Human Interleukin-9: Genomic Sequence, Chromosomal Location, and Sequences Essential for Its Expression in Human T-Cell Leukemia Virus (HTLV)-I–Transformed Human T Cells

By Kerry Kelleher, Kevin Bean, Steven C. Clark, Wah-Ying Leung, Teresa L. Yang-Feng, Janice W. Chen, Pin-Fang Lin, Wen Luo, and Yu-Chung Yang

We have isolated the genomic sequence of human interleukin-9 (IL-9) based on its sequence homology with a human IL-9 cDNA isolated from human T-cell leukemia virus (HTLV)-I–transformed T cells by expression cloning. The entire genomic sequence has been determined and the gene consists of five exons and four introns. The human IL-9 gene is mapped to the long arm of human chromosome 5 at band 5q31-32, a region found to be deleted in a number of patients with acquired 5q− abnormalities and hematologic disorders. Several blocks of transcriptional control sequences have been identified at the 5′-flanking region of the human IL-9 gene that may play an important role in the control of IL-9 gene expression. The 5′-regulatory region of the human IL-9 gene also contains sequences identified in the 5′-flanking regions of other cytokine genes mapped to the long arm of human chromosome 5, including IL-3, IL-4, IL-5, and granulocyte-macrophage colony-stimulating factor and other T-cell growth factor genes including IL-2 and IL-6. The IL-9 gene is constitutively expressed in the HTLV-I–transformed human T cells and the expression of IL-9 in these cells can be further induced by 12-O-tetradecanoyl phorbol 13-acetate. Transient transfection analysis using the plasmid containing the 5′-flanking region of IL-9 gene upstream from the firefly luciferase reporter gene indicated that the 0.9-kb *Sma*I-*Sac*I fragment of the IL-9 gene contains sequences required for the constitutive and activated expression of IL-9 gene in HTLV-I–transformed cells. These results will now allow us to study the regulatory mechanisms of IL-9 gene expression in normal and leukemic human T cells.

INTERLEUKIN-9 (IL-9), also known as T-cell growth factor P40 in the mouse system, was originally isolated based on its ability to support the growth of certain mouse helper T-cell clones.[1,2] Subsequent studies have shown that this cytokine also possesses mast cell enhancing activity and can enhance the mouse mast cell growth elicited by mouse IL-3.[3,4] We have isolated the cDNA clone encoding human IL-9 by expression cloning based on its ability to stimulate the proliferation of a human megakaryocytic leukemic cell line.[5] Our recent studies have shown that human IL-9 can also stimulate erythroid colony formation in vitro.[6,7] Using highly purified progenitor cells, we were able to demonstrate the direct effects of IL-9 to support the growth of $CD34^+DR^+CD33^-$ human erythroid progenitors in serum-free culture conditions.[7] Therefore, IL-9 represents a new cytokine that may play an important role in normal hematopoiesis.

In our initial studies,[5] we showed that IL-9 can be expressed in normal T cells following phytohemagglutinin (PHA) and 12-O-tetradecanoyl phorbol 13-acetate (TPA) induction. It can also be expressed constitutively in the human T-cell leukemia virus (HTLV)-I or -II–transformed human T cells. As a first step in understanding the regulation of the IL-9 gene expression, we have isolated the human IL-9 genomic sequence and compared it with other cytokine genes to identify potential common regulatory sequences. Here we report the entire human IL-9 genomic sequence, its chromosomal mapping position, its 5′ regulatory sequence in comparison with other cytokine genes, and the 5′-flanking sequences required for its expression in HTLV-I–transformed T cells. Our results show that the IL-9 gene is located in a region of the long arm of chromosome 5 known to contain many cytokine genes. Analysis of the 0.9-kb 5′-flanking region, which is sufficient for directing its regulated expression in HTLV-I–transformed T cells, C10MJ2, shows that the human IL-9 gene shares many regulatory elements with other cytokine genes. These results will facilitate further comparison of the regulation of expression of IL-9 with that of other cytokines.

From the Genetics Institute, Cambridge, MA; the Department of Human Genetics, Yale University, New Haven; Bristol Myers Squibb Co, Wallingford, CT; and the Departments of Biochemistry/Molecular Biology and Medicine, Walther Oncology Center, Indiana University School of Medicine, Indianapolis.

Submitted September 25, 1990; accepted November 26, 1990.

Address reprint requests to Yu-Chung Yang, PhD, Indiana University School of Medicine, Walther Oncology Center, 975 W Walnut St, IB540, Indianapolis, IN 46202.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" *in accordance with 18 U.S.C. section 1734 solely to indicate this fact.*

0006-4971/91/7707-0008$3.00/0

MATERIALS AND METHODS

Isolation of genomic human IL-9 clones. A human leukocyte genomic library was purchased from Stratagene (San Diego, CA). The library was screened by hybridizing nitrocellulose filters containing 2×10^6 bacteriophage plaques overnight at 60°C using the [^{32}P]-labeled human IL-9 cDNA as the probe (2×10^6 cpm/mL; specific activity: 2×10^8 cpm/μg DNA). Several positive phage clones were obtained, and the ones containing 5′ and 3′ cDNA sequences were used for further characterization. Phage DNA from one such clone was digested with *Bam*HI, *Eco*RI, *Hin*dIII, *Pst*I, and the combinations of either two of these restriction enzymes. The digested DNAs were analyzed by Southern transfer and probed with total IL-9 cDNA or different oligonucleotide probes. Two *Bam*HI fragments (4 kb and 2 kb) and three *Pst*I fragments (2 kb, 1.5 kb, and 1.3 kb) that hybridized with the total IL-9 cDNA probe were subcloned into the vector IBI24 (International Biotechnologies, Inc, New Haven, CT) for sequence analysis.

