Blackwell Publishing

Asian J Androl 2005; 7 (3): 277–288
DOI: 10.1111/j.1745-7262.2005.00040.x

AJA

·*Original Article*·

Expression of a novel alternative transcript of the novel retinal pigment epithelial cell gene *NORPEG* in human testes

Wa Yuan, Ying Zheng, Ran Huo, Li Lu, Xiao-Yan Huang, Lan-Lan Yin, Jian-Min Li, Zuo-Min Zhou, Jia-Hao Sha

Key Laboratory of Reproductive Medicine, Department of Histology and Embroyology, Nanjing Medical University, Nanjing 210029, China

Abstract

Aim: To identify a novel alternative transcript of the novel retinal pigment epithelial cell gene (*NORPEG*) expressed in the human testis. **Methods:** A human testis cDNA microarray was established and hybridized with cDNA probes from human fetal testes, adult testes and human spermatozoa. Differentially expressed clones were sequenced and analyzed. One of these clones was a short transcript of *NORPEG* which we proceeded to analyze by RT-PCR. **Results:** The novel short alternative transcript of *NORPEG* was isolated and named *sNORPEG*. It was 3486 bp in length and contained a 2952-bp open reading frame, encoding a 110.4-kDa protein of 983 amino acids. Amino acid sequence analysis showed that the *sNORPEG* protein contains six ankyrin repeats and two coiled-coil domains. It shares a high homology with the NORPEG and ankycorbin proteins in both its sequence and motifs. Blasting the human genome database localized *sNORPEG* to human chromosome 5p13.2–13.3. Expression profiles showed that *sNORPEG* was expressed in human fetal testes, adult testes and spermatozoa. Moreover, *sNORPEG* was found to be ubiquitously expressed in human tissues. **Conclusion:** *sNORPEG* is expressed in different developmental stages of the testis and encodes a protein that may have roles in human testis development and spermatogenesis. *(Asian J Androl 2005 Sep; 7: 277–288)*

Keywords: alternative transcript; *NORPEG*; testis development; spermatogenesis

1 Introduction

Actin cytoskeleton structures are essential for a wide variety of cell functions, including cell shape change, cell motility, cell adhesion, cell polarity and cytokinesis. Many actin-associated proteins with functions related to actin dynamics have been isolated and characterized. For example, the recently identified ankycorbin is an actin cytoskeleton-associated protein that may be involved in actin cytoskeleton maintenance and/or reorganization [1]. During testis development and spermatogenesis, actin and actin associated proteins play a crucial role in many important activities, including the Sertoli-germ cell adherens junction dynamics [2], spermiogenesis [3] and acrosomal transformation [4]. The investigation of proteins related to actin dynamics in testes may provide information about testicular function and physiology.

In this study we identified a short alternative transcript of the novel retinal pigment epithelial cell gene (*NORPEG*) in the human testis and named it *sNORPEG*. The discovery was based on the analysis of cDNA probe hybridizations with a human cDNA microarray con-

Correspondence to: Dr Jia-Hao Sha, Key Laboratory of Reproductive Medicine, Nanjing Medical University, Nanjing 210029, China.
Tel/Fax: +86-25-8686-2908
E-mail: shajh@njmu.edu.cn
Received 2004-08-11 Accepted 2005-02-12

structed in this laboratory. We found that *sNORPEG* encodes a protein which is highly homologous with both the NORPEG protein and the recently identified actin-associated protein ankycorbin [1]. Both the NORPEG and ankycorbin proteins are thought to have functions related to the cellular cytoskeleton. This paper describes the novel alternative transcript *sNORPEG* in human testes and discusses its possible roles in testis development and spermatogenesis.

2 Materials and methods

2.1 cDNA microarray construction and hybridization

A human testis cDNA microarray was constructed in order to study gene expression in testis development. A total of 9216 positive phage clones were selected randomly from the Human Testis Insert λ phage cDNA library (HL5503U; Clontech, Palo Alto, CA, USA) and amplified by PCR. The PCR products were spotted onto a membrane to make the human testis cDNA microarray. This microarray was hybridized with ^{33}P-labeled cDNA probes prepared from the mRNA of human fetal testes, adult testes and human spermatozoa. The microarray was scanned by an FLA-3000A plate/fluorescent image analyzer (Fuji Photo Film, Tokyo, Japan). The radioactive signal intensity of each spot was linearly scanned and read using the Array Gauge software (Fuji Photo Film, Tokyo, Japan). After subtraction of the background from an area where no PCR product was spotted, clones with intensities over 10 were considered positive. The hybridization intensities of the corresponding dots from adult and fetal samples were compared. If the intensity comparison between the samples yielded a difference ≥3-fold, then the clones were considered differentially expressed. The cDNA clones showing differential expression patterns between fetal and adult testes were selected and analyzed. Protocols for human testis cDNA microarray construction, adult testis, fetal testis and spermatozoa cDNA probe preparation, hybridization, and signal analysis have been described elsewhere in detail [5–7].

2.2 Sequence identification and analysis

The cDNA clones that were found to be differentially expressed in adult testes and fetal testes were purified with mini-preps (QIAprep Spin Miniprep Kit, Qiagen, Hilden, Germany) and then sequenced by an ABI377 automatic sequencer (Perkin-Elmer, Norwalk, CT, USA). The sequence of the forward sequencing primer was 5'-CCATTGTGTTGGTACCCGGGAATTCG-3' and the sequence of the reverse sequencing primer was 5'-ATAAGCTTGCTCGAGTCTAGAGTCGAC-3'. The results were analyzed with BLAST (http://www.ncbi.nlm.nih.gov/BLAST/) and SMART (http://smart.embl-heidelberg.de/) database programs with the goal of identifying homologous genes and proteins. The nucleotide and the deduced amino acid sequences were also analyzed using Gene Runner software (http://www.generunner.com). Meanwhile, highly homologous proteins were compared by ClustalW (http://www.ebi.ac.uk/clustalw/). The novel *NORPEG* transcript *sNORPEG* was isolated and identified. The promoters of *sNORPEG* and *NORPEG* were analyzed by PROSCAN version 1.7 (http://bimas.dcrt.nih.gov/molbio/proscan/).

