ELSEVIER

Available online at www.sciencedirect.com

ScienceDirect

Progress in Natural Science 18 (2008) 1209–1218

Progress in Natural Science

www.elsevier.com/locate/pnsc

b-*DAZL*: A novel gene in bovine spermatogenesis

Qingbo Zhang [a], Qifa Li [a,*], Jiahuang Li [b], Xinfu Li [a], Zhenshan Liu [a], Dawei Song [a], Zhuang Xie [a]

[a] *Laboratory of Animal Reproductive Genetics and Molecular Evolution, College of Animal Science and Technology, Nanjing Agriculture University, Nanjing 210095, China*

[b] *Pharmaceutic Biotechnology Key Laboratory, College of Life Science, Nanjing University, Nanjing 210093, China*

Received 12 January 2008; received in revised form 20 February 2008; accepted 20 February 2008

Abstract

The *DAZL* gene plays an important role in gametogenesis. Microdeletion and mutation of this gene has been frequently associated with the sterility in both vertebrates and invertebrates. Here, a cDNA containing the complete open reading frame of *DAZL* gene, which encodes putative RRM RNA-binding protein, was isolated from bovine tissues by RT-PCR. The isolated gene was referred to as b-*DAZL*. Computer-based location analysis showed that b-*DAZL* is located near the end of chromosome 1. The b-*DAZL* gene and its vertebrate homologue shared a high nucleotide similarity (90–94%) in the coding region and in the genomic organization, each consisting of 10 exons and 9 introns. Their protein products contained a highly conserved RNA-binding motif, a unique DAZ repeat, and PUM2 and PABP interacting domains. RT-PCR assay revealed that b-*DAZL* was expressed in the testis and ovary. However, no signals were detected in the testes of sterile male cattle–yaks. The microstructure of the testes of sterile males showed very few spermatocytes but mostly somatic cells. These observations are consistent with the typical phenotype of testes with defective *DAZL* expression. Subsequent gene detection studies indicated that transcription arrest might lead to a b-*DAZL* transcript defect in testes of cattle–yaks. Our results suggest that b-*DAZL* may be involved in spermatogenesis, and that transcription arrest of the gene is associated with male sterility in cattle–yaks.
© 2008 National Natural Science Foundation of China and Chinese Academy of Sciences. Published by Elsevier Limited and Science in China Press. All rights reserved.

Keywords: Cattle–yak; b-*DAZL*; Male sterility; Sequence analysis; Expression profile

1. Introduction

The yak of *Bovinae* subfamily is well adapted to high-altitude environments at 3500 m above the sea level [1]. For humans inhabiting these regions, yaks are essential for agricultural production and vital for the sustainment of local communities. Unfortunately, pure bred yaks are poor milk and meat producers. To overcome poor production, cross-breeding of cattle and yaks has resulted in a hybrid animal capable of higher milk and meat production. However, the male offspring resulting from a cattle–yak mating is, in fact, sterile. Although many researches have been focusing on explaining reasons for cattle–yak sterility [2–5], little knowledge has been obtained for this genetic causes.

The *Deleted in Azoospermia-Like* (*DAZL*) gene is a member of the *Deleted in Azoospermia* (*DAZ*) family that consists of three genes: *BOULE*, *DAZL* and *DAZ*. The *DAZL* gene and its homologues encode proteins with a highly conserved RNA-binding motif and a unique DAZ repeat, and all show a germ cell-specific expression pattern [6–11]. The DAZL proteins are believed to play a part in the transcriptional regulation of mRNA expression

* Corresponding author. Tel.: +86 25 8439 5046; fax: +86 25 8439 5314.
E-mail address: liqifa@njau.edu.cn (Q. Li).

1002-0071/$ - see front matter © 2008 National Natural Science Foundation of China and Chinese Academy of Sciences. Published by Elsevier Limited and Science in China Press. All rights reserved.
doi:10.1016/j.pnsc.2008.02.012

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

because they bind RNA homopolymers through the RNA-binding motif and form complexes with either themselves or with their paralogues [12–14]. The *DAZL* gene and its homologues are essential for male and female germ cell development in several species. In *Caenorhabditis elegans*, *DAZL* is an essential factor in female meiosis, and plays an important role in the switch from spermatogenesis to oogenesis (sperm/oocyte switch) [9,15]. In *Xenopus*, inhibition of *Xdazl* (*Xenopus DAZ*-like) leads to defective migration and severe reduction of the primordial germ cell (PGC) population [8,12,16]. The *Dazl* null mouse ($Dazl^{-/-}$) is characterized by severe germ cell depletion and meiotic failure [17,18]. This mutant phenotype can, however, be partially rescued by the human *DAZ* gene [19,20]. In addition, it has been found that *DAZL* deficiency leads to embryonic arrest of germ cell development in XY C57BL/6 mice [21]. In the human, *DAZL* is primarily expressed in embryonic stem cells, corpus lutea, early PGCs in embryonic gonads, and throughout gametogenesis [10,21–27]. It was reported that *DAZL* transcripts are lower in men with spermatogenic failure compared to men with normal spermatogenesis [28]. During spermatogenesis, *DAZL* is expressed in the nucleus of spermatogonia and transfer to the cytoplasm of primary spermatocytes at meiosis [27–30]. In female reproduction, the *DAZL* gene is expressed in the cytoplasm of oogonia and mature oocytes [25,26,31,32], and can be detected in granulosa cells, theca interna and the corpus luteum of the human ovary [22,23].

In this study, we identified and characterized the bovine *DAZL* gene, and mapped it to the bovine chromosome and detected its expression pattern in different tissues by RT-PCR method. We also studied spermatogenic arrest and *DAZL* gene deletion in the seminiferous tubules and testes of cattle–yaks. In particular, we investigated the proliferative activity and abnormal morphology of remaining spermatogonia. Furthermore, the analysis of transcriptional arrest of the *DAZL* gene in cattle–yaks was used to provide further explanation for spermatogenesis failure in the bovine.

2. Materials and methods

2.1. Animals

Cattle and female yaks were provided by the Songpan Bovine Breeding Farm, Sichuan Province. We crossed male cattle with female yaks to produce cattle–yak offspring ($n = 12$). Animals were fed maintenance diets for standard growth. At the age of 15 months, cattle–yaks ($n = 12$), cattle ($n = 8$) and yaks ($n = 8$) were sacrificed for tissues. The testes and ovaries taken from the cattle and yaks showed normal gametogenesis. Ovarian, testicular and epididymal tissues were fixed in Bouin's solution for 3–5 h, and sections were made and stained with hematoxylin–eosin and examined by a phase contrast microscope.

2.2. RNA preparation and the first-strand cDNA synthesis

Total RNA from the cattle, yak and cattle–yak tissues (testis, epididymis, hypothalamus, pituitary, kidney, stomach, liver, spleen, ovary, and uterus) was extracted using Trizol reagent according to the manufacturer's protocol (Invitrogen, Shanghai, China). The first-strand cDNA was synthesized with 2 μg of purified total RNA (pre-treated with DNase I) by RT-PCR according to the manufacturer's protocol (Promega, Shanghai, China). Oligo$(dT)_{15}$ was used as a primer and the reverse transcription reaction was incubated at 42 °C for 1 h in a total volume of 20 μl. The reverse transcription products were diluted to a total volume of 40 μl with sterile ddH_2O.

