BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS **243,** 61–66 (1998)
ARTICLE NO. RC978062

Structure, Chromosomal Location, and Expression Profile of EXTR1 and EXTR2, New Members of the Multiple Exostoses Gene Family

Toshiyuki Saito,*[,1] Naohiko Seki,† Masatake Yamauchi,* Satsuki Tsuji,*
Akiko Hayashi,* Sumie Kozuma,* and Tada-aki Hori*

**Genome Research Group, National Institute of Radiological Sciences, Anagawa 4-9-1, Inage-ku, Chiba 263-8555, Japan; and †Laboratory of Gene Function II, Kazusa DNA Research Institute, Yana 1532-3, Kisarazu, Chiba 292, Japan*

Received December 18, 1997

Hereditary multiple exostoses (EXT) is an autosomal dominant disorder that is characterized by the appearance of multiple outgrowths of the long bones (exostoses) at their epiphyses. Genetical heterogeneities have segregated at least on chromosome 8, 11, and 19 and been designated EXT1, EXT2, and EXT3, respectively. Recently, the responsible genes for EXT1 and EXT2 have been isolated and appeared to define a structurally related gene family. In the present study, we have identified novel genes which share significant sequence homologies with the EXT genes. The predicted protein products of the novel EXT-related genes, EXTR1 and EXTR2 (for *EXT*-related genes 1 and 2), consist of 919 and 330 amino acid residues, respectively. These genes were transcribed ubiquitously in various tissues. Based on PCR-assisted analyses of both a human/rodent mono-chromosomal hybrid cell panel and a radiation hybrid mapping panel, EXTR1 was localized to the chromosome 8p21 region, where loss of heterozygosity has been frequently observed in various tumors, and EXTR2 was assigned to the chromosome 1p21 region, where osteopetrosis, a dominant hereditary disease of bone, has been mapped by genetic linkage analysis, implying that the protein products of these two EXT-related genes, as well as of the EXT genes, have potential tumor suppressor activity.

Hereditary multiple exostoses (EXT) is an autosomal dominant disorder characterized by short stature and the development of bone protuberances at the ends of the long bones (1-3). Three genetic loci for EXT have been identified by genetic linkage analysis at chromosomes 8q24.1 (4, 5), 11p11-13 (6, 7) and 19p (8) and designated EXT1, EXT2 and EXT3, respectively. In recent years, the responsible genes for EXT1 and EXT2 were isolated and appeared to share significant sequence homology with each other; i.e. the EXT genes consist a structurally-related unique gene family (9, 10). Sporadic and exostoses-derived chondrosarcomas have been demonstrated the loss of heterozygosity for the markers in EXT1 and EXT2 loci (11, 12), indicating that the genes responsible for EXTs may be tumor suppressor genes.

Another gene structurally-related to EXT1 and EXT2, a third member of EXT gene family was found more recently by searching the EST (expressed sequence tag) database and its full length cDNA sequence was determined (13). The third gene designated as EXTL was mapped to chromosome 1p36.1, a region where frequent loss of heterozygosity have been detected in various tumor types (14, 15). Thus there is a possibility that mutations of the EXT family genes may lead to tumorgenesis.

It should be important to search for other members of EXT gene family and, if they were, to examine their relationship to various types of tumors. In this study, we identified two kinds of messenger RNAs which share significant sequence homologies with EXT1, EXT2 and EXTL genes by their cDNA cloning and sequencing of the molecules. The entire open reading frames (ORF) of the two members of the EXT gene family were described. We also reported their chromosomal location and expression profile in various human tissues, from which arose the possibility that these new EXT-related genes function as a tumor suppressor gene.

MATERIALS AND METHODS

cDNA cloning and DNA sequencing. 5′- and 3′-rapid amplification of cDNA end (RACE) experiments were performed using prim-

[1] To whom correspondence should be addressed. Fax: 81-43-251-9818. E-mail: t_saito@nirs.go.jp.

0006-291X/98 $25.00

ers, 5′-GGCCCCAGGTCCAGGTCCCCGTTGTCAG-3′ and 5′-AAGCCTCTCCGATAGTGACCTCCTGGCTATG-3′, respectively for EXTR1 and 5′-GTGTCATCATCTACCATCAACACTG-3′ and 5′-ATAATGCAGACGTACAACAGAACAGATC-3′, respectively for EXTR2 with Marathon cDNA amplification kit (Clonetech, USA). RACE products obtained from human adult brain poly(A)$^+$ RNA by multiple independent amplification experiments were cloned and sequenced by the dideoxy chain-termination method with a 377 DNA sequencer (Applied Biosystems, USA) according to the supplier's instructions.

