Cell, Vol. 59, 859–870, December 1, 1989, Copyright © 1989 by Cell Press

Junctional Sequences of T Cell Receptor γδ Genes: Implications for γδ T Cell Lineages and for a Novel Intermediate of V-(D)-J Joining

Juan J. Lafaille, Amy DeCloux, Marc Bonneville, Yohtaroh Takagaki, and Susumu Tonegawa
Howard Hughes Medical Institute at
Center for Cancer Research
and Department of Biology
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Summary

Nucleotide sequences of a large number of V-(D)-J junctions of T cell receptor (TCR) γ and δ genes show that most fetal thymocytes express on their surface one of just two γδ TCRs known to be expressed by epidermal γδ T cells (s-IEL) or intraepithelial γδ T cells associated with female reproductive organs (r-IEL). In contrast, γδ TCRs expressed on adult thymocytes are highly diverse as a result of multiple combinations of gene segments as well as junctional deletions and insertions, indicating that developmental time– and cell lineage–dependent mechanisms exist that control the extent of γδ TCR diversity. In addition, this study revealed a new type of junctional insertion (P nucleotides), which led to a new model of V-(D)-J joining generally applicable to immunoglobulin and TCR genes.

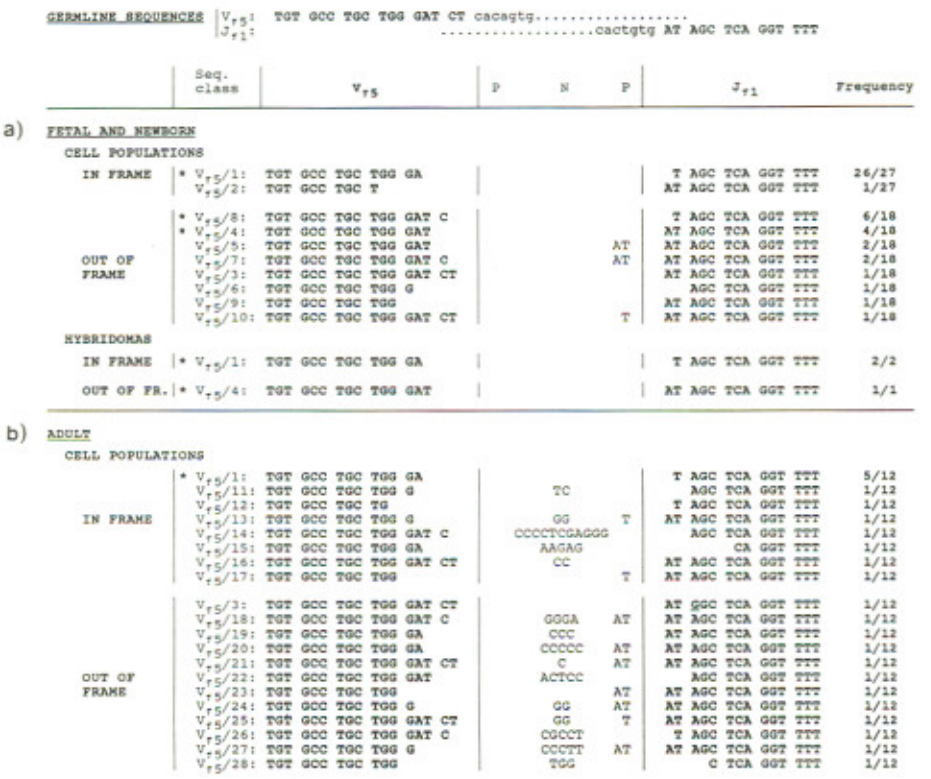
Introduction

T cell receptors (TCRs) play a pivotal role in the immune system by recognizing and distinguishing diverse antigens. Until recently all TCRs were thought to be made up with a heterodimer composed of α and β subunits (Allison et al., 1982; Haskins et al., 1983; Meuer et al., 1983). However, during the search for the genes encoding these TCR subunits, a third gene called γ, which shares a number of characteristics with the TCR α and β genes, was discovered (Saito et al., 1984; Kranz et al., 1985; Hayday et al., 1985). Subsequent studies established that this gene encodes one polypeptide chain of the second TCR, which is composed of the heterodimer γδ (Brenner et al., 1986; Bank et al., 1986; Weiss et al., 1986). Mouse and human γδ TCRs are expressed on a relatively small fraction (less than 5%) of thymocytes and T cells of peripheral lymphoid organs (spleen, lymph nodes, etc.; Bluestone et al., 1987; Borst et al., 1987; Bottino et al., 1988; Itohara et al., 1989). In contrast, T cells bearing the γδ TCR are relatively abundant in a variety of organs containing epithelial cells, such as skin (Koning et al., 1987; Kuziel et al., 1987), small intestine (Goodman and Lefrancois, 1988; Bonneville et al., 1988), and reproductive organs (uterus, vagina, etc.; Itohara et al., submitted). These γδ T cells are mostly in contact with epithelial cells, and we recently proposed to designate these T cells as s-IEL, i-IEL, r-IEL, and t-IEL for these intraepithelial lymphocytes in skin, intestine, reproductive organs, and tongue, respectively (Itohara et al., submitted). The function of γδ T cells is currently unknown, but it is strongly suspected that they play a role in the protection of epithelia (Janeway et al., 1988; Bonneville et al., 1988).

Both γ and δ genes, like immunoglobulin and TCR α and β genes (Tonegawa, 1983; Davis and Bjorkman, 1988), are transmitted genetically as gene segments referred to as V, J, and C or V, D, J, and C, which must be assembled by somatic rearrangement (V-J or V-D-J joining) (Saito et al., 1984; Hayday et al., 1985; Chien et al., 1987a) that generates structural diversity in their products. Three types of diversity were recognized that are attributed to the somatic rearrangement: combinatorial diversity generated by the combination of different gene segments, junctional site diversity generated by the action of a DNA exonuclease that nibbles the terminals of the recombining gene segments, and junctional insertion diversity generated by a terminal transferase–like enzyme that adds template-independent nucleotides, called N, to these terminals (Tonegawa, 1983).

The γδ TCRs expressed on adult thymocytes are diverse (Korman et al., 1988; Takagaki et al., 1989a). However, the γ chains of i-IEL γδ TCR largely lack combinatorial diversity since they are primarily encoded by the $V_7J_1C_1$ γ gene (Bonneville et al., 1988; see Heilig and Tonegawa, 1987, for the nomenclature of mouse γ gene segments), although i-IEL γδ TCR as a whole is diverse as a result of the high degree of variability in the δ subunit (Takagaki et al., 1989b). In contrast to these γδ T subpopulations with diverse γδ TCRs, γδ TCRs borne by the s-IEL and r-IEL are remarkably homogeneous, as they are encoded by the $V_5J_1C_1$ γ and $V_1D_2J_2C$ δ genes (Asarnow et al., 1988; see Elliott et al., 1988, for the nomenclature of δ gene segments) and by the $V_6J_1C_1$ γ and $V_1D_2J_2$ C δ genes, respectively (Itohara et al., submitted), none of which exhibit junctional diversity. Another γδ T cell subpopulation with TCRs of limited structural diversity is composed of fetal thymocytes, which appear in two distinct waves in the developing mouse thymus. The first thymocyte wave peaks around the 15th day of gestation, subsides by the 17th day, and carries TCR encoded by the $V_5J_1C_1$ γ and V_1DJ_2C δ genes (Havran and Allison, 1988; Ito et al., 1989), the same combination of γ and δ genes utilized by s-IEL (Asarnow et al., 1988). The second γδ thymocyte wave, which starts around the 17th day, peaks at birth, and declines thereafter, carries TCR encoded by the $V_6J_1C_1$ γ and V_1DJ_2C δ genes (Ito et al., 1989), the genes that also encode the γδ TCR of r-IEL (Itohara et al., submitted). Whether or not junctional diversity occurs in these fetal γ and δ genes is an interesting, unresolved issue relevant to the possible developmental relationship between the fetal thymocyte populations and s-IEL or r-IEL.

In this study we used the polymerase chain reaction (PCR) method (Saiki et al., 1988) to determine the nucleotide sequences of a large number of V-J γ and V-D-J δ junctions present in both fetal and adult thymocyte populations. The results indicated that the rearranged γ and δ



	Seq. class	$V_{\gamma 5}$	P	N	P	$J_{\gamma 1}$	Frequency
GERMLINE SEQUENCES	$V_{\gamma 5}$:	TGT GCC TGC TGG GAT CT cacagtg.....					
	$J_{\gamma 1}$:	.....cactgtg AT AGC TCA GGT TTT					
a) FETAL AND NEWBORN							
CELL POPULATIONS							
IN FRAME	* $V_{\gamma 5}$/1:	TGT GCC TGC TGG GA				T AGC TCA GGT TTT	26/27
	$V_{\gamma 5}$/2:	TGT GCC TGC T				AT AGC TCA GGT TTT	1/27
OUT OF FRAME	* $V_{\gamma 5}$/8:	TGT GCC TGC TGG GAT C				T AGC TCA GGT TTT	6/18
	* $V_{\gamma 5}$/4:	TGT GCC TGC TGG GAT				AT AGC TCA GGT TTT	4/18
	$V_{\gamma 5}$/5:	TGT GCC TGC TGG GAT			AT	AT AGC TCA GGT TTT	2/18
	$V_{\gamma 5}$/7:	TGT GCC TGC TGG GAT C			AT	AT AGC TCA GGT TTT	2/18
	$V_{\gamma 5}$/3:	TGT GCC TGC TGG GAT CT				AT AGC TCA GGT TTT	1/18
	$V_{\gamma 5}$/6:	TGT GCC TGC TGG G				AGC TCA GGT TTT	1/18
	$V_{\gamma 5}$/9:	TGT GCC TGC TGG				AT AGC TCA GGT TTT	1/18
	$V_{\gamma 5}$/10:	TGT GCC TGC TGG GAT CT			T	AT AGC TCA GGT TTT	1/18
HYBRIDOMAS							
IN FRAME	* $V_{\gamma 5}$/1:	TGT GCC TGC TGG GA				T AGC TCA GGT TTT	2/2
OUT OF FR.	* $V_{\gamma 5}$/4:	TGT GCC TGC TGG GAT				AT AGC TCA GGT TTT	1/1
b) ADULT							
CELL POPULATIONS							
IN FRAME	* $V_{\gamma 5}$/1:	TGT GCC TGC TGG GA				T AGC TCA GGT TTT	5/12
	$V_{\gamma 5}$/11:	TGT GCC TGC TGG G		TC		AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/12:	TGT GCC TGC TG				T AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/13:	TGT GCC TGC TGG G		GG	T	AT AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/14:	TGT GCC TGC TGG GAT C		CCCCTCGAGGG		AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/15:	TGT GCC TGC TGG GA		AAGAG		CA GGT TTT	1/12
	$V_{\gamma 5}$/16:	TGT GCC TGC TGG GAT CT		CC		AT AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/17:	TGT GCC TGC TGG			T	AT AGC TCA GGT TTT	1/12
OUT OF FRAME	$V_{\gamma 5}$/3:	TGT GCC TGC TGG GAT CT				AT GGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/18:	TGT GCC TGC TGG GAT C		GGGA	AT	AT AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/19:	TGT GCC TGC TGG GA		CCC		AT AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/20:	TGT GCC TGC TGG GA		CCCCC	AT	AT AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/21:	TGT GCC TGC TGG GAT CT		C	AT	AT AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/22:	TGT GCC TGC TGG GAT		ACTCC		AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/23:	TGT GCC TGC TGG			AT	AT AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/24:	TGT GCC TGC TGG G		GG	AT	AT AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/25:	TGT GCC TGC TGG GAT CT		GG	T	AT AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/26:	TGT GCC TGC TGG GAT C		CGCCT		T AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/27:	TGT GCC TGC TGG G		CCCTT	AT	AT AGC TCA GGT TTT	1/12
	$V_{\gamma 5}$/28:	TGT GCC TGC TGG		TGG		C TCA GGT TTT	1/12

