Biochem Genet (2010) 48:173–180
DOI 10.1007/s10528-009-9299-2

NOTE

Cloning and Characterization of Novel Isoforms of the *BOULE* Gene in Bats

Lihong Yuan · Xueguo Zuo · Lingjiang He · Paul Racey · Eran Levin · Shuyi Zhang

Received: 16 January 2009 / Accepted: 9 July 2009 / Published online: 28 November 2009

Introduction

Spermatogenesis is a complex physiological process controlled by many genes. The *BOULE* gene, a new member of the *Deleted in Azoospermia (DAZ)* family (which consists of *BOULE*, *DAZ*, and *DAZ*-like *DAZl*), is regarded as the ancestor of the *DAZ* family and a key factor in controlling the meiosis of male germ cells, which can regulate the expression of the *twine* gene and promote progression through meiosis (Eberhart et al. 1996; Karashima et al. 2000; Maines and Wasserman 1999; Xu et al. 2001). Bats account for about 20% of mammals and, during evolution, have evolved many reproductive strategies, including sperm storage, delayed fertilization, delayed implantation, and delayed development (Nowak et al. 1994; Racey and Entwistle 2000). The strategies of sperm storage and delayed fertilization allow many bat species, especially hibernating bats, to achieve synchrony between birth peaks and food availability (Racey 1979). As little is known about the cell biology of bat spermatogenesis, in this study, we obtained *BOULE* gene sequences from four bat species (*Rhinolophus ferrumequinum*, *Myotis ricketti*, *Eonycteris spelaea*, and *Rousettus leschenaultii*) to test the role of the *BOULE* gene in bat

Lihong Yuan and Xueguo Zuo contributed equally to this paper.

L. Yuan
Guangdong Entomological Institute, Guangzhou 510260, China

L. Yuan · X. Zuo · L. He · S. Zhang (✉)
School of Life Sciences, East China Normal University, Shanghai 200062, China
e-mail: syzhang@bio.ecnu.edu.cn

P. Racey
School of Biological Sciences, University of Aberdeen, Aberdeen AB24-2TZ, United Kingdom

E. Levin
Department of Zoology, Tel Aviv University, Tel Aviv 69978, Israel

spermatogenesis. We identified at least two isoforms (named a and b) of the *BOULE* gene in bats. Isoform a is common to the *BOULE* gene of other species, whereas isoform b, which is here identified for the first time, is specific in bats. As for the difference of exons lost, the bat *BOULE* gene isoform b has a premature stop codon in a different position and has different lengths of coding-domain sequence (CDS). Thus isoform b of the *BOULE* gene in some species may have lost the functional domain. Therefore, our study first cloned the multiple transcript variants of the bat *BOULE* gene and identified the novel isoforms. These results may add some useful information for the study of bat spermatogenesis and other reproductive strategies of bats.

Materials and Methods

Sample Collection and RNA Isolation

One individual of each species (*R. ferrumequinum* and *M. ricketti* have sperm storage ability, *E. spelaea* and *R. leschenaultii* do not) was sacrificed as part of a surveillance program for coronaviruses in 2006, and testis tissue was stored at −80°C until RNA extraction. Total RNA was isolated from testes using RNAiso Reagent (TaKaRa, Japan), following the manufacturer's protocol. The concentration of RNA was calculated according to the formula $[\text{RNA}] = 50 \times (\text{OD}_{260} - \text{OD}_{320}) \times$ dilution time, and the quality of total RNA was assessed on a 1% agarose gel.

cDNA Synthesis and RT–PCR Amplification

Two micrograms of total RNA from each sample were used to synthesize cDNA. First, 2 μg total RNA was treated with 2 U RNase-free DNase I (Promega) for 30 min at 37°C to avoid genomic DNA contamination, then converted to cDNA by Superscript III Reverse Transcriptase (Invitrogen), according to the manufacturer's instructions, in a 50 μl reaction mixture containing 500 ng random primer, 1 mM dNTP, 2 mM dithiothreitol, 80 U RNase inhibitor (Promega), 1× First-Strand buffer, and 400 U Superscript III Reverse Transcriptase. The reaction conditions were 65°C for 5 min, followed by ice incubation for 5 min, then 25°C for 5 min, 50°C for 1 h, and finally 70°C for 15 min.

