Biochemical and Biophysical Research Communications **280,** 340–347 (2001)
doi:10.1006/bbrc.2000.4125, available online at http://www.idealibrary.com on IDEAL®

ALEX1, a Novel Human Armadillo Repeat Protein That Is Expressed Differentially in Normal Tissues and Carcinomas

Igor V. Kurochkin,[1] Natsumi Yonemitsu, Shin-ichi Funahashi, and Hitoshi Nomura
Chugai Research Institute for Molecular Medicine, 153-2 Nagai, Niihari, Ibaraki 300-4101, Japan

Received December 6, 2000

Members of the armadillo (arm) repeat family of proteins are implicated in tumorigenesis, embryonic development, and maintenance of tissue integrity. We have cloned cDNA for a novel human arm repeat protein, ALEX1, encoding 453 amino acids. ALEX1 shares significant homology with uncharacterized ORF KIAA0512 and putative protein product of unknown mRNA (GenBank AF211175), designated here as ALEX2 and ALEX3, respectively. The genes encoding ALEX1, ALEX2 and ALEX3 co-localize to the same region in Xq21.33–q22.2. ALEX1 and ALEX2 transcripts are found in all human tissues examined except peripheral blood leukocytes. Expression of ALEX1 and ALEX2 mRNA is lost or significantly reduced in human lung, prostate, colon, pancreas, and ovarian carcinomas and also in cell lines established from different human carcinomas. These genes are, however, normally expressed in cell lines derived from other types of tumors, e.g., sarcomas, neuroblastomas, and gliomas. We speculate that ALEX genes may play a role in suppression of tumors originating from epithelial tissue, i.e., carcinomas.

Key Words: **cloning; *Armadillo;* catenin; carcinoma; tumor suppressor.**

Members of the armadillo (arm) family of proteins relate to the product of the *armadillo* gene of *Drosophila* and have been implicated in a variety of processes such as tumorigenesis, development and maintenance of tissue integrity (1). Their common feature is an amino acid motif of about 42 residues (arm repeat), which has been identified in 6–13 imperfect tandemly repeated units in all members of this family (2). The arm repeats were first noticed in 1989 in the segment polarity gene *armadillo* in *Drosophila* (3). Since then a number of proteins containing homologous motifs have been cloned and sequenced. These include human β and δ-catenins, the tumor suppressor adenoma polyposis coli (APC), the src substrate p120, the import nuclear factor importin, and the GTP/GDP dissociation-stimulating factor smgGDS (1). Binding studies have shown that armadillo-related proteins interact with numerous different binding partners through their arm domains (4–6). The *Drosophila armadillo* gene was first identified as a component of the wingless signaling cascade (7). Similarly, the vertebrate equivalent of armadillo, β-catenin, is a critical component of the Wnt/Wingless growth factor signaling pathway that governs cell fate choice in early embryogenesis. In addition to its role in signaling functions, β-catenin has essential role in the morphogenesis of solid tissues and the subsequent maintenance of tissue integrity. β-Catenin binds to the highly conserved cytoplasmic domain of cadherins and to α-catenin, which binds to actin. The cadherin-catenin complex is a target of regulatory signals that govern cellular adhesiveness and mobility (8). β-Catenin interacts also with another arm protein, the tumor suppressor gene product APC (9). Mutations in APC are associated with familial and sporadic colorectal cancer (10). APC functions to decrease β-catenin stability since APC mutants lacking β-catenin binding site display elevated levels of cytosolic β-catenin and constitutive transcriptional activation of the β-catenin/TCF complex. Thus, arm proteins are involved in structural as well as in signaling functions.

In this study, we describe cloning and sequencing of a cDNA for a novel human arm repeat protein of 49 kDa named ALEX1. Screening the GenBank database revealed two close homologues of ALEX1, designated as ALEX2 and ALEX3. The genes for ALEX1 and its homologues were found to co-localize at the same interval on chromosome X. The expression of ALEX1 and

Sequence data from this article have been deposited with the DDBJ/EMBL/GenBank Data Libraries under Accession Nos. AB039669 and AB039670.

Abbreviations used: ALEX, arm protein lost in epithelial cancers on chromosome X; arm, armadillo; GAPDH, glyceraldehydes-3-phosphate dehydrogenase; ORF, open reading frame; RACE, rapid amplification of cDNA ends; RT-PCR, reverse transcription-PCR.

[1] To whom correspondence should be addressed. Fax: 81-298-30-6270. E-mail: igork@cimmed.com.

0006-291X/01 $35.00

ALEX2 mRNA is either significantly reduced or completely lost in human carcinomas but not in other tumor types. Therefore, we propose to designate new family of closely related proteins described here as ALEX (Arm proteins lost in epithelial cancers on chromosome X).

MATERIALS AND METHODS

Primers. ALEX1 specific primers: F1, 5′-GTGCTCGGGTTAAGAGATTTGTC-3′; F2, 5′-TCACAATGATCTGGTGGTG-3′; F3, 5′-CAACATGACTGTGACTAATC-3′; F4, 5′-GACTGTTAACCAACATGACTGTG-3′; R1, 5′-AGCTCCTTTCACAGTCTC-3′; R2, 5′-ACCCAACCATTACAACCAACATCAG-3′; R3, 5′-GGCCATGTTGTAGCTGGAGCCCTGGTGC-3′; R4, 5′-CAATATTGTTTGGGACAGACTCC-3′. ALEX2/KIAA0512 specific primers: F5, 5′-TAGCAGCACCTACCAAGGTAG-3′; F6, 5′-TGCCTTGCTTCAGAAATCTG-3′; R5, 5′-CCCAGTTCGTCTACTTCAACT-3′; R6, 5′-CTTCCACACTGCAAAATCATG-3′.