Figure 1. Nucleotide Sequences of $V_{\gamma 5}$-$J_{\gamma 1}$ Junctions from Fetal, Newborn, and Adult Thymocytes

Fetal thymocytes were from day 14.5, day 16, and day 17.5. For the thymocyte populations, the sequences were generated from genomic DNA, whereas for the hybridomas they were generated from cDNA. The hybridomas were prepared from thymocytes of C57BL/6 embryos of day 15.5 or 17.5 of gestation. Asterisks indicate the sequences present in the s-IEL clones characterized by Asarnow et al. (1988). See text for definition of P nucleotides. In Figures 1–5, a few nucleotides that can be assigned to either recombining gene segment were arbitrarily placed under one of the two gene segments. The frequency with which a particular sequence was found among DNA clones is listed in the last column.

genes present in fetal thymocytes, in contrast to those in adult thymocytes, have very limited junctional diversity, and that the early and late waves of fetal γδ thymocytes are most probably precursors of s-IEL and r-IEL, respectively. This study provided evidence for the occurrence of a hitherto undescribed type of intermediate of V-J and V-D-J joining that explains the origin of the recurrent, non-germline-coded nucleotides observed in the V-J or V-D-J junctions of some TCR and immunoglobulin genes (Hayday et al., 1985; Wysocki et al., 1986; Milner et al., 1986; Traunecker et al., 1986; McCormack et al., 1989).

Results

Fetal $V_5J_1C_1$ γ Genes Have Highly Homogeneous Junctions and Code for s-IEL γ Chains

We extracted DNA from mouse fetal and newborn C57BL/6 thymocytes, amplified the V_5J_1 γ junctional sequences by the PCR method, cloned the amplified DNA, and determined their nucleotide sequences. The nucleotide sequences in Figure 1a are classified—as in all subsequent figures—according to whether the junctional sequence permits the J segment to be translated in the required frame (in-frame junction) or not (out-of-frame junction). The extra junctional nucleotides (heretofore called N) are labeled N and P in all figures for reasons explained below. The V_5J_1 γ junctional sequences are very homogeneous: 26 out of 27 in-frame junctions are identical to the s-IEL canonical sequence (Asarnow et al., 1988). This same sequence is also expressed on both of the V_5J_1 γ–expressing T hybridomas derived from fetal thymocytes (Figure 1a). The data strongly suggest that the V_5J_1 γ gene–expressing fetal thymocytes are the precursors of s-IEL.

The sequence variability among the out-of-frame junctions is also limited, but not to the same extent as the in-frame junctions. The lack of junctional sequence variability among this population of thymocytes is in sharp contrast to the extensive diversity observed among the γ genes expressed on the $\gamma\delta^+$ T hybridomas derived from adult thymocytes, which preferentially utilize V_4J_1 and V_7J_1 (Takagaki et al., 1989a; Korman et al., 1988).

Fetal and Newborn $V_6J_1C_1$ γ Genes Have Highly Homogeneous Junctions and Code for r-IEL γ Chains

We carried out a similar analysis for the second major type of γ genes expressed on fetal thymocytes, $V_6J_1C_1$. γ genes of this segmental composition have recently been shown to encode the majority of the γ chains expressed on r-IEL (Itohara et al., submitted). As shown in Figure 2a, the junc-

GERMLINE SEQUENCES V_{T6}: TGT GCA TGC TGG GAT A cactcta.....
J_{T1}: .....cactgtg AT AGC TCA GGT TTT

		Seq. class	V_{T6}	P	N	P	J_{T1}	Frequency
a)	FETAL AND NEWBORN							
	CELL POPULATIONS							
	IN FRAME	*	V_{T6}/1: TGT GCA TGC TGG GA		A		T AGC TCA GGT TTT	23/25
			V_{T6}/2: TGT GCA TGC TGG GA		A		AGC TCA GGT TTT	1/25
			V_{T6}/3: TGT GCA T				AT AGC TCA GGT TTT	1/25
	OUT OF FRAME		V_{T6}/4: TGT GCA TGC TGG GAT A				T AGC TCA GGT TTT	12/14
			V_{T6}/5: TGT GCA TGC TGG GAT				C TCA GGT TTT	1/14
			V_{T6}/6: TGT GCA TGC TGG GAT A			T	AT AGC TCA GGT TTT	1/14
	HYBRIDOMAS							
	IN FRAME	*	V_{T6}/1: TGT GCA TGC TGG GA		A		T AGC TCA GGT TTT	15/16
			V_{T6}/2: TGT GCA TGC TGG GA		A		AGC TCA GGT TTT	1/16
b)	ADULT							
	CELL POPULATIONS							
	IN FRAME	*	V_{T6}/1: TGT GCA TGC TGG GA			AT	T AGC TCA GGT TTT	4/11
			V_{T6}/7: TGT GCA TGC TGG GA			AT	AT AGC TCA GGT TTT	3/11
			V_{T6}/2: TGT GCA TGC TGG GA		A		AGC TCA GGT TTT	2/11
			V_{T6}/8: TGT GCA TGC TGG GAT A	T			T AGC TCA GGT TTT	1/11
			V_{T6}/9: TGT GCA TGC TGG GAT		TT	AT	AT AGC TCA GGT TTT	1/11
	OUT OF FRAME		V_{T6}/4: TGT GCA TGC TGG GAT A				T AGC TCA GGT TTT	4/19
			V_{T6}/10: TGT GCA TGC TGG G		GT	AT	AT GGC TCA GGT TTT	1/19
			V_{T6}/11: TGT GCA TGC TGG GAT A	T	CCCCG		A GGA TTT	1/19
			V_{T6}/12: TGT GCA TGC TGG GA		CCG		AT AGC TCA GGT TTT	1/19
			V_{T6}/13: TGT GCA TGC TGG GAT A		A		AT AGC TCA GGT TTT	1/19
			V_{T6}/14: TGT GCA TGC TGG GAT A		GGGGA	AT	AT AGC TCA GGT TTT	1/19
			V_{T6}/15: TGT GCA TGC TGG GAT A		AG		GC TCA GGT TTT	1/19
			V_{T6}/16: TGT GCA TGC TGG GAT		TGG		GC TCA GGC TTT	1/19
			V_{T6}/17: TGT GCA TGC TGG G		TC		AT AGC TCA GGT TTT	1/19
			V_{T6}/18: TGT GCA TGC TGG GA		GGGCCGGGG		AGC TCA GGT TTT	1/19
			V_{T6}/19: TGT GCA TGC		GGGG		AGC TCA GGT TTT	1/19
			V_{T6}/20: TGT GCA TGC TGG G		C		AT AGC TCA GGT TTT	1/19
			V_{T6}/21: TGT GCA		A	AT	AT AGC TCA GGT TTT	1/19
			V_{T6}/22: TGT GCA TGC TGG		AATC		T AGC TCA GGT TTT	1/19
			V_{T6}/23: TGT GCA TGC TGG GA		A	AT	AT AGC TCA GGT TTT	1/19
			V_{T6}/24: TGT GCA TGC TGG G		CCGGGGGG		GC TCA GGT TTT	1/19
			V_{T6}/25: TGT GCA TGC TGG GAT		TC		AGC TCA GGT TTT	1/19
	HYBRIDOMAS							
	IN FRAME	*	V_{T6}/1: TGT GCA TGC TGG GA				T AGC TCA GGT TTT	3/3

Figure 2. Nucleotide Sequences of V_{T6}-J_{T1} Junctions from Fetal, Newborn, and Adult Thymocytes

Fetal thymocytes were from day 17.5. Ten DNA clones all from a newborn thymocyte population were derived from cDNA, while the rest of DNA clones from thymocyte populations were from genomic DNA. All DNA clones from hybridomas were from cDNA. The hybridomas were obtained from newborn thymocytes (A) or from the thymocytes of 5-week-old BALB/c mice (B). Asterisks indicate the sequences present in the r-IEL clones (Itohara et al., submitted). See the legend to Figure 1 for further detail.

tional sequences of the fetal and newborn V_6J_1 γ genes and their cDNA were also highly homogeneous: 23 out of 25 in-frame junctions produced the canonical V_6J_1 sequence expressed on r-IEL. This same sequence is also expressed on 15 out of 16 V_6J_1 γ–expressing T hybridomas derived from fetal thymocytes. The data strongly suggest that the fetal V_6J_1 γ–expressing thymocytes are precursors for r-IEL.

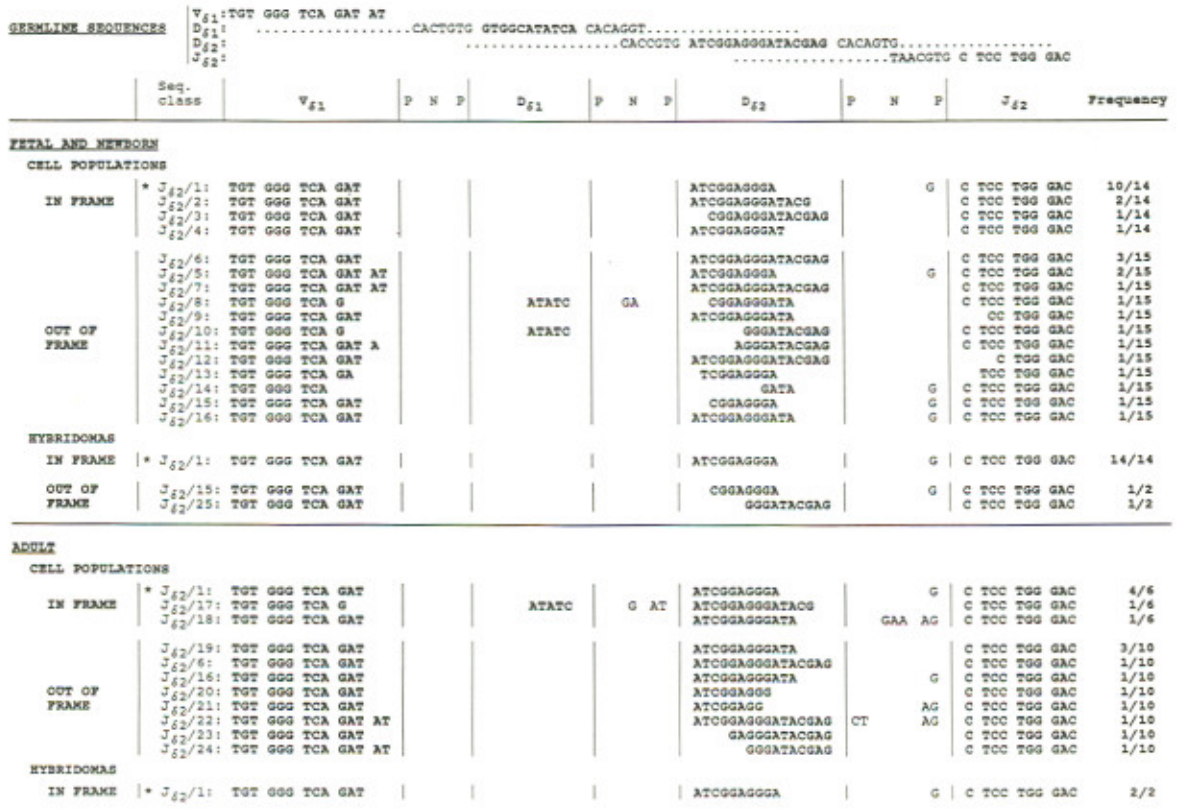
The out-of-frame V_6J_1 junctions are also rather homogeneous: 12 out of 14 are canonical out-of-frame sequences with minimal (1 bp at the V or J end) exonucleolytic nibbling and no N nucleotide insertion.

