Recombination between an expressed immunoglobulin heavy-chain gene and a germline variable gene segment in a Ly 1$^+$ B-cell lymphoma

Robert Kleinfield*, Richard R. Hardy†, David Tarlinton†, Jeffery Dangl†, Leonard A. Herzenberg† & Martin Weigert*

* Institute for Cancer Research, Philadelphia, Pennsylvania 19111, USA
† Department of Genetics, Stanford University, Stanford, California 94305, USA

The early stages of murine B-cell differentiation are characterized by a series of immunoglobulin gene rearrangements which are required for the assembly of heavy(H)- and light(L)-chain variable regions from germline gene segments. Rearrangement at the heavy-chain locus is initiated first and consists of the joining of a diversity (D_H) gene segment to a joining (J_H) gene segment. This forms a DJ_H intermediate to which a variable (V_H) gene segment is subsequently added. Light-chain gene rearrangement follows and consists of the joining of a V_L gene segment to a J_L gene segment: once a productive light-chain gene has been formed the cell initiates synthesis of surface immunoglobulin M (sIgM) receptors (reviewed in ref. 1). These receptors are clonally distributed and may undergo further diversification either by somatic mutation[2,3] or possibly by continued recombinational events[4]. Such recombinational events have been detected in the Ly 1$^+$ B-cell lymphoma NFS-5, which has been shown to rearrange both λ and H-chain genes subsequent to the formation of sIgM ($\mu\kappa$) molecules[5]. Here we have analysed a rearrangement of the productive allele of NFS-5 and found that it is due to a novel recombination event between V_H genes which results in the replacement of most or all of the coding sequence of the initial V_HQ52 rearrangement by a germline V_H7183 gene. Embedded in the V_H coding sequence close to the site of the cross-over is the sequence 5′ TACTGTG 3′, which is identical to the signal heptamer found 5′ of many D_H gene segments[6]. This embedded heptamer is conserved in over 70% of known V_H genes[7–17]. We suggest that this heptamer mediates V_H gene replacement and may play an important part in the development of the antibody repertoire.

The Ly 1$^+$ B-cell lymphoma NFS-5 was generated by inoculating a newborn NFS/N mouse with ecotropic murine leukaemia virus (Cas 2SM)[18]. Recent studies have indicated that this lymphoma is capable of expressing a number of distinct surface immunoglobulin phenotypes during growth *in vitro*[5]. Analysis of primary cultures of NFS-5 on the fluorescence-activated cell sorter (FACS) shows a $\mu^-\kappa^-$ population from which $\mu^+\kappa^-$ cells spontaneously arise. Kappa light-chain synthesis can be induced by treatment with the B-cell mitogen lipopolysaccharide (LPS), to give a $\mu^+\kappa^+$ population. Continued growth of these κ-expressing cells in LPS results in the appearance of a small population, comprising 1–5% of the culture, expressing λ light chains. These λ-expressing cells generally have a $\mu^+\lambda^+$ surface phenotype, however $\mu^+\kappa^+\lambda^+$ co-expressors have also been detected. Hardy *et al.* have cloned a series of cell lines representing each of the various phenotypic forms of NFS-5 (ref. 5). These cloned lines have been designated as follows: 5.3 ($\mu^+\kappa^-$), 5.4κ ($\mu^+\kappa^+$), 5.4κλ ($\mu^+\kappa^+\lambda^+$) and 5.4λ ($\mu^+\lambda^+$). The two λ-expressing lines, 5.4κλ and 5.4λ, which represent progeny derived from a single expanded population of 5.4κ, contain immunoglobulin gene rearrangements not present in either the parental lymphoma or the $\mu^+\kappa^+$ precursor, 5.4κ. Thus, although all lines contain shared rearrangements on both κ alleles, lines differ in the organization of both the λ and H-chain genes. Distinct rearrangements at the λ locus are found in 5.4κλ and 5.4λ; whereas 5.4κλ contains a $V_\lambda1J_\lambda1$ gene, 5.4λ has a rearranged $V_\lambda2J_\lambda2$ gene[5]. As these rearrangements are not evident in either the parental lymphoma or 5.4κ, λ light-chain gene recombination has occurred during growth *in vitro*. In addition, these λ-expressing lines are unusual in that they contain H-chain gene rearrangements not present in the $\mu^+\kappa^+$ precursor, 5.4κ. As both H-chain alleles in 5.4κ consist of complete V_HDJ_H variable regions, these 'secondary' H-chain gene rearrangements have been investigated further.

