GMR Genetics and Molecular Research
Online Journal
ISSN:1676-5680

Cloning and sequencing of the rDNA gene family of the water buffalo (*Bubalus bubalis*)

C.Y. Pang[1]*, T.X. Deng[1]*, D.S. Tang[2], C.Y. Yang[1], H. Jiang[2], B.Z. Yang[1] and X.W. Liang[1]

[1]Guangxi Key Laboratory of Buffalo Genetics, Breeding and Reproduction, Guangxi Buffalo Research Institute, Chinese Academy of Agricultural Science, Nanning, P. R. China
[2]Medical College, Foshan University, Foshan, P. R. China

*These authors contributed equally to this study.
Corresponding author: X.W. Liang
E-mail: liangbri@126.com

Genet. Mol. Res. 11 (3): 2878-2883 (2012)
Received September 21, 2011
Accepted February 1, 2012
Published July 10, 2012
DOI http://dx.doi.org/10.4238/2012.July.10.3

ABSTARACT. The rDNA genes coding for ribosomal RNA in animals are complicated repeat sequences with high GC content. We amplified water buffalo rDNA gene sequences with the long and accurate (LA) PCR method, using LA Taq DNA polymerase and GC buffer, based on bioinformatic analysis of related organisms. The rDNA genes were found to consist of 9016 nucleotides, including three rRNA genes and two internal transcribed spacers (ITS), which we named 18S rRNA, ITS1, 5.8S rRNA, ITS2 and 28S rRNA. We tested and optimized conditions for cloning these complicated rDNA sequences, including specific rules of primer design, improvements in the reaction system, and selection of the DNA polymerase.

Key words: Molecular cloning; GC-rich; rDNA; rRNA; *Bubalus bubalis*

INTRODUCTION

The water buffalo (*Bubalus bubalis*) is one of the major large animals in China, which has a great draft capacity, long working life, docile temperament and rough feeding endurance conditions. To improve the production potential of dairy-type and beef-type buffaloes and its economic value, we should enhance the study of molecular breeding technology, such as *in vitro* fertilization, nucleus transfer, transgenics, etc. Therefore, the genomic study of the water buffalo will provide basic of biology for molecular breeding. To develop the technique of multiple loci gene targeting repeat sequences of buffalo *in vivo* (Tang et al., 2002), i.e., the foreign genes were inserted into the internal transcribed spacers between the 3 rRNA genes, the rRNA gene family of buffalo needs to be cloned and the targeting locus needs to be found by sequencing analysis. In animals, the 18S, 5.8S and 28S rRNA genes are in a series and each is separated by an internal transcribed spacers (ITS), forming transcription units, which are often stretches of G or C, with high GC content and complicated structure (Figure 1). Therefore, the cloning of the complicated rDNA genes is very difficult and we have not found any previous reports about the cloning of the complicated rDNA gene family of water buffalo. In this study, we successfully cloned the complicated rDNA genes of buffalo.

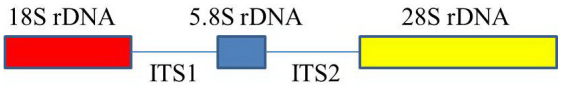

Figure 1. The structure of the rDNA transcription unit in animals.

MATERIAL AND METHODS

Bioinformatic analysis of rRNA genes of *Bubalus*

The rRNA gene family sequence of *Bos taurus* has been submitted to the NCBI under accession No. DQ222453. The water buffalo and *B. taurus* have a close relationship, where they belong to the family Bovidae. It was thus feasible to design primers for the rRNA gene family sequence of water buffalo based on the rRNA gene sequences of *B. taurus*. A contig was assembled with these valuable sequences of the rRNA genes and ITS by using the DNAstar Lasergene SeqMan software. The contig was analyzed by BLAST and FASTA with the non-redundant sequences of GenBank, Protein Data Bank, European Molecular Biology Laboratory and DNA Data Bank of Japan databases. The conserved regions of the contig were found, in which 16 pairs of primers were designed for amplifying the complicated rRNA gene family sequences of water buffalo. Only 6 pairs of primers could effectively amplify its genes and ITS fragments. Five pairs of primers were chosen for the final PCR amplification (Table 1).

Isolation of genomic DNA of water buffalo and PCR amplification

Genomic DNA was extracted from the liver of water buffalo by the TIANamp Marine Animals DNA Kit (TIANGEN, China). The experiment results indicated that the rRNA genes could be amplified effectively using Long and Accurate *Taq* polymerase (LA PCR *Taq*) (TAKARA, Japan) with GC buffer. PCR amplifications were performed in a Biometra ther-

mocycler. Reactions were carried out in 25-µL volumes containing 5 pmol each primer, 100 µM of each dNTP, 12.5 µL 2X GC buffer I or buffer II, 0.25 U LA PCR *Taq* polymerase, and ~100 ng genomic DNA. The primers BRDS1, BRDS2, BRDS3, BRDS4, and BRDS5 were annealed at 60°, 60°, 59°, 59°, and 58°C, respectively (Tang et al., 2006).