2.3. Molecular cloning

The RRM amino acid sequence with human *DAZL* (GenBank Accession No. NW001351) was used as the query probe to search the Bovine Genome Database, National Centre for Biological Information (NCBI), through the tBLASTn algorithm program. The bovine genomic contigs having higher homology with the probe were obtained for the prediction and assembling of putative ORFs from FGENES-M (http://www.softberry.com) and GENESCAN (http://genes.mit.edu/GENSCAN.html). The sequences of the ORFs were further confirmed by ESTs existing in the vertebrate EST division of GenBank. Based on the assembled results, the oligonucleotide primers were designed and synthesized by SNBC (Shanghai) to allow us to clone the cDNA of *DAZL* from bovine tissues. The following primers were used for cloning: P1-1: 5′-AC CCGCCTCTGACTCTCTCC-3′ and P1-2: 5′-TGGTGGT AGCTGAAGACGACTG-3′; P2-1: 5′-CTCCAAACTCA ACCATCTCC-3′, and P2-2: 5′-ACCGTCTGTATGCTTC TGTCC-3′; P3-1: 5′-TTTGCCCAGTGAATGCTCGGT T-3′ and P3-2: 5′-CCCAAGACAGAAAAGCCAGGAA-3′.

Polymerase chain reaction (PCR) was performed in a 10 μl reaction mixture containing 1 μl 10 × buffer, 1 μl of $MgCl_2$, 1.2 μl of 20 mM dNTP, 1 μl RT products, 0.1 μl of *Taq* polymerase (Taraka), 0.6 μl of 10 mM of each oligonucleotide primer, and 4.5 μl of ddH_2O, on a DNA amplification machine (MJ, USA). PCR cycle conditions were 1 cycle of 94 °C for 3 min; 30 cycles of 94 °C for 30 s, 60 °C/60.5 °C/58.5 °C/62 °C for 30 s, and 72 °C for 30 s; and 1 cycle of 72 °C for 5 min. Reaction products were separated by electrophoresis on a 5% polyacrylamide gel. PCR-amplified DNA was purified using the High Pure kit (Roche, Penzburg, Germany) and either directly sequenced or cloned into the pGEM-T Easy Vector (Promega, USA) and then sequenced (SNBC, Shanghai).

2.4. Bioinformatic analysis

Sequence similarity analysis was performed by using BLAST program (http://www.ncbi.nlm.gov/BLAST). The

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

open reading frame was acquired with ORF Finder tool (http://www.ncbi.nlm.nih.gov/gorf/). *DAZL* family protein domains were determined by Interproscan (http://www.ebi.ac.uk/interpro/). The molecular weight of deduced protein of *DAZL* was calculated using the online server program, Protscale (http://au.expasy.org/cgi-bin/protscale.pl). The motif analysis was performed using the online server program, MotifScan (http://myhits.isb-sib.ch/cgi-bin/motif_scan) and NetPhosk 1.0 (http://www.cbs.dtu.dk/services/NetPhosK/). Multiple sequence alignment was performed using ClustalW program. Based on the data from the bovine database of GenBank (http://www.ncbi.nlm.nih.gov/), the genomic organization and map location of *DAZ* family genes were investigated by comparing cDNAs and the corresponding genomic sequences. The transcription factor binding sites (TFBS) search was performed using the MatInspector program (http://www.genomatix.de).

The initial three-dimensional structure of b-*DAZL* was built using threading method with SAM-T06 CASP model [33–35] in SAM_T06_server (http://www.soe.ucsc.edu/research/compbio/SAM_T06/T06-query.html). The structural refinements were accomplished by energy minimization (after 3000 steps, the steepest descent mode, minimization was switched to the 1000-step conjugate gradient method) using software Insight II (Accelrys, San Diego, CA). Consequently, molecular dynamics (MD) simulation was carried out to equilibrate the structure using GROMACS 3.3 package [36,37]. Briefly, we placed a layer of water (SPC216 model) of 0.5 nm around the structure. Then, the system was minimized using 4000 steps of the steepest descent method. After position-restrained simulation was carried out for 200 ps at 300 K, the system was simulated for 5 ns (parameters: Gromos96 force field [38], temperature = 300 K, pressure = 1 atm). The temperature and pressure were ulated using the Berendsen coupling algorithm [39] with a coupling constant of 0.1 and 0.5 ps. Particle Mesh Ewald (PME) was employed to calculate long-range electrostatic interactions.

2.5. Expression profile

Reverse transcription-polymerase chain reaction (RT-PCR) was used to determine the expression profile of the *DAZL* gene in various cattle, yak and cattle–yak tissues mentioned above. The primer sequences for b-*DAZL* were as follows: 5′-CAGCCTCCAACCATGATAAACCC-3′ (sense) and 5′-CATAACTCCTTTGTTCCCCAGCAG-3′ (antisense). The RT-PCR was performed under the conditions of 1 cycle of 94 °C for 3 min; 30 cycles of 94 °C for 30 s, 50 °C for 30 s, and 72 °C for 30 s; and 1 cycle of 72 °C for 3 min for the *GAPDH* control.

2.6. DNA extraction and PCR amplification of b-DAZL gene

DNA extraction was performed as described previously [40]. The entire coding region of b-*DAZL* gene was sequenced, including all exons and flanking regions. The primers 5′-TTTAGAAGTCGGGCAG-3′ and 5′-AAC TATATATTCAACACAGAACCAG-3′ were used for cloning b-*DAZL*. Each PCR contained 2 μl of template DNA, 1 μl of each primer at 20 μM, 2 μl of 1 mM dNTPs, 2 μl of 10× PCR buffer supplied with the enzyme, one unit of *Taq* polymerase (Promega), and ddH_2O up to 20 μl. The PCR was performed for 30 cycles: 1 cycle of 94 °C for 3 min; 30 cycles of 94 °C for 30 s, 60 °C/58.5 °C for 30 s, and 72 °C for 30 s; and 1 cycle of 72 °C for 5 min. Reaction products were used for electrophoresis on 5% polyacrylamide gels. PCR-amplified DNA was purified using the High Pure kit (Roche, Penzburg, Germany) and either directly sequenced or cloned into the pGEM-T Easy Vector (Promega, USA) and then sequenced (SNBC, Shanghai).

3. Results

3.1. Cloning and mapping of b-DAZL gene

It was found that there were three bos genomic contigs (GenBank Accession No. M606246, CB170267 and DT850351) that highly homologous to the query probe RRM amino acids sequence within *DAZL* (GenBank Accession No. NW_001351). Based on the sequences of these contigs, a putative ORF was predicted and assembled. A cDNA clone was identified (with putative ORFs of 885 bp) and designated as b-*DAZL* (Fig. 1). The full length of b-*DAZL* cDNA was 2456 bp and predicted a protein (b-DAZL) of 295 amino acids. The coding region, b-*DAZL*, was highly conserved with about 90–94% conservation of nucleotide sequence to its homologues within vertebrates. The deduced amino acid sequence identity was found to be 90–96% conserved. The complete cDNA and deduced amino acids sequence have been deposited in GenBank/EMBL/DDBJ (GenBank Accession No. EF501823).

The b-*DAZL* gene cDNA sequence was used as the query probe to search bovine genome database of GenBank to localize its position on bos chromosome, and was confirmed by the American UCSC (http://genome.ucsc.edu/cgi-bin/hgBlat). This b-*DAZL* gene was mapped at the terminal of chromosome 1, between NW001493813.1 and NW00149384.1.