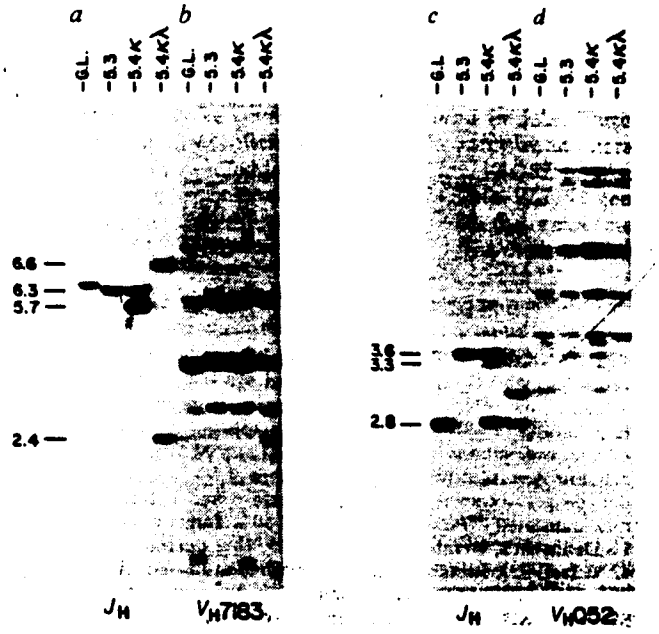

Fig. 1 Context of immunoglobulin variable (V_H) and joining (J_H) gene segments in NFS-5-derived cell lines. Approximately 10 μg of genomic DNA was digested with either *Eco*RI (*a*, *b*) or *Hin*dIII (*c*, *d*), fractionated on 0.8% agarose gels and transferred to nitrocellulose filters[27]; duplicate blots were assayed for hybridization to the ^{32}P-labelled J_H probe, pJ11 (*a*, *c*)[28], and to either V_H7183 (*b*) or V_HQ52 (*d*)[29] gene probe. After hybridization, blots were first washed at low stringency (3×SSC, 68 °C) and then at high stringency (0.2×SSC, 68 °C). The cell lines assayed include 5.3 ($\mu^+\kappa^-$), 5.4κ ($\mu^+\kappa^+$) and 5.4κλ ($\mu^+\kappa^+\lambda^+$). G.L., germline DNA from a NFS/N×NZB RI line homozygous for the NFS/N heavy-chain locus. The J_H region probe, pJ11, consists of a 2.0-kb *Bam*HI/*Eco*RI fragment containing the J_H3 and J_H4 gene segments subcloned into pBR322; the V_HQ52 probe (pV$_H$Q52NHheI) consists of a 0.3-kb *Hha*I fragment subcloned into the *Sma*I site of puC12; the V_H7183 probe (pV$_H$SAPC-15) derives from a 0.9-kb *Eco*RI/*Hae*III restriction fragment subcloned into the *Eco*RI/*Sma*I site of puc12. The 4.1-kb fragment seen in all lanes of *c* and *d* is due to plasmid contamination. Sizes (in kb) of restriction fragments hybridizing to both V_H and J_H probes are given on the left of the autoradiographs. The 3.3-kb fragment in *c* represents a nonproductive V_HDJ_H2 rearrangement (see text, and data not shown). The light hybridization of this band is thought to be due to a *Hin*dIII site between the J_H3 and J_H4 gene segments which cuts the pJ11-hybridizable region into two fragments of 2.8 and 3.3 kb; the 3.3-kb fragment contains the variable region and 400 bases of sequence hyridizable to pJ11.