5′-Rapid amplification of cDNA ends (5′-RACE). The 5′-RACE was performed using Advantage cDNA Polymerase Mix (Clontech) with downstream primer R3 and pAP3neo human heart cDNA library (Takara) as a template. The product obtained was gel-purified, cloned into the vector pCR2.1-TOPO (Invitrogen) and sequenced using T7 primer, M13 reverse primer, and a series of insert-specific primers.

Screening of cDNA library and sequencing. Human testis cDNA library cloned in pCMV-SPORT (Gibco BRL) was divided into 36 pools each containing 3072 clones and gridded onto Nylon filters. The pools were screened by PCR with F1/R1 primers. Filters harboring ALEX1 cDNA were hybridized to the probe generated by PCR using the same primers. The probe was radiolabeled with [α-^{32}P] CTP (3000 Ci/mmol) (Amersham) using the Megaprime DNA Labeling System (Amersham). The longest cDNA thus identified was sequenced on both strands using the ABI dRhodamine Terminator Cycle Sequencing Ready Reaction Kit, and the products were resolved on an ABI Prism 377 Genetic Analyzer (Perkin-Elmer/Applied Biosystems). The sequences were assembled and analyzed using Auto Assembler DNA Sequence Assembly Software (Applied Biosystems). Sequence homology searches were carried out using BLASTN program (11). Protein sequence comparisons were performed using the BLASTP and PROSITE (12, 13). Motif searches were performed with Pfam at the Sanger Center, UK (http://www.sanger.ac.uk/Software/Pfam) (14). Prediction of protein transmembrane regions was performed with SOSUI (http://sosui.proteome.bio.tuat.ac.jp/sosuiframe0.html) (15).

Mapping. To identify the chromosomal location of ALEX1, PCR primers F4 and R4, which amplify a fragment of 460 bp, were used to amplify DNA from each of 24 monochromosomal human/rodent somatic cell lines (16) obtained from Coriell Cell Repositories (Camden, NJ). The regional localization of ALEX1 was identified using the GeneBridge 4 panel of 93 radiation hybrids (17). The panel was screened by PCR, again using primers F4 and R4. A binary code, generated by scoring each hybrid as positive or negative for amplification, was compared against similar codes for markers forming the framework map using a server located at http://carbon.wi.mit.edu:8000/cgi-bin/contig/rhmapper.pl.

Expression analyses. Human Multiple Tissue Northern Blots containing 2 μg of poly(A)$^+$ RNA from specific tissues and ExpressHyb hybridization solution were purchased from Clontech. Hybridization was conducted with DNA probe at 1×10^6 cpm/ml for 1 h at 68°C. The filters were washed with a final stringency of 0.1× SSC, 0.1% SDS at 50°C for 40 min and exposed to Hyperfilm MP (Amersham) using BioMax intensifying screens (Kodak) at −80°C. The probes were generated by PCR, gel-purified, and ^{32}P-labeled as described above. The 227-bp ALEX1 probe was generated using primers F2/R2. The 328-bp ALEX2/KIAA0512 probe was prepared with primers F5/R5. PCR analysis of the expression of ALEX1 and ALEX2/KIAA0512 genes in human normal and cancer tissues was performed with Human I, Human II and Human Tumor Multiple Tissue cDNA panels from Clontech. Primer combinations were F3/R2 for ALEX1 and F6/R6 for ALEX2 mRNA. RT-PCR was utilized to check expression of the genes in various human cell lines. For RT-PCR, 2 μg of total RNA was reverse-transcribed using the Superscript II first-strand cDNA synthesis kit (Gibco BRL) with oligo(dT) primers according to the manufacturer's specifications. Obtained cDNA was subjected to PCR amplification using primers described above. Tissue distribution was also obtained by querying the UniGene database (http://www.ncbi.nlm.nih.gov/UniGene/index.html).

Yeast two-hybrid screening. ALEX1 cDNA fragment was isolated by a two-hybrid screen of a human brain cDNA library (Clontech) in the yeast strain PJ69-4A with three reporter genes, *HIS3*, *ADE* and *LacZ* under the control of GAL4 promoter (18). The bait plasmid (pODB80-PP110) was constructed by the insertion of full-length peroxisome protease PP110 (H108Q) (19) in the polylinker of pODB80.

RESULTS

Cloning of human cDNA encoding a novel arm repeat protein and analysis of the protein. During an attempt to identify proteins interacting with PP110 by the yeast two-hybrid screening, we obtained several candidate cDNAs from normal human brain. Sequencing of the 1290 base pair (bp) cDNA insert from clone B4 revealed an 879 bp-long open reading frame, encoding a novel 293-residue polypeptide. Analysis of the predicted protein sequence against Pfam database identified within it two 42 amino acid armadillo/β-catenin-like repeats. Because of involvement of arm proteins in important cellular functions we decided to clone full-length cDNA for this new arm protein. The insert from clone B4 already contained polyA-tail at its 3′-end. 5′-RACE PCR was employed to obtain the rest of the sequence with human heart cDNA as a template. The largest amplified product contained additional 851 bp to clone B4 bringing total length of the cDNA to 2141 bp (Fig. 1). B4 cDNA was cloned also by hybridization screening of human testis cDNA library. The clone 141D10 had the largest insert length (2087 bp). Its sequence was identical to cDNA obtained by RACE-PCR except that the novel sequence lacked 51-bp fragment from the 5′ end and three nucleotides at positions 131 to 133. The divergence in nucleotide sequence occurs exactly at the junction of exon 1 and exon 2 of the gene (Fig. 1). The putative initiation codon was identified at position 372 and is located within a nucleotide sequence adequate for translation initiation (ACCATGG) (20). The first in-frame stop codon (TAA) was identified at nt 1731, predicting a protein product of 453 amino acids with a calculated molecular weight of 49,178 and a calculated isoelectric point of 9.56. The polyadenylation consensus sequence AATAAA is located 19 bp upstream of the polyadenylation site. Sequence analysis of the ALEX1 polypeptide revealed in

