

Differential gene expression of mammalian *SPO11/TOP6A* homologs during meiosis¹

Mark Shannon^a, Laura Richardson^b, Allen Christian^a, Mary Ann Handel^b,
Michael P. Thelen^{a,*}

^aMolecular and Structural Biology Division, Biology and Biotechnology Research Program, Lawrence Livermore National Laboratory, L452, P.O. Box 808, Livermore, CA 94550, USA

^bDepartment of Biochemistry and Cellular and Molecular Biology, University of Tennessee, Knoxville, TN 37996, USA

Received 25 October 1999; received in revised form 5 November 1999

Edited by Ned Mantei

Abstract As the initiator of DNA double-strand breaks during meiosis in *Saccharomyces cerevisiae*, the SPO11 protein is essential for recombination. Similarity between SPO11 and archaeobacterial TOP6A proteins points to evolutionary specialization of a DNA cleavage function for meiotic recombination. To determine whether this extends to mammals, we isolated and characterized mouse and human *SPO11* cDNAs. Mammalian *SPO11* genes were found to be expressed at high levels only in testis, wherein mouse *Spo11* transcript is restricted primarily to meiotic germ cells and is maximally expressed at mid-pachynema. Mouse *Spo11* is located near the distal end of chromosome 2, while human *SPO11* is found in the homologous position of chromosome 20q13.2–13.3, a region that is amplified in some breast cancers. Sequence homology and differential expression together support a highly conserved role for SPO11 in the enzymatic cleavage of DNA that accompanies meiotic recombination.

© 1999 Federation of European Biochemical Societies.

Key words: SPO11/TOP6A homologs; Meiosis

1. Introduction

Meiosis, producing haploid cells from diploid progenitors, is a crucial step in gametogenesis. During the first meiotic prophase, homologous chromosomes pair and recombine to form stable interhomolog connections. In most instances, these events are required for proper chromosome segregation during the reductional cell division. Recombination also contributes to genetic diversity among gametes by introducing new combinations of alleles in homologous chromosomes via strand exchange events. A considerable body of evidence suggests that meiotic recombination is achieved through a pathway that involves the generation of double-strand breaks (DSBs) in paired homologs as the initial step, a mechanism that was first suggested by Szostak and colleagues [1]. During subsequent steps, DSBs are repaired via homologous recombination involving the intact homolog.

The appearance of transient DSBs during meiotic prophase

I in budding yeast provided the first direct evidence supporting a role for DSBs in meiotic recombination [2–4]. However, the protein that is responsible for meiotic DSB formation, SPO11, was only recently isolated on the basis of its covalent linkage to the 5' ends of DSBs in yeast meiotic cells [5]. SPO11 was shown to be structurally similar to the A subunit of topoisomerase VI (TOP6A), a type II topoisomerase that has been found only in archaea [6,7]. Although this structural similarity is consistent with a role for SPO11 as the enzyme that catalyzes meiosis-specific DSBs, enzymatic activity for yeast or other eukaryotic SPO11 proteins has not been demonstrated.

Sequence homology with the *SPO11* gene product and similar function in meiosis has been noted in *rec12* of fission yeast [8], fly *MEI-W68* [9] and worm *Spo-11* [10], indicating that the underlying mechanism of meiotic recombination is likely to be highly conserved among eukaryotes. However, unlike the situation for budding yeast, chromosome synapsis is not impaired in the *mei-W68* and *spo-11* mutants.

Even though it is commonly accepted that meiotic recombination proceeds by a similar mechanism in mammals, a direct role for DSBs in mammalian meiosis has not been documented. As a first step towards defining the biochemical events that initiate meiotic recombination in mammals, we sought to identify mammalian equivalents of the *Saccharomyces cerevisiae* *SPO11* gene. We isolated both mouse and human *SPO11* homologs as cDNAs and determined chromosomal locations of the genes. The differential expression of these genes in adult testis and meiotic male germ cells is consistent with a proposed role for mammalian *SPO11* in the initial steps of meiotic recombination.

2. Materials and methods

2.1. cDNA isolation

cDNA clones were isolated from a mouse pachytene spermatocyte cDNA library [11] by hybridization of nylon filter plaque lifts to an EST clone (IMAGE no. 1446797) corresponding to the mouse homolog of the *S. cerevisiae* SPO11 protein. Hybridizations were carried out at 42°C under high stringency conditions in 1× hybridization buffer containing 50% formamide, as described previously [12]. cDNA clones were converted to phagemids from rescreened single plaques according to the manufacturer's in vivo excision protocol (Stratagene Inc., La Jolla, CA, USA). A human EST clone (IMAGE no. 2111713) was identified by sequence similarity to mouse *Spo11* cDNA. Further cDNA clones representing portions of the human *Spo11* homolog were obtained with the RACE (rapid amplification of cDNA ends) technique using Marathon-ready adult human testis cDNA as template and an Advantage PCR Core Kit (Clontech, Palo

*Corresponding author. Fax: (1)-925-422 2282.
E-mail: mthelen@llnl.gov

¹ The nucleotide sequences reported here have been submitted to the GenBank database with accession numbers AF126400 (mouse *Spo11* cDNA PA1), AF149309 (mouse *Spo11* cDNA PA7), and AF149310 (human *SPO11* cDNA).