DNA of high relative molecular mass (M_r) from NFS-5.3, 5.4κ and 5.4κλ was digested with either *Eco*RI or *Hin*dIII and analysed by Southern blot hybridization using the J_H-region probe, pJ11 (Fig. 1*a*, *c*). Both H-chain alleles of 5.4κ are rearranged and detected in *Hin*dIII digests as 3.6-kilobase (kb) and 3.3-kb restriction fragments (Fig. 1*c*). Restriction fragments of identical size are detected using a V_HQ52 gene probe (Fig. 1*d*), suggesting that these rearrangements consist of complete V_HDJ_H joins which utilize V_H genes belonging to the V_HQ52 gene family. This conclusion is supported by Northern

Fig. 2 Northern blot analysis of total RNA from NFS-5-derived cell lines using C_μ (*a*), V_HQ52 (*b*) and V_H7183 (*c*) gene probes. Total cellular RNA was prepared from homogenized cells lysed in 6 M urea/3 M LiCl, extracted with phenol/chloroform and ethanol-precipitated[30]. Approximately 10 μg of RNA from each cell line was electrophoresed through a 1.5% agarose/formaldehyde gel[31], transferred to nitrocellulose and hybridized with nick-translated ^{32}P-labelled probe in 50% formamide, 5 × SSC, 1 × Denhardt's, 20 mM $NaHPO_4$ pH 6.5, 10% dextran sulphate at 42 °C (ref. 32). Blots were initially washed at low stringency (2 × SSC, 0.1% SDS, 52 °C), followed by high-stringency washes (0.1 × SSC, 0.1% SDS, 52 °C). Transcript sizes correspond to those expected for mature μ mRNA.

blot analysis which indicates that 5.4κ synthesizes a mature transcript that hybridizes to both C_μ (Fig. 2*a*) and V_HQ52 (Fig. 2*b*) gene probes. This transcript appears to be encoded by the 3.6-kb *Hin*dIII restriction fragment, as the $\mu^+\kappa^-$ cell line, 5.3, which has retained this restriction fragment but deleted the 3.3-kb fragment (Fig. 1*c*), nevertheless continues to produce both mature μ messenger RNA (Fig. 2*a*) and protein[18]. We conclude that the 3.6-kb *Hin*dIII fragment contains the productive H-chain allele. This allele is rearranged to J_H4 because the 5.3 cell line, which is haploid at the J_H locus (see Fig. 1*a*, *c*), has deleted the 2.8-kb germline-encoded *Hin*dIII fragment; this deletion is diagnostic of rearrangements to J_H4 (Fig. 1*c*). The 3.3-kb *Hin*dIII fragment contains a J_H2-rearranged H-chain allele, as determined both by hybridization using J_H-specific probes and detailed restriction mapping of the cloned allele (R.K., unpublished data). This allele appears to be nonproductive as J_H2-containing transcripts cannot be detected in 5.4κ RNA (data not shown).

Continued rearrangement of both the productive and nonproductive H-chain alleles is evident in 5.4κλ. The relationship between these alleles has been determined by Southern blot analysis using J_H1–J_H2 and J_H3–J_H4-specific probes (data not shown). Rearrangement of the nonproductive allele results in a change in the *Hin*dIII fragment size, from 3.3 to 2.8 kb (Fig. 1*c*). The 2.8-kb J_H-hybridizable fragment is also detected using the V_HQ52 probe (Fig. 1*d*), indicating that this allele still contains a complete V_HQ52DJ_H rearrangement. Accompanying this rearrangement, germline genes in both the V_HQ52 (Fig. 1*d*) and V_H7183 (Fig. 1*b*) gene families have been deleted suggesting an event occurring over many kilobases of DNA extending from the J_H region to the germline V_H locus. Rearrangement of the productive allele in 5.4κλ is accompanied by a change in the *Hin*dIII fragment from 3.6 to 3.0 kb (Fig. 1*c*), and a change in the *Eco*RI fragment from 6.3 to 2.4 kb (Fig. 1*a*). As shown in Fig. 1*d*, the 3.0-kb fragment does not hybridize to the V_HQ52 gene probe, despite its derivation from the 3.6-kb V_HQ52-containing variable region (see above). Rather, as shown in the *Eco*RI digests, the productive allele of 5.4κλ, contained on a 2.4-kb fragment (Fig. 1*a*), hybridizes to a V_H7183 gene probe. Thus, secondary rearrangement of the productive allele results in replacement of the initial V_HQ52 gene by a V_H7183 gene. Northern blot analysis of total RNA from 5.4κλ indicates that this replacement is accompanied by a change in V_H RNA expression from V_HQ52 to V_H7183 (Fig. 2*c*).

