Increased frequency of aberrant V(D)J recombination products in core RAG-expressing mice

Sadiqur R. Talukder, Darryll D. Dudley¹, Frederick W. Alt¹, Yousuke Takahama and Yoshiko Akamatsu^{*}

Institute for Genome Research, University of Tokushima, Tokushima 770-8503, Japan and ¹Howard Hughes Medical Institute, Children's Hospital, Harvard Medical School, CBR Institute for Biomedical Research, Boston, MA 02115, USA

Received May 13, 2004; Revised and Accepted July 30, 2004

ABSTRACT

RAG1 and RAG2 play a central role in V(D)J recombination, a process for antigen receptor gene assembly. The truncated 'core' regions of RAGs are sufficient to catalyze the recombination reaction, although with lower joining efficiency than full-length proteins. To investigate the role of the non-core regions of RAGs in the end-joining phase of antigen receptor rearrangement, we analyzed recombination products isolated from core RAG1 and core RAG2 knock-in mice. Here, we report that the truncation of RAGs increases the frequency of aberrant recombination in vivo. Signal joints (SJs) associated with V-to-D recombination of core RAG1 knock-in mice were normal, whereas those of core RAG2 knock-in mice were highly imprecise, containing large deletions and additions, and in some cases coding sequences. In contrast, we found an elevated level of imprecise D-to-J associated SJs for both core RAG1- and RAG2-expressing mice. Likewise, sequences of coding joints (CJs) were also affected by the expression of core RAGs. Finally, sequences found at the junctions of rearranged T-cell receptor loci were highly influenced by differences in rearranging recombination signal sequence pairs. We provide the first evidence that the non-core regions of RAGs have critical functions in the proper assembly and resolution of recombination intermediates in endogenous antigen receptor loci.

INTRODUCTION

During lymphocyte development, immunoglobulin (Ig) and T-cell receptor (TCR) genes are assembled from germline V, D and J gene segments to generate antigen receptor diversity, by a process known as V(D)J recombination (1,2). V(D)J recombination is initiated by RAG1 and RAG2 proteins that introduce double-strand breaks (DSBs) between two coding segments and flanking recombination signal sequences (RSSs). Subsequent repair via the non-homologous end-joining (NHEJ) pathway generates precise signal joints (SJs) and imprecise coding joints (CJs). RSSs are composed of two consensus elements, a heptamer and nonamer, along with an intervening spacer sequence of 12 or 23 bp. Cleavage occurs in a coupled fashion that requires the pairing of a 12 bp RSS with that of a 23 bp RSS. Depending on the transcriptional orientation of the gene segments, recombination results in either deletion or inversion of intervening DNA.

RAG-mediated DNA cleavage produces two types of DNA ends: blunt signal ends (SEs) and covalently sealed hairpin coding ends (CEs). Following DNA cleavage, the ends are held togetherin a cleaved signal complex (CSC), wherein CJ formation is thought to occur, leaving behind a stable signal end complex (SEC) (3–5). The opening and processing of RAG-generated hairpinned CEs are required for CJ formation, and are often accompanied by nucleotide loss and addition. The activity of terminal deoxynucleotidyl transferase (TdT) further increases diversity via N nucleotide addition. In contrast, the blunt SEs within the SEC are usually re-ligated in a precise manner.

Genetic analyses of various mutant cells and mice have revealed important roles for several NHEJ factors in the repair of DSBs introduced by RAGs during V(D)J recombination (1,6). The Ku70/80 heterodimer, XRCC4 and DNA ligase IV are required for the formation of both SJ and CJ. The catalytic subunit of DNA-dependent protein kinase (DNA-PKcs) and Artemis are more important for CJ formation. Recent data suggest that there can be additional component(s) required for V(D)J recombination (7).

Core RAG1 (384–1008 amino acids) and core RAG2 (1– 383 amino acids) have been defined as the minimal regions sufficient for recombination of extrachromosomal substrates (8–12). Although core RAGs are capable of recognizing RSS and perform coupled DNA cleavage *in vitro*, several studies have shown that the missing regions have important functions for normal V(D)J recombinational activity, including suppression of aberrant rearrangements (13–17), regulation of protein degradation (18–20), control of nuclear localization (21,22) and regulation of proper lymphocyte development (23–26).

In addition to these specific effects on V(D)J recombination, studies have shown that the truncated RAGs have decreased overall recombination efficiency compared to the full-length

*To whom correspondence should be addressed. Tel: +1 510 574 1609; Fax: +1 510 574 1500; Email: yakamatsu@pdl.com Present address:

Yoshiko Akamatsu, Protein Design labs, Inc., 34801 Campus Drive, Fremont, CA 94555, USA

Nucleic Acids Research, Vol. 32 No. 15 © Oxford University Press 2004; all rights reserved

RAGs when tested with extrachromosomal substrates (9,11,12). The higher levels of SE recombination intermediates generated using core RAGs suggest that the non-core regions may be required for effective joining, since there is a reduction in the formation of SJ and CJ (27). The identification of RAG mutations within the core regions of RAG1 and RAG2 that affect the second phase of the reaction supports an architectural role of the post-cleavage complex to facilitate the end-joining reaction (28–31). However, a role for RAGs in the end-joining phase of V(D)J recombination at endogenous antigen receptor loci has never been addressed.

To elucidate potential functions of non-core regions of RAGs in end processing and joining of endogenous loci, we have compared the DNA sequences of recombination products isolated from lymphocytes of core RAG1, core RAG2 and wild-type (WT)-RAG expressing mice. We find a reduction in the fidelity of SJ isolated from core RAG1 or core RAG2 mice compared to controls. Expression of core RAG1 alters the features of SJ associated with D-to-J joining but not of V-to-D joining. On the other hand, expression of core RAG2 severely affects the fidelity of V-to-D joining, and to a lesser extent D-J joining. In addition, characteristics of junctional sequences are highly dependent on particular RSS pairings. SJ containing D-coding sequences suggest aberrant cleavage, indicating a possible role of non-core regions of RAGs for controlling proper alignment of the synaptic complex. Our results suggest that the non-core regions of RAGs influence the structure of the synaptic complex and stability of the post-cleavage complex, thereby influencing end-joining products.

MATERIALS AND METHODS

Mice and preparation of genomic DNA

The generation of knock-in mice with the core RAGs have been described (23,24). All the mice used in this study were of the 129SvEv background, and housed under specific pathogenfree conditions. Single cell suspensions of homogenized thymus or spleen were prepared from 8 to 12 weeks old mice. CD4⁺ CD8⁺ (DP) T cells were isolated by staining thymocytes with FITC anti-CD4 and PE anti-CD8 antibodies and sorting on a MoFlo (DakoCytomation). Genomic DNA was prepared as described previously (32).

Analysis of SJs and CJs

An aliquot of 500 ng of genomic DNA for SJ and 150–180 ng for CJ were subjected to 35 cycles of PCR amplification using the following combinations of primers. To amplify SJ associated with V-to-D recombination, V β 14 SIGNAL and D β 1.1 SIGNAL for V β 14-D β 1 SJ; SR1 and SR2 for V δ 5-D δ 2 SJ; and V β 8.3 SIGNAL and D β 1.1 SIGNAL for V β 8.3-D β 1 SJ were used. To amplify SJ associated with D-to-J recombination, SR5 and SR6 for D δ 2-J δ 1 SJ; J β 1-F and D β 1R420 for D β 1-J β 1.1 SJ; and SR14 and SR15 for D β 2-J β 2.1 SJ were used. To amplify V β 14-J β 1.1 CJ, V β 14-S and J β 1.1 CODING primers were used. The primers SR7 and SR8 were used to recover V β 10-J β 2.1 CJ. PCR products containing SJ and CJ were cloned using the TA cloning kit (Invitrogen). For SJ, inserts were identified using PCR and products were digested with ApaLI to detect precise joints. Plasmid DNA was prepared from ApaLI resistant SJ clones and all the CJ clones. Inserts were sequenced using DYEnamicTM ET terminator kit (Amersham Pharmacia Biotech) and analyzed by ABI prism 3100 genetic analyzer (Applied Biosystems).

Oligonucleotides

The following oligonucleotides were used in this study: V β 14 5'-CAGACTTCAACTTGACTATC-3'; Dβ1.1 SIGNAL, SIGNAL, 5'-ACCTTCCTTATCTTCAACTC-3'; SR1, 5'-GC-AAGTCTGGCCTGAACTAA-3'; SR2, 5'-CAACCTGGCA-TGTGACTTTC-3'; Vβ8.3 SIGNAL, 5'-AGCTAGAAACCC-ATCCTGCA-3'; SR5, 5'-CTTGTCCAGTCAACTTCCTG-3'; SR6, 5'-AAGTCATGATGAGCCAGCTG-3'; Jβ1-F, 5'-CT-GTGATGCACACAAAGCGA-3'; Dβ1R 420, 5'-CATTCTG-GATCTAAACAC-3'; SR14, 5'-AAGACCTTGTGAGTCCA-5'-GCCTCATGCAAGGTCAAGAT-3'; CTC-3'; SR15, Vβ14-S, 5'-AGAGTCGGTGGTGCAACTGAACCT-3'; Jβ1.1 CODING, 5'-TGCTTTGTCCGAAGAGAGACCTG-3'; SR7, 5'-GCTTCTCACCTCAGTCTTCA-3'; and SR8, 5'-TGCTAA GGTTTTTCTGCTCCG-3'.

RESULTS

Increased frequency of aberrant SJ in V-to-D recombination in core RAG2 knock-in mice

Unlike CJ, SJ is free from selective forces, leaving the evidence of the joining reaction behind. The majority of SJ isolated from WT lymphocytes do not contain nucleotide additions or deletions (33). Thus, an elevated frequency of imprecise SJ provides evidence of defects in the end-joining process. To compare the sequences of SJ produced in WT, core RAG1 and core RAG2 knock-in mice, we first analyzed SJ formed during V β 14-D β 1 joining (Figure 1 and Table 1). Subcloned PCR fragments containing VB14-DB1 SJ were tested for the presence of a new restriction site, ApaLI, which is generated upon precise head-to-head ligation of RSS heptamers. Genomic thymus DNA was prepared from at least two independent animals. Since no significant difference was observed in the data among these animals, the results were combined to calculate each frequency. As predicted, 90% of clones (43 out of 48 independent clones) isolated from WT thymus were sensitive to ApaLI digestion, demonstrating that the majority of recombination products form precise SJ (Table 1). Similarly, 82% (40 out of 49) of SJ recovered from core RAG1 mice were also precise. On the other hand, only 62% (29 out of 47) of SJ recovered from core RAG2 mice were precise. Chi-square analysis shows that the reduction in the frequency of precise joints in core RAG2 mice is statistically significant (P < 0.01). To further analyze the quality of SJ, ApaLI-insensitive clones were sequenced and compared. Imprecise joints produced in WT and core RAG1 mice contained additions of several nucleotides but no deletions were observed (Figure 1A). On the other hand, imprecise joints produced in core RAG2 mice contained additions and deletions. Two clones contained patches of deletions displaying microhomology at their break points that could have come from either joining

A Vβ14-Dβ1 recombination

VB14 RSS co	oding	N	coding	DB1 RSS
ggtttttgeacagatgtctgecceaccctactcagtgtg(AGA	CTCCAGGCAC)	(GCCCCCTGTCC	C)cacaatgitacagetttatacaaaaaag
Wild type (129 +/+)				
gtttttgeacagatgtctgccccaccctactcagtgtg				cacaatgttacagctitatacaaaaaag
gtttttgeacagatgtctgccccaccctactcagtgtg		CG		cacaatgttacagctttatacaaaaaag
gtttttgeacagatgtctgccccaccctactcagtgtg	CCA	CCCTTA		cacaatgttacagctttatacaaaaaag
gtttitgeacagatgtctgccccaccctactcagtgtg		GG		cacaatgttacagctttatacaaaaaag
gittitgeacagatgtctgeeccaecetactcagigig		GA		cacaatgttacagetttatacaaaaaag
Core RAG1 (1c/1c)				
gtttttgeacagatgtctgccccaccctactcagtgtg				cacaatgttacagctttatacaaaaaag
gtttttgeacagatgtctgccccaccctactcagtgtg		GG		cacaatgttacagctttatacaaaaaag
gtttttgeacagatgtctgccccaccctactcagtgtg	G/	CCC		cacaatgttacagctttatacaaaaaag
gtttitigeacagatgtctgccccaccctactcagtgtg		GA		cacaatgttacagetttatacaaaaaag
ggtttttgeacagatgtctgccccaccctacteagigtg		TC		cacaatgttacagetttatacaaaaaag
gtttttgeacagatgtetgecceaecetacteagtgtg	0	CG		cacaatgttacagetttatacaaaaaag
gtttttgcacagatgtctgccccaccctactcagtgtg	9	GC		cacaatgttacagctttatacaaaaaag
ggtttttgeacagatgtctgccccaccctactcagtgtg	3	GT		cacaatgitacagetttatacaaaaaag
gililigeacagaigteigeeceaceetacteagigig	1	CC		cacaatgttacagetttatacaaaaaag
Core RAG2 (2c/2c)				
ggtttitigeacagatgtctgccccaccctactcagigig				cacaatgitacagetitatacaaaaaag
ggittitgeacagatgtctgccccaccctactcagtgtg		G		cacaatgttacagetttatacaaaaaag
ggtttttgcacagatgtctgccccaccctactcagtgtg(AGA		бT		cacaatgttacagetttatacaaaaaag
ggtttttgcacagatgtctgccccaccctactcagtgtg(AGA	(CTC)	Г		cacaatgttacagetttatacaaaaaag
ggtttttgeacagatgtctgccccaccctactcagtgtg(AGA		÷1	(CCTGTCCC)	cacaatgitacagetttatacaaaaaag
ggtittigeacagatgtctgccccaccctacteagtgtg		TAA		cacaatgttacagetttatacaaaaaag
ggtttttgeacagatgtctgecceaccctactcagtgtg		GAA		cacaatgttacagetttatacaaaaaag
ggtttttgeacagatgtctgccccaccctactcagtgtg		GC		cacaatgttacagctttatacaaaaaag
ggtttttgcacagatgtctgccccaccctactcagtgtg		CCT		cacaatgttacagetttatacaaaaaag
ggtttttgcacagatgtctgccccaccctactcagtgtg		GAC		cacaatgitacagetttatacaaaaaag
gittitgeacagatgtctgccccaccctactcagtgtg		ACT		cacaatgttacagctttatacaaaaaag
ggtttitgeacagatgtctgccccaccctactcagtgtg		TC		cacaatgttacagctttatacaaaaaag
ggtttttgeacagatgtetgeceeacectacteagtgtg		T		cacaatgttacagetttatacaaaaaag
ggittitigeacagatgtctgccccaccctactcagtgtg		GC .		cacaatgttacagetttatacaaaaaag
gtttttgeacagatgtetgecceacectactcagtgtg		βA		cacaatgttacagetttatacaaaaaag
ggtttttgeacagatgtctgccccaccctacteagtgt.(-1)		CCC		cacaatgttacagctitatacaaaaaag
ggtttttgeacagatgtctgccccaccctactcag (- 4)		-		ctttatacaaaaaaag(-13
ggtttttgcaca(-27)		-		gctttatacaaaaaag(-12

