Cloning and expression of the rat interleukin-3 gene

D.R.Cohen, A.J.Hapel* and I.G.Young

Medical Molecular Biology Unit and *Department of Medicine and Clinical Science, John Curtin School of Medical Research, Australian National University, Canberra, 2601, Australia

Received 11 March 1986; Accepted 11 April 1986

ABSTRACT

Genomic clones carrying the rat interleukin-3 (IL-3) gene have been isolated and the nucleotide sequence of the gene determined. Alignment of this sequence with that of the mouse IL-3 gene has allowed the structure of the rat IL-3 gene to be deduced. The intron-exon boundaries are conserved and extensive nucleotide homology (approx 90%) is present in the 5' flanking region and the portion of the gene coding for the signal peptide. Several proposed regulatory sequences are conserved and an analogous element to the tandem repeat in intron 2 of the mouse gene is also present. The predicted amino acid sequence for mature rat IL-3 shows surprisingly low homology (54%) with its murine counterpart, although all four cysteine residues are conserved. The rat IL-3 gene was expressed in monkey COS-1 cells and colony assays established that rat IL-3 is a multi-lineage haemopoietic growth regulator. There was little cross-reactivity of the respective IL-3 species on mouse and rat bone marrow cells suggesting that rat IL-3, in concert with its receptor, has evolved significantly away from the mouse IL-3/receptor system.

INTRODUCTION

Interleukin-3 (IL-3) is a member of the family of colony stimulating factors believed to regulate haemopoiesis $^{1-3}$. It is involved in regulating growth and differentiation of pluripotent stem cells leading to the production of all the major blood cell The major natural source of IL-3 is the antigen-stimutypes. lated T lymphocyte, although it is also produced by a number of continuous cell lines 4-5. IL-3 has been extensively characterized in the murine system, and $cDNA^{6,7}$ and genomic^{8,9} clones have Little is known about IL-3 species in been reported recently. There has been some biological characterization other mammals. of mucosal mast cell growth factor in rats which may be equivato $IL-3^{10-13}$. Pluripotent stem cell factor activity has lent been reported to be produced by the human bladder carcinoma cell line 5637¹⁴. but it is still not established whether an exactly analogous lymphokine to murine IL-3 exists in man. Southern

© IRL Press Limited, Oxford, England.

hybridization analysis of mammalian DNA's, using a murine IL-3 cDNA probe, fails to detect homologous sequences in most mammalian species, even under conditions of relatively low stringency (unpublished data). This apparent low conservation of mammalian IL-3 genes contrasts with genes for other lymphokines such as IL-2, GM-CSF and γ -interferon which cross-hybridize between mouse and human¹⁵⁻¹⁷.

As part of a study of the structure and biological role of IL-3 in other mammals, and in view of the possibility of studying the role of IL-3 using rat experimental models, we felt it was desirable to isolate the rat IL-3 gene and express it in animal cells. This is described in the present work together with the determination of some of the biological properties of recombinant rat IL-3.

MATERIALS AND METHODS

General Methods

(PVGxDA) male laboratory rats were used as the source of DNA for all experiments. Genomic DNA was prepared by homogenization and treatment with sarkosyl and proteinase K, followed by centrifugation in CsCl/ethidium bromide gradients 18,19 . Large scale λ DNA preparations were made from 1 litre cultures in Luria broth, lysed after 6 hr aeration at 37° C. Following addition of $CHCl_3$, phage DNA was purified as described²⁰. Small scale. preparations of λ DNA from plate lysates on Luria + Mg agarose plates were as described⁸, and plasmid DNA minipreparations were made by the rapid boiling method²¹. Large scale plasmid preparations were carried out according to Clewell and Helinski 22 and plasmid DNA was purified using CsCl/ethidium bromide density gradients.

DNA fragments were isolated by electrophoretic separation on low-melting point agarose gels. Transformations were carried out using the high efficiency method of Hanahan 23 .

The primers used for probe preparation and sequencing were synthesized by the phosphoramidite method. Autoradiography was carried out at -70° C using Fuji X-ray or Kodak XAR-5 film with Dupont Lightning Plus intensifying screens.

Construction and Screening of the Rat Genomic Library

Total genomic DNA from rat liver was partially digested with <u>Sau</u>3A and fragments in the size range 9-20kb were purified from low-melting point agarose. These fragments were ligated with λ EMBL3A arms prepared by <u>Eco</u>RI/<u>Bam</u>HI digestion of the phage DNA, and then packaged into bacteriophage particles^{21,24}. Percentage recombinants was determined by <u>spi</u> selection²⁵.

The library was plated out for screening (using <u>E.coli</u> strain ED8655) on L top agarose/L agar in glass baking dishes (20cm x 30cm), approx 250,000 pfu per dish. Nitrocellulose replicas of plates containing plaques were prepared for hybridization as described by Benton and Davis²⁶.

Filters were routinely pre-washed in 50 mM Tris-HC1 pH8/1M NaC1/1mMEDTA/0.1% SDS at 42° C for 1-2 hr, pre-hybridized in 6xSSC/5xDenhardt's solution/ 10mM EDTA/0.5% SDS/50µg/ml salmon sperm DNA at 65°C for 1-4 hr and hybridized in the same solution plus probe at 65°C for 18 hr. Filters were rinsed in 5xSSC/0.1%SDS and then washed twice in 2xSSC/0.1%SDS at 65°C for 45 min.

Probe Preparation

The 467 bp <u>HindIII-Nco</u>I fragment of the murine IL-3 cDNAcontaining plasmid pILM3⁶ was isolated and used as a template for primed synthesis of radioactively labelled cDNA, using a synthetic random decamer as primer. Probe preparation was as described⁸, except that synthesized probe was routinely separated from un-incorporated label and very small labelled fragments by chromatography on Sephadex G-50 (fine).

Southern Hybridization

Restriction endonuclease digests of total genomic DNA ($15 \mu g$ per lane) or recombinant bacteriophage DNA ($0.5-1 \mu g$ per lane) were electrophoresed on 1% agarose/TAE ($40 \mu MTris$ -acetate/lmM EDTA) gels at 25mA for 16 hr. The gels were treated and blotted to nitrocellulose as described²⁷. Pre-washing, pre-hybridization and hybridization of filters was as for genomic library screening. Following hybridization, filters were rinsed briefly in 5xSSC/0.1%SDS at room temperature, and then washed twice in 4xSSC/0.1%SDS at $65^{\circ}C$ for 45 min. Stepwise reductions of salt

concentration (down to 0.1xSSC) were used for additional washes when necessary.

DNA Sequence Analysis

sequence of the HindIII fragment covering the rat IL-3 The gene was determined by the chain termination method of Sanger 28 , using a universal flanking primer²⁹. The DNA was prepared for sequencing by sonication of self-ligated fragment, Τ4 DNA polymerase repair and electrophoresis in low melting point agarose³⁰. Fragments in size ranges 300-500 bp, 500-1000bp and 1000-2000bp were isolated from the gel and cloned into the SmaI site of M13mp10³¹. Competent cells of E.coli strain TG1, prepared by the method of Hanahan 23 , were used for the transfection of the recombinant M13 DNA. Gels were routinely 5% acrylamide with "wedged" bottoms, and were dried to increase resolution³². Compressions were checked using gels containing formamide³³. DNA sequence data was entered directly from 25% autoradiograms into the computer using a digitizer, and assembly and analysis of this data was performed using the computer programs of Staden³⁴⁻³⁶. The Genbank database was searched using the progam of Wilbur and Lipman 37 .

Expression of Rat Interleukin-3

COS-1 cells were seeded at 5×10^5 cells per 60mm diameter petri dish, grown overnight in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% foetal calf serum (FCS), then transfected with the IL-3 expression vectors. 4µg of plasmid DNA plus 16µg of carrier DNA was transfected per 10^6 cells by calcium phosphate precipitation and glycerol shock (15% glycerol for three minutes)^{38,39}. The cells were washed in DMEM containing 10% FCS, and incubated in 5ml of the same medium at 37° C in 5% CO₂. Supernatants were collected at 72 and 144 hours posttransfection and filtered through Amicon 0.2µm filters. Samples were stored at 4° C prior to assay.

