J Hum Genet (1999) 44:185-189

SHORT COMMUNICATION

Naohiko Seki · Atsushi Hattori · Akiko Hayashi
Sumie Kozuma · Nobuyuki Miyajima · Toshiyuki Saito

Cloning, tissue expression, and chromosomal assignment of human MRJ gene for a member of the DNAJ protein family

Received: November 16, 1998 / Accepted: January 22, 1999

Abstract The DnaJ protein family consists of proteins with a highly conserved amino acid stretch called the "J - domain". A cDNA clone encoding a new protein with a J-domain was isolated from a human fetal brain cDNA library. This new member of the DnaJ family of 241 amino acid residues showed 94% identity with mouse Mrj (accession number, AF035962) and 71% identity with mouse Msj-1 (accession number, U95607) along its entire sequence. Reverse transcription - coupled polymerase chain reaction (RT-PCR) analysis showed the messenger RNA was ubiquitously expressed in various human tissues. The chromosomal location of the gene was determined by PCR-based analyses with both a human/rodent monochromosomal hybrid cell panel and a radiation hybrid panel to map on chromosome 11q25 region.

Key words DnaJ · MRJ · MSJ-1 · J-domain · Chromosome 11q25

Introduction

In *Escherichia coli* (*E. Coli*), the heat shock protein DnaJ is known to function together with DnaK and GrpE as a molecular chaperon, that are associated with protein folding, proteolysis, phosphorylation and replication of phage (Caplan et al. 1993, Cyr et al. 1994, Silver and Way 1993). The DnaJ chaperon was initially identified as a heat shock gene product in *E. Coli*, an organism in which the *DnaJ* gene constitutes an operon with the *DnaK* gene, which encodes the prokaryotic analog of the Hsp70 cognate protein of eukaryotes (Ohki et al. 1986). Several *DnaJ* homologs have been identified in organisms ranging from yeast to mouse and human (Lindquist and Craig 1988, Cheetham et al. 1992, Berruti et al. 1998). A sequence comparison among these DnaJ-like proteins showed that a typical DnaJ protein is composed of four modules including a highly conserved 70-amino acid region at the N-terminal end (J-domain), a glycine-rich region (G-domain) of about 30 residues, a domain containing four Cys-X-X-Cys-X-Gly-X-Gly motifs, and a low similarity region at the C-terminal region of 120-170 residues (Bork et al. 1992).

To date, at least seven human DnaJ family genes have been identified in public databases (UniGene, http://www.ncbi.nlm.nih.gov/UniGene/). We searched for other expressed sequence tags (ESTs) showing significant homology to DnaJ family genes against the dbEST database (http://www.ncbi.nlm.nih.gov/dbEST/) and found several ESTs entries. An array of cDNA fragments covering the entire coding region for a new DnaJ-related protein was obtained.

Isolation of cDNA clone for novel human MRJ protein

A total of five ESTs (R39700, AA136981, AA223981, AA223975, and AA804461) were selected to design the primer for the 5′-RACE experiment. Their consensus sequence (AGGAAGAATGCACCAACACATTCGCAT) adjacent to the putative poly (A) was used for the primer. Multiple RACE reactions were performed using testis mRNA according the manufacturer's manual (Clontech, Palo Alto, CA, USA). No nested PCR was performed because a single round of RACE reaction gave an apparent signal of amplified products of about 1.6 kb on an agarose gel. The consensus sequence of the independent RACE products was employed in the definitive structure of MRJ cDNA. The determined nucleotide sequence and deduced

N. Seki · A. Hattori · A. Hayashi · S. Kozuma · T. Saito (✉)
Genome Research Group, National Institute of Radiological Sciences, 4-9-1 Anagawa, Inage-ku, Chiba 263-8555, Japan
Tel. +81-43-206-3135; Fax +81-43-251-9818
e-mail: t_saito@nirs.go.jp

