GENOMICS 7, 491–502 (1990)

cDNA Cloning, mRNA Distribution and Heterogeneity, Chromosomal Location, and RFLP Analysis of Human Osteopontin (OPN)

MARIAN F. YOUNG,* JANET M. KERR,* JOHN D. TERMINE,* ULLA M. WEWER,† MARY GE WANG,‡ O. WESLEY MCBRIDE,‡ AND LARRY W. FISHER*

†*University Institute of Pathological Anatomy, Copenhagen, Denmark, and *National Institute of Dental Research, and ‡National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892*

Received October 6, 1989; revised March 27, 1990

A human osteopontin (OP) cDNA was isolated from a library made from primary cultures of human bone cells. The distribution of osteopontin mRNA in human tissues was investigated by Northern analysis and showed that the human message was predominant in cultures of bone cells and in decidua cells isolated at 6–12 weeks of gestation. Immunohistochemical analysis confirmed that OP expression is high in decidua cells as well as in the endometrial glands of a nonpregnant secretory-phase human uterus. Two variants of the OP message were evident on the basis of DNA sequencing and polymerase chain reaction amplification of bone and decidua cell mRNA. The peptides potentially translated by the variant messages differ by the presence (OP1b) or absence (OP1a) of 14 amino acids at residue 58 of the molecule. The deduced human protein sequence shows a conservation between species in the position of the Arg–Gly–Asp (RGD) cell attachment site. Chromosomal mapping of the osteopontin gene (OPN) using human–rodent cell hybrids demonstrated a location on chromosome 4 in the human genome. *In situ* hybridization of metaphase chromosomes using radiolabeled OP1a as a probe indicated that the gene is located on a region of 4q that is near the centromere. A high-frequency restriction fragment length polymorphism was evident in the DNA from 29 unrelated individuals using the enzyme *Bgl*II. Analysis of total genomic DNA by digestion with several restriction enzymes, Southern blotting, and hybridization with the human osteopontin cDNA indicated that the gene is a single copy with an approximate length of 5.4–8.2 kb. © 1990 Academic Press, Inc.

Sequence data from this article have been deposited with the EMBL/GenBank Data Libraries under Accession No. J04765.

INTRODUCTION

Osteopontin (Opn/2ar/pp69/Spp) is a phosphorylated glycoprotein that is found in rat, porcine, bovine, and human bone (for review see Butler, 1989). Localization of the mRNA (Nomura *et al.*, 1988; Yoon *et al.*, 1987) and protein (Mark *et al.*, 1987, 1988) in mouse and in rat embryos indicates that the gene is expressed in early stages of bone formation by some preosteoblasts and osteoblasts, and in some cells of the marrow. Certain nonbone tissues also appear to express OPN and include the kidney, the sensory epithelium of the embryonic ear, and the metrial gland of the placental decidua. Rat (Oldberg *et al.*, 1986) and human (Somerman *et al.*, 1988) forms of the protein enhance cell adhesion of transformed osteoblasts and gingival fibroblasts, respectively. The cell attachment activity is thought to be mediated by a cellular receptor with high affinity for the amino acid sequence Arg–Gly–Asp (RGD) located midway in the protein (Oldberg *et al.*, 1986). Transcription of the osteopontin gene is sensitive to several environmental factors such as TGF-b (Noda *et al.*, 1988), 1,25-dihydroxyvitamin D_3 (Prince and Butler, 1987), h-*ras* (Craig *et al.*, 1988), and the tumor promoter 12-*O*-tetradecanoylphorbol-13-acetate (Craig *et al.*, 1989). Because secretion of osteopontin is enhanced in a wide variety of transformed mammalian cells it has been implicated as a marker of tumorigenesis (Senger and Perruzzi, 1985). In this report we describe the isolation, mRNA distribution, and location of the OPN gene in the human genome.

MATERIALS AND METHODS

cDNA Library Construction and Screening

A human λ Zap (Stratagene) cDNA library was constructed using poly(A)$^+$ mRNA isolated from primary

0888-7543/90 $3.00
Copyright © 1990 by Academic Press, Inc.
All rights of reproduction in any form reserved.

```
1     GAG GCA GCA GCA GCA GGA GGA GGC AGA GAC AGC ATC GTC GGG ACC AGA CTC GTC TCA GGC

61    CAG TTG CAG CCT TCT CAG CCA AAC GCC GAC CAA GGA AAA CTC ACT ACC ATG AGA ATT GCA
                                                                      Met Arg Ile Ala

121   GTG ATT TGC TTT TGC CTC CTA GGC ATC ACC TGT GCC ATA CCA GTT AAA CAG GCT GAT TCT
4     Val Ile Cys Phe Cys Leu Leu Gly Ile Thr Cys Ala Ile Pro Val Lys Gln Ala Asp Ser

181   GGA AGT TCT GAG GAA AAG CAG CTT TAC AAC AAA TAC CCA GAT GCT GTG GCC ACA TGG CTA
25    Gly Ser Ser Glu Glu Lys Gln Leu Tyr Asn Lys Tyr Pro Asp Ala Val Ala Thr Trp Leu

241   AAC CCT GAC CCA TCT CAG AAG CAG AAT CTC CTA GCC CCA CAG ACC CTT CCA AGT AAG TCC
45    Asn Pro Asp Pro Ser Gln Lys Gln Asn Leu Leu Ala Pro Gln Thr Leu Pro Ser Lys Ser

301   AAC GAA AGC CAT GAC CAC ATG GAT GAT ATG GAT GAT GAA GAT GAT GAT GAC CAT GTG GAC
65    Asn Glu Ser His Asp His Met Asp Asp Met Asp Asp Glu Asp Asp Asp Asp His Val Asp

361   AGC CAG GAC TCC ATT GAC TCG AAC GAC TCT GAT GAT GTA GAT GAC ACT GAT GAT TCT CAC
85    Ser Gln Asp Ser Ile Asp Ser Asn Asp Ser Asp Asp Val Asp Asp Thr Asp Asp Ser His

421   CAG TCT GAT GAG TCT CAC CAT TCT GAT GAA TCT GAT GAA CTG GTC ACT GAT TTT CCC ACG
105   Gln Ser Asp Glu Ser His His Ser Asp Glu Ser Asp Glu Leu Val Thr Asp Phe Pro Thr

480   GAC CTG CCA GCA ACC GAA GTT TTC ACT CCA GTT GTC CCC ACA GTA GAC ACA TAT GAT GGC
125   Asp Leu Pro Ala Thr Glu Val Phe Thr Pro Val Val Pro Thr Val Asp Thr Tyr Asp Gly

540   CGA GGT GAT AGT GTG GTT TAT GGA CTG AGG TCA AAA TCT AAG AAG TTT CGC AGA CCT GAC
145   Arg Gly Asp Ser Val Val Tyr Gly Leu Arg Ser Lys Ser Lys Lys Phe Arg Arg Pro Asp

601   ATC CAG TAC CCT GAT GCT ACA GAC GAG GAC ATC ACC TCA CAC ATG GAA AGC GAG GAG TTG
165   Ile Gln Tyr Pro Asp Ala Thr Asp Glu Asp Ile Thr Ser His Met Glu Ser Glu Glu Leu

661   AAT GGT GCA TAC AAG GCC ATC CCC GTT GCC CAG GAC CTG AAC GCG CCT TCT GAT TGG GAC
185   Asn Gly Ala Tyr Lys Ala Ile Pro Val Ala Gln Asp Leu Asn Ala Pro Ser Asp Trp Asp

721   AGC CGT GGG AAG GAC AGT TAT GAA ACG AGT CAG CTG GAT GAC CAG AGT GCT GAA ACC CAC
205   Ser Arg Gly Lys Asp Ser Tyr Glu Thr Ser Gln Leu Asp Asp Gln Ser Ala Glu Thr His

781   AGC CAC AAG CAG TCC AGA TTA TAT AAG CGG AAA GCC AAT GAT GAG AGC AAT GAG CAT TCC
225   Ser His Lys Gln Ser Arg Leu Tyr Lys Arg Lys Ala Asn Asp Glu Ser Asn Glu His Ser

841   GAT GTG ATT GAT AGT CAG GAA CTT TCC AAA GTC AGC CGT GAA TTC CAC AGC CAT GAA TTT
245   Asp Val Ile Asp Ser Gln Glu Leu Ser Lys Val Ser Arg Glu Phe His Ser His Glu Phe

901   CAC AGC CAT GAA GAT ATG CTG GTT GTA GAC CCC AAA AGT AAG GAA GAA GAT AAA CAC CTG
265   His Ser His Glu Asp Met Leu Val Val Asp Pro Lys Ser Lys Glu Glu Asp Lys His Leu

961   AAA TTT CGT ATT TCT CAT GAA TTA GAT AGT GCA TCT TCT GAG GTC AAT TAA AAG GAG AAA
285   Lys Phe Arg Ile Ser His Glu Leu Asp Ser Ala Ser Ser Glu Val Asn

1021  AAA TAC AAT TTC TCA CTT TGC ATT TAG TCA AAA GAA AAA ATG CTT TAT AGC AAA ATG AAA

1081  GAG AAC ATG AAA TGC TTC TTT CTC AGT TTA TTG GTT GAA TGT GTA TCT ATT TGA GTC TGG

1141  AAA TAA CTA ATG TGT TTG ATA ATT AGT TTA GTT TGT GGC TTC ATG GAA ACT CCC TGT AAA

1201  CAA AAG CTT CAG GGT TAT GTC TAT GTT CAT TCT ATA GAA GAA ATG CAA ACT ATC ACT GTA

1261  TTT TAA TAT TTG TTA TTC TCT CAT GAA TAG AAA TTT ATG TAG AAG CAA ACA AAA TAC TTT

1321  TAC CCA CTT AAA AAG AGA ATA TAA CAT TTT ATG TCA CTA TAA TCT TTT GTT TTT TAA GTT

1381  AGT GTA TAT TTT GTT GTG ATT ATC TTT TGT GGT GTG AAT AAA TCT TTT ATC TTG AAT GTA

1441  ATA AGA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AAA AA
```

