GENOMICS 25, 354–359 (1995)

POM–ZP3, a Bipartite Transcript Derived from Human *ZP3* and a *POM121* Homologue

SIMON KIPERSZTOK,*,†,[1] GAIL A. OSAWA,* LI-FANG LIANG,* WILLIAM S. MODI,‡ AND JURRIEN DEAN*,[2]

*Laboratory of Cellular and Developmental Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892; †Department of Obstetrics and Gynecology, George Washington University, Washington, DC; and ‡Biological Carcinogenesis and Development Program, Program Resources Inc./Dyncorp, NCI-FCRDC, Frederick, Maryland 21702

Received June 10, 1994; revised October 21, 1994

Human POM–ZP3 is a novel bipartite RNA transcript that is derived from a gene homologous to rat POM121 (a nuclear pore membrane protein) and ZP3 (a sperm receptor ligand in the zona pellucida). The 5′ region is 77% identical to the 5′ end of the coding region of rat *POM121* and appears to represent a partial duplication of a gene encoding a human homologue of this rodent gene. The 3′ end of the POM–ZP3 transcript is 99% identical to ZP3 and appears to have arisen from a duplication of the last four exons (exons 5–8) of *ZP3*. Using Northern blots and RT-PCR, POM–ZP3 transcripts were detected in human ovaries, testes, spleen, thymus, lymphocytes, prostate, and intestines. The longest open reading frame encodes a conceptual protein of 210 amino acids, the first 76 of which are 83% identical to residues 241–315 of rat POM121. The next 125 amino acids are 98% identical to residues 239–363 of the 424-amino-acid human ZP3 protein. By fluorescence *in situ* hybridization, genomic fragments of *ZP3* and a human homologue of *POM121* were localized to chromosome 7q11.23. Taken together, these data suggest that partial duplications of human *ZP3* and a *POM121*-like gene have resulted in a fusion transcript, POM–ZP3, that is expressed in multiple human tissues.

INTRODUCTION

The human zona pellucida is an extracellular matrix that surrounds growing oocytes and ovulated eggs. It mediates the relatively species-specific events at fertilization and is subsequently modified to provide a potent postfertilization block to polyspermy. The human zona is composed of three glycoproteins, ZP1, ZP2, and ZP3, with molecular weights of 90–110, 64–76, and 57–73 kDa, respectively (Shabanowitz and O'Rand, 1988). The genes encoding the three human proteins have been characterized. *ZP1* contains 11 exons (79–190 bp), spanning 11 kb, and encodes a 540-amino-acid protein; *ZP2* (19 exons, 45 to 190 bp, spanning 13 kb) encodes a 745-amino-acid protein; and *ZP3* (8 exons, 92 to 330 bp, spanning 18.3 kb) encodes a 424-amino-acid protein (Chamberlin and Dean, 1990; Liang and Dean, 1993; Harris *et al.*, 1994).

The structures of homologous zona proteins have been deduced from cDNA and/or genomic sequences from nine different mammalian species (Ringuette *et al.*, 1988; Chamberlin and Dean, 1989, 1990; Liang *et al.*, 1990; Kinloch *et al.*, 1990; Schwoebel *et al.*, 1991; Liang and Dean, 1993; Koothan-Thillai *et al.*, 1993; Lee *et al.*, 1993; Yurewicz *et al.*, 1993; Harris *et al.*, 1994). All three zona proteins have amino-terminal signal peptides that direct secretion and very hydrophobic domains (23–26 amino acids) near their carboxyl termini. The amino acid sequences are well conserved among mammals within each class of zona proteins. The ZP3 (63–91%) proteins are more identical than the ZP2 (54–70%) and ZP1 (57–68%) proteins. The proteins within each class are approximately the same length (ZP1, 570 amino acids; ZP2, 715 amino acids; ZP3, 425 amino acids). The preservation of the relative positions of the cysteine residues in ZP1 (19 residues), ZP2 (20 residues), and ZP3 (13 residues) in conjunction with overall sequence identity strongly suggests that the three-dimensional structure of the zona proteins within the zona pellucida matrix is similar among mammals.

A recent report (van Duin *et al.*, 1992) described the isolation of a genomic DNA fragment and cDNA clones that contained four exons (or cDNA encoded thereof) virtually identical to exons 5–8 of human *ZP3* (Chamberlin and Dean, 1990). However, a frameshift mutation, caused by an additional guanosine residue in the last exon (5G versus 4G), predicts a truncated protein product that would lack the hydrophobic domain characteristic of zona proteins. Although, the authors attributed their findings to a second, intact *ZP3* locus on chromosome 7, they also raised the possibility of a partial 3′ duplication of the *ZP3* gene.

To further define the gene product of this locus, we

[1] Present address: Department of Obstetrics and Gynecology, School of Medicine, University of Florida, Gainesville, FL 3260-0294.

[2] To whom correspondence should be addressed at Laboratory of Cellular and Developmental Biology, NIDDK, Building 6, Room B1-26, National Institutes of Health, 6 Center Dr. MSC 2715, Bethesda, MD 20892-2715. Telephone: (301) 496-2738.