A. Hattori
Biotechnology and Medical Engineering Field, Aisin Cosmos, Tokyo, Japan

N. Miyajima
Laboratory of Genome Information, Kazusa DNA Research Institute, Chiba, Japan

```
cctggagctgtgaggagattcgggccgtcaccctgcctcccctgcgtcccgccaccggcc    60
gcttctgtcctcggacccattccaacaatctcgtaaaacATGGTGGATTACTATGAAGTT    120
                                          M  V  D  Y  Y  E  V
CTAGGCGTGCAGAGACATGCCTCACCCGAGGATATTAAAAAGGCATATCGGAAACTGGCA    180
L  G  V  Q  R  H  A  S  P  E  D  I  K  K  A  Y  R  K  L  A
CTGAAGTGGCATCCAGATAAAAATCCTGAGAATAAAGAAGAAGCAGAGAGAAAATTCAAG    240
L  K  W  H  P  D  K  N  P  E  N  K  E  E  A  E  R  K  F  K
CAAGTAGCGGAGGCATATGAAGTGCTGTCGGATGCTAAGAAACGGGACATCTATGACAAA    300
Q  V  A  E  A  Y  E  V  L  S  D  A  K  K  R  D  I  Y  D  K
TATGGCAAAGAAGGATTAAATGGTGGAGGAGGAGGTGGAAGTCATTTTGACAGTCCATTT    360
Y  G  K  E  G  L  N  G  G  G  G  G  S  H  F  D  S  P  F
GAATTTGGCTTCACATTCCGTAACCCAGATGATGTCTTCAGGGAATTTTTTGGTGGAAGG    420
E  F  G  F  T  F  R  N  P  D  D  V  F  R  E  F  F  G  G  R
GACCCATTTTCATTTGACTTCTTTGAAGACCCTTTTGAGGACTTCTTTGGGAATCGAAGG    480
D  P  F  S  F  D  F  F  E  D  P  F  E  D  F  F  G  N  R  R
GGTCCCCGAGGAAGCAGAAGCCGAGGGACGGGGTCGTTTTTCTCTGCGTTCAGTGGATTT    540
G  P  R  G  S  R  S  R  G  T  G  S  F  F  S  A  F  S  G  F
CCGTCTTTTGGAAGTGGATTTTCTTCTTTTGATACAGGATTTACTTCATTTGGGTCACTA    600
P  S  F  G  S  G  F  S  S  F  D  T  G  F  T  S  F  G  S  L
GGTCACGGGGGCCTCACTTCATTCTCTTCCACGTCATTTGGTGGTAGTGGCATGGGCAAC    660
G  H  G  G  L  T  S  F  S  S  T  S  F  G  G  S  G  M  G  N
TTCAAATCGATATCAACTTCAACTAAAATGGTTAATGGCAGAAAAATCACTACAAAGAGA    720
F  K  S  I  S  T  S  T  K  M  V  N  G  R  K  I  T  T  K  R
ATTGTCGAGAACGGTCAAGAAAGAGTAGAAGTTGAAGAAGATGGCCAGTTAAAGTCCTTA    780
I  V  E  N  G  Q  E  R  V  E  V  E  E  D  G  Q  L  K  S  L
ACAATAAATGGTAAGGAGCAGCTGCTGCGCTTGGATAACAAGTAAttcaacgcacgcact    840
T  I  N  G  K  E  Q  L  L  R  L  D  N  K  *
taacagaaatgttaaactataacaagcaccatttgaggattaacaggaacatttttttga    900
agatttcaaacgaactcgactttcagtataattgtacctaaagtatttataaacagctca    960
tcggagcctctatttgtcatagacttttgagttgattgttgggaccacataataggacca   1020
ttttttttttgtctttaaaattgttgtaaatctctgtatgcactttgctttttattaaa   1080
cgtactccaaggtgagtcttgactctttagtgtaggacaagattgtacactaacaccagc   1140
atggacctgcttttcattgtgtctgaaatgtgagccacgtagtgtcggcctgctgtgaag   1200
ttaacattgccaggacgattcttctacagaaataatttcaatttttttcagtatttagta   1260
gtgaaagatattaatacattaatggtaatacatttctggtttaatataaattaaggatgt   1320
tttctagttgtgcatgaatgctggcaacttagtaagttttgacaattgtttaaatatgta   1390
atgttaagcttaggtttaaaaaagtaaagctggtaaactgggtctttgtcatttgcttta   1450
aaaaaaaaaaaaaagaaaataaatgcgaatgtgttggtgcattc
```

