Volume 13 Number 8 1985 Nucleic Acids Research

Cloning and sequence analysis of an Ig λ light chain mRNA expressed in the Burkitt's lymphoma cell line EB4

M.L.M.Anderson, L.Brown, E.McKenzie, J.E.Kellow and B.D.Young*

Beatson Institute for Cancer Research, Garscube Estate, Switchback Road, Bearsden, Glasgow, UK

Received 18 February 1985; Accepted 22 March 1985

ABSTRACT

A cDNA library was constructed from the mRNA of the Igλ producing Burkitt's lymphoma cell line, EB4. Overlapping clones encompassing the coding sequence of the Igλ mRNA were isolated and sequenced. The predicted amino acid sequence shows a short hydrophobic leader peptide and a mature polypeptide of 217 residues in which V, J and C regions can be distinguished. The V region belongs to subgroup VI and has greatest homology (80%) with the Amyloid-AR protein. The constant region is the $Kern^-$ Oz^+ isotype. Probing normal human DNA with the subcloned V_λ coding sequence detects one gene at high stringency and a family of 11 members at low stringency. To date, no restriction enzyme site polymorphisms have been detected. The $V_{\lambda VI}$ gene is rearranged on both chromosomes of EB4 and is deleted on both chromosomes in the Burkitt's lymphoma cell line BL2.

INTRODUCTION

The basic unit of an immunoglobulin (Ig) molecule is composed of two heavy and two light polypeptide chains. In any one molecule, the light chains are either both k or both λ type, but never a mixture. The ratio of k:λ chains in human serum Ig is 60:40 (1).

Ig light chains are encoded by three segments, variable (V_L) genes, joining (J_L) sequences and constant (C_L) genes. In germ line cells, these genes are discontinuous and are not transcribed (2,3). Light chain production involves rearrangement of DNA in B lymphocyte precursor cells such that one of many V_L genes is juxtaposed to one of several J_L sequences. This generates an activated light chain gene which is transcribed (2,3). Excision of the intron between the J_L and C_L sequences in the primary transcript gives rise to light chain mRNA.

The human Igk locus has been extensively studied and the V, J and C segments which contribute to k chain production have been defined (2,3). By contrast, less is known about the human Igλ locus. The constant region genes have been mapped and their structure and organisation are known (4,5). However comparatively little is known about the V_λ and J_λ region genes. To

date, only one genomic V_λ gene has been isolated and it cross hybridises to a family of about 10 members (6). A $V_{\lambda I}$ sequence from an Igλ cDNA has recently been reported which detects a family of some 12 members (7). The human Ig V_λ locus must therefore be more complex than the better studied mouse V_λ locus which contains only two members for inbred lines (8-10) and three for feral strains (11).

In order to characterise further the human V_λ locus, we have isolated an Igλ cDNA containing a $V_{\lambda VI}$ sequence and used it to probe human DNA from light chain producing and non producing sources.

MATERIALS AND METHODS

Construction and screening of a cDNA Library

Cytoplasmic RNA was prepared by phenol extraction of the post mitochondrial supernatant of EB4 cells lysed in 0.14 M NaCl, 10mM Tris-HCl, pH7.4, 1.5mM $MgCl_2$, 0.5% NP40. The polyadenylated mRNA was recovered by alcohol precipitation followed by two cycles of adsorbtion and elution from an oligo (dT)-cellulose column (12). Double stranded cDNA prepared by the method of Wickens (13,14), was ligated into the SmaI site of the plasmid pUC8. A cDNA library (50,000 recombinants from 0.5 µg polyadenylated RNA) was made and screened at high density (15) using as a probe, pul 3.2, a subcloned 0.6 kb BglII-EcoRI fragment of the Ig C_λ locus which includes the $C_{\lambda 3}$ gene, Ke^-Oz^+ (4). The DNA from purified positive clones was analysed on Southern blots to determine the insert size and confirm the hybridisation to the Ig C_λ probe.

DNA Sequencing

Inserts from cDNA clones were excised with EcoRI and HindIII and were subcloned into the appropriate sites of replicative form DNA of the bacteriophages M13mp8 and M13mp9. Sequencing was carried out by the method of Sanger (16,17) using (α^{35}S)dATPαS.

Human DNA preparation

DNA from tissue culture cell lines was prepared by the method of Gross-Bellard **et al.** (18).

