Detection of a processed pseudogene of the human MBL-associated serine protease, MASP1

Y. Mochihara, a H. Tazawa, a S. Habuta, b K. Ohthubo, b T. Kozaki, b Y. Ishihara, b Y. Takayama c and M. Kawakami c

a Sumikin Bio-Science, Sagamihara; b SRL, Hachioji; c Department of Molecular Biology, School of Medicine, Kitasato University, Sagamihara (Japan)

Abstract. Southern hybridization analysis of the MASP1 gene using an intron-specific probe detected a single band. An exon-specific probe detected several bands. PCR of genomic DNA using several exon-specific primer sets of MASP1 produced short and long products. Sequence of the shorter products corresponded to the processed pseudogene of MASP1. By fluorescence in situ hybridization, this pseudogene (MASP1P1) was mapped to 1p34.

Ra reactive factor (RaRF) is an anti-microbial lectin present in sera of mammals through cartilaginous fishes (Kawakami et al., 1982). It binds to polysaccharides of various microorganisms including viruses, gram-positive and negative bacteria, fungi, and protozoa. After binding, it activates the C4 and C3 components of complement and triggers the activation of a novel complement cascade (Ji et al., 1988; Matsushita and Fujita, 1992) called the "lectin pathway". Importance of RaRF in defense is evidenced by the fact that genetic mutation of a component of this factor results in a lowered resistance to infections. Human RaRF was found to comprise a polysaccharide-binding component and complement-activating components. The former is comprised of 30-kDa mannan-binding lectin (MBL) and the latter 100-kDa serine proteases P100 (Takayama et al., 1994) that are alternatively called MBL-associated serine proteases (MASP) (Matsushita and Fujita, 1992). Of these serine proteases, MASP-1 activates the C3 while MASP-2 activates C4 components of complement (Thiel et al., 1997). On the other hand MASP-3 downregulates the C4 and C2 cleaving activity of MASP-2 (Dahl et al., 2001). We determined the primary structure of human and mouse MASP-1 proteins and structure of the human MASP-1 gene (Takayama et al., 1994, 1999) and mapped human and mouse MASP-1 genes, MASP1 and Masp1, to human 3q27→q28 and mouse 16B2–B3 chromosomes (Takada et al., 1995). The human MASP2 gene was mapped to the human chromosome 1p36.3→p36.2 (Stover et al., 1999). MASP-3 is generated through alternative splicing of the MASP1 gene (Dahl et al., 2001). In the present study, we demonstrate that the human genome contains a processed pseudogene of MASP1 (MASP1P1). The high sequence homology of this processed pseudogene to exon regions of MASP1 suggests that this pseudogene may be formed by a recent event.

Materials and methods

PCR

Genomic DNA and cDNA were amplified by 35 cycles of PCR at 94°C for 30 s, at 62 or 66°C for 30 s, and 72°C for 90 or 150 s, using primers shown in Table 1. After 2.0 or 3.0% agarose gel electrophoresis, products were extracted from gel using GFX PCR DNA and Gel Band Purification Kit (Amersham Biosciences) and finally purified by gel-filtration using MicroSpin S-300 HR columns (Amersham Biosciences, Tokyo).

Southern hybridization

An intron-specific probe, IS, and an exon-specific probe, ES, were prepared by PCR using human genomic DNA and MASP1 cDNA (Takayama et al., 1994, 1999) as template, respectively. Human genomic DNA prepared from liver was digested with EcoRI or HindIII, fractionated by 0.8% agarose gel electrophoresis, and blotted onto Hybond-N (Amersham Biosciences). Blots were hybridized by overnight incubation at 42°C with probes, which were labeled with [α-32P]dCTP by the random-primed method using Multi-
Fig. 1. cDNA and gene structure of MASP1. (A) cDNA structure. Coding regions of the protease domain and non-protease domains of cDNA are indicated by gray and open boxes, respectively. (B) Gene structure. Gray underlines indicate human genomic clones of MASP1 used as probes for FISH analysis. The numbers on boxes indicate exon numbers of MASP1. Black underlines indicate the expected products of PCR using the primers shown in Table 1.

### Table 1. Sequence of PCR primers

<table>
<thead>
<tr>
<th>Product name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS</td>
<td>5'-GGGTCTCATACAGGGCTG-3'</td>
</tr>
<tr>
<td>ES/c/h</td>
<td>5'-TGGAGTGCGTGACAACTC-3'</td>
</tr>
<tr>
<td>a/f</td>
<td>5'-ACATTGGCGGCTACTACTGC-3'</td>
</tr>
<tr>
<td>b/g</td>
<td>5'-TGGAGTGCGTGACAACTC-3'</td>
</tr>
<tr>
<td>d/l</td>
<td>5'-GAAATGAGTGCCCAGAGCTA-3'</td>
</tr>
<tr>
<td>e/j</td>
<td>5'-GCTGGATCGTGACCGCCG-3'</td>
</tr>
</tbody>
</table>

### DNA sequencing

DNA sequence was determined by the dideoxy chain-termination method, using Cy5 Thermo Sequenase Dye Terminator Kit (Amersham Biosciences) and an automatic DNA sequencer (ALFexpress; Amersham Biosciences).