2.3 Expression profile of sNORPEG in different developmental stages of testis and spermatozoa

The expression profile of *sNORPEG* in a human adult testis (aged 43 years), fetal testis (gestational age ~6 months) and spermatozoa was determined using RT-PCR. Human adult and fetal testes total RNA was isolated using Trizol Reagent (Gibco BRL, Grand Island, New York, USA). Ejaculate spermatozoa from a normal male (WHO, 1999 criteria) were allowed to liquefy for 1 h at room temperature, washed twice in phosphate-buffered saline (pH 7.4) and the total RNA were extracted from the sediment.

Reverse transcription reactions from the extracted RNA samples were performed in 15 μL of reaction mixture. First 2 μL (about 5 μg) total RNA, 1 μL random hexamer primer (0.2 μg/mL, Sangon, Shanghai, China) and 6 μL diethyl pyrocarbonate (DEPC) treated water were mixed and incubated at 70 ºC for 5 min; then 3 μL AMV 5 × buffer, 0.75 μL dNTP (20 mmol/L), 0.25 μL Rnasin (40 U/μL), 1 μL AMV reverse transcriptase (Promega, Madison, USA) and 1 μL DEPC water were added and incubated at 42 ºC for 1.5 h, and then held at 90 ºC for 5 min. PCR was performed with *sNORPEG* specific primers. The specific primers were designed to overpass two introns to prevent contamination of genomic DNA. Primers were as follows: upstream 5'-TGCTGGCTGTATGTTATGC-3', and downstream 5'-GGTAGTATCTTGGGCTGTC-3'. The amplified fragment of *sNORPEG* was 279 bp in size. The upstream primer was located in the exclusive exon of the *sNORPEG* transcript but the downstream primer was homologous with that

of *NORPEG*. Primers were synthesized at BioAsia Company (Shanghai, China). Human β-actin was used as a positive control. The following human β-actin primers were used: upstream 5'-CGGTTGGCCTTGGGGTTCAGGGGG-3', and downstream 5'-ATCGTGGGGGCGCCCCAGGCACCA-3'. The PCR thermal cycling conditions program consisted of an initial denaturation at 94 °C for 5 min, followed by 35 30-s cycles of denaturation at 94 °C, annealing at 54 °C for 30 s, extension at 72 °C for 1 min and an additional extension at 72 °C for 7 min. The PCR products were analyzed by 1.5 % (w/v) agarose gel electrophoresis.

2.4 Expression profile of sNORPEG in different tissues

The expression profile of *sNORPEG* in different tissues was assessed using RT-PCR. Multiple tissue cDNA panels were obtained from the commercial Human Multiple Tissue cDNA (MTC) Panel I and II kit (Cat# K1420-1 and K1421-2, Clontech), which included 16 human tissues (testis, skeletal muscle, liver, pancreas, brain, lung, kidney, heart, placenta, spleen, thymus, prostate, ovary, small intestine, colon and peripheral blood leukocytes). Primers and PCR conditions were the same as that described above. G3PDH was used as a positive control. Its upstream primer was 5'-TGAAGGTCGGAGTCAACGGATTTGGT-3', and downstream primer was 5'-CATGTGGGCCATGAGGTCCACCAC-3'. The desired fragment was 983 bp. PCR conditions were performed according to the manufacturer's instructions as follows: denaturation at 95 °C for 30 s, subsequent annealing and extension at 68 °C for 3 min. The first cycle had a denaturation period of 1 min. The last cycle had an extension period of 3 min at 68 °C. Thirty-six cycles of PCR were performed. The plasmid containing *sNORPEG* from the human testis large insert cDNA library (HL5503U, Clontech) was used as a positive control of PCR amplification. PCR products were analyzed by 2 % (w/v) agarose gel electrophoresis.

2.5 TA clone DNA sequencing

RT-PCR analysis revealed an unexpected fragment in the lung, testis and sperm. The PCR reaction products were gel-purified (QIAquick Gel Extraction Kit, Qiagen) and the purified PCR products were cloned into a pinpoint xa-1 T-vector (Cat# V2610, Promega). Positive clones after transformation were selected and sequenced (sequenced by BioAsia) with a PinPoint Vector Sequencing Primer (sequence 5'-CGTGACGCGGTGCAGGGCG-3', Promega). The publicly available blast program (http://www.ncbi.nlm.nih.gov/BLAST/) was utilized to compare the sequence with expressed sequence tag database (dbEST) and the human genome.

3 Results

3.1 cDNA microarray hybridization

One of the differentially expressed genes was isolated. The hybridization intensities of this gene in fetal and adult testes were 9.88 and 30.90, respectively. The intensity of this gene's expression in the adult testis was about 3-fold stronger than that in fetal testes (Figure 1). The hybridization intensity of the newly identified gene in spermatozoa was 15.54, indicating that it was expressed in germ cells. Subsequent analyses as described below indicated that this novel differentially expressed gene represented a short alternative transcript of *NORPEG* and therefore it was named *sNORPEG*.

3.2 Sequence identification and analysis of sNORPEG

The full cDNA length of *sNORPEG* was 3486 bp and had a 2952-bp open reading frame from 493 bp to 3444 bp, encoding a 110.4-kDa protein of 983 amino acids. The methionine at 493–495 bp was almost certainly the site of initiation because there was an up-stream stop code at 427–429 bp (Figure 2).

Blast analysis showed that *sNORPEG* (GenBank accession number AY317139) was highly homologous with *NORPEG* (GenBank accession number AF155135) and with the *AB037755*, *BC028681* and *AY354204* gene transcripts, all of which are classified as derived from the RAI14 gene (Retinoic Acid Induced gene) in GenBank and belong to the UniGene Cluster Hs. 368605. Blast search in the human genome database showed that the

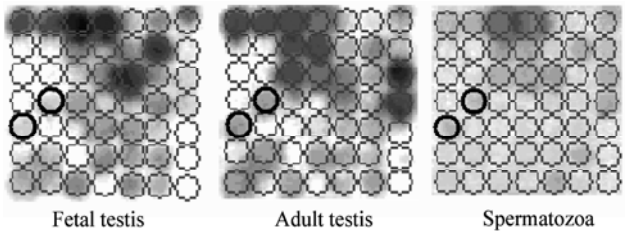

Figure1. cDNA hybridization images showing differential expression of *sNORPEG* in the fetal testis, adult testis and spermatozoa. Black rings indicate *sNORPEG* cDNA. The intensity in fetal testis, adult testis and spermatozoa was 9.88, 30.90 and15.54, respectively.

