Genomic Organization of the Chicken T-Cell Receptor β Chain D-J-C Region

Akiko SHIGETA¹⁾, Masaharu SATO¹⁾, Tsuyoshi KAWASHIMA¹⁾, Hiroyuki HORIUCHI¹⁾, Haruo MATSUDA¹⁾ and Shuichi FURUSAWA^{1)*}

¹⁾Laboratory of Immunobiology, Department of Molecular and Applied Bioscience, Graduate School of Biosphere Science, Hiroshima University, 1-4-4 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8528, Japan

(Received 2 June 2004/Accepted 21 July 2004)

ABSTRACT. We examined overlapping genomic clones containing the chicken T cell receptor (TCR) D β -J β -C β complex, which contains a single diversity segment, four joining segments and four exons that encode the constant region. This sequence comprised 18.3 kb. All four J β sequences possessed typical recombination signal sequences (RSS) with intervening 12-bp spacers at their 5'-ends and splice sites at their 3'-ends. No J β -pseudogenes were identified. TGTTG sequences in the RSS heptamer sequences were well conserved, as is the case in mammals. A chicken repeat 1-like sequence was found in the intron region between J β -1336 and C β , and several small repeat sequences were identified in intron regions throughout this cloned genome. As germline sequences revealed complete J β sequences, the CDR3 (complementarity-determining region) sequences of TCR β from non-immunized splenocytes were analyzed. Non-coding (N) and palindromic (P) nucleotides were frequently observed at the D β -J β recombination sites. There were differences in length of deletion at the 5'-end of each J β . Deletion of the 5'-end of J β -1280 was particularly short when compared with that of J β -1336, but there were no changes in the length of the CDR3 using any of the four J β sequences.

KEY WORDS: chicken, D segment, J segment, recombination, TCR.

J. Vet. Med. Sci. 66(12): 1509–1515, 2004

The expression of antigen-specific receptor genes is a key event in lymphocyte maturation. T cell receptors (TCR) exist as heterodimers, either TCR $\alpha\beta$ or $\gamma\delta$, each of which is specific for peptide antigens associated with either MHC class I or class II molecules. During T cell development in the thymus, the functional diversity of the TCR gene is facilitated by somatic recombination of germline sequences from a limited number of variable (V)- and joining (J)-segments in the TCR α - and γ -chains, and V-, diversity (D)- and J segments in the TCR β - and δ -chains.

The TCR $\alpha\beta$ - and $\gamma\delta$ -proteins of chickens are similar to those of mammals. Previous cDNA and partial genomic studies have found that the TCR β locus has two types of major V β gene families (V β 1, and V β 2), one D β , four J β segments and one C β [5]. In contrast to the Ig gene, V β pseudogenes have not yet been identified, and the four J β segments are similar.

V, D, and J segments are flanked by recombination signal sequences (RSS) that are composed of conserved heptamers and nonamers flanking non-conserved 12- to 23-bp spacers. These RSS sequences are recognized by the RAG1/2 protein, which is followed by the initiation of the gene recombination reaction. For chicken TCR β genes, there is no information on the genome sequence of the four J β genes, their orientation, or the RSS sequences, with the exception of one V β 23-RSS and J β -1340 12-RSS [17]. Thus, the non-coding (N) and palindromic (P) sequences at the recombination sites have not yet been analyzed accurately.

In this report, we describe the 18.3 kb genomic segment

that contains the chicken T cell receptor D β -J β -C β complexes, and demonstrate that N and P nucleotides were observed at the V β -J β and D β -J β recombination sites.

MATERIALS AND METHODS

Cell isolation and extraction of genomic DNA: Genomic DNA was extracted from embryonic chick livers taken at day 16 of embryogenesis. Partially inbred chickens (H-B15 White Leghorn; Bu-1a) were originally supplied by Dr. Vainio (Department of Medical Microbiology, Turku University, Finland), and were bred in our animal facilities. Eggs of the H-B15 strain were incubated in a humidified incubator at 38.5°C. Liver tissue was homogenized, and filtered. Cells were lysed by incubation with 10 mM Tris pH8.0, 150 mM NaCl, 10 mM EDTA pH 8.0, 0.1% SDS buffer containing 100 g/ml of proteinase K (Sigma, U.S.A.) at 55°C for 1 hr followed by incubation at 37°C for 16 hr. Genomic DNA was extracted with phenol/chloroform, followed by ethanol precipitation and spooling with a glass rod.

PCR conditions: Genomic D-J-C sequences were acquired from eight PCR clones. All PCR primers used in this study are listed in Table 1. Primers were designed based on known sequences or the results of primer walking. PCR mixtures (50 μ l) contained 50–100 ng of template DNA, 10 mM Tris, pH 8.3, 50 mM KCl, 1.5 mM MgCl₂, 0.2 mM dNTPs, 0.5 μ M of each primer and 1.5 units of AmpliTaq Gold (Perkin Elmer, U.S.A.) polymerase. Amplification consisted of 10 min at 95°C, followed by 30–40 cycles of 1 min at 94°C, 1 min at the optimal temperature of the primers, and 2–3 min at 72°C. As a final extension step, the reaction mixtures were incubated for 10 min at 72°C. To

* CORRESPONDENCE TO: FURUSAWA, S., Laboratory of Immunobiology, Department of Molecular and Applied Bioscience, Graduate School of Biosphere Science, Hiroshima University, 1-4-4 Kagamiyama, Higashi-Hiroshima, Hiroshima 739-8528, Japan.