### FISH analysis

Prometaphase cells were obtained from phytohemagglutinin-stimulated blood lymphocytes from a normal male, after thymidine synchronization and 5-bromodeoxyuridine incorporation during the last 6.5 h of culture. The EMBL3 phage inserts of the 10-kb BamHI genomic fragment, which contained sixteen exons and the 5' and 3' flanking regions of the MASP1 gene (Takayama et al., 1999), were used as probes (Fig. 1B). Hybridized loci were detected by incubation with fluorescein isothiocyanate (FITC). For R-banding, chromosomes were counterstained with propidium iodide and observed using a microphoto-epifluorescence microscope with B2A filter (Nikon, Japan).

### Results and discussion

To determine number of MASP1 gene copies in the human genome, we analyzed it by Southern hybridization using an intron-specific 281-bp probe, IS (Fig. 1B), and an exon-specific 401-bp probe, ES (Fig. 1A). Only a single band was detected by hybridization with IS (Fig. 2A, left), while several bands were detected with ES (Fig. 2A, right). This suggests that there is an intronless gene having a sequence similar to the exon sequence of MASP1 in the human genome.

Five regions, containing at least one intron (Fig. 1B, f to j) were amplified by PCR, using exon-specific primers shown in Table 1. The sizes of longer products f, g, and j (Fig. 2B, lanes 1, 2 and 5) correspond to the length of the expected products of PCR using the intron-containing MASP1 gene as template (Fig. 1B, f, g and j). Band h was very faint and the expected band i was not detected. This is probably due to low efficiency...
Fig. 2. Southern blot analysis and PCR products of human genomic DNA. (A) Southern blot analysis using intron- and exon-specific probes, IS and ES, respectively. Lanes 1 and 2, EcoRI digest; lanes 3 and 4, HindIII digest; lanes 1 and 3 or lanes 2 and 4 are from the same genomic DNA sample. (B) Agarose gel electrophoresis of PCR products amplified from human genomic DNA using exon-specific primers shown in Table 1. Expected regions of PCR products (a to j) are indicated in Fig. 1. Numbers indicate DNA sizes (bp).

Fig. 3. Sequence of short PCR products. Nucleotides from 1 to 185, 40 to 405, 393 to 1049 and 962 to 1182 were determined on DNA from bands a, c, d and e (Fig. 2B), respectively. Vertical lines indicate potential exon-intron junctions. The sequence has been deposited into the DDBJ/EMBL/GenBank databases under accession number AB036334.

Fig. 4. FISH analysis using human genomic clones of MASP1 (Fig. 1) as probe. Bold and fine arrows indicate FISH signals on chromosomes 1 and 3, respectively.

of PCR on these DNA containing long introns (6.3 and 12.0 kb). Contrarily, shorter products a to e (Fig. 2B, lanes 1 to 5) correspond to the product length of an assumed intronless gene whose sequence is similar to the cDNA (Fig. 1A, a to e). Nineteen samples of human genomic DNA were analyzed by the same PCR method. The bands of shorter products a, b and c were detected by the same PCR analysis on 8 to 19 samples of normal human genomic DNA, so far tested (data not shown).

To confirm that these shorter PCR products are of an intronless gene, the PCR products a, c, d and e were sequenced after separation from the gel and purification by gel filtration. The total length of regions determined was 1,182 bp (Fig. 3). Surprisingly, the determined sequence was completely identical to that of the corresponding region (exons 4 to 13) of human MASP1 cDNA.

FISH analysis was performed using various clones as probes shown in Fig. 1B. When these probes were used, signals were detected on chromosome 3q27 of MASP1 as reported previously (Takada et al., 1995). Furthermore, probe H-16 gave strong signals on 1p34 although they were weaker than those of 3q27 (Fig. 4). No obvious signals were detected when other probes were used. These results indicate that a processed pseudogene of MASP1 is present in human chromosome 1p34.

Recently, various types of processed pseudogenes have been found. For example, a pseudogene can be transcribed to mRNA although a mutation at the ATG start codon prevents its translation into protein (Luo et al., 1999). Another pseudogene is not transcribed though it has promoter elements and a putative poly(A) site. Sequence of the processed MASP1 pseudogene had high homology with its cDNA (100%). This fact suggests the recent generation of this pseudogene in the human genome, or the possibility that it is expressed as RNA.
References


Luo M, Shang Ji, Yang Z, Simkevich CP, Jackson CL, King TC, Rosmarin AG: Characterization and localization to chromosome 7 of hGABPa, a human processed pseudogene related to the GABPa transcrip-