The Major Fetal δ Genes, $V_1D_2J_2C$, are Highly Homogeneous and Encode Both s-IEL and r-IEL δ Chains

Because the V_1DJ_2C δ chain is the major δ gene expressed on the surface of fetal thymocytes (Ito et al., 1989), we investigated the junctional sequences of this δ gene by the PCR-DNA sequencing method (Figure 3a). Both for the genomic DNA clones isolated from thymocytes and for the cDNA clones derived from T hybridomas, the pattern of sequence variation of the V_1DJ_2 junctions was highly restricted (Figure 3a), a finding in stark contrast to the highly diverse V_5DJ_1 δ, V_6DJ_1 δ, and V_7DJ_1 δ junctions contained in the δ genes expressed on the adult thymocytes or i-IEL (Figure 3b) (Takagaki et al., 1989a, 1989b; Korman et al., 1988). Again, in-frame junctions are more homogeneous than out-of-frame junctions. Unlike the δ genes of adult thymocytes or IEL, most of which use both D_1 δ and D_2 δ gene segments (Elliott et al., 1988; Takagaki et al., 1989a, 1989b), only the D_2 δ segment is used by most fetal thymocytes. The canonical in-frame $V_1D_2J_2$ δ junctional sequence observed in fetal thymocytes (Figure 3a) is identical to the sequence of the δ gene expressed on s-IEL (Asarnow et al., 1988) and r-IEL (Itohara et al., submitted), as expected if intraepithelial γδ T cells are derived from fetal or newborn thymocytes.

Developmental Time-Dependent Alterations in the Diversity of the V-J γ and V-D-J δ Junctions

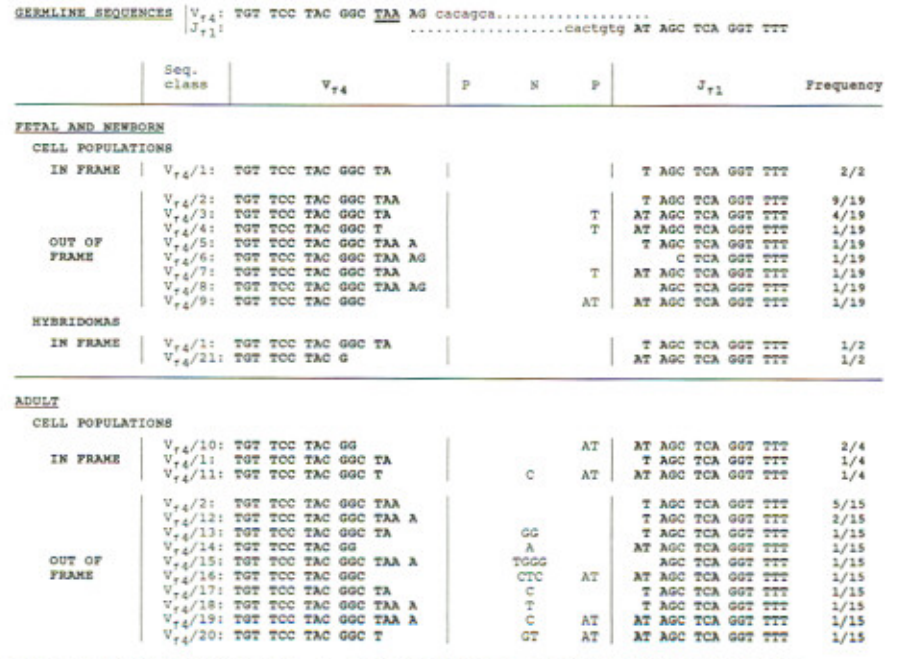
We asked whether the striking difference in diversity of the fetal vs. adult junctional sequences is due to the difference in the developmental state of the thymus per se and is independent of the type of γ and δ genes expressed. For this purpose we amplified the junctional sequences of the rare V_5J_1 γ and V_6J_1 γ genes present in adult thymocytes and those of the equally rare V_4J_1 γ genes present in fetal thymocytes and determined their nucleotide sequences. For comparison we also amplified the junctional sequences of the V_4J_1 γ genes present in adult thymocytes. As shown in Figures 1, 2, and 4, both the exonucleolytic nibbling and



	Seq. class	$V_{\delta1}$	P	N	P	$D_{\delta1}$	P	N	P	$D_{\delta2}$	P	N	P	$J_{\delta2}$	Frequency
GERMLINE SEQUENCES		$V_{\delta1}$: TGT GGG TCA GAT AT													
		$D_{\delta1}$: .....CACTGTG GTGGCATATCA CACAGGT.....													
		$D_{\delta2}$: .....CACCGTG ATCGGAGGGATACGAG CACAGTG.....													
		$J_{\delta2}$: .....TAACGTG C TCC TGG GAC													
a) **FETAL AND NEWBORN**															
CELL POPULATIONS															
IN FRAME	* $J_{\delta2}$/1:	TGT GGG TCA GAT								ATCGGAGGGA			G	C TCC TGG GAC	10/14
	$J_{\delta2}$/2:	TGT GGG TCA GAT								ATCGGAGGGATACG				C TCC TGG GAC	2/14
	$J_{\delta2}$/3:	TGT GGG TCA GAT								CGGAGGGATACGAG				C TCC TGG GAC	1/14
	$J_{\delta2}$/4:	TGT GGG TCA GAT								ATCGGAGGGAT				C TCC TGG GAC	1/14
OUT OF FRAME	$J_{\delta2}$/6:	TGT GGG TCA GAT								ATCGGAGGGATACGAG				C TCC TGG GAC	3/15
	$J_{\delta2}$/5:	TGT GGG TCA GAT AT								ATCGGAGGGA			G	C TCC TGG GAC	2/15
	$J_{\delta2}$/7:	TGT GGG TCA GAT AT								ATCGGAGGGATACGAG				C TCC TGG GAC	1/15
	$J_{\delta2}$/8:	TGT GGG TCA G				ATATC		GA		CGGAGGGATA				C TCC TGG GAC	1/15
	$J_{\delta2}$/9:	TGT GGG TCA GAT								ATCGGAGGGATA				CC TGG GAC	1/15
	$J_{\delta2}$/10:	TGT GGG TCA G				ATATC				GGGATACGAG				C TCC TGG GAC	1/15
	$J_{\delta2}$/11:	TGT GGG TCA GAT A								AGGGATACGAG				C TCC TGG GAC	1/15
	$J_{\delta2}$/12:	TGT GGG TCA GAT								ATCGGAGGGATACGAG				C TGG GAC	1/15
	$J_{\delta2}$/13:	TGT GGG TCA GA								TCGGAGGGA				TCC TGG GAC	1/15
	$J_{\delta2}$/14:	TGT GGG TCA								GATA			G	C TCC TGG GAC	1/15
	$J_{\delta2}$/15:	TGT GGG TCA GAT								CGGAGGGA			G	C TCC TGG GAC	1/15
	$J_{\delta2}$/16:	TGT GGG TCA GAT								ATCGGAGGGATA			G	C TCC TGG GAC	1/15
HYBRIDOMAS															
IN FRAME	* $J_{\delta2}$/1:	TGT GGG TCA GAT								ATCGGAGGGA			G	C TCC TGG GAC	14/14
OUT OF FRAME	$J_{\delta2}$/15:	TGT GGG TCA GAT								CGGAGGGA			G	C TCC TGG GAC	1/2
	$J_{\delta2}$/25:	TGT GGG TCA GAT								GGGATACGAG				C TCC TGG GAC	1/2
b) **ADULT**															
CELL POPULATIONS															
IN FRAME	* $J_{\delta2}$/1:	TGT GGG TCA GAT								ATCGGAGGGA			G	C TCC TGG GAC	4/6
	$J_{\delta2}$/17:	TGT GGG TCA G				ATATC		G AT		ATCGGAGGGATACG				C TCC TGG GAC	1/6
	$J_{\delta2}$/18:	TGT GGG TCA GAT								ATCGGAGGGATA		GAA	AG	C TCC TGG GAC	1/6
OUT OF FRAME	$J_{\delta2}$/19:	TGT GGG TCA GAT								ATCGGAGGGATA				C TCC TGG GAC	3/10
	$J_{\delta2}$/6:	TGT GGG TCA GAT								ATCGGAGGGATACGAG				C TCC TGG GAC	1/10
	$J_{\delta2}$/16:	TGT GGG TCA GAT								ATCGGAGGGATA			G	C TCC TGG GAC	1/10
	$J_{\delta2}$/20:	TGT GGG TCA GAT								ATCGGAGGG				C TCC TGG GAC	1/10
	$J_{\delta2}$/21:	TGT GGG TCA GAT								ATCGGAGG			AG	C TCC TGG GAC	1/10
	$J_{\delta2}$/22:	TGT GGG TCA GAT AT								ATCGGAGGGATACGAG	CT		AG	C TCC TGG GAC	1/10
	$J_{\delta2}$/23:	TGT GGG TCA GAT								GAGGGATACGAG				C TCC TGG GAC	1/10
	$J_{\delta2}$/24:	TGT GGG TCA GAT AT								GGGATACGAG				C TCC TGG GAC	1/10
HYBRIDOMAS															
IN FRAME	* $J_{\delta2}$/1:	TGT GGG TCA GAT								ATCGGAGGGA			G	C TCC TGG GAC	2/2

Figure 3. Nucleotide Sequences of $V_{\delta1}$-D-$J_{\delta2}$ Junctions from Fetal, Newborn, and Adult Thymocytes

Fetal thymocytes were from day 17.5. For the thymocyte populations all clones were from genomic DNA, and for the hybridomas all clones were from cDNA. The hybridomas were prepared from thymocytes of BALB/c embryos of day 15.5 or 17.5 of gestation or of newborn mice (A), or from thymocytes of 5-week-old BALB/c mice (B). Asterisks indicate the sequences present in s-IEL (Asarnow et al., 1988) and r-IEL (Itohara et al., submitted) clones. See the legend to Figure 1 for further detail.