Fig. 1. The human *MRJ* cDNA sequence and predicted protein. The J-motif is *underlined* and the putative polyadenylation signal is *double-underlined*. The sequence have been submitted with accession number AB014888. An *asterisk* denotes the termination codon

amino acid sequence are shown in Fig.1. The cDNA contains a total length of 1485 bp, including a 5′-noncoding region of 100 bp, an open reading frame of 726 bp, and a 3′-noncoding region of 659 bp. The open reading frame encodes a protein of 241 amino acid residues with a calculated molecular weight of 37 kDa. A canonical poly-adenylation signal, AATAAA, was located 21 bp upstream of the poly (A) (Fig.1). The nucleotide sequence data reported here will appear in the DDBJ, EMBL, and GenBank nucleotide sequence databases with the accession number of AB014888. The predicted protein contains a J-motif (residues 46-65) and a canonical J-domain (residues 2-75). Ho-

```
hMRJ        1  MVDYYEVLGVQRHASPEDIKKAYRKLALKWHPDKNPENKEEAERKFKQVAEAYEVLSDAK
mMRJ        1  MVDYYEVLGVQRHASPEDIKKAYRKQALKWHPDKNPENKEEAERKFKQVAEAYEVLSDAK
mMSJ-1      1  MVDYYEVLGVPRQASAEAIRKAYRKLALKWHPDKNPEHKEEAERRFKQVAQAYEVLSDVR

hMRJ       61  KRDIYDKYGKEG.LNGGGGGGSHFDSPFEFGFTFRNPDDVFREFFGGRDPFSFDFF.EDP
mMRJ       61  KRDIYDKYGKEG.LNGGGGGGSHFDSPFEFGFTFRNPDDVFREFFGGRDPFSFDFF.EDP
mMSJ-1     61  KREVYDRCGEVGEVGGGGAAGSPFHDAFQYVFSFRDPAEVFREFFGGHDPFSFDFFGGDP

hMRJ      119  FEDFFGNRRGPRGSRSRGTGSFFSAFSGFPSFGSGFSSFDTGFTSFGSLGHGGLTSFSST
mMRJ      119  FDDFFGNRRGPRGNRSRGAAPFFSTFSGFPSFGSGFPAFDTGFTPFGSLGHGGLTSFSST
mMSJ-1    121  LENFFGDRRSTRGSRSRGAVPFSTSFTEFPGFGGGFASLDTGFTSFGSPGNSGLSSF.SM

hMRJ      179  SFGGSGMGNFKSISTSTKMVNGRKITTKRIVENGQERVEVEEDGQLKSLTINGKEQLLRL
mMRJ      179  SFGGSGMGNFKSISTSTKIVNGKKITTKRIVENGQERVEVEEDGQLKPLTINGKEHLLRL
mMSJ-1    180  SCGGGAAGNYKSVSTSTEIINGKKITTKRIVENGQERVEVEEDGELKSLIINGREQLLRI

hMRJ      239  DNK*
mMRJ      239  DNK*
mMSJ-1    240  NTQ*
```

Fig. 2. Alignment of human MRJ (hMRJ; accession number, AB014888), mouse Mrj (mMRJ; accession number, AF035962), and mouse Msj-1 (mMSJ-1; accession number, U95607). Identities are indicated by *black background* and similar residues are *shadowed*. *Asterisk* denotes the termination codon