0888-7543/95 $6.00

have isolated a near full-length cDNA and have confirmed the presence of the frameshift mutation in a novel gene product, designated POM–ZP3. The POM–ZP3 transcript appears to have arisen from a partial duplication of human *ZP3* (exons 5–8) coupled with additional 5′ exons that encode a protein that is homologous to POM121, a rat nuclear pore membrane protein (Hallberg *et al.,* 1993). Like POM121, POM–ZP3 transcripts are detected in at least eight different human tissues, whereas the zona transcript is restricted to the ovary.

MATERIALS AND METHODS

Characterization of human POM–ZP3 cDNA and human POM121 genomic clones. A human ovarian cDNA library (Liang and Dean, 1993) was screened with a ^{32}P-labeled full-length cDNA of human ZP3 (Chamberlin and Dean, 1990). Positive clones of interest were purified and sequenced as previously described (Liang and Dean, 1993). Oligonucleotide primers were used in polymerase chain reactions (PCR) to generate a 5′ POM–ZP3 unique probe that was labeled with ^{32}P and used in additional screens of the same cDNA library as well as two additional libraries obtained from Clonetech. Restriction enzyme digests, Southern blots, subcloning of DNA fragments, and dideoxy sequencing were performed by standard protocols as previously described (Liang and Dean, 1993). Analysis of DNA and protein sequence was carried out on PCGene (Intelligenetics), BLAST (Altschul *et al.,* 1990), and the sequence analysis software package of the Genetics Computer Group (Devereux *et al.,* 1984).

A lambda human genomic library (Lawn *et al.,* 1978) was screened with a 5′ POM–ZP3 cDNA (180 bp) to isolate a probe for fluorescence *in situ* hybridization. A 4.5-kb *Eco*RI fragment containing two exons (215 and 163 bp) with near identity to the 5′ region (543–920) of POM–ZP3 cDNA was subcloned into Bluescript.

Polymerase chain reactions. Plaque-purified phage were heated to 95°C for 5 min and used in a PCR with oligonucleotides common to ZP3 and POM–ZP3 (5′ GGTTCCCAGTGGAAGGC 3′; 5′ AAGCAGACACAGGGTGG 3′) that flank a frameshift mutation in the last exon (van Duin *et al.,* 1992). The 342-bp product was alkali denatured and dot-blotted in duplicate to nylon membranes (GeneScreen Plus). The membranes were hybridized (37°C, 1× SSC, 5% formamide, 1% SDS, 1 m*M* EDTA) for 1 h with ^{32}P-labeled oligonucleotides, 5′ GTCACCGTGGGGGCCACTGA 3′ or 5′ GTCACCGTGGGGCCACTGAT 3′. After rinsing at room temperature, the blots were washed twice (1× SSC, 45% formamide, 1% SDS, 1 m*M* EDTA) for 10 min at 37°C and autoradiographed.

Oligonucleotide primers were used to amplify the human sequences homologous to rat POM121-specific (5′ CCCAGTGACTGTGAGGATCG 3′ and 5′ CTTCTTCTTCTCTTTGAGGGC 3′) and human *ZP3*-specific (5′ TGACTTTCTCTCTGCGTCTGATGG 3′ and 5′ GAGGCATTCTGGTCTGGTGTCG 3′) sequences from 1 μg of genomic DNA isolated from somatic cell hybrids (human x rodent: mapping panel 2, Coriell Institute for Medical Research, Camden, NJ). The identity of the POM121-like (479 bp) and ZP3 (623 bp, exons 3–4) PCR products was confirmed by Southern blots and hybridization with POM- or ZP3-specific probes.

Northern blots and RT-PCR. Nylon membranes containing human poly(A)$^+$ RNA (2 μg) were obtained from Clonetech and hybridized with ^{32}P-labeled probes specific to the 5′ end of POM–ZP3 (691–907 nt), the 5′ end of human ZP3 (13–719 nt), or the 3′ end of ZP3 (724–1277 nt) that is 99% identical to residues 921–1475 of POM–ZP3. After washing under high stringency conditions (Liang *et al.,* 1990), the membranes were exposed to X-ray film.

RT-PCR was performed with total RNA (1 μg) isolated from human ovarian tissue with RNAzol (Cinna/Biotecx) using a 5′ RACE System kit (GIBCO–BRL). An oligonucleotide (5′ GGAAGACATCCACTGTG 3′) complementary to both POM–ZP3 and ZP3 mRNA was used to make single-stranded cDNA with SuperScript reverse transcriptase. Subsequent PCR with POM–ZP3-specific (5′ AGGATGGTGTGTAGCCC 3′ and 5′ GGTCGGCGGCCGCCACTGTCC 3′) or ZP3-specific (5′ GGAGCTGAGCTATAGGG 3′ and 5′ GGAAGTCCACGATGGTG 3′) primers amplified 196- and 719-bp DNA fragments, respectively. The specificity of the fragments was confirmed by Southern blots with POM–ZP3- or ZP3-specific ^{32}P-labeled probes.