Assays for IL-3

1. <u>Colony forming assays.</u> Colony assays were done using bone marrow cells from WISTAR rats or BALB/c mice, scoring colonies on day 7 of culture³⁹. The entire culture was fixed with 2.5% glutaraldehyde then stained with Lutol Fast Blue-hematoxylin to identify granulocyte-macrophage, erythroid, mixed erythroid and megakaryocytic colonies. Each culture contained 75,000 bone marrow cells in 1.5 ml.

2. Bone marrow cell proliferation assays. Serial two-fold dilutions of test material were made in microtitre plates in 50μ l RPMI 1640, 10% FCS. $5x10^4$ WISTAR rat or BALB/c mouse bone marrow cells were added to each well. Bone marrow cells from PVG, DA, PVG x DA, JC, PVG x JC and Fischer rats were also tested in this assay. Cultures were pulsed overnight on day 3 with [³H]thymidine. Activity is expressed as a reciprocal of the titration endpoint, the endpoint being defined as the supernatant dilution at which activity is no longer detectable⁴⁰.

RESULTS

Southern Hybridization of Rat Genomic DNA

genomic DNA was digested with the restriction enzymes Rat EcoRI and HindIII, electrophoresed on an agarose gel, blotted to nitrocellulose and hybridized with a murine IL-3 cDNA probe. A single hybridizing fragment of appropriate intensity for a single-copy gene was observed in the HindIII digest at approx. 5.8kb (Fig.1). The EcoRI digest gave two bands, suggesting the presence of an EcoRI site within the rat IL-3 gene. The reason the different intensity of the two EcoRI bands is not for clear may be due to a greater representation of sequences derived but from particular regions of the template fragment in the random primer probe.

Screening of the Rat Genomic Library

A library of approx. 10⁶ recombinant phage was generated by cloning 9-20 kb fragments from a partial <u>Sau</u>3A digest of total genomic Rat DNA into λ EMBL3A arms created by <u>Bam</u>HI/<u>Eco</u>RI diges-The complete library was plated out onto 4 baking dishes tion. (approx. 250,000 recombinant phage per plate) and duplicate lifts were taken from each dish. A total of 5 plaques were identified by hybridization to the 467 bp <u>Hin</u>dIII-<u>Nco</u>I fragment of murine IL-3 cDNA labelled with ^{32}P by randomly primed synthesis. These 5 clones were purified and small scale λ DNA preparations were Digestion of the DNA with <u>Eco</u>RI showed the five carried out. clones to be overlapping, each containing one or both of the two expected EcoRI fragments (Fig. 2). Clone $\lambda R3$, which contained



Figure 1 Southern blot analysis of rat genomic DNA and DNA from the genomic clone $\lambda R3$, probed with an $[\alpha - {}^{32}P]$ dATP - labelled murine IL-3 cDNA fragment derived from pILM36. Gel electrophoresis, blotting, probe preparation and hybridization conditions are given in Materials and Methods. Restriction endonuclease abbreviations are E, EcoRI and H, HindIII. Sizes given are in kilobase pairs (kb).

both fragments, was grown in large scale for further characterization. Southern hybridization of enzyme digests of λ R3 DNA shows that it contains the fragments originally seen in the genomic DNA digests (Fig. 1) and allows localizaton of the rat IL-3 gene to a pair of fragments (3.3kb and 1.25kb) from a BamHI/HindIII double digest (Fig.2).

Nucleotide Sequence of the Rat IL-3 Gene

The nucleotide sequence of the 5.8 Kb <u>Hin</u>dIII fragment encompassing the rat IL-3 gene was determined on both strands by the chain termination method. The sequence of this region



Figure 2 Restriction map of the rat IL-3 gene showing the regions covered by the 5 overlapping clones isolated from a λ EMBL3A library of rat genomic DNA. The map of the gene was deduced from the characterization of purified λ R3 DNA. Proposed exons are indicated by solid boxes, and cover the region from the AUG codon believed to be the start site of translation through to the translational termination codon UAA. Restriction endonuclease abbreviations are E, EcoRI and H, HindIII.

(excluding the 5' 1.25kb <u>HindIII/Bam</u>HI fragment) is presented, together with an alignment against the murine IL-3 gene sequence, in Fig.3. The overall homology between the two nucleotide sequences in the coding regions is 76%; the introns and flanking regions (approx. 200 nucleotides either side of the gene) are more highly conserved, with 80% and 90% homology respectively. Structure of the Rat IL-3 Gene

Sufficient conservation of gene structure was present to allow the structure of the rat IL-3 gene to be deduced by alignment with the gene for murine IL-3 (Fig. 3). The rat IL-3 gene consists of five exons interrupted by four introns with conservation of the intron/exon boundaries between the rat and mouse genes. The deduction of the gene structure is greatly assisted by the high sequence homology through these regions. All introns interrupt the coding sequence between codons.

Sequence homology in the region corresponding to the leader sequence and extending 400 nucleotides upstream of the translational start site in murine IL-3 is extremely high. It is therefore likely that translation of rat IL-3 commences at nucleotide position 1374, making the first ten amino acids identical to those in the leader portion of murine IL-3. Transcriptional