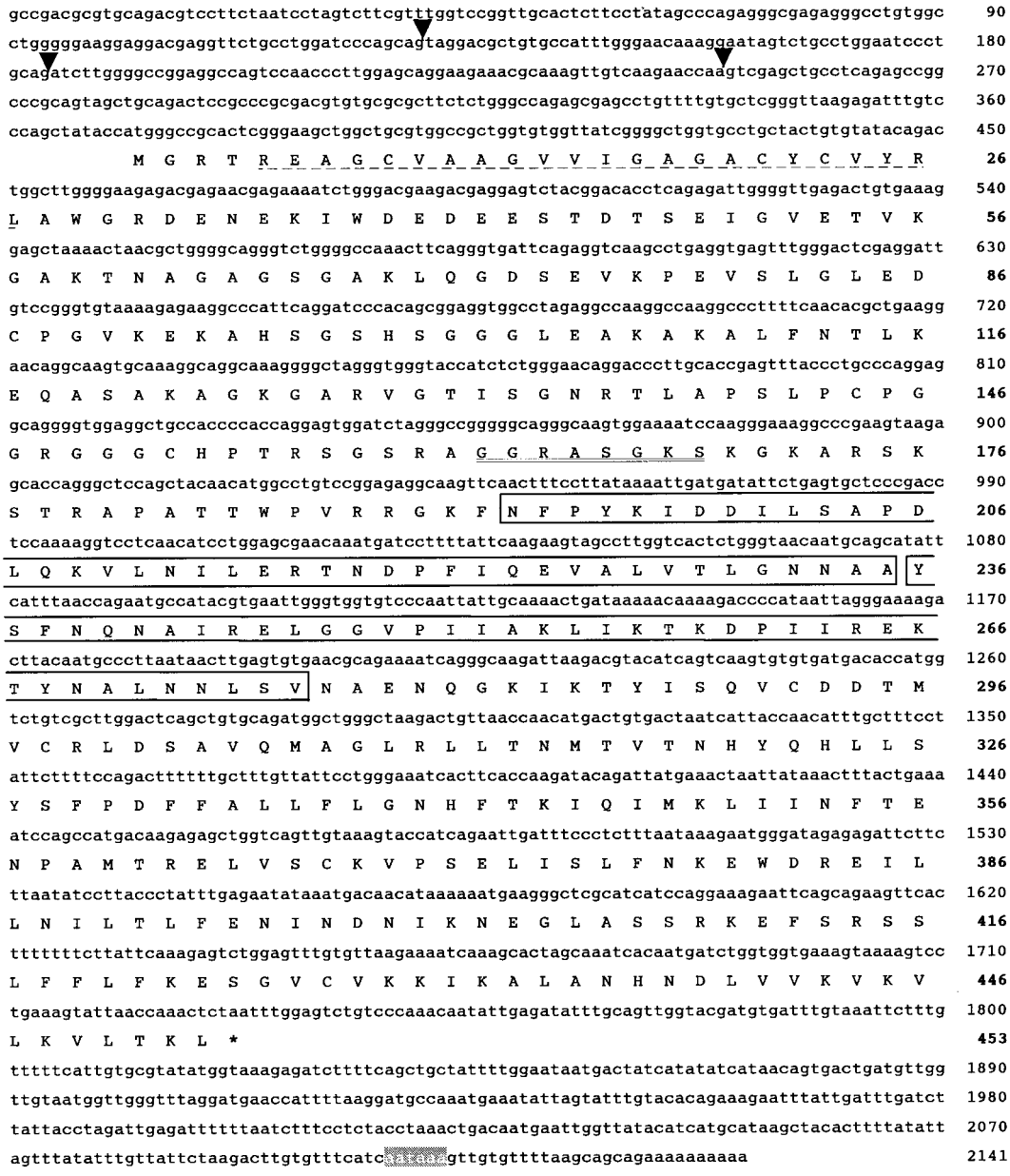

FIG. 1. Nucleotide and deduced amino acid sequence of the human ALEX1 cDNA. Boldface and roman numbers on the right refer to amino acids and nucleotides, respectively. Arm repeats are boxed. A dashed line marks residues of the predicted transmembrane domain. Putative ATP/GTP binding site is double-underlined. Polyadenylation signal AATAAA is shown in gray. Arrowheads indicate the positions of exon boundaries. The nucleotide sequence data reported in this paper have been deposited to DDBJ, EMBL, and GenBank database under the Accession No. AB039670.

addition to two arm repeats, an ATP/GTP-binding site at position 162 to 169. Potential phosphorylation sites present in ALEX1 include eight potential protein kinase C sites and five potential casein kinase II sites. The SOSUI algorithm predicts a transmembrane domain at the N-terminus of the protein (amino acids 5 to 27).