Alto, CA, USA) in accordance with the manufacturer's instructions. 5' and 3' RACE were first carried out with the gene-specific primers hSPO11-5R (5'-ACTCTGGAAAGATAATCTGATGAT) and hSPO11-3F (5'-ATGGCTTGGTCTTCTCCCTTCTGAT), respectively, in combination with the adapter primer AP1. RACE reactions utilized the following method: 5 min at 94°C; 30 s at 94°C; 4 min at 68°C, for 30 cycles. PCR product (1 µl) was reamplified using the same PCR conditions as above, but substituting the primers hSPO11-5NR (5'-ATCAGAAGGGAGAAGACCAAGCCAT) and hSPO11-3NF (5'-ATCATCAGATTATCTTTCCAGAGT) for 5' and 3' RACE reactions, respectively, in combination with the AP2 primer. PCR products were separated on 1% agarose gels and resolved fragments were purified with a Qiaquick kit (Qiagen, Inc., Valencia, CA, USA). Isolated fragments were cloned with a TA cloning kit (Invitrogen, Carlsbad, CA, USA).

2.2. Sequence analysis

cDNA sequences were determined by the dideoxy-termination method [13] in PCR samples containing double-stranded DNA templates and dye-terminator reactions (PE Applied Biosystems, Foster City, CA, USA). Sequence ladders were detected on an ABI373 sequencer, and subsequent sequence data assembly and analysis were performed using the Autoassembler and DNA Strider software packages.

2.3. Spermatogenic cell isolation

Enriched populations of spermatogenic cells were separated by unit gravity sedimentation on 2–4% BSA gradients [14]. Cell types and whole testes were obtained as previously described [15]. Enrichment of each cell type was typically >80% as assessed by light microscopy.

2.4. RNA isolation and Northern blot analysis

RNA was prepared from spermatogenic cell fractions and whole testes [16] and Northern blot analysis was carried out as described [15]. Northern blots of poly(A)⁺ RNA from adult mouse and human tissues (Clontech) were hybridized to randomly primed radiolabeled probes [17] according to our previous protocol [15].

2.5. Interspecific backcross mapping

cDNA probes were mapped in mouse by following the segregation of restriction fragment length variants (RFLVs) in 155 progeny of a previously described *Mus musculus* × *Mus spretus* interspecific backcross (IB; [C3Hf/Mg^{512ENURg}]/+ × *M. spretus*) × C3Hf/R1; this IB has been typed for several hundred markers located on all mouse chromosomes [18]. A mouse *Spo11* cDNA (IMAGE clone 1446797) detected a 6.6 kb *HincII* RFLV in *M. spretus* DNA compared to a 9.4 kb fragment in *M. musculus* and was followed in the IB system. A cDNA probe (IMAGE no. 1510543) for mouse *Cyp24*, detecting 1.8 kb and 6.0 kb *HincII* RFLVs in *M. spretus* and *M. musculus* DNA, respectively, along with a cDNA probe (IMAGE no. 716492) for mouse *Myt1*, detecting 1.8 kb and 2.7 kb *TaqI* RFLVs in *M. spretus* and *M. musculus* DNA, served as a reference markers in this study.

2.6. Fluorescence in situ mapping

A bacterial artificial chromosome (BAC 225K9) clone containing the human *SPO11* gene was identified by hybridization screening of high density human BAC filters (Human Genome Center, Lawrence Livermore National Laboratory) with the human *SPO11* cDNA, and labeled using a DIG-High Prime kit (Roche, Indianapolis, IN, USA) according to the manufacturer's protocol. The resulting digoxigenin-labeled painting probe was hybridized to metaphase chromosome spreads and developed according to standard protocols [19]. The metaphase chromosomes were then counterstained with 4',6-diamidino-2-phenylindole (DAPI) and an antifade solution. Metaphase spreads were observed using a Zeiss Axioskop fluorescence microscope (Carl Zeiss, Inc., Thornwood, NY, USA) with the appropriate filters. Images were captured with a Vysis QUIPS Imaging Analysis System (Vysis, Downers Grove, IL, USA).

