

A *SPGY* copy homologous to the mouse gene *Dazla* and the *Drosophila* gene *boule* is autosomal and expressed only in the human male gonad

Zhihong Shan^{1,+}, Peter Hirschmann^{1,+}, Thomas Seebacher¹, Angela Edelmann¹, Anna Jauch¹, Jane Morell², Peter Urbitsch¹ and Peter H. Vogt^{1,*}

¹Reproductiongenetics in Institute of Human Genetics, University of Heidelberg, D-69120 Heidelberg, Germany and ²Department Reproductionbiology, German Primate Center GmbH, D-37077 Göttingen, Germany

Received August 15, 1996; Revised and Accepted September 20, 1996

We have isolated a series of human testis poly(A) cDNA clones by cross-hybridization to *SPGY1*, a Y gene homologous to *DAZ*. Their sequence analysis revealed an identical nucleotide composition in different 'full-length' clones, suggesting that all were encoded by the same gene. We mapped this gene to the short arm of chromosome 3 and designated it *SPGYLA* (*SPGY* like autosomal). Comparison of the *SPGYLA* cDNA sequence with the cDNA sequences of *DAZ* and *SPGY1* revealed two prominent differences. The tandem repetitive structure of 72 bp sequence units (*DAZ* repeats) is absent. *SPGYLA* contains only one 72 bp sequence unit. Downstream of it, a specific 130 bp sequence domain is present which is absent in *DAZ* and *SPGY1* but present in the mouse gene *Dazla* and in the *Drosophila* gene *boule*. *SPGYLA* encodes an RNA binding protein expressed only in the human male gonad. The data presented give strong evidence that not *DAZ* but *SPGYLA* is the functional human homologue of *Dazla* and *boule*.

INTRODUCTION

In eukaryotes, RNA binding proteins play important roles in various aspects of the splicing process and in the nuclear export of primary transcripts, in cytoplasmic RNA stability and during translation (1). One well characterized group of RNA binding proteins contain one or more copies of the so-called RNA recognition motif (RRM) in their N-terminal regions. The RRM domain consists of ~80 amino-acids (2). Most RRM proteins are expressed constitutively. However, cell type-specific RRM proteins also exist and seem to be functional in the brain (3,4), during sex determination (*Drosophila*; 5,6) and in the germ line (in *Drosophila*, for spermatogenesis, 7; in the mouse, for spermatogenesis and oogenesis, 8,9; in human, for spermatogenesis, 10–12). The recent discovery of the *Drosophila* gene *boule* with a high homology of its RNA binding domain to that of the

human gene *DAZ* (13) suggests that RNA binding proteins in the germ line are highly conserved and might fulfil essential regulative functions during the complex process of spermatogenesis. Gene products of *boule* function during the male meiotic cell cycle at the G2–M transition, and loss of *boule* function results in azoospermia (13). Therefore, it was argued that *DAZ* (because of its distinct sequence homology to *boule*) must have an essential function during the meiotic cell cycle in human spermatogenesis. However, this is inconsistent with the fact that the total deletion of *DAZ* can be inherited from father to son (14) and that deletion of *DAZ* is not always correlated with azoospermia but also with other sterile phenotypes (14,15). Therefore, *DAZ* deletions seem to be compatible with development of motile spermatozoa, and the primary mutation effect in men with deletion of *DAZ* is subfertility but not azoospermia (14,16). These results suggested that the functional human homologue of *boule* had not yet been isolated and that *DAZ* might not have the same function as *boule* in spermatogenesis (17).

The mouse homologue of *boule* is located on chromosome 17 (9). Because of its distinct sequence homology to *DAZ*, it was designated *Dazla* (*DAZ* like autosomal). Therefore, if a human homologue of *boule/Dazla* exists, it is expected to be autosomal and located on the human homologue of mouse chromosome 17, chromosome 6 (17,18).

We have analysed this hypothesis experimentally by screening different cDNA libraries prepared from adult human testis poly(A) RNA with the cDNA probe *SPGY1*, a Y gene homologous to *DAZ* (11). *SPGY1* is member of the *SPGY* (spermatogenesis gene on the Y) locus in distal Yq11 containing several *DAZ*-related genes, in addition to *DAZ* itself (12,16). Their transcripts are characterized by different numbers of a tandem repetitive 72 bp sequence unit (*DAZ* repeat). *DAZ* contains seven such repeats (11), *SPGY1* contains 12 (12,14).

Sequence analysis of *SPGY1* homologous cDNA clones identified a subclass of clones with a common sequence structure absent in *DAZ* or *SPGY1* but present in *boule* and *Dazla*. Two features were most prominent. The 72 bp sequence unit (*DAZ* repeat) was not present in a tandem repetitive structure but as a single sequence unit like in *Dazla* and *boule*. Adjacent to it, a 130 bp sequence domain was analysed which is absent in *DAZ* and

*To whom correspondence should be addressed

+These authors contributed equally to this work

SPGY1 but present in *Dazla* and *boule*. Here we present the isolation of the complete cDNA sequence of this *SPGY* gene and the sequence of its putative translation product. It maps to the short arm of chromosome 3 and therefore is the first *SPGY* copy mapped to an autosome. We designated it as *SPGYLA* (*SPGY* like autosomal). It encodes RNA binding protein only expressed in the human male gonad. Our results strongly suggest that not *DAZ* but *SPGYLA* is the functional human homologue of *Dazla* and *boule* in human spermatogenesis.