	Seq. class	$V_{\gamma4}$	P	N	P	$J_{\gamma1}$	Frequency
GERMLINE SEQUENCES		$V_{\gamma4}$: TGT TCC TAC GGC TAA AG cacagca.....					
		$J_{\gamma1}$: .....cactgtg AT AGC TCA GGT TTT					
a) **FETAL AND NEWBORN**							
CELL POPULATIONS							
IN FRAME	$V_{\gamma4}$/1:	TGT TCC TAC GGC TA				T AGC TCA GGT TTT	2/2
OUT OF FRAME	$V_{\gamma4}$/2:	TGT TCC TAC GGC TAA				T AGC TCA GGT TTT	9/19
	$V_{\gamma4}$/3:	TGT TCC TAC GGC TA			T	AT AGC TCA GGT TTT	4/19
	$V_{\gamma4}$/4:	TGT TCC TAC GGC T			T	AT AGC TCA GGT TTT	1/19
	$V_{\gamma4}$/5:	TGT TCC TAC GGC TAA A				T AGC TCA GGT TTT	1/19
	$V_{\gamma4}$/6:	TGT TCC TAC GGC TAA AG				C TCA GGT TTT	1/19
	$V_{\gamma4}$/7:	TGT TCC TAC GGC TAA			T	AT AGC TCA GGT TTT	1/19
	$V_{\gamma4}$/8:	TGT TCC TAC GGC TAA AG				AGC TCA GGT TTT	1/19
	$V_{\gamma4}$/9:	TGT TCC TAC GGC			AT	AT AGC TCA GGT TTT	1/19
HYBRIDOMAS							
IN FRAME	$V_{\gamma4}$/1:	TGT TCC TAC GGC TA				T AGC TCA GGT TTT	1/2
	$V_{\gamma4}$/21:	TGT TCC TAC G				AT AGC TCA GGT TTT	1/2
b) **ADULT**							
CELL POPULATIONS							
IN FRAME	$V_{\gamma4}$/10:	TGT TCC TAC GG			AT	AT AGC TCA GGT TTT	2/4
	$V_{\gamma4}$/1:	TGT TCC TAC GGC TA				T AGC TCA GGT TTT	1/4
	$V_{\gamma4}$/11:	TGT TCC TAC GGC T		C	AT	AT AGC TCA GGT TTT	1/4
OUT OF FRAME	$V_{\gamma4}$/2:	TGT TCC TAC GGC TAA				T AGC TCA GGT TTT	5/15
	$V_{\gamma4}$/12:	TGT TCC TAC GGC TAA A				T AGC TCA GGT TTT	2/15
	$V_{\gamma4}$/13:	TGT TCC TAC GGC TA		GG		T AGC TCA GGT TTT	1/15
	$V_{\gamma4}$/14:	TGT TCC TAC GG		A		AT AGC TCA GGT TTT	1/15
	$V_{\gamma4}$/15:	TGT TCC TAC GGC TAA A		TGGG		AGC TCA GGT TTT	1/15
	$V_{\gamma4}$/16:	TGT TCC TAC GGC		CTC	AT	AT AGC TCA GGT TTT	1/15
	$V_{\gamma4}$/17:	TGT TCC TAC GGC TA		C		T AGC TCA GGT TTT	1/15
	$V_{\gamma4}$/18:	TGT TCC TAC GGC TAA A		T		T AGC TCA GGT TTT	1/15
	$V_{\gamma4}$/19:	TGT TCC TAC GGC TAA A		C	AT	AT AGC TCA GGT TTT	1/15
	$V_{\gamma4}$/20:	TGT TCC TAC GGC T		GT	AT	AT AGC TCA GGT TTT	1/15

Figure 4. Nucleotide Sequences of $V_{\gamma4}$-$J_{\gamma1}$ Junctions from Fetal, Newborn, and Adult Thymocytes

Fetal thymocytes were from days 14.5 and 17.5. For the thymocyte populations all clones were from genomic DNA, and for the hybridomas all clones were from cDNA. The hybridomas were prepared from thymocytes of C57BL/6 embryos of day 17.5 of gestation. Underlined in the germline $V_{\gamma4}$ sequence is the in-frame stop codon. See the legend to Figure 1 for further detail.

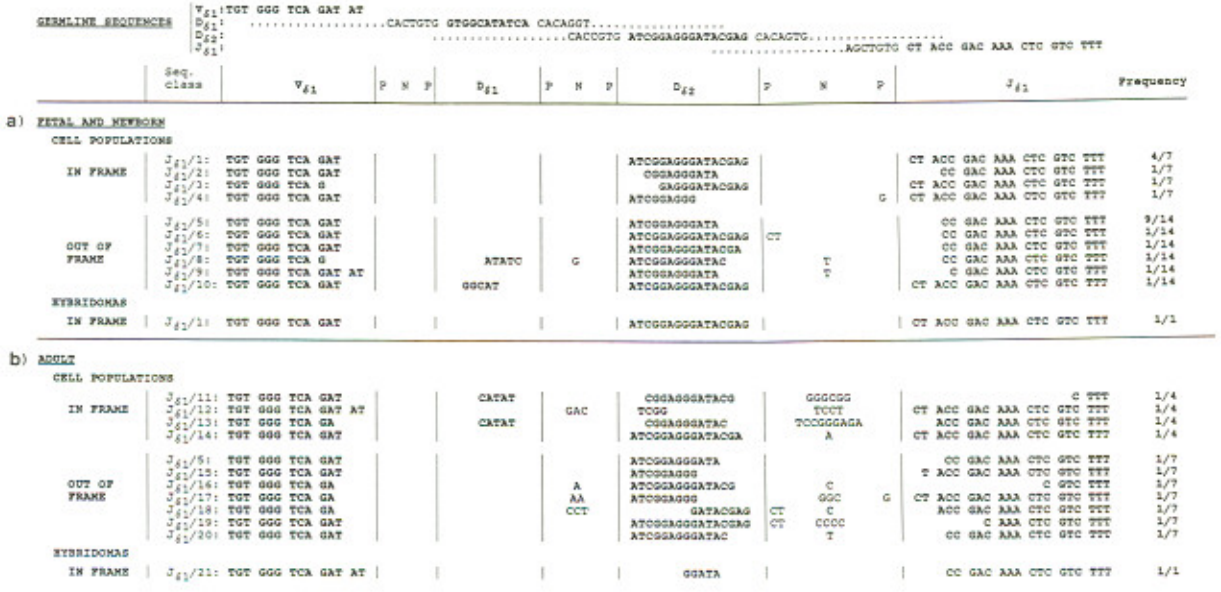


Figure 5. Nucleotide Sequences of $V_{\delta1}$-D-$J_{\delta1}$ Junctions from Fetal, Newborn, and Adult Thymocytes

Fetal thymocytes were from day 17.5. For the thymocyte populations all clones were from genomic DNA, and for the hybridomas all clones were from cDNA. The hybridomas were obtained from newborn C57BL/6 thymocytes (A) or from the thymocytes of 5-week-old BALB/c mice (B). See the legend to Figure 1 for further detail.

N nucleotide insertions are more extensive in the adult V_5J_1, V_6J_1 and V_4J_1 junctions than in their fetal counterparts.

Contrary to this high degree of diversity, the rare fetal V_4J_1 γ junctions exhibit very limited sequence variability (less exonucleolytic nibbling at the V terminals and no N insertions). Taken together, these results strongly suggest that the two major diversifiers of V-J junctional sequences, namely exonucleolytic nibbling and N nucleotide insertional activities, are low in fetal thymocytes and high in adult thymocytes regardless of the V_γ gene segments employed to construct the γ genes.

Rearrangements involving the J_1 δ gene segment are reported to be abundant in fetal thymocytes (Chien et al., 1987b), although analysis of γδ T cell hybridomas indicates that δ genes containing the J_1 δ gene segment are expressed on the surface of only a minor population of fetal thymocytes (Ito et al., 1989). In contrast, J_1 is the heavily preferred J gene segment in the δ genes expressed on adult thymocytes. To determine whether such developmental stage–dependent alteration in the extent of junctional diversity also occurs in δ genes, we compared the diversity of V_1DJ_1 δ junctions in fetal thymocytes with that in adults. As shown in Figure 5, both terminal diversity and N nucleotide insertion diversity are much more pronounced in adult junctions than in fetal junctions. This difference of terminal diversity is particularly evident in the 3′-terminal of the D_2 gene segment and in the 5′-terminal of the J_1 gene segment of out-of-frame D_1J_1 junctions.

A Fetal γδ Thymocyte Lineage Persists in Adult Thymus

While the developmental stage–dependent and V_γ gene segment–independent alteration of the VJ_1 γ junctional sequence diversity is evident, a close examination of the sequences of the adult V_5J_1 γ and V_6J_1 γ junctions reveals an interesting additional feature. Among the rare junctions with no N nucleotides, nearly all consist of the canonical sequence that dominates the corresponding fetal junctions (Figures 1b and 2b). Thus, five out of six N-less in-frame adult V_5J_1 junctions contain the fetal canonical sequences, and all of the N-less V_6J_1 junctions (four out-of-frame, three in-frame) contain the corresponding fetal canonical sequence. The in-frame canonical sequences could be the consequence of cellular selection in the adult thymus, as they seem to be in the fetal thymus (see above). However, the occurrence of the out-of-frame V_6J_1 canonical sequence in multiple DNA clones suggests the persistence of the fetal γδ T cell lineage with low exonucleolytic nibbling and N nucleotide insertion in the adult thymocyte population.

Strong support for this hypothesis comes from the sequences of the rare V_1DJ_2 δ junctions present in the adult thymus. As shown in Figure 3b, the sequences of these adult junctions are as homogeneous as those of the corresponding fetal junctions. Moreover, the occurrence of the $D_{\delta1}$ segment is very rare. This last feature is unique among fetal thymocytes, for it was previously shown that most of the δ genes expressed on adult thymocytes use both $D_{\delta1}$ and $D_{\delta2}$ gene segments (Elliott et al., 1988; Takagaki et al., 1989a).

The Reciprocal Recombination Products of Fetal V-J γ and D-J δ Are Head-to-Head, Flush-Joined Signal Heptamer Sequences

As seen above, coding by a given gene segment (i.e., V, D, or J) ends or starts at one of only a few alternative

Figure 6. Signal Joints of Days 14.5 and 17.5 Fetal Thymocytes

The same DNA preparation used for the study of the coding joints was utilized for the PCR-DNA sequencing of the signal joints generated by $V_{\gamma5}$-$J_{\gamma1}$, $V_{\gamma6}$-$J_{\gamma1}$, and $D_{\delta2}$-$J_{\delta2}$ joinings. The signal heptamers and nonamers are boxed. The frequency with which a given sequence was observed among the DNA clones is indicated in the last column. For comparison, the relevant germline sequences are shown above the sequences of signal joints.

nucleotide positions in the V-J or V-D-J junctions of fetal thymocytes. The "terminal homogeneity" in the in-frame junctions could be attributed to a strong selective pressure operating at the cell level, but this cannot explain the homogeneity observed in out-of-frame junctions. The possibility that a particular out-of-frame junction is passively selected by being present in the same cell that contains the actively selected in-frame junction is ruled out, as the DNA clones analyzed are derived from several embryos. Thus, in addition to cellular selection, there must be a mechanistic cause for the observed terminal homogeneity.

Two mechanisms can be envisaged. The first possibility is that the initial double-stranded DNA cleavage always occurs at the exact border of the signal heptamer, and that the coding but not the heptamer terminal generated is nibbled by the putative exonuclease to a very limited extent. The second possibility is that in fetal thymocytes the cleavage of the phosphodiester backbones occurs at one of a few discrete alternative sites concentrated around the heptamer border, and that both the coding and the heptamer terminals are rejoined with the respective partner without any modification. The two possible mechanisms may be distinguished by determining the sequence of the reciprocal product of the recombination composed of two signal sequences joined head to head. We call the junction contained in the reciprocal recombinant the "signal joint" and the junction composed of two coding gene segments the "coding joint" (Lieber et al., 1988a). For this purpose we cloned, from embryos on days 14.5 and 17.5 of gestation, DNA containing the signal joints of V_5J_1 γ, V_6J_1 γ, and D_2J_2 δ joinings and determined their nucleotide sequences. If the terminal homogeneity is brought about by the single site cleavage combined with limited nibbling, all signal joints should be composed of two intact heptamers joined in flush. On the other hand, if the multiple (preferred) cleavage site model is correct, the signal joints should be heterogeneous and contain insertions or deletions whose sequences are precisely predicted from the sequence of the reciprocal coding joint. Indeed, there are a few precedents in the TCR β and immunoglobulin κ light chain gene rearrangements in which the signal joints contained short insertions apparently composed of the germline nucleotides derived from the recombining gene segments (Malissen et al., 1986; Okazaki et al., 1987; Deev et al., 1987). However, as shown in Figure 6, all of the signal joints studied are composed of two perfect heptamers flush-joined, indicating that in most if not all cases, the processes of V_5J_1 γ, V_6J_1 γ, and D_2J_2 δ joinings in fetal thymocytes occur with a cleavage at the exact boundary between the heptamer and the gene segment. The limited variability in the V-, J-, or D-derived terminal nucleotide in the coding joints must then be due to a limited action of the exonuclease.

