Indian Journal of Biotechnology
Vol 1, January 2002, pp. 96-100

Development of Genetic Markers Specific for Highly Endangered Species *Rhinoceros unicornis*: Strategies and Implications in Conservation

Vipra Kapur, Anu Bashamboo and Sher Ali*

Molecular Genetics Laboratory, National Institute of Immunology, Aruna Asaf Ali Marg, New Delhi 110 067, India

Indian rhino, *Rhinoceros unicornis*, is a highly endangered species and unless collective efforts are made for its protection, the future of this species will continue to remain bleak. Application of restriction survey and minisatellite associated sequence amplification (MASA) for identification, cloning and sequencing of the potential genomic fragments from this species, led to the development of species specific DNA marker. These markers are envisaged to be useful for ascertaining the origin of rhino's biological samples in the event of poaching and illegal trafficking. This approach may be used for conservation and management of other endangered species.

Keywords: endangered species, *Rhinoceros unicornis*, confined gene pool, biodiversity, conservation genetics, DNA probes

Introduction

Globally, there are five surviving species of rhinos- the Javan (*Rhinoceros sondaicus*), Sumatran (*Dicermocerus sumatrensis*), Indian (*Rhinoceros unicornis*), and South African black (*Diceros bicornis*) and white (*Ceratotherium simum*). The one-horned Indian rhino, *R. unicornis*, was once distributed along the plains of Indus, Ganges and Brahmaputra river systems from the foothills of Hindukush range of Himalayas in the West to Myanmar in the East. Rampant poaching, degradation of habitat due to encroachment, lack of management, inbreeding within a small population, are some of the causes of extirpation of this species. Furthermore, the absence of forest corridors between the isolated populations in some parts of the world (such as between Jaldapara Wildlife Sanctuary in West Bengal and Kaziranga National Park in Assam, India) hampers free-flow of the genetic material This leads to stagnation of the confined gene pool. The world population of wild *R. unicornis* is estimated to be about 2000, of which 1465 are known to be present in India. Literature search on rhino shows conspicuous absence of data on its genome organization and thus very few reports are available on this species.

Earlier studies on mtDNA between South African white rhino and Indian rhino (Xu & Arnason, 1997), have shown divergence time of about 27 million years before present which is almost close to other report of about 26 million years before present based on allozyme studies (Merenlender *et al*, 1989). Cross-hybridising micro satellite markers from *D. bicornis* (Cunningham *et al*, 1999) are inadequate in the context of population dynamics or development of genetic markers. A genetically robust population with a sound reproductive health requires appropriate sex ratios, healthy gene pool, secure environment, adequate and diversified food supply and complete protection of the animals from the perils of poaching (Ali & Hasnain, 1999). Prolonged inbreeding within a small population quells genetic variability leading to loss of hybrid vigour contributing towards extinction of the species (Ali *et al*, 1998). Genetic analysis of endangered species is envisaged to be useful not only for understanding the biology of these animals but also for augmenting the possible conservation efforts (Ali *et al*, 1998; Ali & Hasnain, 1999). Varying causative factors including microbial and viral diseases for a given population of rhinos need to be addressed and resolved independently. Similarly, a curb on menace of illegal trafficking or indiscriminate killing requires not only protection of the animals in their habitat but also proof for ascertaining the origin of biological samples in the event of poaching. Higher eukaryotic genomes seem to show characteristic enrichment of different kinds of DNA repeats (Cox & Mirkin, 1997; Patrushev, 1997). Similar information on the organisation of satellite DNA and VNTR loci

*Author for correspondence:
Tel: 6103799, Fax: 011-616 2125; 617 7626
E-mail: sheralib5@hotmail.com;sherali@nii.res.in

in species and sub-species of rhinoceroses would be of relevance for the development of desired genetic markers (Ali *et al*, 1986; John *et al*, 1996; Raina *et al*, 1996; Mattapallil & Ali, 1997, 1999; John & Ali, 1997; Ray *et al*, 1999).