Nucleotide sequence analysis of the IL-9 gene. The nucleotide sequence of each of the *Bam*HI and *Pst*I subclones was determined by the dideoxy chain termination method on supercoiled templates with synthetic oligonucleotide primers.[8,9]

In situ hybridization. A 2-kb *Pst*I fragment of human IL-9 genomic sequence (containing part of the second exon and the

```
1    CCCGGGAGGC AGAGGTTGCA GCGAACCTGG ATTGTGCCAC CGCACTCCAG CCTGGTGACA GAGGGAGACT CTGTCTCAAA AAAAAAAAAA AAAAAAAAG AAATGTTACT AGTATGTAGT   120

121  AAGTTCTCAG TAAATGTTAG CTACTATACT CTTTCAAGTG CTGGGTTTTT ACTTGATGTC ATACAGTGTT ATATAAGATC TCCAAAGATA CTGAGGAGTC CTCAAGGCCA ATTTTAACAA   240

241  GCATGGTTGC CGCATTCTTG TGCTTATAGT TGAACATTTC TTCTTTCAGA CACTTGCACA AAGTGATACT TCTAAGATGC ATTTGCATTA GGTGGCAAAC TTCATCCTGG GTATGAAAAA   360

361  CATTGAGATT TGGGAATAAA GCATAGTAAG ACTGAGGTTG CAATTACTAA AGGAAAACCC CAACAGAGAT AAGTGAAGTT CTGCAATATC ATGCACCCTC CCCCAACCCG CTCTGTCTCC   480

481  CCAGGCCCCC CTTCGTTAGA ACACCCATGA CTGGCTATAT TATATCAGCA TTTCCCATAA TGTAAAAAGG GAAAATACAG ACCTGGGCGT TCATGGAAAG TATTCTAACT CTCACAACCA   600

601  GAATCCCTGT CTTTGAATTT TTTTTCTTGG TTTTTAGATC TTTAACTTTT CCTTCAGCAT TTCAGTACTC AACTTTTTGA AAATCATCTT TTCTGAGGAA TGATATTTCC TGGCACAGCA   720

721  TCATCTCTGT CAAGTGACTC AGTTTGATTT TTTTGTTTGT TAGTATAAAG TGGCCCCAAC TTACAGAGAA AAAGTGGGCT CTTGGTATCA GTTTGATGTC AGGGTTTTTC CGTGTTTGAG   840

841  AGGGAGCTTT AAATACCACT CGATTTGAAG GTGTCTGCAA GCGAGCTCCA GTCCGCTGTC AAG ATG CTT CTG GCC ATG GTC CTT ACC TCT GCC CTG CTC CTG TGC TCC   948
                                                                          M  L  L  A  M  V  L  T  S  A  L  L  L  C  S

949  GTG GCA GGC CAG GGG TGT CCA ACC TTG GCG GGG ATC CTG GAC ATC AAC TTC CTC ATC AAC AAG ATG CAG GTAGGCTGCA GGGGGAGCCC ATGGGAAAGA          1047
     V  A  G  Q  G  C  P  T  L  A  G  I  L  D  I  N  F  L  I  N  K  M  Q

1048 CAGCTACTGA CAAAGTGAAA TATGTATGAG GATGAAAAAA CTCGGGGCTG ACTAAAGGTT CTTATCTCTC TATCTACTTT AG GAA GAT CCA GCT TCC AAG TGC CAC TGC AGT   1159
                                                                                              E  D  P  A  S  K  C  H  C  S

1160 GCT AAT GTGAGTGAAT GCTCTTTAAG AACTTTCCAA ATTAATTTTA ATTTTCACAT CTGGAATCTT CACTCTGAAA TTTCCCTTGC AG GTG ACC AGT TGT CTC TGT TTG GGC   1271
     A  N                                                                                                 V  T  S  C  L  C  L  G

1272 ATT CCC TCT GTAAGTATAG TGAAATAACA TAATGTTGAC CTTGGATTTT TTTGGTTTGT TTTTAAGTAA AAATAAGTTG CTTTATTTAA TATTTAATGT TATACATTGT          1380
     I  P  S

1381 TGCTTAATTT AATTGTTACA GATTAGTATT CCCTGTTAAA ACCACATTGT TACAAATTAT TCCCTTTTAA AACTACGATC TTGAAATCCT ATATTATGAA CATTTCTTTG TATTTAATTA   1500

1501 ACTTTATGCC TCTTGAGAAG TTTGAACACT TTTCAACATT AAAAAAAGAA TCCTGAATAT CTTTTTAGAT AGGTGGCCAT GTGCACAATT AAATAAAACT GGAACTAAGG ATATAATAAT   1620

1621 TGCTGTAGCT CATATCATAT TGCTTTCTAA CTCATTTACT GATAACTCTA GAGTTGTGAA ACAATGTAAA TAAAATGACA ACTCCTTATC TTTCATCTGT CATGAATGAT CTATGCGCTA   1740

1741 TACCTCCCCC TCCCTGCCTC CTCCCTTCCT CCCCACCACC CTGTTGTCTG TCTAGCTGAT TAGAGTGACT GTTGGTTTGA ATGCTGCCCT CTGGGCAGGT AGAGGATCTG AGGTTGTGAG   1860

1861 TGGAAGGAGG GCTTCCAGAG GGCCACTGCC CACTACGGCA GGAAGGATGG GTGGCAGGAA AGTTCTGATT CCTAATTCAA ACTCCTGGTT AGGGTGAGGA GGAGGCACTT CTCCAAGGTG   1980

1981 CAGTGCTTTA TTCTTTCTCA TGCAAGGCCT GGGAGAATCT GAAGAATCTG AGCTTCTTGC CCTGGCTAGG GTAAGACATC GCACCCATCG CGGTCCATCC ATTAGATGAG AAGAGGATAG   2100

2101 AGTGCCTTCT GGGCAGGAAC CAGGCAGACA GCACAGCCCC TGTCCCTTGG AGTACCGTCC ATGTTTTTAG CTGCTGCTGA AATACCAGCT GCATTCAATT GTCACATCCC ATTAGCTGGT   2220

