J Hum Genet (1999) 44:52–56

SHORT COMMUNICATION

**Manabu Futamura · Hiroyuki Nishimori
Takayuki Shiratsuchi · Shigetoyo Saji · Yusuke Nakamura
Takashi Tokino**

Molecular cloning, mapping, and characterization of a novel human gene, *MTA1-L1*, showing homology to a metastasis-associated gene, *MTA1*

Received: June 19, 1998 / Accepted: August 5, 1998

Abstract Through large-scale sequencing of clones randomly selected from libraries of human cDNAs, we have isolated a novel human gene encoding a product with 59.6% identity in amino acid sequence to human MTA1, a protein associated with tumor invasion and metastasis. This cDNA, named *MTA1-L1* (*MTA1 like 1*), consists of 2736 nucleotides with an open reading frame encoding 668 amino acids. A single 3.0-kb transcript of *MTA1-L1* was expressed ubiquitously on Northern blots. Structural analysis of the *MTA1-L1* gene revealed 18 exons spanning 8.1 kb of genomic DNA. We assigned the *MTA1-L1* locus to chromosomal band 11q12–13.1 by fluorescence in situ hybridization.

Key words Human genome project · cDNA library screening · Cancer · Metastasis · *MTA1*

Introduction

A metastasis-associated gene, *MTA1*, was originally isolated by differential screening of a cDNA library using a rat mammary adenocarcinoma metastatic system (Toh et al. 1994, 1995). To evaluate the relevance of the human homologue of *MTA1* in progression of carcinoma in humans, Toh et al. (1997) examined levels of *MTA1* mRNA in colorectal and gastric carcinomas and showed that overexpression of the *MTA1* gene correlated with tumor invasion and the presence of metastases. They suggested that an increase in *MTA1* mRNA might serve as an indicator for assessing the malignant potential of colorectal and gastric carcinomas.

As part of our human genome project, we have been determining nucleotide sequences of randomly selected cDNA clones and comparing them with a public database to identify novel genes that are likely to possess biologically important functions and/or might be associated with human diseases (Isomura et al. 1996; Tanaka et al. 1996). Here we describe the isolation, mapping, and characterization of a novel gene, *MTA1-L1*, whose product shows significant homology to *MTA1*.

M. Futamura · H. Nishimori · T. Shiratsuchi · Y. Nakamura · T. Tokino (✉)
Laboratory of Molecular Medicine, Human Genome Center, Institute of Medical Science, The University of Tokyo, 4-6-1 Shirokanedai, Minato-ku, Tokyo 108-8639, Japan
Tel. +81-3-5449-5372; Fax +81-3-5449-5433
e-mail: tokinta@ims.u-tokyo.ac.jp

M. Futamura · S. Saji
Second Department of Surgery, Gifu University, Gifu, Japan

Materials and methods

Isolation of MTA1-L1 cDNA

The 5′ portions of cDNA clones randomly selected from a human adult-heart cDNA library were sequenced as part of the human genome project (Sudo et al. 1994, and unpublished data). Using FASTA (Pearson and Lipman 1988) and Basic Local Alignment Search Tool (BLAST, NCBI database) programs against a public database, we identified a cDNA fragment highly homologous to *MTA1*. To isolate its full-length cDNA, we screened the same library using the fragment as a probe, and determined nucleotide sequences of positive clones with an ABI Prism 377 DNA sequencer (Applied Biosystems, Foster City, CA, USA). Computer-assisted analysis of the deduced amino acid sequence was performed using MotifFinder (Kyoto University, Japan) and PSORT (Osaka University, Japan) software programs.