The degenerated primer pairs PF1 (5′GRC GCA ARC AWC AAA YCA GAT GCA AAC AGA 3′) and PR1 (5′AGY TGG AMT AGA GCT GCC CAA TTG TCT TAA3′), according to the conservative part of the nucleotide sequences of the *BOULE* gene from several mammals, were synthesized to obtain the completed CDS of the bat *BOULE* gene. Using the first-strand cDNAs of bats as templates, PCR was performed at 94°C first predenatured for 5 min, 30 cycles of 94°C for 30 s, 56°C for 30 s, 72°C for 1 min, and finally 72°C extended for 10 min.

PCR products were isolated from 1% agarose gel and purified using the agarose gel DNA purification kit version 2.0 (TaKaRa, Japan), followed by ligation with the pGEM-T-easy vector (Promega, USA), then transformed into the DH5α competence

cell (TaKaRa, Japan). The identity and orientation of each clone were verified by the universal M13 (−47)/(−48) primer, and then each was sequenced from both directions on an ABI 3730A automated DNA sequencer. To avoid artifacts, multiple clones were sequenced for every specimen.

Sequence Alignment and Structure Analysis of the *BOULE* Gene

Fourteen *BOULE* CDS sequences, from *Homo sapiens* (NM_033030), *Pan paniscus* (AJ717405), *Pan troglodytes* (XM_516011), *Callithrix jacchus* (AJ717407), *Saguinus oedipus* (AJ717406), *Saimiri sciureus* (AJ717408), *Macaca mulatta* (XM_001086915), *Macaca fascicularis* (AB074454), *Microcebus murinus* (AJ746579), *Canis familiaris* (XM_545580), *Bos taurus* (NM_001102115), *Equus caballus* (XM_001500223), *Rattus norvegicus* (XM_001067043), and *Mus musculus* (AF272859), were obtained from GeneBank. The putative coding region of *BOULE* of *Pongo pygmaeus* was retrieved from the UCSC database (http://genome.ucsc.edu).

Nucleotide sequences were aligned using Clustal X 1.81 (Thompson et al. 1997) and used to generate amino acid alignments with Mega 4.0 (Tamura et al. 2007).

Results

A search for sequences on the NCBI database found eight clones of the *BOULE* gene from four bat species (GenBank acc. nos, FJ541190- FJ541197).

Sequence alignments indicated that there were at least two transcript variants for each bat *BOULE* gene, named isoform a and b. Further analysis showed that transcript variants of the bat *BOULE* gene differ mainly in the 3′ end of their CDS, due to selective splicing (Table 1). The CDS of four bat species isoform a were encoded by 10 exons (a–j) and were identical to *H. sapiens* isoform 2. However, for *E. spelaea* isoform a, a 14 bp (AGGAGTGGGGAGTA) insert in exon i results in an open reading frame shift, leading to a premature stop of translation and a change of the amino acid sequence in exon i. Moreover, the CDS of *M. ricketti and R. leschenaultii* isoform b consists of nine exons, and the lack of exon i does not affect the open reading frame. Isoform b of *R. ferrumequinum* had three exons and *E. spelaea* had two exons. Exon d in *R. ferrumequinum* and exon b in *E. spelaea* are lost, their open reading frames are changed, and they have a stop codon in exon e and c, respectively (Fig. 1).