Table 1. Primer sequences used in PCR amplification

Location	PCR polymerase	Annealing temperature(°C)	Primer sequence	Cloning Enzyme	Reference
D β	A ^{a)}	58	1F5'-ccgcccgcgTTTTCATCAGTCATAGGCAACC-3'	Eag I	[15] ^{d)}
			1R5'-ccctcgagGCAAGAGAAAGCCGTAATAACC-3'	Xho I	[15] ^{d)}
D β -J β 1280	A	54	2F5'-gcctcgagTTCTGTTACCCCTCAAAATGGTTC-3'	Xho I	
			2R5'-gcaagcttTGATTTTCGGTGATGGCACG-3'	Hind III	[7] ^{d)}
J β 1280-J β 4.2	A	58	3F5'-gcggatccCGGTGATGGCACGAAAC-3'	BamH I	[7] ^{d)}
			3R5'-ccctcgagTGTCCCAGTGCCAAAGATCA-3'	Xho I	
J β 4.2-J β 1340	Z ^{b)}	57	4F5'-agcggccgcGTTTCAGAGAACGGACAGTAGCAGGGAGGT-3'	Eag I	
			4R5'-ccgctcgagTTTGTGAGAAGGTGAGGGAGATGACAAGCAC-3'	Xho I	
J β 1340-J β 1336	A	58	5F5'-gcggatccTTTGGACAGGGCACTCG-3'	BamH I	[7] ^{d)}
			5R5'-gcaagcttCTTTGTTCCTTCTCCAAAA-3'	Hind III	[7] ^{d)}
J β 1336-C β	Z	60	6F5'-cccgcccgcTTTGGAGAAGGAACAAAAG-3'	Eag I	
C β	A	56	7F5'-gcgaattcCACCATCAAAGCAAGAGATTCAAGAGAA-3'	EcoR I	[5]
			7R5'-acaggtcgacGTACCAAAGCATCCCCATCACAAT-3'	Sal I	[5]
C β intron C	Ad ^{c)}	68	8F5'-cccgcccgcCAGAGAAGTGCTACAGCTGGGAAGTTG-3'	Eag I	
			8R5'-ccctcgagTTGGTGAAGATTGCCGTGTAGTTTGT-3'	Xho I	
V β -J β 1280	A	54	F5'-TTGTCCGACTGGGAGACTCT-3'		[17] ^{d)}
J β 4.2			R5'-TGATTTTCGGTGATGGCACG-3'		[7] ^{d)}
J β 1340			R5'-TGTCCTCCAGTGCCAAAGATCA-3'		[7] ^{d)}
J β 1340			R5'-TGAACCTTTGGACAGGGCACT-3'		[7] ^{d)}
J β 1340			R5'-CTTTTGTCTCTCTCCAAAA-3'		[7] ^{d)}

a) AmpliTaq Gold. b) Z-Taq. c) Advantage 2. d) Referred to sequences when setting the primer.

amplify longer regions, PCR mixtures (50 μ l) contained 100 ng of template DNA, 10 $\times$ Z-Taq buffer, 0.2 mM dNTPs, 0.2 μ M of each primer and 1.25 units of Z-Taq (TaKaRa, Japan). Amplification consisted of 30–40 cycles of 30 s at 98°C, 10 s at the optimal temperature of the primers, and 2 min at 72°C. To amplify the C β intron, PCR mixtures (50 μ l) contained 100 ng of template DNA, 10 $\times$ Advantage2 buffer, 0.2 mM dNTPs, 0.2 μ M of each primer and 1.25 units of Advantage 2 (Clontech, U.S.A.). Amplification consisted of 3 min at 94°C, followed by 30 cycles of 30 s at 94°C and 3 min at 68°C.

Subcloning and sequences. In order to analyze D β -J β -C β genomic sequences, PCR products were purified using low-melt agarose following digestion with an appropriate restriction enzyme tagged to the 5'-end of each primer, and were then ligated into a pBluescript II SK- (STRAT-AGENE, U.S.A.) vector using Ligation kit ver.1 (TaKaRa, Japan), and transformed into XL1-blue. Deletion mutants were re-cloned if the insert length was longer than 1 kb. To analyze CDR3 (complementarity-determining region) diversification, V β -J β PCR products were purified using low-melt agarose, directly ligated to a pGEM-T Easy vector (Promega, U.S.A.) and transformed into JM109. Cloned *E. Coli* were cultured and purified plasmid DNA was used as a template for sequencing. Sequencing was performed on an ABI 373 or 3100 automatic sequencer. GENETYX software was used to analyze the sequences.