To frozen cord blood was added an equal volume of 5M guanidinium thiocyanate, 50 mM Tris-HCl pH 7.0, 50mM EDTA, 5% mercaptoethanol. The blood was thawed at room temperature with gentle mixing. An equal volume of isopropanol was added and the DNA which precipitated immediately, was isolated by centrifugation, then washed in 70% ethanol and dried. The DNA was purified from protein and RNA by conventional procedures (18).

Restriction digests, Gel electrophoresis, Transfer and Hybridisation of DNA

These procedures are described in (6).

Probes - In all cases inserts excised from plasmids were used as probes.

pul 3.2 A 0.6 kb EcoRI-Bgl II fragment carrying the Igλ $C_{\lambda 3}$ gene (4)

pLB1.3 A 0.3 kb Alu I- Alu I fragment carrying the $V_{\lambda VI}$ gene (this paper)

pHV0.6 A 0.6 kb BamHI-BglII fragment subcloned from pHV4A and carrying a genomic Ig V_0 gene (6). (The predicted amino and sequence of this gene does not fall into any of the subgroups so far described. It was formerly tentatively assigned to a new subgroup VII (6) however, since it is not known if it is ever transcribed, the designation of subgroup 0 seems more appropriate.)

RESULTS AND DISCUSSION

Characterisation of λ chain cDNA by cloning and sequencing

The EB4 cell line is derived from malignant lymphocytes of a patient suffering from Burkitt's lymphoma (19) and has been shown to produce Igλ light chains (20). To isolate a V sequence, we constructed a cDNA library from EB4 polyadenlyated cytoplasmic RNA and screened it with a genomic Ig C_λ gene probe. Although it was not known which of the six C_λ genes is used for mRNA synthesis in EB4 cells, the C_λ genes differ in sequence by only a few nucleotides (4), so any one should hybridise to all C_λ sequences. A total of 8 positively scoring recombinants was isolated. To determine if these recombinant clones contained a V_λ sequence, the inserts were recloned into the phage vectors M13mp8 and M13mp9 and sequenced.

Fig. 1 shows restriction maps and the sequencing strategy of 4

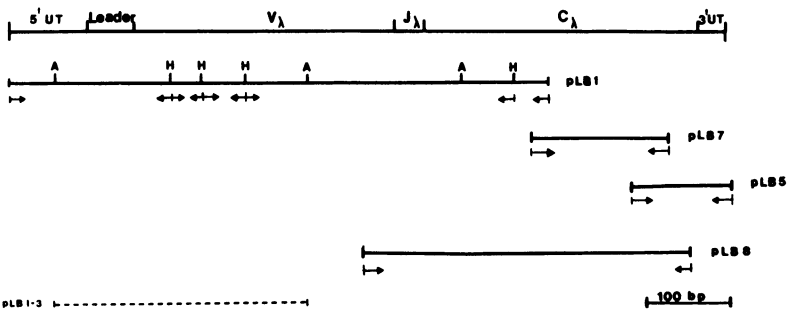

Fig. 1. Restriction map of cDNA clones representing human Ig λ light chain gene and the strategy for DNA sequencing. The DNA restriction fragments indicated by arrows were sequenced by the procedure of Sanger (16,17). Overlapping sequences from both strands were determined. A = AluI, H = HpaI. The dotted line shows the 0.3 kb subcloned AluI-AluI fragment, pLB1.3.