The present report deals with the development of DNA based genetic marker(s) specific to *R. unicornis* genome. The study involves identification of a genomic fragment employing restriction survey followed by its cloning and sequencing (Chattopadhyay *et al*, 2001). In addition, minisatellite associated sequence amplification (MASA) (Ali *et al*, 1999; Bashamboo & Ali, 1999, 2001) was conducted for identification of species specific DNA fragment from *R. unicornis* genome. These genomic fragments cloned, sequenced and used as a probe in Southern analysis were found to be unique to *R. unicornis*. Significance of this approach towards management and conservation of the other endangered species is highlighted.

Experimental Approaches

Collection of Blood Samples and DNA Isolation

Details of blood collection from one-horned rhino, *R. unicornis* (four males and two females), from Jaldapara Wildlife Sanctuary, used in our study have been reported earlier (Ali *et al*, 1999). DNA samples from *D bicornis* were kindly provided by Dr Colleen O'Ryan, Department of Biochemistry, University of Cape Town, Cape Town, South Africa. For inter-species MASA, Southern and Zoo-blot hybridizations, DNA samples from several other species were used and details have been reported earlier (Ali *et al*, 1999).

Conventional Restriction Survey, Identification of the Desired Fragment, Cloning and Sequencing

Conventional restriction survey is a potential approach for identification of the desired repeat fragment for cloning and sequencing and for ascertaining its organization and biological functions. A vertebrate genome is full of repetitive DNA (Cox & Mirkin, 1997) which are highly diverse encompassing a wide variety of repeat motifs including short and long ones (Ali & Gangadharan, 2000). This can be uncovered by digesting total genomic DNA with restriction enzyme and resolving on the agarose gel. Different enzymes generate varying patterns owing to their unique restriction recognition abilities. Thus, by using a set of 10-20 restriction enzymes specific for recognizing 4-6 base pairs, a unique DNA fragment from a desired species may be identified and characterized (Chattopadhyay *et al*, 2001). The sequences may be analyzed for their biological function and the resultant clone may be used as probe for its evolutionary conservation or species specificity.

MASA Reaction

It is conducted with or without radioisotopes depending upon the required resolutions and objectives. Both cold and labelled MASA reactions are conducted in a 25 µl volume. The cold MASA is free from radio-reporters and resolved on the agarose gel. A typical labelled MASA reaction contains approximately 25 ng of genomic DNA each as template, 20 picomoles of primer, 0.25 units of *Taq* DNA polymerase (Bangalore Genei, India), 2.5 mM $MgCl_2$, 200 µM of dTTP, dCTP, dGTP each and 0.5 µl of [α^{35}S] dATP (specific activity 1250 Ci/mmol, NEN, Boston, USA), 50 mM KCl, 20 mM Tris-HCl (*p*H 8.3), 0.1% Triton X-100 and an equal volume of mineral oil on a Thermal Cycler (Perkin Elmer, Cetus). The DNA is first heat denatured at 96°C for 2 min and amplified for 35 cycles comprising subsequent steps of denaturation at 94°C for 1 min, annealing at 60°C for 1 min and extension at 72°C for 1 min. On completion of the cycles, the reaction mixture is further incubated at 72°C for 5 min. Following this, approximately 5 µl of the amplified product is resolved on a 40 cm long, 3.5% non-denaturing polyacrylamide gel in 1xTBE (Tris-Borate EDTA) buffer for 8 hr at 400V using constant voltage. After the electrophoresis, the gel is dried for 2 hr at 80°C and exposed to X-ray film. Autoradiography is conducted according to standard protocol (Sambrook *et al*, 1989).

Restriction Survey and Generation of Species Specific Clone pSS(R)2

Following restriction survey using a battery of enzymes, an *Eco*RI specific fragment of about 906-bp from *R. unicornis* genome was identified. Cloning of this fragment and analysis (Accession No. Y-13565) of the resultant contig pSS(R)2 revealed presence of several eukaryotic transcriptional elements such as MALT box, NF-E1, Poly-A signal, lariat consensus sequences, TATA box, translational initiation sequences and several stop codons. Translation of the contig showed EGF-like domain cysteine pattern and Bowman-Birk serine protease inhibitor signatures.

