Isolation and Characterization of the Human Interleukin-9 Receptor Gene

By Ming-shi Chang, Greg Engel, Chris Benedict, Rita Basu, and Jennifer McNinch

To better understand the regulation of interleukin-9 (IL-9) receptor expression, we have isolated the genomic clone of the human IL-9 receptor based on its sequence homology with a human IL-9 receptor cDNA isolated from the human megakaryocyte cell line UT-7. The entire genomic structure has been determined. The human IL-9 receptor gene consists of 10 exons spread over approximately 13.7 kb of DNA. The nucleotide sequence of the coding region from the genomic DNA is identical to our cDNA clone. Several blocks of transcriptional control sequence have been identified at the 5' noncoding region of the IL-9 receptor gene that may play an important role in the regulation of the IL-9 receptor gene. A fusion gene containing 659 bp of human IL-9 receptor 5' noncoding region linked to the firefly luciferase gene directed expression of luciferase activity in human embryonic kidney 293 cell line.

INTERLEUKIN-9 (IL-9) was originally identified by its ability to support the growth of helper T cells, but its potential targets have been shown to include mast cells, megakaryoblastic leukemia cells, fetal thymocytes, murine erythroid progenitors, and human erythroid and myeloid precursors. Involved in IL-9 in tumorigenesis has been suggested by the observation that murine helpet T cells become tumorigenic after transfection with IL-9 cDNA.

The murine IL-9 receptor was recently characterized. A single class of receptor of 64 kD was determined on a T-cell clone. Murine and human IL-9 receptor cDNAs have been isolated by expression cloning from the murine T-cell clone TS1 and the human megakaryoblastic leukemia cell line Mo7E. We report here an independent isolation of the human IL-9 receptor cDNA and cloning and analysis of the complete human IL-9 receptor genomic DNA.

MATERIALS AND METHODS

Isolation of human IL-9 receptor cDNA clone. Human IL-9 receptor cDNA clone was isolated from the cDNA library of the human megakaryocyte cell line, UT-7 cells. A mixture of degenerate synthetic oligonucleotides encoding the WSXWS motif (Trp-Ser-X [Glu, Pro, Ala, Ser, Asp]-Trp-Ser) was used as a hybridization probe. The oligonucleotide sequences are the mixture of

A TAA
TGG-AQF-GCG-TGG-AQF C C

The probes were end-labeled with 32P and hybridized to the cDNA library in the hybridization solution of 20% formamide, 5 x SSC, 5 x Denhardt’s solution, 0.05% sodium pyrophosphate and 0.1% sodium dodecyl sulfate (SDS) at 39°C for 16 hours. This WSXWS motif is the consensus sequence for the cytokine receptors. To avoid reisolation of known proteins with WSXWS motif, we searched in the gene bank for all proteins with nucleotide sequences identical to the human megakaryocyte cell line UT-7. Forty positive clones that hybridized with both 5'-end and 3'-end probes were isolated, and phage DNA was prepared using the Magic Lambda Preps DNA Purification System with some modification (Promega, Madison, WI).

Preparation of λ phage DNA for sequencing. λ Phage DNA was prepared by the Magic Lambda Preps DNA purification system with some modification to obtain cleaner λ DNA for use as a sequencing template. After λ DNA was eluted from the resin by following the protocol of the isolation kit, it was further purified by proteinase K (200 μg/mL) treatment, phenol-chloroform extraction, and ethanol precipitation.

Nucleotide sequence analysis of the IL-9 receptor gene. For the genomic DNA sequence, λ DNA prepared as described above was used as a template and synthetic oligonucleotide was used as primer. DNA sequences were determined by using the Prism Ready Reaction Dye Deoxy Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) along with the 373A DNA sequencer (Applied Biosystems).

Construction of fusion genes and assay for their promoter activity. An 848-bp fragment (−2,431 bp to −1,584 bp as shown in Fig 4) of IL-9 receptor intron 1 region was generated by polymerase chain reaction (PCR) with two oligonucleotide primers (primer A1 and primer A2) with HindIII site at the 5' end and NheI site at the 3' end. The amplified fragment was cloned into the pGLOW-enhancer plasmid vector containing the entire coding sequence of firefly luciferase along with the SV40 enhancer (Promega) to generate a pIL-9R-int-enhancer fusion gene. The 659-bp fragment (−5,700 bp to −5,042 bp in Fig 4) containing IL-9 receptor exon 1 region was similarly amplified by PCR with two oligonucleotides, primer B1 and primer B2, and subcloned into PGL-2 enhancer vector to generate a pIL-9R-Pro-enhancer fusion gene. These two fusion genes were used along with the promoterless pGLOW-enhancer plasmid in...
the transfection of human embryonic kidney 293 cells or NIH3T3 cells. The luciferase activities of transfectants harvested 72 hours after transfection were analyzed according to the protocol of the luciferase assay system (Promega).

