Vol. 208, No. 1, 1995 BIOCHEMICAL AND BIOPHYSICAL RESEARCH COMMUNICATIONS
March 8, 1995 Pages 360-367

STRUCTURE OF THE GENE ENCODING THE α SUBUNIT OF THE HUMAN INTERLEUKIN 3 RECEPTOR*

Hiroshi Kosugi[1,2], Yasushi Nakagawa[1], Tomomitsu Hotta[2], Hidehiko Saito[2], Atsushi Miyajima[3], Ken-ichi Arai[1], and Takashi Yokota[1]§

[1]*Department of Molecular and Developmental Biology, The Institute of Medical Science, The University of Tokyo, 4-6-1, Shirokanedai, Minato-ku, Tokyo 108, Japan*

[2]*The First Department of Internal Medicine, Nagoya University School of Medicine, 65 Tsurumai-cho, Shouwa-ku, Nagoya 466, Japan*

[3]*Department of Molecular Biology, DNAX Research Institute of Molecular and Cellular Biology, Palo Alto, California 94304*

Received January 23, 1995

Interleukin 3 is a cytokine that stimulates proliferation and differentiation of hematopoietic progenitor cells. Its receptor consists of two subunits, an interleukin 3-specific α subunit and a β subunit shared by garanulocyte-macrophage colony stimulating factor and interleukin 5 receptors. In this paper, we determined the genomic structure of the α subunit of the human interleukin 3 receptor, which spans approximately 40 kb and has 12 exons. We found that the genomic structures of the α subunits of the human interleukin 3 and granulocyte-macrophage colony stimulating factor receptors are very similar. They possess a unique additional intron in the 'C domain', which is absent in the α subunit of the interleukin 5 receptor. These results suggest a shared evolutionary pathway of these two genes.

Pleiotropy and redundancy are the feature of the cytokine action (1-4). Interleukin 3 (IL-3) exhibits many biological activities common to granulocyte-macrophage colony stimulating factor (GM-CSF) and interleukin 5 (IL-5) (3). Evidence indicates that the shared receptor subunit and signaling molecules in the downstream signal transducing pathway such as Ras, MAP kinase, and JAK kinases contribute to this feature (5,6). The receptor for IL-3 consists of two subunits: an IL-3-specific α subunit

*This work was supported in part by a grant-in-aid for Creative Basic Research (Human Genome Program) from The Ministry of Education, Sience, and Culture, Japan. The DNAX Research Institute is supported by Schering-Plough Corp.

§ To whom correspondence should be addressed.
Fax: 81-3-5449-5424; E-mail: tyokota@ims.u-tokyo.ac.jp.

0006-291X/95 $5.00

(IL-3Rα) and a β subunit shared by GM-CSF and IL-5 receptors. All of these receptor subunits belong to the cytokine receptor superfamily (CRSF) (2,7). Members of this family have a characteristic structural feature in their extra-cellular segments; tandem array of domains each of which is folded as a sandwitch of two antiparallel β sheets similar to the fibronectin type III domain (Fn3-like domain). The CRSF members can be further classified into class 1 and class 2 receptors. Class 1 receptors include most receptor subunits of the hematopoietic cytokines and contain special motifs; four conserved cysteine residues in the C-terminal Fn3-like domain ('C domain') and a 'WSxWS' motif in the N-terminal domain ('N domain') (7).

We previously (8) characterized the gene structure of the human GM-CSF receptor α subunit (hGMRα), in comparison with genes for other CRSF members. This led us to propose a model for gene evolution of the CRSF. We also pointed out that IL-3Rα, IL-5Rα, and GMRα have common genomic as well as protein structures and can be classified into subfamily of the class 1 CRSF. In this paper, we have characterized the structure of the hIL-3Rα gene and found that, besides shared features of the CRSF members, it contains an additional intron unique to the hIL-3Rα and the hGMRα genes. This implies that these two genes are one of the closest siblings in the CRSF.

MATERIALS AND METHODS

Oligonucleotides

Synthetic oligonucleotides used in this study are summarized in Table I. They were all synthesized based on the cDNA sequence (9).