В V_{β8.3}-D_{β1} recombination

Vβ8.3 RSS	coding	N	coding Dβ1 RSS	r
acttictgtgcaaagggggggggggggggggggggggggg	FCACTGCTGC	GC) (GCCCCCTGTCCC)cacaatgttacagctttatacaaaaaag	
Wild Type (129+/+)				
actitetgtgcaaagggggggggggggggggggggggggggggg			cacaatgttacagctttatacaaaaaag	47
actitctgtgcaaagggggggggggggggggggggggggg		AT	cacaatgitacagetttatacaaaaaag	1
actitetgtgcaaagggggggggggggggggggggggggggggg		GG	cacaatgttacagctttatacaaaaaag	1
actttetgtgcaaagggggggggggggggggggggggggg		С	cacaatgitacagetitatacaaaaaag	1
Core RAG1 (1c/1c)				
actitetgtgcaaagggggggggggggggggggggggggggggg			cacaatgitacagctitatacaaaaaag	42
actitetgtgcaaagggggggggggggggggggggggggggggg		CCC	cacaatgttacagctttatacaaaaaag	1
actitetgtgcaaagggggggggggggggggggggggggggggg		GGG	cacaatgttacagctttatacaaaaaag	1
actttctgtgcaaaggggggggggggggggggggggggg		CAA	cacaatgttacagctttatacaaaaaag	1
actiticigtgcaaaggggggggggggggggggggggggg		CC	cacaatgttacagctttatacaaaaaag	1
actiticigigcaaaggggggggggggggggggggggggg		AA	cacaatgttacagctttatacaaaaaag	1
actitctgtgcaaagggggggggggggggggggggggggg		TC	cacaatgttacagctttatacaaaaaag	1
actttetgtgcaaagggggggggggggggggggggggggg		G	cacaatgttacagctttatacaaaaaag	1
Core RAG2 (2c/2c)				
actitetgtgcaaagggggggggggggggggggggggggggggg			cacaatgttacagctttatacaaaaaag	32
actttctgtgcaaaggggggggggggggggggggggggg		ATT	(TGTCCC)cacaatgttacagctttatacaaaaaag	2
actttctgtgcaaaggggggggggggggggggggggggg	8	TCGT	(CCCCCTGTCCC)cacaatgttacagctttatacaaaaaag	1
actttctgtgcaaaggggaggaagccacac(-9)		-	(CCCCTGTCCC)cacaatgttacagetttatacaaaaaag	1
actttctgtgcaaaggggggggggggggggggggggggg		GA	(CCCTGTCCC)cacaatgttacagetttatacaaaaaag	1
actttctgtgcaaaggggggggggggggggggggggggg)	AAA	(CCTGTCCC)cacaatgttacagctttatacaaaaaag	1
actttctgtgcaaaggggaggaagccaca(-10)		TG	(CCCTGTCCC)cacaatgttacagctttatacaaaaaag	1
actttctgtgcaaaggggggggggggggggggggggggg		т	(CTGTCCC)cacaatgttacagctttatacaaaaaag	1
actitetgtgcaaagggggggggggggggggggggggggggggg		CCC	cacaatgttacagctttatacaaaaaag	1
actitetgtgcaaaggggaggaagccacacateaetgtg		CC	cacaatgttacagetttatacaaaaaag	1
actitctgtgcaaagggggggggggggggggggggggggg	C	TTGGC	Gacaatgttacagctttatacaaaaaag	1
actitetgtgcaaagggggggggggggggggggggggggggggg	1	AGCCT	cacaatgitacagcittatacaaaaaag	1
actitetgtgcaaaggggaggaagccacacateactgtg		TGG	cacaatgttacagetttatacaaaaaag	1
actttetgtgcaaaggggaggaagccaca(-10)		AAT	aatgttacagctttatacaaaaaag(-3)	1
actttetgtgcaaaggggaggaagccaca(-10)		AT	caatgttacagetttatacaaaaaag(-2) 1
actitetgtgcaaagggggggggggggggggggggggggggggg		Α	agetttatacaaaaaag(-1	1) 1

partner. Furthermore, three clones contained additional sequences identical to coding sequences originally located adjacent to the RSS.

V β 14 rearranges via inversional joining, whereas other V β s undergo deletional recombination, therefore we next asked if the quality of SJ could be affected by transcriptional orientation of RSSs. V β 8.3-D β 1 SJ, produced by deletional recombination, were readily recovered by PCR amplification of total thymus DNA, which contains excision circles formed by deletional recombination. We detected a significant decrease in the frequency of precise Vβ8.3-Dβ1 SJ derived from core RAG2 mice (67%) compared to WT or core RAG1 mice (94 and 86%, respectively; Table 1). VB8.3-DB1 SJ was similar in feature to those isolated from V β 14-D β 1, with nucleotide deletions and additions occurring, and some containing homologous sequences (Figure 1B). However, the frequency of deletions