Recurrence of a Specific Mono- and Dinucleotide in the Coding Joints

The N nucleotides are supposed to be generated by template-independent additions of nucleotides at the terminal of combining gene segments by the action of the terminal transferase or a terminal transferase–like enzyme (Alt and Baltimore, 1982; Desiderio et al., 1984). Reflecting the substrate preference of this enzyme, N nucleotides tend to contain GC or CG pairs more often than expected from the random frequency, but otherwise they are highly variable. However, when we surveyed the nucleotide sequences of a large number of γ and δ coding joints, we noticed that specific mononucleotides and dinucleotides recur at an unusually high frequency in the region where N nucleotides normally appear. These nucleotides are designated as P in Figures 1–5 because, as we shall see below, they have entirely different origins from N nucleotides.

Discussion

Homogeneous γ and δ Coding Joints in Fetal Thymocytes

In this study we determined the nucleotide sequences of a large number of DNA clones containing the V-(D)-J junctions of TCR γ and δ genes present in the fetal and adult mouse thymocytes. In agreement with earlier analyses of γ and δ genes expressed on the surface of T hybridomas derived from adult thymocytes (Takagaki et al., 1989a; Korman et al., 1988), coding joints present in adult thymocytes are highly variable; most reveal extensive exonucleolytic nibbling and N nucleotide insertion. In contrast, both types of diversity are highly limited in the coding joints of γ and δ genes present in fetal or newborn thymocyte populations. The drastic shift in the extent of junc-

tional diversity in fetal vs. adult thymocytes is not simply due to the difference in cellular selection, because the shift is also evident among the out-of-frame coding joints. Thus, these results strongly suggest that the activity of both terminal transferase–like enzyme (Rothenberg and Triglia, 1983; Elliott et al., 1988) and exonuclease is lower in fetal γδ thymocytes than in adult γδ thymocytes.

The effect of the highly preferential usage of a limited set of gene segments (Havran and Allison, 1988; Ito et al., 1989; Itohara et al., 1989) and the restricted junctional diversity is that just two primary structures form the majority of γδ TCR expressed on fetal and newborn thymocytes.

Fetal γδ Thymocytes Are Probably Precursors of γδ s-IEL and γδ r-IEL

Since the overwhelming majority of early fetal γδ thymocytes and γδ s-IEL share a unique TCR, it is very likely that the former is the precursor of the latter. By the analogous reason, γδ r-IEL seems to be derived from late fetal γδ thymocytes. While the thymocytes containing these γ and δ genes continue to be present in the adult thymus, they form a very minor subpopulation. Thus, whatever the reason the γδ TCRs of s-IEL and r-IEL need to be homogeneous, the fetal thymus seems to be designed to be the major supplier of these homogeneous γδ TCRs.

Although the γ and δ coding joints present in fetal thymocytes are remarkably homogeneous, the in-frame coding joints are distinctly more homogeneous than the out-of-frame coding joints. This strongly suggests that the fetal thymus has a mechanism for selecting cells expressing the canonical γ and δ genes on their surface. Thus, the fetal thymus seems to be designed at both the genetic and cellular levels to produce effectively the γδ thymocytes carrying the s-IEL and r-IEL TCRs.

It is of interest to note that the two major kinds of γδ TCRs expressed on fetal thymocytes, namely the s-IEL and r-IEL γδ TCRs share identical amino acid sequences in both the V-J γ and V-D-J δ junctional regions. These subregions of TCR subunits, thought to be equivalent to the CDR3 (complementarity determining region, number 3) of immunoglobulin light and heavy chains (Novotny et al., 1986), have been suggested to be primarily responsible for the determination of the (antigen-derived) peptide binding specificity of the TCR (Davis and Bjorkman, 1988). If this is indeed the case, sharing of the identical γ and δ CDR3 regions suggests that the ligand for s-IEL and r-IEL γδ TCRs may include a common peptide.

In addition to the aforementioned relationships between fetal thymocytes and s-IEL or r-IEL, both the pattern of the use of V, D, and J gene segments and the pattern of junctional diversity suggest that γδ T cells in peripheral lymphoid organs (spleen, lymph node, etc.) and i-IEL are derived from two major subpopulations of adult thymocytes. In Figure 7 we summarize the relationships of various $C_{\gamma 1}$-expressing γδ T cell subpopulations.

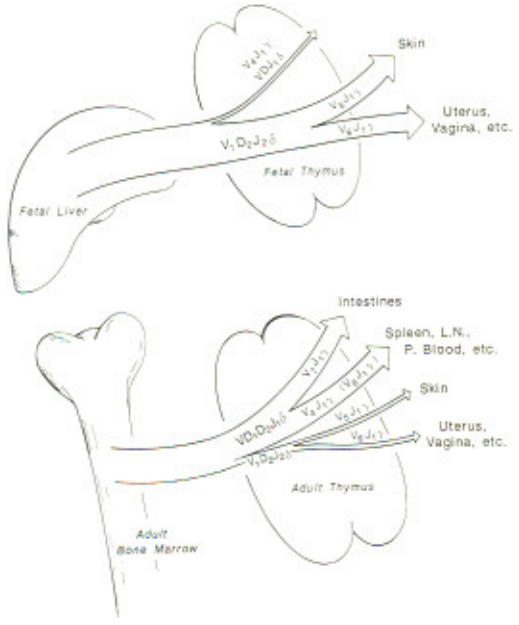


Figure 7. Development of Various γδ T Cell Subsets

The arrows indicate the proposed pathways followed by γδ T lymphocytes expressing the $C_{\gamma 1}$ region in fetal and adult mice.

P Nucleotides—A New Type of Insert in γ and δ Coding Joints

The present study revealed that some γ and δ coding joints contain recurrent mono- or dinucleotides (called P nucleotides for palindrome, see below) as an insert: the dinucleotide AT or mononucleotide T, and the dinucleotide AG and mononucleotide G appear at a remarkably high frequency as the sole insert or at the 3′ end of the insert in the VJ_1 γ coding joints and D_2J_2 δ coding joints, respectively. Furthermore, this study indicated that all coding joints with a P nucleotide(s) utilize the full coding capacity of the J gene segment involved. In fact, a few previous reports, including one from our laboratory, noted this unusual property of inserts in V_4J_1 γ and V_2J_2 γ coding joints (Hayday et al., 1985; Traunecker et al., 1986).

Anticipating that this characteristic of inserts may actually be a general rule, we examined the nucleotide sequences of the other coding joints determined in this study as well as the published sequences of VJ γ, VD δ, and DJ δ coding joints. This search revealed several additional di- and mononucleotides that are recurrent in the 5′ or 3′ end of the inserts. These are AT or T at the 3′ end of the inserts in V_7J_1 coding joints, AG or G at the 3′ end of the inserts in DJ_1 δ coding joints, and CT or C at the 5′ end of the inserts in D_2J_1 δ and D_2J_2 δ coding joints (see legend to Table 1 for references).

Combining the data on recurrent nucleotides obtained in this study with those in the literature, we could extract the following as the general rules. First, the nucleotides that recur are specific to the terminal of the neighboring gene segment: AT or T for the 5′-terminals of J_1 γ and J_2 γ, AG or G for the 5′-terminals of J_1 δ and J_2 δ, and CT or C for the 3′-terminal of D_2 δ. Second, the recurring mononucleotide assigned to a given terminal of a gene segment is always the same as the terminal-proximal mononucleotide of the recurring dinucleotide assigned to the same terminal. Third, as already noted above, whenever

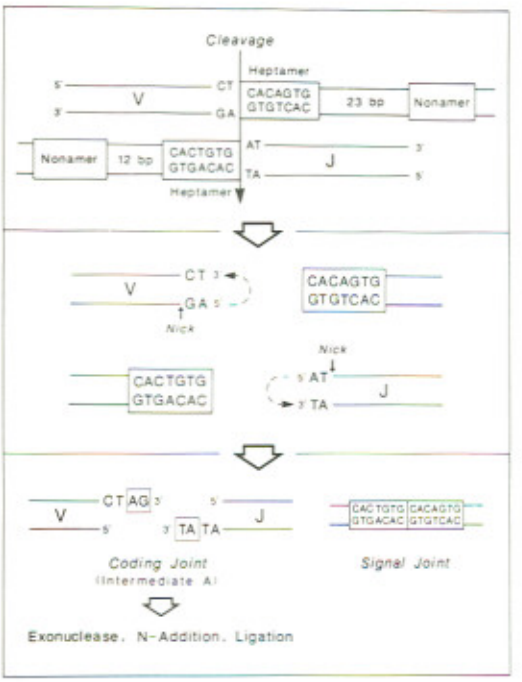


Figure 8. P Nucleotide Model of V-J (V-D-J) Joining
See text for explanations.

the nucleotide(s) occurs in the inserts, the corresponding neighboring gene segment appears in its full sequence in the coding joint. Finally, the recurrent dinucleotide and the immediately adjacent dinucleotide, which belong to the corresponding gene segment, form a tetranucleotide palindrome. For instance, both $J_1\gamma$ and $J_2\gamma$ gene segments start with an AT and the tetranucleotide ATAT formed by this AT, and the recurrent AT present just upstream is a palindrome. Similarly, the $D_2\delta$ gene segment ends with an AG and the tetranucleotide AGCT formed by this AG, and the recurrent CT present just downstream is a palindrome.

A New Model of V-(D)-J Joining

We wish to propose a new model of V-(D)-J joining that would explain the origin of the recurrent P nucleotides (Figure 8). According to this model, a recombining gene segment is first cleaved precisely at the gene segment–proximal border of the signal heptamer (Alt and Baltimore, 1982). While the resulting heptamer terminal is rejoined to its counterpart to form a signal joint, the coding terminal is obligatorily modified before being ligated to its counterpart. The first modification is cleavage of the terminal dinucleotide (P dinucleotide) from the 5′ end of one strand, followed by its flipping and joining to the 3′ end of the other. This modification will generate a single-stranded tail composed of a tetranucleotide palindrome (intermediate A).

The single-stranded tail may then be converted to a duplex by the action of a DNA polymerase. If a pair of terminals thus generated is ligated, the resulting coding joint will have a tetranucleotide insert composed of the P dinucleotides of the two joining gene segments. This type of coding joint, however, is rare—only 1 (clone $J_{\delta 2}$/22 in Figure 3b) out of 396 coding joint sequenced in this study was of this type. In an overwhelming majority of cases at least one of the two coding terminals is further modified by exonucleolytic nibbling. If the nibbling is restricted to the very terminal mononucleotide, then the P nucleotide will be retained in the coding joint, but if the nibbling proceeds beyond the terminal dinucleotide on the intermediate A, then no P nucleotide will make its way through to the coding joint. Another modification that frequently occurs before ligation is an addition of N nucleotides by terminal transferase or a similar enzyme (Alt and Baltimore, 1982; Desiderio et al., 1984; Landau et al., 1987).

