Microsatellite loci isolation from river buffalo using enriched partial genomic libraries

L.P. Rodrigues Venancio¹, M.E.J. Amaral¹, M.P.C. Schneider², A.L. Silva²

¹UNESP – Universidade Estadual Paulista, IBILCE, Depto. Biologia, São José do Rio Preto, SP, Laboratório de Genômica Comparativa, 15054-000, Brazil
²Centro de Ciências Biolóxicas, Depto. Genética, Laboratório de Polimorfismo de DNA - LPDNA, Universidade Federal do Pará, Belém, Brazil

Corresponding author: L.P. Rodrigues Venancio, Universidade Estadual Paulista (UNESP - São José do Rio Preto, São Paulo, Brazil) - Email: larissa_biorp@yahoo.com.br

ABSTRACT: The extensive use of buffalo in agriculture, especially in developing countries, begs for genetic resources to evaluate and improve traits important to local and regional economies. Brazil presents the largest water buffalo populations in the New World, with 1.1 million heads including swamp and river types. To design rational breeding strategies for optimum utilization and conservation of available genetic variability in the Brazilian buffalo's population, it is essential to understand their genetic architecture and relationship among various breeds. This depends, in part, on the knowledge of their genetic structure based on molecular markers like microsatellites. In the present study, we developed six enriched partial genomic libraries for river buffalo using selective hybridization methods. Genomic DNA was hybridized with six different arrays of repeat motif, 5’ biotinylated - (CA)₁₅, (CT)₁₅, (AGG)₈, (GAAA)₈, (GATA)₈, (AAAAC)₈ – and bound to streptavidin coated beads. The cloning process generated a total of 1920 recombinant clones. Up to date, 487 were directly sequenced for the presence of repeats, from which 13 have been positive for presence of repeats as follows: 9 for di-nucleotide repeats, 3 for tri-nucleotide repeats and 1 for tetra-nucleotide repeat. PCR primer pairs for the isolated microsatellites are under construction to determine optimum annealing temperature. These microsatellites will be useful for studies involving phylogenetic relationships, genome mapping and genetic diversity analysis within buffalo populations worldwide.

Key words: Microsatellites isolation, Enriched partial genomic libraries, River buffalo.

INTRODUCTION - The water buffalo (Bubalus bubalis) contributes immensely to the agricultural economy through milk, meat, hides and draught power (Kumar et al., 2006). There are about 158 million buffaloes in the world and a larger part of the human population depends on domestic river buffalo than on any other livestock species in the world (FAO & UNEP, 2000). Scientific resources are limited in many of the countries where buffalo are economically important livestock and as a consequence genome research has not been supported at he level of some the others species (Womack, 2005). During the nineteenth and twentieth centuries water buffalo were imported into Brazil and population have increased so greatly that Brazil maintains the largest river buffalo...
population in the Americas (Sena et al., 2003). There are approximately 1.1 million buffaloes in Brazil, including swamp (Carabao breed) and river types (Murrah, Jafarabadi, Mediterranean breeds) (ANUALPEC, 2006). To design rational breeding strategies for optimum utilization and conservation of available genetic variability in the Brazilian buffalo's population, it is essential to understand their genetic architecture and relationship among various breeds. This depends, in part on the knowledge of their genetic structure based on molecular markers like microsatellites (Gonçalves et al., 2006). Microsatellites or simple sequence repeats (SSRs) are tandemly repeated motif of 1-6 bases pairs found in all eukaryotic and prokaryotic genomes present in coding and noncoding regions (Tóth et al., 2000). Microsatellites are powerful genetic markers and extremely valuable for genome mapping, forensic DNA studies, paternity testing, population genetics and conservation/management of biological resources (Jarne and Lagoda, 1996). Microsatellites loci can be isolated de novo from a species of interest or utilized for related species in which it has been previously characterized. This strategy of using a particular microsatellites in different but related species saves time in the genetic analysis process. However, several studies using this strategy has described a decrease in the PIC (Polymorphism Information Content) and an increase of null alleles in the population. Considering this scenario, in the present study we developed six enriched partial genomic libraries using selective hybridization methods for microsatellites isolation from river buffalo.