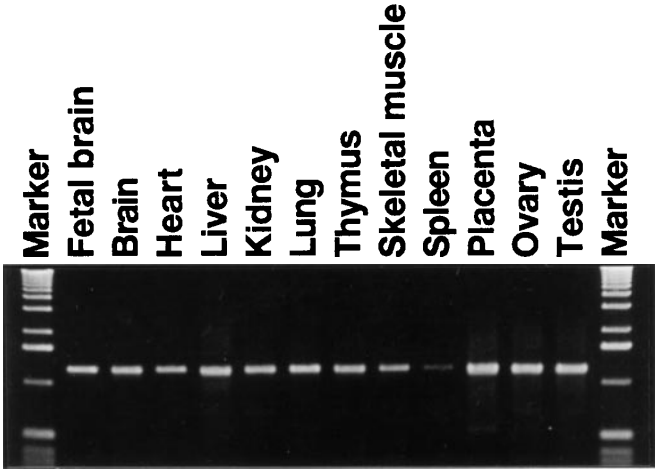

Fig. 3. Tissue distribution analysis of human *MRJ* transcript using reverse transcription-coupled polymerase chain reaction. The 12 tissues examined are indicated *above each lane*

mology searches indicated that the deduced amino acid sequence shares homology with mouse Mrj (accession number, AF035962, 94% identity, 96% similarity) and mouse Msj-1 (accession number, U95607, 71% identity, 80% similarity) (Fig.2). From these results, we concluded that the obtained cDNA encodes the human MRJ protein. The protein contains the two structural features of the J-domain at the amino terminus region and the glycine-rich region distal to the J-domain, but not the cysteine-rich repeats in the middle of the protein. The J-domain is a conserved sequence of about 70 amino acids, thought to be a site of functional interaction between the DnaJ-like protein and the Hsp70 homolog. While the J-domain is present in all DnaJ-like proteins, the other sequence features, the G-domain and cysteine-rich sequence, are less well conserved (Lindquist and Craig 1988).

Tissue expression of the human *MRJ* gene

Since the mouse *Mrj* gene has been reported only in the nucleotide databases, no information on its expression profile is available. We examined the distribution of human MRJ transcript in various human tissues by reverse-transcription -coupled chain reaction (RT-PCR) according to a previously described method (Saito et al. 1998, Seki et al. 1998). The primer set used for the RT-PCR covered the whole protein coding region of the messenger (5′-CTGGAGCTGTGAGGAGATTC-3′ and 5′-CTGTAGAAGAATCGTCCTGGC-3). No amplification was detected from human genomic DNA due to multiple exons (data not shown). A clear common signal of the expected size (1228-bp-long) was detected in all the tissues examined (Fig. 3), indicating that the *MRJ* gene is ubiquitously transcribed in various tissues and would be involved in the basic housekeeping function of cells. In contrast, a very close relative of *MRJ*, mouse Msj-1, is specifically expressed in testicular germ cells (Berruti et al. 1998). It remains to be investigated how these two highly conserved isologs share or alternate their function in cells.

Chromosomal assignment of the human *MRJ* gene

Chromosomal assignment of the *MRJ* gene was done by PCR analysis of a human/rodent somatic cell hybrid panel, and a radiation hybrid panel as described previously (Saito et al. 1997, 1998, Seki et al. 1997, 1998). The specific amplified PCR primers were designed at the 3′-untranslated region of the gene (5′-ACCTGCTTTTCATTGTGTCTG-3′, 5′-CTGTAGAAGAATCGTCCTGGC-3′, the PCR product size was 116 bp). First, a specific amplified product for human was detected only from the hybrid containing