Analysis of subset sequences in the 5' region from 1-165 nt indicated possible coding potential (test code value=0.97). Translation of the complementary strand from 906 to 706 nt and 190 to 2 nt showed more than 7 kDa proteins rich in non-polar residues. Thus, it is tempting to speculate that pSS(R)2 could be a part of coding region or adjacent to a functional gene. However, it is difficult to make unequivocal comments on the expression of pSS(R)2 since logistic constraints did not allow to conduct Northern or RT-PCR analyses. The pSS(R)2 used for zoo blot hybridization was found to cross hybridize only with *R. unicornis* genomic DNA clearly indicating that the same is a species specific fragment and can be used as potential probe (Fig. 1). This approach may be adopted for the generation of probes useful for other endangered species as well.

Genesis and Significance of MASA

The genesis of MASA and its significance have been reported earlier (Ali *et al,* 1999; Bashamboo & Ali, 1999, 2001). MASA conducted with 16 nt primer (CACCTCTCCACCTGCC) corresponding to 33.15 repeat loci (Ali & Wallace, 1988) with genomic DNA from different species showed species specific band profile (Fig. 2) substantiating that the primer sequences have evolved and organized in a unique species specific manner. Using same primer, MASA with DNA samples from *R. unicornis and D. bicornis* was conducted in an independent study. Though few bands were common but the overall band profile was found to be different for both the species (Fig. 3a). In a similar reaction, another 18 nucleotide long $(TGTC)_{4,2}$ primer was used that also showed species specific band pattern to both species (Fig. 3b). This suggests that using appropriate primers, a large number of species may be discriminated. This approach is similar to that of RAPD (Welsh *et al,* 1990) or AP-PCR (Williams *et al,* 1990) except that primers used for MASA are not arbitrary and have relatively fewer annealing sites in the target DNA.

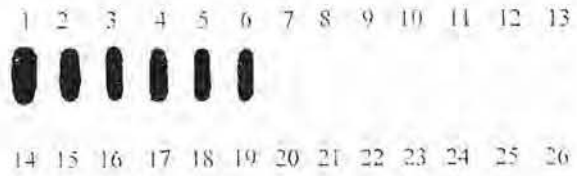

Fig. 1—Evolutionary uniqueness of pSS(R)2 clone from *R. unicornis* based on slot blot hybridization using about 200 ng genomic DNA from different organisms. Spots represent samples: 1-6 *R. unicornis*; 7, human; 8-9, goat; 10, sheep; 11, buffalo; 12, cattle; 13, pig; 14, kangaroo; 15, rabbit; 16, rat; 17, mouse; 18, catfish; 19, camel; 20, bird; 21, house cricket; 22, *E. coli*; 23-24, *D. bicornis;* 25-26 control (2xSSC). Note the exclusive signals in the *R. unicornis* and its absence in all the other samples. (Reproduced from GENE, 228; 133-42, 1999, with permission of Elsevier Science).

As shown earlier (Fig. 2), MASA reaction conducted simultaneously with several unrelated genomic DNA samples using a single primer shows species-specific band profiles. Using this approach, individual band may be cloned and sequenced from a desired species. This approach provides a feasible option for the development of species-specific markers. In combination with an appropriate primer, MASA can discriminate not only amongst different species but also offers a reliable tool to walk along the chromosome since each fragment originates from a specific chromosome (Bashamboo *et al,* 2001).

Following this approach, 688 base pair band specific to *R. unicornis*, was identified, cloned and sequenced (Accession N0. AF-296689). The resultant clone, pSG5, was found to have an island of "GT" repeats showing homology with microtubule associated serine/threonine protein kinase gene indicating its possible coding potential Such dimeric repeats

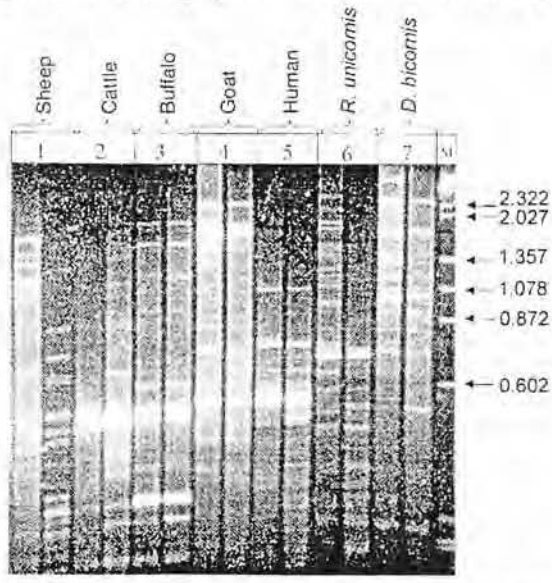

Fig. 2—Minisatellite associated sequence amplification (MASA) with genomic DNA from several species including *R. unicornis* and *Diceros bicornis* using 16 base long oligo primer representing consensus sequence (5' CACCTCTCCACCTGCC 3') of 33.15 minisatellite repeat loci, resolved on 2% agarose gel showing species specific band profile. Molecular size marker ΦX174 in kb (M) is given on the right.