RNA blot analysis. Northern blot filters containing adult human poly (A)$^+$ RNAs (2 μg/lane) (Clonetech Laboratories, USA) were hybridized with 6.2 kb EXTR1 cDNA or 1.4 kb cDNA fragment of EXTR2 containing the entire open reading frame reconstructed from the RACE products. The cDNA fragments were labeled with [α-^{32}P]-dCTP and used as a hybridization probe. The northern blot filters were washed following the manufacturer's instruction. The RNA amount and integrity of each lane was evaluated by a control hybridization of the same blots with a ^{32}P-labeled β-actin probe.

Chromosome mapping. PCR-based chromosome mapping was carried out basically following our previous papers (16, 17). The human-rodent monochromosomal somatic cell hybrid panel (Mapping panel #2) and the radiation hybrid panel (Genebridge 4) were purchased from the National Institute of General Medicine Service, Coriell Cell Repositories and Research Genetics, USA, respectively. For EXTR1, primers used for PCR amplification correspond to the n. t. 130 to 150 n. t. (5′-GGCAAAGGCATCATAAGAAGC-3′) and the n. t. 330 to n. t. 350 (5′-TCAAACGATAAAAGGTGCCAC-3′). For EXTR2, primers used for PCR amplification correspond to the n. t. 44 to n. t. 64 (5′-TGGAGAAGGATAGATGCACGC-3′) and the n. t. 170 to n. t. 190 (5′-TGTGATGTTTCTCGAGTGCAG-3′). PCR was carried out in a final volume of 10 μl containing 1 $\times$ LA-PCR buffer (Takara, Japan), 2 μM each primer, 200 μM each dNTP, 50 ng template DNA and 0.01 units of LA-Taq DNA polymerase (Takara, Japan). Temperature and time profile were 30 cycles of 95°C for 20 sec. and 62°C for 1 min. Human, mouse and hamster genomic DNAs were also included as controls.

RESULTS

Cloning and structure of the novel EXT-related genes. The amino acid sequences deduced from the cDNAs of EXT1, EXT2, and EXTL were compared with the EST database (GenBank release, April, 1997) using tBLASTN program. Several ESTs were found to be similar but not identical to the three known genes of EXT family after translated into amino acid sequence. The ESTs were finally categorized into two groups as for independent transcripts, meaning that at least two extra members might exist. To obtain the full length structure of the putative new members of the EXT gene family, which we designate as EXTRs here for *EXT-r*elated genes, 5′ and 3′-RACE (*r*apid *a*mplification of *c*DNA *e*nd) experiments were performed using primers designed from the ESTs' sequences.

The amino acid sequences of EXTR proteins predicted from the common cDNA sequences among the independent RACE products are compared and aligned with the known EXT family proteins in Fig. 1. The gene having the longer ORF was designated as EXTR1 and the other gene was termed EXTR2. Their nucleotide sequence data will appear in the DDBJ, EMBL and GenBank nucleotide sequence databases with the following accession numbers AB007042 and AB009284, respectively. For EXTR1, a single long open reading frame (ORF) was identified, starting at nucleotide position 594 and extending to a stop codon at the position 3351. A typical polyadenylation signal (AATAAA) was found 16-21 bases upstream the 3′ poly (A) stretch (refer AB007042). The predicted EXTR1 protein consists of 919 amino acid residues and its calculated molecular weight is 105 kilodaltons. For EXTR2, the largest ORF was capable of coding for a protein of 330 amino acid residues, initiating at nucleotide position 289 and terminating at the position 1281. The calculated molecular mass of EXTR2 protein is 37 kilodaltons. Both proteins contain hydrophobic residues more than 50% of total.

The amino acid sequence homology among the EXT family proteins was better conserved in the carboxyl half than in the amino terminus side. There were few short gaps of homology in the carboxyl terminal side whereas many long gaps in the amino terminal side. The conserved region in the carboxyl side may involve in the common function of the EXT family proteins.

Detection of the EXTR mRNAs in various human tissues. Northern blots of poly (A)+ RNA from various human tissues were hybridized with the radio-labeled cDNAs of EXTR1 or EXTR2 and analyzed the gene expression profiles among tissues. The autoradiograms of the blots are shown in Fig. 2.

