THE UNIVERSITY OF CALGARY

An Attempt to Isolate an Interleukin-2 Related Gene Using a Synthetic Oligonucleotide Probe

by

Gregory S. Storwick

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTERS OF SCIENCE

DEPARTMENT OF MEDICAL BIOCHEMISTRY

CALGARY, ALBERTA

MARCH, 1986

ISBN 0-315-30010-8

THE UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled, "An Attempt to Isolate an Interleukin-2 Related Gene Using a Synthetic Oligonucleotide Probe," submitted by Gregory Storwick in partial fulfillment of the requirements for the degree of Master of Science.

Dr. D.I. Hoar, Supervisor
Departments of Medical Biochemistry
and Paediatrics

Dr. H. van de Sande
Department of Medical Biochemistry

Dr. L. Gedamu
Department of Biology

Dr. D. Matheson
Department of Paediatrics

March 18, 1986.
Date

Dr. D. Hart
Department of Microbiology and
Infectious Diseases

ABSTRACT

A synthetic oligonucleotide probe was made to isolate the cloned sequence for the T cell activator, interleukin-2 (IL-2), from a human cosmid library. The twenty five base oligonucleotide was homologous to a region within the third exon of the published human IL-2 sequence (Fujita et al., 1983). When Southern transfers of Eco RI digested human genomic DNA were probed with the synthetic probe under high stringency conditions (hybridization and washing at 43°C.) five bands appeared on the autoradiograph. Besides the expected 3.7 kb IL-2 Eco RI fragment, four others ranging in size from 4.2 kb to 14 kb were seen. Since it has been reported that only a single copy of the IL-2 gene exists within the genome, the oligonucleotide was thought to be hybridizing to the IL-2 gene as well as some related sequences.

The human cosmid pool library was screened with the synthetic IL-2 probe to try and isolate some of these related sequences. The 4.2 kb hybridizing fragment seen in the genomic screen was present in the two cosmid pools, 7 and 11. Two identical cosmids containing the 4.2 kb hybridizing Eco RI fragment were isolated from cosmid pool 11.

The 4.2 kb fragment was isolated from one of the large cosmids and subcloned into the plasmid vector pUC8. The probe was found to hybridize to a 350 bp region delineated by Hinc II and Pst I. The small fragment was subcloned into pUC12 to sequence the region surrounding the probe.

The sequence revealed a discrepancy between the hybridizing region and the oligonucleotide probe. The probe is 76% homologous (19/25 bases) to the only possible hybridizing region only if a guanosine residue loops out when the probe binds to the genomic sequence. This type of mispairing with oligonucleotides has been previously reported in the literature (Panabieres et al., 1982). The sequence of the synthetic oligonucleotide probe was verified to rule out a missed base during synthesis. The probe was therefore probably not long enough and the stringency of the conditions high enough to detect specific sequences.

As expected the comparison of the DNA and amino acid sequence to that of the IL-2 gene identified no significant homology except at the region thought to be hybridizing to the probe. A search of a gene sequence data bank with the sequence of the 350 kb segment failed to identify homology to any other known sequences.

ACKNOWLEDGEMENTS

I would like to foremost thank my supervisor Dr. David Hoar for his continued patience, encouragement and helpful suggestions throughout my masters project. His philosophy of teaching reinforces valuable human attributes such as independent thought, patience and above all persistence.

I would like to especially thank Dr. Steve Krawetz, Bob Winkfein and Bernd Kalisch for their much needed help over the last two years.

The support of one's peers is also very important to sucessful laboratory research. Leo Dimnik, Debbie Starozik, and Katherine Tynan provided that and good company during my stay in David's lab.

The financial support of the Heritage Foundation for Medical Research and the William H. Davies Foundation is gratefully acknowledged. These scholorship funds greatly emphasize that the driving force behind medical research is people.

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Tables

Figures

LIST OF ABBREVIATIONS

A	absorbance
A	adenine
ATP	adenosine 5'-triphosphate
bp	base pair
BSA	bovine serum albumin
C	cytosine
^{o}C.	degree centigrade
Ci	curie
cDNA	DNA complementary to RNA
cpm	counts per minute
CTP	cytosine 5'-triphosphate
d	deoxyribo
dd	dideoxyribo
Denhardt's	0.02% BSA, 0.02% PVP 40, 0.02% ficoll
DNA	deoxyribonucleic acid
DTT	dithiothreitol
E. coli	Escherichia coli
EDTA	ethylene diamine tetracetate
g	gram
G	guanine
GTP	guanosine 5'-triphosphate
HEPES	N-2-Hydroxyethylpiperazine-N'-2-ethanesulfonic acid
kb	kilo base pairs
LB	Luria broth
l	liters
M	molar
m	mili (10^{-3})
mA	miliamps
min.	minutes
mRNA	messenger RNA
n	nano (10^{-9})
p	pico (10^{-12})
p	plasmid
PNK	phosphonucleotide kinase
RNA	ribonucleic acid
rpm	revolutions per minute
SDS	sodium dodecyl sulfate
T	thymine
TEMED	N,N,N',N' tetramethyl ethylene diamine
tk	thymidine kinase
Tris	tris-(hydroxymethyl) amino methane
u	units
u	micro (10^{-6})
UV	ultraviolet
v	volts

INTRODUCTION

The promise of a cure for diseases from the common cold to cancer has led to a flurry of activity in research on a group of molecules responsible for regulation within the immune system. Lymphokines or monokines, depending on the cell of origin, are growing classes of these hormone-like proteins that act on specific reticuloendothelial cells to influence their growth, maturation and function.

Interleukin-2 (IL-2) or T-cell growth factor is characterized as a lymphokine secreted by lectin or antigen stimulated T cells (Morgan et al., 1976, Gillis et al., 1978). Although it has been assigned other important immunoregulatory functions its key role appears to be the initiation and maintenance of T cell clonal expansion (Gillis and Smith, 1977). Interleukin-2 has been recently reviewed by Kendall Smith in the Annual Reviews of Immunology (Smith, 1984). Smith also earlier reviewed the subject in an article published in Immunological Reviews (Smith, 1980),

Interleukin-2 was first isolated from lectin stimulated human peripheral blood and bone marrow T cell cultures but has since been purified to homogeneity from a

variety of other tissues (Smith, 1980, Robb and Smith, 1981, Maretta et al.,1981). It has also been shown to be present in a number of other mammalian systems (Smith, 1980, Kashima et al., 1980). The human protein has a molecular weight of approximately 15,000 daltons (Robb and Smith, 1981) and a varying degree of reported glycosylation (Maeda et al.,1983).

The first cloning and nucleotide sequence of the cDNA for IL-2 from the human Jurkat leukemic T cell line was reported by Taniguchi's group in 1983 (Taniguchi et al., 1983). The cDNA codes for a protein of 153 amino acids including a signal sequence of 20 amino acids that does not appear in the secreted protein. The genetically engineered cDNA has been expressed in E. coli (Devos et al., 1983, Rosenberg et al., 1984) and cultured monkey cells (Taniguchi et al., 1983) to produce a biologically active protein product. More work on the molecular biology of IL-2 has led to the cloning and sequencing of the chromosomal gene and its flanking regulatory sequences (Fujita et al.,1983). The IL-2 gene has since been assigned to chromosome 4q. (Seigel, 1984).

In this project an attempt was made to isolate and characterize DNA sequences related to the human IL-2 gene. A synthetic oligonucleotide probe, corresponding to a coding region of the human IL-2 gene, was made to isolate the cloned sequence for IL-2 from a human cosmid library. Upon screening Southern blots (Southern et al., 1975) of

Eco RI digested human genomic DNA it was found that the oligonucleotide probe had specific homology to other sequences as well as the expected fragment for IL-2. It may be possible that these other homologous sequences are other cytokines that share a common domain with IL-2. Support of this hypothesis comes from the similarity in the cellular origin of cytokines, their target membrane receptors and their function. These concepts are disscussed below.

It was these other sequences that were of interest as the IL-2 gene had already been cloned (Fulita et al., 1983). Isolation and characterization of the homologous sequences would determine if they indeed were related to IL-2 and fell into one of the broad categories of cytokines.

Lymphokines

A list of some of the best characterized lymphokines together with the cells they act on is presented in table 1. (as adapted from Panem, 1983). Most of the soluble mediators originate from mononuclear or T cells after specific external activation. They in turn induce autoregulatory and/or growth control actions on other cells within the immune and reticuloendothelial system.

Interleukin-2 and interferon-gamma have been the most extensively studied lymphokines to date in part due to

early studies which suggested a possible role for these factors in cancer therapy. Since the cloning and expression of the first lymphokine gene, interferon-gamma (Gray et al.. 1982) several others have followed including interleukin-1 (Auron et al., 1984), lymphotoxin (Gray et al., 1984) and tumor necrosis factor (Pennica et al., 1984). Work on lymphotoxin and tumor necrosis factor is progressing rapidly, especially by the pharmaceutical and biotechnology industries, as both kill tumor cells in experimental systems and so have potential in tumor therapy (Sikora, 1984).

Lymphokine	Cell of Origin	Target Cell	Regulatory Function
Interleukin 1 (IL-1)	Monocytes/Macrophages	T cells	T cell activation
Interleukin 2 (IL-2)	T cells	T cells	Maturation and Proliferation
Interleukin 3 (IL-3)	T cells	Some cytotoxic T lymphocytes	T cell Growth
Interferon-gamma (IFN-Y)	T cells	Numerous	Numerous
Colony stimulating factor (CSF)	T cells and macrophages	Granulocytes	Stimulates growth of Granulocytes and Macrophages
B cell growth factor (BCGF)	T cells	B cells	Stimulates growth of B cells
T cell replacement factor (TRF)	T cells	B cells	Stimulates immuno-globulin production
Tumour necrosis factor (TNF)	Macrophages	Tumour cells	Lysis of tumour cells
Lymphotoxin	T cells	Tumour cells	Lysis of tumour cells

Table 1. Some of the best characterized lymphokines and their properties

The interrelationship between several known lymphokines involved in the growth regulatory activity of lymphoid cells is illustrated in figure 1 (as adapted from Maizel, 1984). Stimulation of mononuclear cells or macrophages by antigens or other exogenous stimuli (lectins) produces IL-1 which initiates the cascade of events involved in a cellular immune response. IL-1 acts on antigen primed T cells to trigger the synthesis and secretion of IL-2, colony stimulating factor (CSF) and interferon-gamma (also may be called macrophage activating factor, Feldmann, 1985). One of most important responses from the activated T lymphocytes is the production of IL-2 which induces long term proliferation and clonal expansion of T cells. A large population of T-helper, T-suppressor and cytotoxic T lymphocytes is produced to aid B cells and macrophages in the immune response. Colony stimulating factor and interferon-gamma act back on primed macrophages to stimulate their growth and maturation. B cell growth factor (BCGF) and T cell replacement factor (TRF) are produced by T-helper cells to aid in the proliferation and differentiation of B cells to form antibody secreting end stage plasma cells. In a classical immune response the specific antibody produced attaches to the antigen and aids macrophages and cytotoxic T cells in its destruction. This is a simplification of a yet undefined process that undoubtably involves many other still illusive soluble mediators and cell-cell interactions.

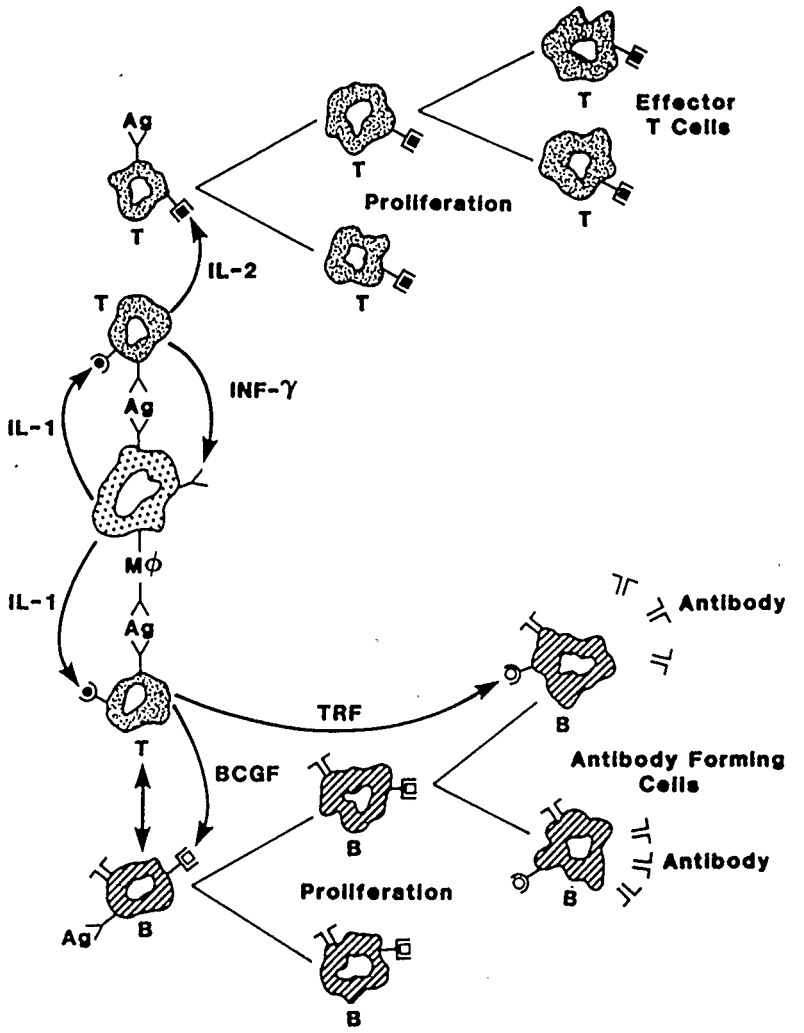

Figure 1. The interrelationship of some of the known lymphokines. Il-1 produced from antigen stimulated macrophages acts on T cells which subsequently produce other lymphokines important in an immune response; T, T cell; B, B cell, MO, Macrophage.

Interleukin-2 Receptors

Similar to other immunoregulatory soluble mediators, IL-2 binds to high affinity receptors on the surface of antigen stimulated T cells (Cantrell and Smith, 1983). From SDS-PAGE gel electrophoresis and immunoprecipitation studies the receptor appears to be a 60,000 dalton membrane glycoprotein devoid of intrachain disulfide bonds (Smith, 1984). The presence of the receptor on the surface of the cell is a transient response and only occurs shortly after a separate antigen receptor is triggered. To prolong the proliferative response the antigen must be continually present to maintain the synthesis of the IL-2 receptors. The number of receptors on the surface of the cell depends upon the amount of antigen present. The rate of T cell proliferation is also dependent upon the concentration of IL-2 available to the cells. This unique IL-2 hormone receptor system ensures that only antigen stimulated T cells within a population of lymphocytes respond by clonal expansion because unstimulated T cells do not express IL-2 receptors.

There is evidence for cross reactivity between the cellular receptors of the immune system and other hormones. Thyroid stimulating hormone and chorionic gonadotropin have been shown to effectively compete with interferon-beta for the binding to receptors on human cells. This commonality may be due to conserved homologous regions within the

protein hormones (Revel, 1979, Grollman et al., 1978). Conservation in functional sequences between different species is evident as many lymphokines will act across species. One example is human IL-2 which will mimic the action of murine IL-2 on murine lymphocytes (Feldmann, 1985). However this action is unidirectional as murine IL-2 will not stimulate human lymphocytes.