Homology of human ALEX1 protein with other proteins. A computer sequence search of public databases using the BLASTP program revealed that the full-length ALEX1 protein product is most closely related to a previously uncharacterized ORF KIAA0512 of 632 amino acids, designated here ALEX2. The homology region is confined to the C-terminal parts of the proteins with 51% identity and 72% similarity over a length of 234 amino acids (Fig. 2A). In addition, BLAST searches of dbEST database identified a number of highly homologous ESTs that assembled in a contig of 2,036 nt (THC257925) at the dbTHC database (TIGR, http://www.tigr.org/). A conceptual translation

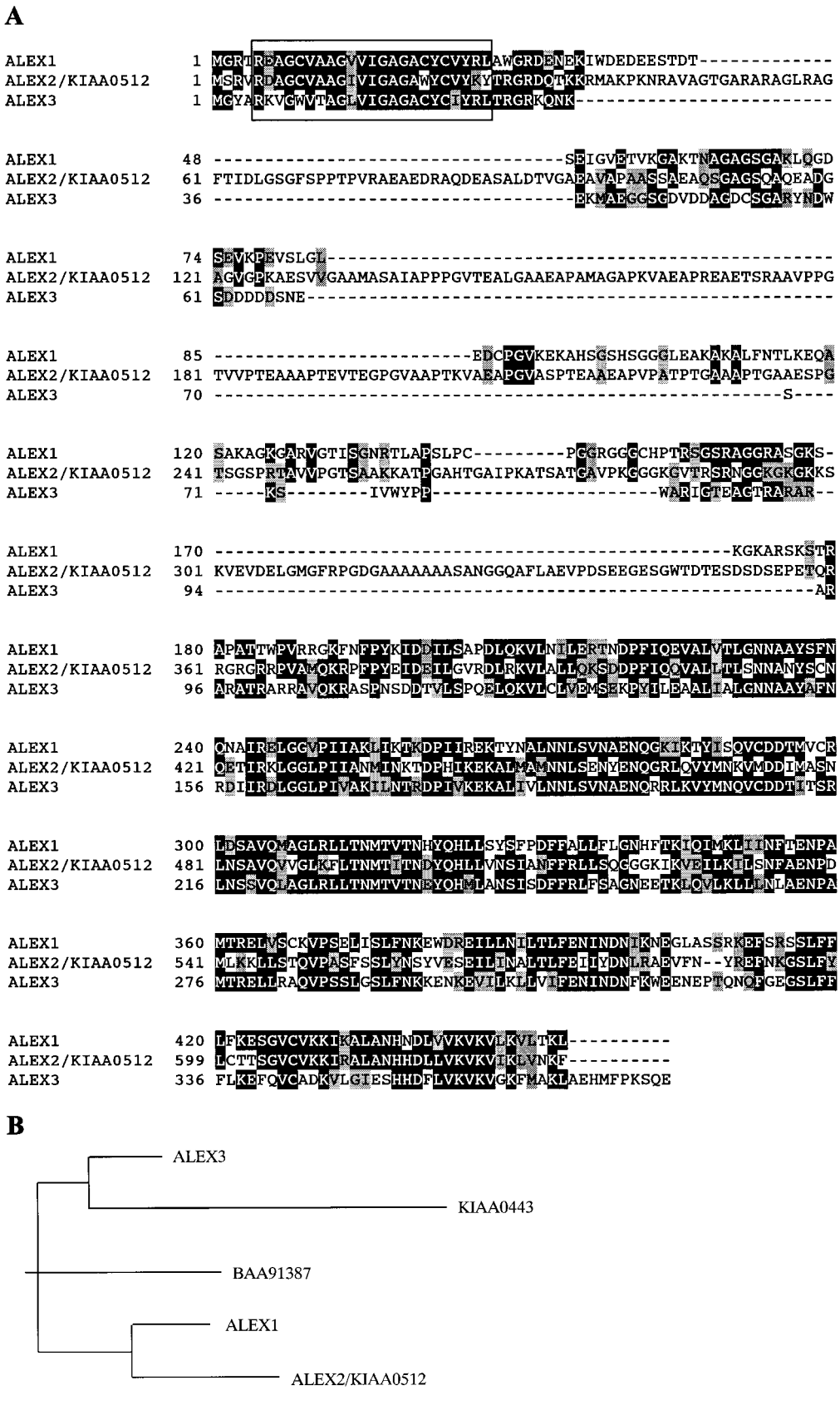

FIG. 2. (A) Multiple alignments of the human ALEX1, ALEX2 (KIAA 0512), and ALEX3 predicted protein products. Sequences were aligned by ClustalW and shaded using Boxshade 3.21. Black residues are identical and gray residues are similar. The putative transmembrane domain is boxed. (B) Phylogenetic relationship among currently known ALEX1 homologues.

of the THC257925 revealed an ORF encoding putative 342 amino acid protein with a strong homology to ALEX1 and ORF KIAA0512. To obtain a full-length cDNA for this homologue, a human testis cDNA library was screened with a probe based on available ESTs. The clone with the longest insert (2,025 bp) yielded an ORF of 379 amino acids with the first ATG initiation codon at nt 481 and a translation stop codon at nt 1618. Predicted protein, designated ALEX3, shares higher degree of homology to ALEX1 than to ALEX2 with 41% identity and 56% similarity to ALEX1 over entire length (Fig. 2A). While our work was in progress, Nicolas *et al*. (21) deposited in GenBank cDNA sequence for a new protein, a putative binding partner of SH3 domain of alpha-fodrin (GenBank AF211175). The deduced protein is identical to the ALEX3 ORF. The cDNA sequence reported earlier (1640 bp) (21), however, has a shorter 5′-UTR as compared with the ALEX3 cDNA. The N-terminal region of ALEX3 like that one in ALEX1 and ALEX2 contains a potential transmembrane domain (aa 7 to 29). Homology search revealed that ALEX1 polypeptide is also similar, although more weakly, to the C-terminal part of a large hypothetical protein KIAA0443 of 1395 amino acids (31% identity and 55% similarity over a length of 225 amino acids) and unnamed protein product (GenBank BAA91387) of 236 amino acids (33% identity and 55% similarity over a length of 195 amino acids). ALEX1 sequence was analyzed also by PSI-BLAST program. The PSI-BLAST method significantly increases the database-search sensitivity and allows the detection even of very subtle sequence similarities at a statistically significant level (12). Second iteration with E-value of 0.01 revealed that ALEX1 exhibits significant homologies to members of the arm repeat family. ALEX1 is most closely related to human importin α and yeast vacuolar protein 8 (Vac8p). Less significant homology is shared with armadillo, β-catenin and plakoglobin. The homology region is confined to the two arm repeats present in ALEX1.