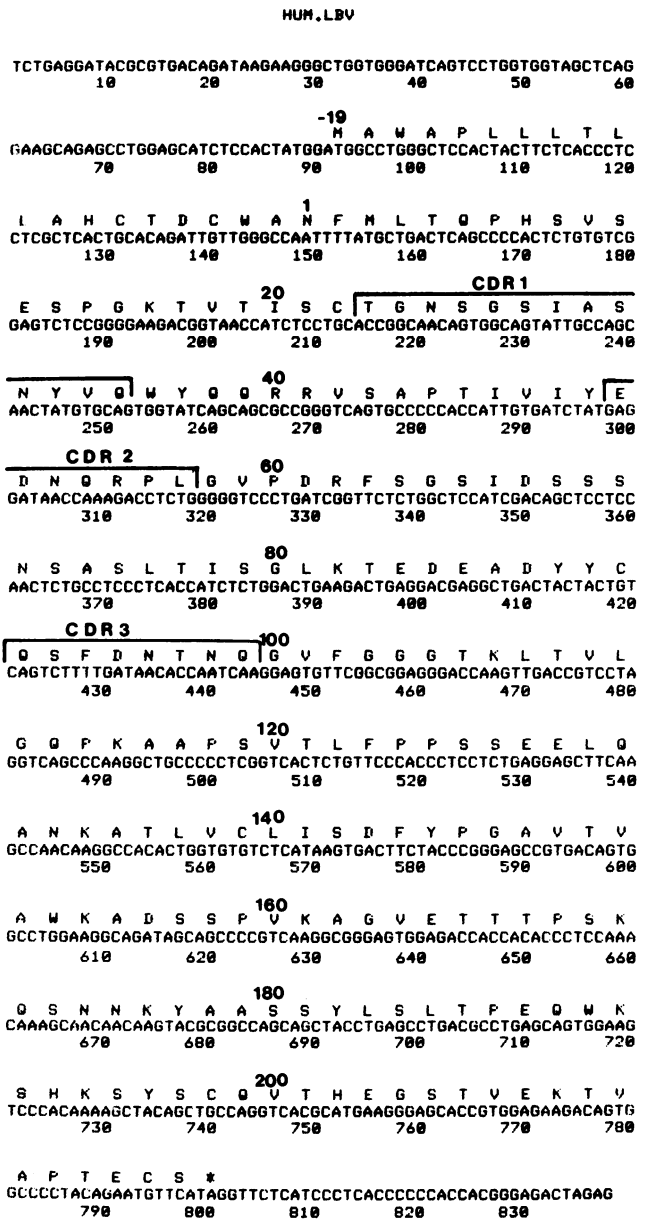

Fig. 2. Nucleic acid sequence of the human Ig λ cDNA. The predicted amino acid sequence is shown above the coding region. Figures below the nucleic acid sequence refer to the nucleotide number and figures above the amino acid sequence refer to the amino acid number. Minus numbers denote amino acids in the leader peptide. CDR denotes complementary determining region.

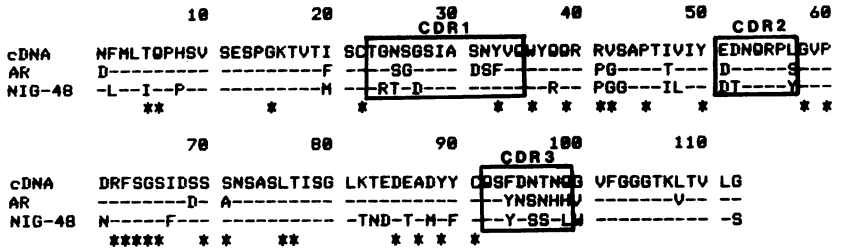

Fig. 3. Comparison of the predicted amino acid sequence encoded by the LBV cDNA V region and members of human subgroup VI. A dash in the sequences of AR (Amyloid-AR) and NIG 48 indicates identity at that position to the cDNA V sequence. CDR denotes complementarity determining region. Stars denote positions of low variability (21-23).

overlapping subclones which encompass the coding region. A compilation of the sequence information is presented in Fig. 2. We shall refer to the compiled sequence as LBV. There is an open reading frame starting at nucleotide 91 and extending for 700 bases. This predicts an amino acid sequence in which a leader peptide, and the V, J and C regions characteristic of Ig light chains can be distinguished.

A 19 amino acid long hydrophobic peptide encoded by nucleotides 91-147 precedes the V region. This peptide is probably a signal peptide which is cleaved from precursor Ig chains as they pass through the cell membrane. While mouse V_λ leader peptides have been known for several years (10,9), human λ signal peptides have only recently been predicted from the nucleotide sequences of a genomic V_λ (pHV4A) and a cDNA $V_{\lambda I}$ clone (6,7). The mouse and human leader peptides have the same length, 19 amino acids, and are hydrophobic in nature, although the amino acid sequences are not identical. The LBV leader peptide has the same amino acid at 12/19 positions as the cDNA V_λ sequence and at 9/19 positions as the genomic pHV4A V_λ sequence. There is a striking conservation of the first 9 nucleotides in the leader sequences of the LBV, pHV4A and mouse Ig $V_{\lambda I}$ DNAs, but the significance of this is not known.

The complete V domain of the mature light chain polypeptide is 112 amino acids long and is encoded by nucleotides 148-483. The three complementarity determining regions (CDR1,2,3) characteristic of V regions can be distinguished and the appropriate amino acid is generated at 23/26 positions designated as low variant; Fig. 3 (21-23).