Fluorescence in situ hybridization. Bluescript plasmids with the 4.5 kb of human genomic DNA homologous to *POM121* or a 20-kb fragment containing *ZP3* sequences (pZTD-19A, intron 2 to the 3′ end) (Chamberlin and Dean, 1990) were labeled and fluorescence *in situ* hybridization was carried out as previously reported (Tory *et al.,* 1992).

RESULTS

Full-Length cDNA of POM–ZP3

A previously constructed human ovarian cDNA library (Liang and Dean, 1993) was screened with a ^{32}P-labeled full-length human ZP3 cDNA. DNA from plaque-purified clones was used as a substrate for a polymerase chain reaction with oligonucleotide primers that bracketed the previously reported frameshift mutation (created by the insertion of a guanosine residue in the last exon) (van Duin *et al.,* 1992). The PCR products were hybridized with ^{32}P-labeled oligonucleotides containing either 5Gs (5′ GTCACCGTGGGGGCCACTGA 3′) or 4Gs (5′ GTCACCGTGGGGCCACTGAT 3′). Six clones with (5Gs) and nine clones without (4Gs) the frameshift mutation were identified and the presence of 5Gs was confirmed by DNA sequencing. Of the six mutant clones, two had an additional 40 and 194 bp 5′ of the sequences that have near identity with those encoded by exons 5–8 of human *ZP3.*

The additional 5′ sequences were used as a probe to rescreen the same library as well as two commercially available cDNA human ovarian libraries. Four additional cDNA clones were isolated that had both the novel 5′ sequence and the 3′ frameshift mutation. The DNA sequence of the longest cDNA insert (1.5 kb) was determined. Despite additional screening, no clones with inserts extending more 5′ were isolated.

Comparison of POM–ZP3, ZP3, and POM121-like Transcripts and Proteins

The sequence of the POM–ZP3 transcript was deduced from the 1496-bp cDNA[3] (Fig. 1). Assuming a 150- to 200-nt poly(A) tail, it is likely that the reported 1496 bp represent a near full-length cDNA clone of the 1.6-kb POM–ZP3 mRNA (see below). The 576 nt at its 3′ end are 99% identical to the sequences encoded by exons 5 to 8 of the human *ZP3* gene (Fig. 2). Five nucleotides that differ between the two genes are indicated in Fig. 1. All five differences were observed in four independent isolates from a previously described ovarian cDNA library (Liang and Dean, 1993). Only the last three differences (including the frameshift substitution) were observed in clones isolated from a commercial library, in agreement with an earlier report using the same source (Koothan-Thillai *et al.,* 1993). The 922-

