

Restricted Immunoglobulin Junctional Diversity in Neonatal B Cells Results from Developmental Selection Rather Than Homology-based V(D)J Joining

By Akhilesh Pandey,[‡] Larry W. Tjoelker,[‡]
and Craig B. Thompson^{*‡}

From the Departments of ^{*}Internal Medicine and [‡]Microbiology and Immunology, Howard Hughes Medical Institute, University of Michigan, Ann Arbor, Michigan 48109-0650

Summary

The mechanism by which coding ends are joined during immunoglobulin (Ig) recombination is poorly understood. Recently, short sequence similarities (2–6 bp) observed at the ends of certain variable (V), diversity (D), and joining (J) gene segments of Ig have been correlated with limited junctional diversity observed in coding exons assembled from these elements. However, it is unclear whether these sequence homologies play any direct role in favoring coding joint formation by influencing the V(D)J recombination process. In this report, we demonstrate that coding sequence similarities do not influence the position of coding joints during V(D)J recombination *in vivo*. Instead, during embryonic development, B cells with certain joining products undergo progressive selection. Developmental selection is completed before exposure to external antigens and appears to be determined by the amino acid sequence encoded by the coding joint. We conclude that the nucleotide sequences of the coding regions do not play a major role in directing V(D)J recombination. Instead, we propose that limited Ig junctional diversity results from prenatal developmental selection of B cells based on the protein sequence of their surface Ig antigen-binding site. Sequence identities at the ends of coding segments may have evolved because they increase the likelihood that a selectable antigen-binding site is created during a random recombination process.

The need to respond to a wide variety of foreign antigens requires that the vertebrate immune system produce a large repertoire of Ig and TCR. Production of a diverse set of Ig and TCR is facilitated by the fact that the genes that encode each of these receptors are composed of families of discrete germline elements (1, 2). During development, individual members of variable (V), diversity (D), and joining (J) families in the B and T cells must recombine to create a translatable antigen-binding exon for each chain of their individual antigen receptor. The developmental process by which rearrangement occurs has been termed V(D)J recombination. V(D)J recombination is a form of site-specific recombination whose specificity is determined by recombination signal sequences (RSS)¹ consisting of a palindromic heptamer and an AT-rich nonamer element that flank the coding sequences of the individual V, D, and J elements (1). During V(D)J recombination, two RSS elements are paired so that one of the pair has a 12-bp spacer between the heptamer and nonamer elements, whereas the other has a 23-bp spacer between the heptamer and nonamer elements. Two recombi-

nation products are produced during a V(D)J recombination event (Fig. 1). One product is the precise end-to-end joining of the two recombination signal sequences. The other recombination product results from the joining of the coding sequences of the two Ig gene segments themselves (3). It takes two serial recombination events to generate an Ig H chain or a functional TCR β or δ chain. First, recombination occurs between a D and a J element followed by the joining of a V element to the previously recombined DJ element (4–6). For Ig L chain and TCR α and γ chains, a functional gene product is created directly by the recombination of individual V and J gene elements (7, 8).

The recombination of the individual coding elements of Ig and TCR genes creates diversity by several mechanisms. First, in many species there are families of V, D, and J gene elements encoded in the germline (1). The random recombination of individual elements results in considerable combinatorial diversity. Second, unlike the signal joints that are usually created by precise joining of the heptamer elements of the two recombination signal sequences, joining of individual V, D, and J elements is imprecise (4, 9, 10). Coding ends are subject to both base loss and addition before ligation resulting in diversity of the sequences at the joining po-

¹ Abbreviation used in this paper: RSS, recombination signal sequences.

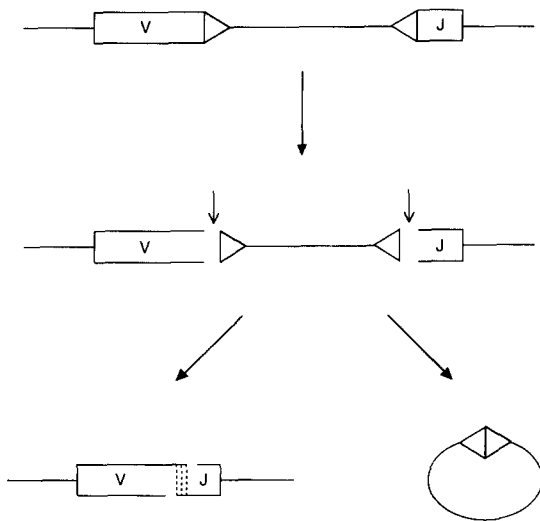


Figure 1. Schematic representation of V(D)J joining process. (Open boxes) V_i and J_i coding segments, (triangles) recombination signal sequences. Endonucleolytic cleavage followed by ligation leads to the formation of a coding joint and a circular episome containing a signal joint. Formation of the signal joint occurs by precise end-to-end ligation of the signal ends. In contrast, as described in the text, the position of coding joint ligation is highly variable and has been suggested to occur by pairing of exposed single-stranded regions followed by filling in and ligation.

sitions of the individual elements (3). Finally, recombination appears to activate Ig gene transcription, and this may play a role in the generation of further diversity in some species by rendering the locus susceptible to gene conversion and/or somatic mutation (11–13).

At present, relatively little is known about how coding joints are brought together during V(D)J recombination (for review see reference 14). As noted above, there is a considerable variability in the precise position of joining. In most coding joints there is nucleotide loss from both coding ends before end ligation. Coding ends are also subjected to two forms of nucleotide addition. P nucleotides are nucleotides added only to full-length coding sequences and are complementary to the sequence at the ends of coding elements (15, 16). In contrast, N nucleotides can be added to intact or processed ends, and although enriched for G/C bases, are otherwise random in sequence composition (for review see reference 17). The average number of N nucleotides added during V(D)J recombination increases during development (18–21). Since coding joints involve the joining of two coding sequences in a random fashion, they are frequently out-of-frame. In fact, two thirds of joints formed randomly would be expected to be out-of-frame. This high percentage of out-of-frame joints may be tolerated simply because of the high degree of diversity created through the formation of coding joints.

