

A Mouse Gene Encoding a Novel Member of the WD Family of Proteins Is Highly Conserved and Predominantly Expressed in the Testis (*Wdr13*)

AMRITHA SURESH, VANYA SHAH, DEEPA SELVI RANI, BHUPENDRA N. SINGH, GOGINENI UMAPRASAD, SUBBAYA SUBRAMANIAN, SATISH KUMAR, AND LALJI SINGH*

Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad, India

ABSTRACT *Wdr13*, a novel member of the WD family of proteins and the mouse homolog of *WDR13* is localized to the locus XA1.1 and is predominantly expressed in the testis. The expression begins at the early stages of gonadal development and is maintained throughout the adult life with a predominant expression in the germ cells of adult testis. RNA in situ hybridization on the testis and brain sections indicated a cytoplasmic expression of the transcript. The alternatively spliced transcripts of the gene are generated by different methods and showed a differential pattern of expression, suggesting functional diversity. The expression of the gene in the unfertilized egg and in the neural stem cells indicated the functional significance of the gene from the early stages of development. The nuclear localization of the mouse *WDR13* protein suggested a regulatory function. Evolutionary analysis of the gene indicated an extensive functional conservation across diverse species. Comparison of the genomic organization of the different homologs revealed a varied organization in the invertebrate homolog and the retention of the functionally significant introns in the same. *Mol. Reprod. Dev.* 72: 299–310, 2005. © 2005 Wiley-Liss, Inc.

Key Words: *Wdr13*; WD family of proteins; testis; germ cells; alternative splicing

INTRODUCTION

WD (Trp-Asp) repeats constitute a class of structurally related motifs that are found in a large number of proteins, which do not share any apparent functional properties. WD proteins are known to have 4–16 repeats (van de Voorn and Ploegh, 1992), and several WD proteins of varied and unknown functions have been identified in human, mouse, *D. melanogaster*, *A. thaliana*, *S. cerevisiae*, and *C. elegans* (Duronio et al., 1992; Neer et al., 1994; Smith et al., 1999). A WD-repeat motif is comprised of 44–60 amino acid residues and typically contains a tryptophan-aspartic acid (WD) dipeptide at its carboxy terminus and a glycine-histidine (GH) dipeptide, 11–24 residues downstream from its amino terminus (van de Voorn and Ploegh, 1992). One of the best-characterized members of the family is the β -

subunit of the heterotrimeric G-proteins. Structurally, the proteins with WD repeat motif are believed to act as a platform for the interaction of different proteins and are grouped into two major classes; one in which the proteins are composed almost entirely of the repeat motifs and a second in which the repeats are restricted to the C-terminal domain (Spevak et al., 1993). Members of the first class serve as β -subunits in the heterotrimeric G-protein complexes that transduce receptor-generated signals. Mammalian β -transducins and *S. cerevisiae* Ste4 fall into this class. WD proteins of the second class are involved in diverse functions such as microtubule-dependent processes (*Cdc20*), catabolite repression (*Tup1*), regulation of the RAS-cAMP pathway (*Msi1*), RNA splicing (*Prp4*), DNA replication (*Cdc4*) and neurogenesis (*E-spl*) (Spevak et al., 1993). With such functional heterogeneity the biological role of most of the WD-repeat proteins is difficult to understand. Further, it is not clear in most cases whether functionality, whatsoever is known, should be attributed to the repeat domain itself or to the amino or carboxy terminal extensions of the protein. However, as the WD-repeat family of proteins is believed to mediate interactions between different proteins, it is likely that functionally similar proteins have similar binding partners.

In our earlier studies we isolated and characterized a novel human gene encoding WD-repeat containing protein, *WDR13* (Singh et al., 2003). The *WDR13* protein showed a nuclear localization and promoter analysis of this gene in transgenic mouse revealed a testis-specific upregulation pointing out to a significant role in the testis. The mouse ortholog of the gene has been identified and here, we report the isolation and characterization of the full-length mouse cDNA, hereafter referred to as *Wdr13*.

Amritha Suresh and Vanya Shah contributed equally to the paper.

*Correspondence to: Dr Lalji Singh, Centre for Cellular and Molecular Biology, Uppal Road, Hyderabad 500 007, India.

E-mail: lalji@cemb.res.in

Received 4 April 2005; Accepted 22 June 2005

Published online 10 August 2005 in Wiley InterScience (www.interscience.wiley.com).

DOI 10.1002/mrd.20362

MATERIALS AND METHODS

Screening of the Mouse Genomic and cDNA Library

A male mouse (strain 129) genomic library in Lambda FIX II phage (Stratagene (Loyola, CA, USA), no. 946308) was titrated, plated, and screened according to the manufacturer's protocol. Three independent positive plaques were identified after tertiary screening with the human *WDR13* cDNA probe (Singh et al., 2003). Phage DNA was purified and a 13.3 kb genomic insert was excised and cloned into pBluescript KS+/- for further mapping and sub-cloning. A mouse testis cDNA library in Lambda Zap (Stratagene) was screened by plaque hybridization method described above using the same probe. Eleven plaques were identified as true positive after tertiary screening. cDNA inserts were excised in vivo into pBluescript SK+/- (Stratagene) according to the manufacturer's instructions. DNA fragments from the clones were used either directly or after deletion sub-cloning for DNA sequencing by the Dye Terminator cycle sequencing method using an automated DNA Sequencer (ABI PRISM 377).

Chromosomal Localization of *Wdr13* Gene

The mouse genomic clone containing *Wdr13* gene was labeled with biotin 16-dUTP by random priming (Roche Applied Sciences, Mannheim). Hundred nanograms of the labeled DNA was used as a probe for single hybridization reaction. Mouse metaphase spreads were prepared using bone marrow culture. Hybridization was done at 37°C for 16 hr in a moist chamber as described in Roche FISH laboratory manual. Slides were observed under Zeiss fluorescence microscope and the images were captured and analyzed using the CytoVision program. In silico physical mapping of *Wdr13* gene was performed using mouse genome BLAST and UniGene database search.

Database Search and Sequence Analysis

The sequence generated was screened for vector contamination using VecScreen of NCBI (National Centre for Biotechnology Information, Bethesda) server. The cDNA and genomic DNA sequences were submitted in GenBank under accession nos. AF353243 and AF353244, respectively. Nucleotide and protein BLAST (<http://www.ncbi.nlm.nih.gov/BLAST>) were used for sequence homology searches. Mouse UniGene database at NCBI was searched to identify a UniGene build corresponding to the queried cDNA sequence. ORF analysis was performed using ORF finder program on NCBI. Putative protein sequence was searched for homology against Swissprot database. Presence of various consensus signatures in putative protein was analyzed with PSORT II. PROSITE (Bairoach et al., 1997) database was searched to identify potential structural motifs and patterns. The WD-repeat motifs present in the sequence were identified on the protein structure analysis (PSA) server of BioMolecular Engineering Research Centre (BMERC).

5' Rapid Amplification of cDNA Ends

In order to obtain the full-length cDNA sequence, the positive cDNA clone was analyzed by the 5' RACE kit (Life Technologies, Gaithersburg, MD, USA) as per instructions of the manufacturer. Total RNA (5 µg) from mouse testis was reverse transcribed using gene-specific antisense primer and amplified by PCR with abridged anchor primer and another nested gene specific antisense oligonucleotide primer (GSP2). The amplified product was diluted and reamplified by universal amplification primer and gene specific nested primer containing 5' (CAU)₄ sequences. The latter were designed in a way suitable for UDG cloning system. RT-PCR products were subcloned into vector pAMPI using Clone AMP pAMPI system (Life Technologies). The product obtained was purified and sequenced completely and incorporated into the cDNA sequence.

Fluorescence in Situ Hybridization

The RNA probe specific to *Wdr13* was synthesized by in vitro transcription using DIG RNA labeling kit (Roche Applied Sciences) as specified by the manufacturers. A plasmid clone encompassing exons 2–4 when digested with BamHI and transcribed with T3 RNA polymerase gave the antisense probe and the sense probe was obtained by transcribing the XhoI digested plasmid clone. RNA probe was quantified using the DIG detection kit and around 10 ng of the probe was used for in situ hybridization. Tissues were dissected out from adult mice and fixed in 4% paraformaldehyde followed by treatment with 30% sucrose solution and 15-micron sections were prepared using the cryostat onto gelatin-coated slides. The prehybridization and hybridization procedures were performed according to the nonradioactive in situ hybridization application manual (Roche). The fluorescent signals were detected using Fluorescence Enhancer Kit (Roche) and the slides were observed under confocal microscope.