RESULTS

Isolation of human IL-9 receptor cDNA. To isolate the novel receptor specific for the growth or differentiation factor for megakaryocytes, we generated a cDNA library from human megakaryoblastic cell line UT-7. A mixture of degenerate synthetic oligonucleotides encoding for WSXWS motif, a cytokine-receptor consensus sequence, was used as a hybridizing probe. We isolated several novel cDNA clones whose sequence encoded open reading frame with WSXWS motif. One of these clones was found to be 99% identical to that of the published IL-9 receptor cDNA, Ph9RA3, in the coding region. There are two differences in the sequences between our U41 cDNA clone and the published cDNA. In Ph9RA3, amino acid 331 is Arg, whereas in U41, it is Gly. The genomic sequence is identical to the U41 cDNA sequence at this position. The second difference occurs at the 5' end of the cDNA sequences. Alignment of the 5' end of the nucleotide and amino acid sequences of U41 and Ph9RA3 are shown in Fig 1. In U41, there is an insertion and the published human IL-9 receptor Ph9RA3 cDNA (bottom). Vertical lines indicate identical nucleotides (A) or amino acids (B). The translation initiation codon is boxed for each sequence. The signal peptide is underlined.

Fig 1. Comparison of two cDNA sequences for the human IL-9 receptor. Comparison of the nucleotide (A) and amino acid (B) sequences of the 5' end of the human IL-9 receptor U41 cDNA (top) and the published human IL-9 receptor Ph9RA3 cDNA (bottom). Vertical lines indicate identical nucleotides (A) or amino acids (B). The translation initiation codon is boxed for each sequence. The signal peptide is underlined.

Fig 2. Genomic organization of the human IL-9 receptor gene. (A) Restriction enzyme mapping of the gene. Only BamHI (B), ClaI (C), HindIII (H), KpnI (K), SmalI (S), XbaI (Xb), and XhoI (Xh) sites are shown. (B) Schematic diagram of the human IL-9 receptor gene. The length and position of each exon are indicated (B). Restriction sites are shown only in the introns. (C) Relationship of the genomic DNA to the cDNA (—) and protein (—). The exoplasmic domain (EXO), transmembrane domain (TM), and cytoplasmic domain (CYT) are shown. The dashed lines indicate (from left to right) the start of the cDNA, the start of the coding region, the transmembrane domain, the end of the coding region, and the end of the cDNA.
ISOLATION OF INTERLEUKIN-9 RECEPTOR GENE

Fig 3. Sequences of the intron-exon junctions of the human IL-9 receptor gene. Exon sequences are shown in uppercase letters, and intron sequences in lowercase letters. The encoded amino acid and its position are indicated below the exon sequences. The length of each intron is shown in parentheses.

The extracellular domain is encoded by exon 3 to exon 7. The transmembrane domain is encoded by exon 8 and the cytoplasmic domain is encoded by exons 9 and 10. The sequences of intron-exon junction conform to the consensus sequence of eukaryotic splice junction (Fig 3). This assignment of the exon-intron boundaries is based on our cDNA clone U41 isolated from UT-7 cells. Alternatively, in the published cDNA, PH9RA3, exon 2 would be skipped and would be part of intron 1.

Search for potential regulatory sequences. Many sequences regulating transcription, as well as the nuclear factors with which they interact, have been extensively characterized. Figure 4 shows the genomic sequences of the 5' noncoding region of the IL-9 receptor gene. We have identified two putative TATA boxes and several potential regulatory elements in this region. In the intron 2, there are NJK-B (-1,407 bp), SP-1 (-1,177 bp), glucocorticoid response element (-937 bp) and AP-1 (-245 and -310 bp). In the intron 1, there are regulatory sequences of AP-2 (-2,409 bp), Oct-2 (-2,248 bp), AP-4 (-1,648 bp) and one putative TATA box (1,273 bp). In the exon 1, there are AP-3 (-5,226 bp) and SP-1 (-5,211 bp). Upstream of the cDNA start site (-5,245 bp), there is another TATA box. Therefore, the region from -5,700 to -5,245 could be the potential promoter region.
Promoter activity. There are two putative TATA boxes in the 5' noncoding region. One is located at intron 1 (−2,173 bp), the other is located upstream of the cDNA start site (−5,145 bp). To determine the ability of the putative regulatory elements to function as a promoter, we constructed two fusion genes in which a firefly luciferase reporter gene was placed downstream under the control of the regulatory elements. One fusion gene, PIL-9R-pro-enhancer, contains a regulatory fragment of 659 bp (nucleotide -5,700 to -5,042) including partial intron 1 sequences. The other fusion gene, PIL-9R-int-enhancer, contains a regulatory fragment of 848 bp (nucleotide -2,431 to -1,584) including partial intron 1 sequences. The dotted line indicates the 2-kb intron that has not been sequenced.