Despite the great variability in the generation of coding joints, several examples of limited coding joint diversity have been observed (15, 22–25). In these cases, it has been postulated that short stretches of sequence identity increase the likelihood of recombination occurring at a given position (24, 25). This has led to the proposal that the presence of com-

plementary sequences on exposed single-stranded tails of the two coding elements are used to facilitate coding end alignment and joining (14, 24, 25). However, it seems paradoxical that a recombination mechanism that appears to have evolved to maximize heterogeneity would involve an intermediate that limits its diversity. Therefore, an alternative hypothesis for the observed limitation in junctional diversity is that there is developmental selection for specific coding joint sequences.

Recently, during our studies of B cell development in the Muscovy duck, we encountered an example of limited junctional diversity. 53 of 64 rearranged V_{L1} - J_L gene products isolated from posthatch ducks had identical sequences at the site of VJ joining. The ends of V_{L1} and J_L contained a 9-bp region of sequence similarity in which 7 out of 9 bp were identical. Alignment of these two segments of homology could result in the creation of the preferential joint observed in B cells of adult ducks. Unlike mammals, recombination in avian B cells occurs only during a limited period of development (16). Therefore, we were able to design experiments to address whether this preferential joint was the result of the V(D)J recombination process or of developmental selection that occurred subsequent to V(D)J recombination. In the Muscovy duck, Ig L chain recombination occurs only in lymphoid progenitor cells arising in the spleen between days 20 and 30 of embryonic development. Upon sequencing the initial joints created in the spleen on day 21 of embryonic life, we found that there was no bias for the creation of V_{L1} - J_L joints as a result of the sequence homology at the V_{L1} and J_L coding ends. Instead, we were able to document that at sequential times during embryonic development, there was a progressive increase in the proportion of B cells that had Ig L chains containing the V_{L1} - J_L joint observed in adult ducks. In fact, this joint was observed in over 94% of V_{L1} - J_L joints observed in B cells from the spleen and bursa at day 35 of embryogenesis (day of hatch). This developmental selection was not limited to the V_{L1} - J_L product. In the other functional V-region gene (V_{L5}) that we studied, two V_{L5} - J_L recombination joints were observed to increase in frequency during development, both of which encoded the same amino acid product across the V_{L5} - J_L joint. Based on these results, we conclude that sequence homology does not play a major role in promoting the formation of the coding joint during V(D)J recombination. Instead, developmental selection is the reason why cells with a particular joint predominate at birth. It appears that selection occurs as result of the protein product of V(D)J recombination rather than as a result of the process of recombination itself. This selection occurs before exposure of the embryo to exogenous antigen and therefore must be mediated by intrinsic developmental factors and/or antigens.

Materials and Methods

Cell Isolation and DNA Extraction. Muscovy ducks were killed at days 20, 21, 25, 26, and 30 of embryogenesis, at the day of hatch (day 35 of embryogenesis), and at 1, 2, and 4 wk of age. At each of these times, the bursa and spleen were removed and the lymphoid cells isolated as described previously (26). Cells were lysed

in RSB buffer (10 mM Tris, pH 7.4; 10 mM NaCl; and 5 mM MgCl₂) containing 0.5% NP-40, and nuclei were isolated. The nuclei were then solubilized in 0.5% SDS, 300 mM NaCl, 10 mM Tris, pH 7.4; and 5 mM EDTA. Proteinase K (250 mg/ml) was used to digest proteins and the DNA was precipitated in 2.5 M ammonium acetate and 50% (vol/vol) isopropanol. It was resuspended in TE (10 mM Tris, pH 7.4; and 1 mM EDTA) and then reprecipitated as above. The resulting precipitate was then recovered by centrifugation, washed twice with 70% ethanol containing 50 mM NaCl, and then resuspended to ~1 mg/ml.

PCR and Southern Blot Analysis. PCR (27) was used to clone the coding joints and to amplify the signal joints created during V(D)J recombination as described previously (16). Oligonucleotide primers were synthesized on a DNA synthesizer (model 380B; Applied Biosystems, Foster City, CA). The primers for V_{L1} and J_L segments had restriction sites for EcoRI or HindIII to allow directional cloning into pGEM-3Z.

35 cycles of amplification (1 min at 94°C, 2 min at 72°C) were performed in a thermal cycler (Perkin-Elmer Cetus Instruments, Norwalk, CT) followed by a final 10 min. extension. The PCR primers used for signal joint assays were: no. 2081-GGGGGAATTCCTCCATCCCACCACTTTCCATCCTGGTGCA; no. 3544-GGGGGAATTCGGACAGCAGGCATGGCCACCTGCCGCACTG; no. 5299-TTGGAAGCTTGACAGGCACAGCCACTTG-CAGCACAGAGCTG; and no. 7009-TGGAAGCTTGCCCCAAGGGCATTGCTGAGTCA.

The PCR primers used to amplify rearranged IgL clones were chosen so that the rearranged joints amplified were not subject to bias by subsequent gene conversion within the primer sequence. For this purpose, the primers for various V_L gene segments were located in FR1 region or 5' of it. The primers were: for V_{L1}: GGGGGAATTCCTGTGGGAGCTGTGAGCGCTGCGTAGTACTC and GCGCGAATTCCTGGAGCCGAAGTCGGTGAATCCGGGAGAGC; for ϕ V₃: GGGGGAATTCACATTTCTCCTCCAGTGTCTCAGCAGGGCTG; for V_{L5}: GGGGGAATTCCTGGAGCCGCAGTCAGTGAACCTGGGAGGA; for J_L: TTGGAAGCTTAGGCCATGAAAGACCGAGGCGAGGTCAGCG and TTGGAAGCTTGCGACAGACCCAGCCACCGCCAAGTCCG. The PCR products were digested with EcoRI and HindIII and cloned into pGEM-3Z.

Southern blot analysis was carried out as described previously (12). The signal joints were confirmed by hybridizing with a radiolabeled J_L RSS-specific fragment.

