Enumeration and characterization of DJH structures in mouse fetal liver

Yenhui Chang, Christopher J.Paige and Gillian E.Wu¹

Department of Immunology, Medical Sciences Building, University of Toronto, Toronto, Ontario Canada M5S 1A8 ¹Corresponding author

Communicated by K.Rajewsky

The primary immunoglobulin (Ig) repertoire in the mouse develops during fetal life in the liver. The first Ig gene rearrangement-the joining of a DH to a JH gene segment-contributes largely to the diversity found in CDR3, as well as potentially encoding the $D\mu$ protein which is believed to function in the development of a B cell. In this report, the number of DJH joins in two mouse strains, C57BL/6 and BALB/c, were enumerated from days 12 to 16 of fetal development. It was found that the number of DJH structures increased from <300 per liver on day 12 to >700 000 (C57BL/6) and 300 000 (BALB/c) on day 16. Each JH gene segment was used approximately equally on each day examined. When the DJH structures were examined by cloning and sequencing it was found that the DJH reading frame (RF) usage (with respect to JH) was not random-RF1 was used 70% of the time. Morevover, a single D gene segment, DFL16.1, was used in >50% of all joins reinforcing the notion that the fetal repertoire is restricted in its antigen binding potential.

Key words: B cell development/fetal liver repertoire/ immunoglobulin gene rearrangement

Introduction

In embryonic mice, hematopoiesis is initiated in the yolk sac blood islands and subsequently localizes to the fetal liver. The first genetic events associated with B cell differentiation, rearrangement of the immunogloblin (Ig) gene segments, is initiated during this period (Melchers, 1979; Paige, 1983; Kincade, 1987). To date, the low frequency of rearranged Ig gene structures in fetal tissue has made it difficult to monitor them. PCR now makes it possible to obtain such information. In this report, we enumerate and describe the DJH Ig gene segments during these early stages of the developmental process.

Rearrangements in the heavy chain locus utilize three clusters of gene segments: variable (VH), diversity (DH) and joining (JH) (reviewed in Tonegawa, 1983; Alt *et al.*, 1986). In the mouse, it is estimated that there are well over 100 VH gene segments which have been grouped into at least 13 families based on DNA sequence similarity. The DH gene segments have been similarly grouped into at least three families: DQ52 (one member), DFL (two members) and DSP (at least ten members) (Kurosawa and Tonegawa, 1982; Wood and Tonegawa, 1983; Ichihara *et al.*, 1989). There

© Oxford University Press

may be other, as yet undetected, DH gene segments. There are four JH gene segments named JH1 to JH4. In the mouse, four JH gene segments lie within 1.3 kb of DNA (Sakano *et al.*, 1980; Alt *et al.*, 1986). The D segments are scattered over \sim 80 kb of DNA, the closest (DQ52) lying 700 bp 5' to the JH cluster. The 5' and 3' borders of each DH segment and the 5' borders of each JH gene segment are flanked by conserved recombination signal sequences (RSS) which consist of a palindromic heptamer, a 12 bp spacer (DH), or a 23 bp spacer (JH) and a conserved nonamer (reviewed in Tonegawa, 1983).

A VDJ gene segment which encodes the variable region of an IgH molecule is formed by the joining of one segment from each of the VH, DH and JH gene clusters. Since the first event in Ig gene rearrangement is the joining of DH to JH (Alt et al., 1986; Rolink and Melchers, 1991) and this determines a large part of the antigen binding repertoire, we have directed our initial analysis to the DJH structure. DJH is a stable intermediate in the process leading to a complete VDJ gene segment. The intermediate may function solely as a substrate for VH to DJH joining. Alternatively, evidence demonstrating the presence of a 'D μ ' protein on the cell surface has led to suggestions that this protein may influence B cell development (Tsubata et al., 1991). The DH to JH joining event is usually imprecise with associated deletions and 'N' insertions. The DNA that results encodes most of the third complementarity determining region (CDR3) of the mature Ig heavy chain region. This region has the highest variability in the Ig molecule (Kabat et al., 1987), contributing to a large part of the diversity found in the primary Ig repertoire in the mouse. This function makes it important to determine what governs the selection and recombination of particular JH and DH gene segments.

Using a PCR assay to quantify DJH rearrangements in DNA, we have examined these rearrangements in the developing hematopoietic tissue where the primary repertoire is generated. The results of this analysis leads to five conclusions: (i) from days 12 to 16 of gestation the number of DJH rearrangement increases ~ 10-fold each day such that at day 16 ~ 1 in 80 cells have DJH rearrangements; (ii) the usage of each JH segment is approximately equal and does not change during the days examined; (iii) the usage of the DH genes is not random with the DFL gene family being used most frequently; (iv) all DJH joins examined had associated deletions, however, with one possible exception, no 'N' additions were found; and (v) the joint structures were such that the resulting reading frame usage was not random with respect to JH.

Results

Development of a standard PCR assay for DJH rearrangement

In order to detect rearrangements of DH and JH gene segments in genomic DNA, we designed DH and JH



Fig. 1. Genomic map of the DH locus (adapted from Kurosawa and Tonegawa, 1982 and Ichihara et al., 1989) showing oligomer primer position and expected products. Not drawn to scale.



Fig. 2. Southern blot analysis of PCR amplification of DNA with the DSF/JH4 primer pair from the standard A-MuLV lines probed with JH4. Number above each lane refers to the number of targets of each DJH rearrangement. Thus, in the lanes marked 10 there are 10 JH1, 10 JH2, 10 JH3 and 10 JH4 targets. (A) 10-5000 range in number of targets. (B) 0-100 range with five separate amplifications of 10 targets. M = molecular weight marker, 1 kb ladder (BRL).

oligonucleotide primers based on a consensus sequence derived from a comparison of sequences found in Genbank (version 60.0) They are described in Materials and methods. The two DH primers selected, designated as DSF and DQ52, are immediately 5' of the corresponding DH gene segments. The JH4 primer is immediately 3' of JH4. The DSF/JH4 oligomer pair amplifies the products of recombinations between both DSP and DFL gene segments and all four JH elements. The DQ52/JH4 oligomer pair amplifies the DNA products of recombination between DQ52 and all four JH segments, as well as the unrearranged germline DNA between DQ52 and JH4 DNA (Figure 1). Neither primer pair amplifies DNA containing VDJH rearrangements. The specificity of the primer pairs was verified with cell lines and plasmid DNA containing known DJH rearrangements, as described in Material and methods.