3. Results

3.1. Cloning of mouse and human *SPO11* homologs

We obtained a mouse EST clone (IMAGE no. 1446797) after a database search indicated that this cDNA encodes

segments of amino acid sequence with appreciable homology to the *S. cerevisiae* SPO11 protein. To complete the 5' coding portion of the putative mouse *Spo11* homolog, we obtained additional cDNA clones corresponding to mouse *Spo11* by screening a pachytene spermatocyte cDNA library with the mouse EST sequence as a probe. Two cDNAs (*mSpo11*-PA7 and *mSpo11*-PA1) of approximately 1.6 kb were chosen for further sequence characterization. Analysis of *mSpo11*-PA7 revealed a long open reading frame that encodes a predicted protein consisting of 358 amino acids. An additional stretch (472 nt) of untranslated sequence is present at the 3' end of the cDNA and is followed by a poly(A)⁺ tail, suggesting that the complete 3' end of the gene is represented in this cDNA. The coding region of *mSpo11*-PA1 contains an additional 13 amino acids within the N-terminal one-third of the protein, suggesting that the two cDNAs resulted from alternative mRNA splicing events.

A human EST clone was identified using the mouse *Spo11* sequence to query the dbEST database. Since only the middle portion of the *SPO11* coding region was present, we used the RACE technique [20] to obtain the cDNA ends. A composite of three overlapping cDNAs, spanning 1.8 kb, encodes a protein containing 396 amino acids. Comparison of the predicted SPO11 proteins from mouse and human revealed an overall amino acid identity of 82% between common portions. The difference in length between the proteins is due to the presence of 38 additional amino acids near the N-terminus of human SPO11. Like the situation for *mSpo11* cDNAs PA1 and PA7, this difference in structure is perhaps due to an alternative mRNA splicing event in human. The mouse and human SPO11 amino acid sequences are only 20–30% identical with related proteins from *S. cerevisiae*, *Schizosaccharomyces pombe*, *Drosophila melanogaster*, *Caenorhabditis elegans*, and also with the TOP6A protein sequences from *Sulfolobus shibatae* and other archaeobacteria (Fig. 1). Regional similarity is much higher, particularly in the five short stretches indicated.

3.2. Expression of the mouse *Spo11* and human *SPO11* genes in adult tissues

The adult tissue-specific pattern of expression of the mouse *Spo11* gene was determined by Northern blot analysis. A probe corresponding to the 3' end of mouse *Spo11* detected an approximately 1.6 kb transcript only in adult testis (Fig. 2A). Several larger transcripts (approximately 3.0 kb, 4.4 kb and 9 kb in length) were also found at low levels in testis and a few other tissues (Fig. 2A), and in whole embryos at 7–17 days post coitum (data not shown). The makeup and biological importance of these alternative transcripts are uncertain. However, it is possible that they are derived from different promoters or alternative mRNA splicing events. Northern blot analysis of human tissues also revealed high steady-state levels of *SPO11* transcript only in adult testis (Fig. 2C). These data are consistent with a conserved role for the related genes in male germ cell development.

3.3. Spermatogenic stage-specific expression of mouse *Spo11*

As a first step in defining the role of *Spo11* in spermatogenic cells, we examined the pattern of gene expression during postnatal testis development. Since the first round of spermatogenesis occurs synchronously during testis maturation, it is possible to correlate changes in gene expression with the first appearance of certain cell types in the testis [14]. Northern