To gain insight into the mechanism of this V_H gene replacement, we have sequenced the expressed variable regions of 5.4κ and 5.4κλ using RNA primer extension techniques[19]. Analysis of 5.4κ yields a single V_HQ52DJ_H4 sequence (Fig. 3), indicating that the remaining V_HQ52 rearrangement (the 5.7-kb *Eco*RI fragment in Fig. 1) in this cell line does not accumulate RNA to any significant extent and is therefore nonproductive at the phenotypic level. Analysis of the H-chain mRNAs in 5.4κλ similarly yields a single V_H7183DJ_H4 sequence. No V_HQ52 transcript is detected in these cells by Northern analysis (Fig. 2*b*), indicating that the remaining V_HQ52 rearrangement (the 6.6-kb EcoRI fragment) is not expressed. As the productive V_H7183 rearrangement in 5.4κλ derives from a V_HQ52 rearrangement, V_H gene replacement has occurred during propagation *in vitro*. This conclusion is supported by sequence identity of the *N*, D_H and J_H segments expressed in 5.4κ and 5.4κλ. The *N* sequences are of particular interest as they are created *de novo* during the *V*–*D* recombinational process. Identity of *N* sequences confirms the common origin of these alleles (see Fig. 3). The variable gene segments of these alleles, however, show marked differences in sequence. Comparison of the 5.4κ sequence with a representative germline V_HQ52 sequence (M141; ref. 16) confirms that this gene is a member of the V_HQ52 gene family. A similar comparison of the 5.4κλ sequence with a representative germline V_H7183 sequence (V_HE4.30; ref. 15) establishes that the productive allele of 5.4κλ is derived from a member of the V_H7183 gene family. Recombination, therefore, results in replacement of most or all of the coding sequence of

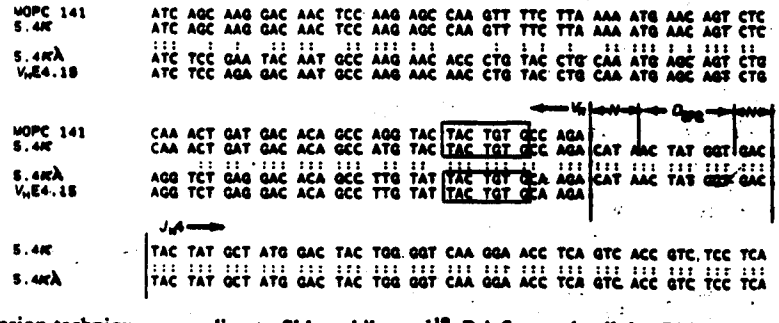

Fig. 3 Nucleotide sequence comparison between the expressed alleles of 5.4κ and 5.4κλ. Sequence homologies are shown and reference V_HQ52 (MOPC 141)[16] and V_H7183 (V_HE4.15)[15] germline sequences are given to indicate V_H gene family relationships (>80% sequence homology)[30]. Because of identity in the terminal three bases of the V_H segments, the exact site of the recombination cannot be determined precisely; however, four bases 5′ of the V_HD_H junction, a single germline-encoded base difference between 5.4κ (cytosine) and 5.4κλ (adenine) indicates the 5′ boundary of the recombination. Base sequences identical to the recombination signals found 5′ of many *D* elements (5′TACTGTG 3′)[17] are boxed.