С V₈₅-D₈₂ recombination

n

43

	coding N	coding Dδ2 RSS	
ggtttgggtacaggeteeetgggcacetgeaceagtg	(ATACCCCGA) (CT	CGTATCCCTCCGAT)cacggtgctacagagetttgcaaaaace	
Wild Type (129+/+)			
ggtttgggtacaggeteeetgggeacetgeaceacagtg		caeggtgctacagagetttgcaaaaaee	5
ggtttgggtacaggeteeetgggcaeetgeaecacagtg	CCGC		-
	GCO		
ggttiggglacaggeteeetgggcacetgeaccacagtg	CTC		
ggtttgggtacaggeteeetgggeacetgeacacagtg	CGC		
ggtttgggtacaggeteeetgggeacetgeaceacagtg			
ggtttgggtacaggeteeetgggeacetgeaccacagtg			
ggtttgggtacaggeteeetgggeacetgeaceacagtg			
ggtttgggtacaggeteeetgggeaeetgeaecacagtg		cacggtgctacagagctttgcaaaaacc	
ggtttgggtacaggctccctgggcacctgcaccagtg	GG	cacggtgctacagagctttgcaaaaacc	
ggtttgggtacaggeteeetgggcacetgeaccacagtg	G	cacggtgctacagagctttgcaaaaacc	
gtttgggtacaggeteeetgggeacetgeaccacagtg	т	caeggtgctacagagetttgcaaaaacc	
ggtttgggtacaggeteeetgggeacetgeaceacagtg	C	cacggtgctacagagctttgcaaaaacc	
Core RAG1 (1c/1c)			
ggtttgggtacaggeteeetgggeacetgeacacagtg		caeggtgetacagagetttgeaaaaace	1
gittgggtacaggeteeetgggcaeetgcaecacagig	GG		1
gillggglacaggeteeelgggcacetgeaccacagig		T(ATCCCTCCGAT) cacggtgctacagagetttgcaaaaace	
gtttgggtacaggeteeetgggcacetgcaccacagtg			
gtttgggtacaggeteeetgggeacetgeaceacagtg	GGC1		
gtttgggtacaggeteeetgggcacetgeaccacagtg	CGG/		
gtttgggtacaggeteeetgggeacetgeaccacagtg	GGCC		
gtttgggtacaggeteeetgggeacetgeaccacagtg	GA/		
gtttgggtacaggeteeetgggcacetgcaccacagtg	CCC		
gtttgggtacaggeteeetgggcacetgcaccacagtg	GC	cacggtgctacagagctitgcaaaaacc	
gtttgggtacaggeteeetgggeacetgeaccacagtg	CC	cacggtgctacagagetttgcaaaaacc	
ggttiggglacaggetecetgggeacetgeaceacagtg	CG	caeggtgctacagagctttgcaaaaacc	
gtttgggtacaggetecetgggcacetgeaccacagtg	Т	caeggtgctacagagetttgcaaaaacc	
Core RAG2 (2c/2c)			1
ggttiggglacaggeteeetgggeacetgeacacagig	175	caeggtgctacagagetttgcaaaaace	
ggtttgggtacaggeteeetgg(-		<u>gca</u> aaaace(-19)	
		T)TTCGCCTTCCGGGC cacggtgctacagagetttgcaaaaacc	
gtttgggtacaggeteeetgggcacetgeae(
gtttgggtacaggeteectgggea(-	14) GTGGC		
gtttgggtacaggctccctgggcac(-	 CTGCA 	CCActacagagctttgcaaaaace(-7)	
gtttgggtacaggetecetgggcacetgcaceae(-4) CTAT.	AAA .acggtgctacagagctttgcaaaaace(-1)	
gttiggglacaggeteeetgggeacetgeacea(
gtttgggtacaggeteeetgggeacetgeaceagtg	GGGG		
gttigggtacaggeteeetgggcacetgcaccacagtg	CG		
gtttgggtacaggetecetgggeacetgeacea(
gtttgggtacaggeteeetgggeacetgeaccacagtg			
		T cacggtgctacagagctttgcaaaaacc	
gtttgggtacaggeteeetgggcacetgeaceacagtg	GC		
gtttgggtacaggeteeetgggcacetgcaceacagt.(Ctgctacagagctttgeaaaaace(-5)	
gtttgggtacaggeteeetgggeacetgeacacagt.(gtttgggtacaggeteeetgggeacetgeacacagtg	-1) CC GC	Ctgctacagagctttgcaaaaace(-5) Ccggtgctacagagctttgcaaaaace(-2)	
gttigggtacaggeteedgggcaeetgeaecagt.(gttigggtacaggeteeetgggcaeetgeaecagtg gttigggtacaggeteeetggg	-1) CC G0 -16) G0	Ctgctacagagctttgcaaaaace(-5) Ccggtgctacagagctttgcaaaaace(-2) Gggtgctacagagctttgcaaaaace(-3)	
gttiggglacaggetecetgggeacetgeaceacagl.(gttiggglacaggetecetgggeacetgeaceacagig gttiggglacaggetecetggg	-1) CC -16) GC -16) GC TC	Ctgctacagagctttgcaaaaace(-5) Ccggtgctacagagctttgcaaaaace(-2) Gggtgctacagagctttgcaaaaace(-3) Gggtgctacagagctttgcaaaaace(-2)	
gttiggglacaggeteetggeaeetgeaecaeagl.(gttiggglacaggeteetggeaeetgeaecaeagig gttiggglacaggeteetggg	-1) CC G(-16) G(-2) A/	Ctgctacagagctttgcaaaaacc(-5) Cggtgctacagagctttgcaaaaacc(-2) Gggtgctacagagctttgcaaaaacc(-3) Gggtgctacagagctttgcaaaaacc(-3) Agtgctacagagctttgcaaaaacc(-4)	
gtttigggtacaggetecetgggeacetgeaceaegt. gttigggtacaggetecetgggeacetgeaceaegtg gtttigggtacaggetecetggg	-1) CC G(-16) G(-2) A/ -5) A/	XC	
gttigggtacaggeteetgeacetgeaceacagt. gttigggtacaggeteetggeacetgeaceacagt gttigggtacaggeteetggeacetgeaceacagt gttigggtacaggeteetggeacetgeaceacag. gttigggtacaggeteeetggeacetgeaceacag	-1) CC G (-16) G (-2) Au -5) Au	Ctgetacqugettigeanance(-5) Cggtgetacqugettigeanance(-2) Gggtgetacqugettigeanance(-2) Aggtgetacqugettigeanance(-4) Aggtgetacqugettigeananace(-4) Aggtgetacqugettigeananace(-1) T cneggtgetacqugettigeananace(-1)	
gttigggtacaggeteetgeacetgeaceacagt. gttigggtacaggeteetggeacetgeaceacagt gttigggtacaggeteetggeacetgeaceacagt gttigggtacaggeteetggeacetgeaceacag. gttigggtacaggeteeetggeacetgeaceacag	-1) CC G(-16) G(-2) Au -2) Au -5) Au G(Ctgttacquegttigeanance(-5) Cggtgetacquegttigeanance(-2) Gggtgetacquegttigeanance(-2) Gggtgetacquegttigeanance(-2) Agtgetacquegttigeanance(-1) Agtgetacquegttigeanance(-1) Tggtgetacquegttigeanance(-1) Cggtgetacquegttigeanance(-1)	
gttfiggthanggetectgggeactgeaceargif. gttfiggthanggetectgggeactgeaceargif. gttfiggthanggetectgggeactgeaceargif. gttfiggthanggetectgggeactgeaceargif. gttfiggthanggetectgggeactgeaceargi. gttfiggthanggetectgggeactgeaceargi. gttfiggthanggetectgggeactgeaceargi.	-1) CC G (-16) G (-2) Au -5) Au	XC	
gttfuggtaneggetectggenetigeneragig gttfuggtaneggetectgggenetigenet	-1) CC GG -16) GG -2) A/ -5) A/ GG GG C	CCtgetaccagecturgeanance(-5) CCggtgetacaggecturgeanance(-2) GCggtgetacaggecturgeanance(-2) GCggtgetacaggecturgeananace(-4) Aggtgetacaggecturgeananace(-1) Aggtgetacaggecturgeananace(-1) CCggtgetacaggecturgeananace(-1)	
tittiggtasegencerg	-1) CC -16) GG -16) GG -2) AJ -5) AJ -5) AQ -C -C -C -C -C -C -C -C -C -C -C -C -C	Ctgttacquagettigeanance(-5) Cggtgetacquagettigeanance(-2) Gggtgetacquagettigeanance(-2) Gggtgetacquagettigeanance(-2) Agtgetacquagettigeanance(-1) Agtgetacquagettigeanance(-1) Tggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeanance(-2) Cggtgetacquagettigeananance(-2) Cggtgetac	
gtttggtanagetoctgggoodgacaaaggg gtttggtagtanagetoctgggoodgacaaaggg gtttggtagtanagetoctgggoodgacaaaggg gtttggtagtanagetoctgggoodgacaaaggg gtttggtagtanagetoctgggoodgacaaaggg gtttggtanagetoctgggoodgacaaaggg gtttggtanagetoctgggoodgacaaaggg gtttggtanagetoctgggoodgacaaaggg gtttggtanagetoctgggoodgacaaaggg gtttggtanagetoctgggoodgacaaaggg gtttggtanagetoctgggoodgacaaaggg	-1) CC -16) GG -2) Au -5) Au -5) Au -5) C -10 -10 -10 -10 -10 -10 -10 -10 -10 -10	Cggtgetnenggettigenangee(-3) Gggtgetnenggettigenangee(-2) Gggtgetnenggettigenangee(-2) Gggtgetnenggettigenangee(-2) Aggtgetnenggettigenangee(-4) Aggtgetnenggettigenangee(-4) Cggtgettigenangee(-4) Cggtgettigenangee(-4) Cggtgettigenangee(-4) Cggtgettigenangee(-4) Cggtgettigenangee(-4) Cggtgettigenangee(-4) Cggtgettigenangee(-4) Cggtgettigenangee(-4) Cggtgettigenangee(-4) Cggtgettigenangee(-4) Cggtgettigenangee(-4) Cggtgettig	
titiggtangenergenergenergenergenergenergenergen	-1) CC -16) GG -2) Au -5) Au -5) Au -5) C -10 -10 -10 -10 -10 -10 -10 -10 -10 -10	Ctgttacaggettgtcanance(-5) Cggtgtacaggettgtcanance(-2) Gggtgtacaggettgtcanance(-2) Gggtgtacaggettgtcanance(-2) Aggtgtacaggettgtcanance(-1) Aggtgtacaggettgtcanance(-1) Cggtgtacaggettgtcanance(-1) Cggtgtacaggettgtcanance(-2) Cggtgtacaggettgtacaggettgtcananace(-2) Cggtgtacaggettgtacaggettgtcananace(-2) Cggtgtacaggettgtcananace(-2) Cggtgtacaggettgtcananace(-2) Cggtgtacaggettgtacaggettgtcananace(-2) Cggtgtacaggettgtacaggettgtcananace(-2) Cggtgtacaggettgtacaggettgtacaggettgtacaggettgtacanance(-2) Cggtgtacaggettgtacaggettgtacanance(-2) Cggtgtacaggettgtacaggettgtacanance(-2) Cggtgtacaggettgtacaggettgtacanance(-2) Cggtgtacaggettgtacanance(-2) Cggtgtacaggettgtacananace(-2) Cggtgtacaggettgtacananace(-2) Cggtgtacaggettgtacananace(-2) Cggtgtacaggettgtacananace(-2) Cggtgtacaggettgtacananace(-2) Cggtgtacaggettgtacananace(-2) Cggtgtacaggettgtacananace(-2) Cggtgtacaggettgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgtacananace(-2) Cggtgt	
gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug gttuggtanagetoctgggootgacacagug	-1) CC -16) GC -16) GC -2) AA -5) AA C G G G C C C C C C C C C C C C C C C	Ctgetacagaettigeanance(-5) Cggtgetacagaettigeanance(-2) Gggtgetacagaettigeanance(-2) Gggtgetacagaettigeanance(-2) Agtgetacagaettigeanance(-2) Tggtgetacagaettigeananace(-2) Cggtgetacagaettigeananace(-2) Cggtgetacagaettigeananace(-2) Cggtgetacagaettigeananace(-1) Cggtgetacagaettigeananace(-1) gtgetacagaettigeananace(-1) gtgetacagaettigeananace(-1) gtgetacagaettigeananace(-1) gtgetacagaettigeananace(-1) gtgetacagaettigeananace(-1) gtgetacagaettigeananace(-1)	
utilitygitasegetecetygesecteseenesse gettiggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse utilityggtasegetecetygesecteseenesse	-1) CCC G G G -16) G G -2) AJ -5) AJ - C G G C C - C - C - C - C - C - C - C - C - C	Ctgttacaggettgtcaanaec(-5) Cggtgtacaggettgtcaanaec(-2) Gggtgtacaggettgtcaanaec(-2) Gggtgtacaggettgtcaanaec(-2) Agtgtacaggettgtcaanaec(-1) Tggtgtacaggettgtcaanaec(-1) Tggtgtacaggettgtcaanaec(-1) Cggtgtacaggettgtcaanaec(-2) Cggtgtacaggettgtcaanaec(-2) Cggtgtacaggettgtcaanaec(-2) ggtgtacaggettgtcaanaec(-2) ggtgtacaggettgtcaanaec(-2) ggtgtacaggettgtcaanaec(-2) ggtgtacaggettgtcaanaec(-2) ggtgtacaggettgtcaanaec(-2) ggtgtacaggettgtcaanaec(-2) ggtgtacaggettgtcaanaec(-2) ggtgtacaggettgtcaanaec(-2) 	
gtttggtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggttagacagtoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg gtttggttagacagtoctgggootgaceaeagtg gtttggtagtanagetoctgggootgaceaeagtg	-1) CCC G G G -16) G G -2) AJ -5) AJ C G G C -5) A 2) T - - - -	Ctgetaccagectificananace(-5) Cggtgetacagectificananace(-2) Gggtgetacagectificananace(-2) Gggtgetacagectificananace(-2) Aggtgetacagectificananace(-1) Tggtgetacagectificananace(-2) Cggtgetacagectificananace(-2) Cggtgetacagectificananace(-2) Cggtgetacagectificananace(-1) gtgetacagectificananace(-1) gtgetacagectificananace(-1) gtgetacagectificananace(-1) gtgetacagectificananace(-3) gtgetacagectificananace(-3) gtgetacagectificananace(-3) gtgetacagectificananace(-3) gtgetacagectificananace(-3) gtgetacagectificananace(-4) gtgetacagectificananace(-4)	
ttittittittittittittittittittittittitti	-1) CCC G G G -16) G G -2) AA -5) AA - C G C C - C - C - C 	Ctgttacaggettgtcanance(-5) Cggtgtacaggettgtaanance(-2) Gggtgtacaggettgtaanance(-2) Gggtgtacagggettgtaanance(-2) Aggtgtacagggettgtaanance(-1) Aggtgtacagggettgtaanance(-1) Cggtgtacagggettgtaanance(-1) Cggtgtacagggettgtaanance(-1) Cggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1)	
gtttggtansgetoctggenoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage g	-1) CCC G G -16) G (-16) G (-16) G (-10) G	CCtgetaccagectiftecanance(-5) CCgetgetacagectiftecanance(-2) GCgetgetacagectiftecanance(-2) GCgetgetacagectiftecanance(-2) Agetgetacagectiftecanance(-1) Agetgetacagectiftecanance(-1) CCgetgetacagectiftecanance(-1) CCgetgetacagectiftecanance(-1) CCgetgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) getgetacagectiftecanance(-1) 	
gtttggtansgetoctggenoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage gtttggtansgetoctganorage g	-1) CCC G G -16) G (-16) G (-16) G (-10) G	Ctgttacaggettgtcanance(-5) Cggtgtacaggettgtaanance(-2) Gggtgtacaggettgtaanance(-2) Gggtgtacagggettgtaanance(-2) Aggtgtacagggettgtaanance(-1) Aggtgtacagggettgtaanance(-1) Cggtgtacagggettgtaanance(-1) Cggtgtacagggettgtaanance(-1) Cggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1) ggtgtacagggettgtaanance(-1)	
uttituggtasegencetorggeneetige	-1) CCC -16) GG -16) GG -17 -2) AA -5) AA -7 -7 -7 -1) -7 -1) -7 -1) -7 -1) -7 -1) -7 -1) -7 -1) -7 -1) -7 -1) -7 -1) -7 -1 -1) -7 -1 -1 -1 -1 -1 -1 -1 -1 -1 -1	Ctgttacaggettgtcaansec(-5) Gggtgtacaggettgtaansec(-2) Gggtgtacaggettgtaansec(-2) Gggtgtacaggettgtaansec(-2) Aggtgtacaggettgtaansec(-1) Aggtgtacaggettgtaansec(-1) Cggtgtacaggettgtaansec(-1) Cggtgtacaggettgtaansec(-1) ggtgaansec(-1) 	
gettinggtanaggetectgezencegnacemaging gettinggtanaggetectgezencegnacemaning gettinggtanaggetectgezencegnacemaning gettinggtanaggetectgezencegnacemaning gettinggtanaggetectgezencegnacemaning gettinggtanaggetectgezencegnacemaning gettinggtanaggetectgezencegnacemaning gettinggtanaggetectgezencegnacemaning gettinggtanaggetectgezencegnacemaning gettinggtanaggetectgezencegnacemaning	-1) CCC G G -16) G G -16) G G T -2) A - - -2) C - - - - - - - - - - - - -	CCtgetaccagecittgeaaaace(-3) CCgetgetacagecittgeaaaace(-2) GCgetgetacagecittgeaaaace(-2) GCgetgetacagecittgeaaaace(-2) Agetgetacagecittgeaaaace(-1) TCgetgetacagecittgeaaaaace(-2) CCgetgetacagecittgeaaaaace(-2) CCgetgetacagecittgeaaaaace(-2) CCgetgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) getgetacagecittgeaaaaace(-1) 	
<pre>gttfuggtangegetectggeneet</pre>	-1) CCC -16) GG -16) GG -17 -2) AA -5) AA -7 -7 -7 -1) -7 -1) -7 -1) -7 -1) -7 -1) -7 -1) -7 -1) -7 -7 -7 -7 -7 -7 -7 -7 -7 -7	Ctgttacaggettgtcaansec(-5) Cggtgtacaggettgtcaansec(-2) Gggtgtacaggettgtcaansec(-2) Gggtgtacaggettgtcaansec(-2) Aggtgtacaggettgtcaansec(-2) Aggtgtacaggettgtcaansec(-1) Tggtgtacaggettgtcaansec(-2) Cggtgtacaggettgtcaansec(-2) Cggtgtacaggettgtcaansec(-2) ggtgtacaggettgtcaansec(-2) ggtgtacaggettgtcaansec(-2) ggtgtacaggettgtcaansec(-2) ggtgtacaggettgtcaansec(-2) ggtgtacaggettgtcaansec(-2) ggtgtacaggettgtcaansec(-2) ggtgtacaggettgtcaansec(-2) ggtgtacaggettgtcaansec(-3) ggtgtacaggettgtcaansec(-3) ggtgtacaggettgtcaansec(-3) ggtgtacaggettgtcaansec(-3) ggtgtacaggettgtcaansec(-3) ggtgtacaggettgtcaansec(-3) ggtgtacaggettgtgaansec(-3) ggtgttgtacaggettgtgaansec(-3) ggtgttgtacaggettgtgaansec(-3) 	
gettiggtanaggetectgeacetgeacetgeacet gettiggtanaggetectgeacetgeacetgeacet gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectgeacetgeacetag gettiggtanaggetectag g	-1) CCC -16) G -16) G -16, G -16, G -17,	CCtgetaccagectiftecanance(-5) CCgetgetacagectiftecanance(-2) GCgetgetacagectiftecanance(-2) GCgetgetacagectiftecanance(-2) Agetgetacagectiftecanance(-1) TC caeggtgetacagectiftecanance(-1) CCgetgetacagectiftecanance(-1) CC caeggtgetacagectiftecanance(-1) CC caeggtgetacaggectiftecanance(-1) getgetacaggectiftecanance(-1) getgetacaggectiftecanance(-1) getgetacaggectiftecanance(-1) getgetacaggectiftecanance(-1) getgetacaggectiftecanance(-1) getgetacaggectiftecananace(-1) getgetacaggectiftecananace(-1) getgetacaggectiftecananace(-1) getgetacaggectiftecananace(-1) getgetacaggectiftecananace(-1) getgetacaggectiftecananace(-1) 	
puttinggationaggiotochgggenetigeneensee guttinggationaggiotochggeneensegut guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochgeneenselsenseense guttinggationaggiotochgeneenselsenseense guttinggationaggiotochgeneenselsenseense guttinggationaggiotochgeneenselsenseense guttinggationaggiotochgeneenselsenseense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense	-1) CCC -16) GG -16) GG -16) GG -16) GG -10 -10 -11)	Ctgttacaggettgtcaansec(-5) Cgtgtacaggettgtaansec(-2) Ggtgtacaggettgtaansec(-2) Ggtgtacaggettgtaansec(-2) Agtgtacaggettgtaansec(-2) Agtgtacaggettgtaansec(-1) Tgtgtacaggettgtaansec(-2) Cgtgtacaggettgtaansec(-2) Cgtgtacaggettgtaansec(-2) Cgtgtacaggettgtaansec(-2) gtgtacaggettgtaansec(-2) gtgtacaggettgtaansec(-2) gtgtacaggettgtaansec(-2) gtgtacaggettgtaansec(-2) gtgtacaggettgtaansec(-2) gtgtacaggettgtaansec(-2) gtgtacaggettgtaansec(-2) gtgtacaggettgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtgtaansec(-2) gtaansec(-2) gtgtaansec(-2) gtaansec(-2	
puttinggationaggiotochgggenetigeneensee guttinggationaggiotochggeneensegut guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochggeneenselsenseense guttinggationaggiotochgeneenselsenseense guttinggationaggiotochgeneenselsenseense guttinggationaggiotochgeneenselsenseense guttinggationaggiotochgeneenselsenseense guttinggationaggiotochgeneenselsenseense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense guttinggationaggiotochgeneense	-1) CCC -16) GG -16) GG -16) GG -16) GG -10 -10 -11)	CCtgetaccagecittgeanance(-5) CCgetgetacagecittgeanance(-2) GCgetgetacagecittgeanance(-2) GCgetgetacagecittgeanance(-2) Agetgetacagecittgeanance(-2) Agetgetacagecittgeanance(-1) CCgetgetacagecittgeanance(-1) CCgetgetacagecittgeananace(-1) CCgetgetacagecittgeananace(-1) getgetacagecittgeananace(-1) getgetacagecittgeananace(-1) getgetacagecittgeananace(-1) getgetacagecittgeananace(-1) getgetacagecittgeananace(-1) getgetacagecittgeananace(-1) getgetacagecittgeananace(-1) getgetacagecittgeananace(-1) getgetacagecittgeananace(-1) getgetacagecittgeananace(-1) getgetacagecittgeananace(-1) 	
gettingstanaggetectiggeactigaacaaagu gettingstanaggetectiggeacatigaacaaagu	-1) CCC -16) GG -16) GG -16) GG -16) GG -17 -2) AA - - -1) - - - - - - - - - - - - - -	CCtgetaccagectiftgeamanee(-5) CCgetgetacagectiftgeamanee(-2) GCgetgetacagectiftgeamanee(-2) GCgetgetacagectiftgeamanee(-2) Agetgetacagectiftgeamanee(-2) Agetgetacagectiftgeamanee(-1) CCgetgetacagectiftgeamanee(-2) CCgetgetacagectiftgeamanee(-2) CCgetgetacagectiftgeamanee(-2) CCgetgetacagectiftgeamanee(-2) CCgetgetacagectiftgeamanee(-2) getgetacagectiftgeamanee(-1) getgetacagectiftgeamanee(-1) getgetacagectiftgeamanee(-1) getgetacaggetiftgeamanee(-1) getgetacaggetiftgeamanee(-3) getgetacaggetiftgeamanee(-3) getgetacaggetiftgeamanee(-3) 	
mettingsthengsotochggenoctigaacenaengst gettingsthengsotochggenoctigaacenaengst gettingsthengsotochggenoctigaacenaengst gettingsthengsotochggenoctigaacenaengst gettingsthengsotochggenoctigaacenaengst gettingsthengsotochggenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenocti	-1) CCC -16) GG -16) GG -16) GG -16) GG -10 -22) AA -10 -10 -11) -2 -11) -2 -11) -2 -11) -2 -11) -2 -23) -2 -23) -2 -27) -2 -27) -2 -27) -2 -27) -2 -27) -2 -27) -2 -2 -27) -2 -2 -27) -2 -2 -27) -2 -2 -2 -2 -2 -2 -2 -2 -2 -2	Ctgttacapactitgcanance(-5) Ggtgtacapactitgcanance(-2) Ggtgtacapactitgcanance(-2) Ggtgtacapactitgcanance(-2) Agtgtacapactitgcanance(-1) Tgtgtacapactitgcanance(-1) Tgtgtacapactitgcanance(-1) Cgtgtacapactitgcanance(-1) Cgtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcananace(-1) gtgtacapagactitgcananace(-1) gtgtacapagactitgcananace(-1) gap	5
gettinggitanaggiocotggenociganeenaget gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget. gettinggitanaggiocotggenociganeenaget.	-1) CCC -16) GG -16) GG -16) GG -16) GG -17 -22) A - - - - - - - - - - - - -	CCtgctaccagectiftcananace(-5) CCgtgtcacagectiftcananace(-2) GGgtfctacagectiftcananace(-2) GGgtfctacagectiftcananace(-2) Agtfctacagectiftcananace(-2) Agtfctacagectiftcananace(-2) TC cacggtgtacagectiftcananace(-2) CC cacggtgtacagectiftcananace(-2) gtfctacagectiftcananace(-2) gtfctacagectiftcananace(-2) gtfctacagectiftcananace(-2) gtfctacagectiftcananace(-2) gtfctacagectiftcananace(-2) gtfctacagectiftcananace(-2) gtfctacagectiftcananace(-2) gtfctacagectiftcananace(-2) gtfctacagectiftcananace(-3) gtfctacagectiftcananace(-3) gtfctacagectiftcananace(-3) 	5
mettingsthengsotochggenoctigaacenaengst gettingsthengsotochggenoctigaacenaengst gettingsthengsotochggenoctigaacenaengst gettingsthengsotochggenoctigaacenaengst gettingsthengsotochggenoctigaacenaengst gettingsthengsotochggenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenoctigaacenaense gettingsthengsotochgenocti	-1) CCC -16) GG -16) GG -16) GG -16) GG -10) GG -22) AA -10 -11)	Ctgttacapactitgcanance(-5) Ggtgtacapactitgcanance(-2) Ggtgtacapactitgcanance(-2) Ggtgtacapactitgcanance(-2) Agtgtacapactitgcanance(-1) Tgtgtacapactitgcanance(-1) Tgtgtacapactitgcanance(-1) Cgtgtacapactitgcanance(-1) Cgtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcanance(-1) gtgtacapagactitgcananace(-1) gtgtacapagactitgcananace(-1) gtgtacapagactitgcananace(-1) gap	

Figure 1. Sequences of SJs formed by V-D recombination. (A) PCR fragments containing SJs formed by V β 14-D β 1 recombination were compared among WT (RAG1^{+/+} RAG2^{+/+}), RAG1 core (RAG1^{c/c} RAG2^{+/+}) and RAG2 core (RAG1^{+/-} + RAG2^{c/c}) mice. The original sequences are indicated by lowercase letters and possible insertions are indicated in capital letters. Heptamer and nonamer sequences are in **boldface** letters. Nucleotide additions containing more than three continuous bases identical to either CEs were considered as coding sequence and indicated in parentheses. Nucleotide additions assigned as N additions with 2 or 3 nt homology with either CE are italicized. Numbers of deleted bases are shown in parentheses at the right end of each RSS. Microhomologies are underlined. 'n' indicates the number of junctions with the indicated sequence. (B) SJs formed by VB8.3-DB1 recombination. (C) SJ formed by $V\delta 5$ -D $\delta 2$ recombination.

and presence of coding sequences were elevated when $5'D\beta1$ RSS was paired with V β 8.3 RSS (Table 1). Thus, the characteristics of SJ appeared sensitive to the differences in RSS partners.