2221 GTGAAAAGGC TTTTCCTCAC TCTGCACTTT CAGACTTACA AGCCTTGAAG CCGGGAAGCA CCCGTTGAAA AGAACATTCA GAGCCGACTA TTTCAGGGCC CAGAGCCCTC ATGTTTCCTG   2340

2341 GATGTAACAT ACAGGAAGTC TCCTCCAGGG GATGTCACTG TGGAAAAATG GCATCCCCTT TAAATACGGG AGATCACTTC CTACATTGGC AAGGGACCTG TCTAAAAATA ATGCAAGTTT   2460

2461 GAGTAATGGT GATTAAATAA AAATCATCTC TATTATATTG CTCTTTGTGA TATATTTCCA AAGCTGTCCT CAGAATATTT CTTTGAATAA ATCCTTACTA TTTACCAG GAC AAC TGC   2577
                                                                                                                          D  N  C

2578 ACC AGA CCA TGC TTC AGT GAG AGA CTG TCT CAG ATG ACC AAT ACC ACC ATG CAA ACA AGA TAC CCA CTG ATT TTC AGT CGG GTG AAA AAA TCA GTT GAA   2676
     T  R  P  C  F  S  E  R  L  S  Q  M  T  N  T  T  M  Q  T  R  Y  P  L  I  F  S  R  V  K  K  S  V  E

2677 GTA CTA AAG AAC AAC AAG TGT CCA GTAAGTTTGT TTTCATATGT GATATGTTCC TGTTGGTGAT TTCTATGTGA ATGGTGATGC CAACCCTGTT TGAACACAAA AGGATGATAA   2790
     V  L  K  N  N  K  C  P

2791 AGTTGGAATT GGTAGTTCAA GGTTGATAAA AGACATCTAA GAATTTTAAT CAGAAGTAAT ATAATTAAAG TGAGATCCAC TGAAACAATA GAATTAAAGT GAGATAGATC ATTGTTCCTG   2910

2911 ACGAGGCCAT TTACTTCTCT CTACTATGGA ATAATGAAAG AATCCTTTCT GAGTGTAATT AGAAGCTACA ATCTAGAGAA TCAGGGATGT AGCTCACATA ATACTAAATT ATCCTAGAGA   3030

3031 TTCAATGTAC TAACTGAATG GATGTTGTTA ACAGGGATTT TTTTTCCTG TTGGTTAAGG AGGTTTTGTT TTGTTTTGGA GACAGAGTCT TGCTCTGTTG CCCAGGCTGG AGTGCAGTGG   3150

3151 TGCCATCTGA GCTCACTGCA GCCTCTGCCT CCCGGGTTCA AGTGATTATC CTGCCTCAGC CTCCCGAGTA GCTGGCATTA CAGGTGCGTG CCACCATGCC TGGCTAATTT TTGTATTTTT   3270

3271 AATAGAGATG GGGTTTCACC ATGTTGGCCA GGTTGCTCTC CAACTCCTGA ACTCAAGTGA TTTGCCCGCC TTGACCTCCC AAAGTGCTGG GATGACAGGT GTGAGCCACC ATGCCTGGCC   3390

3391 TGCATTAAGG AGGTATTTAA AGGGCAATGC ACCCAGGTCA AGGTGGAAGC TTGCTACTCA TCCTGAATGC CCATCCACAC ATTCTTTTCT TCAGCATATA CCCTAGTCCC TGACAGCAGA   3510

3511 CTGGGATGGC AAGTTGGGTA GAGGTGACCT CCCTCTGTTT TTTGGGTATT AGCATCTCCA CACAAGATCC TAGAAGGCTG AAAGCCCTGA GCTCAGCTGT TTAGCTGCAT GCGTTTCTAC   3630

3631 CATCAATGGC ATCTAGTTCT AAGTGCTTAA TATATGCTGT CTCACTGAAT AAATACATAC CTTAGGGACA ATTATTCAAT TTATTACTCT CAGTGAGGTT AACTAATTTG CCTAAGGCTG   3750

3751 CATATTTGAT AAGTGGCAGA GCTGAGATTT GAACTCAGGC CTATATGACC TCAGAGCCCC ACTCTTAGCC ATTGTACTGT CAAATGACCT TGGAAAGACA ACCTAAAAGG ATAATGATAC   3870

3871 AATTTTAGGC CTCAAAGAGT CCCCAGAAAA GGCTTTCTCT AATGCAGAGA TTTAGGGCCA CTTAATAGGG GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTAAAG   3990

3991 ACCCCTGAAA TCCAATTTGA GGTCAACCAC CTATGCTGTC TTTACACCAC ATGAGCTAGC CTGGACCTGC CCACCTATTT GCTCTGTGTC TCAAGCCACT TCCCTTCCCA TCCCCACAAT   4110

4111 CCTCACCACC GACTCTGGCT CTTGGCAGGT AGGCTTCTGG GGCTGCTTGG CTCTACATCA TTTGAGTCAC TCTGTCCTTA TCAACTTTCA TCCCCACAG TAT TTT TCC TGT GAA   4224
                                                                                                                  Y  F  S  C  E

4225 CAG CCA TGC AAC CAA ACC ACG GCA GGC AAC GCG CTG ACA TTT CTG AAG AGT CTT CTG GAA ATT TTC CAG AAA GAA AAG ATG AGA GGG ATG AGA GGC   4320
     Q  P  C  N  Q  T  T  A  G  N  A  L  T  F  L  K  S  L  L  E  I  F  Q  K  E  K  M  R  G  M  R  G

4321 AAG ATA TGAAGATGAA ATATTATTTA TCCTATTTAT TAAATTTAAA AAGCTTTCTC TTTAAGTTGC TACAATTTAA AAATCAAGTA AGCTACTCTA AATCAGTATC AGTTGTGATT   4436
     K  I

4437 ATTTGTTTAA CATTGTATGT CTTTATTTTG AAATAAATAC ATATGTGGAA AAAACAACAT GAGCTGGTCT CTTGGCAATT ATTCATTTCT TGCTGCTCAG AACAAAGAAA GCTACAAGTG   4556

4557 TTGTTAAGGG GAAGAATAGA TCAGAGACTC CTGTAGGAGT CTCTGTGATA AGACTCCTGA TGCTGAATAC AGACCCTCAG GCTCATAGGC TGTGGCTGGA GCTGCAG              4663
```