RESULTS

Cloning of the D β to C β region of the chicken TCR β locus: Overlapping genomic clones containing the chicken

TCR β locus (D β -J β -C β gene complexes) were examined. To amplify the possible intervening sequences, chicken genomic DNA, which was extracted from embryonic chick livers at day 16 of embryogenesis, was used as a template for PCR. The sequenced region contained a single D β segment, four J β segments, and four exons encoding the C β region. Contents are schematically illustrated in Fig. 1 (GenBank accession number AB092341). This region spanned a total of 18.3 kb. The four J β segments were located at J β -1280, J β -4.2, J β -1340 and J β -1336. The three introns between J β -1280 and J β -4.2, J β -4.2 and J β -1340, and J β -1340 and J β -1336 were approximately 1.6 kb, 6.6 kb and 1.9 kb in length, respectively. The D β gene was located approximately 1.2 kb upstream of J β -1280, and the C β gene was located approximately 3.6 kb downstream of J β -1336.

The sequenced region contained the complete J β gene, and genomic sequences and predicted amino acid sequences are shown in Fig. 2. J β -4 exists in two putatively allelic forms, J β -4.1 and J β -4.2, which differ by a single substitution and have been described in a previous report [7]. The present study confirms that H-B15 chickens use the J β -4.2 allele. All four chicken J β sequences were flanked on the 5'-end by RSS and on the 3'-end by exon/intron splice sites. Although J β 1340 was reported previously [17], this is the first paper to confirm the presence of the other 3 chicken J β RSS. These RSS consist of a nonamer-12-bp spacer-heptamer motif. Each of RSS site was defined based on similarities to consensus sequences. Homology to the consensus sequences for the nonamer and heptamer was 56% and 75% respectively. The heptamer-GTG sequences were particularly highly conserved. In contrast, the nonamer had T-rich core sequences. Because the rearranged TCR genome is

Fig. 1. Structural features of the H-B15 germline TCR β chain D-J-C region. A. Genomic organization of the TCR β chain D-J-C region. Closed triangles represent 12-RSS (nonamer-12-bp spacer-heptamer). Open triangle represents 23-RSS (heptamer-23-bp spacer-nonamer). B. Location of PCR clone sequences. C. Location of single repeat sequences and CR1 sequences.

Fig. 2. A. Nucleotide sequences of the germline H-B15 TCR $J\beta$ segments. Coding regions of $J\beta$ is indicated in capital letters. RSS nonamer and heptamer sequences are underlined. Boxes indicate the 5' splice donor location. B. Complete amino acid sequences of the H-B15 TCR $J\beta$ segments.

known to contain an N nucleotides addition and partially deleted $J\beta$ sequences, the $J\beta$ sequences gained from cDNA are probably not accurate. Germline sequences revealed complete $J\beta$ sequences at the 5'-end. Pseudogenes that contained $J\beta$ -like sequences with incomplete RSS were not detected.

A copy of the CR1 type-repeat was located between the $J\beta$ -1336 and $C\beta$ inverted transcriptional orientation. The transcriptional orientation of the CR1 repeat is indicated in Fig. 1. The CR1 repeat encodes a retroviral reverse transcriptase pseudogene.

Germline sequences also revealed chicken TCR $C\beta$ exon-intron organization. Exon-intron organization as compared with that in other species is shown in Fig. 3. $C\beta$ was encoded by four exons. The intron-exon splice sites (--- $\downarrow$ GT.....AG $\downarrow$ ---) and their flanking sequences in the chicken TCR genes were closely related to the consensus

splicing signals (--- $\downarrow$ GTAAGT.....CAG $\downarrow$ ---). Exon-intron organization of the chicken $C\beta$ was very similar to that in human, mouse, and cow. Exon length was also very similar among these species, particularly the length of exon 2 (18 bp) and exon 3 (107 bp), which were same in all four species. The length of intron A was shorter, but the length of intron B and intron C was much longer when compared with the other three spaces. Two types of repeat sequence were located in intron C; one of these was a simple GAAA repeat (35 times), and the other was a simple TTCTC repeat (31 times).

Presence of conserved RSS sequence: Comparison of the $J\beta$ RSS flanking sequences of chicken to the nonamer and heptamer consensus sequences of trout, human, mouse and rat (Fig. 4) revealed numerous differences in the nonamer sequence among the species. However, the last three nucleotides (GTG) of the heptamer were well conserved in all of

Fig. 3. Exon-intron structural features of chicken, mouse (Gen Bank accession number: AE000665), human (M12510, M14157) and cow (AF453325) C β . Exons are indicated by closed boxes.

the above species.