To determine if the reduced ability of core RAGs to mediate precise SJ extended to other TCR loci, we analyzed V δ 5-D δ 2 SJ that occur by inversional joining (Figure 1C and Table 1). A total of 83 and 81% of clones recovered from thymus DNA of WT and core RAG1 knock-in mice, respectively, contained precise V δ 5-D δ 2 SJ. All imprecise joints were found to

Table	1.	Comparison	of	V-D SJs	
-------	----	------------	----	---------	--

	WT	Core RAG1	Core RAG2
Vβ14-Dβ1			
Precise joints	43/48 (90%)	40/49 (82%)	29/47 (62%)
Imprecise joints	5/48 (10%)	9/49 (18%)	18/47** (38%) ^a
N addition	5/5 (100%)	9/9 (100%)	15/18 (83%)
Coding sequences	0/5 (0%)	0/9 (0%)	3/18 (17%)
Deletion	0/5 (0%)	0/9 (0%)	3/18 (17%)
Microhomology	0/5 (0%)	0/9 (0%)	2/18 (11%)
Vβ8.3-Dβ1			
Precise joints	47/50 (94%)	42/49 (86%)	32/48 (67%)
Imprecise joints	3/50 (6%)	7/49 (14%)	16/48** (33%) ^a
N addition	3/3 (100%)	7/7 (100%)	15/16 (94%)
Coding sequences	0/3 (0%)	0/7 (0%)	8/16 (50%)
Deletion	0/3 (0%)	0/7 (0%)	12/16 (75%)
Microhomology	0/3 (0%)	0/7 (0%)	0/16 (0%)
Vδ5-Dδ2			
Precise joints	55/66 (83%)	54/67 (81%)	15/56 (27%)
Imprecise joints	11/66 (17%)	13/67 (19%)	41/56** (73%) ^a
N addition	11/11 (100%)	13/13 (100%)	23/41 (56%)
Coding sequences	0/11 (0%)	1/13 (8%)	1/41 (2%)
Deletion	0/11 (0%)	0/13 (0%)	37/41 (90%)
Microhomology	0/11 (0%)	0/13 (0%)	13/41 (32%)

Number of each joint out of total number analyzed was shown. The frequency of each joint is shown in parentheses. Significant differences versus WT mice, when compared by chi-squared test, are indicated by ** (P < 0.01). Junctions were sequences from two independent animals, and the results were combined to calculate the number of each case.

^aDifferences from WT mice are still statistically significant even when the data containing coding sequences are excluded, to exclude possible non-12/23 HJ from SJ (see Discussion).

contain nucleotide additions, with no deletions observed. In contrast, only 27% of V δ 5-D δ 2 SJ isolated from core RAG2 thymus DNA were precise (Table 1). Strikingly, 90% (37 of 41) of imprecise V δ 5-D δ 2 SJ from core RAG2 thymus DNA contained deletions (average = -17.1 ± 2.8 bp). Furthermore, 13 imprecise joints exhibited microhomologies at the junction. These results indicate that the non-core region of RAG2, but not the non-core region of RAG1, is important for precise V δ 5-D δ 2 SJ formation.

These findings demonstrate that a recombinase complex containing truncated RAG2 more frequently produces imprecise SJ than does full-length RAG2. In contrast, core RAG1 does not affect the characteristics of SJ produced by VD joining. Furthermore, the quality of SJ produced in the presence of truncated RAG2 was not dependent on the genomic configuration of RSS pairs, as we detected reduced fidelity in SJ formed during both deletional and inversional recombination. However, all three different sets of VD joints tested from core RAG2 knock-in mice showed distinct preferences for modifications: VB14-DB1 SJ displayed fewer deletions (3 out of 18 imprecise SJ) though with a high frequency of N additions (15 out of 18). V β 8.3-DB1 SJ often contained DB1 coding sequences (8 out of 16 imprecise SJ), nucleotide additions (15 out of 16) and deletions (12 out of 16). Most VS5-DS2 SJ had deletions (37 out of 41 imprecise SJ), and frequently contained short stretches of microhomology (13 out of 41) (Table 1). Thus, the characteristics of end processing mediated by core RAG2 when producing imprecise SJ seemed to be affected by the combination of different RSSs.

Increased frequency of aberrant SJ in D-to-J recombination in core RAG1 and core RAG2 knock-in mice

To determine if non-core regions of RAGs differentially influence V-to-D and D-to-J joining steps, we next analyzed SJ associated with D β 1-J β 1.1, D β 2-J β 2.1 and D δ 2-J δ 1 joining, all produced by deletional recombination (Figure 2 and Table 2). PCR amplification of total thymus DNA isolated from WT mice demonstrated 14, 18 and 15% of D β 1-J β 1.1, D β 2-J β 2.1 and D δ 2-J δ 1 SJ, respectively, were imprecise joints (Table 2). In contrast, we detected a 2-4-fold increased frequency (for D\beta1-J\beta1.1, D\beta2-J\beta2.1 and D\beta2-J\beta1 SJ: 29, 80 and 43% for core RAG1 and 26, 69 and 22% for core RAG2, respectively) of imprecise SJ from both core RAG1- and core RAG2-expressing mice. The higher frequency of imprecise joints associated with core RAG1 expression was statistically significant for D β 2-J β 2.1 and D δ 2-J δ 1 SJ (Table 2). Although the overall frequency of D β 1-J β 1.1 and D δ 2-J δ 1 imprecise joints were elevated in core RAG2 expressing mice, only the level of D\u00e32-J\u00e32.1 imprecise joints was statistically significant.

Thus, we conclude that the expression of core RAG1 affected the accuracy of recombination associated with DJ joining in general. Expression of core RAG2 also influences DJ associated SJ, however, it seems to be more sensitive to different RSS pairs. Most imprecise D β 1-J β 1.1 and D δ 2-J δ 1 joints contained N nucleotide additions, although some deletions were detected in a few D β 1-J β 1.1 SJ isolated from both core RAG1 and core RAG2 knock-in mice. D β 2-J β 2.1 SJ was even more severely affected, containing deletions, N nucleotide additions and sequences homologous to coding gene segments at the junction (Figure 2B). The long D β 2 coding sequences attached to modified J β 2.1 SE strongly suggests that the involvement of aberrant cleavage during D β 2-J β 2.1 recombination.

Non-core regions of RAGs influence the quality of CJ in vivo

To ask if non-core regions of RAGs play a critical role in normal CJ formation, we next analyzed CJ associated with V β 14-D β -J β 1.1 and V β 10-D β -J β 2.1 (Figure 3). The V β 14 RSS was selected as an example of an RSS that rearranges by inversion. V β 10 RSS was chosen as an example of an RSS whose recombination with downstream gene segments results in the deletion of genomic DNA, since it showed the closest match with consensus 23-RSS of VH/VB origin (25). VB14-Dβ-Jβ1.1 CJ was analyzed from genomic DNA of splenocytes using PCR (Figure 3A). To exclude effects biased by the negative/positive selection of productive rearrangements (34), VB14-DB-JB1.1 CJ was also analyzed from genomic DNA of CD4⁺CD8⁺ (DP) thymocytes (Figure 3B). V β 10-Dβ-Jβ2.1 CJ was analyzed from DP thymocytes (Figure 3C). Frequency and average length of nucleotide modifications are summarized in Table 3.

Although the sequences of CJ were heavily modified in all clones analyzed, CJ isolated from core RAG1 and core RAG2 knock-in mice frequently contained a higher degree of modification than those of WT mice (Table 3). The average length of V β 14 segment deletions were larger in core RAG1 and core RAG2 than in WT (-2.86, -3.34 and -1.79 bp, respectively), although we did not observe a difference in the frequency of such deletions. Similarly, the average length of