We expected the DSF/JH4 and the DQ52/JH4 amplification products to be five distinct sizes, ranging from 2 kb to ~120 bp. These products result from the amplification of the germline unrearranged DQ52-JH (2 kb) locus and from the amplification of DH gene segments joined to each of the four JHs (1.3-0.12 kb). Because PCR amplifies fragments of different sizes to different extents, we developed a standard assay for each expected product which allowed us to quantify the range of products expected from fetal liver DNA amplification. The required DNA standards were identified by screening DNA from our previously characterized collection of A-MuLV transformed pre-B cell lines (Atkinson *et al.*, 1991) as described in Material and methods.

DNA from the four selected standard lines was mixed in equal amounts and serially diluted into a constant amount



Fig. 3. Southern blot analysis of PCR amplification of DNA from C57BL/6 fetal livers using the DSF/JH4 primer pair. The numbers in the lane headings are the age in days of the fetus from which the DNA was made. Amplifications of DNA from standard lines were run in parallel as in Figure 2. (A) Complete analysis. (B) Analysis of five day 12 amplifications. The large band in the second lane marked * is the size of a germline DQ52-JH4 fragment. This band was not seen in other amplifications.

of DNA from a subcloned VDJ/VDJ A-MuLV cell line (CB32.12), which has no detectable DJH product upon PCR amplification. PCR conditions were identified in which the amount of amplified product was proportional to the amount of target DNA between 50 and 5000 PCR target sites when analysed by Southern blotting with a JH4 probe (Figure 2A). The limit of sensitivity of our assay was found to be 10 of each of the four targets per μ g DNA (1 μ g $\approx 1.7 \times 10^5$ cells). This conclusion was verified in parallel by repetitive amplification of a number of samples containing 10 targets/ μ g DNA. All four DJH bands were amplified, although not in every reaction (Figure 2B). Thus 10 targets/ μ g DNA is the limit of sensitivity in our standard assay.

Quantification of DJH rearrangements in C57BL/6 mouse fetal liver

In order to quantify precisely the number of DJH targets in a given sample of fetal liver DNA, the assay system standards were run in parallel in each experiment. DNA was obtained from pools of six to eight C57BL/6 fetal livers isolated on days 12 to 16 of gestation and amplified as described. Figure 3A shows a representative Southern blot of PCR reaction products using the DSF/JH4 primer pair. This experiment is tabulated in Table I as the number of rearrangements per fetal liver, estimated from the standard curve derived from densitometry tracings in this experiment. With the DSF/JH4 primer pair on day 12, there were <40DJH rearrangements per μg DNA, which is the limit of sensitivity for standard assay conditions. Multiple repetitions of identical PCR reactions with day 12 fetal liver DNA extend the range of the assay and we subsequently found ~33 DJHs per μg DNA (Figure 3B). Since there are $\sim 1.0 \times 10^6$ cells in a fetal liver on day 12, this indicated that ~ 200 DJHs per fetal liver are detectable with the DSF/JH4 primer pair. From days 13 to 15, the number of DJH rearrangements per fetal liver increases about 10-fold

each day. On day 16, the increase is ~ 2 -fold more than on day 15, signifying perhaps increasing numbers of VDJH rearrangements (which are not amplified). By day 16 about one in 80 to 100 cells in the C57BL/6 fetal liver has a DSF to JH rearrangement.

Using the DQ52/JH4 primer pair, we found that rearrangements of DQ52 were detectable on day 12 and increased to ~420 rearrangements per μ g DNA on day 16 (6.8×10⁴/liver) (Table I). Since there is one DQ52 gene segment and at least 12 DSP and DFL segments, this result indicates that DQ52 usage conforms to predictions based on family size.

Quantification of DJH rearrangements in BALB/c fetal liver

A representative Southern blot analysis of amplifications of day 12 to 16 BALB/c fetal DNA (obtained as described for Figure 3) with the DSF/JH4 primer pair is shown in Figure 4. The number of DSFJH and DQ52JH rearrangements are tabulated in Table I.

The number of DSFJH rearrangments per fetal liver increases from ~ 200 on day 12 to 3.2×10^5 on day 16. Similarly, DQ52JH rearrangements increase in number from days 13 to 16 where there are ~ 160 rearrangements per μg DNA (2.6×10⁴/liver). No DQ52JH rearrangements were detectable on day 12. In an attempt to understand this absence, we further investigated the sensitivity of the primer pair. We amplified DNA with a DQ52JH2 rearrangement (70Z/3 DNA) at decreasing target numbers diluted with CB32.12 DNA (which has no DJH targets) or day 12 fetal liver. As shown in Figure 5, 10 copies of the 70Z/3, DQ52JH2 rearrangements can be amplified. However, in the parallel experiment with day 12 fetal liver DNA alone, no targets are detectable. This control experiment verifies that our assay has the sensitivity to detect 10 DQ52JH2 rearrangements per µg DNA and thus the absence of rearrangements reflects the absence of targets. Based on this

Strain	DJH rearrangement per μg DNA ^a										
Day of Gestation		DFS	/JH4 prime	r pair		DQ52/J	H4 primer				
	DJHI	DJH2	DJH3	DJH4	total	DJH1	DJH2	DJH3	DJH4	total	
C57BL/6											_
12	≤10	≤10	≤10	≤10	≤40	<2	<2	<2	<2	<8	$< 2.8 \times 10^{2}$
13	21	28	68	25	141	<2	2	4	<2	≤10	$\leq 4.4 \times 10^3$
14	104	137	150	100	491	4	4	4	4	16	3.1×10^{4}
15	510	505	690	540	2245	25	25	25	25	100	2.8×10^{5}
16	850	880	1280	860	3870	100	100	120	100	420	7.3×10^{5}
BALB/c											
12	≤10	≤10	≤10	≤10	≤40	ND ^c	ND	ND	ND	<8	$< 2.8 \times 10^{2}$
13	81	61	68	38	248	<2	4	<2	<2	≤10	$\leq 7.6 \times 10^3$
14	150	195	275	190	810	25	25	30	25	105	5.7×10^{4}
15	210	225	328	145	908	25	25	25	25	100	1.2×10^{5}
16	400	405	590	505	1900	50	30	50	30	160	3.5×10^{5}

Table I. Number of DJH rearrangements in mouse fetal liver

^aAssuming there is 6 pg of DNA per cell, 1 μ g of DNA corresponds to 1.7 \times 10⁵ cells.