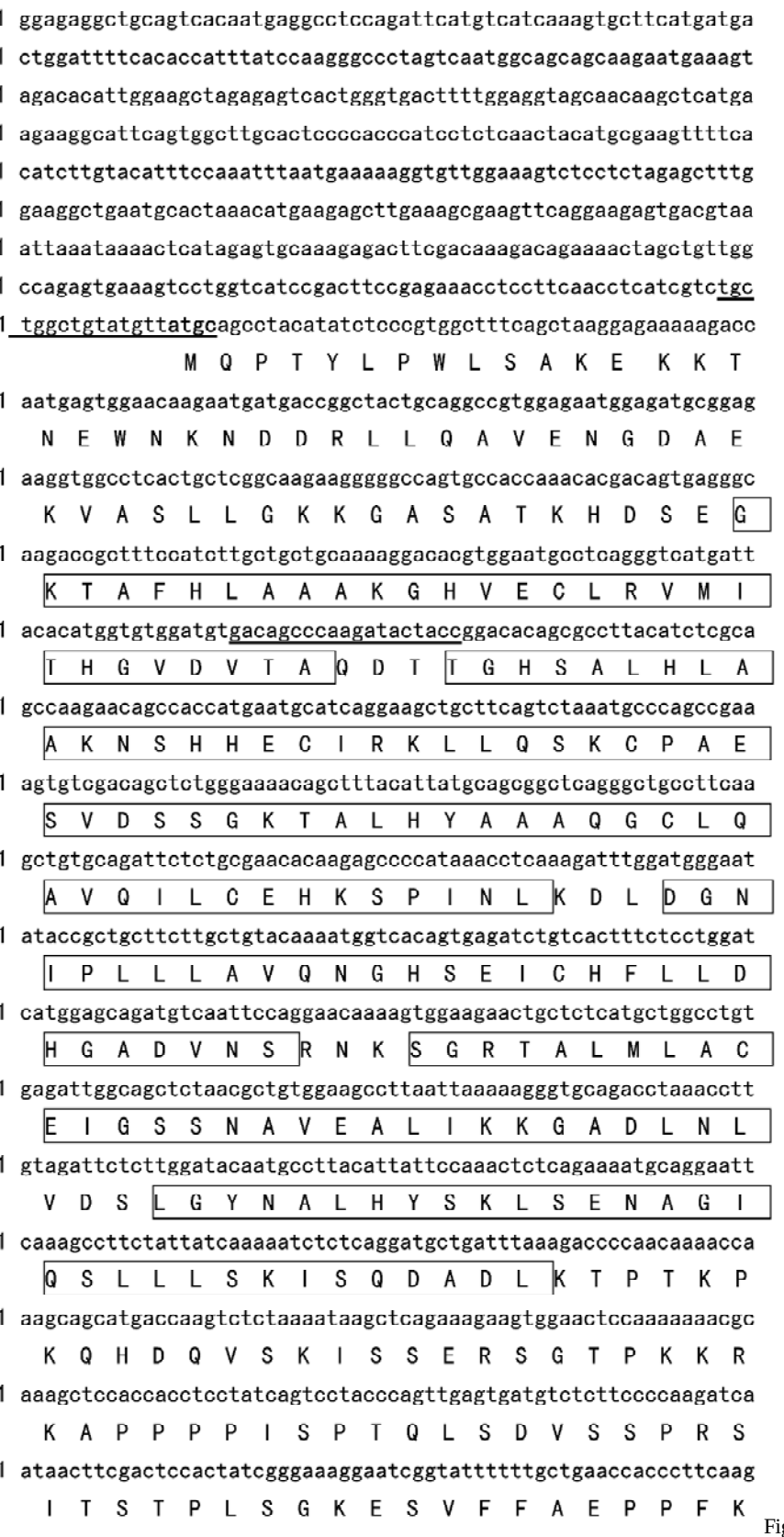

Figure 2 (to be continued).

(continued)