Two discrete signals of approximately 6.2 and 4.7 kb in length were detected for EXTR1 (Fig. 2, top panel). The size of the cloned EXTR1 cDNA (6172 bases long) showed good accordance with the larger signal of 6.2 kb, indicating that the cDNA represents almost the full structure of EXTR1 mRNA. Although it is not known whether the shorter signal was derived from alternative form of EXTR1 mRNA or another related gene transcript, it is more likely that we detected the shorter form of EXTR1 mRNA because we employed a very stringent hybridization/washing conditions. The reason why the larger one was selectively cloned is probably due to the RNA source tissue employed in our RACE reaction. The source tissue was a adult brain and it appeared to express dominantly the larger transcript (see Fig. 2). If an RNA fraction prepared from other tissues such as a liver, the smaller form of EXTR would be obtained. EXTR1 was expressed in all tissues except ovary in which very slight signals were seen. In brain, skeletal muscle and testis, the larger messenger of 6.2 kb gave more intense signal than that of the 4.7 kb. In contrast, the signal of the 4.7 kb was stronger than that of the 6.2 kb in heart, liver, thymus and prostate. Comparable intensity of both the messenger was shown in kidney, placenta, spleen, small intestine, colon and leukocyte.

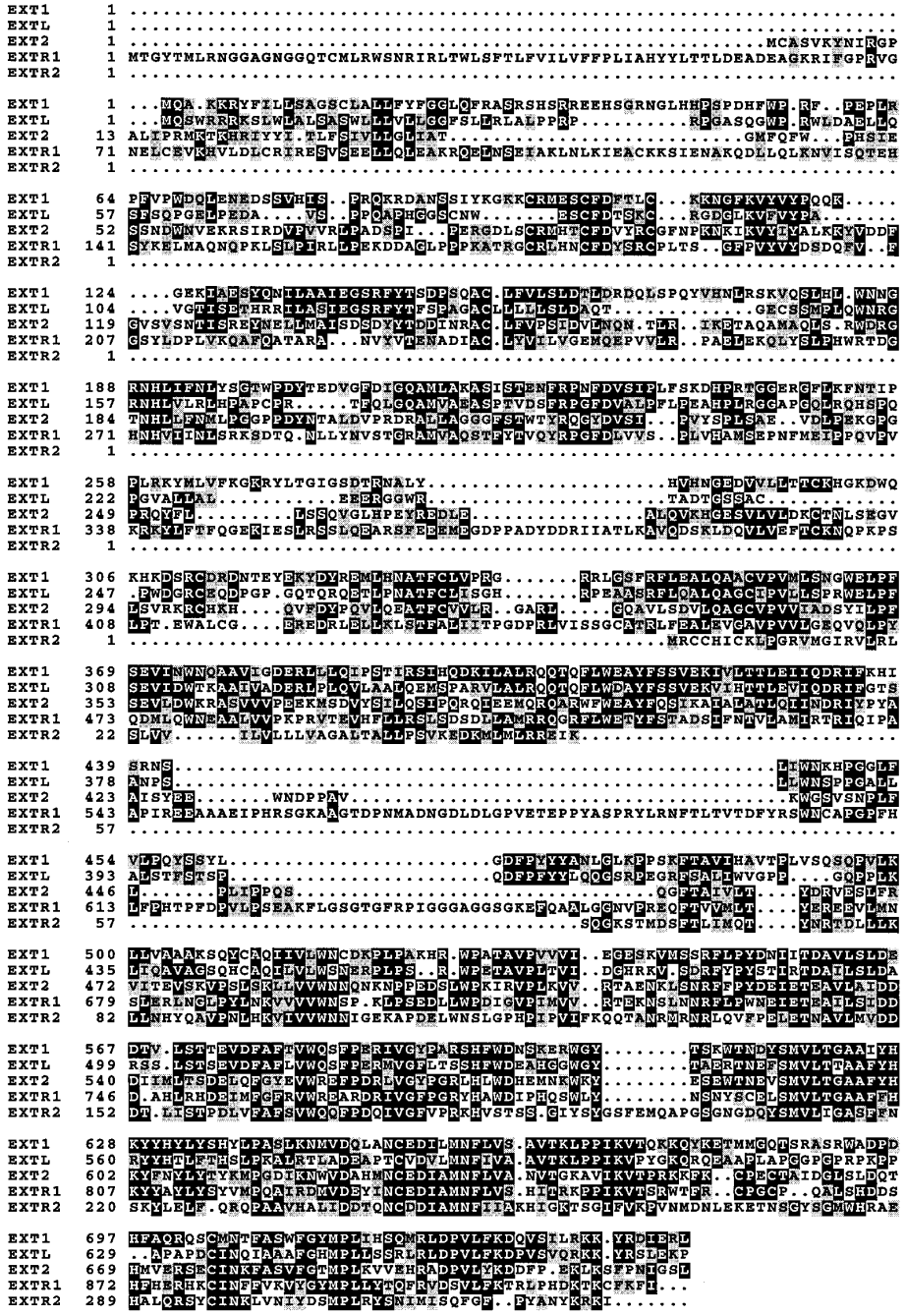

FIG. 1. Amino acid sequence alignment of the EXTR1 and EXTR2 proteins with other EXT family proteins. The consensus cDNA sequence was determined from multiple RACE products obtained from independent experiments and the amino acid sequence of the EXTR proteins were deduced