Northern blot analysis

Northern blots containing poly(A)+ RNA from 16 normal human tissues were purchased from Clontech (Palo Alto, CA, USA). The blots were hybridized with the random-primed, ^{32}P-labeled cDNA probe according to the supplier's recommendations. The blots were washed with 0.1 × stan-

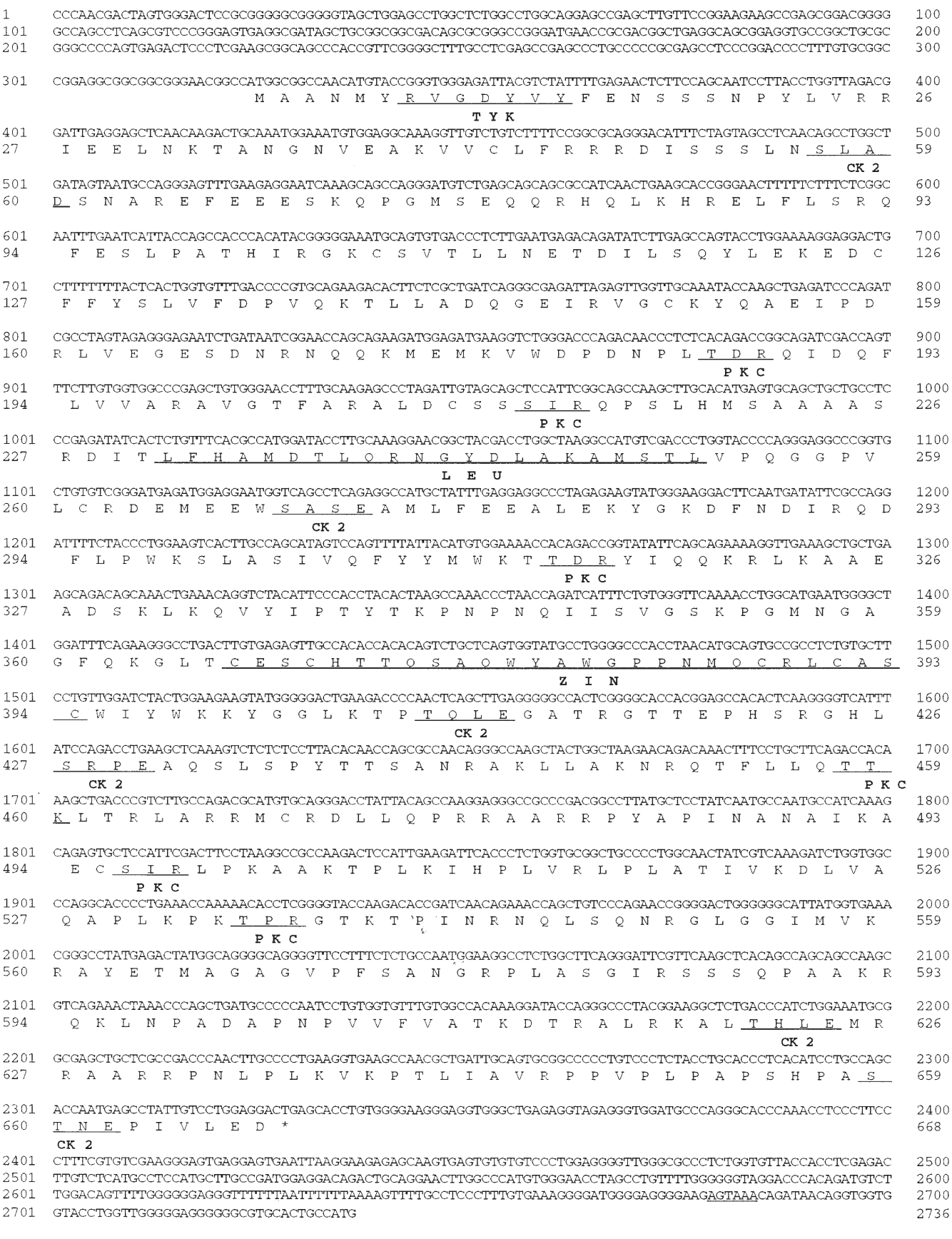

Fig. 1 Characterization of full-length *MTA1-L1* cDNA. Nucleotide sequence of the *MTA1-L1* cDNA (accession number AB016591) and the deduced amino acid sequence of the MTA1-L1 protein (one-letter code) are shown. Polyadenylation signal (AGTAAA) is *underlined*, and termination codon (TGA) is indicated by an *asterisk*. Possible phosphorylation sites for tyrosine kinase (TYK), protein kinase C (PKC), and casein kinase 2 (CK 2) are indicated, as well as a putative leucine zipper (LEU) and a zinc-finger (ZIN) DNA-binding motif

dard saline citrate (SSC), 0.1% sodium doclecylsulfate (SDS) at 65°C, and exposed for autoradiography at −80°C for 72 h.