Fig 4. Sequence of the 5' noncoding region of the IL-5 receptor gene (A through C). The 5' upstream sequence along with the first 30 bases of coding region are shown. Putative TATA-box are boxed and the potential regulatory elements are underlined. Two sets of oligonucleotide primers (primer A1 and A2, primer B1 and B2) used as PCR primers to amplify the fragment put upstream of the luciferase gene for promoter activity analysis are overlined. The boundaries of intron 1, intron 2, and the start of cDNA (not arrow) are indicated. The dotted line indicates the 2-kb intron that has not been sequenced.
Table 1. Activity of the Human IL-9R Promoter in 293 Cells

<table>
<thead>
<tr>
<th>Plasmid DNA</th>
<th>Luciferase Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>No PMA</td>
<td>293</td>
</tr>
<tr>
<td>With PMA</td>
<td>NIH 3T3</td>
</tr>
<tr>
<td>PGL-2-enhancer</td>
<td>6,810</td>
</tr>
<tr>
<td>No PMA</td>
<td>15,173</td>
</tr>
<tr>
<td>With PMA</td>
<td>246</td>
</tr>
<tr>
<td>PIL-9R-pro-enhancer</td>
<td>114,471</td>
</tr>
<tr>
<td>No PMA</td>
<td>358,457</td>
</tr>
<tr>
<td>With PMA</td>
<td>276</td>
</tr>
<tr>
<td>PIL-9R-int-enhancer</td>
<td>15,100</td>
</tr>
<tr>
<td>No PMA</td>
<td>43,038</td>
</tr>
<tr>
<td>With PMA</td>
<td>226</td>
</tr>
</tbody>
</table>

The 293 cells or NIH3T3 cells were seeded at 8 x 10⁵ cells per well in the six-well plate. Twelve hours later, cells were transfected with promoterless control vector (PGL-2-enhancer) or PIL-9R-pro-enhancer plasmid DNA that contained the 659-bp fragment (nt -5,700 to -5,042) including exon 1 sequence or PIL-9R-int-enhancer that contained 848-bp fragment (nt -2,431 to -1,584) including partial intron 1 sequences. Forty-eight hours after transfection, cells were treated with PMA at the concentration of 10 ng/mL or without PMA treatment. Twenty-four hours later, cells were lysed for luciferase activity analysis. Reproducible results have been obtained by using the transfection methods of either calcium phosphate or lipofectamine. Data presented are the average of the duplicate samples transfected with 2 μg of plasmid DNA by using the calcium phosphate method. The number represents the activity from 1.6 x 10⁶ cells.

Furthermore, the promoter activity can be induced by PMA treatment (Table 1). The result is consistent with our earlier observation that mRNA expression level of IL-9 receptor can be induced by PMA treatment in the UT-7 cells (data not shown). These data show that the promoter activity of the 659-bp fragment in the IL-9 receptor gene is cell-specific and can be induced by PMA.

DISCUSSION

We have isolated the genomic clone of the IL-9 receptor and determined the genomic sequence and structure of the gene. The genomic structure consists of 10 exons. Exons 3 through 7 encode the extracellular domain, exon 8 encodes the transmembrane domain, and exons 9 and 10 encode the cytoplasmic domain. The genomic structures of several members of the cytokine receptor superfamily have been reported. Comparison of the genomic structure of the IL-9 receptor with other members of the family shows a considerable homology of their genomic structure. As shown in Fig 5, the extracellular domains of most cytokine receptors are encoded by 4 exons. Two of them contain the four conserved cysteine residues and the exon next to the transmembrane domain exon contains the WSXWS motif. The structural homology supports the theory that the IL-9 receptor and other members of the cytokine receptor superfamily are derived from a primordial gene by gene duplication.