Expression Analysis of *Wdr13*

Mouse multiple tissues Northern blot was obtained from Ambion Inc., USA (cat no. 3100-1) and hybridized with ³²P labeled mouse *Wdr13* cDNA spanning exon 1–9. Hybridization and post-hybridization washings were performed essentially as per manufacturers instructions. Blot was deprobed and hybridized with actin probe in a similar manner. Mice embryos of different stages ranging from 10.5 to 15.5 dpc were used for RNA isolation. The gonadal ridges were dissected out and analyzed separately. Tissues were dissected out from 1- to 8-week old mice for expression analysis. RNA was isolated using Trizol (GIBCO-Invitrogen, USA) according to manufacturer's instructions. RT-PCR was performed (RNA-PCR core kit of PE, Applied Biosystem, CA, USA) using exon specific primers under conditions that included initial denaturation at 94°C for 3 min followed by 35 cycles, each at 92°C for 30 sec, 65°C for 1 min, and 72°C for 2 min with last extension at 72°C for 5 min. The 5' end of the gene was analyzed by

using different primers to ascertain the presence of alternatively spliced products. The regions spanning exons 1–3 were amplified using the forward primer from exon 1 (E1F-5'/GTGGGAGTTCGCGACATC-3') and reverse primer from exon 3 (E3R-5'/TCCTCATAGAC-TGCACGGTTC-3'). The thermal cycling conditions were as described above with annealing at 62°C for 30 sec. PCR was also done with forward primer from exon 1 (E1F) and the reverse primer from intron 2 (R2-5'/CCTCTACTGACCTGGAACAG 3') at similar conditions with the annealing at 60°C. The products were purified and sequenced completely. The 3'UTR (untranslated region) of the *Wdr13* transcripts was amplified and primers specific to the 3' untranslated region; (UTR/F-5'/CAGATGAGCTTGCACAAA-3' and UTR/R-5'/CCCCCAAATATGCAAACAAT-3') were used. The thermal cycling conditions followed were as before with the annealing at 55°C. *GAPDH* primers were used as control to check the integrity of RNA and efficiency of RT-PCR reaction.

Separation of Testicular Cells

Separation of testicular cells was done according to the protocol previously described (Koga et al., 1998). The testis was dissected out and carefully decapsulated. The tubules were gently unraveled and were incubated in 1X PBS solution with 0.1% collagenase, at 33°C for 30 min. The reaction was stopped by the addition of PBS with 1 mM EDTA, the tubules were transferred to a conical tube and allowed to stand for 10 min. The supernatant was transferred to a fresh tube and the sedimentation process was repeated thrice. The supernatant was then taken and cells pelleted down at 600 g for 10 min and were used as the Leydig cell fraction. The remaining tubules were again incubated in PBS-collagenase solution at 33°C for 15 min. The tubules were gently pipetted in a PBS-EDTA solution and allowed to stand for 15 min. The sedimentation process was repeated twice and the supernatant was used as the germ cell fraction. The sedimented tubules were vigorously pipetted and allowed to sediment for 15 min. The supernatant containing the residual germ cells was discarded and the tubules were pelleted at 600 g for 10 min. The pelleted tubules were used as the Sertoli cell fraction. RNA was isolated from the fractions and used for the expression studies.

Cloning of Mouse *Wdr13* ORF in an Expression Vector

The regions encompassing the exons 1–3 of the gene was amplified by RT-PCR using the primers MTCPB (5'/CCGCGGCTCCAGGGAATGGCC3') that had a BamHI site at its 5' end and ME3R (5'/TCCTCATAGACTGCACGGTTC3'). The thermal cycling conditions followed were the same as mentioned earlier with annealing at 62°C. A BamHI-XhoI fragment was recovered after digesting the amplicon and directionally cloned into the pBSKS vector. This insert was eluted and three-way ligated with a XhoI-SacI fragment from a mouse testis cDNA clone encompassing exon 3–9 and the BamHI-SacI digested pBSKS vector. Desired recom-

binants were identified and the orientation of the cloned fragment as well as the correct open reading frame was ascertained by sequencing. The reconstituted mouse ORF was further cloned into the bacterial expression vector pET 21C(+) (Novagen, Madison, WI, USA). The target plasmid was established in BL21 (DE3) and over-expression was carried out according to the protocol provided by Novagen Technical Bulletin.

Sub-Cellular Localization of *Wdr13*

The ORF from pET 21C(+) vector was sub-cloned into pEGFPC2 (BD Biosciences Clontech, USA) using BamHI-SalI sites. The same insert was also sub-cloned into pEGFPN2 vector using BamHI-HindIII sites. The sub-cellular localization studies were conducted in HeLa (Human cervical cell carcinoma), NIH3T3 (mouse fibroblast) and CHO (Chinese hamster ovary) cell lines. Cell culture and transfection experiments were carried out as described previously (Singh et al., 2003). The expression was observed after 36–48 hr. The coverslips were mounted onto slides in the medium containing 99% glycerol with 1XPBS and 5 µg/ml DAPI stain. The slides were observed using confocal microscope and Zeiss fluorescence Microscope.

Phylogenetic Analysis

The chimpanzee *WDR13* homolog was identified by using primers specific to the human sequence. The homologs from other organisms were identified from the respective EST databases by BLAST search using the *WDR13* cDNA. The complete cDNAs, wherever available, were translated using the ORF analysis (NCBI). The genome sequences were retrieved from the databases using the specific cDNAs and the organization of the gene was identified by alignment with the cDNA sequences. Putative introns were identified by locating the termini of the ORFs, with the presence of the consensus 5'/GT/AG3' sequence and the presence of an intron in a similar position in the human homolog was considered a validation. The nucleotide and amino acid sequences were aligned using CLUSTAL W algorithm (Thompson et al., 1994). Neighbor joining (NJ) trees were constructed using p-distance, nonsynonymous mutation rates in the MEGA package with 1,000 bootstrap replications (Kumar et al., 1994).

RESULTS

Identification of the *Wdr13* Gene, its Sequencing and Structural Organization

Screening of mouse testis cDNA library with human *WDR13* cDNA yielded 11 positive cDNA clones, which were independently characterized, and their identity confirmed by sequence analysis. Two of these clones, namely CD7 and CD1, were identified as major cDNA clones. The genomic clone of 13 kb, positive to these cDNA encompassing the total *Wdr13* gene, was identified and completely sequenced. The vector sequence database was used to remove the contaminating vector sequences. Alignment of CD7, CD1, and genomic

sequences identified a total of nine exons, which were spread over 8.0 kb of genomic DNA. All the exon and intron junctions were marked by the presence of GT and AG signatures, respectively. CD7 encompassed the region that extended from exon 2 to 9 with a portion of the intron 1 sequence and showed the presence of a canonical polyadenylation signature (AATAAA) at the 3' end of the cDNA sequence. CD1 included the region that extended from exon 1 to 3 and was incomplete towards its 3' end. Alignment of both these clones constituted the complete cDNA that corresponded to the human *WDR13* transcript (Fig. 1). Another cDNA clone named CD3 was similar to CD7 but showed the presence of a part of intron 2 sequence towards its 5' end. This suggested the existence of alternatively spliced forms of the *Wdr13* transcript.

Chromosomal Localization of *Wdr13* Gene

Fluorescence in situ hybridization mapping of *Wdr13* revealed that the gene was located on the proximal region of the X chromosome close to centromere (Fig. 2A). In silico mapping confirmed the physical position of the gene on to X chromosome in the mouse. ESTs database search identified several ESTs similar to the *Wdr13* cDNA sequence. They assembled together into a single UniGene cluster Mm181968 that cytogenetically mapped to XA1.1 (Fig. 2B).

Expression Analysis of *Wdr13*

A multi-tissue Northern blot was probed using a region spanning exons 2–9 of the *Wdr13* cDNA revealed a single transcript of about 4 kb in majority of the tissues analyzed except in testis where it showed a transcript of 2 kb (Fig. 3A). The 4 kb transcript observed in the other tissues corresponded to the complete size of the transcript including the 5' and 3' UTRs. Reprobing with the actin probe confirmed the quality of RNA being used in this experiment (Fig. 3B).

In order to identify the localization of the transcript in the tissues, RNA in situ hybridization was done on the testis and the brain sections. The cryosections of the mouse testis showed a cytoplasmic signal in the cells within the seminiferous tubule with the expression being higher in the germ cells (Fig. 4A). The nucleus, counterstained with DAPI was completely devoid of the signal in all the cells. A similar cytoplasmic signal was observed in the brain sections (Fig. 4C). Controls using the sense probe (Fig. 4B,D) showed negative results. Analysis in the female brain and ovary sections showed similar cytoplasmic expression (Fig. 4E,F).

