Biochemical and Biophysical Research Communications **259,** 60–66 (1999)
Article ID bbrc.1999.0728, available online at http://www.idealibrary.com on IDEAL®

Cloning and Characterization of SOB1, a New Testis-Specific cDNA Encoding a Human Sperm Protein Probably Involved in Oocyte Recognition

A. Lefevre,*[,1] C. Duquenne,* M. F. Rousseau-Merck,† E. Rogier,* and C. Finaz*
**INSERM U 355, Maturation Gamétique et Fécondation, IPSC, 32 rue des Carnets, 92140, Clamart, France; and †INSERM U 509, Institut Curie, 26 rue d'Ulm, 75005 Paris, France*

Received April 19, 1999

A human sperm-oocyte binding protein, SOB1, was purified by two dimensional gel electrophoresis and sequenced. This protein was selected because it was recognized by a monoclonal antibody that inhibited the binding of human sperm to zona-free hamster oocytes. The sequences of the tryptic peptides were used to design degenerate primers. These were used to amplify a specific fragment from human testis cDNA by the polymerase chain reaction. This 1233 bp fragment was extended in 3′ and 5′ by RACE to obtain the 3 kb full length SOB1 cDNA. Sequence analysis indicated that the deduced open reading frame encodes a 853 amino acid protein, with a molecular mass of 94.7 kDa. This is a new testis-specific cDNA. It is 27, 32.8 and 34.4% homologous to three sperm proteins, HI, Fsc1 and AKAP82 respectively. A single 3kb transcript was demonstrated only in the testis by northern blot analysis. It is a single copy gene, well conserved among mammals and located on human chromosome 12 at band p13.

The union of the spermatozoon and the oocyte requires a series of specific interactions that include gamete recognition, adhesion, signaling and fusion. Mammalian sperm first bind to the glycoprotein components of the egg coat or zona pellucida. This initial interaction triggers the acrosome reaction which allows activated sperm to penetrate the zona pellucida. Sperm proteins are then believed to bind to partners on the egg plasma in a highly cell specific process that probably involves several ligand-receptor couples located on the surfaces of gametes.

Several approaches have been used to identify the mechanisms underlying this last step. Some groups have employed a genetic approach based on the hypothesis that androgen-dependent epididymal proteins enable the spermatozoa to recognize and penetrate the egg (1, 2). A number of gene products have been identified but it is difficult to assign physiological roles to them.

Others have used polyclonal and monoclonal antibodies (mAb) directed against sperm surface proteins and able to inhibit the binding of sperm to the vitelline membrane of the oocyte. Primakoff et al. used the mAb PH-30 to isolate the α/β heterodimer fertilin from guinea pig sperm (3). The β subunit has a disintegrin domain which could interact with integrin-like receptors on the egg (4). Fertilin belongs to a large family of proteins containing cystein-like and disintegrin-like domains, some of them are present only in the male reproductive tract (5). The analogous human gene was cloned in an attempt to determine the role of fertilin β in human fertilization. The deduced amino acid sequence was shown to be 56% identical to guinea pig fertilin β (6, 7). Recent encouraging results indicate that the β subunit has a contraceptive effect in male guinea pigs suggesting that fertilin is directly involved in gamete interaction (8). But similar experiments in rabbits were not conclusive (9).

Several other human proteins may be involved in sperm-oocyte interaction, as they have been characterized using mAbs that inhibited human sperm attachment in the hamster egg binding assay (10, 11, 12). The gene coding for SP-10 in human and several mammals has been cloned but the protein is finally thought to be involved in sperm-zona secondary binding (13). *Via* an elegant alternative strategy, Diekman et al used human antisperm auto-antibodies associated to unexplained infertility to screen a human testis cDNA library and identified an outer dense fiber protein (14). But the molecular basis of sperm-oocyte interaction are still far from having been resolved in human in spite of the importance of the results.

[1] To whom correspondence should be addressed. E-mail: annick.lefevre@inserm.ipsc.u-psud.fr.

Abbreviations: RT-PCR, reverse transcriptase-polymerase chain reaction; RACE, rapid amplification of cDNA ends.