3.2. Analysis of the b-DAZL gene sequence

The b-*DAZL* gene consisted of 12 exons interrupted by 11 introns (Fig. 2(a)). To provide additional data to support further analysis, we performed a similar search for b-*DAZL* gene homologues in vertebrates whose genomic structure could be studied. Thus, the structural data for *DAZL* obtained in this study allowed a comparative analysis of the genomic organization of all four *DAZL* genes. Among the 10 exons within the region of *DAZL* gene we found that 9 exons were identical in size, intron–exon boundaries, and position and sequence. The

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

```
                                                                                TACCCGCC
9    TCTGACTCTCTCCGTCGGGTCTGCTGCCGCTCCCTCTGCCCAGTCCGGCCGGGCAGCGGGCACGAGCAGACCCAACAG
87   CAGCTCCAAGCGCGGTGCCCGAGGCCAGGCTCTGCTTCGTCCTCGAGCCCCTCTGAGTGAACACTGGCTGTCGCCATC
165  ATGTCTGCTGCAAATCCTGAGACTCCAAACTCAACCATCTCCAGAGAGGCCAGCACCCAGTCGTCTTCAGCTACCACC
      M  S  A  A  N  P  E  T  P  N  S  T  I  S  R  E  A  S  T  Q  S  S  S  A  T  T
243  AGCCAAGGCTATGTTTTACCAGAAGGCAAAATCATGCCAAACACTGTTTTTGTCGGTGGAATTGATGTTAGGATGGAT
      S  Q  G  Y  V  L  P  E  G  K  I  M  P  N  T  V  F  V  G  G  I  D  V  R  M  D
321  GAAACAGAAATTAGAAGTTTCTTTGCTAGATATGGTTCAGTAAAAGAAGTGAAGATAATCACGGATCGAACTGGTGTG
      E  T  E  I  R  S  F  F  A  R  Y  G  S  V  K  E  V  K  I  I  T  D  R  T  G  V
399  TCCAAAGGCTATGGATTTGTTTCATTTTATAATGACGTGGATGTGCAGAAGATAGTAGAATCACAGATAAATTTCCAT
      S  K  G  Y  G  F  V  S  F  Y  N  D  V  D  V  Q  K  I  V  E  S  Q  I  N  F  H
477  GGTAAAAAGCTGAAACTGGGCCCTGCAATCAGGAAACAAAATTTGTGTGCTTATCATGTGCAGCCACGTCCTTTGGTT
      G  K  K  L  K  L  G  P  A  I  R  K  Q  N  L  C  A  Y  H  V  Q  P  R  P  L  V
555  TTTAATCCTCCTCCTCCACCACAATTTCAGAGTGTCTGGAGTAATCCAAACGCTGAAACTTACATGCAGCCTCCAACC
      F  N  P  P  P  P  P  Q  F  Q  S  V  W  S  N  P  N  A  E  T  Y  M  Q  P  P  T
633  ATGATAAACCCTATAACTCAGTATGTTCAGGCTTATCCTCCTTATCCAAGTTCACCAGTTCAGGTTATCACTGGATAT
      M  I  N  P  I  T  Q  Y  V  Q  A  Y  P  P  Y  P  S  S  P  V  Q  V  I  T  G  Y
711  CAGCTGCCTGTTTATAATTATCAGATGCCACCGCAGTGGCCTGCTGGGGAACAAAGGAGTTATGTTATTCCTCCGGCT
      Q  L  P  V  Y  N  Y  Q  M  P  P  Q  W  P  A  G  E  Q  R  S  Y  V  I  P  P  A
789  TATACAACCATTAACTACCACTGTAATGAAGTTGACACCGGAGCTGACGTTTTGCCCAGTGAATGCTCGGTTCATGAA
      Y  T  T  I  N  Y  H  C  N  E  V  D  T  G  A  D  V  L  P  S  E  C  S  V  H  E
867  GCTACTCCGTCCTCTGGAAATGGCCCACAAAAGAAATCTGTGGACAGAAGCATACAGACGGTGGTATCTTGTCTATTT
      A  T  P  S  S  G  N  G  P  Q  K  K  S  V  D  R  S  I  Q  T  V  V  S  C  L  F
945  AATCCAGAGAATAGACTGAGAAACTCCGTTGTCACTCAAGATGACTACTTCAAGGATAAAAGAGTTCATCACTTTAGA
      N  P  E  N  R  L  R  N  S  V  V  T  Q  D  D  Y  F  K  D  K  R  V  H  H  F  R
1023 AGAAGTCGGGCAGTGCTTAAGTCTGTTTGATCCTCTCTGCTAACTTTCTAGTCTCATAGGAAGTCACAGGTTTTGAAT
      R  S  R  A  V  L  K  S  V  *
1101 AGTAAGAGACTGAAAGTTTTCCACTGTTATAGAAATTTAATTCTGAATTTTGATAAATCACACTCAGACTTTCTATAT
1179 GTTATATTAGATTGCTTAGTTTTATTTTACATTGAAACTTTTTCATTAGACATTTACTTAGAAACTGGTTCTGTGTTG
1257 AATATATAGTTAAAAGTTAAAAAATAATTGAGACTGAAAAGAATAACTTTAAGATATATCTGTAAGGATTTTTACCTG
1335 ACATTTTAAGAGTTTAAAAACAAATGCTGATTTTTTGGGAAAAAAAATTGTTTTAGAAAACCTATTTTGAAAGGTCAG
1413 AATTTTGATAGTTTAAATACAAGCACTTCACTTCTCCAACAAGTACCTATGAACAGTACAGTATCTGCAATATTGTGC
1491 TTTTATGATTTGTCTAGAAATTTACCACAAAAGCAGAGTTTTTAAAACTGAATTTTTAATCAGTGAAACTCAATGTAT
1569 AGTTAGCTTTATTGAAGTTTTCCTTATCTAAACCCAGTAAAACAGATTCTAAACAAACAGTCCAATCAGTGGGTCTTA
1647 TGTTTATTACTTCAAAATACTTTATCTTTTATCTAGAATCCACACATAGAGGTATCTGATGGGATGGTAATTAGGATT
1725 ACTAAAATTCTGGGCCAAATTTTTTTTTAAAAGAATCTAAAACAAACTTAACTTTTCTAGGTACATAAACTTCTTAGT
1803 TGAGTTACCATTCTTTTTTCTTTTACTTTTTTTTCCCCCTTGAAATGTTAAACTCAATGGCTGTATTTTAAATTGTGC
1881 AAAATATTGGTATTAAAGAAGGCTGAAACTTTGTCACTTAAAAGTTAATCAGAGTATCTGAAAGGAATTGTTTTTATA
1959 AGAAACATTAAAATATTAGCTATTTGTTAAAGAACAGATAGATTTTAATTTTTTTTTAGCCTGTTATTTTTCTTTCTT
2037 TTTTTTTTTTTTTAAGTAAAATATGTCCAGGAGAGTCAGAGTAGTTTTGATGTTTCTAAACCACAAAAGTTGTTTAGT
2115 AATAAGCATATCCCCAAAACACACTTGAGTTAGTTAACAGTCCATTTTTAGATGAGGTGGCTCCTCCTCAGACATTGT
```

Fig. 1. Nucleotide and deduced amino acid sequences of the bovine b-*DAZL* gene. The start code ATG and the stop codon TGA are shaded in gray.

other one exon had only 0–6 bp difference in size but was identical in position with other highly homologous sequences (Fig. 2(b)). Also the promoter fragment was analyzed by MatInspector program based on a database of transcription factors. More than five potential transcription factor binding sites (TFBS) associated with reproduction and embryonic development were identified, indicating that b-*DAZL* may have a function in the bovine reproductive system (Table 1).