DNA Sequencing. Dideoxynucleotide sequencing reactions were performed using double-stranded DNA template with a Sequenase kit according to the supplier's protocol (U.S. Biochemical Corp., Cleveland, OH), and a primer specific for the SP6 promoter site of pGEM-3Z.

Results

Bias for One V_{L1}-J_L Joint in Muscovy Ducks. Muscovy ducks have a single functional J_L and at least five functional and 32 pseudogene V_L segments in the IgL locus (28; Tjoelker, L. J., L. M. Carlson, and C. B. Thompson, manuscript in preparation). To activate IgL gene expression, a functional V_L gene segment must undergo rearrangement with a J_L segment. When the two coding segments are joined in-frame, subsequent transcription and processing lead to a translatable IgL transcript. Using PCR primers specific for V_{L1} and J_L, we sequenced the productive V_{L1}-J_L joints from 1-4-wk-old Muscovy ducks. It is surprising that one particular joint was found in 53 out of 64 (83%) total clones and 53 out of 58 (91%) clones with in-frame V_{L1}-J_L joints (Fig. 2). This predominant joint has the nucleotides GAT TTT GGA at the V_{L1}-J_L junction. Since the codon TTT is present at both of the coding ends, the Ts could belong to either the V_L or the J_L end depending on the crossover point. Hence, theoretically, the frequency of this joint could be the summation of the frequencies of four different joints. Even in that case, each of these four joints is present at a high frequency (13/64) indicating a bias for ligation occurring in this region. On examining the ends of V_{L1} and J_L germline sequences, we found that there was a 9-bp-long region in which 7 bp were identical in both the coding ends (Fig. 3 A). Base pairing of the complementary strands of the V_{L1} and J_L elements could thus align these two elements during V(D)J recombination. As shown in Fig. 3 B, this could account for the preferential joint that we observed.

V(D)J Recombination Occurs in the Spleen during a Brief Developmental Period. The germline configuration of the 3' end of the IgL locus in the Muscovy duck is shown in Fig. 4 (28). Since one of the products of V(D)J recombination is

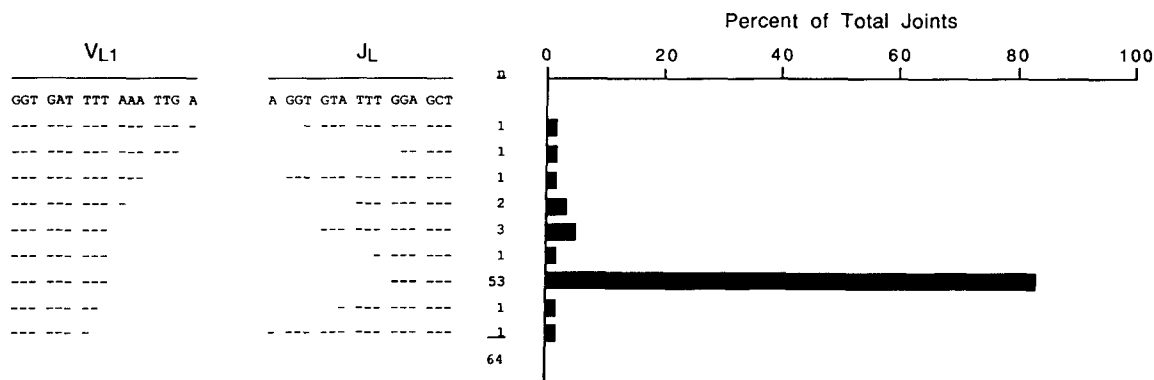


Figure 2. Nucleotide sequences of V_{L1}-J_L joints from adult Muscovy ducks. The nucleotide sequences of the V_{L1}-J_L joints of rearranged genes are compared with the germline sequences obtained from the Muscovy duck. (Dashes) Identity to the germline. Rearranged IgL genes were cloned from splenic and bursal tissues of 1-4-wk-old Muscovy ducks after PCR amplification. (n) Number of clones isolated with the particular V_{L1}-J_L joint. (Right) Histogram showing the frequency of joints as a percentage of total joints cloned.

A

```

      VL1                                JL
      *** *                               *** *
GAC AAC AGT GGT GAT TTT AAA TTG A   A GGT GTA TTT GGA GCT GGG ACC ATG
-----

```

B

```

      VL1                                JL
5'   GAC AAC AGT GGT GAT TTT           GCT GGG   3'
      ||| | |||
3'   CTG TTG           CCA CAT AAA CCT CGA CCC   5'

```

Figure 3. Sequence of the preferred joint in adult Muscovy ducks. (A) Germline sequence of V_{L1} and J_L genes in the Muscovy duck IgL locus is shown. (Dashes) Identity to the germline sequence. (*) Nucleotides common to both V_{L1} and J_L. (B) Possible alignment of complementary sequences in V_{L1} and J_L during the process of recombination.

a circular episome containing the signal joint, we examined lymphoid tissues at various developmental times for ongoing VJ recombination by assaying for signal joints. In this assay, primers that mediate divergent DNA synthesis in the germline, would, upon formation of the circular episome, prime convergent synthesis across the signal joint (16).

Primers specific for the flanking DNA of various V_L gene segments and J_L were designed (Fig. 4). These primers were used in PCR on DNA samples from the bursa and spleen of 26- and 30-d embryos (Fig. 5 a). The PCR products were stained with ethidium bromide and visualized (Fig. 5 a, top). To confirm that these bands indeed contained the episome, they were subjected to electrophoresis, transferred to a Southern blot, and hybridized with a J_L RSS-specific oligonucleotide probe (Fig. 5 a, bottom). Signal joints were detected at days 26 and 30 in the embryonic spleen. No signal joints were detected in the bursa at either developmental time suggesting that recombination does not occur in the bursa. As a control for the quantitation of DNA and the PCR reaction, we also performed PCR on the same DNA aliquots using primers from the 5' and 3' flanks of each V_L gene segment. No reproducible differences between DNA samples were observed in control PCR from either tissue or any time point.

