VETERINARSKI ARHIV 92 (5), 577-590, 2022 DOI: 10.24099/vet.arhiv.1736

Association between single-nucleotide polymorphisms in *ARID4A* gene and sperm quality of Chinese water buffalo

Guanhong Lu[1,2#], Chuanyu He[3#], Xi Yang[1#], Yunyun Cheng[1], Rui Yang[1], Jie Song[1], Siyao Wang[1], Chunli Wang[1], Ze Ma[1], Jiapeng Li[1], Tianqi Feng[1], Linlin Hao[1]*, and Changhong Li[2]*

[1]College of Animal Science, Jilin University, Jilin, China
[2]College of Life Sciences, Baicheng Normal University, Jilin, China
[3]TECON Biological Co., Ltd., Urumqi, China

LU, G., C. HE, X. YANG, Y. CHENG, R. YANG, J. SONG, S. WANG, C. WANG, Z. MA, J. LI, T. FENG, L. HAO, C. LI: Association between single-nucleotide polymorphisms in *ARID4A* gene and sperm quality of Chinese water buffalo. Vet. arhiv 92, 577-590, 2022.

ABSTRACT

ARID4A (AT-rich interaction domain 4A) is closely related to animal sperm quality traits. In the present study, the association between *ARID4A* gene polymorphisms of Chinese water buffalo (*Bubalus bubalis*) with sperm quality traits was examined, including ejaculate volume, sperm concentration, post-thaw sperm motility, and sperm abnormality of buffalo semen. Seven single-nucleotide polymorphisms (SNPs) of *ARID4A* gene were detected in 156 Chinese water buffaloes by Sanger sequencing and identifying overlap peaks. Among the SNPs, six were associated with at least one sperm quality trait. In brief, g.21192G>C, g.21285C>G, and g.21364A>G could be used as potential markers for selecting semen with low sperm abnormality, high ejaculate volume, sperm concentration, and sperm motility. Furthermore, 10 haplotypes (H1: -CTCGG, H2: GTGGCA, H3: GCGGCA, H4: GCTGCA, H5: GCTCGA, H6: GTGGGG, H7: GCTCCG, H8: -CGGGA, H9: GCGGCG, and H10: GTTGCA) were formed by the six SNPs through linkage disequilibrium analysis, and then 14 different combined haplotypes were collected. Correlation analysis showed that the combined H1H2 haplotype had the highest genotype frequency. Notably, the combined H1H2 haplotype had low sperm concentration, low sperm motility, and high sperm abnormality. The combined H2H3 haplotype could be used as a potential molecular marker for selecting semen with high sperm motility. In general, we illustrated a significant correlation between SNPs in *ARID4A* and sperm quality traits of Chinese water buffalo, which may be useful in the marker-assisted selection of buffalo breeding. This study was the first to analyze the genetic polymorphisms of *ARID4A* and association with sperm qualities of Chinese buffalo.

Key words: Chinese water buffalo (Bubalus bubalis); ARID4A; sperm quality; single-nucleotide polymorphisms

*Corresponding author:
Linlin Hao, College of Animal Science, Jilin University, 5333 Xi'an Road, Changchun, Jilin 130062, China, E-mail: haolinlin@jlu.edu.cn, Tel: +86 136 3055 2320
Changhong Li, College of Life Sciences, Baicheng Normal University, Baicheng, Jilin, China, E-mail: 178163246@qq.com, Tel: +86 133 3175 0780
# These authors contributed equally to this work

ISSN 0372-5480
Printed in Croatia

Introduction

Water buffalo can be divided into two types of domestic water buffalo: swamp buffalo and river buffalo (YINDEE et al., 2010). China has the second largest swamp buffalo population and the richest germplasm resources in the world. Buffalo has been a major farm animal in agricultural production (ZHANG et al., 2016). In addition, buffalo milk contains higher milk fat and protein content than dairy cow milk. Buffalo meat is also regarded as a Mediterranean diet because of its high nutritional and organoleptic characteristics (CAI et al., 2017; TAMBURRANO et al., 2019). Thus, enlarging the scale of good buffalo population and strengthening the construction of the breeding core group are important. Compared with other domesticated bovids, buffalo, whose populations are decreasing, has poor reproductive efficiency. Therefore, developing the genetic potential of buffalo reproductive characters is challenging (MOKHBER et al., 2018).

Sperm quality is a primary factor affecting animal production performance. The selection of high-quality buffalo semen is the prerequisite for assisted reproduction, in which ejaculate volume, sperm concentration, sperm malformation rate, and sperm motility are important parameters for sperm quality evaluation (CHENG et al., 2017; WIEDEMANN et al., 2018). Furthermore, semen quality evaluation can greatly improve the fertility and economics of buffalo herds.

ARID4A is a homologous member of *DRID* family genes, also known as retinoblastoma-binding protein 1, which regulates cell proliferation by directly binding related proteins (CAJUSO et al., 2014). Moreover, it can regulate Sertoli cell function and male fertility by enhancing the androgen receptor (ROTGERS et al., 2014; WANG et al., 2017; WU et al. 2013). Therefore, the *ARID4A* gene can be used as a candidate gene to evaluate sperm quality.

With the development of molecular genetic technology, the method of screening the polymorphic loci of candidate genes has attracted considerable attention in genetics and breeding. In determining the genetic susceptibility of idiopathic male infertility, the previous study investigated the association between the four SNPs in DNA mismatch repair genes and idiopathic male infertility. Furthermore, its genetic etiology was investigated on the basis of the interactions of SNPs (ZHAO et al. 2019). Eight SNPs of the *GnRHR* gene were detected in buffalo, which indicated that the polymorphic loci of *GnRHR* can affect sperm quality (WANG et al., 2017). Another study has demonstrated that SNPs in the *ARID4A* gene promoter region could affect semen quality of bulls (YANG et al., 2018). However, few studies have reported the *ARID4A* gene as a molecular marker of sperm quality in Chinese water buffalo.

This study aimed to assess the association between SNPs of the *ARID4A* gene with semen quality of Chinese water buffalo and identify the molecular markers of sperm quality, which can provide a theoretical guidance for buffalo breeding.

Material and methods

Animal blood collection and isolation of DNA. In this research, 156 normal mature water buffaloes were selected from Livestock and Poultry Breed Improvement Station of Guangxi Zhuang Autonomous Region. Five milliliters of blood obtained from the jugular vein of each buffalo (n=156) was collected aseptically in a tube containing EDTA dipotassium salt and stored at 4 °C. The genomic DNA was isolated from buffalo blood samples in accordance with the standard TIANamp Genomic DNA kit protocol and then completely dissolved in Tris-EDTA buffer solution. The DNA concentration was measured using a spectrophotometer (Thermo Fisher, Massachusetts, USA, Cat. #NanoDrop8000) (SASAKI et al., 2019). Blood collection of water buffaloes was performed by a well-trained technician from Breed Improvement Station of Guangxi Zhuang Autonomous Region to minimize subjective variation. In addition, all related animal experiments were approved by the Ethics Committee on the Use and Care of Animals of Jilin University (Changchun, China).