DB1-JB1.1 recombination

A P	β1-Jβ1.1 recombin	ation			
	Dβ1 RSS getteccatagaattgaatcaccgtg	coding N	coding (GTGTTTG)ca	Jβ1.1 RSS cagtgccataggatgaggagaaaaat	n
Wild Type			AT 100 000		
	getteecatagaattgaatcaccgtg		ca	cagtgccataggatgaggagaaaaat	43
ggtttttgtaaag	getteecatagaattgaateacegtg	GAAGO	G ca	cagtgccataggatgaggagaaaaat	1
ggtttttgtaaag	getteecatagaattgaatcaccgtg	AAGC		cagtgccataggatgaggagaaaaat	1
	getteecatagaattgaateacegtg getteecatagaattgaateacegtg	GGGT GCCA		cagtgccataggatgaggagaaaaat	1
ggtttttgtaaa	getteccatagaattgaatcaccgtg	GGC		cagtgccataggatgaggagaaaaat cagtgccataggatgaggagaaaaat	i
ggtttttgtaaas	getteccatagaattgaatcaccgtg	GCC	ca	cagtgccataggatgaggagaaaaat	1
ggtttttgtaaag	getteecatagaaltgaatcaccgtg	GA	ca	cagtgecataggatgaggagaaaaat	1
Core RAG	i1 (1c/1c) rgcttoccatagaattgaatcaccgtg		ca	cagtgccataggatgaggagaaaaat	34
	getteecatagaattgaatcaccgtg	CCCGT	C ca	cagtgccataggatgaggagaaaaat	1
	rgetteecatagaattgaatcaccgtg	CCGA		caglgccataggatgaggagaaaaat	1
	getteecatagaattgaateacegtg	GGTG		cagtgccataggatgaggagaaaaat	1
	getteccatagaattgaatcaccgtg	TCGA		cagtgecataggatgaggagaaaaat	1
	igetteecatagaattgaatcacegtg igetteecatagaattgaatcacegtg	AAG		cagtgccataggatgaggagaaaaat cagtgccataggatgaggagaaaaat	- 6
	getteccatagaattgaatcaccgtg	GGG		cagtgccataggatgaggagaaaaat	- i
ggtttttgtaaag	getteecatagaattgaateaccgtg	GC		cagtgecataggatgaggagaaaaat	1
	getteecatagaattgaatcaccgtg	GG	ca	cagtgccataggatgaggagaaaaat	1
	getteecatagaattgaatcaccgtg	CG		cagtgccataggatgaggagaaaaat	- 3
ggttttgtaaaj	getteecatagaattgaatcaccgtg	2) CC GGCG		cagtgccataggatgaggagaaaaat	1
ggttfffefaaa	rgetteecatagaattgaatcaceg(- rgetteecatagaattgaatcaceg(-			cagtgccataggatgaggagaaaaat cagtgccataggatgaggagaaaaat	
	getteecatagaattgaatcacegtg	-, -		gatgaggagaaaaat(-	
Core RAG			niciae de		
	getteecatagaattgaateacegtg getteecatagaattgaateacegtg	GGGA	T ca	cagtgccataggatgaggagaaaaat cagtgccataggatgaggagaaaaat	35
	igetteccatagaattgaatcacegtg igetteccatagaattgaatcacegtg	CCGA	C ca	cagtgccataggatgaggagaaaaat cagtgccataggatgaggagaaaaat	1
ggtttttgtaaaj	getteccatagaattgaatcaccgtg	GC TC	ca	cagtgccataggatgaggagaaaaat	i
	getteccatagaattgaatcaccgtg	GCCA		cagtgecataggatgaggagaaaaat	1
ggtttttgtaaag	getteecatagaattgaatcaccgtg	AGA	ca	cagtgccataggatgaggagaaaaat	1
ggtttttgtaaaj	getteecatagaattgaatcaccgtg	CC		cagtgccataggatgaggagaaaaat	1
	getteccatagaattgaatcaccgtg	GG	ca	cagtgecataggatgaggagaaaaat	
	getteecatagaattgaateacegtg getteecatagaattgaateacegtg	GC T		cagtgecataggatgaggagaaaaat cagtgecataggatgaggagaaaaat	1
	getteecatagaattgaatcaccgtg	ċ		cagtgccataggatgaggagaaaaat	. i
ggtttttgtaaaj	getteecatagaattgaatca(.agtgccataggatgaggagaaaaat(-3	
ggtttttgtaaag	getteecatagaattgaatcaccgtg			gatgaggagaaaaat(-)	13) 1
Dβ2 F		oding N		ding Jβ2.1 RSS	an atta
	lccagttgaatcattgtg (GCCCCC			ding Jβ2.1 RSS GTTA)cacagcagaaaagggctaccaa	gaatte
tttgtaaaageacetet d Type (129+) tttgtaaaageacetet	tecagtigaat cattgtg (GCCCCC /+) tecagtigaat cattgtg	CCAGTCCC)		GTTA)cacagcagaaaagggetaccaa	gaatte
ttigtaaaageaeetet d Type (129+/ ttigtaaaageaeetet ttigtaaaageaeetet	tecagitgaat cattgtg (GCCCCC /+) tecagitgaatcattgtg tecagitgaatcattgtg	CCCAGTCCC)	(Ai	GTTA)cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa	gaatte gaatte
tttgtaaaageaeetet d Type (129+) tttgtaaaageaeetet tttgtaaaageaeetet tttgtaaaageaeetet	tecagitgaateattgtg (GCCCCC /+) tecagitgaateattgtg tecagitgaateattgtg tecagitgaateattgtg	CCAGTCCC)	(A) T	GTTA)cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa	gaatte gaatte gaatte
ttgtasaageacetet d Type (129+, ttgtasaageacetet ttgtasaageacetet ttgtasaageacetet ttgtasaageacetet ttgtasaageacetett	tccagtigaatcattgtg (GCCCCC (+) tccagtigaatcattgtg tccagtigaatcattgtg tccagtigaatcattgtg tccagtigaatcattgtg rcagtigaatcattgtg	CCAGTCCC) CC AGGC GGC1 GGC1 GGC1	(Al	(JTTA) cacagcagaaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa	gaatte gaatte gaatte gaatte gaatte gaatte
ittgtaaaagcacetet d Type (129+, ittgtaaaagcacetet ittgtaaaagcacetet ittgtaaagcacetet ittgtaaagcacetetti ittgtaaagcacetetti ittgtaaagcacetetti	tccagtigaatcattgig (GCCCCC /+) tccagtigaatcattgig tccagtigaatcattgig tccagtigaatcattgig tccagtigaatcattgig tccagtigaatcattgig	CCAGTCCC) CC AGGC GGCI GGCI AGT	(Al	GTTA)cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa	gaatte gaatte gaatte gaatte gaatte gaatte gaatte
ttgtaaagcacetet d Type (129+, ttgtaaagcacetet ttgtaaagcacetet ttgtaaagcacetet ttgtaaagcacetet ttgtaaagcacetett ttgtaaagcacetett ttgtaaagcacetett ttgtaaagcacetett	tccagtiganteattgtg (GCCCCC (+) tcccagtiganteattgtg tcccagtiganteattgtg tccagtiganteattgtg tccagtiganteattgtg tccagtiganteattgtg tccagtiganteattgtg	CCAGTCCC) AGGC GGCI GGCI AGT CCC	(Al	GTTA)cacagcaganangggctaccan cacagcaganangggctaccan cacagcaganangggctaccan cacagcaganangggctaccan cacagcaganangggctaccan cacagcaganangggctaccan cacagcaganangggctaccan cacagcaganangggctaccan	gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte
ttigtaaagcacetet d Type (129+, ttigtaaagcacetet ttigtaaagcacetet ttigtaaagcacetet ttigtaaagcacetet ttigtaaagcacetet ttigtaaagcacetet ttigtaaagcacetet ttigtaaagcacetet	tccagtigaatcattgig (GCCCCC /+) tccagtigaatcattgig tccagtigaatcattgig tccagtigaatcattgig tccagtigaatcattgig tccagtigaatcattgig	CCAGTCCC) CC AGGC GGCI GGCI AGT	(Al	GTTA)cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa cacagcagaaaagggctaccaa	gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte
tttglaaaageacetet d Type (129+, tttglaaaageacetet tttglaaaageacetet tttglaaaageacetet tttglaaaageacetet tttglaaaageacetet tttglaaaageacetet tttglaaaageacetet tttglaaaageacetet tttglaaaageacetet tttglaaaageacetet tttglaaaageacetet	tocagtigaateattgtg tocagtigaateattgtg tocagtigaateattgtg tocagtigaateattgtg tocagtigaateattgtg coagtigaateattgtg tocagtigaateattgtg tocagtigaateattgtg tocagtigaateattgtg tocagtigaateattgtg tocagtigaateattgtg tocagtigaateattgtg	CCAGTCCC) CC AGGC GGCT AGT CCC GCT	(Al	CITTA)cacageaganaagggetaceaa cacageaganaagggetaceaa cacageaganaagggetaceaa cacageaganaagggetaceaa cacageaganaagggetaceaa cacageaganaagggetaceaa cacageaganaagggetaceaa cacageaganaagggetaceaa cacageaganaagggetaceaa cacageaganaagggetaceaa	gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte
Ittgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaaaageaectet ttgtaageaectet ttgtaageaect	tocagtigansentigig (GCCCCC (cagtigansentigig tocagtigansentigig tocagtigansentigig tocagtigansentigig tocagtigansentigig tocagtigansentigig tocagtigansentigig tocagtigansentigig tocagtigansentigig tocagtigansentigig	CC AGGG GGC GGC CC CC CC CC CC CC CC CC CC C	(Al	CITTA) cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa	gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte
ttgtnaageaectet ttgttaageaectet ttgtta	tocagitgaateattgig (GCCCCC (4) tocagitgaateattgig tocagitgaateattgig cocagitgaateattgig cocagitgaateattgig cocagitgaateattgig tocagitgaateattgig tocagitgaateattgig tocagitgaateattgig tocagitgaateattgig tocagitgaateattgig tocagitgaateattgig tocagitgaateattgig tocagitgaateattgig	CCCAGTCCC) CCC AGGC GGCT GGCT CCC GCT AC CCCCAGTCC) TTAAGG CCCCAGTCC) TTAAGG	(Al	CITTA)cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean	gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte
ttgtnanagesectet d Type (129+, ttgtnanagesectet ttgtnanagesectet ttgtnangesectet ttgttat ttgttat ttgttat ttgttat ttgttat ttgttat ttgttat ttgttat ttgttat ttgttat ttgttat ttgttat ttgttat ttgttat	iccogriganicatiggi (GCCCCC (ccogriganicatiggi iccog	CCCAGTCCC) CCC GGC GGC GGC GGC GCC GCC ACC CCCAGTCC) TTAAG CCCAGTCC) TTAAG CCCAGTCC) TTAAG	(Al T AAG G	CITTA)cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean	gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte
ttgtaaageacetet ttgtaaageacetet	tocagitgaatentigig (GCCCCC (+) (coagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig (GCCCCC cocagitgaatentigig (GCCCCC cocagitgaatentigig (GCCCCC	CCAGTCCC) CCC GGC GGC AGT CCC GCT AC CCCAGTCC) THAAG CCCAGTCC) TAAA CCCAGTCC) TAAA CCCAGTCC) TAAA CCCAGTCCC TAT	(Ai T SAAG G G G	CITTA) cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean cacageaganangggetacean	gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte
ttgtanangeacetet ttgtanangeacette ttgtanangeacettet ttgtanangeacetet	tocagitgaatentigig (GCCCCC (+) (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC)	CCAGTCCC) CC AGGC GGC AGGC AGT CCC GCT AC CCCAGTCC) TAAAA CCCAGTCC) TAAAA CCCAGTCC) TAAAA CCCAGTCC) TAAAA CCCAGTCCC GGC AGGC CCCAGTCCC GGC CCCAGTCCC GGC CCCAGTCCC GGC CCCAGTCCC GGC	(Ai T T AAG G G G A	CITTA)cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean geaganangggotacean geaganangggotacean	gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte
ttgtanangeacetet ttgtanangeacette ttgtanangeacettet ttgtanangeacetet	tocagitgaatentigig (GCCCCC (+) (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC) (coagitgaatentigig (CCCCC)	CCAGTCCC) CC AGGC GGC GGC GGC GGC CCAGTCC) CCCAGTCC) CCCAGTCC) CCCAGTCCC)	(Al T - - - - A A A	CITTA)cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean cacageaganangggotacean 	gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte
ttgtanaageioett ttgtanaageacett ttgtan	tocagitgaancentigig (GCCCCC (+) tocagitgaancentigig cocagitgaancentigig cocagitgaancentigig cocagitgaancentigig cocagitgaancentigig cocagitgaancentigig cocagitgaancentigig cocagitgaancentigig cocagitgaancentigig (GCCCCC tocagitgaancentigig (GCCCCC tocagitgaancentigig (GCCCCC tocagitgaancentigig (GCCCCC tocagitgaancentigig (GCCCCC tocagitgaancentigig (GCCCCC tocagitgaancentigig (GCCCCC	CCAGTCCC) CC AGG GGC AGG GGC AGT CCC GCT AC CCAGTCC) TTAAG CCCAGTCC) TTAAG CCCAGTCC) TAAA CCCAGTCCC) GC CCCAGTCCC) GC CCCAGTCCC) GC CCCAGTCCC) TCG	(Al T - - - - A A A	CITTA) cacageaganangggetacean	gaatte gaatte
ttgtanangeöcett ttgtanangescht	iccaginganenting (GCCCC) (ccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaging iccacing iccaging iccacing icca	CCAGTCCC) CCCAGTCCC) CCCAGTCC) CCCAGTCC) CCCAGTCC) CCCAGTCC) CCCAGTCCC CCCAGTCCC) CCCAGTCCC CCCAGTCCC) CCCAGTCCC CCCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCC C	(At T AAAG iAG GG A A	CITTA)cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa 	gaatic ga
ttgtanagedeett ttgtanagedett ttgta	tocagitgaatentigig (GCCCCC (r) (coagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig cocagitgaatentigig (GCCCCC cocagitgaatentigig (GCCCCC cocagitgaatentigig (GCCCCC cocagitgaatentigig (GCCCCC cocagitgaatentigig (GCCCCC cocagitgaatentigig (GCCCCC cocagitgaatentigig (GCCCCC cocagitgaatentigig (GCCCCC cocagitgaatentigig (GCCCCC) cocagitgaatentigig (GCCCCC) cocagitgaatentigig (GCCCCC) cocagitgaatentigig (GCCCCC)	CCAGTCCC) CCCAGTCCC) GGCI GGCI ACT CCCAGTCC) TTAAG CCCAGTCC) TAAA CCCAGTCC) TAAA CCCAGTCC) TAAA CCCAGTCC) TAAA CCCAGTCC) TAAA CCCAGTCCC) TAA CCCAGTCCC) GCC CCCAGTCCC) TAA	(Al AAAG AAG G G G G G A A	CITTÀ)cacageaganangggetacean cacageaganangggetacean	gaatic gaatic
