J Hum Genet (2001) 46:70–75

ORIGINAL ARTICLE

**Haruhito Harada · Hisaki Nagai · Nobuya Mine
Yoshie Terada · Hiromichi Fujiwara · Iwao Mikami
Michiko Tsuneizumi · Aya Yabe · Kumi Miyazaki
Takashi Yokota · Issei Imoto · Johji Inazawa
Mitsuru Emi**

Molecular cloning, tissue expression, and chromosomal assignment of a novel gene encoding a subunit of the human signal-recognition particle

Received: September 18, 2000 / Accepted: November 10, 2000

Abstract Human cancers derived from breast, esophageal, or ovarian tissues frequently show allelic losses on chromosome band 17q25. Moreover, a locus responsible for hereditary focal nonepidermolytic palmoplantar keratoderma, a condition associated with esophageal cancer (TOC; tylosis with oesophageal cancer), has been mapped to the same band. During efforts to sequence, by shotgun methods, a 1-Mb target region that we had defined as the DNA segment harboring the putative tumor suppressor gene(s) involved in these events, we identified a novel cDNA. The full-length cDNA is 2495 bp long and is expressed predominantly in skeletal muscle, heart, kidney, and placenta. The predicted product, a 627-amino-acid protein, exhibited significant sequence homology to the canine 68-kd subunit of the signal recognition particle that has been implicated in the transport of secreted and membrane proteins to the endoplasmic reticulum for proper processing. We confirmed the location of this gene at chromosome 17q25.1 by radiation-hybrid mapping and by fluorescence in situ hybridization.

Key words Signal-recognition particle · 17q25 · Tylosis with oesophageal cancer · Breast cancer · hSRP68

Introduction

Frequent observations of allelic loss in chromosomal band 17q25.1 in a variety of cancers in humans have suggested that one or more tumor suppressor genes are present in that region. We recently defined a commonly deleted region in breast cancer within a 1-cM interval flanked by D17S1603 and D17S1839 on chromosome 17q25.1 (Fukino et al. 1999). Moreover, a genetic locus for hereditary focal non-epidermolytic palmoplantar keratoderma, a condition associated with cancer of the esophagus (TOC; tylosis with oesophageal cancer), lies in an equally narrow, adjacent region flanked by D17S1839 and D17S785 (Risk et al. 1999).

On the basis of this evidence we hypothesized that genes mutated in the germline DNAs of patients affected with this syndrome may also be targets of allelic loss in sporadic tumors. Therefore we combined results of linkage studies in affected families with data on deletions in tumors and narrowed the critical region containing the genetic defect(s) within one megabase of DNA at 17q25.1. We constructed a physical map of this segment, using yeast artificial chromosome (YAC) and bacteriophage P1-derived artificial chromosome (PAC) genomic clones. To identify genes present within that region, we have been performing a large-scale, shotgun analysis of genomic DNA sequences in the vicinity. We describe here the identification in the region under investigation of a novel cDNA encoding a subunit of the human signal-recognition particle (SRP).

H. Harada · H. Nagai · N. Mine · Y. Terada · H. Fujiwara · I. Mikami · M. Tsuneizumi · A. Yabe · K. Miyazaki · T. Yokota · M. Emi (✉)
Department of Molecular Biology, Institute of Gerontology, Nippon Medical School, 1-396 Kosugi-cho, Nakahara-ku, Kawasaki 211-8533, Japan
Tel. +81-44-733-5230; Fax +81-44-733-5192
e-mail: memi@nms.ac.jp

I. Imoto · J. Inazawa
Department of Molecular Cytogenetics, Medical Research Institute, Tokyo Medical and Dental University, Tokyo, Japan

Materials and methods

Shotgun sequencing of PAC DNA

A PAC clone (105a14) from the 1-Mb target region on 17q25.1 was fragmented by sonication; 1- to 5-kb frangments were recovered and subcloned into pBluescript II SK(−) plasmid (Stratagene, La Jolla, CA, USA) (Ikegawa et al. 1999). Plasmid DNAs were isolated from 1000 randomly chosen subclones, and their DNA inserts were sequenced from both ends with an ABI PRISM 377 automated sequencer (Foster City, CA, USA). Contigs from the sequence data were assembled using Sequencher software (Hitachi Software, Yokohama, Japan). After removal of repetitive sequence elements with the RepMask program (Genome Net Database Service, Tokyo, Japan),

we searched for homologies using basic local alignment search tool (BLAST) and expressed sequence tags database (dbEST) programs (Tamari et al. 1999).