Structural analysis of the MTA1-L1 gene

A genomic cosmid clone (cosMTA1-L1) was isolated by screening a human cosmid library with a clone of *MTA1-L1* cDNA (Sambrook et al. 1989). Sequencing of the human insert confirmed the cDNA sequence and determined the structure of the entire *MTA1-L1* gene, including the promoter region and the introns.

Fluorescence in situ hybridization (FISH)

We performed FISH using the cosmid clone as a probe, as previously described (Inazawa et al. 1993). For denaturation of the G-banding pattern, we prepared metaphase chromosomes by thymidine synchronization and bromodeoxyuridine release. Before hybridization, metaphase cells were stained with Hoechst 33258 (Sigma, St. Louis, USA) and irradiated with ultraviolet light. The probe was labeled with biotin-16-2′-deoxy-uridine-5′-triphosphate (dUTP) (Boehringer, Mannheim, Germany) by nick translation and hybridized to the denatured chromosomes. Hybridization signals were detected with FITC-avidin (Boehringer). Precise assignments of the signals were determined by visualization of the replicated G-bands.

Results and discussion

Among the cDNA clones randomly selected from an adult-heart cDNA library, one revealed significant homology to the human *MTA1* gene. To isolate a full-length cDNA, we screened the same cDNA library (1×10^6 clones) using the