from their cDNA sequences. The amino acid sequences of EXT1 (S79639), EXT2 (U62740), EXTL (U67191), EXTR1 (AB007042), and EXTR2 (AB009284) were analyzed and aligned by GCG software. The predicted amino acids are represented in one-letter designation. The most identical amino acids at each position are black boxed and similar ones are shadowed.

A single signal of 3.4 kb was obtained using EXTR2 probe in all tissues examined except leukocyte where the gene was hardly transcribed (Fig. 2, middle panel). Its expression was relatively constant among tissues whereas weak signals were seen in liver, lung and thymus. Although the EXTR2 cDNA cloned in the present study was 1.4 kb and the structure of approximately 2 kb of the mRNA remains to be determined, a single ORF surrounded with in-frame stop codons was identified in the present cDNA sequence (refer Acc. No. AB009284).

The ubiquitous expression of EXTR1 and EXTR2 in

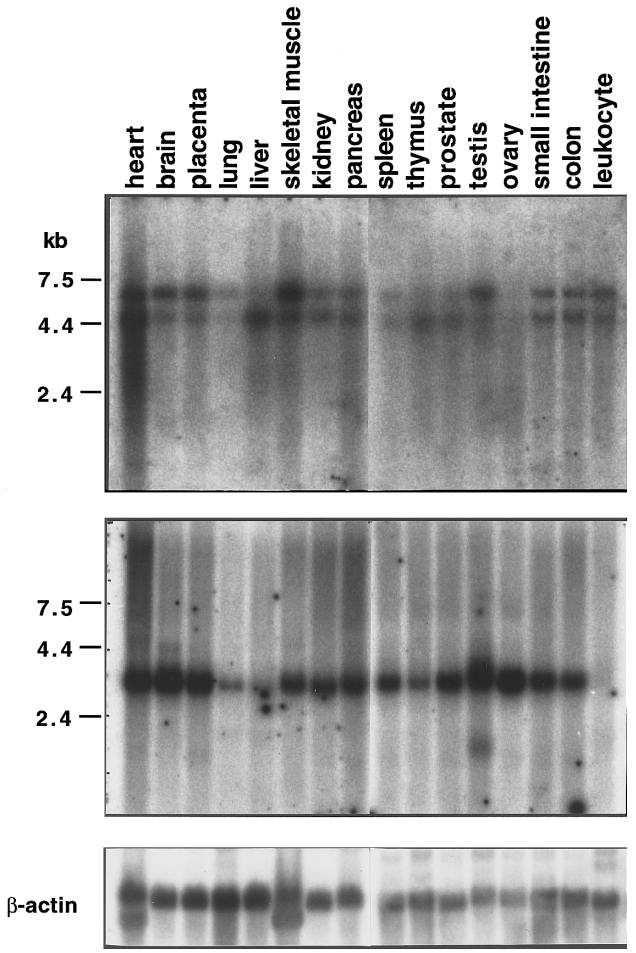

FIG. 2. Northern blot analysis of EXTR1 and EXTR2 mRNAs in various human tissues. The ^{32}P-labeled cDNA fragments of EXTR1 or EXTR2 containing the entire open reading frame reconstructed from the RACE products were used as hybridization probes. Two transcripts of 6.2 and 4.7 kb were detected in various tissues for EXTR1 (top). A single-sized mRNA of 3.4 kb was identified for EXTR2 (middle). The marker positions are indicated on the left side of the panel in kilobases (kb). As a control transcript, beta-actin mRNA was detected (bottom).

most types of tissues implies their important role(s) in a wide spectrum of cells, though the pattern of signal intensity was different from each other.