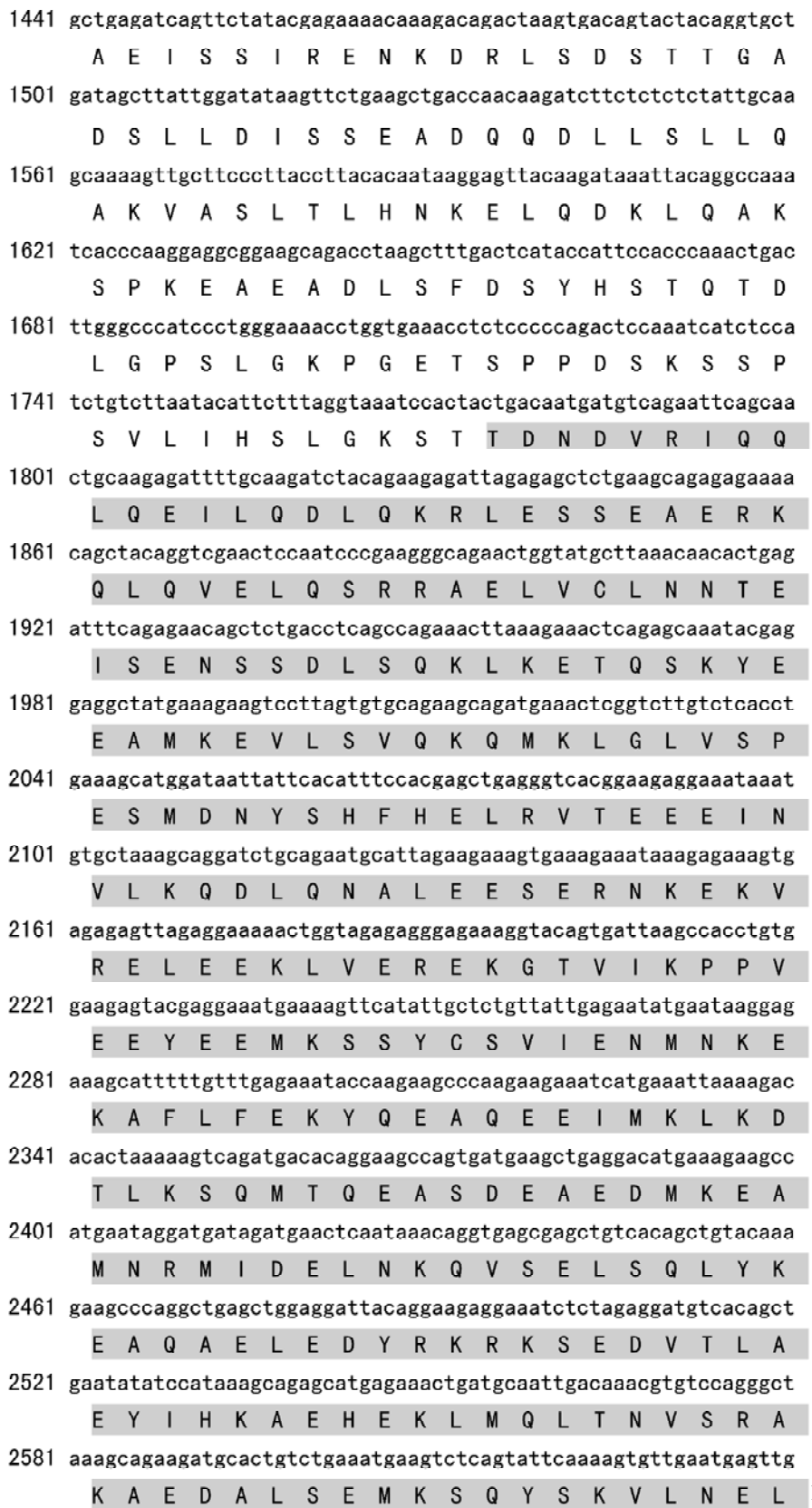

Figure 2 (to be continued).

(continued)

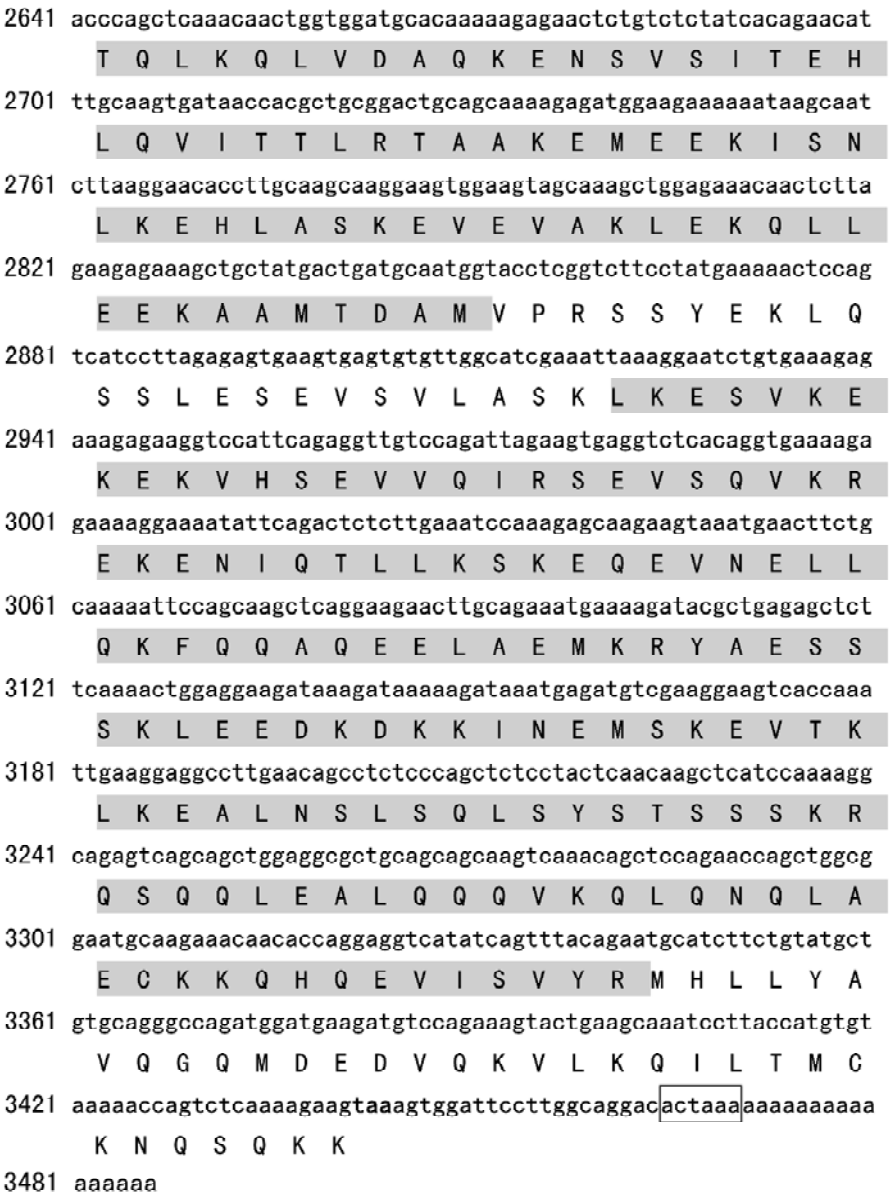

Figure 2. Nucleic acid and deduced amino acid sequences of the cDNA for *sNORPEG*. Underlining shows the specific PCR primers for the determination of the expression profile. The upstream primer is located in the specific region of *sNORPEG*. The downstream primer is homologous with that of *NORPEG*. The initiation and stop codons are in bold type. Poly A signal site is boxed. Six ankrin repeats are boxed and the coiled-coil domain is in shadow. Ankyrin repeats: amino acids (aa) 55-84, aa 88-120, aa 121-150, aa 154-183, aa 187-216, aa 220-250. Coiled-coil domain: aa 428-786, aa 810-950.

RAI14 gene consists of 23 exons and 22 introns and is localized to human chromosome 5p13.2–13.3. Splicing comparison of *sNORPEG* with its homologous genes indicated that *sNORPEG* had 20 exons. Exon 7 was its unique exon at the 5' terminus. The last exon of *sNORPEG* (103 bp) at the 3' terminus was shorter than that of *NORPEG* (1927 bp). Sequence analysis (http://l25.itba.mi.cnr.it/~webgene/wwwHC_polya.html) indicated that while *NORPEG* had a typical poly A signal at its 3' terminal (AATAAA, nt 4880–4885), *sNORPEG* had an atypical poly A signal (ACTAAA, nt 3465–3470) (Figure 2). Promoter scan software predicted that *sNORPEG* and

NORPEG had different putative promoter regions. One of the predicted promoters (P1) was located at position -250 bp to -1 bp upstream of the 5' terminal of *NORPEG*. The transcripts of *NORPEG* and *AB037755* could be driven by the P1 promoter. Another predicted promoter (P2) was located between -982 bp and -732 bp upstream of the 5' end of *sNORPEG*. The expression of *sNORPEG*, *BC028681* and *AY354204* could be initiated at the P2 promoter (Figure 3A). The above analysis indicated that *sNORPEG* (3486 bp) was a short alternative transcript of *NORPEG* (4925 bp).