Polymerase chain reaction (PCR)

Exons 7, 8, 9, 10, 11, and 12 were mapped by PCR using subcloned genomic DNAs as templates. Lengths of introns were determined by amplification with pairs of oligonucleotide primers corresponding to the sequences included in adjacent exons. PCR was performed in a thermal cycler (Astec) for 30 cycles (denatured for 1 min at 94°C, annealed for 1 min at 54°C, extended for 5 min at 72°C), using Taq DNA polymerase (Thermalase, Kodak). For PCR to determine the lengths of introns 7 to 12, the reactions were carried out in a P-J 9600 (Perkin-Elmer) for 35 cycles (denatured for 20 sec at 98°C, annealed/extended for 10 min at 68°C) according to Barnes' method (10,11), using Ex Taq DNA polymerase (TaKaRa).

Isolation and subcloning of the lambda phage clones for the hIL-3Rα gene

Three lambda phage libraries of human genomic DNA were screened. Library A was prepared from size-fractionated (10-14 kb) human placental DNA, which was digested completely with *Bam*HI and isolated from gel. We thereafter ligated the DNA with EMBL3 vector (Stratagene) and packaged the product using GIGA PACK II GOLD (Stratagene). Libraries B and C were obtained from American Type Culture Collection (ATCC). Library B is enriched for X-chromosome (Charon 35, 10-21-kb, *Sau*3AI partial, ATCC 57750) and library C is specific for Y-chromosome (Charon 40, 9-24 kb, *Mbo*I partial, ATCC 57780). These libraries were screened with probes from the hIL-3 full-length cDNA or its digestion products. Positive clones were subcloned into pBluescript II KS (+) or SK (-) vector.

P1 phage clones

Since the contig of the lambda phage clones obtained lacked the region spanning the exon 9, P1 phage library was screened by PCR using 91S and 91A primers by Genome Systems Inc. and three clones were isolated. By PCR, we determined the included exons in each of these clones.

DNA sequencing

Exon-intron junctions were sequenced in a 373A DNA sequencer (ABI) according to the manufacturer's instructions. 1-1.5 μg of double-stranded plasmid DNA was used as a template. Primers used in this procedure are listed in Table I.