To estimate the time of peak recombination in the spleen, the same method of analysis was repeated on DNA samples from 20-, 25-, and 30-d embryonic spleen and 2-wk posthatch spleen (Fig. 5 b). In these results, which are representative of five separate experiments, we observed that recovery of signal joints by PCR was greatest at day 25. Lower levels of signal joint recovery were observed at days 20 and 30. No recovery of signal joints was observed at hatch (data not shown) or subsequent time periods. Taken together, these data suggest that recombination occurs during a brief developmental period in the spleen of the Muscovy duck. Based on our analysis, we conclude that recombination begins by day 20 and peaks around day 25 of embryonic development in the Muscovy duck.

V_{L1}-J_L Joints at Day 21 in the Spleen. We wanted to look at the V_{L1}-J_L joints before a time at which selection could have occurred. Thus we sequenced V_{L1}-J_L rearrangements from 21-d spleen. At this time, no gene conversion had occurred and hence it was impossible to predict if two identical clones out of one PCR reaction were independent sequences or amplified products of one DNA sequence. This problem was circumvented by performing seven separate PCR reactions on 21-d spleen DNA samples. We found that 24 out

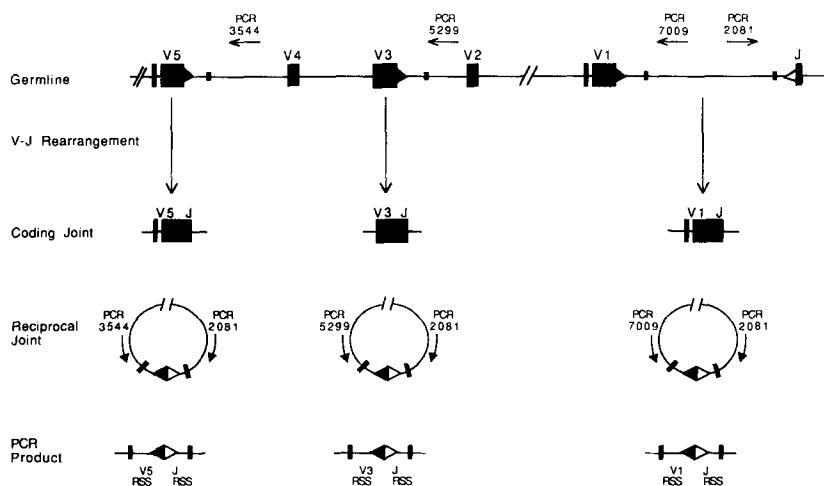


Figure 4. Detection of V(D)J recombination in the spleen. Illustration of the strategy used to detect signal joints by PCR amplification. A schematic representation of the 3' end of the V_L region cluster and J_L region of the Muscovy duck IgL locus. (■) V_L and J_L coding segments. (▶) V region recombination signal sequences (RSS). (▷) J_L region RS. (Arrows) Orientation of the PCR primers used. V2, V3, and V4 are pseudogenes.

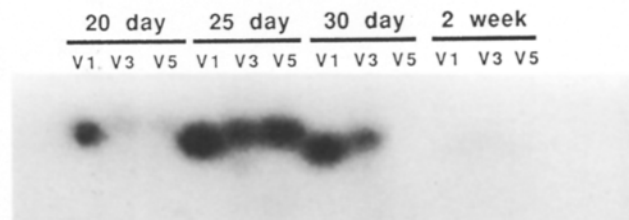
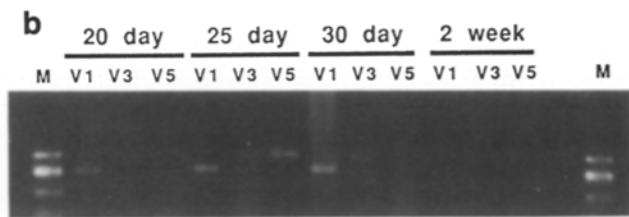
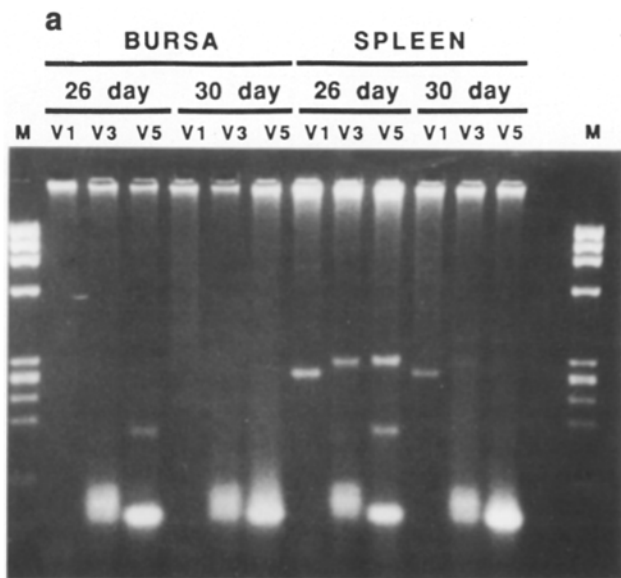


Figure 5. Presence of signal joints during development. (A) Genomic DNA from 26- and 30-d embryonic bursa and spleen was assayed for the presence of signal joints formed by the site-specific joining of V_{L1} , V_{L3} , or V_{L5} RSS with the J_L RSS element by PCR as depicted in Fig. 4. PCR products were stained with ethidium bromide (*top*) and Southern blotted and hybridized with a J_L region RSS-specific oligonucleotide probe (*bottom*). Size markers (M) are fragments of phage ϕ X DNA digested with *Hae*III. Differences in the size of the PCR products represent variation in the distance of each V_L -specific PCR primer from the 3' end of each V_L element. (B) Genomic DNA from 20-, 25-, and 30-d embryonic spleens and 2 wk posthatch spleen was assayed as described above. Size markers (M) are fragments of phage ϕ X DNA digested with *Hae*III. As controls, all DNA samples *A* and *B* were also subjected to PCR using a 5' V_{L1} , 5' V_{L3} , or 5' V_{L5} specific oligonucleotide paired with the oligonucleotide specific for the 3' region of that V segment as depicted in Fig. 4. No reproducible difference in the control PCR reactions were observed in any sample (data not shown).

of 51 (46%) joints were in-frame, indicating that little if any selection had occurred. We found 25 different types of joints (Fig. 6). As seen in other species, several of these joints contained one or two basepair P elements but none contained N nucleotide addition. However, only 4 out of 51 (8%) of these joints were of the type GAT TTT GGA that predominate in adult ducks (Fig. 6, *arrow*). Ignoring identical joints from a single PCR reaction essentially gives similar results. 14 out of 36 (39%) were in-frame and 4 out of 36 (11%) were of the type GAT TTT GGA. Thus, by either method of analysis it is clear that no single joint predominates at day 21.