FIG. 1. DNA and deduced protein sequence of human osteopontin cDNA. DNA sequences are on the top line and translated protein sequences below. The thick underline indicates the location of the analogous NH_2-terminal sequence for a human bone matrix-derived osteopontin (7). The arrows point to potential sites for N-linked glycosylation and the box surrounding the sequence Gly–Arg–Gly–Asp–Ser indicates the presumed site for cell attachment. Dashed lines represent three potential poly(A) adenylation consensus sequences. The solid triangle indicates where 42 additional bp of DNA (14 amino acids) are present in the cDNA reported by Keifer *et al.* (12).

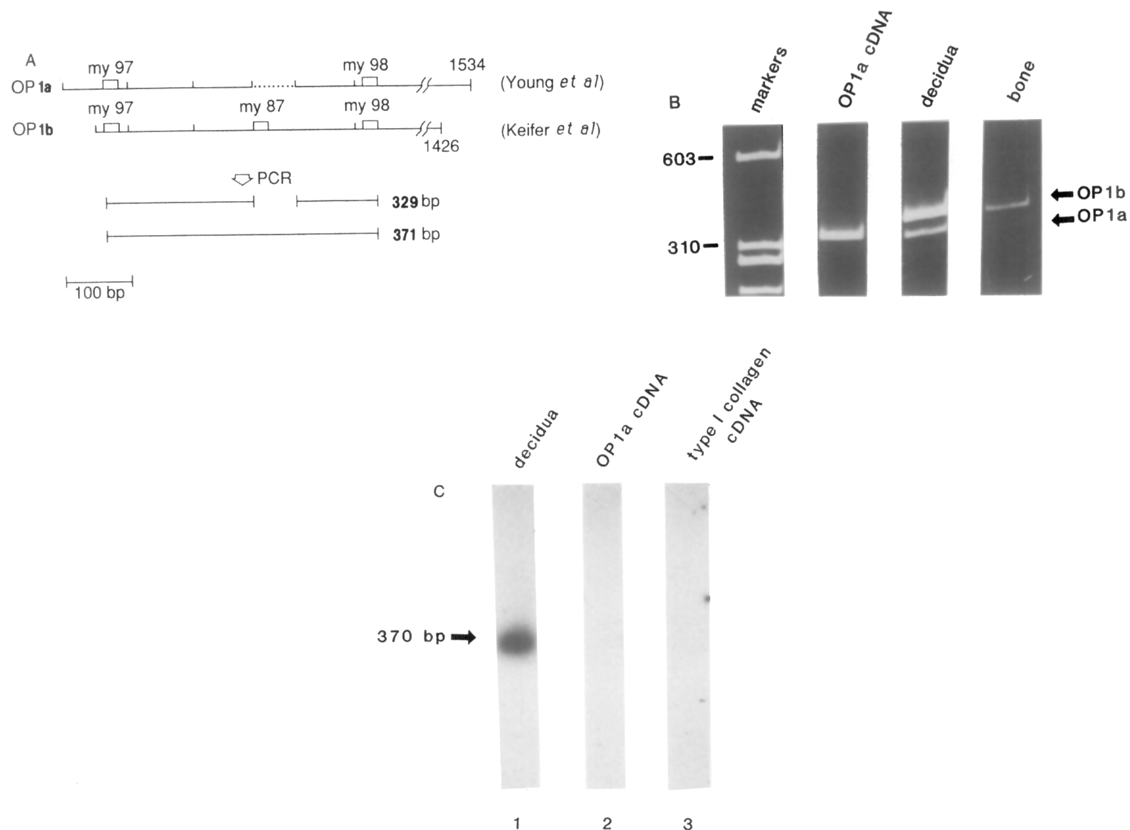

FIG. 2. Polymerase chain reaction (PCR) amplification of osteopontin mRNA sequences. (**A**) Schematic diagram depicting strategy for the amplification. OP1a is the cDNA clone isolated in this lab and OP1b the cDNA isolated by Keifer *et al.* (12). (**B**) Analysis of PCR-amplified material by acrylamide gel electrophoresis and ethidium bromide staining. OP1a cDNA was used as control and size reference for the PCR reaction. (**C**) Hybridization of PCR-amplified decidua mRNA (lane **1**) using oligonucleotide MY87 which contains a sequence encoded by the "gap" region OP1b (see **A**). PCR-amplified OP1a cDNA (lane **2**) and type I collagen cDNA (lane **3**) were used as negative controls.

cultures of human bone cells as described previously (Fisher *et al.*, 1989). The amplified cDNA library was screened using a cDNA encoding the rat osteopontin gene (gift of G. Rodan and K. Yoon, Merck Sharp and Dohme Laboratories, West Point, PA) under low-stringency conditions. Briefly, insert DNA was labeled with ^{32}P by nick-translation (Amersham) and hybridized in a solution containing 0.12% bovine serum albumin, 0.02% Ficoll, 0.02% polyvinylpyrrolidine, 6× SSC (1× is 15 m*M* sodium citrate and 150 m*M* sodium chloride), 0.1% sodium dodecyl sulfate, and 0.1 mg/ml denatured salmon sperm DNA for 16 h at 60°C. Nitrocellulose filters were washed free of unbound probe three times for 15 min with 2× SSC, 0.1% SDS and twice for 10 min with 0.2× SSC, 0.1% SDS at 68°C and autoradiographed at −70°C with an intensifying screen.

Polymerase Chain Reaction Amplification of Osteopontin mRNA Sequences

Approximately 2.0 μg of total bone and decidua cell RNA (see below for details on the source of the cells)

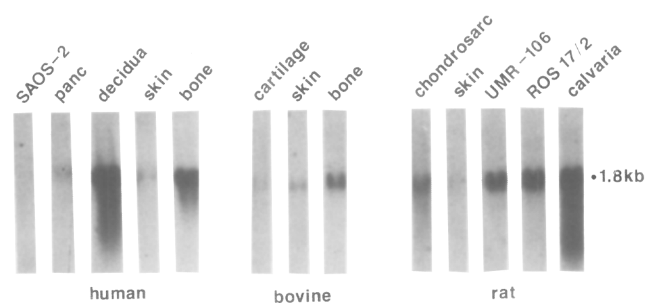

FIG. 3. Northern analysis of RNA from human, bovine, and rat tissues. Insert DNA was isolated from human osteopontin clone OP1a (OP10), radiolabeled with ^{32}P and hybridized to 10 μg of total RNA that had been previously electrophoresed in 1.2% agarose gels and transferred to nitrocellulose. X-ray films were exposed to nitrocellulose for 16 h using intensifying screens. A detailed description of the source of each RNA is outlined under Materials and Methods. The length of osteopontin mRNA is shown to the right of the last lane in kilobases.

was incubated in the presence of oligonucleotide MY98 (5′AGAGTCGTTCGAGTCAATGG3′) and reverse transcriptase using a reaction procedure that has been described previously (Shimokawa *et al.*, 1986). cDNA