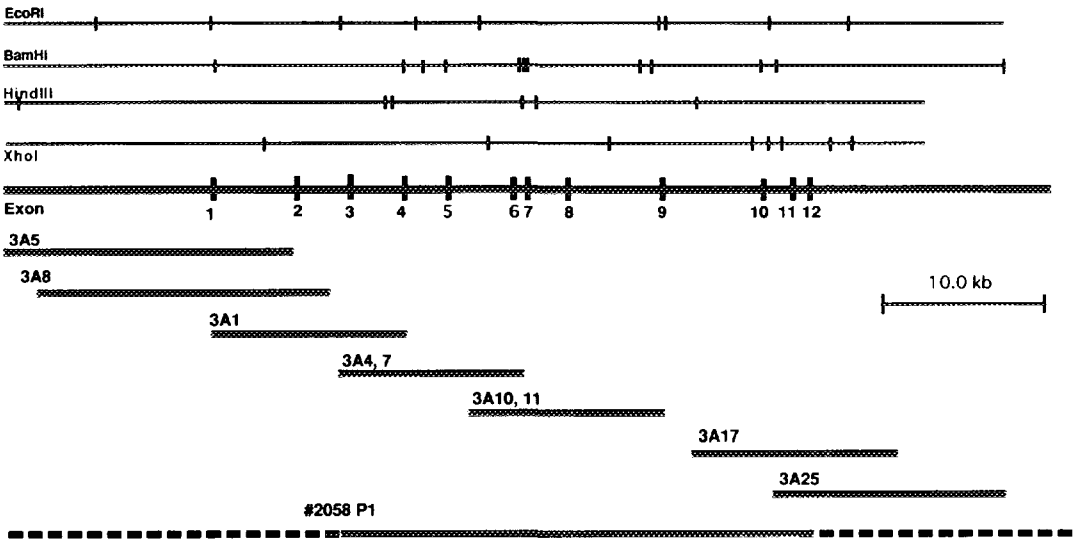
Table I
Oligonucleotides used

exon	primer	sense/antisense	sequence
1	P2	A (69-40)	5'-ttc gtg gtt cca gtt gag aag atc tgg ggt-3'
	P1	A (99-70)	5'-aga tca aag tgc tgt gtg tgg aag aaa cag cc-3'
2	285	S (112-130)	5'-gca cct ctg tcc tgc gtt c-3'
	2A	A (206-189)	5'-ctt cgt ttg cag gag aca-3'
3	10S	S (246-265)	5'-gca aag gct cag cag ttg ac-3'
	10A	A (265-246)	5'-gtc aac tgc tga gcc ttt gc-3'
	286	S (255-273)	5'-cag cag ttg acc tgg gac c-3'
	287	A (304-285)	5'-aca cac tcg ata tcg gtc ac-3'
4	3S	S (390-408)	5'-acc gtc cga gtg gcc aac c-3'
	3A	A (408-390)	5'-ggt tgg cca ctc gga cgg t-3'
	GSP2	A (434-415)	5'-gaa gag gat cca cgt gga ga-3'
5	4S	S (507-527)	5'-agc tgc agc tgg gcg gta ggc-3'
	4A	A (527-507)	5'-gcc tac cgc cca gct gca gct-3'
6	11S	S (615-635)	5'-gat gct cag gga aca cgt atc-3'
	11A	A (635-615)	5'-gat acg tgt tcc ctg agc atc-3'
	7S	S (709-726)	5'-gcg cag cct tcg gta tcc-3'
7	12S	S (803-823)	5'-gac aca ttc ctt tat gca ctg-3'
	12A	A (823-803)	5'-cag tgc ata aag gaa tgt gtc-3'
8	81S	S (879-899)	5'-aga atg cag cct gta atc aca-3'
	81A	A (909-889)	5'-tga cct gtt ctg tga tta cag-3'
	5S	S (891-912)	5'-gta atc aca gaa cag gtc aga-3'
	5A	A (912-891)	5'-tct gac ctg ttc tgt gat tac-3'
9	91S	S (931-951)	5'-tac tca atc ctg gaa cgt aca-3'
	91A	A (995-975)	5'-gct caa gaa ttc ata cac tct-3'
	13S	S (984-1003)	5'-gaa ttc ttg agc gcc tgg ag-3'
	13A	A (1003-984)	5'-ctc cag gcg ctc aag aat tc-3'
10	283	S (1064-1081)	5'-gtc gct gct gat cgc gct-3'
	284	A (1120-1101)	5'-cag atc acg aag aca cag ac-3'
11	6S	S (1128-1148)	5'-tat ctg gtg atg cag aga ctc-3'
	6A	A (1148-1128)	5'-gag tct ctg cat cac cag ata-3'
	282	A (1200-1181)	5'-cgt ttt gga agc tgt cac cg-3'
12	8S	S (1278-1295)	5'-act tga gac tgg ggt tca-3'
	8A	A (1295-1278)	5'-tga acc cca gtc tca agt-3'
	288	S (1405-1423)	5'-ggc atg gga gat gcc tgt g-3'

RESULTS

Genomic Southern blot analysis revealed two positive *Bam*HI fragments ranging in size between 10-14 kb. Based on this result, library A was constructed and screened. To facilitate the characterization of the hIL-3Rα gene which maps to the X-Y pseudoautosomal region (PAR), two chromosome-specific libraries B and C were also used. We screened total of 1.0×10^5, 1.3×10^6, and 1.5×10^6 clones of libraries A, B, and C, respectively, with various combinations of probes and obtained positive clones.

A



B

Exon No.	size(bp)	codon phase	Intron donor	size(kb)	Acceptor
1	108		TTAAG gtaaggtccc	5.3	tctcctccag CAGGC
2	102	1	G GAA G gtaagaactg	3.0	accgttttag AT CCA
3	129	0	TG CCG gtaaatcata	2.9	tctctcttag GCA GT
4	145	1	G AAC A gtgagaaaaa	2.5	gtgaacccag GT GGG
5	133	2	GCC AA gtaagtgtgc	4.9	cttaccgcag C AGG C
6	85	1	G ATT G gtgagtagcc	1.8	tcgttgctag AG ATA
7	126	0	AA AAG gtaaactttc	2.0	tgtgttgcag AGA AT
8	27	0	AA CAG gtagtgctct	6.0	caaaccacag GTC AG
9	115	1	C TTC G gtgttccgat	6.4	ttacccctag AG TGC
10	106	2	AGA AG gtgagccctc	2.0	ttcctccgag G TAT C
11	82	0	AG CTG gtatgttgtt	1.1	gtctctgcag GTG GT
12	252				

Fig.1.