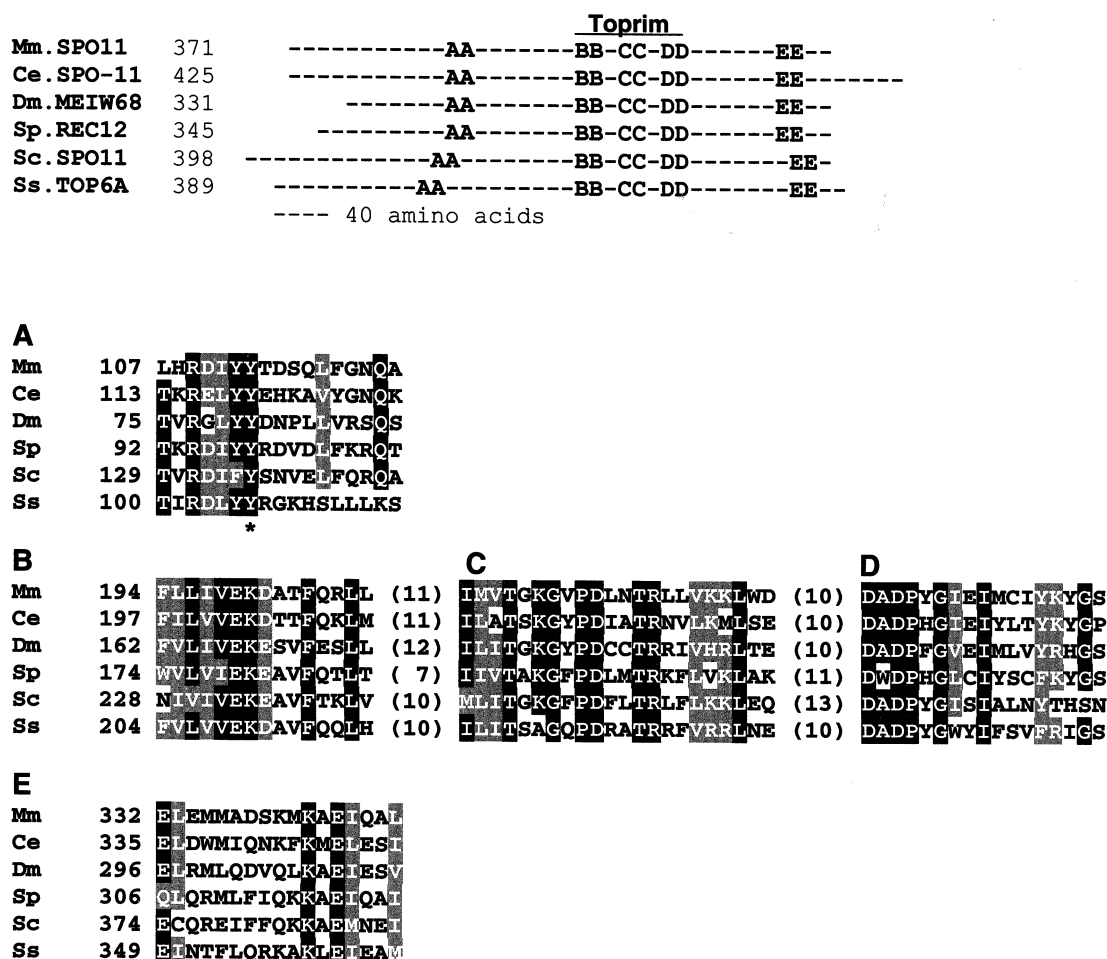


Fig. 1. Sequence alignment of SPO11/TOP6A homologs. Predicted protein sequences were compared using the BLOCKS program [26] and formatted using Boxshade 3.21 (www.ch.embnet.org). A schematic alignment of the proteins with conserved blocks A–E is shown at the top, centered around the Toprim domain [23]. The aligned amino acid sequences comprising each block are shown at the bottom. Amino acids that are identical in five out of six proteins are darkly shaded, and conservative substitutions are lightly shaded. Block A contains the conserved tyrosine residue (denoted by an asterisk) implicated in DNA strand cleavage in *S. cerevisiae* [7], and block D contains the DXD motif within the Toprim domain. Numbers indicate length of proteins (top) and positions within the proteins (bottom); numbers in parentheses indicate spacing between blocks.