RESULTS

A *SPGY* copy with specific cDNA structure is located on chromosome 3

During analysis of the genomic sequence structure of *SPGY* genes by DNA blot experiments, we found that not all *SPGY* genes could be located in the *AZFc* region in distal Yq11. Azoospermic men with deletion of *AZFc* retained a cross-hybridizing genome fragment also present in the female genome (12). The same was observed in blots with the *DAZ* probe (11). We were able to identify this non-Y gene copy by analysis of a series of cDNA clones isolated by homology to *SPGY1*. They were characterized by the absence of all but one of the tandem repetitive 72 bp sequence units (*DAZ* repeats) found in *DAZ* (11) and in *SPGY1* (12,14), and the presence of a specific sequence unit of 130 nucleotides (sequence position 762–891 in full-length cDNA clone CT355 submitted to the GenBank database; accession no. U66726) which is absent in the sequence of *DAZ*, *SPGY1* and other cDNA clones with tandem repetitive 72 bp sequence units.

All cDNA clones containing this sequence structure (one 72 bp sequence unit plus the specific 130 bp sequence unit downstream of it) were identical, so presumably were derived from one *SPGY* gene. It was mapped by fluorescence *in situ* hybridization (FISH) analysis to the short arm of chromosome 3 (Fig. 1). As other *SPGY* copies (analysed so far) were mapped to the *AZFc* region in distal Yq11 (14,16), the new autosomal copy was designated as *SPGYLA* (*SPGY* like autosomal). We scrutinized the mapping position of *SPGYLA* by PCR analysis of a panel of human–hamster hybrid cell lines containing overlapping subsets of human chromosomes. With the *SPGYLA*-specific primer pair (*LA1-for/LA1-rev*), a PCR product was present in all hybrid cells containing human chromosome 3.

When we searched for homology of the *SPGYLA*-specific 130 bp sequence domain in different databases, a sequence of the same length was identified in the mouse gene *Dazla* (sequence position 932–1061 in *Dazla* clone pSG4; accession no. U38690) and a sequence of similar length in the *Drosophila* gene *boule* (position 1362–1485 in *boule* sequence; accession no. U51858). This homology was confirmed at the protein level. Translation of *SPGYLA* into its putative protein sequence revealed homology to the protein of *Dazla* along its whole length of 295 amino acids (Fig. 2a). Comparison of these proteins with the *boule*-encoded protein revealed a high sequence similarity in the regions of their RNA recognition motif (RRM) but also downstream of it, e.g. along the 24 amino acids encoded by the conserved 72 bp sequence unit (*DAZ* repeat) as described already for *boule/Dazla* by Eberhart *et al.* (13). In Figure 2b and c, we display these homologous sequence domains together with that of the *DAZ* protein. We conclude that *SPGYLA* most likely also encodes an RNA binding protein.



Figure 1. Fluorescence *in situ* hybridization of the *SPGY* copy CT355 to human male metaphase chromosomes. CT355 hybridizes exclusively to the short arms of two metacentric chromosomes identified as homologues of chromosome 3 by DAPI banding (see picture insert). The chromosomes were stained by propidium iodide.

SPGYLA is expressed only in adult human testis

Gene copies of the *SPGY* locus in distal Yq11 are expressed only in human testis (12,16), but *Dazla* is expressed in the mouse female gonad as well (9). We analysed the tissue expression pattern of *SPGYLA* by RNA blot and RT–PCR analyses. In the RNA blot experiment, we used a probe of *SPGYLA* containing only its 3′-untranslated region (UTR) in order to exclude potential cross-hybridizations to testis transcripts of other *SPGY* genes. A strong hybridization signal was visible with the *SPGYLA* 3′ UTR probe to a testis poly(A) RNA population of 2.9 kb length (Fig. 3a). This length is similar to the length of the isolated cDNA clone CT355 (2924 bp), suggesting that it corresponds to a full-length testis transcript of the corresponding *SPGYLA* gene. Additional homology to testis RNA of 2.1 kb was indicated by a weak second hybridization signal. RNA pools prepared from male kidney and male liver in the same experiment did not reveal any cross-hybridization signals. To confirm the testis specificity of *SPGYLA* transcription in males and to analyse its presence in female tissues, we performed RT–PCR with the *SPGYLA*-specific primer pair (*LA-1for/LA-1rev*). An RT–PCR product was only visible in testis tissue and was absent in all female tissues analysed (Fig. 3b). The tissue expression pattern of *SPGYLA* therefore resembles the pattern of *DAZ* and *SPGY1*, but not the pattern of the mouse gene *Dazla*.