Semen sample collection and sperm quality trait assessment. Our laboratory primarily focused on the basic phenotypes of semen samples,

including sperm volume per ejaculate, post-thaw cryopreserved sperm motility (%), abnormal rate of sperm, and sperm concentration ($\times 10^8$/mL). These phenotypes were all determined in accordance with the guidelines of the World Health Organization. Semen samples in accordance with the Frozen Bovine Semen standard (GB/T 4143-2008, China) were collected by professionals. The semen samples were collected from each water buffalo in 3 to 6-day intervals using an artificial vagina. The ejaculates were preserved in a 37 °C water bath before assessing the quality traits. The ejaculate volume was evaluated by collecting sperm into a sperm collection bottle and measuring its capacity. The sperm concentration was measured using a Sefi Medical Instruments-Makler. The abnormality rate of sperm was studied by Giemsa staining of fixed smears (KOVACS and FOOTE, 1992). In addition, stained slides were examined using a light microscope at 400× and 1000× magnification to determine sperm morphology, and then the abnormality rate of sperm was calculated. Fresh sperm was diluted with glycerol egg-yolk citrate and collected in straws, and then the sample was stored in liquid nitrogen for 5–7 days. Afterward, we randomly obtained two straws from each buffalo semen sample, which was thawed at 37 °C for 20 s. The post-thaw sperm motility was subjectively observed under a phase-contrast microscope with a heating plate (37 °C) at 400× magnification. The collection of sperm samples and the assessment of sperm quality traits of water buffaloes were conducted by a well-trained technician to minimize subjective variation.

Genomic DNA pool construction and ARID4A gene SNP screening. Based on the NCBI and Ensembl Databank, the potential regions with SNP were obtained from the published reference sequence of cattle *ARID4A* gene (Genbank NC _037555.1). Five pairs of primers (Table 1) were designed to amplify the potential regions of SNPs by Primer Premier 5.0 (Premier Biosoft International, USA). Equal amounts of genomic DNA were collected from each water buffalo and mixed, which is known as DNA pool, and this mixture served as the polymerase chain reaction (PCR) template. PCR was performed in a 20 μL reaction mixture consisting of 60 ng of template DNA, 0.5 μL of forward and 0.5 μL of reverse primer (10 μM), 2 μL of 10 ×PCR buffer, 2 μL of 10 mM dNTPs, and 1 unit of Taq polymerase (TakaRa Biotechnology, Dalian, China). A PCR program was set in accordance with the following conditions: 10 min pre-denaturation at 95 °C, followed by 30 cycles of 95 °C for 30 s, annealing at 56 °C (an optimized temperature) for 30 s, extending at 72 °C for 90 s; final extension at 72 °C for 10 min. The PCR products were separated using 1% agarose gel electrophoresis after staining with 0.5 μg/mL of ethidium bromide. The gels were visualized and documented using a Gel documentation system. Then, amplified PCR products of the DNA pool were sequenced by Comate Bioscience Co., Ltd. (Jilin, China). The amplifications were conducted in a T100™ Thermal Cycler (Bio-Rad, USA), and Geneious bioinformatics 9.1.5 was used (Biomatters Ltd., Auckland, New Zealand) to identify SNPs in the amplified fragments.

Table 1. Primer pair used for the AT-rich interaction domain 4A(A*RID4A)* gene amplification

Primers	Sequences (5' -3')	Position in gene†	Product lengths (bp)	Annealing temperature (°C)
ARID4A 1	F: AACTTTTCATTGAGACTTGGGGTG	3070~3093	350	59
	R: AGTTCCTAATGTGCCTTTCCTACAG	3395~ 3419		
ARID4A 2	F: AAGGAAGTAGACATTCTCAATTTGC	20330~20354	1354	56
	R: TGGGAAGCTCCTGTAATCTCTTAC	21660~21683		
ARID4A 3	F: ACACCAAGAAACCAGTAACAGG	33195~33216	510	56
	R: TAGCCTGGCTGAAGCATAAC	33684~33703		

Table 1. Primer pair used for the AT-rich interaction domain 4A(A*RID4A)* gene amplification (continued)

ARID4A 4	F: TCCTGTTCATCTGATAGTGAAACAG	48123~48147	1234	60
	R: AGGGTACTTAACAGATCAAAGCTCC	49332~49356		
ARID4A 5	F: TATAGATCAGAGATGCTGTATGTTG	55118~55142	1050	58
	R: TGAATAGTGAGAGATATTCAGAGGG	56143~56167		

F was upstream primer. R was downstream primer.
†Based on reference sequence of *ARID4A* gene

SNP genotyping assay. PCR was performed with each water buffalo genomic DNA using the abovementioned method to determine the genotypes of SNPs. The sequences of the *ARID4A* gene fragment were analyzed on the basis of sequencing alignment and screening overlapping peaks. Each buffalo was evaluated for genetic diversity parameters such as genotype frequency and allelic frequency, which were measured directly on the basis of the genotyping results. The genetic diversity parameters such as gene heterozygosity [He], gene homozygosity [Ho], and polymorphism information content (PIC) were calculated using a PIC calculator (PIC_CALC).

Correlation and haplotype analysis. The correlation of SNPs in buffalo *ARID4A* gene was tested using ANOVA in SPSS v19.0 (IBM, USA). The genotype frequency distribution of each SNP was detected for deviation from Hardy–Weinberg equilibrium (HWE, $P<0.05$). Under the premise of HWE, linkage disequilibrium (LD) and haplotypes were verified using HaploView software 3.32 (Daly Lab at the Broad Institute Cambridge, USA, ver. 3.32).

Statistical analysis. Phenotypic dates of every sperm quality trait were collected in triplicate, and each value represents the average of at least three measurements. ANOVA was used to validate the relationship between SNP loci and sperm quality. P value<0.05 was considered significant, and all the data were presented as means±SEM.

Results

SNP detection of ARID4A gene in water buffalo. Compared with the reference sequence, 14 SNPs (g.21124Indel G, g.21188C>T, g.21190G>T, g.21192G>C, g.21285C>G, g.21364A>G, g.21536A>G, g.33314G>C, g.33380G>C, g33619G>A, g.55307G>A, g.55315A>G, g.55316G>T, and g.55851G>C) were detected during DNA pool of 156 water buffaloes (Figure 1), which were all novel in GenBank. Among these SNPs, seven were located in intron 11–12, whose overlapping peaks were dense and evident, and such SNPs may have a concentrated and intense function. Therefore, we focused on the seven SNPs (g.21124Indel G, g.21188C>T, g.21190G>T, g.21192G>C, g.21285C>G, g.21364A>G, and g.21536A>G) at intron 11–12 (Figure 2).