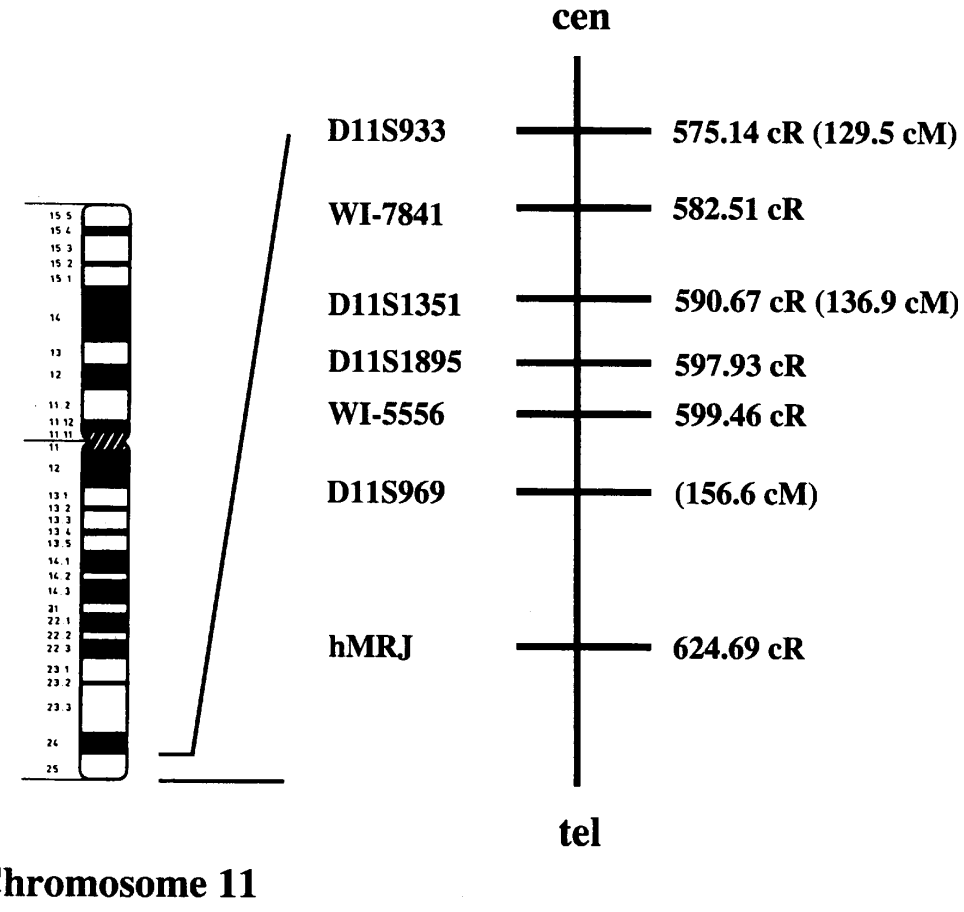

Fig. 4. Chromosomal placement of the human *MRJ* gene at a relative distance to framework markers on the WICGR (Whitehead Institute for Biomedical Research/MIT Center for Genome Research) radiation hybrid map of the human genome (http://carbon.wi.mit.edu:8000/cgi-bin/contig/phys_map). The approximate corresponding cytogenetic location of the gene on the long arm of the telomeric region of chromosome 11 is indicated. Distances of the markers are in centirays (cR) and centimorgans (cM) from the top of the chromosome 11 linkage group

human chromosome 11 (data not shown). Then, we performed further mapping analysis, using a PCR-based radiation hybrid panel (Genebridge 4; Research Genetics, Huntsville, AL, USA) with the same primers as those used in the assay for the human/rodent somatic cell hybrid panel. Statistical analysis of the radiation hybrid data was performed using the RHMAPPER software package (http://carbon.wi.mit.edu:8000/cgi-bin/contig/rhmapper.pl). The data vector for the human *MRJ* gene was 1010110110 0000110110 0100010010 0000000111 0110111001 1110110010 0010111010 0001000010 1101001000 101. Although the telomeric region contains few markers and the linkage framework of the radiation hybrid panel has relatively low resolution, the consequent report indicated that the gene was placed to 25.23 cR distal from WI-5556 (lod>3.0), which was cytogenetically mapped to the 11q25 region (Fig.4). A significant linkage of the following diseases is found to be assigned to this chromosomal region by referring to Online Mendelian Inheritance in Man (OMIM, http://www. ncbi. nlm. nih.gov/omim) and the Entrez database (http://www.ncbi. nlm.nih.gov/Entrez). A hereditary disease, histiocytosis with joint contractures and sensorineural deafness (Moynihan et al. 1998) was recently mapped to the telomeric region of chromosome 11. Several putative tumor suppressor genes for anal canal carcinoma (Muleris et al. 1987), desmoplastic infantile ganglioglioma (Park et al. 1996), oral cancer (Uzawa et al. 1996), sporadic breast cancer (Koreth et al. 1997), and subcutaneous sacrococcygeal myxopapillary ependymoma, (Sawyer et al. 1998) have also been assigned to this region. The *MRJ* gene positioned to this chromosomal region could be a useful marker for, and contribute toward, loss of heterozugosity studies and genetic linkage analysis of this genomic locus.