(R)	10	20	30	40	50	60	70	80
GATCCAGG	AGACCAGTCT	CCTAGTACCA	GGTCTGCTTG	CCTAAACTTG	GAGTATAAGA	GCCATAGACA	CTGTCTCTTC	GATCAGTCCTTG
CCCCCACC	100	110	120	130	140	150	160	170
	CCCTGCTGTT	GCACCCTTAT	CTTTCACCCT	ATTGCTCCTG	Cattgaagac	AGAAGCACCC	AGTTTCCCCT	GCCTCAGCATAA
CTTGCTAG	190	200	210	220	230	240	250	260
	CCTTCATTTC	CTCGTGCTGG	TCACATCACA	CCACAACCCG	ACCCAAACCC	IGGIIIICICI	ACCATGCCCC	IGCTTCCCTGCA
CCCCAGGC	280	290	300	310	320	330 ·	340	350
	TTGTCACACT	CATCTTCTAC	CAAAACTCCA	GCTTTGTGCT	GTGGCCTGTC	AACCTGTCCC	ATGGAAAAGG	GGGCCACCCCAT
CCTTCAGG	370	380	390	400	410	420	430	440
	GACTGTCCCC	IGGCTCTCCA	CACTCCTGGC	TTTGCCACTT	ICTCTCTAGC	IGTGGTTTCT	CAGGTCCTTT	GAGAACTTCCCA
TAACTGTC	460	470	480	490	500	510	520	530
	CCTGTITCCT	ICCCACCTCT	-GTAGGCCTG	AGCTGCAAAC	CAGCTCCCAC	TCCACCCAGG	CTCCAGGGCC	GACTGGGATTTA
GATCCCTC	550	560	570	580	590	600	610	620
	AATATGGCTT	ICCTTCAGGG	AGTAGTTCTC	FTCTCTCCTC	FTGCCCTCCC	GGCTCAAACT	TGTCCATGCC.	ACCTGCTACACC
GTCCTGCA	640	650	660	670	680	690	700	710
	GCTCCCAGTA	ACTAACACTA	TTCTCAAGGC	CCACCTTTGT	CCTAGGTCC	CTAAGCCTAA	TTATCTGAGT	TATCAGAAGGAT
GGCCTAGT	730	740	750	760	770	780	790	800
	GTTTGCAGTC	ATATCTCCAT	CAAGGGTTCT	GTCCTCTAGA	IGTGGGCCTT	AGCGCATTGC	CTTACTGCAC	TGAGACTAGACC
AGTGAAGG	820	830	840	850	860	870	880	890
	AGTGAGCTGA	ACTCCATATC	CACCTGCAAG	GAATAAGGGT	CAATGGGAAG	GCTGCCTAGA	GGGAGAGGGA	GCTCTAGCTACC
AGCGGCCA	910	920	930	940	950	960	970	980
	GAGGACTAGC	CCACCCATGG	ACGTTTAACC	ATGTGCCAGA	ATGCCTACCA	IGTTCAAGTT	TGCCCCAGTG	ACCCTGGTGGCC
САСТААТА	1000	1010	1020	1030	1040	1050	1060	1070
	GTGGTGGCCC	ACAGTCAGGG	GCAGATTTGT	ACAAGGGATG	GTAGGAAGAG	GTTCCAGTGC	ACAGAAACCC	CAAGCTGGCTCG
	(M)	CAGTCAGGG	TCAAGTTTGT	GCAAGGGATG	GTAGGATGAG	ATTCCACTGC	ATAGAAAGCC	CAAGCTGGCTCA
GAGCCAGG	CTACTTCCTC	CCACCACCTG	TTTCCACTCG	GTCCATCTCT.	1130 ATGACAAAGG	AAGAAGATGG	CCTTTGAATA	AGCAGTCTTTCT
GAGCCAGG	1180	1190	1200	1210	1220	1230	1240	1250
TCCCATGT	CGATAATTTT CGATAATCTT	GAGTACTAGA :::::::::: GAGTACTAGA	AAACGATGAA 	TAAGTCTGTG 111111 1111 TAAGTTTGTG	STTIGCTATG STTIGCTATG	GAGGTTCCAT IIIIIIIII GAGGTTCCAT	GTCAGATAAA GTCAGATAAA GTCAGATAAA	GCTGCTTCTGAT
GCCTGCCC	1270	1280	1290	1300	1310	1320	1330	1340
	TTCCCCCCAT	GCCCTGCCTG	GGGCCCGCCC	CGCCCCTCTC	GATGAATATA	IATAAGGTGA	AGGCTCCTGT	GGCTTCTTCAGA
GCCTGCCC	TCCCCCCT	GCCCCGCCGG	GC-CCCGCCC	CACCCCTCTC	TGAATACA	TATAAGGTGA	AGGCTCCTGT	GGCTTCTTCAGA
ACTCTTTG	1360	1370 F	1380	1390	1400	1410	1420	1430
	GAGGACCAGA	ACGAGACAAT	GGTTCTTGCC	AGCTCTACCA	CCAGCATCCT	CTGTATGCTG	CTCCCGCTCC	IGATGCTCTTCC
ACCCCTTG	GAGGACCAGA	ACGAGACAAT		AGCTCTACCA	CCAGCATCCA	CACCATGCTG	CTCCTGCTCC	TGATGCTCTTCC
ACCAGGGA	1450	1460	1470	1480	1490	1500	1510	1520
	CTCCAGATTT	CAGACAGGGG	CTCAGATGCC	CACCATTTAC	TCAGGACGTT	GGATTGCAGG	ACTATTGCCT	TGGAGATTTTGG
ACCTGGGA	CTCCAAGCTT	CAATCAGTGG	CCGGGATACC	LEZO	1580	GAATTGCAGC	1600	AGGAGATTATAG
TGAAGCTC	CCAGTGAGTA	GCTGGCTGAG	GTTAGCCTGG	-GCAGGCTGG	CTTCAACAGG	TGCCTCGGAC	CAATAAGCCT	CATGATTCTTTC
	xon 1	1640	1650	1660	1670		1690	1700
TTTTAGTA	TCCTCAGG	TATCTGGACT	CAATAATAGT	GACGACAAAG	CCAATCTGAG	GGTAAGAGCC	CIGCICITIGG	GCATTCTTGGGT
1710	1720	exon 2->	1740	1750	-ex 1760	on 2 -	1780	1790
TCCATCTO	TCTCCTGCCT	GGGTGACTTT	AGCCATGTCA	CTGCACCCTG	CTITIGCTTCC	GTTTTCACAT	CTATCTCAGT	GGGGTTATTAAG
1800 1810 1820 1830 1840 1850 1860 1870 1880 GAAATCATCAGATGACTCTCTGAGCCTCAGTCTGTGCCACAGCCAGC								
1890	1900	ISON	10 19	20 19	GCAATAATGA	40 19	950 19	060 1970
AACIGIGA	GTGAACCCTC	CCACAG	GCCTCTGGCI	CCACTTCAG	TGGGGATGC	LATGGGGATGG LII CAT	LLATUGACUA	::::