Presence of N and P nucleotides at the rearranged sites of V β -D β and D β -J β : The present J β germline sequence makes it possible to evaluate the addition by N nucleotides, and the deletion by exonuclease. Rearranged V β -J β from each of the four J β DNA sequences were cloned from non-immunized splenocytes, and were sequenced. The joining sites of D β -J β were then analyzed (Fig. 5). In addition, the modified nucleotide length in the process of V(D)J rearrangement was shown (Fig. 6). The range of CDR3 was defined as 13 bp from the 3'-end of V β 1 to 18 bp from 5'-end of each J β . An average of 3.8 ± 3.0 (mean $\pm$ SD) N nucleotides were detected at the V β -D β junction and while 4.3 ± 3.4 N nucleotides were identified at the D β -J β junction. Few P nucleotides were observed at the N nucleotide addition sites. P nucleotides can arise from palindromic pairing of germline nucleotides at the ends of gene segments. Several nucleotide deletions were detected at the boundaries of these gene segments. An average of 4.9 ± 3.4 nucleotides were deleted from the 3'-end of V β , 3.1 ± 2.6 and 2.9 ± 2.7 nucleotides from the 5'- and 3'-ends of D β , and 3.8 ± 3.3 nucleotides at the 5'-end of J β . The average remaining D β length was 8.0 ± 3.2 nucleotides. After 3' deletion in V β , 5' and 3' deletion in D β and N and P nucleotide addition at the V β -D β and D β -J β junctions, there were no differences in modified nucleotide length between the 4 types of J β . In contrast, there was a difference in deletion length between J β -1280 and J β -1336, which exhibited average deletion lengths of 1.4 ± 2.4 nucleotides and 5.9 ± 3.1 nucleotides, respectively ($P < 0.0001$). However, there were no differences in the length of the CDR3 irrespective of which of the four J β sequences was used.

DISCUSSION

Avian TCR proteins have been reported to closely resemble those of mammals [7]. Three avian TCR subclasses can

be identified in the chicken with anti-chicken TCR antibodies. One of these is the $\gamma\delta$ type TCR [16], the other two are $\alpha\beta$ type TCR containing either V β 1 or V β 2 [4]. V β 1 and V β 2 T cells differ in their ontogeny [5, 7, 13] and function [5], with germinal center T cells expressing V β 1 [5]. Therefore, V β 1 T cells are thought to play an important role in acquired immune function.

Germline sequences of the chicken TCR β gene were examined in this study. Sequence regions contained a single D β segment, four J β segments, and four exons that encode the C β region. The murine TCR β locus contains a duplication of the D-J-C region, seven J β 1 (two of which were pseudogenes) sequences and seven J β 2 (one of which was a pseudogene) sequences [3]. The present study confirmed that there are four J β sequences between D β and C β , and that there were no J β or D β pseudogenes. Although the chicken contains only one copy of the D-J-C region, the four chicken J β segments spanned about 10 kb, and these are longer than the mouse J β clusters, which span about 2 kb. There were numerous differences in sequence between mammal and chicken J β , although the basic structure of genomic organization is similar; the exon-intron structure of C β is identical in chicken and mammals. These data suggest that the evolutionary course of avian TCR β loci diverged before that of mammals.

All four chicken J β sequences were flanked on the 5'-ends by RSS. Each RSS site was defined based on similarity to consensus sequences. In comparison with mouse, human, rat and rainbow trout J β -RSS, the heptamer-GTG sequences were highly conserved. Although the heptamer-spacer-nonamer motif is well conserved, base substitutions are rather common at certain nucleotides in natural RSS. Akamatsu *et al.* [1] reported that the mutations in the first three nucleotides of the heptamer severely reduced joining rates, and the first residue was found to be particularly essential. As is the case with the present data in chickens, the above nucleotides were well conserved. Therefore, the