Methods. Sequences were obtained using RNA primer extension techniques according to Shlomchik *et al.*[19]. Briefly, total cellular RNA was extracted in guanidinium isothiocyanate[33], and poly(A)$^+$ RNA was selected on oligo-(dT)-cellulose columns[32]. Synthesis of complementary DNA was primed with 50 ng of the 5′-end-labelled oligonucleotide (5′GCAGGAGAGGAGGGGGA 3′) homologous to μ constant-region exon 1. Labelled primer was annealed to 80–120 μg of poly(A)$^+$ RNA, and reverse transcription carried out in 100 mM Tris(8.3), 140 mM KCl, 10 mM $MgCl_2$, 0.5 mM $dNTP_s$, 10 mM dithiothreitol, 120 U RNasin, 40 U reverse transcriptase for 2 h at 42 °C. The cDNA was obtained by fractionation on a 5% polyacrylamide/7M urea gel for 2–3 h at 30 V cm^{-1}. Gel slices were chopped and eluted overnight at 37 °C in 0.5 M NH_4Ac, 1 mM EDTA. Full-length cDNA was ethanol-precipitated and washed extensively before sequencing by modified chemical degradations, as described by Rubin and Schmid[34] and Bencini *et al.*[35].

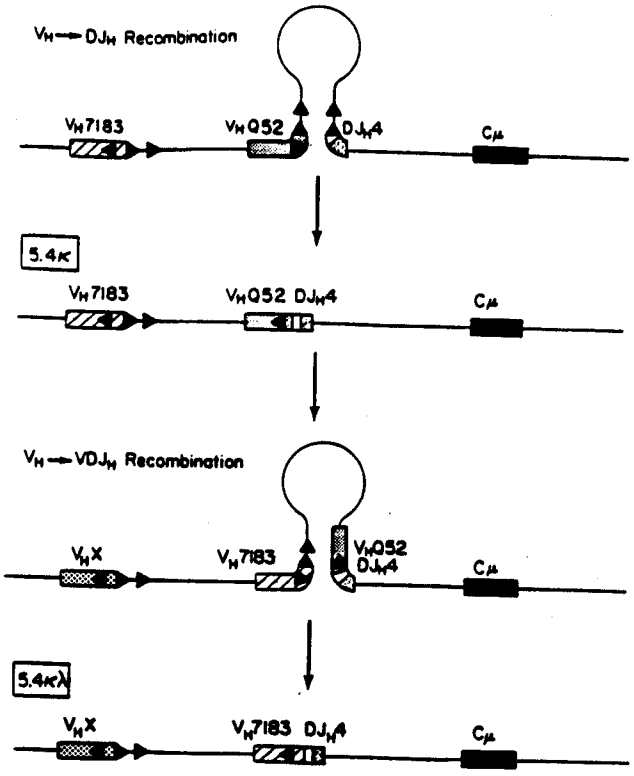

Fig. 4 Model of the V_H gene recombinations leading to the formation of the expressed H-chain allele of 5.4κλ. Heptameric and nanomeric signal sequences are indicated by black arrows, and component gene segments are indicated by patterned overlays. Both N and D_{SP2} sequences are depicted together in the figure as an open box. Top panel shows the initial V_HQ52 to DJ_H4 recombination, involving interaction of signal sequences found 3′ of the V_HQ52 gene with those 5′ of the DJ_H4 complex to form a hypothetical stem-loop structure. Excision at the base of the stem and fusion of V_HQ52 gene to DJ_H4 completes the formation of the intact variable region of 5.4κ. Subsequent recombination between an upstream germline V_H7183 gene segment and the productive V_HQ52 rearrangement is postulated to occur via interaction of the heptameric and nonameric signal sequences 3′ of the V_H7183 gene with the embedded heptamer of the rearranged V_HQ52 gene, resulting in the replacement of the V_HQ52 coding region by a V_H7183 gene segment. This recombination is precise, and does not result in the addition or deletion of bases: the 5′ portion of the newly formed variable region is contributed by the incoming V_H7183 gene, whereas the 3′ portion, consisting of N, D_{SP2} and J_H4 sequences, has been retained from the recombinations which initially formed the allele.

the resident V_HQ52 gene with that of the incoming V_H7183 gene. This recombination has occurred so as to maintain a continuous open reading frame. Because of the identity of the 3′-terminal three bases of the germline V_H gene segments (both gene families use an AGA for this codon), the precise 3′ point of the recombination cannot be identified. However, a single germline-encoded base difference between 5.4κ and 5.4κλ delineates the 5′ boundary of the recombination four bases 5′ to the V-D junction.