Blast protein analysis showed that the sNORPEG protein was highly homologous with the NORPEG (98 % identity) and ankycorbin (84 % identity) proteins. The ankycorbin protein was encoded by a mouse *NORPEG*-homologous gene. The sNORPEG, NORPEG and ankycorbin proteins were found to contain 983, 980 and 979 amino acid residues, respectively. As shown in Figure 4, the sequence from the 13th amino acid residue to the end of the sNORPEG protein matched the sequence from the 16th amino acid residue to the end of the NORPEG protein.

Analysis of the amino acid sequence using SMART software (http://smart.embl-heidelberg.de/) revealed that

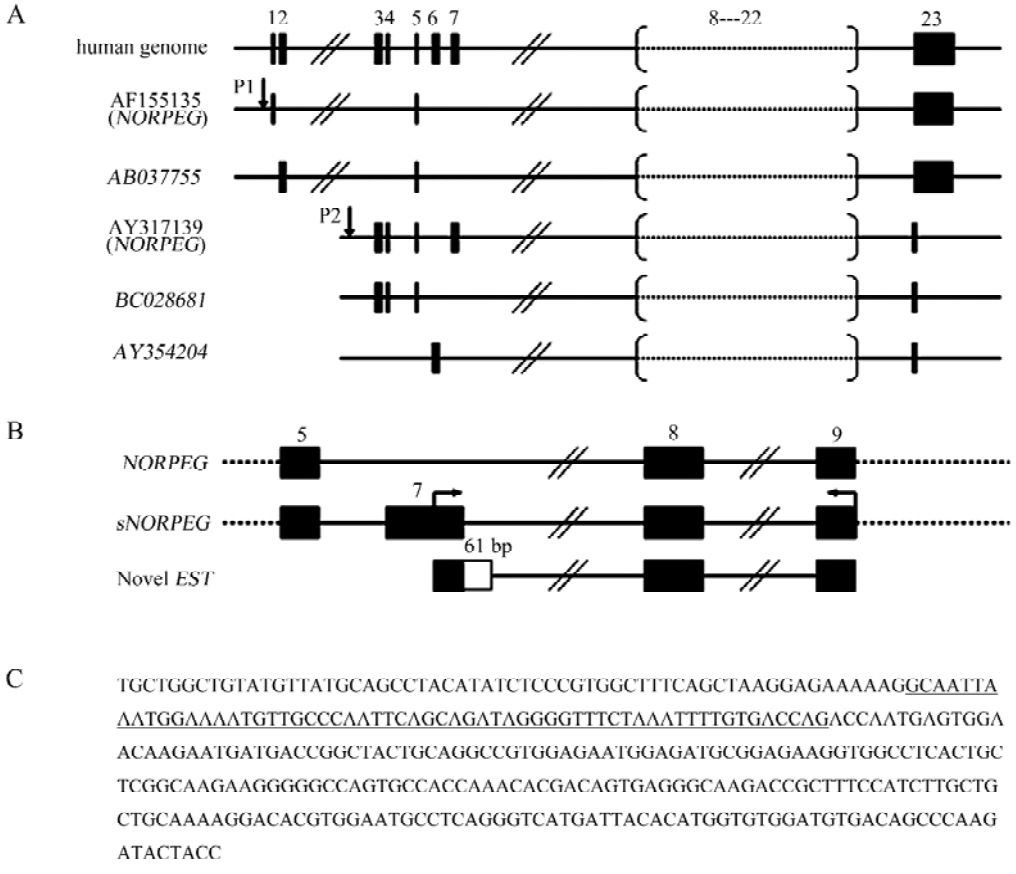

Figure 3. (A) Transcript and splicing comparison of *sNORPEG* with its homologous genes. Homologues originate from one gene and consist of 23 exons. *sNORPEG* contains 20 exons and exon 7 is its specific exon. *NORPEG* consists of 18 exons and exon 1 is its specific exon. Arrows indicate the loci of the two putative promoters, P1 and P2. Exons are represented by rectangles; horizontal lines represent intron sequences. Identical exons (8-22) are bracketed and omitted. Introns >10 kb are broken by diagonal hatch marks. (B) Comparison of the novel *EST* with *NORPEG* and *sNORPEG*. The arrows indicate the primers used for the detection of *sNORPEG*. The exon annotation follows that of panel A. The black rectangles represent the exon. Other exons are omitted. The white rectangle represents the part of the *EST* which was different from the other sequences. (C) The sequence of the novel *EST* is listed. The 61bp extra sequence of the novel *EST* is underlined.

```
A    MQPTYLPWLS AKEKK---TNEWNKNDDRLLQAVENGDAEKVASLLGKKGASATKHDSEGKT 58
B    ----------MKSLKAKFRKSDTNEWNKNDDRLLQAVENGDAEKVASLLGKKGASATKHDSEGKT 55
C    ----------MKSLKAKFRKSDTNEWNKNDDRLLQAVENGDAEKVASLLGKKGASATKHDSEGKT 55
                *  **   *  ********************************************

A    AFHLAAAKGHVECLRVM I THGVDVTAQDTTGHSALHLAAKNSHHEC I RKLLQSKCPAESV 118
B    AFHLAAAKGHVECLRVM I THGVDVTAQDTTGHSALHLAAKNSHHEC I RRLLQSKCPAESV 115
C    AFHLAAAKGHVECLKVMVTHGVDVTAQDSSGHSALHVAAKNGHPEC I RKLLQYKSPAENI 115
     ************** ** ********** ****** **** * **** *** * ***

A    DS SGKTALHYAAAQGCLQAVQI LCEHKSP I NLKDLDGN I PLLLAVQNGHSEI CHFLLDHG 178
B    DS SGKTALHYAAAQGCLQAVQI LCEHKSP I NLKDLDGN I PLLLAVQNGHSEI CHFLLDHG 175
C    DNSGKTALHYAAAQGCLQAVQLLCEHKSP I NLKDLDGN I PLLV AVQNGHSEACHFLLDHG 175
     ********************** ********************  ********  ********

A    ADVNSRNK SGRTALMLACE IGSSNAVEAL IKKGADLNLVDSLGYNALHYSKLSENAG I QS 238
B    ADVNSRNK SGRTALMLACE IGSSNAVEAL IKKGADLNLVDSLGYNALHYSKLSENAG I QS 235
C    ADVNSRDKNGRTALMLACETGSSNTVDAL IKKGADLS LVDSLGHNALHYSKLSENAGI QN 235
     ******  *  ********** **** *  ********* ******  ****************

A    L L LSK I SQDADL KTPT KPKQHDQVSK I SSE RSGTPKKRK APPPP I SPTQL SDVSSPRS I T 298
B    L L LSK I SQDADL KTPT KPKQHDQVSK I SSE RSGTPKT RK APPPP I SPTQL SDVSSPRS I T 295
C    L L LSK I SQDADL KTPT KPKQHDQVSK I SSE RSGTPKKRK APPPP I SPTQL SDVSSPRS I T 295
     ********** * ********************** **  ********************

A    STP LSGKESVFFAE PPFKAE I SS I RENKDR LSDSTTGADSLLD I SSEADQQDL L S L LQAK 358
B    STP LSGKESVFFAE PPFKAE I SS I RENKDR LSDSTTGADSLLD I SSEADQQDL L S L LQAK 355
C    STP LSGKESVFFAE APFKAE I SS I QENKDR LSDSTAGADSLLD I SSEADQQDL L V L LQAK 355
     ************  **********  ***** ***** ******************  * ****

A    VASLT LHNKE LQDK LQAKSPK---- EAEAD LSFDSYHSTQTD LGPS LGKPGET SPPDSKSSP 416