A copy of *SPGYLA* translocated to the Y chromosome during primate evolution

Male-specific genomic DNA fragments cross-hybridizing to the *Dazla* probe were only observed in the genome of human and gorilla (9). As *SPGYLA* and *Dazla* sequences are homologous to

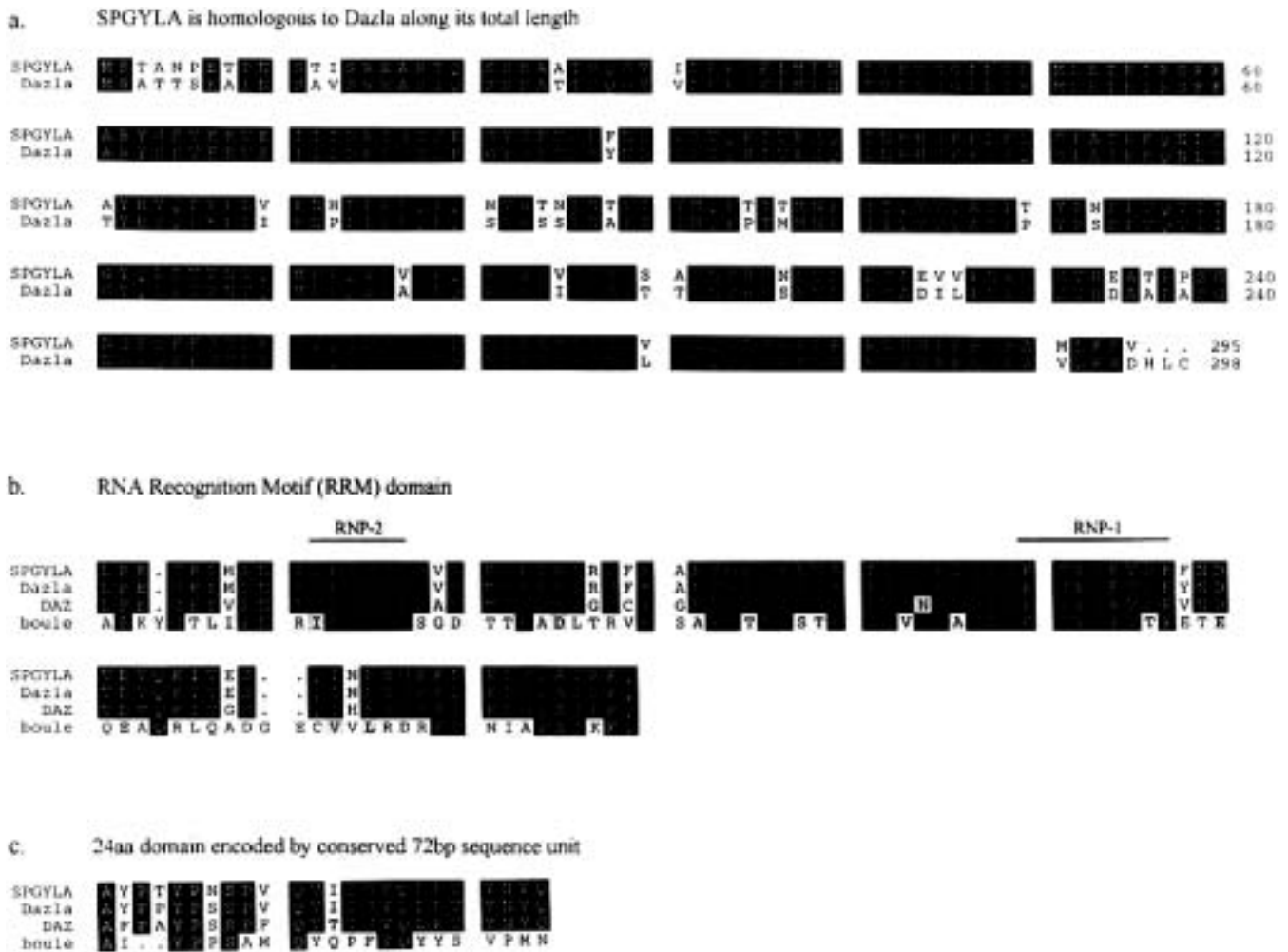


Figure 2. (a) Comparison of the putative translation products encoded by *SPGYLA* and *Dazla*. The *SPGYLA*-encoded protein is homologous to the mouse *Dazla*-encoded protein (89% identity and 93% similarity) along the whole sequence. Homology of the RRM domains (b) and 24 amino acid domains (c) encoded by the conserved 72 bp sequence units (*DAZ* repeat) between *SPGYLA*, *Dazla*, *DAZ* and *boule*. The 24 amino acid domain of the *DAZ*-encoded protein was translated from the 72 bp consensus sequence unit (11). The highly conserved protein boxes of the RRM domain: RNP-1 (an octapeptide) and RNP-2 (a hexapeptide) are marked. Both peptide boxes are identical in the *SPGYLA*-, *Dazla*- and *DAZ*-encoded proteins, and in the *boule*-encoded protein with two conservative amino acid exchanges. Protein sequences were aligned by the software tool PRETTYBOX highlighting identical amino acids by black boxes. Conservative amino acid exchanges are marked by grey boxes according to a matrix scale based on the Dayhoff table (in *Atlas of Protein Sequence and Structure*, 1979, pp. 353–358).

each other, we expected a similar pattern of cross-hybridization in zoo-DNA blots with the human *SPGYLA* probe. Consequently, male-specific hybridization fragments were observed in the genome of chimpanzee (*Pan troglodytes*) and gorilla (*Gorilla gorilla*) (data not shown). However, in a zoo-DNA blot containing genomic DNA samples of a new world monkey (Platyrrhini: *Callithrix jacchus*) and of two old world monkeys (Catarrhini: *Macaca mulatta*, *Macaca nemestrina*), male-specific *SPGYLA* cross-hybridizing fragments were only present in the DNA of the old world monkeys but not in the DNA of the new world monkey *Callithrix* (Fig. 4). Cross-hybridization signals in the mouse genome, the cattle genome and the wallaby genome (*Macropus eugenii*) marked the evolutionary conservation of the *SPGYLA* sequence structure.

