THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 277, No. 16, Issue of April 19, pp. 13518–13527, 2002
Printed in U.S.A.

An Imprinted *PEG1/MEST* Antisense Expressed Predominantly in Human Testis and in Mature Spermatozoa*

Received for publication, January 16, 2002
Published, JBC Papers in Press, January 30, 2002, DOI 10.1074/jbc.M200458200

Tao Li‡, Thanh H. Vu‡§, Kok-Onn Lee¶, Youwen Yang, Chuyen V. Nguyen‖, Huy Q. Bui, Zhi-Lan Zeng, Binh T. Nguyen, Ji-Fan Hu, Susan K. Murphy, Randy L. Jirtle**, and Andrew R. Hoffman§**

*From the Medical Service, Veterans Affairs Palo Alto Health Care System and the Department of Medicine, Stanford University School of Medicine, Palo Alto, California 94304, the ¶Department of Medicine, National University of Singapore, Singapore 119074, the ‖Department of Food and Nutrition, Japan Women's University, Tokyo 113, Japan, and the **Departments of Radiation Oncology and Pathology, Duke University Medical Center, Durham, North Carolina 27710*

***PEG1* (or *MEST*) is an imprinted gene located on human chromosome 7q32 that is expressed predominantly from the paternal allele. In the mouse, *Peg1/Mest* is associated with embryonic growth and maternal behavior. Human *PEG1* is transcribed from two promoters; the transcript from promoter P1 is derived from both parental alleles, and the transcript from P2 is exclusively from the paternal allele. We characterized the P1 and P2 transcripts in various normal and neoplastic tissues. In the normal tissues, *PEG1* was transcribed from both promoters P1 and P2, whereas in six of eight neoplastic tissues, *PEG1* was transcribed exclusively from promoter P1. Bisulfite sequencing demonstrated high levels of CpG methylation in the P2 region of DNA from a lung tumor. In the region between P1 and P2, we identified a novel transcript, *PEG1-AS*, in an antisense orientation to *PEG1*. *PEG1-AS* is a spliced transcript and was detected as a strong 2.4-kilobase band on a Northern blot. *PEG1-AS* and *PEG1* P2-sense transcript were expressed exclusively from the paternal allele. Fragments of DNA from within the 1.5-kilobase region between *PEG1-AS* and the P2 exon were ligated to a pGL3 luciferase reporter vector and transfected into NCI H23 cells. This DNA exhibited strong promoter activity in both the sense and antisense directions, indicating that *PEG1-AS* and P2 exon share a common promoter region. Treatment of the transfected DNA fragments with CpG methylase abolished the promoter activity. Of interest, *PEG1-AS* was expressed predominantly in testis and in mature motile spermatozoa, indicating a possible role for this transcript in human sperm physiology and fertilization.**

Genomic imprinting is an epigenetic modification in the germ line leading to preferential expression of one of the two parental alleles in a parent-of-origin-specific manner. Although the precise mechanisms underlying genomic imprinting have not been fully elucidated, both differentially methylated regions (DMR)[1] and imprinted antisense transcripts have been observed in association with imprinting. Based on the parent-of-origin of the transcribed allele, imprinted genes can be categorized into two groups, maternally expressed genes and paternally expressed genes (PEG). Using subtraction hybridization between cDNAs from normal and parthenogenetic mouse embryos, Kaneko-Ishino *et al.* (1) identified eight *Peg*s including *Peg1* and *Peg3*, two imprinted genes associated with abnormal maternal behavior in the *Peg*-deficient females when the disrupted Pegs were paternally transmitted (2, 3). *Peg1* (or *Mest*, a mesoderm-specific transcript) belongs to the α/β-hydrolase family (1). Human *PEG1/MEST* is also an imprinted gene expressed from the paternal allele and located on chromosome 7q32 (4).

Human *PEG1* is transcribed from two alternative promoters, resulting in the transcription of two isoforms, of which only one (isoform 1) is imprinted (5). Paternal expression of isoform 1 occurs in conjunction with an unmethylated CpG island in exon 1 of the paternal allele, whereas the corresponding CpG island in the maternal allele is fully methylated (6). Imprinting of *PEG1* is lost in lymphocytes and lymphoblastoid cell lines (6) and frequent loss of imprinting (LOI) of *PEG1* is observed in invasive breast cancer (7) and in colorectal cancer (8).

In this report, we have analyzed transcripts of *PEG1* in various normal and neoplastic tissues and have described a novel, imprinted paternally expressed antisense transcript, *PEG1-AS*, within the two promoter regions. The allelic expression of both sense and antisense transcripts may be regulated by a common DMR, but the tissue distribution of these transcripts varies dramatically. No antisense transcripts were observed in some malignant tissues including breast, colon, lung, ovary, and pancreas, and surprisingly, *PEG1-AS* was found most abundantly in the testis and in mature sperm.

EXPERIMENTAL PROCEDURES

Human Tissue and RNA Extraction—Normal fetal tissues of 6–12 weeks of gestation were obtained from the Central Laboratory for Human Embryology Tissue, University of Washington, Seattle, WA. Fresh motile sperm were collected from healthy adult volunteers. Total nucleic acid was prepared as described previously (9). RNA and DNA were extracted from fresh-frozen tissues using Tri reagent (Sigma) or by

* This study was supported by National Institutes of Health Grant DK36054 and by the Research Service of the Department of Veterans Affairs. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ These authors contributed equally to this work.

§ To whom correspondence may be addressed: Bldg. 101, Rm. B2–125, Veterans Affairs Palo Alto Health Care System, 3801 Miranda Ave. Palo Alto, CA 94304. Tel.: 650-493-5000 (ext. 63185); Fax: 650-856-8024; E-mail: thanhvu@stanford.edu or arhoffman@stanford.edu.

[1] The abbreviations used are: DMR, differential methylation region; *PEG1*, paternally expressed gene 1; *MEST*, mesoderm-specific transcript; *IGF2*, human insulin-like growth factor II gene; LOI, loss of imprinting; RT, reverse transcription; RACE, rapid amplifications of cDNA ends; bp, base pairs(s); MR, methylation and restriction; kb, kilobase(s); nt, nucleotide(s).

using an RNA preparation kit from Qiagen (Valencia, CA).

Reverse Transcription—Reverse transcription (RT) was performed with murine leukemia reverse transcriptase (Invitrogen) using both random hexamers and d(T)17 primers as described previously (9). To eliminate any residual genomic DNA, total RNAs (or total nucleic acids) were treated with DNase I (Invitrogen) for 45 min (2 units/1 μg of RNA) and then extracted with phenol-chloroform before reverse transcription. Various human cDNAs from poly(A^+) RNA were also obtained from CLONTECH (Palo Alto, CA).

5′- and 3′-Rapid Amplifications of cDNA Ends (RACE) and DNA Sequencing—RACE-PCRs were performed using Marathon-Ready cDNA of human fetal brain tissue according to the suggested protocol (CLONTECH).

The PCR products were subcloned using a TOPO-TA cloning kit (Invitrogen). DNA sequencing was performed on an ABI 377 sequencer using Big-Dye terminator chemistry (PerkinElmer Life Sciences). Sequences of the primers are as follows: p#463 (5′-GTT GAG ATG GGC ATG GAA CCC AGA TCT-3′), p#464 (5′-GAA CCC TGA GTA CAG AAG TCA CCT CA-3′), p#465 (5′-TGT GCC AGA GGT TCT GAT GAT AGG CT-3′), p#466 (5′-AGG CGG TAG TTT CCT CAG TGT CCG TG-3′), p#479 (5′-CAA GAG CCA TCA TTG TAG ATT CA-3′).

Expression of PEG1 Sense and PEG1 Antisense Transcripts—Transcripts specific for each promoter (P1a and P1b, isoform 2; P2, isoform 1) were amplified using promoter-specific primers (p#441 for P1a and P1b, p#467 for P2) and a common primer p#444 in exon 2. *PEG1-AS* was amplified by primers p#481 and p#566. The amplification was performed for 35 cycles at 95 °C for 20 s, 65 °C for 20 s, and 72 °C for 1 min and a final extension at 72 °C for 5 min. The PCR products were labeled by [^{32}P]dCTP and analyzed on a 5% polyacrylamide-urea gel.

Sequences of the primers used for the expression analysis are as follows: p#441 (5′-GAG TCC TGT AGG CAA GGT CTT ACC T-3′), p#467 (5′-GCT GCT GGC CAG CTC TGC ACG GCT G-3′), p#444 (5′-CTT GCC TGA AGA CTT CCA TGA GTG A-3′), p#566 (5′-CCT CCT TCA AGG AAG TTT TTC CTA ACT CCT-3′), and p#481 (5′-AGC CTA TCA TCA GAA CCT CTG GCA C-3′).

Northern Blot Analysis of PEG1 Sense and Antisense—Multiple tissue Northern blots containing poly(A^+) RNAs isolated from human adult tissues were obtained from CLONTECH. The Northern filter was hybridized with single strand oligonucleotide DNA probes that hybridize specifically to the sense (or antisense) mRNAs. The oligonucleotide probes were end-labeled with [γ-^{32}P]ATP (Amersham Biosciences) and hybridized to the filter for 2 h at 40 °C in a Rapid-hyb buffer (Amersham Biosciences). The filter was also hybridized to a cDNA probe labeled by PCR amplification (primers p#647, 5′-ATC CAG AGT TTT TGG AGC TGT ACA GGA-3′ and p#619 (5′-GAG AGA ATT ATG TTT GAT GGC CAG GAC CT-3′) using [^{32}P]dCTP) at 65 °C. The filter was subsequently washed three times for 10 min each with 0.2 × SSC (1× SSC = 0.15 M NaCl and 0.015 M sodium citrate) containing 0.1% SDS at 60 °C and analyzed by a PhosphorImager (Molecular Dynamics). The sequences of the oligonucleotide nucleotide probes are as follows: PEG1 sense mRNA, p#575 (5′-TGT ACT CAG GGT TCT TCC AAA CAG GAA TAT GCT TTC-3′); PEG1-*AS* RNA, p#578 (5′-TGC GGC ACT G CG CTT GCG AGG CGC AGC TGC CGC AGA GGA GGT GCC-3′); β-actin, p#5900 (5′-CGG ATG TCM ACG TCA CAC TTC ATG A-3′) and p#775 (5′-GGA AGC TTA TCA AAG TCC TCG GCC ACA-3′).