Thus, according to the model, the origin of the recurrent nucleotides (P nucleotides) observed in coding joints is the terminal dinucleotide from the 5′ end of one strand of the recombining gene segment, and P nucleotides are not the substrate nucleotides of a terminal transferase. Because of this origin, P nucleotides are in fact germline-derived and specific to the joining end of the recombining gene segment. They are not necessarily G-rich, as are N nucleotides. Furthermore, the model explains why P nucleotides are always preceeded or followed in the coding joints by the corresponding gene segment in its full coding capacity. When and only when no exonucleolytic nibbling occurs on the intermediate A will the P dinucleotide remain in the final recombination product; when and only when the exonuclease nibbles away only 1 nucleotide from the intermediate A will the P mononucleotide appear in the recombinant. In either case the recombining gene segment itself cannot be nibbled by the exonuclease.

The Model Is Generally Applicable to TCR and Immunoglobulin Genes

In addition to the cases already described above, there are numerous cases of P nucleotide insertion in the published sequences of γ and δ genes (Table 1). Indeed, among the terminals of all known mouse γ and δ gene segments only two, namely the 3′ ends of $V_{\delta 6}$ and $V_{\delta 7}$, did not provide any evidence that their P nucleotides were actually present in the junctions. However, in these cases sequenced junctions are rare, and sequenced junctions utilizing the full coding capacity of the gene segments are even rarer. Furthermore, the data in the literature show that the P nucleotide insertion model applies to human γ and δ genes (Table 1).

As summarized in Table 1, there is clearly evidence of P nucleotides in the junctions of TCR, both α and β, and immunoglobulin genes. For instance, Roth et al. (1988) determined the sequences of four $V_{\alpha 58}J_{\alpha 1/2}$ coding joints in which the $V_{\alpha 58}$ gene segment is fully retained. Among these four coding joints, two have the P dinucleotide TG, and the third has the P mononucleotide T following CA, the terminal dinucleotide of $V_{\alpha 58}$. Similarly, in one of the three D-$J_{2.5}$ TCR β coding joints in which the full $J_{\beta 2.5}$ gene segment is utilized, the P dinucleotide TT precedes the starting J dinucleotide AA, and in the other two cases the P mononucleotide T is found in the same position (Barth et al., 1985; Behlke et al., 1985).

Table 1. Summary of P Nucleotides in the Coding Joints of TCR α, β, γ, δ, and Immunoglobulin Genes

Gene Segment	P Dinucleotide	Palindromic Outcome	P in Joints	Notes	References
Mouse TCR γ/δ Genes					
$J_{\gamma1}$	**AT**	**AT/AT** AGCT . .	+ +		1, 2, 3
$J_{\gamma2}$	**AT**	**AT/AT** AGCT . .	+ +		4, 5, 6
$J_{\delta1}$	**AG**	**AG/CT** ACCG . .	+ +	Low proportion	7
$J_{\delta2}$	**AG**	**AG/CT** CCTG . .	+ +		1, 8
5′ of $D_{\delta1}$	**AG**	**AG/GT** GGCA . .	+ +		7, 8
5′ of $D_{\delta2}$	**AT**	**AT/AT** CGGA . .	+ +		1, 2, 3, 7, 8
3′ of $D_{\delta2}$	**CT**	. . TACG **AG/CT**	+ +		1, 2, 3, 7, 8
$V_{\gamma4}$	**CT**	. . CTAA **AG/CT**	**C**	Only 2 cases analyzed	1
$V_{\gamma5}$	**AG**	. . GGAT **CT/AG**	c	Only 6 cases analyzed	1
$V_{\gamma6}$	**TA**	. . GGGA **TA/TA**	+	Canonical out-of-frame is **C**	1
$V_{\gamma7}$	**CC**	. . GGCT **GG/CC**	+ +		2, 3
$V_{\delta1}$	**AT**	. . AGAT **AT/AT**	**C**		1
$V_{\delta4}$	**CG**	. . GGAG **CG/CG**	+ +		3, 8
$V_{\delta5}$	**AT**	. . GGGT **AT/AT**	**C**	4 cases analyzed	7
$V_{\delta6}$	**AG**	. . GGAA **CT/AG**	–	5 cases analyzed	3, 7
$V_{\delta7}$	**CC**	. . CTAT **GG/CC**	–	2 cases analyzed	3
Human TCR γ/δ Genes					
$J_{\gamma}1$ or 2	**TC**	**TC/GA** ATTA . .	+	1 case analyzed	9
$J_{\gamma}P2$	**AT**	**AT/AT** AGTA . .	+ +		9
$J_{\delta1}$	**GT**	**GT/AC** ACCG . .	+ +		10, 11
5′ of $D_{\delta1}$	**TC**	**TC/GA** AATA . .	+ +		11
5′ of $D_{\delta2}$	**GG**	**GG/CC** TTCC . .	+ +		10
5′ of $D_{\delta3}$	**GT**	**GT/AC** TGGG . .	+ +		11
3′ of $D_{\delta1}$	**AC**	. . AATA **GT/AC**	**C**	2 cases analyzed	10
3′ of $D_{\delta2}$	**GT**	. . TCCT **AC/GT**	+	3 cases analyzed	10
3′ of $D_{\delta3}$	**CG**	. . GATA **CG/CG**	+ +		10, 11
$V_{\gamma}2$	**CC**	. . GACG **GG/CC**	+ +		9
$V_{\gamma}3$	**CC**	. . GACA **GG/CC**	+	1 case analyzed	9
$V_{\gamma}4$	**CC**	. . GATG **GG/CC**	+	1 case analyzed	9
$V_{\gamma}9$	**CA**	. . GAGG **TG/CA**	+	3 cases analyzed	9
$V_{\delta1}$	**AG**	. . GGAA **CT/AG**	c	1 case analyzed	10
$V_{\delta3}$	**GG**	. . GACA **CC/GG**	+ +		11
Mouse TCR α/β Genes					
$J_{\beta1.3}$	**AA**	**AA/TT** CTGG . .	+	2 cases analyzed	12
$J_{\beta2.3}$	**CT**	**CT/AG** TGCA . .	+ +		13
$J_{\beta2.5}$	**TT**	**TT/AA** CCAA . .	+ +		14
5′ of $D_{\beta1.1}$	**CC**	**CC/GG** GACA . .	+ +		15, 16
5′ of $D_{\beta2.1}$	**CC**	**CC/GG** GACT . .	+	2 cases analyzed	16, 17
3′ of $D_{\beta1.1}$	**GC**	. . GGGG **GC/GC**	+ +		18
3′ of $D_{\beta2.1}$	**GC**	. . GGGG **GC/GC**	+ +		19
$V_{\alpha58}$	**TG**	. . TGAG **CA/TG**	+ +		20
$V_{\beta8.1}$	**CA**	. . GTGA **TG/CA**	–	1 case analyzed	14
$V_{\beta8.2}$	**CA**	. . GTGA **TG/CA**	+ +		21
Immunoglobulin Genes					
$J_{\kappa4}$	**AT**	**AT/AT** TCAC . .	+ +		22
5′ $D_{FL16.1}$	**AA**	**AA/TT** TATT . .	+ +		23
5′ $D_{FL16.2}$	**AA**	**AA/TT** TATT . .	+ +		24
3′ of D_{Q52}	**GT**	. . TGGG **AC/GT**	+ +		25
3′ of D_{FL16}	**GT**	. . AGCT **AC/GT**	+	3 cases analyzed	23
3′ of D_{SP2}	**GT**	. . AACT **AC/GT**	+ +		26
V_HID_{CR}	**TC**	. . GCAA **GA/TC**	+ +		27
$V_{H186.2}$	**TC**	. . GCAA **GA/TC**	+	Anti-NP hybrid use Tyr position 99	28

Only those gene segment terminals that appear in their full coding capacity are included in this analysis. In cases where the P dinucleotide was not found in the inserts of coding joints, the number of gene segment terminals that appear in their full coding capacity is indicated in the Notes column.
+ +, P dinucleotide observed as insert.
+, Only P mononucleotide observed as insert.
C, Junctional sequences are compatible with the presence of P dinucleotide but could also be accounted for by the germline sequences.
c, Junctional sequences are compatible with the presence of P mononucleotide but could also be accounted for by the germline sequences.
–, No real or potential case of P nucleotides found as inserts.
References: (1) This work; (2) Takagaki et al., 1989a; (3) Takagaki et al., 1989b; (4) Kranz et al., 1985; (5) Hayday et al., 1985; (6) Traunecker et al., 1986; (7) Elliott et al., 1988; (8) Lacy et al., 1989; (9) Huck et al., 1988; (10) Loh et al., 1988; (11) Loh et al., 1989; (12) Hedrick et al., 1984; (13) Behlke et al., 1986; (14) Barth et al., 1985; (15) Behlke et al., 1985; (16) Blackman et al., 1986; (17) Goverman et al., 1985; (18) Saito et al., 1984; (19) Patten et al., 1984; (20) Roth et al., 1988; (21) Urban et al., 1988; (22) Lewis et al., 1985; (23) Kurosawa and Tonegawa, 1982; (24) Clarke et al., 1985; (25) Alt et al., 1982; (26) Rocca-Serra et al., 1983; (27) Wysocki et al., 1986; and (28) Boersch-Supan et al., 1985.

An interesting example of a P dinucleotide in immunoglobulin heavy chain genes is seen in the V-D junctions of several independently derived arsonate binding hybridomas (Wysocki et al., 1986; Milner et al., 1986). It was previously observed that these heavy chain genes utilize the same, complete V_H gene segment that ends with the dinucleotide GA. While the inserts of these heavy chain genes are diverse, they all start with an identical dinucleotide, TC, that immediately follows the V_H segment–derived GA. Because the TC in these gene comprises the first two nucleotides of a Ser codon, and because Ser codons are highly redundant, it is unlikely that the recurrence of the Ser codon containing the TC dinucleotide is attributed to cellular selection. This then seems to be a perfect example of a P dinucleotide.

No insertion has been reported in the coding joints of immunoglobulin light chain genes except for one case. Interestingly, the insertion of this exceptional joint, the dinucleotide AT, seems to be the P dinucleotide of the full $J_{\kappa 4}$ gene segment that immediately follows the insertion with an AT dinucleotide (Lewis et al., 1985).

P nucleotides are also evident in the recombinants produced by transfection of immature B cell lines with plasmids containing artificial recombination substrates in which no immunoglobulin or TCR gene sequence (called "coding flank") was linked to the heptamer-nonomer signal sequence (Lieber et al., 1988a).

Finally, a model that explains the recurrence of a nongermline coded, specific mononucleotide (C or A) in the junction of chicken immunoglobulin light chain genes was proposed (McCormack et al., 1989). According to this model the recombinase complex contains a C or A nucleotide and exchanges the C nucleotide for the signal sequence 3′ of the V_L segment and the A nucleotide for the signal sequence 5′ of the J_L segment. This model, however, does not specify the source of the recombinase-bound C and A nucleotides and is not generally applicable to other numerous recurrent nucleotides described in this paper. Since the chicken V_L gene segment ends with TG, and the J_L segment starts with TG, the expected P dinucleotides are both CA. Thus, the recurrent C, A, or CA can be explained within the framework of our model.

A Possible Function of the P Nucleotides

As shown in Figure 6 and indicated in the model (Figure 8), the signal joints are almost always composed of flush-joined signal heptamers, and no P nucleotides are observed. The latter seems to be the case even in those rare signal joints that do contain inserts (Lieber et al., 1988a). The presence of P nucleotides in the coding joints and its absence in the signal joints are most dramatically demonstrated in the "hybrid joints" produced by the combination of a coding flank with a signal sequence belonging to another coding flank. In the hybrid joints analyzed by Lewis et al. (1988), seven had intact coding flanks, and three of these seven joints had dinucleotides that can be interpreted as the P nucleotides derived from the respective coding flank. In contrast, although nearly all of the hybrid joints (19 cases) contained an intact heptamer, none had a dinucleotide attributed to its P nucleotides.