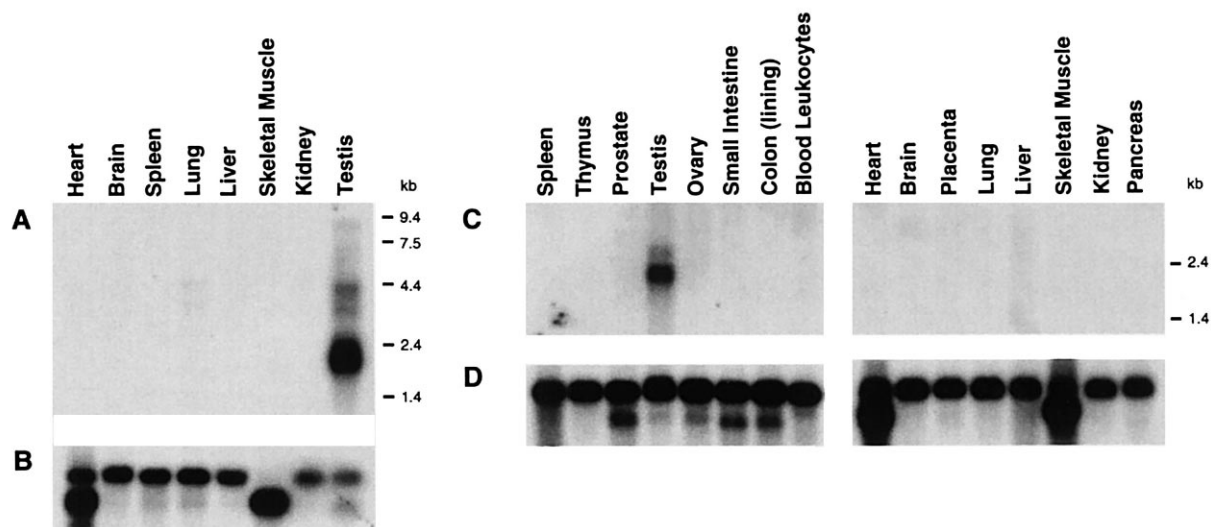


Fig. 2. Mouse *Spo11* and human *SPO11* gene expression in adult tissues. Northern blots containing 5 µg of poly(A)⁺ RNA from adult tissues were hybridized to cDNA probes for (A) mouse *Spo11* and (C) human *SPO11*. In each case the blots were rehybridized to a β-actin cDNA probe (B and D) as a control for RNA integrity and loading.

blot analysis indicated that *Spo11* is not expressed at detectable levels at day 7 of age when spermatogonia and early spermatocytes are the only germ cell types present in the developing testis (Fig. 3A). Levels of *Spo11* transcript are low at day 12, but increase by day 17. High steady-state levels of *Spo11* mRNA are maintained throughout the remainder of testis development. The dramatically elevated level of *Spo11* transcript at day 17 coincides with the peak in the proportion

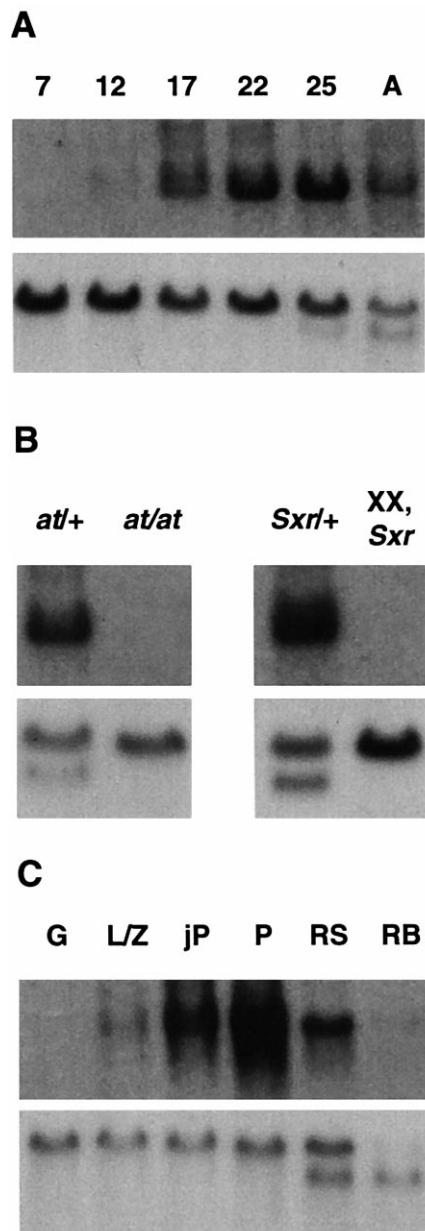


Fig. 3. Expression of *Spo11* in developing testes and male germ cells. A Northern blot containing total RNA (10 µg each) was hybridized with a *Spo11* cDNA probe. Samples were derived from: (A) testes at 7, 12, 17, 22, 27, and >60 (Adult) days of postnatal development and (B) testes from germ cell-deficient mice (*atlat*, *XX^{Sxr}*), and (C) isolated populations of type A and B spermatogonia (G), leptotene/zygotene spermatocytes (LZ), juvenile pachytene spermatocytes (jP), mid-pachytene spermatocytes (P), round spermatids (RS), and residual bodies from elongating spermatids (RB). The blot was rehybridized to a β -actin cDNA probe (lower panels), as in Fig. 2.