B    VASLT LHNKE LQDK LQAKSPK---- EAEAD LSFDSYHSTQTD LGPS LGKPGET SPPDSKSSP 413
C    VASLT LHNKE LQDK LQAKSPKDK EAEAD LSFQSFHSTQTD LAPS PGKASD I PSSDAKSSP 415
     **********************    ********  * ******* **  **         * ****

A    SV L IHSLGKSTTDNDVR I QQLQE ILQDLQKRLESSEAERKQLQVELQSRRAE L VCLNNTE 476
B    SV L IHSLGKSTTDNDVR I QQLQE ILQDLQKRLESSEAERKQLQVELQSRRAE L VCLNNTE 473
C    --PVEHPAGT STTDNDV I I RQLQDSLHDLQKRLESSEAEKKQLQDELQSQRTD T LCLNNTE 474
           *   **********  *  ***   *  ************  ****  **** *       ******

A    I SENSSDLSQK LKETQSKYEEAMKEVLSVQKQMK LGLVSPESMDNYSHFHELRVTEEE I N 536
B    I SENSSDLSQK LKETQSKYEEAMKEVLSVQKQMK LGLVSPESMDNYSHFHELRVTEEE I N 533
C    I SENGSDLSQKLKETQSKYEEAMKEVLSVQKQMK LGLLSQESADGYSHLREA -PADED I D 533
     **** ********************************** ***  *  **  *  ***   *       *  *

A    VLKQDLQNALEESERNKEKVRELEEKLVEREKGTVI KPPVEEYEEMKSSYCSV I ENMNKE 596
B    VLKQDLQNALEESERNKEKVRELEEKLVEREKGTVI KPPVEEYEEMKSSYCSV I ENMNKE 593
C    TLKQDLQKAVEESARNKERVRELETKLAEKEQAEATKPPAEACEE L RSSYCSV I ENMNKE 593
      ******  *  *** **** ***** **  *  *        *** *    **   ************

A    KAFLFEKYQEAQEEIMKLKDTLKSQMTQEASDEAEDMKEAMNRMIDELNKQVSELSQLYK 656
```

Figure 4 (to be continued).

(continued)

```
B    KAFLFEKYQEAQEEIMKLKDTLKSQMTQEASDEAEDMKEAMNRMIDELNKQVSELSQLYK 653
C    KAFLFEKYQQAQEEIMKLKDTLKSQMPQEAPDDSGDMKEAMNRMIDELNKQVSELSQLYR 653
     ********* ***************  **** *********************

A    EAQAELEDYRKRKSLEDVTAEY IHKAEHEKLMQLTNVSRAKAEDALSEMKSQYSKVLNEL 716
B    EAQAELEDYRKRKSLEDVTAEY IHKAEHEKLMQLTNVSRAKAEDALSEMKSQYSKVLNEL 713
C    EAQAELEDYRKRKSLEDA--AEY IHKAEHERLMHVSNLSRAKS EEALSEMKSQYSKVLNEL 712
     ***************** ********** * *  * **** * ***************

A    TQLKQLVDAQKENSVS I TEHLQV I TTLRTAAKEMEEKISNLKEHLAS KEVEVAKLEKQL L 776
B    TQLKQLVDAQKENSVS I TEHLQV I TTLRTAAKEMEEKISNLKEHLAS KEVEVAKLEKQL L 773
C    TQLKQLVDAHKENSVS I TEHLQV I TTLRTT AKEMEEKISALTGHLANKEAEVAKLEKQL A 772
     ********* ******************** ********* *  *** ** *********

A    EEKAAMTDAMVPRSSYEKLQSSLESEV SV LASKLKESVKEKEKVHSEVVQ I RSEVSQVKR 836
B    EEKAAMTDAMVPRSSYEKLQSSLESEV SV LASKLKESVKEKEKVHSEVVQ I RSEVSQVKR 833
C    EEKAA VSDAMVPKSSYEKLQASLESEVNALATKLKESVR EREKAHSEV AQVRSEVSQARR 832
     *****  ***** ******* ******  ** ****** * ** **** * ******  *

A    EKENIQTLLKSKEQEVNELLQKFQQAQEELAEMKRYAESSSKLEEDKDKK I NEMSKEVTK 896
B    EKENIQTLLKSKEQEVNELLQKFQQAQEELAEMKRYAESSSKLEEDKDKK I NEMSKEVTK 893
C    EKDNIQTLLKAKEQEVTALVQKFQRAQEELAGMRRCSETSSKLEEDKDEK I NEMTREVLK 892
     **  ******* *****  * **** ****** *  *  * ********* * ****  ** *

A    LKEALNSLSQLSYSTSSSKRQSQQLEALQQQVKQLQNQLAECKKQIIQEV I SVYRMIILLYA 956
B    LKEALNSLSQLSYSTSSSKRQSQQLEALQQQVKQLQNQLAECKKQHQEV I SVYRMHLLYA 953
C    LKEALNSLSQLSYSTSSSKRQSQQLDLLQQQVKQLQNQLAECKKHHQEV I SVYRMHLLYA 952
     ************************  ****************  ***************

A    VQGQMDEDVQKVLKQ I LTMCKNQSQKK 983
B    VQGQMDEDVQKVLKQ I LTMCKNQSQKK 980
C    VQGQMDEDVQKVLKQ I LTMCKNQSQKK 979
     ***************************
```

Figure 4. Amino acid alignment of the sNORPEG, NORPEG and ankycorbin proteins. The sNORPEG protein had 84% identity with ankycorbin and 98% identity with the NORPEG protein. *Marks identical amino acids. A: sNORPEG protein; B: NORPEG protein; C: Ankycorbin protein

the sNORPEG and NORPEG proteins had identical domains. They both contain six ankyrin repeats in the N-terminal region and two coiled-coil domains in the C-terminal region (Figure 5). Ankycorbin similarly has six ankyrin repeats and a long coiled-coil domain [1].

