu KC, Wang JP, Chen XL, Chiang TY (2004). Isolation and characterization of microsatellite loci in *Acrossocheilus paradoxus* (Cyprinidae) using PCR-based isolation of microsatellite arrays (PIMA). Conserv. Genet. 5: 13-115

Joshi CG, Rank DN, Brahmakshtr BP, Patel AV, Vataliya PH, Muraleedharan P, Khoda VK, Solanki JV (1998). RAPD analysis by PCR using arbitrary primers in different animal species. Indian. Vet. J. 75: 1029-1031.

Kianzad D (2000). A case study on Buffalo recording and breeding in Iran. In: Animal Recording for Improved Breeding and Management Strategies for Buffaloes, pp. 37-44. Slovenia, Bled

Li YC, Korol AB, Fahima T, Beiles A, Nevo E (2002). Microsatellites: Genomic distribution, putative functions and mutational mechanisms: A review. Mol. Ecol. 11: 2453-2465.

Lin HD, Lee TW, Lin FJ, Lin CJ, Chiang TY (2008). Isolation and characterization of microsatellite loci in the endangered freshwater fish *Pararasbora moltrechti* (Cyprinidae) using PCR-based isolation of Microsatellite arrays (PIMA). Conserv. Genet. 9: 945-947

Lunt DH (2005). PCR Isolation of Microsatellite Arrays (PIMA), Molecular Ecology Laboratory, Department of Biological Sciences, University of Hull, HULL, HU6 7RX, UK.

Lunt DH, Hutcchinson WF, Carvalho GR (1999). An efficient method for PCR-based isolation of microsatellite arrays (PIMA). Mol. Ecol. 8: 891-894.

Miller SA, Dykes DD, Polesky HF (1988). A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 16: p. 1215.

Mirhoseinie SZ, Vahidie SMF, Gharehyazie B (2005). Survey of efficiency of six microsatellite loci in Iranian indigenous cattle and buffalo populations. Iran. J. Biotechnol. 3: 41-47.

Navani N, Jain PK, Gupta S, Sisodia BS, Kumar S (2001). A set of cattle microsatellite DNA markers for genome analysis of riverine buffalo (*Bubalus bubalis*). Anim. Genet. 33: 149-154.

Nei M (1978). Estimation of average heterozygosity and genetic distance from a small number of individvals. Genetics, 89: 583-590.

Rassmann K, Schlötterer C, Tautz D (1991). Isolation of simple sequence loci for use in polymerase chain reaction-based DNA fingerprinting. Electrophoresis, 12: 113-111.

Salah M, Abdel-Rahman, Elsayed EH (2007). Genetic Similarity Among the Three Egyptian Water Buffalo Flocks Using RAPD-PCR and PCR-RFLP Techniques. Res. J. Agric. Biol Sci. 6: 1045-1046.

Sanches A, JR PMG (2006). Microsatellites loci isolated in the freshwater fish Brycon hilarii. Mol. Ecol. Notes. 6: 1045-1046.

Schnabel RD, Ward TJ, Derr JN (2000). Validation of 15 microsatellites for parentage testing in North American bison, Bison bison, and domestic cattle. Anim. Genet. 31: 360-366.

Seyedabadi H, Amirinia C, Banabazi MH, Emrani H (2006). Parentage verification of Iranian Caspian horse using microsatellites markers. Iran. J. Biotechnol. 4: 260-264.

Slatkin M (1995). Hitchhiking and associative overdominance at a microsatellite locus. Mol. Biol. Evol. 12: 473-480.

Supajit S, Benchamart M, Ancharlie NA, Sakol PRSD, Kanokporn T (2008). Use of cattle microsatellite markers to assess genetic diversity of Thai Swamp buffalo (*Bubalus bubalis*). Asian-Aust. J. Anim. Sci. 21: 177-180.

Takezaki N, Nei M (1996). Genetic Distances and Reconstruction of Phylogenetic Trees from Microsatellite DNA. Genetics, 144: 389-399

Tavakolian J (1998). Overview on animal and poultry genetic resources of Iran, Animal science Research institute, Karaj, Iran

Thuillet AC, Bataillon T, Sourdille P, David JL (2004). Factors affecting polymorphism at microsatellite loci in bread wheat [*Triticum aestivum* (L.) Thell]: effects of mutation processes and physical distance from the centromere. Theor. Appl. Genet. 18: 368-377

Yeh FC, Boyle T, Rongcai Y (1999). POPGENE version 1.31. Microsoft window-based Freeware for population genetic analysis.

Zane L, Bargelloni L, Patarnello T (2002). Strategies for microsatellite isolation: a review. Mol. Ecol. 11: 1-16.