Discussion

Gametogenesis involves specification of germ cell fate, mitotic replication of early germ cell populations, meiotic and postmeiotic development, and complex regulation at the levels of transcription and translation (Urano et al. 2005). The *BOULE* gene family consists of three RNA-binding proteins, BOULE, Daz, and Daz-like (Dazl), that regulate germ cell development and differentiation (Houston and King 2000). *BOULE* encodes a critical conserved switch that regulates

Table 1 Location of exons in the amino acid sequences of human and bat *BOULE* transcript isoforms

Isoform	Acc. no.	Exon	Animo acid sequence
Homo sapiens 2	NM_033030	2	1–43(43aa)
		3	44–73(30aa)
		4	74–92(19aa)
		5	93–117(25aa)
		6	118–160(43aa)
		7	161–184(24aa)
		8	185–200(16aa)
		9	201–243(43aa)
		10	244–276(33aa)
		11	277–283(7aa)
Myotis ricketti a	FJ541192	a	1–43(43aa)
Rhinolophus ferrumequinum a	FJ541194	b	44–73(30aa)
Rousettus leschenaultii a	FJ541196	c	74–92(19aa)
		d	93–117(25aa)
		e	118–160(43aa)
		f	161–184(24aa)
		g	185–200(16aa)
		h	201–243(43aa)
		i	244–276(33aa)
		j	277–283(7aa)
Eonycteris spelaea a	FJ541190	a	1–43(43aa)
		b	44–73(30aa)
		c	74–92(19aa)
		d	93–117(25aa)
		e	118–160(43aa)
		f	161–184(24aa)
		g	185–200(16aa)
		h	201–243(43aa)
		i	244–273(30aa)
		j	–
Myotis ricketti b	FJ541193	a	1–43(43aa)
Rousettus leschenaultii b	FJ541197	b	44–73(30aa)
		c	74–92(19aa)
		d	93–117(25aa)
		e	118–160(43aa)
		f	161–184(24aa)
		g	185–200(16aa)
		h	201–243(43aa)
		i	–
		j	277–283(7aa)

Table 1 continued

Isoform	Acc. no.	Exon	Animo acid sequence
Rhinolophus ferrumequinum b	FJ541195	a	1–43(43aa)
		b	44–73(30aa)
		c	74–92(19aa)
		d	–
		e	93–95(3aa)
		f	–
		g	–
		h	–
		i	–
		j	–
Eonycteris spelaea b	FJ541191	a	1–43(43aa)
		b	–
		c	44–49(6aa)
		d	–
		e	–
		f	–
		g	–
		h	–
		i	–
		j	–

progression of germ cells through meiosis in men (Xu et al. 2003). The full-length of the human *BOULE* gene is 2046 bp, including a coding region (849 bp), an untranslated region (325 bp), introns (872 bp), and a DAZ repeat, and the human BOULE protein consists of 283 amino acids with a molecular weight of 31.3 kDa (Xu et al. 2003). As an RNA-binding protein like other members of the DAZ family, *BOULE* comprises an RNA recognition motif (RRM) including two ribonucleoprotein signal motifs (RNP1 and RNP2) and a DAZ repeat (Reynolds and Cooke 2005) (Fig. 1). The translational or transcriptional induction of *BOULE* required an RNA-binding protein or transcriptional factor. In *Drosophila*, *BOULE* is a post-transcriptional regulator of a *CDC25* homolog called *twine,* which is required for the G2–M transition in the meiotic cell cycle during spermatogenesis (Eberhart et al. 1996). *Twine* encodes CDC25-type phosphates and activates the maturation promoting factor (MPF), consisting of the cdc2/cyclinB complex (Sigrist et al. 1995). So we hypothesize that the *BOULE* gene may participate in the regulation of spermatogenesis of bats. To test whether and how the *BOULE* gene plays a role in spermatogenesis, we cloned the *BOULE* gene from four bat species, two of them with the ability to store sperm.