Table 1. The primers used to amplify the rDNA genes in *Bubalus bubalis*.

PCR product	Primer	Primer sequences (5'→3')	Direction	Annealing temperature °C
BRDS1	P1	TACCTGGTTGATCCTGCCAGTAGCATATGC	F	60
	P2	GTTCACCTACGGAAACCTTGTTACGACTT	R	
BRDS2	P3	CGTGACATCAGATCAGACGTGGCAACC	F	60
	P4	CAGGGCTAGTTGATTCAGCAGGTGAGTTGT	R	
BRDS3	P5	CGGTGGATCACTCGGCTCGTGCGTCGAT	F	60
	P6	AGCCTTAGATGTTTACCACCCACTT	R	
BRDS4	P7	GCCGTTCTTAGTTGGTGGAGCGATTT	F	59
	P8	GCACGAGCCGAGTGATCCACCGCTAAGAGT	R	
BRDS5	P9	AAGCCGCTGTGGAGCAATGAAGGTGAAG	R	58
	P10	GGACAAACCCTTGTGTCGAGGGCTGACT	F	

F = forward; R = reverse.

Cloning and identification of the rRNA genes of buffalo

The PCR products were separated on 1.0% agarose gel electrophoresis and purified from the gel according to the instruction of the E. Z. N. A. Gel Extraction kit (OMEGABIO-TEK, USA). The PCR products were cloned to pMD19-T Simple vector (TAKARA) and transformed into *Escherichia coli* DH5α. The positive colonies were picked out and were then identified by PCR with the original primers. The positive plasmids were further tested by sequencing by Invitrogen (Guangzhou, China).

RESULTS

PCR amplification

Five pairs of primers were designed in the conserved regions of the rRNA genes of *B. taurus*, namely P1/P2, P3/P4, P5/P6, P7/P8, and P9/P10, for amplifying the rRNA gene family of water buffalo. Specific PCR products were obtained as seen by agarose/TAE gel electrophoresis (Figure 2). The sizes of the products were consistent with those expected, namely 1,856, 1880, 1240, 1640, and 3320 bp, respectively.

Cloning and identification of repeat sequences of the rRNA genes of buffalo

Colonies in which the rRNA genes could be amplified using the original primers were considered positive. A specific plasmid was purified from the positive colonies as seen by agarose/TAE gel electrophoresis. The sizes of the plasmids were consistent with those expected, i.e., 4500, 4500, 3800, 4300, and 5900 bp, respectively.

Sequencing of the repeat sequences of the rRNA genes of buffalo

The positive plasmids were confirmed by PCR and sequencing. The sizes of 5 rRNA

gene fragments BRDS1, BRDS2, BRDS3, BRDS4, and BRDS5 in water buffalo were 1856, 1880, 1240, 1640, and 3320 bp, respectively. The complete rRNA gene family of 9016 bp was obtained from the 3 rRNA gene fragments and 2 ITS gene fragments, using SeqMan/DNAstar. The sizes of 18S rRNA, ITS1, 5.8S rRNA, ITS2, and 28S rRNA were 1869, 1465, 157, 973, and 4552 bp, respectively. The average GC content of the complete sequence was as high as 61.3%. The GC content of 18S rRNA, 5.8S rRNA, 28S rRNA, ITS1, and ITS2 was as high as 56.1, 51.6, 58.1, 67.6, and 79.9%, respectively. The results indicated that the rRNA genes of water buffalo are DNA repeat sequences containing mostly stretches of G or C, with complicated structure (GenBank accession No. JN412502).