Fig 1. Complete nucleotide sequence of human IL-9 gene and predicted amino acid sequences. Nucleotides are numbered starting at the *Sma*I site. The "TATA" box, ATTTA motifs, and the presumed polyadenylation signal sequences are underscored. The differences between the present and published sequences are: base 304 is T instead of G, base 2156 is C instead of A, base 2776 is A instead of G, and there is a 6-base insertion, GTGTGT, at nucleotide 3981.

entire third and fourth exons) was nick-translated to a specific activity of 3×10^7 cpm/μg with [^{3}H] dATP and [^{3}H] dCTP. In situ hybridization to human metaphase chromosomes and autoradiography were performed according to published procedures.[10,11]

Construction of phuIL-9-LUC plasmid and assay for IL-9 promoter activity. The 0.9-kb *Sma*I-*Sac*I (*Sma*I site: base 1 and *Sac*I site: base 883 in Fig 1) fragment of IL-9 5′-flanking sequence was cloned into the polylinker region of pXP2,[12] a plasmid containing the entire coding sequence of firefly luciferase to generate phuIL-9-LUC plasmid. Transient transfection analysis in the HTLV-I–transformed T cells, C10MJ2,[13] was performed using electroporation technique described previously.[14] C10MJ2 cells were induced immediately after transfection using different combinations of the following conditions: TPA, 10 ng/mL; PHA-P, 10 μg/mL; A23187, 50 ng/mL; anti-CD3, 10 ng/mL. The luciferase activities of transfectants harvested 24 to 48 hours after transfection were analyzed accordingly.[15] Protein quantitation of various transfectants was performed using the Micro BCA Protein Reagent Assay Kit from Pierce (Rockford, IL).

RESULTS

Organization and nucleotide sequence of the human IL-9 gene. The human IL-9 genomic clone was isolated from a human leukocyte genomic library using the human IL-9 cDNA as a probe. The complete nucleotide sequence of the IL-9 gene has been determined and is shown in Fig 1. The sequence reported here agrees with the published IL-9 sequence,[16] except that the current sequence contains more 5′ and 3′ information and there are a few base differences as indicated in the legend of Fig 1. The intron/exon structure of this sequence was deduced from the cDNA sequence and the gene consists of five exons and four introns. AT-rich sequences that are important in controlling the stability of the respective mRNAs have been found within the 3′-untranslated region of the genes for many lymphokines and growth factors.[17] The AT-rich sequences can also be identified in the human IL-9 gene between the translation termination codon and the polyadenylation signal, and comprise four copies of the sequence ATTTA. A "TATA"-like sequence is located 50 nucleotides upstream from the translation initiation codon ATG. Our transient transfection analysis (shown below) and the published S1 mapping analysis[16] clearly demonstrated that this "TATA" sequence is most likely to be recognized by RNA polymerase II in activated normal T cells and in the HTLV-I–transformed cells.