0006-291X/99 $30.00

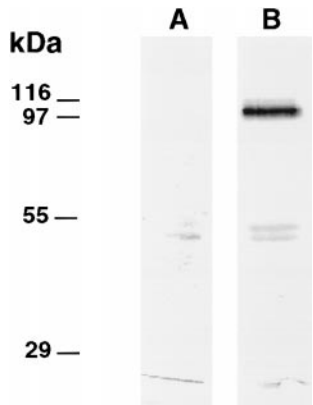

FIG. 1. Western blots of human spermatozoa. Protein extracts (10 μg per lane) were separated by SDS-PAGE and incubated with SOB1 mAb (B) or a negative hybridoma supernatant (control) (A). The major protein identified (100 kDa) was SOB1. Molecular weight standards are indicated on the left.

We have cloned and sequenced SOB1 a human sperm protein. It was purified by two-dimensional gel electrophoresis and then microsequenced. This protein was selected because monoclonal antibody (mAb) anti-SOB1 significantly inhibits human sperm binding to zona-free hamster oocytes. The 3000 bp cDNA was obtained by RT-PCR using degenerate oligonucleotides and RACE-PCR. Its open reading frame codes for a 853 amino acid protein. SOB1 is a new gene whose expression is testis-specific. It is located on human chromosome 12p13 and conserved among mammals.

MATERIALS AND METHODS

Cloning of Human SOB1

Degenerate primers based on SOB1 peptides 14, 17c, 22b and 23 (Table I) were used for PCR with human testis cDNA as template. Lymphocyte, liver and epididymis cDNAs were used as controls. The single 1233 bp testis-specific fragment obtained was amplified with primers 22b and 23 (Table II). It was cloned into vector pcR-II (TA cloning kit, Invitrogen) and sequenced on both DNA strands (Genome Express, Paris, France). 5′ and 3′ RACE was performed using the Marathon cDNA Amplification Kit from Clontech, according to the instructions of the manufacturer. The anchor primers provided and the specific primers designed from the sequenced 1233 bp fragment (Table II) were used for PCR with the Advantage Klen Taq polymerase from Clontech. The resulting products were cloned into pcR-II vector and sequenced on both DNA strands.

Chromosomal Location of the Human SOB1 Gene

In situ hybridization. High resolution chromosome preparations were obtained from phytohemaglutinin-stimulated blood cell cultures of two healthy men after methothrexate synchronization. A 2.2 kb genomic probe was amplified using primers N22 and 9Q (Table II) and the Advantage Genomic PCR Kit from Clontech. It was labeled by nick-translation with biotinylated 11dUTP (Boehringer Mannheim, Indianapolis, IN) and mixed with about a 50 fold excess of human Cot-1DNA. This mixture was used for *in situ* hybridization with normal human metaphase chromosomes, as previously described (15). The slides were counterstained with DAPI and examined under a fluorescence Leica DMBR microscope. Images were acquired with an NU 200 CCD camera (Photometrics, Tucson AZ) and analysed with Smart Capture software (Digital Scientific, Cambridge).

Human–rodent somatic cell hybrids. The DNA from a NIGMS panel of DNA isolated from 24 human-rodent somatic cell hybrids, each retaining one intact human chromosome (Coriell Institute for Medical Research, NJ), was digested with EcoRI. These samples were hybridized with the 1233 bp probe labeled with α-^{32}P dCTP by random priming in a Southern blot.

Southern Blots

Genomic DNA was isolated by standard methods (16), digested with restriction endonucleases, separated on 0.8% (w/v) agarose gels in TBE, blotted onto Hybond-N (Amersham) and hybridized with the 1233 bp probe labelled with α-^{32}P dCTP (Amersham). Hybridization was carried out overnight at 65°C in 10 X SSPE, 10 X Denhardt's solution, 0.1% SDS, 100 μg/ml denatured salmon sperm DNA. Membranes were washed with SSC solutions of increasing stringency up to 0.5 X SSC, 0.1% SDS at 65°C and exposed to X-ray film for 36 h.