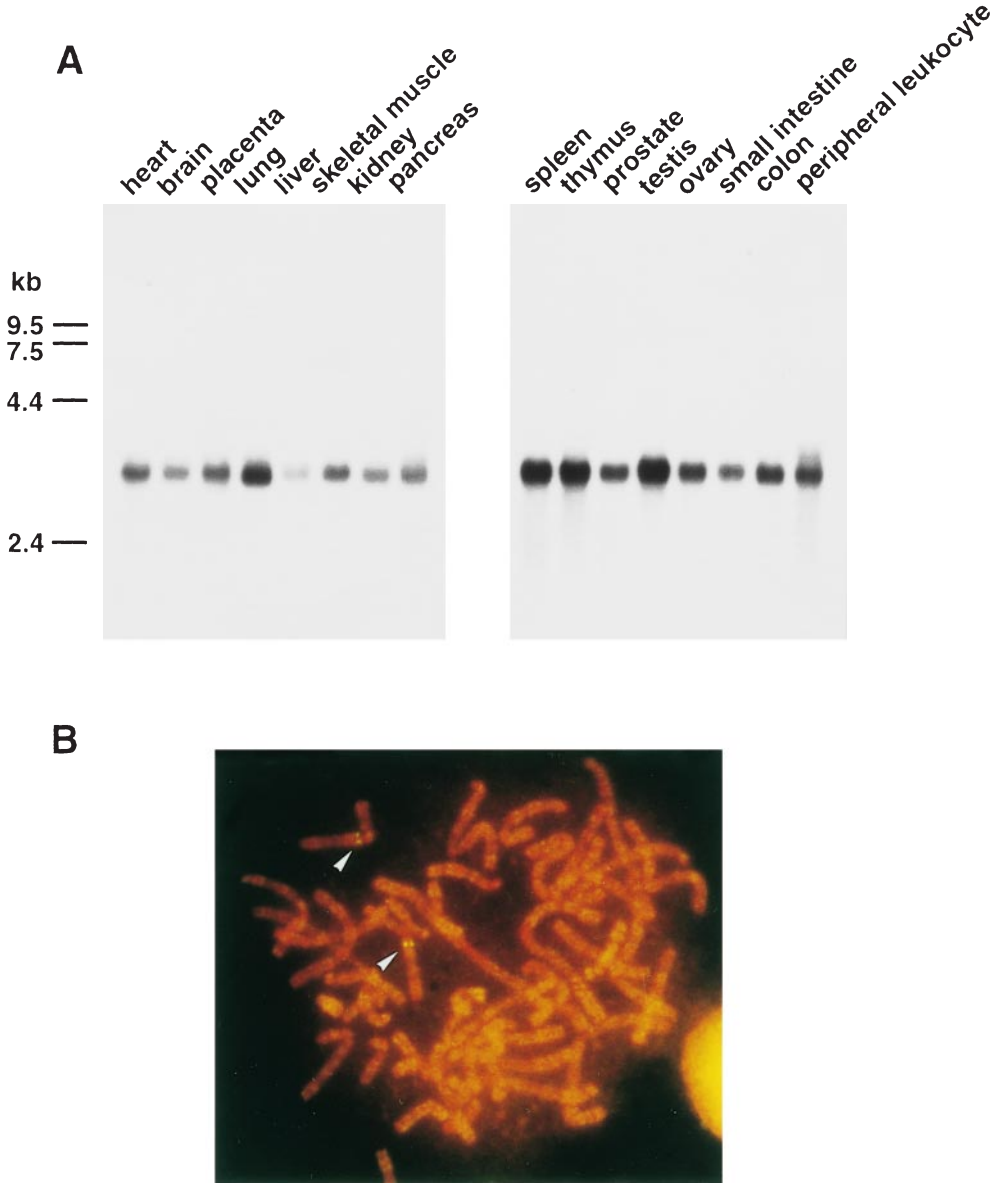

Fig. 2 **A** Multiple-tissue Northern blot of *MTA1-L1* mRNA. **B** Chromosomal localization of the *MTA1-L1* gene. Metaphase chromosomes were stained with propidium iodide, showing doublet signals (*arrowheads*)