[3] Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. U10099.

```
NA
1    GCGCCGGCGGCTGCGGCGGCTGGAGCAGGCGAGCGGCGGCGGCCGATAGCGAGTGTCAGG
61   GCCGGCCGGGGGCGGCGCTTCTCGGCCTGTCGCTGGTCGGCCTCCTACTGTACCTCGTGCC
121  TGCTGCGGCTGCGCTGGCCTGGCTGGCCGTGGGGACTACCGCGGCCTGGTGGGGACTGAG
181  CCGCGAGCCCCGAGGTTCGCGCCCCTTGTCCTCCTTCGTTCAGAAGGCGCGACATCGGCG
241  AACACTGTTCGCTTCGCCTCCGGCCAAGTCGACAGCCAACGGAAACCTCCTAGAGCCGCG
301  GACCCTGCTCGAAGGACCTGACCCTGCCGAACTGCTCCTCATGGGCAGTTACCTGGGCAA
361  GCCCGGGCCGCCGCAGCCCGCCCCCGCTCCGGAGGGCCAGGACCTGCGGAATAGGCCTGG
421  CCGCCGCCCGCCCGCCCGGCGCCGCGCTCCACACCGCCCTCCCCGCCGACCCATCGCGTT
481  CACCACTTTTACCCCTCTCTCCCCACTCCTCTTCTCCGACCCTCCGGGAGGCCTTCCCCA
541  CGGGATCGTGGGACTTTACCAGATCGGTTTGTAATAACACCTCGAAGACGCTATCCGATC
601  CATCAGGCCCAGTATTCCTGTCCGGGGGTACTTCCCACAGTGTGCTGGAATGGTTATCAC
661  AAGAAGGCTGTGCTGTCCCCTCGCAACTCCAGGATGGTGTGTAGCCCAGTGACTGTGAGG   AA
                                  M  V  C  S  P  V  T  V  R          9

721  ATCGCCCCTCCTGACAGAAGATTTTCGCGTTCTGCGATACCAGAGCAGATAATCAGCTCA
       I  A  P  P  D  R  R  F  S  R  S  A  I  P  E  Q  I  I  S  S   29

781  ACACTGTCCTCACCATCAAGTAATGCCCCAGACCCATGTGCAAAGGAGACTGTACTGAGT
       T  L  S  S  P  S  S  N  A  P  D  P  C  A  K  E  T  V  L  S   49

841  GCCCTCAAAGAGAAGAAGAAGAAAAGGACAGTGGAGGAAGAAGACCAAATATTCCTTGAT
       A  L  K  E  K  K  K  K  R  T  V  E  E  E  D  Q  I  F  L  D   69
                  POM121◄┬►ZP3
901  GGCCAGGAAAATAAAAGAAGCTGTCTTGTCGACGGTCTCACTGATGCCTCTTCTGCATTC
       G  Q  E  N  K  R  S  C  L  V  D  G  L  T  D  A  S  S  A  F   89

961  AAAGTTCCTCGACCCGGGCCAGATACACTCCAGTTCACAGTGGATGTCTTCCACTTTGCT
       K  V  P  R  P  G  P  D  T  L  Q  F  T  V  D  V  F  H  F  A  109

1021 AATGACTCCAGAAACATGATATACATCACCTGCCACCTGAAGGTCACCCTAGCTGAGCAG
       N  D  S  R  N  M  I  Y  I  T  C  H  L  K  V  T  L  A  E  Q  129

1081 GACCCAGATGAACTCAACAAGGCCTGTTCCTTCAGCAAGCCTTCCAACAGCTGGTTCCCA
       D  P  D  E  L  N  K  A  C  S  F  S  K  P  S  N  S  W  F  P  149

1141 GTGGAAGGCCCGGCTGACATCTGTCAATGCTGTAACAAAGGTGACTGTGGCACTCCAAGC
       V  E  G  P  A  D  I  C  Q  C  C  N  K  G  D  C  G  T  P  S  169

1201 CATTCCAGGAGGCAGCCTCGTGTCGTGAGCCAGTGGTCCACGTCTGCTTCCCGTAACCGC
       H  S  R  R  Q  P  Ⓡ  V  Ⓥ  S  Q  W  S  Ⓣ  S  A  S  R  N  R  189

1261 AGGCATGTGACAGAAGAAGCAGATGTCACCGTGGGGGCCACTGATCTTCCTGGACAGGAG
       R  H  V  T  E  E  A  D  V  T  V  Ⓖ  Ⓐ  Ⓣ  Ⓓ  Ⓛ  Ⓟ  Ⓖ  Ⓠ  Ⓔ  209

1321 TGGTGACCATGAAGTAGAGCAGTGGGCTTTGCCTTCTGACACCTCAGTGGTGCTGCTGGG
       Ⓦ                                                          210

1381 CGTAGGCCTGGCTGTGGTGGTGTCCCTGACTCTGACTGCTGTTATCCTGGTTCTCACCAG
1441 GAGGTGTCGCACTGCCTCCCACCCTGTGTCTGCTTCCGAATAAAAGAAGAAAGCAA
```

FIG. 1. POM–ZP3 cDNA and deduced protein sequence. The nucleic acid (NA) sequence of the cDNA is numbered 1 to 1496 on the left, and the amino acid (AA) sequence of the longest open reading frame (printed in single-letter code) is numbered 1–210 on the right. The division of the cDNA and conceptual POM–ZP3 protein into regions homologous to rat POM121 and human ZP3 is indicated at nucleic acid residue 920. A solid triangle between residue 72 and 73 indicates a 32-nt deletion in POM–ZP3 compared to rat POM121. The initiator ATG and the termination codon TGA of the longest open reading frame are overlined. The 5 nucleic acid residues that differ between the conceptual POM–ZP3 and human ZP3 are underlined, and the amino acids that differ between the two are circled.

nt sequences at the 5′ end of POM–ZP3 diverge from those of ZP3 exactly at the 5′ terminus of exon 5 of human *ZP3*. These 5′ sequences of POM–ZP3 are 77% identical to the rat POM121 transcript (Fig. 2) that encodes a pore membrane protein (Hallberg *et al.*, 1993).

Although the 5′ residues of POM–ZP3 mRNA are well conserved with those that encode the first 240

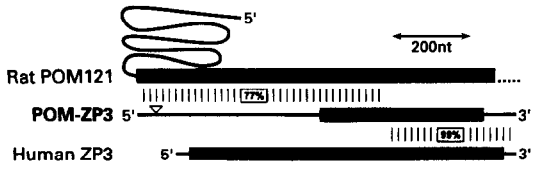

FIG. 2. Comparison of human POM–ZP3 with rat POM121 and human ZP3. Schematic representation of the 1496-bp bipartite human POM–ZP3 mRNA with a 5′ end that is similar to rat POM121 and a 3′ end that has near identity with human ZP3. Although the 5′ region is 77% identical to the open reading frame of rat POM121, a 32-bp deletion (triangle) in the POM–ZP3 transcript introduces multiple stop codons. Rectangles indicate open reading frames in each of the three proteins. The longest open reading frame in POM–ZP3 encodes a 210-amino-acid protein.