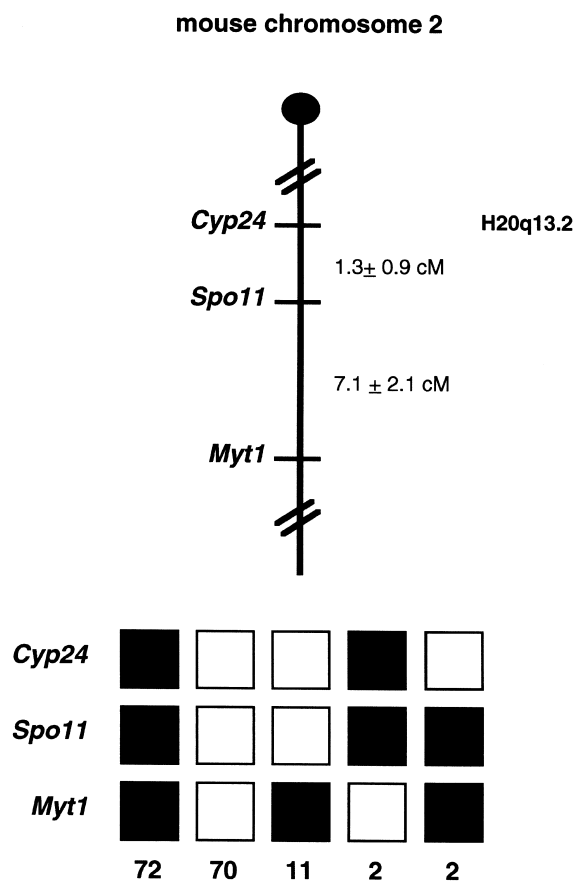


Fig. 4. Chromosomal localization of the mouse *Spo11* gene. Inter-species backcross analysis was used to determine that *Cyp24*, *Spo11*, and *Myt1* are tightly linked in distal Mmu2. Numbers to the right of the chromosome (top panel) represent the calculated distances between loci (cM) and their standard errors. Each column (bottom panel) represents a type of parental or recombinant chromosome. The numbers of backcross animals inheriting a particular type of chromosome are indicated below each column. Black boxes denote *M. spretus* alleles, while white boxes denote C3H alleles.

of pachytene spermatocytes in the testis. Notably, the 1.6 kb transcript is absent from germ cell-deficient testes of *XXSxr* and *atlat* mice (Fig. 3B), indicating that the gene is specifically expressed in testicular germ cells.

To determine more precisely the germ cell type-specific expression of *Spo11*, steady-state levels of transcript were measured in isolated spermatogenic cell populations (Fig. 3C). *Spo11* mRNA is not present in spermatogonia, but appears at low levels in leptotene/zygotene spermatocytes. A marked increase in steady-state levels of *Spo11* mRNA occurs in juvenile pachytene spermatocytes and mid-to-late pachytene spermatocytes compared to earlier germ cell types. A low level of *Spo11* mRNA is also found in round spermatids. These data support the hypothesis that *Spo11* has a specific function in mammalian meiosis.

3.4. Mapping of mouse *Spo11* and human *SPO11* genes

By correlating the distribution of *Spo11* *M. spretus* and *M. musculus* alleles with segregation patterns of other genes typed in an IB panel, we mapped *Spo11* within the distal end of mouse chromosome 2 (Mmu2), in a region lying between the *Cyp24* and *Myt1* genes (Fig. 4). This interval of Mmu2

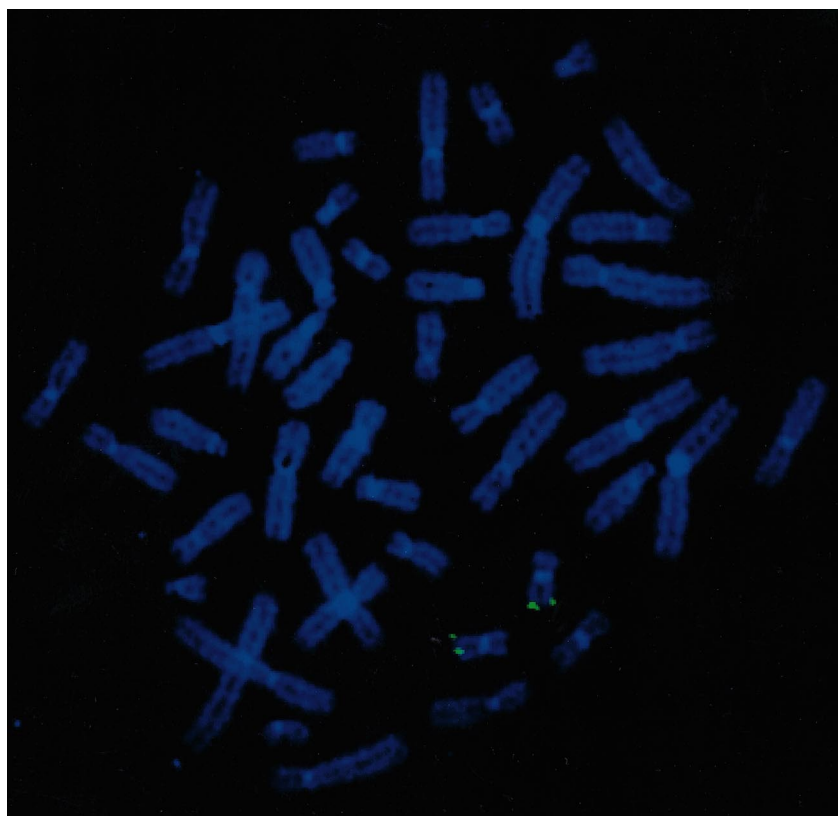


Fig. 5. Chromosomal localization of human *SPO11*. A BAC clone (225K9) containing the *SPO11* gene was hybridized to metaphase chromosome spreads that were counterstained with DAPI. *SPO11* localizes to the distal end of chromosome 20q.

has extensive homology with human chromosome 20q13.2–13.3 [21].