Northern Blots

Total RNA (10 μg) was isolated from various human tissues and electrophoresed on a formaldehyde-1.2% agarose gel. The separated mRNA was transferred to a nylon membrane and hybridized overnight at 65°C with the 1233 bp probe labelled with α-^{32}P dCTP. The membrane was then washed with SSC solutions of increasing stringency (up to 0.1 X SSC, 0.1% SDS at 65°C) and exposed to X-ray film for 48 h.

Western Blots

Proteins from human ejaculated sperm (10 μg) were separated under reducing conditions by SDS-polyacrylamide gel electrophoresis on a 9% (w/v) gel (17) and electrophoretically transferred to nitrocellulose membranes. The membranes were saturated with 10% skim milk in PBS for 2 h at room temperature, incubated for 2 h at room temperature with SOB1 mAb supernatant (diluted 1:2), washed again with PBS-0.05% Tween incubated for 1 h at room temperature with peroxidase-conjugate goat antimouse IgG (1:1000) (SIGMA) and washed with PBS-Tween. Bound peroxidase activity was detected with 1.5 mM 3.3-diaminobenzidine in 0.05 M phosphate

TABLE I

Sequences of SOB1 Tryptic Peptides

Peptide	Sequence
14	G F S V D Y Y N T
17a	L C V I I A K
17b	S F F Y K
17c	G V Y Q S L Y M G N E P T P T
22a	L L Q L S A A A V D
22b	S T A E F Q D V R F
23	E G L T L W H K
26	L H E R P L A S S P X R L Y E
39	G S V G E V L Q S V L R Y E

Note. Trypsin digestion of SOB1 was carried out followed by reverse phase-high performance liquid chromatography of the resulting peptides. Amino acid sequencing of the selected peptides was done by automated Edman degradation by Dr. J. D'Alayer (Institut Pasteur, Paris).

FIG. 2. Nucleotide sequence and deduced amino acid sequence of human SOB1. The nine amino acid sequences determined from tryptic digest fragments of SOB1 are underlined. The regions incorporated into the sense and antisense primers used to produce the 1233 bp probe and the 2.2 kb genomic probe are underlined. Upper numbers refer to nucleotide positions. The putative polyadenylation signal is underlined. This sequence was deposited in GenBank under accession number U85715.

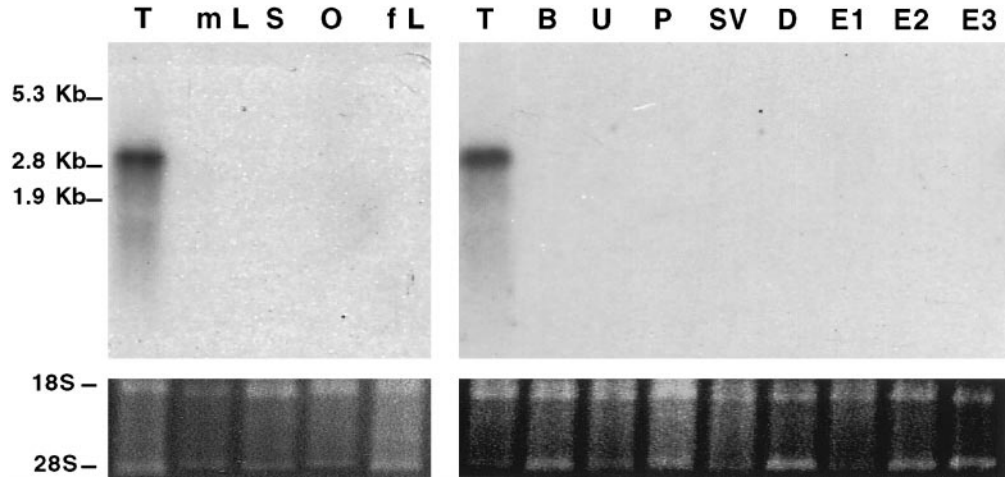

FIG. 3. Tissue distribution of SOB1 mRNA. Northern blot of total RNA (10 μg per lane) from human testis (T), male and female liver (mL and fL), spleen (S), ovary (O), bladder (B), ureter (U), prostate (P), seminal vesicle (SV), deferent duct (D), head (E1), corpus (E2) and cauda (E3) epididymis hybridized with ^{32}P labeled 1233 bp cDNA probe. A single 3 kb band was found only in the testis sample. Equivalent amounts of RNA were assessed by the abundance of 18 and 28s RNAs.

buffer pH 7.4, 0.03% H_2O_2. Controls were incubated with the second antibody only.