3.3. Analysis of deduced amino acids sequence and protein structure

3.3.1. Analysis of deduced amino acids sequence

The predicted protein product of b-*DAZL* gene comprises 295 amino acid residues, and no signal peptide cleavage sites were found. The deduced amino acid sequence was compared to its homologues, and the results indicated that the amino acid sequence and length were extensively

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

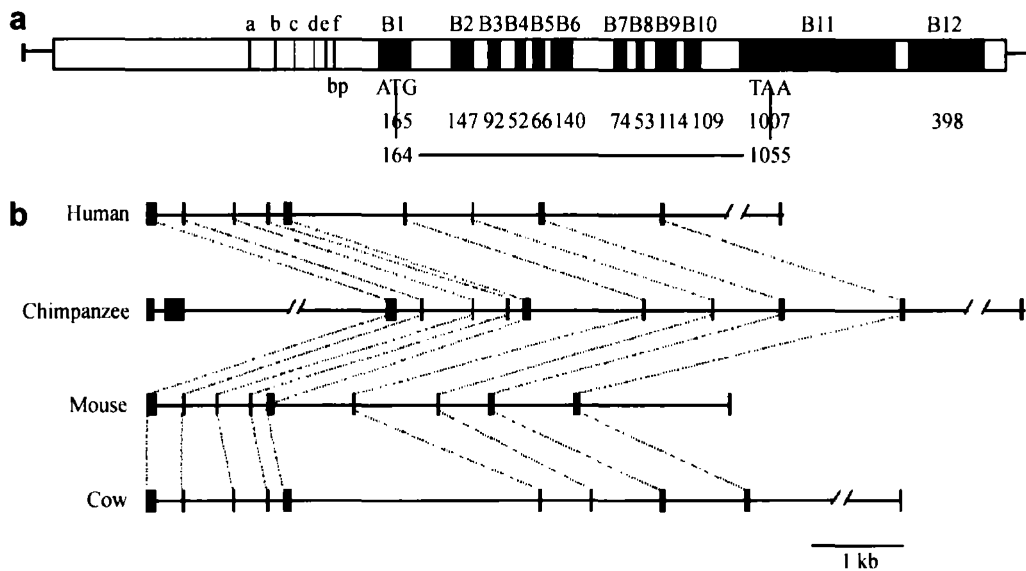

Fig. 2. Schematic structure of the mammalian *DAZL* gene. B1–B12 indicate exons, and the coding region is indicated. (a) Genomic organization of bovine b-*DAZL*; (b) comparison of the genomic organization of ORF of *DAZL* in mammals.

Table 1
Probable promoter sequences of b-*DAZL* gene which are associated with reproduction and embryonic development

Name	Start	End	Score	Promoter sequence
GREF/PRE.01	1165	1183	0.84	aaaatagttttTGTTctct
GREE/ARE.02	1572	1590	0.89	gttgttcattcaGTCCtgt
	2389	2407	0.89	attttgcatgatGTACtct
ZF35/ZF35.01	1788	1800	0.96	gggggcAAGAgtt
OCT1/OCT1.06	1374	1388	0.81	tttatcgtAATTtct
OCTP/OCT1P.01	1828	1840	0.86	tgaATATtcaggg
	2305	2407	0.86	ttaATATgctgtc
	2433	3445	0.86	acaATATgcagcc

conserved in the vertebrates (53–96% identity) and mammalian (90–96% identity) DAZ-like proteins (Table 2). Primary sequence and motif analysis further helped to depict features of b-*DAZL*. As shown in Fig. 3, the amino acid residues 33–121 constitute an RRM motif, DAZ repeat sites were from Ala165 to Gln194, and amino acid residues 126–206 constitute a Pro-rich region. From the comparison, we were able to define two well-characterized functional domains, the amino acid residues 99–166 and 121–171 as the PABP and PUM2 interacting domain, respectively (Fig. 3). The sequence alignment also demonstrated that the amino acid sequence within RNP2 and RNP1 showed no variability and the majority of amino acid substitutions in these homologues occurred outside the RRM and DAZ repeat. We observed that four amino acids changed within the two well-characterized functional domains, the RRM domain and the DAZ repeat. However, the changes that occurred were conserved.

3.3.2. Analysis of the protein structure

During molecular dynamics simulation (DM), root mean square deviations (RMSDs) of Cα atoms from DAZL protein reached plateau at about 2 ns (Fig. 4(a)), indicating that

Table 2
Comparison of homology between b-DAZL and other mammalian DAZ-like proteins

No.	Origin	GenBank No.	Identity (%)
1	*Microcebus murinus*	CAG34242	95
2	*Mus musculus*	NP_034151	93
3	*Rattus norvegicus*	XP_001057401	93
4	*Macaca mulatta*	XP_001084406	92
5	*Pan paniscus*	CAG30560	91
6	*Homo sapiens*	NP_001342	91
7	*Macaca fascicularis*	CAA68233	91
8	*Pan troglodytes*	XP_516314	91
9	*Monodelphis domestica*	XP_001380250	92
10	*Saguinus oedipus*	CAG30561	91
11	*Callithrix jacchus*	Q9BGN8	90
12	*Saimiri sciureus*	CAG30562	91
13	*Canis familiaris*	XP_534251	96
14	*Cynops pyrrhogaster*	BAD38676	72
15	*Gallus gallus*	NP_989549.1	67
16	*Ambystoma mexicanum*	AAK58846	73
17	*Xenopus laevis*	AAH71023.1	57
18	*Xenopus tropicalis*	AAH67947.1	55
19	*Rana pipiens*	AAV30542.1	53

the structure was stable after this time. Fig. 4(b) shows the average structure of b-DAZL protein in 2–5 ns from molecular dynamics simulation. We can see that the structure is mainly composed of a RRM domain, several loops and a loose C-terminal region. At N-terminal b-*DAZL* folded into a compact αβ structure comprising a four-strand antiparallel β-sheet packed against two α-helices, which is the characteristic of RNP-type RBDs (Fig. 5(a)). Special structural features of the RRM domain (residues 33–121) include a βαββαβ secondary structure element which forms a four-strand antiparallel β-sheet packed against the two perpendicularly oriented α-helices (Fig. 5(a)). Amino acids of RNP2 (residues 43–47) and RNP1 (residues 80–87) were juxtaposed on the two central β strands (β3 and β1) of the folded domain (Fig. 5(b)). The structural arrangement in