A tentative phylogenetic relationship among all known ALEX1 homologues has been deduced. Figure 2B depicts the phylogenetic tree of the ALEX family constructed by the neighbor-joining method (22), based on the sequence alignment carried out with the ClustalW program (23) by eBioinformatics (http://bionav.ebioinformatics.com). A tentative phylogenetic relationship between ALEX1 homologues suggests that the ALEX gene family evolved from a common ancestor, which has undergone two successive duplications. ALEX1 is more closely related to ALEX2/KIAA0512 than to other homologues (Fig. 2B).

Mapping of ALEX1 gene. For mapping of ALEX1, primers F4 and R4 were used to screen a panel of a monochromosomal hybrid cell lines. The expected 460-bp product was amplified specifically from cell line GM/NA06318B, assigning ALEX1 to human chromosome X (data not shown). Analysis of the GeneBridge 4 radiation hybrid panel with the same primers located ALEX1 gene at the interval 0.30 cR from the marker DXS7495 (WI-8717) and 2.43 cR from the marker DXS1231. This places ALEX1 gene to a band Xq21.33–q22.2. The search of the UniGene collection at NCBI established that ALEX2 matches to the cluster Hs.48924 (marker stSG22124), while ALEX3 matches to the cluster Hs.172788 (marker stSG13135) and cDNA for homologous putative protein BAA91387 matches to the cluster Hs.83530 (marker stSG2028) (http://www.ncbi.nlm.nih.gov/UniGene). All these markers reside at the same interval, DXS990-DXS1059, which is mapped to a band Xq21.33–q22.2. BLASTN analysis revealed that the gene encoding KIAA0443 matches to a sequence from clone 769N13 derived from Xq22.1–23 (GenBank AL035427). Thus, all ALEX1 homologues are clustered at the same region on chromosome X. Comparison of cDNA and genomic sequences of ALEX2, ALEX3 and BAA91387 revealed that the coding regions of these genes reside entirely in a single exon.

Genomic structure of ALEX1 gene. BLASTN analysis against GenBank and EMBL databases revealed that the entire ALEX1 cDNA sequence is present in the clone U61B11 from the human chromosome X-specific cosmid library (GenBank Z73913). We used this genomic sequence to determine the exon/intron boundaries in ALEX1 gene. The gene is 4.2 kbp in length and is composed of four exons, ranging in size from 54 to 1892 bp, with the coding region residing entirely in a single exon (exon 4) (Fig. 1). All splice junctions conform to the consensus sequence for splice donor/acceptor sites (gt/ag).

Analysis of ALEX1 and ALEX2 expression in human normal and cancer tissues. To investigate the presence of ALEX1 and ALEX2 transcripts in human tissues, Northern blots containing poly(A)$^+$ RNAs prepared from a variety of tissues were hybridized with the probes derived from the region of least homology between the two cDNA. As can be seen in Fig. 3, when ALEX1 probe was used, a clear single band corresponding to a transcript of about 2.2-kb was detected in the majority of the tissues. The size of the transcript indicates that the cloned cDNA represents the full-length ALEX1 mRNA. When the blots were hybridized with a probe specific for ALEX2, a transcript of about 2.7-kb was detected in most tissues. An additional transcript of 1.4 kb is prominently present in the skeletal muscle, testis and placenta, and a weak band of about 7 kb is seen in the brain. The analysis revealed that ALEX1 and ALEX2 mRNA have a remarkably similar pattern of tissue expression. Both mRNAs were high in ovary, heart, testis, prostate, brain, spleen and colon. The transcripts were only barely detectable in liver and

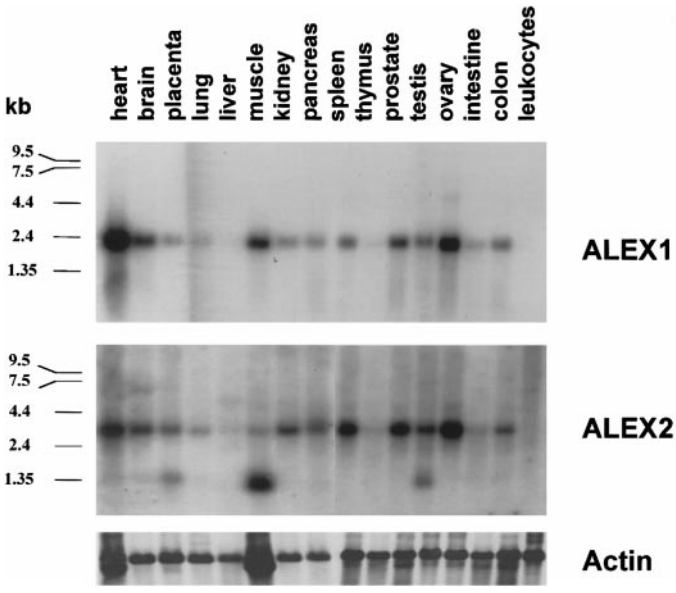

FIG. 3. Northern blot analysis of ALEX1 and ALEX2 mRNA expression. Northern blot containing 2 μg poly(A)$^+$ RNA from human adult tissues (Clontech) was hybridized to the probe from the coding region of the ALEX1 and ALEX2. Molecular size markers are shown on the left.

thymus. The expression of both mRNAs in peripheral blood leukocytes was below the limits of detection. Even highly sensitive PCR approach using first-strand cDNA as a template failed to detect ALEX1 and ALEX2 transcripts in leukocytes (not shown). Tissue distribution of ALEX3 was obtained by querying the UniGene database. Matching ESTs were detected in cDNA libraries constructed from various embryonic and adult tissues including heart, brain, placenta, lung, muscle, kidney, spleen, prostate, ovary, and colon. It is noteworthy that the ALEX3 cDNA did not match any ESTs prepared from liver, thymus and leukocyte. This suggests an expression pattern of ALEX3 similar with that of ALEX1 and ALEX2/KIAA0512.