Organization of the hIL-3Rα gene and the conserved gene structure in the CRSF members.

A. Fine restriction endonuclease cleavage map of the hIL-3Rα gene. *Eco*RI, *Bam*HI, *Hind*III,and *Xho*I sites are shown. No lambda clone was found which includes exon 9. This exon is included only in a P1 clone (2058). Clone 3A1 and clones 3A4, 5, 7, 8, 10, and 11 were isolated from library A and library B, respectively. They were screened with a full-length cDNA probe excised by *Xho*I from pCEV4 expression vector. Clones 3A17 and 25 were isolated from library C with a labeled 490-bp *Eco*RI-fragment of the cDNA.

B. Exon-Intron junctions of the hIL-3Rα gene and its intron phase.
From exon 1 to exon 4, exon-intron junctions were determined mainly by sequencing subcloned DNA using universal primers. From intron 7 to intron 12, their lengths were determined by 'long and accurate PCR' (10,11).

Since the contig of the lambda phage clones was found to lack the region spanning the exon 9, we obtained P1 phage clones 2057, 2058, and 2059 by PCR.

The isolated lambda and P1 phage clones cover the entire coding sequence of the hIL-3Rα gene (Fig. 1A). The restriction endonuclease cleavage map is consistent with the result of the genomic Southern blot analysis (Fig. 2). The gene spans approximately 40 kb and contains 12 exons. All exon-intron junctions follow the GT-AG rule (12) (Fig. 1B). The translation initiation site is located on exon 2. The three Fn3-like domains are

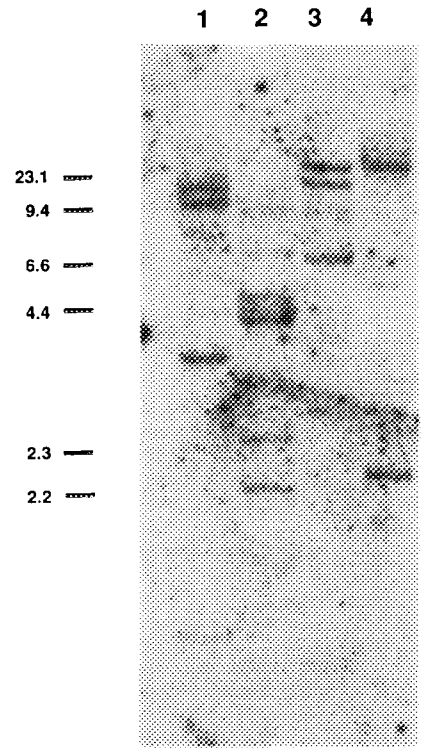


Fig.2.
Genomic Southern hybridization of the hIL-3Rα gene.
The restriction endonuclease cleavage map of the hIL-3Rα gene shown in Fig. 1A was consistent with this result. Ten micrograms of genomic DNA extracted from human placenta of healthy volunteer was digested with *Bam*HI (lane 1), *Eco*RI (lane 2), *Hind*III (lane 3), and *Xho*I (lane 4), respectively. The full-length cDNA of the hIL-3Rα was used as a probe. DNA size-marker is shown (kb).

encoded by total of seven exons. The two conserved Fn3-like domains of the CRSF are encoded by five exons; the 'N domain' (N-terminal domain with conserved four cysteine residues) is encoded by exons 5 and 6 and the 'C domain' (C-terminal domain with a 'WSxWS' motif) is encoded by exons 7, 8, and 9. Another Fn3-like domain ('N' domain') which is specific to IL-3Rα, IL-5Rα, and GMRα, is encoded by exons 3 and 4.