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

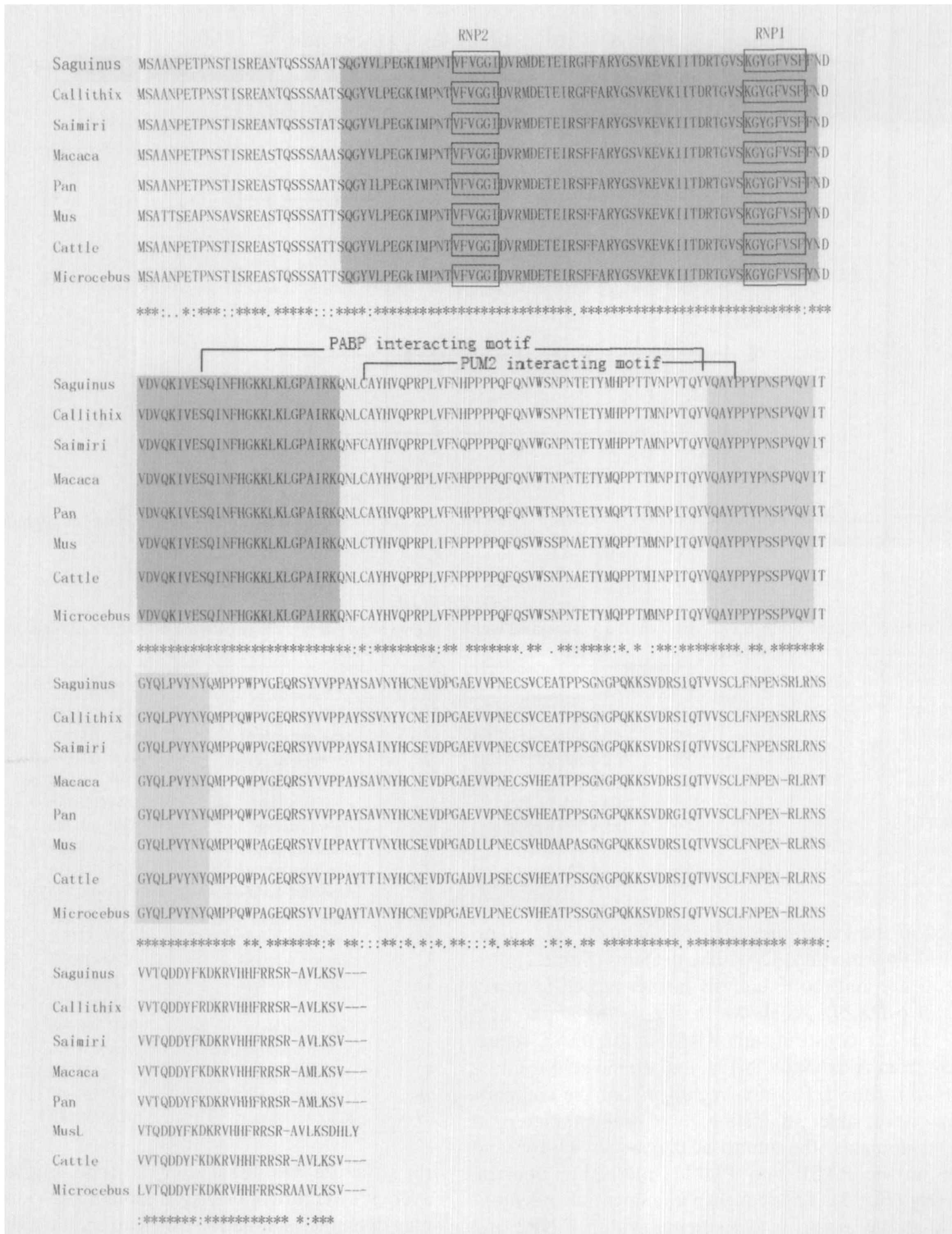

Fig. 3. Conservation of the DAZL proteins in mammals. Alignment of b-DAZL protein sequence with *Callithrix* (Q9BGN8), *Microcebus* (CAG34242), *Mus* (NP_034151), *Pan* (XP_516314), *Saguinus* (CAG30561) and *Saimiri* (CAG30562). The gray box indicates the RRM domain, the light-gray box represents the DAZ repeat, and the highly conserved RNP2 and RNP1 motifs in the RRM are boxed. Asterisks indicate identical amino acids among the homologues. Closed circles represent identical amino acids among the homologues.

the RRM domain was a ring of the aromatic residues of β2 and β3 which is suitable for stacking interaction with RNA bases (Fig. 5(a)). The characteristic of this structure coincided with the other RNA-binding proteins, indicating that b-*DAZL* might function in the same way as RNA-binding protein. Residues 99–166 which comprised β4 and a loop were defined as the PABP interacting domain (Fig. 4(b)). The DAZ repeat showed the feature of coils and α-helices (Fig. 4(b)). A long proline-rich region (PRR; residues 126–207), which occur widely in prokaryotes and eukaryotes, was present in the C-terminal of b-*DAZL*. Despite its wide distribution, the functions of PRRs are often unclear. They prefer β-turns and random coils, rarely found in α-helices and β-sheets (Fig. 4(b)).

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

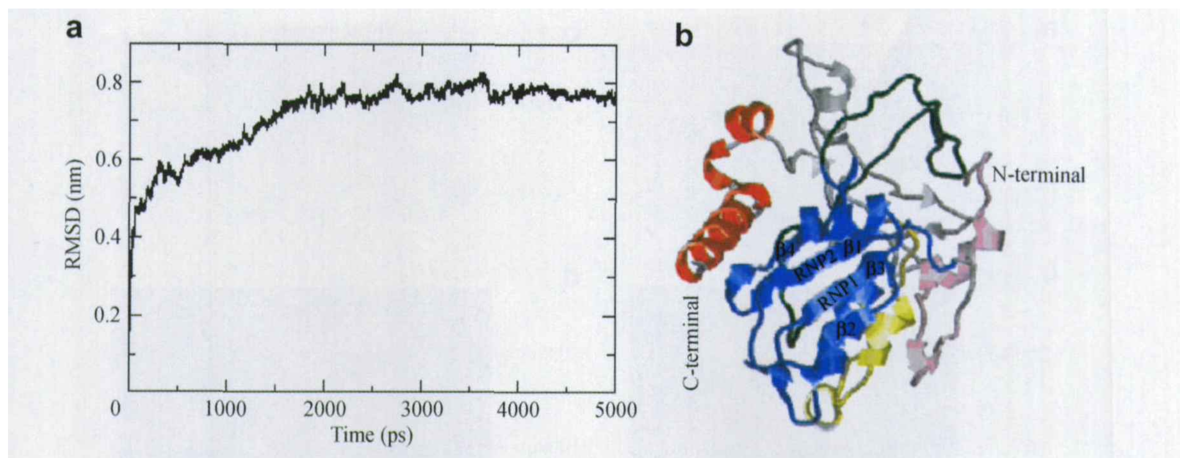

Fig. 4. Structure of b-*DAZL* from the simulation displayed in ribbon style. (a) Root mean square deviations (RMSDs) of Cα atoms from b-DAZL protein reached plateau at about 2 ns; (b) structure of b-DAZL protein from MD simulation.

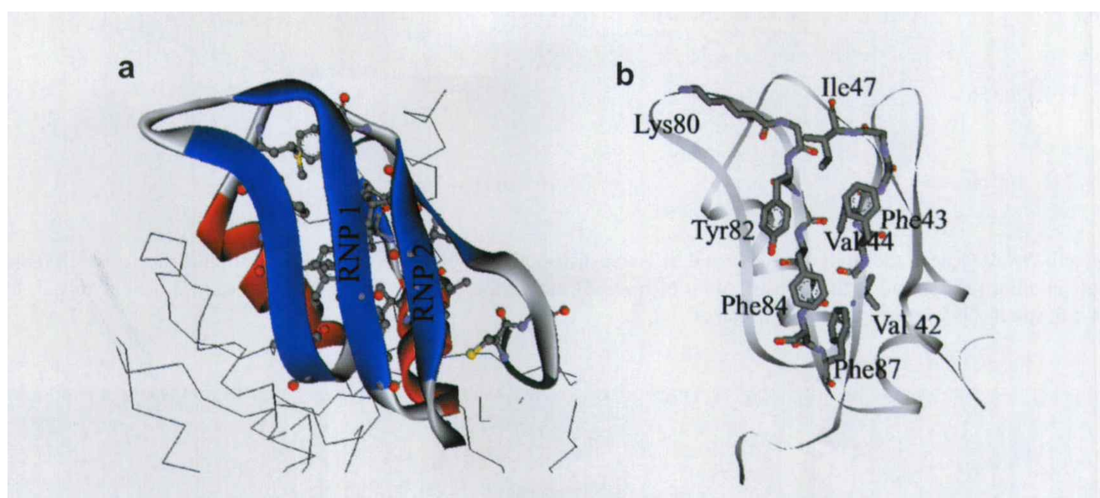

Fig. 5. RRM domain of b-*DAZL*. (a) RRM domain displayed in ribbon, hydrophobic core of this domain represented in stick and ball style, helices colored in red and β-sheets colored in cyan; (b) residues in RNP1 and RNP2 represented in stick style.