The expression of ALEX1 and ALEX2 mRNAs was also examined in human tumor-derived samples using RT-PCR. The relative levels of the ALEX1 and ALEX2 transcripts were compared to those of the internal control gene GAPDH. The results are shown in Fig. 4A. No expression of ALEX1 and ALEX2 mRNA was detectable in two lung carcinoma samples, in a prostatic adenocarcinoma sample, and in a colon adenocarcinoma sample. In addition, the ALEX2 transcripts were not detectable in pancreas and ovarian carcinomas, while ALEX1 mRNA was significantly reduced in these samples as compared to normal tissues. The same conclusions were made when RT-PCR assays were performed at different cycle numbers and with variously diluted templates (not shown). To corroborate the results of RT-PCR, we utilized Northern blot analysis. Total RNA prepared from tumor and normal parts of the same ovaries excised from four patients were hy-

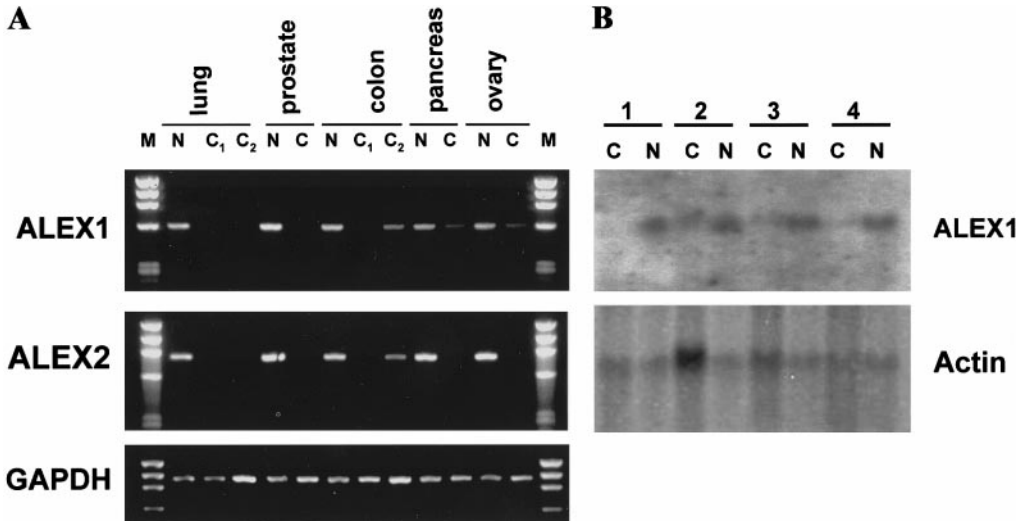

FIG. 4. Tissue expression pattern of ALEX1 and ALEX2 mRNA in normal compared with tumor samples. (A) cDNAs prepared from both normal and tumor samples (Multiple Tissue cDNA Panels, Clontech) were used as templates for PCR with the primers specific for ALEX1 and ALEX2 (see Materials and Methods). PCR products were then resolved on a 2% agarose/EtBr gel. N, normal tissues. Lung carcinomas: C_1, Human Lung Carcinoma (LX-1); C_2, Human Lung Carcinoma (GI-117). Prostate carcinoma: C, Human Prostatic Adenocarcinoma (PC3). Colon carcinomas: C_1, Human Colon Adenocarcinoma (CX-1); C_2, Human Colon Adenocarcinoma (GI-112). Pancreatic carcinoma: C, Human Pancreatic Adenocarcinoma (GI-103). Ovarian carcinoma: C, Human Ovarian Carcinoma (GI-102). DNA size markers (ϕX174 digested with *Hae*III) are shown in lane M. (B) Northern blot analysis of ALEX1 in human tumor and normal ovaries. Blot contains 20 μg of total RNA prepared from tumor and normal tissue excised at the same operational site (Human Ovarian Tumor Blot, Invitrogen). Donors: 1, serous papillary cystadenocarcinoma of left ovary (age, 48 years old); 2, serous cystadenocarcinoma (30 years); 3, granulosa-theca cell tumor (42 years); 4, adenocarcinoma (28 years). C, cancer tissue; N, normal tissue.