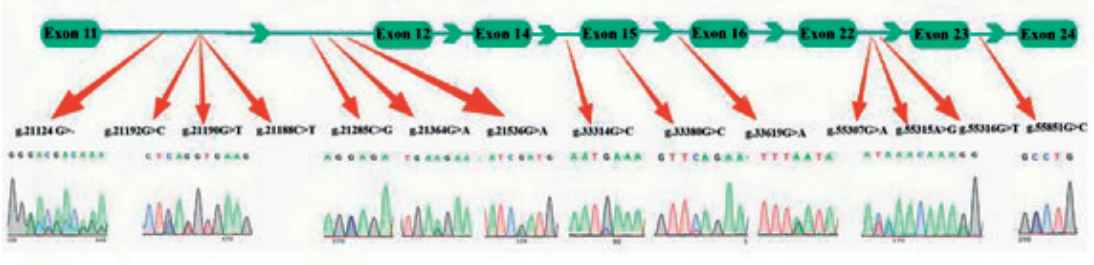

Fig. 1. SNPs located in *ARID4A* gene of water buffalo

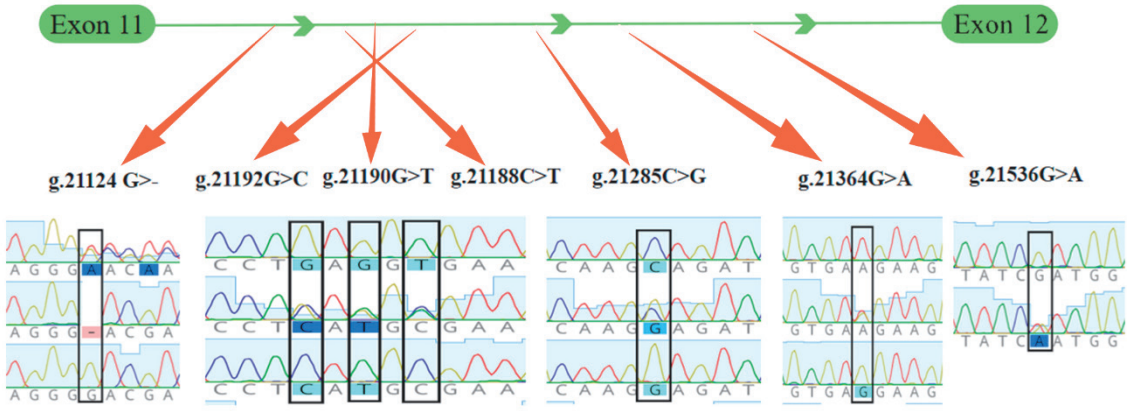

Fig. 2. SNPs located in intron 11-12 *ARID4A* gene of water buffalo.

SNP genotyping and diversity analysis. The genotype of each individual was detected, and the genotype frequency of each buffalo was measured directly on the basis of individual sequencing data (Table 2). PIC values of the seven SNPs are also summarized (Table 2), of which g.21536A>G has a PIC value of 0.287. This result shows a medium polymorphic frequency based on the PIC value (when the PIC value is between 0.25 and 0.5). The PIC values of the other six SNPs were 0.532, 0.551, 0.517, 0.519, 0.528, and 0.537, all of which had high polymorphisms (a PIC value>0.5 represents a high polymorphism frequency). Meanwhile, the values of Ho (gene homozygotes) and He (gene heterozygotes) were all around 0.5 in this study, indicating that the Chinese water buffalo samples selected in this study had a certain extent of polymorphisms.

Table 2. Genetic diversity parameters of AT-rich interaction domain 4A (*ARID4A)* gene

SNP	Genotype/ samples	Genotypic frequency (%)	Allele	Allelic frequency (%)	†PIC[a]	‡Ho[b]	§He[c]	Genbank accession No.
g.21124Indel G	GG/56 -/ -/ 24 G/-/76	35.90 15.38 48.72	I¶ -	60.53 39.47	0.532	0.51	0.49	Novel
g.21188C>T	CC/60 TC/68 TT/28	38.46 43.59 17.95	C T	60.26 39.74	0.551	0.56	0.44	Novel
g.21190G>T	GT/88 TT/28 GG/40	56.41 17.95 25.64	G T	53.85 46.15	0.517	0.44	0.56	Novel
g.21192G>C	GG/48 GC/84 CC/24	30.77 53.85 15.38	G C	57.69 42.31	0.519	0.46	0.54	Novel

Table 2. Genetic diversity parameters of AT-rich interaction domain 4A (*ARID4A)* gene (continued)

g.21285C>G	CC/44 GC/84 GG/28	28.20 53.85 17.95	C G	55.13 44.87	0.528	0.46	0.54	Novel
g.21364A>G	AA/48 AG/80 GG/28	30.77 51.28 17.95	A G	56.41 43.59	0.537	0.49	0.51	Novel
g.21536A>G	GG/84 AG/68 AA/0	55.26 44.74 0	A G	22.37 77.63	0.287	0.55	0.45	Novel

Abbreviations:[†]PIC[a]: polymorphism information content; [‡]Ho[b]: gene homozygosity, [§]He[c]: gene heterozygosity

Association studies of SNPs and sperm quality traits. The correlation between the seven SNPs in the *ARID4A* gene and semen quality traits was analyzed by the Chi-squared test, and the results are shown in Table 3. Determining whether each SNP locus is associated with at least one semen quality is not difficult. The results indicated that the g.21124 Indel G of Chinese water buffaloes with GG genotype had higher ejaculate volume (9.82±0.50, P<0.05), higher sperm concentration (12.53±0.25, P<0.05), and higher post-thaw sperm motility (41.36±0.28, P<0.05) than that of the other genotypes, whereas that of buffaloes with G-genotype had higher sperm abnormalities (15.58±0.35, P<0.05). For g.21188C>T, the buffaloes with TC genotype had higher sperm abnormalities (15.17±0.36, P<0.05) than those with CC and TT genotype, but no significant difference was found between this SNP and the other traits. For g.21190G>T, GG genotype was the dominant genotype of high post-thaw sperm motilities (41.30±0.36, P<0.05), whereas GT genotype was the dominant genotype of high sperm abnormalities (15.00±0.33, P<0.05). For g.21192G>C, when the genotype was GG, the ejaculate volume (9.83±0.58, P<0.05) and sperm concentration (12.64±0.29, P<0.05) were the highest, whereas GC genotype had the highest sperm abnormalities (15.48±0.33, P<0.05). For g.21285C>G, when the genotype was GG, the ejaculate volume (7.14±0.82, P<0.05) and sperm concentration (11.35±0.48, P<0.05) were the lowest, whereas the GC genotype had the highest sperm abnormalities. For g.21364A>G, the buffaloes with AA genotype had the highest ejaculate volume (9.29±0.58, P<0.05), sperm concentration (12.94±0.23, P<0.05), and post-thaw sperm motilities (41.58±0.27, P<0.05), whereas the buffaloes with AG genotype had the highest sperm abnormalities (15.05±0.36, P<0.05). For g.21536A>G, the ejaculate volume (8.95±0.34, P<0.05) and sperm concentration (12.74±0.20, P<0.05) of water buffaloes with AG genotype were higher than those with GG genotype, whereas buffaloes with homozygous AA genotype were not found.

