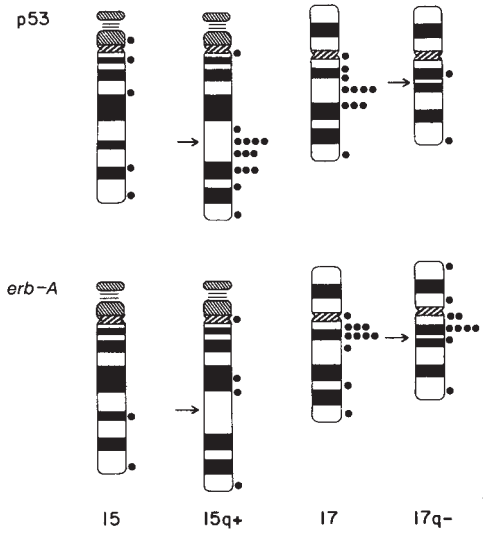

Fig. 2 Distribution of labelled sites on normal chromosome homologues 15 and 17, and on the translocation derivatives 15q+ and 17q−, in metaphase cells from an APL patient (no. 1) with a t(15;17), which were hybridized to the p53 or *erb-A* probes. One hundred metaphase cells were examined for each probe. Arrows identify the breakpoint junctions on the rearranged homologues. The results of these hybridizations indicated that the p53 gene is translocated to chromosome 15, whereas *erb-A* remains on chromosome 17. The radiolabelled probes were nick-translated to specific activities of 2.9×10^7 (p53) and 4.9×10^7 d.p.m. μg^{-1} (*erb-A*). The *erb-A*-specific probe (c-HerbA-1) was a 1.6-kb human *erb-A* cDNA clone[11].

exclude the possibility that the gene is activated by the translocation, as a rearrangement may have occurred outside the regions homologous to the probe (protein-coding region). Such a situation has been reported for the c-*abl* proto-oncogene in the t(9;22) associated with chronic myelogenous leukaemia[16]. Similarly, in Burkitt's lymphoma, rearrangements are usually seen within the introns of the protein-coding regions of c-*myc* in the common t(8;14); however, such rearrangements occur outside this region in the variant translocations t(2;8) and t(8;22)[17,18]. Thus, in the latter translocations, rearrangements of c-*myc* cannot be detected with a probe containing only protein-coding sequences.

The location of two genes (the p53 gene and *erb-A*), each with transforming capabilities, adjacent to the t(15;17) breakpoint raises the question as to which gene is involved in the pathogenesis of APL. In a reciprocal translocation such as t(15;17) there are two rearranged junctions, and there is no *a priori* reason for choosing either the 15q+ or the 17q− chromosome as containing the critical junction. The analysis of complex translocations that involve three chromosomes can reveal which of the junctions that are seen in the simple translocation is conserved. Such an analysis has been performed for t(15;17) as well as t(8;14) and t(9;22)[8]; in the latter two translocations, the 14q+ and 22q− (Philadelphia, Ph^1) chromosomes contained the conserved junctions. Subsequent molecular analysis has shown that each of these chromosomes contains the critical genes [c-*myc* and immunoglobulin heavy chain on 14q+ (refs 17, 18), and c-*abl* and *bcr* (breakpoint cluster region) on Ph^1 (refs 16, 19)]. For t(15;17), the 15q+ chromosome junction was conserved in three clearcut complex translocations[20,21]. In two other reports of variant translocations, however, either the karyotypes permit several interpretations[22] or they have not been published[23].

Thus, most of the data available support the notion that the 15q+ chromosome contains the critical junction, and this would suggest that the p53 gene has a critical involvement in the transformation of normal myeloid cells into leukaemic cells whose maturation is arrested at the promyelocyte stage. Further investigation of the role of the p53 gene in t(15;17) awaits molecular cloning of the junction fragment and analysis of the p53 gene product in these malignant cells.

We thank Carlo M. Croce for the gift of the *erb-A* probe, Rafael Espinosa and Alanna Harden for technical assistance, and Debbie Gifford and Grace Kim for help in preparing the manuscript. This work was supported in part by the US Department of Energy, contract DE-ACO2-80EV10360; USPHS grant CA16910 from the NCI; and by the University of Chicago Cancer Research Foundation (to J.D.R.). M.M.LeB. is a Fellow of the Leukemia Society of America. C.A.W. is a Fellow of the Damon Runyon-Walter Winchell Cancer Foundation. M.O. is a Scholar of the Leukemia Society of America.

Note added in proof: We have recently confirmed the localization of p53 to 17q12-q24 by hybridizing a human p53 cDNA probe to normal metaphase cells. In addition, a smaller yet significant cluster of grains was observed on 17p11-p13; this labelling was not seen with the mouse probe.

Received 9 April; accepted 24 June 1985.