References

Berruti G, Perego L, Borgonovo B, Martegani E (1998) MSJ-1, a new member of the DNAJ family of proteins, is a male germ cell-specific gene product. Exp Cell Res 239: 430-441

Bork P, Sander C, Valencia A, Bukau B (1992) A module of the DnaJ heat shock proteins found in malaria parasites. Trends Biochem Sci 17: 129

Caplan AJ, Cyr DM, Douglas MG (1993) Eukaryotic homologues of *Escherichia coli* dnaJ: a diverse protein family that functions with hsp70 stress proteins. Mol Biol Cell 4: 555-563

Cheetham ME, Brion JP, Anderton BH (1992) Human homologues of the bacterial heat-shock protein DnaJ are preferentially expressed in neurons. Biochem J 284: 469-476

Cyr DM, Langer T, Douglas MG (1994) DnaJ-like proteins: molecular chaperones and specific regulators of Hsp70. Trends Biochem Sci 19: 176-181

Koreth J, Bakkenist CJ, McGee JO (1997) Allelic deletions at chromosome 11q22-q23.1 and 11q25-qterm are frequent in sporadic breast but not colorectal cancers. Oncogene 14: 431-437

Lindquist S, Craig EA (1988) The heat-shock proteins. Annu Rev Genet 22: 631-677

Moynihan LM, Bundey SE, Heath D, Jones EL, McHale DP, Mueller RF, Markham AF, Lench NJ (1998) Autozygosity mapping, to chromosome 11q25, of a rare autosomal recessive syndrome causing

histiocytosis, joint contractures, and sensorineural deafness. Am J Hum Genet 62: 1123-1128

Muleris M, Salmon RJ, Girodet J, Zafrani B, Dutrillaux B (1987) Recurrent deletions of chromosomes 11q and 3p in anal canal carcinoma. Int J Cancer 39: 595-598

Ohki M, Tamura F, Nishimura S, Uchida H (1986) Nucleotide sequence of the *Escherichia coli* dnaJ gene and purification of the gene product. J Biol Chem 261: 1778-1781

Park JP, Dossu JR, Rhodes CH (1996) Telomere associations in desmoplastic infantile ganglioglioma. Cancer Genet Cytogenet 92: 4-7

Saito T, Seki N, Ishii H, Ohira M, Hayashi A, Kozuma S, Hori T (1997) Complementary DNA cloning and chromosomal mapping of a novel phosphatidylinositol kinase gene. DNA Res 4: 301-305

Saito T, Seki N, Ohira M, Hayashi A, Kozuma S, Hattori A, Hori T (1998) Assignment of the ZIP kinase gene to human chromosome 19p13.3 by somatic hybrid analysis and fluorescence in situ hybridization. J Hum Genet 43: 209-211

Sawyer JR, Miller JP, Ellison DA (1998) Clonal telomeric fusions and chromosome instability in a subcutaneous sacrococcygeal myxopapillary ependymoma. Cancer Genet Cytogenet 100: 169-175

Seki N, Nimura Y, Ohira M, Saito T, Ichimiya S, Nomura N, Nakagawara A (1997) Identification and chromosome assignment of a human gene encoding a novel phosphatidylinositol-3 kinase. DNA Res 4: 355-358

Seki N, Hayashi A, Abe M, Araki R, Fujimori A, Fukumura R, Kozuma S, Ohira M, Hori T, Saito T (1998) Chromosomal assignment of the gene for human DNA-PKcs interacting protein (KIP) on chromosome 15q25.3-q26.1 by somatic hybrid analysis and fluorescence *in situ* hybridization. J Hum Genet 43: 275-277

Silver PA, Way JC (1993) Eukaryotic DnaJ homologs and the specificity of HSP70 activity. Cell 74: 5-6

Uzawa K, Suzuki H, Komiya A, Nakanishi H, Ogawara K, Tanzawa H, Sato K (1996) Evidence for two distinct tumor-suppressor gene loci on the long arm of chromosome 11 in human oral cancer. Int J Cancer 67: 510-514