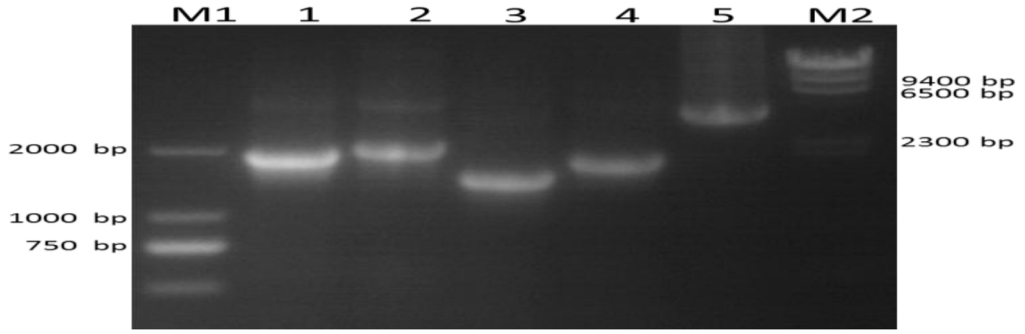

Figure 2. Ethidium bromide-stained agarose gel showing the PCR products of BRDS1, BRDS2, BRDS3, BRDS4, and BRDS5. *Lane M1* = 2-kb ladder; *lane 1* = 1856-bp product amplified using the primers P1/P2; *lane 2* = 1880-bp product amplified using the primers P3/P4; *lane 3* = 1240-bp product amplified using the primers P5/P6; *lane 4* = 1640-bp product amplified using the primers P7/P8; *lane 5* = 3320-bp product amplified using the primers P9/P10; *lane M2* = λ *Hin*dIII digest.

Analysis of homology and evolution of the rRNA gene family of buffalo

The homology analysis of 18S, 5.8S and 28S rRNA genes between related animals displayed 79.2~99.0% similarity, especially between *B. bubalis* and *Odocoileus virginianus*, *B. taurus* and *Budorcas taxicolor*, with as high as 99.0% based on the homology of the18S rRNA gene. The 18S, 5.8S and 28S rRNA genes were more significant than ITS1 and ITS2 for homology comparison among animals (data not shown). The results in the present study showed that there are more variations in ITS1 and ITS2 between the related animals. Among the available nuclear sequences, ITS sequences have been used successfully in studying phylogenetic and genomic relationships of plants at lower taxonomic levels (Nieto et al., 2004; Liu et al., 2006; Peng et al., 2010), and the evolution of the rRNA genes in organisms is highly conservative. The variations in the rRNA genes occur slowly and are more conservative than in the whole genome of the organism. This phenomenon leads to the retention of more sequences from ancestors in rRNA genes. Therefore, the rRNA genes can be used for the analysis of evolutionary relationships. The evolutionary tree of 18S rRNA, 5.8S rRNA and 28S rRNA genes (Figure 3 only shows 18S rRNA gene) showed that *B. bubalis* and *O. virginianus* are the closest, belonging to *Artiodactyla*.

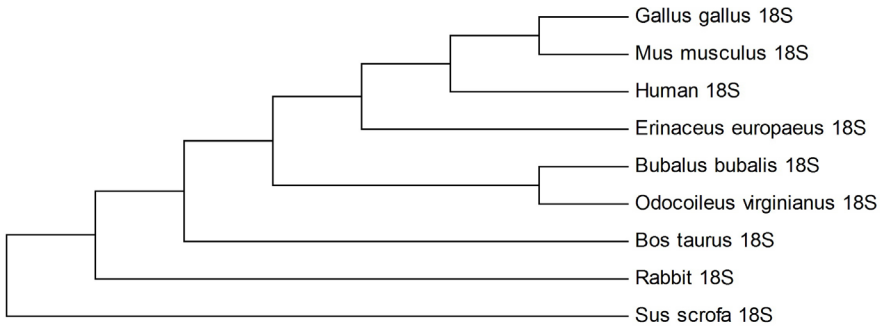

Figure 3. The evolutionary relation analysis of 18S rRNAs.

DISCUSSION

To amplify the complicated sequences with high GC-content, the conditions for the cloning of complicated DNA sequences are of the utmost importance. PCR primers are usually 18-30 nucleotides in length, giving them high specificity. However, this theoretical prediction may not always be true in diverse and complicated biological genomes. With the development of sequencing technology and rapid cost reduction, many genomes have been sequenced and annotated in databases. The pre-designed primer can thus be used to decrease non-specific amplification and test PCR specificity including the target location and amplicon size in one or multiple target genome(s) by *in silico* PCR. It can identify the mismatches in primer binding sites due to known single nucleotide polymorphisms and/or unwanted amplicons from a homologous gene or a pseudogene (Yu and Zhang, 2011).

There were 5 pairs of primers in the present study that could amplify the specific products from the genomic DNA of water buffalo. The results revealed that the design of the primers for complicated repeat sequences should pay attention to the following special rules in addition to the common principles. 1) To avoid the sequence with continuous G or C repeats were used to design the binding loci of the primers, especially the primer 3'-end, i.e., the designed primer containing three continuous G or C at the primer 3'-end could easily make mismatch between primers and GC-rich regions of templates. 2) Because the primer is GC-rich, the annealing temperature should not be below 58°C. 3) The GC content in the middle of the primers can be fairly higher, while at the 3'-end relatively lower, to avoid the four bases A, C, T, and G being in a cluster, which could increase the hybrid stability between the primers and target sequences. 4) According to the assumption of assuring the specificity and validity of PCR amplification, shorter primers are better to increase the efficiency of amplification.