		nonamer GGTTTTGT	spacer bp	heptamer CACTGTG
Chicken	J β 1280	AC--A-GC-	12	--TA---
	J β 4	----C-AA-	12	T-G----
	J β 1340	AA--C-G--	12	T-----
	J β 1336	AAG---GA-	12	GG-----
Human	J β 1.1	-A----CAC	12	-----
	J β 1.2	CC----A-A	12	TTA----
	J β 1.3	-----GAA	12	GG-----
	J β 1.4	-----CC-	12	TGT-----
	J β 1.5	--G---GCC	12	-----
	J β 1.5	--G---A-	12	AG-----
	J β 2.1	--AA--C--G	12	-----
	J β 2.2	-----GC-C	12	GG-----
	J β 2.3	-----	12	GG-----
	J β 2.4	A-----C	12	GG-----
	J β 2.5	-----	11	GG-C----
	J β 2.6	-----C	12	GG-----
	J β 2.7 ψ	-----GCA-	12	--T-C---
Mouse	J β 1.1	TA-----C-	13	-----
	J β 1.2	CCA-A--CG	12	TGA-----
	J β 1.3	-----GAA	12	GG-----
	J β 1.4	A-----ACC	12	TGT-----
	J β 1.5	--AG-----	13	T-----
	J β 1.6	-----ACC	12	AG-----
	J β 1.7 ψ	-C-CCA-T-	12	GG-----
	J β 2.1	AA--C--G	12	TG-----
	J β 2.2	-----G-C	12	G-----
	J β 2.2	A-----	12	GG-----
	J β 2.4	A-----	12	GG-----
	J β 2.5	A-----	12	GG-----
	J β 2.6 ψ	-----C-C-	12	GGT-----
	J β 2.7	-----G---	12	-T-----
Rat	J β 1.1	AT---CCTC	12	-----
	J β 1.2	TA--AGA-A	12	TGA-----
	J β 1.3	-----GAA	12	GG-----
	J β 1.4	A-----ACA	12	TGT-----
	J β 1.5	-AG---GCC	12	T-----
	J β 1.6	-----ACC	12	A-----
	J β 2.1	-AA--CGTG	12	TG-----
	J β 2.2	-----G-C	12	G-----
	J β 2.3	-T---AC-	12	GG-----
	J β 2.4	A-----N-	12	GG-----
Trout	J β 1	A--AAAGA-	12	-T-----
	J β 2	CTAAA-A--	11	-----
	J β 3	CA--AC-C	12	-----
	J β 4	TA-----G	13	-GA-----
	J β 5	-T-A---C-	12	--G-----
	J β 6	-T---C---	11	TGG-----
	J β 7	ACA-ACA-G	11	--TA-----
	J β 8	-C-----A	13	T--AC---
	J β 9	-TAA-G-AG	12	-T-----
	J β 11	TAC-AGA-C	12	T-----

Fig. 4. Alignment of J β 12-RSS of chicken, human [18], mouse [8], rat [20] and trout [6]. Canonical RSS [9] are shown in top of the list. Dashes indicate identity with the canonical RSS.

last three nucleotides (GTG) in the heptamer are also likely to be essential in chickens. It is reported that the four J β sequences were equally used among TCR derived from the thymus, spleen and intestine [7]. This strongly suggests that the RSS present in the four chicken J β sequences are equally functional.

In the recombination of RSS, there is what is known as a "12/23 rule", and chicken D β and four J β also possessed RSS with 12- or 23-bp spacers. Because the 5'-end of D β has a 12-spacer RSS, the 3'-end of V β has a 23-spacer RSS. The 5'-ends of the four J β sequences had 12-spacer RSS, and thus recombination between the 3'-end of V β and the 5'-end of J β is possible according to the 12/23 rule. However, such recombination was not observed in this study. It was also reported in a murine system that little or no direct V β to J β joining was observed [2], despite the 12/23 com-

patibility of V β 23-RSS and J β 12-RSS. These data suggest the presence of some unique mechanism for the recombination of these segments.

Some TCR β rearrangement studies in chicken have shown that TCR β sequences have a large number of N sequences in the CDR3 at the V-D and V-J junctions. In this study, differences were observed in the length of the 5'-end deletions in the J β sequences. However, there were no changes in the length of the CDR3 using any of these four J β sequences. It is known that T cell precursors enter the chick thymus in three waves during ontogeny [7]. In the progeny of the third wave, more nucleotides were deleted at the 5'-end of the J β segments, but the overall size of the CDR3 was conserved by a concomitant increase in N nucleotide addition at the V-D-J junctions during rearrangement [7]. The present results confirmed this phenomenon. This suggests that the length of CDR3 in TCR β is essential for T cell function.

The coding sequences of D β , J β and C β , including the intron sequences of each, were first revealed in this study. The present data may prove to be useful in the study of circular DNA derived from recombination sites.

REFERENCES

1. Akamatsu, Y., Tsurushita, N., Nagawa, F., Matsuo, M., Okazaki, K., Imai, M. and Sakano, H. 1994. Essential residues in V(D)J recombination signals. *J. Immunol.* **153**: 4520-4529.
2. Bassing, C. H., Alt, F. W., Hughes, M. M., D'Auteuil, M., Wehrly, T. D., Woodman, B. B., Gartner, F., White, J. M., Davidson, L. and Sleckman, B. P. 2000. Recombination signal sequences restrict chromosomal V(D)J recombination beyond the 12/23 rule. *Nature (Lond.)* **405**: 583-586.
3. Bosc, N. and Lefranc, M.P. 2000. The mouse (*Mus musculus*) T cell receptor beta variable (TRBV), diversity (TRBD) and joining (TRBJ) genes. *Exp Clin Immunogenet.* **17**: 216-228.
4. Chen, C.L., Cihak, J., Losch, U. and Cooper, M.D. 1988. Differential expression of two T cell receptors, TcR1 and TcR2, on chicken lymphocytes. *Eur. J. Immunol.* **18**: 539-543.
5. Cooper, M. D., Chen, C. L., Bucy, R. P. and Thompson, C. B. 1991. Avian T cell ontogeny. *Adv. Immunol.* **50**: 87-117.
6. De Guerra, A. and Charlemagne, J. 1997. Genomic organization of the TcR beta-chain diversity (Dbeta) and joining (Jbeta) segments in the rainbow trout: presence of many repeated sequences. *Mol. Immunol.* **34**: 653-662.
7. Dunon, D., Schwager, J., Dangy, J. P., Cooper, M. D. and Imhof, B. A. 1994. T cell migration during development: homing is not related to TCR V β 1 repertoire selection. *EMBO J.* **13**: 808-815.
8. Gascoigne, N.R., Chien, Y., Becker, D.M., Kavalier, J. and Davis, M.M. 1984. Genomic organization and sequence of T-cell receptor beta-chain constant- and joining-region genes. *Nature (Lond.)* **310**: 387-391.
9. Hesse, J.E., Lieber, M.R., Mizuuchi, K. and Gellert, M. 1989. V(D)J recombination: a functional definition of the joining signals. *Genes. Dev.* **3**: 1053-1061.
10. Hughes, M. M., Tillman, R. E., Wehrly, T. D., White, J. M. and Sleckman, B. P. 2003. The B12/23 restriction is critically dependent on recombination signal nonamer and spacer sequences. *J. Immunol.* **171**: 6604-6610.