The Molecular Biology of Interleukin-2

The molecular biology of interleukin-2 has progressed rapidly since Taniguchi et al. (1983) first reported the cloning, nucleotide sequence and expression of a cDNA coding for IL-2. A selected cDNA library was constructed from a lectin induced, high IL-2 producing, Jurkat leukemic T cell line. The mRNA was fractionated on a sucrose density gradient and each fraction tested for IL-2 activity after *in vivo* translation by *Xenopus laevis* oocytes. Fractions which stimulated the incorporation of tritiated thymidine by an IL-2 dependent T cell line were chosen for library screening. This dramatically increased the probability of finding IL-2 positive clones.

The DNA sequence found coded for a polypeptide 152 amino acids long containing a signal sequence of 20 amino acids. The cDNA when expressed in cultured monkey COS (SV40 transformed simian cells) cells produced a biologically

active protein of 133 amino acids. This corresponds to a molecular weight of 15,000 daltons which is in agreement with the values obtained by fractional gel filtration (Robb and Smith, 1981).

Previous heterogeneity observed in the molecular weight was found to be due to differences in sialation and gylcosylation. O-linked sugars (O-Ser, O-Tyr) attached to the molecule during post translational modification altered the molecular weight and electrophoretic mobility of the protein on an isoelectric focusing gel (Smith, 1984, Robb et al., 1983).

About the same time Fiers group (Devos et al., 1983) using similar techniques also isolated a cDNA for IL-2. The DNA sequence isolated from mitogen stimulated splenocytes was appropriately engineered and inserted into E. coli. The biologically active bacterially derived IL-2 indicated that glycosylation was not necessarily important for biological activity.

From the published IL-2 cDNA sequence other groups isolated IL-2 clones from cDNA libraries using synthetic oligonucleotide probes. Rosenberg et al. (Rosenberg et al., 1984) isolated an IL-2 cDNA from human peripheral blood lymphocytes (PBL's) which when inserted into E. coli was expressed in high concentration. They again found no functional differences between native and recombinant interleukin-2 molecules.

Using the cDNA as a probe Taniguchi (Fujita et al.,

1983) cloned and sequenced the IL-2 chromosomal gene and its flanking sequences from human peripheral blood lymphocytes. The entire gene is divided into three introns and four exons centered within two adjacent Eco RI fragments. The first exon contains a 5' untranslated region and codes for the first 49 amino acids including the presumptive 20 amino acid secretory signal peptide. The second exon, which codes for the next 20 amino acids. is separated from the first exon by a short intervening sequence of 91 base pairs. A large intron (2292 base pairs) separates the third exon which codes for the next 48 amino acids. The final exon, which is followed by a termination codon codes for 36 amino acids. The expected promoter and termination elements are present in respective flanking sequences. Upstream from the translation initiation site 77 base pairs is the promoter sequence TATAAA. Transcription was found to occur 53 base pairs 5' to the translation initiation codon (Holbrook et al., 1984).

Other groups have since cloned and sequenced the chromosomal gene for IL-2. Degrave et al. (1983), Mita et al. (1983) and Holbrook et al. (1984) used the available IL-2 cDNA clone to screen bacteriophage Charon 4A human gene libraries. Holbrook's group screened three genomic libraries cloned into Charon 4A (human spleen, placental and peripheral blood lymphocyte) but were only successful in isolating IL-2 clones from the peripheral blood lymphocytes. Three overlapping clones were isolated which

contained the entire IL-2 gene. All four groups have published similar results with only minor discrepencies within the sequence and always within an intervening sequence (Holbrook et al., 1984).

Data reported groups who have isolated the human IL-2 gene indicates that there is only a single copy of the IL-2 gene within the human genome. Taniguchi's group (Fujita etal., 1983) reported that only a single species of mRNA coding for IL-2 could be identified in both a leukemic T cell line and peripheral blood lymphocytes. Using Southern blot analysis of cellular DNA's from normal and malignant lymphocytes Holbrook et al. (Holbrook et al., 1984) demonstrated that the predicted restriction map for the cloned genomic IL-2 was obtained. This indicated that there was only a single copy of the IL-2 gene and that no rearrangement of the genomic sequence took place in the malignant lines they studied.

There is little or no reported homology between IL-2 and other lymphokines except within a small 5' flanking region of gamma interferon (Gray and Goeddel, 1982). A 62% homology extending 50 nucleotides upstream from the "TATA" box suggests, at best, conservation in sequences that might be involved in the regulation of these two lymphokines as they are coordinately induced by similar stimuli (Mita et al., 1983). In addition there is a limited amount of homology between the human IL-2 gene and the adult T cell leukemic virus LTR region (Holbrook et al., 1984).

Significant homology between the sequence of murine IL-2, cloned by Kashima et al. (1985) and human IL-2 accounts for the fact that human IL-2 is active on mouse lymphocytes. The fact that mouse IL-2 does not act on human lymphocytes is probably due to a stretch of 12 glutamine residues present in mouse IL-2 but not in human IL-2 (Feldmann, 1985).

Research Rationale

The synthetic oligonucleotide probe, 25 nucleotides long, is homologous to a region within the third exon of the published human IL-2 sequence (shown in figure 2, Fujita et al., 1983). This corresponds to positions 293-317 of the reported sequence of the IL-2 cDNA (Taniguchi et al., 1983). The probe consisting of the bases: A CTC AAA CCT CTG GAG GAA GTG CTA, has a G-C content of 48% and is complementary to the interleukin-2 mRNA.

When Southern blots of Eco RI digested human genomic DNA, isolated from peripheral blood lymphocytes, were probed with the ^{32}P labelled probe, multiple bands appeared on the autoradiograph (see figure 3). Since it has been reported that only a single copy of the IL-2 gene exists within the genome (Fujita et al., 1983), the oligonucleotide probe may be hybridizing to the IL-2 gene and some related sequences.

From the previously published restriction map of the region surrounding the IL-2 gene (Figure 2) the probe should hybridize to a 3700 base pair Eco RI restriction fragment. When genomic Southerns were probed under high stringency conditions (hybridization, 0.9 M NaCl, and washing, 2X SSC, at 43 °C., see materials and methods) the 3700 base pair band showed up as well as four others ranging in size from 4200 to 14,000 base pairs. When the Southerns were probed at a lower stringency (hybridization at 43°C., 0.9 M NaCl, washing in 2X SSC at room temperture) up to 15 hybridizing bands were seen on the autoradiograph again including the expected 3700 base pair Eco RI fragment (figure 4).

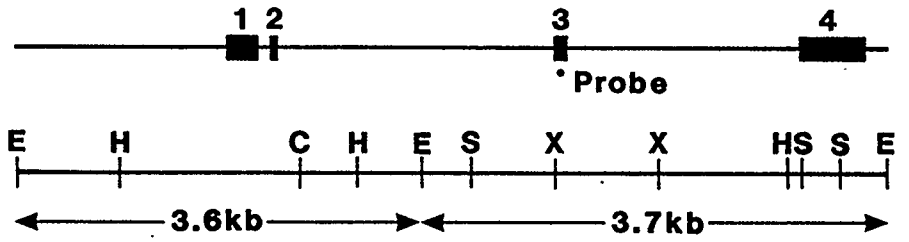

Figure 2. Restriction map of IL-2. The sequence is represented from 5' to 3' going left to right. The exons are indicated by the numbered black rectangles. The location of the oligonucleotide probe within the third exon is also shown. E, Eco RI; H, Hind III; C, Cla I; S, Stu I; X, Xba.

The stability between nucleic acid hybrids is affected by the ionic strength (M, in mol/liter) of the medium, base composition (% G+C), the length of the hybridizing fragment (n) and the concentration of destabilizing agents like formamide (Meinkoth and Wahl, 1984). The melting temperature (t_m) or temperature at which 50% of the DNA hybrids are dissociated is given by:

(1) $$t_m = 81.5°C. + 16.6LogM + 0.41(\%G+C) - 500/n - 0.61(\%formamide).$$

This pertains to nucleic acid fragments longer than 50 nucleotides. Shorter duplexes, especially between oligonucleotides 14-20 base pairs in length, show a decreased stability (Meinkoth and Wahl, 1984). The T_d or temperature at which 50% of these shorter duplexes dissociate can be estimated by multiplying the number of G and C residues by four and adding it to twice the number of A and T nucleotides:

(2) $$t_d = 4(G+C) + 2(A+T).$$

To detect perfectly matched molecules a hybridization temperature of 5°C. below the t_d is used. Thus a perfect match between the oligonucleotide probe, 25 bases long, and the IL-2 gene could be detected at a theoretical hybridization temperature of 69°C. (4(6+6)+2(8+5)-5°C.).

Mismatches between bases decreases the stability of the duplex formed. It has been estimated that mismatches within shorter duplexes of less than 20 base pairs decreases the t_m by appproximately 5-10°C. for every mismatched pair (Wallace et al., 1981). This is significantly more than if the duplex was just shorter by one base pair. With the IL-2 25 mer oligonucleotide probe, a hybridization temperature of 43°C. would only allow for five theoretical mismatches. Therefore only sequences with 20/25 matches would hybridize to the probe.

Another method to determine hybridization and washing conditions when using synthetic oligonucleotide probes is given by Lathe (1985). He proposes that hybridization should be done at 5 to 25 °C. below the t_m of the hybrid so that all possible matches can be formed. The temperature of a high stringency washing step, under salt conditions of 2X SSC, determines the degree of homology allowed between the probe and the sequence. The high stringency washing temperature (t_w) is determined by:

(3) $$t_w = 94 - 820/l - 1.2(100-h)$$

The value of 94 is the melting temperature for long random DNA. For reduced length of a hybrid the expression -820/l is used, where l is the probe length in nucleotides. The expression -1.2(100-h) is the change in the t_m per percent non-homology, where h is the percent homology

between the two sequences.

For the 25 mer IL-2 oligonucleotide, a hybridization temperature of 43°C. (25°C. below the t_m) and a high stringency wash at 43°C. (in 2X SSC) would allow hybrids with only 85% or greater homology to the probe. This corresponds to matches of greater than 21/25 nucleotides. At a lower washing temperature of 25°C. up to eight mismatches would be allowed, which explains the presence of additional bands when genomic Southerns were probed at a lower stringency.

In theory, an oligonucleotide of 16 bases should be uniquely represented within the human genome. This is arrived at from the formula where the number of hybridizing sequences (N) is equal to twice the complexity of the gene pool (C) divided by 4 to the power of the probe length (h) (Wood et al., 1984).

(4) $$N = C(2)/4^h.$$

In practice however at least a 19 base oligonucleotide probe is needed to have a high probability of recognizing an unique sequence from the human genome (Conner et al., 1983, Lathe, 1985).

In view of the conditions used (hybridization and washing at 43°C.) to identify the interleukin-2 3700 base pair Eco RI fragment, the other bands should represent unique sequences with a high degree of homology to the

probe and possibly IL-2. Other evidence for relatedness of the sequences to the probe and possibly interleukin-2 come from other experiments.

When genomic Southern blots using other restriction enzymes (Hind III, Xba I and Bam H1) were analyzed with the probe, multiple bands emerged as well as the expected ones for IL-2. This included the anticipated 3600 base pair Hind III fragment and the 1000 base pair Xba I band. In another experiment, mouse Eco RI cut genomic DNA Southern blots were analyzed with the probe. A strongly hybridizing band of 2500 base pairs was seen along with other less intense signals.

Research Goal

From this point on the project focused on the isolation and characterization of 4.2 kb Eco RI fragment hybridizing to the probe. The oligonucleotide will be used to screen a human cosmid library from placental DNA in an attempt to identify the 4.2 kb fragment. Sequencing of the fragment would determine the amount of homology to the oligonucleotide probe, IL-2 and to any other lymphokine of which the sequence is known. Northern analysis of transformed human T and B cell lines may identify the expression of the fragment in these tissues.

MATERIALS AND METHODS

Preparation of the Oligonucleotide Probe

The oligonucleotide probe, 25 bases long corresponding to a region within the third exon of the IL-2 gene was synthesized on a Applied Biosystems Model 380A synthesizer using phosphoramidite chemistry (Beaucage and Caruthers, 1981). A final yield of 104 A_{260} units was dissolved in 500 ul of water and further purified by preparative polyacrylamide gel electrophoresis. Six A_{260} units were applied to a 2 cm well on a 20% polyacrylamide gel containing 7M urea (Maniatis et al.,1982). The 34 x 40 cm, 1.5 mm thick gel was run for 4 hours at 600 volts. Short wave U.V. shadowing was used to visualize the major band corresponding to 25 base pairs (Applied Biosystems, 1983). The oligonucleotide within the polyacrylamide gel was excised and subsequently electroeluted into Spectrapor dialysis tubing (Spectrum Medical Industries, mw cut off 2000) in 1X Tris-Borate (TEB) buffer (10X TEB = 0.9M Tris borate, pH 8.3, 0.9M Boric acid, 0.02M EDTA, pH 8.0) at 50 mA for 2 hours. The eluted material was dialysed against 2 liters of water, 25 mM EDTA overnight at 5°C.

The sample was subsequently lyophilized to dryness,

resuspended in 100 ul of water, spun at 12,000 rpm to remove insoluble matter and relyophilized. The resulting pellet was resuspended in 200 ul of water to yield a final concentration of approximately 1 ug/ul.

The oligonucleotide was phosphorylated using T4 polynucleotide kinase (PNK) and crude [gamma ^{32}P] labelled ATP purchased from ICN (greater than 7000 Ci/mmole) to yield a specific activity of 1 x 10^9 cpm/ug (Meinkoth and Wahl, 1984). A hundred nanograms (10-20 pmoles) of DNA was added to 5 ul of 2X kinase buffer (140 mM Tris-HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM DTT), 4 units of PNK, 70 uCi (3-4 molar excess) of [gamma ^{32}P] ATP and reacted at 37°C for 30 minutes.

Unincorporated label was removed from the labelled material by gel filtration through a Bio-Gel (Bio-Rad) P4 column 4 mm x 100 mm long according to the methods of Yousaf et al. (1984). The column was equilibrated and eluted with buffer containing 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 10 mM EDTA, 0.1% SDS. The column was monitored using a hand held Geiger counter. The eluted probe contained in the first peak was used directly for Southern analysis of genomic and cosmid DNA.

Preparation of Genomic and Cosmid DNA

Genomic DNA from human peripheral blood lymphocytes

was isolated essentially RNA free by a procedure from the Molecular Diagnostic Laboratory of the Alberta Children's Hospital (MDL Method 001). Hypotonic tris buffered ammonium cloride (0.4 M NH_4Cl, 0.02 M Tris-HCl, pH 7.65) was used to lyse the non-nucleated red cells so that the nucleated cell fraction could be isolated intact. Four volumes of the NH_4Cl:Tris solution warmed to 37°C. were added to 1 volume of whole blood and incubated at 37°C. for 5 minutes. The solution was centrifuged at 2000 rpm for 10 minutes and the resultant supernatant was aspirated off. The pellet was resuspended in 10 mls of saline (0.85 % NaCl), recentrifuged and redissolved in 2 mls of high TE (100 mM Tris, pH 8.0, 40 mM EDTA). The cells were immediately lysed by injecting 2 mls of lysis mixture (100 mM Tris, pH 8.0, 40 mM EDTA, 0.2 % SDS) into the solution with a 5 ml syringe. Phenol/chloroform extractions were done to remove proteins and organic matter. The nucleic acids were precipitated in the presence of 1/10 volume 4 M ammonium acetate by adding an equal volume of isopropanol. The precipitated DNA from 50 mls of blood was collected on the end of a curved pasteur pipette, washed with 70% ethanol, dried and redissolved in 1 ml of low TE (10 mM Tris-HCl, pH 8.0, 1 mM EDTA).

The human cosmid library, a generous gift of Dr. J. Collins, was prepared by cloning approximately 45 kilobase (kb) Msp I partial digestion fragments of human placental DNA into a derivative of pHC 79 (Hohn and Collins, 1980).