Cloning of full-length cDNA

A full-length cDNA corresponding to a sequence that showed homology to the 3′ portion of canine *SRP68* was cloned by the 5′ rapid amplification of cDNA ends (5′ RACE), method for which mRNA from normal human adult lung was reverse-transcribed by procedures described previously (Harada et al. 1999). For the 5′ RACE experiments, four gene-specific primers were designed on the basis of sequence corresponding to the 3′ portion of the cDNA. 5′ RACE reactions were performed with nested polymerase chain reactions (PCRs) after the first amplification (Kitamura et al. 1999). Fragments amplified by RACE were separated by electrophoresis in agarose gel, eluted, and subcloned into TA cloning vector (AdvanTAge PCR cloning kit; Clontech, Palo Alto, CA, USA) (Emi et al. 1999).

Northern-blot analysis

A blotted membrane containing polyA RNA from various normal human tissues was obtained from Clontech. Hybridization of the membrane was carried out according to procedures described elsewhere (Futamura et al. 1999), using a full-length cDNA as a probe.

Fluorescence in-situ hybridization (FISH)

The PAC clone carrying the human *SRP68* gene was used as a probe for FISH in the manner described by Mukae et al. (1998). The PAC DNA was labeled with biotin-16-deoxyuridine triphosphate (dUTP) by nick-translation and hybridized to denatured chromosomes at a final concentration of 25 ng/ml in 50% formamide, 10% dextran sulfate, 2 × standard saline citrate (SSC), 0.2 μg/ml Cot-1 DNA (GIBCO/BRL, Rockville, MD, USA), 2 μg/ml salmon sperm DNA, and 2 mg/ml *Escherichia coli* tRNA. Hybridization signals were detected with fluorescein isothiocyanate-avidin (Boehringer Mannheim, Mannheim, Germany). Metaphase cells were counterstained with 4′,6-diamidino-2-phenylindole, and the slides were examined through a Nikon epifluorescent microscope (Nikon, Tokyo, Japan) equipped with a charge-coupled device (CCD) camera (Photometrics, Tucson, AZ, USA). Images were captured with Quips (Vysis, Tokyo, Japan) software and processed with Adobe Photoshop 3.0 software.

Radiation-hybrid mapping

We refined the chromosomal assignment of the novel gene by PCR analysis, using the G3 RH mapping panel from the Stanford Human Genome Center (Boehnke et al. 1991). Typing was carried out using a pair of primers located in the contig obtained from genomic sequencing of PAC 105a14 (PAC105RH1F: 5′-GGTAAGAGCTCTAACTCGTA-3′, PAC105RH1R: 5′-CGTATGTAGCATCGGTCACA-3′). Results of maximum-likelihood analysis and RH mapping were obtained after submitting the raw scores to the Stanford Human Genome Center (http://www-shgc.stanford.edu).

Results

Cloning a cDNA from genomic sequence in PAC 105a14

We constructed contigs of genomic sequence from YACs and PACs within the 1-Mb segment covering the critical region of 17q25. Shotgun sequencing of one of the PAC clones (105a14) and a search of the Genbank archive revealed that a 1393-bp portion of this clone, sequence 0212, was 88% homologous to the 3′ end of the mRNA sequence of the 68-kd subunit of canine signal-recognition particles. The human counterpart, sequence 0212, was used to clone a full-length cDNA by the 5′ RACE method, using normal human adult lung as the source of mRNA. Four sequence-specific primers were prepared on the basis of sequence 0212, and several overlapping 5′ RACE experiments were carried out sequentially to verify the 5′ terminus. The full-length cDNA was 2495 bp long, consisting of 12 bp of 5′-untranslated sequence, 1881 bp of coding sequence, 580 bp of 3′-untranslated sequence, and a 22-bp poly(A) tail (Fig. 1).