3.4. Expression profile of b-DAZL gene

The expression of b-*DAZL* gene in various adult F1 cattle–yak, cattle and yak tissues was investigated. The results showed that b-*DAZL* was expressed exclusively in cattle and yak testes and ovaries (Fig. 6(a–d)). In female cattle–yaks with normal oogenesis, b-*DAZL* signals could be detected, but not in testes where only somatic cells and primordial germ cells were found (Fig. 6(e and f)).

3.5. Histological examination of the testes

The tissue sections from testes of eight cattle with complete spermatogenesis contained multiple germ cells per testis cord, and the chromatin of these germ cells, as expected, was diffused rather than condensed (Fig. 7(a and b)). Many sperm cells appeared in the epididymis (Fig. 7(c)). In contrast, the tissue sections from testes of six F1 cattle–yaks contained few identifiable germ cells or spermatocytes and some testes only contained somatic cells (Fig. 7(d and e)). The germ cells displayed a degenerated morphology and were loosely attached to the seminiferous epithelium (Fig. 7(d and e)). In addition, a few germ cells with chromatin condensed at the edge of the nucleus, typical of apoptotic cells, were observed in cattle–yak testes (Fig. 7(e)). No sperm appeared in the epididymis (Fig. 7(f)).

3.6. Detection of b-DAZL gene in cattle–yak by PCR

A 260 bp fragment representing the b-*DAZL* gene was amplified from male cattle–yaks. All samples yielded positive amplification of the b-*DAZL* fragment, confirming the existence of the b-*DAZL* gene in cattle–yaks (Fig. 8).

4. Discussion

4.1. b-DAZL gene is a DAZL homologue

Numerous genes are required for gametogenesis in mammals. However, until now, no genes have been shown to be necessary for gametogenesis in the bovine. To further analyze the reasons for cattle–yak sterility, b-*DAZL*, the homologue of *DAZL* gene was identified through *in silico* cloning and RT-PCR. b-*DAZL* contains an RRM domain,

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

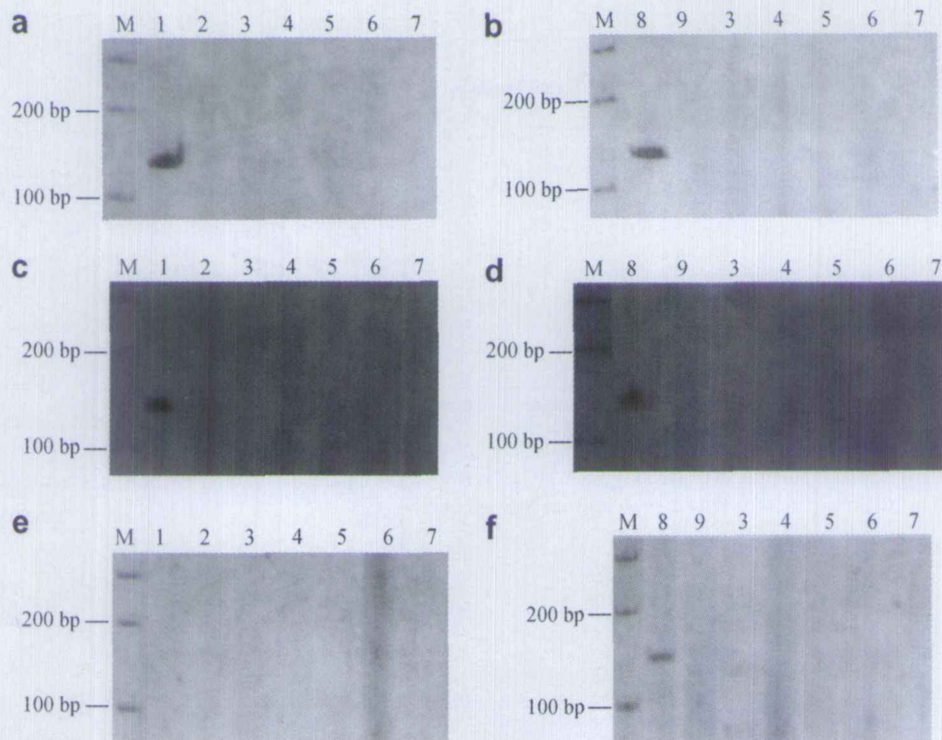

Fig. 6. Expression of b-*DAZL* gene in the adult bovine. (a and b) Tissue distribution of b-*DAZL* transcripts in adult cattle; (c and d) tissue distribution of b-*DAZL* transcripts in adult yak; (e and f) tissue distribution of b-*DAZL* transcripts in adult cattle–yak. 1, Testis; 2, epididymis; 3, hypothalamus; 4, pituitary; 5, kidney; 6, spleen; 7, stomach; 8, ovary; 9, uterus.