Chromosome mapping of EXTR1 and EXTR2. Chromosome assignment of EXTRs was performed using a human/rodent monochromosomal hybrid cell panel and a radiation hybrid mapping panel with primer sets for the 5′-untranslated region (UTR) of the genes.

For EXTR1, an apparent single amplified product was detected only in the lane of the hybrid containing human chromosome 8 (Fig. 3, top panel). 3′-UTR specific primers gave the same mapping data as in the figure (data not shown). We determined the further subchromosomal location of the gene by the radiation hybrid mapping method. PCR analysis of the radiation hybrid panel was performed with the same primers used in the above assay. The radiation hybrid data was statistically processed and analyzed by the RHMAPPER software package (http://www-genome.wi.mit.edu/cgi-bin/contig/rhmapper.pl). The data vector for EXTR1 gene was 0000000101 0010000000 0001110111 0100100010 0101111000 1000001100 0110010010 1001101101 1101000000 110 and the consequent report indicated the gene was mapped between markers WI-1172 and WI-8496, both of which have been cytogenetically mapped to 8p21. The position is 5.02 cR proximal from WI-1172. The same mapping position was predicted also when using a primer set designed in the 3′-UTR sequence (data not shown). Therefore, the gene was judged to map on 8p21.

For EXTR2, a single-sized amplified product was obtained only in the lane of the hybrid containing human

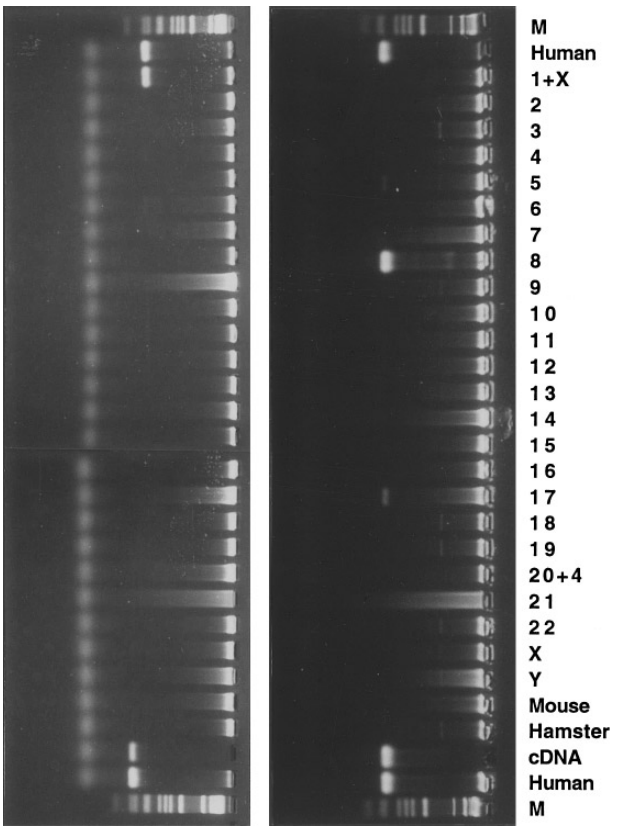

FIG. 3. Chromosome mapping of EXTR genes. PCR screening of a human-rodent somatic cell hybrid panel was performed to map EXTR genes to a specific human chromosome with their 5′-UTR specific primer pairs. Details of the experimental conditions are described under Materials and Methods. The top and bottom panels represent the chromosome which harbors EXTR1 or EXTR2, respectively. The numbers on the top of each lane indicate the human chromosome contained in each somatic cell hybrid. Human, mouse, and hamster genomic DNAs and the EXTR cDNA were included as controls.

chromosome 1 (Fig. 3, bottom panel). Subchromosomal location of the gene was analyzed by the radiation hybrid mapping method. The statistical processing of the data vector (0000010000 1100001000 1010000100 000-1000210 2201000000 1100100000 0000010010 000000-0000 0100011010 020) indicated that EXTR2 was mapped between markers D1S206 and WI-10815. The position is placed at 7.36 cR from D1S206. Both the markers have been cytogenetically mapped to 1p21.