Allelic Expression of PEG1 Antisense—For genotyping, amplification was performed for 35 cycles (95 °C for 20 s, 65 °C for 20 s, and 72 °C for 1 min and a final extension at 72 °C for 5 min) using primers p#478 (5′-CAA CAG GAA TTC CAA GTC TGG CAC TC-3′) and p#483 (5′-CTG CCT AAG ATC TGG GTT CCA TGC C-3′). The PCR products were digested with *Sty*I (New England Biolabs, MA) to reveal an undigested allele (137 bp, T allele) and a digested allele (65 and 72 bp, C allele). The cDNAs from informative samples were first amplified by 38 cycles using cross-intron primers p#481 and p# 483. The 306 bp PCR products were diluted 5-fold and then labeled by five PCR cycles using end-labeled primer p#478. The labeled PCR products were diluted 5-fold and then digested with *Sty*I. The digested C allele showed only one end-labeled band (65 bp), whereas the undigested T allele showed a 137-bp band.

DNA Methylation Analysis by Methylation and Restriction (MR) PCR—We analyzed DNA methylation at the promoter regions of *PEG1* sense and *PEG1-AS* by a modified method of the combined bisulfite restriction analysis (10). The combined bisulfite restriction analysis method includes the digestion of PCR products of sodium bisulfite-treated genomic DNAs with a CpG-containing restriction enzyme (CGCG, *Bst*UI; or TCGA, *Taq*I). The digested products were separated by gel electrophoresis, blotted on a nylon filter, and hybridized to a specific probe. We simplified the combined bisulfite restriction analysis by directly labeling the PCR products with ^{32}P and analyzing the digested products using a PhosphorImager. The method involved only amplification and restriction digestion (MR-PCR).

Genomic DNAs were treated with sodium bisulfite to convert all unmethylated cytosines (but not methylated cytosines) to uracils according to a published protocol (11). Methylated cytosines retained after bisulfite treatment were PCR-amplified as cytosines, which were accessible to restriction digestion. Unmethylated cytosines that were converted to uracils (PCR-amplified as thymines) were undigested by the restriction enzyme. Therefore, in MR-PCR, undigested fragments represent the unmethylated genomic DNAs, whereas amplified products from methylated DNAs are completely digested.

We designed primers to amplify bisulfite-treated DNAs encompassing restriction sites (*Bst*UI or *Taq*I) and employed a touchdown PCR program. The PCR program had a decreasing annealing temperature from 70 to 55 °C for the first 30 cycles (95 °C for 2 min, 30 cycles at 95 °C for 15 s, 70 to 55 °C for 1 min). The DNA was further amplified at a 55 °C annealing temperature (18 cycles of 95 °C for 15 s, 55 °C for 1 min, and a final extension at 70 °C for 10 min). The ^{32}P-labeled products were diluted 5-fold and then digested with the restriction enzymes overnight. Restriction enzyme-treated products were analyzed by 5% polyacrylamide-urea gel electrophoresis, and the relative amounts of undigested PCR products (percent of unmethylated DNA) were quantified by a PhosphorImager. The sequences of the primers are as follows: p#548 (5′-GTG TTY GTT ATG TGA ATA GTA GAA AGT AG-3′), p#549 (5′-AAT AAC ATT TTA ATA ATT ATT AAC TAC CT-3′), p#550 (5′-ATG TAA AAG TAT TTA GAT TAG GTY GAA-3′), p#551 (5′-CCA AAA ATT CTC GAC CTT CAC CCT ATT CCC-3′), p#552 (5′-AGG TTA TAA AGA GTT TAA ATT TAT TTG TGG-3′), p#553 (5′-CTA ACC CCT AAT AAA TAC TAC TTA AAT-3′), p#554 (5′-GTG TTA GAG GTT TTG ATG ATA GGT TTA TAG-3′), p#555 (5′-CAC GAA ATA TCT ATA TTC ATT TTC TAA CC-3′), p#556 (5′-GTT GTT GTG GTT AGG YGT TTG GTA TGT TGA-3′), p#557 (5′-CCA CTA TAA CCA AAA TTA CAC AAA ATC C-3′), p#558 (5′-GTT TGG TGG TGG GTT TAA TAG AGT TTG TTG-3′), and p#559 (5′-TAA ATA CCC CAA CTC TTT CCT TAA ACC ACC-3′). The methylation status of all cytosines (CpG and CpN) was also determined by bisulfite sequencing according to our previously published protocol (11).

Luciferase Constructs and Transient Transfection Assays—The DNA fragments in the 5′-flanking region of *PEG1*-P2 and *PEG1-AS* that were obtained by PCR amplification of human genomic DNA were cloned into the pGL3-basic luciferase reporter vector (Promega, Madison, WI). The primers were designed to include restriction sites for *Xho*I and *Sac*I to allow directional cloning. For *PEG1-AS*, six constructs containing the putative antisense promoter fragments (287, 362, 417, 674, 887, and 1193 bp, constructs #1–6, respectively) were generated by using primer p#826 and one of the six primers (p#952, p#829, p#953, p#830, p#949, and p#948). Three *PEG1* P2 sense constructs were also generated by PCR using primer p#825 and one of the three primers (p#954, p#823, and p#822). The sense constructs 7, 8, and 9 contained P2 promoter fragments extended to −309, −580, and −1506 bases, respectively, upstream of the P2 transcription site. To methylate all CpGs in the constructs, plasmids (3, 4, 6, 8, and 9) were treated with *Sss*I methylase (New England Biolabs, MA) to create methylated constructs 3*, 4*, 6*, 8*, and 9*. Target DNAs from constructs 4 and 8 were purified, methylated by *Sss*I methylase, and ligated to the luciferase reporter to make constructs 4m and 8m (only target DNA was methylated) by standard molecular techniques.

NCI-H23 cells (ATCC#CRL-5800) were cultured in 48-well plates at a density of 40,000 cells/well on day 1. Cultures were transiently transfected on day 2 using GenePORTER 2 transfection reagent (Gene Therapy System, San Diego, CA) according to the suggested protocol. Briefly, for each well, hydrated GenePORTER 2 (2.5 μl) was diluted with 10 μl of serum-free medium. Reporter DNA (500 ng) and pSV-β-galactosidase plasmid (25 ng) were diluted with 12.5 μl of DNA diluent and then incubated for 5 min at room temperature. The DNA solutions were added to the diluted GenePORTER-2 reagent and incubated at room temperature for 10 min to form GenePORTER 2/DNA complexes. The complexes were added to the cells growing in 200 μl of serum-free culture medium. After a 4-h incubation at 37 °C, equal volumes of medium containing 20% serum were added, and the assay was performed 48 h after transfection.

The luciferase activities were determined in the linear range for each assay, and the activity of the co-transfected "control" pSV-β-galactosidase provided an internal control. Luciferase and β-galactosidase activities were assayed according to the instructions provided with the reagent kit for each enzyme (Promega). The normalized luciferase activity was expressed as fold of activity over the control vector pGL3-basic.

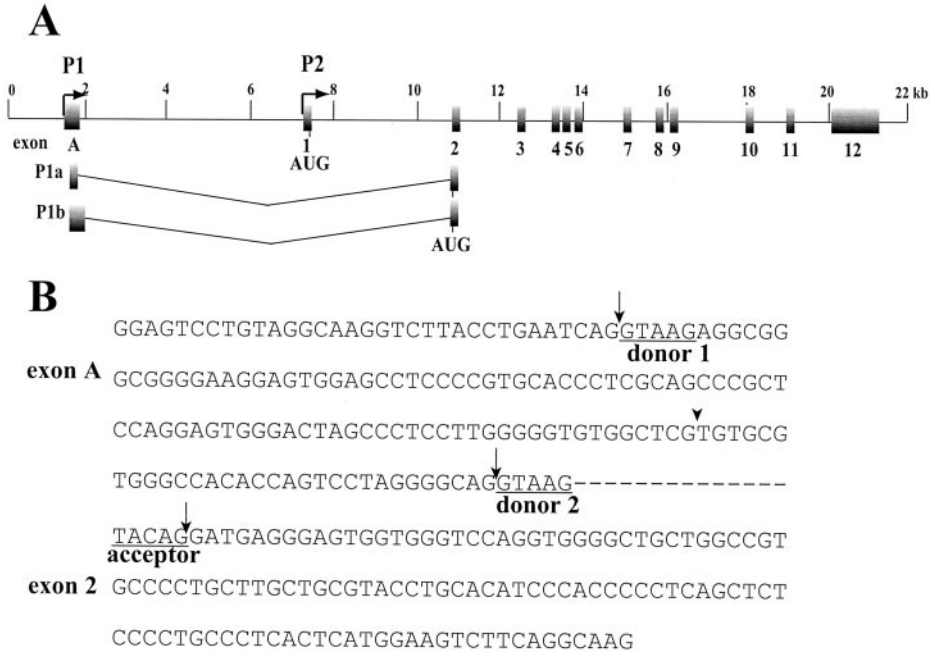

FIG. 1. **Alternative splicing in P1 promoter exon of *PEG1*.** *A*, map of P1 and P2 transcripts. *PEG1* is transcribed from two alternative promoters, P1 and P2. Alternative splicing in exon A results in P1a and P1b. Exons are identified by the *number* below each *box*. The diagram was drawn to scale. *B*, DNA sequencing identified a new splicing donor site (donor site 2). The acceptor site is in exon 2.