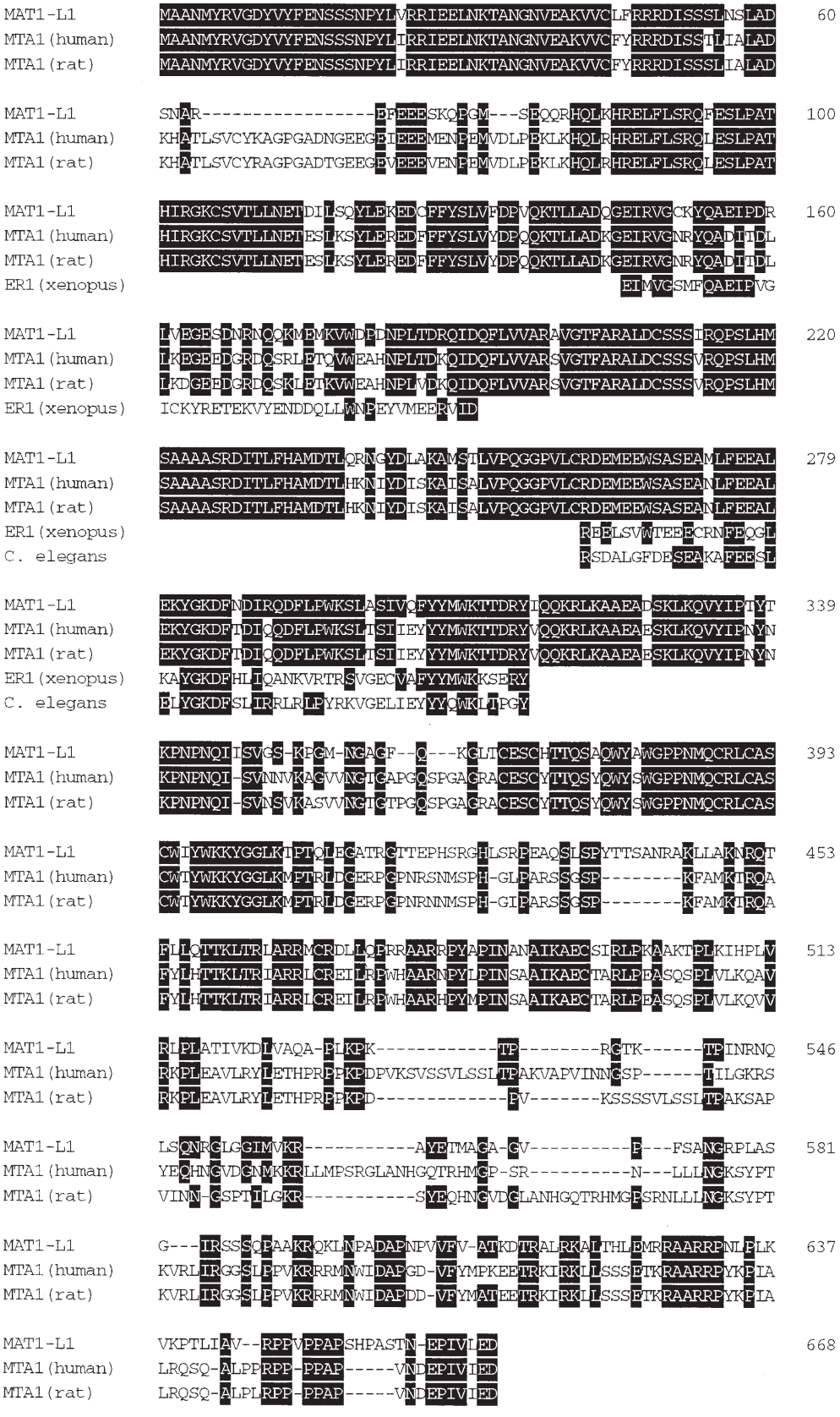

Fig. 3 Comparison of MTA1-L1 amino acid sequence with human (accession number U35113) and rat (accession number U02522) MTA1 proteins, with *Xenopus* ER1 (accession number AF015454), and with the similar-to-MTA1 protein of *Caenorhabditis elegans* (accession number U41264). Identities are indicated in *black*; gaps introduced for maximal alignment are marked with *dashes*

cDNA fragment as a probe, and isolated ten cDNA clones. Alignment of overlapping clones revealed that the cDNA, which we named *MTA1-L1 (MTA1 like 1)*, consisted of 2736 nucleotides whose open reading frame (2004 nucleotides) encoded a 668-amino acid protein (Fig. 1). Since Northern-blot analysis with a *MTA1-L1* cDNA probe identified a single 3.0-kb transcript in all the adult human tissues examined (Fig. 2A), we estimated that the nucleotide sequence for *MTA1-L1* shown in Fig. 1 encompassed almost the full-length cDNA. The expression pattern of this novel gene was ubiquitous and strong not only in testis but also in lung, spleen, and thymus in the human tissues, in contrast to that of *MTA1* in the rat, in which expression is strong in the testis but weak in other normal tissues, including brain, heart, lung, liver, and kidney. The functions of MTA1 and MTA1-L1 are probably redundant, but these two genes may be expressed in a cell-specific manner.

From a human genomic cosmid library we obtained one cosmid (cosMTA1-L1) containing the *MTA1-L1* gene. Sequencing of the human insert revealed that the *MTA1-L1* gene spanned approximately 8.1 kb of genomic DNA and consisted of 18 exons (database accession number for the entire genomic sequence, AB012922). Using cosMTA1-L1 as a probe, we performed FISH to determine the chromosomal location of *MTA1-L1*. Specific doublet signals were observed on both copies of chromosome 11 at band q12–13.1 (Fig. 2B).