Search for potential regulatory sequences. Many sequences governing transcriptional control and the nuclear factors with which they interact have been extensively characterized.[18] As summarized in Table 1, many sequences related to these signals can be found in the 5′-flanking region of the IL-9 gene. We identified recognition sites for several TPA-inducible transcriptional factors, including AP-1, AP-2, AP-3, AP-5, and NF-kB, at the 5′-flanking region of IL-9 gene. The presence of recognition sequences for these TPA-inducible activation proteins may account for the TPA inducibility of IL-9 in the HTLV-I–transformed T cells. Previous studies have suggested that the cellular transcriptional factors mediating the tax transactivation of the 21-bp repeat of HTLV-I long terminal repeat (LTR) most likely are also responsive to cAMP. In this regard, it is interesting to find in the IL-9 gene the sequence TGAGGT, which is part of the cAMP-response element and has been found to be important for tax-mediated transactivation.[19]

In addition to these sequences, we also identified a perfect octamer sequence ATTTGCAT, an SP1 site, and a CK-1–like sequence. Other sequences including the glucocorticoid response element and several copies of the interferon enhancer and interferon-inducible elements have also been found in the 5′-noncoding region of IL-9 gene.

Chromosomal location of the human IL-9 gene. In situ hybridization of the human IL-9 sequence to chromosome preparations showed a single site of specific labeling on the distal long arm of human chromosome 5. Twenty-one of 80 metaphase cells (26.3%) exhibited silver grains over bands q31-32. Among a total of 266 grains scored, 22 (8.3%) were found at this site (Fig 2) and no other chromosomal site was labeled above background. This portion of human chromosome 5 has been referred to as the 5q− critical region because it is frequently deleted in patients with a myelodysplastic syndrome or acute nonlymphocytic leukemia arising either de novo or secondarily to chemotherapy for prior malignancy.[20]

Comparison of the 5′-flanking region of the IL-9 gene with that of other cytokine genes. A computer search was performed to look for sequence similarities between human

Table 1. Potential Transcriptional Regulatory Sequences in the 5′-Flanking Region of the IL-9 Gene

Consensus Sequences for Eucaryotic Transcriptional Control Elements	Actual Sequences and Nucleotide Numbers* in IL-9 Gene
AP-1	
(TGACTCA)	TGACTCA (735)
AP-2	
(CCCCCAGGC)	TCCCCAGGC (478)
SP1	
(GGGCG)	GGGCG (565)
Octamer	
(ATTTGCAT)	ATTTGCAT (321)
NF-kB	
(GGGACTTTCC)	GGGTTTTTCC (822)
AP-3	
(GGGTGTGGAAAG)	GGGTATGAAAAA (349)
AP-5	
(CTGTGGAATG)	CTGAGGAATG (693)
CK-1	
(GAGATTCCAC)	AAGATCTCCA (195)
Interferon inducible element	TTTCC (531,648,706)
(TTTCC or GGAAA)	GGAAA (575)
Tax response element	
(TGACGT)	TGAGGT (393)
Interferon enhancer	
GT	AAGTGA (301,431,733)
AA GA	AAAGGA (409)
AG	AAGGGA (547)
Glucocorticoid response element	
(AGAACA)	AGAACA (498)

*Nucleotide numbers were assigned according to the numbering system in Fig 1.

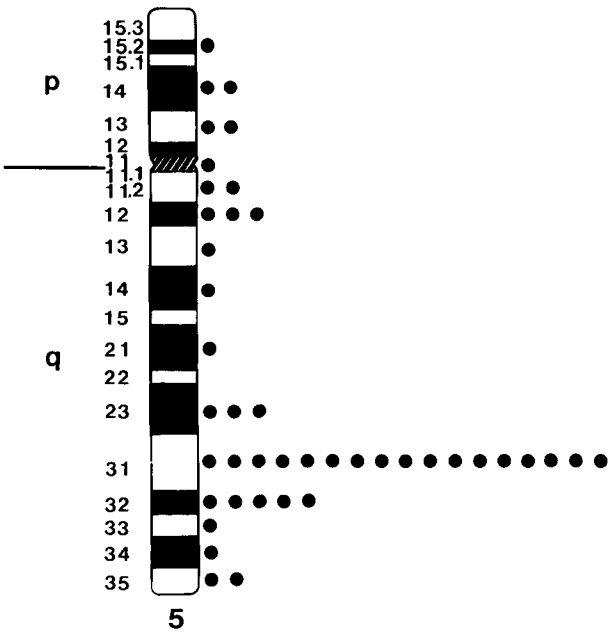

Fig 2. Autoradiographic silver grain distribution along chromosome 5 after in situ hybridization with the human IL-9 genomic sequence.

IL-9 5′-noncoding region and that of genes either mapped to the 5q critical region (IL-3, IL-4, IL-5, and granulocyte-macrophage colony-stimulating factor [GM-CSF])[20,21] or shown to possess T-cell growth factor functions (IL-2 and IL-6). The analysis failed to find overall sequence similarities in the 5′-flanking regions of these cytokine genes. However, limited sequence similarities were detected among these sequences, and the regions where these similarities are statistically significant are shown in Fig 3. Interestingly, several transcriptional control sequences found in the IL-9 gene as shown in Table 1 are also present in these cytokine genes (data not shown). The significance of these transcriptional control sequence blocks shared among different cytokine genes in the coordinated control of their expression requires further investigation.