Sequences of the primers used for the luciferase constructs are as follows: p#826 (5′-AACCTCTCGAGGTGCAGAGTTGCAGAGCCGCG-GAGG-3′), p#952 (5′-ACGTGAGCTCGAATCACAATGCAAGGGCCCA-G-3′), p#829 (5′-CAGCCGAGCTCGCTGCATGCGCCGGTCCGCGCTG-CA-3′), p#953 (5′-ACAGGAGCTCGACGTGGTCAGCATGCCAGACG-3′), p#830 (5′-GGCGCGAGCTCCCATCCCTCGTTCGAAGCGTGGGT-3′), p#949 (5′-ACCCTGAGCTCGAGATTCTGCTGCAGAAACC-3′), p#948 (5′-CCACGAGCTCCGCGCTGCCGCGGCAACCAGCAC-3′), p#825 (5′-GGTTCTCGAGGCAGCGCGGAGCGAGTGGGCACCGACT-3′), p#954 (5′-TGCTGTGAGCTCCCGTGGCCTTAACTCATCAGG-3′), p#823 (5′-GCTTGGAGCTCCCTAGCCGCTTGCTCGTGCCCTTG-3′), and p#822 (5′-ACTTGAGCTCTGCCTCTGAGGGCCTTACACGTTA-3′). The underlined sequences represent the restriction sites.

RESULTS

Multiple Human PEG1 Transcripts—The human *PEG1* gene is transcribed from two alternative promoters, P1 and P2 (Fig. 1*A*). The novel exon A was previously demonstrated by the analysis of *PEG1* expression in two patients with maternal uniparental disomy 7 (12). The two *PEG1* isoforms 1 and 2 (deriving from promoters P2 and P1, respectively) have recently been characterized further by Kosaki *et al.* (5). Promoter-specific imprinting has previously been demonstrated in the human *IGF2* gene (13, 14) and in the human (and murine) *GNAS* gene (15, 16). Paternally expressed antisense transcripts, *IGF2-AS* and *GNAS-AS*, were found in both imprinted regions (15–18).

The *PEG1* sense transcripts encode for a predicted *PEG1* protein, a member of an α/β-hydrolase fold family (1). We identified sense transcripts from the two promoters, P1 and P2, by performing RT-PCR using primers that are specific for promoter exon A and exon 1 and a common primer in exon 2 (Fig. 1*A*). In human ovary and spleen, the RT-PCR amplification yielded the expected major products (P1, 155 bp; P2, 230 bp) and a minor P1–274-bp product (data not shown). We sequenced the P1–274-bp product, compared the P1 sequence with the human genomic sequence (GenBank™ accession number AC007938) and found alternative splicing sites in exon A (Fig. 1*B*). Although the major P1 transcript (P1a or isoform 2) utilizes the splicing donor site 1, the novel transcript (P1b) utilizes donor site 2. Because the extended exon A contains no AUG codon, both P1a and P1b transcripts would utilize the translation codon AUG in exon 2. The P1a and P1b transcripts therefore code for a putative *PEG1* protein that is nine amino acids shorter than a putative *PEG1/MEST* protein from the P2 transcript (isoform 1), which has an AUG codon in the exon 1.

PEG1 Antisense Transcript in the P1-P2 Region—We used the BLAST program to search for ESTs within the *PEG1* human genomic sequence (GenBank™ accession number AC007938, nt 140,000–167,000) and identified six novel ESTs (AW161444, AW163454, AI188165, BE090096, AW157393, and AW156980) in the 5.7-kb DNA region between exon A and exon 1 (Fig. 2*A*, *arrows*). Two IMAGE clones (2783984 and 2782723) had EST sequences homologous to the 3290 bp genomic sequence ((GenBank™ accession number AC007938, nt 146372–143083). Comparison with the genomic sequence indicates that the 5′-ESTs (AW161444 and AW163454) are spliced transcripts spanning an 872-bp intron sequence (GenBank™ accession number AC007938, nt 145386–146259). The presence of the consensus donor and acceptor splicing sequences demonstrates that the novel transcript is in fact in antisense orientation to the *PEG1* sense transcripts from promoters P1 (a and b) and P2. Furthermore, the presence of RNA-splicing sites in the transcript argues against any artifact from genomic DNA cloning.

Characterization of the PEG1 Antisense Transcript by RACE and Northern Blotting—We performed RACE using human brain Marathon Ready cDNA (CLONTECH) in an attempt to obtain a full-length *PEG1-AS* transcript. In the 3′-RACE, primers p#463, p#464, and primer p#479 extended the sequence 1.8 kb downstream from the AW161444 sequence (Fig. 2*A*). In the 5′-RACE, primers p#465 and p#466 failed to extend the sequence further than the AW161444 sequence. DNA sequencing from the RACE products confirmed the 2418-bp sequence of the *PEG1-AS* transcript (as predicted from the two IMAGE clones). The *PEG1-AS* cDNA sequence overlapped the six EST sequences identified by BLAST analysis. The *PEG1-AS* gene is interrupted by a single 872-bp intron. There is a short open reading frame of up to 122 amino acids, but the predicted peptides lack homology to any known proteins.

To confirm the expression of *PEG1-AS* in various human tissues, we hybridized a poly(A)$^+$ multiple tissue Northern blot (CLONTECH) with a 45-bp single-stranded oligonucleotide probe that hybridizes specifically to the antisense transcript. As shown in Fig. 2*B*, *PEG1-AS* was detected as a strong band of ~2.4 kb in testis, prostate, and spleen and as a weaker

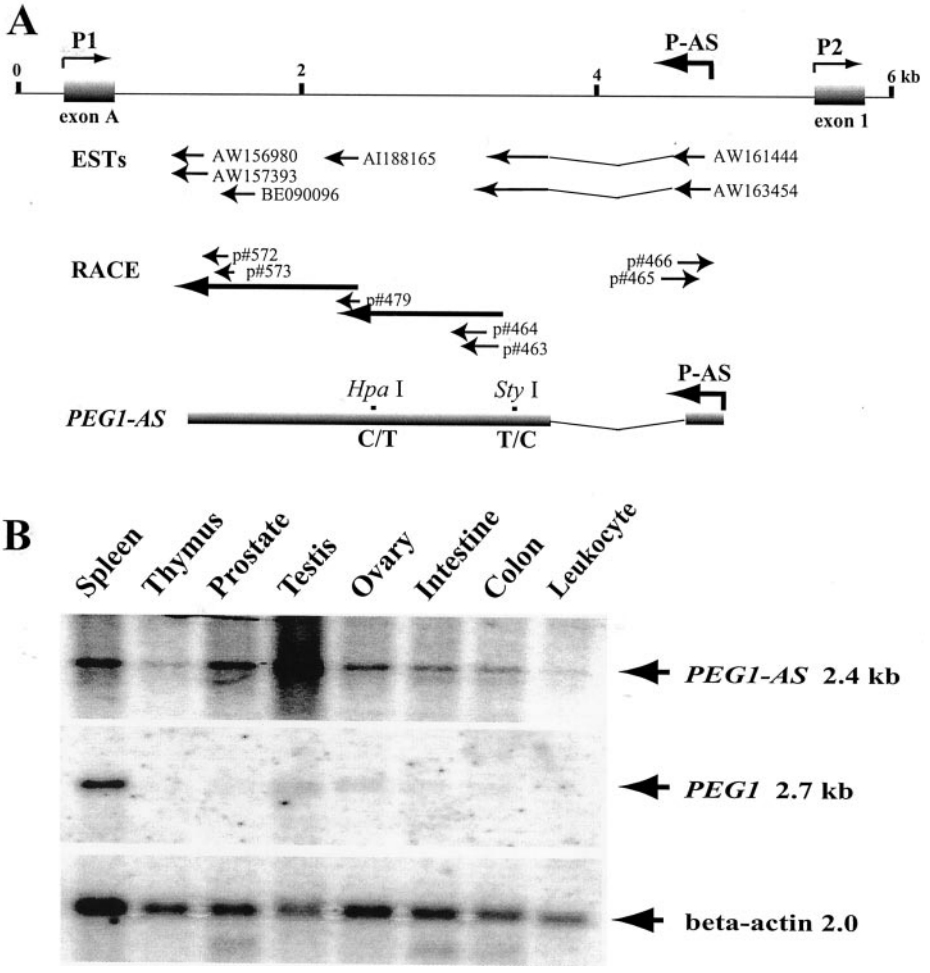

FIG. 2. **Characterization of *PEG-1-AS* transcript.** *A*, diagram showing the location of the *PEG1-AS*, ESTs, and primers for RACE analysis. Human genomic sequence (GenBank™ accession number AC007938 from nt 140001) was analyzed by BLAST analysis. Six ESTs (*arrows*) AW161444, AW163454, AW188165, BE090096, AW157393, and AW156980 were identified. The AW161444 and AW163454 ESTs were spliced transcripts. The *PEG1-AS* sequence was generated from the EST sequences and sequence from RACE. Two polymorphic sites (*Hpa*I and *Sty*I) are indicated. *B*, Northern blot analysis of *PEG1-AS* expression. The multiple tissue blots (CLONTECH) from adult human tissues were probed with a strand-specific oligonucleotide, p#575, for detecting antisense transcripts. The oligonucleotide was end-labeled by [γ-^{32}P] ATP and T4 protein kinase. *PEG1* sense-transcripts were probed with a standard cDNA probe. The same blot was re-hybridized with β-actin probe to verify RNA loading.