1980		1990	2000	2010	2020	2030	2040	2050
CTICIGIGAC	TGIGIU			CCAAAACIGA		TCCACTICCA	ACCAGCCTAA	JACATTACCATT
CCTCTGTGAC	TTTTGTGTC	TITIGCTIT	TCTTCCTCCT	CCAAAACTG	AGTITGTGTTT	TCTACTTCC	CCAGCCCCA	GACATTACTATT
2060	2070	2080	2090	2100	2110			2120
TGCAGTTATI	TTCCCAGCT	CTAGTTAGA	TACAATGGTT	CIGITICGI	TTTATTGTT	G		CAAGCGTGC
				111 1 11		:		
TGTAGTTATT	FICC	-TAGTITGA	TACAATAGT	ATGTCTTGT	TTATTGTT	GGACCTAAC	ATGAAGTTCT	FIGCAAGAGIGC
2130	2140	2150	2160	2170	2180	2190	2200	2210
TGAGTGTCT	ACGTTTCCCC	TCCCTAGGG	ACATGATGA	GTCTGTAGG	TTTTCTTCTAC	ATATCTAGA	AGTTCTTAAT	TAAATTAAAGCA
TGAATGTCT	TATTTCCCC	TCCCTAGGA	ACATGTTGG	GCCTATAGG	TTTTCTTCTA	GTATCGAGA	ATTCTTAAT	TAAATTAAAGCA
TTGGGGTTG	2230 CGATTTACC	2240 TCAGTGGTA	2250 GAGTGCTTG	2260 "CTAGCAAGC	2270 Craageccetter	2280	2290 CAGCTCCGA	2300
		10110100111				011011100		
TTGG								
2310	2320	2330	2340	2350	2360	2370	2380	2390
AGAAAAAGAA	AAAAAATTA	AAGCATTAA	CCITGGTGT	TGGCATCTT	GGCATAAGTA	TTCCCTTGG	CCAACCTTCT	GCCTTTTCTAGA
			COTCOCTT		11 11: Coorden TN		CAACCUTCT	CONTRACTA
				IGOCATCII	J0IA		CONNECTICE	GCCTTTTCINGA
2400	2410	2420	2430	2440	2450	2460	2470	2480
GCTIGICIG	JAGAGATATG	TTTCCCTTA	AAAACAGAC	GATCIGCII	AGAGCCPICAC		GGCIGCCAGG	GGTTAAGACCIG
GCTTTTCTG	GAGGGATGTG	TTTTCCTCA	AGTACAGAC	AATCTGCTT	GATCCTTCAC	CAGCTCACA	GGCTGCCAGG	AGTTAAGACCTG
				0500	2540	2552	2560	2520
GTGCTCAGG	AGAAACAGGC	CCTTGTCTG	GGATGTGCC	Z550 TAGCTTT-A	GCCCCAGGATA	GGAAAGGAC	CAGGAGTAAG	GCTGTTCAAAGA
			111 1 1					
GTGCTTGGG	AGAAACAGGC	CCTIGICIG	AGATATACA	CTAGCTTTTA	GCCCCAGGATA	TG-AAGGGA	CAGGAATAAG	GCTGTTCAAAGA
2580	2590	2600	2610	2620	2630	2640	2650	2660
AACCTCTAA	CAGCAGTCAC	ACCTCCCCA	GCTCTCACC	ICCCCAGCTC:	TCACCTCCCCAC	CTCTCACCT	CTCCCGCTCT	CACCTCCCCAGC
			COTOTOATO					: C
ANICICAAN.	INGUNGGUNU	ACTICCCIA	OCICIONIC.	ICICCAGCIC.	CACIFICCAN			•
2670	2680	2690	2700	2710	2720	2730	2740	~ ~
TCTCACCTC	TCCAGCTCTC	ACTTCCCCA	GTTCTCACC	ICCCCAGCIC	FCACCTCTCCA	CICICACCI	CCCCAGCICI	
TCTCACCTC	CCCAGCTCTI	TCCTTTCCA	GTTCTCACC	ICCCIGCCIC	TCACCTCTCTG	ATCTCATCT	CCCTGGCTCT	CAA
		2990	2900	2910	2920	2930	2940	2950
(approx.	140 nucl.) TCCAGGO	CTGATCTCT	ACAACTCTCA	CCTCCCTGATT	TCACCTCCC	CATTTCTCAC	CTTCCCAACTCT
(uppion		,						
2960	2970	2980	2990	3000	3010	3020	3030	3040
TACCTCCCT	AGTICICACO	TCCTCTGAI	CTCACCITC	CTAGCTCTTA	CCCCCTCAGTT	TCACTCCCC	AAGCTCTCAC	CTCCCCAGTICT
							21.00	2120
3050	3060	3070	3080	3090 ന്നപ്പാലം	3100 CCTCCCCAGTT	3110 TCACCTCTC	AAACTCTTAT	TACCTCCCTGGC
111111	IIIIII I	: ::::::	11111111					
CTCCCA	AGCTCTTTC	ATTTCCAGT	CTCACCT					
3140	3150	3160	3170	3180	3190	3200	3210	3220
TTTCCAGTT	CTTACCACC	CAGCTCTC	ACATCCCTAG	ATTTTACACA	AAGATTCTTTG	CCCTCTTAGA	ATAGTACCTI	GCGGAGAGTAAA
								TCCCACACTAAA
	:	00000000	CATCOTO	COMPACATA	A & C & TYTY TYTY	CCCTCTTAGA	ATAAGAGCTT	
	C	CAGCTCTC	ACATCCCTGG	GCCTTACATA	AAGATTTCTTG	CCCTCTTAGA	ATAAGAGCTI •exon 3-⇒	
	C	CAGCTCTC	ACATCCCTGG	GCCTTACATA	AAGATTTCTTG	CCCTCTTAGA	ATAAGAGCTI •exon 3-►	> 2210
3230	3240	3250	ACATCCCTGG 3260 3GAGTTTGAT	GCCTTACATA 3270 TCTCAGGACA	AAGATTTCTTG 3280 CAACGGACATC	CCCTCTTAGA	ATAAGAGCTI • exon 3->	3310
3230 CCTGGACGA	3240 ATTCCTAAA	3250 AAGCCAAGA	3260 GAGTTTGAT	GCCTTACATA 3270 TCTCAGGACA	AAGATTTCTTG 3280 CAACGGACATC	CCCTCTTAGA 3290 AAGTCCAAAC		3310
3230 CCTGGACGA 1111 1 1 CCTGTCCAA	3240 ATTCCTAAAA	3250 AAGCCAAGA	3260 3GAGTTTGAT 11 1 111 AGAAGTGGAT	GCCTTACATA 3270 TCTCAGGACA 11 11111 CCTGAGGACA	3280 CAACGGACATC : : ::: GATACGTTATC	3290 AAGTCCAAAC	ATAAGAGCTI • exon 3-> TTCAGGTGCC	3310 TGGGGCTGCTGA
3230 CCTGGACGA 1111 1 1 CCTGTCCAA	3240 ATTCCTAAA IIIIIIIIIIIIIIIIIIIIIIIIIIIII	3250 AAGCCAAGA AAGCCAAGA	3260 3260 3GAGTTTGAT 11 1 111 AGAAGTGGAT	GCCTTACATA 3270 TCTCAGGACA 11 11111 CCTGAGGACA	3280 CAACGGACATC : : ::: GATACGTTATC	3290 AAGTCCAAAC	ATAAGAGCTI exon 3-	3310 TGGGGCTGCTGA
3230 CCTGGACGA :::: : : CCTGTCCAA 3320	3240 ATTCCTAAA IIII I I ATTCGTGGA 3330	3250 3250 AAGCCAAGA AAGCCAAGG	3260 3260 3GAGTTTGAT 11 1 11 AGAAGTGGAT 33	GCCTTACATA 3270 TCTCAGGACA CCTGAGGACA 40 33	3280 CAACGGACATC : : ::: GATACGTTATC 50 336	3290 AAGTCCAAAC AAGTCCAAAC AGTCCAATC AGTCCAATC CO 337	ATAAGACTI • exon 3- TTCAGTGCC TTCAGGTGCC on 3 0 33E	3310 TEGGGCTGCTGA TEGGAGCTGCTGA
3230 CCTGGACGA IIII III CCTGTCCAA 3320 GGGCTTAGG	3240 ATTCCTAAAA ATTCCTAAAA ATTCGTGGA 3330 GATCCCAC	3250 AAGCCAAGA AAGCCAAGA	3260 3260 36AGTTTGAT 11 1 11 AGAAGTGGAT 33 ACC	GCCTTACATA 3270 TCTCAGGACA 11 11111 CCTGAGGACA 40 33 TTTCCCAGGA	AAGATTTCTTG 3280 CAACGGACATC : : ::: GATACGTTATC 50 336 TTTTTCGAGAG	3290 AAGTCCAAAC AAGTCCAAATC AAGTCCAAATC AAGTCCAATC AAGTCCAATC 337 GAAAGTAGAG	ATAAGACCTI • exon 3- TTCAGTGCC TTCAGGTGTC on 3 0 33E CTTTCTCCTI	3310 TGGGGCTGCTGCAGA TGGGGCTGCTGCAGA 10 3390 TCTA-GAGGGAT
3230 CCTGGACGA IIII II CCTGTCCAA 3320 GGGCTTAGG	3240 ATTCCTAAA IIIIIIIIII ATTCGTGGA 3330 GATCCCAC	3250 AAGCCAAGA AAGCCAAGA AAGCCAAGG CTCC CTCC SGGGCTCCC	3260 3260 3360 3360 337 33 33 33 33 33 33 33 33 33 33 33 33	GCCTTACATA 3270 TCTCAGGACA IIIIII CCTGAGGACA 40 33 ITTCCCAGCC IIIIIIIII	ANGATTTCTTG 3280 CAACGGACATC : : :: GATACGTTATC 50 336 TTTTTGGAGAG :::::::::::	3290 AAGTCCAAAC AAGTCCAAAC CAAGTCCAAAC CAAGTACAAG GAAAGTACAGG	ATAAGAGCTI • exon 3- TTCAGGTGCO TTCAGGTGCO 0 33E CTTTCTCCTI CCTTTCTCCTI	3310 TGGGGCTGCTGA TGGGGCTGCTGA TGGAGCTGCAGA 3390 TCTA-GAGGGAT
3230 CCTGGACGA 1111 1 CCTGTCCAA 3320 GGGCTTAGG 111 1 1 GGGTTGGG	3240 ANTTCCTARAN ANTTCCTARAN ANTTCGTGGAN 3330 GATCCCAC GTGGGGGGGGTC	3250 AAGCCAAGA AAGCCAAGA AAGCCAAGG CTCC GGGCTCCG	3260 3260 3360 33777037 34 34 34 34 35 35 35 35 35 35 35 35 35 35 35 35 35	GCCTTACATA 3270 TCTCAGGACA 11 11111 CCTGAGGACA 40 33 ITTCCCAGCC 11111111 ITTCCCAGCC	ANGATTTCTTG 3280 CAACGGACATC GATACGTTATC 50 336 TTTTTCGAGAG	3290 AAGTCCAAAC AAGTCCAAAC AAGTCCAATC AAGTCCAATC AAGTCGAATGAGG 337 GAAAGTAGGG 300 337 GAAAGTAGGG	ATAAGAGCTT • exon 3- TTCAGTGCC TTCAGGTGCC 0 33E CTTTCTCCTT CCTTCTCCCAC	3310 TGGGGCTGCTGA TGGGGCTGCTGA TGGAGCTGCAGA 3390 TCTA-GAGGGAT TTTAGGAGGGAT