The vector pHC 79-2 cos/tk contains a Herpes Simplex Virus thymidine kinase (tk+) gene and tandom cos sites for efficient in-vivo packaging (Lindenmeir et al., 1982). After in-vitro packaging and infection of the lysogenic strain E. coli 1400 (supE, supF, hsdS$^-$, met$^-$, recA56, lambda L512, Cami and Kourilsky, 1978) the total pool was divided into twelve separate aliquots each containing 20,000-30,000 clones.

The in-vivo packaging system of Lindenmeir was employed to obtain the human placental cosmid library as a cosmid lysate (Lindenmeir and Hauser et al., 1982). Transfection of the lysogen BHB 3064 (HB101, lambda^{i4}, cIts, red3, Sam7) containing the temperature inducible lambda i434 prophage was done as previously described by Lindenmeir (1982). Fresh BHB 3064 was grown in NZ-broth plus maltose (5g NaCl, 2g $MgCl_2$, 10g NZ-amineA , 4% maltose per liter) at 32°C. to late log phase. One hundred microliters of the cosmid lysate from each pool was added to 100 ul of TM (50 mM Tris, pH 7.5, 10 mM $MgCl_2$) plus 100 ul of BHB 3064 and incubated at 32°C. for 20 minutes. 5 mls of L-broth plus 50 ug ampicillin/ml was added to each tube and the solutions incubated with vigorous shaking at 32°C. overnight.

These starter cultures were used to innoculate 500 mls of M9 salts media plus 50 ug ampicillin/ml (Clewell and Helinski, 1970) which were then grown at 32°C. to an A_{600} of 0.6. Cloramphenicol (170 ug/ml) and 10 mls of 20%

glucose were added and the cultures further incubated for 16 hours.

The cosmid DNA was prepared by a modification of the procedure of Godson and Vapnek (Godson and Vapnek, 1973). The bacterial pellet was resuspended in 30 mls of STE (8% sucrose, 100 mM Tris, pH 8.0, 40mM EDTA) mixed with 80 mg of lysozyme and left at room temperature. After ten minutes 4 mls of 0.2% SDS in 100 mM Tris, pH 8.0, 40 mM EDTA was injected into the solution and left at 37°C. until clear. After clearing, 4 mls of 5 M potassium acetate was added and the solution put on ice for 30-40 minutes. The chilled tubes were centrifuged at 17,000 rpm for 45 minutes at 4°C. in a Sorval SS34 rotor. The supernatant was collected and mixed with 0.6 volumes of absolute isopropanol. The resulting precipitate was collected by centrifugation at 10,000 rpm and air dried.

The pellets were dissolved in 19 mls of 100 mM Tris, pH 8.0, 4 mM EDTA and added to 20.4 g of solid CsCl. Ethidium bromide (2.4 mls of 5 mg/ml) was added before the solutions were dispensed into two 80Ti tubes, heat sealed and centrifuged in an 80Ti rotor on a Beckman L8-80 ultracentrifuge for 44 hours at 45,000 rpm.

Following puncture of the top of the tube, the lower band, representing the covalently closed cosmid DNA, was collected through a 21 gauge hypodermic needle inserted into the side of the tube.

Three to four N-butanol (water saturated)

extractions were done to remove ethidium bromide and then water added to bring up the volume to four times the original. Precipitation of the nucleic acids was done by adding 1/10 volume of 4 M ammonium acetate and 2.5 volumes of 95% ethanol, freezing in liquid nitrogen and centrifuging to collect the pellet. The precipitate was washed with 70% ethanol to remove salt and recentrifuged. The cosmid DNA from each pool was redissolved in an appropriate volume of low TE (10 mM Tris pH 8.0, 0.1 mM EDTA).

Southern Transfer and Colony Screening

Digestion of the cosmid pool (1 ug) and genomic (10 ug) DNA by the restriction enzyme Eco-R1 (1 unit/ug) was done in high salt buffer (100 mM NaCl, 50 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM DTT) at 37°C. for 5 hours (Maniatis et al., 1982). The completed reactions were stopped by the addition of gel loading dye (0.025% bromophenol blue, 0.025% xylene cyanol FF, 50% glycerol, 0.1% SDS) and loaded onto a 1% agarose gel formed by melting 2.5 g of powdered agarose into 250 mls of TEA buffer (0.04 M Tris-acetate, 0.001 M EDTA, 0.5 ug/ml ethidium bromide). The agarose gel was run in TEA buffer at 120 mA for 4 hours.

The electrophoresed DNA was transfered to nitrocellulose by the passive diffusion technique

established by Southern (Southern, 1975). To facilitate the transfer of high molecular weight DNA the gel was first acid hydrolyzed for 2 x 15 min. in 0.25 M HCl (Wahl et al., 1979), and then subsequently denatured (2 x 15 min. in 1.5 M NaCl, 0.5 M NaOH) and neutralized (2 x 30 min. in 1.5 M NaCl, 0.5 M Tris-HCl, ph 7.4). The gel was then placed on filter paper (Watmann 3MM), cut to the size of the gel, which was placed on two shammy soaked in 20X SSC (3 M NaCl, 0.3 M NaCitrate). The nitrocellulose soaked in 2X SSC along with two sheets of filter paper was layered on top followed by a 1-2 inch stack of paper towels. Transfer took place in 20X SSC overnight after which the filter was floated in 2X SSC and baked in a vacuum oven at 80°C. for 1-2 hours.

The filters were prehybridized in hybridization buffer (0.9 M NaCl, 6 mM EDTA, 90 mM Tris-HCl, pH 7.5, 0.1% SDS) at 43°C. for one to three hours (Meinkoth and Wahl, 1984). Homo-chromatography mix II (Jay et al., 1974) was added as carrier nucleic acid at a concentration of 100 ug/ml.

The filters were hybridized for 16 hours in hybridization buffer containing the probe at a concentration of approximately 10 ng/ml or 5 x 10^6 cpm/ml. High stringency hybridization was done at a hybridization temperature of 43°C. followed by four, 15 minute, high stringency washes in 2X SSC at 43°C. Low stringency hybridization was also done at 43°C. but washed at room temperature (25°C.) for 4 x 15 min. in 2X SSC.

The nitrocellulose filters were dried at room temperature, wrapped in plastic wrap and applied as a sandwich between two x-ray films flanked by two Cronex Quanta III intensifying screens. After 16 hours the first film was developed to check for signal intensity. The second film was developed according to the results of the first.

Cosmid pool 11 was chosen for colony screening to isolate the desired 4.2 kb clone shown to be present in that pool when the cosmid pools were screened. Screening was done in a similar fashion to the method published by Grosveld (Grosveld and Dahl, 1981) for the isolation of beta-globin related genes from a human cosmid library. Fresh BHB 3064 was innoculated (as previously described) with the cosmid lysate and grown at the permissive temperature of 32°C. overnight. The transformants were appropriately diluted and plated onto six 150 mm ampicillian (50 ug/ml) MacConkey agar plates such that there were approximately 5000 colonies per plate (30,000 in total) and incubated at 32°C. for 16 hours. Nitrocellulose lifts of the original plates were grown on chloramphenicol (170 ug/ml) MacConkey agar plates for a further 16 hours. The original plates were further incubated for 6 hours at 32°C. and then sealed and stored in the cold until the results of the hybridization screening reaction were available.

The nitrocellulose replica filters were denatured and

lysed by floating for 4 minutes on 0.5 M NaOH according to the procedure of Grosveld (Grosveld and Dahl, 1981). They were immediatly floated for another 4 minutes on 1 M Tris, pH 7.4, 1.5 M NaCl and washed by shaking in 2X SSC, 0.1% SDS. Cell debris was removed by gently wiping with a soft tissue. Following washing, the filters were baked for 1-2 hours at 80 °C. in a vacuum oven.

The hybridization procedure was the same as previously described for the Southern blots. Hybridization was at 43°C. for 16 hours at a probe concentration of 5 x 10^6 cpm/ml. Positive colonies were picked from the original plates and transfered in duplicate to grided ampicillin (50 ug/ml) LB agar plates. After incubation at 32°C. overnight nitrocellulose lifts were done as described above and the filters hybridized at 55°C. Cosmid DNA preparations were done on colonies that were positive on both plates.

Restriction Mapping and Subcloning

The two positive cosmid DNA clones from pool 11 (5 ug) were cut with the restriction enzymes Eco-R1, Bam-H1 and Hind III to delineate the region that hybridized to the probe. Eco-R1 and Bam-H1 digestions were done in high buffer as previously mentioned whereas the Hind III reactions (1 unit enzyme/ug DNA) were in medium buffer (50 mM NaCl, 10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$, 1 mM DTT) at

37°C. for 4 hours. The completed reactions stopped with loading dye were loaded onto a 1% agarose gel and run at 100 mA for 4 hours. The gel was transferred to nitrocelllulose as described previously (Southern, 1975) and hybridized to the oligonucleotide probe at 43°C. overnight.

The positively hybridizing 4.2 kb Eco-R1 fragment was subcloned into the plasmid vector pUC-8 (Viera and Messing, 1982). Twenty micrograms of the cosmid pool 11 clone was cut with Eco-R1 and run on a 1% agarose gel until the bands in the 3-8 kb region were well separated. Agarose in front of the 4.2 kb Eco-R1 fragment was cut away and replaced with a small piece of Spectrapor dialysis tubing (mw cut-off 2000) to form a well which was filled with 1X running buffer. The gel was run at 70 mA for three ten minute periods with the buffer being collected and replaced at each interval. At the end of 30 minutes the current was reversed for 2 seconds and the buffer collected for a final time. Several N-butanol extractions were done to reduce the volume and remove any traces of ethidium bromide. To precipitate the DNA 1/20 volume of 4 M ammonium acetate and 2.5 volumes of cold 95% ethanol were added, the solution frozen and then spun at 12,000 rpm for 15 minutes. The pellet was collected, dried and resuspended in low TE.

The plasmid cloning vector pUC 8 was linearized by digestion with Eco-R1 and phosphatased with calf intestinal alkaline phosphatase (Boehringer Manheim) to remove the

terminal 5' ends (Maniatis et al.,1982). After a phenol/chloroform extraction and precipitation 200 ng of the vector were mixed together with a threefold molar excess of the 4.2 kb Eco-R1 fragment, 1X ligation buffer (10X = 0.06 M Tris-HCl, pH7.5, 50 mM $MgCl_2$, 50 mM DTT, 10 mM ATP), 10 units of T4 ligase and incubated at 4°C. for 8 hours. After ligation 2.5 ul were used to transform E. coli HB 101 (F^-, hsdS20, recA13, ara-14, proA2, lacY1, galK2, rpsL20, xyl-5, utl-1, supE44, lambda$^-$) according to the procedure of Hanahan (Hanahan,1983).

Colonies which exhibited antibiotic resistence were picked with a sterile toothpick and transferred to ampicillin (50 ug/ml) LB agar plates in duplicate. Nitrocellulose lifts were done on the plates after incubation at 37°C. overnight. The filters were processed as previously mentioned and probed at 43°C. with the oligonucleotide probe. One liter plasmid preparations were done on the positive colonies according to the procedure as previously described for cosmids (Godsen and Vapnek, 1973).

Probing of Genomic and Cosmid DNA with the 4.2 kb subclone

Southern transfer blots of Eco RI, Bam H1 and Hind III cut genomic DNA and Eco RI cut cosmid DNA were probed using the 4.2 kb Eco RI fragment isolated from cosmid pool 11 as a probe. The probe was radiolabeled with the large

fragment of DNA polymerase I, using random oligonucleotides as primers according to the procedure of Feinberg and Volgelstein (1983). Approximately 100 ng of the heat denatured isolated 4.2 kb insert was added to 10 ul of 5X oligo-labeling buffer (5X = Solution A:B:C: mixed in the ratio of 100:250:150, A: 1.25 M Tris-HCl, pH 8.0, 0.125 M $MgCl_2$, 0.01 M beta mercaptoethanol B: 2 M HEPES, pH 6.6 with NaOH C: Hexanucleotides at 90 OD_{260}/ml in 3 mM Tris, pH 7.0, 0.2 mM EDTA), 5 ul of 20X cold dNTP's (0.2 mM ATP, GTP, CTP and TTP), 10 ul of bovine serum albumin (2 mg/ml), 7.5 ul of [alpha-^{32}P] labelled dCTP and 2 units of the large fragment of DNA polymerase I. The volume was made up to 50 ul and the solution left at room temperature for 2-4 hours. Unincorporated label was removed by gel filtration through a Sephadex G-50 (Pharmacia) column equilibrated and eluted with HENS buffer (10 mM Hepes, 1 mM EDTA, 10 mM NaCl, 0.1% SDS). THe column was monitered using a hand held Geiger counter and the eluted probe contained in the first peak was used directly for Southern hybridizations.

The nitrocellulose filters were prehybridized at 43°C. for a minimum of 6 hours in prehybridization buffer (50% formamide, 5X SSC, 50 mM Na Phosphate, pH 7.2, 5X Denhardts, 0.25 mg/ml of heat denatured salmon sperm DNA and 0.1% SDS). Hybridization was done at 43°C. for 16 hours with a probe concentration of 1 x 10^6 cpm/ml in hybridization buffer (50% formamide, 5X SSC, 20 mM Na Phosphate, 1X Denhardts, 0.1 mg/ml heat denatured salmon

sperm DNA, 0.1% SDS and 10%, w/v, dextran sulfate). Following hybridization, four 15 minute washes were done in 2X SSC, 0.1% SDS at 43°C. followed by four 15 minute high stringency washes in 0.2X SSC, 0.1% SDS at 43°C.. The nitrocellulose filters were air dried and set up for autoradiography as described previously.

Subcloning of the 350 bp Hybridizing Region

Further detail of the hybridizing region was obtained by digestion with the restriction enzymes Pst I, Pvu II, Hinc II, and Hind III. A small piece, 350 bp in length, delineated by Pst I and Hinc II was chosen for additional subcloning and subsequent sequencing.

The small 350 bp fragment was isolated from an agarose gel, ligated into pUC-12 (cut with Pst I and Sma I) and transformed into HB101 using the same procedure as for the larger 4.2 kb fragment. Instead of using hybridization techniques to screen for positive clones, plasmid preparations were done on five colonies picked at random from the plate.

The single colonies were grown up in 10 mls of LB broth plus ampicillin (50 ug/ml) at 37°C. overnight. Chloramphenicol (170 ug/ml) was added and the culture further incubated at 37°C. for 16 hours. The bacteria were pelleted by spinning at 2000 rpm for 10 minutes,

resuspended in 0.5 mls of 50 mM Tris, pH 8.5, 50 mM EDTA, 15% sucrose and 1 mg/ml of lysozyme. After standing at room temperature for 10 minutes, 10 ul of 10% SDS was added and the tubes inverted to mix. Fifty microliters of 5 M potassium acetate was added before cooling on ice for 30 minutes. The sediment was precipitated in a microfuge tube for 15 minutes in the cold at 12,000 rpm. The DNA contained in the supernatant was decanted into a new tube, phenol/chloroform extracted and ethanol precipitated. The pellet was redissolved in 50 ul of 10 mM Tris, pH 7.5, 1 mM EDTA and ribonucleased with 10 ug/ml of heat treated RNase A at 37°C. for one hour. Phenol/chloroform extractions were again done before precipitating the DNA with 1/10 volume 4 M ammonium acetate and 2 volumes of cold 95% ethanol.