Haplotype analysis. Among the 156 water buffalo samples, seven SNP genotypes were loaded into HaploView software 3.32 to calculate their LD relationships, of which g.21536A>G was not following the Hardy–Weinberg Law, and the other six SNPs formed a strong linkage block (Figure 3). The block generated 10 haplotypes: H1: -CTCGG, H2: GTGGCA, H3: GCGGCA, H4: GCTGCA, H5: GCTCGA, H6: GTGGGG, H7: GCTCCG, H8: -CGGGA, H9: GCGGCG, and H10: GTTGCA. The 10 haplotypes combined a total of 14 genotypes (Table 4). Among these genotypes, H1H2 (30.77%), HIH1 (15.38%), H2H2 (12.82%), H1H3 (12.82%), and H2H3 (5.13%) had higher genotype frequency. We focused on these five genotypes. Chi-squared test analysis indicated that the buffaloes with combined H2H2 genotypes had the highest sperm concentration (12.83±0.38), and the buffaloes with combined H1H2 genotypes had the highest sperm abnormalities (15.67±0.46). In addition, H2H3 had the highest post-thaw sperm

motilities (43.00±0.16), but no evident difference in ejaculate volume was found among the assorted haplotypes. Considering that H2H2, H1H2, and H2H3 could affect sperm quality traits, we further investigated whether the haplotypes of the three genotypes (H2H2, H1H2, and H2H3) will affect gene structures.

Table 3. Least squares mean (LSM) and standard errors (SE) for sperm quality traits of different genotypes in 156 Chinese water buffaloes

SNP	Genotype/ Samples	Ejaculate volume (mL)	Sperm concentration ($\times 10^8$/mL)	Post-thaw sperm motilities (%)	sperm abnormalities (%)
g.21124IndelG	GG/56	9.82±0.50^{b}	12.53±0.25^{b}	41.36±0.28^{b}	12.64±0.29^{b}
	-/ -/ 24	7.25±0.95^{a}	11.49±0.55^{a}	40.50±0.23^{ab}	11.50±0.20^{a}
	G/-/76	7.23±0.34^{a}	11.78±0.26^{ab}	40.00±0.41^{a}	15.58±0.35^{c}
g.21188C>T	CC/60	8.11±0.59	12.20±0.33	40.67±0.27	13.40±0.36^{b}
	TC/68	8.19±0.40	11.67±0.24	40.17±0.44	15.17±0.36^{c}
	TT/28	8.20±0.53	12.42±0.36	41.28±0.42	11.86±0.36^{a}
g.21190G>T	GT/88	8.12±0.42	11.84 ±0.22	40.5±0.35^{ab}	15.00±0.33^{b}
	TT/28	7.24±0.82	12.02±0.54	39.71±0.42^{a}	12.43±0.47^{a}
	GG/40	8.9±0.44	12.37±0.33	41.30±0.36^{b}	12.50±0.31^{a}
g.21192G>C	GG/48	9.83±0.58^{b}	12.64±0.29^{b}	41.25±0.30	12.33±0.28^{a}
	GC/84	7.47±0.32^{a}	11.79±0.23^{ab}	40.19±0.39	15.48±0.33^{b}
	CC/24	7.25±0.95^{a}	11.49±0.55^{a}	40.50±0.23	11.50±0.20^{a}
g.21285C>G	CC/44	9.85±0.63^{b}	12.93±0.28^{b}	41.00±0.31	12.55±0.28^{a}
	GC/84	7.62±0.32^{a}	11.74±0.24^{a}	40.29±0.39	15.14±0.35^{b}
	GG/28	7.14±0.82^{a}	11.35±0.48^{a}	40.71±0.22	12.29±0.41^{a}
g.21364A>G	AA/48	9.29±0.58^{b}	12.94±0.23^{b}	41.58±0.27^{b}	12.92±0.31^{a}
	AG/80	7.67±0.36^{a}	11.60±0.26^{a}	40.00±0.41^{a}	15.05±0.36^{b}
	GG/28	7.64±0.84^{a}	11.57±0.46^{a}	40.43±0.20^{ab}	12.29±0.41^{a}
g.21536A>G	GG/84	7.30±0.46^{a}	11.55±0.27^{a}	41.05±0.24^{ab}	14.10±0.33^{ab}
	AG/68	8.95±0.34^{b}	12.74±0.20^{b}	40.12±0.44^{ab}	13.76±0.38^{ab}
	AA/0	-	-	-	-

Unless indicated otherwise, data are presented as means±SEM.
Values with a, b superscripts within the same row of the same locus denote significant difference (<0.05), the values with ab denote have no significant difference with the values of a and b (>0.05).