Previous reports suggest that the hGMRα gene (*CSF2RA*) is located within 190 kb of the hIL-3Rα gene (*IL3RA*) (13,14). However we did not find hGMRα coding sequences within 15 kb upstream and 8 kb downstream of the hIL-3Rα gene in our genomic clones.

DISCUSSION

We isolated genomic clones that cover the entire coding sequences of the hIL-3Rα gene by screening chromosome-specific libraries and P1 phage library. The gene

spans approximately 40 kb, which is one of the largest among known cytokine receptor genes (8,15-22).

In our previous paper, we compared the exon organization and the corresponding intron phases (23) for the available genes encoding members of the CRSF (8). An intron phase is defined as the position where it disrupts the coding sequence. For example, when intron is located between the first and the second nucleotides of a codon, this intron is called phase 1. We found that in all the known members of the CRSF, the intron phases are conserved around the tandem array of the two Fn3-like domains; the boundaries of each Fn3-like domain are defined by phase 1 introns. Furthermore, the phases of the intervening introns are 2 for the 'N domain' and 0 for the 'C domain', respectively. This '1-2-1-0-1 rule' suggests that a common origin of Fn3-like domains with the phase 2 or 0 introns exists among the CRSF members, and that exon shuffling plays an important role in the evolution of these genes. In this study, we confirmed that this '1-2-1-0-1 rule' also applies to the hIL-3Rα gene (Fig. 3A).

In addition to the above two Fn3-like domains that are common to all the CRSF members, IL-3Rα, IL-5Rα, and GMRα contain yet another Fn3-like domain in their N-termini[1]. We have found that the boundaries of this 'N' domain' are also defined by phase 1 introns and that the intron within this domain is of phase 0. Based on this conservation of the exon-intron organization as well as the protein structure among these three genes (8,24), we assume that the IL-3Rα, the IL-5Rα, and the GMRα genes have a common origin (Fig. 3B).

Moreover, by comparing the genomic structures of the three genes in detail, we found an intriguing difference between the IL-3Rα/GMRα and the IL-5Rα genes. The hIL-3Rα and the hGMRα genes contain an additional intron in their 'C domain', whereas the mIL-5Rα gene does not (Fig. 3A). This feature is conserved in the IL-3Rα and the GMRα genes, and is not restricted to human since this intron is also found in the mouse IL-3Rα gene (25). This intron is of phase 0 for the hIL-3Rα, the mIL-3Rα, and the hGMRα genes and its position in the protein corresponds to the 4th and the 5th β strand (7,26,27). It is tempting to speculate that the IL-3Rα and the GMRα genes have diverged very recently in the evolution of the CRSF and share evolutionary process apart from the IL-5Rα gene (Fig. 3B).

The 5' untranslated regions of the hGMRα and the hIL-3Rα genes are encoded by three and two exons, respectively, judging from the reported cDNA. To confirm this difference, we tried to determine the transcription start site of the hIL-3Rα gene by primer extension and 5' RACE, but no distinct cDNA end was detected. This may be due to the high GC content of the putative 5' untranslated region. For example, we have sequenced 580 bp of the 5' flanking region from the reported cDNA end and found SINES-homologous region (28) with the GC content of more than 80%. This sequence

---

[1] J.F.Bazan, personal communication.