3230 CCTGGACGA IIII III CCTGTCCAA 3320 GGGCTTAGG IIII III GGGTTGGGG 340 CTGACTET	3240 ANTTCCTANAN IIIIIIIIII ANTTCGTGGAN 3330 GATCCCAC IGTGGGGGGGTC 0 341 CCTTUGTGTGT	3250 AAGCCAAGM AAGCCAAGM AAGCCAAGG CTCC GGGGCTCCGT 10 TTCTCCACAC	3260 3260 3360 33777647 33 34 34 34 34 35 35 36 36 36 37 37 37 37 37 37 37 37 37 37 37 37 37	3270 3270 TCTCAGGACA 11 11111 CCTGAGGACA 40 33 TTTCCCAGCC 1111111 TTTCCCAGCC 430 3	AAGATTTCTTG 3280 CAACGGACATC 1 1 11 GATACGTTATC 50 336 TTTTTGGAGAG 1111111111 TTTTTGGAGAG 440 34	3290 AAGTCCAAAC AAGTCCAAAC AGGCAACCAATC AAGTCCAATC AAGTCAAG 337 GAAAGTAGAG 50 AGGGACTCTG	ATARAGETTI exon 3-> TTCAGTECC TTCAGTECT 0 338 CTTTCTCCT CTTCTCCCT	3310 TTGGGGCTGCTGCA TTGGGGCTGCAGA 10 3390 TCTA-GAGGGAT 11 1111 TTTTAGGAGGAT 170 3480 14CCTCCTTCA
3230 CCTGGACGA 1111 1 1 CCTGTCCAA 3320 GGGCTTAGG 111 1 1 GGGTTGGGG 340 CTGACTGTG CTGACTGTG	3240 ATTCCTAAN 3330 3330 3330 GTCCCAC	3250 AAGCCAAGA AAGCCAAGA AAGCCAAGG CTCC iiii SGGGCTCCGI L0 CTCTCCACAC	ACATCCCTGG 3260 GGGTTTGAT 11 111 AGAAGTGGAT 33 ACC 111 TGCCTTTACC 5 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3 3	3270 3270 TCTCAGGACA 111111 CCTGAGGACA 40 33 TTTCCCAGCC 111111 TTTCCCAGCA 430 3 TGTTGTATTC 11111	AAGATTTCTTG 3280 CAACGGACATC i iii GATACGTTATC 50 336 TTTTTGGAGAG 440 34 CTGCAGCTCCG 440 34	2290 аластссалас аластссалас аластссалас аластссалас аластссалас аластасалас аластасаас баластасас 50 аластасас	ATARAGETTI • exon 3-> TTCAGTECC TTCAGETECC 0 33E CTTCCCCT CTTCCCCA 34 TCGTAAGCTO	3310 TTCGGACCTGCTGCA TTCGGACCTGCAGA 10 3390 TCTA-GACGGAT TTTTAGGACGGAT TTTTAGGACGGAT 10 3480 MACCTCCCTTCCA
3230 CCTUGACGA 3320 GGGCTTAGG 111 : 11 GGGTTGGGG 340 CTGACTOTG CTGACTATG	3240 ATTCCTAAAI ATTCCTAAAI ATTCCTAAAI ATTCCTAGA GATCCCAC : GTCCCCCCC : GTCCCCCCT 0 341 CCTTGGTGT	3250 AAGCCAAGA AAGCCAAGA AAGCCAAGG CTCC SGGGCTCCG7 10 CTCTCCACAC	3260 3260 3360TTGAT 44 46AAGTGGAT 33 ACC 111 10CCCTTACC 7 3AAACTTAAG	GCCTTACATA 3270 TCTCAGGACA 11 11111 CCTGAGGACA 40 33 TTTCCCAGGA 11111111 TTTCCCAGGA 430 3 TGTTGTATATA 1111111	AAGATTTCTTG 3280 CAACGGACATC : : ::: GATACGTTATC 50 336 TTTTTTGGAGAG 440 34 CTGCAGCTCGG CTACATCTGCG	3290 AAGTCCAAAC AAGTCCAAAC → AAGTCCAATC → 6X(0) 337 GAAAGTAGAG 50 AGGGACTCTC 50 AGGGACTCTCC	ATARCACCTI • exon 3- TTCACGTCCC TTCACGTCCC TTCACGTCCCAC 0 338 CTTTCACCTCCCAC 34 TCGTAAGCTC CCGTAAGCTCC	3310 TGGGGCTGCTGCA TGGGGCTGCAGA 10 3390 TCTA-GAGGGAT TTTAGGAGGGAT 10 3480 ACCTCCTTCCCA CCCTTCCC
3230 CCTGGACGA 1111 1 CCTGTCCAA 3320 GGCCTTAGG 340 CTGACTGTG 11111 1 CTGACTGTG	3240 ATTCCTAAAI ATTCCTAAAI ATTCGTGGA GATCCCAC: GGTGGGGGGGTC 0 341 CCTTGGTGTC	3250 AGCCAAGA AAGCCAAGA AAGCCAAGG CTCC 3GGGCTCCG 10 CTCTCCACAC	3260 3260 3360TTCAT 44 46 46 47 47 47 47 47 47 47 47 47 47 47 47 47	GCCTTACATA 3270 TCTCAGGACA 11 11111 CCTGAGGACA 40 33 TTTCCCAGCC 11111111 TTTCCCAGCA 430 3 TGTTGTATTCC ↓ 1111111 TGTTGCTGC ►	AAGATTTCTTG 3280 CAACGGACATC 1 111 GATACGTATC 50 336 1111111111 111110GAGAG 440 34 CTWCAGCTGCG 111111111	3290 AAGTCCAAAG AAGTCCAAAG AAGTCCAAAG AAGTCCAAAG GAAGTCCAAAG GAAGTCCAAAG GAAGTCCAAAG GAAGTCCAAAG GAAGTCCAAAG GAAGTCCAAAG GAAGTCCAAGG SGAAGTAGAG HATGACTCCG GACCCACTCCG GACCACTCCG		3310 TTGGAGCTGCTGA TTGGAGCTGCAGA TTTTAGAGCGAT TTTTAGGAGGAT TTTTAGGAGGAT TTTTAGGAGGAT TTTTAGGAGGAT TTTTAGGACGAT
3230 CCTGGACCA CCTGGTCCAA 3320 GGGCTTAGGG 111 1 1 GGGTTGGGG 340 CTGACTATG	3240 ATTCCTRAAA ATTCGTGGA GATCCCAC GTGGGGGGGT 0 3430 CCTTGGTGTC CCTTGGTGTC	3250 AAGCCAAGA AAGCCAAGA CTCC :::: 3006CTCCG 10 TCCTCCACACA CTCTCCACACA 3500	ACATCCCTGG 3260 CGAGTTTCAT AGAAGTGGAT 33 ACC ACC 	GCCTTACATA 3270 TCTCAGGACA 1 11111 CCTGAGGACA 40 33 TTTCCCAGCA 40 33 TTTCCCAGCA 430 3 TTTCCCAGCA 430 3 CGTTGTACTCC TGTTGCCTGCC → 3520 COMBINING	3280 CAACGGACATC i iii GATACGGACATC i iii GATACGTATC 50 336 TTTTTTGGAGAG 400 34 CTGCAGCTCGG iiii iiii CTGCAGCTCGG 23530 CTGCAGCTCGG	3290 AAGTCCAAAC AAGTCCAAAC AAGTCCAAAC 30 337 GAAAGTAGAG 50 AAGGGACTCTGG AATGACTCTGG 4 0X00 50 AAGGGACTCTGG 50 AAGGGACTCTGG 6 0X00 50 6 0X00 50 7 0X00 50 7 0X00 7	ATARCAGCTI • exon 3 → TTCAGTCCC · · · · · · · · · · · · · · · · · · ·	3310 3310 TrGGGGCTGCTGA TrGGGGCTGCAGA 3390 TCTA-GGAGCTGCAGA TCTA-GGAGGAT TCTA-GGAGT TCTA-GGAGGAT TCTA-GGAGT TCTA
3230 CCTGGACCA 1111 1 CCTGTCCAA 3320 GGGCTTAGGG 340 CTGACTGTGG 1111 1 CTGACTATG TGTGTCC	3240 MTTCCTRAAN MTTCCTRAAN 1111 1 1 MTTCGTGGA 3330 GATCCCAC CTTGGGGGGTC 0 341 CCTTGGTGTC CCTTGGTGTC CCTTGGTGTC 3490 3490 1111111	3250 AAGCCAAGA AAGCCAAGA AAGCCAAGA AAGCCAAGA CTCC J J J J J J TCTCCACACA 3500 CTCCTCGACACA 	асатссстоя 3260 3260 336 33 33 33 33 34 50 50 50 50 50 50 50 50 50 50	GCCTTACATA 3270 TCTCAGGACA 1 11111 CCTGAGGACA 40 33 TTTCCCAGGCA 430 3 TTTCCCAGCA 430 3 TTTCCCAGCA 430 3 TTTCCCAGCA 50 50 50 50 50 50 50 50 50 50	ANGATTTCTTG 3280 CAACGGACATC i iii GATACGTATC 50 336 TTTTTGGAGAG 440 34 440 34 CTOCAGCTCGG iiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiiii	3290 AAGTCCAAAC AAGTCCAAAC AAGTCCAAAC AGTCCAAAT AAGTCCAAAT AAGTCCAAAT AGAAAGTAGGG 50 AGCGACTCTC AGC	ATAGAGCTI exon 3- TTCAGTGCC TTCAGTGCC TTCAGTGCC 0 33E CTTTCTCCT CCTTCTCCCAC TCGGTAAGCTC 4 3550 TCTTCACATC	3310 3310 TTGGAGCTGCTGA TTGGAGCTGCAGA 10 3390 TTTA-GAGGGAT 111 111 TTTTAGAGGGAT 111 111 TTTTAGGAGGAT 111 111 TTTAGAGGGAT 111 111 TTTAGAGGGAT 111 111 TTTAGAGGGAT 111 111 TTTAGAGGGAT 111 111 TTTAGAGGGAT 111 111 TTTAGAGGGAT 111 111 TTTAGAGGGAT 111 111 TTTAGAGCTGCAGA 111 111 111 111 TTTAGAGCTGCAGA 111 111 111 111 111 111 TTTAGAGCTGCAGA 111 111 111