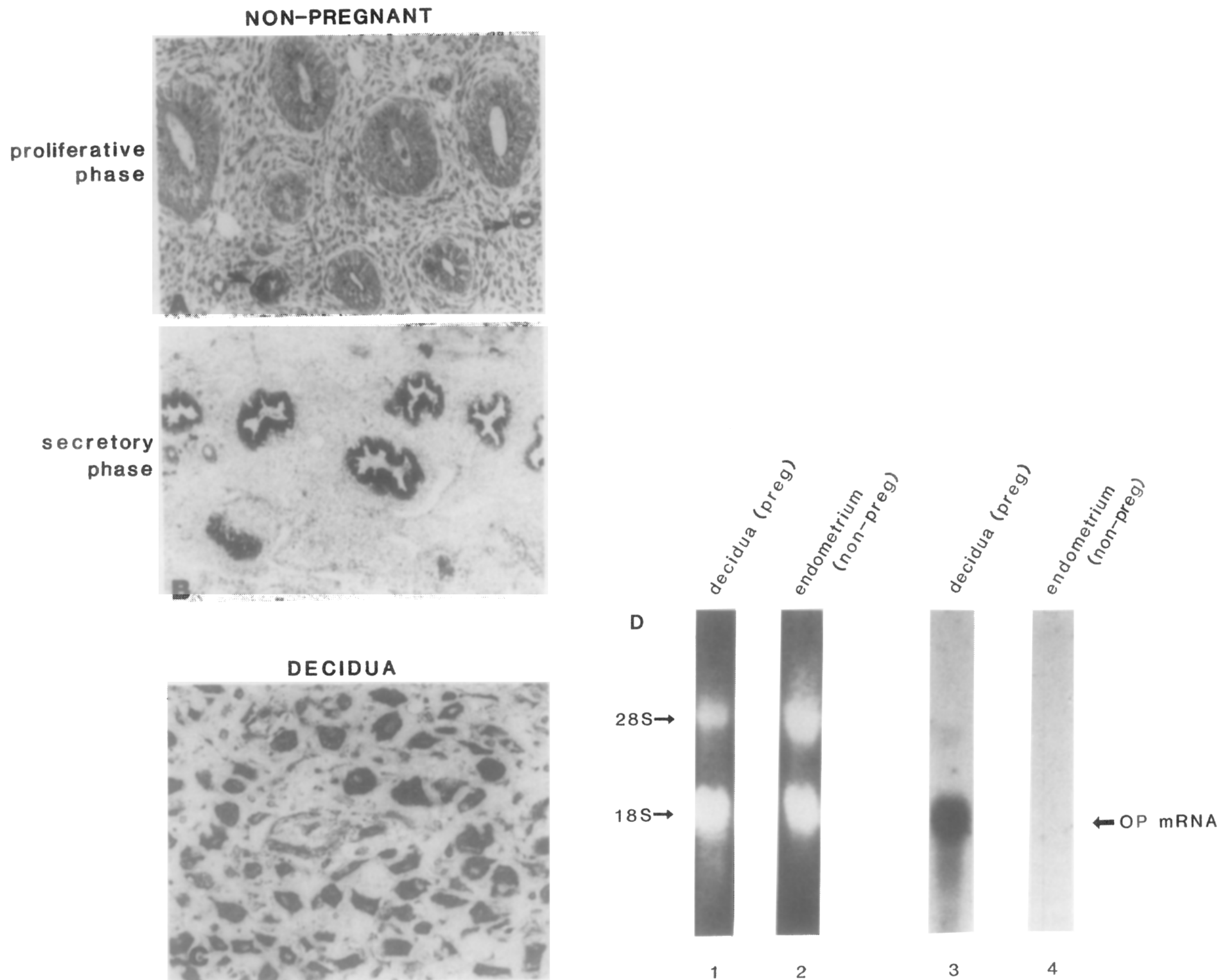

FIG. 4. Osteopontin protein expression in cycling nonpregnant human endometrium and in the decidua (8 weeks of pregnancy). (**A**) Proliferative phase. Endometrial glands and endometrial stromal cells are devoid of immunoreactivity. The blood vessels are slightly immunoreactive (arrows). ×140. (**B**) Secretory phase. Endometrial glands exhibit positive cytoplasmic immunoreactivity while endometrial stromal cells are devoid of immunoreactivity. The blood vessels are slightly immunoreactive. ×64. (**C**) Decidua cells. Intense immunoreactivity is evident in all of the cells. ×200. All sections were counterstained with hematoxylin. (**D**) Osteopontin mRNA expression in proliferative endometrium (see panel **A**) and decidua (panel **C**). Lanes **1** and **2** show ethidium bromide staining of the mRNA bound to nitrocellulose and lanes **3** and **4** the same membrane after hybridization with a ^{32}P-labeled OP1a cDNA probe.

encoding osteopontin was amplified by addition of a second oligonucleotide MY97 (5′CAGTTGCAG-CCTTCTCAGCC3′) using the components of a kit purchased from Perkin–Elmer/Cetus. Annealing was performed at 55°C for 2 min, extension at 72°C, and denaturation at 94°C for 35 cycles with a final extension time of 7 min at 74°C. Reaction products were analyzed by separation in 7% nondenaturing polyacrylamide gels and stained with ethidium bromide. For "in gel" hybridization studies, PCR-generated material was separated in 1.2% agarose gels and denatured in 1.5 *M* NaCl, 0.5 *M* NaOH and neutralized in 0.5 *M* Tris–HCl, 3.0 *M* NaCl as described by Ausubel (1989). An oligonucleotide MY87 (5′GGTTTCCTTCAG-AGGACACAGC3′) was radiolabeled with γ-[^{32}P]ATP (NEN 002) and T4 polynucleotide kinase and incubated with the gel in a solution of 6× SSC, 20 m*M* NaH_2PO_4, pH 7.4, 5× Denhardt's, 0.1% SDS, and 250 μg/ml denatured salmon sperm DNA for 16 h at 55°C. Nonspecifically bound probe was removed from the gel by incubation in 6× SSC, 0.1% SDS at 25°C and then in 6× SSC at 55°C. Specifically bound probe was visualized by autoradiography.

DNA Sequencing

Plaque-purified cDNAs were isolated by standard procedure (Maniatis *et al.,* 1982) and a 1.5-kb *Xho*I–*Xba*I fragment of the OP1a cDNA (plasmid OP10) was subcloned into M13mp18 and M13mp19 (Messing *et*

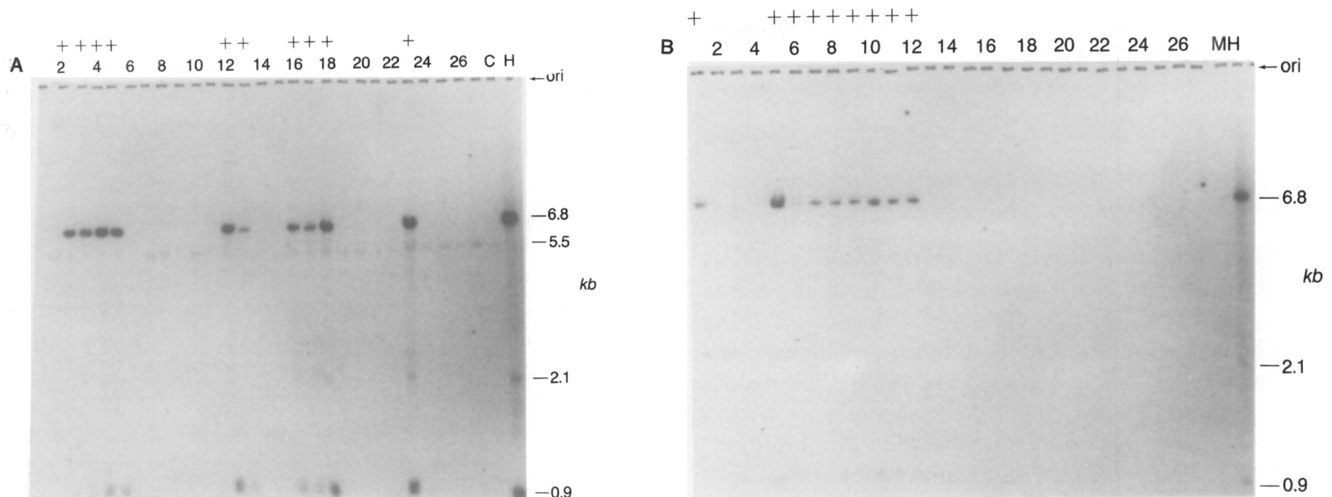

FIG. 5. Southern hybridization of representative human–hamster (**A**) and human–mouse (**B**) somatic cell hybrid DNA *Eco*RI digests with a 0.9-kb *Eco*RI fragment of human osteopontin cDNA. A different hybrid cell DNA is present in each numbered lane. Parental Chinese hamster (C), mouse (M), and human placental (H) DNAs are also shown. The presence of hybridizing human sequences (+) is indicated above the lanes.

al., 1981) vectors and sequenced by the dideoxy chain termination method (Sanger *et al.*, 1977). DNA sequences generated from the −20-bp universal primer (New England Biolabs) were used to construct synthetic oligonucleotides shown in Fig. 1. Additional oligonucleotides were constructed based on the sequences generated from each new set of oligonucleotides until each strand was fully sequenced. DNA and deduced protein sequences were analyzed by the University of Wisconsin Genetics Computer Group (UWGCG).