The human *BOULE* has three major species of transcripts (B1, B2, and B3), which differ only in their 5′ ends, specifically in exon 1. Among these isoforms, B2 plays a major role in meiotic completion (Kostova et al. 2007). Sequencing results showed that there were at least two transcript variants in each bat species. Isoform a

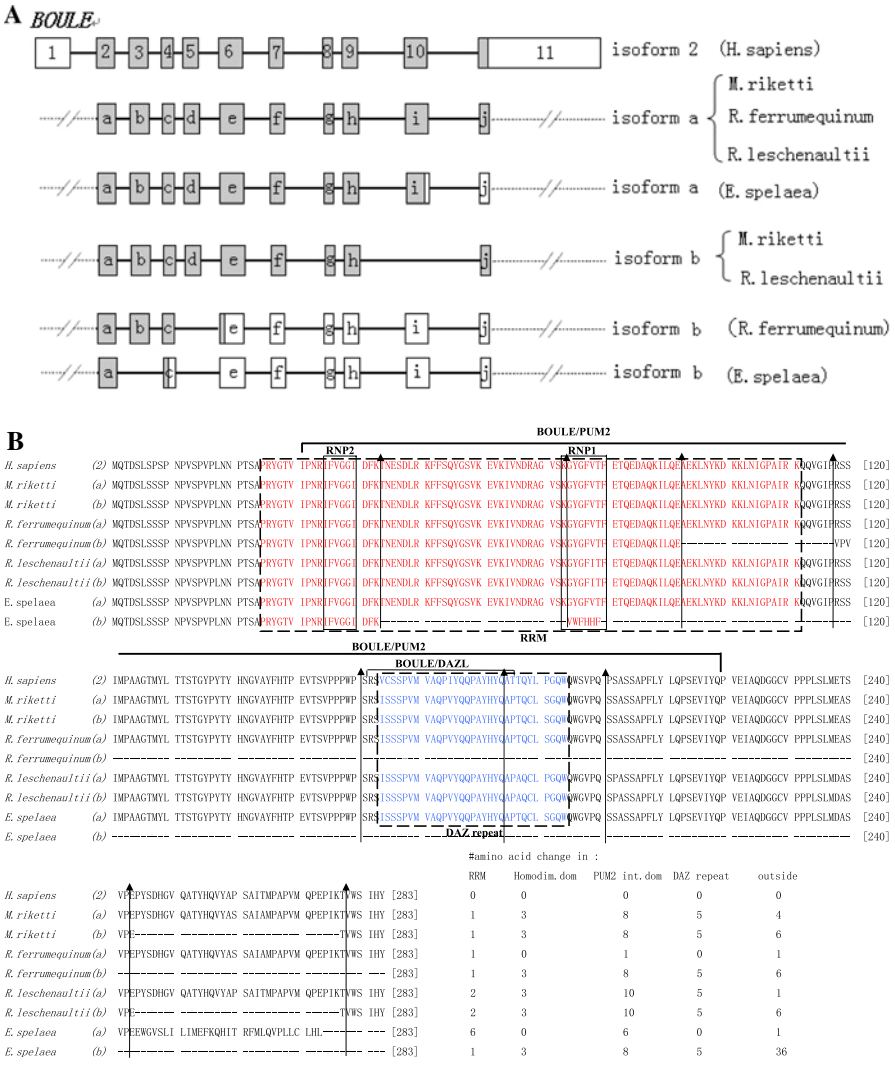

Fig. 1 *BOULE* transcript isoforms (**A**) and alignments (**B**) based on the amino acid sequences of bat and human *BOULE* genes. **A** *Numbers* within boxes indicate human exons, *letters* indicate bat exons. *Gray shading* indicates a translated region. Untranslated regions are not shaded. *Dotted lines* indicate unknown regions in bat *BOULE* transcript isoforms. **B** RNA recognition motif (RRM) is highlighted in *red* and *dotted box*; the DAZ repeat in *blue* and *dotted box*. The highly conserved RNP2 and RNP1 motifs in the RRM are boxed, and the number of amino acid changes in each domain is listed at the bottom of the figure. The region required for homodimerization and interaction with the DAZ protein and PUM2 protein (Urano et al. 2005) is indicated by a *line* above the human sequence. *Solid arrows* indicate potential splice sites

in four bat species was similar to the human *BOULE* B2, but isoform b is totally different from human *BOULE* B2 (Fig. 1). All bat *BOULE* isoforms have complete RRM domains, except for isoform b in *E. spelaea* and *R. ferrumequinum*. Members