DISCUSSION

Because loss of heterozygosity of nearby markers for EXT1 and EXT2 in sporadic and exostoses-derived chondrosarcomas were frequently observed (11, 12), both EXT genes are considered to consist a certain class of tumor suppressor genes. It is proposed that as a result of one allele of EXT genes, the dominant phenotype of bony protrusion occurs and additional mutation of the other allele will cause progression of tumors (9-12). In addition to these structurally related EXT genes, EXTL, another member of the EXT gene family was recently reported and it was mapped to chromosome 1p36.1 where chromosomal aberrations are frequently observed in a variety of tumors (13-15), suggesting that the third EXT isolog may also be a tumor suppressor.

Thus, the genes sharing some features of the EXT family have the possibility to act as a tumor suppressor gene. EXTR1 and EXTR2 in the present study, the forth and fifth members of the EXT gene family, are new isologs which share structural similarity with the other members. It must be worth examining their potential activity as a tumor suppressor.

After the two novel EXT isologs were found and isolated in the this work, a total of five family members came to be aligned. Structural similarity among them existed throughout the entire sequence and a highly conserved long contiguous region was revealed in their carboxyl terminal halves. The sequence is unique to the EXT family proteins and shows no homology to any other known proteins. The conservation may reflect the importance of the region (tentatively called EXT domain here) for a yet-unknown putative common activity of the proteins. We are now in an attempt to survey any other EXT isologs using PCR primers designed for the EXT domain.

EXTR1 was mapped to chromosome 8p21 in which frequent loss of heterozygosity has been observed in a variety of tumor types and existence of putative tumor suppressor gene(s) are considered (18-25). Therefore it would be interesting to survey EXTR1 mutations in various tumors.

EXTR2 was assigned to chromosome 1p21 where osteopetrosis, a dominant heritable disease of bone, has been genetically mapped (26). This disease is characterized by excessive formation of dense trabecular bone and calcified cartilage, especially in long bones, leading to obliteration of marrow spaces and to anemia. The bone disorders, hereditary multiple exostoses 1 and 2 are also inherited in a dominant manner and caused by mutation of the corresponding EXT genes. Thus the chromosomal localization of EXTR2 to the osteopetrosis locus proposes the possibility that mutation of the gene may lead to malformation and/or metabolic dysfunction of bone. EXTR2 should be a strong candidate gene for the disease osteopetrosis.

At present, EXT1 and EXT2 have been identified to be responsible genes for hereditary multiple exostoses which are genetically assigned to the EXT loci at 8q24.1 (4, 5) and 11p11-13 (6, 7), respectively. Although, for the remaining locus 19p (8), EXTR1 at 8p21 and EXTR2 at 1p21 are excluded as the candidate gene, these novel EXT isologs are to be subjected to extensive analyses on their mutation in various tumors and bone disorders because of their structural similarities to the EXT genes and their chromosomal mapping to the disease loci.

REFERENCES

1. Solomon, L. (1963) *J. Bone Joint Surg.* **45B,** 292–304.
2. Hennekam, R. C. (1991) *J. Med. Genet.* **28,** 262–266.
3. Schmale, G. A., Conrad, E. U., III, and Raskind, W. H. (1994) *J. Bone Joint Surg. Am.* **76,** 986–992.
4. Cook, A., Raskind, W., Blanton, S. H., Pauli, R. M., Gregg, R. G., Francomano, C. A., Puffenberger, E., Conrad, E. U., Schmale, G., Schellenberg, G., Wijsman, E., Hecht, J. T., Wells, D., and Wagner, M. J. (1993) *Am. J. Hum. Genet.* **53,** 71–79.
5. Ludecke, H. J., Wagner, M. J., Nardmann, J., La Pillo, B., Parrish, J. E., Willems, P. J., Haan, E. A., Frydman, M., Hamers, G. J., Wells, D. E., and Bernhard, H. (1995) *Hum. Mol. Genet.* **4,** 31–36.
6. Wu, Y. Q., Heutink, P., de Vries, B. B., Sandkuijl, L. A., van den Ouweland, A. M., Niermeijer, M. F., Galjaard, H., Reyniers, E., Willems, P. J., and Halley, D. J. (1994) *Hum. Mol. Genet.* **3,** 167–171.
7. Wuyts, W., Van Hul, W., Wauters, J., Nemtsova, M., Reyniers, E., Van Hul, E. V., De Boulle, K., de Vries, B. B., Hendrickx, J., Herrygers, I., Bossuyt, P., Balemans, W., Fransen, E., Vits, L., Coucke, P., Nowak, N. J., Shows, T. B., Mallet, L., van den Ouweland, A. M., McGaughran, J., Halley, D. J., and Willems, P. J. (1996) *Hum. Mol. Genet.* **5,** 1547–1557.
8. Le Merrer, M., Legeai-Mallet, L., Jeannin, P. M., Horsthemke, B., Schinzel, A., Plauchu, H., Toutain, A., Achard, F., Munnich, A., and Maroteaux, P. (1994) *Hum. Mol. Genet.* **3,** 717–722.
9. Ahn, J., Ludecke, H.-J., Lindow, S., Horton, W. A., Lee, B., Wagner. M. J., Horsthemke, B., and Wells, D. E. (1995) *Nature Genet.* **11,** 137–143.
10. Stickens, D., Clines, G., Burbee, D., Ramos, P., Thomas, S., Hogue, D., Hecht, J. T., Lovett, M., and Evans, G. A. (1996) *Nature Genet.* **14,** 25–32.
11. Hecht, J. T., Hogue, D., Strong, L. C., Hansen, M. F., Blanton, S. H., and Wagner, M. (1995) *Am. J. Hum. Genet.* **56,** 1125–1131.