^bThe number of cells per liver (taken from Paige *et al.*, 1984) is: day 12, 1.0×10^6 ; day 13, 5.0×10^6 ; day 14, 10.5×10^6 ; day 15, 20.2×10^6 ; day 16, 29.0 × 10⁶.

^cND - not detectable.

Data are the average from three experiments-See Materials and methods.



Fig. 4. Southern blot analysis of PCR amplification of DNA from BALB/c fetal livers. Legend as in Figure 3.

control result, day 12 BALB/c fetal livers have <10 DQ52JH rearrangements per μ g of DNA or <70 per fetal liver.

On day 16, there are a total of 3.5×10^5 DSFJH plus DQ52JH rearrangements in a BALB/c fetal liver, compared with 7.3×10^5 for a same sized C57BL/6 fetal liver. As described in Material and methods, these differences indicate that there are more DJH structures in C57BL/6 than in BALB/c at this stage of development.

JH usage in C57BL/6 and BALB/c fetal livers

The intensity on autoradiograms of the JH amplification bands of the DNA standards was, as expected, different for the four amplification products. In order to determine the JH usage of the DJH rearrangements, the intensities of the experimental DJH bands were compared with the standards.

Densitometry analyses of the JH usage indicate that for both C57BL/6 and BALB/c fetal liver DNA, the relative JH usage ranges from ~ 20 to 35% and does not change

1894

significantly during days 14 to 16 of gestation. By day 16, the JH usage for C57BL/6 is: JH1, 22%; JH2, 23%; JH3, 32%; and JH4, 23%. For BALB/c, the JH usage is: JH1, 22%; JH2, 23%; JH3, 33%; and JH4, 23%. That is, all four JH gene segments are used at comparable frequencies in both C57BL/6 and BALB/c. JH3 was used at a slightly higher frequency than expected for random usage (33% versus the expected 25%) on all days examined.

DH usage in BALB/c 16 day fetal liver

The DH gene sequences from GenBank (version 60.0) and published reports are listed in Figure 6. These DH genes are all from the BALB/c strain. Thus, to identify with confidence DH and JH sequences, we concentrated on analysing this strain. The DNA sequences of 40 DSFJH products from day 16 BALB/c fetal liver DNA were determined after cloning of the amplified targets in pBlueScribe (Stratagene). DJH1, DJH2, DJH3 and DJH4 products were cloned separately and their sequences are



Fig. 5. Southern blot analysis of PCR amplification of DNA from 70Z/3 and day 12 BALB/c fetal liver using the DQ52/JH4 primer pair. Lane 1, 1 μ g CB32.12 DNA; lane 2, 600 pg 70Z/3 DNA plus 1 μ g CB32.12 DNA; lane 3, 60 pg 70Z/3 DNA plus 1 μ g CB32.12 DNA; lane 4, 600 pg 70Z/3 DNA plus 1 μ g of day 12 fetal liver DNA digested with *Sst*I and *Hinc*II as described in Materials and methods; lane 5, 60 pg 70Z/3 DNA plus 1 μ g of digested day 12 fetal liver DNA; lane 6, 1 μ g digested day 12 fetal liver DNA.

presented in Figure 7. The nucleotides that have been deleted as a result of the joining process are included and so marked in the figure. Although many of the potential DFL and DSP targets in the genome were found in the DJH structures, there was a clear overusage of DFL16.1. DFL16.1 is used in 8/11 of the JH1 clones (73%) and 3/7 (43%), 5/9 (56%) and 6/13 (46%) of the clones with JH2, JH3 and JH4, respectively (Table II). Taking the DQ52 usage into account, 50% of the DJH joins used DFL16.1, 14% DFL16.2, 10% DQ52, 10% DSP2.9, 7% DSP2.11. The remainder of the DSP gene segments were used <5% of the time.

Deletions at DJH joins

The DJH joining event is usually imprecise with nucleotide additions and deletions commonly found at the join (Tonegawa, 1983; Meek et al., 1989; Gu et al., 1990; Meek, 1990; Rolink et al., 1991). All the DJH joins sequenced here had deletions in either DH or JH, or both. There were no DJH joins without deletions and, as has been the case with other analyses of DJH joins from fetal liver (with one possible exception), no N insertions were found. The deletions ranged from two to 15 nucleotides, the average being 6.7 deletions per DJH join. Examination of the sequences flanking the DJH and JH gene segments reveal that the terminal 3' of all the DFL and DSP gene segments is either CTAC or CGAC; likewise, the 5' sequences of three of the JH segments are similar-JH1 is CTAC, JH2 is ACTAC and JH4 is ATTAC. (JH3 bears no similarity being TGCCT.) DJH structures using JH1, JH2 and JH4 which have homologies had an average of 4.8, 6.0 and 6.7 deletions per join on average. Structures using JH3, which has no sequence homology, had 9.4 deletions per join. JH3 structures are also notable in that a G was present at the D-J border of one sequence (clone 20, Figure 7). This G could be an 'N' insertion, (not found previously in fetal liver VDJ joins) or the final G of the heptamer of the JH3 RSS

Dsp 2.1	TCT ACT ATG GTA ACT AC	(1)
Dsp 2.2	TCT ACT ATG ATT ACG AC	(2)
Dsp 2.3	TCT ACT ATG GTT ACG AC	(2)
Dsp 2.4	TCT ACT ATG GTT ACG AC	(2)
Dsp 2.5	TCT ACT ATG GTA ACT AC	(2)
Dsp 2.6	CCT ACT ATG GTT ACG AC	(2)
Dsp 2.7	CCT ACT ATG GTA ACT AC	(2)
Dsp 2.8	CCT AGT ATG GTA ACT AC	(2)
Dsp 2.9	TCT ATG ATG GTT ACT AC	(3)
Dsp 2.10	TCT ACT ATA GGT ACG AC	(4)
Dsp 2.11	CCT ACT ATA GGT ACG AC	(5)
Dsp 2.X	CCT ACT ATA GTA ACT AC	(1)
DFL 16.1	TTT ATT ACT ACG GTA GTA GCT AC	(2)
DFL 16.2	TTC ATT ACT ACG GCT AC	(2)
DQ52	CAACTGGGAC	(2)

Fig. 6. Collection of DH gene sequences published in the references indicated (1, Gu *et al.*, 1991; 2, Kurosawa and Tonegawa, 1982; 3, Lawler *et al.*, 1987; 4, Feeney, 1990; 5, this paper).