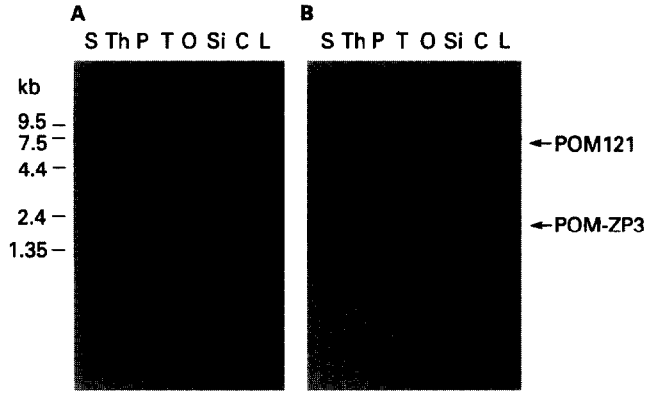

FIG. 3. Tissue-specific expression: Northern blot of eight different human tissues. (**A**) Autoradiograph of a Northern blot hybridized with a ^{32}P-labeled probe from the 5′ end of POM–ZP3 (691–907 nt). Each lane contains 2 μg of poly(A)$^+$ RNA from the following human tissues: S, spleen; Th, thymus; P, pancreas; T, testis; O, ovary; Si, small intestine; C, colon; L, lymphocytes. Numbers to the left are molecular size markers (kb). (**B**) Autoradiograph of the blot in **A** after it was stripped and rehybridized with a ^{32}P-labeled probe from the 3′ end of human ZP3 that is 99% identical to residues 921–1475 of POM–ZP3.

amino acids of rat POM121, a 32-bp deletion in POM–ZP3 results in stop codons in all three reading frames (Figs. 1 and 2). The longest open reading frame in POM–ZP3 begins at 694 nt with an AUG that conforms to the consensus sequence for an initiator methionine, ANNAUGG (Kozak, 1991). This open reading frame terminates at 1323 nt and could encode a 210-amino-acid polypeptide (13% acid, 14% basic, 8% aromatic, 46% hydrophobic) with a calculated molecular mass of 23,196 Da. The first 76 amino acids of this conceptual protein are 83% identical with rat POM121 (amino acids 241–315) and the next 125 amino acids have a 98% identity to a carboxyl region of ZP3. Although the nucleic acid identity between the POM–ZP3 and the ZP3 transcripts continues to the poly(A) tail, an additional guanosine at 1294 nt in POM–ZP3 results in premature termination of the polypeptide chain. Thus, the conceptual POM–ZP3 would be truncated with a unique 9-amino-acid terminus.

Tissue-Specific Expression

The approximate size and tissue distribution of POM–ZP3 mRNA was determined by Northern blot analysis. A 5′ ^{32}P-labeled POM–ZP3 probe (217 bp, restricted to the region homologous to rat POM121) was hybridized to poly(A)$^+$ RNA (2 μg) from eight different human tissues. Two transcripts (1.6 and 6.0 kb) of comparable abundance were detected (Fig. 3A). When the blot was stripped and rehybridized with a probe specific to the 3′ end of POM–ZP3 (containing 554-bp sequences homologous to ZP3), the 1.6-kb but not the 6.0-kb transcript was detected (Fig. 3B). When stripped and rehybridized a third time with a probe specific to the 5′ end of human ZP3 (exons 1–4), a very faint band (1.5 kb) was observed only in the lane containing ovarian RNA (data not shown).

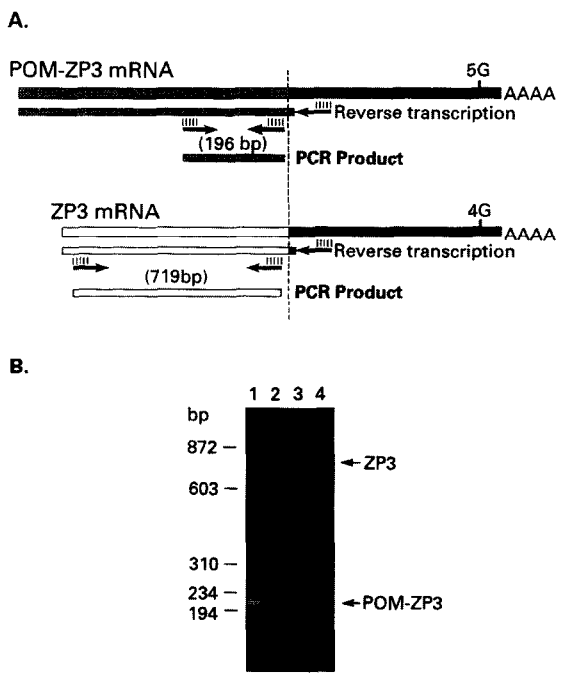

FIG. 4. Expression of ZP3 in the ovary. (**A**) Reverse transcription of human ovarian poly(A)$^+$ RNA was primed with an oligonucleotide specific to a region of identity between ZP3 (799–815 nt) and POM–ZP3 (996–1012 nt). The resultant single-stranded cDNA was used in parallel polymerase chain reactions with oligonucleotides to amplify cDNA fragments specific to ZP3 (719 bp) or POM (196 bp). (**B**) The cDNA products of the polymerase chain reactions were separated by gel electrophoresis and visualized with ethidium bromide. Each lane contains a 25-μl PCR reaction: lane 1, reverse transcription followed by PCR with oligonucleotides to POM sequences; lane 2, reverse transcription followed by PCR with oligonucleotides specific to ZP3; lane 3, same as lane 1 but in the absence of reverse transcriptase; lane 4, same as lane 2 but in the absence of reverse transcriptase.