A number of mechanisms such as gene conversion[4] and homologous recombination[20] may be invoked to explain this V_H gene replacement. We consider yet a third model of V_H-V_H recombination by a mechanism analogous to V_H-D_H recombination. Embedded in the V_H coding region, close to the site of the recombination, is the sequence 5′ TACTGTG 3′[21], which is identical to the signal heptamer found 5′ of many D_H gene segments[17]. We suggest that this heptamer functions as a pairing element in mediating V_H-V_H recombination (see Fig. 4). Here a V_HQ52 gene is paired by its 3′ signal sequences to the signal sequences found 5′ of DJ_H4; shown as a stem-loop structure. Endonucleolytic cleavage at the base of the stem, and fusion of the V_HQ52 gene segment to this DJ_H4 intermediate results in the formation of the productive allele found in 5.4κ. Further rearrangement of this allele is postulated to occur by an interaction between the embedded heptamer of the V_HQ52 gene and the signal sequences 3′ of one of the germline V_H7183 genes still present on the chromosome. This results in the replacement of the V_HQ52 gene by a V_H7183 gene to give the productive allele found in 5.4κλ.

Although conventional immunoglobulin gene recombination is correlated with the presence of both heptameric and nonameric signal sequences, recombination at isolated heptamers has also been reported. A precedent for heptamer-mediated recombination has been obtained at the κ locus, where rearrangements of both $V_κ$ and $J_κ$ gene segments to an isolated heptamer in the $J_κ$-$C_κ$ intron have been reported[20,22]. Moreover, in many λ-producing plasmacytomas and hybridomas, deletion of the $C_κ$ gene segment is mediated by a recombinational event between this intron heptamer and heptamer-nonamer sequences located 3′ of the $C_κ$ locus (such $C_κ$ deletion has not, however, been observed in 5.4κλ)[23,24]. It has been suggested that such a deletion may be an important preamble to λ gene recombination[23], implying that these heptamer-mediated recombinations may have biologically important consequences.

Yancopoulos *et al.*[15] and Perlmutter *et al.*[25] have recently described the preferential usage of V_H genes from the J-most proximal V_H gene family early in fetal development. This preference is not evident in the adult repertoire where the V_H gene families appear to be used at frequencies reflecting the relative complexity of each gene family[26]. We suggest that V_H gene replacement mediated by the embedded heptamer may represent one mechanism leading to the random V_H usage observed in the adult repertoire. Indeed, the strong conservation of this heptameric sequence within the FR3 region where it encodes Tyr and Cys residues at positions 91 and 92 supports this hypothesis: over 70% of V_H genes which have been examined, including all functional members of the J-proximal V_H families (for example, V_H7183, V_HQ52 and V_HS107)[7-17], have been found to contain this heptameric sequence. Thus, conservation is evident not only at the amino-acid sequence level, but also at the nucleotide level where little, if any, third base degeneracy is observed at both the Tyr and Cys codons. Such a high degree of conservation, even among highly diverged V_H gene families, suggests that this heptameric sequence has an important role in the somatic diversification of the antibody repertoire.

We thank G. Lennon and M. Shapiro for useful discussions, J. Brill Dashoff for expert technical assistance and M. Shlomchik for sequencing reagents. This work was supported in part by NIH grants GM-20964-13, CA-04681 and HD01287. R.R.H. was supported by the American Cancer Society, California Division (Fellowship S-12-84).

Received 6 May; accepted 28 May 1986.