It glanagescett I Type (128-, type 128-, type 128-,	iccaginganenting (GCCCC) (ccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaginganenting iccaging iccacing iccaging iccacing icca	CCAGTCCC) CCCAGTCCC) CCCAGTCC) CCCAGTCC) CCCAGTCC) CCCAGTCC) CCCAGTCCC CCCAGTCCC) CCCAGTCCC CCCAGTCCC) CCCAGTCCC CCCAGTCCC) CCCAGTCCC CCCAGTCCC) CCCAGTCCCC CCCAGTCCC CCCAGTCCC CCCAGTCCCC CCCAGTCCC CCCAGTCCCC CCCAGTCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCC CCCAGTCCCCCCCCCC	(At T AAAG AG GG A A A	CITTA)cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa cacageagaaaagggotaceaa 	gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte gaatte
ttgtanagedeett ttgtanagedett ttgta	iccaginganicatigg (GCCCC) (4) iccaginganicatigg iccaging iccaging iccaging iccanicatigg iccaging iccanicatigg iccaging iccanicatigg iccaging iccanicatigg iccaging iccanicatigg iccaging iccanicatigg iccaging iccanicatigg iccaging iccanicatigg iccaging iccanicatigg	CCAGTCCC) CC AGG GGC AGG GGC AGT CCC GGC AC CCAGTCC) TAAAG CCCAGTCC) TAAAG CCCAGTCC) TAAAC CCCAGTCC) TAAAG CCCAGTCC) GGC CCCAGTCCC GGC CCCAGTCCC TAC CCCAGTCCC TAC CCCAGTCCC TAC CCCAGTCCC TAC CCCAGTCC GGC CCCAGTCCC TAC CCCAGTCC GGC CCCAGTCC GGC CCCAGTCC GGC CCCAGTCC GGC CCCAGTCC GGC CCCAGTCC GGC CCCAGTCC GGC CCCAGTCC GCC CCCAGTCC GCC CCCAGTCC CCCAGCCCC TAC CCCAGTCC GGC CCCAGTCCC TAC	(Al AAAG AAG G G G G G G G G G G G G G G	CITTA)cacageagaaaagggotaceaa cacageagaaagggotaceaa cacageagaaagggotaceaa cacageagaaagggotaceaa cacageagaaagggotaceaa cacageagaaagggotaceaa cacageagaaagggotaceaa cacageagaaagggotaceaa cacageagaaagggotaceaa	gaatic ga
ttgtanagedeett ttgtanagedett t	iccaginganicating (GCCCC) (4) iccaginganicating constraints iccaginganicating constraints iccaginganicating constraints iccaginganicating constraints iccaginganicating constraints iccaginganicating constraints iccaginganicating (GCCCC) iccaginganicating (GCCCC) iccaginganicating (GCCCC) iccaginganicating (GCCCC) iccaginganicating (GCCC) iccaginganicating (GCC) iccaginganicating (GCCC) iccaginganicating (GCC) iccagi	CCAGTCCC) CC AGG GGC1 GGC1 AGT GGC1 AGT CCCAGTCC) THAAG CCCAGTCC) TAAAC CCCAGTCC) TAAAC CCCAGTCC) TAAAC CCCAGTCC) TAAAC CCCAGTCC) TAAAC CCCAGTCC) GGC CCCAGTCCC) TAA CCCAGTCC) GGC CCCAGTCCC) TAA CCCAGTCC) GGC CCCAGTCCC) TAA CCCAGTCCC) TAA CCCAGTCCCC TAA CCCAGTCCCC TAA CCCAGTCCCC TAA CCCAGTCCCC TAA CCCAGTCCCC TAA CCCAGTCCCC TAA CCCAGTCCCC TAA CCCAGTCCCC TAA CCCAGTCCCCC TAA	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacageaganangggetacean cacageaganangggetacean	gaatic ga
ttgtanagecert ttgtanagecett ttgtan	iccaginganenting (GCCCC) (caginganenting for the second s	CCAGTCCC) CCCAGTCCC) CCCAGTCC] CCCAGTC] CCCAGTC]	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ/cacageaganangggetacean cacageagangggetacean	gaatic gaatic
	iccagitgaancentigig (GCCCCC (a) (ccagitgaancentigig (ccagitgaancentigig (ccagitgaancentigig (ccagitgaancentigig (ccagitgaancentigig (ccagitgaancentigig (ccagitgaancentigig (ccagitgaancentigig (CCCCCC (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC) (ccagitgaancentigig (CCCCCC)	CCAGTCCC) CCCAGTCCC) CCCAGTCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC] CCCAGTCC] CCCAGTCCC] CCCAGTCCC CCCAGTCC] CCCCAGTCC] CCCCAGTCC] CCCCAGTCC] CCCCAGTCC] CCCCAGTCCC CCCCAGTCC] CCCCCCCCC CCCCCC CCCCCCCCC CCCCCC	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacageaganangggetacean cacageaganangggetacean	gaatic ga
ttglanangeocett ttglanangeocett	iccaginganenting (GCCCC) (caginganenting for the second s	CCAGTCCC) CCCAGTCCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC] CCCAGTCCCCCCCC] CCCCAGTCCC] CCCCAGTCCC] CCCAGTCCC] CCCAGTCCC] CCCAGTCCC]	(At T AAAG AAG G G G G A A A S G S G S S S S	CTTA)cacageagaaaagggctaceaa cageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaaagggctaceaa cacageagaaagggctaceaa cacageagaaagggctaceaa cacageagaaagggctaceaa cacageagaaagggctaceaa cacageagaaagggctaceaa cacageagaaagggctaceaa	gaatic gaatic
ttgtanagecett ttgtanagecett	iccaginganenting (GCCCC) (c) (c) (c) (c) (c) (c) (c) (c	CCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCCC) CCCAGTCCCC CCCAGTCCCC) CCCAGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacageaganangggetacean cacageaganangggetacean	gaatic gaatic
	iccaginganenting (GCCCC) (r) iccaginganenting iccaging	CCAGTCCC) CCAGTCCC) CCAGTCC) CCCAGTCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) GCC CCCC CCC CCC CCC CCC C	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacageaganangggetacean cacageaganangggetacean	gaatte gaatte
ttgtanagecett ttgtanagecett	iccaginganenting (GCCCC) (c) (c) (c) (c) (c) (c) (c) (c	CCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC CCCAGTCCC) CCCAGTCCCC CCCAGTCCC CCCAGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacageaganangggetacean cacageaganangggetacean	gaatic gaatic
	iccagingancentiging (GCCCCC (r) (ccagingancentiging (ccagingancentiging (ccagingancentiging (ccagingancentiging (ccagingancentiging (ccagingancentiging (ccagingancentiging (ccagingancentiging (ccagingancentiging (CCCCCC)	CCAGTCCC) CCCAGTCCC) CCCAGTCC) CCCAGTCCC) CCCAGTCCC) GCCCAGTCCC) GCCCCGCC GCCCCG GCCCAGTCCC) GCCCCGCC GCCCCGCC GCCCCG GCCCGTC GCCCCG GCCCGTCC GCCCCG GCCCGTC GCCCCC GCCCCG GCCCGTC GCCCCG GCCCCG GCCCCG GCCCCG GCCCCG GCCCCG GCCCCG GCCCC GCCCCG GCCCC GCCCC GCCCC GCCCC GCCCC GCCCC GCCC GCCCC GCCC GCCCC GCC GCCC GCC	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacageaganangggetucean cacageaganangggetucean	gaatic gaatic
It Jype (12) I Jy	iccaginganentinging (GCCCCC) (caginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccacinging iccacingin	CCAGTCCC) CCCAGTCCC) CCCAGTCC) CCCAGTCCC) CCCAGTCCC CCCAGTCCC CCCAGTCC CCCAGTCC CCCAGTCC CCCAGTCC CCCAGTCC CCCAGTCC CCCAGTCC CCCAGTC CCCAGTCC CCCAGTC CCCAGTC CCCAGTC CCCCAGTC CCCCA	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacageaganangggetacean eacageaganangggetacean	gaatic gaatic
ttgtanagecett ttgtanagecett	iccaginganentigg (GCCCCC (caginganentigg (caginganentigg (caginganentigg (caginganentigg (caginganentigg (caginganentigg (caginganentigg (caginganentigg (CCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCC (caginganentigg (GCCCCCCCCCCCCCCCCC) (Caginganentigg (GCCCCCC (caginganentigg (GCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	CCAGTCCC) CCCAGTCCC) CCCAGTCC) GCC CCCAGTCC) GCCCAGTCC) GCCCAGTCC GCCCAGTC GCCCC GCC G	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacageaganangggetucean cacageaganangggetucean	gaatic gaatic
ttgtanagecett ttgtanagecett	iccaginganentigg (GCCCCC (caginganentigg iccaging iccaginganentigg iccaging	CCAGTCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCCCCC CCCAGTCCC) CCCCCCCCC CCCCCCCCCCC CCCCCCCCCC	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacageaganangggetacean cacageagananggetacean cacageagananggetacean cacageagananggetacean cacageagananggetacean cacageaganangganangganangganangganangganang	gaatic gaatic
Ittgaaaagcoccit Ittga	iccaginganenting (GCCCC) (c) (c) (c) (c) (c) (c) (c) (c	CCAGTCCC) CCCAGTCCC) CCCAGTCC CCCAGTCC) CCCAGTCC CCCAGTCC) CCCAGTCC CCCAGTCCC) CCCAGTCC CCCAGTCCC) CCCAGTCC CCCAGTCCC) CCCAGTCCC CCCAGTCC CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacageaganangggetacean cacageaganangggetacean	gaalic gaatic
titgtaaageceett titgtaaageceet	iccaginganentigg (GCCCCC (caginganentigg (caginganentigg (caginganentigg (caginganentigg (caginganentigg (caginganentigg (caginganentigg (caginganentigg (CCCCCC (caginganentigg (CCCCCC) (caginganentigg (CCCCCC) (caginganentigg (CCCCC)	CCAGTCCC) CCCAGTCCC) CCCAGTCC) GCC CCCAGTCCC) GCC CCCAGTCC GCC CCCAGTCC GCC CCCAGTCC GCC CCCAGTC GCC CCC C	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ/cacageaganangggetacean cageaganangggetacean cacageaganangggtacean cacageaganangggtacean cacageaganangggtacean cacageaganangggtacean cacageaganangggtatean cacageagananggtatean cacageagananggtatacean	gaalic gaatic
ttgtanagecett tt	iccaginganentinging (GCCCCC) (caginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccaginganentinging iccacinging iccacing	CCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC CCCAGTCCC) CCCAGTCCC CCCAGTCCC) CCCAGTCCC CCCAGTCCC CCCCGCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCAGTCCC CCCCGCC CCCAGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ/cacageaganangggetacean cacageaganangggetacean	gaalic ga
titgtaaagcecttittitgtaaagcecttittitgtaaagcecttittitgtaaagcecttittitgtaaagcecttittitgtaaagcecttittitgtaaagcecttittitgtaaagcecttittitgtaaagcecttittitgtaaagcecttittitgtaaagcecttittittittaaagcecttittittittaaagcecttittittittaaagcecttittittaaagcecttittittittaaagcecttittittaaagcecttittittittaaagcecttittittittaaagcecttittittittaaagcecttittittittaaagcecttittittittaaagcecttittittittaaagcecttittittittaaagcecttittittittittittittittittittittittitti	iccaginganentigg (GCCCCC (caginganentigg iccaging iccaginganentigg iccaginganentigg iccaginganentigg iccaginganentigg iccaging iccacing iccaging iccaging iccaging iccacing iccaging iccac	CCAGTCC) CCCAGTCCC) CCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	(At T AAAG AAG G G G G A A A S G S G S S S S	CTTTA)cacageagaaaagggetaceaa cacageagaaagggetaceaa cacageagaaagggaaaggetaceaa cacageagaaagggetaceaa cacageagaaagggetaceaa cacageagaa	gaalic gaalic
ttgtanagescett ttgtan	iccaginganenting (GCCCC) (caginganenting (GCCCC) (caginganenting (GCCCC) (caginganenting (GCCC) (caginganenting (GCCC) (c	CCAGTCC) CCCCCCCCCCCCC CCCCCCCCCCCCCCCCCCC	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ/cacageaganangggetacean cacageaganangggetacean	gaatic gaatic
titgtanagecett titgtanagecett	iccaginganenting (GCCCC) (c) (c) (c) (c) (c) (c) (c) (c	CCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCCC CCCAGTCCC) CCCAGTCCCC CCCAGTCCCC) CCCAGTCCCC CCCAGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	(At T AAAG AAG G G G G A A A S G S G S S S S	CITTÀ)cacagenganangggetacena enengengananggge	gaatic gaatic
titgtanagecett titgtanagecett	iccaginganentigg (GCCCCC (caginganentigg iccaging iccaginganentigg iccacing iccacinganentigg iccacin	CCAGTCCC) CCCAGTCCC] CCCCAGTCCC] CCCCCCCCC] CCCCCCCCCCCC] CCCCCCCC	(At T AAAG G G G G A A A	CTTTA)cacageagaaaagggotaceaa cacageagaaagggotaceaa cacageagaaagggo	gaatic gaatic
Ittigaaagaceetti Ittigaaagacetti	iccaginganenting (GCCCC) (c) (c) (c) (c) (c) (c) (c) (c	CCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCC) CCCAGTCCCC CCCAGTCCC) CCCAGTCCCC CCCAGTCCCC) CCCAGTCCCC CCCAGTCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC	(At T T AAG G G G A A A A A	CITTÀ)cacagenganangggetacena enengengananggge	gaatic gaatic