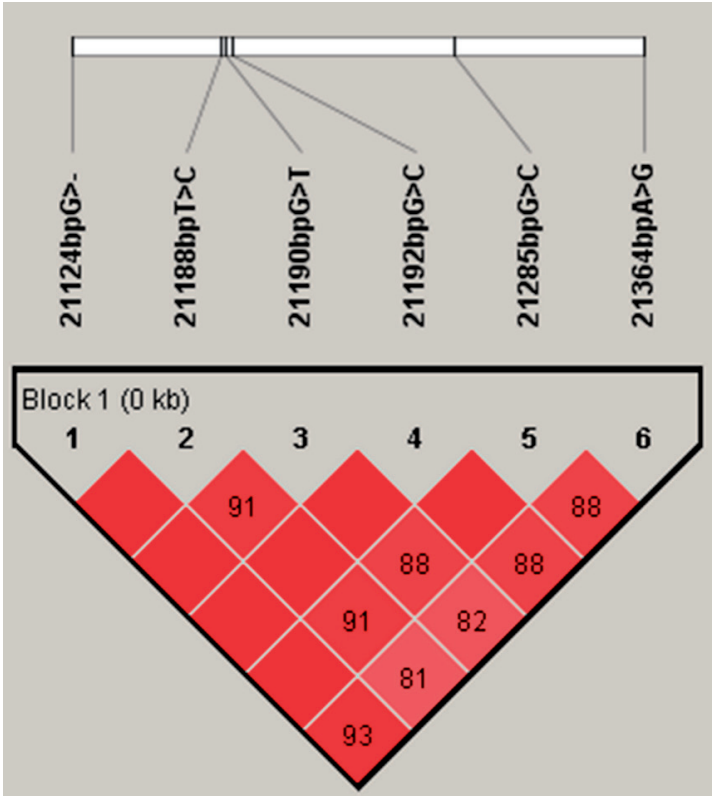

Fig. 3. Single nucleotide polymorphisms (SNPs) and Linkage disequilibrium (LD) pattern of *ARID4A* gene in the Chinese water buffalo (LD blocks are marked with triangles).

Table 4. Effect of different combinations of three single nucleotide polymorphisms (SNPs) on sperm quality traits in 156 water buffalo

Genotype/ samples	Genotypic frequency(%)	Ejaculate volume(mL)	Sperm concentration ($\times 10^8$/mL)	Post-thaw sperm motilities (%)	Sperm abnormalities (%)
H1H1/24	15.38	7.25 ± 0.95^{a}	11.49 ± 0.55^{abc}	40.50 ± 0.23^{bcd}	11.50 ± 0.20^{ab}
H2H2/20	12.82	7.78 ± 0.72^{a}	12.83 ± 0.38^{bcd}	41.00 ± 0.50^{bcd}	12.60 ± 0.40^{ab}
H1H2/48	30.77	7.33 ± 0.49^{a}	11.64 ± 0.32^{abc}	39.83 ± 0.59^{bc}	15.67 ± 0.46^{cde}
H1H3/20	12.82	7.14 ± 0.58^{a}	11.70 ± 0.49^{abc}	41.00 ± 0.41^{bcd}	14.60 ± 0.66^{bcd}
H3H3/4	2.56	8.80 ± 0.09^{ab}	15.27 ± 0.22^{d}	43.00 ± 0.09^{cd}	11.00 ± 0.29^{a}
H1H4/4	2.56	7.20 ± 0.21^{a}	15.19 ± 0.17^{d}	35.00 ± 0.25^{a}	18.00 ± 0.14^{e}
H2H3/8	5.13	8.10 ± 0.31^{a}	12.22 ± 0.20^{bc}	43.00 ± 0.16^{cd}	13.00 ± 0.28^{abc}

Table 4. Effect of different combinations of three single nucleotide polymorphisms (SNPs) on sperm quality traits in 156 water buffalo (continued)

H2H5/4	2.56	9.60±0.39^{ab}	11.76±0.05^{abc}	44.00±0.44^{d}	12.00±0.34^{ab}
H2H6/4	2.56	9.60±0.27^{ab}	9.36±0.12^{a}	44.00±0.31^{d}	10.00±0.20^{a}
H2H9/4	2.56	12.90±0.23^{bc}	9.22±0.24^{a}	38.00±0.17^{ab}	12.00±0.17^{ab}
H2H10/4	2.56	8.90±0.03^{ab}	13.44±0.16^{cd}	40.00±0.28^{bc}	10.00±0.26^{a}
H3H4/4	2.56	15.35±1.76^{c}	13.98±0.21^{cd}	41.00±0.38^{bcd}	14.50±0.57^{bcd}
H5H8/4	2.56	6.50±±0.07^{a}	10.50±0.13^{ab}	42.00±0.22^{cd}	17.00±0.33^{de}
H6H7/4	2.56	10.00±0.10^{ab}	12.07±0.26^{bc}	40.00±0.13^{bc}	17.00±0.12^{de}

Unless indicated otherwise, data are presented as means±SEM.
† Values with a, b superscripts within the same row in the same locus denote significant difference (p<0.05), the values with ab denote have no significant difference with the values with both a and b (p>0.05). H1: -CTCGG, H2: GTGGCA, H3: GCGGCA, H4: GCTGCA, H5: GCTCGA, H6: GTGGGG, H7: GCTCCG, H8: -CGGGA, H9: GCGGCG, and H10: GTTGCA

Effect of ARID4A haplotypes on gene structure. We predicted the DNA folding form and minimum free energy of the fragment (from 21104 to 21384 nt) containing six SNPs by using the Mfold Web server (http://unafold.rna.albany.edu/?q=mfold/DNA-Folding-Form) to study the effect of *ARID4A* haplotypes on gene structure. The results indicated that the DNA structure of H1, H2, and H3 haplotypes changed because of the mutation and the minimum free energy. For example, the minimum free energy of H3 (−26.15 kcal/mol) increased by 1.41 kcal compared with H2 (−24.74 kcal/mol). Furthermore, the minimum free energy of H3 was more stable than that of H2 (Figures 4A, B, and C).

Discussion

Buffalo is a large-scale economic animal in the tropical and subtropical regions. It has strong physical strength, gentle temperament, resistance to rough feeding, strong disease resistance, and long service life. Although reproductive biotechnology methods have been applied to buffalo, most methods are not as effective as in bovine (PARNPAI et al., 2016). Despite the widespread use of artificial insemination (AI), semen characteristics determine the conception rate (MATHEVON et al., 1998). Therefore, identifying the molecular markers related to semen quality is necessary to improve the breeding level of water buffalo and enlarge the scale of good buffalo population.

Semen quality indicators are used as surrogate indicators of male fertility in clinical androgenology, reproductive toxicology, epidemiology, and risk assessment (DRUET et al., 2009; ZINAMANN et al., 2000). At present, a large number of studies have been conducted to explore the relationship between genetic polymorphism and semen quality (FORTES et al., 2013; HERING et al., 2014), which contribute to the identification of genetic markers and provide comprehensive understanding of genomic selection strategies to assist cattle breeding by improving fertility.