Northern Blot Analysis

Total RNA was isolated from the human bone cell tumor line SAOS-2 (gift from M. Weiss, University of Pennsylvania, Philadelphia, PA), a human pancreatic tumor line (ATCC CRL 1469), human decidua (6–12 weeks of pregnancy) and nonpregnant endometrial tissues (University of Copenhagen, Denmark), a human dermal skin cell line 1106, and primary cultures of bone cells (Gehron Robey and Termine, 1985) derived from the trabecular bone of a femoral head of human donors undergoing hip replacement (gift of P. Robey, NIDR, NIH, Bethesda, MD). Bovine articular cartilage RNA was obtained from an adult steer after digestion of the tissue with collagenase, and bovine bone and dermal skin fibroblast RNA were isolated from primary cultures derived from a calf at 3–5 months of gestation. A detailed characterization of the phenotype of the bone cell cultures has been described previously (Gehron Robey *et al.*, 1985; Gehron Robey and Termine, 1985). Rat RNA was isolated after collagenase treatment of a Swarm chondrosarcoma, from intact skin and calvaria of 19-day-old rat embryos and from the osteogenic sarcoma cell lines ROS 17/2 and UMR 106-01 (gift of T. J. Martin, St. Vincent's Institute of Medical Research, Australia). Five-microgram aliquots of total RNA were electrophoresed in 1.2% agarose gels containing 2.2 *M* formaldehyde and transferred to nitrocellulose filter paper as described (Maniatis *et al.*, 1982). cDNA isolated free from vector was labeled by nick-translation (Amersham kit) to a specific activity of 10^8 cpm/μg DNA using [^{32}P]dCTP and hybridized to filter-bound DNA for 16 h at 37°C as described (Fisher *et al.*, 1989). Filters were washed three times in 2× SSC and 0.1% SDS at 25°C and analyzed by autoradiography.

Immunohistochemical Analysis of Osteopontin

Human decidual tissue samples (20 cases) were collected at the termination of pregnancy (6–12 weeks) by carefully dissecting away placental and fetal material. Cycling nonpregnant endometrial tissue was investigated by analyzing five cases of proliferative- and five cases of secretory-phase tissue with no obvious pathological changes (tissue was obtained at the University of Copenhagen, Denmark). Tissue specimens were fixed in cold ethanol/acetic acid (99:1, v/v) at 4°C overnight and embedded in paraffin. Deparaffinized and rehydrated sections were processed for immunostaining using the unlabeled peroxidase–antiperoxidase (PAP) technique essentially as described (Wewer *et al.*, 1988). The primary antiserum LF-7 was used which is directed against the intact protein molecule isolated from bone (Fisher *et al.*, 1987). It is not known at this time if the antisera is directed against OP1a, OP1b, or both. Antisera was diluted 1:200 and incubated with the spec-

imens for 4 h at room temperature. The secondary antibodies were used at a dilution of 1:50 for 30 min each. In control sections, the specific antiserum was replaced with preimmune serum.

Chromosome Mapping and Restriction Fragment Length Polymorphism Analysis

A *Xba*I–*Xho*I full-length fragment of OP1a cDNA was labeled with ^{32}P by nick-translation as described above and hybridized to Southern blots containing DNA from human–rodent cell hybrids as described previously (McBride *et al.*, 1982a,b, 1983). The results were confirmed by hybridizing the same blots with a 0.9-kb osteopontin 5′ cDNA probe. The human–hamster hybrids consisted of 27 primary clones and 14 subclones (16 positive of 41 total) and the human–mouse hybrids represented 14 primary clones and 40 subclones (33 positive of 54 total). Hybrid cells were analyzed for the presence of all human chromosomes except Y by standard isozyme analysis, by Southern analysis using probes from previously localized genes, and, often, by cytogenetic analysis. DNA was isolated from hybrid cell lines, digested with *Eco*RI, and size fractionated in a 0.7% agarose gel by electrophoresis, and partially depurinated fragments were transferred to positively charged nylon membranes in 0.5 *M* NaOH (Reed and Mann, 1985). The membranes were hybridized at 42°C with ^{32}P-labeled probes in a solution of 50% formamide, 5× SSPE (1× is 0.15 *M* NaCl, 0.01 *M* sodium phosphate, 0.001 *M* EDTA (pH 7.4), 5× Denhardt's, 10% dextran sulfate, 0.2% SDS, and denatured herring sperm DNA at 200 μg/ml. Membranes were washed at 55°C in 0.1× SSC containing 0.2% SDS. For DNA restriction fragment length polymorphism analysis, DNA was isolated from the peripheral leukocytes of 29 unrelated normal individuals, digested with restriction endonucleases, size fractionated by gel electrophoresis, transferred to nylon membranes, and hybridized with radiolabeled probes as described above.

In Situ Hybridization

In situ hybridization experiments were performed using peripheral blood lymphocytes from a normal male (46;X) which were cultured for 72 h at 37°C in RPMI 1640 supplemented with 15% fetal bovine serum, phytohemagglutinin (0.5 μg/ml), and antibiotics. Cultures were then synchronized by addition of BUdR (100 μg/ml) for 16 h prior to washing and resuspension in fresh medium containing thymidine (2.5 μg/ml) and incubated for an additional 5.5 h (Bhatt *et al.*, 1988) with colcemid (0.05 μg/ml) present during the final 20 min. The cells were centrifuged, swollen, and fixed, and air-dried metaphase spreads were prepared by standard procedures (Harper and Saunders, 1981). After treatment with RNase A (100 μg/ml) for 1 h at 37°C, the chromosomal DNA was denatured for 3 min in 0.07 *N* NaOH in 64% ethanol (Singh *et al.*, 1977; Landegent *et al.*, 1985). Radiolabeled probe was prepared by nick-translation of OP1a plasmid DNA with [^{3}H]TTP and [^{3}H]dCTP. The probe was mixed with hybridization solution (50% formamide, 5% dextran sulfate, 2× Denhardt's solution, 2× SSC, 5 m*M* EDTA, 20 m*M* sodium phosphate, pH 6.4, and 200 μg/ml sheared herring sperm carrier DNA), heat denatured, applied to slides (2.5×10^5 cpm probe/slide), and hybridized for 20 h at 42°C. Slides were washed in 50% formamide–2× SSC (pH 7.0) for 10 min at 42°C and in 2× SSC at 42°C to remove the nonspecifically bound probe and coated with a 50% solution of NTB2 nuclear track emulsion (Kodak, Rochester, NY). The slides were stored dessicated at 4°C for 10 days and then developed, stained (0.25% Wright stain), and photographed. The slides were destained and chromosomal banding was obtained by staining with 33258 Hoechst (150 μg/ml) for 30 min and exposure to uv illumination for 30 min after rinsing. The slides were again stained with Wright stain and the same metaphase spreads were rephotographed (Bhatt *et al.*, 1988). Only metaphase spreads containing a grain on chromosomes 4 or 5 were analyzed.