A RT-PCR assay was used to confirm the presence of ZP3 transcripts in the human ovary. An oligonucleotide primer common to both POM–ZP3 (residues 996–1012) and ZP3 (residues 799–815) was used in a reverse transcriptase reaction with total human ovarian RNA. Sequences of 5′ cDNA were then amplified by PCR in separate reactions using either ZP3- or POM-specific oligonucleotide primers (Fig. 4A). The ZP3 primers from exons 1 and 4 amplify a 719-bp cDNA fragment; the POM primers amplify a 196-bp cDNA fragment. The PCR products were separated by gel electrophoresis and visualized with ethidium bromide. Both the ZP3- and POM-specific bands were detected in human ovarian tissue; neither was present as PCR products in reactions in which reverse transcriptase had been omitted (Fig. 4B). The identity of the ZP3 and the POM PCR products was confirmed by Southern blot analysis using probes specific to the 5′ end of ZP3 or the 5′ end of POM–ZP3 (data not shown).

From these data, we conclude that the 1.6-kb POM–ZP3 transcript is present in eight human tissues including the ovary, where it is expressed along with the previously described ZP3. In addition, there is a 6.0-kb transcript that contains the novel 5′ sequences of POM–ZP3 but not the 3′ sequences shared with ZP3. This larger transcript most likely represents the human homologue of rat POM121 or a closely related family member.

Chromosome Localization

The chromosomal localization of the locus that encodes POM–ZP3 was determined by fluorescence *in situ* hybridization. Plasmid clones with 4.5 kb of genomic DNA insert containing sequences homologous to rat *POM121* (see Materials and Methods) or 20 kb containing *ZP3* sequences were hybridized to human chromosomes. Thirty-eight paired signals were detected on chromosomes at 7q11.23 (Fig. 5) in 30 of 41 metaphase cells probed with the *POM121*-like fragment; 34 paired signals were detected on chromosomes at 7q11.23 in 31 of 43 metaphase cells probed with *ZP3*. No significant background was detected at any other chromosomal position with either probe. Chromosomal localization was confirmed by PCR of genomic DNA from human x rodent somatic cell hybrids. Sets of PCR primers specific to the POM or to ZP3 sequences produced expected PCR products of 479 and 623 bp, respectively, with DNA from a cell line known to contain human chromosome 7 but not with DNA from cells lines lacking this chromosome. The identities of the PCR products were confirmed by Southern blot analysis using gene-specific probes (data not shown).

DISCUSSION

The 1.6-kb POM–ZP3 mRNA is composed of 5′ sequences (920 nt) that are highly similar to rat *POM121* and 3′ sequences (576 nt) with near identity to human *ZP3*. POM–ZP3 mRNA is a relatively abundant transcript present in at least eight different human tissues as well as a human leukemia cell line, K562 (data not shown). *POM121* (or a highly similar family member) is also widely transcribed, whereas *ZP3* transcripts are detected in only human ovarian tissue as has been reported for other zona pellucida genes.

A 32-nt deletion near the 5′ end of the POM–ZP3 transcript shifts the reading frame from that of the POM121 mRNA, and the presence of stop codons would result in short (20–107 amino acids) polypeptides. The longest open reading frame begins at 693 nt and encodes 210 amino acids. The first 76 amino acids are homologous to rat POM121 but do not encode the N-terminal hydrophobic domain hypothesized to anchor POM121 in the nuclear pore membrane and do not contain any pentapeptide motifs (XFXFG) associated with nucleoporin proteins. The next 125 amino acids share near identity with human ZP3 but an additional guanosine residue causes a frameshift in the open reading frame so that the following 9 amino acids differ from ZP3. A stop codon results in a truncated protein lacking the hydrophobic domain associated with all zona proteins. POM121, a nuclear pore membrane protein, and ZP3, a zona pellucida protein involved in fertilization, have no sequence or functional similarities.