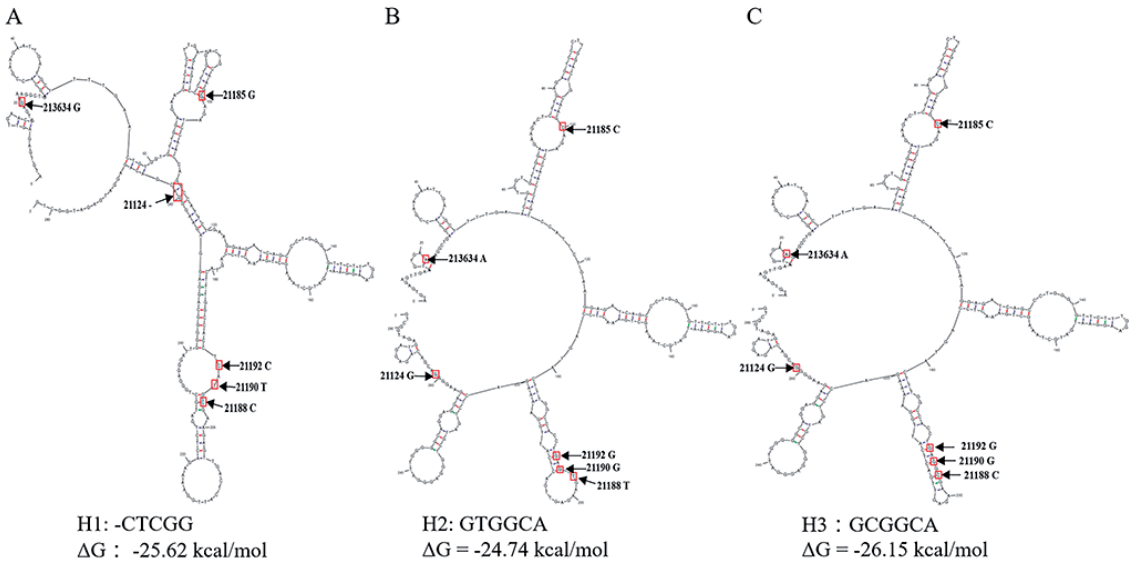

Fig. 4. Gene structure prediction. (A, B and C) gene structure and minimum free energy of the fragment (from 21104 to 21384 nt) containing the six SNPs of *ARID4A* gene in the Chinese water buffalo. The red rectangle contains six SNPs.

ARID4A is functionally associated with the AR and RB pathways that regulate the meiosis of spermatogenic, which may also have an impact on male fertility. ARID4A is primarily expressed in the testis, and *ARID4A* knockout mice showed spermatogenic stasis during meiosis of meiotic spermatocytes and postmeiotic haploid spermatids and progressive loss of male fertility (WU et al., 2013). Therefore, the *ARID4A* gene could be used as a candidate gene to evaluate sperm quality traits. The SNP within the intron region of the Cluster-of-differentiation antigen 9 gene was significantly associated with sperm quality. Furthermore, SNP may be closely associated with the expression of the gene (GUNAWAN et al., 2011; KAEWMALA et al., 2011).

Seven SNPs were detected to identify the association of SNP-assisted markers with sperm quality traits in *ARID4A*. All detected SNPs with high polymorphism were located in intron 11–12 of the *ARID4A* gene. Association studies were performed to identify the genotypes of the seven SNPs using certain semen quality data (semen volume, concentration, motility, and deformity) of 156 Chinese water buffaloes. The results revealed that these polymorphisms had a large number of genetic variations and high selection potential. At present, these polymorphism loci were all novel SNPs, indicating that research about polymorphism in water buffalo is rare. Association analysis revealed that these seven SNPs had a significant correlation with at least one sperm quality trait in Chinese water buffalo. In addition, these seven SNPs were all correlated with ejaculate volume. Therefore, our research results suggest that the *ARID4A* gene is a powerful potential selective marker of Chinese buffalo sperm traits. After comprehensive consideration, the GG genotype of g.21192G>C, the CC genotype of g.21285C>G, and the AA genotype of g.21364A>G could be used as a marker for selecting semen with low sperm abnormality, high ejaculate volume, sperm concentration, and sperm motility.

Based on the abovementioned results, the screened SNPs were located in introns. The ability of introns to stimulate gene expression has been demonstrated in a wide range of organisms. It can not only enhance the expression by affecting the rate of transcription, nuclear export, and stability of the transcript but also improve mRNA translation efficiency (AKUA et al., 2010; SAMADDER et al., 2008). In addition, it may affect the specific activity

of the gene (HUANG et al., 2015). Thus, although SNPs exist in introns, a significant relationship is observed between intron SNPS and traits.

Based on previous reports, the genotypic effect of single SNP was affected by other SNPs, whereas the genotypic combination effect indicated the interactions among SNPs (ZHENG et al., 2011). Inheritance effects of genotype combinations are superior to single SNPs (FALLIN et al., 2001). At present, relationship analysis indicated that the combined H1H2 genotype had the highest genotype frequency, whereas genotypes carrying H1 haplotype, including H1H1, H1H2, and H1H3, had low sperm concentration. This phenomenon suggests that H1 haplotype may lead to an adverse phenotype indicating sperm quality, which should be avoided in future breeding work. In particular, the combined H1H2 genotype had low sperm motility and high sperm abnormality. In addition, the combined H2H3 genotype could be used as a molecular marker for selecting semen with high sperm motility. A SNP was obtained in intron 6, which changed the primary structure of the *CD9* gene (KAEWMALA et al., 2011). The role of introns in regulating gene expression levels or tissue-specific expression patterns has been increasingly reported (JIANG et al., 2000; PAGANI and BARALLE, 2004). In this study, the H1, H2, and H3 haplotypes have different gene structures (Figures 4A, B, and C). The change in gene secondary structure will affect the expression of gene and its function. For example, a previous study suggested that different GH gene haplotypes affected the synthesis and secretion of GH protein (CHENG et al., 2016). This study has also a limitation. Semen quality is related not only to genetic factors, but also to physical factors, seasonal factors, and feeding factors. Consequently, further studies are needed to confirm the relationship between semen quality and these factors.

Conclusion

This study illustrates the association between SNPs of the *ARID4A* gene and sperm quality traits in water buffalo. Thus, the use of SNP-assisted markers (particularly the haplotype carrying H1) in predicting sperm quality and selecting buffalo can improve herd fertility and production. *ARID4A* is associated with male fertility, but it is rarely reported. As far as we known, this study is the first to report the relationship between SNPs of the *ARID4A* gene and the quality of Chinese buffalo sperm. These results provide a theoretical basis for the selection of high-quality genetic resources in Chinese water buffalo.

Acknowledgments
This study was supported by the Education Department of Jilin Province (JJKH20210008KJ).

Conflict of interest
We certify that there is no conflict of interest with any financial organization regaing the material discussed in the manuscript.

Author contributions
Linlin Hao and Changhong Li conceived and designed the experiments. Guanhong Lu, Chuanyu He, Xi Yang, Ze Ma and Jiapeng Li and Tianqi Feng performed experiments. Yunyun Cheng, Siyao Wang, Guanhong Lu, Chunli Wang, Rui Yang and Jie Song assessed the experiments and provided the data analysis. Guanhong Lu and Linlin Hao wrote the manuscript. All authors read and approved the manuscript.