have been reported in coding and non-coding sequences (Dokholyan *et al*, 2000). In our recent study, the pSG5 clone showed eukaryotic transcriptional and promoter elements in the upstream regions to the putative ORF of the contig. The presence of ORF is in accordance with the data obtained from the Blast search (http://www.ncbi.nlm.nih.gov/cgi-bin/BLAST/nph-newblast) showing homology of pSG5 from nucleotide 1-358 with microtubule associated serine/threonine kinase gene. In addition, several eukaryotic transcriptional elements such as MALT Box (GGAKGGA), NF-E1 (WGATAMS), lariat consensus sequence (YNYTRAY), TATA box (TATAAAA), T cell element (GGGRTTTMA) and translational initiation sequence were detected. The relative position of ORF with island of "GT" repeats is shown in the schematic representation (Fig. 4).

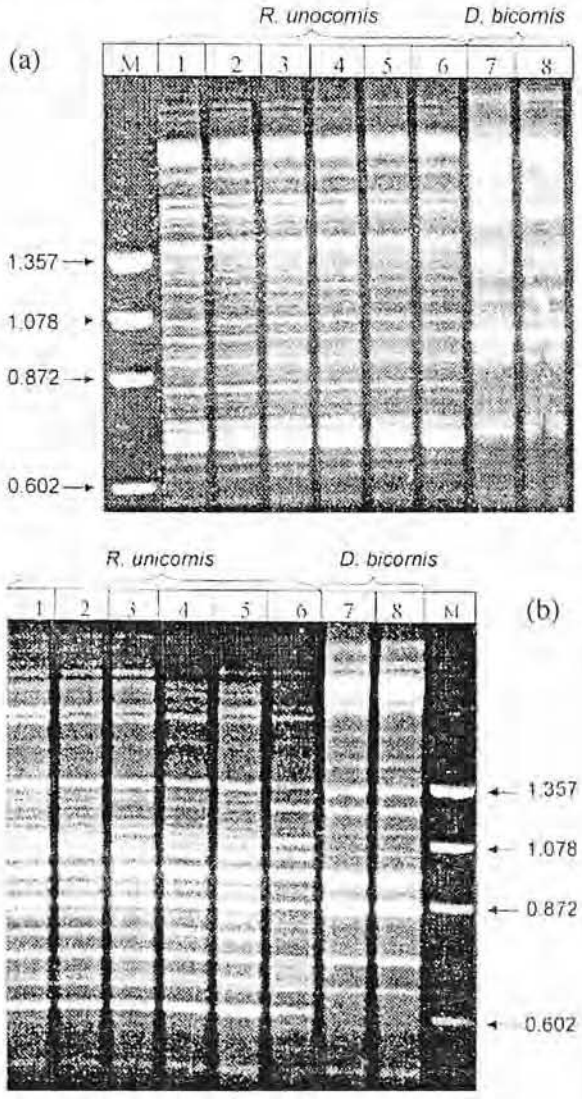

Fig. 3—Minisatellite associated sequence amplification (MASA) with genomic DNA from *R. unicornis* and *Diceros bicornis* using 16 base long oligo primer representing consensus sequence (5' CACCTCTCCACCTGCC 3') of 33.15 minisatellite repeat loci (a) and 18 base long primer $(TGTC)_{4.2}$, (b), both resolved on 2% agarose gel. Note the differences in the band profiles between the two species seen in both the reactions. Molecular size marker ΦX174 in kb is given on the left and right, respectively. (Reproduced from GENE, 228: 133-42, 1999, with permission of Elsevier Science).

pSG5 and pSSR(2), unique to Rhinoceros unicornis, are Related to Repeat Sequences