The amino acid sequence of the complete V domain was compared by

computer to all known human Ig V_λ chains. Greatest homology was observed with Amyloid-AR and NIG-48 (80% and 73% respectively), the only two members of subgroup VI for which the complete sequence is known (22). Since V chains are assigned to the same subgroup if they share 70% or greater homology (22,23), we conclude that EB4 cells produce subgroup VI V_λ chains. A characteristic feature of the fully sequenced members of subgroup VI, which is not shared by other subgroups, is the presence of three extra amino acids, X-Asp-X, at positions 66-68. This feature is also present in the LBV protein sequence. Of the partially sequenced members of subgroup VI, only GIO has complete homology with LBV. However, since only 24 amino acids of GIO have been determined, we do not know if LBV is identical to GIO.

The thirteen amino acids at the carboxy terminal end of the complete V domain are highly conserved and are encoded by a J segment of DNA (22,2,3). Within the complete V domain of LBV, nucleotides 445-483 corresponding to amino acids 100-112 derive from a J segment of DNA. To date apart from that of a processed pseudogene (24), no human genomic J_λ DNA sequence has been reported. The J sequence used in EB4 λ mRNA production shares identity at 34/39 positions at the nucleotide level and 12/13 positions at the amino acids level with the J sequence recently described for a human $V_{\lambda I}$ cDNA clone (7). The amino acid difference between these two sequences is at position one. This is known to be a hot spot for mutation (22) because it occurs at the join of V and J DNA where the position of recombination is flexible and thus contributes to V chain diversity (2,3).

There are six human λ constant region genes clustered in a 40 kb stretch of DNA of which three have been sequenced (4). The constant region expressed in EB4 cells (amino acids 113-217 encoded by nucleotides 484-801) is most homologous to the $C_{\lambda 3}$ gene which encodes the isotype $Kern^-Oz^+$. The only differences (at nucleotides 600,636,700,789 in Fig. 2) are in the third base position of codons and therefore the predicted amino sequence is identical to the $Kern^-Oz^+$ isotype.

Nucleotides 1-90 of LBV constitute the 5' non coding sequence. It has little homology to the mouse or either of the two other human λ 5' flanking regions so far described (9,6,7). Human λ chains may be more variable in their 5' untranslated sequences than their k counterparts for it has been shown that two members of the same subgroup are identical for 388 nucleotides immediately preceding the coding region (25). On the other hand, it is possible that flanking regions of members of the same λ subgroup are very

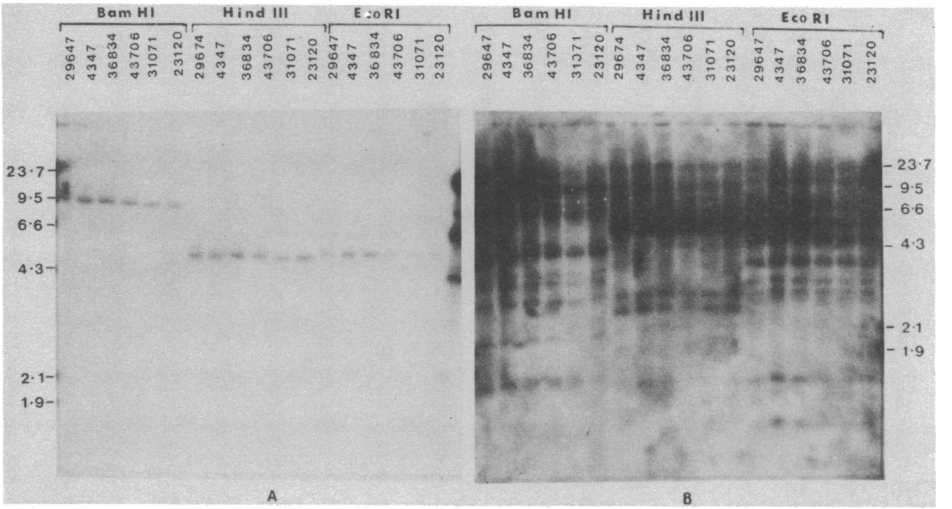

Fig. 4. Hybridisation of subcloned 0.3 kb $V_{\lambda VI}$ probe to human DNA. Ten micrograms of DNA from each of six individuals was restricted with BamHI, HindIII or EcoRI. Five micrograms of each digest was fractionated on duplicate 1% agarose gels and transferred to GeneScreen filters. Hybridisation of the probe was under A) high stringency or B) low stringency conditions as described in (6). Lanes 1-6, BamHI digests: lanes 7-12, HindIII digests: lanes 13-18, EcoRI digests. Numbers above lanes identify the individuals; numbers to the left indicate the length in kb of HindIII/λ DNA and HaeIII/ ϕX174 DNA markers.

similar, but differ from those of other subgroups. When other members of the same subgroups are characterised, it will be possible to resolve this question.