Subsequently, some strategies for improving regular PCR were also useful in long and accurate (LA) PCR. 1) There has been an appreciable increase in the PCR product of large DNA fragments with 8% glycerin. 2) The GC-rich regions in DNA sequences with high denaturing temperature can hardly be amplified successfully. To increase the efficiency of amplification of GC-rich regions, 1-10% dimethyl sulfoxide (DMSO) is added as an alternative. Theoretically, the mechanism of DMSO can give rise to instability of DNA structure by dehydration of DNA, lowering the melting point and denaturing temperature, so as to increase the efficiency

of amplification (Li et al., 2002). It has been shown that glycerin and DMSO used in combination to increase the efficiency of amplification are better than singly (He and Li, 2001). 3) A low concentration of KCl can be helpful for the annealing of primers to promote the extension of the DNA strand. 4) LA PCR has its own peculiarity; while the Mg^{2+} concentration in conventional PCR is 2.5 mM, LA PCR needs a higher Mg^{2+} concentration (5 mM). Another characteristic of the LA PCR system is the requirement of high pH, that is, pH 8.8-9.2. It is helpful to maintain the stability of DNA templates with high pH. Increased pH leads to higher amplification efficiency maybe because of lower depurination of the DNA template. It was confirmed that the GC buffer I and GC buffer II (TAKARA) are suitable for the amplification of complicated repeat sequences with high GC content.

To amplify the complicated repeat sequences with high GC content, high-fidelity DNA polymerase should be the first choice. With regard to the efficiency and fidelity of amplifying long DNA fragments, especially fragments longer than 10 kb, LA PCR *Taq* polymerase is superior to other polymerases. When the DNA templates have complicated secondary structures such as GC-rich regions and repeat sequences, LA PCR *Taq* polymerase with the GC buffer I or GC buffer II can be used to amplify effectively the special products.

Finally, studies of different denaturation times and temperatures showed that there was a negative correlation between the size and output of the amplified fragments and denaturation time. If the denaturation time and temperature are kept in the lowest range, the former will be positively correlated with the extension time. According to the above study, the optimal cycle parameters were set as followed: denaturation at 94°C for 30 s, annealing at 58°-64°C for 30 s and extension at 72°C for 1-3 min.

ACKNOWLEDGMENTS

Research supported by the Natural Science Foundation of Guangxi, China (grant #2010GXNSFA013096) and Guangxi Foundation for Distinguished Professor, China.

REFERENCES

Chen SW and Wang XZ (1996). The PCR methods of amplifying the large fragments of target DNA. *Biotechnology* 6: 1-2.

He ZY and Li YQ (2001). The role of dimethyl sulfoxide on polymearase chain reaction for human *apoE* gene. *J. Jinan Univ. (Nat. Sci.).* 22: 104-107.

Li AL, Jiang T, Ma ZY and Jia JZ (2002). Several methodologies on improving the PCR products. *Biotechnol. Bull.* 6: 33-35.

Liu Q, Ge S, Tang H, Zhang X, et al. (2006). Phylogenetic relationships in *Elymus* (Poaceae: Triticeae) based on the nuclear ribosomal internal transcribed spacer and chloroplast *trnL*-F sequences. *New Phytol.* 170: 411-420.

Nieto FG, Gutierrez LB and Fuertes AJ (2004). Fine-scale geographical structure, intra-individual polymorphism and recombination in nuclear ribosomal internal transcribed spacers in *Armeria* (Plumbaginaceae). *Ann. Bot.* 93: 189-200.

Peng YY, Baum BR, Ren CZ, Jiang QT, et al. (2010). The evolution pattern of rDNA ITS in Avena and phylogenetic relationship of the *Avena* species (Poaceae: Aveneae). *Hereditas* 147: 183-204.

Tang D, Shi J, Tian Y and Gu W (2002). Multiple site gene targeting in somatic cells and desired site integration of genes. *J. Foshan Univ. (Nat. Sci. Edit.).* 20: 64-68.

Tang D, Jiang H, Zhang Y, Li Y, et al. (2006). Cloning and sequencing of the complicated rDNA gene family of *Bos taurus*. *Czech J. Anim. Sci.* 51: 425-428.

Yu B and Zhang C (2011). *In silico* PCR analysis. *Methods Mol. Biol.* 760: 91-107.