(CAATGTG). A G in this position has also been found in other Ig gene sequences (Kabat *et al.*, 1987).

Reading frame usage of DJH joins

Although most germline DH gene segments analysed have promotors and ATG initiation codons in their 5' sequences (the exception being Dsp2.3) (Reth and Alt, 1984; Gu et al., 1991), the deletions or insertions that accompany DJH joining result in changes in the reading frame context such that only one of the translation products is in frame with respect to $JH-C\mu$. This frame is conveniently referred to as reading frame 2 (RF2) and encodes the $D\mu$ protein. With the DJH reading frame standardized conventionally relative to JH, 28/40 of the joins examined were in RF1, 3/40 in RF2 and 9/40 in RF3 (Table II). RF1 was used 100% of the time in JH1 structures, 82% of the time in the JH4 structures and 43% and 33% of the time in JH2, and JH3 structures respectively; RF2 was only present in DJH2 structures. Interestingly, five out of six of the RF2 DJH structures analysed by Gu et al. (1990) were also in JH2.

The high frequency of potentially productive joins (36/40, 90%), is mainly the result of the overusage of RF1 (which has no termination codons in any DH segment) and the underusage of RF3 (which has termination codons in about one-half of the forward reading frames). Thirty-three structures in RF1 and RF3 could be translated into a polypeptide that would terminate in the C μ 1 exon resulting in a 5–6 kDa translated product. These products, if made, would not be detected by any antibodies that are currently in use because the polypeptide sequences lack the antigenic epitopes found in JH-C μ .

Discussion

Number of DJH rearrangements

In this report we examined DJH gene rearrangements in the fetal livers of two strains of mice during days 12 to 16 of development. The standard assay developed for these studies

	Clone #	DE sequence	Deleted Seg	Jehoes J	JE	DE	f of Del.	RF
	11,13-16	TT TAT TAC TAC GGT AGT AG	CTAC		C TAC TGG TAC TTC GAT	DFL16.1	4	1
	37,43	TT TAT TAC TAC GOT AGT AG	CTAC	CTACTG	G TAC TTC GAT	DEL16 1	10	1,
.1111	33			0		Den2 0		
UHI	34		CTAC	C		DSp2.9		
	44	TC TAT GAT GGT	TACTAC	c		Dap2.9		
				•		- Coperty	Av.= 4.8	-
	- 26		ACTINCTING			DETICI		<u> </u>
	58	TT CAT TAC TAC GGC T	AGIAGCIAC		AC TAC TTT GAC TAC TGG	DFL16.2	2	
	55	TT TAT TAC TAC GGT AGT AGC T	лс АС	ACT	AC THE THE GAC TAC TOG	DFL16.1	5	
	55			home common			10	
UHZ	53	THE ATT ACT AGG GET AC		ACTACITICA		De 510.2	10	
	61		AC	AL.		Dsp2.11	10	2
	32		ACIAC	ACTAC		DSp2.7	10	,
	51		~			Di BIU.I	Av. = 6	
				~	07 Bb0 B00			
	18			COGING		Dap2.2	,	
	20	T TTA TTA CTA CGG	TAGTAGCTAC		Gree TGG TTT GET TAC TGG	DFL16.1	10	3
	21,23	TC TAT GAT GGT TAC		CENGG	TTT GCT TAC TGG	Dsp2.9	5	1
JH3	22	TT TAT TAC TAC GGT AGT AGC T	AC		CC TGG TTT GCT TAC TGG	DFL16.1	2	1
	24	T TTA TTA CTA CGG <u>TAG</u> T	AGCTAC	CCTGGT	TT GCT TAC TGG	DFL16.1	12	3
	25	т ста ста	TOGTAACTAC	œ	TGG TTT GCT TAC TGG	DSP2. ^b	12	3
	63,65	T TTA TTA CTA CGG T	AGTAGCTAC	CCTGGT	TT GCT TAC TGG	DFL16.1	15	3
							Av.= 9.4	
^b This DH	could come from gene segment is	ambiguous and could be DSP2.1, 2.2, 2.3, 2	2.4, 2.5 or 2.10					
							8	$\left - \right $
	2	TT CAT TAC TAC COC TAC	č	ATTACTAT	GCT ATG GAC TAC TOG	DFL16.2	8	
	3	TT TAT TAC TAC GGT AGT AGC TA	с	ATTACTA	T GCT ATG GAC TAC TGG	DFL16.1	8	
	4	TC TAC TAT GAT TA	GCAC	ATTACTA	T GCT ATG GAC TAC TOG	Dap2.2	11	1
	5	C CTA CTA TAG GT	ACGAC	ATTA	C TAT GCT ATG GAC TAC TGG	Dsp2.11	9	3
	6	TT TAT TAC TAC GGT AGT AGC TAC		ATTAC	TAT GCT ATG GAC TAC TGG	DFL16.1	5	1
	7		<u> </u>	2	T THE THE COT ATE CAC THE TOT	DET.16.2	2	
014	,	TT TAT TAC TAC COT ACT ACT TA	c c	2	T TAC TAT OCT ACG GAC TAC TOG	DFL16 1	2	
	9	TT TAT TAC TAC GGT AGT AGC TAC	Ĩ	ATTACTAT	GCT ATG GAC TAC TGG	DFL16.1	8	
	10	TT TAT TAC TAC GGT AGT AGC TAC		ATTAC	TAT GCT ATG GAC TAC TOG	DFL16.1	5	1
	35	TC TAC TAT GGT TAC GAC		ATTACTAT	GCT ATG GAC TAC TGG	Dsp2.3	8	1
	30	T TTA TTA CTA CGG TAG	TAGCTAC	ATTAC	TAT GCT ATG GAC TAC TOG	DFL16.1	12	3
	27	CC TAC TAT AGG TAC G	AC		AT TAC TAT GCT ATG GAC TAC TGG	Dsp2.11	2	1
						•	Av.= 6.8	

Fig. 7. DNA sequences of the DJH structures analysed. The nucleotides that were deleted as a result of the joining process are indicated. In some cases deleted nucleotides could have come from either DH or JH sequences. In frame termination codons are underlined.