12. Raskind, W. H., Conrad, E. U., Chansky, H., and Matsushita, M. (1995) *Am. J. Hum. Genet.* **56,** 1132–1139.
13. Wise, C. A., Clines, G. A., Massa, H., Trask, B. J., and Lovett, M. (1997) *Genome Res.* **7,** 10–16.
14. Schwab, M., Praml, C., and Amler, L. C. (1996) *Genes Chromosom. Cancer* **16,** 211–229.
15. Weith, A., Brodeur, G. M., Bruns, G. A., Matise, T. C., Mischke, D., Nizetic, D., Seldin, M. F., van Roy, N., and Vance, J. (1996) *Cytogenet. Cell Genet.* **72,** 114–144.
16. Saito, T., Seki, N., Ishii, H., Ohira, M., Hayashi, A., Kozuma, S., and Hori, T. (1997) *DNA Res.* **4,** 301–305.
17. Seki, N., Nimura, Y., Ohira, M., Saito, T., Ichimiya., S., Nomura, N., and Nakagawara., A. (1997) *DNA Res.* **4,** 355–358.
18. Kagan, J., Stein, J., Babaian, R. J., Joe, Y.-S., Pisters, L. L., Glassman, A. B., von Eschenbach, A. C., and Troncoso, P. (1995) *Oncogene* **11,** 2121–2126.
19. Tanaka, K., Kikuchi-Yanoshita, R., Muraoka, M., Konishi, M., Oshimura, M., and Miyaki, M. (1996) *Oncogene* **12,** 405–410.
20. Sunwoo, J. B., Holt, M. S., Radford, D. M., Deeker, C., and Scholnick, S. B. (1996) *Genes Chromosom. Cancer* **16,** 164–169.
21. Yaremko, M. L., Kutza, C., Lyzak, J., Mick, R., Recant, W. M., and Westbrook, C. A. (1996) *Genes Chromosom. Cancer* **16,** 189–195.
22. Becker, S. A., Zhou, Y.-Z., and Slagle, B. L. (1996) *Cancer Res.* **56,** 5092–5097.
23. Scholnick, S. B., Haughey, B. H., Sunwoo, J. B., El-Mofty, S. K., Baty, J. D., Piccirillo, J. F., and Zequeira, M. R. (1996) *J. Natl. Cancer Instit.* **88,** 1676–1682.
24. Nihei, N., Ichikawa, T., Kawana, Y., Kuramochi, H., Kugoh, H., Oshimura, M., Hayata, I., Shimazaki, J., and Ito, H. (1996) *Genes Chromosom. Cancer* **17,** 260–268.
25. He, W. W., Sciavolino, P. J., Wing, J., Augustus, M., Hudson, P., Meissner, P. S., Curtis, R. T., Shell, B. K., Bostwick, D. G., Tindall, D. J., Gelmann, E. P., Abate-Shen, C., and Carter, K. C. (1997) *Genomics* **43,** 69–77.
26. Hul, W. V., Bollerslev, J., Gram, J., Hul, E. V., Wuyts, W., Benichou, O., Vanhoenacker, F., and Willems, P. J. (1997) *Am. J. Hum. Genet.* **61,** 363–369.