The 0.9-kb Sma*I-*Sac*I fragment of the IL-9 5′-flanking region is sufficient for its expression in T cells.* To confirm that the IL-9 gene we identified is functional and contains all the sequences responsible for its expression, we constructed a plasmid, phuIL-9-LUC, containing the 0.9-kb *Sma*I to *Sac*I fragment of IL-9 5′-sequences upstream from a firefly luciferase reporter gene. The plasmid DNA was transfected into the HTLV-I–transformed T cells, C10MJ2, and the expression of luciferase activities in these cells was analyzed 24 to 48 hours after transfection under various induction conditions (Table 2). Transfection analysis showed that the 0.9-kb *Sma*I to *Sac*I fragment of the IL-9 gene acts as an active promoter for the downstream reporter gene in the HTLV-I–transformed T cells. This region includes not only sequences essential for the basal level expression but also DNA elements required for the activated expression of the IL-9 gene in the HTLV-I–transformed T cells. These results are consistent with our earlier observation that IL-9 can be constitutively expressed in the HTLV-I–transformed T cells and the expression can be further induced with T-cell activator TPA.[5] Previous studies have shown that IL-9 is not expressed in uninduced normal human T cells but the expression can be induced by PHA, A23187, and anti-CD3 and the effects elicited by these T-cell activators can be further enhanced by TPA.[16] Unlike its promoter activity in normal T cells, the activation of IL-9 promoter sequences in the HTLV-I–transformed cells is constitutive but can be further induced with a single inducer, TPA. T-cell activators such as PHA, A23187, and anti-CD3 did not further activate IL-9 promoter activity in the HTLV-I–transformed T cells.

DISCUSSION

We have isolated and determined the genomic sequence encoding human IL-9. The gene has also been regionally mapped by in situ hybridization to 5q31-32, in the critical region of human chromosome 5 which is often missing in patients with various hematologic disorders and 5q− abnormalities. Our results agree with previously published data,[22] but further refine the localization of the IL-9 gene to the

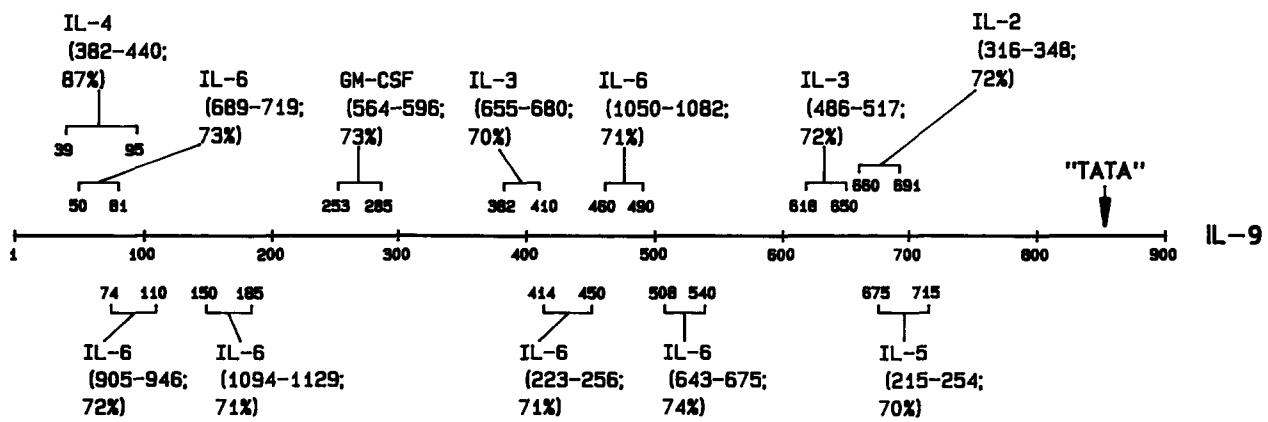

Fig 3. Comparison of the 5′-flanking sequence of human IL-9 gene with that of human IL-2,[26] IL-3,[27] IL-4,[28] IL-5,[29] IL-6,[30] and GM-CSF.[31] Numbers in parentheses represent the nucleotide numbers for each cytokine gene followed by the percent of homology shared by IL-9 and the respective gene in the specified regions. Nucleotide numbers for the transcriptional start site of each cytokine are: IL-9, 880; IL-2, 430; IL-3, 680; IL-4, 1100; IL-5, 510; IL-6, 1200; and GM-CSF, 620.

Table 2. Activation of the Human IL-9 Promoter in C10MJ2 Cells Transfected With phuIL-9-LUC Plasmid Under Different Stimuli

Plasmid	Inducer	Luciferase Activity*
(A) pXP2	—	154
	TPA + PHA	152
(B) phuIL-9-LUC	—	1,545
	PHA	1,748
	A23187	1,054
	Anti-CD3	1,571
	TPA	23,562
	TPA + PHA	25,767
	TPA + A23187	24,405
	TPA + Anti-CD3	26,331

Reproducible results have been obtained using different preparations of plasmid DNAs with either diethylaminoethyl (DEAE)-dextran-mediated transfection or electroporation method. Data presented here are representative of four similar experiments.