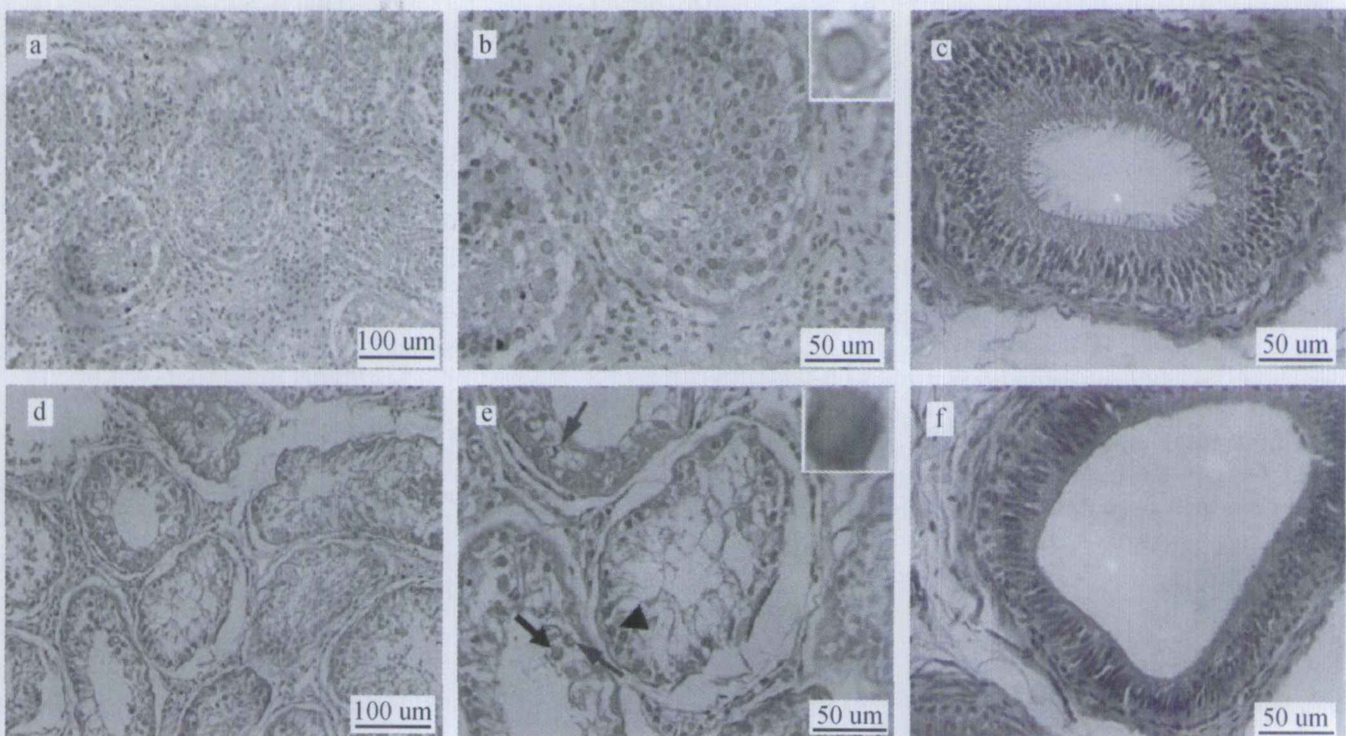

Fig. 7. Histology of testis and epididymal from b-$DAZL^{+/+}$ (a–c) cattle and b-$DAZL^{-/-}$ (d–f) cattle–yak. (a) A common feature of the testis from b-$DAZL^{+/+}$ male cattle; (b) seminiferous tubular section showing multiple germ cells; (c) epididymal section showing many sperm cells; (d) a common feature of the testis from b-$DAZL^{-/-}$ male cattle–yak; (e) seminiferous tubular section showing primordial germ cells (PGCs) (arrowhead), some of which are in apoptosis (arrow), and oocytes (broad arrow); (f) epididymal section showing no germ cells.

one copy of the DAZ repeat and a C-terminal domain. As expected, the bovine sequence is most closely related to the mammal *DAZL* genes at both the nucleotide and protein levels. From sequence alignment we found that the RRM

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

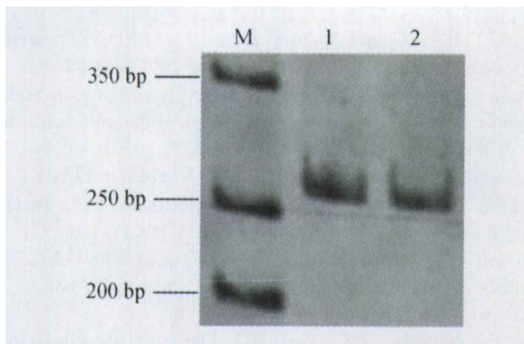

Fig. 8. b-*DAZL* gene detection in male cattle–yaks. 1 and 2, products of b-*DAZL* gene; M, DNA marker.

domain is highly conserved in all organisms, especially within the RNP1 and RNP2 peptide motifs. The interacting domains with Pumilio/PABP protein within b-*DAZL* homologues are also highly conserved.

The RRM domain and the *DAZ* repeat in the b-DAZL protein are located conservatively. The presence of a four amino acids insertion that is not conserved in *Saguinus*, *Callithix*, *Saimiri*, *Macaca*, *Pan*, *Mus*, *Bovinae* and *Microcebus* suggests that the event occurred within bovine evolution. Two changes are present within the RRM domain but DAZL protein's function is conserved in gametogenesis. The other two changes that occurred within the DAZ repeat also have no effect on b-*DAZL* function. Even the great divergence of the DAZ repeat between *BOULE* and *DAZL* has not reduced its ability to rescue the *BOULE* defect [12,20,41]. This suggests that the DAZ repeat is not critically important, or that a smaller motif is actually the functional unit.

4.2. Expression of the b-DAZL gene in cattle–yak

DAZL is expressed in both the testis and ovary [8,24,25,42], and some researchers detected its expression in embryonic gonads [22]. We found a significantly higher *DAZL* mRNA level in bovines with normal gametogenesis. It is not surprising that we found no signals in testes with a spermatogenic defect, because *DAZL* is expressed in primary spermatocytes and spermatogonia. The absence of b-*DAZL* transcripts may be attributed to the presence of no germ cells or only primordial germ cells in cattle–yak testis. Several spermatogonia with the degenerated morphology were contained in seminiferous tubules, which may confirm that b-*DAZL* expresses in spermatogonia and is critical for spermatogonia formation and development. The comparison of the gene arrangement and expression pattern of b-*DAZL* with those of other *DAZL* genes suggests a possible function as a crucial regulator of spermatogenesis.

4.3. Microstructure of testes from cattle and cattle–yaks

Research has provided substantial evidence for the requirement of the *DAZL* gene in germ cell formation and development. In invertebrates, *dazl* gene in *C. elegans* is only expressed in ovary and its defect results in female sterility caused by blocked oogenesis [9]. In vertebrates, depletion of *Xdazl* mRNA in *Xenopus* oocytes leads to subsequent generations lacking primordial germ cells, indicating that *Xdazl* might be involved in spermatogonia formation and development [8,12,43]. $DAZL^{-/-}$ mice are sterile in both sexes, lacking all spermatozoa or oocytes [17–19]. In this study, we found that cattle–yak testes hosted a great majority of the seminiferous tubules, primordial germ cells (PGCs) were the only germ cells presented there, many of which showed an abnormal morphology. Moreover, only a few PGCs developed further, and even these cells underwent apoptosis and never developed beyond the spermatocyte stage. Our results may indicate that *DAZL* genes are required for spermatogonia development and that the defects will result in disability of the cells to differentiate into spermatocytes.

4.4. Significance of b-DAZL gene existence in cattle–yaks' genome

The b-*DAZL* gene was detected in both the male and female cattle–yak genomes. The fact that no transcription signals of b-*DAZL* detected in spermatogenic defective cattle–yaks may be due to transcript arrest. The transcript arrest of b-*DAZL* might cause a defect in the spermatogenesis control mechanism in male cattle–yaks. The difference in the control mechanism of sexual development, accompanied by gender difference, might explain spermatogenesis arrest in male cattle–yaks.

In conclusion, we have demonstrated that the b-DAZL protein might be required for normal fertility in the male cattle–yak and may act as a crucial regulator of gametogenesis because its function is so well conserved across a number of species. In addition, the expression pattern and transcript level of the b-*DAZL* gene in cattle, yaks and cattle–yaks with either complete gametogenesis or a spermatogenic defect were shown. The level of b-*DAZL* gene transcripts was low or extremely low in male cattle–yaks with spermatogenesis arrest. These results, combined with gene detection, suggest that b-*DAZL* gene might be involved in spermatogenesis in the bovine testis and its transcription arrest might result in male cattle–yak infertility.

Acknowledgements

This work was supported by National Natural Science Foundation of China (Grant No. 30500360) and Youth Science and Technology Innovation Fund for young scientists of Nanjing Agriculture University (KJ05011).

References

[1] Li QF, Li YX, Zhao XB, et al. Sequencing cytochrome b gene of mitochondrial DNA in yak and researching its origin and taxonomic status. Acta Vet Zootech Sin 2006;37(11):1118–23.