Since human chromosome 20q13.2–13.3 is a common site of genomic amplification and instability in certain tumor cell lines [22], the chromosomal position of human *SPO11* was determined by FISH mapping with a BAC clone containing the *SPO11* gene. Hybridization signal for *SPO11* was detected at a single site near the distal end of human chromosome 20q as determined by DAPI counterstaining of metaphase chromosome spreads (Fig. 5).

4. Discussion

The aim of this study was to begin to understand the mechanisms underlying meiotic recombination by cloning and characterizing mammalian homologs of the *S. cerevisiae* *Spo11* gene. We identified cDNAs through database searches corresponding to mouse and human homologs of yeast *SPO11*. Two mouse cDNAs, each approximately 1.6 kb, were sequenced and found to encode complete, but slightly different, open reading frames of 358 and 371 amino acids. The two predicted protein forms, which differ by 13 amino acids within the N-terminal one-third of the proteins, most likely result from alternative mRNA splicing events. Three overlapping cDNA clones corresponding to human *SPO11* were found to encode a 396 amino acid protein that is approximately 82% identical to the mouse protein. It is noteworthy that the human protein contains an additional internal stretch of 38 amino acids near the N-terminal end that is absent from both of the predicted mouse *SPO11* sequences.

However, since a single 5' end-containing cDNA clone for *SPO11* was analyzed in this study, we cannot rule out the possibility that other mRNA variants are also present in human that are colinear with the mouse coding sequences.

The two mammalian proteins share only 20–30% identity with other members of the *SPO11*/TOP6A family. However, three regions of higher homology are present within the related proteins. One region (block A in Fig. 1) contains a tyrosine that is present in all functional homologs and, when mutated in budding yeast, results in the loss of DSB activity [6]. A second region (blocks B–D) shares homology with the Toprim domain common to topoisomerase and DNA primase enzymes, which may participate in cleavage and/or rejoining reactions by coordinating Mg^{2+} ions [23]. A third region (block E) with higher than average homology is located near the C-terminal end of *SPO11*, but is not appreciably similar to any previously described motif, and its function is presently unknown.

The mouse *Spo11* gene was mapped to the distal end of mouse chromosome 2 within an interval that is homologous to a large portion of human chromosome 20q13.2–13.3. Localization of human *SPO11* confirmed this position near the telomere of chromosome 20q, which is a common site of genomic amplification in certain breast cancers and is associated with genomic instability in some tumor cell lines [22]. Although *SPO11* does not appear to be expressed at appreciable levels in somatic tissues, it will be interesting to determine whether *SPO11* expression is elevated within these cell lines, especially given the proposed function of *SPO11* as a DNA cleavage enzyme.

Among adult tissues, transcripts for the mouse and human genes were detected at high levels only in adult testis as determined by Northern blot analysis. In mouse testis, the predominant transcript size is approximately 1.6 kb, but several transcripts of larger size are also present in this tissue and several others, including lung and kidney. Whether these alternative transcript forms encode functional SPO11 protein remains to be determined experimentally.

Spo11 expression in male germ cells was investigated by determining the relative steady-state levels of transcript in normal testes, germ cell-deficient testes and isolated populations of spermatogenic cells. *Spo11* mRNA was not detected in testes lacking germ cells or in spermatogonia, but was found in leptotene/zygotene spermatocyte, pachytene spermatocyte, and round spermatid populations. Among these cell types, the relative proportion of *Spo11* transcript was found to be highest in pachytene spermatocytes. The first appearance of *Spo11* transcript in leptotene/zygotene spermatocytes may indicate that recombination-initiating DSBs are formed at this early stage of meiosis in mammals. This notion is consistent with data regarding the timing of DSB formation during early meiotic prophase in budding yeast [24]. It is also noteworthy that Mre11, which is thought to function in both non-homologous end-joining (NHEJ) and homologous recombination (HR) pathways for DSB repair, has been immunolocalized within mouse leptotene/zygotene spermatocyte nuclei, while Ku70, which appears to function only in NHEJ, is absent at this stage of meiotic prophase [25]. These observations have led to the hypothesis that the presence or absence of Ku70 within the spermatocyte nucleus may determine whether Mre11-mediated repair of DSBs proceeds by NHEJ or HR. If so, DSBs generated in early meiotic spermatocytes, likely by the DNA cleaving action of SPO11, would be preferentially repaired via HR. However, this hypothesis does not adequately account for the high steady-state level of *Spo11* transcript that is found in pachytene spermatocytes. Moreover, the finding that chromosome synapsis is not impaired in fly *mei-W68* and worm *spo-11* mutants argues that recombination-initiating DBSs may not occur until after chromosome synapsis is completed in higher eukaryotes. One possibility is that SPO11 acts in both the early and middle stages of meiotic prophase. Additional studies, including the localization of SPO11 protein in mouse spermatocyte nuclei, will provide insight into the timing and sequence of events that initiate meiotic recombination in mammals and will clarify the role of SPO11 in this crucial developmental process.