The complete lack of signal heptamer–derived P nucleotides in both the signal and hybrid joints strongly suggests that the P nucleotide addition step occurs only at coding terminals as indicated in Figure 8. While the function of this specific biochemical step is far from obvious, the role of the resulting terminal with a single-stranded tail may be to promote the initiation of the exonucleolytic attack, which in turn contributes to the diversification of coding joints. For instance, the exonuclease involved, although capable of nibbling double-stranded DNA, may require a single-stranded tail for efficient binding. This hypothesis implies that the P nucleotide addition is a prerequisite for the exonucleolytic nibbling.

Finally, if the P nucleotide addition is indeed an obligatory event for the generation of coding joints, but not for the generation of signal joints, it may be the step at which the immunoglobulin and TCR gene rearrangement is defective in the SCID mice (Lieber et al., 1988b; Okazaki et al., 1988; Malynn et al., 1988).

Experimental Procedures

T Cell Hybridomas

Hybridomas were prepared by fusing BW5147 thymoma cells with $CD4^-CD8^-$ thymocytes of various stages of development as described previously (Ito et al., 1989). All hybridomas express γδ TCR as assessed by immunoprecipitation of ^{125}I surface-labeled cells with anti-CD3 and anti-γ antibodies (see Ito et al., 1989, for details).

Thymocytes

C57BL/6 mice (Jackson Laboratory) were mated at night, and females were examined the next morning. Day 0 of embryonic development was considered to be the day a vaginal plug was found.

Thymi of several mice were removed by dissection and homogenized. Thymocytes were then washed once with phosphate buffered saline. One fetal thymus typically gave 5×10^4 (day 14.5), 3×10^5 (day 16), or 3×10^6 (day 17.5) cells.

Nucleic Acids

Total cellular DNA was extracted by the SDS/proteinase K/phenol method according to Maki et al. (1981), and total RNA was prepared using the guanidinium thiocyanate/CsCl method (Chirgwin et al., 1979).

cDNA was synthesized with 10 μg of total RNA from the hybridomas or thymocytes, using 100 pmol of the C_γ or $J_{\gamma 1}$ primer, 40 U of human placental ribonuclease inhibitor (Amersham), and 20 U of Reverse Transcriptase (Amersham). Half of the product of the cDNA synthesis was heated at 65°C for 15 min and used directly for the PCR.

PCR

The reaction was performed using T. aquaticus polymerase (Saiki et al., 1988) and a Perkin-Elmer-Cetus thermal cycler. Each cycle consists of incubations at 94°C for 1 min, followed by 55°C for 2 min and 72°C for 3 min. Before the first cycle, a 2 min 94°C denaturation step was included, and after the 25th cycle the extension at 72°C was prolonged for 7 min. Reactions with genomic DNA from total thymocytes included 2 μg of EcoRI-digested DNA, 50 pmol of each phosphorylated primer, 0.2 mM each dNTP, 50 mM KCl, 10 mM Tris–HCl (pH 8.3), 1.5 mM $MgCl_2$, 0.01% gelatin, and 2.5 U of Taq polymerase.

PCR Primers

For coding joints:

$J_{\gamma 1}$:	STP.100:	5′-CAGAGGGAATTACTATGAGC-3′
$V_{\gamma 4}$:	STP.073:	5′-TGTCCTTGCAACCCCTACCC-3′
$V_{\gamma 5}$:	STP.094:	5′-TGTGCACTGGTACCAACTGA-3′
$V_{\gamma 6}$:	STP.107:	5′-GGAATTCAAAAGAAAACATTGTCT-3′
C_δ: (for cDNA)	STP.074:	5′-GTCATTTTCAGGTTCTGGTAG-3′
$V_{\delta 1}$:	STP.096:	5′-GAATGGAACTAATGCTCTGT-3′
$J_{\delta 1}$:	STP.101:	5′-TTGGTTCCACAGTCACTTGG-3′
$J_{\delta 2}$:	STP.097:	5′-CCAACTTACGGGGCTCCAC-3′

For signal joints:

$J_{\gamma 1}$:	STJL.027:	5′-GCCTGTTTCTCAGGACATAAATA-3′
$V_{\gamma 5}$:	STJL.029:	5′-AGGACTAGGGCACCAAGGGGATA-3′
$V_{\gamma 6}$:	STJL.028:	5′-AGCCTAGGTGTCTGTGCAGGTGA-3′
$J_{\delta 2}$:	STJL.031:	5′-CAGGCCCAGGGCTGGTCCCAG-3′
$D_{\delta 2}$:	STJL.030:	5′-AGGCCTGGGAGACGTTCTTCA-3′

DNA Sequencing

Ten percent of the PCR products were fractionated by electrophoresis in 5.5% polyacrylamide gel, and the appropriate fragments were extracted from the gel in 0.5 M NH_4Ac/1 mM EDTA. The fragments were then ligated in the SmaI site of the vector M13mp19 (Norrander et al., 1983) and sequenced using the dideoxynucleotide chain termination method (Sanger et al., 1977) with Sequenase (US Biochemicals), following the manufacturer's instructions.

Acknowledgments

We thank Edvins Krecko for excellent technical assistance. We also thank Charles Steinberg and David Gerber for helpful comments and Elly Basel for secretarial help. This work was supported by grants from Howard Hughes Medical Institute, National Institutes of Health, American Cancer Society, and Ajinomoto Co., Ltd. to S. T. as well as the NIH CORE P30-CA14051. J. J. L. is supported by Conselho Nacional de Pesquisas (CNPq), Brazil. We also thank FUJI Photo Film Co. for the use of the BA100-BioImage Analyzer.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received August 9, 1989; revised September 20, 1989.

References

Allison, J., McIntyre, B., and Block, D. (1982). Tumor-specific antigen of murine T lymphoma defined with monoclonal antibody. J. Immunol. *129*, 2293–2300.

Alt, F. W., and Baltimore, D. (1982). Joining of immunoglobulin heavy chain gene segments: implications from a chromosome with evidence of three D–J_H fusions. Proc. Natl. Acad. Sci. USA *79*, 4118–4122.

Alt, F. W., Rosenberg, N., Enea, V., Siden, E., and Baltimore, D. (1982). Multiple immunoglobulin heavy-chain gene transcripts in Abelson Murine Leukemia virus–transformed lymphoid cell lines. Mol. Cell. Biol. *2*, 386–400.

Asarnow, D. M., Kuziel, W. A., Bonyhadi, M., Tigelaar, R. E., Tucker, P. W., and Allison, J. P. (1988). Limited diversity of γδ antigen receptor genes of Thy-1$^+$ dendritic epidermal cells. Cell *55*, 837–847.

Bank, I., DePinho, R. A., Brenner, M. B., Cassimeris, J., Alt, F. W., and Chess, L. (1986). A functional T3 molecule associated with a novel heterodimer on the surface of immature human thymocytes. Nature *322*, 179–181.

Barth, R. K., Kim, B. S., Lan, N. C., Hunkapiller, T., Sobieck, N., Winoto, A., Gershenfeld, H., Okada, C., Hansburg, D., Wiessman, I. L., and Hood, L. (1985). The murine T-cell receptor uses a limited repertoire of expressed V_β gene segment. Nature *316*, 517–523.

Behlke, M. A., Spinella, D. G., Chou, H. S., Sha, W., Hartl, D. L., and Loh, D. Y. (1985). T-cell receptor beta-chain expression: dependence on relatively few variable region genes. Science *229*, 566–570.

Behlke, M. A., Chou, H. S., Huppi, K., and Loh, D. Y. (1986). Murine T-cell receptor mutants with deletions of beta-chain variable region genes. Proc. Natl. Acad. Sci. USA *83*, 767–771.

Blackman, M., Yagüe, J., Kubo, R., Gay, D., Coleclough, C., Palmer, E., Kappler, J., and Marrack, P. (1986). The T cell repertoire may be biased in favor of MHC recognition. Cell *47*, 349–357.

Bluestone, J. A., Pardoll, D., Sharrow, S. O., and Fowlkes, B. J. (1987). Characterization of murine thymocytes with CD3-associated T-cell receptor structures. Nature *326*, 82–84.

Boersch-Supan, M. E., Agarwal, S. A., White-Scharf, M. E., and Imanishi-Kari, T. (1985). Heavy chain variable region. Multiple gene segments encode anti-4-(hydroxy-3-nitrophenyl)acetyl idiotypic antibodies. J. Exp. Med. *161*, 1272–1292.

Bonneville, M., Janeway, C. A., Jr., Ito, K., Haser, W., Ishida, I., Nakanishi, N., and Tonegawa, S. (1988). Intestinal intraepithelial lymphocytes are a distinct set of γδ T cells. Nature *336*, 479–481.

Borst, J., van de Griend, R. J., van Oostveen, J. W., Ang, S. L., Melief, C. J., Seidman, J. G., and Bolhuis, R. L. H. (1987). A T-cell receptor γ/CD3 complex found on cloned functional lymphocytes. Nature *325*, 683–688.

Bottino, C., Tambussi, G., Ferrini, S., Ciccone, E., Varese, P., Mingari, M. C., Moretta, L., and Moretta, A. (1988). Two subsets of human T lymphocytes expressing gamma/delta antigen receptor are identifiable by monoclonal antibodies directed to two distinct molecular forms of the receptor. J. Exp. Med. *168*, 491–505.

Brenner, M. B., McLean, J., Dialynas, D. P., Strominger, J. L., Smith, J. A., Owen, F. L., Seidman, J. G., Ip, S., Rosen, F., and Krangel, M. S. (1986). Identification of a putative second T-cell receptor. Nature *322*, 145–149.

Chien, Y.-h., Iwashima, M., Kaplan, K. B., Elliott, J. F., and Davis, M. M. (1987a). A new T-cell receptor gene located within the alpha locus and expressed early in T-cell differentiation. Nature *327*, 677–682.

Chien, Y.-h., Iwashima, M., Wettstein, D. A., Kaplan, K. B., Elliott, J. F., Born, W., and Davis, M. M. (1987b). T-cell receptor δ gene rearrangements in early thymocytes. Nature *330*, 722–727.

Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J., and Rutter, W. J. (1979). Isolation of biologically active ribonucleic acid from sources enriched in ribonuclease. Biochemistry *18*, 5294–5299.

Clarke, S. H., Huppi, K., Ruezinsky, D., Staudt, L., Gerhard, W., and Weigert, M. (1985). Inter- and intraclonal diversity in the antibody response to influenza hemagglutinin. J. Exp. Med. *161*, 687–704.

Davis, M. M., and Bjorkman, P. J. (1988). T-cell antigen receptor genes and T-cell recognition. Nature *334*, 395–402.

Deev, S. M., Combriato, G., Klobeck, H. G., and Zachau, H. G. (1987). Reciprocal recombination products of V_κ-J_κ joining reactions in human lymphoid cell lines. Nucl. Acids Res. *15*, 1–14.

Desiderio, S. V., Yancopoulos, G. D., Paskind, M., Thomas, E., Boss, M. A., Landau, N., Alt, F. W., and Baltimore, D. (1984). Insertion of N regions into heavy-chain genes is correlated with expression of terminal deoxytransferase in B cells. Nature *311*, 752–755.