Data availability
The data used to support the findings of this study are available from the corresponding author upon request.

References

AKUA, T., I. BEREZIN, O. SHAUL (2010): The leader intron of AtMHX can elicit, in the absence of splicing, low-level intron-mediated enhancement that depends on the internal intron sequence. BMC Plant Biol. 10, 93.
DOI: 10.1186/1471-2229-10-93

CAI, X., Q. LIU, X. ZHANG, Y. REN, X. LEI, S. LI, D. SHI (2017): Identification and analysis of the expression of microRNA from lactating and nonlactating mammary glands of the Chinese swamp buffalo. J Dairy Sci. 100, 1971-1986.
DOI:10.3168/jds.2016-11461

CAJUSO, T., U. A. HANNINEN, J. KONDELIN, A. E. GYLFE, T. TANSKANEN, R. KATAINEN, P. VAHTERISTO (2014): Exome sequencing reveals frequent inactivating mutations in ARID1A, ARID1B, ARID2 and ARID4A in microsatellite unstable colorectal cancer. Int J Cancer 135, 611-623.
DOI:10.1002/ijc.28705

CHENG, Y., J. GU, H. XUE, Q. LI, M. LIANG, N. WANG, L. HAO (2016): Identification of Four SNPs in LHB Gene and Their Associations with Sperm Qualities of Chinese

Buffaloes. Anim Biotechnol 28, 168-173.
DOI:10.1080/10495398.2016.1244068

CHENG, Y., S. LIU, C. LU, Q. WU, S. LI, H. FU, G. WANG, C. LV, L. NIE, Y. ZHANG, H. YU, L. HAO (2016): Missense mutations in the signal peptide of the porcine GH gene affect cellular synthesis and secretion. Pituitary, 19, 362-369.
DOI: 10.1007/s11102-016-0713-6.

DRUET, T., S. FRITZ, E. SELLEM, B. BASSO, O. GERARD, L. SALAS-CORTES, A. EGGEN (2009): Estimation of genetic parameters and genome scan for 15 semen characteristics traits of Holstein bulls. J Anim Breed Genet. 126, 269-277.
DOI:10.1111/j.1439-0388.2008.00788.x

FALLIN, D., A. COHEN, L. ESSIOUX, I. CHUMAKOV, M. BLUMENFELD, D. COHEN, N. J. SCHORK (2001): Genetic analysis of case/control data using estimated haplotype frequencies: application to APOE locus variation and Alzheimer's disease. Genome Res. 11, 143-151.
DOI:10.1101/gr.148401

FORTES, M. R., A. REVERTER, M. KELLY, R. MCCULLOCH, S. A. LEHNERT (2013): Genome-wide association study for inhibin, luteinizing hormone, insulin-like growth factor 1, testicular size and semen traits in bovine species. Andrology 1, 644-650.
DOI:10.1111/j.2047-2927.2013.00101.x

GUNAWAN, A., K. KAEWMALA, M. J. UDDIN, M. U. CINAR, D. TESFAYE, C. PHATSARA, K. SCHELLANDER (2011): Association study and expression analysis of porcine ESR1 as a candidate gene for boar fertility and sperm quality. Anim Reprod Sci. 128, 11-21.
DOI:10.1016/j.anireprosci.2011.08.008

HERING, D. M., K. OLENSKI, A. RUSC, S. KAMINSKI (2014): Genome-wide association study for semen volume and total number of sperm in Holstein-Friesian bulls. Anim Reprod Sci. 151, 126-130.
DOI:10.1016/j.anireprosci.2014.10.022

HUANG, M., H. M. WANG, Y. GUO, J. PING, M. CHEN, D. XU, H. WANG (2015): Single nucleotide polymorphism of CYP3A4 intron 2 and its influence on CYP3A4 mRNA expression and liver enzymatic activity in human liver. J Huazhong Univ Sci Technolog Med Sci. 35, 502-507.
DOI:10.1007/s11596-015-1460-z

JIANG, Z., J. COTE, J. M. KWON, A. M. GOATE, J. Y. WU (2000): Aberrant splicing of tau pre-mRNA caused by intronic mutations associated with the inherited dementia frontotemporal dementia with parkinsonism linked to chromosome 17. Mol Cell Biol. 20, 4036-4048.
DOI:10.1128/mcb.20.11.4036-4048.2000

KAEWMALA, K., M. J. UDDIN, M. U. CINAR, C. GROSSE-BRINKHAUS, E. JONAS, D. TESFAYE, K. SCHELLANDER (2011): Association study and expression analysis of CD9 as candidate gene for boar sperm quality and fertility traits. Anim Reprod Sci, 125, 170-179.
DOI:10.1016/j.anireprosci.2011.02.017

KOVACS, A., R. H. FOOTE (1992): Viability and acrosome staining of bull, boar and rabbit spermatozoa. Biotech Histochem 67, 119-124.
DOI:10.3109/10520299209110020

MATHEVON, M., M. M. BUHR, J. C. DEKKERS (1998): Environmental, management, and genetic factors affecting semen production in Holstein bulls. J Dairy Sci. 81, 3321-3330.
DOI:10.3168/jds.S0022-0302(98)75898-9

MOKHBER, M., M. MORADI-SHAHRBABAK, M. SADEGHI, H. MORADI-SHAHRBABAK, A. STELLA, E. NICOLZZI, J. L. WILLIAMS (2018): A genome-wide scan for signatures of selection in Azeri and Khuzestani buffalo breeds. BMC Genomics 19, 449.
DOI:10.1186/s12864-018-4759-x

PAGANI, F., F. E. BARALLE (2004): Genomic variants in exons and introns: identifying the splicing spoilers. Nat Rev Genet. 5, 389-396.
DOI:10.1038/nrg1327

PARNPAI, R., Y. LIANG, M. KETUDAT-CAIRNS, T. SOMFAI, T. NAGAI (2016): Vitrification of buffalo oocytes and embryos. Theriogenology 86, 214-220.
DOI:10.1016/j.theriogenology.2016.04.034

ROTGERS, E., A. RIVERO-MULLER, M. NURMIO, M. PARVINEN, F. GUILLOU, I. HUHTANIEMI, J. TOPPARI (2014): Retinoblastoma protein (RB) interacts with E2F3 to control terminal differentiation of Sertoli cells. Cell Death Di. 5, e1274.
DOI:10.1038/cddis.2014.232

SAMADDER, P., E. SIVAMANI, J. LU, X. LI, R. QU (2008): Transcriptional and post-transcriptional enhancement of gene expression by the 5' UTR intron of rice rubi3 gene in transgenic rice cells. Mol Genet Genomics 279, 429-439.
DOI:10.1007/s00438-008-0323-8