The open reading frame, nucleotides 13 to 1893, encoded a peptide of 627 amino acids. Because the sequence surrounding the initiation codon agreed with the Kozak rule (Kozak 1984), the first ATG most likely represents the transcription-start site. A polyadenylation signal, ATTAAA, was present at nucleotides 2331–2336. Analysis of the predicted amino-acid sequence by the BLAST program revealed homologies to signal-recognition particles of *Schizosaccharomyces pombe* (23%), *Saccharomyces cerevisiae* (21%), *Caenorhabditis elegans* (31%), and *Canis familiaris* (96%; Fig. 2). We concluded that the novel gene at 17q25 encodes the human counterpart of the 68-kb subunit of many signal-recognition-particle complexes (hSRP68).

Expression of hSRP68 in human tissues

To examine the size and expression pattern of the complete transcript, we hybridized the full-length cDNA probe to a multiple-tissue Northern blot, which indicated a transcript of about 2.8 kb (Fig. 3). This transcript was expressed ubiquitously, but highly in skeletal muscle and preferentially in heart, kidney, and placenta. The full-length cDNA clone appeared to account for most of the size of the transcript detected by Northern-blot analysis, when the poly(A) tail was taken into account.

Chromosomal localization

PAC genomic clone 105a14 served as the probe to localize the *hSRP68* gene by fluorescent in situ hybridization on