Elliott, J. F., Rock, E. P., Patten, P. A., Davis, M. M., and Chien, Y.-h. (1988). The adult T-cell receptor δ-chain is diverse and distinct from that of fetal thymocytes. Nature *331*, 627–631.

Goodman, T., and Lefrancois, L. (1988). Expression of the gamma/delta T-cell receptor on intestinal CD8$^+$ intraepithelial lymphocytes. Nature *333*, 855–858.

Goverman, J., Minard, K., Shastri, N., Hunkapiller, T., Hansburg, D., Sercarz, E., and Hood, L. (1985). Rearranged β T cell receptor genes in a helper T cell clone specific for lysozyme: no correlation between V_β and MHC restriction. Cell *40*, 859–867.

Haskins, K., Kubo, R., White, J., Pigeon, M., Kappler, J., and Marrack, P. (1983). The major histocompatibility complex–restricted antigen receptor on T cells. I. Isolation with a monoclonal antibody. J. Exp. Med. *157*, 1149–1169.

Havran, W. L., and Allison, J. P. (1988). Developmentally ordered appearance of thymocytes expressing different T-cell antigen receptors. Nature *335*, 443–445.

Hayday, A. C., Saito, H., Gillies, S. D., Kranz, D. M., Tanigawa, G., Eisen, H. N., and Tonegawa, S. (1985). Structure, organization, and somatic rearrangement of T cell gamma genes. Cell *40*, 259–269.

Hedrick, S. M., Nielsen, E. A., Kavaler, J., Cohen, D. I., and Davis, M. M. (1984). Sequence relationships between putative T-cell receptor polypeptides and immunoglobulins. Nature *308*, 153–158.

Heilig, J. S., and Tonegawa, S. (1987). T-cell γ gene is allelically but not isotypically excluded and is not required in known functional T-cell subsets. Proc. Natl. Acad. Sci. USA *84*, 8070–8074.

Huck, S., Dariavach, P., and Lefranc, M. P. (1988). Variable region genes in the human T-cell rearranging gamma (TRG) locus: V–J junction and homology with the mouse genes. EMBO J. *7*, 719–726.

Ito, K., Bonneville, M., Takagaki, Y., Nakanishi, N., Kanagawa, O., Krecko, E., and Tonegawa, S. (1989). Different γδ T-cell receptors are expressed on thymocytes at different stages of development. Proc. Natl. Acad. Sci. USA *86*, 631–635.

Itohara, S., Nakanishi, N., Kanagawa, O., Kubo, R., and Tonegawa, S. (1989). Monoclonal antibodies specific to native murine T cell receptor γδ: analysis of γδ T cells in thymic ontogeny and peripheral lymphoid organs. Proc. Natl. Acad. Sci. USA *86*, 5094–5098.

Janeway, C. A., Jr., Jones, B., and Hayday, A. (1988). Specificity and function of T cells bearing γδ receptors. Immunol. Today *9*, 73–76.

Koning, F., Stingl, G., Yokoyama, W. M., Yamada, H., Maloy, W. L., Tschachler, E., Sevach, E. M., and Coligan, J. E. (1987). Identification of a T3-associated gamma delta T cell receptor on Thy-1$^+$ dendritic epidermal cell lines. Science *236*, 834–837.

Korman, A. J., Galesic, S. M., Spencer, D., Kruisbeek, A. M., and Raulet, D. (1988). Predominant variable region gene usage by γ/δ T cell receptor–bearing cells in the adult thymus. J. Exp. Med. *168*, 1021–1039.

Kranz, D. M., Saito, H., Heller, M., Takagaki, Y., Haas, W., Eisen, H. N., and Tonegawa, S. (1985). Limited diversity of the rearranged T-cell γ gene. Nature *313*, 752–755.

Kurosawa, Y., and Tonegawa, S. (1982). Organization, structure, and assembly of immunoglobulin heavy chain diversity DNA segments. J. Exp. Med. *155*, 201–218.

Kuziel, W. A., Takashima, A., Bonyhadi, M., Bergstresser, P. R., Allison, J. P., Tigelaar, R. E., and Tucker, P. W. (1987). Regulation of T-cell receptor γ-chain RNA expression in murine Thy-1$^+$ dendritic epidermal cells. Nature *328*, 263–266.

Lacy, M. J., McNeil, L. K., Roth, M. E., and Kranz, D. M. (1989). T-cell receptor δ-chain diversity in peripheral lymphocytes. Proc. Natl. Acad. Sci. USA *86*, 1023–1026.

Landau, N. R., Schatz, D. G., Rosa, M., and Baltimore, D. (1987). Increased frequency of N-region insertion in a murine pre-B-cell line infected with a terminal deoxynucleotidyl transferase retroviral expression vector. Mol. Cell. Biol. *7*, 3237–3243.

Lewis, S., Gifford, A., and Baltimore, D. (1985). DNA elements are asymetrically joined during the site-specific recombination of kappa immunoglobulin genes. Science *228*, 677–685.

Lewis, S. M., Hesse, J. E., Mizuuchi, K., and Gellert, M. (1988). Novel strand exchanges in V(D)J recombination. Cell *55*, 1099–1107.

Lieber, M. R., Hesse, J. E., Mizuuchi, K., and Gellert, M. (1988a). Lymphoid V(D)J recombination: nucleotide insertion at signal joints as well as coding joints. Proc. Natl. Acad. Sci. USA *85*, 8588–8592.

Lieber, M. R., Hesse, J. E., Lewis, S., Bosma, G. C., Rosenberg, N., Mizuuchi, K., Bosma, M. J., and Gellert, M. (1988b). The defect in murine severe combined immune deficiency: joining of signal sequences by not coding segments in V(D)J recombination. Cell *55*, 7–16.

Loh, E. Y., Cwirla, S., Serafini, A. T., Phillips, J. H., and Lanier, L. L. (1988). Human T-cell receptor delta chain: genomic organization, diversity, and expression in populations of cells. Proc. Natl. Acad. Sci. USA *85*, 9714–9718.

Loh, E. Y., Elliott, J. F., Cwirla, S., Lanier, L. L., and Davis, M. M. (1989). Polymerase chain reaction with single-sided specificity: analysis of T cell receptor delta chain. Science *243*, 217–220.

Maki, R., Roeder, W., Traunecker, A., Sidman, C., Wabl, M., Raschke, W., and Tonegawa, S. (1981). The role of DNA rearrangement and alternative RNA processing in the expression of immunoglobulin delta genes. Cell *24*, 353–365.

Malissen, M., McCoy, C., Blanc, D., Trucy, J., Devaux, C., Schmitt-Verhulst, A.-M., Fitch, F., Hood, L., and Malissen, B. (1986). Direct evidence for chromosomal inversion during T-cell receptor β-gene rearrangements. Nature *319*, 28–33.

Malynn, B. A., Blackwell, T. K., Fulop, G. M., Rathbun, G. A., Furley, A. J. W., Ferrier, P., Heinke, L. B., Phillips, R. A., Yancopoulos, G. D., and Alt, F. W. (1988). The *scid* defect affects the final step of the immunoglobulin VDJ recombinase mechanism. Cell *54*, 453–460.

McCormack, W. T., Tjoelker, L. W., Carlson, L. M., Petryniak, B., Barth, C. F., Humphries, E. H., and Thompson, C. B. (1989). Chicken Ig_L gene rearrangement involves deletion of a circular episome and addition of single nonrandom nucleotides to both coding segments. Cell *56*, 785–791.

Meuer, S. C., Fitzgerald, K. A., Hussey, R. E., Hodgdon, J. C., Schlossman, S. F., and Reinhertz, E. L. (1983). Clonotypic structures involved in antigen-specific human T-cell function. Relationship to the T3 molecular complex. J. Exp. Med. *157*, 705–719.

Milner, E. C. B., Meek, K. D., Rathbun, G., Tucker, B., and Capra, J. D. (1986). Are anti-arsonate antibody N-segments selected at both the protein and the DNA level? Immunol. Today *7*, 36–40.

Norrander, J., Kempe, T., and Messing, J. (1983). Construction of improved M13 vectors using oligonucleotide-directed mutagenesis. Gene *26*, 101–106.

Novotny, J., Tonegawa, S., Saito, H., Kranz, D. M., and Eisen, H. N. (1986). Secondary, tertiary, and quaternary structure of T-cell-specific immunoglobulin-like polypeptide chains. Proc. Natl. Acad. Sci. USA *83*, 742–746.

Okazaki, K., Davis, D. D., and Sakano, H. (1987). T cell receptor β gene sequences in the circular DNA of thymocyte nuclei: direct evidence for intramolecular DNA deletion of V-D-J joining. Cell *49*, 477–485.

Okazaki, K., Nishikawa, S., and Sakano, H. (1988). Aberrant immunoglobulin gene rearrangement in SCID mouse bone marrow cells. J. Immunol. *141*, 1348–1352.

Patten, P., Yokota, T., Rothbard, J., Chien, Y.-h., Arai, K., and Davis, M. M. (1984). Structure, expression and divergence of T-cell receptor beta-chain variable regions. Nature *312*, 40–46.

Rocca-Serra, J., Matthes, H. W., Kaartinen, M., Milstein, C., Théze, J., and Fougereau, M. (1983). Analysis of antibody diversity: V-D-J mRNA nucleotide sequence of four anti-GAT monoclonal antibodies. A paucigene system using alternate D-J recombinations to generate functionally similar hypervariable regions. EMBO J. *2*, 867–872.

Roth, M. E., Lacy, M. J., McNeil, L. K., and Kranz, D. M. (1988). Selection of variable-joining region combinations in the alpha chain of the T cell receptor. Science *241*, 1354–1358.

Rothenberg, E., and Triglia, D. (1983). Clonal proliferation unlinked to terminal deoxynucleotidyl transferase synthesis in thymocytes of young mice. J. Immunol. *130*, 1627–1633.

Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B., and Erlich, H. A. (1988). Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. Science *239*, 487–491.

Saito, H., Kranz, D. M., Takagaki, Y., Hayday, A. C., Eisen, H. N., and Tonegawa, S. (1984). Complete primary structure of heterodimeric T-cell receptor deduced from cDNA sequences. Nature *309*, 757–762.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA *74*, 5463–5467.

Takagaki, Y., Nakanishi, N., Ishida, I., Kanagawa, O., and Tonegawa, S. (1989a). T cell receptor-γ and -δ genes preferentially utilized by adult thymocytes for the surface expression. J. Immunol. *142*, 2112–2121.

Takagaki, Y., DeCloux, A., Bonneville, M., and Tonegawa, S. (1989b). γδ T cell receptors on murine intestinal intra-epithelial lymphocytes are highly diverse. Nature *339*, 712–714.

Tonegawa, S. (1983). Somatic generation of antibody diversity. Nature *302*, 575–581.

Traunecker, A., Oliveri, F., Allen, N., and Karjalainen, K. (1986). Normal T cell development is possible without "functional" gamma chain genes. EMBO J. *5*, 1589–1593.

Urban, J. L., Kumar, V., Kono, D. H., Gomez, C., Horvath, S. J., Clayton, J., Ando, D. G., Sercarz, E. E., and Hood, L. (1988). Restricted use of T cell receptor V genes in murine autoimmune encephalomyelitis raises possibilities for antibody therapy. Cell *54*, 577–592.

Weiss, A., Newton, M., and Commie, D. (1986). Expression of T3 in association with a molecule distinct from the T-cell antigen receptor heterodimer. Proc. Natl. Acad. Sci. USA *83*, 6998–7002.

Wysocki, L., Manser, T., and Gefter, M. L. (1986). Somatic evolution of variable region structures during an immune response. Proc. Natl. Acad. Sci. USA *83*, 1847–1851.