RESULTS

DNA and Deduced Protein Sequence of Human Osteopontin

DNA sequencing of a 1.5-kb human osteopontin cDNA revealed an open reading frame of 900 nucleotides and 5′ and 3′ untranslated regions of 108 and 481 bases, respectively (Fig. 1). The predicted protein has a prepeptide sequence composed of 16 amino acids that are primarily hydrophobic in nature. The amino-terminal sequence of bone matrix-derived human osteopontin has been determined previously (Fisher *et al.*, 1987) and is in general agreement with the deduced sequence of the cDNA (bold underline, Fig. 1). The two amino acids that differ from the cDNA occur at secreted protein positions 10 and 13 (26 and 29 of Fig. 1). Position 10 was originally reported as a glutamic acid rather than a serine suggesting that this serine may have been post-translationally modified. Position 13 was reported as glutamic acid rather than as the glutamine seen in the cDNA sequence. This may have been the result of deamination of the glutamine during purification. The predicted molecular weight of a protein translated from this mRNA is 33,842.8 Da. A Gly–Arg–Gly–Asp–Ser (GRGDS) sequence, present at amino acid residues 144–148, is presumed to be the site of cell attachment activity (Oldberg *et al.*, 1986). Three potential poly(A) attachment signals (AATAA) are found in the 3′ untranslated region, the most 3′ being

TABLE 1

Segregation of the SPP-1/OP Gene with Human Chromosome 4

Human chromosome	Gene/chromosome				% Discordancy
	+/+	+/−	−/+	−/−	
1	28	21	6	40	28
2	24	25	2	44	28
3	33	16	2	44	19
4	49	0	0	46	0
5	22	27	3	43	32
6	28	21	20	26	43
7	17	32	22	24	57
8	23	26	13	33	41
9	25	24	7	39	33
10	15	34	4	42	40
11	29	20	6	40	27
12	26	23	13	33	38
13	19	30	16	30	48
14	19	30	23	23	56
15	19	30	27	19	60
16	19	30	17	29	49
17	32	17	25	21	44
18	31	18	19	27	39
19	20	29	8	38	39
20	28	21	12	34	35
21	33	16	29	17	47
22	20	29	12	34	43
X	29	20	19	27	41

Note. Tabulation of the presence of the absence of SPP-1 gene in cell hybrids. The detection of the human osteopontin gene is correlated with the presence of each human chromosome in the group of somatic cell hybrids described under Materials and Methods. Discordancy represents the presence of the gene in the absence of the chromosome (+/−) or the absence of the gene despite the presence of the chromosome (−/+); the sum of these numbers divided by total hybrids examined (×100) represents percentage discordancy.

located 24 bp from the poly(A) tail of this osteopontin message.

A comparison of the deduced protein sequence among rat (Oldberg *et al.*, 1986), mouse (Craig *et al.*, 1989), and porcine (Wrana *et al.*, 1989) counterparts shows that the position of the GRGDS cell attachment sequence is highly conserved. A preponderance of Asp (D) residues which are located at amino acid positions 72–81 (in the human) are also conserved in position between species as well as the first potential N-linked glycosylation consensus sequence Asn–Xaa–Ser (position 65).

A comparison of the DNA sequence of OP1a with that of a human osteopontin cDNA sequence reported by Keifer *et al.* (1989) indicated that in addition to minor length differences at the 3′ and 5′ ends, the two cDNA clones differed internally at a single location as well; the osteosarcoma-derived cDNA contained an additional 42 bp of sequence located at base 280 of our bone cell-derived clone (see Fig. 1). To determine whether normal bone and other normal osteopontin-enriched tissues possess one or both forms of the mRNA, RNA from bone and decidua cells was amplified by polymerase chain reaction (PCR) using oligonucleotides that flanked the "gap" sequence (see Fig. 2A). As predicted from our cDNA, both bone and decidua contain an RNA species that comigrates precisely with material amplified using OP1a cDNA as template (Fig. 2B). A larger, apparently more predominant message with the exact length predicted from an amplification of the sequence of Keifer *et al.* (1989) was also apparent in both cell types. To demonstrate that the larger species is indeed the same as that reported by Keifer *et al.* (1989) PCR-amplified material was separated by agarose gels and hybridized with an oligonucleotide encoding the "gap" sequence (see Fig. 2A). Strong hybridization to the amplified decidua and bone (not shown) material was evident indicating that both species of message (OP1a and OP1b) are present in human bone and in decidua cells.

Relative Osteopontin Distribution in Human, Bovine, and Rat Tissue

RNA from human, rat, and bovine tissues was analyzed for its relative abundance in several bone and nonbone tissues (Fig. 3). The size of the human osteopontin mRNA is approximately 1.8 kb and is conserved in length between bovine and rat tissues. In the human, osteopontin mRNA is abundant in primary cultures of bone but is low in the human osteogenic sarcoma cell line SAOS-2 and in the pancreatic tumor line CRL 1469. Striking is the abundance of osteopontin in the decidua of the placenta. No osteopontin mRNA was detected in nonpregnant proliferative-phase endometrium (Fig. 4D). Osteopontin mRNA is scarce in the skin from all species analyzed. Strong cross-hybridization of the human cDNA to mRNA from the rat osteogenic sarcoma cell lines ROS 17/2 and UMR 106-01 as well as to a rat chondrosarcoma and rat calvaria tissue under stringent conditions confirmed the high degree of homology noted at the nucleic acid level between the two species.

Osteopontin Protein Expression in Human Endometrium and Decidua

In human pregnancy, the endometrium (epithelium and stroma) is termed decidua. The large epitheloid endometrial stromal cells, or the decidua cells, constitute the major cellular component. These decidual cells exhibited an intense cytoplasmic immunoreactivity against polyclonal antiserum raised against human bone osteopontin (Fig. 4C). To determine whether os-

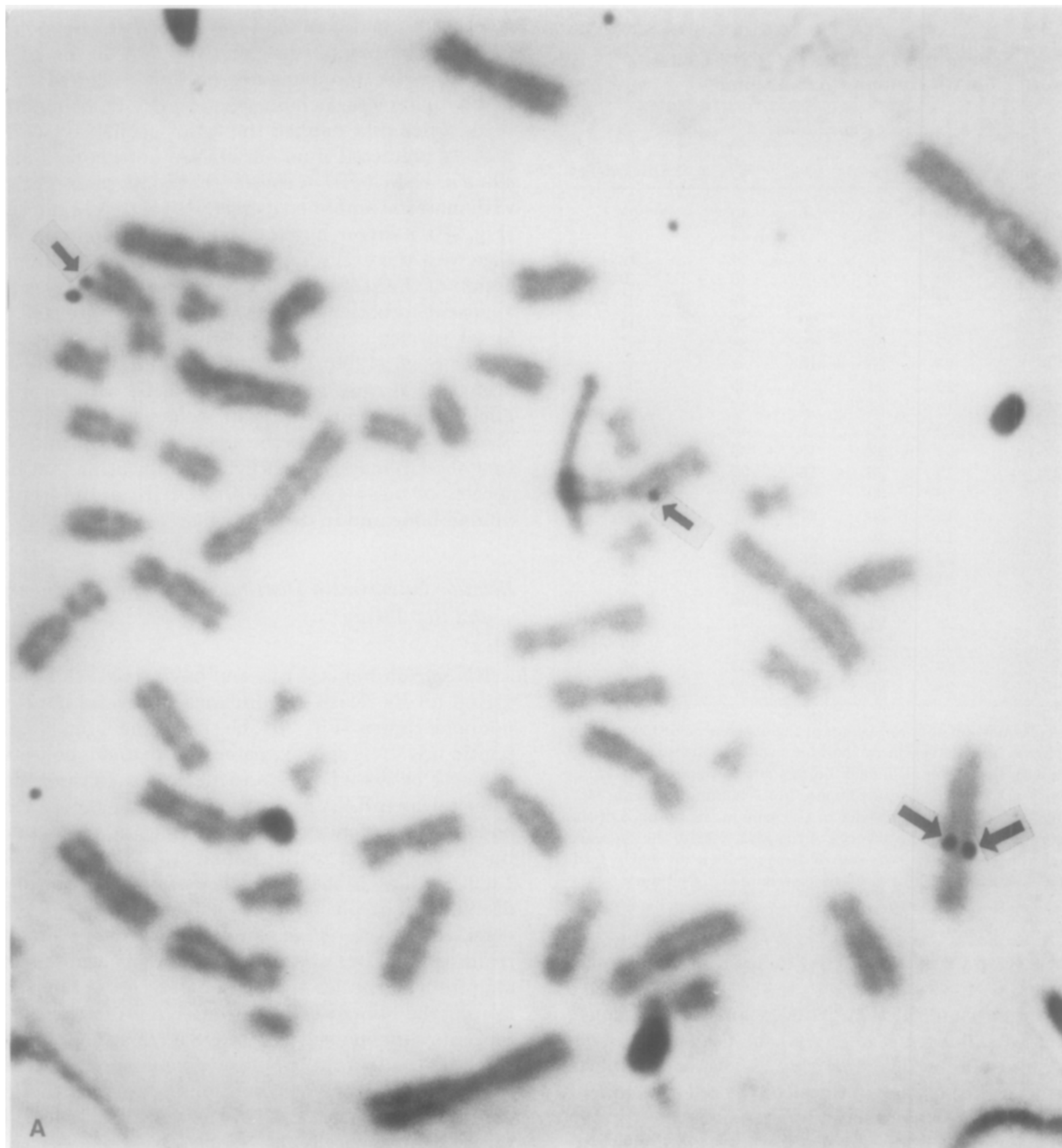

FIG. 6. *In situ* hybridization. (**A**) Typical metaphase spread showing a grain on each chromatid of the long arm of chromosome 4 (broad arrows) and nonspecific grains on two other chromosomes (thin arrows). (**B**) The same metaphase spread as shown in (**A**) after chromosome banding. (**C**) Distribution of grains on chromosomes 4 and 5 in 62 metaphase spreads examined.