DNA Sequencing

Sequencing using the dideoxy chain termination method established by Sanger (Sanger et al., 1977) was done following a modification of the procedure using AMV reverse transcriptase (Life Sciences Inc.) by Karanthanasis (1982). Lyophylized [alpha-^{32}P] labelled dCTP (20 uCi, 2000 Ci/mmole) was resuspended in 5 ul of dNmix (0.5 mM dGTP, 0.5 mM dTTP, 0.5 mM dATP), 1 ul of cold dCTP (0.05 mM) and 0.8 ul of 10X RT buffer (0.5 M Tris-HCl, pH 8.5, 0.5 M Nacl, 50 mM $MgCl_2$, 20 mM DTT). This solution was mixed with

1 ug of the plasmid DNA that was first digested with Pvu II and annealed with a commercially available M13 primer (3-4 molar excess) by heating at 100°C. for 3 minutes and cooling on ice for five minutes. Three microliters of the template + primer + [alpha-^{32}P] dCTP solution was added to four eppendorf tubes marked A, C, G and T. To each tube 1 ul of the appropriate dideoxy-deoxy nucleotide mix (62.5 uM ddATP or 62.5 uM ddGTP or 5 uM ddCTP or 12.5 uM dd TTP all from Pharmacia) was added followed by 1 ul of RTmix (1 ul AMV reverse transcriptase, 18 units/ul, 0.5 ul 10X RT buffer, 3.5 ul water). The reactions were incubated at 43°C. for 15 minutes after which 3 ul of all chase (1.25 mM dATP, dGTP, dTTP, 1.125 mM dCTP in 1X RT buffer) was added and the reactions further incubated at 43°C. for 15 minutes. The reactions were stopped by the addition of 7 ul of stop solution (90% formamide in 1X TBE buffer, 0.04% xylene cyanol FF, 0.04% bromophenol blue) and heated at 100°C. for 3 minutes before loading onto a polyacrylamide sequencing gel.

The sequencing gels were 85 cm x 20 cm x 0.4 mm, 6% polyacrylamide buffer gradient gels. (Biggin and Gibson, 1983) containing 8 M urea. The bottom gel which contained 10 g urea, 2 g sucrose, 3 mls 40% acrylamide (38:2 acrylamide:bis-acrylamide), 5 mls 10X TBE (0.4 M boric acid, 0.25 M EDTA, 0.75 M Tris, pH 7.2), 10 ul TEMED, 120ul 10% ammonium persulfate and 4.6 mls of water was poured first and allowed to settle. The upper gel (40 g urea, 12

mls of 40% acrylamide, 4 mls 10X TBE, 20 ul TEMED, 500 ul 10% ammonium persulfate and 38 mls water) was layered on top by pouring down the opposite side. The gel was allowed to polymerize for 1 hour and then pre-electrophoresed at 200 volts for 2 hours. After loading 3 ul of each A, G, C, and T reaction the DNA was electrophoresed at 2500 volts for 6 hours.

On occasion 8% non-gradient polyacyramide gels were run to further identify certain regions. The gel was poured and run in a similar fashion as the gradient gels but contained 40 g urea, 17 mls 40% acrylamide, 4 mls 10X TBE, 20 ul TEMED, 500 ul 10% ammonium persulfate and 33 mls of water.

The sequence was analyzed using three programs designed for DNA sequence data analysis. Entering and editing of all sequence data was done using a program by Schwindinger and Warner (1984). homology to the oligonucleotide probe as well as other fragments was checked using a program by Mount and Conrad (1984). The dot matrix program used to compare the sequence to that of other human sequences was that of Lagrimini et al. (1984)

The MicroGenie DNA and NBRF (National Biomedical Research Foundation) protein data banks were also searched for homologous sequences.

Northern Analysis

RNA from continuous human B and T cell lines, generous gifts of Dr. D. Matheson and Dr W. Dixon, were screened with the oligonucleotide and the isolated 4.2 kb fragment to determine the presence of an RNA transcript. The human B lymphoblastoid cell line, CESS, (Muraguchi et al., 1981) and the human leukemic T cell line Jurkat (Schneider et al., 1977) were used.

The suspension cultures were maintained by the addition or replacement of media every 1-2 days. The media was the alpha modification of Eagle's media (Stanners et al., 1971) filter sterilized and supplemented with 12% fetal calf serum (Flow Laboratories). The pH was maintained at 7.15 by the addition of 20 mM HEPES buffer. The cells were incubated at 37°C. in humidified air containing 5% CO_2.

Total RNA was prepared using guanidine thiocyanate and LiBr according to the procedure of Krawetz et al. (1986). Approximately 2 x 10^7 cells (20 mls), spun down at 2000 rpm for 5 minutes, were lysed in 2 mls of homogenization solution (4 M guanidine thiocyanate, 2% 2-mercaptoethanol, 0.1 M Tris-HCl, pH 7.2, 2 M LiBr, 50 mM N-lauryl sarcosine) and transferred to a Corex tube. Acetic acid (0.025 volumes) and absolute ethanol (0.5 volumes) were added and the RNA precipitated overnight at -20°C. The RNA was pelleted by centrifugation at 10,000 rpm for 15

min. at -10°C. and redissolved in 1 ml of resuspension solution (6 M guanidine-HCl, 10 mM DTT, 20 mM EDTA, 0.1M Tris-HCl, pH 7.2). After removal of the insoluble material by centrifugation the RNA was precipitated by the addition of 0.06 volumes 3 M sodium acetate (NaAc), pH 5.2 and 0.5 volumes absolute ethanol at -20°C. overnight. The resulting pellet was resuspended in a minimal volume of 0.3 M NaAc and precipitated again by adding 2.5 volumes of ethanol and storing overnight at -20°C. The RNA was again peletted as above, washed with absolute ethanol, dried, and redissolved in an appropriate amount of distilled H_2O.

For reasons of safety, agarose gels containing formaldehyde as opposed to methyl mercury were used for RNA electrophoresis and Northern analysis (Maniatis et al., 1982). The gel was prepared by first melting 0.5 g of agarose in 37 mls of sterile water. Before pouring into a gel tray, 8 mls 37% formaldehyde and 5 mls 10X running buffer (filter sterilized 0.4 M morpholinopropanesulfonic acid, MOPS, pH 7.0, 100 mM NaAc, 10 mM EDTA) were added. The sample (in 4.5 ul) was prepared by adding 2 ul of running buffer, 3.5 ul 37% formaldehyde, 10 ul deionized formamide and incubating at 55°C. for 15 minutes. Two microliters of loading buffer (50% glycerol, 1mM EDTA, 0.4% bromophenol blue) were added and the sample loaded onto the gel.

The gel was run at 50 mA for approximately 1.5 hours and then transferred directly to nitrocellulose without

staining and visualization. After electrophoresis the gel was washed for 5 min. in several changes of water and then soaked for 45 min. in 50 mM NaOH, 10 mM NaCl. It was then neutralized in 0.1 M Tris-HCl, pH 7.5 for 45 min. and soaked in 20X SSC for 1 hour. Transfer took place overnight in 20X SSC as described previously for Southern blots. The next day the filter was air dried and baked at 80°C. for 1 hour.

The filters were prehybridized, hybridized, washed and autoradiographed as previously described for Southern blots using both double stranded and oligonucleotide probes.

After autoradiography the Northern transfers were stained to visualize the RNA transferred to the nitrocellulose. The dried filters were soaked in 5% acetic acid for 15 minutes and transferred to a solution containing 0.5 M NaAc, pH 8.2, and 0.04% methylene blue. After 5-10 minutes the filters were rinsed with water and dried. The presence of the ribosomal 28S and 18S RNA's served as markers.

Sequencing the Oligonucleotide Probe

The sequence of the oligonucleotide probe was confirmed using a modified Maxam-Gilbert procedure for oligonucleotides 5 to 120 bases long (Applied Biosystems

Inc., 1984). With this technique it was possible to sequence all 25 bases of the probe.

The purified oligonucleotide (300 ng) was phosphorylated with [gamma^{32}P] labelled ATP as previously described. Unincorporated label was removed by gel filtration through a Bio-Gel P4 column. The column was equilibrated and eluted with low TE buffer (10 mM Tris-HCl, pH 8.0, 1 mM EDTA) to minimize the final salt concentration. The labelled probe was collected and precipitated by the addition of 1/20 volume 4 M ammonium acetate and 2.5 volumes of 95% ethanol. The pellet was collected by centrifugation, dried on a Speed-Vac concentrator and resuspended in 30 ul of water.

A portion of the sample was partially cleaved at the pyrimidine bases (C+T) by treatment with hydrazine. Ten microliters of the labelled DNA was added to 10 ul of water, 30 ul hydrazine, vortexed and incubated at 37°C. for 20 minutes. The reaction was stopped with 200 ul of 0.3 M sodium acetate (pH 6.0 at 0°C.), 10 ul tRNA (1 mg/ml), and 1.25 ml ethanol (-20°C.).

A portion of the sample was partially cleaved at the purine bases (G+A) by treatment with formic acid. Ten microliters of DNA was added to 50 ul of 88% formic acid, vortexed and incubated at 37°C. for 20 minutes. The reaction was stopped as above.

A portion of the sample was partially cleaved at the cytosine bases (C) by treatment with hydrazine in the

presence of salt. Five microliters of the DNA was added to 15 ul of 5 M NaCl, 30 ul hydrazine, vortexed and incubated at 25°C. for 30 minutes. The reaction was stopped as previously descibed.

Finally, a portion of the sample was partially cleaved at the guanine bases (G) by treatment with dimethyl sulfate. Five microliters of the DNA was added to 200 ul of 50 mM sodium cacodylate buffer (pH 8.0, 1 mM EDTA), 1 ul dimethyl sulfate, mixed and incubated at 25°C. for 20 minutes. This reaction was stopped with 50 ul of a 1.5 M sodium acetate (pH 7.0), 1.0 M mercaptoethanol buffer, 10 ul tRNA (1 mg/ml) and 1.25 ml of 95% ethanol (20°C.).

The salts and reagents from each of the four tubes were removed by selective precipitation with ethanol. The tubes were placed in liquid nitrogen until frozen, thawed and centrifuged at 14,000 rpm for 15 minutes. The supernatants were decanted and discarded and the pellets rinsed by adding 70% ethanol, centrifuging and again decanting the supernatant. The pellets were redissolved in 250 ul of 0.3 M sodium acetate (pH 6.0) and reprecipitated by adding 1 ml of 95% ethanol. The pellets were collected and rinsed with 70% ethanol as above.

To cleave the chains at the positions where the bases were removed all four reactions were treated with piperidine. The pellets were dissolved in 100 ul of 1 M piperidine and incubated at 90°C. for 30 minutes. The reagent was removed by lyophilizing overnight on a

Speed-Vac concentrator. The remaining piperidine was removed by adding 10 ul of water and relyophilizing twice.

The reactions were prepared for loading onto a polyacrylamide gel by dissolving the material in each tube in 10 ul of 80% formamide, 10 mM NaOH, 1 mM EDTA, 0.1% Xylene cyanol FF and 0.1% bromophenol blue. The solutions were heated for two minutes at 90°C. and loaded immediately.

The 20% polyacrylamide gel, containing 7 M urea, was run at 1600 volts for three hours and visualized by autoradiography.

RESULTS

Probe of Human Genomic DNA

Southern blots of Eco RI cut human genomic DNA derived from peripheral blood lymphocytes (PBL's) were screened with the oligonucleotide probe at a hybridization and wash temperature of 43°C. Five distinct hybridizing bands were seen on the autoradiograph (figure 3). The expected IL-2 3.7 kb Eco RI restriction fragment was present as well as four others of sizes 4.2, 4.6, 5.8 and 14 kb. Although the signal intensity of each fragment varied between different autoradiographs, the 4.2 and 5.8 bands were consistently more intense. This difference in intensity could be explained by a difference in copy number of the various DNA fragments. Subsequent data suggests that this may be due to varying degrees of homology between the probe and the hybridizing sequences.

When Eco RI cut genomic Southern blots were probed under low stringency conditions (hybridization at 43°C., washing in 2X SSC at 25°C.) additional signals were seen on the autoradiograph (figure 4). Although the background was high no difference was observed in the restriction patterns between the Eco RI cut placental DNA and the PBL genomic

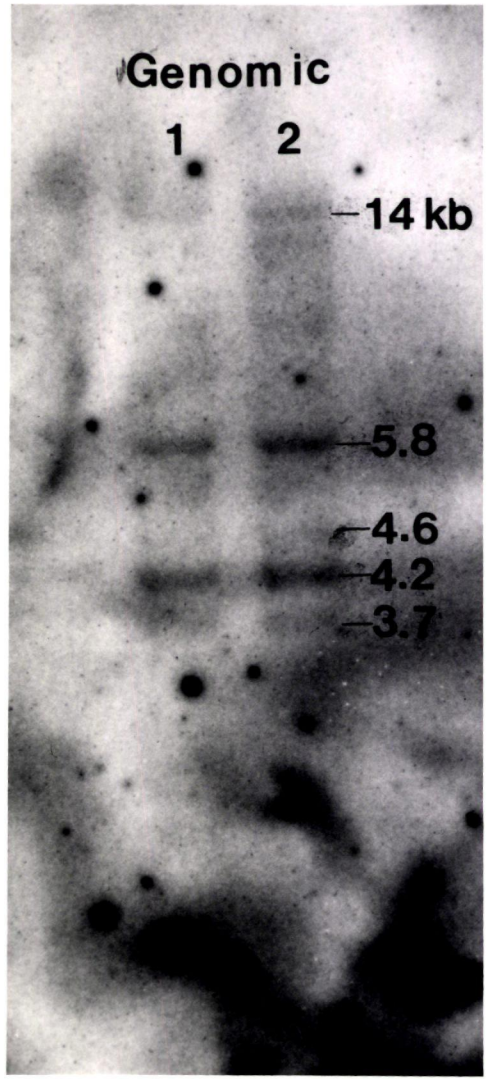

Figure 3. Southern blots of Eco RI digested peripheral blood lymphocyte (PBL) genomic DNA (10 ug) screened with the oligonoucleotide probe at a hybridizing and washing (in 2X SSC) temperature of 43°C. Sizes of the five hybridizing fragments are marked in kilobases (kb). Tracks 1 and 2 are from the same individual.