1. Yancopoulos, G. & Alt, F. W. *A. Rev. Immun.* 4, 339–368 (1986).
2. Weigert, M., Cesari, I. M., Yonkovich, S. J. & Cohn, M. *Nature* 228, 1045–1047 (1970).
3. McKean, D. *et al. Proc. natn. Acad. Sci. U.S.A.* 81, 3180–3184 (1984).
4. Dildrop, R., Brüggemann, M., Radbruch, A., Rajewski, R. & Beyreuther, K. *EMBO J.* 5, 635–640 (1982).
5. Hardy, R. R. *et al. Proc. natn. Acad. Sci. U.S.A.* 83, 1438–1442 (1986).
6. Kurosawa, Y. & Tonegawa, S. *J. exp. Med.* 155, 201–218 (1982).
7. Bothwell, A. L. *et al. Cell* 24, 625–637 (1981).
8. Crews, S., Griffin, J., Huang, H., Calame, K. & Hood, L. *Cell* 25, 59–66 (1981).
9. Givol, D. *et al. Nature* 292, 426–430 (1981).
10. Kaartinen, M., Griffiths, G. M., Martham, A. F. & Milstein C. *Nature* 304, 320–324 (1981).
11. Loh, D. Y., Bothwell, A. L. M., White-Scharff, M. E., Imanishi-Kari, T. & Baltimore, D. *Cell* 33, 85–93 (1983).
12. Near, R. I. *et al. Proc. natn. Acad. Sci. U.S.A.* 81, 2167–2171 (1984).
13. Sims, J. *et al. Science* 216, 309–311 (1982).
14. Kaartinen, M. *et al. J. Immun.* 130, 937–945 (1983).
15. Yancopoulos, G. D. *et al. Nature* 311, 727–733 (1984).
16. Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. & Tonegawa, S. *Nature* 286, 676–683 (1980).
17. Early, P., Huang, H., Davis, M., Calame, K. & Hood, L. *Cell* 19, 981–992 (1980).
18. Davidson, W. F. *et al. J. Immun.* 133, 744–753 (1984).

19. Shlomchik, M. J. *et al. J. exp. Med.* (in the press).
20. Seidman, J. G. & Leder, P. *Nature* **286**, 779–783 (1980).
21. Early, P., Nottenburg, C., Weissman, I. & Hood, L. *Molec. cell. Biol.* **2**, 829–836 (1982).
22. Kelley, D. E. *et al. Molec. cell Biol.* **5**, 1660–1675 (1985).
23. Durdik, J., Moore, M. W. & Selsing, E. *Nature* **307**, 749–752 (1984).
24. Siminovitch, K. A., Bakhshi, A., Goldman, P. & Korsmeyer, S. J. *Nature* **316**, 260–262 (1985).
25. Perlmutter, R. M., Kearney, J. F., Chang, A. P. & Hood, L. E. *Science* **227**, 1597–1600 (1985).
26. Dildrop, R., Krawinkel, V., Winter, E. & Rajewsky, K. *Eur. J. Immun.* **15**, 1154–1156 (1985).
27. Southern, E. M. *J. molec. Biol.* **98**, 503–517 (1975).
28. Marcu, K. B., Banerji, J., Penncavage, N. A., Long, R. & Arnheim, N. *Cell* **22**, 187–196 (1980).
29. Brodeur, P., Riblet, R. *Eur. J. Immun.* **14**, 922–930 (1984).
30. Auffray, C. & Rougeon, F. *Eur. J. Biochem.* **107**, 303–314 (1980).
31. Thiele, C. J., Reynolds, C. P. & Israel, M. A. *Nature* **313**, 404–406 (1985).
32. Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
33. Chirgwin, J., Przybyla, A., MacDonald, R. & Rutter, W. *Biochemistry* **18**, 5294–5299 (1979).
34. Rubin, C. & Schmid, C. *Nucleic Acids Res.* **8**, 4613–4619 (1980).
35. Bencini, D., O'Donovan, G. & Wild, J. *Biotechniques* **4** (January/February 1984).