teopontin is expressed in nonpregnant endometrial tissue, samples from the proliferative and secretory phase of the normal menstrual cycle were examined. In the proliferative phase, both endometrial glands and endometrial stromal cells were devoid of osteopontin immunoreactivity (Fig. 4A). The blood vessels exhibited variable immunoreactivity. In the secretory phases, the endometrial glands became immunoreactive. During most of the menstrual cycle the endometrial stromal cells were devoid of immunoreactivity (Fig. 4B); however, late in the secretory phase they became slightly positive (not shown).

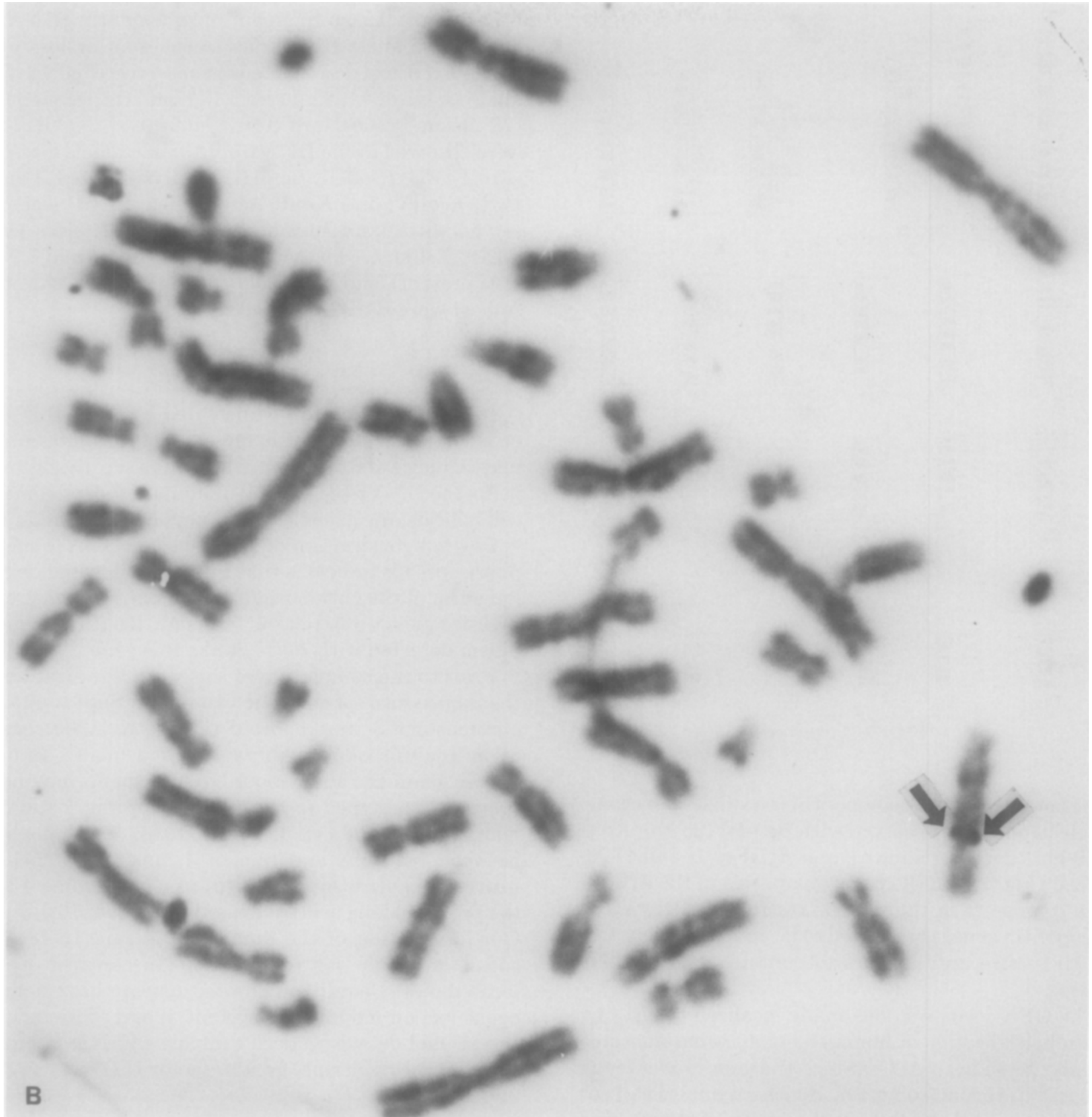

FIG. 6—*Continued*

Chromosomal Mapping and Analysis of the Human Osteopontin/SPP-1 Gene

The human osteopontin (OPN) gene was chromosomally mapped by Southern analysis of a panel of human–rodent hybrid cell DNAs. A 0.9-kb *Eco*RI fragment containing most of the coding sequence was used to probe the hybrid cell DNAs under stringent conditions. The probe identified 0.9-, 2.1-, and 6.8-kb bands in *Eco*RI digests of human DNA and all three bands were detected in the same hybrid cell DNAs. (Figs. 5A and 5B). Very weakly cross-hybridizing 5.5-kb hamster and 6.6-kb mouse bands were resolved from the human bands in these digests. Analysis of the total array of hybrids (Table 1) allowed unambiguous localization of the osteopontin gene to human chromosome 4. The gene segregated discordantly (>19%) with all other human chromosomes. Analysis of the size and number of the hybridizing restriction fragments detected in human DNA digested with 12 different restriction en-

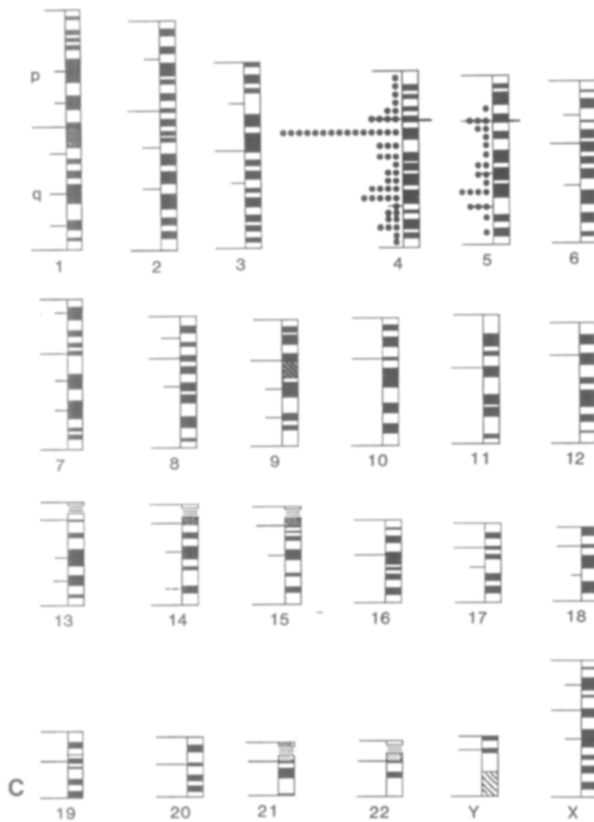

FIG. 6—*Continued*

zymes indicated that osteopontin is a single-copy gene. Single bands were observed with *Bam*HI (25 kb), *Kpn*I (11.9 kb) (data not shown), and *Bgl*II (9.1- and 11-kb alleles) (Fig. 7). On the basis of analysis of the hybridizing fragments in the various restriction digests, the osteopontin gene is estimated to be approximately 5.4–8.2 kb in size and appears to contain at least three introns.