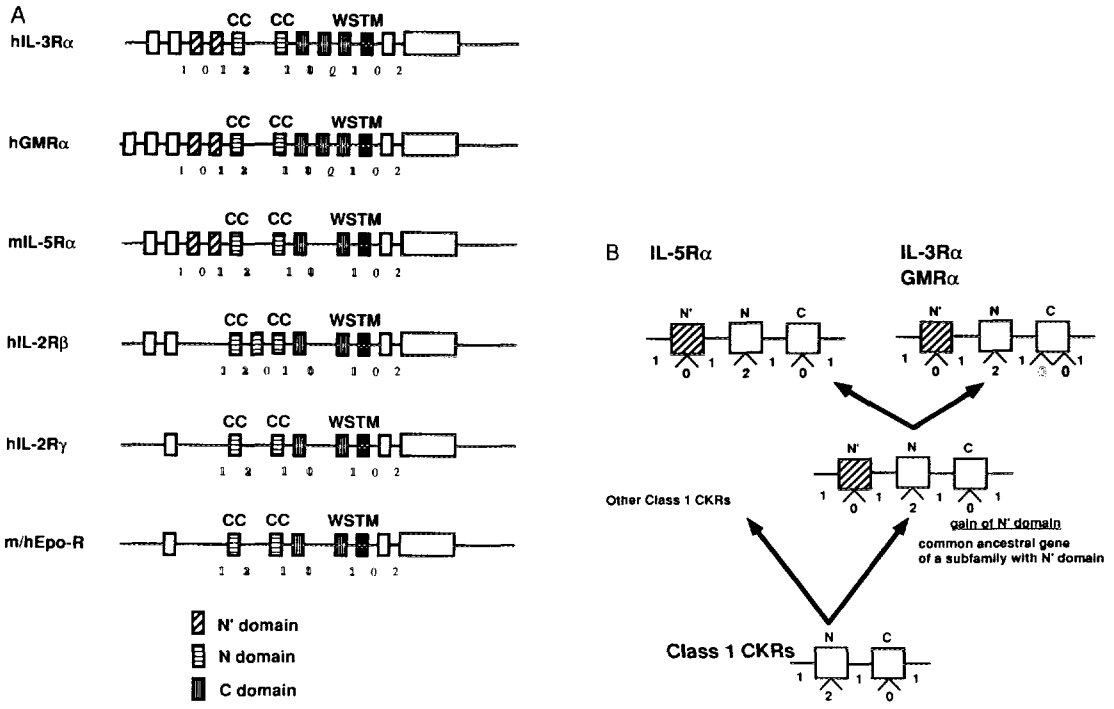


Fig.3.
A. Schematic representation of genomic structure of the CRSF members.
'1-2-1-0-1 rule' of intron phasing is well-conserved among the CRSF members including the hIL-3Rα gene. Another Fn3-like domain ('N' domain') exists in its N terminus of hIL-3Rα, hGMRα , and mIL-5Rα, respectively, and intron phase around this domain is 1-0-1. It should be noted that the hIL-3Rα and the hGMRα genes possess an extra intron of phase 0 (underlined italic) in the 'C domain'.
B. A possible model for gene evolution of the CRSF.
A putative common ancestral gene for the CRSF have evolved from a prototype exon encoding an Fn3-like domain by duplication and intron insertion. After class 1 cytokine receptors (CKRs) had evolved, a common ancestral gene for the IL-3Rα, the GMRα, and the IL-5Rα genes emerged by acquiring the third Fn3-like domain. These three receptors can be classified into a subfamily because of N' domain at their N-termini. The IL-3Rα and the GMRα genes possess an additional intron in the 'C domain', whereas the IL-5Rα gene does not.

would interfere with the reverse transcription used in the above two procedures. Therefore, we cannot exclude the possibility that the hIL-3Rα gene contain one more exon in its 5' end.

In conclusion, we isolated the genomic clones for the hIL-3Rα gene and compared its structure with gene structures for other CRSF members. From this analysis, we have obtained a novel insight pertaining to the origin of three α subunit genes encoding IL-3R, GMR, and IL-5R. We also found that exon-intron structure of the hIL-3Rα gene is remarkably similar to that of the hGMRα gene among these three genes, which implies that the IL-3Rα and the GMRα genes might have evolved separately from the IL-5Rα gene (Fig. 3B).

Acknowledgments: We thank the staff in the Laboratory of Molecular Genetics (The Institute of Medical Science, The University of Tokyo (IMSUT)) for oligonucleotide synthesis and the Human Genome Center (IMSUT) for homology-search. We thank J.F.Bazan (DNAX) for communicating unpublished results; M.Muramatsu, S.Watanabe, N.Koyano-Nakagawa, and A.Muto (IMSUT) for helpful discussion.