SASAKI, S., E. MURAKI, Y. INOUE, R. SUEZAWA, H. NIKADORI, Y. YOSHIDA, T. KOJIMA (2019): Genotypes and allele frequencies of buried SNPs in a bovine single-nucleotide polymorphism array in Japanese Black cattle. Anim Sci J. 90, 1503-1509.
DOI:10.1111/asj.13293

TAMBURRANO, A., B. TAVAZZI, C. A. M. CALLA, A. M. AMORINI, G. LAZZARINO, S. VINCENTI, P. LAURENTI (2019): Biochemical and nutritional characteristics of buffalo meat and potential implications on human health for a personalized nutrition. Ital J Food Saf. 8, 8317.
DOI:10.4081/ijfs.2019.8317

WANG, G., L. HAO, Y. CHENG, S. LI, Y. ZHANG, C. LV, J. ZHANG (2017): Effects of GnRHR polymorphisms on sperm quality in Chinese water buffalo. Anim Reprod Sci. 186, 37-43.
DOI:10.1016/j.anireprosci.2017.09.001

WANG, H., Z. DING, Q. M. SHI, X. GE, H. X. WANG, M. X. LI, L. C. XU (2017): Anti-androgenic mechanisms of Bisphenol A involve androgen receptor signaling pathway. Toxicology 387, 10-16.
DOI:10.1016/j.tox.2017.06.007

WIEDEMANN, I., A. MAEHLMEYER, S. JANSEN, A. R. SHARIFI, C. KNORR (2018): SNP g.1007A>G within the porcine DNAL4 gene affects sperm motility traits and percentage of midpiece abnormalities. Reprod Domest Anim. 53, 401-413.
DOI:10.1111/rda.13120

WU, R. C., M. JIANG, A. L. BEAUDET, M. Y. WU (2013): ARID4A and ARID4B regulate male fertility, a functional link to the AR and RB pathways. Proc Natl Acad Sci U S A 110, 4616-4621.
DOI:10.1073/pnas.1218318110

YANG, C., J. WANG, J. LIU, Y. SUN, Y. GUO, Q. JIANG, C. WANG (2018): Functional haplotypes of ARID4A affect promoter activity and semen quality of bulls. Anim Reprod Sci. 197, 257-267.
DOI:10.1016/j.anireprosci.2018.08.038

YINDEE, M., B. H. VLAMINGS, W. WAJJWALKU, M. TECHAKUMPHU, C. LOHACHIT, S. SIRIVAIDYAPONG, J. A. LENSTRA (2010): Y-chromosomal variation confirms independent domestications of swamp and river buffalo. Anim Genet. 41, 433-435.
DOI:10.1111/j.1365-2052.2010.02020.x

ZHANG, Y., Y. LU, M. YINDEE, K. Y. LI, H. Y. KUO, Y. T. JU, J. S. BARKER (2016): Strong and stable geographic differentiation of swamp buffalo maternal and paternal lineages indicates domestication in the China/Indochina border region. Mol Ecol. 25, 1530-1550.
DOI:10.1111/mec.13518

ZHAO, X., C. MU, J. MA, X. DAI, H. JIAO (2019): The association of four SNPs in DNA mismatch repair genes with idiopathic male infertility in northwest China. Int. J Immunogenet 46, 451-458.
DOI:10.1111/iji.12448

ZHENG, X., Z. JU, J. WANG, Q. LI, J. HUANG, A. ZHANG, C. WANG (2011): Single nucleotide polymorphisms, haplotypes and combined genotypes of LAP3 gene in bovine and their association with milk production traits. Mol Biol Rep. 38, 4053-4061.
DOI:10.1007/s11033-010-0524-1

ZINAMAN, M. J., C. C. BROWN, S. G. SELEVAN, E. D. CLEGG (2000): Semen quality and human fertility: a prospective study with healthy couples. J Androl. 21, 145-153.

Received:18 September 2021
Accepted: 20 December 2021

LU, G., C. HE, X. YANG, Y. CHENG, R. YANG, J. SONG, S. WANG, C. WANG, Z. MA, J. LI, T. FENG, L. HAO, C. LI: Povezanost polimorfizma pojedinačnog nukleotida gena *ARID4A* i kvalitete sperme kineskog vodenog bivola. Vet. arhiv 92, 577-590, 2022.

SAŽETAK

Gen *ARID4A* (engl *AT-rich interaction domain 4A*) usko je povezan s kvalitetom sperme. U ovom je radu istraživana povezanost polimorfizma gena *ARID4A* u kineskih vodenih bivola (*Bubalus bubalis*) s kvalitetom sperme, uključujući volumen ejakulata, koncentraciju sperme, pokretljivost spermija nakon odmrzavanja i abnormalnost spermija u sjemenu bivola. U 156 kineskih vodenih bivola otkriveno je sedam polimorfizama pojedinačnog nukleotida (SNPs) gena *ARID4A* Sangerovim sekvenciranjem i identifikacijom preklopljenih vrhova. Među SNP-ovima njih je šest bilo povezano s barem jednim svojstvom kvalitete spermija. Ukratko, g. 21192G>C, g. 21285C>G i g. 21364A>G mogu se upotrijebiti kao potencijalni markeri za selekciju sjemena s niskom abnormalnošću spermija, većim volumenom ejakulata, većom koncentracijom i pokretljivošću spermija. Nadalje, šest SNP-ova formiralo je 10 haplotipova (H1: -CTCGG, H2: GTGGCA, H3: GCGGCA, H4: GCTGCA, H5: GCTCGA, H6: GTGGGG, H7: GCTCCG, H8: -CGGGA, H9: GCGGCG i H10: GTTGCA) analizom povezanosti nepodudarnosti te je ustanovljeno 14 različitih kombiniranih haplotipova. Analiza korelacije pokazala je da kombinirani haplotip H1H2 ima najveću učestalost. Kombinirani haplotip H1H2 imao je najmanju koncentraciju sperme, slabu pokretljivost seprmija i znatnu abnormalnost spermija. Kombinirani haplotip H2H3 može se upotrijebiti kao potencijalni molekularni marker za odabir sjemena s većom pokretljivošću. Općenito je pokazana znakovita korelacija između SNP-ova u *ARID4A* i kvalitete sperme kineskog vodenog bivola, što može biti korisno u selekciji bivola potpomognutoj markerima. Ovo je prvo istraživanje koje je analiziralo genske polimorfizme *ARID4A* i njihovu povezanost s kvalitetom sjemena kineskih vodenih bivola.

Ključne riječi: kineski vodeni bivol (*Bubalus bubalis*); *ARID4A*; kvaliteta spermija; polimorfizmi pojedinačnog nukleotida