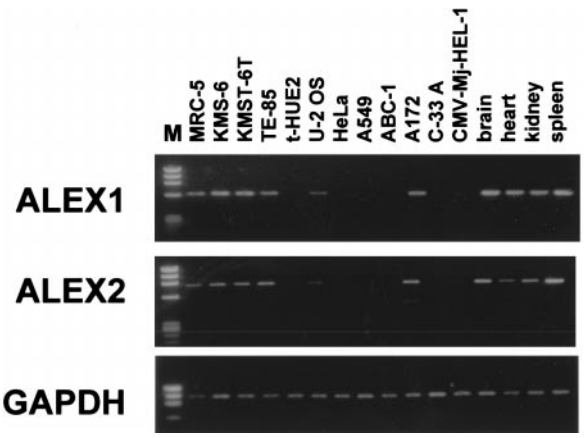

FIG. 5. Expression of ALEX1 and ALEX2 mRNA in human normal compared with transformed cell lines. cDNAs prepared from various human cell lines and normal tissues were used as templates for PCR with the same primers combinations used in Fig. 4A. PCR products were then resolved on a 2% agarose/EtBr gel. Normal tissues (Human MTC Panel, Clontech): brain, heart, kidney, and spleen. Cell lines of human origin: MRC-5, normal fetal lung diploid fibroblasts; KMS-6, normal diploid fetal fibroblasts; KMST-6T, neopastically transformed KNS-6 fibroblasts; TE-85, tumorigenic osteosarcoma; t-HUE2, immortal cell line established from endothelial cell line ECV304; U-2 OS, osteosarcoma; HeLa, cervix adenocarcinoma; A549, lung carcinoma; ABC-1, non-small cell lung carcinoma; A172, malignant glioma; C-33 A, cervical carcinoma; CMV-Mj-HEL-1, CMV-transformed embryo lung fibroblasts. DNA size markers (ϕX174 digested with *Hae*III) are shown in lane M.

bridized with the ALEX1 specific probe. We could detect ALEX1 mRNA only in normal but not in tumor parts of the same ovaries in all four patients under the study (Fig. 4B).

Expression pattern of ALEX1 and ALEX2 mRNA in human tumor-derived cell lines. Using RT-PCR, we also examined ALEX1 and ALEX2 mRNA expression in various established human cancer cell lines (Fig. 5). Four normal tissues and two normal diploid fibroblast cell lines were included as controls. Both transcripts are expressed in the glioblastoma-derived cell line A172 and the osteosarcoma cell line TE-85. Low levels of ALEX1 and ALEX2 mRNA were found in the osteosarcoma cell line U-2 OS. However, no signal was detected in the immortal endothelial cell line t-HUE2, the cervix adenocarcinoma cell line HeLa, the lung carcinoma cell line A549, the non-small cell lung carcinoma cell line ABC-1, the cervical carcinoma cell line C-33, and CMV-transformed embryo lung fibroblast cell line CMV-Mj-HEL-1. Northern blot analysis failed to detect ALEX1 mRNA also in HeLa S3 and colorectal adenocarcinoma cell line SW480 (not shown). ALEX1 and ALEX2 mRNA was analyzed in four additional cell lines. These included the normal breast epithelial cell line HBL-100, the breast adenocarcinoma cell line MDA-MB-468, and the neuroblastoma cell lines SH-SY5Y and IMR-32. ALEX1 and ALEX2 mRNA were detected in HBL-100 and both neuroblastoma cell lines but not in MDA-MB-468 (not shown). Thus, ALEX1 and ALEX2 mRNA are expressed in human sarcoma, glioblastoma and neuroblastoma-derived cell lines but are not detectable in cell lines originating from human carcinomas.

DISCUSSION

This study describes the cloning and nucleotide sequence of a cDNA encoding a novel human arm repeat protein ALEX1. BLAST analysis revealed significant similarities of ALEX1 polypeptide to uncharacterized ORF KIAA0512 and hypothetical protein encoded by unknown mRNA (GenBank AF21117), designated as ALEX2 and ALEX3, respectively. ALEX1 protein is more distantly related to two other putative proteins, KIAA0443 and BAA91387. While ALEX1 contains two arm repeats, only single arm repeat is detected in ALEX2 and ALEX3 protein sequences. So far, however, no proteins with less than six arm repeats have been described (1). Thus, ALEX1 and its homologues are not classical members of arm repeat family of proteins and constitute rather a new family of proteins. The genes for ALEX1 and its homologues are located on the same chromosomal interval Xq21.33–q22.2 suggesting that the ALEX genes evolved as a result of amplification and divergence of a common ancestral gene related to classical members of arm repeat family. Analysis of public databases allowed determining exon-intron structure of the ALEX genes. An interesting feature of these genes is that their entire coding regions reside in a single exon. The proportion of genes with intronless protein coding regions is thought to be low—at most 5% (24). This feature of the ALEX genes might have importance in not having differential and aberrant splicing that will result in higher transcriptional fidelity in the coding region.

Expression studies demonstrate that ALEX1 and ALEX2 are expressed in all human tissues examined, except leukocytes (Fig. 3). The tissue distribution of ALEX1 and ALEX2 transcripts does not provide an immediate clue as to physiological function of the encoded protein products. The lack of expression of these genes in peripheral blood leukocytes, however, suggests that ALEX1 and ALEX2 proteins might be involved in cell-to-cell signaling or establishment of cell polarity. Indeed, of all tissues examined only leukocytes exist as single cells and only leukocytes lack any long-term established polarity and display highest motility. On the other hand, tissues with maximal ALEX mRNA expression, brain, heart and skeletal muscle, consist of cells with the most striking polarized organization of any cell type in the body.

All three members of the ALEX family contain at their N-termini a hydrophobic cluster predicted to target these proteins to the membrane. Noteworthy, none of the previously described arm proteins has a transmembrane domain. Although classical members of arm

repeat family are not transmembrane proteins, armadillo homologues plakoglobin and β-catenin localize near the plasma membrane where they function in connecting adhesive junctions to the cytoskeleton (25, 26). It remains to be established whether ALEX proteins are recruited to the cell membranes and play a role in cell-cell contacts.