REFERENCES

1. Arai, K., Lee, F., Miyajima, A., Miyatake, S., Arai, N. and Yokota, T. (1990) Annu. Rev. Biochem. 59, 783-836.
2. Miyajima, A., Kitamura, T., Harada, N., Yokota, T. and Arai, K. (1992) Annu. Rev. Immunol. 10, 295-331.
3. Miyajima, A., Mui, A. L.-F., Ogorochi, T. and Sakamaki, K. (1993) Blood 82, 1960-1974.
4. Kishimoto, T., Taga, T. and Akira, S. (1994) Cell 76, 253-262.
5. Darnell, J. E., Kerr, I. M. and Stark, G. R. (1994) Science 264, 1415-1421.
6. Quelle, F. W., Sato, N., Witthuhn, B. A., Inhorn, R. C., Eder, M., Miyajima, A., Griffin, J. D. and Ihle, J. N. (1994) Mol. Cell. Biol. 14, 4335-41.
7. Bazan, J. F. (1990) Proc. Natl. Acad. Sci. USA. 87, 6934-6938.
8. Nakagawa, Y., Kosugi, H., Miyajima, A., Arai, K. and Yokota, T. (1994) J. Biol. Chem. 269, 10905-10912.
9. Kitamura, T., Sato, N., Arai, K. and Miyajima, A. (1991) Cell 66, 1165-1174.
10. Barnes, W. M. (1994) Proc. Natl. Acad. Sci. USA 91, 2216-2220.
11. Cheng, S., Fockler, C., Barnes, W. M. and Higuchi, R. (1994) Proc. Natl. Acad. Sci. USA 91, 5695-5699.
12. Breathnach, R. and Chambon, P. (1981) Annu. Rev. Biochem. 50, 349-383.
13. Milatovich, A., Kitamura, T., Miyajima, A. and Francke, U. (1993) Am. J. Human. Genet. 53, 1146-1153.
14. Kremer, E., Baker, E., D'Andrea, R. J., Slim, R., Philips, H., Moretti, P. A. B., Lopez, A. F., Petit, C., Vadas, M. A., Sutherland, G. R. and Goodall, G. J. (1993) Blood 82, 22.
15. Shibuya, H., Yoneyama, M., Nakamura, Y., Harada, H., Hatakeyama, M., Minamoto, S., Kono, T., Raymond, W. and Tanigichi, T. (1990) Nucleic Acids Res. 18, 3697-3703.
16. Penny, L. A. and Forget, B. G. (1991) Genomics 11, 974-980.
17. Gorman, D. M., Itoh, N., Jenkins, N. A., Gilbert, D. J., Copeland, N. G. and Miyajima, A. (1992) J. Biol. Chem. 267, 15842-15848.
18. Wrighton, N., Campbell, L. M., Harada, N., Miyajima, A. and Lee, F. (1992) Growth Factors 6, 103-118.
19. Pleiman, C. M., Gimpel, S. D., Park, L. S., Harada, H., Taniguchi, T. and Ziegler, S. F. (1991) Mol. Cell. Biol. 11, 3052-3059.
20. Noguchi, M., Adelstein, S., Cao, X. and Leonard, W. J. (1993) J. Biol. Chem. 268, 13601-13608.
21. Ito, Y., Seto, Y., Brannan, C. I., Copeland, N. G., Jenkins, N. A., Fukunaga, R. and Nagata, S. (1994) Eur. J. Biochem. 220, 881-91.
22. Mignotte, V., Vigon, I., Boucher de Crevecoeur, E., Romeo, P. H., Lemarchandel, V. and Chretien, S. (1994) Genomics 20, 5-12.
23. Sharp, P. A. (1984) Cell 23, 643-646.
24. Imamura, F., Takaki, S., Akagi, K., Ando, M., Yamamoto, K., Takatsu, K. and Tominaga, A. (1994) DNA Cell Biol. 13, 283-292.
25. Miyajima, I., Levitt, L., Hara, T., Bedell, M. A., Copeland, N. G., Jenkins, N. A. and Miyajima, A. (1995) Blood (in press).
26. De Vos, A. M., Ultsch, M. and Kossiakoff, A. A. (1992) Science 255, 306-312.
27. Somers, W., Ultsch, M., De Vos, A. M. and Kossiakoff, A. A. (1994) Nature 372, 478-481.
28. Britten, J., Baron, W. F., Stout, D. B. and Davidson, E. H. (1988) Proc. Natl. Acad. Sci. USA 85, 4770-4774.