Both the fragments generated by MASA reaction and restriction survey used for Southern blot hybridisation in independent studies were found to be unique to *R. unicornis* species (Ali *et al*, 1999). Repetitive sequences have been reported to expand or shrink depending on the evolutionary forces (John *et al*, 1996). However, the actual molecular mechanism and biological significance of such expansion or shrinkage remain unknown. Our analysis of these clones showed sites for *Alu*I enzyme in pSG5 and pSSR(2) contigs. In earlier studies on the bubaline genome, *Alu*I sites were detected in a satellite tagged transcribing sequence that showed programmed modulation in the germline (semen DNA) samples (Chattopadhyay *et al*, 2001). This was construed to be a possible mechanism of transcriptional inactivation of the sequnces in the meiocytes during spermatogenesis. In the pSG5 contig, presence of *Alu*I sites conferring sequence modulation in germline seems to be an attractive possibility (Fig. 5). However, owing to logistic constrain, it was not feasible to conduct similar analysis on rhino germline DNA samples.

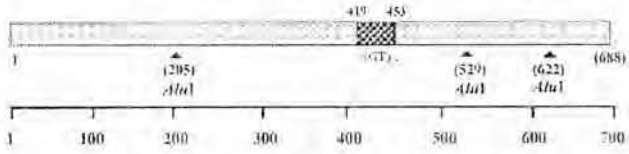

Fig. 4—Schematic representation of pSG5 genomic DNA inserts from *R. unicornis* generated by minisatellite associated sequence amplification showing three *Alu*I recognition sites in each and an island of "GT" repeat.

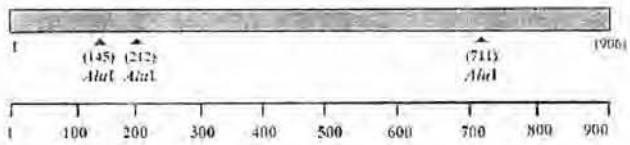

Fig. 5—Schematic representation of pSSR(2) genomic DNA inserts from *R. unicornis* generated by restriction survey showing three *Alu*I recognition sites (see text for details and Ali *et al*, 1999).

Irrespective of the germline sequence modulation or coding potential of *R. unicornis* derived contigs, the same may be used as a reliable marker for ascertaining the origin of biological sample(s) in the event of poaching or illegal trafficking of the animals. Thus, DNA based genetic marker is likely to complement the already existing approaches such as the use of ribosomal gene from the mtDNA for species identification.

Acknowledgement

This work was supported by a grant no. SP/SO/D-51/93 from the Department of Science and Technology, New Delhi to SA and a core grant from the Department of Biotechnology, Govt of India to the National Institute of Immunology, New Delhi.

References

Ali S & Gangadharan S, 2000. Differential evolution of coding and non-coding sequences in related vertebrates: Implications in probe design. *Proc Indian Natl Sci Acad Part B Biol Sci.*, **66**, 49-67.

Ali S & Hasnain S E, 1999. Conservation, management and documentation of livestock genetic resources and biodiversity: a biotechnology perspective. *Curr Sci*, **77**, 1154-1158.

Ali S & Wallace R B, 1988. Intrinsic polymorphism of variable number tandem repeat loci in the human genome. *Nucleic Acids Res*, **16**, 8487-8496.

Ali S, Ansari S, Ehtesham N Z, Azfer M A, Homkar U, Gopal R & Hasnain S E, 1998. Analysis of the evolutionarily conserved repeat motifs in the genome of the highly endangered central Indian Swamp deer *Cervus duvauceli branderi*. *Gene*, **223**, 361-367.

Ali S, Azfer M A, Bashamboo A, Mathur P K, Malik P K, Mathur V B, Raha A K & Ansari S, 1999. Characterisation of a species specific repetitive DNA from a highly endangered wild animal *Rhinoceros unicornis* and assessment of genetic polymorphism by microsatellite associated sequence amplification (MASA). *Gene*, **228**, 33-42.

Ali S, Müller C R & Epplen J T, 1986. DNA fingerprinting by oligonucleotide probes specific for simple repeats. *Hum Genet*, **74**, 239-243.

Bashamboo A & Ali S, 1999. DNA typing and microsatellite associated sequence amplification (MASA) in detecting loss or gain of alleles. *Indian J. Biochem. Biophys*, **36**, 312-317.