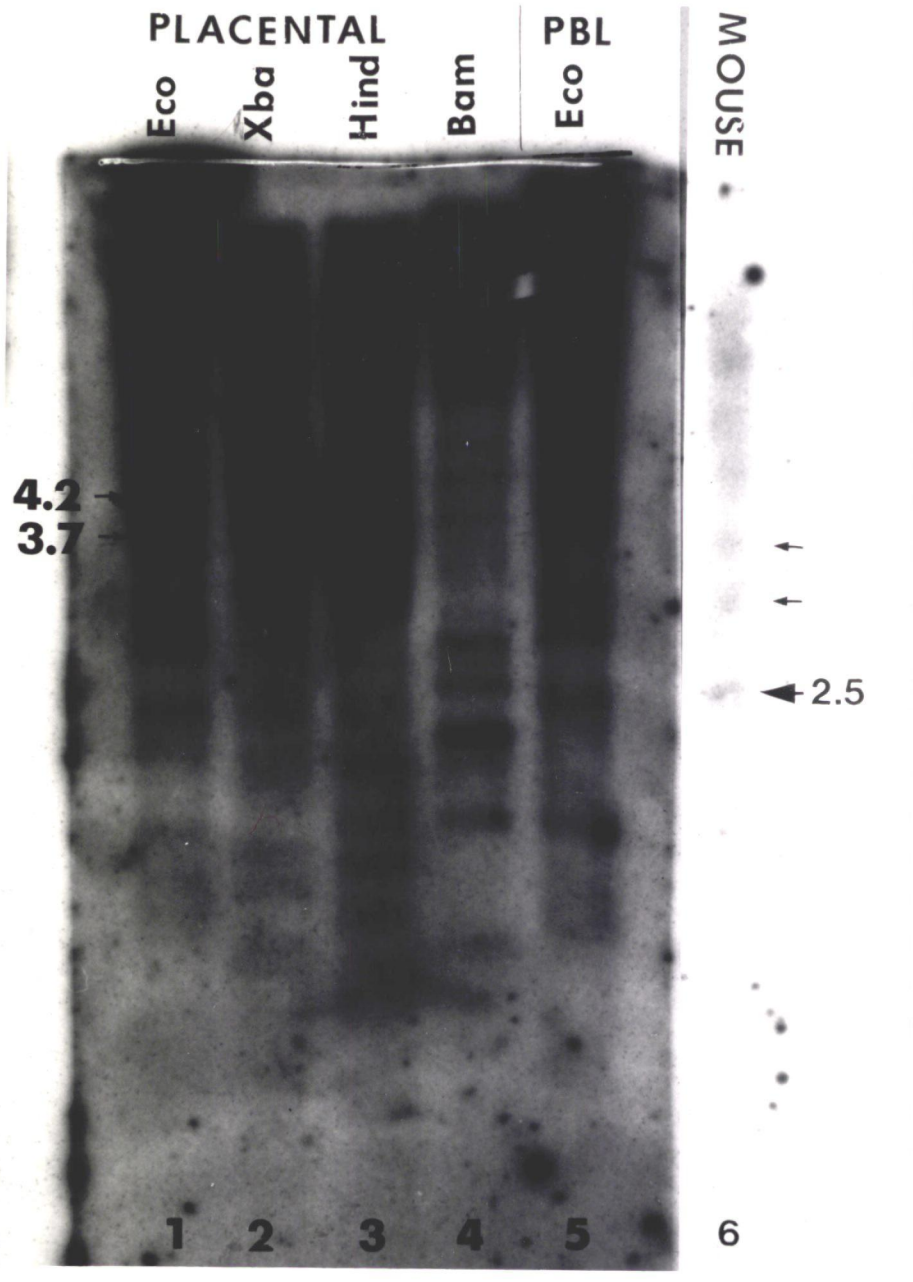

Figure 4. Placental, PBL and mouse LMTK$^-$ genomic DNA (10 ug) cut with different restriction enzymes, Southern blotted and probed with the oligonucleotide at a lower stringency (hybridization at 43°C. and washing at 25°C., 2X SSC). Tracks 1-4 are placental genomic DNA digested with Eco Xba I, Hind III and Bam HI. PBL genomic DNA is shown in track 5 to compare to that of placental DNA. Mouse LMTK$^-$ DNA cut with Eco RI is shown in track 6. Note the presence of the 2500 bp hybridizing band.

DNA. Placental DNA cut with Xba I, Hind III and Bam HI also showed multiple hybridizing bands when probed with the oligonucleotide under low stringency conditions. Although not clear in figure 4, there are bands corresponding to the expected IL-2 3600 bp Hind III fragment and the 1000 bp Xba I fragment. Southern blots of murine LMTK$^-$ genomic DNA cut with Eco RI and probed with the synthetic probe at a lower stringency showed the presence of a 2500 bp hybridizing fragment along with less intense signals. This was not an unexpected result due to the homology between the human and mouse IL-2 molecules. Under the low stringency conditions this band may also represent background hybridization.

Probing of the Cosmid Pools

A human placental cosmid library was screened with the oligonucleotide probe to try and isolate clones of some of the other hybridizing fragments seen in the genomic screens. When the twelve Eco RI digested cosmid pools were screened at 43°C., hybridizing fragments 4.2 kb in length were seen in cosmid pools 7 and 11 (figure 5). Although there was some background hybridization and faint bands at 4.6 and 5.8 kb, no other intense signals were seen in any of the other cosmid pools. The signal from pool 7 was significantly less intense than that of pool 11 indicating a decreased representation of that fragment in that pool.

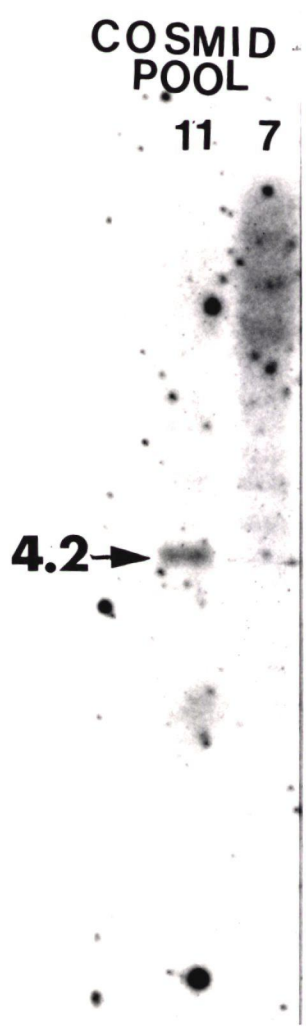

Figure 5. A Southern transfer of cosmid pool 7 and 11 DNA (1 ug), cut with Eco RI and probed with the oligonucleotide at a hybridization and washing temperature of 43°C. Note the strong 4.2 kb hybridizing fragment present in pool 11 and the weak signal in pool 7.

Cosmid pool 11 was chosen for colony screening.

The pool was grown up, appropriately diluted and plated onto six large LB agar plates such that there was approximately 5000 colonies per plate. This represented the whole pool which contained about 20,000-30,000 different clones. Nitrocellulose lifts were screened at 43°C. with the oligonucleotide and positive colonies transfered in duplicate to smaller LB agar plates. The nitrocellulose filters from the smaller plates, when screened under the same high stringency conditions, showed the presence of two hybridizing colonies at positions 46 and 72 on the plate (figure 6). The two colonies were picked for cosmid DNA preparations.

Restriction Mapping of the Positive Clones

The two cosmid DNA clones were digested with the restriction enzymes Eco RI, Bam HI, Hind III and Xba I and electrophoresed on an agarose gel. After Southern transfer and hybridization to the synthetic probe the autoradiogram (figure 7) showed that 46 and 72 were the same clone. It was also observed that the clones contained a 4.2 kb Eco RI fragment, the same size observed in the cosmid pool and genomic screens. Cosmid clone 46 was chosen for analysis.

Detail of the region surrounding the hybridizing segment was obtained by digesting the cosmid with Eco RI,

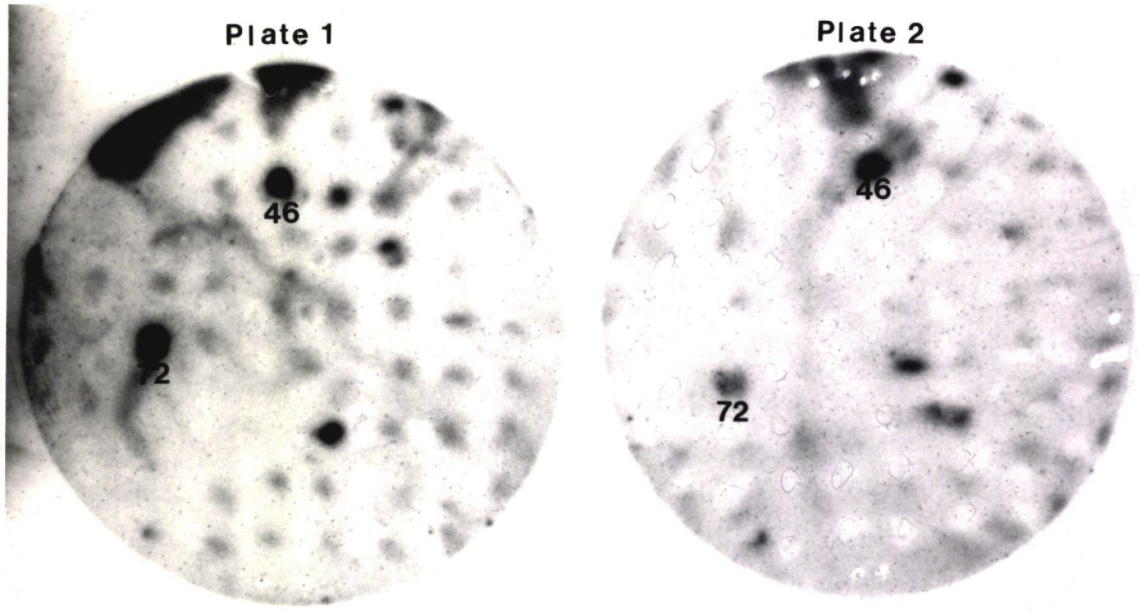

Figure 6. Autoradiograms of nitrocellulose lifts from the duplicate plates of positive colonies picked from the primary screen of cosmid pool 11. After probing with the oligonucleotide at 43°C. the colonies at positions 46 and 72 were positive on both plates and subsequently picked for restriction mapping of the clones.

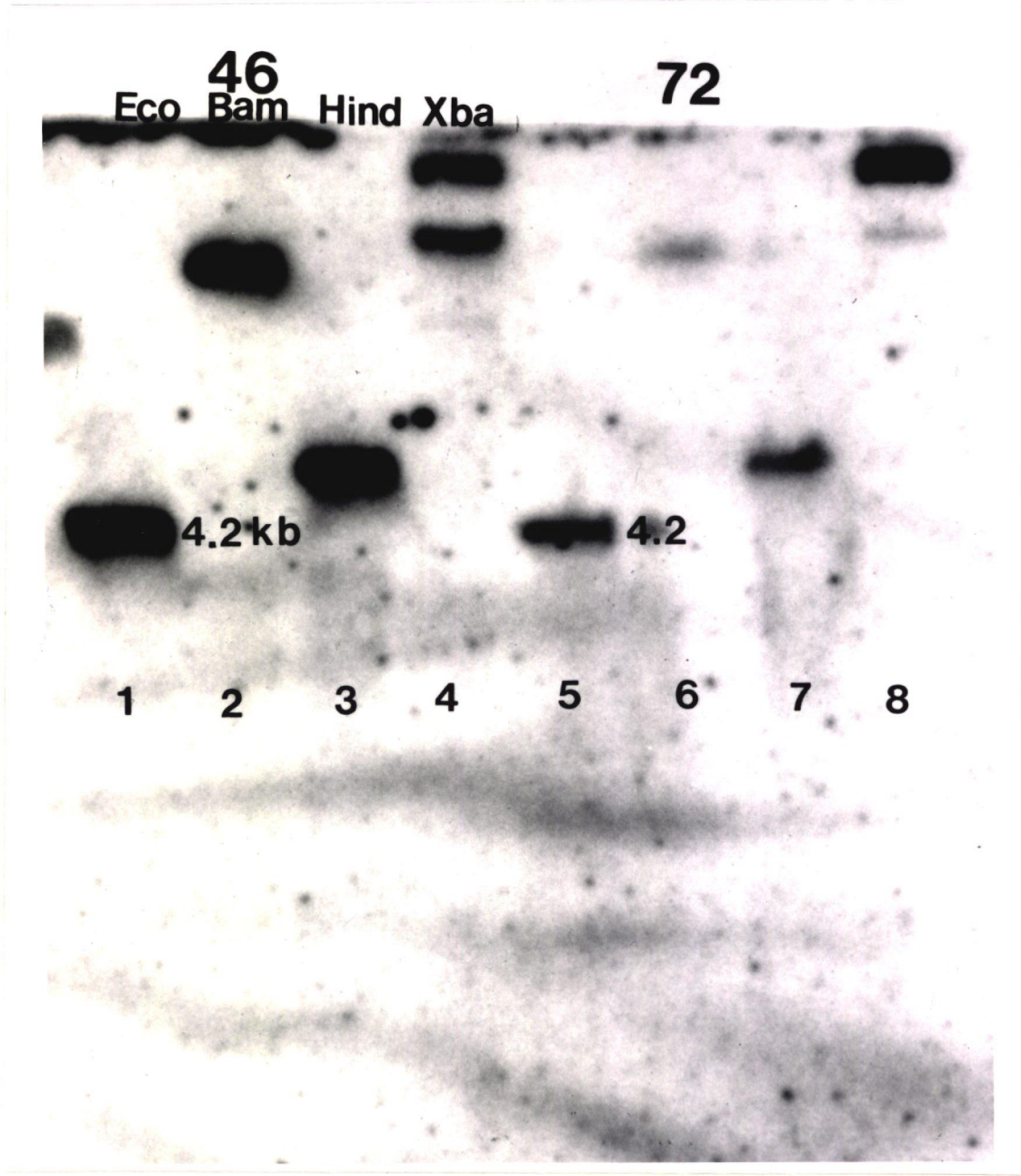

Figure 7. Digestion of the positive cosmid pool 11 clones 46 and 72 with Eco RI, Bam HI, Hind III and Xba I (tracks 1-4, 5-8). The autoradiogram shows identical hybridization patterns for the two clones indicating that they are the same clone. The difference in intensity between the two clones is due to the amount of DNA loaded onto the gel and the completness of the individual digestion reactions.

Hinc II, Hind III, Xba I, Pst I and Bam HI. The Southern blots were probed with the oligonucleotide and set up for autoradiography. A partial restriction map is shown in figure 8. The probe hybridized to a region of 350 bp delineated by Hinc II and Pst I. Besides the 4.2 Eco RI fragment the probe hybridized within a 5.0 kb Hind III fragment and a 14 kb Bam HI fragment.

Backprobing of Genomic and Cosmid Pool DNA

The 4.2 kb Eco RI fragment was subcloned into pUC 8 (figure 9) and the isolated insert used to probe genomic DNA and the cosmid pool library. A Southern blot of human genomic DNA from different individuals (placental and PBL sources) cut with Eco RI, Hind III and Bam HI was probed with the 4.2 kb probe (figure 10). As expected a 4.2 kb Eco RI, 5.0 kb Hind III and a 14 kb Bam HI fragment lit up with the probe. Across four individuals no difference was noted in the Eco RI fragment indicating an absence of a restriction fragment length polymorphism (RFLP) within these individuals. Extensive searches for RFLP's were not undertaken with this probe.

The Eco RI cut cosmid pools were also screened with the 4.2 kb probe. Signals at 4.2 kb were seen in pools 7, 11 and the lane of human genomic DNA which served as a control (figure 11). A second weakly hybridizing genomic

RESTRICTION MAP OF COSMID POOL 11 CLONE #46

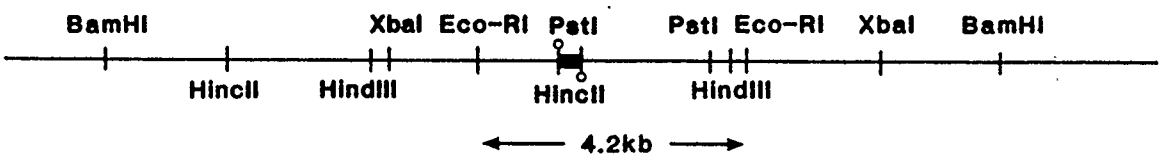

Figure 8. Restriction map of the cosmid pool 11 clone 46, surrounding the region hybridizing to the probe. The shaded section is the 350 bp region hybridizing to the probe. One centimeter is approximately equal to 1.2 kb.