Remarkably, significant reduction or loss of ALEX1 mRNA in primary ovarian carcinoma samples has been demonstrated using Northern blot hybridization (Fig. 4B). Through RT-PCR, ALEX1 and ALEX2 transcripts were found to be lost or significantly reduced in carcinomas derived from several tissues types (lung, prostate, colon, pancreas, ovary) (Fig. 4A) and also in cell lines established from carcinomas (Fig. 5). ALEX1 and ALEX2 mRNA were, however, detectable in cells lines originating from neuroblastomas, gliomas or sarcomas (Fig. 5). Carcinomas, the cancers of epithelial tissue, represent about 85% of all human tumors. The loss of expression specifically in carcinomas suggests that the ALEX genes locus at chromosome band Xq21.33–q22.2 is linked to this type of malignancy. This locus may contain, therefore, either tumor suppressor gene(s) or, the ALEX proteins might be themselves involved in regulation of normal cell growth and as such play role in tumor suppression. A major focus of our current studies is comprehensive functional analysis of the ALEX genes and status of these genes in clinical material and established cancer cell lines.

ACKNOWLEDGMENTS

We thank Dr. M. Jones for providing gridded human testis cDNA library and Dr. R. Wadhwa for RNA samples isolated from various human cell lines.

REFERENCES

1. Hatzfeld, M. (1999) The armadillo family of structural proteins. *Int. Rev. Cytol.* **186,** 179–224.
2. Peifer, M., Berg, S., and Reynolds, A. B. (1994) A repeating amino acid motif shared by proteins with diverse cellular roles. *Cell* **76,** 789–791.
3. Riggleman, B., Wieschaus, E., and Schedl, P. (1989) Molecular analysis of the armadillo locus: Uniformly distributed transcripts and a protein with novel internal repeats are associated with a *Drosophila* segment polarity gene. *Genes Dev.* **3,** 96–113.
4. Ozawa, M., Terada, H., and Pedraza, C. (1995) The fourth armadillo repeat of plakoglobin (gamma-catenin) is required for its high affinity binding to the cytoplasmic domains of E-cadherin and desmosomal cadherin Dsg2, and the tumor suppressor APC protein. *J. Biochem. (Tokyo)* **118,** 1077–1082.
5. Rubinfeld, B., Souza, B., Albert, I., Munemitsu, S., and Polakis, P. (1995) The APC protein and E-cadherin form similar but independent complexes with alpha-catenin, beta-catenin, and plakoglobin. *J. Biol. Chem.* **270,** 5549–5555.
6. Troyanovsky, R. B., Chitaev, N. A., and Troyanovsky, S. M. (1996) Cadherin binding sites of plakoglobin: Localization, specificity and role in targeting to adhering junctions. *J. Cell Sci.* **109,** 3069–3078.
7. Wieschaus, E., and Riggleman, R. (1987) Autonomous requirements for the segment polarity gene armadillo during *Drosophila* embryogenesis. *Cell* **49,** 177–184.
8. Kinch, M. S., Clark, G. J., Der, C. J., and Burridge, K. (1995) Tyrosine phosphorylation regulates the adhesion of ras-transformed breast epithelia. *J. Cell. Biol.* **130,** 461–471.
9. Su, L. K., Vogelstein, B., and Kinzler, K. W. (1993) Association of the APC tumor suppressor protein with catenins. *Science* **262,** 1734–1737.
10. Kinzler, K. W., and Vogelstein, B. (1996) Lessons from hereditary colorectal cancer. *Cell* **87,** 159–170.
11. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215,** 403–410.
12. Altshul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res.* **25,** 3389–3402.
13. Bairoch, A., Bucher, P., and Hofmann, K. (1997) The PROSITE database, its status in 1997. *Nucleic Acids Res.* **25,** 217–221.
14. Bateman, A., Birney, E., Durbin, R., Eddy, S. R., Finn, R. D., and Sonnhammer, E. L. (1999) Pfam 3.1: 1313 multiple alignments match the majority of proteins. *Nucleic Acids Res.* **27,** 260–262.
15. Hirokawa, T., Boon-Chieng, S., and Mitaku, S. (1998) SOSUI: Classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14,** 378–379.
16. Dubois, B. L., and Naylor, S. L. (1993) Characterization of NIGMS human/rodent somatic cell hybrid mapping panel 2 by PCR. *Genomics* **16,** 315–319.
17. Walter, M. A., Spillet, D. J., Thomas, P., Weissenbach, J., and Goodfellow, P. N. (1994) A method for constructing radiation hybrid maps of whole genomes. *Nat. Genet.* **7,** 22–28.
18. James, P., Halladay, J., and Craig, E. A. (1996) Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144,** 1425–1436.
19. Authier, F., Bergeron, J. J. M., Ou, W.-J., Rachubinski, R. A., Posner, B. I., and Walton, P. A. (1995) Degradation of the cleaved leader peptide of thiolase by a peroxisomal proteinase. *Proc. Natl. Acad. Sci. USA* **92,** 3859–3863.
20. Kozak, M. (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res.* **12,** 857–872.
21. Nicolas, G., Galand, C., and Lecomte, M.-C. (1999) http://www.ncbi.nlm.nih.gov/entrez/.
22. Saitou, N., and Nei, M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4,** 406–425.
23. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** 4673–4680.
24. Gentles, A. J., and Karlin, S. (1999) Why are human G-protein-coupled receptors predominantly intronless? [letter]. *Trends Genet.* **15,** 47–49.
25. Franke, W. W., Goldschmidt, M. D., Zimblemann, R., Mueller, H. M., Schiller, D. L., and Cowin, P. (1989) Molecular cloning and amino acid sequence of human plakoglobin, the common junctional plaque protein. *Proc. Natl. Acad. Sci. USA* **86,** 4027–4031.
26. McCrea, P., Turck, C., and Gumbiner, B. (1991) A homologue of the armadillo protein in Drosophilla (plakoglobin) associated with E-cadherin. *Science* **254,** 1359.