← CDR3 →					
Vβ1	P or N	Dβ	P or N	Jβ-1280	clone (total 23)
GCTAAGCAAGATA		GGGACAGGGGGATC		ACAGTAACATGATTTTCG	1
GCTAA-----	CGG	-----AGGG-----	T	ACAGTAACATGATTTTCG	1
GCTAAGCAAGATA		-----CAG-----	CTA	-----TCATGATTTTCG	1
GCTAAGCA-----	GATC	-----GACAGGG-----	TTGCGTTTATGGG	-CAGTAACATGATTTTCG	2
GCTAAG-----	-	-----GGGGG-----	GT	ACAGTAACATGATTTTCG	1
GCTAA-----	C	-----CAGGGGGAT-----	AGGT	ACAGTAACATGATTTTCG	1
GCTAAG-----	GAC	-----GA-----	CC	ACAGTAACATGATTTTCG	1
GCTAAGC-----	GCC	GGGACAGGG-----	-	---GTAACATGATTTTCG	1
GCTAAGCGA-----	-	-----CAGGGGGG-----	ACTGT	ACAGTAACATGATTTTCG	1
GCTAAGCAAGA-----	-	-----CAGGGGGATC-----	GT	ACAGTAACATGATTTTCG	1
GCT-----	TCCGTC	-----GGGGGATC-----	GTACCC	-----TGATTTTCG	1
GCTAAGCAAGATA	-	-----GG-----	CTCTA	---TAACATGATTTTCG	1
GCTAAG-----	-	-----GGGGATC-----	GGT	ACAGTAACATGATTTTCG	1
GCTAAG-----	A	-GGACAGGGG-----	GT	ACAGTAACATGATTTTCG	1
GCTAA-----	GT	-----ACAGGGGGATC-----	G	ACAGTAACATGATTTTCG	1
GCTAAGCAAGATA	CACT	-----GGGG-----	-	---GTAACATGATTTTCG	1
GCTAAGCAAGA-----	CGTC	-----GACAGGGG-----	T	ACAGTAACATGATTTTCG	1
GCTAAGCAAG-----	CACAG	-----GGGGGATC-----	GT	ACAGTAACATGATTTTCG	1
GCTAA-----	TTCCTTTT	-GGACAGGG-----	ATT	ACAGTAACATGATTTTCG	1
GCTAAGCAAGA-----	GAG	GGGACAGGGG-----	-	---AGTAACATGATTTTCG	1
GCTAAGCAAGA-----	CTCC	GGGACAGGGG-----	GT	ACAGTAACATGATTTTCG	1
GCTAAG-----	TCGATGGGTA	GGGAC-----	GATCGTCGG	ACAGTAACATGATTTTCG	1
GCTAAGCAAGAT-----	TTCTTGCA	-----GGGGATC-----	GCCTCGT	ACAGTAACATGATTTTCG	1
Jb-4.2					
GCTAAGCAAG-----	GTC	-----GG-----	CTCC	AATAACGAAAGACTGATC	(total 22)
GCTAAGCA-----	TGA	-----GGATC-----	GACG	-----CGAAAGACTGATC	2
GCTAAG-----	-	-GGACAGGG-----	ACC	---TAACGAAAGACTGATC	1
GCTAA-----	CCA	-----GGGGATC-----	GAGG	---ACGAAAGACTGATC	3
GCTAA-----	AAAGA	-GACAGG-----	ACATT	-ATAACGAAAGACTGATC	1
GCTAA-----	TCCC	GGGACAGGG-----	TCGTAT	-----GACTGATC	1
GCTAAGCAAGATA	-	-----CAGGGG-----	-	---TAACGAAAGACTGATC	1
GCTAAGCAAGATA	C	---AC-----	TCGTCT	-----GAAAGACTGATC	1
GCTAAGCAAG-----	CA	-----GGGGGATC-----	GGCGGG	-----AAGACTGATC	1
GCTAAGCAAG-----	ATT	GGGACAGG-----	TCTCCAG	---AACGAAAGACTGATC	1
GCTAAGCAAG-----	GAGACG	-----AGGG-----	CTTCGGGAC	---ACGAAAGACTGATC	1
GCTAAGCAAGATA	-	-----GGGGGATC-----	TTCAT	---CGAAAGACTGATC	1
GCTAAGCAAGATA	TGACCTTC	-----CAGGGGG-----	GCG	---ATAACGAAAGACTGATC	1
GCT-----	CGAAC	-GACAGGGG-----	-	---ATAACGAAAGACTGATC	1
GCTAAGCAAGATA	C	GGGACAGGGG-----	G	---CGAAAGACTGATC	1
GCTAAGCAAG-----	AGTATTTGA	GGGACAGGGG-----	TTTTT	---ACGAAAGACTGATC	1
Jβ-1340					
GCTAAGCA-----	C	-----G-----	-	CAAAACACCACTGAACT	(total 21)
GC-----	CC	GGGACAGGGGG-----	TTCACC	-----CCACTGAACT	1
GCTAAG-----	TCGGA	-----GGGGATC-----	CTT	-----CCACTGAACT	3