Male-specific genomic *SPGYLA* fragments were not only present in macaques but also in other Catarrhini. Examination of corresponding blot patterns in baboons (*Papio hamadryas*; an old world monkey) and tamarins (*Saguinus fuscicollis*; a new world

monkey) presented the same basic result. Cross-hybridization to male-specific fragments was only present in the DNA of the old world monkey but not in the DNA of the new world monkey. These results suggest that during primate evolution translocation of a *SPGYLA* copy to the Y chromosome had occurred after divergence of the lineage of Platyrrhini and before divergence of the lineage of Catarrhini. This time interval ranges between 36 and 55 million years ago (19).

DISCUSSION

SPGYLA is the human homologue of *Dazla* and *boule*

It is well known that genes expressed in different tissues create testis-specific RNAs by alternative splicing. Moreover, genes expressed in testis and transcribed pre-meiotically might be translated only post-meiotically (20–22). This indicates that in the male germ line, the availability of functional gene products

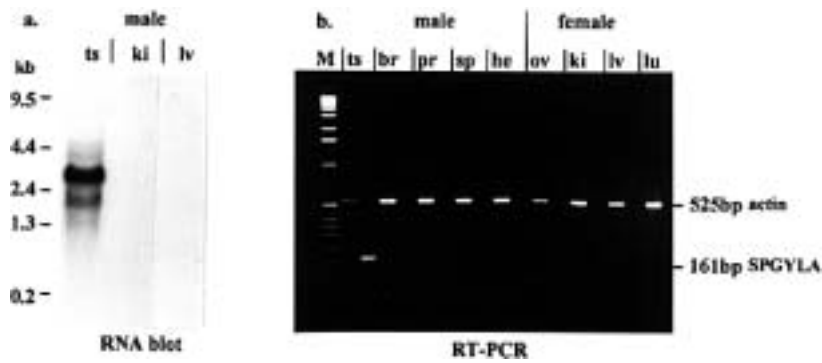


Figure 3. Analysis of *SPGYLA* expression in human tissues: (a) RNA blot hybridization performed with a *SPGYLA*-specific 3' UTR probe prepared as described in Materials and Methods. The RNA blot contained poly(A) RNA of human testis (ts), of male kidney (ki) and of male liver (lv). A strong hybridization signal was observed only in testis tissue to a RNA population of 2.9 kb; a second weak testis RNA signal was obtained at 2.1 kb. RNA marker lengths are indicated at the left in kb (GIBCO-BRL: 15620). (b) RT-PCR with the *SPGYLA*-specific primer LA1-for/LA1-rev was performed on RNA samples of different male tissues (ts: testis; br: brain; pr: prostate; sp: spleen; he: heart) and female tissues (ov: ovary; ki: kidney; lv: liver; lu: lung). The first lane on the left (M) shows fragments of a 1 kb DNA ladder marker (Gibco-BRL: 15615). *SPGYLA* homologous transcripts (identified by the 161 bp RT-PCR product) were only observed in testis tissue. In a parallel control experiment, the same RNA samples were analysed for transcripts of the constitutively expressed human β -actin gene. Primers actin-forward/actin-reverse amplified a homologous RNA product (identified by the 525 bp RT-PCR product) in each tissue analysed.

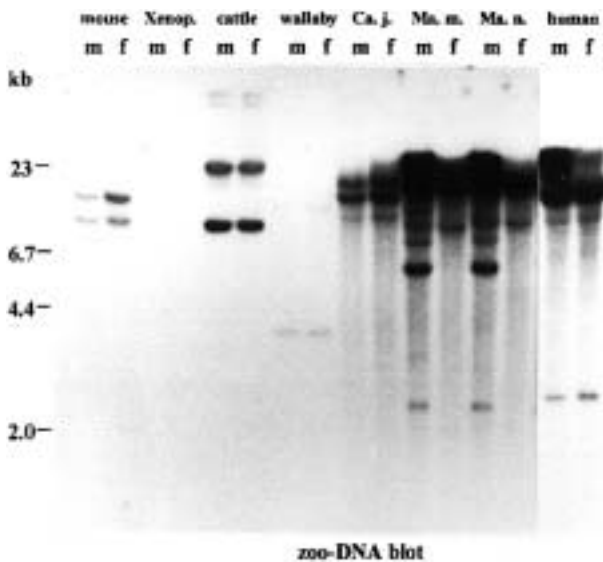


Figure 4. Hybridization pattern of *SPGYLA* probe CT355 to a zoo-DNA blot. Male (m) and female (f) genomes of mouse, *Xenopus*, cattle, *Macropus eugenii* (wallaby), *Callithrix jacchus* (Ca. j.), *Macaca mulata* (Ma. m.), *Macaca nemestrina* (Ma. n.), and human were restricted with *Bam*HI. Cross-hybridization of CT355 is observed to genomic DNA fragments of all species except *Xenopus*. The weak signals in the wallaby genomes are due to a low amount of DNA in these lanes. Male-specific cross-hybridization fragments are observed only in the genomes of both macaques and of human. Hybridization stringency was normal as described earlier (14). Filter wash after hybridization was done in $2\times$ SSC, 0.5% SDS at 65°C . The human blot filter was exposed for a shorter time in order to visualize the male-specific *Bam*HI fragment. Positions of λ -*Hind*III marker fragment lengths are the same and indicated at the left in kb.

might be regulated primarily by post-transcriptional processes. Testis-specific RNA binding proteins are strong candidates to be involved in such processes. They were described for *Drosophila*, (7,13), mouse (8,9,23) and human (10–12). All contain at least one copy of the highly conserved RRM in the N-terminal part of their protein structure. RRM domains consist of ~80 amino acids.