1. Klein, G. (ed.) *Advances in Viral Oncology* Vol. 2 (Raven, New York, 1982).
2. Eliyahu, D., Raz, A., Gruss, P., Givol, D. & Oren, M. *Nature* **312**, 646-649 (1984).
3. Parada, L. F., Land, H., Weinberg, R. A., Wolf, D. & Rotter, V. *Nature* **312**, 649-651 (1984).
4. Harper, M. E. & Saunders, G. F. *Chromosoma* **83**, 431-439 (1981).
5. Le Beau, M. M., Westbrook, C. A., Diaz, M. O. & Rowley, J. D. *Nature* **312**, 70-71 (1984).
6. Rowley, J. D. *et al. Int. J. Cancer* **20**, 869-872 (1977).
7. Larson, R. A. *et al. Am. J. Med.* **76**, 827-841 (1984).
8. Rowley, J. D. *Science* **216**, 749-751 (1982).
9. Zakut-Houri, R. *et al. Nature* **306**, 594-597 (1983).
10. Mitelman, F. *Cytogenet. Cell Genet.* **36**, 1-516 (1983).
11. Dayton, A. I. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 4495-4499 (1984).
12. Sheer, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 5007-5011 (1983).
13. Sheer, D. *et al. Ann. hum. Genet.* (in the press).
14. Matlashewski, G. *et al. EMBO J.* **3**, 3257-3262 (1984).
15. Wolf, D. & Rotter, V. *Proc. natn. Acad. Sci. U.S.A.* **82**, 790-794 (1984).
16. Heisterkamp, N. *et al. Nature* **306**, 239-242 (1983).
17. Croce, C. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 6922-6926 (1983).
18. Erikson, J. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 7581-7585 (1983).
19. Groffen, J. *et al. Cell* **36**, 93-99 (1984).
20. Bernstein, R., Mendelow, B., Morcom, G. & Bezwoda, W. *Br. J. Haemat.* **46**, 311-314 (1980).
21. Ohyashiki, K. *et al. Cancer Genet. Cytogenet.* **14**, 247-255 (1985).
22. Yamada, K. *et al. Cancer Genet. Cytogenet.* **9**, 93-99 (1983).
23. Van den Berghe, H. *et al. Cancer* **43**, 558-562 (1979).

Unusual organization and diversity of T-cell receptor α-chain genes

Adrian C. Hayday*, Don J. Diamond, Gary Tanigawa, Joseph S. Heilig, Virginia Folsom, Haruo Saito & Susumu Tonegawa

Center for Cancer Research and Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

T lymphocytes recognize cell-bound antigens in the molecular context of the self major histocompatibility complex (MHC) gene products through the surface T-cell receptor(s). The minimal component of the T-cell receptor is a heterodimer composed of α and β subunits, each of relative molecular mass (M_r) ~ 45,000 (refs 1–3). Recently, complementary DNA clones encoding these subunits have been isolated and characterized along with that of a third subunit of unknown function, termed γ (refs 4–9). These studies revealed a primary structure for each subunit that was clearly similar to that of immunoglobulin and indicated a somatic rearrangement of corresponding genes that are also immunoglobulin-like. Recently, the analysis of the sequence organization of the T-cell receptor β-chain[10–16] and T-cell-specific γ-chain gene families[17] has been reported. We now present an initial characterization of the murine T-cell receptor α-chain gene family, and conclude that although it is clearly related to the gene families encoding immunoglobulins[18], T-cell receptor β-chains[10–16] and also T-cell γ-chains[17], it shows unique characteristics. There is only a single constant (*C*) region gene segment,

* Present address: Department of Biology, Yale University, New Haven, Connecticut 06511, USA.

which is an exceptionally large distance (~20–40 kilobases (kb) in the cases studied here) from joining (*J*) gene segments. In addition, the *J* cluster and the variable (*V*) segment number seem to be very large. Finally, in the case studied here, a complete α-chain gene shows no somatic mutation and can be assembled directly from V_α, J_α and C_α segments without inclusion of diversity (D_α) segments.

We reported previously[6] the isolation of a cDNA (pHDS58) from a cloned cytotoxic T lymphocyte (2C) that encodes the T-cell receptor α-chain[18]. The sequence of this cDNA clone showed that, by analogy with immunoglobulin genes, the cDNA can be divided into *V*, *J* and *C* regions[6]. The whole α-cDNA, or its fragments containing one or more regions, was used to probe *Eco*RI-digested BALB/c embryo DNA. The results can be summarized as follows (Fig. 1). First, there are at least four fragments (20, 15, 9 and 2.2 kb) carrying V_{2C} or V_{2C}-like sequences, indicating that the pHDS58 *V* region is derived from a family of closely homologous germline *V* DNA segments. Second, the 2C *V* probe hybridized most intensely with the 2.2-kb fragment, suggesting that this fragment contains the germline counterpart of the pHDS58 *V* region. Third, the pHDS58 *J* region is encoded in an *Eco*RI fragment of ~800 base pairs (bp). Fourth, *Eco*RI fragments of 6.2 and 3.9 kb contain the genetic information for the 3′ and 5′ ends of the *C* region, respectively, suggesting that, unlike the β- and γ-genes, the BALB/c embryo contains only one *C* gene segment for the α-chain. In addition, when we analysed BALB/c embryo DNA using *V*-region probes derived from several other α-chain cDNAs, we detected different hybridization patterns composed of multiple bands (data not shown), suggesting that, unlike the β-chain gene family, the α-chain gene family contains many germline V_α DNA segments.

The 2.2-kb *Eco*RI fragment containing the germline counterpart of the pHDS58 V_α region was cloned (clone λ2.2) and sequenced. The sequence is identical to that derived from the corresponding region of pHDS58, indicating that somatic mutation did not occur in the formation of the complete α gene (Fig. 2*a*). The same conclusion has been reached for a β-chain[12] and for T-cell γ genes[17].