GGCAGGGGCA AGATGGCTGC TGAGAAGCAG GTCCCAGGCG GCGGCGGCGG CGGCGGCAGT
GGCGGCGGCG GTGGCAGTGG CGGCGGCGGT AGCGGCGGTG GACGTGGTGC CGGAGGGGAA
GAAAATAAAG AAAACGAACG CCCTTCGGCC GGATCGAAGG CAAACAAAGA ATTTGGGGAT
AGCCTGAGTT TGGAGATTCT TCAGATTATT AAGGAATCCC AGCAGCAGCA TGGTTTACGG
CATGGAGATT TTCAGAGGTA CAGGGGCTAC TGTTCCCGTA GACAAAGACG TCTTCGAAAA
ACACTCAACT TCAAGATGGG TAACAGACAC AAATTCACAG GGAAGAAAGT GACTGAAGAG
CTTCTGACCG ATAATAGATA CTTGCTTCTG GTTCTGATGG ATGCTGAAAG AGCCTGGAGC
TACGCCATGC AGCTGAAACA GGAAGCCAAC ACTGAACCCC GAAAACGGTT TCACTTGTTA
TCTCGCCTAC GCAAAGCCGT GAAGCATGCA GAGGAATTGG AACGCTTGTG TGAGAGCAAT
CGCGTGGATG CCAAGACCAA ATTAGAGGCT CAGGCTTACA CAGCTTACCT CTCAGGAATG
CTACGTTTTG AACATCAAGA ATGGAAAGCT GCCATTGAGG CTTTTAACAA ATGCAAAACT
ATCTATGAGA AGCTAGCCAG TGCTTTCACA GAGGAGCAGG CTGTGCTGTA TAACCAACGT
GTGGAAGAGA TTTCACCCAA CATCCGCTAT TGTGCATATA ATATTGGGGA CCAGTCAGCC
ATCAATGAAC TCATGCAGAT GAGATTGAGG TCTGGGGGCA CTGAAGGTCT CTTGGCTGAA
AAATTGGAGG CTTTGATCAC TCAGACTCGA GCCAAACAGG CAGCTACCAT GAGTGAAGTG
GAGTGGAGAG GGAGAACGGT TCCAGTGAAG ATTGACAAAG TGCGCATTTT CTTATTAGGA
CTGGCTGATA ACGAAGCAGC TATTGTCCAG GCTGAAAGCG AAGAAACTAA GGAGCGCCTG
TTTGAATCAA TGCTCAGCGA GTGTCGGGAC GCCATCCAGG TGGTTCGGGA GGAGCTCAAG
CCAGATCAGA AACAGAGAGA TTATATCCTT GAAGGAGAGC CAGGGAAGGT GTCTAATCTT
CAATACTTGC ATAGCTACCT GACTTACATC AAGCTATCAA CGGCAATCAA GCGTAATGAG
AACATGGCCA AAGGTCTGCA GAGGGCTCTG CTGCAGCAGC AGCCAGAGGA TGACAGCAAG
CGCTCACCCC GGCCCCAGGA CCTGATCCGA CTCTATGACA TCATCTTACA GAATCTGGTG
GAATTGCTCC AGCTTCCTGG TTTAGAGGAA GACAAAGCCT TCCAGAAAGA GATAGGCCTC
AAGACTCTGG TGTTCAAAGC TTACAGGTGT TTTTTCATTG CTCAGTCCTA TGTGCTGGTG
AAGAAGTGGA GCGAAGCCCT TGTCCTGTAT GACAGAGTCC TGAAATATGC AAATGAAGTA
AATTCTGATG CTGGCGCCTT CAAGAACAGC CTAAAGGACC TGCCTGATGT GCAAGAGCTC
ATCACTCAAG TGCGGTCAGA GAAGTGCTCC CTGCAGGCCG CAGCCATCCT TGATGCAAAC
GACGCTCATC AAACAGAGAC CTCCTCCTCC CAAGTCAAGG ACAATAAGCC TCTGGTTGAA
CGGTTTGAGA CATTCTGCCT GGACCCTTCC CTTGTCACCA AGCAAGCCAA CCTTGTGCAC
TTCCCACCAG GCTTCCAGCC CATTCCCTGC AAGCCTTTGT TCTTTGACCT GGCCCTCAAC
CATGTGGCTT TCCCACCCCT TGAGGACAAG TTGGAACAGA AGACCAAGAG TGGCCTCACT
GGATACATCA AGGGCATCTT TGGATTCAGG AGCTAACCAG GCTCTTCCTC GGGGGCGGGG
GAGATTCTGA CTCTTAATCT GTATTGTGAG AAAATCCCAG CAAGTTCCAT GATATTAAAT
CCAGGTCTGC ATTGGCCCGG GGCAAGAGTT TAACATCTTC GGCCCTGCAT TCCTACATCT
TGTGTCTGTA CACGTTCTTA AGCAGCGTGT CAGGAGAGCA CCCTGTTGTC TTCTGGTAAA
TGTGTGCAGG GTCATCCTGT CTCCTGTACC TCCTGGGAAA GGGGCCGCTG CTGTCTGGTG
CCCTGTGAGC TGTGATTGAT TGCCTTTGGT CAAGTAATGC GTTCAGGAGT CCACACCAGG
CACAGATGGG GCCTTGAAAC GCTTTGTCAT GCTTCTTCAG TACCATGGAT TTGAAATGAA
CTCATYCTTG CTGTGAGCAT TCCAGGARCC CTTGAGAAGT TATCTATACT ATTAAACTGG
SAACGTYACC CCAAAAATAC GGGAGGCTTA TTTCCCTTAC CTCCAAGGAA CCTAAAARTT
TAAACAACCR AAAGCTTTTT TTTARAAAAA ATTTTCCCTC CAAADGTTGA ACCATTTTTT
TTAAAGGGGG TYCAAAAAAA AAAAAAAAAA AAAAA

Fig. 1. cDNA sequence of the *hSRP 68* gene; *boldface type* indicates the coding region. Sequence homologous to an archived canine expressed sequence tag (EST) in GenBank is *underlined*

Fig. 2. Comparison of SRP68 protein sequences among five species. *SCH.POM.1, Schizosaccharomyces pombe* (GenBank accession no. CAA20671); *SAC.CER.1, Saccharomyces cerevisiae* (Brown et al. 1994); *CAE.ELE.1, Caenorhabditis elegans* (Washington University Genome Sequencing Center, 1998); *CAN.SPE.1, Canis familiaris* (Herz et al. 1990); and *Homo sapiens* (*Hom.SAP. 2*; this report). Homologies are evident only in the carboxy-terminal regions