Two so-called RNP boxes (RNP-1, an octapeptide; RNP-2, a hexapeptide) can be matched in RRM domains of RNA binding proteins in yeast as in human. This suggests that RRM domains have a very ancient origin and that the mode of their interaction with specific RNAs is highly conserved. The specificity of the RNA–protein binding reaction is thought to be achieved by auxiliary sequence blocks in the RRM domain and the C-terminal part of the RRM protein structure (2,24). By these criteria, we suppose that the protein of *SPGYLA* presented here is the functional human homologue of the protein encoded by the mouse gene *Dazla* and most likely also of the protein encoded by the *Drosophila* gene *boule*.

SPGYLA- and *Dazla*-encoded proteins have similar molecular weights (*SPGYLA*, 33 178 Da; *Dazla*, 33 312 Da) and the sequences are 89% identical (93% similar) along their total length (Fig. 2a). In contrast to this, the *DAZ*-encoded protein (with one consensus 72 bp sequence unit encoding 24 amino acids) has a higher molecular weight (44 300 Da) and is homologous to the *Dazla*-encoded protein only along 201 amino acids. Their C-terminal peptides are different because of the presence of the 130 bp sequence unit in the protein-coding region of *SPGYLA* and *Dazla* and its absence in *DAZ*. Therefore, our supposition is that not *DAZ* but *SPGYLA* is the functional human homologue of the mouse gene *Dazla*.

A functional homology between the *boule*- and *SPGYLA*-encoded proteins is also most likely but, due to the large evolutionary distance between *Drosophila* and human, it is difficult to judge it by sequence comparison. Both are RNA binding proteins with different conserved domains including the RRM domain (Fig. 2b). Moreover, the cDNA structure of *boule* is homologous to the cDNA structure of *Dazla* and *SPGYLA*. All three contain one homologous 72 bp sequence unit and the 130 bp sequence unit downstream of it. A functional homology of all three proteins is therefore most likely.

SPGYLA is located on the short arm of chromosome 3

Mapping of human genes homologous to mouse genes of mouse chromosome 17 show that they are clustered on the short and long

arm of human chromosome 6 (18). Therefore, it was suggested recently that the functional human homologue of *Dazla/boule* should also be located on chromosome 6 (17). However, FISH analysis of the location of *SPGYLA* on human chromosomes does not support this prediction. A distinct hybridization signal was only observed on the short arm of chromosome 3 (Fig. 1). No additional FISH signals were observed on chromosome 6. This result was confirmed by mapping of *SPGYLA* on a human–hamster hybrid cell panel. It is intriguing to note that, despite its sequence homology to *SPGY* genes in distal Yq11, the *SPGYLA* probe does not hybridize on the Y chromosome. We explain this result by the presence of a 1.9 kb long 3' UTR domain in the *SPGYLA* cDNA structure. This sequence region is specific for this gene and represented 66% of the length of the *SPGYLA* FISH probe used.

No evidence exists for the presence of a male fertility locus on chromosome 3 by genetic or molecular linkage analysis (25). Comparative mapping of DNA loci of human chromosome 3 to mouse chromosomes indicates that this human chromosome is composed of chromatin domains of different mouse chromosomes (3, 6, 9, 14 and 16) (25), and now also of 17.

Function of *SPGYLA* in human spermatogenesis

SPGYLA is only expressed in the human male gonad. Its transcription product is polyadenylated and has a length of 2.9 kb. The second weak cross-hybridizing RNA signal (2.1 kb) observed in the testis RNA blot is difficult to explain at the moment. Cross-hybridization to testis transcripts of other Y chromosomal *SPGY* genes was excluded by using a *SPGYLA*-specific 3' UTR probe. Moreover, *DAZ* and *SPGYI* hybridized to testis RNA of 3.5 kb (11,12). It is interesting to note that the mouse homologue *Dazla* also hybridized to two different testis RNA populations (9). Our result might therefore reflect the use of an alternative polyadenylation site in *SPGYLA* transcripts. A second possibility may be that a divergent *SPGYLA* copy is present on the same or another autosome or on the Y chromosome as discussed below.

The expression pattern of *SPGYLA* is different from the pattern of its mouse homologue *Dazla* which is expressed also in the female gonad (9). We have examined the expression of *SPGYLA* in human ovaries of different ages (all before menopause) in order to exclude a possible age-dependent inactivation of the *SPGYLA* gene in the female gonad. However, we were not able to detect any *SPGYLA* transcripts in any ovary tissue by RT-PCR. We cannot exclude expression of *SPGYLA* in certain ovarian stages not available from human females for medical reasons. We also cannot exclude the possibility that the ovarian tissues analysed do not show *SPGYLA* transcription because of their pathological aetiology.

The function of *Dazla* in the male and female gonad is not yet known. The function of *boule* most likely occurs during the male meiotic cell cycle at the G2–M transition (13). Loss of *boule* results in azoospermia. Therefore, if *SPGYLA* is the functional homologue of *boule*, its function is expected to be at human male meiosis, and *SPGYLA* mutations should cause azoospermia.