The Predominant V_{L1} - J_L Joint Is Due to Selection. To determine how a predominant V_{L1} - J_L recombination joint arises, we next examined the relative numbers of individual V_{L1} - J_L joints at different times during development. To this end,

we sequenced V_{L1} - J_L joints from days 21, 26, and 35 (hatch) bursa and spleen using PCR. The results are tabulated in Fig. 7. A clear increase in the presence of the GAT TTT GGA joint occurs throughout embryonic development. For example, at day 21, there were 25 types of joints in the spleen, whereas four types were found at day 26, and only two types were found at day 35. Concomitantly, the proportion of joints with GAT TTT GGA increased from 4/51 (8%) at day 21, to 19/28 (68%) at day 26, and to 16/17 (94%) at hatch. As we have previously described, most avian B cells undergo VJ rearrangement of only one of their IgL alleles (16). Thus, cells that produce an out-of-frame joint cannot proceed with further B cell development. As a result, it might be expected that an increase in GAT TTT GGA joints which are in-frame might

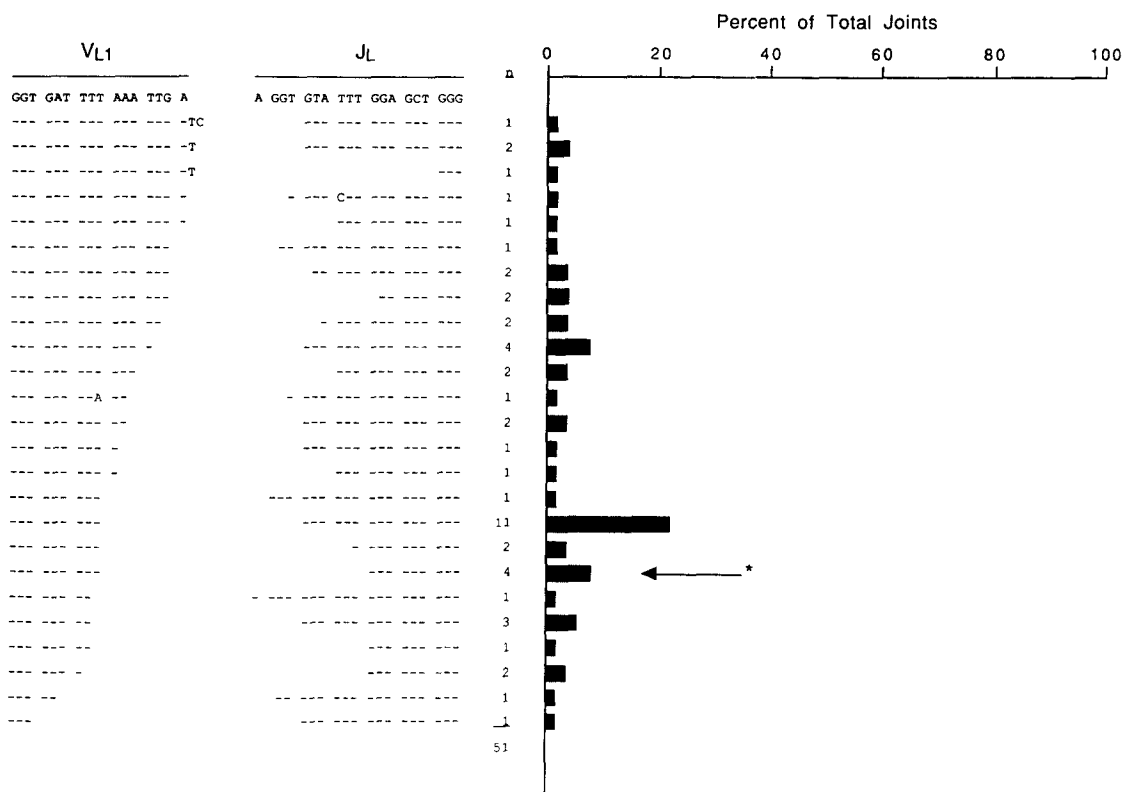


Figure 6. Nucleotide sequences of V_{L1} - J_L joints from 21-d embryonic spleen. The nucleotide sequences of rearranged V_{L1} - J_L genes are compared with the germline sequence. (Dashes) Identity to the germline sequence. Several clones have P nucleotides present at the coding joint, but no N nucleotides were observed. The other base pair changes are consistent with PCR incorporation errors. Rearranged IgL genes were cloned from the spleen of 21-d-old Muscovy duck embryos after PCR amplification. (n) Number of clones isolated with a particular V_{L1} - J_L joint. (Right) Histogram showing the frequency of the joints as a percentage of total joints cloned. (\leftarrow *) Joint predominant in adult ducks.

result just from loss of cells with out-of-frame joints. However, developmental selection of the GAT TTT GGA is just as apparent when the data are expressed as a percentage of in-frame joints. The number of joints with GAT TTT GGA as a percentage of all in-frame joints in the spleen increased

from 4/24 (17%) at day 21, to 19/26 (73%) at day 26, and to 16/16 (100%) at hatch. Essentially the same phenomenon was observed in the bursa.