Figure 3 Nucleotide sequence of the rat (R) IL-3 gene covering the 3.3 Kb BamHI and 1.25 Kb BamHI/<u>Hin</u>dIII fragments, shown aligned against the sequence of the murine (M) IL-3 gene. TATA box and poly-adenylation signals are underlined. Intron/exon boundaries are shown as deduced for the murine gene⁸,9.

control signals (including the potential enhancer hexanucleotide CCGCCC) described by Campbell <u>et al</u>⁸ and Miyatake <u>et al</u>⁹ are identical, although the proposed CAAT sequence would be a further 7 nucleotides upstream of the TATA box in the rat gene than is the case in murine IL-3.

Nucleotide sequence homology of the 3' end of the coding region is once again high, and there is an in-frame translational termination codon at position 3733. A single AATAAA sequence is found in the 3' untranslated region, 25 nucleotides nearer the

3650

TAA stop codon than in murine IL-3. This sequence is likely to be involved in polyadenylation of the $mRNA^{41,42}$.

Other conserved features include two sets of tandemly repeated sequences. The first, a 14-15 bp element repeated about 12 times in intron 2 of murine IL-3, is also found in intron 2 of rat IL-3, except that in the latter case the repeat appears to extend for a further 400-450 nucleotides ie 25-30 more repeating units than in the murine gene. The complete sequence through these repeating units was not determined due to the difficulty in The amount of "missing" obtaining clones through this region. sequence was deduced by visual estimation from sequencing ge1 autoradiograms together with careful estimation of restriction fragment lengths. The consensus sequence of these repeats in the rat gene. AGCTCTCACCTCCC, is identical to the consensus sequence of the murine IL-3 intron 2 repeats⁸. The 9-bp inverted repeat which flanks these tandem repeats in the mouse IL-3 gene is not conserved in the rat gene. The second repetitive element occurs between the termination codon and the polyadenylation signal (nucleotides 3872-3922), and consists of nine copies of the sequence TATT⁸.

One further feature of interest is the presence of sequences the 5' untranslated region which show homology to murine Bl, in B2 and new Alu type repeating elements (data not shown). These elements represent major classes of short repeats dispersed throughout the mouse genome $4^{3,44}$, and may each be present in up 50,000 copies per haploid genome. Sequences homologous to to rat poly (RY) repeats 45 are also found in this 5' region. The function of these sequences is unknown; there are suggestions ranging from a possible role in regulation of transcription/mRNA transport/translation $^{46-48}$ to the notion of "junk DNA" or pseudogenes⁴⁹.

Amino Acid Sequence of Rat IL-3

The deduced amino acid sequence of rat IL-3 is shown in Fig. 4 compared to murine IL-3. Overall homology is 59% which reduces to 54% in the "mature" protein ie. after the removal of the putative leader sequence 50,51 , which is the most highly conserved part of the protein. Rat IL-3 has two potential N-glycosylation sites : the second of these corresponds to one of

			homology
exon 1	(R) (M)	MVLASSTTSILCHLLPLINLFHQGLQISDRGSDAHLLPTLDCRTIALEILVKLP AAAAAAAAA MVLASSTTSIHTMLLLLMLFHLGLQASISGRDTHRLTRTLMCSSIVKEIIGKLP 1	67%
	(R)	VSGLINNSDDKANLR	
exon 2	(M)	EPELKTDOBGPSLR	29%
AY 00 9	(R)	NSTLRRVNLDEFLKSQEEFDSQDTTDIKSKLQ	
8701 3	(M)	MKSFRROMLSKFVESQGEVDPEDRYVIKSNLQ	53%
	(R)	KLKCCIPAAASDSV	
exon 4	(M)	AA AA AA AAA KLNCCLPTSANDSA	57%
exon 5	(R)	LPGVVNKDLDDEKKKLRFVVIHLKDLQPVSVSRPPQPTSSSDNFRPMTVEC	
0101 0	(M)	LPGVFIRDLDDFRKKLRFYMVHLNDLETVLTSRPPQPASGSVSPNRGTVEC	63%

overall homology 59%

Figure 4 Comparison of predicted coding sequences of the exons of rat (R) and murine (M) IL-3 genes. Potential N-glycosylation sites are underlined. The arrow indicates the most likely site of cleavage of the signal peptide 50,51.

the 4 potential sites in murine IL-3. All four cysteine residues are conserved.

Expression of Rat IL-3 in Monkey COS-1 Cells

The 5.8kb HindIII fragment carrying the entire rat IL-3 gene (including transcriptional/translational control signals) was cloned into the <u>Hin</u>dIII site of the expression vector $pSV2-neo^{52}$. Two clones were isolated, one for each possible orientation of the IL-3 gene relative to the SV40 promoter: pILR1 contained the IL-3 insert in the same orientation as the SV40 promoter and pILR2 had the insert in the reverse direction. Purified plasmid DNA from both constructions was used to transfect monkey COS-1cells. Supernatants from the transfected cells were collected at 72 and 144 hours, based on previous experience with this expression system for murine IL-3. Murine IL-3 was prepared at the same time, using pSV2-neo vectors containing either genomic (pILM13) or cDNA (pILM4) murine IL-3 inserts. Supernatants from untransfected COS-1 cells were used as controls.

Colony Stimulating Activity of Rat IL-3

Recombinant rat IL-3 from both pILR1 and pILR2 stimulated

Growth Factor	Source	Bone Marrow Proliferation(a			
		Mouse	Rat		
Mouse I1-3	WEHI-3 CM	1025	<10		
Recombinant mouse IL-3	COS-1	. 250 360 90	<10 <10 <10		
Recombinant rat IL-3	COS-1	<10 10 50	725 510 630		
Mouse GM-CSF	purified ^(b)	3100	210		

Table 1: Bone Marrow Proliferation Activity of Rat and Murine IL-3

(a) Proliferation was measured as described in Methods using bone marrow cells from either Balb/c mice or WISTAR rats. Activity is expressed as the reciprocal of the titration endpoint. COS-1 conditioned medium showed no detectable activity.
(b) Purified murine GM-CSF was obtained from Genzyme (Boston, Mass.) Lot No. 01531 and was verified to be GM-CSF by its ability to stimulate growth of FDC-P1 but not 32Dc1-23 cells.

colony formation only with rat bone marrow cells, while expressed murine IL-3 stimulated colony formation only with mouse cells. Bacterially expressed murine IL-3 was also inactive on the rat bone marrow cells.