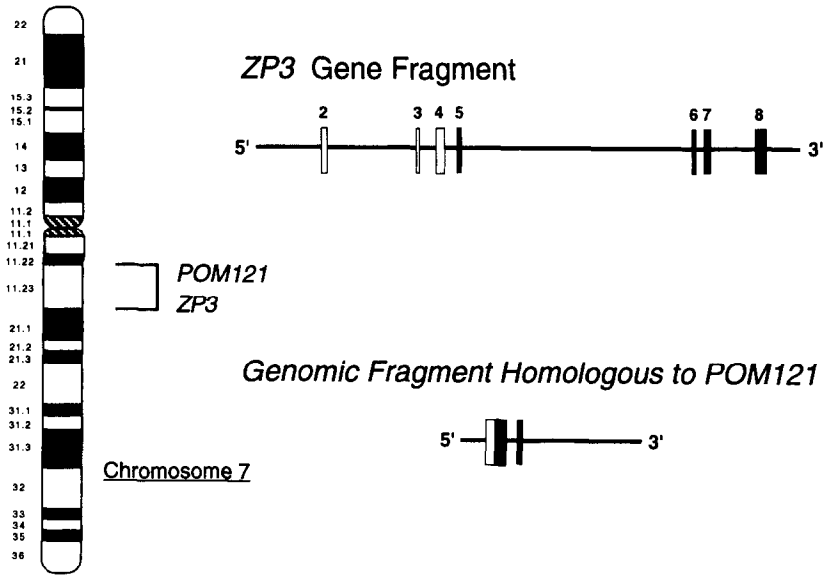

FIG. 5. Idiogram of human chromosome 7 illustrating the location of the *POM–ZP3* locus. Plasmid clones of human genomic DNA inserts containing sequences homologous to rat *POM121* (4.5 kb) or to human *ZP3* sequences (20 kb) were used in fluorescence *in situ* hybridization. The exon sequence with near identity to POM–ZP3 cDNA is indicated by the filled-in areas.

In preliminary studies using antibodies specific to the unique carboxyl terminal peptide of the conceptual POM–ZP3, no protein was detected in a variety of human tissue culture lines (Marcel van Duin, Oss, pers. comm., Sept. 1994). Taken together, these data suggest either that POM–ZP3 transcripts are not expressed as protein or that if expressed, POM–ZP3 does not play a functional role in the zona pellucida or in nuclear pore formation.

Given the high degree of identity between the two portions of POM–ZP3 and their ancestral genes, it appears that the gene encoding the POM–ZP3 transcript arose quite recently. If the POM–ZP3 transcript is not translated into protein, there should be no functional countervailing pressure against mutation. Using a molecular clock rate of 1.1 mutations/nucleotide site/million years (Bailey *et al.,* 1991), the five nonsynonymous (i.e., causing an amino acid substitution) nucleotide changes in the 576 bp of similarity between POM–ZP3 and ZP3 suggests that the two genes diverged approximately 4 (3.97 ± 1.78) million years ago (Li and Graur, 1991). In concert with this calculation, rodents, whose ancestors diverged from those of humans 65 million years ago, have at least one *POM121* locus and a single-copy *ZP3* locus, but do not express a *POM–ZP3* homologue (Philpott *et al.,* 1987; Ringuette *et al.,* 1988; Chamberlin and Dean, 1989; Hallberg *et al.,* 1993). It will be of interest to determine the presence or absence of the transcript in a species more closely related to humans, such as chimpanzee.

In situ fluorescence hybridization located *POM–ZP3* to human chromosome 7q11.23 in a region known to be syntenic with the location of mouse *Zp-3* on mouse chromosome 5, 9.2 cM distal to *Gus* (Lunsford *et al.,* 1990). The localization of rodent *POM121* has not been reported. A simple interpretation of these data is that a 5′ portion of a *POM121* homologue and exons 5–8 of human *ZP3* have been duplicated and juxtaposed so as to transcribe a novel POM–ZP3 transcript. Although the mechanism of such a duplication remains to be elucidated, the *ZP3* locus contains at least 16 *Alu* sequences (one of which is just upstream of exon 5), a concentration in excess of the average of one every 5–10 kb (Chamberlin and Dean, 1990). It may be that these repetitive elements have played a role in the creation of *POM–ZP3,* as has been postulated for other duplications (Hu and Worton, 1992). Future characterization of the *POM–ZP3* locus will clarify its structure and should provide insights into how it arose.

ACKNOWLEDGMENTS

We thank Dr. Marc Reitman for the molecular clock calculations and Dr. Marcel van Duin for sharing unpublished observations. We appreciate the critical reading of the manuscript by Drs. John Hanover, Marc Reitman, and Robert T. Simpson.

REFERENCES

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* **215:** 403–410.

Bailey, W. J., Fitch, D. H., Tagle, D. A., Czelusniak, J., Slightom, J. L., and Goodman, M. (1991). Molecular evolution of the psi eta-globin gene locus: Gibbon phylogeny and the hominoid slow-down. *Mol. Biol. Evol.* **8:** 155–184.

Chamberlin, M. E., and Dean, J. (1989). Genomic organization of a sex specific gene: The primary sperm receptor of the mouse zona pellucida. *Dev. Biol.* **131:** 207–214.

Chamberlin, M. E., and Dean, J. (1990). Human homolog of the mouse sperm receptor. *Proc. Natl. Acad. Sci. USA* **87:** 6014–6018.