DNA clones containing the various genomic DNA sequences of the J_α-C_α region in the germline configuration are summarized in Fig. 3*a* (see legend for details of the derivation of these genomic clones). All genomic clones isolated with phage λ vectors are from BALB/c and the cosmid clone 11E is from C57BL/6. However, the Southern blot analysis (Fig. 3*b*) indicates that there is no gross length polymorphism between these two strains in the relevant region.

We determined the exon–intron structure of the C_α sequence by DNA sequencing (Fig. 2*c*). Apart from the overall homology with the *C*-gene segments of immunoglobulins, T-cell receptor, β-chain and T-cell-specific γ genes, a striking feature of the α-chain *C* region is its shortness. This is primarily reflected in the first exon which encodes a *C* domain of only 87 amino acids. By contrast, the analogous exons of the T-cell receptor β-chain genes and of the immunoglobulin heavy-chain genes and T-cell γ genes are, respectively, 124 (ref. 10) and 110 amino acids[17,19]. The second *C*-segment exon is similar to that of the T-cell receptor β-chain and T-cell γ-gene *C* segments in encoding only a short, cysteine-containing peptide of similar size, but not homologous in sequence, to the exon that encodes the hinge region of IgG[20]. The third exon primarily encodes the putative transmembrane domain and the proposed cytoplasmic tail, and the fourth exon encodes only 3′ untranslated region. Terminal exons that encode only untranslated region are a feature of some class II MHC genes[21], but are without precedent among immunoglobulin, T-cell receptor β-chain[10,11] and T-cell γ genes[17].

The *J*-segment sequence (Fig. 2*b*) derived from clone λ*9.1 (Fig. 3) clearly represents the *J* segment used to assemble the α-chain gene (pHDS58) active in 2C cells. It demonstrates two features that are highly conserved among *J* segments: the nucleotide sequence GGNNNNGGN that encodes 'G—G',

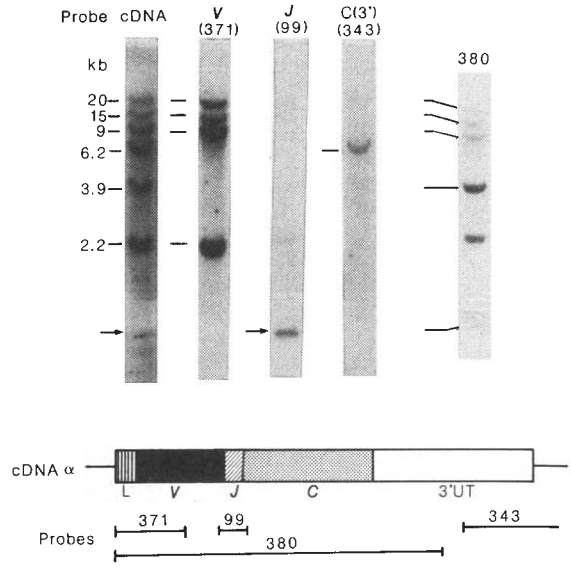

Fig. 1 *Eco*RI-cleaved BALB/c embryo DNA electrophoresed through 0.9% agarose gels, transferred to nitrocellulose (BA85)[32], and hybridized to probes derived from cDNA pHDS58 as indicated in the diagram. Probes were labelled by nick translation[33], and the sizes (kb) of the fragments detected are indicated. All samples, except those probed with probe 380, were run on the same gel, which was subsequently divided into strips. The arrow indicates a fragment of ~800 bp, which is detectable with *J* probe (no. 99) and faintly with other *J*-containing probes. Fragment sizes were determined from co-electrophoresis of phage λ DNA cut with *Hin*dIII and *Sal*I.

which is invariant in functional *J* segments outside the γ-gene family, and the presence 5′ to the *J*-coding region of the conserved nonamer and heptamer sequences implicated in *V*–*J* rearrangement (Fig. 4)[22]. The sequences, separated by 12 bp, are essentially complementary to a heptamer and nonamer that occur immediately 3′ to the V_α segment separated by 23 bp (Figs 2*a*, 4). The spacings of 23 bp in the sequences 3′ to V_α and of 12 bp in the sequences 5′ to J_α are common among the three T-cell-specific rearranging genes, α, β and γ, and suggest (according to an empirical rule[23,24]) that these *V* segments can join directly to the respective *J* segments without the involvement of the third (*D*) segment (see below).