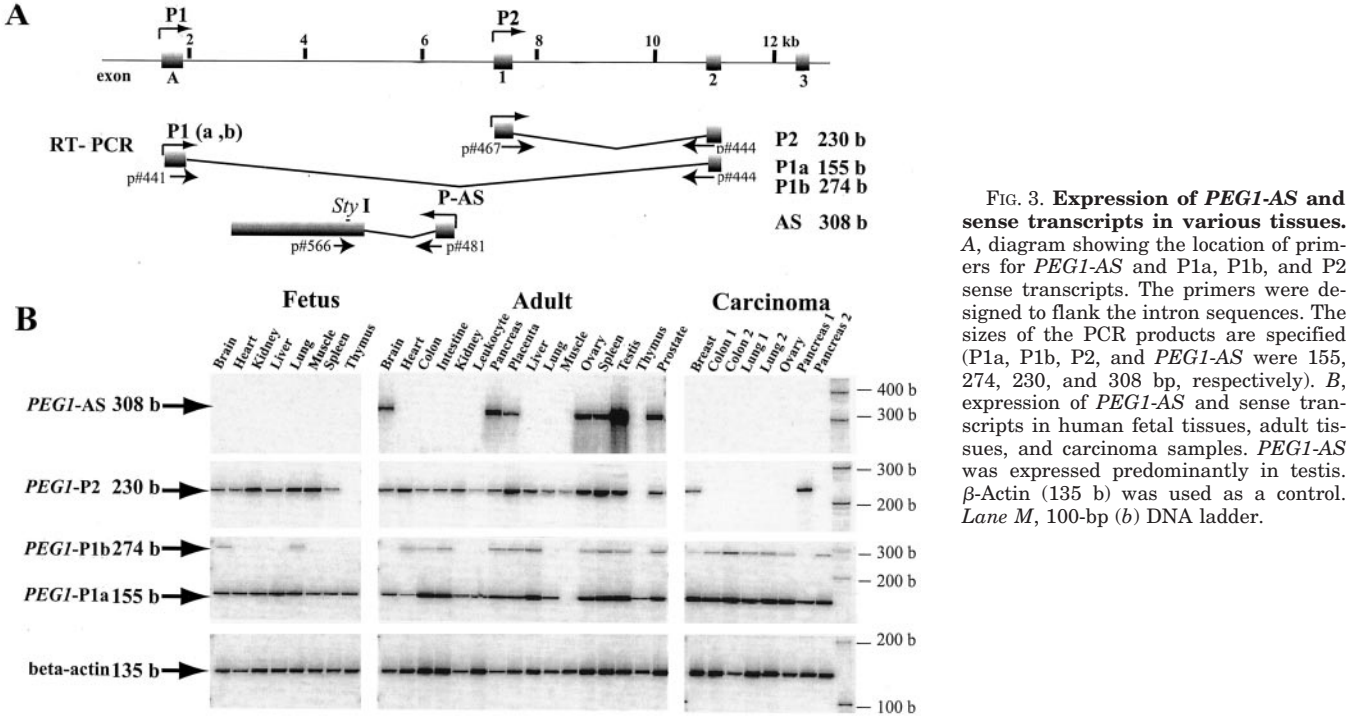

FIG. 3. **Expression of *PEG1-AS* and sense transcripts in various tissues.** *A*, diagram showing the location of primers for *PEG1-AS* and P1a, P1b, and P2 sense transcripts. The primers were designed to flank the intron sequences. The sizes of the PCR products are specified (P1a, P1b, P2, and *PEG1-AS* were 155, 274, 230, and 308 bp, respectively). *B*, expression of *PEG1-AS* and sense transcripts in human fetal tissues, adult tissues, and carcinoma samples. *PEG1-AS* was expressed predominantly in testis. β-Actin (135 b) was used as a control. *Lane M*, 100-bp (*b*) DNA ladder.

~2.4-kb band in other tissues. A cDNA probe spanning part of exons 11 and 12 detected *PEG1* sense transcripts as a strong 2.7-kb band in spleen and as a weaker 2.7-kb band in testis, ovary, intestine, and colon. A strand-specific oligonucleotide probe in exon 1 also detected a strong 2.7-kb P2 sense transcript in spleen (data not shown), whereas strand-specific oligonucleotide probes for β-actin confirmed similar loading of RNA in all lanes (Fig. 2*B*).

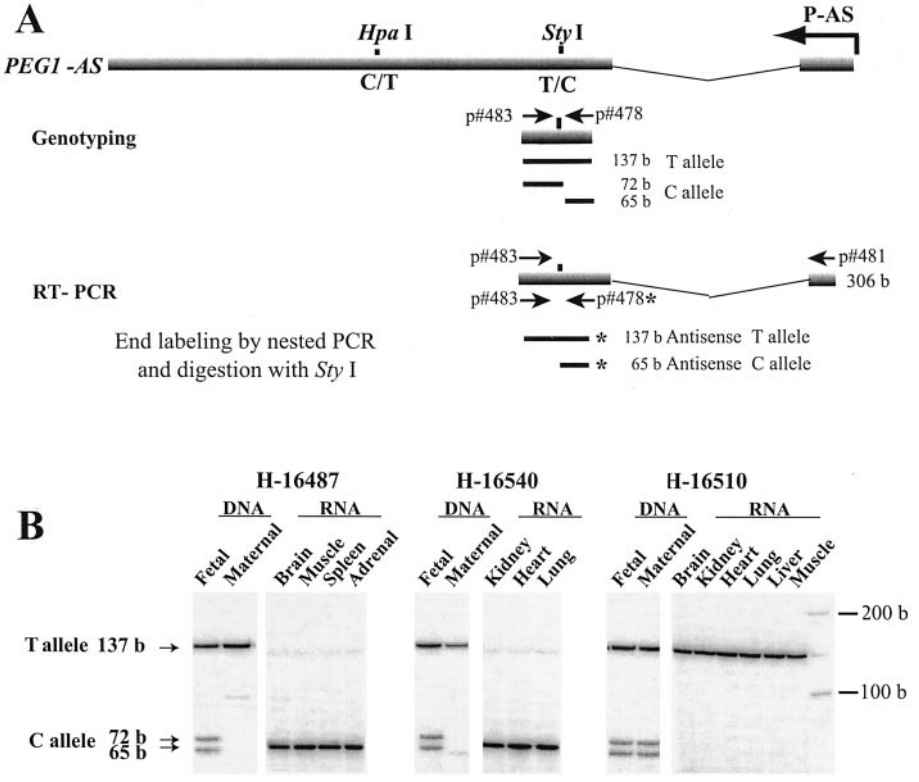

FIG. 4. **Allelic expression of *PEG1-AS*.** *A*, map of the *PEG1-AS* and location of primers. The polymorphic site *Sty*I (GenBank™ accession number AC007938, positions145137 C/T) was used for genotyping and for allelic expression analysis of *PEG1-AS*. Genomic DNA was amplified by primers p#483 and p#478. Digestion with *Sty*I revealed an undigested allele (137 bp, T allele) and a digested allele (72 and 65 bp, C allele). Allelic expression was analyzed by RT-PCR using cross-intron primers (p#483 and p#481), which yielded a 306-bp PCR product. The amplified product was labeled by a nested PCR using end-labeled primer p#478. Only one of the two digested bands was labeled by primer p#478 (65 bp, C allele). *B*, allelic expression of *PEG1-AS* in fetal tissues and maternal decidua. Both fetal subjects H-16487 and H-16540 were heterozygous (T and C alleles), whereas maternal decidua were T/T homozygous. In both subjects, *PEG1-AS* was expressed predominantly from the C allele (65 bp), indicating a paternal expression. In subject H-16510, both fetal and its maternal tissue DNA were T/C heterozygous; *PEG-1-AS* was, however, transcribed from only one allele (T allele). *Lane M*, 100-bp (*b*) DNA ladder.

Tissue-specific Expression of PEG1-AS—Because Northern blots indicated that the expression of *PEG1-AS* is predominant in one tissue (testis *versus* seven other tissues), we further analyzed the tissue-specific expression of *PEG1-AS* by RT-PCR in a larger panel comprising 16 adult tissues, 8 fetal tissues, and 8 neoplasms. The cDNAs were made from purified mRNAs (poly(A)$^+$ fraction, CLONTECH). *PEG1-AS* was amplified using primers (p#481 and p#566) crossing the 872-bp intron. The amplified product, a 308-bp band, is therefore a genuine full-processed transcript of *PEG1-AS* (Fig. 3), whereas contaminating genomic DNA, if any exists, would present as a band of 1180 bp.

There was large variation in the level of *PEG1-AS* expression (Fig. 3). Under the log phase of amplification (high annealing temperature of 65 °C), *PEG1-AS* was detected predominantly in the testis and (in lesser amounts) in prostate, ovary, spleen, pancreas, brain, and placenta. This pattern of tissue-specific expression is similar to that observed by Northern blotting using a single strand-specific probe (Fig. 2*B*), which confirms the predominant expression of *PEG1-AS* in testis among various adult tissues. Under these PCR conditions, no detectable *PEG1-AS* was seen in eight neoplasms and in eight fetal tissues. However, *PEG1-AS* could be amplified in fetal tissues by using a nested PCR (see Fig. 4).

PEG1 Sense Transcripts in Normal and in Neoplastic Tissues—As a cross control of *PEG1-AS* expression, we also analyzed *PEG1* sense transcripts from the two promoters P1 (P1a, 155 bp; P1b, 274 bp) and P2 (230 bp) and β-actin across the same cDNA panel. The sense transcripts were amplified using primers (crossing intron) in the alternative promoter exons (exon A and exon 1) and a primer in the common exon 2 (Fig. 3). We designed the primers with similar T_m values so that we could perform semi-quantification of the relative levels of P1 and P2 transcripts in a multiplex PCR assay; we have previously used a similar multiplex PCR procedure to quantify multiple transcripts of the *IGF2* (9). Although β-actin (amplified at 25 cycles) showed similar loading of cDNAs across the panel, levels of expression of *PEG1* sense transcripts varied. We noted that 1) both P1a and P2 transcripts were the major ones observed in various fetal and adult tissues, whereas P1b, which was amplified by the same P1a primer set, was a minor transcript, and (2) the relative levels of P1a/P2 varied in a number of tissues. Interestingly, in six of the eight neoplasms (two colon, two lung, one ovary, and one pancreas) tested, *PEG1* is exclusively transcribed from promoter P1 (exon A). However transcripts from promoter P2 were also detected in other tumor tissues (1 breast, 1 pancreas, and 10 Wilms' tumor; Fig. 3*B* and data not shown). It has recently been shown that the *PEG1* transcript from exon A (isoform 2) is expressed biallelically in contrast to the exclusively paternal expression from exon 1 (5) and that frequent loss of imprinting (LOI) of *PEG1* has been observed in invasive breast cancer (7). Our present data indicate that the mechanism for this LOI in some neoplastic tissues may be the switching of promoter usage from promoter P2 to the exclusive expression from promoter P1 (exon A). We further analyzed the CpG methylation of the P2 promoter region in the human Hep 3B cell line (ATCC HB-8064), the Hs 1Tes cell line (a fibroblast from normal embryonic testis, ATCC CRL-7002), and a pair of lung tumors and adjacent normal tissues. Both MR-PCR and bisulfite genomic sequencing revealed hyper-DNA methylation in the Hep 3B cell line and in the lung tumor tissue (see Fig. 7). This suggests that hypermethylation of CpG in the P2 promoter region during tumorigenesis may be responsible for the silencing of the *PEG1* P2 sense transcript in some neoplastic tissues.