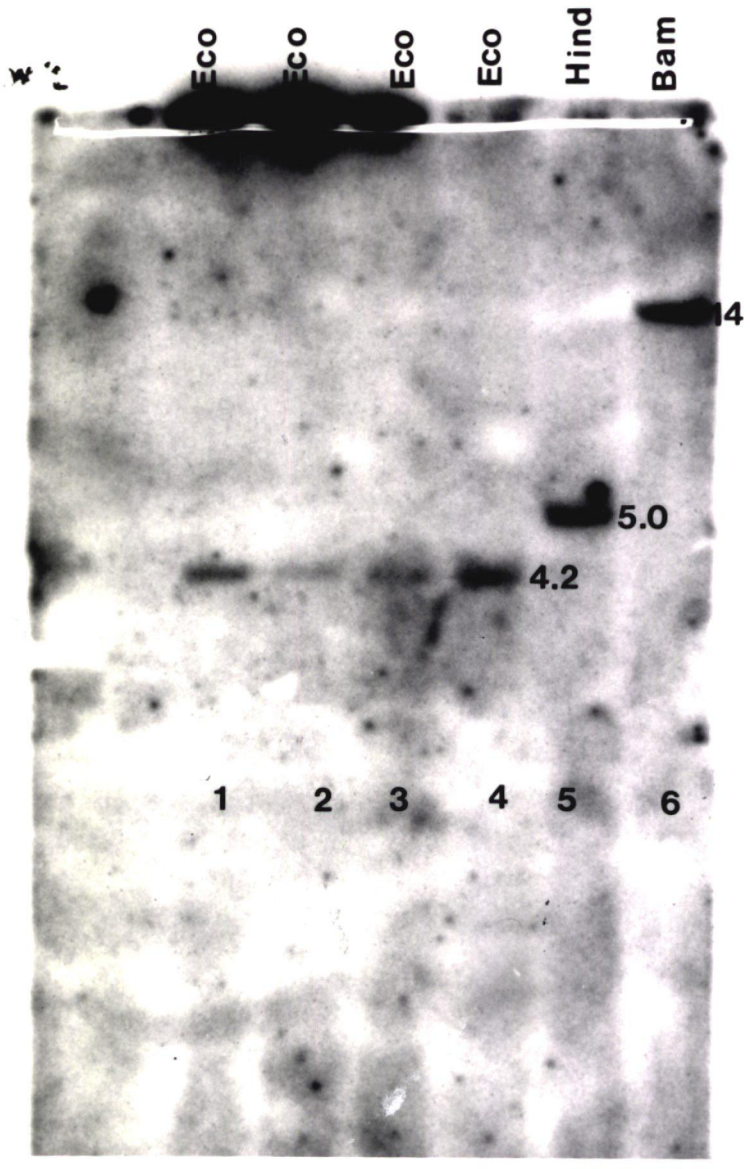

Figure 10. A Southern blot of human genomic DNA (10 ug) screened with the 4.2 kb Eco RI fragment. Tracks 1-4 are genomic DNA from different individuals (both placental and PBL sources). Tracks 5 and 6 are genomic DNA cut with Hind III and Bam HI respectively.

RESTRICTION MAP OF 4.2kb INSERT IN pUC8

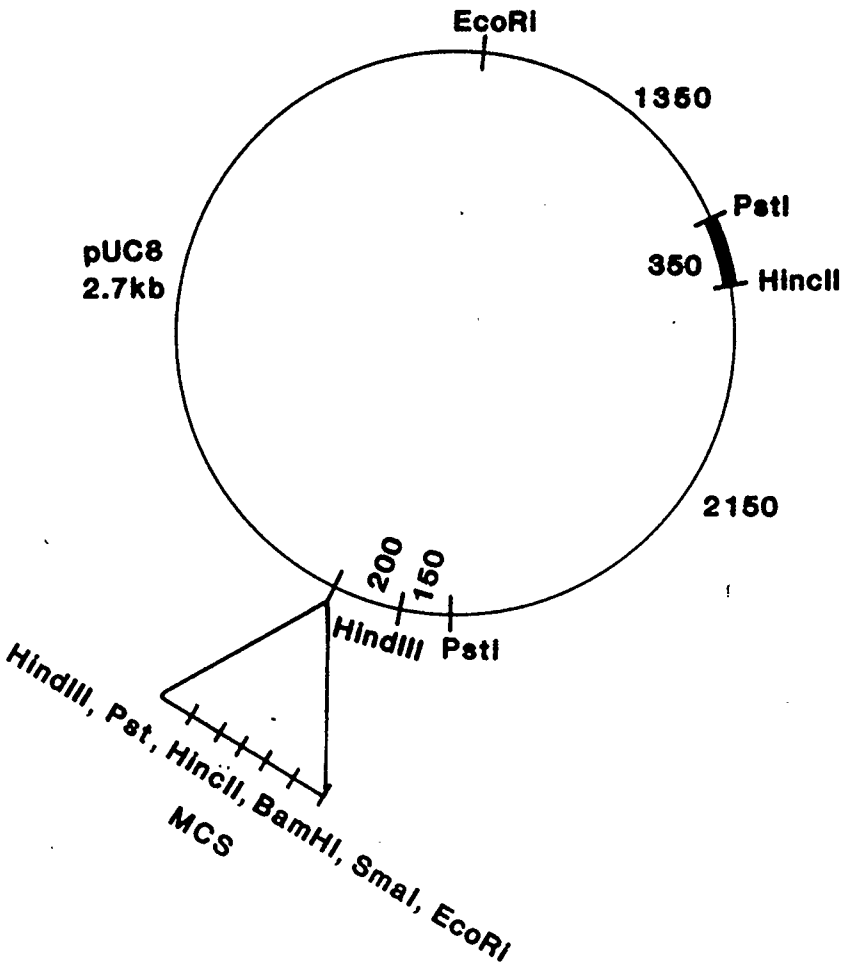

Figure 9. Restriction map of the 4.2 Eco RI restriction fragment subcloned into pUC 8. Sizes are in base pairs unless otherwise marked. The 350 bp region delineated by the probe is also shown. MCS; multiple cloning site.

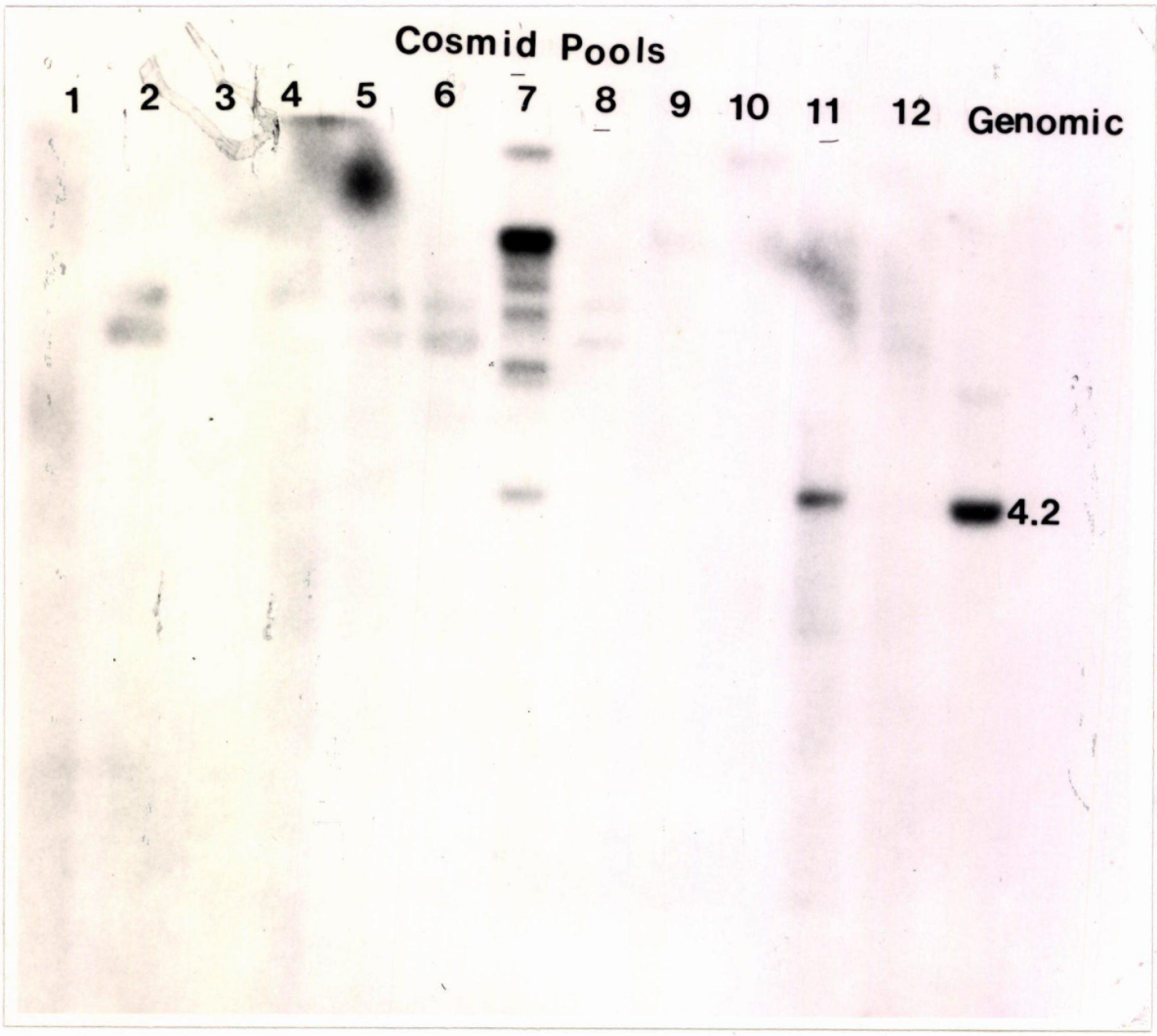

Figure 11. Southern blots of the twelve cosmid pools (1 ug) and genomic DNA (10 ug) cut with Eco RI and probed with the 4.2 kb Eco RI fragment isolated from cosmid clone 46. Note the presence of the strong hybridizing bands 4.2 kb in length in cosmid pools 7 and 11 and human genomic DNA.

signal was also seen that may be represented in pool 7 as well. However pool 7 contained a partial digest as larger fragments within the pool also hybridized to the probe.

The hybridizing 350 bp Hinc II/Pst I fragment was isolated from the 4.2 kb fragment and subcloned into pUC 12.

Isolation of the 350 bp HincII/Pst I Restriction Fragment

The 4.2 kb Eco RI subclone when cut with Hinc II and Pst I releases two identical fragments of about 350 bp in length. One arises from the hybridizing region and the other due to the Hinc II site within the polylinker of pUC 8.To differentiate between the two fragments a triple digest of the 4.2 kb subclone was done using Hind III, Hinc II and Pst I. Hind III cut the non-hybridizing Hinc II/Pst I fragment in half leaving the hybridizing fragment intact.

The 350 bp Hinc II/Pst I hybridizing insert was then ligated to pUC 12 which had been digested with Pst I and Sma I. Sma I was chosen because the Hinc II site was too close to the Pst I site within the polylinker. Both Hinc II and Sma I leave blunt ended restriction fragments. A map of the 350 bp insert within pUC 12 is shown in figure 12.

Ligation of Sma I site to the Hinc II site destroys both sites leaving the next intact restriction site within the polylinker as Sst I. The M13 universal primer anneals

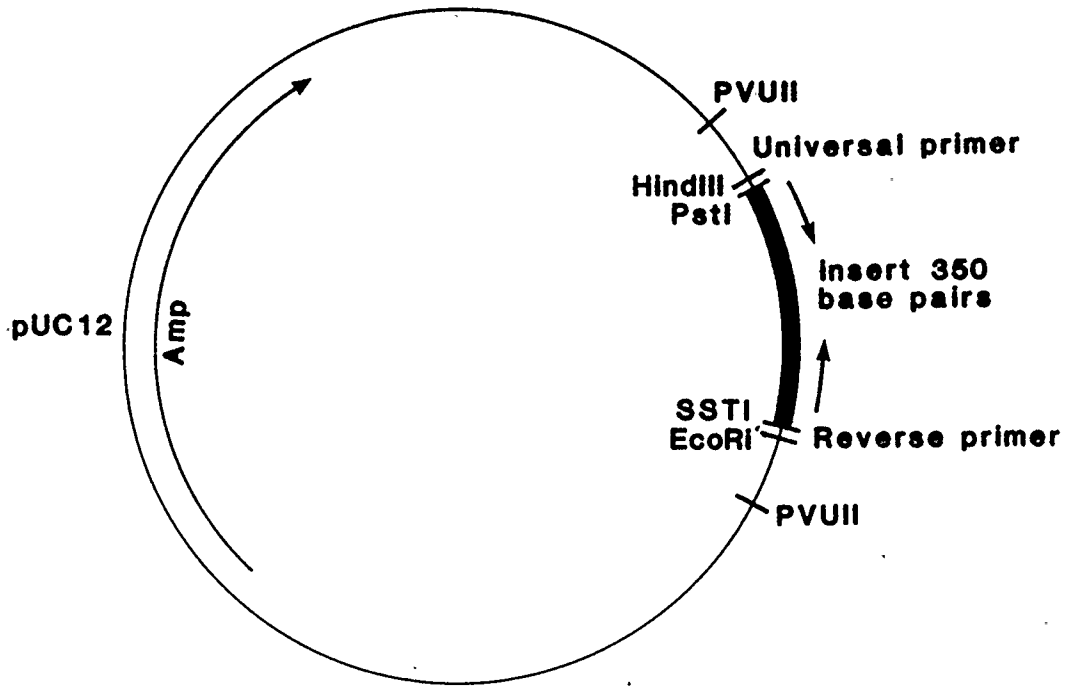

Figure 12. Restriction map of the 350 bp Hinc II/Pst I frament subcloned into pUC 12. The location of the M13 sequencing primers is shown.

upstream from the Hind III site and allows sequencing through it and the Pst I site. The M13 reverse primer allows sequencing of the opposite strand through the Eco RI site.

Sequence Analysis

The dideoxy chain termination method of Sanger (Sanger and Nicklen, 1977) was employed to sequence the 350 bp insert in both directions. Using the M13 universal (figure 13) and reverse (figure 14) primers, approximately 250-300 bases were read in each direction which allowed for confirmation of most of the sequence (figure 15). The known sequences of the pUC 12 multiple cloning site when sequenced helped in orientation. The sequence shown in figure 15 was derived using the universal primer sequencing through the pUC 12 Hind III and Pst I sites. The Eco RI site at the end of the inserted sequence is also shown.

No clear cut homology to the probe is seen. However if a guanosine residue is deleted from the oligonucleotide probe sequence (or conversely inserted into the isolated sequence) the probe can be made homologous to a region starting at position 286 of the sequence (figure 15). If the sequences are aligned such that the missed base is accounted for, 20 out 25 bases hybridize to the probe (figure 16).