RESULTS AND DISCUSSION

Characterization of SOB1

Analysis of the antigens recognized by antibodies to sperm proteins has been successfully used to identify sperm molecules involved in oocyte recognition (3, 11, 12, 18–22). The anti-SOB1 mAb was selected in a preliminary screening because of its ability to inhibit the binding of human sperm to zona free hamster oocytes (62% inhibition). The corresponding antigen, SOB1, had an apparent molecular weight of 100 kDa when total proteins from human ejaculated sperm were subjected to SDS-PAGE followed by Western blotting (Fig. 1). SOB1 was purified to homogeneity by two dimensional gel electrophoresis, and the purified protein was digested with trypsin. The sequences of 9

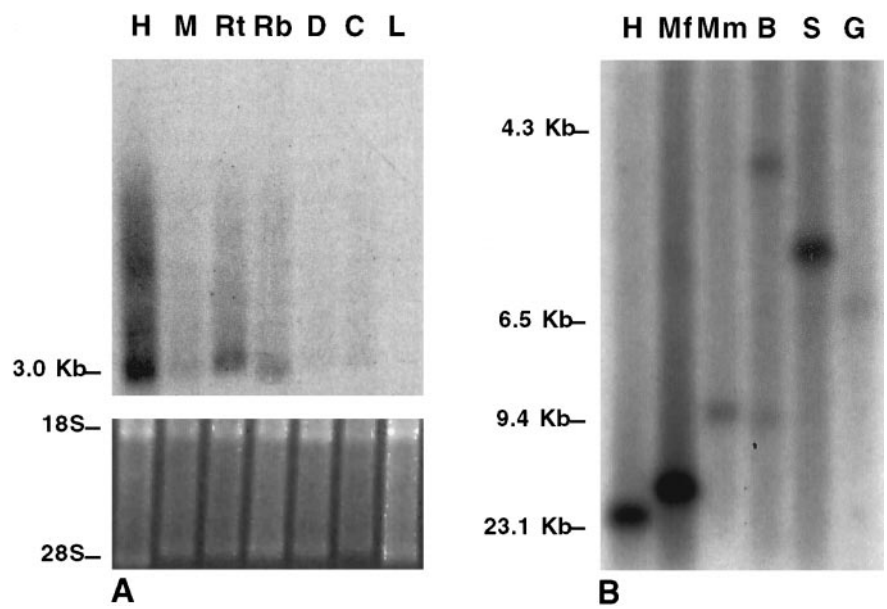

FIG. 4. Conservation of SOB1 among mammals. (A) Northern blot of total testicular mRNA from human, mouse, rat, rabbit, dog, cat, and lemur hybridized with ^{32}P labeled 1233 bp cDNA probe. Equivalent amounts of RNA were assessed by the abundance of 18 and 28s RNAs. (B) Southern blot hybridization of genomic DNA from different mammals. 10 μg genomic DNA from human (H), macaque (Mf), mouse (Mm), boar (B), stallion (S), and goat (G) were digested with EcoRI, electrophoresed, blotted onto Hybond N^+ and hybridized with SOB1 1233 bp cDNA probe.