Spatial expression analysis of *Wdr13* indicated a higher level of expression in testis and brain (Fig. 5A). The lower bands indicate the corresponding *GAPDH* levels (Fig. 5A). In accordance with the predominant expression of the gene in the testis, the pattern of expression in the different testicular cells was analyzed. The germ cell fraction showed a predominantly higher expression compared to the fractions of Sertoli cells and Leydig cells (Fig. 5B). The lower panel shows the corresponding *GAPDH* levels as controls (Fig. 5B). Analysis

```

gctatggggggggggcggtgtgctacggcaactggagaggggggggtcatgtctgaactgctgctgt
gagtaccgggagtgctgccaggaaagcagggtggtgatgaccacgctaacgcccagtgggga
tttcgacatcacctgtgtggaggacttctgtctatggcaactgtcacctgtgtggcggaacta
gattttctctgctgggacgctgacattccaagccctgcccctgagagctcccgcaggcagacaac
aaagaaggaagccagggaATGgCCGCGGTGTGGCAGCAAGTATTAGCAGTGGACGCGAGGTACAAC
      M A A V W Q Q V L A V D A R Y N
GCCTACCGTACACCAACGTTTCCACAGTTTCGGACCCAGTATATCCCGCGGCAGCCAGCTGCTT
A Y R T P T F P Q F R T Q Y I R R R S Q L L
CGGAGAAATGCCAAAGCTGGTCAACCCCGCAGCATTGCTCGGCAGTACCTGAGGCTACGGGGCCAG
R E N A K A G H P P A L R R Q Y L R L R G Q
CTATTGGGCCAAACGGTACGGGCCACTCTCAGAACCAGGCAGTGTCTGTCCTATAGCAACAGCATT
L L G Q R Y G P L S E P G S A R A Y S N S I
GTCCGAGCAGCCGAACAACCCCTTGATCGAATGGAGGACTTTGAAGATGACCCAGACCCCTGGG
V R S S R T T L D R M E D F E D D P R A L G
GCTCGAGGACACCCCGATCTGTACGCCGAGGTTCTACCAGCTGCAGGCACAAATGAACCGTGCA
A R G H R R S V S R G S Y Q L Q A Q M N R A
GTCTATGAGACAGGCCTCTGGCAGTGTGGTACCACCTGCGGTGGCAGAGCAAGTCCGGCCATG
V Y E D R P P G S V V P T S V A E A S R A M
CGCCGGACACGCTCGCTGAGTGAGAACTATGCCCTTTGCAGGCATGACCATGTTTTTGACCAACAC
A G D T S L S E N Y A F A G M Y H V F D Q H
GTGGATGAGGCGTCCCAAGGGTGGCCTTCGCCAATACGACCGGCACCCCTGGCCTGCTGCTCC
V D E A V P R V R P A N D D R H R L A C C S
CTGGACCGCAGCATCTCCCTGTGCCAGTGGTGGCTGCCCCACCCACTGTGCTCCATGTGCTACGG
L D G S I S L C Q L V P A P P T V L H V L R
GGCCATACACCTGGCTCTCGGACTTCCGCTGGTCCCTCTCCANATGACATCTGTGTCCACCTCC
G H T R G V S D F A F A G M Y H V F D Q H
CTCGATGCCACCATGCGCATCTGGCCCTCCGAGGACGGCCGCTGCATCCGTGAGATCCCTGACCC
L D A T M R I W A S E D G R C I R E I P D P
GATGGCCCGAATGCTCTGCTGCACCTTCCAGCCAGTCAACAACAACCTCACTGTGTGGGGAAAC
D G A E L L C C T F Q P V N N N L T V V G N
GCCAAGCACAACTGCATGTCAATCAATCTCCACAGGCAAGAAGTGAAGGGTGGCTCCAGCAAG
A K H N V H V M N I S T G K K V K G G S S K
CTCACCGGCGGCTCCTCGCCCTGTCTTTGATGCCCTGGTGGCTGCTCTGGCCAGCGATGAC
L T G R V L A L S F D A P G R L L W A G D D
CGCGGAGTGTCTTCTCCTTTCTTTGACATGGCCACAGGAAAGCTGACCAAGGCAAGCGGACTA
R G S V F S F L F D M A T G K L T K A K R L
GTAGTGCATGAAGGCGCCCTGTAACCAAGCAITTTCTGCCGCTCTGGTTCAGCCGTGAAGCACGG
V V H E G S P V T S I S A V N K L Q A Y S R E A F
GACCCCTCTCTGCTCATCAATGCTCCCTCAACAAGCTGCTACTCTACAGAGTGGTGGACAACGAA
D P S L L I N A C L N K L L L Y R V V D N E
GGGCGCTACAGCTGAAGAGAAGCTTCCCAATGAACAGAGTTCACCCCTGTACGAGTATCTTC
G A L Q L K R S F P I E Q S S H P V R S I F
TGCCCCCTCAITGCTCCTCCGCCAGGGGCTGTGTGGTGGACAGGAGTGAAGATATGTGGTTCAC
C P L M S F R Q G A C V V T G S E D M C V H
TTCTTTGACGTGGAGCGGGCAGCCAGGCTGTCTTAAACAGKCTGACGGGCCACGCGCCCGCTG
F F T V E R A A K A A V N K L Q G H S A P V
CTTGACGTGAGCTTCAACTGTGACGAGAGTCTGCTGGCCCTCCAGTATGACGAGTGGCAGTGCATC
L D V S F N C D E S L L A S S D A S G M V I
GTCTGGAGACGAGCAAAAGTAgggttctgtcacctgtgtgtgccactgatccctcctcctca
V W R R E Q K *
ctccagcttggcgtataaattaaattcctgtggttgtgtga

```

Fig. 1. cDNA sequence of *Wdr13*. The sequence indicated in capitals denotes the ORF of the gene with the corresponding amino acid sequence. The amino acid sequence in bold indicates the six WD motifs. The sequences in small letters denote the 5' and 3' UTRs. The three translational start sites are indicated in bold. The stop codon is marked by a star and the polyadenylation signal is also shown in bold.

of the female tissues indicated a higher level of expression in the brain and kidney (Fig. 5C). The corresponding *GAPDH* levels are indicated in Figure 5D. The temporal expression during various stages of testicular development from gonadal ridges and embryos of mice ranging from 10.5 to 15.5 dpc and from tissues of 1- to 8-week-old mice was studied. RT-PCR analysis was done

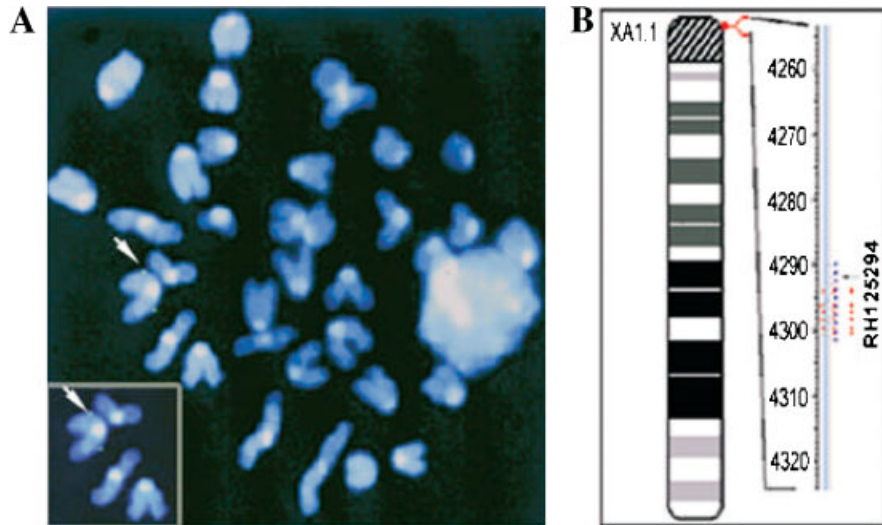


Fig. 2. Mapping of *Wdr13*. **A:** Fluorescence in situ hybridization shows the localization of *Wdr13* onto the mouse X chromosome. The magnified view of the X chromosome with the signal is shown in the inset. **B:** shows the in silico mapping of *Wdr13* to the XA1.1 region.

using primers encompassing exons 1–3. The *Wdr13* transcript was expressed from the initial stages of gonadal ridge development (10.5 dpc) but was found to be significantly low in the gonadal ridge at 15.5 dpc. Expression was also found in embryos of all stages from 10.5 to 15.5 dpc and in testis of mice from 1 to 8 weeks (Fig. 5E). The corresponding *GAPDH* expressions as controls are shown in the lower panels (Fig. 5E).

RT-PCR Analysis to Study the Alternatively Spliced Variants of *Wdr13*

Northern analysis of the *Wdr13* transcripts showed varied sized transcripts in different tissues suggesting the possibility of alternative splicing from this locus. This was also evident from the database search as the UniGene build Mm 181968 enlisted several ESTs, derived from different mouse tissues, which were of varied sizes. Importantly, they build up to a 4.2 kb transcript (TC689367, www.tigr.org) corroborating to the most prevalent transcript detected by Northern analysis.