Devereux, J., Haeberli, P., and Smithies, O. (1984). A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res.* **12:** 387–395.

Hallberg, E., Wozniak, R. W., and Blobel, G. (1993). An integral membrane protein of the pore membrane domain of the nuclear envelope contains a nucleoporin-like region. *J. Cell Biol.* **122:** 513–521.

Harris, J. D., Hibler, D. W., Fontenot, G. K., Hsu, K. T., Yurewicz, E. C., and Sacco, A. G. (1994). Cloning and characterization of zona pellucida genes and cDNAs from a variety of mammalian species: ZPA, ZPB, and ZPC gene families. *DNA Sequence* **4:** 361–393.

Hu, X., and Worton, R. G. (1992). Partial gene duplication as a cause of human disease [see comments]. *Hum. Mutat.* **1:** 3–12.

Kinloch, R. A., Ruiz-Seiler, B., and Wassarman, P. M. (1990). Genomic organization and polypeptide primary structure of zona pellucida glycoprotein hZP3, the hamster sperm receptor. *Dev. Biol.* **142:** 414–421.

Koothan-Thillai, P., van Duin, M., and Aitken, R. J. (1993). Cloning, sequencing and oocyte-specific expression of the marmoset sperm receptor protein, ZP3. *Zygote* **1:** 93–101.

Kozak, M. (1991). Structural features in eukaryotic mRNAs that modulate the initiation of translation. *J. Biol. Chem.* **266:** 19867–19870.

Lawn, R. M., Fritsch, E. F., Parker, R. C., Blake, G., and Maniatis, T. (1978). The isolation and characterization of a linked δ- and β-globin gene from a cloned library of human DNA. *Cell* **15:** 1157–1174.

Lee, V. H., Schwoebel, E., Prasad, S., Cheung, P., Timmons, T. M., Cook, R., and Dunbar, B. S. (1993). Identification and structural characterization of the 75-kDa rabbit zona pellucida protein. *J. Biol. Chem.* **268:** 12412–12417.

Li, W-H., and Graur, D. (1991). "Fundamentals of Molecular Evolution," Sinauer Associates, Sunderland, MA.

Liang, L-F., Chamow, S. M., and Dean, J. (1990). Oocyte-specific expression of mouse Zp-2: Developmental regulation of the zona pellucida genes. *Mol. Cell. Biol.* **10:** 1507–1515.

Liang, L-F., and Dean, J. (1993). Conservation of the mammalian secondary sperm receptor genes results in promoter function of the human homologue in heterologous mouse oocytes. *Dev. Biol.* **156:** 399–408.

Lunsford, R. D., Jenkins, N. A., Kozak, C. A., Liang, L-F., Silan, C. M., Copeland, N. G., and Dean, J. (1990). Genomic mapping of murine Zp-2 and Zp-3, two oocyte-specific loci encoding zona pellucida proteins. *Genomics* **6:** 184–187.

Philpott, C. C., Ringuette, M. J., and Dean, J. (1987). Oocyte-specific expression and developmental regulation of ZP3, the sperm receptor of the mouse zona pellucida. *Dev. Biol.* **121:** 568–575.

Ringuette, M. J., Chamberlin, M. E., Baur, A. W., Sobieski, D. A., and Dean, J. (1988). Molecular analysis of cDNA coding for ZP3, a sperm binding protein of the mouse zona pellucida. *Dev. Biol.* **127:** 287–295.

Schwoebel, E., Prasad, S., Timmons, T. M., Cook, R., Kimura, H., Niu, E-M., Cheung, P., Skinner, S., Avery, S. E., Wilkins, B., and Dunbar, B. S. (1991). Isolation and characterization of a full-length cDNA encoding the 55 kDa rabbit zona pellucida protein. *J. Biol. Chem.* **266:** 7214–7219.

Shabanowitz, R. B., and O'Rand, M. G. (1988). Characterization of the human zona pellucida from fertilized and unfertilized eggs. *J. Reprod. Fertil.* **82:** 151–161.

Tory, K., Latif, F., Modi, W., Schmidt, L., Wei, M. H., Li, H., Cobler, P., Orcutt, M. L., Delisio, J., and Geil, L. (1992). A genetic linkage map of 96 loci on the short arm of human chromosome 3. *Genomics* **13:** 275–286.

van Duin, M., Polman, J. E., Verkoelen, C. C., Bunschoten, H., Meyerink, J. H., Olijve, W., and Aitken, R. J. (1992). Cloning and characterization of the human sperm receptor ligand ZP3: Evidence for a second polymorphic allele with a different frequency in the Caucasian and Japanese populations. *Genomics* **14:** 1064–1070.

Yurewicz, E. C., Hibler, D., Fontenot, G. K., Sacco, A. G., and Harris, J. (1993). Nucleotide sequence of cDNA encoding ZP3 alpha, a sperm-binding glycoprotein from zona pellucida of pig oocyte. *Biochim. Biophys. Acta* **1174:** 211–214.