Determination of Copy Number of $V_{\lambda VI}$ genes

To determine the extent of polymorphism and to estimate the copy number of the $IgV_{\lambda VI}$ genes, we probed digests of several human DNAs with a subcloned 0.3 kb Alu-AluI fragment, pLB1.3, which contains only leader and V_{λ} coding sequences (Fig. 1). The results obtained when DNA samples from cord blood of six unrelated individuals were analysed at two different stringencies are shown in Fig. 4. For HindIII, BamHI and EcoRI digests only one band was observed at high stringency. At lower stringency, 11 bands were observed with each digest. In an analysis of 18 different people, we found no evidence of polymorphism for any of the above enzymes. This means that here, gene counting by Southern blot analysis is not complicated by restriction fragment length polymorphisms. The above results suggest that there is one gene per haploid complement with sufficient homology to be detected at high

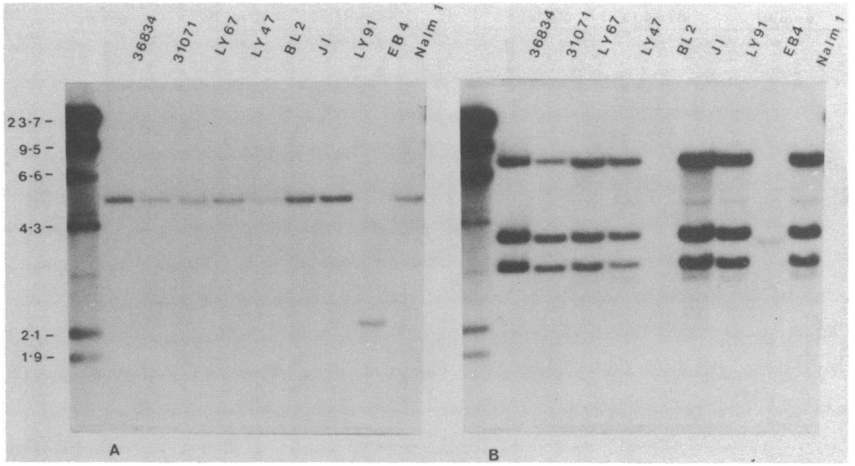

Fig. 5. Hybridisation of subcloned V_λ probes to DNA from Burkitt's lymphoma cell lines. Human DNAs were digested with EcoRI, fractionated on 0.8% agarose gels and transferred to filters. Hybridisation and washing were at high stringency conditions. 36834 and 31071, cord blood DNA. The type of Ig light chain produced and translocation in Burkitt's lymphoma lines is: LY47 (-, 8:22); LY67 (λ, 8:22); BL2 (λ, 8:22); JI (k, 8:2); LY91 (k, 8:2); EB4 (λ, 8:14) (26). The NALM-1 cell line carries a 9:22 chromosomal translocation (30).
A - $V_{\lambda VI}$ probe (pLB 1.3), B - genomic V_λ probe (pHV0.6), both under high stringency conditions..

stringency of hybridisation and that there is a family of at least 11 members which cross hybridise with the $V_{\lambda VI}$ sequence. When the same blots were hybridised at low stringency to a subcloned genomic $V_{\lambda 0}$ probe pHV0.6 (6), about 10 bands were observed which did not overlap with those hybridising to pLB1.3 (data not shown). At a minimum estimate then, the human V_λ locus must consist of at least 20 members. When several hybridization probes are available for each subgroup, it will be possible to make more accurate estimates. The number of V_λ genes will be known definitely only when all the genes have been cloned and sequenced, since only by sequencing can we distinguish between potentially functional genes and pseudogenes.