RT-PCR analysis was used to identify the alternatively spliced transcripts of *Wdr13*. Analysis using the primers encompassing exons 1–3 indicated the presence of the transcript in the original form (AF353243) with a predominant expression in the testis (Fig. 5A). RT-PCR using primers from exon 1 and the intronic region observed in the CD3 clone showed the retention of the entire intron 1 and a part of intron 2 sequence (Fig. 6). Sequence analysis showed the usage of a different splice site resulting in the retention of the intronic sequences. Several ESTs from different tissues like neural retina (BG804614), 15-day-old embryo (BY110885, BY107845), testis (BU963044), limbs (BQ964168), also showed the retention of the intron 1 sequence. The spatial expression pattern of the CD3 transcript differed in tissues with the testis showing a relatively lesser expression (Fig. 7A). Temporal expression analysis indicated a higher expression of this transcript in the 14.5 and 15.5 dpc embryos (Fig. 7B). Expression levels were very low in the gonadal ridges of the early stage embryos. The corresponding *GAPDH* expressions as controls are shown in the lower panels of Figure 7A,B. This evidence thus points out to the

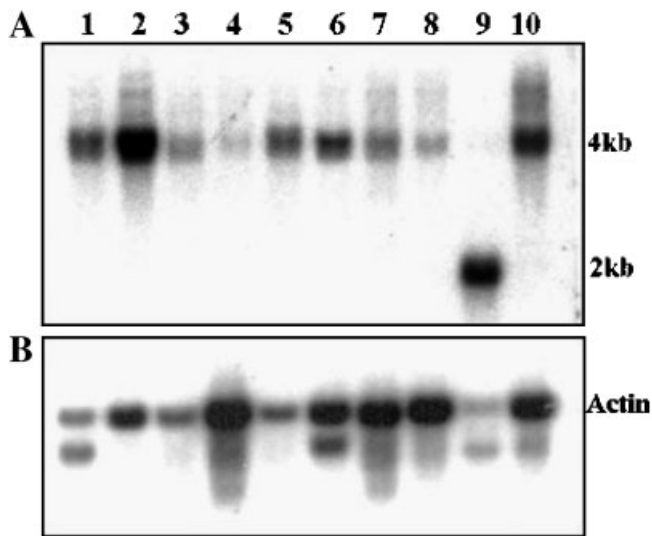


Fig. 3. Expression Analysis by Northern hybridization. **Panel A** shows the expression of *Wdr13* in different tissues; lanes (1) heart, (2) brain, (3) liver, (4) spleen, (5) kidney, (6) embryo, (7) lung, (8) thymus, (9) testis, (10) ovary. **Panel B** shows the corresponding hybridization with actin probe. Note predominant expression of the gene *Wdr13* in the brain and the testis and the difference in the size of the transcript.

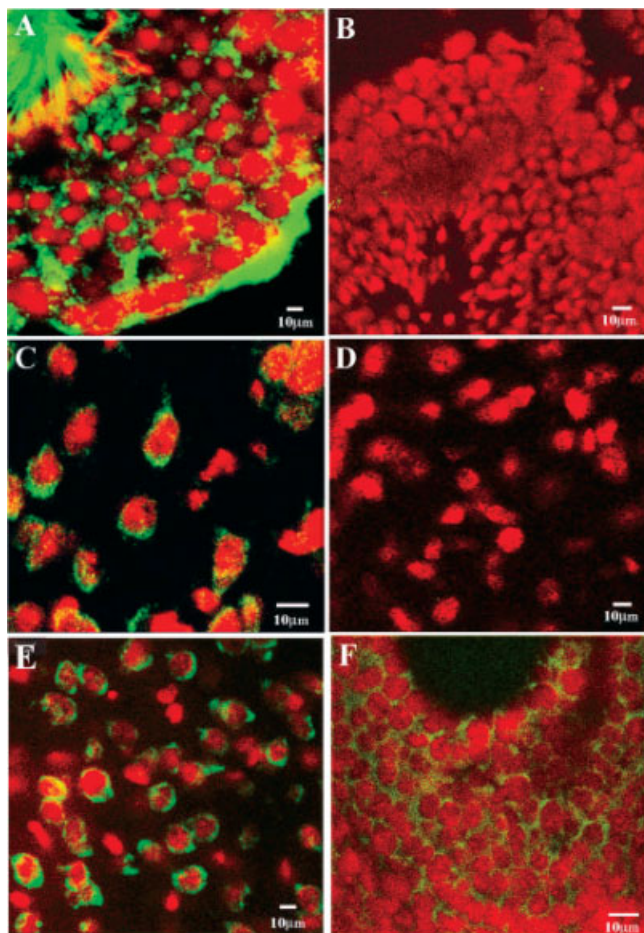


Fig. 4. Fluorescence in situ hybridization using antisense RNA probes showing the cytoplasmic localization of the transcript on sections of testis (A) male brain (C), female brain (E), ovary (F) and sense probe showing complete absence of hybridization on sections of testis (B), and male brain (D).

presence of alternatively spliced transcripts with a different expression pattern.

Analysis by RT-PCR also revealed the presence of transcripts with different lengths of 3'UTR. A 1.3 kb 3'UTR region was detected in tissues like brain, heart, ovary, and liver; however, it was not found in kidney and testis (Fig. 7C). The corresponding *GAPDH* expression is shown in Figure 7D. Database search indicated the presence of this extended 3'UTR region in a cDNA clone from mouse 8-day-old embryo (AK017549). However, a smaller 3'UTR region of around 100 bp was found in the CD3 cDNA clone while the CD7 clone had a 3'UTR of 585 bp. Sequence analysis revealed the presence of canonical polyadenylation signature in each of these transcripts.

Analysis of the mouse cDNA library database (NIA Mouse cDNA Project) gave evidence of the expression of the gene from the stem cell stage onwards. cDNA clones were obtained from neural stem cells, unfertilized eggs, and mouse 8-day-old brain, mouse 8- and 12-day-old

embryo and in the neonate kidney (Table 1). These clones were incomplete at their 5' and 3' ends and with the exception of the clones from the newborn kidney, all clones showed 99–100% identity with the *Wdr13* transcript. Pairwise alignment of the cDNA clones from the kidney (B0832G07-5, B0828F10-5) with the *Wdr13* sequence indicated the retention of the 5th intron sequence partially in both the clones. The expression of the gene in the stem cells and the unfertilized egg along with its maintenance throughout the developmental stages of the mouse and the human and its specific expression in the germ cells, testify the functional significance of the gene from the early stages of development.

Analysis of the Predicted Protein Sequence and Identification of WD-Repeats

5'RACE resulted in the identification of the 5' end of the cDNA and the ORF analysis revealed a longest ORF encoding for 485 amino acids with an estimated molecular mass of 53 kDa. Comparison of human and mouse cDNA showed 87% homology; however, the mouse protein varied from the human counterpart at only four amino acids. The start codon (ATG), flanked by the Kozak consensus sequence, was present in the first exon of cDNA sequence leaving 286 bases of untranslated region (UTR) towards its 5' end (Fig. 1). Termination codon TAG was present in exon 9. Primary sequence analysis of the putative protein predicted this to be a basic protein with a pI value of 9.4. Putative protein analysis of the CD7 and CD3 transcript indicated that they code for a protein of 393 and 363 amino acids, respectively, with 1.2–1.4 kb of the 5' end of the cDNA acting as 5'UTR. These proteins are devoid of the N-terminal amino acids with the C-terminal being conserved. PROSITE database search identified regions bearing significantly high level of homology with WD-repeat motifs. Six such potential WD-repeats with significantly high profile probabilities were found to be present in both the encoded proteins indicating the functional significance of these motifs (Fig. 8). BLASTP searches with the protein sequence did not identify any significant match except WD-repeat encoding region.

Over-Expression and Sub-Cellular Localization of the *Wdr13* Protein

The complete ORF of *Wdr13* encoding a protein of 485 amino acids (53 kDa) was sub-cloned in pET21C plasmid vector and over expressed in BL21 (DE3) bacterial strain. The over expressed protein was found in the insoluble fraction. The protein was eluted from the SDS-PAGE gel and used for raising antibodies. We attempted to raise antibodies in different animals like rabbits, chicken, and guinea pigs over a period of 8–10 months; however, we did not succeed in raising antibodies with good titer.