Similar Selection Occurs in V_{L5} - J_L Joints. We were interested in addressing whether this process of selection was limited to V_{L1} or if it also involved other functional V_L gene rearrangements. Therefore, we sequenced V_{L5} - J_L joints from the bursa and spleen of 4-wk-old Muscovy ducks. As shown in Fig. 8, two types of joints predominated. These two joints

		V_{L1} - J_L Joints					
		Spleen			Bursa		
		Total Joints	Independ. Joints	# of PCRs	Total Joints	Independ. Joints	# of PCRs
21 day	GAT TTT GGA	4/51	4/36	}	6/18	2/11	}
	Types of Joints		25	}	10		}
	In Frame	24/51	14/36	}	6/18	2/11	}
26 day	GAT TTT GGA	19/28	13/20	}	19/25	18/24	}
	Types of Joints		4	}	5		}
	In Frame	26/28	18/20	}	25/25	24/24	}
35 day (Hatch)	GAT TTT GGA	16/17	16/17	}	23/28	23/28	}
	Types of Joints		2	}	4		}
	In Frame	16/17	16/17	}	25/28	25/28	}

Figure 7. Summary of V_{L1} - J_L joints cloned from the bursa and spleen of Muscovy ducks at different developmental times. The number of GAT TTT GGA joints and in-frame joints are expressed as a proportion of the total or independent joints. Joints were characterized as independent on the basis of unique sequence with the V_{L1} - J_L joint, distinct gene conversion tracts within the V_{L1} region, or derivation from separate PCR reactions.

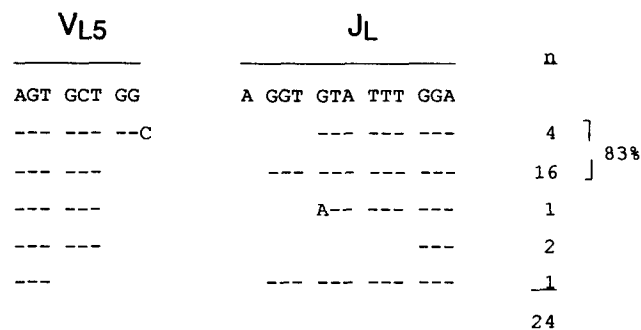


Figure 8. Nucleotide sequences of in-frame V_{L5} - J_L joints from bursa and spleen of 4-wk-old Muscovy ducks. (Dashes) Identity to the germline sequence. Together, the first two joints comprise 83% of all joints in this data set. The top sequence includes a P nucleotide within the coding joint.

together formed 83% of all in-frame joints in the 4-wk-old ducks. However, these two joints comprised only 16% of independent joints isolated from spleen cells at day 21 of development. Both of these joints code for the same amino acids and hence for the same protein product.

Discussion

The process of coding joint formation results in the generation of considerable diversity (3, 8). The junctional diversity created by V(D)J recombination occurs in the CDR3 of the TCR and Ig genes (1). Analysis of the crystal structure of Ig has shown that the CDR3 region forms a major portion of the antigen-binding pocket (8, 29). Thus, junctional diversity results in variations in the CDR3 region of the antigen-binding pocket of both Ig and, by analogy, TCR (9). Therefore, it was surprising when a number of examples of limitations in junctional diversity were reported. Limited junctional diversity has been observed primarily in the neonatal immunologic repertoire. The Ly-1 B cell lineage and γ/δ T cell lineage, both of which develop before birth, have been shown to have V(D)J recombination products with restricted junctional diversity (14, 24, 30, 31). Gu et al. (24) proposed that sequence homologies present at the coding ends are utilized during V(D)J recombination to align the coding sequences before ligation. They suggested that V(D)J coding joint formation occurred in a manner similar to the mechanism of illegitimate recombination proposed by Roth and Wilson (32). In this model, illegitimate recombination of coding joints is initiated by exonuclease digestion of both 5' to 3' or both 3' to 5' strands of the coding ends followed by alignment of the single-strand protrusions. Joining of ends is then mediated by stretches of sequence homology. Gu et al. (24) noted that the coding joints of γ/δ T cells or Ly-1 B cells observed early during ontogeny all involve potential alignment of the most appropriate sequence similarities at the ends of the recombining segments.

It has recently been noted that N segment addition is developmentally regulated with N segment addition beginning before birth and increasing until maturity in a variety of vertebrate species (13, 19–21). Based on these observations, several authors have suggested that homologous joining is a generalized rule for the formation of coding joints (14, 24, 25). They argue that the only reason junctional restriction is limited to neonatal cells is that with the increase in N segment addition observed after birth, the sequence homologies used for alignment are shifted to the N segments, and therefore, are not apparent in the final product. This suggests that the use of homology in the alignment of coding joints is not restricted to neonatal development but is a more general rule of V(D)J recombination. These observations have led to the widely held assumption that homology-mediated pairing provides an important means for the alignment of coding ends for coding joint formation.

During our studies of Muscovy duck B cell development, we came across an example of extreme junctional limitation. This restriction in junctional diversity appeared to be predicted by sequence similarities at the ends of the V_{L1} and J_L coding

sequences. Therefore, we designed experiments to test whether alignment of coding joints using regions of sequence similarity could account for restricted joint formation. Avian species are an excellent system in which to test this hypothesis. All the rules of avian V(D)J recombination established to date are the same as those of other vertebrate species. However, unlike mammals, studies in the chicken showed that V(D)J recombination is restricted to a very short developmental period during early embryogenesis, allowing us to separate events that occur during recombination from those that occur during subsequent B cell development (16). In the present studies, we were able to demonstrate that the duck has a similarly restricted period of V(D)J recombination during embryogenesis. Unlike other avian species, the Muscovy duck has five functional V gene segments. We have previously characterized two of these as a part of our initial identification of the Muscovy duck locus (28). One of these (V_{L1}) was noted to have sequence similarities to the J_L gene segment which were not observed in the other functional V_L region (V_{L5}). We reasoned that if a homology-based joining process is involved in V(D)J recombination, one would expect to observe a restricted set of coding joints in the case of V_{L1} , whereas coding joints containing V_{L5} would be diverse at the time recombination is initiated. As documented in our results, this is not the case. Initial V(D)J recombination of V_{L1} is random. Instead, B cells with Ig L chains encoded by V_{L1} segments undergo developmental selection which rapidly restricts junctional diversity. This selection in VJ coding joints appears to occur in the spleen or during migration to the bursa of Fabricius as virtually all of the functional V_{L1} - J_L recombination events observed in the bursa are of the GAT TTT GGA type. Selection for the GAT TTT GGA joint occurs very rapidly during development. Although the GAT TTT GGA joint accounts for only 8% of V_{L1} joints at day 21, the GAT TTT GGA joint accounts for 94% of the V_{L1} joints at birth. V_{L5} -containing joints in neonatal birds also show limited junctional diversity. In this case, two distinct joints together account for 83% of joints observed in postnatal ducks. Although these two joints are different in position, they result in the same amino acid sequence in the CDR3 domain, suggesting that selection is occurring at the level of the encoded protein.