Table II. The utilization of DII segments in PCR clones												
ЪН	Number of clones sequenced	DH segments used								RF		
		DFL16.1	DFL16.2	Dsp2.9	Dsp2.2	Dsp2.4 ^a or Dsp2.3	Dsp2.11	Dsp2.7	Dsp2 ^b	I	II	III
јні	11	8	1	2	0	0	0	0	0	11	0	0
JH2	7	3	2	0	0	0	1	1	0	3	3	1
JH3	9	5	0	2	1	0	0	0	1	3	0	6
JH4	13	6	3	0	1	1	2	0	0	11	0	2
Total	40	22	6	4	2	1	3	1	1	28	3	9
RF I		16	5	4	1	1	1	0	0			
RF II		0	1	0	0	0	1	1	0			
RF III		6	0	0	1	0	1	0	1			

^aEither Dsp2.4 or Dsp2.3, sequences are the same.

^bEither Dsp2.1, Dsp2.2, Dsp2.3, Dsp2.4, Dsp2.5 or Dsp2.10.

allowed us to quantify the number of DJH rearrangements, whereas previously it has not been possible to make these estimates. On day 12, there were < 280 DJH rearrangements per liver. By day 16, this frequency had increased to 350 000 per liver in BALB/c and to 730 000 per liver in C57BL/6. This difference in the two strains parallels the differences found in the frequency of B cell precursors (Paige, 1983) and adds to the evidence that the strain differences in B cell

development affect very early differentiation events. The number of DJH structures increased ~1000-fold per liver from days 12 to 16. At the same time the number of cells per liver increased ~30-fold, from ~ $1 \times 10^{6} - 29 \times 10^{6}$.

One important, yet unresolved aspect of lineage commitment deals with the potential for expansion and self renewal of cells at different stages (Rolink and Melchers, 1991). The relative contribution of undifferentiated precusors

(a cell with no Ig gene rearrangements) versus differentiated (a cell with a DJH gene rearrangement) to the resultant B cell pool is not known. Our findings indicate that there must be a large number of cells which are germline on both alleles that feed into the DJH pool during the fetal development because a 1000-fold increase in DJH structures cannot be accounted for by cell division of DJH/germline cells (assuming a maxiumum 8 h cell division time). These results argue that a majority of the precursor cells in days 12 to 15 fetal liver are one or two divisions away from the first Ig gene rearrangement but have not yet undergone the DJH rearrangement and reinforce the evidence for only a limited number of cell divisions during progenitor to B cell differentiation (Melchers *et al.*, 1989; Strasser *et al.*, 1989).

JH gene segment usage

Our analysis determined that the usage of individual JHs ranged from ~20 to 35%. This usage was found for both of the two strains tested, BALB/c and C57BL/6. Other laboratories have reported non-random JH usage in pre-B cells. Gu *et al.* (1990, 1991) analysed PCR amplified cDNAs from neonatal pre-B cells (B220⁺, IgM⁻) by a similar method and found preferential usage of JH1 (57%), although the usage of pre-B cells from the adult was random (27%). In contrast, Feeny (1990) sequenced PCR amplified cDNA products from newborn liver RNA and found less usage of JH1 (7%). These populations, although not the same as our fetal liver samples, are the most comparable pre-B populations that have been reported.

Our demonstration of random JH usage suggests secondary rearrangement must be rare. There is evidence that secondary DH to JH joining can occur after the primary DH to JH joining event (e.g. Maeda et al., 1989). These rearrangements are possible because of the configuration of DH and JH gene segments allows a chromosome that has DHs 5' and JHs 3' of the DJH join to undergo secondary DJH rearrangements utilizing the remaining gene segments (Reth et al., 1986). Primary joins might be expected to contain JH1, but joins containing JH1 would be absent in secondary joins. Thus, in a population with significant numbers of secondary joins, JH1 should be used less often than JH2, JH3 or JH4. Our results which demonstrate no loss of JH1 usage with either the DSF or the DQ52 gene segments argue that secondary recombinations are not common in the fetal liver. This is not to say that secondary rearrangements do not occur, only that they do not seem to be a common component of the DJH population in fetal liver at given time. Because there is an ~ 10 -fold increase in the number of DJH structures each day in our assay, most of the structures present are newly generated. Thus, the secondary rearrangements that have occurred would be masked by the new primary DJH rearrangements. The finding that the majority of the DJH1 structures we sequenced were DFL16.1 is intriguing as these structures cannot undergo further rearrangement. (A joining involving DFL16.1 results in the deletion of all other known DH, which precludes secondary DJH joining as presently understood on this chromosome).

DH gene segment usage

A single gene segment, DFL16.1, was used in more than half of the DJH structures we sequenced. Since much of CDR3 comes from the D gene segment, this degree of restriction imposed on the diversity of the Ig repertoire is surprising. Seven of the DFL16.1 JH1 structures were identical. Since these structures were from the same amplification as the other 33 DJH clones sequenced, we believe that they reflect the frequency of this product in the starting fetal liver population. Gu *et al.* (1990, 1991) also found this same DJH1 join in half of the products they sequenced and the identical structure was similarly isolated by Feeny (1990). Moreover, experiments using both cell lines and normal tissues have, in general, shown overusage of DFL16.1 and DQ52 (Lawler *et al.*, 1987; Suzuki *et al.*, 1989; Feeny, 1990; Gu *et al.*, 1990, 1991; Tsukada *et al.*, 1990; Rolink *et al.*, 1991).