*Luciferase activities were adjusted based on the protein concentrations of each cellular extract, pCMV-β-gal, a plasmid encoding β-galactosidase driven by cytomegalovirus (CMV) promoter, was also used as internal control in some of the experiments. The numbers represented the activities from 5 × 10^6 C10MJ2 cells transfected with 2.5 μg of the plasmid DNA.

same mapping position known to contain the cluster of cytokine genes including IL-3, IL-4, IL-5, and GM-CSF. Interestingly, preliminary studies have shown that human IL-9 cDNA probe did not hybridize with YAC clones containing either GM-CSF/IL-3 or IL-4/IL-5 gene cluster, indicating that the IL-9 gene is not closely linked to these genes (W. Neuman and C. Westbrook, unpublished results, September 1990). Furthermore, mouse IL-9 gene was mapped to chromosome 13,[22] whereas mouse IL-3, IL-4, IL-5, and GM-CSF genes were localized to chromosome 11. Mouse homologues of human genes on the long arm of chromosome 5 are distributed on mouse chromosomes 11, 13, and 18. The mapping position of IL-9 appears to be outside the known conserved segment (5q11-5q13) between human chromosome 5 and mouse chromosome 13. Because of the clustering of several growth factor genes on human 5q and the overlapping biologic activities of some of these cytokines, it is tempting to speculate that some of these genes may be coordinatedly expressed. Although there are similarities in the transcriptional control sequences of these cytokine genes, the role (if any) of these *cis* DNA elements in controlling the expression of each cytokine gene needs to be established. Nevertheless, the close proximity and the sequence similarities in the regulatory region suggest a distant evolutionary relationship among at least some of these cytokines. Because several of the cytokines such as GM-CSF, IL-3, IL-4, and IL-9 are capable of supporting cells within the same hematopoietic lineage (eg, erythroid progenitors), it is also possible that temporal control mechanisms are involved in the expression of these genes. The availability of different regulatory sequences and the ability to reintroduce them into different cell types will allow a detailed analysis of the regulatory mechanisms involved in the expression of these genes.

We have introduced the 0.9-kb 5′-flanking region of human IL-9 gene into the HTLV-I–transformed T cells and showed that this DNA segment can govern the expression of the downstream reporter gene in a manner identical to the expression of IL-9 gene in the HTLV-I–transformed T cells. One of the 3′ open reading frames of HTLV-I encodes a 40-Kd nuclear protein, tax, which is responsible for the transactivation from the viral LTR sequences.[23] Recent studies have shown that constitutive expression of many cytokine genes in HTLV-I–transformed T cells is due to the transactivation of their promoter regions by tax.[24,25] It is possible that the HTLV-I tax protein can interact directly or indirectly with the cellular transcriptional machinery and promote the deregulated expression of various T-cell growth factor genes, including IL-9, that may contribute to the transformation process. Our Northern analysis (reference 5 and unpublished results, July 1990) and the transient transfection experiments indicated that the *Sma*I to *Sac*I fragment of the IL-9 gene contains sequences responsible for the gene activation in the HTLV-I–transformed T cells. The presence of the NF-kB site and part of cAMP response element, sequences found to be associated with tax-mediated transactivation, make it tempting to speculate that tax may be involved in the constitutive expression of IL-9 in these cells.

Transfection analysis using deletion or substitution mutants of the IL-9 5′-flanking region will further dissect the *cis*- and trans-DNA elements required for the IL-9 expression in T cells. This information will also be useful in uncovering the regulatory mechanisms involved in the coordinated control of different growth factor productions in the complicated cytokine network.

ACKNOWLEDGMENT

We acknowledge the DNA synthesis group of Genetics Institute for synthesizing oligonucleotides and Dr Steve Nordeen for providing pXP2 vector. We also thank Scott Cooper for preparing the artwork and Becki Robling for assistance in preparing the manuscript.

REFERENCES

1. Van Snick J, Goethals A, Renauld J-C, Van Roost E, Uyttenhove C, Rubira MR, Moritz RL, Simpson RJ: Cloning and characterization of a cDNA for a new mouse T cell growth factor (P40). J Exp Med 169:363, 1989

2. Uyttenhove C, Simpson RJ, Van Snick J: Functional and structural characterization of P40, a mouse glycoprotein with T-cell growth factor activity. Proc Natl Acad Sci USA 85:6934, 1988

3. Moeller J, Hultner L, Schmitt E, Breuer M, Dormer P: Purification of MEA, a mast cell growth-enhancing activity, to apparent homogeneity and its partial amino acid sequence. J Immunol 144:4231, 1990

4. Hultner L, Druez C, Moeller J, Uyttenhove C, Schmitt E, Rude E, Dormer P, Van Snick J: Mast cell growth enhancing activity (MEA) is structurally related and functionally identical to the novel mouse T cell growth factor P40/TCGFIII. Eur J Immunol 20:1413, 1990

5. Yang Y-C, Ricciardi S, Ciarletta A, Calvetti J, Kelleher K, Clark SC: Expression cloning of a cDNA encoding a novel human

hematopoietic growth factor: Human homologue of murine T-cell growth factor P40. Blood 74:1880, 1989

6. Donahue RE, Yang Y-C, Clark SC: Human P40 T-cell growth factor (interleukin-9) supports erythroid colony formation. Blood 75:2271, 1990

7. Lu L, Leemhuis T, Srour EF, Yang Y-C: Effects of human interleukin (IL)-9 on highly enriched erythroid progenitors in normal human bone marrow under serum-containing and serum-depleted culture conditions. J Immunol (submitted)

8. Sanger F, Nicklen S, Coulson AR: DNA sequencing with chain-terminating inhibitors. Proc Natl Acad Sci USA 74:5463, 1977

9. Chen EY, Seeburg PH: Supercoil sequencing: A fast and simple method for sequencing plasmid DNA. DNA 4:165, 1985

10. Harper ME, Saunders GF: Localization of single-copy DNA sequences on G-banded chromosomes by in situ hybridization. Chromosoma 83:431, 1981

11. Yang-Feng TL, Floyd-Smith G, Nemer M, Drouin J, Francke U: Human pronatriodilatin gene is located on the distal short arm of human chromosome 1 and on mouse chromosome 4. Am J Hum Genet 37:1127, 1985