The gene encoding the ligands of these receptors may also be derived from a common ancestor. The interleukins and colony-stimulating factors have similar exon-intron structure and share the same elements regulating gene expression. This phenomenon suggests that hematopoietic growth factors and their receptors have evolved from a single primordial growth factor gene and a single primordial receptor gene. The current complex system of factors and receptors reflects the evolution of the immune system as a whole.

Recent studies have shown that a section of the cytoplasmic domain that includes the serine-rich region appears to have a critical role for mitogenic signal transduction. This critical region shows the highest degree of homology between IL-2Rβ subunit and murine erythropoietin receptor. Both human and murine IL-9 receptors also contain a high percentage of serine and proline in their cytoplasmic domain. The proline- and serine-rich region in all three receptors are encoded by the last exon. This suggests that the last exon of these three receptors may involve the mitogenic signal transduction.

Comparison of our cDNA sequence with that of the published cDNA showed that there are some discrepancies in the 5′ end. The difference between the U41 cDNA sequence and the published cDNA sequence could be caused by the differential splicing. In the process of screening the cDNA library from UT-7 cells, we isolated several full-length cDNA clones of different sizes. Nucleotide sequence analysis of some clones showed that there are differential splicings in the 5′-end or 3′-end noncoding region. Based on PCR analysis, we have observed that the IL-9 receptor is expressed on a wide range of cell types, including B cells, T cells, neutrophils, megakaryocytes, and human kidney cell line 293 (data not shown). It is possible that the expression of the IL-9 receptor gene is differentially regulated in different cell types or in different developmental stages of a hematopoietic cell lineage. The high degree of divergence in their 5′ ends or 3′ ends may account for the different level of gene regulation in the various cell lines.

The assignment of the translation initiation codon in U41 cDNA is different from that of the published Ph9RA3 cDNA. Renaud et al assigned the translation initiation site at the first methionine (Fig 1B). Our hydrophobicity analysis of their signal peptide sequences did not match with a signal peptide profile. However, if the translation initiation site in Ph9RA3 cDNA was assigned to the second methionine (same methionine as U41 cDNA), the following 22 amino acids fit perfectly with the hydrophobicity profile. Therefore, the mature protein of IL-9 receptor should be glutamic acid instead of valine. Assignment of either first methionine or second methionine as a translation start site follows Kozak's rules on +4 position (ATGG), but neither of them follow the rules on −3 position.

The fusion gene, PIL-9R-pro-enhancer, containing 659 bp (nucleotide −5,700 to −5,042) showed a 16.8-fold increase in luciferase activity over the control vector. Whereas the fusion gene, PIL-9R-int-enhancer, containing 848 bp from intron 1 showed a 2.2-fold increase luciferase activity over the control vector. Therefore, the 659-bp fragment is 7.5-fold more active than the 848-bp fragment. It appears that the 659-bp region contains the promoter, and the TATA box identified in the intron 1 may not be the true start site. However, this 848 bp may contain an enhancer.

We have sequenced additional 500 bp of further 5′ end of the gene (data not shown in Fig 4) and identified two Alu repeats. We know of no case that Alu repeats are present in the first exon of a human gene. This further supports our belief that the transcription start site is located in the 659-bp region and the TATA box at position −5,580 bp is the best candidate.
Fig 5. Cytokine receptor family. Schematic representation of the various members of the cytokine receptor gene family that have a striking similarity in the genomic structure are shown. Shaded boxes represent exons containing protein coding sequences. 3' untranslated exon sequences are not shown. Amino acid residues encoded by individual exons (rounded off to the nearest amino residue), are indicated below the exon. 5' untranslated DNA (UT), signal peptide (SIG), conserved cysteine (CC), WS motif (WSXWS), and transmembrane domain (TM) are depicted for each gene.

ACKNOWLEDGMENT

We acknowledge Drs Norio Komatsu and John Adamson for providing the UT-7 cells. We thank our colleagues at Amgen, especially Drs Frank Calzone and Frank Martin, for helpful discussion and critical reading of the manuscript, Joan Bennett for preparing the manuscript, Jennifer Keysor and Rick Pinal for artwork, Edward Shatzen for the luminometer, and the DNA synthesis group for the oligonucleotides.

REFERENCES


Isolation and characterization of the human interleukin-9 receptor gene

MS Chang, G Engel, C Benedict, R Basu and J McNinch