The reasons for this preferential usage are probably complex. It has been suggested that the reason for the biased usage of DQ52 is the proximity of it to the JH locus (700 bp). The over usage of DFL16.1, the most 5' DH gene thus far identified, is difficult to reconcile with this hypothesis, suggesting that mechanisms other than proximity are significant. It has also been postulated that differences in RSS may contribute to the high usage of DQ52 and DFL16.1 (Feeny, 1990) since subtle differences in RSS may result in a different frequency of recombination (Hesse et al., 1990; Ramsden and Wu, 1992). This is unlikely, as DQ52 and DFL16 RSS differ from the commonly accepted consensus RSS by one nucleotide each, while other DH gene segments have a perfect consensus sequence and are used less frequently. A third possibility is that a particular configuration of chromosomal DNA promotes recombination. If JH is 'open', it is reasonable to believe that the nearby DQ52 gene segment also may be open and thus active in recombination. This mechanism would also require that DFL16.1 be in a similar open configuration.

Yet another possibility has been suggested by the observation that there are short stretches of homology between the 3' end of the DFL16.1 coding region and the 5' end of the JHs (JH1 in particular). However, upon closer examination of the DFL16.1 usage and the 5' JH homology, one finds little correlation between the presence of homology and usage. JH1 (5'-CTAC) uses DFL16.1, 8/11 (73%); JH2 (5'-ACTAC) 3/7 times (43%); JH3, (5'-TGCCT) 5/9 (56%); and JH4 (5'-TTAC), 6/13 (46%). Thus DFL16.1 joins to the JH with no homology (JH3) at about the same frequency as the other JHs. More problematic is the observation that DH gene segments DSP2.1, DSP2.5, DSP2.7, DSP2.8, DSP2.9, DSP2.X and DFL16.2 also have CTAC at their 3' border. These segments are used 0%, 0%, 2.5%, 0%, 10%, 0% and 15% respectively, in our study. It is also notable that DQ52 does not have this sequence homology yet it is used in 10% of all DJH structures. Thus, it is unlikely that short stretches of sequence homology play a role in DH gene usage although it may (as discussed below) influence RF usage.

Thus, none of these mechansims appear to explain fully the preferential usage of DFL16.1 and DQ52. It has been suggested that D regions may have a function other than contributing to diversity (Cohn, 1990). The non-random usage of DFL16.1, the numerous deletions and general lack of N additions in the fetus is consistant with a function for D other than solely as a source of diversity.

Deletions in DJH joins

In general, site specific recombination systems require exact joining of the DNA that participates in the reaction (Sadowski, 1986; Craig, 1988). This is generally not the

case in Ig gene recombination. Although joins of the recombination signal sequences (signal joins) seem to have little variation (Lewis *et al.*, 1988; Lieber *et al.*, 1988), the coding joins invariably have insertions and/or deletions of nucleotides (Tonegawa, 1982; Hesse *et al.*, 1987; Gu *et al.*, 1991). These variations at gene segment borders can add diversity to CDR3 and may be particularly important in the early Ig repertoire in light of the restricted DH usage of fetus. In the work presented here we found deletions in all DJ joins. The number of base pairs deleted ranged from two to 15, with a median of seven. It is notable that the most deletions were found in DJH3 structures where there are no short stretches of sequence homology.

Reading frame usage

Unlike other genes, most D gene segments can be translated in three reading frames and in both directions. The reading frame of D in a DJH gene segment is fixed, in that it is dependent on an ATG ~ 60-100 nt 5' of DHs (Reth and Alt, 1984). This RF is conventionally referred to as RF2 and is the reading frame that would produce the 'D μ ' polypeptide. In this analysis only DJH joins in the 'forward' direction were examined. Given the potential for joining in most frames, it is surprising that RF1 is used so frequently (70% in this study). Feeney (1990) found a similar biased usage of RF1 in both the newborn (84/98, 86%) and the adult RNA (158/215, 73%). Gu *et al.* (1990, 1991) has also found more prevalent use of RF-1 in adult B cells (24/34, 70%) than in pre-B cells (7/15, 47%).

Rajewski and his colleagues have presented compelling evidence that RF2, the frame that encodes $D\mu$ protein, may be selected against (Gu et al., 1991). The paucity of structures in RF3 is explained by the presence of termination codons in half of the reading frames. Although this would explain their rarity in VDJ joins, this explanation cannot be used to explain their underuse (9/40) in DJH structures in our study (where expression of DJH is not required). Instead, our results support the notion that RF1 usage is promoted by the short stretches of sequence homology at the 3' border of DH and the 5' border of JH (Gu et al., 1991). In this hypothesis, when a DJH join occurs, one of the pair of homologous sequences is deleted and this results (most frequently) in a join in RF1. Of the joins analysed here, most (20/24) with such sequence homologies are in RF1, whereas those with no sequence homologies are in RF1 and RF3 about equally. An example of joins with sequence homologies are the DJH1 joins, (the sequence homology is CTAC). 100% of the joins are in RF1. The converse may be seen by inspecting the DJH3 structures where there are no short sequence homologies. Three joins are in RF1; six are in RF3.

The DJH transcripts could encode 5-6 kDa polypeptides when the joins are in RF1 and RF3. These potential translation products, would not have the ability to be expressed on the cell surface (as D μ can be) nor would they have the common C μ sequences (found in D μ products). At first inspection, their heterogeneity might dissuade one from thinking that such polypeptides would have a function. However, the biased usage of DFL16.1, especially with JH1, reduces the heterogeneity to a level where a function for DJH structures in RF1 and RF3 might be considered.

Materials and methods

Mice and cell lines

C57BL/6 and BALB/c mice were purchased from Jackson Laboratories (Bar Harbor, ME) and maintained at the animal colony of the Ontario Cancer Institute. Timed pregnancies were established as previously described, with day 0 of gestation counted as the day of mating. Livers were removed from fetuses at days 12 to 16 of gestation (Paige *et al.*, 1984). Six to eight fetal livers from one mother were pooled.

DNA preparation

Single cell suspensions were prepared from fetal liver using standard procedures (Paige *et al.*, 1984). Genomic DNA was isolated from fetal liver cells or cultured cells as described previously using the 'spooling' method to recover precipated DNA (Atkinson *et al.*, 1991)

Oligonucleotide primers

The DSF primer is 5'-AGGGATCCTTGTGAAGGGATCTACTACT-GTG-3'. It is a 31mer extending from the 5' end of the nonamer of DSP/DFL through the spacer to the 3' end of the heptamer and contains no DH coding sequences. It is specific for DFL and DSP gene segments, differing from the published sequences of each member of the two families in three positions (including an embedded *Bam*H1 site).