SPGYLA and the *SPGY* locus

It is attractive to speculate that Y genes of the *SPGY* locus in distal Yq11 containing a repetitive 72 bp sequence unit (like *DAZ* and

SPGYI) might have evolved from an functional ancestor *SPGYLA* copy on the Y chromosome with one 72 bp sequence unit. We were able to show that each 72 bp sequence unit is a separate exon unit in the *SPGY* locus (Vogt *et al.*, in preparation). However, molecular analysis of all *SPGY* genes in distal Yq11 is an essential prerequisite to distinguish between *SPGY* gene structures containing repetitive 72 bp exon units and *SPGY* genes containing only one 72 bp exon unit like *SPGYLA*. This work is in progress.

Our analysis of the presence of *SPGYLA* in different primate species suggests that a copy of *SPGYLA* was translocated to the Y chromosome at a time interval between 36 and 55 million years ago, after divergence of Platyrrhini and before divergence of Catarrhini. Although, we have not yet confirmed the presence of *SPGYLA* on the Y chromosome of the macaques and baboon by FISH analysis, the male-specific fragments of their genomic DNA blot patterns can be explained most easily by the presence of at least one *SPGYLA* copy on their Y chromosomes. The presence of a *SPGYLA* copy on the human Y chromosome, therefore, also cannot yet be excluded. However, because we detect no FISH signal with the *SPGYLA* probe on the human Y chromosome, its sequence structure is expected to be divergent from that of the autosomal *SPGYLA* gene described here.

MATERIALS AND METHODS

DNA analysis

Hybridization of filter discs (duplicates) prepared from plated testis poly(A) cDNA libraries (Clontech: HL1010b; HL3024b) with the probe *SPGYI* was performed as described by Quertermous (26). Sequence analyses of selected cDNA clones were performed on the Pharmacia A.L.F. express using the Sanger dideoxy chain termination method (Autoread kit, Pharmacia) and the strategy of primer walking on overlapping subfragments from both DNA strands. The sequence of the full-length clone CT355 was submitted to the GenBank database; (accession no. U66726). For sequence alignment analyses, we used the PRETTYBOX and PRETTYPLOT software tools written by Peter Rice, Informatics Division, The Sanger Centre, Cambridge, UK, and part of the HUSAR software package of the German Cancer Institute, Heidelberg.

Genomic DNA samples from peripheral blood lymphocytes of human individuals and animals, and DNA samples of hybrid cell lines were prepared using standard protocols. Conditions used for restriction, blotting, probe labelling and hybridization were the same as described previously (14). Genomic PCR for analysis of the specific 130 bp sequence region in CT355 (*SPGYLA*) was performed with 200 ng of DNA and the primer pair: LA-1for (5'-ATGCCACCACAGTGGCCTGTTG-3') and LA-1rev (5'-AACACCCGGTAAAGGTCTCCCA-3') using the Expand™ Long Template PCR system (Boehringer). Conditions for amplification were: 10× 1 min at 97°C, 1 min at 65°C, 5 min at 68°C followed by 15 cycles of 1 min at 94°C, 1 min at 65°C, 5 min at 68°C. The final elongation time was 5 min at 68°C. The template was denatured for 5 min at 98°C.

RNA analysis

RNA samples were extracted from the different tissues indicated in the text using the standard guanidinium isothiocyanate procedure (27) or ordered from Clontech (male tissues: brain, cat.

no. 64020; heart, cat. no. 64025-1; spleen, cat. no. 64034-1; prostate, cat. no. 64038; female tissues: kidney, cat. no. 63047; liver, cat. no. 64022-1; lung, cat. no. 64023). Poly(A) RNA was purified on oligo(dT) columns (Pharmacia). For RNA blot experiments, RNA samples (20 µg) were electrophoresized on 1% agarose gels containing 2.2 M formaldehyde as denaturing agent. Before blotting, RNAs in the gel were partially hydrolysed in 50 mM NaOH/10 mM NaCl and neutralized in 0.1 M Tris/10× SSC buffer (pH 7.5). Blots were pre-hybridized (5 h) and hybridized (overnight) in a 50% formamide buffer (5× SSPE, 10× Denhardt blocking solution, 2% SDS); 20× SSPE is: 3 M NaCl, 0.2 M NaH₂PO₄, 0.2 M EDTA (pH 7.4). For the RNA blot experiment we used a *SPGYLA*-specific 3' UTR probe in order to exclude potential cross-hybridizations to testis transcripts of other *SPGY* genes. The 3' UTR probe was prepared by amplifying the last 1644 nucleotides of CT355 (positions 1277–2921) by PCR using the CT355 forward primer 355P1 (5'-GAACTAAGATT-TATCTGCAAGG-3') and a vector primer (λ -gt11 rev: 5'-TTGACACCAGACCAACTGGTAATG-3'). Cycling conditions: 25× 1 min at 94°C, 1 min at 55°C, 5 min at 72°C. Final extension was for 10 min at 72°C. Denaturation was for 4 min at 94°C. After hybridization, filters were washed under stringent conditions (2× SSC, 0.1% SDS, 65°C). RT-PCR experiments were performed from cDNA aliquots prepared with the SUPERSRIPT™ II system of Gibco-BRL from all RNA samples indicated in the text. The LA-1for/LA-1rev primer pair was used for analysis of RNA tissue specificity with the following cycling conditions: 30× 1 min at 94°C, 1 min at 65°C, 1.5 min at 72°C. The final extension was for 5 min at 72°C. Denaturation was for 5 min at 94°C. The length of the RNA amplification product is 161 bp (see Fig. 2b). As a positive control in all RT-PCR experiments, we used actin-forward (5'-ATCCTGCGTCTGGACCTG-3') and actin-reverse (5'-GCTGATCCACATCTGCTG-3'), which are two primers of a human β -actin gene (28), under the same cycling conditions. They are expected to amplify an RNA product of 525 bp in each tissue analysed.