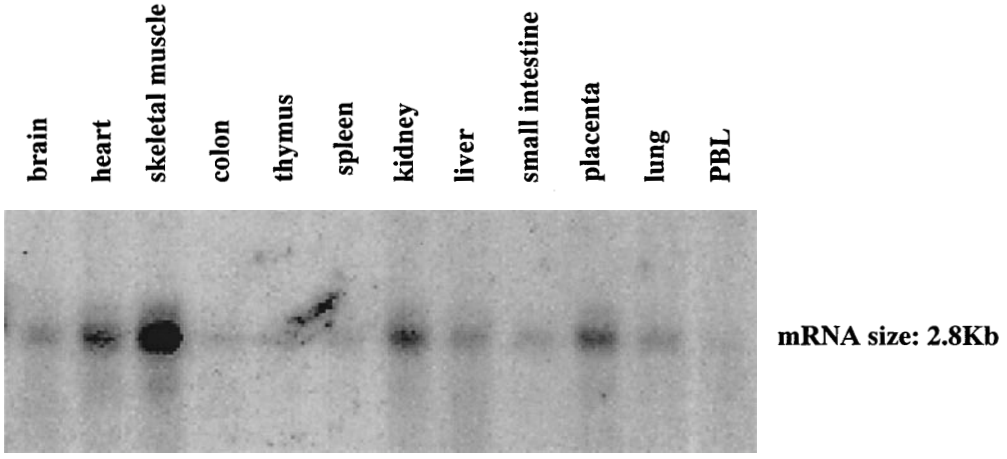

Fig. 3. Northern-blot analysis of the *hSRP68* gene in normal adult tissues. Each lane contains approximately 1 μg of polyA RNA. *PBL*, Peripheral blood lymphocytes

metaphase chromosomes. Clear fluorescent signals were present on chromosomal band 17q25 (data not shown). We mapped the gene more precisely using the G3 RH mapping panel of 83 hybrid cell lines from the Stanford Human Genome Center (Boehnke et al. 1991). The data vector for the *hSRP68* gene was 0110001110 1100RR0000 0111101011 1011110000 0011100011 1011101101 1001110010 0111101100 101, and the consequent report indicated that the gene was mapped 17q25.1 through linkage to two markers (D17S1839 and D17S2113) that had been mapped cytogenetically to 17q25.1. The position of this gene is 3 cR proximal to marker SHGC-30509 (dbSTS accession no. G27143) by linkage with a logarithm of differences (LOD) score of more than 21. These two findings strongly suggest that the *hSRP68* gene exists on chromosomal segment 17q25 and support our physical mapping data indicating that the PAC 105a14 insert represents genomic DNA located at 17q25.1.

Discussion

We have described here the isolation, tissue expression, and chromosomal mapping of a cDNA encoding the 68-kd subunit of the human signal recognition particle, after sequencing a PAC clone from within a 1-Mb segment of DNA at 17q25.1. This region is a target of allelic losses in breast and esophageal cancers, as well as being the inferred location for genetic mutations that confer predisposition to TOC. Together with other subunits of the SRP complex, the 68-kd molecule participates in the translation of membrane and secreted proteins to the endoplasmic reticulum for proper processing. The high degree of evolutionary conservation that this subunit exhibits across widely divergent species appears to indicate that the SRP complex plays an essential role in the intracellular trafficking of protein molecules in all living organisms.

While this manuscript was being prepared, a human cDNA sequence also homologous to canine *SRP68* was submitted directly to Genbank by J. C. Politz. That sequence is 97% homologous to the cDNA isolated in the present study, the major distinction being that both our cDNA and the canine cDNA for the SRP 68-kd subunit encode glutamic acid at codon 174, aspartic acid at codon 534, and alanine at codon 535, while Politz's cDNA encodes lysine, glutamate, and serine, respectively, at those positions. Also, the Politz sequence encodes a peptide shorter by eight amino acids (619 total), as it lacks Ser-Gly-Gly-Gly-Gly residues at codons 16–20 and Gly-Gly-Ser residues at codon 25–27, when compared with the product predicted from our cDNA sequence. Clarification of whether Politz's cDNA and ours derive from two related but distinct genes, are spliced differently from a single gene, or harbor sequence differences as polymorphisms, awaits characterization of Politz's cDNA through chromosomal mapping and examination of its expression pattern. Meanwhile, we plan to evaluate the exon-intron boundary sequences and to pursue functional analysis of our cDNA as steps toward discovering whether breast and/or esophageal cancers might show altered expression, or whether patients with TOC carry germline mutations of this gene.