3.3 Expression profiles of sNORPEG

Expression profiles in different developmental stages of the testis and spermatozoa showed that *sNORPEG* is expressed in fetal testes, adult testes and spermatozoa (Figure 6). Multi-tissue PCR data indicated that *sNORPEG* is widely expressed in human tissues (Figure 7). In addition, unexpected 340 bp band was detected and sequenced (Figure 3C). Blast searches revealed that it was a novel EST of *NORPEG* (Figure 3B, EST id number: 25994518; GenBank accession number: CK433905).

Figure 5. Schematic drawing of the sNORPEG protein structure. The domain organization of the sNORPEG protein was determined with the Web-based tool, Smart program.

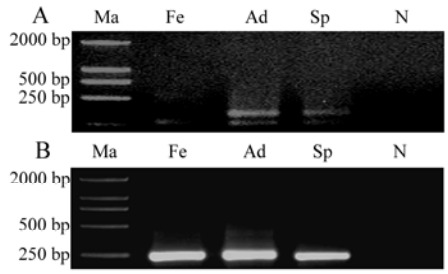

Figure 6. Expression profile of (A): sNORPEG and (B): β-actin as control in fetal testis, adult testis and spermatozoa. *sNORPEG* was expressed in fetal testis, adult testis and spermatozoa with a PCR product of 279 bp in size. An unexpected fragment about 340 bp was expressed in adult testis and spermatozoa. Ma, marker; Ad, adult testis; Fe, fetal testis; Sp, spermatozoa; N, negative control.

4 Discussion

In the present study, a testis cDNA microarray was used to identify genes related to testis development and spermatogenesis. A novel short alternative transcript of *NORPEG* was cloned and identified with this method and given the name *sNORPEG*. Bioinformatics analysis and experimental results suggest that *sNORPEG* may play a role in testis development and spermatogenesis.

Sequence analysis showed that the sNORPEG protein contains six ankyrin repeats and two coiled-coil domains. The ankyrin repeat is one of the most common protein sequence motifs. It comprises approximately 33 amino acids and occurs in at least four consecutive copies [8–10]. Ankyrin repeats have been found in proteins as diverse as Cdk inhibitors, signal transduction and transcriptional regulators, cytoskeletal organizers, developmental regulators, and toxins [9]. It is generally assumed that the ankyrin repeats play an important role in protein-protein interactions [11]. The C-terminal domain contains the coiled-coil domain which is a highly versatile protein folding motif related to protein-protein interaction [12, 13]. The coiled-coil domain exists in some actin-binding proteins, such as tara [14], tropomyosin [15] and KRAP [16]. Thus it is likely that these two conserved domains are involved in mediating protein-protein interactions of the sNORPEG protein with its partner proteins.

The sNORPEG protein shows 84 % identity with ankycorbin and 98 % identity with NORPEG. These three homologous proteins have in common six ankyrin repeats and coiled-coil domains, suggesting that the sNORPEG protein may have a function similar to that of the NORPEG and ankycorbin proteins. Prior evidence suggests that the NORPEG protein is associated with the cytoskeleton [17]. Likewise, ankycorbin is highly concentrated at cortical actin cytoskeleton structures in terminal web and cell-cell adhesion sites and stress fibers. Ankycorbin appears to be an actin cytoskeleton-associ-

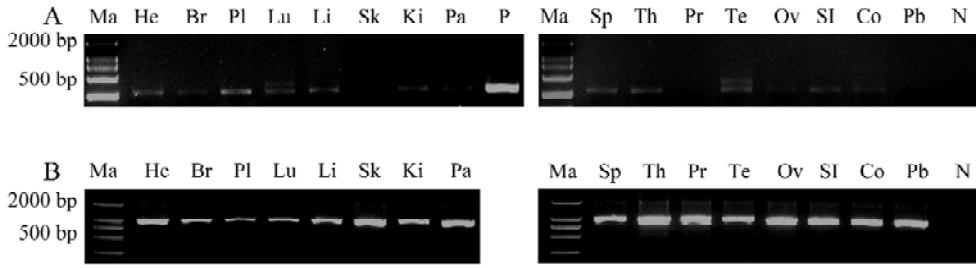

Figure 7. Tissue distribution of *sNORPEG* (A) and *G3PDH* as control (B) after electrophoresis. *sNORPEG* was widely expressed in human tissues with a PCR product of 279 bp in size. An unexpected fragment was detected in lung and testis which was 340 bp in size. Ma, marker; He, heart; Br, brain; Pl, placenta; Lu, lung; Li, liver; Sk, skeletal muscle; Ki, kidney; Pa, pancreas; P, sNORPEG plasmid as positive control; Sp, spleen; Th, thymus; Pr, prostate; Te, testis; Ov, ovary; SI, small intestine; Co, colon; Pb, peripheral blood leukocyte; N, negative control.

ated protein and may be involved in actin cytoskeleton maintenance and/or reorganization [1]. The homology of the sNORPEG protein to the NORPEG and ankycorbin proteins is consistent with our hypothesis that its function may be related to actin cytoskeleton dynamics and that it may play a role in the actin-related events that occur during testis development and spermatogenesis.