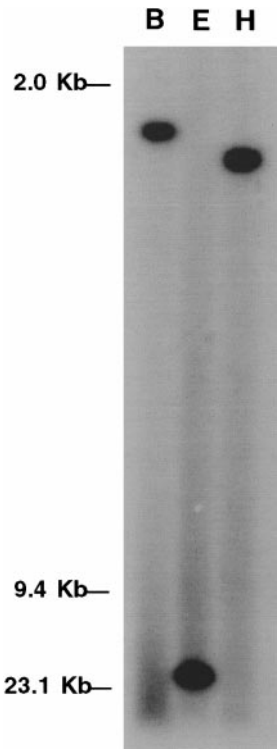

FIG. 5. Southern blot hybridization of human genomic DNA. 10 μg of genomic DNA were digested with Bam HI (B), EcoRI (E), or Hind III (H), electrophoresed, blotted onto Hybond N^+ and hybridized with SOB1 1233 bp cDNA probe.

tryptic peptides were determined (Dr. J. D'Alayer, Institut Pasteur, Paris) (Table I). They showed no similarity with any previously reported protein, and were therefore used to design suitable primers for cloning SOB1.

Cloning of SOB1 and Determination of Nucleotide Sequence

When degenerate sense and antisense primers based on SOB1 peptides 14, 17c, 22b and 23 (Table I) were used for PCR treatment of human testis cDNA, a number of components were obtained. Human liver, lymphocyte and epididymis cDNA were used as control tissues. There was only one testis-specific product, obtained with the primer pair representing peptide 22a sense and peptide 23 antisense. The sequence of this 1233 bp product indicated that it was a SOB1 related product, as it contained the sequences encoding SOB1 peptides 22a, 14, 17c, 17b and 23a. This 1233 bp fragment was subsequently used to design the oligonucleotide primers used in RACE to isolate the 5′ and 3′ sequences of the cDNA. The full length cDNA was 3000 bp long (Fig. 1) and had an untranslated 5′ region (1-202 bp) which contained one stop codon preceding the first ATG of the open reading frame. This codon had an A at position −3 and a G at +4, consistent with

TABLE II

Primers Utilized for RT-PCR, 3′ and 5′ RACE and the Generation of a Genomic Probe

Degenerate primers based on peptide sequence[a]	
22b S	ACi GCi GAR TTY CAR GAY GT
23 AS	TTR TGC CAi AGN GTi AGN CCY TC
Primers for 3′ RACE PCR	
S68 S	CC CTC ATC TCC ATG AAA GGA ATG GGT G
S266 S nested	CC AAA TCA GAG GAG GAG ACT TGT GAG
Primers for 5′ RACE PCR	
T1 AS	TCT TTC TGG AGT GCC CTT GGT GGT GG
T2 AS nested	ACC TTT GTG TGG GTC TCC TGA GTT GG
Primers for generating the 2.2 kb genomic probe	
N22 S	ACA GCA GAG TTC CAA GAT GTT CGG
9Q AS	CTG AAC TGT GTC CGT TAG GTT GTG

Note. S and AS designed sense and antisense primers, respectively.

[a] Degenerate sense and antisense primers based on sequences obtained for peptides 22b and 23b (Table I). The abbreviations used for nucleotides are: i for inosine; R for A or G; Y for T or C; N for A, C, G or T.

a Kosak consensus sequence for translation (23). The 3′ end, had an untranslated region (236 bp) after the stop codon at 2761-2764 bp; it contained a polyadenylation consensus (24).

The open reading frame encoded a protein of 853 amino acid with a deduced molecular mass of 94.793 and an estimated isoelectric point of 5.7. The difference between the apparent and calculated molecular weights may be due to post-translational modifications such as glycosylation or phosphorylation. The nine tryptic peptides (Table I) are contained in the open reading frame and, as anticipated, each one is preceded by a Lys residue (Fig. 2). The deduced SOB1 protein was predicted to have 4 N-glycosylation sites, 10 protein kinase C phosphorylation sites, 15 casein kinase II phosphorylation sites, one tyrosine kinase phosphorylation site, 8 N-myristoylation sites and 1 amidation site.