The OPN locus was regionally localized to 4q13 by *in situ* hybridization of human metaphase chromosome spreads (see Figs. 6A–C). Since the gene was already assigned to chromosome 4 by Southern analysis of somatic cell hybrids, only metaphase spreads containing hybridization to chromosome 4 and 5 were examined. As shown, 40% (22/55 total) of the grains were localized to the band 4q11–21 and 27% (15/55 total) were localized to the band 4q13. In contrast, there was a random distribution of grains on chromosome 5 (23 total) and all other chromosomes (62 total) representing nonspecific background.

Restriction Fragment Length Polymorphism Analysis

DNAs isolated from peripheral leukocytes of 29 normal unrelated individuals were digested with *Bgl*II, size fractionated by agarose gel electrophoresis, and hybridized to a 0.9-kb fragment of osteopontin cDNA. A simple two-allele RFLP due to the gain or loss of a single *Bgl*II restriction site was observed (Fig. 7). The allele sizes are A1:A2, 11 kb:9.1 kb and the frequencies (78 chromosomes) are A1:A2, 0.64:0.36. No RFLPs were observed in the DNA from 10 individuals digested with *Eco*RI, *Hin*dIII, *Bam*HI, *Xba*I, *Sac*I, *Taq*I, *Pvu*II, *Pst*I, *Eco*RV, and *Kpn*I. A two-allele RFLP was observed with *Msp*I (not shown); the allele sizes were B1:B2, 6.4 kb:5.1 kb (constant bands of 1.4 and 7.2 kb also present) and the allele frequencies (76 chromosomes) were B1:B2, 0.12:0.88.

DISCUSSION

In this investigation we isolated a cDNA that encodes human osteopontin and studied the distribution and sequence of its mRNA in human tissue. In addition, using the DNA from a panel of human–rodent somatic cell hybrids, the gene has been mapped to chromosome 4. *In situ* hybridization of radiolabeled cDNA to metaphase spreads indicates that the gene is located on the long arm of the chromosome 4 near the centromere. A high-frequency restriction fragment length polymorphism detected with *Bgl*II flanks this gene.

Prior to the completion of this work, localization of the mouse form of the gene (known as Spp) to mouse chromosome 5 was reported (Fet *et al.*, 1989) based upon analysis of recombinant inbred mouse strains. Eleven genes and anonymous DNA segments have been previously assigned to mouse chromosome 5 and human chromosome 4 (Lalley *et al.*, 1988). The homology between mouse 5 and human 4 appears to include nearly the entire short arm (p) of chromosome 4 and the proximal long arm (q) extending from 4p16.3–4q21. Our localization of the human OPN gene to 4q13 is consistant with this relationship. In contrast, genes located more distally on chromosome 4q have homologous loci on mouse chromosome 3, and these include the alcohol dehydrogenase complex (ADH2) at 4q21–q25 and the epidermal growth factor (EGF) at 4q25.

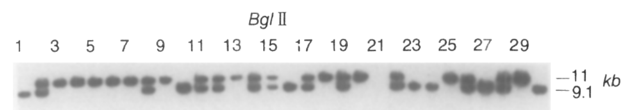

FIG. 7. Detection of a *Bgl*II restriction fragment length polymorphism (RFLP) with an osteopontin/SPP-1 cDNA probe. DNAs isolated from the peripheral leukocytes of 29 normal unrelated individuals were digested with *Bgl*II and size fractionated by electrophoresis in 0.7% agarose gels and the transferred DNA was probed with a 0.9-kb *Eco*RI fragment from the 5′ untranslated and coding regions of the human osteopontin cDNA OP10 (see Fig. 1). Allelic bands of 9.1 and 11 kb were observed. Both alleles are present in lanes **2, 8, 11, 12, 14, 15, 17, 19, 22, 26,** and **28,** whereas homozygotes are present in the remaining lanes. No human DNA is present in lane **21.**

The cDNA sequence of osteopontin determined from primary cultures of human bone cells (this study) differs from a human osteopontin cDNA isolated from a human osteosarcoma (Keifer *et al.*, 1989); the latter sequence predicts 14 additional amino acids at residue 58 of the protein. Polymerase chain reaction amplification of osteopontin mRNA from normal bone and decidual cells in conjunction with specific oligonucleotide hybridization of amplified DNA (to OP1b) indicated that these two forms of the message (OP1a and OP1b) exist in both of these cell types. Because gene mapping analysis indicates the presence of only one gene (Fet *et al.*, 1989; and this study) it is likely that these messages are generated by a post-transcriptional mechanism such as differential RNA splicing. Two models might be predicted for the generation of OP1a and OP1b messages. The first model predicts that the sequences of the "gap" are encoded by an individual exon. A second possibility is that "gap" sequences are part of an exon that encodes additional protein sequences located 3′ or carboxy-terminal to the "gap." This latter possibility seems more likely due to the presence of a splice consensus sequence (AG) at the 3′ terminus of the "gap" which could result in a "cryptic" splice of this exon. It is possible that the multiple forms of osteopontin recently detected in normal and transformed rat bone cells (Kubota *et al.*, 1989) may, in part, be generated by a polymorphism in the osteopontin mRNA.

A study of the distribution of osteopontin mRNA expression in human tissues indicates a preponderance of the message in cultured bone cells and in the decidua. These observations agree well with those of Nomura *et al.* (1988) who, by *in situ* hybridization, found substantial amounts of osteopontin mRNA in the developing limb bones and in the metrial gland cells of the murine decidua. Immunohistochemical analysis of decidual and nonpregnant endometrial tissues confirmed that the source of expression of osteopontin in the human is the cells of the decidua and, in secretory-phase endometrium, the cells of the secretory gland as well. Secretory glands from the proliferative-phase endometrium, on the other hand, are completely devoid of osteopontin mRNA or protein. The changes in expression of osteopontin in the endometrium during normal cycling and during pregnancy implicate a potential hormonal element in the regulation of this gene. The function of the osteopontin(s) in these tissues is not entirely clear but is likely to be associated with its ability to enhance cell attachment *in vitro*. Indeed, the cell attachment sequence Gly–Arg–Gly–Asp–Ser which is located midway in the protein is conserved in position in four species examined. A highly acidic stretch composed exclusively of aspartic acid residues is also conserved in position and has been speculated to be the site of mineral binding within the protein (Oldberg *et al.*, 1986). Because the crystallographic or NMR-derived tertiary structure of osteopontin is not known at this time, it is difficult to speculate what role the presence or absence of 14 amino acids at residue 58 of the protein will have on its three-dimensional folding pattern. However, it should be noted that the two predicted variants of osteopontin differ by two serine residues (SSEETDD) that are potential sites for O-linked glycosylation as well as protein phosphorylation. It is interesting to speculate that these or other post-translational modifications modulate the function(s) of osteopontin in normal and transformed tissues.