Acknowledgments This work was supported in part by special grants for Strategic Advanced Research and for Scientific Research on Priority Area on "Cancer" and "Genome Science" and a Grant-in-Aid for Scientific Research on Priority Area (C) from the Ministry of Education, Science, Sports, and Culture of Japan; by a Research Grant from the Ministry of Health and Welfare of Japan; and by a Research for the Future Program Grant of The Japan Society for the Promotion of Science.

References

Boehnke M, Lang K, Cox DR (1991) Statistical methods for multipoint radiation mapping. Am J Hum Genet 49:1174–1188

Brown JD, Hann BC, Medzihradszky KF, Niwa M, Burlingame AL, Walter P (1994) Subunits of the *Saccharomyces cerevisiae* signal recognition particle required for its functional expression. EMBO J 13:4390–4400

Emi M, Keicho N, Tokunaga K, Katsumata H, Souma S, Nakata K, Taguchi K, Ohishi N, Azuma A, Kudoh S (1999) Association of diffuse panbronchiolitis with microsatellite polymorphism at the human interleukin 8 (IL-8) locus. J Hum Genet 44:169–172

Fukino K, Iida A, Teramoto A, Sakamoto G, Kasumi F, Nakamura Y, Emi M (1999) Frequent allelic loss at the TOC locus on 17q25.1 in primary breast cancers. Genes Chromosom Cancer 24:345–350

Futamura M, Nishimori H, Shiratsuchi T, Saji S, Nakamura Y, Tokino T (1999) Moleclar cloning, mapping, and chracterization of a novel human gene, MTA1-L1, showing homology to metastasis-associated gene, MTA1. J Hum Genet 44:52–56

Harada Y, Ozaki K, Suzuki M, Fujiwara T, Takahashi E, Nakamura Y, Tanigami A (1999) Complete cDNA sequence and genomic organization of human pancreas-specific gene homologous to *Caenorhabditis elegans* sel-1. J Hum Genet 44:330–336.

Herz J, Flint N. Stanley K, Frank R, Dobberstein B (1990) The 68 kDa protein of signal recognition particle contains a glycine-rich region also found in certain RNA-binding proteins. FEBS Lett 276:103–107.

Ikegawa S, Isomura M, Koshizuka Y, Nakamura Y (1999) Cloning and characterization of human and mouse PROSC (proline synthetase co-transcribed) genes. J Hum Genet 44:337–342

Kitamura Y, Nakata T, Minobe K, Shimizu K, Tanaka S, Fujimori M, Yokoyama Y, Ito K, Onda M, Emi M (1999) Ret/PTC3 is the most frequent form of gene rearrangement in papillary thyroid carcinomas in Japan. J Hum Genet 44:96–102

Kozak M (1984) Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. Nucleic Acids Res 12:857–872

Mukae N, Enari M, Sakahira H, Fukuda Y, Inazawa J, Toh H, Nagata S (1998) Molecular cloning and characterization of human caspase-activated DNase. Proc Natl Acad Sci USA 95:9123–9128

Risk JM, Ruhrberg C, Hennies H, Mills HS, Di Colandrea T, Evans KE, Ellis A, Watt FM, Bishop DT, Spurr NK, Stevens HP, Leigh IM, Reis A, Kelsell DP, Field JK (1999) Envoplakin, a possible candidate gene for focal NEPPK/esophageal cancer (TOC): the integration of genetic and physical maps of the TOC region on 17q25. Genomics. 59:234–242

Tamari M, Daigo Y, Nakamura Y (1999) Isolation and characterization of a novel serine threonine kinase gene on chromosome 3p22–21.3. J Hum Genet 44:116–120

Washington University Genome Sequencing Center (1998) Genome sequence of the nematode *C. elegans*: a platform for investigating biology. The *C. elegans* Sequencing Consortium. Science 282:2012–2018