PEG1-AS Is Imprinted and Expressed Exclusively from the Paternal Allele—To investigate the allelic expression of *PEG1-AS*, we first searched for single nucleotide polymorphisms in the antisense sequence. Comparison of sequences from our RACE results and the EST sequences from (GenBank™ yielded two putative polymorphisms: a 145137 C/T transition creating a *Sty*I site and a 144370 T/C transition creating a *Hpa*I site ((GenBank™ accession number AC007938). We confirmed the two single nucleotide polymorphisms by PCR and

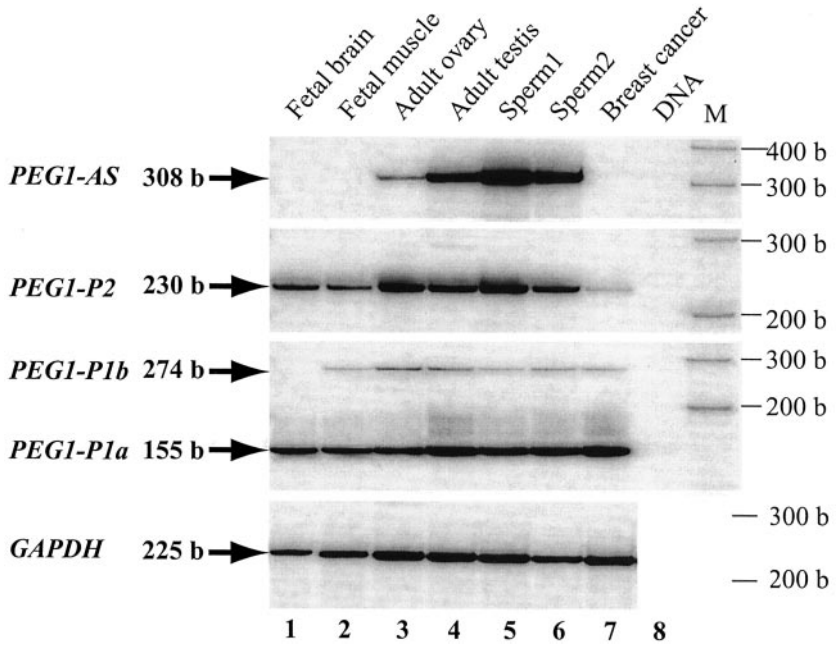

FIG. 5. **Predominant expression of *PEG1-AS* in adult testis and in mature sperm.** Expression of *PEG1-AS* and sense transcripts in motile sperm from two subjects was compared with *PEG1* expression in other tissues (fetal brain, fetal muscle, ovary, testis, and breast carcinoma). The primers specific for each transcript (*PEG1-AS*-308 bp, P2–230 bp, P1a-155 bp, P1b-274 bp) and RT-PCR conditions are described in the legend for Fig. 3. *DNA*, human genomic DNA as a negative control. *Lane M*, 100-bp (*b*) DNA ladder. *GAPDH*, glyceraldehyde-3-phosphate dehydrogenase.

restriction digestion (data not shown) and established the frequency of these alleles in a normal population. The frequency of the 145137 C and T alleles was 95 and 5%, respectively, and the frequency of the 144370 T and C alleles was 72 and 28%, respectively (in 110 chromosomes studied).

We chose the single nucleotide polymorphism 145137 C/T (*Sty*I site) that is located close to an exon/intron border of the *PEG1-AS* for the allelic expression study (Fig. 4). We genotyped DNA from a collection of 37 pairs of fetal and maternal (decidua) samples using the primers p#483 and p#478 and found three heterozygous fetuses. The mothers of two subjects (H-16487 and H-16540) were homozygous for the haplotype, and the mother of one subject (H-16510) was heterozygous for the haplotype (Fig. 4, DNA). We analyzed allelic expression of *PEG1-AS* in all available tissues from the three fetuses by using a nested RT-PCR. The cDNAs were first amplified by primers p#481 and p#483, crossing the 872-bp intron, and subsequently amplified by an internal, ^{32}P-end-labeled primer, p#478. Digestion of the PCR products with *Sty*I revealed two alleles, an undigested 137-bp T allele and a digested, end-labeled 65-bp C allele (Fig. 4*A*). In both informative fetuses, H-16487 and H-16540 (122 days and 80 days of gestation, respectively), *PEG1-AS* was expressed exclusively from the digested C allele in all tissues including brain, muscle, spleen, adrenal, kidney, heart, and lung (Fig. 4*B*, H-16487 and H-16540). Because the corresponding maternal decidua tissues were homozygous for the T allele, the expressed C allele must be derived from the paternal allele. In fetus H-16510 (86 days of gestation), *PEG1-AS* was expressed exclusively from the undigested 137-bp T allele in all tissues. However because the maternal tissue was heterozygous, the parental origin of the expressed T allele could not be deduced (Fig. 4*B*, H-1650). Thus, human *PEG1-AS* is imprinted in all tissues from the three subjects and expressed from the paternal allele in the two parentally informative subjects.

Biallelic Expression of PEG1-AS in a Population of Mature Sperm from a Single Donor—We were curious that the expression of *PEG1-AS* was relatively testis-specific, whereas the sense transcripts were ubiquitously expressed in all tissues that were examined (Fig. 3). Recent studies indicate the presence of various mRNAs detected by PCR amplification in human mature sperm, including *N*-cadherin, heat shock proteins, protamines, β-actin, and short and long interspersed repetitive elements (SINEs and LINEs) (19, 20). We detected high levels of the *PEG1-AS* transcript, comparable with those detected by Northern blotting in adult testis and in mature sperm from six different donors (two shown in Fig. 5, *lanes 5* and *6*). Interestingly, *PEG1*-sense transcripts were also observed at levels similar to those in adult testis and in other tissues (Fig. 5, compare *lanes 5* and *6* to *1–4* and *7*). Mock PCR using genomic DNA as template showed no detectable bands, which excluded the presence of pseudo-gene products and further confirmed the PCR products as the genuine spliced RNA transcripts (Fig. 5, *lane 8*).

Various mRNA species present in mature sperm are likely to be carried over from the early stages of spermatogenesis. If imprinting of *PEG1-AS* persists in mature sperm, this could be revealed by detection of only one *PEG1-AS* allele in a population of sperm from an informative donor. We identified four informative subjects by genotyping DNA from sperm using the *Hpa*I and the *Sty*I polymorphisms. Although the parental origin of each allele is unknown, transcripts originating from both parental alleles of *PEG1-AS* were observed in a population of sperm from each single donor (data not shown). This result is consistent with the previous finding that imprinting marks (of the *H19* and *PEG1* gene) in embryonic germ cells are erased very early, before mid-gestation (21, 22). The *PEG1* gene remains unmethylated in all stages of spermatogenesis including spermatogonia and primary and secondary spermatocytes (22). Transcription of the unmethylated *PEG1* gene in male germ cells is therefore expected from both parental alleles.

Methylation Analysis of PEG1 Gene in Sperm and in Other Tissues—Methylation analysis of a CpG island spanning the promoter exon 1 (P2) has been performed in both human and mouse. For the sense transcript derived from P2, the expressed paternal allele is unmethylated; in contrast, the silenced maternal allele is fully methylated (6, 23). Because *PEG1-AS* is most likely transcribed from the 1142-bp region (GenBank™ accession number AC007938, nt 146372–147514) 5′ of exon 1, we further extended the methylation analysis in this region and in the promoter P1 region.

We used the MR-PCR method to estimate the percentage of methylated alleles (digested fraction by *Bst*UI or *Taq*I restriction) in 7 PCR products in the 2.5-kb region of *PEG1-AS* and