FISH analysis

Fluorescence *in situ* hybridization with CT355 on human male metaphase chromosomes was carried out as described by Köhler and Vogt (29). Both probes were biotinylated by nick translation after a limited DNase I fragmentation. Hybridization signals were amplified once using a biotinylated anti-avidin antibody (Vector Laboratories). Fluorescent signals were analysed in a Zeiss Axiophot microscope with an aligned filter set for DAPI (LP450-490, BP365, FT395) and FITC (LP515-565, BP450-490, FT510). The Axiophot was coupled to a cooled charge-coupled device (CCD) camera for further image processing using the software package 'Gene JOIN' as described by Ried *et al.* (30).

ACKNOWLEDGEMENTS

We thank Weiyun Chen at the Biocomputing Unit of the German Cancer Institute for generous support in sequence analyses with the HUSAR software package. Margaret Delbridge (La Trobe University, Bundoora, Australia) is thanked for the wallaby DNA sample. Luda Diatchenko (Clontech Laboratories, Inc, Palo Alto, CA) is thanked for valuable RNA samples of human male and female tissues. Volker Wetzels (Karlsruhe, Germany) is thanked for valuable ovarian tissue samples. We are indebted to Karl-

Heinz Grzeschik (University of Marburg, Germany) for providing us with DNA samples of his human–hamster hybrid cell panel. Wolfgang Hennig (PUDONG, Kranenburg, Germany) is thanked for support in automated sequence analysis. Annemarie Wiegenstein is thanked for excellent photographic assistance and Christine Clayton for critically improving our English grammar and expression. Zhihong Shan is a recipient of a Ph.D. fellowship of the C.C. Wu Cultural & Education Foundation Fund, Hong Kong. This work was supported by a grant from the DLR (Bundesministerium für Bildung und Forschung) to P.H.V. (01KY95074).

REFERENCES

- Dreyfuss, G., Matunis, M.J., Pinol-Roma, S. and Burd, C.G. (1993) hnRNP proteins and the biogenesis of mRNA. *Annu. Rev. Biochem.*, **62**, 289–321.
- Kenan, D.J., Query, C.C. and Keene, J.D. (1991) RNA recognition: towards identifying determinants of specificity. *Trends Biochem. Sci.*, **16**, 214–220.
- Robinow, S., Campos, A.R., Yao, K.-M. and White, K. (1988) The *elav* gene product of *Drosophila*, required in neurons, has three RNP consensus motifs. *Science*, **242**, 1570–1572.
- Levine, T.D., Gao, F., King, P.H., Andrews, L.G. and Keene, J.D. (1993) Hel-N1: an autoimmune RNA-binding protein with specificity for 3' uridylate-rich untranslated regions of growth factor mRNAs. *Mol. Cell. Biol.*, **13**, 3494–3504.
- Amrein, H., Gorman, M. and Nöthiger, R. (1988) The sex determining gene *tra-2* of *Drosophila* encodes a putative RNA binding protein. *Cell*, **55**, 1025–1035.
- Bell, L.R., Maine, E.M., Schedl, P. and Cline, T.W. (1988) Sex-lethal, a *Drosophila* sex determination switch gene, exhibits sex-specific RNA splicing and sequence similarity to RNA binding proteins. *Cell*, **55**, 1037–1046.
- Karsch-Mizrachi, I. and Haynes, S.R. (1993) The *Rb97D* gene encodes a potential RNA-binding protein required for spermatogenesis in *Drosophila*. *Nucleic Acids Res.*, **21**, 2229–2235.
- Kleene, K.C., Wang, M.-Y., Cutler, M., Hall, C. and Shih, D. (1994) Developmental expression of poly(A) binding protein mRNAs during spermatogenesis in the mouse. *Mol. Reprod. Dev.*, **39**, 355–364.
- Cooke, H.J., Lee, M., Kerr, S. and Ruggiu, M. (1996) A murine homologue of the human *DAZ* gene is autosomal and expressed only in male and female gonads. *Hum. Mol. Genet.*, **5**, 513–516.
- Ma, K., Inglis, J.D., Sharkey, A., Bickmore, W.A., Hill, R.E., Prosser, E.J., Speed, R.M., Thomson, E.J., Jobling, M., Taylor, K., Wolfe, J., Cooke, H.J., Hargreave, T.B. and Chandley, A.C. (1993) A Y chromosome gene family with RNA-binding protein homology: candidates for the azoospermia factor *AZF* controlling human spermatogenesis. *Cell*, **75**, 1287–1295.
- Reijo, R., Lee, T.-Y., Salo, P., Alagappan, R., Brown, L.G., Rosenberg, M., Rozen, S., Jaffe, T., Straus, D., Hovatta, O., de la Chapelle, A., Silber, S. and Page, D.C. (1995) Diverse spermatogenic defects in humans caused by Y chromosome deletions encompassing a novel RNA-binding protein gene. *Nature Genet.*, **10**, 383–393.
- Maiwald, R., Seebacher, T., Edelmann, A., Hirschmann, P., Köhler, M.R., Kirsch, S. and Vogt, P.H. (1996) A human Y-chromosomal gene mapping to distal Yq11 is expressed in spermatogenesis. In Affara *et al.* (eds), *Report of the Second International Workshop on Y Chromosome mapping, Asilomar 1995*. *Cytogenet. Cell Genet.*, **73**, 33–76.
- Eberhart, C.G., Maines, J.Z. and Wasserman, S.A. (1996) Meiotic cell cycle requirement for a fly homologue of human *Deleted in Azoospermia*. *Nature*, **381**, 783–785.
- Vogt, P.H., Edelmann, A., Kirsch, S., Henegariu, O., Hirschmann, P., Kiesewetter, F., Köhn, F.M., Schill, W.B., Farah, S., Ramos, C., Hartmann, M., Hartschuh, W., Meschede, D., Behre, H.M., Castel, A., Nieschlag, E., Weidner, W., Gröne, H.-J., Jung, A., Engel, W. and Haidl, G. (1996) Human Y chromosome azoospermia factors (*AZF*) mapped to different subregions in Yq11. *Hum. Mol. Genet.*, **5**, 933–943.
- Reijo, R., Alagappan, R.K., Patrizio, P. and Page, D.C. (1996) Severe oligozoospermia resulting from deletions of azoospermia factor gene on Y chromosome. *Lancet*, **347**, 1290–1293.
- Vogt, P.H. (1996) Human Y chromosome function in male germ cell development. *Adv. Dev. Biol.*, **4**, 193–258.