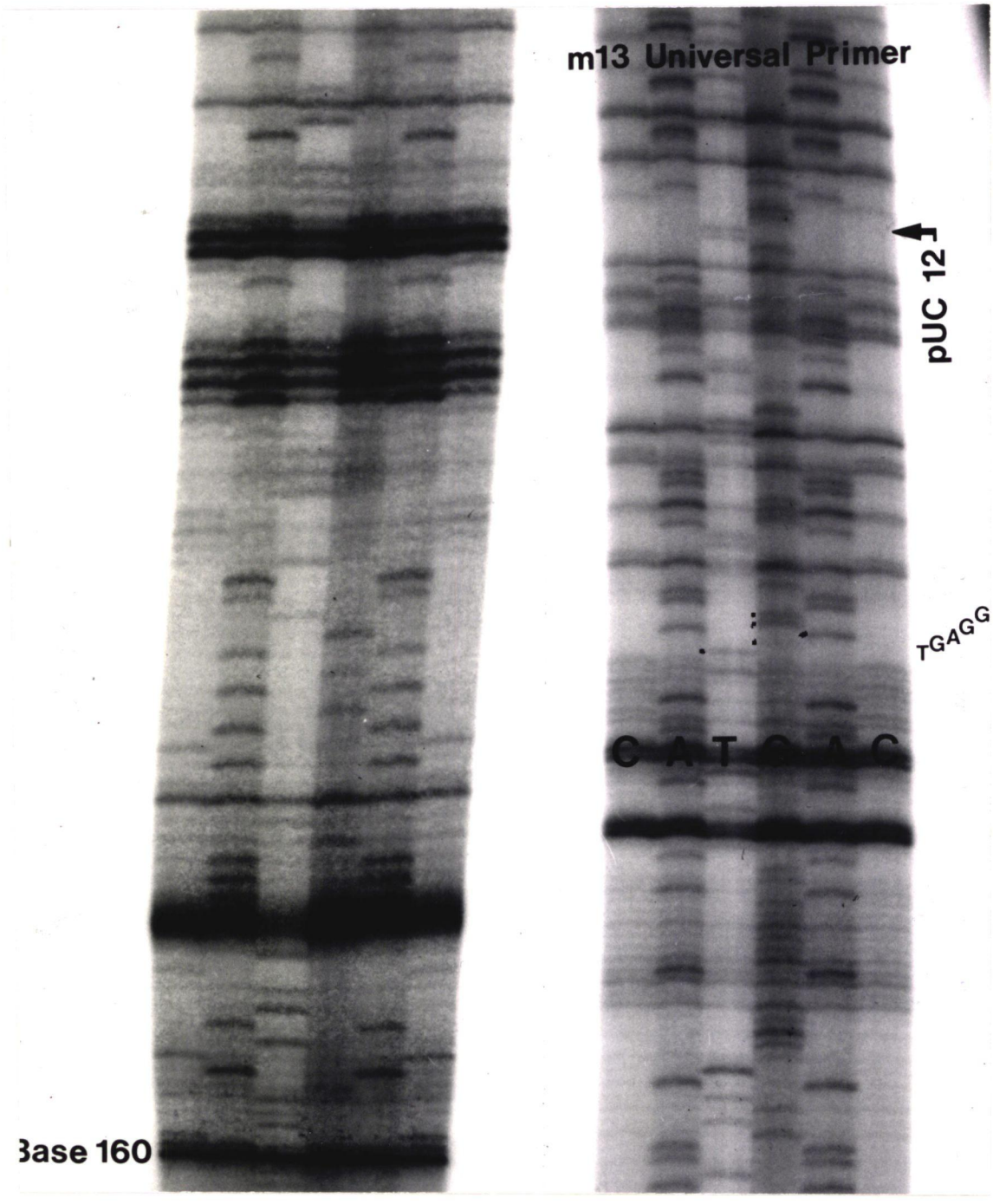

Figure 13. A portion of DNA sequencing gel of the 350 bp fragment showing the region thought to be hybridizing to the probe. The figure is the strand sequenced by the universal M13 primer starting at base 160 and running up to the Eco RI site of pUC 12. The area highlighted and lettered shows the ambiguous region that contains only one guanosine residue insted of the expected two.

Figure 14. The DNA sequencing gel showing the oppposite strand of the 350 bp fragment as sequenced by the reverse primer in the direction of Eco RI to Pst I. The area lettered again shows the region which hybridizes to the oligonucleotide probe and clearly shows the presence of only one C.

5'

10	20	30	40	50	60
GGACCCCTGG	TCTGTCCATG	TGTGTGTATA	GAAAGTCCTG	GTGAGTTTGC	TGGTGTGTGC
70	80	90	100	110	120
CTCTGTTCAA	ATCCCGAGGG	CCTAATACTG	TCCGTGGAGG	TGCCTTTCTT	CCTAGGTTAG
130	140	150	160	170	180
GGCTCTTCTG	GGTTAGCCTT	CGTATCTCTT	TGGCAATTCC	TTTGTTTCCT	TTGACTATTC
190	200	210	220	230	240
TCAAAAGTGA	ACAGAGAGTA	ATCCCTTTGC	AGGTGTGACT	TCTTTGATTT	GAGCCAAAAT
250	260	270	280	290	300
GTTAGGAACA	TAAGTGTATG	GGGGCGAGGG	GGGGGGATTC	CTGATTCCAA	CACCTCTGAG
310	320	330	340	-1	-10
GAAGTGCTTA	GCAGAAGCCT	GCTGGGTATT	AGAACATCCA	GGGTTGGGAA	TTCGTAA

(pUC 12 Eco R1 site)

3'

Figure 15. DNA sequence of the 350 bp HincII/Pst I insert as derived from the universal primer. The area homologous to the probe is underlined. The negative numbers indicated the start of the plasmid sequence.

```
                  5'                                3'
350 bp fragment- TCCAACACCTCT GAGGAAGTGCTT
IL-2 Probe     - ACTCAAACCTCTGGAGGAAGTGCTA
                   *  * ****** ***********
                  5'                                3'
```

Figure 16. Comparison of the region from the isolated sequence thought to be hybridizing to the IL-2 oligonculeotide probe. The guanosine residue is not present within the isolated sequence.

Comparing the oligonucleotide sequence to both strands of the 350 bp sequence reveals no other significant homology. The probe is 60% homologous (15/25) to other regions within the sequence but this is not significant because it is also 60% homologous to regions within the pUC plasmids. Under the stringent conditions used the probe was found not to hybridize to any plasmid sequences hence the only significant homology is that in figure 16. The region thought to be hybridizing to the probe was clearly evident on both sequencing gels, especially with respect to the single gaunosine residue at position 298.

homology to Interleukin-2 and Other Sequences

When the isolated sequence was compared to the published IL-2 sequence the only significant homologous region is that which is thought to be hybridizing to the probe. Dot matrix analysis using the program of Lagrimini

(1983) showed no other significant homology between the 350 bp sequence and the entire interleukin-2 chromosomal gene.

The 350 bp sequence was transcribed in the reading frame that gave amino acids corresponding to that of IL-2 for the hybridizing region. If a guanosine residue was inserted into the isolated sequence, six amino acids were the same as interleukin-2. Other than that no other amino acids corresponded between the two sequences.

In order to detect possible homologies between the 350 bp sequence and other known DNA sequences or proteins, both the Microgenie and the NBRF data banks were searched. No significant homolgies were evident from either search.

Northern Analysis

Total cellular RNA, poly(A^+) and (A^-), from two human lymphoblastoid cell lines (CESS and Jurkat) was probed using the oligonucleotide and the 4.2 kb subclone. Approximately 15 ug of total RNA from each cell line was gel electrophoresed and transfered to nitrocellulose. No significant hybridization was detected with either probe, indicating no evidence of a transcription product.

The presence of nucleic acid on the nitrocellulose was detected by post hybridization staining with methylene blue. The 28S and 18S ribosomal RNA's served as markers and positive controls.

The oligonucleotide showed some hybridization with the CESS B cell RNA in the region of the ribosomal RNA's. This probably represents background hybridization and is not significant. To determine this poly (A^+)-enriched mRNA's could be isolated and rescreened with the oligonucleotide probe.

Sequencing of the Oligonucleotide Probe

To verify that the IL-2 oligonucleotide probe was synthesized correctly, it was sequenced using a modification of the Maxam-Gilbert technique. The method allowed all 25 bases of the probe to be sequenced on a single gel. It was clear from the autoradiogram that the sequence that was synthesized was indeed the one that corresponded to the exon of Il-2. It was especially clear that in the region of contention there was not a guanosine residue missing from the synthesized sequence (figure 17). This would have explained the aberrant duplex formation.

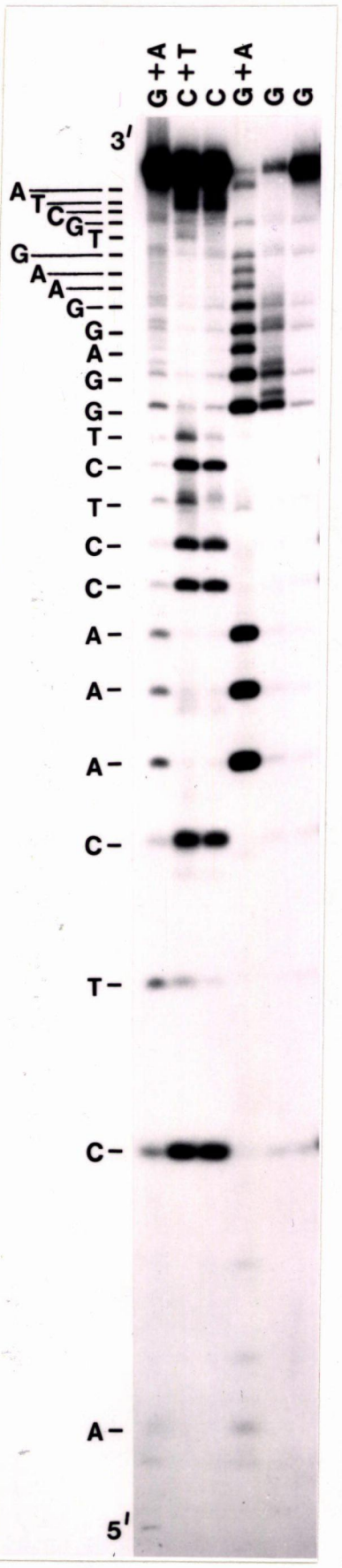

Figure 17. Maxam-Gilbert sequencing gel of the synthetic oligonucleotide showing clearly that the synthesized sequence is correct.

Discussion

When the synthetic oligonucleotide, made to isolate the cloned sequence for IL-2, was used to probe human Eco R1 cut genomic Southerns multiple bands appeared on the autoradiograph. The fragments ranging in size from 3.7 kb to 14 kb had consistent, but varying, hybridization intensities. A weakly hybridizing 3.7 kb fragment was thought to be the reported IL-2 restriction fragment. The strongly hybridizing 4.2 kb fragment as well as the others were thought to be genes related to IL-2 since IL-2 is part of a large class of molecules called lymphokines.

A human cosmid library was screened in an attempt to isolate some of these possibly related sequences. A strong band of 4.2 kb was seen in cosmid pool 11. That pool was chosen for colony screening and two identical clones containing the 4.2 kb Eco RI fragment were isolated. The hybridizing fragment was then subsequently subcloned into a plasmid vector and restriction mapped. The probe was found to hybridize to a 350 bp region delineated by Hinc II and Pst I. This region was subcloned and sequenced in both directions.

The sequence identified only one possible hybridizing segment. This region was 80% homologous to the probe only if a gap of one base was introduced in the middle of the genomic sequence. The hybridizing region was unambiguous in both sequencing directions and the experiments were repeated to confirm the sequence.

The possible explanation that there was an error in the synthesis of the probe sequence was ruled out by Maxam-Gilbert sequencing of the oligonucleotide. The results verified that the probe sequence was synthesized correctly.

A possible explanation for the mismatch between the IL-2 probe and the isolated sequence is that the guanosine residue bulges out when the duplex is formed. The mismatch within the middle of the oligonucleotide would confer instability on the duplex.

From the published literature on oligonucleotide hybridization, it was found that mismatches decrease the stability of the duplex formation (Lathe, 1985). For an oligonucleotide of 25 bases a hybridization and washing temperature of 43°C. should select for homolgies of only 85% or greater (Lathe, 1985). This allows for only four or possibly five mismatches. If the guanosine residue does loop out, 19 out of the 25 bases would hybridize with each other. The unpaired bases would include five single base pair mismatches and one unpaired base bulging out. It might be expected that the unpaired base may contribute more

instability to the duplex than that of a single base pair mismatch. In theory it would seem unlikely that given the stringency of the conditions used, the duplex would be allowed to form.

However the recommended stringent wash temperature (t_w) as calculated by Lathe is only an empirical approximation. The true value very much depends on indivdual circumstances. The destablizing effect of mismatched base pairs depends on, among other things, where the mismatch occurs, the G/C content of the probe, and what bases are involved in the mismatch. As quoted by Lathe, "the determination of stringent wash temperature is given as a guideline and is not intended to provide a final solution to the most appropriate temperature and salt conditions for hybridization".

As a note for future considerations, the t_m for oligonucleotide duplex formation can be better approximated by hybridization in the presence of tetramethylammonium chloride (Wood et al., 1985). This eliminates the differential stability between G/C and A/T base pairs and subsequently eliminates the variable of percent G/C.

In view of the margin of error of the published technique, it is likely that the mispairing of the IL-2 probe and the 4.2 kb fragment would be allowed to occur. It is worth noting that the position of the extraneous base within the duplex is an important factor in its stability. There is a matched stretch of 11 residues before and seven

residues after the suspect guanosine which would help in keeping the oligonucleotide bound.

The problem of mispairing during oligonucleotide hybridizations due to internal residues looping out has been observed before (Panabieres et al., 1982). It was reported that a 13 mer oligonucleotide designed to specifically initiate on a lactic dehydrogenase mRNA also primed with high efficiency the synthesis of DNA corresponding to the mitochondrial 16S ribosomal RNA. This could only be explained by the bulging out of a single base within the oligonucleotide probe bound to the 16S ribosomal RNA.

This explanation, however, still leaves some fundamental questions unanswered. Perhaps the most important one is the difference in intensity between the 3.7 and 4.2 kb bands seen when the oligonucleotide was used to probe human genomic blots. If the probe is in fact an IL-2 probe, why isn't the supposed 3.7 kb IL-2 band of equal or greater intensity than that of the 4.2 kb band? This is especially puzzling considering the mispairing found within the hybridizing region of the 4.2 kb band.

Another inconsistency is the absence of the IL-2 band when screening the placentally derived cosmid pools. There is the possibility that the IL-2 gene is not represented within any of the 12 cosmid pools. Holbrook et al. (1984) was unable to isolate the IL-2 gene from a placental library using a much larger cDNA probe. Perhaps the IL-2

gene within placental DNA does not lend itself to be easily cloned.

Northern analysis of human T cell RNA from the Jurkatt line, a known IL-2 producer, failed to reveal the presence of an IL-2 transcript when probed with the oligonucleotide probe. This result is not conclusive as the cell cultures were not mitogen stimulated before the RNA was isolated. It has been shown that IL-2 is only produced in quantitative amounts if the cells are first stimulated with a lectin such as concanavalin A (Gillis et al., 1978). The level of expression of the IL-2 gene in the unstimulated cells may have been below the level of sensitivity of the assay.

The isolation of genomic clones using oligonucleotides

This project has brought to light some important considerations regarding the isolation of genomic clones using oligonucleotide probes. It is evident from the results that the probe may not have been long enough and the stringency of the conditions high enough to rule out incorrect hybrids. Due to the complexity of the human genome perhaps a 19 mer, as stated earlier, is not long enough to recognize unique sequences. This is especially true if a stable duplex is allowed to form that contains a single nucleotide bulging out.

The problem of picking up unwanted signals can be approached in one of two possible ways. First, the

oligonucleotide could be used to first screen a cDNA library derived from a high IL-2 producing cell line. The complexity of a cDNA library is significantly less than that of a genomic library and thus so is the chance of picking up aberrent clones. The isolated cDNA could then be used to screen a genomic library to isolate genomic clones.

In fact this method was used by Holbrook et al. (1984). They synthesized an 18 mer oligonucleotide probe to isolate an IL-2 cDNA from a Jurkatt cDNA library. The resulting cDNA clone was then used to screen a genomic library and they identified three overlapping IL-2 clones containing the entire IL-2 gene.

The other solution is to make the oligonucleotide longer and use higher stringency conditions. This method is particularily appropriate when a high product producing cell line or tissue is not available. Wood et al. (1984) used this technique to isolate the human factor VIII gene from an X-chromosome enriched genomic library. A 36 mer oligonucleotide was used to hybridize against the genomic library under high stringency washing conditions of 1 X SSC and 46°C. They isolated 15 clones representing eight distinct overlapping recombinants. On genomic blot hybridization they observed only a single Eco RI or Bam HI band. Thus the conditions they used were stringent enough to identify only the unique fragments they were interested in.