By the criteria of the conserved 'G–G' coding triplets and the conserved heptamer and nonamer sequences, we searched for additional *J* segments and found a second (J_1) *J* segment >700 bp downstream of the first one, and a third one (J_2) ~0.5 kb downstream of the second (Figs 2*b*, 3*a*). A fourth, J_{C10} used by the α-chain of a helper T-cell hybridoma, was mapped by Southern blotting to ~4.0 kb downstream of J_2. Three more *J* segments (J_{C11}, J_{B8C3} and J_{DFL12}) were mapped on another genomic clone, λ8.2, whose DNA occupies a region upstream of the C_α gene segments (Fig. 3*a*). Thus, we have mapped seven different J_α segments over a distance of ~20 kb. More *J* segments may exist in the 19-kb region between J_{DFL} and C_α, and possibly also 5′ to J_{2C}. By comparison, the two T-cell receptor J_β clusters are spread over 1.1 and 1.9 kb (refs 10, 12), the immunoglobulin J_H cluster over 1.36 kb (refs 23, 25) and the immunoglobulin J_κ cluster over 1.4 kb (refs 22, 26). In each of these cases the *J* segments are relatively evenly spaced ~50–500 bp apart. Given that the first seven J_α segments we have identified were all different and that our analysis already points to clustering of some J_α segments (J_{2C}, J_1 and J_2; and J_{C11}, J_{B8C31} and J_{DFL}), the α-gene family must contain many more J_α segments. If there are 20 J_α segments, the probability that the first seven analysed would be different is significant (30%). Our analysis of *J* segments shows those used by cytotoxic T lymphocytes and by T-helper cells to be interspersed one with another, as with T-cell receptor β-chain segments[10,12], implying