17. Burgoyne, P.S. (1996) Fruit(less) flies provide a clue. *Nature*, **381**, 740–741.
18. Copeland, N.G., Jenkins, N.A., Gilbert, D.J., Eppig, J.T., Maltais, L.J., Miller, J.C., Dietrich, W.F., Weaver, A., Lincoln, S.E., Steen, R.G., Stein, L.D., Nadeau, J.H. and Lander, E.S. (1993) A genetic linkage map of the mouse: current applications and future prospects. *Science*, **262**, 57–66.
19. Martin, R.D. (1993) Primate origins: plugging the gaps. *Nature*, **363**, 223–233.
20. Baker, B.S. (1989) Sex in flies: the splice of life. *Nature*, **340**, 521–524.
21. Hake, L.E., Alcivar, A.A. and Hecht, N.B. (1990) Changes in mRNA length accompany translational regulation of the somatic and testes-specific cytochrome c genes during spermatogenesis in the mouse. *Development*, **110**, 249–257.
22. Kleene, K.C., Distel, R.J. and Hecht N.B. (1984) Translational regulation and deadenylation of a protamine mRNA during spermiogenesis in the mouse. *Dev. Biol.*, **105**, 71–79.
23. Elliott, D.J., Ma, K., Kerr, S.M., Thakrar, R., Speed, R., Chandley, A.C. and Cooke, H. (1996) An *RBM* homologue maps to the mouse Y chromosome and is expressed in germ cells. *Hum. Mol. Genet.*, **5**, 869–874.
24. Burd, C.G. and Dreyfuss, G. (1994) Conserved structures and diversity of functions of RNA-binding proteins. *Science*, **265**, 615–621.
25. Naylor, S.L., Carritt, B., Boileau, C., Beroud, C., Alexander, C., Allderdice, P., Alimov, A., Ashworth, T., Bonifas, J., Bugert, P., Buys, C.H.C.M., Chipperfield, M.A., Deng, G., Drabkin, H., Gemmill, R.M., Grompe, M., Joensuu, T., Jonasdottir, A., Gizatullin, R., Krols, L., Leach, R.J., Lott, S.T., Killary, A., Martinsson, T., Messiaen, L., O'Connell, P., Opalka, B., Plaetke, R., Sankila, E.-M., Smith, D.I., Strachen, T., van den Berg, A. and Zabarovsky, E. (1996) Report of the sixth international workshop on human chromosome 3 mapping 1995. *Cytogenet. Cell Genet.*, **72**, 255–270.
26. Quertermous, T. (1989) Plating and transferring bacteriophage libraries. In Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A. and Struhl, K. (eds), *Current Protocols in Molecular Biology*. John Wiley & Sons, New York, Vol. 1, pp. 6.1.1–6.1.4.
27. Chirgwin, J.M., Przybyla, A.E., MacDonald, R.J. and Rutter, W.J. (1979) Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. *Biochemistry*, **18**, 5294–5299.
28. Ponte, P., Ng, S.-Y., Engel, J., Gunning, P. and Kedes, L. (1984) Evolutionary conservation in the untranslated regions of actin mRNAs: DNA sequence of a human β -actin cDNA. *Nucleic Acids Res.*, **12**, 1687–1696.
29. Köhler, M.R. and Vogt, P.H. (1994) Interstitial deletions of repetitive DNA blocks in dicentric human Y chromosomes. *Chromosoma*, **103**, 324–330.
30. Ried, T., Landes, G., Dackowski, W., Klinger, K. and Ward, D.C. (1992) Multicolor fluorescence *in situ* hybridization for the simultaneous detection of probe sets for chromosome 13, 18, 21, X and Y in uncultured amniotic fluid cells. *Hum. Mol. Genet.*, **1**, 307–313.