Acknowledgements: The authors are grateful to Dr. Lisa Stubbs for mouse mapping resources and advice. This research was partly funded

by a grant to M.P.T. from the U.S. Department of Energy under Contract No. W-7405-ENG-48 with Lawrence Livermore National Laboratory, and by NIH Grant HD31376 to M.A.H.

References

- [1] Szostak, J.W., Orr-Weaver, T.L., Rothstein, R.J. and Stahl, F.W. (1983) *Cell* 33, 25–35.
- [2] Game, J.C., Sitney, K.C., Cook, V.E. and Mortimer, R.K. (1989) *Genetics* 123, 695–713.
- [3] Sun, H., Treco, D., Schultes, N.P. and Szostak, J.W. (1989) *Nature* 338, 87–90.
- [4] Cao, L., Alani, E. and Kleckner, N. (1990) *Cell* 61, 1089–1101.
- [5] Keeney, S., Giroux, C.N. and Kleckner, N. (1997) *Cell* 88, 375–384.
- [6] Bergerat, A., Gadelle, D. and Forterre, P. (1994) *J. Biol. Chem.* 269, 27663–27669.
- [7] Bergerat, A., de Massy, B., Gadelle, D., Varoutas, P.C., Nicolas, A. and Forterre, P. (1997) *Nature* 386, 414–417.
- [8] Lin, Y. and Smith, G.R. (1994) *Genetics* 136, 769–779.
- [9] McKim, K.S. and Hayashi-Hagihara, A. (1998) *Genes Dev.* 12, 2932–2942.
- [10] Dernburg, A.F., McDonald, K., Moulder, G., Barstead, R., Dresser, M. and Villeneuve, A.M. (1998) *Cell* 94, 387–398.
- [11] Caldwell, K.A., Wiltshire, T. and Handel, M.A. (1996) *Mol. Reprod. Dev.* 43, 403–413.
- [12] Stubbs, L., Poustka, A., Baron, A., Lehrach, H., Lonai, P. and Duboule, D. (1990) *Genomics* 7, 422–427.
- [13] Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463–5467.
- [14] Bellve, A.R., Cavicchia, J.C., Millette, C.F., O'Brien, D.A., Bhatnagar, Y.M. and Dym, M. (1977) *J. Cell Biol.* 74, 68–85.
- [15] Shannon, M., Lamerdin, J.E., Richardson, L., McCutchen-Maloney, S.L., Hwang, M.H., Handel, M.A., Stubbs, L. and Thelen, M.P. (1999) *Genomics* (in press).
- [16] Chomczynski, P. and Sacchi, N. (1987) *Anal. Biochem.* 162, 156–159.
- [17] Feinberg, A.P. and Vogelstein, B. (1983) *Anal. Biochem.* 132, 6–13.
- [18] Doyle, J., Hoffman, S., Ucla, C., Reith, W., Mach, B. and Stubbs, L. (1996) *Genomics* 35, 227–230.
- [19] Christian, A.T., Garcia, H.E. and Tucker, J.D. (1999) *Mamm. Genome* 10, 628–631.
- [20] Frohman, M.A. (1993) *Methods Enzymol.* 218, 340–356.
- [21] Peters, J., Church, D., Zuberi, A.R. and Selley, R. (1999) *Mamm. Genome* 10, 941.
- [22] Collins, C. et al. (1998) *Proc. Natl. Acad. Sci. USA* 95, 8703–8708.
- [23] Aravind, L., Leipe, D.D. and Koonin, E.V. (1998) *Nucleic Acids Res.* 26, 4205–4213.
- [24] Padmore, R., Cao, L. and Kleckner, N. (1991) *Cell* 66, 1239–1256.
- [25] Goedecke, W., Eijpe, M., Offenberger, H.H., Aalderen, M. and Heyting, C. (1999) *Nature Genet.* 23, 194–198.
- [26] Henikoff, S., Henikoff, J.G., Alford, W.J. and Pietrovski, S. (1995) *Gene* 163, GC17–26.