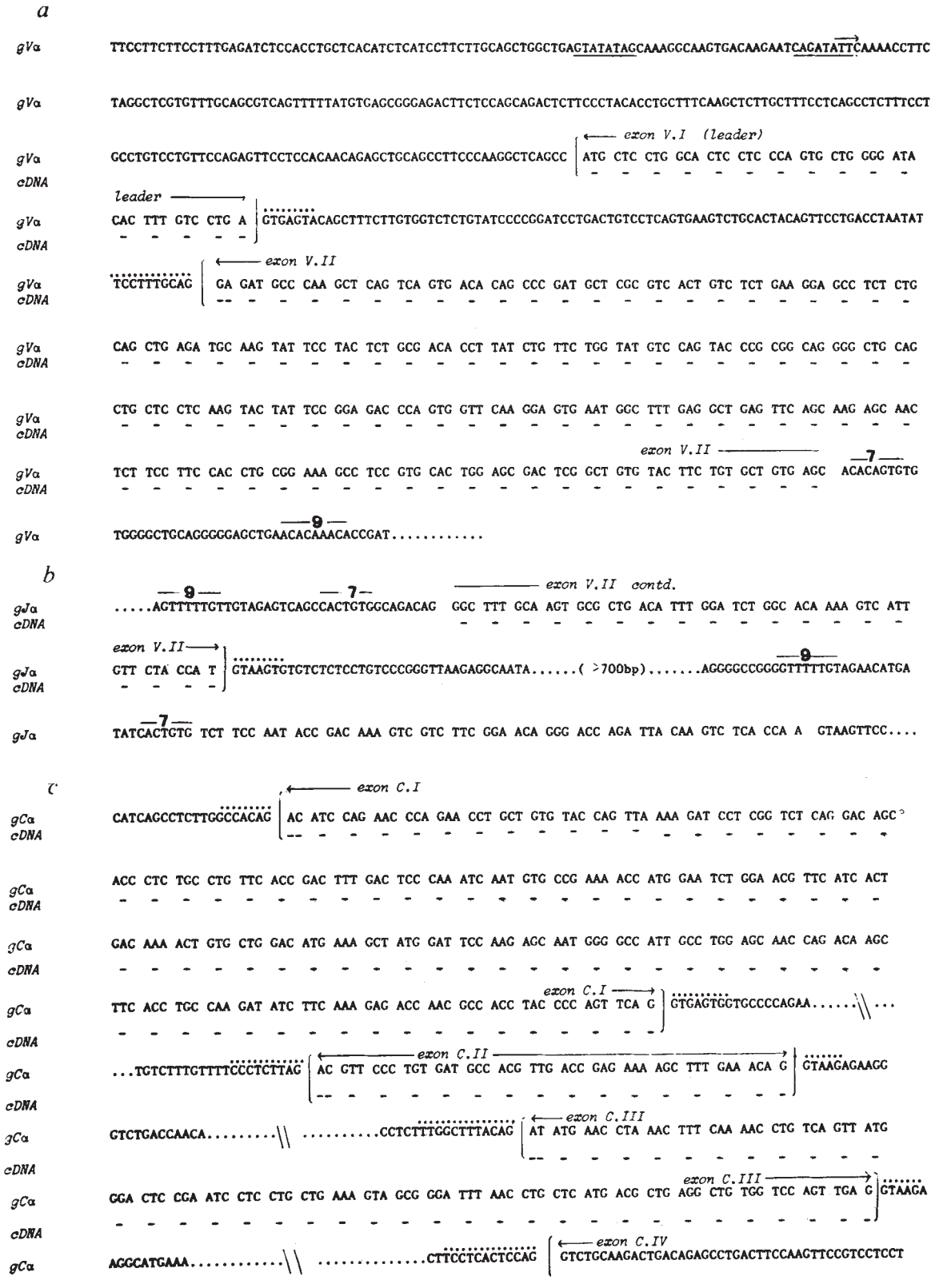

Fig. 2 Comparison of T-cell receptor α-chain sequences and cDNA α-sequences. *a*, Comparison of germline *V* sequence (gV_α) determined using clone λ2.2 as described in ref. 34 and cDNA pHDS58 sequence[6] throughout the *V* segment. The perfect homology of the two sequences is indicated (for the protein coding regions only) by dashes, but extended upstream of that for all the available pHDS58 sequence[6]. A candidate transcriptional start site is superscripted with an arrow, being within a sequence (underlined) resembling Py---PyAPyPyPyPy[35]. Between 20 and 30 bp upstream of this is a candidate TATA box (underlined)[35]. The exon *V.I* denotes the putative leader region although the transcript of this region is contiguous with that of upstream untranslated sequence. The splice donors and receptors are overlined with dots. The proposed heptamer and nonamer *V* segment signal sequences at the 3′ end of *V* are indicated. *b*, Comparison of germ line J_α sequence (gJ_α) and pHDS58 sequence through the *J* segment. The J_α segment completes the α-chain *V* region, that is, exon *V.II*. 3′ of this is shown sequence of another putative J_α segment (J_1) from downstream in the germline. *c*, Comparison of germline C_α sequence (gC_α) and pHDS58 sequence throughout the *C* segment. Homology is shown by dashes only for the protein coding sequences. Exon *C.IV* encodes only untranslated RNA.

The germline sequence of this continues for over 500 bp with the same sequence as was determined for pHDS58 (data not shown).