In these experiments, rat IL-3 produced the full range of colony types expected for this factor on rat bone marrow cells, although there was a much greater proportion of megakaryocytes eosinophils amongst the colonies generated than in the and corresponding assays of murine IL-3 on murine cells. (A similar elevated response of megakaryocytes and eosinophils has been observed when assaying bacterially expressed murine IL-3). The biological properties of recombinant rat IL-3 will be reported in detail elsewhere (A.J.Hapel, D.R.Cohen and I.G.Young, In preparation).

Proliferation of Rat and Mouse Bone Marrow Cells in IL-3

Bone marrow cells are generally as sensitive as FDC-P1 cells for detecting IL-3 and GM-CSF activities in the mouse when used as at 10^5 cells per well in the standard microtitre cell proliferation assay. As there are no established cell line assays for rat IL-3 we used bone marrow cells from a variety of rats and BALB/c mice to measure the growth factor activity of expressed rat and mouse IL-3.

The data in Table 1 show that rat IL-3 had good activity in promoting the proliferation of WISTAR rat bone marrow cells and slight activity on murine cells, while expressed murine IL-3 was active only on cells from murine bone marrow. This is in contrast to murine GM-CSF which has significant activity in the rat system. A similar result has been obtained using bone marrow cells from other strains of rats, except that PVG, PVG x JC and PVG x DA rats gave end-points about 4-8 fold greater than WISTAR, Fischer and JC rats. More extensive characterization of the biological relationship between rat and murine IL-3 will be described elsewhere (A.J.Hapel, D.R.Cohen and I.G.Young, In preparation).

DISCUSSION

Previous work on the biology and molecular biology of IL-3 has been predominantly with the mouse system. It is clear from the work described above that the rat genome carries an IL-3 gene which is closely related in overall structure to that found in the mouse genome. The cloning and sequence analysis of the rat IL-3 gene has enabled a detailed comparison to be made between The flanking regions (200 nucleotides the rat and mouse genes. side of the gene) and introns of the mouse and rat IL-3 either genes show 90% and 80% nucleotide homology, respectively, with the intron/exon junctions fully conserved. The portion of exon encoding the putative signal peptide is also highly conserved 1 Surprisingly, the nucleotide homology between the coding (90%). This is reflected in the amino regions is only 76%. acid homology which is only 54% for mature rat and mouse IL-3 (59% including the signal peptide). This is atypical of other lymphokines characterized to date. For example, γ -IFN shows an amino acid homology of 87% between mouse and rat 53 , while IL-2 and GMmouse : human amino acid homologies of 60% and 54% CSF show respectively^{15,16,54}. The very high conservation of the signal peptide region observed between rat and mouse IL-3 is also not typical of the other lymphokines 15, 16, 53, 54.

The conservation of the two sets of tandomly repeated units The TATT sequence, repeated nine times is also of interest. between the translational termination codon and the poly-adenylation signal in both murine and rat IL-3, is also found in a similar position in all IL-2, GM-CSF and Y-IFN genes sequenced to date^{15-17,53-58} The consensus sequence of the other repeat, a 14-15 bp element in intron 2, is identical to that of the tandem repeats in intron 2 of mouse IL-3. This element shows homology to a human BKV enhancer sequence homolog 8,59 , and also to a portion of the consensus sequence for a 33-bp human mvoglobin gene tandem repeat^{8,60}. Similar repeats are found in the human insulin gene⁶¹ and globin gene complex⁶² and are associated with "hypervariable regions" of DNA.

Expression experiments using monkey COS-1 cells have demonstrated that the rat IL-3 gene encodes a multi-lineage haemopoietic growth regulator, which appears to have an analogous biological role to mouse IL-3⁶³. The low amino acid homology between rat and mouse IL-3 correlates with the demonstration that rat and mouse IL-3 show little cross-reactivity. Other lymphokines and CSF's are somewhat varied in this regard. For example, human IL-2 is about 5 times more active on mouse cells than is murine IL-2 on human cells 15 , whereas human and murine GM-CSF's do not cross-react either way¹⁶. Rat Y-IFN is active on mouse cells but not on human cells 53 , and human M-CSF can stimulate growth of murine macrophages⁶⁴. The low sequence homology and poor cross-reactivity suggest that rat IL-3, together with its receptor, have evolved significantly away from the murine IL-3/receptor system.

The availability of the rat IL-3 gene together with recombinant rat IL-3 should assist in studying the biological role of IL-3 in rat experimental models such as adjuvant-induced arthritis⁶⁵, graft rejection⁶⁶ and megakaryocyte proliferation^{67,68}. The rat gene sequence also provides useful information for continuing studies on the regulatory elements controlling IL-3 expression.

ACKNOWLEDGEMENTS

We would like to thank G. Mayo for synthesis of oligonucleotides, R.M. Johnson for some DNA preparations, A. George for preparation of packaging extracts and K. Ridgway, W. Tucker and P. Townsend for technical assistance. We are also grateful H. Campbell for helpful discussions and to G. Quinn to for programming assistance.

REFERENCES

- Burgess, A.W., Camakaris, J. and Metcalf, D. (1977) J. Biol. 1. Chem. 252, 1998-2003.
- Howard, M., Burgess, A.W., McPhee, D. and Metcalf, D. (1980) 2. Cell <u>18</u>, 993-999. Burgess, A.W., Metcalf, D., Russell, S.H.M. and Nicola, N.A.
- 3. (1980) Biochem. J. <u>185</u>, 301-314. Ihle, J.N., Lee, J.C. and Rebar, L. (1981) J. Immun. <u>127</u>,
- 4. 2565-2570.
- Ihle, J.N., Rebar, L., Keller, J., Lee, J.C. and Hapel, A.J. (1982) Immun. Rev. <u>63</u>, 5-32. Fung, M.C., Hapel, A.J., Ymer, S., Cohen, D.R., Johnson, R.M., Campbell, H.D. and Young, I.G. (1984) Nature 5.
- 6. 307, 233-237.
- Yokota, T., Lee, F., Rennick, D., Hall, C., Arai, N. 7. Mosmann, T., Nabel, G., Cantor, H. and Arai, K-I. (1984)
- Proc. Natl. Acad. Sci. USA <u>81</u>, 1070-1074. Campbell, H.D., Ymer, S., Fung, M.C. and Young, I.G. (1985) Eur. J. Biochem. <u>150</u>, 297-304. Miyatake, S., Yokota, T., Lee, F. and Arai, K-I. (1985) 8.
- 9. Proc. Natl. Acad. Sci. USA, <u>82</u>, 316-320.
- Haig, D.M., McKee, T.A., Jarrett, E.E.E., Woodbury, R. and 10. Miller, H.R.P. (1982) Nature 300, 188-190.
- Haig, D.M., McMenamin, C., Gunneberg, C., Woodbury, R. and 11. Jarrett, E.E.E. (1983) Proc. Natl. Acad. Sci. USA 80, 4499-4503.
- Abbud-Filho, M., Kupiec-Weglinski, J.W., Araujo, J.L., Heidecke, C.D., Tilney, N.L. and Strom, T.B. (1984) 12. J. Immunol. <u>133</u>, 2582-2586.
- 13. McMenamin, C., Jarrett, E.E.E. and Sanderson, A. (1985) Immunol. <u>55</u>, 399-403.
- Welte, K., Platzer, E., Lu, L., Gabrilove, J.L., Levi, E., Mertelsmann, R. and Moore, M.A.S. (1985) Proc. Natl. Acad. 14. Sci. USA <u>82</u>, 1526-1530.
- Yokota, T., Arai, N., Lee, F., Rennick, D., Mosmann, T. and Arai, K-I. (1985) Proc. Natl. Acad. Sci. USA <u>82</u>, 68-72. 15.
- Cantrell, M.A., Anderson, D., Cerretti, D.P., Price, V. McKereghan, K., Tushinski, R.J., Mochizui, D.Y., Larsen, A., Grabstein, K., Gillis, S. and Cosman, D. (1985) Proc. Natl. 16. Acad. Sci. USA 82, 6250-6254.
- Gray, P.W. and Goeddel, D.V. (1983) Proc. Natl. Acad. Sci. 17. USA 80, 5842-5846.
- Gross-Bellard, M., Oudet, P. and Chambon, P. (1973) Eur. J. 18. Biochem. <u>36</u>, 32-38.
- Blin, N. and Stafford, D.W. (1976) Nucl. Acids Res. 3, 19. 2303-2308.
- Yamamoto, K.R., Alberts, B.M., Benzinger, R., Lawhorne, L. and Treiber, G. (1970) Virol. <u>40</u>, 734-744. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) 20.
- 21.