The DQ52 primer is 5'-GCGGAGCACCACAGTGCAACTGGGAC-3'. It is a 26mer, specific for DQ52, extending from the within the 5' spacer through the heptamer and contains all of the coding sequence of DQ52. It is a unique sequence in GenBank.

The JH4 primer is 5'-AAAGACCTGCAGAGGCCATTCTTACC-3'. It is a 26mer containing sequences in the J-C intron immediately 3' of JH4 exon excepting that the ninth nucleotide was changed from a C to a G to obtain a *Pst*I site. It is a unique sequence in GenBank.

The oligonucleotides were synthesized on a DNA synthesizer (Applied Biosystems) and purified by using NENSORB PREP cartridges (Du Pont).

Standard PCR assay

PCR reactions were performed in siliconized 500 ml Eppendorf tubes in a volume of 100 ml in 10 mM Tris-Cl (pH 8.3 at 25°C), 50 mM KCl, 1.5 mM MgCl₂, 0.01% (w/v) gelatin and containing 1µg DNA, 200 µM of each dNTP, 0.5 µM of each oligonucleotide primer and 2.5 units Taq polymerase (Perkin-Elmer Cetus). The reaction was overlaid with oil. A Perkin-Elmer Cetus DNA thermalcycler was used. Each cycle consisted of 1 min denaturation at 94°C, 1.5 min annealing at 60°C and 2 min polymerization at 72°C. The cycle was repeated 30 times. The polymerization step was extended an additional 3 s in each cycle. The final polymerization step was extended an additional 10 min.

In standardizing the assay we found that if the PCR were allowed to continue for 40 or 60 cycles, more products were detectable when using 10 standard targets. However, with these conditions the assay was no longer linear for higher numbers of targets and thus not useful for experimental samples. The upper limit was determined to be 5000 haploid genomes. Thus, dilutions of experimental DNA samples in CB32.12 DNA had to be made before amplification in order to determine the number of rearrangements when there were more than 5000 targets.

A standard assay for DJH1 to DJH4 rearrangements was established as follows: from a survey of 40 A-MuLV lines which by Southern analysis were VDJ on one allele and DJ on the other (Aktinson *et al.*, 1991), four cell lines with four different JH (JH1–JH4) rearrangements in DJH chromosome were identified with the DSF/JH4 primer pair. They are: CB178, which has a DJH1 rearrangement; CB101, which has a DJH2 rearrangement; CB173, which has a DJH3 rearrangement; and CB165, which has a DJH4 rearrangement. Each cell line used a DH gene segment from either the DSP2 or DFL16 gene families. In a similar fashion, 55 VDJ/DJ A-MuLV cell lines were screened with the DQ52/JH4 primer pair and six lines with DQ52JH rearrangements were identified. One line with JH1; four with JH2; zero with JH3; and one with JH4.

The absence of a line containing DQ52JH3 rearrangements prevented us from setting up a complete standard assay for DQ52JH rearrangements. Thus, in order to quantify DQ52 rearrangements, we used a cell line, 70Z/3, which has a DQ52JH2 rearrangement on one allele and compared the amplification of this DQ52JH2 product to that of the amplification of the DSFJH2 product from CB101. We found that the DQ52JH2 product was five times the amount of the DSFJH2 product, a result which parallels the observations of others, with different primer pairs and is not unexpected Analyses of DQ52JH rearrangements were complicated by the proximity of DQ52 to JH1 (700 bp) which amplified a germline unrearranged target. Since the vast majority of the DNA in fetal liver is in the germline configuration, the amplified germline band overshadowed the rearranged targets and rendered estimates useless. In order to remove the germline target we took advantage of restriction sites for *HincII* and *SsI* which are found between DQ52 and JH1 but not between JH1 and JH4. Exhaustive double digestions of DNA with *HincII* and *SsI* reduces any amplifiable germline DNA to minimal levels. The remaining undigested DNA is still amplified with the DQ52/JH4 primer pair but at a level that is not problematic (Figure 5). Nevertheless, these complications with DQ52 usage. Estimates of DQ52 usage shown in Table I reflect this reduced accuracy.

To verify that the DSF primer does not amplify the DFL16 family preferentially, genomic DNA from A-MuLV cell lines with known DJH rearrangements was amplified. Six cell lines were used: CB4 (DFL16.1 JH2 rearrangement); CB82 (DSPJH3) rearrangement); CB43, (DSPJH2); CB134 (DFL16.1JH3); CB172 (DSPJH3); and CB135 (DFL16.1JH2). The DNAs from these lines were mixed in pairs at 1:1 ratios in the amounts of 500 ng or 100 ng each per reaction. The degree of amplification of each DJH rearrangement was independent of the DH segment used, verifying that the DSF primer does not amplify DFL16.1 preferentially (data not shown).

Southern hybridization analysis

One-tenth of each PCR amplification reaction was loaded on a 1.5% agarose gel (Sigma) and electrophoresed in TAE buffer (Maniatis *et al.*, 1982). Gels were Southern blotted on Zeta Probe nylon membrane (BioRad) using a vacuum blotting unit (Vacugene, LKB Pharmacia). The filters were prehybridized for 4 h in 5× SSC, 0.5% SDS, 1× Denhardt's and 200 µg/ml sheared Salmon sperm DNA (Sigma) at 65°C. Filters were hybridized in fresh solution containing, in addition, ³²P-labelled probe at 1×10⁶ c.p.m./cm². Hybridization was carried out at 65°C for 16 h. The probe used for hybridization was a 1.2 kb genomic *Hind*III–*Eco*RI fragment containing the JH4 gene segment (from pGW78). Filters were washed in 2×SSC, 0.1% SDS four times at 42°C, each time for 30 min. These conditions were chosen because reduced washing temperatures allowed detection of DJH4 rearrangements that have only 50 nt of homology to the JH4 probe. Autoradiography was done without intensifying screens.