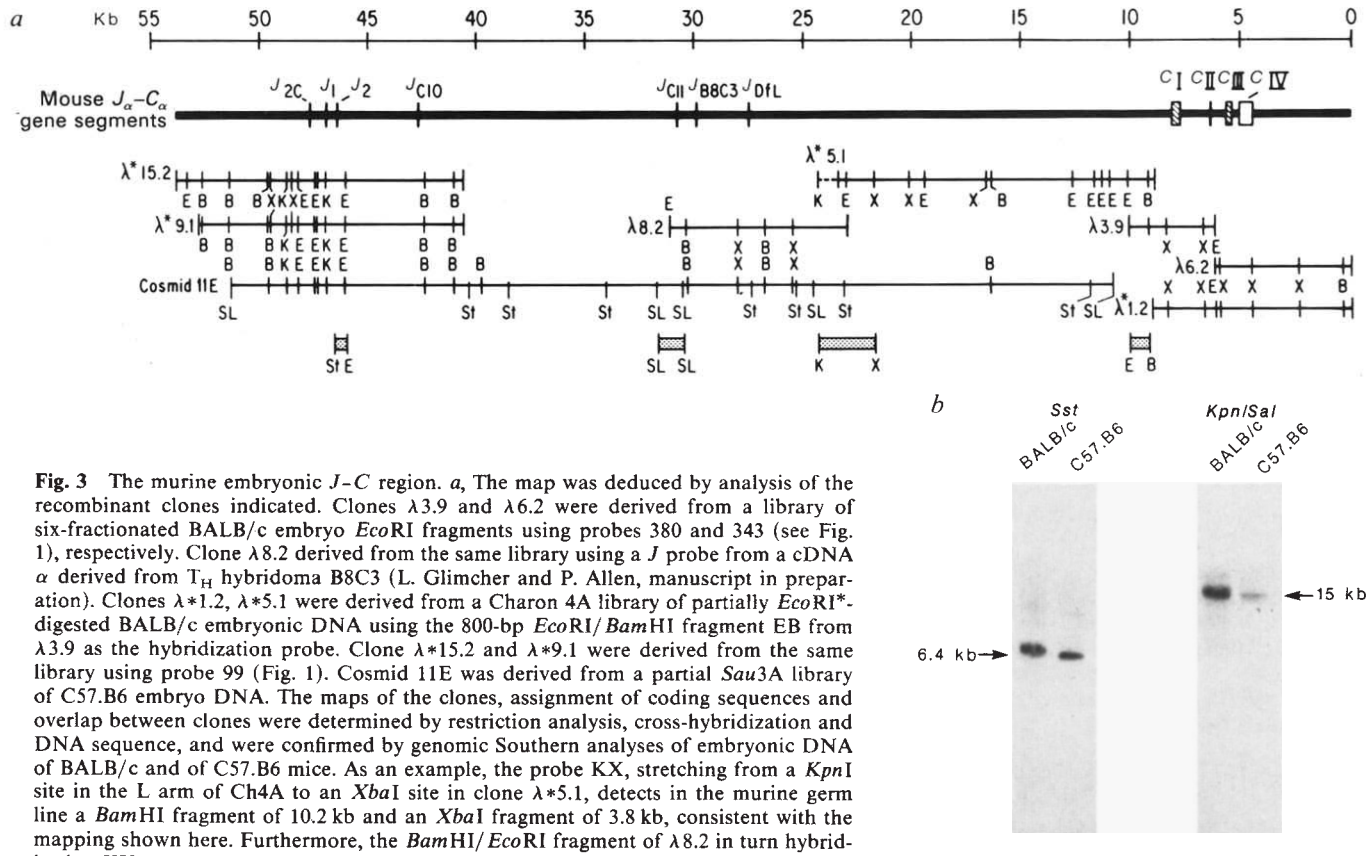

Fig. 3 The murine embryonic *J–C* region. *a*, The map was deduced by analysis of the recombinant clones indicated. Clones λ3.9 and λ6.2 were derived from a library of six-fractionated BALB/c embryo *Eco*RI fragments using probes 380 and 343 (see Fig. 1), respectively. Clone λ8.2 derived from the same library using a *J* probe from a cDNA α derived from T_H hybridoma B8C3 (L. Glimcher and P. Allen, manuscript in preparation). Clones λ*1.2, λ*5.1 were derived from a Charon 4A library of partially *Eco*RI*-digested BALB/c embryonic DNA using the 800-bp *Eco*RI/*Bam*HI fragment EB from λ3.9 as the hybridization probe. Clone λ*15.2 and λ*9.1 were derived from the same library using probe 99 (Fig. 1). Cosmid 11E was derived from a partial *Sau*3A library of C57.B6 embryo DNA. The maps of the clones, assignment of coding sequences and overlap between clones were determined by restriction analysis, cross-hybridization and DNA sequence, and were confirmed by genomic Southern analyses of embryonic DNA of BALB/c and of C57.B6 mice. As an example, the probe KX, stretching from a *Kpn*I site in the L arm of Ch4A to an *Xba*I site in clone λ*5.1, detects in the murine germ line a *Bam*HI fragment of 10.2 kb and an *Xba*I fragment of 3.8 kb, consistent with the mapping shown here. Furthermore, the *Bam*HI/*Eco*RI fragment of λ8.2 in turn hybridized to KX, as well as to the 400-bp fragment that lies between the left-most *Eco*RI site in the insert of λ*5.1, and the L-arm. Restriction sites are: E (*Eco*RI), B (*Bam*HI), X (*Xba*I), K (*Kpn*I), SL (*Sal*I) and St (*Sst*). Maps are not necessarily complete, notably for *Eco*RI and *Xba*I on cosmid 11E. In addition to B8C3, J_α segments were localized for T_H hybridomas C10 and C11 (Glimcher and Allen, in preparation), and for cytotoxic T-lymphocyte clone DFL.12 (ref. 36). Exons are boxed and shaded if translated. Germline $V_{2C\alpha}$, cloned in λV2.2, lies upstream at an as yet undetermined distance from J_{2C}. *b*, Southern analysis of BALB/c and C57.B6 genomic DNA using cosmid 11E-derived probes. Embryonic DNA from BALB/c and C57.B6 mice was co-electrophoresed on 1.2% agarose gels after restriction with *Sst* (St) or *Kpn* (K) and *Sal* (SL). Southern transfers of the electrophoresed DNA were probed with either the 1.2-kb *Sal* fragment (dotted line in *a*) for St-cut DNA or the 0.6-kb St-E fragment (dotted line in *a*) in the case of SL–K-cut DNA. Methods are further described in Fig. 1 legend.

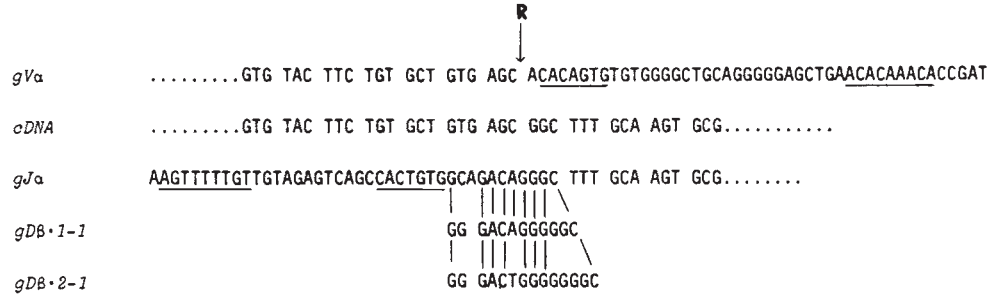

Fig. 4 Comparison of germline V_α(gV_α), germline J_α(gJ_α) and cDNA pHDS58 sequences indicates recombination of V_α to J_α at point R. The germline g*J* has homology at its 5′ end to sequences of Dβ.1-1 and Dβ.2-1 (refs 13, 14). The proposed heptamer and nonamer sequences 3′ to *V* and 5′ to *J*, respectively, are underlined.

that there is no functional clustering of *J* segments in either gene family.

One consequence of the large *J* cluster is the presence of an unusually long intron between *V*-(*D*)-*J* and *C* segments in at least some assembled α genes. For example, in the 2C α gene this intron is ~40 kb, which is without precedent among immunoglobulin or immunoglobulin-like genes characterized to date. Very large introns have recently been discovered in the *Ubx* and *Antp* loci of *Drosophila melanogaster*[27]: in *Drosophila*, the large intron may be differentially processed, whereas here the primary function of the large intron may be to accommodate many *J* segments. Because of the similarity between both the gene segments that comprise T-cell receptor and immunoglobulin genes and the overall organization of these gene families, transcriptional activity at the promoter of a rearranged V_α gene segment may be enhanced in an analogous manner to that of rearranged immunoglobulin V_H and V_κ segments. However, if a tissue-specific enhancer (analogous to the immunoglobulin enhancer) does exist in the *J–C* intron of the T-cell receptor α-chain genes, it must be activating the promoter associated with the rearranged V_α gene segment at a distance of at least 20 kb in 2C cells. Alternatively, the existence of such a large J_α cluster may have excluded the activation of V_α segment transcription by this mechanism.