Computer-assisted analysis of the amino acid sequence predicted that MTA1-L1 would contain neither an N-terminal signal sequence for transfer into the endoplasmic reticulum nor a hydrophobic domain characteristic of transmembrane proteins (Kyte and Doolittle 1982). However, several potential phosphorylation sites were present (Fig. 1). A data-base homology search revealed the similarity of predicted amino acid sequences of the MTA1-L1 gene product to the MTA1 proteins of rat (58.6% identity) and human (59.6%) (Fig. 3). The deduced MTA1-L1 protein contains 13 dispersed cysteine residues that may be capable of forming intra- and/or inter-molecular disulfide bonds to stabilize the protein structure; 11 of these cysteine residues are conserved in rat and human MTA1. The potential phosphorylation sites and a zinc-finger domain as a DNA-binding motif (amino acid residues 367-394) are imperfectly conserved between MTA1 and MTA1-L1. However, a putative leucine-zipper motif in MTA1-L1 (amino acid residues 231-252) is not present in MTA1 (Fig. 1).

The MTA1-L1 product also contains a region (amino acid residues 146-191) showing 40% homology to the *Xenopus* ER1 (early response 1) protein, and another region (amino acid residues 262–316) with 30% similarity to ER1 and to the *Caenorhabditis elegans* MTA1-like sequence (Fig. 3). *ER1* had been isolated as a novel gene whose expression increased in *Xenopus* embryo explants during mesoderm development induced by fibroblast growth factor (Paterno et al. 1997). ER1 is thought to function as a transcription factor; as its essential region for transactivation is conserved in MTA1-L1 (Fig. 3), MTA1-L1 may also be a transcription factor. In any event, the data reported here suggest that *MTA1-L1* is a member of a conserved multi-gene family and that it could play a role in cancer invasion and metastasis.

Acknowledgements We thank Keiko Okui for technical assistance.

This work was supported in part by a special Grant for Strategic Advanced Research on Cancer from the Ministry of Education, Culture, Sports, and Science of Japan.

References

Inazawa J, Saito H, Ariyama T, Nakamura Y (1993) High resolution cytogenetic mapping of 342 new cosmid markers including 43 RFLP markers on human chromosome 17 by fluorescence in situ hybridization. Genomics 17: 153–162

Isomura M, Okui K, Fujiwara T, Shin S, Nakamura Y (1996) Isolation and mapping of RAB2L, a human cDNA that encodes a protein homologous to RalGDS. Cytogenet Cell Genet 74: 263–265

Kyte J, Doolittle RE (1982) A simple method for displaying the hydropathic character of a protein. J Mol Biol 157: 105–132

Paterno GD, Li Y, Luchman HA, Ryan PJ, Gillespie LL (1997) cDNA cloning of a novel, developmentally regulated immediate early gene activated by fibroblast growth factor and encoding a nuclear protein. J Biol Chem 272: 25519–25595

Pearson WR, Lipman DJ (1988) Improved tools for biological sequence comparison. Proc Natl Acad Sci USA 85: 2444–2448

Sambrook J, Frietsch EF, Maniatis T (1989) Molecular cloning: A laboratory manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Sudo K, Chinen K, Nakamura Y (1994) 2058 expressed sequence tags (ESTs) from a human fetal lung cDNA library. Genomics 24: 276–279

Tanaka T, Inazawa J, Nakamura Y (1996) Molecular cloning of a human cDNA encoding putative cysteine protease (PRSC1) and its chromosome assignment to 14q32.1. Cytogenet Cell Genet 74: 120–123

Toh Y, Pencil SD, Nicolson GL (1994) A novel candidate metastasis-associated gene, *mta1*, differentially expressed in highly metastatic mammary adenocarcinoma cell lines. J Biol Chem 269: 22958–22963

Toh Y, Pencil SD, Nicolson G (1995) Analysis of the complete sequence of the novel metastasis-associated candidate gene, *mta1*, differentially expressed in mammary adenocarcinoma and breast cancer cell lines. Gene 159: 97–104

Toh Y, Oki E, Tokunaga E, Ohno S, Maehara Y, Nicolson GL, Sugimachi K (1997) Overexpression of the *mta1* gene in gastrointestinal carcinomas: Correlation with invasion and metastasis. Int J Cancer 74: 459–463