REFERENCES

1. AUSEBEL, F. M. (1989). "Current Protocols in Molecular Biology," Vol. 1, pp. 2.910–2.9.12, Wiley, New York.
2. BHATT, B., BURNSS, J., FLANNERY, D., AND MCGEE, J. O. D. (1988). Direct visualization of single copy genes on banded metaphase chromosomes by non-isotopic *in situ* hybridization. *Nucleic Acids Res.* **16:** 3951–3961.
3. BUTLER, W. T. (1989). The nature and significance of osteopontin. *Connect. Tissue Res.* **23:** 123–126.
4. CRAIG, A. M., NEMIR, M., MUKHERJEE, B. B., CHAMBERS, A. F., AND DENHARDT, D. T. (1988). Identification of the major phosphoprotein secreted by many rodent cell lines as 2ar/osteopontin: Enhanced expression in H-*ras*-transformed 3T3 cells. *Biochem. Biophys. Res. Commun.* **157:** 166–173.
5. CRAIG, A. M., SMITH, J. H., AND DENHARDT, D. T. (1989). Osteopontin, a transformation-associated cell adhesion phosphoprotein, is induced by 12-*O*-tetradecanoylphorbol 13-acetate in mouse epidermis. *J. Biol. Chem.* **264:** 9682–9689.
6. FET, V., DICKINSON, M. E., AND HOGAN, B. L. M. (1989). Localization of the mouse gene for secreted phosphoprotein 1 (*Spp-1*) (2ar, osteopontin, bone sialoprotein 1, 44-kDa bone phosphoprotein, tumor-secreted phosphoprotein) to chromosome 5, closely linked to *Ric (Rickettsia resistance)*. *Genomics* **5:** 375–377.
7. FISHER, L. W., HAWKINS, G. R., TUROSS, N., AND TERMINE, J. D. (1987). Purification and partial characterization of small proteoglycans I and II, bone sialoproteins I and II and osteonectin from the mineral compartment of developing human bone. *J. Biol. Chem.* **262:** 9702–9708.
8. FISHER, L. W., TERMINE, J. D., AND YOUNG, M. F. (1989). Deduced protein sequence of bone small proteoglycan I (biglycan) shows homology with proteoglycan II (decorin) and several nonconnective tissue proteins in a variety of species. *J. Biol. Chem.* **264:** 4571–4576.
9. GEHRON ROBEY, P., KERSHNER, J. A., CONN, K. M., AND TERMINE, J. D. (1985). Biosynthesis of non-collagenous proteins by bone cells *in vitro*. *In* "Current Advances in Skeletogenesis" (A. Ornoy, A. Harrell, and J. Sela, Eds.), Elsevier, New York.
10. GEHRON ROBEY, P., AND TERMINE, J. D. (1985). Human bone cells *in vitro*. *Calcif. Tissue Int.* **37:** 453–460.
11. HARPER, M. E., AND SAUNDERS, G. F. (1981). Localization of single copy DNA sequences on G-banded human chromosomes by *in situ* hybridization. *Chromosoma* **83:** 431–439.
12. KIEFER, M. C., BAUER, D. M., AND BARR, P. J. (1989). The cDNA and derived amino acid sequence for human osteopontin. *Nucleic Acids Res.* **17:** 3306.
13. KUBOTA, T., ZHANG, Q., WRANA, J. L., BER, R., AUBIN, J. E.,

BUTLER, W. T., AND SODEK, J. (1989). Multiple forms of spp (secreted phosphoprotein, osteopontin) synthesized by normal and transformed bone cell populations: Regulation by TGF-b. *Biochem. Biophys. Res. Commun.* **162:** 1453–1459.

14. LALLEY, P. A., DAVISSON, M. T., GRAVES, J. A. M., O'BRIAN, S. J., RODERICK, T. H., DOOLITTLE, D. P., AND HILLYARD, A. L. (1988). Report of the committee on comparative mapping (in HGM 9.5). *Cytogenet. Cell Genet.* **49:** 227–235.

15. LANDEGENT, J. E., JANSEN IN DE WEL, N., VAN OMMEN, G. J. B., BAAS, F., DE VIJLDER, J. J., VAN DIUJN, P., AND VAN DER PLOEG, M. (1985). Chromosomal localization of a unique gene by nonautoradiographic *in situ* hybridization. *Nature (London)* **327:** 175–177.

16. MANIATIS, T., FRITSCH, E. F., AND SAMBROOK, J. (1982). "Molecular Cloning: A Laboratory Manual," Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

17. MARK, M. P., PRINCE, C. W., OOSAWA, T., GAY, S., BRONKERS, A. L. J. J., AND BUTLER, W. T. (1987). Immunohistochemical demonstration of a 44kdal phosphoprotein in developing rat bones. *J. Histochem. Cytochem.* **35:** 707–716.

18. MARK, M. P., PRINCE, C. W., GAY, S., AUSTIN, R. L., AND BUTLER, W. T. (1988). 44kDal bone phosphoprotein antigenicity at ectopic sites in newborn rats: Kidney and nervous tissue. *Cell Tissue Res.* **251:** 23–30.

19. MCBRIDE, O. W., HEITER, P. A., HOLLIS, G. F., SWAN, D., OTEY, M. C., AND LEDER, P. (1982a). Chromosomal location of human kappa and lambda immunoglobulin light chain constant region genes. *J. Exp. Med.* **155:** 1480–1490.

20. MCBRIDE, O. W., BATTEY, J., HOLLIS, G. F., SWAN, D., SEIBENLIST, U., AND LEDER, P. (1982b). Localization of human variable and constant region immunoglobulin heavy chain genes on subtelomeric bandq32 of chromosome 14. *Nucleic Acids Res.* **10:** 8155–8170.

21. MCBRIDE, O. W., SWAN, D. C., TRONICK, S. R., GOL, R., KLIMANIS, D., MOORE, D. E., AND AARONSON, S. A. (1983). Regional chromosomal localization of N-*ras*, K-*ras*-1, K-*ras*-2, and *myb* oncogenes in human cells. *Nucleic Acids Res.* **11:** 8221–8236.

22. MESSING, J., CREA, R., AND SEEBERG, P. H. (1981). A system for shotgun DNA sequencing. *Nucleic Acids Res.* **9:** 309–321.

23. NODA, M., YOON, K., PRINCE, C. W., BUTLER, W. T., AND RODAN, G. A. (1988). Transcriptional regulation of osteopontin production in rat osteosarcoma cells by type beta transforming growth factor. *J. Biol. Chem.* **263:** 13,916–13,921.

24. NOMURA, S., WILLS, A., EDWARDS, D. R., HEATH, J. K., AND HOGAN, B. L. M. (1988). Developmental expression of 2ar (osteopontin) and SPARC (osteonectin) RNA as revealed by in situ hybridization. *J. Cell Biol.* **106:** 441–450.

25. OLDBERG, A., FRANZEN, A., AND HEINEGARD, D. (1986). Cloning and sequence analysis of rat bone sialoprotein (osteopontin) cDNA reveals an Arg–Gly–Asp cell-binding sequence. *Proc. Natl. Acad. Sci. USA* **83:** 8819–8823.

26. PRINCE, C. W., AND BUTLER, W. T. (1987). 1,25-Dihydroxyvitamin D_3 regulates the biosynthesis of osteopontin, a bone-derived cell attachment protein, in clonal osteoblast-like osteosarcoma cells. *Collagen Relat. Res.* **7:** 305–313.

27. REED, K. C., AND MANN, D. A. (1985). Rapid transfer of DNA from agarose gels to nylon membranes. *Nucleic Acids Res.* **13:** 7207–7221.

28. SANGER, F., NICKLEN, S., AND COULSON, A. (1977). DNA sequencing with chain terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74:** 5463–5467.

29. SENGER, D. R., AND PERRUZZI, C. A. (1985). Secreted phosphoprotein markers for neoplastic transformation of human epithelial and fibroblastic cells. *Cancer Res.* **45:** 5818–5823.

30. SHIMOKAWA, H., SOBEL, M. E., SASAKI, M., TERMINE, J. D., AND YOUNG, M. F. (1987). Heterogeneity of amelogenein mRNA in the bovine tooth germ. *J. Biol. Chem.* **262:** 4042–4047.

31. SINGH, L., PURDOM, K. F., AND JONES, K. W. (1977). Effect of different denaturing agents on the detectability of specific DNA sequence of variant base compositions by *in situ* hybridization. *Chromosoma* **60:** 377–389.

32. SMITH, J. H., AND DENHARDT, D. T. (1987). Molecular cloning of a tumor promoter-inducible mRNA found in JB6 mouse epidermal cells: Induction is stable at high, but not low, cell densities. *J. Cell. Biochem.* **34:** 13–22.

33. SOMERMAN, M. J., PRINCE, C. W., SAUK, J. J., FOSTER, R. A., AND BUTLER, W. T. (1987). Mechanism of fibroblast attachment to bone extracellular matrix: Role of 44 kilodalton bone phosphoprotein. *J. Bone Mineral Res.* **2:** 259–265.

34. SOMERMAN, M. J., FISHER, L. W., FOSTER, R. A., AND SAUK, J. J. (1988). Human bone sialoprotein I and II enhance fibroblast attachment *in vitro*. *Calcif. Tissue Int.* **43:** 50–53.

35. WEWER, U. M., ALBRECHTSEN, R., FISHER, L. W., YOUNG, M. F., AND TERMINE, J. D. (1988). Osteonectin/SPARC/BM-40 in human decidua and carcinoma tissues characterized by *de novo* formation of basement membrane. *Amer. J. Pathol.* **132:** 345–355.

36. WRANA, J. L., ZHANG, Q., AND SODEK, J. (1989). Full length cDNA sequence of porcine secreted phosphoprotein-I (SPP-I, osteopontin). *Nucleic Acids Res.* **17:** 10,119.

37. YOON, K., BUENAGA, R., AND RODAN, G. (1987). Tissue specificity and developmental expression of rat osteopontin. *Biochem. Biophys. Res. Commun.* **148:** 1129–1136.