One interesting feature of the J_{2C} segment is its unusual length of 63 bp (Fig. 2*b*)[10,12,25]. At its 5′ end, the *J* sequence GCAGACAGGGC resembles the sequences of the T-cell receptor β-chain gene *D* segment[13,14] (Fig. 4). Thus, the J_{2C} segment can be regarded as a *D*-type element of 11 bp pre-fused in the germ line to a *J*-type element of 52 bp. In both IgH and T-cell receptor β-gene families, single *D* elements occur in relatively close proximity to the respective *J* clusters[13,14,28]. It is possible that the analogous D_α segment does not exist, having already fused to J_α to generate the J_{2C} segment in the germ line. The implication is that, albeit at low frequency, cells other than lymphocytes can effect *D–J* recombination.

Another interesting feature of the J_{2C} segment is its direct joining with the germline V_{2C} segment in 2C cells. The fusion of the germline V_{2C} and J_{2C} segments should create a novel V_α probe reactive *Eco*RI fragment of 1.9 kb in 2C cells (probe 371)

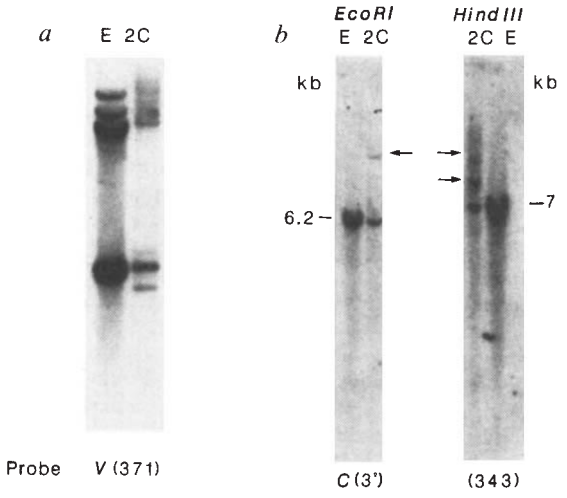

Fig. 5 *a*, *Eco*RI-cleaved DNA of BALB/c embryo, E and cytotoxic T-lymphocyte clone 2C cells, electrophoresed through 0.9% agarose, and analysed as described in Fig. 1. *b*, *Eco*RI-cleaved and *Hin*dIII-cleaved DNA of BALB/c embryo, E and 2C cells, analysed as described in Fig. 1 and hybridized with probe 343 (see Fig. 1). Bands strongly detected in 2C DNA but not in embryo DNA are arrowed.

that would not be present either in germline DNA, B-cell DNA or DNA of T cells that have not made this particular *V-J* combination. Such a fragment is detected in 2C DNA (Fig. 5*a*). In IgH, T-cell receptor β-chain or T-cell γ-gene families, extra nucleotides are commonly found between the fused *V* and *J* segments[18]. These insertions are either template-independent or the result of the participation of a third type of germline DNA segment, *D*, in the fusion process[13,14,28,29]. Although the 12/23-bp spacer rule permits direct *V-J* joining for the T-cell β-receptor and T-cell γ-gene families, no such case has yet been clearly demonstrated. By contrast, the germline V_{2C} and J_{2C} sequences presented here can fully account for the sequence of pHDS58 in the *J* region (Fig. 4), strongly suggesting that a direct *V-J* joining has taken place. However, α-cDNA sequences derived from other cytotoxic T-lymphocyte clones and helper T-cell hybridomas (unpublished data) as well as the sequences of other germline J_α segments strongly suggest that complete α genes are also assembled with the participation of *D* segments.

Probe 343 (Fig. 1) derives solely from the 3′ end of C_α, and detects in BALB/c embryo DNA a single *Eco*RI fragment of 6.2 kb and a single *Hin*dIII fragment of ~7 kb. Surprisingly, when this probe is directed against *Eco*RI- or *Hin*dIII-cleaved 2C cell DNA, it detects additional fragments (Fig. 5*b*) which, because of their reproducibility, are apparently not caused by partial digestion of 2C cell DNA. The nature of the gene rearrangement in 2C cells relative to the BALB/c germ line that generates these novel fragments is unclear. Candidates include chromosomal rearrangements 3′ to the C_α locus, or amplification of germline sequences which, when in single copy, are undetectable with probe 343. Note that in the experiments shown in Fig. 5*b*, an excess of digested BALB/c embryo DNA was loaded in an effort to detect in the germline single-copy, unrearranged versions of the novel bands present in 2C cells.

In conclusion, our characterization of murine gene segments encoding the T-cell receptor α-chain indicates a multi-gene family that is comprised of many *V*-gene segments, a large array of *J*-gene segments and a single C_α gene. Furthermore, our data strongly suggest that a given pair of V_α and J_α segments can join either directly, without the participation of a D_α segment, or indirectly, through one or more D_α segment(s). Thus, the scope for combinatorial diversity of the α-gene family seems to be significantly greater than that of IgH, Igκ and T-cell receptor β- and γ-gene families. By contrast, the data presented here for the 2C α gene and those reported earlier by others for β genes[12] and by us for γ genes[17] did not give any indication of somatic mutation, an important source of diversity for immunoglobulin genes[18]. This apparent lack of somatic mutation in T-cell receptor genes may be related to the relatively large combinatorial diversity seen in the α-gene family, the latter partially compensating for the former. However, apparent somatic mutation of both α- and β-chain T-cell receptor genes was noted on *in vitro* passage and subcloning of an interleukin-2-secreting T-cell hybridoma[30]. The role of somatic mutation in diversifying the T-cell receptor repertoire[31] remains an open question.