nucleotide additions at the VD junction of V β 14-D β -J β 1.1 CJ was also larger in core RAG2 samples than WT (+3.87 and +2.72 bp, respectively). On the other hand, no significant differences were observed in the frequency or degree of modifications of V β or VD junctions for V β 10-D β -J β 2.1 CJ between core RAG1, core RAG2 and WT samples. Instead, the length of $D\beta$ coding sequence was frequently

re RAG2 (2c/2c)		
ittigtaaaagcaccicitccagitgaatcattgig	GG cacag	engaaaagggctaccangnatte
tttigtaaaageacetetteeagttgaateattgtg		cagaaaagggctaccaagaatte
ttttgtaaaagcacetetteeagttgaateattgtg (GCCCCC ttttgtaaaagcacetetteeagttgaateattgtg (GCCCCC	CCAGICCC)CCATACTAGCTAAT cacag	cagaaaagggctaccaagaatte
ittigtaaaagcaccicticcagitgaatcattgig (GCCCCC	CCAGTCCC) GTCCC cacag	cagaaaagggctaccaagaatte cagaaaagggctaccaagaatte
ttttgtaaaagcacctettecagttgaateattgtg (GCCCCC	CCAGTC) TCTTG cacas	cagaaaagggctaccaagaatte
tttigtaaaagcacctetteeagttgaateattgig (GCCCCC	CCAGTCC) AGG	agaaaagggctaccaagaatte(-6)
ttttgtaaaagcacctcttccagttgaatcattgtg (GCCCCC	CCAGTCCC) GGG cacag	cagaaaagggctaccaagaatte
ttttgtaaaagcacctetteeagttgaateattgtg (GCCCCC	CCAGTCCC) GGGg	cagaaaagggctaccaagaatte(-4)
tttigtaaaagcacctcttccagttgaatcattgtg (GCCCCC) TCG cacag	cagaaaagggctaccaagaatte
ittigtaaaagcacciciiccagiigaaicatigig (GCCCCC		cagaaaagggclaccaagaatte
tttgtaaaagcacctcttccagttgaatcattgtg (GCCCCC	CCAGTCCC) Gg	cagaaaagggctaccaagaattc(-4)
tttgtaaaagcacctetteeagttgaateattgtg (GCCCCC tttgtaaaagcacctetteeagttgaateattgtg (GCCCCC	CCAGTC)	aaaagggctaccaagaatte(-8)
ittigtaaaageaceceteteeagtigaateatigtg (GCCCCC	(CAG)ag	cagaaaagggctaccaagaatte(-3) cagaaaagggctaccaagaatte
ittigtaaaagcaccicticcagitgaatcattgig		aaaagggctaccaagaatte(-8)
tttfgtaaaagcacctcttccagttgaat (-7)		cagaaaagggclaccaagaatte
tttigtasaageacetetteeagttgaateattgtg	TGGGG cacas	cagaaaagggctaccaagaatte
Ittigtaaaagcaccicticcagtigaateatigig	GGTTT cacag	cagaaaagggctaccaagaatte
ittigtaaaagcaccicitecagitgaatcatigig	ACTTA cacag	cagaaaagggctaccaagaatte
ittigiaaaageaccicticcagtigaateatigig		cagaaaagggclaccaagaatte
tttigtaaaagcaccicticcagttgaatcattgig		cagaaaagggctaccaagaatte
tttgtaaaagcacctettecagttgaateattgtg		cagaaaagggctaccaagaatte
ittigtaaaagcaccicticcagitgaalcattgtg		cagaaaagggctaccaagaatte
ittigtaaaagcaccicticcagitgaateatigtg		cagaaaagggctaccaagaatte
ittigtaaaagcaccicticcagttgaatcattgtg		cagaaaagggctaccaagaatte
Ittigtaaaageaccicticcagtigaateattgtg		cagaaaagggctaccaagaatte
ittigtaaaagcaccicticcagitgaateattgig ittigtaaaagcaccicticcagitgaateattgig		cagaaaagggctaccaagaatte
ittigtaaaagcaccicticcagtigaatcattgig		cagaaaagggetaccaagaatte cagaaaagggetaccaagaatte
ittigtaaaagcacetetteeagttgaateattgtg		cagaaaagggctaccaagaatte
C Dδ2-Jδ1 recombinati	on	
Dô2 RSS	coding N coding Jδ1	RSS n
agtttttgtacaggteectatggggtttgcaacactgtg (
0 0 00 000 0 00		•
Wild Type (129+/+)		
agttitigtacaggtocotatggggttigeaacactgtg	cacagefactgag	gccatteeaaaaaee 41
agtittigtacaggicectalggggtitgcaacacigig(CTCGTA) - cacagetactgag	gccatteeaaaaaee 1
agtititigtacaggteectatggggtttgcaacactgtg		ccatteeaaaaaee 1
agtttttgtacaggtccctatggggtttgcaacactgtg		gccattecaaaaaee 1
agtttttgtacaggtccctatggggtttgcaacactgtg		gecattecaaaaace 1
agtttttgtacaggtccctatggggtttgcaacactgtg		gccattecanaaaec 1
agtittigtacaggtccctatggggtttgcaacactgtg		gccatteeaaaaaee 1
agtttttgfacaggteectatggggtttgcaacactgtg		gecatteeaaaaaee 1
Core RAG1 (1c/1c)		
agtititgtacaggtecetatggggtitgeaacactgtg	cacagefactore	
agttittgtacaggteectatggggttigeaacactgtg		
agttittgtacaggteectatggggtttgeaacactgtg		ccatteeaaaaaee 26 cccatteeaaaaaee 2
	CC cacagctactgag	ccattecaaaaacc 2
	CC cacagetactgag CCTCCT cacagetactgag	gccattecaaaaace 2 gccattecaaaaace 2
agtttttgtacaggteeetatggggtttgeaacactgtg	CC cacagetactgag CCTCCT cacagetactgag AGGCA cacagetactgag	gecattecaaaaaee 2 gecattecaaaaaee 2 gecattecaaaaaee 1
agtititgtacaggteeetatggggtitgeaacactgtg agtititgtacaggteeetatggggtitgeaacactgtg	CC cacagetactagg CCTCCT cacagetactagg AGGCA cacagetactagg AGGACT cacagetactagg	continue and a contin
agttitigtacaggteeetatggggttgeaacactgtg agttitigtacaggteeetatggggttgeaacactgtg agttitigtacaggteeetatggggttigeaacactgtg	CC cacagetactgag CCTCCT cacagetactgag AGGCA cacagetactgag AGGACT cacagetactgag GGGGG cacagetactgag	contecaaaaace 2 contecaaaaace 2 contecaaaaace 1 contecaaaaace 1 contecaaaaace 1
agtittigtacaggteectatggggtttgeaacactgtg agtittigtacaggteectatggggtttgeaacactgtg agtittigtacaggteectatggggtttgeaacactgtg agtittigtacaggteectatggggtttgeaacactgtg	CC cacagetactaga CCTCCT cacagetactaga AGGCA cacagetactaga AGGACT cacagetactaga GGGGG cacagetactaga GGGCC cacagetactaga	gecattecaaaaace 2 gecattecaaaaace 2 gecattecaaaaace 1 gecattecaaaaace 1 gecattecaaaaace 1 gecattecaaaaace 1 gecattecaaaaace 1 gecattecaaaaace 1
agtittigtacaggteectalggggtitgeaacactgig agtittigtacaggteectalggggtitgeaacactgig agtittigtacaggteectalggggtitgeaacactgig agtittigtacaggteectalggggtitgeaacactgig agtittigtacaggteectalggggtitgeaacactgig	CC cacagetacipag CCTCCT cacagetacipag AGGCA cacagetacipag AGGACT cacagetacipag GGGGG cacagetacipag GGGCC cacagetacipag GGCA cacagetacipag	gecattecaaaaace 2 gecattecaaaaace 2 gecattecaaaaace 1 gecattecaaaaace 1 gecattecaaaaace 1 gecattecaaaaace 1 gecattecaaaaace 1 gecattecaaaaace 1
agttttigtacaggtoctatagggtitigenancatgig agtttilgtacaggtoctatagggtitigenacatgig agtttilgtacaggtoctatagggtitigenacatgig agttttigtacaggtoctatagggtitigenacatgig agtttilgtacaggtoctatagggtitigenacatgig agtttilgtacaggtoctatagggtitigenacatgig	CC cnaegatricega CCTCCT cnaegatricega AGGCA cnaegatricega GGGGG cnaegatricega GGGCC cnaegatricega GGCA cnaegatricega CCCC cnaegatricega	gecattecaaaaaee 2 gecattecaaaaaee 1
agttttigtacaggtoctatgggtittgcanacatgig agtttigtacaggtoctatgggtittgcanacatgig agtttigtacaggtoctatggggtittgcanacatgig agttttigtacaggtoctatggggtittgcanacatgig agttttigtacaggtoctatggggtittgcanacatgig agttttigtacaggtoctatggggtittgcanacatgig agttttigtacaggtoctatggggtittgcanacatgig agttttigtacaggtoctatggggtittgcanacatgig	CC cacagetaciega CCTCCT cacagetaciega AGGCA cacagetaciega GGGGGG cacagetaciega GGGCC cacagetaciega GGCA cacagetaciega GGCA cacagetaciega CCCC cacagetaciega CCCC cacagetaciega	gecattecananace 2 gecattecananace 2 gecattecananace 1 gecattecananace 1 gecattecananace 1 gecattecananace 1 gecattecananace 1 gecattecananace 1 gecattecananace 1
agttttfgtacagtocctatggggtttgcaaactigfg agttttfgtacagtocctatggggtttgcaactigfg agttttfgtacagtocctatggggtttgcaacatigfg agtttttfgtacagtocctatggggtttgcaactigfg agttttfgtacagtocctatggggtttgcaacatigfg agttttfgtacagtocctatggggtttgcaacatigfg agttttfgtacagtocctatggggtttgcaacatigfg agttttfgtacagtocctatggggtttgcaacatigfg	CC cnaegatrictaga CCTCCT cnaegatrictaga AGGCA cnaegatrictaga GGGGG cnaegatrictaga GGGCC cnaegatrictaga CCCC cnaegatrictaga CCCC cnaegatrictaga CCCC cnaegatrictaga GGCA cnaegatrictaga CCCC cnaegatrictaga	gecattecennance 2 gecattecennance 1
agtttiigiacagiocotaggggttigeanacatgig agtttiigiacagiocotaggggttigeanacatgig agtttiigiacagiocotaggggttigeanacatgig agtttiigiacagiocotaggggttigeanacatgig agttitigiacagiocotaggggttigeanacatgig agttitigiacagiocotaggggttigeanacatgig agttitigiacagiocotaggggttigeanacatgig agttitigiacagiocotaggggttigeanacatgig agttitigiacagiocotaggggttigeanacatgig agttitigiacagiocotaggggttigeanacatgig	CC cacagedacipag CCTCCT cacagedacipag AGGCA cacagedacipag GGGGGG cacagedacipag GGGGCC cacagedacipag GGCA cacagedacipag CCCC cacagedacipag CCCC cacagedacipag CCCC cacagedacipag GACT cacagedacipag TAAA cacagedacipag	2ccattecannance 2 2ccattecannance 2 2ccattecannance 1 2ccattecannance 1 2ccattecannance 1 2ccattecannance 1 2ccattecannance 1 2ccattecannance 1 2ccattecannance 1 2ccattecannance 1 2ccattecannance 1
agttttfgtacagtocctatggggtttgcaaactigfg agttttfgtacagtocctatggggtttgcaactigfg agttttfgtacagtocctatggggtttgcaacatigfg agtttttfgtacagtocctatggggtttgcaactigfg agttttfgtacagtocctatggggtttgcaacatigfg agttttfgtacagtocctatggggtttgcaacatigfg agttttfgtacagtocctatggggtttgcaacatigfg agttttfgtacagtocctatggggtttgcaacatigfg	CC cacagedacipag CCTCCT cacagedacipag AGGCA cacagedacipag GGGGGG cacagedacipag GGGGCC cacagedacipag GGCA cacagedacipag CCCC cacagedacipag CCCC cacagedacipag GACT cacagedacipag TAAA cacagedacipag CGAT cacagedacipag CGAT cacagedacipag CGAT cacagedacipag CTCA cacagedacipag	2ccatteeannance 2 2ccatteeannance 2 2ccatteeannance 1 2ccatteeannance 1 2ccatteeannance 1 2ccatteeannance 1 2ccatteeannance 1 2ccatteeannance 1 2ccatteeannance 1 2ccatteeannance 1 2ccatteeannance 1
attittigiacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi attittigacagiocontagggittiganaeetigi	CC cacagetacipag CCTCCT cacagetacipag AGGCA cacagetacipag AGGCG cacagetacipag GGGGG cacagetacipag GGGCC cacagetacipag CCCC cacagetacipag CCCC cacagetacipag CCCC cacagetacipag CCCC cacagetacipag CACT cacagetacipag TAAA cacagetacipag CCAT cacagetacipag CCAT cacagetacipag CCAT cacagetacipag	geomterannance 2 geomterannance 2 geomterannance 1 geomterannance 1
ettittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi	CC cnaegatricipag CCTCCT cnaegatricipag AGGCA cnaegatricipag GGGGG cnaegatricipag GGGGCC cnaegatricipag GGCA cnaegatricipag CCCC cnaegatricipag GCCA cnaegatricipag GCCC cnaegatricipag GCCC cnaegatricipag GACT cnaegatricipag CCTC cnaegatricipag GACT cnaegatricipag CAAT cnaegatricipag AAT cnaegatricipag AAT cnaegatricipag	ocontreannance 2 geontreannance 2 geontreannance 1 geontreannance 1
ettittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi aptittigiacagiocottagggttigueneetigi	CC cacagedaciegas CCTCCT cacagedaciegas AGGCA cacagedaciegas GGGGGG cacagedaciegas GGGGGG cacagedaciegas GGGCA cacagedaciegas CCCC cacagedaciegas CCCC cacagedaciegas CCCC cacagedaciegas CCCC cacagedaciegas CCTC cacagedaciegas CCAT cacagedaciegas CTCA cacagedaciegas AAT cacagedaciegas AT cacagedaciegas CCC cacagedaciegas CCC cacagedaciegas CCC cacagedaciegas CCC cacagedaciegas CCC cacagedaciegas CCC cacagedaciegas CCC cacagedaciegas CCC cacagedaciegas CCC cacagedaciegas	ocontreannance 2 geontreannance 2 geontreannance 1 geontreannance 1
attittigiacagiocotaggggttiganaedtig attittigacagiocotaggggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig attittigacagiocotagggttiganaedtig	CC cnaegatrictaga CCTCCT cnaegatrictaga AGGCA cnaegatrictaga GGGGG cnaegatrictaga GGGGCC cnaegatrictaga GGCA cnaegatrictaga CCCC cnaegatrictaga GCCA cnaegatrictaga CCCC cnaegatrictaga CCCC cnaegatrictaga CCCC cnaegatrictaga CCCC cnaegatrictaga CCAT cnaegatrictaga CCAT cnaegatrictaga CCAT cnaegatrictaga CCAT cnaegatrictaga CCCC cnaegatrictaga CCCC cnaegatrictaga CCCC cnaegatrictaga CCC cnaegatrictaga CCC cnaegatrictaga CCC cnaegatrictaga	geomterannance 2 geomterannance 2 geomterannance 1 geomterannance 1
ettiittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi aptittijtavagiocottagggttijouwachigi	CC cacagedactaga CCTCCT cacagedactaga AGGCA cacagedactaga GGGGGG cacagedactaga GGGGCC cacagedactaga GGGCA cacagedactaga CCCC cacagedactaga GACT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCA cacagedactaga CCA cacagedactaga CCA cacagedactaga CCA cacagedactaga CCA cacagedactaga CCC cacagedactaga	ocontreannance 2 geontreannance 2 geontreannance 1 geontreannance 1
attiffigia.cgiloccitaggggffigianaecitgig attiffigia.cgiloccitagggffigianaecitgig attiffigia.cgiloccitagggffigianaecitgig attiffigia.cgiloccitagggffigianaecitgig attiffigia.cgiloccitagggffigianaecitgig attiffigia.cgiloccitagggffigianaecitgig attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi attiffigia.cgiloccitagggffigianaecitgi	CC cacagedactaga CCTCCT cacagedactaga AGGCA cacagedactaga GGGGGG cacagedactaga GGGGCG cacagedactaga GGCA cacagedactaga CCCC cacagedactaga CCCC cacagedactaga CCCC cacagedactaga GACT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCAT cacagedactaga CCCC cacagedactaga CCC cacagedactaga CCC cacagedactaga	geomterannance 2 geomterannance 1 geomterannance 1
attittigiacagicoctatggggttiganaetigg attittigiacagicoctatggggttiganaetigg attittigiacagicoctatggggttiganaetigg attittigiacagicoctatggggttiganaetigg attittigiacagicoctatgggttiganaetigg attittigiacagicoctatgggttiganaetigg attittigiacagicoctatgggttiganaetigg attittigiacagicoctatgggttiganaetigg attittigiacagicoctatgggttiganaetigg attittigiacagicoctatgggttiganaetigg attittigacagicoctatgggttiganaetigg attittigacagicoctatgggttiganaetigg attittigacagicoctatgggttiganaetigg attittigacagicoctatgggttiganaetigg attittigacagicoctatgggttiganaetigg attittigacagicoctatgggttiganaetigg attittigacagicoctatggggttiganaetigg attittigacagicoctatggggttiganaetigg attittigacagicoctatggggttiganaetigg attittigacagicoctatggggttiganaetigg attittigacagicoctatggggttiganaetigg attittigacagicoctatggggttiganaetigg	CC cacagetaicteag CCTCCT cacagetaicteag AGGCA cacagetaicteag GGGGGG cacagetaicteag GGGGGC cacagetaicteag GGCA cacagetaicteag CCCC cacagetaicteag CCCC cacagetaicteag CCCC cacagetaicteag CCCC cacagetaicteag CCCC cacagetaicteag CCCC cacagetaicteag CCCC cacagetaicteag CCCC cacagetaicteag CCCC cacagetaicteag CCC cacagetaicteag	ocontreannance 2 gocantreannance 1 gocantreannance 1
attittigiacagiocotaggggttiganaedigg attittigacagiocotaggggttiganaedigg attittigacagiocotagggttiganaedigg attittigacagiocotaggggttiganaedigg attittigacagiocotaggggttiganaedigg attittigacagiocotaggggttiganaedigg attittigacagiocotaggggttiganaedigg attittigacagiocotaggggttiganaedigg attittigacagiocotaggggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotagggttiganaedig attittigacagiocotaggttiganaedig attittigacagiocotaggttiganaedig attittigacagiocotaggttiganaedig attittigacagiocotagg	CC cacagetactega CCTCCT cacagetactega AGGCA cacagetactega GGGGGG cacagetactega GGGGCC cacagetactega CCCC cacagetactega CCC cacagetactega	geontreannance 2 geontreannance 1 geontreannance 1
attiffigia.agiccetatgggffigianaestigi attiffigia.agiccetatgggffigianaestigi	CC catagetaictaga CCTCCT catagetaictaga AGGCA catagetaictaga GGGGGG catagetaictaga GGGGGC catagetaictaga GGCA catagetaictaga CCCC catagetaictaga CCCC catagetaictaga CCCC catagetaictaga CCCC catagetaictaga CCCC catagetaictaga CCCC catagetaictaga CCCC catagetaictaga CCC catagetaictaga	geomterannance 2 geomterannance 1 geomterannance 2
attittigiacagiocotaggggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotaggggttigeanaetigi attittigiacagiocotaggggttigeanaetigi attittigiacagiocotaggggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotagggttigeanaetigi attittigiacagiocotaggggttigeanaetigi attittigiacagiocotaggggttigeanaetigi attittigiacagiocotaggggttigeanaetigi	CC cacagetactega CCTCCT cacagetactega AGGCA cacagetactega GGGGGG cacagetactega GGGGCC cacagetactega CCCC cacagetactega CCC cacagetactega CCC cacagetactega CCC cacagetactega CCC cacagetactega CCC cacagetactega	geontreannance 2 geontreannance 2 geontreannance 1 geontreannance 2 geontreannance 2 geontreannance 2 geontreannance 2
attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatgggttiganaetigi attittigiacagicoctatggggttiganaetigi attittigiacagicoctatggggttiganaetigi attittigiacagicoctatggggttiganaetigi attittigiacagicoctatggggttiganaetigi attittigiacagicoctatggggttiganaetigi attittigiacagicoctatggggttiganaetigi attittigiacagicoctatggggttiganaetigi attittigiacagicoctatggggttiganaetigi attittigiacagicoctatggggttiganaetigi attittiganagicoctatggggttiganaetigi attittiganagicoctatggggttiganaetigi attittiganagicoctatggggttiganaetigi attittiganagicoctatggggttiganaetigi attittiganagicoctatggggttiganaetigi attittiganagicoctatgggttigattiganaetigi	CC catagetaictaga CCTCCT catagetaictaga AGGCA catagetaictaga GGGGGG catagetaictaga GGGGGG catagetaictaga GGGCA catagetaictaga CCCC catagetaictaga CCTC catagetaictaga CCCC catagetaictaga CCCC catagetaictaga CCCC catagetaictaga CCCC catagetaictaga CCCC catagetaictaga CCC catagetaictaga	geomterannance 2 geomterannance 1 geomterannance 2 geomterannance 2 geomterannance 2 geomterannance 2 geomterannance 2 geomterannance 2 geomterannance 2 geomterannance 2 geomterannance 2
attittigiacagiocotaggggttiganaedtigi attittigiacagiocotagggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi	CC cacagetaicteag CCTCCT cacagetaicteag AGGCA cacagetaicteag GGGGGG cacagetaicteag GGGGCC cacagetaicteag GGCA cacagetaicteag GGCA cacagetaicteag GGCA cacagetaicteag GCCC cacagetaicteag CCCC cacagetaicteag CCA cacagetaicteag CCC cacagetaicteag CCA cacagetaicteag CCC cacagetaicteag	geontecannance 2 geontecannance 2 geontecannance 1 geontecannance 2 geontecannance 2 geontecannance 2 geontecannance 2 geontecannance 2 geontecannance 2 geontecannance 2
attittigiacagiocotaggggttiganaedigi attittigiacagiocotaggggttiganaedigi attittigiacagiocotaggggttiganaedigi attittigiacagiocotaggggttiganaedigi attittigiacagiocotagggttiganaedigi atti	CC c cacagetacigag CCTCCT cacagetacigag AGGCA cacagetacigag GGGGG cacagetacigag GGGGG cacagetacigag GGCA cacagetacigag CCCC cacagetacigag CCCC cacagetacigag CCCC cacagetacigag CCCC cacagetacigag CCCC cacagetacigag CCCC cacagetacigag CCCC cacagetacigag CCC cacagetacigag CCC cacagetacigag CCC cacagetacigag CCC cacagetacigag CCC cacagetacigag CCC cacagetacigag CC cacagetacigag CC cacagetacigag CC cacagetacigag CC cacagetacigag CC cacagetacigag CCC cacagetacigag CCACA cacagetacigag CCAACA cacagetacigag CCAACA cacagetacigag CCACC cacagetacigag CCACC cacagetacigag CCACC cacagetacigag CCACC cacagetacigag CCACC cacagetacigag CCACC cacagetacigag CCACC cacagetacigag CCACC cacagetacigag CCACC cacagetacigag	geomterannance 2 geomterannance 1 geomterannance 2 geomterannance 1 geomterannance 1 geomterannance 1 geomterannance 1 geomterannance 1
attittigiacagiocotaggggttiganaedtigi attittigiacagiocotagggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi attittigiacagiocotaggggttiganaeattigi	CC c cacagedactega CCTCCT cacagedactega AGGCA cacagedactega GGGGGG cacagedactega GGGGCC cacagedactega CCCC cacagedactega CCC cacagedactega CCCAC cacagedactega GCGGAA cacacgedactega GCCAC cacagedactega GCCAC cacagedacteg	geontecannance 2 geontecannance 2 geontecannance 1 geontecannance 2 geontecannance 2 geontecannance 2 geontecannance 2 geontecannance 2 geontecannance 2 geontecannance 2

Cor gettii getti

Figure 2. Sequences of SJs formed by D-J recombination. (A) PCR fragments containing SJs formed by D\beta1-J\beta1.1 recombination were compared among WT (RAG1^{+/+}RAG2^{+/+}), core RAG1 (RAG1^{c/c} RAG2^{+/+}) and core RAG2 (RAG1^{+/-} RAG2^{c/c}) expressing mice. Details are as in Figure 1 legend. (**B**) SJ formed by DB2-JB2.1 recombination. (C) SJ formed by D82-J81 recombination.

shorter in VB10-DB-JB2.1 CJ when core RAG1 was expressed, with 50% of clones (15 out of 30) containing 4 or less nucleotides of potential DB coding sequence origin. Only 10% (3 out of 30) of CJ isolated from WT samples contained DB segments consisting of four or fewer nucleotides, and only 20% (6 out of 30) in core RAG2 samples. In addition, core RAG1 D β segments in V β 10-D β -J β 2.1 CJ had an average length deletion of -8.83 bp, compared to an average of -6.03 bp in WT D β segments. Thus, core RAG1 expression resulted in more frequent and extensive deletions within the D β segment of V β 10-D β -J β 2.1 CJ. In contrast, we observed no significant differences in the frequency or length of DB within VB14-DB-JB1.1 CJ of core RAG1, core RAG2 and WT samples. The frequency of nucleotide additions at DJ junctions of VB14-DB-JB1.1 CJ