Densitometry

Densitometry was used to measure relative intensities of autoradiographic bands for each set of PCR amplifications. Several exposures of Southern blots were scanned to obtain the data in Table I. For each exposure the densitometry values of the standards were plotted and the values for the experimental bands determined from the plots. Table I contains the data from the average of three separate experiments determined in this way. To determine the significance of the differences we found between the number of DJH structures in C57BL/6 and BALB/c fetal liver, we amplified dilutions of fetal liver DNA from the two strains over a 100- to 800-fold range and calculated the C57BL/6: BALB/c ratios over these ranges. The densitometry ratio of C57BL/6:BALB/c was 2.26. When data from all experiments were pooled, the ratio was 2.03.

Cloning and sequencing

The amplified products were purified from Nusieve agarose (FMC Bioproducts) gels and cloned into pBlueScribe either by blunt ligation or by utilizing the *Bam*HI and *PstI* enzyme sites contained in the primers. Sequencing was performed using the double stranded method with the T7 Sequencing kit (Pharmacia) and the reverse and universal sequencing primers.

Acknowledgements

We thank Michael Atkinson, Jakub Celler, Ana Cumano, Dale Ramsden and Charley Steinberg for helpful discussion. We thank Lynne Omoto for excellent preparation of the manuscript. This work was supported by grants to C.J.P. and G.E.W. from the Medical Research Council of Canada and the National Cancer Institute of Canada. G.E.W. is an MRC scholar.

References

- Alt,F.W., Blackwell,T.K., DePinho,R.A., Reth,M.G. and Yancopoulos,G. (1986) *Immunol. Rev.*, **89**, 5.
- Atkinson, M.J., Michnick, D.A., Paige, C.J. and Wu, G.E. (1991) J. Immunol., 146, 2805-2812.
- Cohn, M. and Langman, R.E. (1990) Immunol. Rev., 115, 7-147.
- Craig, N.J. (1988) Annu. Rev. Genet., 22, 77-105.
- Feeny, A.J. (1990) J. Exp. Med., 172, 1377-1390.
- Gu,H., Förster,I. and Rajewsky,K. (1990) EMBO J., 9, 2133-2140.
- Gu,H., Kitamura,D. and Rajewsky,K. (1991) Cell, 65, 47-54.
- Hesse, J.E., Lieber, M.R., Gellert, M. and Mizuuchi, K. (1987) Cell, 49, 775-783.
- Hesse, J.E., Lieber, M.R., Mizuuchi, K. and Gellert, M. (1989) *Genes Dev.*, **3**, 1053-1061.
- Ichihara, Y., Hayashida, H., Miyazawa, S. and Kurosawa, Y. (1989) Eur. J. Immunol., 19, 1849-1854.
- Kabat, E.A., Wu, T.T., Reid-Miller, M., Perry, H.M. and Gottesman, K.S. (1987) *Sequences of Proteins of Immunological Interest*, Vol. 1, 5th ed. US Department of Health and Human Services, pp. 339–494 and 1350–1441.
- Kincade, P.W. (1987) Adv. Immunol., 41, 181-267.
- Kurosawa, I. and Tonegawa, S. (1982) J. Exp. Med., 155, 201-218.
- Lawler, A.M., Lin, P.S. and Gerhart, P.J. (1987) Proc. Natl. Acad. Sci. USA, 84, 2454-2459.
- Lewis,S.M., Hesse,J.E., Mizuuchi,K. and Gellert,M. (1988) Cell, 55, 1099-1107.
- Lieber, M.R., Hesse, J.E., Lewis, S., Bosma, G.C., Rosenberg, N., Mizuuchi, K., Bosma, M.J. and Gellert, M. (1988) Cell, 55, 7-16.
- Maeda, T., Sugiyama, H., Tani, Y. and Kishimoto, S. (1989) J. Immunol., 142, 3652-3656.
- Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning, A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Meek,K. (1990) Science, 250, 820-823.
- Meek,K.D., Hasemann,C.A. and Capra,J.D. (1989) J. Exp. Med., 170, 39-57.
- Melchers, F. (1979) INSERM Symp., 10, 281-289.
- Melchers, F., Strasser, A., Bauer, S.R., Kudo, A., Thalmann, P. and Rolink, A. (1989) Cold Spring Harbor Symp. Quant. Biol., 54, 183-189.
- Misener, V., Jongstra-Bilen, J., Young, A.J., Atkinson, M.J., Wu, G.E. and Jongstra, J. (1990) J. Immunol., 145, 905-909.
- Misener, V., Downey, G.P. and Jongstra, J. (1991) Int. Immunol., 3, 1129-1136.
- Paige, C.J. (1983) Nature, 302, 711.
- Paige, C.J., Gisler, R.H., McKearn, J.P. and Iscove, N.N. (1984) Eur. J. Immunol., 14, 979-987.
- Ramsden, D.A. and Wu, G.E. (1991) Proc. Natl. Acad. Sci. USA, 88, 10721-10725.
- Reth, M.G. and Alt, F.W. (1984) Nature, 312, 418-423.
- Reth, M.G., Jackson, S. and Alt, F.W. (1986) EMBO J., 5, 2131-2138.
- Rolink, A. and Melchers, F. (1991) Cell, 66, 1081-1094.
- Rolink, A.G., Kudo, A., Karasuyama, H., Kikuchi, Y. and Melchers, F. (1991) *EMBO J.*, **10**, 327-336.
- Sadowski, P. (1986) J. Bacteriol., 165, 341-347.
- Sakano, H., Maki, R., Kurosawa, Y., Roeder, W. and Tonegawa, S. (1980) *Nature*, **286**, 676-683.
- Strasser, A., Rolink, A. and Melchers, F. (1989) J. Exp. Med., 170, 1973-1986.
- Suzuki, H., Abe, M., Nishikawa, S.-i., Nakayama, E. and Shiku, H. (1989) Int. Immunol., 1, 643-646.
- Tonegawa, S. (1983) Nature, 302, 575.
- Tsubata, T., Tsubata, R. and Reth, M. (1991) Eur. J. Immunol., 21, 1359-1363.
- Tsukada, S., Sugiyama, H., Oka, Y. and Kishimoto, S. (1990) J. Immunol., 144, 4053-4059.
- Wood, C. and Tonegawa, S. (1983) Proc. Natl. Acad. Sci. USA, 80, 3030-3034.
- Wu,G.E., Atkinson,M.J., Ramsden,D.A. and Paige,C.J. (1990) Semin. Immunol., 2, 207-216.

Received on November 4, 1991; revised on February 3, 1992