GCTAAGCAAGAT-----	CTT	-----GG-----	C	---AACACACCACTGAACT	1
GCTAA-----	C	-GACAGGGGGATC-----	G	CAAAACACCACTGAACT	1
GCTAAG-----	-	-----GGGGATC-----	GTCTAGGGTCGG	---ACACACCACTGAACT	1
GCTAAGCAAG-----	GGGG	-----AGGGG-----	GTTT	---AACACACCACTGAACT	1
GCTAAG-----	G	-----AGGGGGATC-----	GAAGGACT	---AACACACCACTGAACT	3
GCTAAGCAAGA-----	ATCTCC	GGGACAGGGGG-----	GTTTG	CAAAACACCACTGAACT	1
GCTAAG-----	TAGG	GGGACAGGGGGATC-----	-	CAAAACACCACTGAACT	1
GCTAAGCAAGAT-----	CACACT	-----CAGGG-----	AGGGCGG	-----ACCACTGAACT	1
GCTAAGCAAGA-----	AGAGGTG	GGGACAGGG-----	AAAC	-----CACCCTGAACT	3
GCTAAG-----	GG	-----CAGGGGGAT-----	ATTGAGACGG	---AACACACCACTGAACT	1
GCTAAGCAAG-----	TGTCGA	-----GGGGATC-----	GAAGGTTG	CAAAACACCACTGAACT	1
GCTAAGCAAG-----	GGAATGTGAACG	-----GGGGGAT-----	TCG	CAAAACACCACTGAACT	1
Jβ-1336					
GCTAAGCAAGA-----	C	-----G-----	-	TTATGTTAATATCCAGTA	(total 30)
GCT-----	TCGG	-----ACAGGGGGATC-----	-	-----TTAATATCCAGTA	1
GCTAAGCAAGAT-----	GCTA	GGGACAGGG-----	ACGG	-----GTTAATATCCAGTA	1
GCTAAGCAAGATA	TTGACCC	GGGAC-----	CC	-----TAATATCCAGTA	1
GCTAAG-----	ATATATG	-----AGGGGGATC-----	CCCGA	-----GTTAATATCCAGTA	1
GCTAAGCAAGATA	TAGA	-GGACAGGG-----	CGACCTCTTCATA	-----TAATATCCAGTA	1
GCTAAG-----	T	-GGACAGGG-----	C	-----TTAATATCCAGTA	1
GC-----	CC	-----G-----	TTCGAA	TTATGTTAATATCCAGTA	1
GCTAAGC-----	TCAC	-GACAGG-----	ATA	-----TAATATCCAGTA	1
GCTAAGC-----	CTAGGG	GGGACAGGGGG-----	CAAACTACGACATG	-----A	1
GCTAAG-----	ATACCA	-----GGGGATC-----	GC	-----TATCCAGTA	1
GCT-----	TCGTCTCCGAGC	GGGACAGGG-----	CGTTTCGATCGG	-----TATCCAGTA	1
GCTAA-----	AGACT	-----CAGGGGGATC-----	GG	-----ATATCCAGTA	1
GCTAA-----	TACGGG	GGGACAGGG-----	A	-----AATATCCAGTA	1
GCTAAGCAAGA-----	-	-----CAGGGGG-----	GGTC	-----TAATATCCAGTA	1
GCTA-----	GTC	-GGACAGGGGGATC-----	GACATC	-----TAATATCCAGTA	1
GCT-----	TCGGGGCCC	GGGACAGGGGGATC-----	CTA	-----TAATATCCAGTA	1
GCTAAGCAAG-----	TCCTA	-----ACAGGGGGATC-----	GTCT	-----ATATCCAGTA	2
GCTAAGCAAG-----	-	-----GGGGATC-----	GACGA	-----TAATATCCAGTA	1
GCTAA-----	C	-----CAGGGGGATC-----	GAATCTAC	-----GTTAATATCCAGTA	1
GCTAAGCAAG-----	-	GGGACAGGGGGATC-----	AAAGGG	-----TAATATCCAGTA	1
GCTAA-----	TTAATCAG	-----GGGGATC-----	CAG	-----GTTAATATCCAGTA	1
GCTAAGCAAGATA	C	-GGGACAGGGGGAT-----	-	-----TAATATCCAGTA	1
GC-----	CGCTC	-GGACAGGGGGAT-----	ACGACCC	-----TAATATCCAGTA	1
GCT-----	TTCTGTACC	GGGACAGGGGG-----	TGT	-----TTAATATCCAGTA	1
GCTAAGC-----	GGGC	GGGACAGGG-----	TCGAAGGG	-----ATATCCAGTA	1
GCTAAGCAAGA-----	CTTAACC	-GGACAGGG-----	TTACT	-----TATCCAGTA	1
GCTAAGCA-----	GGTCG	-----GGGGAT-----	GAATACCGTA	-TATGTTAATATCCAGTA	1
GCTAAGCAAG-----	GGGA	GGGACAGGGGGATC-----	GTTTC	-TATGTTAATATCCAGTA	1