The sub-cellular localization of the 53 kDa protein was detected in different cell lines such as HeLa, CHO, and NIH3T3. The complete ORF (485 amino acids) cloned in pEGFP2 and pEGFPN2 was transfected into the cell lines mentioned above. We noticed that while

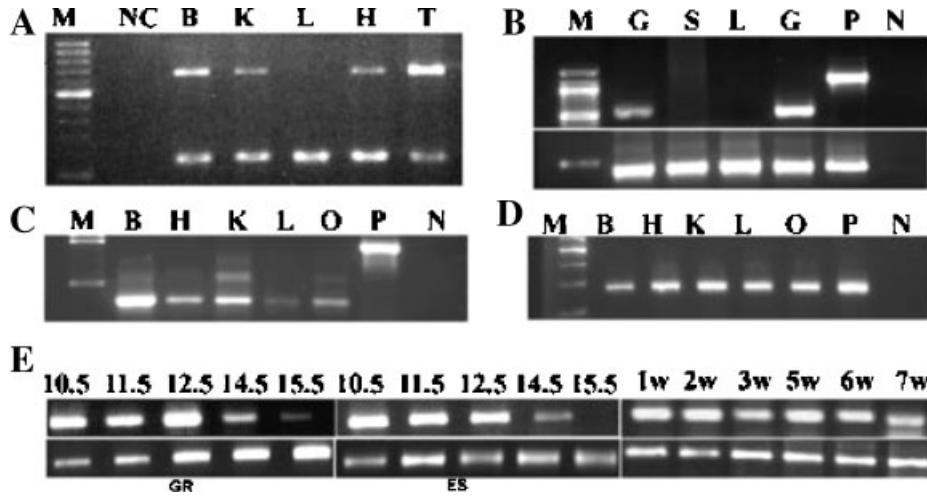


Fig. 5. Spatial and temporal expression of *Wdr13*. **A:** Expression in different tissues by RT-PCR. The **upper band** shows the expression of *Wdr13* in different tissues while the lower band denotes GAPDH expression. The lanes are represented as marker (M), PCR without template as negative control (NC), brain (B), kidney (K), liver (L), heart (H), and testis (T). **B:** Expression of *Wdr13* is predominantly high in the germ cells compared to the other testicular cells. The lanes represented are marker (M), germ cells (G), Sertoli cells (S), Leydig cells (L), amplification of the *Wdr13* cDNA clone as positive (P) and PCR without

template as negative controls (N). The **lower panel** denotes the GAPDH amplification. **C:** Expression in the different tissues of adult female mice. The lanes are represented as marker (M), brain (B), heart (H), kidney (K), liver (L), and ovary (O) with *Wdr13* cDNA clone as positive (P) and PCR in absence of template as negative (N) controls. The **panel D** denotes the GAPDH levels. **E:** Temporal expression of *Wdr13* in the gonadal ridges, the early embryonic stages (from stages 10.5 to 15.5 dpc) and in 1- to 7-week-old mice. The lower panel denotes the corresponding GAPDH amplification.

the localization of the GFP protein was observed in the nucleus as well as the cytoplasm (Fig. 9A–C), the *Wdr13-GFP* fusion product localized specifically in the nucleus in all the three cell lines studied (Fig. 9D–F). The nuclear localization was confirmed by colocalization using DAPI (Fig. 9G–I). Sequence analysis to identify potential nuclear localization signals (NLS) did not reveal any known signature sequences essential for nuclear localization. This signifies the presence of novel nuclear localization signals in the *Wdr13* protein.

Evolutionary Conservation of *Wdr13*

In an effort to understand the evolutionary conservation of the gene, the *WDR13* homolog from chimpanzee was isolated and sequenced using human specific primers (AY888002). Comparative sequence analysis of the coding region with that of the human sequence indicat-

ed a 99% identity and sequence analysis revealed an ORF that encoded a protein identical to the human protein. In silico mapping of the gene localized the gene to the X chromosome. The genome organization was similar to the human gene and the intronic sequences showed a 98% identity. The cDNA homologous to *Wdr13* (XM_228783) was identified from the database in the rat brain. It was 94% identical to the mouse transcript. The rat complete gene identified from the genome (NW_048035.1), spanned 13 kb of the sequence with an organization similar to the human and mouse genes. The gene was localized to Xq13 and the size of the transcript reported (4 kb) corroborated with the 4 kb transcript identified in the mouse. The ORF of 1457 bp encoded a protein (485 amino acids) that showed a 99% identity with the mouse protein. All the six WD motifs were conserved in the rat protein, pointing out to a possible functional conservation of the gene. It is noteworthy that the gene is localized on the X chromosome in all these organisms as is seen in the human (Xp11.23). Comparative analysis of the X chromosome map also revealed that the position of the gene is conserved with respect to the flanking genes like *OATL1*, *HDAC6*, and *SUV39H1* in the human, mouse and rat systems.

Analysis of the available protozoan genomes did not result in the identification of any homolog. Among the other genomes, homologs were identified from *Danio rerio* (CK361970), *Fugu rubripipes*, *Xenopus tropicalis* (BX733253), *X. laevis* (CB591223), *Bos taurus* (CB452613), *Sus scrofa* (BP158083), *Ciona intestinalis* (BW314224), and *Apis mellifera* (XM_396208). Neighbor joining trees were hence constructed using the p-distance between the cDNA and the translated protein

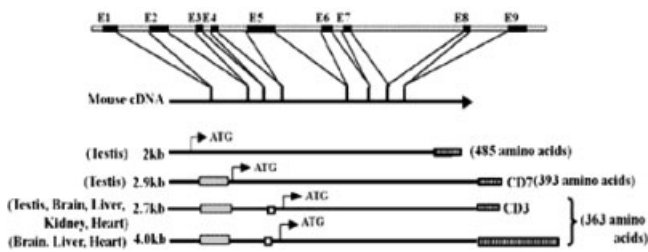


Fig. 6. Alternative Splicing of *Wdr13*. The figure shows different alternative spliced forms of the gene. The dotted boxes indicate the introns that are retained in the transcripts and the three initiation codons are also shown. The patterned boxes point out the varied lengths of the 3'UTR. The tissues in which the transcripts are predominantly expressed and the different proteins synthesized are also mentioned.

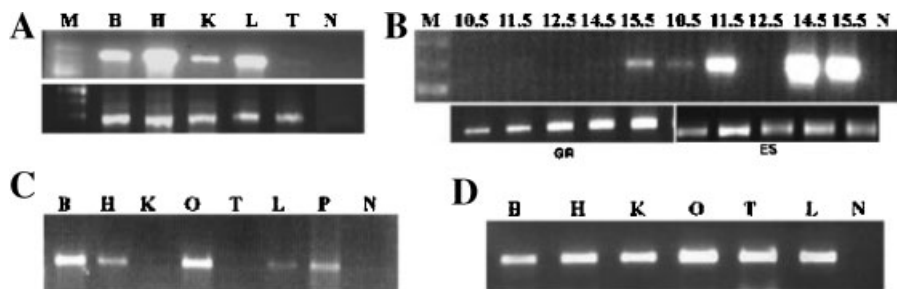


Fig. 7. Spatial and temporal expression of alternatively spliced transcripts. **A:** Indicates spatial expression of the CD3 transcript; the lanes are marker (M), brain (B), heart (H), kidney (K), liver (L), testis (T) and amplification in absence of template as negative control (N). The lower panel shows GAPDH expression. The **upper panel** shows the expression of this transcript in the gonadal ridge and the early

embryonic stages and the **lower panel** shows the corresponding GAPDH levels. **C:** Shows the expression of the 3'UTR and **(D)** the corresponding GAPDH expression. The lanes are brain (B), heart (H), kidney (K), ovary (O), testis (T), and liver (L) with amplification using the human genomic DNA as positive control (P) and in the absence of template as negative control (N).

sequences with 1,000 bootstrap applications. A comparative analysis of the NJ trees obtained from the cDNA and the translated protein sequences (Fig. 10A,B) clearly indicated a higher and extensive conservation at the protein level. The clustering of the mammalian homologs clearly indicated a functional conservation of the gene. This is also particularly evident in the *WDR13* homologs identified from the arthropod and the ascidean. In these cases the cDNA sequences that show practically no identity, translate into a 45 and 51% similarity at the protein level, respectively, signifying a conserved function that is mediated by the WD motif. All the translated proteins were subjected to analysis by the NCBI conserved domain search for the presence of WD motifs. Potential WD40 motifs were identified in all the proteins and almost all of them being confined to the C-terminal region of the protein as is observed in the case of the human protein and hence classifying them in the second group of WD family of proteins. Sequence analysis of the WD motifs indicated a greater degree of conservation as compared to the rest of the protein. The WD motif identified in the arthropod *A. mellifera*, on the other hand showed extensive variation.

The genomic organization of the previously uncharacterized genes from rat, puffer fish, and honeybee were identified by alignment of the ESTs with the corresponding genomic contigs maintaining maximum similarity with the human and mouse homologs. The invertebrate *A. mellifera* (order: hymenoptera, class: Insecta) showed

a genomic organization of four exons and three introns spanning 2.2 kb of the genome (Fig. 11). This gene has an ORF of 1.3 kb and encodes a protein of 435 amino acids. The honeybee *WDR13* is the only evidence of the existence of the gene in the invertebrate groups. Analysis of the *Bombyx mori* (order: Lepidoptera) genome also did not reveal any significant homology to the *WDR13* gene. Though the Hymenopterans and Lepidopterans are members of the same group, Endopterygota, the former is believed to have diverged from the rest of the group during the process of evolution. This suggests a recent origin of the gene in the invertebrates, approximately 270 MYA, when the divergence of these two insect orders is believed to have taken place.