of the DAZ family can stimulate translation inhibition by interacting with poly(A)-binding proteins (PABPs), and deletion of the RRM domain will completely abrogate the interaction with PABP (Collier et al. 2005). Moreover, RRM and the DAZ repeat of *BOULE* are required for interaction with Pumilio-2 (PUM2). By binding with Pumilio, BOULE can relieve the repression of Pumilio on a B cyclin in order to promote meiotic G2/M translation (Urano et al. 2005). Thus, the loss of RRM or the DAZ repeat of bat *BOULE* may disrupt its control of the transcription and translation of target genes, resulting in the progress of spermatogenesis being disrupted.

The multiple transcript isoforms of *BOULE* gene, including the lost functional domains suggest a new regulatory mechanism in bat spermatogenesis, and these sequences may aid in the understanding of the reproductive strategies of bats.

Acknowledgments This work was funded by a grant awarded to Shuyi Zhang under the Key Construction Program of the National 985 Project and 211 Project. Lihong Yuan is supported by the Scientific Research Starting Foundation of Guangdong Entomological Institute (2-4610) and Guangdong Academy of Sciences Scientific Research Fund (qnjj20091).

References

Collier B, Gorgoni B, Loveridge C, Cooke HJ, Gray NK (2005) The DAZL family proteins are PABP-binding proteins that regulate translation in germ cells. EMBO J 24:2656–2666

Eberhart CG, Maines JZ, Wasserman SA (1996) Meiotic cell cycle requirement for a fly homologue of human deleted in Azoospermia. Nature 381:783–785

Houston DW, King ML (2000) A critical role for Xdazl, a germ plasm-localized RNA, in the differentiation of primordial germ cells in *Xenopus*. Development 127:447–456

Karashima T, Sugimoto A, Yamamoto M (2000) *Caenorhabditis elegans* homologue of the human azoospermia factor DAZ is required for oogenesis but not for spermatogenesis. Development 127:1069–1079

Kostova E, Yeung CH, Luetjens CM, Brune M, Nieschlag E, Gromoll J (2007) Association of three isoforms of the meiotic BOULE gene with spermatogenic failure in infertile men. Mol Hum Reprod 13:85–93

Maines JZ, Wasserman SA (1999) Post-transcriptional regulation of the meiotic Cdc25 protein Twine by the Dazl orthologue Boule. Nat Cell Biol 1:171–174

Nowak RM, Walker EP, Kunz TH, Pierson ED (1994) Walker's bats of the world. Johns Hopkins University Press, Baltimore

Racey PA (1979) The prolonged storage and survival of spematozoa in Chiroptera. J Reprod Fertil 56:391–402

Racey PA, Entwistle AC (2000) Life-history and reproductive strategies of bats. In: Crighton E, Krutzsch PH (eds) Reproductive biology of bats. Academic Press, NY, pp 363–414

Reynolds N, Cooke HJ (2005) Role of the DAZ genes in male fertility. Reprod Biomed Online 10:72–80

Sigrist S, Ried G, Lehner CF (1995) Dmcdc2 kinase is required for both meiotic divisions during *Drosophila* spermatogenesis and is activated by the Twine/cdc25 phosphatase. Mech Dev 53:247–260

Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Mol Biol Evol 24:1596–1599

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG (1997) The Clustal X Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res 25:4876–4882

Urano J, Fox MS, Reijo Pera RA (2005) Interaction of the conserved meiotic regulators, BOULE (BOL) and PUMILIO-2 (PUM2). Mol Reprod Dev 71:290–298

Xu EY, Moore FL, Pera RA (2001) A gene family required for human germ cell development evolved from an ancient meiotic gene conserved in metazoans. Proc Natl Acad Sci USA 98:7414–7419

Xu EY, Lee DF, Klebes A, Turek PJ, Kornberg TB, Reijo Pera RA (2003) Human *BOULE* gene rescues meiotic defects in infertile flies. Hum Mol Genet 12:169–175