	<u>Molecular Cloning: a laboratory manual</u> , Cold Spring Harbor
	Laboratory, Cold Spring Harbor, New York.
22.	Clewell, D.B. and Helinski, D.R. (1970) Biochem. 9, 4428-
~ ~	4440.
23.	Hanahan, D. (1983) J. Mol. Biol. <u>166</u> , 557-580.
24.	Scalenghe, F., Turco, E., Edstrom, P., Pirrotta, V. and Melli, M. (1981) Chromosoma 82, 205-216.
25.	Zissler, J., Singer, E. and Schaefer, F. (1971) in The
23.	bacteriophage lambda, ed. A.D. Hershey p.455. Cold Spring
	Harbor Laboratory, Cold Spring Harbor, New York.
26.	Benton, W.D. and Davis, R.W. (1977) Science 196, 180-182.
27.	Meinkoth, J. and Wahl, G. (1984) Anal. Biochem. <u>138</u> , 267-
20	Zor, E. Nichlas, C. and Coulers, A.D. (1077) Dece. Natl
20.	Sanger, F., Nicklen, S. and Coulson, A.K. (1977) Proc. Nati.
	Acad. Sci. USA <u>74</u> , 5463-5467.
29.	Duckworth, M.L., Gait, M.J., Goelet, P., Hong, G.F.,
	Singh, M. and Titmas, R.C. (1981) Nucl. Acids Res. 9,
	1681–1706.
30.	Deininger, P.L. (1983) Anal. Biochem. 129, 216-223.
31	Massing I (1983) Matheda in Engunal $101 20-78$
33	$ \begin{array}{c} \text{Appendix} \\ \text{Constraint} \\ Constr$
52.	450-457.
33.	Sanger, F., Coulson, A.R., Hong, G.F., Hill, D.F. and
	Peterson, G.F. (1982) J. Mol. Biol. 162, 729-773.
34.	Staden, R. (1980) Nucl. Acids Res. $8.3673-3694$.
35	Staden R (1982) Nucl Acids Res $\overline{10}$ (4731-4751)
36	Staden = (102/) Nucl. Acids Res. 12, 400 502
27	Staten, K. (1964) Nucl. Actus Kes. 12, 499-505.
37.	USA. 80, 726-730.
38.	Parker, B.A. and Stark, G.R. (1979) J. Virol. 31, 360-369.
39.	Hapel, A.J., Fung M.C., Johnson, R.M., Young, T.G.,
	Johnson G and Matcalf D (1985) Blood 65 1453-1459
10	Lafester VI Deves CI Al Ade A. Harsey
40.	Latterty, K.J., Flowse, S.J., Al-Adra, A., Walten, H.,
	Vasalli, J. and Keich, E. (1980) Aust. J. Exp. Biol. Med.
	Sci. <u>58</u> ,533-544.
41.	Montell, C., Fischer, E.F., Caruthers, M.H. and Berk, A.J.
	(1983) Nature 305 600-605.
42.	Nevins, J.R. (1983) Annu. Rev. Biochem. 52, 441-466.
43.	Kramerov, D.A., Grigorian, A.A., Ryskov, A.P., and
	Georgiev, G.P. (1979) Nucl. Acids Res. 6 697-713
4.4	Kalb V F Classon S King D and Lineral I B (1983)
44.	Nucl. Aside Des. 11, 2127, 210/
	Nucl. Acids Res. 11, $21/7 - 2184$.
43.	Alonso, A., Kuenn, B. and Fischer, J. (1983) Gene 20 ,
	303-306.
46.	Thimmappaya, B., Weinberger, C., Schnieder, R.J. and
	Schenk, T. (1982) Cell 31, 543-551.
47.	Walter, P. and Blobel, \overline{G} , (1982) Nature 299, 691-698.
48	Carlson, D.P. and Ross, J. (1983) Cell 34, 857-864.
40.	Kramerov D A Tillih S V Prekov A P and
42.	$\begin{array}{c} \text{All and } 0 \text{ , } p \text{ , } n \text{ , } n \text{ , } 111110 \text{ , } 5 \text{ , } n \text{ , } y \text{ , } n ,$
50	Georgiev, G.r. (1903) MUCI. ACIDS Kes. 13, 0423-043/.
50.	Conton, P.J., Luk, K.H., Park, L.S., March, C.J., Hopp, T.P.
	and Urdal, D.L. (1985) J. Immunol. <u>135</u> , 328-332.
51.	Clark-Lewis, I., Kent, S.B.H. and Schrader, J.W. (1984)
	J. Biol Chem. 259, 7488-7494.
52.	Southern, P.J. and Berg, P. (1982) J. Mol. Appl. Genet.
-	1. 327-341.
53.	Dijkema, R., Van der Meide, P.H., Pouwels, P.H.,

Caspers, M., Dubbeld, M. and Schellekens, H. (1985) EMBO. J. <u>4</u>, 761-767. Kashima, N., Nishi-Takoaka, C., Fujita, T., Taki, S., Yamada, G., Hamuro, J. and Taniguchi, T. (1985) Nature <u>313</u>, 54. 402-404. Gray, P.W. and Goeddel, D.V. (1982) Nature, 298, 859-863. Gough, N.M., Gough, J., Metcalf, D., Kelso, A., Grail, D., 55. 56. Nicola, N.A., Burgess, A.W. and Dunn, A.R. (1984) Nature 309, 763-767. 57. Taniguchi, T., Matsui, H. Fujita, T., Takoaka, C., Kashima, N., Yoshimoto, R. and Hamuro, J. (1983) Nature 302, 305-310. Fujita, T., Takoaka, C., Matsui, H. and Taniguchi, T. (1983) Proc. Natl. Acad. Sci. USA <u>80</u>, 7437-7441. 58. Rosenthal, N., Kress, M., Gruss, P. and Khoury, G. (1983) 59. Science 222, 749-755. Jeffreys, A.J., Wilson, V. and Thein, S.L. (1985) Nature 60. 314, 67-73. 61. Bell, G.I., Selby, M.J. and Rutter, W.J. (1982) Nature 295, 31-35. Goodbourn, S.E.V., Higgs, D.R., Clegg, J.B. and 62. Weatherall, D.J. (1983) Proc. Natl. Acad. Sci. USA 80, 5022-5026. Iscove, N. and Roitsch, C. (1985) in <u>Cellular and Molecular</u> <u>Biology of Lymphokines</u>, ed. C. Sorg and A. Schimpl p.397-63. 424. Academic Press, New York. Stanley, E.R.. Chen, D-M. and Lin, H.S. (1978) Nature <u>274</u>, 64. 168-170. Crowe, W.E., Battisto, J.R. and Smith, R.N. (1985) Arthritis 65. Rheum. 28, 537-541. Lowry, R.P. and Forbes, R.D.C. (1985) Transplant. Proc. 17, 66. 1963-1965. Ebbe, S. and Stohlman, F. Jr. (1965) Blood 26, 30-35. 67. 68. Rolovic, Z., Jovanovic, T., Stankovic, Z. and Marinkovic, N. (1985) Blood <u>65</u>, 60-64.