Based on our data, we conclude that homology plays no significant role in V(D)J recombination. The frequent occurrence of short sequence similarities in the coding ends of recombination events that display junctional limitation does not imply a causal role. Rather, we suggest that the sequence similarities have been selected during evolution to increase the likelihood of a particular protein being formed during an essentially random recombination process. Junctional restriction in the neonatal repertoire suggests that junctional selection occurs before exposure to external antigens. This selection appears to be based on protein sequences in the CDR3 domain of the Ig L chain and suggests that such selection may involve the antigen-binding pocket. This embryonic selection process is not restricted to avian B cells since Ig genes of neonatal murine B cells display similar junctional restriction (20, 24).

The immunologic repertoire of neonatal γ/δ T cells, neonatal Ly-1 B cells, and primary avian B cells have all been suggested to be directed against self-antigens (31, 33–35). Based on our studies, it seems possible that the development of these neonatal cell lineages requires selection by self-antigen in a manner analogous to the way α/β T cells undergo selection by self-encoded MHC molecules during thymic development (36, 37). This hypothesis is supported by several observations. The fact that murine dendritic epidermal T cell clones expressed identically rearranged γ and δ genes suggests that these cells undergo significant selection during neonatal development (15, 22). Similarly, Feeney (30) has shown that there is preferential development of a particular V_HD_HL16.1J_H1 junctional sequence during the neonatal period.

Considerable evidence suggests that the B cells that develop during embryogenesis are long-lived cells which are not constantly regenerated from stem cell precursors (for review see reference 13). One mechanism for the maintenance of such cells is constant reselection by exposure to self-antigens. In contrast, in mammalian B cells that develop postnatally, N nucleotide addition during V(D)J recombination prevents

creation of the recombination products that undergo selection during the prenatal period. Thus, B cells developing postnatally do not display junctional restriction in the absence of exogenous antigens. In the absence of antigenic stimulation, these adult B cells undergo fairly rapid turnover. This suggests that primary B cells with Ig receptors directed against external antigens must be constantly regenerated in the absence of intervening antigen exposure. Thus, the adult or non-Ly-1 B cells may represent an adaptation of the immune system to allow the generation of Ig responses to external antigens. In contrast, the neonatal B cell repertoire is generated in a manner which is dependent to some degree on the recognition of self-antigens. Potentially these self-antigens may be involved in the induction of stress responses, and thus, serve as signals to the primitive immune system to initiate an inflammatory response. Our data suggest that a search for the endogenous ligands of avian B cells, mammalian Ly-1 B cells, and γ/δ T cells may provide significant insights into our understanding of the processes that shape the neonatal immune repertoire.

We thank Louise Carlson for expert technical assistance, and Jeanelle Pickett for preparing the manuscript. We thank our colleagues John Lowe, Stacy Ferguson, Emma Masteller, Pam Reynolds, Antti Mansikka, and Marty Gellert for many stimulating discussions and thoughtful critiques of the manuscript.

This work was supported by National Institutes of Health grant CA-48023.

Address correspondence to Craig B. Thompson, Howard Hughes Medical Institute, 1150 W. Medical Center Drive, MSRB I, Room 3510, Ann Arbor, MI 48109-0650.

Received for publication 16 September 1992.

References

1. Tonegawa, S. 1983. Somatic generation of antibody diversity. *Nature (Lond.)* 302:575.
2. Toyonaga, B., and T.W. Mak. 1987. Genes of the T-cell antigen receptor in normal and malignant T cells. *Annu. Rev. Immunol.* 5:585.
3. Blackwell, T.K., and F.W. Alt. 1989. Molecular characterization of the lymphoid V(D)J recombination activity. *J. Biol. Chem.* 264:10327.
4. Sakano, H., R. Maki, Y. Kurosawa, W. Roeder, and S. Tonegawa. 1980. Two types of somatic recombination are necessary for the generation of complete immunoglobulin heavy-chain genes. *Nature (Lond.)* 286:676.
5. Malissen, M., C. McCoy, D. Blanc, J. Trucy, C. Devaux, A.-M. Schmitt-Verhulst, F. Fitch, L. Hood, and B. Malissen. 1986. Direct evidence for chromosomal inversion during T-cell receptor β -gene rearrangements. *Nature (Lond.)* 319:28.
6. Korman, A.J., J. Maruyama, and D.H. Raulet. 1989. Rearrangement by inversion of a T-cell receptor δ variable region gene located 3' of the δ constant region gene. *Proc. Natl. Acad. Sci. USA.* 86:267.
7. Bernard, O., N. Hozumi, and S. Tonegawa. 1978. Sequences of mouse immunoglobulin light chain genes before and after somatic changes. *Cell.* 15:1133.
8. Davis, M., and P. Bjorkman. 1988. T-cell antigen receptor genes and T-cell recognition. *Nature (Lond.)* 334:395.
9. Lewis, S., A. Gifford, and D. Baltimore. 1985. DNA elements are asymmetrically joined during the site-specific recombination of κ immunoglobulin genes. *Science (Wash. DC)* 228:677.
10. Lieber, M.R., J.E. Hesse, K. Mizuuchi, and M. Gellert. 1988. Lymphoid V(D)J recombination: nucleotide insertion at signal joints as well as coding joints. *Proc. Natl. Acad. Sci. USA.* 85:8588.
11. Reynaud, C.-A., V. Anquez, H. Grimal, and J.-C. Weill. 1987. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. *Cell.* 48:379.
12. Thompson, C.B., and P.E. Neiman. 1987. Somatic diversification of the chicken immunoglobulin light chain gene is limited to the rearranged variable gene segment. *Cell.* 48:369.
13. Rolink, A., and F. Melchers. 1991. Molecular and cellular origins of B lymphocyte diversity. *Cell.* 66:1081.
14. Lieber, M.R. 1991. Site-specific recombination in the immune system. *FASEB (Fed. Am. Soc. Exp. Biol.) J.* 5:2934.
15. Lafaille, J.J., A. DeCloux, M. Bonneville, Y. Takagaki, and S. Tonegawa. 1989. Junctional sequences of T cell receptor $\gamma\delta$