The genomic organization of the *WDR13* is conserved across all the vertebrate phyla with the homologs from human, chimpanzee, mouse, rat, and puffer fish showing the presence of nine exons and eight introns. The *T. rubripipes* gene (puffer fish) *WDR13* spans 4.2 kb with nine exons and eight introns with an ORF of 1.4 kb encoding a protein of 468 amino acids. The genomic organization being similar to that of the mammalian genes, the sizes of the intron vary considerably leading to the smaller size of the *T. rubripipes* gene. A similar trend is observed in the other homologs wherein the organization remains the same but the introns vary considerably in their size and sequence, which was evident by the variation in the size of the complete gene; Apis ~ 2 kb, Fugu ~ 4 kb, mouse 8.0 kb, rat 13 kb,

TABLE 1. EST Clones Identified From the Different Developmental Stages of Mouse and Human

Stage (mouse)	EST (accession number)	Stage (human)	EST (accession number)
Neural stem cell	CA880847	Embryonic stem cell	CN301722.1
Neural stem cell	CA883823	Embryonic trophoblast	CD390179.1
Neural stem cell (Undifferentiated)	K0974B02-5N	Fetal Brain	CR594799.1
Unfertilized egg	B0932D01	Fetal liver	CR602306.1
Unfertilized egg	B0932D01-5		
8-day-old brain	X14943		
8-day-old embryo	AK017549		
12-day-embryo	BY721651		
New-born kidney	B0832G07-5, B0828F10-5		

	S	L	C	I	L	SSN	VLGS	L	IYR
WD Consensus	AD	I	AL	W	DND	IVTAG	SVRLFN		
	GHXXKV	XSVXFX	[0-?]X	PDG	[0-?]X	KLASGS	XDX	TIKVVD	
WD1 215-245	GHTRGV	SDFAWS	L	SND	ILVSTS	LDA	TMRIWA		
WD2 258-290	PDGAEI	LCCTFQ	PV	NNN	L	TVVGNA	KHN	VHVMNI	
WD3 302-332	KLTCRV	LALSFD	A	PGR	LLWAGD	DRG	SVFSFL		
WD4 349-385	HEGSPV	TSISAR	SWV	SRE	ARDP	SLLINA	CLN	KLLLYR	
WD5 403-437	QSSHPV	RSIFCP	LMSF	RQG	A	CVVTGS	EDM	CVHFFD	
WD6 451-481	GHSAPV	LDVSNF	CDE	S	LLASSD	ASG	MVIVVR		

Fig. 8. WD motifs in the *Wdr13* protein sequence and their correspondence with the WD-repeat consensus. The amino acid sequence in bold shows the WD repeat consensus with the alternative amino acids for each position listed above in order of their frequency of occurrence. The six WD motifs in *Wdr13* and their corresponding positions are shown in comparison with the consensus sequence. The secondary structures, loop (↪), β-strand (→) and turn (⊥) taken up by the different regions of the motif are also shown.

chimpanzee ~ 7 kb, and human 7.3 kb (Fig. 11). Multiple alignments of the different introns showed higher degrees of identity between the homologous genes with a pattern that followed the general evolutionary trend. The identity levels between the different homologs were: Human-chimp (87–99%), mouse-rat (30–92%), human-mouse (26–77%) and human-fugu (1–35%). Pairwise alignment of the intron 1 sequences between the human and mouse genes also indicated specific regions (ranging from 50 to 100 bp) wherein the sequences showed 90–95% conservation. No identity was observed between the introns of the honeybee gene and introns from the chordate homologs. Within the introns of the chordate genes, the highest conservation across the species was

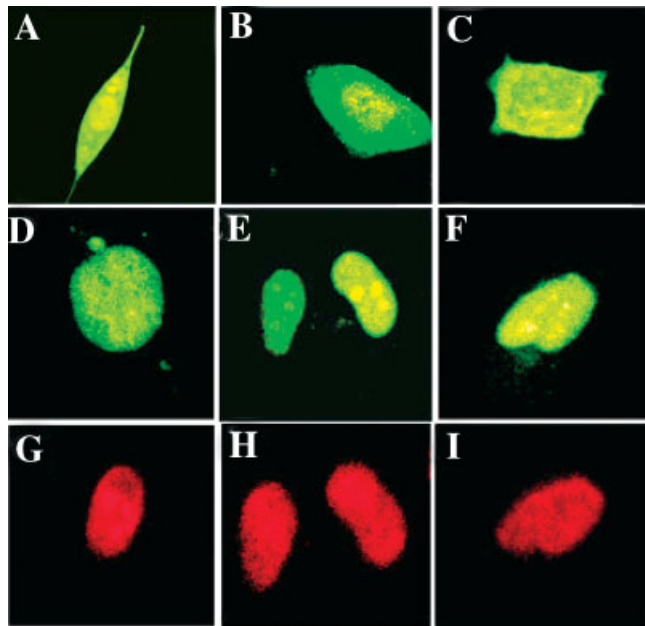


Fig. 9. Sub-cellular localization of the GFP fusion protein. The localization was observed in three cell lines NIH3T3 (A, D, G), CHO (B, E, H), and HeLa (C, F, I). The figures denote the localization of GFP protein in the nucleus as well as in the cytoplasm (A, B, C), the GFP-*Wdr13* fusion protein only in the nucleus (D, E, F) and the DAPI control showing the nucleus (G, H, I).

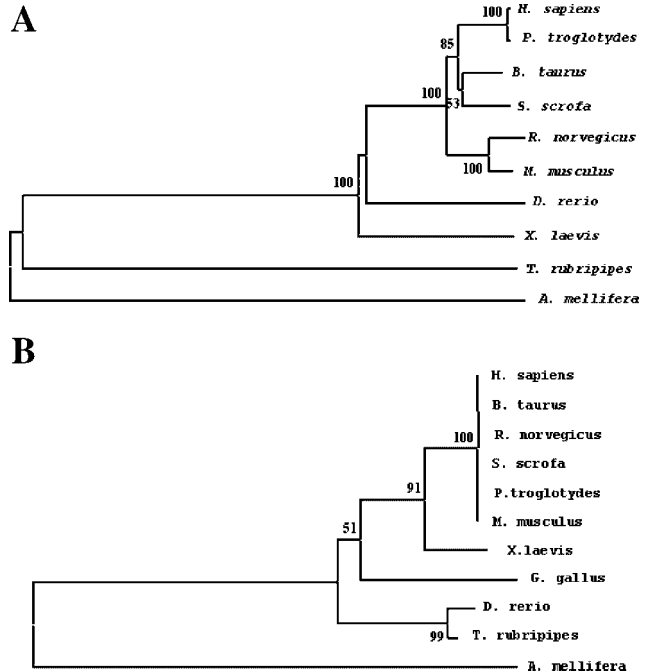


Fig. 10. Evolutionary analysis of *WDR13*. (A) shows the NJ tree constructed using the cDNA sequences and (B) shows the NJ tree constructed using the protein sequences. A comparison between the two trees points out to the extensive functional conservation of the gene. The clustering pattern distinctly illustrates the functional similarity of the protein within the mammalian class. The numbers at the various nodes are the bootstrap support values obtained by neighbor joining method.

observed in the intron 6 (35–100%), intron 8 (14–99%), and intron 4 (22–97%). Out of these, the introns 4 and 6 correspond to the introns in the invertebrate gene indicating the conservation with respect to their sequence and position (Fig. 11).

DISCUSSION

We have described the isolation and characterization of a novel mouse gene designated as *Wdr13* encoding a novel member of WD-repeat family of proteins. The characterization of this gene was carried out since it was predominantly expressed in the mouse testis and brain. The human ortholog showed testis-specific promoter activity that was further indicative of its functional significance. The mouse *Wdr13* showed 87% homology at cDNA level and differed at protein level at four amino acid positions. The conservation of the six WD motifs in both the genes also attributed a functional importance to these motifs.