Fig. 5. Features of junctional diversity. Nucleotide sequence alignment of Vβ1-Dβ and Dβ-Jβ junction sites is shown. DNA clones from non-immunized spleen lymphocytes are shown for each of the 4 Jβ segments. From left to right, 3'-end of Vβ1, inserted N and P nucleotides, Dβ, N and P, and 5'-end of Jβ. P nucleotides are indicated by bold letters. Analyzed clone numbers are listed in the right side. Nucleotides deleted by exonuclease activity are indicated by dashes.

Fig. 6. Nucleotide length modified through V(D)J rearranged processes for each of the four Jβ segments. Values are the mean ± SD. Total CDR3 nucleotide length is listed on the right side.

- Kamper, S. M. and McKinney, C. E. 2002. Polymorphism and evolution in the constant region of the T-cell receptor beta chain in an advanced teleost fish. *Immunogenetics* **53**: 1047–1054.
- Kanari, Y., Nakagawa, R., Arakawa, H. and Yamagishi, H. 1998. Variable gene segment-specific N-insertions at the signal joint of T-cell receptor V beta-D beta recombinations. *Immunol Lett.* **61**: 151–155.
- Lahti, J.M., Chen, C.L., Tjoelker, L.W., Pickel, J.M., Schat, K.A., Calnek, B.W., Thompson, C.B. and Cooper, M.D. 1991. Two distinct alpha beta T-cell lineages can be distinguished by the differential usage of T-cell receptor V beta gene segments. *Proc. Natl. Acad. Sci. U. S. A.* **88**: 10956–10960.
- Malissen, M., Minard, K., Mjolsness, S., Kronenberg, M., Gorman, J., Hunkapiller, T., Prystowsky, M.B., Yoshikai, Y., Fitch, F., Mak, T.W., et al. 1984. Mouse T cell antigen receptor: structure and organization of constant and joining gene segments encoding the beta polypeptide. *Cell* **37**: 1101–1110.
- McCormack, W.T., Tjoelker, L.W., Stella, G., Postema, C.E. and Thompson, C.B. 1991. Chicken T-cell receptor beta-chain diversity: an evolutionarily conserved D beta-encoded glycine turn within the hypervariable CDR3 domain. *Proc. Natl. Acad. Sci. U. S. A.* **88**: 7699–7703.
- Sowder, J.T., Chen, C.L., Ager, L.L., Chan, M.M. and Cooper, M.D. 1998. A large subpopulation of avian T cells express a homologue of the mammalian T gamma/delta receptor. *J Exp Med.* **167**: 315–322.
- Tjoelker, L. W., Carlson, L. M., Lee, K., Lahti, J., McCormack, W. T., Leiden, J. M., Chen, C. L., Cooper, M. D. and Thompson, C. B. 1990. Evolutionary conservation of antigen recognition: the chicken T-cell receptor β chain. *Proc. Natl. Acad. Sci. U. S. A.* **87**: 7856–7860.
- Toyonaga, B., Yoshikai, Y., Vadasz, V., Chin, B. and Mak, T.W. 1985. Organization and sequences of the diversity, joining, and constant region genes of the human T-cell receptor beta chain. *Proc. Natl. Acad. Sci. U. S. A.* **82**: 8624–8628.
- Tunnacliffe, A., Kefford, R., Milstein, C., Forster, A. and Rabbitts, T.H. 1985. Sequence and evolution of the human T-cell antigen receptor beta-chain genes. *Proc. Natl. Acad. Sci. U. S. A.* **82**: 5068–5072.
- Williams, C.B., Blankenhorn, E.P., Byrd, K.E., Levinson, G. and Gutman, G.A. 1991. Organization and nucleotide sequence of the rat T cell receptor beta-chain complex. *J. Immunol.* **146**: 4406–4413.