**MATERIAL AND METHODS** - Six enriched microsatellite libraries for *Bubalus bubalis* were constructed and screened, using modified protocols described by Gautschi et al., 2000 and Guillemaud et al., 2000. Genomic DNA was digested overnight at 37°C with *Sau3AI* restriction enzyme (Amersham Bioscience). The resulting fragments were separated using agarose gel electrophoresis. Those fragments ranging from 300 to 1000 bp were purified using DNA GFX PCR Purification Kit (GE Healthcare) according to the manufacturer’s instructions, and ligated at 16°C overnight into *Sau3AI* adaptors (AdapF: 5’CTCTTGCTTACGCTGACTC3’ / AdapR: 5’GATCGAGTCCACGCTAAGCAAGAGCACA3’) using T4 DNA ligase at 16°C overnight. After denaturation, single-strand DNA fragments were enriched by hybridization with 5'-biotinylated, 3'-aminated (CA)\textsubscript{15}, (CT)\textsubscript{15}, (AGG)\textsubscript{8}, (GAAA)\textsubscript{8}, (GATA)\textsubscript{8}, (AAAAC)\textsubscript{8} oligonucleotides bound to streptavidin-coated magnetic beads (Dynabeads M-280 streptavidin, Dynal Biotech-Invitrogen). Each set of fragments captured by the probes were eluted in 100µL TE buffer prior to the PCR amplification using adapter-specific primers (AdapF). The size-selected fragments were cloned into pCR2.1 by using the TA cloning kit (Invitrogen). The resulting colonies were grown ~5 hours in liquid media and used as templates for PCR by using vector primers. PCR product was used for DNA sequencing by using standard protocols of the ThermoSequenase II dye terminator cycle sequencing kit (Amersham Bioscience). Sequencing reactions were analyzed by using the capillary sequencer MegaBACE 1000 (Amersham Bioscience). The sequences were edited and aligned using BIOEDIT software (Hall, 1999) for primer design.

**RESULTS AND CONCLUSIONS** - The cloning process generated a total of 1920 recombinant clones. Up to date, 487 (25.37%) clones were directly sequenced for the presence of repeats, from which 13 (2.66%) have been positive follows: 9 for di-nucleotide repeats, 3 for tri-nucleotide repeats and 1 for tetra-nucleotide repeats. Microsatellites are classified as...
perfect, imperfect, interrupted or composite according with the type of the repeat sequence (Oliveira et al, 2006). Among the isolated microsatellites, seven were classified as perfect with the repeat sequence not interrupted by any base different from the motif (e.g. CACA-CACACACACACACACACACACA) (Figure 1). Three were classified as interrupted, since we found a small no match sequence of 3 to 4 nucleotides within the repeated motif (e.g. CACA-CACACTTACACA). Three microsatellites were considered imperfect, because we found a pair of bases no match inside of the repeated motif (e.g. TGTGTCCTGTG). PCR primer pairs for all isolated microsatellites are under construction. Although, the results are still preliminary, the selected microsatellites will be useful for studies involving phylogenetic relationships, genome mapping and genetic diversity analysis within river buffalo populations worldwide. Considering that only 25.37% of the total clones were screened from the library and characterized, we verified that the library construction was successful regarding the identification of new microsatellites for this specie.

Figure 1. A) Chromatogram showing (CA)_{12} repeat from an microsatellite isolated, named BBUMS001. B) Complete sequence in FASTA format from the microsatellite BBUMS001. The sequence in red shows (CA)_{12} repeat.

ACKNOWLEDGMENTS - This study was supported by the CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) and FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo) (proc. FAPESP 06/54618-1). We are grateful to Dr. Evonnildo C. Gonçalves for his important suggestions and M. Silvanira R. Barbosa and Soraya S. Andreade for their technical assistance.