Evidences from Northern and RT-PCR analysis indicated the presence of alternatively spliced products that have a differential expression pattern. Alternative splicing accounts for increase in the coding capacity of around 35% of genes by allowing for the generation of multiple protein isoforms with distinct function (Breitbart et al., 1987). The alternative splice forms of *Wdr13* are generated by the frequently documented mechanisms such as shift in the 5' or 3' splice sites,

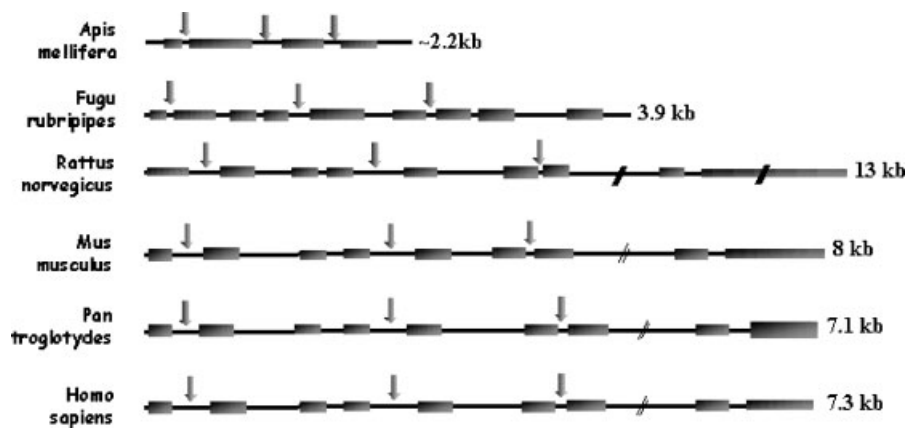


Fig. 11. Evolutionary organization of *WDR13*. The genomic organization of the different *WDR13* genes and their corresponding sizes are indicated. The boxes represent exons while the black lines represent the intronic regions. The arrows indicate the introns in the invertebrate gene (1, 2, and 3) that correspond to the positions 1, 4, and 6 in the chordate homologs. Note that the invertebrate homolog shows a different genomic organization while the chordate genes vary only in the size of the complete genes.

retention of introns and use of alternative polyadenylation signals (Black, 2000). The CD3 transcript is created by the utilization of different splice sites resulting in the retention of the intronic sequences. Intronic sequences are known to influence the expression of the splice variants by regulating mRNA related processes like transport, translation, and decay (Hir Le et al., 2003). The inclusion of the intronic sequences may hence account for the differential expression pattern of the CD3 transcript. ESTs from different tissues as well as RT-PCR analysis indicated the presence of the extended 3'UTR in tissues like brain, liver, and ovary and its absence from the testis and kidney. It is noteworthy to mention that the 3'UTR sequence of the transcript harbors the AU rich elements (ARE). The AREs are implicated in the stability and decay of various transcripts. Absence of ARE elements in the testis transcript implies that these transcripts may have a longer half-life period compared to transcripts present in other tissues as seen in the case of the testis-specific isozyme of lactate dehydrogenase in rodents (Salehi-Ashtiani and Goldberg, 1995). The differential spatial and temporal expression pattern of the varied splice variants, specifically in gonadal development points out to distinct functions in the same.

Another significant observation is the identification of the human *WDR13* from a variety of cancers as indicated by the compiled expression information available in the database (Hs. 12142). The gene has been identified on the basis of differential expression in cancers like colon adenocarcinoma (BC002507), neuroblastoma (CR601751.1, CR603392.1), Ewing's Sarcoma (AA730971), Larynx sarcoma (AJ403132), pancreas ductal carcinoma (BC067094), signet ring-cell carcinoma (AK000570), ovary teratocarcinoma (BI905079), and lung tumors (BI905079). The characters exhibited by *WDR13* gene such as a predominant upregulation in the testis, being mapped to the X chromosome and a heterogeneous expression pattern in the cancerous

tissues indicated a profound similarity to the members of the Cancer-Testis group of antigens (Scanlan et al., 2002). A number of the CT antigens have been identified and the function of a major number of these genes is unknown while some of them like *SCP1* and *OY-TES* are implicated in sperm maturation and function (Tureci et al., 1998; Ono et al., 2001). An important observation in this regard is that 10% of the genes on the human X chromosome are found to be CT antigens (Ross et al., 2005). The significance of these observations with respect to the involvement of the gene in the tumor progression process and other aspects needs to be thoroughly analyzed.

The predominant expression in the testis, specifically in the germ cells suggests the necessity of the gene during their differentiation or maintenance. Though various genes have been implicated in different stages of germ cell differentiation, the complete process of their maintenance is still elusive. The expression pattern of *Wdr13* is similar to that exhibited by the candidate genes downstream to *Sry*, as in the case of *Sox9* (Kent et al., 1996) and hence points out to a wider range of function for the gene. *Sry*, expressed in the short window of 10.5–12.5 dpc (Gubbay et al., 1990; Hacker et al., 1995) initiates the entire male sex determination process by triggering the Sertoli cell differentiation, but the actual cascade and genes involved in it are still to be identified. The higher level of gene expression observed in the brain and its identification from neural stem cells gains significance in the light of the identification of the cDNA from the brain encoding the memory related protein (Kawai et al., 2001), the sequence of which is similar to that of *Wdr13*. The female brain also shows a predominant expression as compared to the ovary indicating that the gene performs a distinct function in the brain. Sex-specific differential expression in the brain of genes associated with the sex determining process was observed in the *Sry* transcript (Mayer et al., 2000), *SRY* and *ZFY* (Mayer et al., 1998), and *Usp9y* and *Ube1x*

(Xu et al., 2002). The sex differences in nongonadal phenotypes are generally attributed to the action of gonadal hormones but the dimorphic expression of these genes tell a different story. The human homolog also showed a comparatively higher expression in the brain apart from kidney and heart. A recent report implicated a translocation t(X;2)(p11.2;q37) in an individual with a developmental delay accompanied by seizures, infantile hypotonia, obesity, and livedo reticularis (a purplish network-patterned discoloration of the skin) (Leach et al., 2002). The gene disrupted by the first breakpoint was diacylglycerol kinase delta (*DGKD*) that is considered responsible for the seizure phenotype (Rodriguez de Turco et al., 2001). The second breakpoint was identified within the *WDR13* gene, implicating the gene in the normal functioning of the brain; the details of the mechanisms involved are still to be understood.

Phylogenetic analysis of *WDR13* gave evidences of the conservation of the gene between the different chordate and invertebrate homologs with the conservation at the protein level being significantly high. This clearly points out to similar functional constraints being exerted on them; otherwise a varied functional selective pressure would have resulted in accumulation of nonsynonymous mutations as compared to the synonymous ones. This has been observed during the study of other genes like *DAX1* that had a higher amino acid substitution rate as compared to *SRY* and *SOX9* in the primate species, which was attributed to the varied functional constraints of *DAX1* (role in adrenal development and its recent entry in the sex determination pathway) (Patel et al., 2001). Another significant observation that strengthens the idea of functional conservation is the identification of WD motifs in all the homologs from *A. mellifera* to *P. troglodytes*.

The fewer number of introns in the ancient *A. mellifera* homolog and the subsequent increase in the number of introns in the higher organisms support the theory of late addition of the introns into the genes of higher organisms (Palmer and Logsdon, 1991). The presence of the specific introns of *WDR13* (1,4,6) in the primitive invertebrate gene and their subsequent conservation in the higher organisms probably qualifies them as being ancient introns. Similar examples have been observed in genes like aspartate aminotransferase isoenzyme wherein introns are in identical positions between the cytosolic and mitochondrial forms providing evidence of the conservation of these introns in the ancient gene (Obaru et al., 1988). It may be argued that the insertion of the introns at these positions might be a chance occurrence since other invertebrate homologs with a similar exon–intron organization have not been identified. Search of the available databases in this regard did not yield any result. A significant observation in this context is that the introns that have their position and sequence conserved provide a substrate for increased diversity in mechanisms associated with mRNA processing. The conservation of the *WDR13* intron positions coupled with the fact that the first intron is involved in alternative splicing strengthens this view-

point. Similar studies have resulted in the identification of transcriptional enhancers in the intronic sequences of immunoglobulins (Church et al., 1985). Hence, conservation of *WDR13* intronic sequences across phyla could be because of the functional role they play in the regulation of the gene expression.

WD proteins have been assigned important roles in various processes in murines and other organisms, some of them implicated in sex determination. The family of corepressors that include *Drosophila* Groucho (*Gro*) and the human transducin-like enhancer of split proteins (*TLE*) are essential for a number of processes like segmentation, eye development and in sex determination (Chen and Courey, 2000). The *WDR10* gene and its mouse homolog showed a predominant expression in the late spermatids in the testis and the expression was confined to the somatic cells in the ovary (Gross et al., 2001). The presence of the WD motif in these proteins made them a prime candidate for mediating interactions between different proteins involved in the sex determination process. A recent report showed that a duplication of the Xp11.23, to which the human *WDR13* is localized, in combination with an intact *SRY* gene lead to sex reversals in females (Ito et al., 1996). This X-chromosome region, Xp11.23 in human and XA1.1 in mouse are referred to as gene rich with a number of candidate genes for diseases like Retinitis Pigmentosa type 2, Wiskott–Aldrich syndrome and X-linked nephrolithiasis assigned to the same (Coleman et al., 1994). Recent studies also reveal that the region Xp11.2-qtter is conserved across mammalian orders with syntenic regions in the chick and fish genomes (Kohn et al., 2004). These observations further show the functional importance of the genes localized to this region. Hence, the present study and the studies of human *WDR13* strongly suggest a function for the gene in the testicular tissue, specifically in the germ cells as indicated by the expression profile and its testis-specific regulation. Its necessity for the maintenance of other cell types is also evident in its identification from the different cancerous tissues. Phylogenetic analysis reveals an extensive conservation of the gene function while the presence of the WD motifs in the different homologs suggests the significance of these motifs. Identification of the interacting proteins by the yeast two-hybrid system, which is in progress in our laboratory, would throw a valuable light on the functioning of the mouse gene, *Wdr13*.