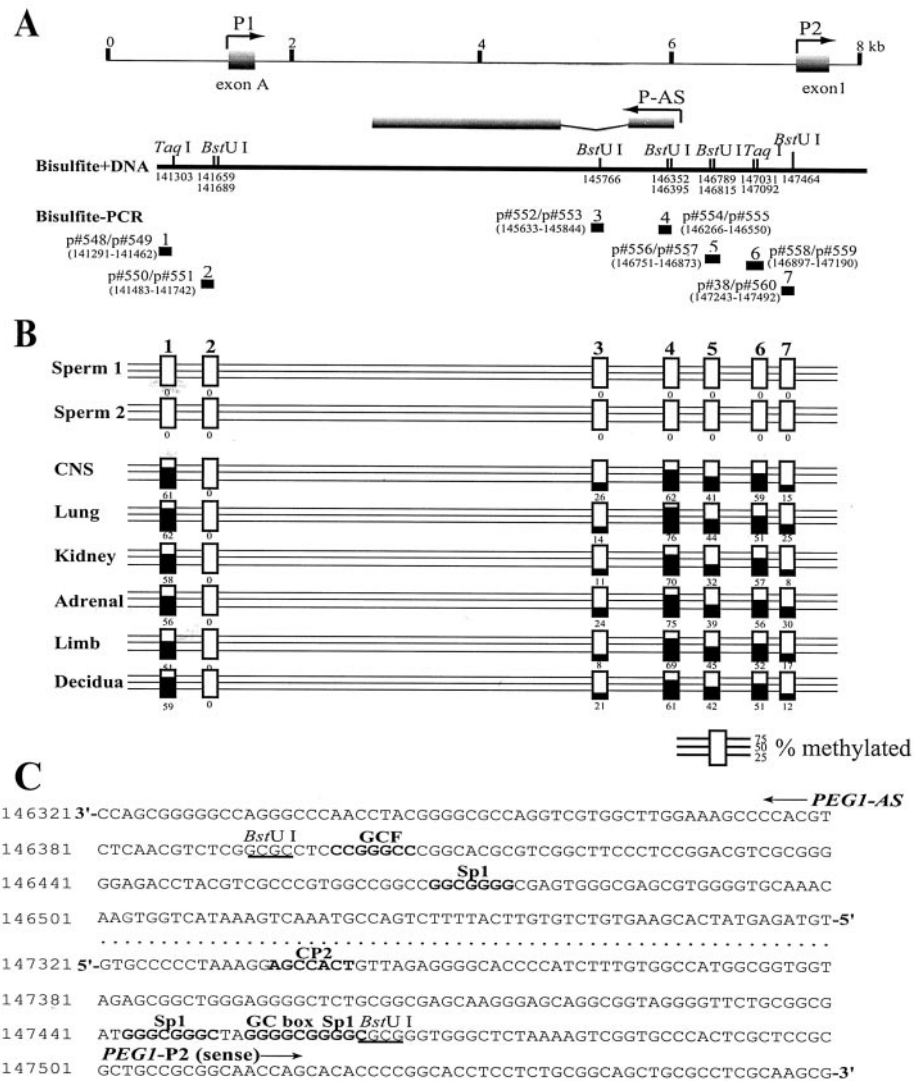

FIG. 6. **Methylation profile of the human *PEG1* gene in sperm, fetal tissues, and maternal decidua.** *A*, map of the P1-P2 and *PEG1-AS* region of the *PEG1* gene. The transcripts from P1, P2, and *PEG1-AS* are indicated by *arrows*. A partial map of the *Bst*UI and *Taq*I sites (site location, nt) on the bisulfite-modified DNA is shown. Seven PCR amplicons (*1–7*) were designed to amplify the modified DNA encompassing these sites. The primer numbers (sequence described under "Experimental Procedures") and the nucleotide location (GenBank™ accession number AC007938) are indicated. *B*, methylation profile of the human *PEG1* gene. The seven PCR products (*1–7*) were digested with *Bst*UI (or *Taq*I) and analyzed on a 5% polyacrylamide-urea gel. Methylated CpG is not modified by bisulfite, and therefore, it remains as a CpG sequence and is digested. Methylation at each site was represented as the percent of the digested bands to the total of the digested and undigested bands quantified by a PhosphorImager. *White box*, unmethylated site, 0%. *Partially filled box*, partially methylated. The *number* at the *bottom* of each *box* represents % of CpG methylation. *C*, putative transcription factor motifs in the *PEG1-AS* and the P2 region (GenBank™ accession number AC007938, nt 146321–147560). One GC box (GGGGCGGGGC), two Sp1 binding sites (GGGCGGGC and GGGGCGGGGC) and one CP2 site (AGCCACT) were observed at ~50 and ~180 bp upstream of the *PEG1* P2 sense transcript, respectively. One GCF transcriptional regulator (CCGGGCC) and a Sp1 binding site (GGGGCGG) were found at ~20 and ~90 bp upstream of the *PEG1-AS* promoter, respectively. Methylation of the *Bst*UI sites (*underlined*) adjacent to the GCF (*PEG1-AS*) or the GC box (*PEG1*-P2 sense) was shown in *B*, *boxes 4* and *7*, respectively. Note that the sequence of the *PEG1-AS* was displayed from 3′ to 5′.

exon 1 and the 0.5-kb P1 region (Fig. 6*A*). The results are shown as a string of boxes with each box representing the methylation status of CpG (CGCG, *Bst*UI and TCGA, *Taq*I) in each PCR fragment. As shown in Fig. 6, sperm DNAs were completely unmethylated in all seven PCR fragments (*white boxes*). DNAs from various fetal tissues and from maternal decidua were partially methylated in the 2.5-kb region (*boxes 3–7*), which contains a DMR. The promoter region of P1 (exon A) was completely unmethylated in all tissues (Fig. 6*B*, *box 2*), consistent with biallelic expression from P1 (*isoform 2*), whereas the region farther upstream of P1 was, unexpectedly, partially methylated (*box 1*).

We analyzed the 1.2-kb putative promoter sequence in the *PEG1-AS* and the P2 region (GenBank™ accession number AC007938, nt 146321–147560) for potential transcription factor motifs (Fig. 6*C*). Although TATA and CAAT sequence motifs were absent in the promoter region, a GC box (GGGGCGGGGC), two Sp1 binding sites (GGGCGGGC and GGGGCGGGGC), and a CP2 site (AGCCACT, a developmental signal response factor) were found at ~50 and ~180 bp, respectively, upstream of the *PEG1* P2 sense transcript. In the *PEG1-AS* promoter region, a GCF (CCGGGCC, a GC-rich binding transcriptional regulator) and an Sp1 binding site (GGGGCGG) were found at ~20 and ~90 bp, respectively, upstream (the sequence of the *PEG1-AS* is displayed from 3′ to 5′). The methylation status of the GCF (*PEG1-AS*) and the GC-box/Sp1 sites (*PEG1* P2 sense) is depicted in the amplicon boxes 4 and 7, respectively (Fig. 6*B*). In all tissues except for sperm, the GCF site (*Bst*UI, nt 146385, *box 4*) was hypermethylated (61–76% methylation), whereas the GC-box/Sp1 site (*Bst*UI, nt 147464, *box 7*) was hypomethylated (8–30% methylation).

Bisulfite Sequencing of a DMR within PEG1-AS and PEG1 P2 Promoter Region—We further analyzed the methylation status of all CpG sites in box 5 and box 6 by bisulfite sequencing. Primers p#556 and p#559 encompassing the DNA region of box 5–6 were used to amplify bisulfite-treated genomic DNAs. The 450-bp PCR fragments were cloned and sequenced by standard methods (11). The methylation status of 25 CpGs in these PCR fragments (GenBank™ accession number AC007938, nt 146751–47190) was depicted as a string of *open* and *filled beads* (Fig. 7). The MR-PCR assay indicated partial methylation in an embryonic lung DNA (44 and 51% methylation in *box 5* and *box 6*, respectively, Fig. 6*B*). A pair of lung tumors and normal adjacent tissues that demonstrated a silencing of P2 transcript in the tumor (*versus* normal) tissue was hypermethylated in the tumor (35% methylation in normal adjacent tissues and 70% in tumor, by MR-PCR assay). Bisulfite sequencing revealed the hypermethylation of all 25 CpG sites in the lung tumor (Fig. 7). A fibroblast Hs 1 cell line from normal embryonic testis that had detectable P2 sense and *PEG1-AS* transcripts was hypomethylated at CpG 1–11 sites, whereas the Hep 3B cell line (undetectable P2 sense and *PEG1-AS* transcripts) was hypermethylated at all CpG sites

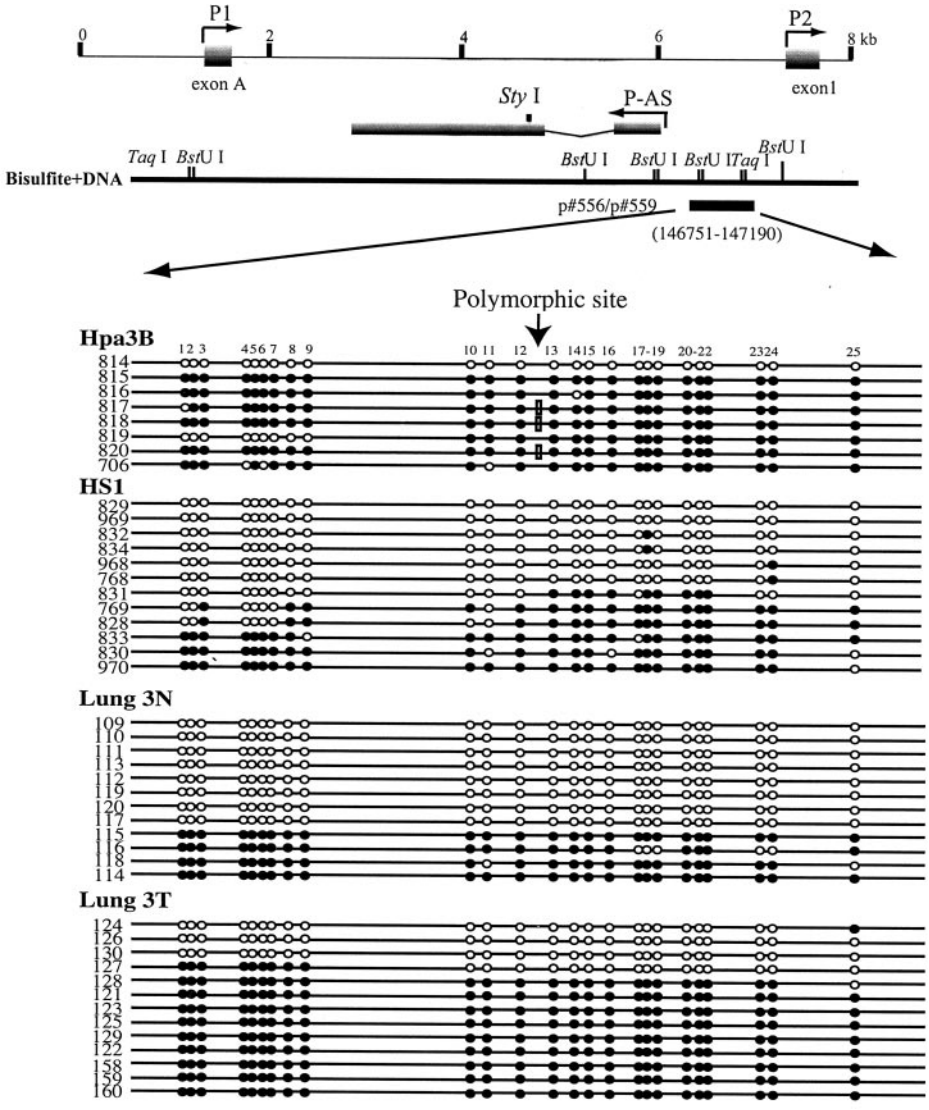

FIG. 7. **CpG methylation of the DMR region.** *Top*, partial map of the human *PEG1* gene showing exon A and exon 1 and the *PEG1-AS* transcript. *Bottom*, CpG methylation status. The 450-bp PCR fragments (GenBank™ accession number AC007938, position 146751–147190) were amplified with primers p#556 and p#559, cloned, and sequenced. Each *line* represents an individual clone. Methylated CpGs are represented by *filled beads*, and unmethylated CpGs are represented by *open beads*. *Open boxes* indicate the single nucleotide polymorphism site (nt 146966 A → G) in Hep 3B cell line.