Actin filaments are concentrated in specific regions of spermatogenic cells and Sertoli cells. In spermatogenic cells they occur in intercellular bridges and in the subacrosomal space. In Sertoli cells they are abundant in the ectoplasmic specializations and in regions adjacent to the tubulobulbar processes of spermatogenic cells [18]. In the testis there exists an important cell-cell actin-based adherens junction, the dynamics of which are important in permitting the timely movement of germ cells across the epithelium [19]. Actin and actin-associated proteins are involved in regulating the Sertoli-germ cell actin-based adherens junction assembly and disassembly. It is possible that the sNORPEG protein may play an important role in this process. sNORPEG was not only expressed in human fetal testes, adult testes and spermatozoa, but also ubiquitously expressed in other human tissues. These findings suggest that the sNORPEG protein may also be involved in actin cytoskeleton dynamics in these tissues.

Alternative transcripts can be generated during gene expression by using promoters and transcription factors that activate transcription at different start sites upstream or downstream of the usual site, by incorporation of alternate exons, by germ cell-specific splicing events, and by using alternate initiation sites for polyadenylation [20]. Compared with *NORPEG*, *sNORPEG* has a different putative promoter and a shorter 3'-untranslated region. The *sNORPEG* transcript may be generated by an alternative promoter usage and an alternate polyadenylation signal. Our study revealed an additional unknown alternatively spliced variant of *NORPEG* that was only expressed in the lung, adult testis and spermatozoa. The restricted expression of this unknown splice variant suggests that its function would be specific to the lung and testis. The diversity of these alternative transcripts gives us new insight into the complex genetic regulation of *NORPEG*.

In summary, a novel mRNA transcript, *sNORPEG*, was identified which has several important conserved domains. The homologous motif properties of the protein encoded by *sNORPEG* and its expression in human fetal testes, adult testes and spermatozoa suggest that it may participate in actin cytoskeleton dynamics during testis development and spermatogenesis. Further study will be required to elucidate the functional role and regulatory mechanisms of the sNORPEG protein in testicular development and spermatogenesis.

Acknowledgment

This work was supported by grants from China National 973 Program (No. G1999055901), National Natural Science Foundation of China (No. 30425006), the Foundation of Science and Technology of Jiangsu Province, China (No. BG2003028) and the Foundation of National Ministry of Science and Technology, China (No. 2004CCA06800).

References

1 Peng YF, Mandai K, Sakisaka T, Okabe N, Yamamoto Y, Yokoyama S, *et al*. Ankycorbin: a novel actin cytoskeleton-associated protein. Genes Cells 2000; 5: 1001–8.
2 Lui WY, Lee WM, Cheng CY. Sertoli-germ cell adherens junction dynamics in the testis are regulated by RhoB GTPase via the ROCK/LIMK signaling pathway. Biol Reprod 2003; 68: 2189–206.
3 Bouchard MJ, Dong Y, McDermott BM Jr, Lam DH, Brown KR, Shelanski M, *et al*. Defects in nuclear and cytoskeletal morphology and mitochondrial localization in spermatozoa of mice lacking nectin-2, a component of cell-cell adherens junctions. Mol Cell Biol 2000; 20: 2865–73.
4 Scarlett CJ, Lin M, Aitken RJ. Actin polymerisation during morphogenesis of the acrosome as spermatozoa undergo epididymal maturation in the tammar wallaby (*Macropus eugenii*). J Anat 2001; 198 (Pt 1): 93–101.
5 Sha J, Zhou Z, Li J, Yin L, Yang H, Hu G, *et al*. Identification of testis development and spermatogenesis-related genes in human and mouse testes using cDNA arrays. Mol Hum Reprod 2002; 8, 511–7.
6 Sha JH, Zhou ZM, Li JM, Lin M, Zhu H, Zhu H, *et al*. Expression of a novel bHLH-Zip gene in human testis. Asian J Androl 2003; 5: 83–8.
7 Wang H, Zhou Z, Xu M, Li J, Xiao J, Xu ZY, *et al*. A spermatogenesis-related gene expression profile in human spermatozoa and its potential clinical application. J Mol Med 2004; 82: 317–24.
8 Bork P. Hundreds of ankyrin-like repeats in functionally diverse proteins: mobile modules that cross phyla horizontally? Proteins 1993; 17: 363–74.
9 Sedgwick SG, Smerdon SJ. The ankyrin repeat: a diversity of interactions on a common structural framework. Trends Biochem Sci 1999; 24: 311–6.

10 Albert S, Despres B, Guilleminot J, Bechtold N, Pelletier G, Delseny M, *et al*. The EMB 506 gene encodes a novel ankyrin repeat containing protein that is essential for the normal development of Arabidopsis embryos. Plant J 1999; 17: 169–79.
11 Mosavi LK, Minor DL Jr, Peng ZY. Consensus-derived structural determinants of the ankyrin repeat motif. Proc Natl Acad Sci USA 2002; 99: 16029–34.
12 Burkhard P, Stetefeld J, Strelkov SV. Coiled coils: a highly versatile protein folding motif. Trends Cell Biol 2001; 11: 82–8.
13 Blake DJ, Tinsley JM, Davies KE, Knight AE, Winder SJ, Kendrick-Jones J. Coiled-coil regions in the carboxy-terminal domains of dystrophin and related proteins: potentials for protein-protein interactions. Trends Biochem Sci 1995; 20: 133–5.
14 Seipel K, O'Brien SP, Iannotti E, Medley QG, Streuli M. Tara, a novel F-actin binding protein, associates with the Trio guanine nucleotide exchange factor and regulates actin cytoskeletal organization. J Cell Sci 2001; 114: 389–99.
15 Singh A, Hitchcock-DeGregori SE. Local destabilization of the tropomyosin coiled coil gives the molecular flexibility required for actin binding. Biochemistry 2003; 42: 14114–21.
16 Inokuchi J, Komiya M, Baba I, Naito S, Sasazuki T, Shirasawa S. Deregulated expression of KRAP, a novel gene encoding actin-interacting protein, in human colon cancer cells. J Hum Genet 2004; 49: 46–52.
17 Kutty RK, Kutty G, Samuel W, Duncan T, Bridges CC, E1-Sherbeeny A, *et al*. Molecular characterization and developmental expression of NORPEG, a novel gene induced by retinoic acid. J Biol Chem 2001; 276: 2831–40.
18 Vogl AW. Distribution and function of organized concentrations of actin filaments in mammalian spermatogenic cells and Sertoli cells. Int Rev Cytol 1989; 119: 1–56.
19 Cheng CY, Mruk DD. Cell junction dynamics in the testis: Sertoli-germ cell interactions and male contraceptive development. Physiol Rev 2002; 82: 825–74.
20 Eddy EM. Male germ cell gene expression. Recent Prog Horm Res. 2002; 57: 103–28.