Identification of related genes with oligonucleotides

The original hypothesis that the 4.2 kb fragment may have been related to the IL-2 sequence has been shown by this project to be incorrect. The concept, however, that related sequences may have among themselves homologous conserved regions is still valid. It is entirely possible that one of the other signals picked up on the oligonucleotide probe of the genomic blots may be related to IL-2. Admittedly it would be very serendipitous to choose a region at random that is conserved between a family of molecules. The logical approach would be one of the opposite tack.

If there is evidence that a family of related genes existed and that they shared a conserved region, then it would be reasonable to postulate that an oligonucleotide could be used to identify other sequences within the family. There are many examples of gene families present in the literature. They may be functionally or structurally related as with the globin genes or come from a common cell of origin as is the case with lymphokines.

There are also examples of gene families which fulfill the second criteria of a shared conserved region. If protiens are organized into functional domains then it

is possible that these functional units are conserved among molecules with similar function. A good example exists within the genes of the human major histocompatibility complex (Das et al., 1982). Sequence analysis shows a close relationship between intron-exon organization and putative protein domains. This relationship between structural organization extends down to the level of the amino acid sequence. Other examples of conservation of functional domains are present within the globin genes, collagen genes and the human clotting factors.

Closer to the scope of this project, no evidence, other than their functional association, has been presented to show that IL-2 belongs to a individual family of molecules. However work on IL-1 indicates that there is a family of interleukin-1 proteins with identical amino acids tending to cluster in particular regions (Marx, 1985). So within the lymphokine group of molecules there is an example of functionally related genes that may also be related on the molecular level. In retrospect, perhaps we should have initiated the project with an oligonucleotide homologous to an IL-1 molecule.

This project illustrated some of the important problems that can be encountered when using oligonucleotides as probes. To eliminate the isolation of erroneous sequences, hybridization conditions and probe length must be carefully considered in relation to the complexity of the libraries to be screened.

REFERENCES

Applied Biosystems Inc. (1983) Purification of crude synthetic oligonucleotides by preparative gel electrophoresis, Feb.1.

Applied Biosystems Inc. (1984) The evaluation and purification of synthetic oligonucleotides, Issue 13, Nov.9.

Auron, P. E., Webb, A. C., Rosenwasser, L. J., Mucci, S. F., Rich, A., Wolff, S. M. and Dinarello, (1984) Nucleotide sequence of human monocyte interleukin 1 precursor cDNA. Proc. Natl. Acad. Sci. USA, 81:7907-7911.

Beaucage, S. L. and Caruthers, M. H. (1981) . Tetrahedron Letters, 22:1859-1862.

Biggin, M. D., Gibson, T. J. and Hong, G. F. (1983) Buffer gradient gels and 35S label as an aid to rapid DNA sequence determination. PNAS, 80:3963-3965.

Cami, B. and Kourilsky, P. (1978) Screening of cloned recombinant DNA in bacteria by in situ colony hybridization. Nuc. Acids. Res., 5:2381-2390.

Cantrell, D. A. and Smith, K. A. (1983) Transient expression of interleukin 2 receptors, consequences for T cell growth. J. Exp. Med., 158:1895-1911.

Clewell, D. E. and Helinski, D. R. (1970) Properties of a supercoiled deoxyribonucleic acid-protein relaxation complex and strand specificity of the relaxation event. **Biochemistry**, 9:4428-4440.

Conner, B. J., Reyes, A. H., Morin, C., Itakura, K., Teplitz, R. L. and Wallace, R. B. (1983) Detection of sickle cell B-globin alleles by hybridization with synthetic oligonucleotides. **Proc. Natl. Acad. Sci. USA**, 80:278-282.

Das, H.K., Biro, S., Cohen, H., Erlich, H., VonGrabam, V., Lawrance, P., Lemaux, P., McDivitt, H.O., Shulz, M.L., Sood, A., Weissman, S.M. (1982) Use of synthetic probes complementary to human HLA-DR alpha and beta as extension primers for the isolation of 5' specific genomic clones., PNAS,79:1893-1897

Degrave, W., Tavernier, J., Duerinck, F., Devos, R. and Fiers, W. (1983) Cloning and structure of the human interleukin-2 gene. **EMBO Journal**, 2:2349-2353.

Devos, R., Plaetinck, G., Cheroutre, H., Simons, G., Degrave, W., Tavernier, J. and Fiers, W. (1983) Molecular cloning of human Il-2 cDNA and its expression in E. coli. **Nuc. Acids Res.**, 11:4307-4323.

Evans, C. H., Heinbaygh, J. A. and Dipaolo, J. (1983) Comparative effectiveness of lymphotoxin anticarcinogenic and tumor cell growth inhibitory activities. **Cellular. Immunol.**, 76:295-303.

Feinberg, A. P. and Vogelstein, B. (1983) A technique for

radiolabeling DNA restriction endonuclease fragments to a high specific activity. **Analytical Biochemistry**, 132:6-13.

Feldmann, M. (1985) Lymphokines and interleukins emerge from the primeval soup. **Nature**, 313:351-353.

Friedman, R. M. and Vogel, R. S. (1983) Interferon, with special emphasis on the immune system. **Adv. Immunology**, 34:97-140.

Fujita, T., Takaoka, C., Matsui, H. and Taniguchi, T. (1983) Structure of the human Il-2 gene. **Proc. Natl. Acad. Sci. USA**, 80:7437-7441.

Gillis, S., Ferm, M., Ou, W. and Smith, K. A. (1978) T cell growth factor: parameters of production and a quantitative microassay for activity. **J. Immunology**, 120:2027-2032.

Gillis, S., Mochizuki, P., Conlon, P., Hefeneider, S., Henney, C. and Watson, J. D. (1982) Molecular characterization of Il-2. **Immunol. Rev.**, 12:450-502.

Gillis, S. and Smith, K. A. (1977) Long term culture of tumour specific cytotoxic T cells. **Nature**, 268:154-156.

Godson, G. N. and Vapnek, D. (1973) A simple method of preparing large amounts of 174 RFI super-coiled DNA. **Biochim. Biophys. Acta**, 299:516-521.

Gray, P. W., Aggarwal, B. B., Benton, C. V., Bringman, T. S., Henzel, W. J., Jarrett, J. A., Leung, D. W., Moffat, B., Ng, P., Svedersky, L. P., Palladino, M. A.

and Nedwin, G. E. (1984) Cloning and expression of cDNA for human lymphotoxin, a lymphokine with tumour necrosis activity. **Nature**, 312:721-724.

Gray, P. W. and Goeddel, D. V. (1982) Structure of the human immune interferon gene. **Nature**, 298:859-860.

Gray, P. W., Leung, D. W., Pennica, D., Yelverton, E., Najarian, R., Simonsen, C. C., Derynick, R., Sherwood, P., Wallace, D. M., Berger, S. L., Levinson, A. D. and Goeddel, D. V. (1982) Expression of human immune interferon cDNA in E. coli and monkey cells. **Nature**, 295:503-506.

Grollman, E. F., Lee, G., Ramos, S., Lazo, P. S., Kaback, R., Friedman, R. M. and Kohn, L. D. (1978) Relationships of the structure and function of the interferon receptor to hormone receptors and establishment of the antiviral state. **Cancer Research**, 38:4172-4185.

Grosveld, F., Dahl, H., deBoer, E. and Flavell, R. A. (1981) Isolation of B-globin from a cosmid library. **Gene**, 13:227-237.

Hanahan, D. (1983) Studies on transformation of Escherichia coli with plasmids. **J. Mol. Biol.**, 166:557-580.

Hauser, H., Gerhard, G., Bruns, W., Hochkepple, H., Mayr, U. and Collins, J. (1982) Inducibility of human beta interferon gene in mouse L cells. **Nature**, 297:650-654.

Hohn, B. and Collins, J. (1980) A small cosmid for

efficient cloning of large DNA fragments. **Gene**, 11:291-298.

Holbrook, N. J., Lieber, M. and Crabtree, G. R. (1984) DNA sequence of the 5' flanking region of the human interleukin 2 gene: homolgies with adult T-cell leukemic virus. **Nuc. Acids Res.**, 12:5005-5010.

Holbrook, N. J., Smith, K. A., Fornance, A. J., Comeau, C. M., Wiskocil, R. L. and Crabtree, G. R. (1984) T cell growth factor: complete nucleotide sequence. **Proc. Natl. Acad. Sci. USA**, 81:1634-1638.

Howard, M. (1983) Interleukines for B lymphocytes. **Surv. Immunol. Res.**, 2:210-212.

Jay, E., Bambara, R., Padmanabhan, R. and Wu, R. (1974) DNA sequence analysis: a general, simple and rapid method for sequencing large oligodeoxynucleotide fragments by mapping. **Nuc. Acids Res.**, 1:331-334.

Karanthanasis, S. (1982) M13 DNA sequencing using reverse transcriptase. **BRL Focus 4**, #3:6-7.

Kashima, N., Takaoka, C., Fujita, T., Taki, S., Yamada, G., Hamuro, J. and Taniguchi, T. (1985) Unique structure of murine interleukin-2 as deduced from cloned cDNA's. **Nature**, 313:402-404.

Krawetz, S. A., States, J. C. and Dixon, G. H. (1985) Isolation and fractionation of total nucleic acids from tissues and cells., **Journal of Biochem. Biophys. Techniques**, 12:29-36

Lagrimini, M., Brentano, S. and Danelson, J. (1984) A DNA

sequence analysis package for the IBM PC. Nuc. Acids Res., 12:605-614.

Lathe, R. (1985) Synthetic oligonucleotide probes deduced from amino acid sequence data, theoretical and practical considerations. J. Mol. Biol., 183:1-12.

Lindenmaier, W., Hauser, H., deWilke, I. and Schutz, G. (1982) Gene shuttling: moving of cloned DNA into and out of eukaryotic cells. Nuc. Acids Res., 4:1243-1256.

Liu, D. Y. (1984) Macrophage directed lymphokines. Surv. Immunol. Res., 3:154-160.

Maizel, A. L. and Lachman, L. B. (1984) Control of human lymphocyte proliferation by soluble factors. Lab. Invest., 50:369-377.

Maniatis, T., Fritsch, E. F. and Sambrok, J. (1982) Molecular Cloning: A laboratory Manual. Cold Spring Harbor Laboratory, New York, New York

Marx,J. L. (1985) Interleukin-1 genes are cloned.,Science, 228:1076-1077

Meinkoth, J. and Wahl, G. (1984) Hybridization of nucleic acids immobilized on solid supports. Annals. of Biochem., 138:267-275.

Mita, S., Maeda, S., Oboro, K., Nishino, N., Shimada, K., Hirano, K. and Ogawa, H. (1983) Isolation and characterization of the human Il-2 gene. Biochem. Biophys. Res. Comm., 117:114-121.

Molecular Diagnostics Laboratory, Division of genetics, Alberta Childrens Hospital, Calgary, Alberta, Method

001

Moretta, A., Colombatti, M. and Chapuis, B. (1981) Human spleen as a source of T cell growth factor. Cin. exp. Immunol., 44:262-269.

Morgan, D. A., Ruscetto, F. W. and Gallo, R. (1976) Selective in vitro growth of T lymphocytes from human bone marrow. Science, 193:1007-1009.

Mount, D. W. and Conrad, B. (1984) Microcomputer programs for DNA sequence analysis. Nuc. Acids Res., 16:31-37.

Muraguichi, A., Kishimoto, T., Miki, T., Kaieda, A., Tyoshizaki, K. and Yamamura, Y. (1981) T cell replacement factor induced IgG secretion in a human B blastoid cell line and demonstration of acceptors for TRF. J. Immun., 127:412-418.

Panem, S. (1983) The emergent importance of lymphokines. JAMA, 249:166-171.

Pennica, D. P., Nedwin, G. E., Hayflick, J. S., Seeburg, P. H., Derynick, R., Palladino, M. A., Kohr, W. J., Aggarwal, B. B. and Goeddel, D. V. (1984) Human tumour necrosis factor: precursor structure, expression and homolgy to lymphotoxin. Nature, 312:724-728.

Panabieres, I., Fort, P., Piechaezyk, M., Blanchard, J., Jeanteur, P. (1982) A warning on the use of synthetic DNA primers for initiation of reverse transcription on RNA templates: unexpected initiation at a mismatched nucleotide., Gene, 19:321-326

Robb, R. J. and Smith, K. A. (1981) Heterogeneity of human T cell growth factor due to glycosylation. Mol. Immunol., 18:1087-.

Rosenberg, S. A. and Grimm, E. A. (1984) Biological activity of recombinant human Il-2 produced in E. coli. Science, 233:1412-1414.

Sanger, F., Air, B. G., Barrell, N. L., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A. and Smith, M. (1977) Nucleotide sequence of bacteriophage phi X 174 DNA. Nature, 265:687-695.

Schneider, U., Schwenk, H. and Bornkamm, G. (1977) Characterization of EBV-genome negative null and T cell lines derived from children with acute lymphoblastic leukemia and leukemic transformed non-Hodgkin lymphoma. Int. J. Cancer, 19:621-626.

Schwindinger, W. and Warner, J. (1984) DNA sequence analysis on the IBM PC. Nuc. Acids Res., 12:601-604.

Seigel, L. J. (1984) Gene for TCGF: location on human chromosome 4q and feline chromosome 3. Science, 223:175-178.

Sikora, K. (1984) Cancer toxin genes cloned. Nature, 312:699-700.

Smith, K. A. (1980) T cell growth factor. Immunol. Rev., 51:337-357.

Smith, K. A. (1984) Interleukin 2. Ann. Rev. Immunol., 2:319-333.

Southern, E. M. (1975) Detection of specific sequences

among DNA fragments separated by gel electrophoresis. J. Mol. Biol., 98:503-517.

Stanners, C. P., Elicine, C. L. and Green, H. (1971) Two types of ribosome in mouse-hamster hybrid cells. Nature, 230:52-54.

Taniguchi, T., Matsui, H., Fujita, T., Takaoka, C., Kashima, N., Yoshimoto, R. and Hamuro, J. (1983) Structure and expression of a cloned cDNA for human Il-2. Nature, 302:305-310.

Viera, J. and Messing, J. (1982) The pUC plasmids, an M13mp7 derived system for insertion mutagenesis and sequencing with synthetic universal primers. Gene, 19:259-268.

Wahl, G. M., Stern, M. and Stark, G. R. (1979) Effective transfer of large DNA fragments from agarose gels to DBM paper and rapid hybridization by using dextran sulfate. Pro. Natl. Acad. Sci., 76:3683-3689.

Wallace, R. B., Johnson, M. J., Hirose, T., Miyaki, T., Kanashima, E. H. and Itakura, K. (1981) The use of synthetic oligonucleotides as hybridization probes. II. Hybridization of oligonucleotides of mixed sequences to rabbit B-globin DNA. Nuc. Acids Res., 9:879-894.

Watson, J. D., Mochizuki, D. Y. and Gillis, S. (1983) Molecular characterization of Il-2. Federation proc., 42:2747-2752.

Wood, W.I., Capon, D.J., Simonsen, C., Eaton, D.,

Gitschier, J., Kept, B., Seeburg, P., Smith, D., Hollingshead, P., Wion, K., Delwart, E., Tuddenham, E., Vehar, G., Lawn, R. (1984) Expression of active human factor VIII from recombinant DNA clones., Nature, 312:330-336

Wood, W.I., Gitschier, J., Lasky, L., Lawn, R. (1985) Base composition-independent hybridization in tetramethylammonium chloride. A method for oligonucleotide screening of highly complex gene libraries., PNAS, 82:1585-1588

Yousaf, S., Carroll, A. and Clarke, B. (1984) A new and immproved method for 3'-end labelling DNA using gamma -32P dATP. Gene, 27:309-313.