12. Nordeen SK: Luciferase reporter gene vectors for analysis of promoters and enhancers. BioTechniques 6:454, 1988

13. Arya SK, Wong-Staal F, Gallo R: T-cell growth factor gene: Lack of expression in human T-cell leukemia-lymphoma virus-infected cells. Science 223:1086, 1984

14. Sherman PA, Basta PV, Moore TL, Brown AM, Ting JP-Y: Class II box consensus sequences in the HLA-DR alpha gene: Transcriptional function and interaction with nuclear proteins. Mol Cell Biol 9:50, 1989

15. DeWet JR, Wood KV, DeLuca M, Helinski DR, Subramani S: Firefly luciferase gene: Structure and expression in mammalian cells. Mol Cell Biol 7:725, 1987

16. Renauld J-C, Goethals A, Houssiau F, Merz H, Van Roost E, Van Snick J: Human P40/IL-9: Expression in activated CD4+ T cells, genomic organization and comparison with the mouse gene. J Immunol 144:4235, 1990

17. Shaw G, Kamen R: A conserved AU sequence from the 3′ untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 46:659, 1986

18. Jones NC, Rigby PW, Ziff EB: Trans-acting protein factors and the regulation of eucaryotic transcription: Lessons from studies on DNA tumor viruses. Genes Dev 2:267, 1988

19. Giam C-Z, Xu Y-L: HTLV-I tax gene product activates transcription via pre-existing cellular factors and cAMP responsive element. J Biol Chem 264:15236, 1989

20. Le Beau MM, Chandrasekharappa SC, Lemons RS, Schwartz JL, Larson RA, Arai N, Westbrook CA: Molecular and cytogenetic analysis of chromosome 5 abnormalities in myeloid disorders: Chromosomal localization and physical mapping of IL-4 and IL-5. Cancer Cells, vol 7: Molecular Diagnostics of Human Cancer. Cold Spring Harbor, NY, Cold Spring Harbor Laboratory, 1989, p 53

21. Yang Y-C, Kovacic S, Kriz R, Wolf S, Clark SC, Wellems TE, Nienhuis A, Epstein N: The human genes for GM-CSF and IL-3 are closely linked in tandem on chromosome 5. Blood 71:958, 1988

22. Mock BA, Krall M, Kozak CA, Nesbitt MN, McBride OW, Renauld J-C, Van Snick J: IL-9 maps to mouse chromosome 13 and human chromosome 5. Immunogenetics 31:265, 1990

23. Yoshida M, Seiki M: Recent advances in the molecular biology of HTLV-I: Trans-activation of viral and cellular genes. Annu Rev Immunol 5:541, 1987

24. Miyatake S, Seiki M, Malefijt RD, Heike T, Fujisawa J, Takebe Y, Nishida J, Shlomai J, Yokota T, Yoshida M, Arai K, Arai N: Activation of T cell-derived lymphokine genes in T cells and fibroblasts: Effects of human T cell leukemia virus type I p40x protein and bovine papillomavirus encoded E2 protein. Nucleic Acids Res 16:6547, 1988

25. Nimer SD, Gasson JC, Hu K, Smalberg I, Williams JL, Chen ISY, Rosenblatt JD: Activation of the GM-CSF promoter by HTLV-I and -II tax proteins. Oncogenes 4:671, 1989

26. Fujita T, Takaoka C, Matsui H, Taniguchi T: Structure of the human IL-2 gene. Proc Natl Acad Sci USA 80:7437, 1983

27. Yang Y-C, Clark SC: Molecular cloning of a primate cDNA and the human gene for interleukin 3. Lymphokines 15:375, 1988

28. Arai N, Nomura D, Villaret D, Malefict RD, Seiki M, Yoshida M, Minoshima S, Fukuyama R, Maekawa M, Kudoh J, Shimizu N, Yokota K, Abe E, Yokota T, Takebe Y, Arai K: Complete nucleotide sequence of the chromosomal gene for human IL-4 and its expression. J Immunol 142:274, 1989

29. Campbell HD, Tucker WQJ, Hort Y, Martinson ME, Mayo G, Clutterbuck EJ, Sanderson CJ, Young IG: Molecular cloning, nucleotide sequence and expression of the gene encoding human eosinophil differentiation factor (interleukin-5). Proc Natl Acad Sci USA 84:6629, 1987

30. Yasukawa K, Hirano T, Watanabe Y, Muratani K, Matsuda T, Nakai S, Kishimoto T: Structure and expression of human B cell stimulatory factor 2 (BSF-2) gene. EMBO J 6:2939, 1987

31. Miyatake S, Otsuka T, Yokota T, Lee F, Arai K: Structure of the chromosomal gene for granulocyte-macrophage colony stimulating factor: Comparison of the mouse and human genes. EMBO J 4:2561, 1985

1991 77: 1436-1441

Human interleukin-9: genomic sequence, chromosomal location, and sequences essential for its expression in human T-cell leukemia virus (HTLV)-I-transformed human T cells

K Kelleher, K Bean, SC Clark, WY Leung, TL Yang-Feng, JW Chen, PF Lin, W Luo and YC Yang

Blood (print ISSN 0006-4971, online ISSN 1528-0020), is published weekly by the American Society of Hematology, 2021 L St, NW, Suite 900, Washington DC 20036.