A BLAST search of GenBank revealed that the nucleotide sequence of SOB1 was 99% identical to two

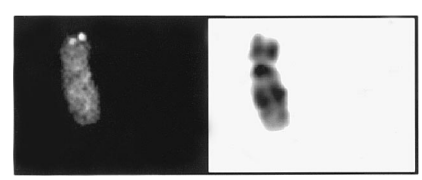

FIG. 6. Chromosomal location of the human SOB1 gene by *in situ* hybridization. FITC signals were obtained with the 2.2 kb biotinylated genomic probe on both chromatids of chromosome 12 at band 12p13.

newly deposited sequences (accession numbers AF087003 and AF093408). All three are likely to be for the same gene. No data are available on these two sequences. SOB1 is also 32.9 and 33.4% identical to two proteins of the fibrous sheath, AKAP82 and Fsc1 respectively (25, 26). AKAP 82 is a phosphorylated protein synthesized as a 97 kDa precursor, pro-AKAP82, in mouse and man (25, 27). The pro-mAKAP82 and Fsc1 cDNAs are essentially identical, except in their 5′-untranslated region. The two transcripts arise from alternative splicing of a single gene. AKAP82 is believed to be an anchoring protein for the subcellular localization of protein kinase A in the sperm fibrous sheath (28). In addition, SOB1 is 27% identical to another sperm-specific protein, HI, having a regional homology to the domain of PKA anchoring protein (29). These two proteins, AKAP82 and HI, are believed to be involved in sperm motility. However, the similarity of their sequence to that of SOB1 sheds no light on the role of SOB1, since it has no A-kinase-anchoring-like domain.

Expression of the SOB1 Gene and Conservation through Evolution

Northern blots of various human tissues were hybridized with the 1233 bp initially cloned fragment to examine tissue specific distribution of SOB1 mRNA. As shown in Fig. 3, SOB1 demonstrated a unique transcript of 3 kb in the testis only (particularly, SOB1 was not expressed in the other tissues of the genital tract), as might be expected if the gene encodes a protein important in sperm differentiation.

SOB1 appears to be conserved among mammalian species as it is also present in several rodents, farm animals and the macaque monkey (Fig. 4). But, the SOB1 gene was not found in more distant monkey species such as the lemur.

Southern hybridization also showed that SOB1 is a single copy gene (Fig. 5).

Chromosome Mapping

Chromosome preparations from metaphase cells were examined by *in situ* hybridization with a biotinylated human SOB1 genomic probe to determine chromosomal location of SOB1 gene (see Table II). A total of 30 metaphase cells were examined. A signal on chromosome band 12p13 (Fig. 6) was detected 50 times in the 60 chromosomes 12 examined. No other recurrent signal was found on any other chromosome. These results were confirmed by Southern blot analysis of Eco RI digested DNA from the human-rodent somatic cell hybrids since the specific human DNA fragment was present only in the NA10868A hybrid containing human chromosome 12 and absent from the others (data not shown).

The chromosomal locations obtained by *in situ* hybridization and human-rodent cell hybrids thus establish that the human SOB1 gene is located on chromosome 12 at band p13. This was further confirmed by a search of the GenBank data base, which found SOB1 located in its entirety on BAC (Bacterial Artificial Chromosome) genomic clone RPCI11-500M8 (accession no AC005832), which is from human chromosome 12p13.3. By contrast, it is to note that AKAP82 is located on Xp11.2. Incidentally, FISH analysis demonstrated amplification of this 12p13 region in testicular germ-cell tumors (30).

Thus, we have identified the full length cDNA of the SOB1 gene. SOB1 appears to be a new gene expressed only in the testis. Further studies are required to elucidate the function of this gene. Binding studies with recombinant protein will be used to assess directly its postulated role in the recognition of the oocyte membrane.

ACKNOWLEDGMENTS

We thank Dr. A. Aurias for helpful discussions. This work was supported in part by a grant from Ministère de la Recherche et de la Technologie (ACCSV 4).