- genes: implications for $\gamma\delta$ T cell lineages and for a novel intermediate of V-(D)-J joining. *Cell*. 59:859.
16. McCormack, W.T., L.W. Tjoelker, L.M. Carlson, B. Petryniak, C.F. Barth, E.H. Humphries, and C.B. Thompson. 1989. Chicken IgL gene rearrangement involves deletion of a circular episome and addition of single nonrandom nucleotides to both coding segments. *Cell*. 56:785.
 17. Lewis, S., and M. Gellert. 1989. The mechanism of antigen receptor gene assembly. *Cell*. 59:585.
 18. Elliott, J.F., E.P. Rock, P.A. Patten, M.M. Davis, and Y. Chien. 1988. The adult T-cell receptor δ -chain is diverse and distinct from that of fetal thymocytes. *Nature (Lond.)*. 331:627.
 19. Meek, K. 1990. Analysis of junctional diversity during B lymphocyte development. *Science (Wash. DC)*. 250:820.
 20. Feeney, A.J. 1990. Lack of N regions in fetal and neonatal mouse immunoglobulin V-D-J junctional sequences. *J. Exp. Med.* 172:1377.
 21. McCormack, W.T., L.W. Tjoelker, G. Stella, C.E. Postema, and C.B. Thompson. 1991. Chicken T-cell receptor β -chain diversity: an evolutionarily conserved D β -encoded glycine turn within the hypervariable CDR3 domain. *Proc. Natl. Acad. Sci. USA*. 88:7699.
 22. Asarnow, D.M., W.A. Kuziel, M. Bonyhadi, R.E. Tigelaar, P.W. Tucker, and J.P. Allison. 1988. Limited diversity of $\gamma\delta$ antigen receptor genes of Thy-1⁺ dendritic epidermal cells. *Cell* 55:837.
 23. Ichihara, Y., H. Hayashida, S. Miyazawa, and Y. Kurosawa. 1989. Only D_{H16}, D_{H2}, and D_{H52} gene families exist in mouse immunoglobulin heavy chain diversity gene loci, of which D_{H16} and D_{H2} originate from the same rimordial D_H gene. *Eur. J. Immunol.* 19:1849.
 24. Gu, H., I. Forster, and K. Rajewsky. 1990. Sequence homologies, N sequence insertion and J_H gene utilization in V_H-DJ_H joining: implications for the joining mechanism and the ontogenetic timing of Ly1 B cell and B-CLL progenitor generation. *EMBO (Eur. Mol. Biol. Organ.) J.* 9:2133.
 25. Feeney, A.J. 1992. Predominance of V_H-D-J_H junctions occurring at sites of short sequence homology results in limited junctional diversity in neonatal antibodies. *J. Immunol.* 149:222.
 26. Thompson, C.B., P.B. Challoner, P.E. Neiman, and M. Grou-dine. 1986. Expression of the *c-myb* proto-oncogene during cellular proliferation. *Nature (Lond.)*. 319:374.
 27. Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scharf, R. Higuchi, G.T. Horn, K.B. Mullis, and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science (Wash. DC)*. 239:487.
 28. McCormack, W.T., L.M. Carlson, L.W. Tjoelker, and C.B. Thompson. 1989. Evolutionary comparison of the avian IgL locus: combinatorial diversity plays a role in the generation of the antibody repertoire in some avian species. *Int. Immunol.* 1:332.
 29. Chothia, C., and A.M. Lesk. 1987. Canonical structures for the hypervariable regions of immunoglobulins. *J. Mol. Biol.* 196:901.
 30. Feeney, A.J. 1991. Predominance of the prototypic T15 anti-phosphorylcholine junctional sequence in neonatal pre-B cells. *J. Immunol.* 147:4343.
 31. Allison, J.P., and W.L. Havran. 1991. The immunobiology of T cells with invariant $\gamma\delta$ antigen receptors. *Annu. Rev. Immunol.* 9:679.
 32. Roth, D., and J. Wilson. 1988. Illegitimate recombination in mammalian cells. In *Genetic Recombination*. R. Kucherlapati and G.R. Smith, editors. ASM Press, Washington, DC. 621-653.
 33. Hayakawa, K., and R.R. Hardy. 1988. Normal, autoimmune and malignant CD5⁺ B cells: the Ly-1 B lineage? *Annu. Rev. Immunol.* 6:197.
 34. McCormack, W.T., L.W. Tjoelker, C.F. Barth, L.M. Carlson, B. Petryniak, E.H. Humphries, and C.B. Thompson. 1989. Selection for B cells with productive IgL gene rearrangements occurs in the bursa of Fabricius during chicken embryonic development. *Genes & Dev.* 3:838.
 35. Forster, I., H. Gu, W. Muller, M. Schmitt, D. Tarlinton, and K. Rajewsky. 1991. CD5 B cells in the mouse. *Curr. Top. Microbiol. Immunol.* 173:247.
 36. Sprent, J., D. Lo, E.-K. Gao, and Y. Ron. 1988. T cell selection in the thymus. *Immunol. Rev.* 101:173.
 37. Benoist, C., and D. Mathis. 1989. Positive selection of T cell repertoire: where and when does it occur? *Cell*. 58:1027.