Characterisation of the $V_{\lambda VI}$ gene in lymphoid cell lines

The Ig loci in cord blood DNA are in the germ line or unrearranged form. In order to examine the $V_{\lambda VI}$ gene arrangement in Ig producing cells, we probed a Southern blot of EcoRI digested DNAs with ^{32}P-labelled pLB1.3 DNA (Fig. 5A). DNA samples were from cord blood (non producing), Burkitt's lymphoma cell lines (k or λ producers) and the NALM-1 cell line (a myeloid

line with a 9:22 translocation). In cord blood DNA (lanes 2 and 3), the probe detected a single germline band of 5.0 kb. With DNA of the k producing lines, JI and LY91, the same 5.0 kb band was observed indicating that the $V_{\lambda VI}$ gene is not rearranged in these cells. This is what is expected since it appears that during B lymphocyte differentiation, there is an ordered sequence of light chain gene rearrangement during B lymphocyte differentiation. Kappa genes are rearranged first and only if this is aberrant and fails to give rise to k light chains, does the λ locus become rearranged (27).

Since EB4 cells express $V_{\lambda VI}$ sequences, the $V_{\lambda VI}$ gene must be rearranged on at least one copy of chromosome 22. As seen in Fig. 5A, lane 9, the 5.0 kb band is missing and instead a single band of 2.2 kb is detected. This is probably the productively rearranged gene. The other so-called "excluded" allele appears to be deleted or, less likely, may also be rearranged to the same size of fragment.

BL2 (a λ producer) gives a weaker signal with the pLB1.3 probe than control samples even although the same amount of DNA was applied to the gel. This suggests that one allele of the pLB1.3 gene has been deleted. The other λ producing cell line examined, LY67, must express a different V_λ gene from EB4 cells since the pLB1.3 gene is in the germ line form.

An identical blot prepared at the same time was probed with the $V_{\lambda 0}$ probe, pHV0.6 (6) to compare its pattern of hybridisation to that of pLB1.3. Fig. 5B shows that the three strongly hybridising bands detected in the control cord blood samples are missing in EB4 DNA showing that both copies of chromosome 22 have a rearranged λ locus. A new faint band of 3.7 kb is present in EB4 DNA and may represent a $V_{\lambda 0}$ gene on the non productively rearranged chromosome.

The LY47, LY67 and BL2 cell lines analysed in Fig. 5 are all Burkitt's lymphoma cell lines with 8:22 chromosomal translocations. The breakpoint on chromosome 22 (22q11) is in the same cytogenetic band to which IgλC and λV genes have been assigned (28,29, Anderson **et al.** manuscript in preparation). The $V_{\lambda VI}$ and the pHV0.6 genes are not rearranged in LY47 and LY67 so they can not be very close to the breakpoint. However, it is striking that there are no bands hybridising to pHV0.6 in BL2 DNA. There has been loss of V_λ sequences from both the productively rearranged chromosome 22 and the copy of chromosome 22 which is involved in the 8:22 translocation. The same conclusion was recently reported based on analysis of BL2 DNA with a $V_{\lambda I}$

probe (7). There is no alteration in the size of restriction fragments containing V_λ genes in NALM-1 cells which have a 9:22 translocation involving the 22q11 band (30). This suggests that the breakpoint is not very close to these genes.

Concluding Remarks

The above results show that EB4 cells produce λ light chains of subgroup VI and subtype $Ke^{-}Oz^{+}$. Subgroup VI was initially recognised through peptide analysis of proteins extracted from the splenic fibrils of patients with primary amyloidosis (31,32). In the normal population, subgroup VI represents only 5% of the V sequences in λ chains, but it is the predominant subgroup associated with the amyloid condition (33). It is perhaps surprising that a subgroup which is such a minor contributor to λ chains in normal cells should be one of the first to have its mRNA sequence determined. Further analyses will determine if subroup VI is a predominant contributor to Burkitt's lymphoma light chains.

ACKNOWLEDGEMENTS

We thank G. Stewart for gifts of frozen human cord blood and G.M. Lenoir for Burkitt's Lymphoma cell lines. This work was supported by grants from the Leukaemia Research Fund and the Cancer Research Campaign.