REFERENCES

1. Frayne, J., Jury, J., Barker, H., Perry, A., Jones, R., and Hall, L. (1998) *Mol. Hum. Reprod.* **4,** 429–437.
2. Kirchoff, C. (1998) *Rev. Reprod.* **3,** 86–95.
3. Blobel, C. P., Wolfsberg, T. G., Turck, C. W., Myles, D. G., Primakoff, P., and White, J. M. (1992) *Nature* **356,** 248–252.
4. Bronson, R. A., and Fusi, M. F. (1996) *Mol. Hum. Reprod.* **2,** 153–168.
5. Frayne, J., Jury, J. A., Barker, H. L., and Hall, L. (1997) *Mol. Reprod. Dev.* **48,** 159–167.
6. Gupta, S. K., Alves, K., O'Neil Palladino, L., Mark, G. E., and Hollis, G. F. (1996) *Biochem. Biophys. Res. Com.* **224,** 318–326.
7. Vidaeus, C. M., Von Kap-Herr, C., Golden, W. L., Eddy, R. L., Shows, T. B., and Herr, J. C. (1997) *Mol. Reprod. Dev.* **46,** 363–369.
8. Ramarao, C. S., Myles, D. G., White, J. M., and Primakoff, P. (1996) *Mol. Reprod. Dev.* **43,** 70–75.
9. Hardy, C. M., Clarke, H. G., Nixon, B., Grigg, J. A., Hinds, L. A., and Holland, M. K. (1997) *Biol. Reprod.* **57,** 879–886.
10. Naz, R. K., Morte, C., Garcia-Framis, V., Kaplan, P., and Martinez, P. (1993) *Biol. Reprod.* **49,** 1236–1244.
11. Wright, R. M., Suri, A. K., Kornreich, B., Flickinger, C. J., and Herr, J. C. (1993) *Biol. Reprod.* **49,** 316–325.
12. Lefevre, A., Martin Ruiz, C., Chokomian, S., Duquenne, C., and Finaz, C. (1997) *Mol. Hum. Reprod.* **3,** 507–516.
13. Coonrod, S. A., Herr, J. C., and Westhusin, M. E. (1996) *J. Reprod. Fert.* **107,** 287–297.
14. Diekman, A. B., Olson, G., and Goldberg, E. (1998) *Mol. Reprod. Dev.* **50,** 284–293.
15. Desmaze, C., Zucman, J., Delattre, O., Melot, T., Thomas, G., and Aurias, A. (1994) *Cancer. Gent. Cytogenet.* **74,** 13–18.
16. Sambrook, J., Fitsch, E. F., and Maniatis, T. (1989) Molecular Cloning. A Laboratory Manual, Cold Spring Harbor Laboratory Press.
17. Laemmli, U. K. (1970) *Nature* **15,** 680–685.
18. Lathrop, W. F., Carmichael, E. P., Myles, D. G., and Primakoff, P. (1990) *J. Cell. Biol.* **11,** 2939–2949.

19. Naz, R. K., and Zhu, X. (1997) *Mol. Reprod. Dev.* **48,** 449–457.
20. Zhu, X., and Naz, R. K. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 4704–4709.
21. Martin Ruiz, C., Duquenne, C., Treton, D., Lefèvre, A., and Finaz, C. (1998) *Mol. Reprod. Dev.* **49,** 286–297.
22. Zhu, X., and Naz, R. K. (1998) *Mol. Reprod. Dev.* **51,** 176–183.
23. Kozak, M. (1987) *Nucleic Acids Res.* **15,** 8125–8132.
24. Juretia, N., and Theus, M. (1991) *FEBS Letters* **290,** 4–8.
25. Carrera, A., Gerton, G., and Moss, S. (1994) *Dev. Biol.* **165,** 272–284.
26. Fulcher, K., Mori, C., Welch, J., O'Brien, D., Klapper, D., and Edy, E. (1995) *Biol. Reprod.* **52,** 41–49.
27. Turner, R., Johnson, L., Haig-Ladewig, L., Gerton, G., and Moss, S. (1998) *J. Biol. Chem.* **273,** 32135–32141.
28. Visconti, P., Johnson, L., Oyaski, M., Fornes, M., Moss, S., Gerton, G., and Kopf, G. (1997) *Dev. Biol.* **192,** 351–363.
29. Mohapatra, B., Verma, S., Shankar, S., and Suri, A. (1998) *Biochem. Biophys. Res. Com.* **244,** 540–545.
30. Henegarin, O., Vance, G. H., Heiber, D., Pera, M., and Heerema, N. A. (1998) *J. Mol. Med.* **76,** 648–655.