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net

[2] Zhang XJ. An observation and study of testicular histology in hybrid bull copulated by yak and cattle. Acta Vet Zootech Sin 2001;32(4):314–8.
[3] Zhang RC, Duan LC. A study on the sperm production of species hybrid bells between yaks and cattle. Acta Vet Zootech Sin 1991;22(3):231–4.
[4] Guo AP. A comparative study on chromosomes of oxen (*Bos taurus*), Yaks (*Bos grunnieus*) and their hybrids. J Genet Genomics 1983;10(2):137–43.
[5] Ji QM. Advances in research of yak resources in China.J Nat Resour 2001;16(6):564–9.
[6] Reijo R, Lee TY, Salo P, et al. Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. Nat Genet 1995;10:383–93.
[7] Ruggiu M, Speed R, Taggart M, et al. The mouse Dazla gene encodes a cytoplasmic protein essential for gametogenesis. Nature 1997;389:73–7.
[8] Houston DW, King ML. A critical role for Xdazl, a germ plasm-localized RNA, in the differentiation of primordial germ cells in *Xenopus*. Development 2000;127(3):444–56.
[9] Karashima T, Sugimoto A, Yamamoto M. *Caenorhabditis elegans* homologue of the human azoospermia factor DAZ is required for oogenesis but not for spermatogenesis. Development 2000;127(5): 1069–79.
[10] Xu EY, Moore FL, Pera RA. A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in metazoans. PNAS 2001;98(13):6983–8.
[11] Tung JY, Luetjens CM, Wistuba J, et al. Evolutionary comparison of the reproductive genes, DAZL and BOULE, in primates with and without DAZ. Dev Genes Evol 2006;216(3):158–68.
[12] Houston DW, Zhang J, Maines JZ, et al. A *Xenopus* DAZ-like gene encodes an RNA component of germ plasm and is a functional homologue of *Drosophila* boule. Development 1998;125(2):171–80.
[13] Tsui S, Dai T, Roettger S, et al. Identification of two novel proteins that interact with germ-cell-specific RNA-binding proteins DAZ and DAZL1. Genomics 2000;65(3):266–73.
[14] Yen PH. Putative biological functions of the DAZ family. Int J Androl 2004;27(3):125–9.
[15] Otori M, Karashima T, Yamamoto M. The *Caenorhabditis elegans* homologue of deleted in azoospermia is involved in the sperm/oocyte switch. Mol Biol Cell 2006;17(7):3147–55.
[16] Padmanadhban K, Richter JD. Regulated Pumilio-2 binding controls RINGO/Spy mRNA translation and CPEB activation. Gene Dev 2006;20(2):199–209.
[17] McNeilly JK, Saunders PT, Taggart M, et al. Loss of oocytes in Dazl knockout mice results in maintained ovarian steroidogenic function but altered gonadotropin secretion in adult animals. Endocrinology 2000;141(11):4284–94.
[18] Schrans-Stassen BH, Saunders PT, Cooke HJ, et al. Nature of the spermatogenic arrest in Dazl 2/2 mice. Biol Reprod 2001;65(3):771–6.
[19] Slee R, Grimes B, Speed RM, et al. A human DAZ transgene confers partial rescue of the mouse Dazl null phenotype. PNAS 1999;96(14):8040–5.
[20] Vogel T, Speed RM, Ross A, et al. Partial rescue of the Dazl knockout mouse by the human DAZL gene. Mol Hum Reprod 2002;8(9):797–804.
[21] Lin Y, Page DC. Dazl deficiency leads to embryonic arrest of germ cell development in XY C57BL/6 mice. Dev Biol 2005;288(2):309–16.
[22] Cauffman G, Van de Velde H, Liebaers I, et al. DAZL expression in human oocytes, preimplantation embryos and embryonic stem cells. Mol Hum Reprod 2005;11(6):405–11.
[23] Pan HA, Tsai SJ, Chen CW, et al. Expression of DAZl protein in the human corpus luteum. Mol Hum Reprod 2002;8(6):540–5.
[24] Seligman J, Page DC. The Dazh gene is expressed in male and female embryonic gonads before germ cell sex differentiation. Biochem Biophys Res Commun 1998;245(3):878–82.
[25] Dorfman DM, Genest DR, Reijo Pera RA. Human DAZL1 encodes a candidate fertility factor in women that localizes to the prenatal and postnatal germ cells. Hum Reprod 1999;14(10):2531–6.
[26] Brekhman V, Itskovitz-Eldor J, Yodko E, et al. The DAZL1 gene is expressed in human male and female embryonic gonads before meiosis. Mol Hum Reprod 2000;6(5):465–8.
[27] Ruggiu M, Saunders PT, Cooke HJ. Dynamic subcellular distribution of the DAZL protein is confined to primate male germ cells. J Androl 2000;21(3):470–7.
[28] Lin YM, Chen CW, Sun HS, et al. Expression patterns and transcript concentrations of the autosomal DAZL gene in testes of azoospermic men. Mol Hum Reprod 2001;7(11):1015–22.
[29] Reijo RA, Dorfman DM, Slee R, et al. *DAZ* family proteins exist throughout male germ cell development and transit from nucleus to cytoplasm at meiosis in humans and mice. Biol Reprod 2000;63(5): 1490–6.
[30] Lin YM, Chen CM, Sun HS, et al. Presence of DAZL transcript and protein in mature human spermatozoa. Fertil Steril 2002;77(3): 626–9.
[31] Nishi S, Hoshi N, Karahara M, et al. Existence of human DAZLA protein in the cytoplasm of human oocytes. Mol Hum Reprod 1999;5(6):495–7.
[32] Tsai MY, Chang SY, Lo HY, et al. The expression of DAZL1 in the ovary of the human female fetus. Fertil Steril 2000;73(3):627–30.
[33] Hughey R, Krogh A. Hidden Markov models for sequence analysis: extension and analysis of the basic method. CABIOS 1996;12(2): 95–107.
[34] Karplus K, Barrett CR. Hidden Markov models for detecting remote protein homologies. Bioinformatics 1998;14(10):846–56.
[35] Krogh A, Brown M, Mian IS, et al. Hidden Markov models in computational biology: applications to protein modeling. JMB 1994;235:1501–31.
[36] Berendsen HJC, van der Spoel D, van Drunen R. GROMACS: a message-passing parallel molecular dynamics implementation. Comp Phys Comm 1995;91:43–56.
[37] Lindahl E, Hess B, van der Spoel D. GROMACS 3.0: a package for molecular simulation and trajectory analysis. J Mol Mod 2001;7(8):306–17.
[38] van Gunsteren WF, Billeter SR, Eising AA, et al. Biomolecular simulation: the GROMOS96 manual and user guide. Zurich, Switzerland: Hochschulverlag AG an der ETH Zurich; 1996.
[39] Berendsen H, Postma J, van Gunsteren W, et al. Molecular dynamics with coupling to an external bath. J Comput Phys 1984;81(8): 3684–90.
[40] Dianna S, Irma O. Single-step PCR for detection of brucell from blood and milk of infected animals. Clin Microbiol 1995;23(3): 3087–90.
[41] Xu EY, Lee DF, Klebes A, et al. Human BOULE gene rescues meiotic defects in infertile flies. Hum Mol Genet 2003;12(2):169–75.
[42] Xu H, Li M, Gui J, et al. Cloning and expression of medaka dazl during embryogenesis and gametogenesis. Gene Expr Patterns 2007;7(3):332–8.
[43] Mita K, Yamashita M. Expression of *Xenopus* Daz-like protein during gametogenesis and embryogenesis. Mech Dev 2000;94(1–2):251–5.

?1994-2018 China Academic Journal Electronic Publishing House. All rights reserved. http://www.cnki.net