Bashamboo A & Ali S, 2001. Minisatellite associated sequence amplification (MASA) of the hypervariable repeat marker 33.15 reveals a male specific band in humans. *Mol Cell Probes*, **15**, 89-92.

Chattopadhyay M, Gangadharan S, Kapur V, Azfer M A, Prakash B & Ali S, 2001. Satellite tagged transcribing sequences in the *Bubalus bubalis* genome undergo programmed modulation in the meiocytes: possible implication in transcriptional inactivation. *DNA Cell Biol*, **20**, 587-593.

Cox R & Mirkin S M, 1997. Characteristic enrichment of DNA repeats in different genomes. *Proc Natl Acad Sci USA*, **94**, 5237-5242.

Cunningham J, Harley E H & O'Ryan C, 1999. Isolation and characterisation of minisatellite loci in black rhinoceros (*Diceros bicornis*). *Electrophoresis*, **20**, 1778-1780.

Dokholyan N V, Buldyrev S V, Havlin S & Stanley H E, 2000. Distributions of dimeric repeats in non-coding and coding DNA sequences. *J Theor Biol*, **202**, 273-282.

Jeffreys A J, Wilson V & Thein S L, 1985. Hypervariable minisatellite regions in human DNA. *Nature (London)*, **314**, 67-73.

John M V & Ali S, 1997. Synthetic DNA based genetic markers reveal intra and inter-species DNA sequence variability in the *Bubalus bubalis* and related genomes. *DNA Cell Biol*, **16**, 369-378.

John M V, Parwez I, Sivaram M V S, Mehta S, Marwah N & Ali S, 1996. Analysis of VNTR loci in fish genomes using synthetic oligodeoxyribonucleotide probes. *Gene*, **172**, 191-197.

Mattapallil M J & Ali S, 1997. Parentage assessment of an IVF CALF from *Bubalus bubalis* by DNA fingerprinting. *Gene*, **206**, 209-214.

Mattapallil M J & Ali S, 1999. Analyses of conserved microsatellite sequences suggest closer relationship between water buffalo *Bubalus bubalis* and sheep *Ovis aries DNA*. *Cell Biol*, **18**, 513-519.

Merenlender A M, Woodruff D S, Ryder O A, Kock R & Vahala J, 1989. Allozyme variation and differentiation in African and Indian rhinoceroses. *J Hered*, **80**, 377-382.

Patrushev L I, 1997. Altruistic DNA. About protective functions of the abundant DNA in the eukaryotic genome and its role in stabilising genetic information. *Biochem Mol Biol Int*, **41**, 851-860.

Raina A, Sulaiman I M, Das P, Ehtesham N Z, Ali S, Dogra T D & Hasnain S E, 1996. Characterization of a human alphoid satellite DNA sequence: Potential use in assessing genetic diversity in Indian populations. *Gene*, **173**, 247-250.

Ray P, Gangadharan S, Chattopadhyay M, Bashamboo A, Bhatnagar S, Malik P K & Ali S, 1999. A bubaline derived satellite DNA probe uncovers generic affinities of gaur with other bovids. *J Biosci*, **24**, 295-299.

Sambrook J, Fritsch E F & Maniatis T, 1989. Molecular cloning a Laboratory Manual, (2nd edn), Cold Spring Harbor Laboratory Press, New York.

Vos P, Hogers R, Bleeker M, Reijans M van de, Lee T, Hornes M, Frijters A, Pot J, Peleman J & Kuiper M, 1995. AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res*, **23**, 4407-4414.

Welsh J & McClelland M, 1990. Fingerprinting genomes using PCR with arbitrary primers. *Nucleic Acids Res*, **18**, 7213-7218.

Williams J G, Kubelik A R, Livak K J, Rafalski J A & Tingey S V, 1990. DNA polymorphisms amplified by arbitrary primers are useful as genetic markers. *Nucleic Acids Res*, **25**, 6531-6535.

Xu X & Arnason U, 1997. The complete mitochondrial DNA sequence of the white rhinoceros, *Ceratotherium simum* and comparison with the mtDNA sequence of the Indian rhinoceros, *Rhinoceros unicornis*. *Mol Phylogenet Evol*, **7**, 189-194.