We thank Howard Siegel for technical assistance and Eleanor Basel for preparation of the manuscript. We also thank Drs D. M. Kranz, H. N. Eisen, L. Glimcher and W. Häas for T-cell clones and hybridomas. D.D. is a fellow of the Leukemia Society; J.S.H. is supported by an NIH predoctoral training grant 2T32CA09255; V.F. is supported by a fellowship from the Cancer Research Institute. The work was supported by AI-17879, CA-28900 (S.T.) and CA14051 (a core grant to S. Luria).

Note added in proof: After submission of this manuscript, we learned that two other groups have also found an unusually large number of germline *J* DNA segments[37,38].

Received 30 April; accepted 16 July 1985.

1. Allison, J. P., McIntyre, B. W. & Bloch, D., *J. Immun.* **129**, 2293-2300 (1982).
2. Haskins, K. *et al. J. exp. Med.* **157**, 1149-1169 (1983).
3. Meuer, S. C. *et al. J. exp. Med.* **157**, 705-719 (1983).
4. Hedrick, S. M. *et al. Nature* **308**, 153-158 (1984).
5. Yanagi, Y. *et al. Nature* **308**, 145-149 (1984).
6. Saito, H. *et al. Nature* **312**, 36-39 (1984).
7. Chien, Y.-H. *et al. Nature* **312**, 31-35 (1984).
8. Sim, G. K. *et al. Nature* **312**, 771-775 (1984).
9. Saito, H. *et al. Nature* **309**, 757-762 (1984).
10. Gascoigne, N. *et al. Nature* **310**, 387-391 (1984).
11. Mallissen, M. *et al. Cell* **37**, 1101-1110 (1984).
12. Chien, Y.-H., *et al. Nature* **309**, 322-326 (1984).
13. Kavaler, J., Davis, M. M. & Chien, Y.-H., *Nature* **310**, 421-423 (1984).
14. Siu, G. *et al. Nature* **311**, 344-350 (1984).
15. Siu, G. *et al. Cell* **37**, 393-401 (1984).
16. Sims, J. E., Tunnacliffe, A., Smith, W. J. & Rabbitts, T. H. *Nature* **312**, 541-545 (1984).
17. Hayday, A. C. *et al. Cell* **40**, 259-269 (1985).
18. Tonegawa, S. *Nature* **302**, 575-581 (1983).
19. Kabat, E. A., Wu, T. T., Bilofsky, H., Reid-Miller, M. & Perry, H. *US Dept Hlth hum. Serv. Publ.* 80-2008 (1983).
20. Sakano, H. *et al. Nature* **277**, 627-633 (1979).
21. Steinmetz, M. & Hood, L. *Science* **222**, 727-733 (1983).
22. Sakano, H., Hüppi, K., Heinrich, G. & Tonegawa, S. *Nature* **280**, 288-294 (1979).
23. Sakano, H., Maki, R., Kurasawa, Y., Roeder, W. & Tonegawa, S. *Nature* **286**, 676-683 (1980).
24. Early, P. *et al. Cell* **19**, 981-992 (1980).
25. Tonegawa, S. *et al. Cold Spring Harb. Symp. quant. Biol.* **40**, 839-857 (1981).
26. Max, E. E., Seidman, J. G. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3450-3454 (1979).
27. Gehring, W. J. *Cell* **40**, 3-5 (1985).
28. Sakano, H., Kurosawa, Y., Weigert, M. & Tonegawa, S. *Nature* **290**, 562-565 (1981).
29. Kurosawa, Y. *et al. Nature* **290**, 565-570 (1981).
30. Augustin, A. & Sim, G. K. *Cell* **39**, 5-12 (1984).
31. Jerne, N. K. *Eur. J. Immun.* **1**, 1-9 (1971).
32. Southern, E. *J. molec. Biol.* **98**, 503 (1975).
33. Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. molec. Biol.* **113**, 237 (1977).
34. Maxam, A. & Gilbert, W. *Meth. Enzym.* **65**, 499 (1980).
35. Breathnach, R. & Chambon, P. *A. Rev. Biochem.* **50**, 349 (1981).
36. Haas, W., Mathur-Rochat, J., Pohlit, H., Nabholz, M. & von Boehmer, H. *Eur. J. Immun.* **10**, 828-834 (1980).
37. Yoshikai, Y. *et al. Nature* **316**, 837-840 (1985).
38. Winoto, A. *et al. Nature* **316**, 832-836 (1985).

Genomic organization of the genes encoding mouse T-cell receptor α-chain

Astar Winoto, Shelley Mjolsness & Leroy Hood

Division of Biology, California Institute of Technology, Pasadena, California 91125, USA

The vertebrate immune system uses two kinds of antigen-specific receptors, the immunoglobulin molecules of B cells and the antigen receptors of T cells. T-cell receptors are formed by a combination of two different polypeptide chains, α and β (refs 1–3). Three related gene families are expressed in T cells, those encoding the T-cell receptor, α and β, and a third, γ (refs 4–6), whose function is unknown. Each of these polypeptide chains can be divided into variable (V) and constant (C) regions. The V_β regions are encoded by V_β, diversity (D_β) and joining (J_β) gene segments that rearrange