ACKNOWLEDGMENTS

The authors gratefully acknowledge the help of Nandini R., Thangaraj K., and Lakshmi Rao at various stages of experiments. Fellowship to A.S. and S.S. from CSIR is duly acknowledged.

REFERENCES

- Bairoach A, Bucher P, Hofmann K. 1997. The PROSITE database, its status. *Nucleic Acids Res* 25:217–221.
- Black DL. 2000. Protein diversity from alternative splicing: A challenge for bioinformatics and post-genome biology. *Cell* 103:367–370.

- Breitbart RE, Andreadis A, Nadal-Ginard B. 1987. Alternative splicing: A ubiquitous mechanism for the generation of multiple protein isoforms from single genes *Annu Rev Biochem* 56:467–495.
- Chen G, Courey AJ. 2000. Groucho/TLE family proteins and transcriptional repression. *Gene* 16:1–16.
- Church GM, Ephrussi A, Gilbert W, Tonegawa S. 1985. Cell-type-specific contacts to immunoglobulin enhancers in nuclei. *Nature* 313:798–801.
- Coleman MP, Nemeth AH, Campbell L, Raut CP, Weissenbach J, Davies KE. 1994. A 1.8-Mb YAC contig in Xp11.23: Identification of CpG islands and physical mapping of CA repeats in a region of high gene density. *Genomics* 21:337–343.
- Duronio RJ, Gordon JI, Boguski MS. 1992. Comparative analysis of the β -transducin family with identification of several new members including PWP1, a non-essential gene of *S. cerevisiae* that is divergently transcribed from NMT1. *Proteins: Structure, function and Genetics* 13:41–56.
- Gross C, De Baere Lo A, Chang W, Messiaen L. 2001. Cloning and characterization of human WDR10, a novel gene located at 3q21 encoding a WD-repeat protein that is highly expressed in pituitary and testis. *DNA Cell Biol* 20:41–52.
- Gubbay J, Collignon J, Koopman P, Capel B, Economu A, Munsterberg A, Vivian N, Goodfellow P, Lovell-Badge R. 1990. A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* 346:245–250.
- Hacker A, Capel B, Goodfellow P, Lovell-Badge R. 1995. Expression of Sry, the mouse sex determining gene. *Development* 121:1603–1614.
- Hir Le H, Nott A, Moore MJ. 2003. How introns influence and enhance eukaryotic gene expression. *Trends Biochem Sci* 28:215–220.
- Ito T, Kaino Y, Hirai H, Kida K, Nakahori Y, Nakagome Y. 1996. Sex reversal in a child with duplication of sex reversing locus on the short arm of the X chromosome (Xp). *J Pediatr Endocrinol Metab* 9:429–433.
- Kawai J, Shinagawa A, Shibata K, Yoshino M, Itoh M, et al. 2001. Functional annotation of a full-length mouse cDNA collection. *Nature* 409:685–690.
- Kent J, Wheatley SC, Andrews JE, Sinclair AH, Koopman P. 1996. A male specific role for SOX9 in vertebrate sex determination. *Development* 122:2813–2822.
- Koga M, Tanaka H, Yomogida K, Tsuchida J, Uchida K, Kitamura M, Sakoda S, Matsumiya K, Okuyama A, Nishimune Y. 1998. Expression of selenoprotein-P messenger ribonucleic acid in the rat testis. *Biol Reprod* 58:261–265.
- Kohn M, Kehrer-Sawatzki H, Vogel W, Graves JAM, Hameister H. 2004. Wide genome comparisons reveal the origins of the human X chromosome. *TIGS* 20:598–603.
- Kumar S, Tamura K, Nei M. 1994. MEGA: Molecular evolutionary genetics analysis software for microcomputers. *Comput Appl Biosci* 10:189–191.
- Leach T, Michaud S, Quade BJ, Ligon AH, Michelson AM, Maas RL, Korf BR, Herrick SR, Ferguson HL, Morton CC. 2002. Characterization of a t(X;2)(p11.2;q37) de novo associated with congenital anomalies: Positional identification of disrupted genes. ASHG. Development of Genome Anatomy Project Session Abstract.
- Mayer A, Lahr G, Swaab DF, Pilgrim C, Reisert I. 1998. The Y-chromosomal genes *SRY* and *ZFY* are transcribed in adult human brain. *Neurogenetics* 1:281–288.
- Mayer A, Mosler G, Just W, Pilgrim C, Reisert I. 2000. Developmental profile of Sry transcripts in mouse brain. *Neurogenetics* 3:25–30.
- Neer EJ, Schmidt CJ, Nambudripad R, Smith TF. 1994. The ancient regulatory-protein family of WD-repeat proteins. *Nature* 371:297–300.
- Obaru K, Tsuzuki T, Setoyama C, Shimada K. 1988. Structural organization of the mouse aspartate aminotransferase isoenzyme genes. Introns antedate the divergence of cytosolic and mitochondrial isoenzyme genes. *J Mol Biol* 200:13–22.
- Ono T, Kurashige T, Harada N, Noguchi Y, Saika T, Niikawa N, Aoe M, Nakamura S, Higashi T, Hiraki A, Wada H, Kumon H, Old LJ, Nakayama E. 2001. Identification of proacrosin binding protein sp32 precursor as a human cancer/testis antigen. *Proc Natl Acad Sci U S A* 98:3282–3287.
- Palmer JD, Logsdon JM, Jr. 1991. The recent origins of introns. *Curr Opin Genet Dev* 1:470–477.
- Patel M, Dorman SK, Zhang YH, Huang BL, Arnold AP, Sinsheimer SJ, Vilain E, McCabe ERB. 2001. Primate DAX1, SRY, and SOX9: Evolutionary stratification of sex-determination pathway. *Am J Hum Genet* 68:275–280.
- Rodriguez de Turco EB, Tang W, Topham MK, Sakane F, Marcheselli VL, Chen C, Taketomi A, Prescott SM, Bazan NG. 2001. Diacylglycerol kinase epsilon regulates seizure susceptibility and long-term potentiation through arachidonoyl-inositol lipid signaling. *Proc Natl Acad Sci USA* 98:4740–4745.
- Ross TM, Grafam DV, Coffey AJ, Scherer S, McLay K, et al. 2005. The DNA sequence of the human X chromosome. *Nature* 434:325–337.
- Salehi-Ashtiani K, Goldberg E. 1995. Post-transcriptional regulation of primate Ldhc mRNA by its AUUUA-like elements. *Mol Endocrinol* 9:1782–1790.
- Scanlan MJ, Gure AO, Jungbluth AA, Old LJ, Chen YT. 2002. Cancer/testis antigens: An expanding family of targets for cancer immunotherapy. *Immunol Rev* 188:22–32.
- Singh BN, Suresh A, UmaPrasad G, Subramanian S, Sultana M, Goel S, Kumar S, Singh L. 2003. A highly conserved human gene encoding a novel member of WD family of proteins (WDR13). *Genomics* 81:315–328.
- Smith TF, Gaitatzes C, Saxena K, Neer EJ. 1999. The WD-repeat: A common architecture for diverse functions. *Trends Biochem Sci* 24:181–185.
- Spevak W, Keiper BD, Stratowa C, Castanon MJ. 1993. *Saccharomyces cerevisiae* cdc15 mutants arrested at a late stage in anaphase are rescued by *Xenopus* cDNAs encoding N-ras or a protein with beta transducin repeats. *Mol Cell Biol* 13:4953–4966.
- Thompson JD, Higgins DG, Gibson TJ. 1994. CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.
- Tureci O, Sahin U, Zwick C, Koslowski M, Seitz G, Pfreundschuh M. 1998. Identification of a meiosis-specific protein as a member of the class of cancer/testis antigens. *Proc Natl Acad Sci U S A* 95:5211–5216.
- van de Voorn L, Ploegh HL. 1992. The WD-40 repeat. *FEBS Lett* 307:131–134.
- Xu J, Burgoyne PS, Arnold AP. 2002. Sex differences in sex chromosome gene expression in mouse brain. *Hum Mol Genet* 11:1409–1419.