(Fig. 7). A polymorphic site in the Hep 3B cells separated the two parental alleles, which suggests the spreading of methylation in this region of the normally unmethylated parental allele.

CpG Methylation in the DMR Abolishes Promoter Activity of PEG1-AS and PEG1 P2—To assay the DMR within the *PEG1-AS* and P2 promoter exon for promoter activity we inserted various PCR fragments in both orientations (sense and antisense) into the pGL3-basic luciferase reporter vector and transfected them into the NCI-H23 cells. The H23 cell line was chosen because it provided high transfection efficiency and high luciferase activity with these constructs. Normalized luciferase activity (against pSV40β-gal internal control) increased ~23-fold with antisense construct 1 containing 287 bp of the putative antisense promoter. This fragment contained the putative GCF and Sp1 sites (Fig. 6*C*). The promoter activity increased to ~50-fold with longer constructs (362 and 417 bp) but leveled off with fragments longer than 674 bp (Fig. 8, *A* and *B*). In contrast, the P2 sense constructs demonstrated higher promoter activity (~97-fold with 309 bp upstream of the transcription site). The luciferase activity of the sense construct increased to 165-fold with a 1.5-kb construct that encompassed the antisense promoter (in a reverse orientation).

To determine whether methylation in the DMR region affects the promoter activity, the antisense constructs 3, 4, and 6 and the P2 sense constructs 8 and 9 were methylated with *Sss*I methylase, which methylates all CpG sites within the constructs (including the luciferase reporter plasmid). The luciferase activity of these methylated constructs was reduced drastically to the control pGL3-basic level in the transient transfection assay (Fig. 8*B*). To verify that CpG methylation within the DMR core region but not the methylation in the reporter plasmid abolished the luciferase activity, the purified 674 bp (antisense construct 4) and 580 bp (P2 sense construct 8) fragments were methylated before ligation to the unmethylated pGL3 luciferase plasmid. The two DMR-specific methylated plasmids also exhibited luciferase activity at 0.1 and 0.3 times the level of pGL3-basic.

DISCUSSION

Imprinting and LOI of PEG1 Sense Transcripts—We have analyzed both sense and antisense transcripts from the human *PEG1* gene on chromosome 7q32. The locus is clustered with another gene, *γ2-COP* (COPG2, coatomer protein complex), that overlaps with one part of exon 12 in the 3′-untranslated region of the *PEG1* gene in a tail-to-tail orientation (24). Conflicting reports concerning the imprinted status of *COPG2* might be explained by the partial overlap of the two adjacent genes (24). To add further complexity to this region, *C1T1*, an antisense of *COPG2*, originates from intron 20 of *COPG2* and is imprinted and paternally expressed in both human and mouse (24, 25).

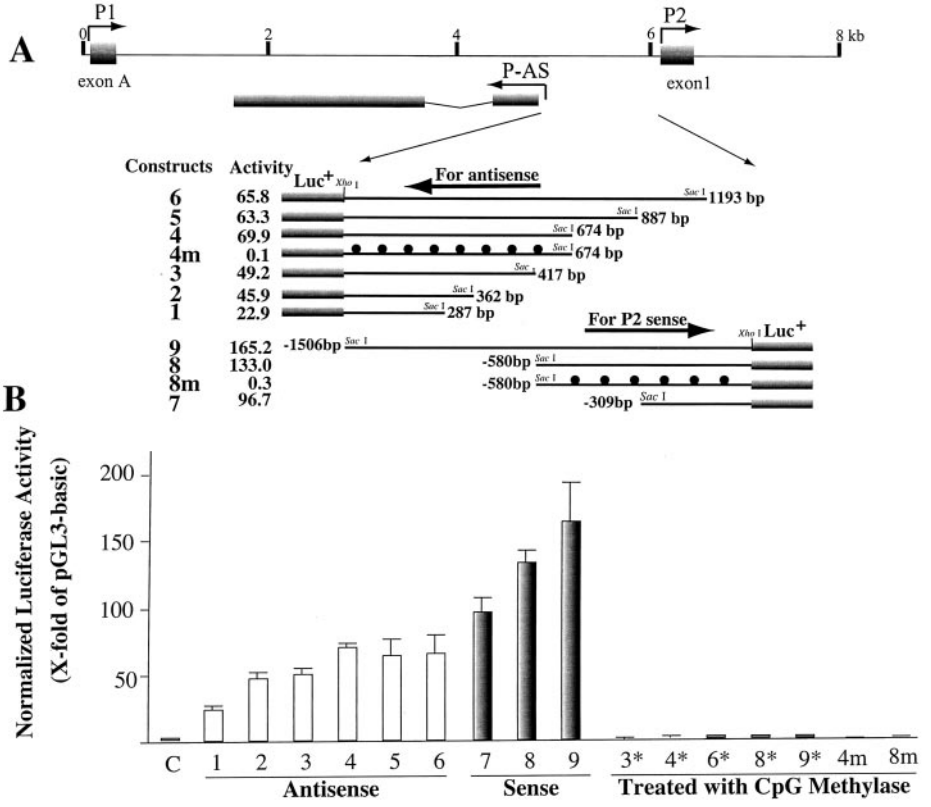

FIG. 8. **Promoter activity by luciferase reporter assay.** *A*, diagrams of *PEG1* P2 sense and *PEG1-AS* genomic fragments and the promoter-luciferase constructs. Both the *PEG1* P2 sense and *PEG1-AS* fragments were PCR-amplified from same DNA region. Constructs 1–6 were *PEG1-AS* promoter constructs, and constructs 7–9 were P2 sense constructs. Constructs 4m and 8m had the target DNA fragments methylated by *Sss*I methylase before ligation to the reporter vector. *B*, luciferase activity assay of *PEG1* P2 sense and *PEG1-AS* constructs. The normalized luciferase activity is represented as fold of control vector pGL3-basic. Constructs 3*, 4*, 6*, 8*, and 9* were methylated (both target DNA and reporter gene) before transfection. Constructs 4m and 8m had only the target DNA fragments methylated. The data (mean ± S.E.) are derived from three separate transfection experiments in which triplicate wells were assayed.

Our report focuses on transcripts expressed from the 5′ region of the *PEG1* exon 1 that lies 112 kb upstream from the overlapping region and, therefore, are not likely to be associated with the *COPG2* transcription unit. We have confirmed expression of *PEG1* sense transcripts from the two promoters P1 and P2 in various tissues and have characterized a novel imprinted antisense transcript *PEG1-AS* located within the two promoter exons. We have verified the promoter activity of a region between *PEG1-AS* and *PEG1*-P2 sense using luciferase reporter constructs, and we have demonstrated that methylation of the DMR by CpG methylase abolished the promoter activity in both sense and antisense orientations. We suggest that the DMR contains an imprinting element(s) of both the *PEG1-AS* and *PEG1* P2 sense transcripts. Both the sense and antisense transcripts are transcribed from the same paternal allele, and both share the same DMR in the promoter region (head-to-head conformation). Despite sharing the 1.5-kb promoter region, although in opposite orientation, the *PEG1-AS* and the P2 sense transcripts differ in their pattern of tissue-specific expression.

We also have confirmed the imprinting of P2 (isoform 1) and biallelic expression of P1 (data not shown) and reported the splicing variants from P1. The simultaneous contribution of the imprinted and non-imprinted promoters to the sense transcription of *PEG1* complicates the analysis of its allelic expression. In fact, leaky expression from the maternal allele was observed in the human *PEG1* (4) but not in the mouse *Peg1* (1), which has only one promoter (5). It is possible that the biallelic *PEG1* expression (LOI) in human lymphobastoid cells is due to switching promoter usage to the unmethylated promoter P1 (5, 12). Altering the imprinting status to LOI by switching promoter usage to the non-imprinted promoter P1 was also observed in the human *IGF2* gene in human adult liver (13). Interestingly, as in the case of the mouse *Peg1*, the mouse *Igf2* gene also has no homologous human P1 and shows no LOI in mouse liver.

LOI of human *IGF2* was first observed in Wilms' tumor (26, 27) and, subsequently, in many other tumors. The mechanisms of LOI of *IGF2* in tumors have been ascribed to the coordinate LOI of all three imprinted *IGF2* promoters (9, 28). In contrast, LOI of the human *PEG1* has only been observed in invasive breast cancer (7) and in colorectal mucosa, in which both *PEG1* and *IGF2* showed LOI (8). In our present study, the switching of promoter usage to P1 in various tumors may indicate the common occurrence of LOI of *PEG1* in tumorigenesis and, furthermore, implicate promoter usage switching as the mechanism of LOI for the *PEG1* gene in tumorigenesis. Methylation of the P2 promoter region during tumorigenesis is the likely mechanism underlying the switching of promoter usage since the methylated promoter construct demonstrated no activity in the luciferase reporter assay (Fig. 8).