*Present address: Imperial Cancer Research Fund, Lincoln's Inn Fields, London

REFERENCES

1. Kabat, E. (1976) Structural Concepts in Immunology and Immunochemistry (Holt, Rinehart and Winston, New York) p232.
2. Wall, R. and Kuehl, M. (1983) Ann. Rev. Immunol. **1**, 393-422.
3. Honjo, T. (1983) Ann. Rev. Immunol. **1**, 499-528.
4. Hieter, P.A., Hollis, G.F., Korsmeyer, S.J., Waldemann, T.A. and Leder, P. (1980) Nature, **304**, 536-540.
5. Taub, R.A., Hollis, G.F., Hieter, P.A., Korsmeyer, S.J., Waldmann, T.A. and Leder, P. (1980) Nature, **304**, 172-174.
6. Anderson, M.L.M., Szajnert, M.F., Kaplan, J.C., McColl, L. and Young, B.D. Nuc. Acid. Res. **12**, 6647-6661.
7. Tsujimoto, Y. and Croce, C.M. (1984) Nuc. Acid Res. **12**, 8407-8414.
8. Brack, C., Hirama, M., Lenhard-Schuller, R. and Tonegawa, S. (1978) Cell, **15**, 1-14.
9. Bernard, O., Hozumi, N. and Tonegawa, S. (1978) Cell **15**, 1133-1144.
10. Tonegawa, S., Maxam, A.M., Tizard, R., Bernard, O. and Gilbert, W. (1978) Proc. Natl. Acad. Sci. (U.S.A.) **75**, 1485-1489.
11. Scott, C.L. and Potter, M. (1984) J. Immunol. **132**, 3883-2643.
12. Aviv, H. and Leder, P. (1972) Proc. Natl. Acad. Sci. (U.S.A.) **69**, 1408-1412.
13. Buell, G.N., Wickens, M.P., Payvar, F. and Schimke, R.T. (1978) J. Biol. Chem. **253**, 2471-2482.

14. Wickens, M.P., Buell, G.N. and Schimke, R.T. (1978) J. Biol. Chem. 253, 2483-2495.
15. Gergen, J.P., Stern, R.H. and Wensink, P.C. (1979) Nucl. Acid. Res. 7, 2115-2136.
16. Sanger, F., Nicklen, S. and Coulson, A.R. (1978) Proc. Natl. Acad. Sci. (U.S.A.) 74, 5463-5467.
17. Sanger, F. and Coulson, A.R. (1978) FEBS Letts, 87, 107-110.
18. Gross-Bellard, M., Oudet, P. and Chambon, P. (1973). Eur. J. Biochem. 36, 32-38.
19. Epstein, M.A., Barr, Y.M. and Achong, B.G. (1966) Brit. J. Cancer 20, 475-479.
20. McIntosh, R.V., Cohen, B.B., Steel, C.M., Read, H., Moxley, M. and Evans, H.J. (1983) Int. J. Cancer, 31, 275-279.
21. Wu, T.T. and Kabat, E.A. (1970) J. Exp. Med. 132, 211-250.
22 Kabat, E.A., Wu, T.T., Bilofsky, H., Reid-Miller, M. and Perry, H. (1983) Sequences of Proteins of Immunological Interest, U.S. Department of Health and Human Services. N14.
23. Dayhoff, M.O. (1972) Atlas of Protein Sequence and Structure 5, D229-D233.
24. Hollis, G.F., Hieter, P.A., McBride, O.W., Swan, D. and Leder, P. (1982) Nature 296, 321-325.
25. Gustav-Klobeck, Combriato, G. and Zachau, H.G. (1984) Nucl. Acid. Res. 12, 6995-7006.
26. Lenoir, G.M., Preud'homme, J.L., Bernheim, A. and Berger, R. Nature, 298, 474-476.
27. Korsmeyer, S.J., Hieter, P.A., Ravetch, J.V., Poplack, D.G., Leder, P. and Waldmann, T.A. (1981) Leukaemia Markers (Academic Press), 85-97.
28. Rowley, J. (1982) Science 216, 749-751.
29. Goyns, M.H., Young, B.D., Guerts van Kessel, A., deKlein, A., Grosveld, G., Bartram, C.R. and Bootsma, D. (1984) Leukaemia Res. 8, 547-553.
30. Minowada, J., Tsubota, T., Greaves, M.F. and Walters, T.R. (1977) J. Natl. Cancer. Inst. 59,83-88.
31. Sletten, K., Husby, G. and Natvig, J.B. (1974) Scand. J. Immunol. 3, 833-836.
32. Skinner, M., Benson, M.D. and Cohen, A.S. (1975) J. Immunol. 114, 1433-1435.
33. Solomon, A., Frangione, B. and Franklin, E.C. (1982) J. Clin. Invest. 70, 453-460.