Imprinting and LOI of PEG1-AS in Germ Cells—We have shown that both *PEG1-AS* and the sense transcript from promoter P2 are derived exclusively from the paternal allele and are likely to be governed by a common imprinting element(s) in the DMR encompassing exon 1 (Fig. 6). In an analogous mouse (and human) *Gnas* locus, a common DMR for both sense (*Xlas*) and antisense (*Gnas-as* or *Nespas*) transcripts is located in a ~2-kb region encompassing the *Xlas* exon (15, 16). In both *PEG1* and *Gnas* genes, methylation of the maternal DMR corresponds to the silencing of the maternal allele in both sense and antisense transcripts.

In gametogenesis, the observation of early erasure of methylation and the maintenance of the unmethylated status of both allelic DMRs in embryonic germ cells may indicate biallelic expression of *PEG1-AS* in spermatocytes. Consistent with the unmethylated status of biallelic DMRs of the *PEG1*gene in male germ cells, we observed biallelic expression of the *PEG1* sense transcript in a population of mature sperm from a single donor. Biallelic expression of *IGF2*, a paternally expressed gene, was demonstrated in pachytene spermatocytes in adult testis (29, 30). It is likely that other *PEG* genes would also demonstrate LOI (biallelic expression) in spermatocytes (due to unmethylated DMRs) and would be silenced in the female germ

cell (due to methylated DMRs). The converse prediction would be that maternally expressed genes (*MEG*) would be biallelically expressed in a population of female germ cells and would be silenced in the male germ cells.

Functional PEG1-AS?—We have demonstrated that *PEG1-AS* is an abundant transcript of ~2.4 kb in human testis by Northern blotting using a strand-specific probe. This is in contrast to other imprinted antisense transcripts that are difficult to detect by Northern blotting due to low levels of expression and/or heterogeneity in the transcript sizes (for example, *Gnas-as* or *Nespas*) (16, 18). In contrast to a majority of imprinted antisense genes with continuous, unprocessed transcripts, the *PEG1-AS* gene has an intronic sequence. *PEG1-AS* contains multiple small open reading frames but with weak Kozak consensus sequences for initiation of translation. Therefore, *PEG1-AS* is likely to function as a non-coding RNA. Non-coding RNA is a distinct form of RNA that differs from the other three kinds of common RNA (mRNA, ribosomal RNA, and transfer RNA), which are all components of protein synthesis pathways. Non-coding RNA may act on a second RNA template in splicing, editing, and post-transcriptional modifications or act on a DNA template to extend its telomere (telomerase activity). Recent studies indicate a novel function of non-coding RNA in imprinted regions (for a recent review, see Kelley and Kuroda (31)). Antisense and non-coding transcripts in imprinted regions may act in *cis* near the transcription site. Antisense RNAs could function as chromosomal RNA by remodeling (or repackaging) local chromatin structure or by capturing or interfering with the transcription of the sense mRNA transcript.

It is known that oocytes contain large amounts of maternal RNA that are actively transcribed in meiotic prophase I. In parallel, during spermatogenesis transcription may be active in primary spermatocytes. The RNA transcripts are associated with ribonucleoproteins and are well preserved through meiosis, a process that takes about 24 days in man. During spermatogenesis, a 9-week process in man, newly synthesized proteins from the preserved RNA stock help to transform the spermatid to a mature sperm. At the final stage of maturation, spermatozoa pinch off most of their cytoplasm containing the remaining RNA stock. It is intriguing therefore to find abundant RNA transcripts of *PEG1-AS* in mature sperm.

Our finding of both parental transcripts of *PEG1-AS* gene in a population of mature sperm from a single donor has two alternative explanations. One possibility is that regardless of the parental origin of the *PEG1-AS* allele in each haploid sperm, there are transcripts from one allele in each sperm, which results in equal expression of both alleles when we study a sperm population. Another possibility is that both parental *PEG1-AS* alleles are expressed in the male germ cells, and both are retained in each mature sperm. It will be necessary to study *PEG1-AS* allelic expression in individual sperm by allele-specific RNA-fluorescence *in situ* hybridization, a powerful technique that has been used to reveal changes in *Igf2* imprinting during pre-implantation mouse development (30). Although the function of RNA transcripts in mature sperm is unknown, the presence of the unique and abundant *PEG1-AS* RNA in motile spermatozoa may argue for its role in sperm physiology and/or fertility. This may be a novel function of a non-coding RNA in human fertilization.

Acknowledgment—We thank the Central Laboratory for Human Embryology Tissue, University of Washington, Seattle, for fetal tissues.

REFERENCES

1. Kaneko-Ishino, T., Kuroiwa, Y., Miyoshi, N., Kohda, T., Suzuki, R., Yokoyama, M., Viville, S., Barton, S. C., Ishino, F., and Surani, M. A. (1995) *Nat. Genet.* **11,** 52–59
2. Lefebvre, L., Viville, S., Barton, S. C., Ishino, F., Keverne, E. B., and Surani, M. A. (1998) *Nat. Genet.* **20,** 163–169
3. Li, L., Keverne, E. B., Aparicio, S. A., Ishino, F., Barton, S. C., and Surani, M. A. (1999) *Science* **284,** 330–333
4. Kobayashi, S., Kohda, T., Miyoshi, N., Kuroiwa, Y., Aisaka, K., Tsutsumi, O., Kaneko-Ishino, T., and Ishino, F. (1997) *Hum. Mol. Genet.* **6,** 781–786
5. Kosaki, K., Kosaki, R., Craigen, W. J., and Matsuo, N. (2000) *Am. J. Hum. Genet.* **66,** 309–312
6. Riesewijk, A. M., Hu, L., Schulz, U., Tariverdian, G., Hoglund, P., Kere, J., Ropers, H. H., and Kalscheuer, V. M. (1997) *Genomics* **42,** 236–244
7. Pedersen, I. S., Dervan, P. A., Broderick, D., Harrison, M., Miller, N., Delany, E., O'Shea, D., Costello, P., McGoldrick, A., Keating, G., Tobin, B., Gorey, T., and McCann, A. (1999) *Cancer Res.* **59,** 5449–5451
8. Nishihara, S., Hayashida, T., Mitsuya, K., Schulz, T. C., Ikeguchi, M., Kaibara, N., and Oshimura, M. (2000) *Int. J. Oncol.* **17,** 317–322
9. Vu, T. H., and Hoffman, A. (1996) *J. Biol. Chem.* **271,** 9014–9023
10. Xiong, Z., and Laird, P. W. (1997) *Nucleic Acids Res.* **25,** 2532–2534
11. Vu, T. H., Li, T., Nguyen, D., Nguyen, B. T., Yao, X. M., Hu, J. F., and Hoffman, A. R. (2000) *Genomics* **64,** 132–143
12. Cuisset, L., Le Stunff, C., Dupont, J. M., Vasseur, C., Cartigny, M., Despert, F., Delpech, M., Bougnere, P., and Jeanpierre, M. (1997) *Ann. Genet.* **40,** 211–215
13. Vu, T. H., and Hoffman, A. R. (1994) *Nature* **371,** 714–717
14. Ekstrom, T. J., Cui, H., Li, X., and Ohlsson, R. (1995) *Development* **121,** 309–316
15. Hayward, B. E., and Bonthron, D. T. (2000) *Hum. Mol. Genet.* **9,** 835–841
16. Wroe, S. F., Kelsey, G., Skinner, J. A., Bodle, D., Ball, S. T., Beechey, C. V., Peters, J., and Williamson, C. M. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 3342–3346
17. Okutsu, T., Kuroiwa, Y., Kagitani, F., Kai, M., Aisaka, K., Tsutsumi, O., Kaneko, Y., Yokomori, K., Surani, M. A., Kohda, T., Kaneko-Ishino, T., and Ishino, F. (2000) *J. Biochem.* (*Tokyo*) **127,** 475–483
18. Li, T., Vu, T. H., Zeng, Z. L., Nguyen, B. T., Hayward, B. E., Bonthron, D. T., Hu, J. F., and Hoffman, A. R. (2000) *Genomics* **69,** 295–304
19. Miller, D. (2000) *Mol. Reprod. Dev.* **56,** 259–264
20. Goodwin, L. O., Karabinus, D. S., and Pergolizzi, R. G. (2000) *Mol. Hum. Reprod.* **6,** 487–497
21. Davis, T. L., Yang, G. J., McCarrey, J. R., and Bartolomei, M. S. (2000) *Hum. Mol. Genet.* **9,** 2885–2894
22. Kerjean, A., Dupont, J. M., Vasseur, C., Le Tessier, D., Cuisset, L., Paldi, A., Jouannet, P., and Jeanpierre, M. (2000) *Hum. Mol. Genet.* **9,** 2183–2187
23. Lefebvre, L., Viville, S., Barton, S. C., Ishino, F., and Surani, M. A. (1997) *Hum. Mol. Genet.* **6,** 1907–1915
24. Yamasaki, K., Hayashida, S., Miura, K., Masuzaki, H., Ishimaru, T., Niikawa, N., and Kishino, T. (2000) *Genomics* **68,** 330–335
25. Lee, Y. J., Park, C. W., Hahn, Y., Park, J., Lee, J., Yun, J. H., Hyun, B., and Chung, J. H. (2000) *FEBS Lett.* **472,** 230–234
26. Feinberg, A. P. (1993) *Nat. Genet.* **4,** 110–113
27. Ogawa, O., Eccles, M. R., Szeto, J., McNoe, L. A., Yun, K., Maw, M. A., Smith, P. J., and Reeve, A. E. (1993) *Nature* **362,** 749–751
28. Taniguchi, T., Sullivan, M. J., Ogawa, O., and Reeve, A. E. (1995) *Proc. Natl. Acad. Sci. U. S. A.* **92,** 2159–2163
29. Szabo, P. E., and Mann, J. R. (1995) *Genes Dev.* **9,** 1857–1868
30. Ohno, M., Aoki, N., and Sasaki, H. (2001) *Genes Cells* **6,** 249–259
31. Kelley, R. L., and Kuroda, M. I. (2000) *Cell* **103,** 9–12