Δ Vβ14-Dβ-Jβ1.1 recombination (splenocyte)

		mer Dβ1	7 mer	7 mer Jβ1.1	Frame
tgtgcctggagtct(CA	CACTG) (C P/N	ATTGTG)gggacaggggg	CACGGTG) (C. P/N	ACTGTG)caaacacagaagtct	
Wild Type (129			P/N		
tgtgcctggagtet	TGCAC		-	caaacacagaagtet	20
tgtgcctggagtc.		aggggg.		cagaagtet	
tgtgcctggagtet	AG	aggg	TTC	ct	-
tgtgcctggagtc.	AAGAGACGTC	ggg	-	cacagaagtct	-
tgtgcctggagtc.	С	gacagg		aacacagaagtet	
tgtgcctggagtct	AG T	agggg	А	agaagtet	
tgtgcctggagtct	C	aggg	TT	caaacacagaagtet	+ - + +
tgtgcctggag	CTTC	cagggggc	TGCGAG	acacagaagtet	
tgtgcctggagtct	GGC	gggacaggggg.		cacagaagtct	
tgtgcctggag tgtgcctggagtc.	CGGG	acaggg	-	caaacacagaagtet	-
tgtgcctgg	T	acag	AA	caaacacagaagtet	+
tgtgcctggagt	CCC	cagg		caaacacagaagtet	
tgtgcctggagtct	AGA	ggg	-	aacacagaagtet	-
tgtgcctgga	A	gggacagggggc	G	aacacagaagtet	+
tgtgcctggagtct	AG	ggc	CGA	caaacacagaagtet	-
tgtgcctggagtc.	CTTT	acagggggc	-	agaagtct	-
tgtgcctggagtc.	CA	gggacagg		cacagaagtct	
tgtgcctggagtct	AC	·····gggg	-	agaagtet	
tgtgcctgga	AG	agg	-	caaacacagaagtet	
tgtgcctggagtc.	GG	cagggg	AT	cacagaagtct	-
tgtgcctggagtc.	CCCA	.ggacag	TC TTC	caaacacagaagtet	1
tgtgcctggagtc.	GG	acag	TC <u>TTG</u> TT	caaacacagaagtet	1
tgtgcctggagtc. tgtgcctggagtct	AG GA	gggacaggg		cacagaagtet caaacacagaagtet	
igigeeiggagiei	2004	gacaggg		caacacagaageet	
Core RAG1(1c	/1c)				
tgtgcctggagtct	AG AA	gga	GGAG	acacagaagtct	
tgtgcctgga	-	gggacaggg	TTG	caaacacagaagtet	+
tgtgcctgg	CCC	ggg	GCC G	caaacacagaagtet	+
tgtgcctggagtct	AG G	ggga	GG	agaagtct	-
tgtgcctggagtct	AG T		-	.aaacacagaagtct	-
tgtgcctggagtct	AGA GG	gggggc	A	agaagtet	-
tgtgcctggagtct	-	gggacag	AGGG	acacagaagtet	-
tgtgcctggagtct		**********	CC TTG	caaacacagaagtct	
tgtgcctggagtct	000	ga	G	agaagtet	
tgtgcctggag	_CCC A	ggg	CA G AA	caaacacagaagtet	
tgtgcctggagtet	Δ	cag	AA	.aaacacagaagtet	1
tgtgcctggag tgtgcctggag	AG	cagggg	CCTTTGG	.aaacacagaagtet cacagaagtet	+++
tgtgcctggagtc.	AA	ga	GC	acacagaagtet	
tgtgcctggagtct	AG TT	gc	ACTT	aacacagaagtet	-
tgtgcctggagtct	AG TGG	ggggg.	G	acacagaagtct	
tgtgcctggagtct	AA	gggaca	-	acacagaagtet	
tgtgcctggagtc.	CGC	.ggacag	TG	aacacagaagtct	-
tgtgcctggagtc.	AC	cagg	-	acacagaagtct	-
tgtgcctggagtct	AGA GT	ggc	G AGGC	cacagaagtet	-
tgtgcctggagtct	CC	gacag	т	caaacacagaagtet	+
tgtgcctggagtct	GT	ggg	-	caaacacagaagtet	
tgtgcctggagtct	GA	acaggg	CCG	acacagaagtct	
tgtgcctggagtc. tgtgcctggagtct	CATC AG TTGCACT	TAaca	A CG	caaacacagaagtet gaagtet	
1515001550500	<u>Ao</u> Hocker			gaagee	
Core RAG2 (20	2/2c)				
tgtgcctggag	CCCTCGT CC	<u>C</u> ggg	-	acagaagtct	+
tgtgcctggagtc.	GGC	agggg	TTG	aacacagaagtet	-
tgtgcctggagte.	-	acagg	AGGG	acacagaagtct	
tgtgcctggag.	CCTAGAA	agggg		agaagtet	+
tgtgcctggagtc.	С	gggac	GGC	caaacacagaagtet	-
tgtgcctggagtct	GG	·····ggggg.	AC	cacagaagtet	-
tgtgcctggagtct	Δ	cag	G	aacacagaagtet	
tgtgcctggag	GC TCCC	gggacaggggg.	production of the second	cacagaagtct	+
tgtgcctgga tgtgcctggagtct	TCCC CT CC	gggacagggggc gggacag	GC TTGGG	aacacagaagtet aacacagaagtet	+
tgtgcctgg	GG	gggacagggg.		aacacagaagtet	+
tgtgcctggag	CCTA	gacagggg		agaagtet	+
tgtgcctggagtct	CTGGG	gggac	GG	cacagaagtet	-
tgtgcctggagtct	AG CAGGG		-	cagaagtct	
tgtgc				cagaagtet	
tgtgcctggag	CCTA	gacagggg		agaagtet	+
tgtgcctggagtct	Α	caggg	12	cacagaagtct	-
tgtgcctggagtct	A CAG	·····ggggg.		agaagtet	
tgtgcctggagt	TGG	·····ggggg.	-	aacacagaagtct	-
tgtgcctgg	GATCA C	gggacagggggc		acagaagtet	-
tgtgcctggagtct	AG CC		-	caaacacagaagtet	
tgtgcctggagtct	AGAC A	gggc	-	agaagtct	-
tgtgcctggag	C <u>CCC</u>	gggacaggggg.	AT	caaacacagaagtet	+
tgtgcctggagtc. tgtgcctggag	CCC	gggacagg	A	caaacacagaagtet cacagaagtet	

was significantly elevated when core RAG1 was expressed (50 and 85% for WT and core RAG1, respectively), although we did not observe a difference in the size of such additions. The increased frequency of nucleotide additions at DJ junctions was less apparent in V β 10-D β -J β 2.1 CJ. However, core RAG2 V β 10-D β -J β 2.1 CJ exhibited more extensive deletion of J β coding sequences compared to WT (-3.93 and -2.6 bp, respectively). V β 10-D β -J β 2.1 CJ from core RAG1 T cells also included deletions of J β (-3.81 bp), although these were not statistically different from WT. V β 14-D β -J β 1.1 CJ showed no significant difference among samples in the frequency or extent of deletions of J β sequences. Therefore, though expression of core RAGs

Β Vβ14-Dβ-Jβ1.1 recombination (DP thymocyte)

Vβ14	7 mer	7 mer	Dβ1	7 mer	7 mer	Jβ1.1	Fram
tgtgcctggagtc			TG)gggacaggggg		CACTGTG)C	aaacacagaagtct	
	P	/N		P/N			
Wild Type (129+/+)						
tgtgcctggagtct		G	gggg	AGG		agaagtet	
tgtgcctggagtc.		C	gggaca	TG		caaacacagaagtet	
tgtgcctggagtct		GTAA	.ggac	-		cagaagtet	-
tgtgcctggagtc.		C	caggg	AAA		.aaacacagaagtet	
Igtgcctggagtcl		A	agg	AT		cagaagtet	-
Igtgcctggagtcl		-	ggggg.	G		agaagtet	2
tgtgcctggagtc.		CC	ggga	T		caaacacagaagtet	
igtgcctggagtc.		CT	gacaggggg.			caaacacagaagtet	
igtgcctggagtc.		G	gggacag	А		caaacacagaagtet	
Igtgcctggagtcl		A		TG		caaacacagaagtet	
		-	agggg gaca	TG		caaacacagaagtet	
igtgcctggagtc.							- 0
tgtgcctggagt		G	caggg	С		acacagaagtet	-
igtgcctggagtci		G	ggga	-		.aaacacagaagtet	
lgtgcctgga		CCTC	acagggg			agaagtet	+
tgtgcctggagtct	A	G	.ggacagggg	5.0		caaacacagaagtet	-
Core RAG1	(1c/1c)						
igtgcctggagt		т	gggaca	ATAGAC	'G	aacacagaagtct	23
Igtgcctggag		A	acaggg	С		caaacacagaagtct	+
gtgcctggag		G	gggacagggg	TGTTG	G	acacagaagtet	+
gtgcctggagtcl		JA.	gggacagggg.		ð	caaacacagaagtet	_
gtgcctgga		-	ggc	AC G		caaacacagaagtet	+
tgtgcctggagtcl		G	gggg	AGG		agaagtet	
Igtgcctggagtci		ïC	.ggaca	CATO		caaacacagaagtet	2
igtgcctggagtc.		C		ACGG			
			cagggg	AGG		aacacagaagtet	+
tgtgcctgga		TC	cagggg			acacagaagtet	1
lgtgcctgga			ggg	TAC		caaacacagaagtet	
tgtgcctgga		200	gggacagg	AAG		acacagaagtet	+
tgtgcctgga		GC	.ggacaggg	CG		acagaagtct	+
gtgcctgga		СТ	acaggg	TG		caaacacagaagtet	-
tgtgcctggag	c	CC	caggggg.	AT		cacagaagtct	+
igtgcctggagt	1	ACT	aggggg.	AGG		agaagtct	-
Core RAG2	(20/20)						
		CC	00002	12		casacacagaagtet	60x
gtgcctggagtcl		C	······ggggg	-		caaacacagaagtet	+
gtg		C	aca			cacagaagtet	+
gtgcctggagtc.	GC	0	.ggacagggggc	G		caaacacagaagtet	
lgtgcctgg			******	G		caaacacagaagtet	+
lgtgcctggagtc.		CG		-		acacagaagtct	
gtgcctggagtc.		-	ggc	ACT		caaacacagaagtet	-
gtgcctggagtcl		G	gggg	AGG		agaagtet	-
gtgcctggagt		AG	ggggc	<u>G</u> GGG		cacagaagtct	-
gtgcctgga		-	gacag	AA		agaagtct	+
gtgcctggagtcl		C	gggg	TA		acagaagtet	
tgtgcctggagtc.		AGA	gc	-		acacagaagtet	-
Igtgcctggagtcl	AG	AG	ggggg.	GT		caaacacagaagtct	-
Igtgcctggagtcl	L A	TA CC	ggga			caaacacagaagtct	-
tgtgcctgga		AA	acag	-		.aaacacagaagtct	+
tgtg	CTTCT	ACCTC	gg	-		aacacagaagtet	

influenced the overall quality of CJ, the characteristics of these modifications tended to be associated with specific V-D and D-J junctions.

DISCUSSION

It has been proposed that the post-cleavage complex facilitates DNA repair by serving as a scaffold for the four broken ends generated by RAG-mediated cleavage (28-31). RAG proteins have also been shown to affect the formation of precise SJ in a yeast system (35). Although studies using artificial substrates are useful for understanding the molecular mechanism of V(D)J recombination, they may not completely recapitulate all aspects of antigen receptor rearrangement in vivo. Here, we provide the first evidence that the non-core regions of RAGs have critical functions in the proper assembly and resolution of recombination intermediates in endogenous loci. RAG2 truncation affects SJ associated with VD joining more severely than that of DJ joining, whereas core RAG1 affects DJ joining, as shown by previous studies focused on the efficiency of VD and DJ rearrangement associated with lymphocyte development (23,24). Moreover, we demonstrate a variety of CJ features that are influenced by the truncation of RAGs.

Since DSBs are the most lethal form of cellular DNA damage, there are at least three different pathways to repair a DSB, including classical NHEJ, microhomology-directed end-joining (MHEJ, or alternative NHEJ) and homologous recombination repair (HRR) (36–39). Although DSBs

C Vβ10-Dβ-Jβ2.1 recombination (DP thymocyte)

Vβ10 7 me gccagcagctaaga(CACA		mer Dβ2 ATTGTG)gggactggggggc(C		' mer Jβ2.1 TGCTGTG)taactatgetgage	Frame
	P/N	ATTOTO/gggactggggggggg	P/N	10c1010)unicungergage	
Wild Type (129+/+)			080		
gecageage	AA CCT	actggggggg.	GTG	.aactatgctgage	+
gccagcagctaa gccagcagcta	cci	tgg	ACAAAT	.aactatgctgagc taactatgctgagc	+
gcca	TTCGC	·····gggggg.	C	actatgctgage	+
gccagc	AGTCC	gactggggg	-	tatgctgagc	
gccagcagctaa	-	gggac	А	.aactatgctgagc	-
gccagcagetaa	ATT	ggggggg.	GTGA	.aactatgctgagc	-
gccagcageta	TCC	gactgggggg	-	.aactatgctgagc	+
gccagcagctaa	-	gggactggggggg.	-	ctatgctgagc	-
gccagcagc	С	gggact	т	ctatgctgagc	+
gccagcage	TTC	gggactgg	AC	tgctgagc	
gccagcagc	G <u>C</u>	gggactggggggg.	ATC	gctgagc	+
gccagcage	CCT CC	gggac	GG	. aactatgetgage	+
gccagcagc	CT CCCGT	gactg	GTTA	taactatgetgage	+
gccagcagct gccagcagct	TC	gggactgg	0	tatgctgagc	2
gccagcagct	CC	gggac	AGA	.aactatgetgage	2
gccagcagcta	T	.ggac	AAAT	taactatgetgage	+
gccagcag	AATCC	actgggg	TT	ctatgctgagc	-
gccagcagc	GGA CC	gggactgg	TT	ctatgctgagc	- 20
gccagcagct	TCGTC	ctggggg	-	. aactatgctgagc	23
gccagcagc	GGC	gggactggggg	т	actatgctgagc	-
gccagcagc	GC	gggactggggg	т	taactatgctgagc	+
gcca	CC	gggac	A	.aactatgctgagc	+
gccagcag	-	gactgggggg	CC	taactatgctgagc	+
gccagcagc	CTC	gactgggggggc	TT	atgctgagc	-
gccagcag	AA	gactggggggg.		.aactatgctgagc	+
gccagcagct,	CCTAGA	gggg	AAGGG	ctatgctgage	-
gccagcagct	GT	actgggggg	-	.aactatgctgagc	+
gccagcag	AG CC	gggactggggggg.	-	.aactatgetgage	
Core RAG1 (1c/1c)					
gccagcagct	CA	gac	AGA	.aactatgctgagc	+
gccagcagcta	TCA	ggg	AAGGG	atgctgage	+
gccagcag	-	.ggac	AGA	.aactatgctgagc	+
gccagcagct	CGACA	ggggg	A	tatgetgage	+
gccagcagct	CCA	gggg	A	tatgctgagc	+
gccagcag	-	.ggac	AGAA	actatgctgage	+
gccagcagct	CTAGCCA	ggg	AGGG	tatgctgagc	+
gccagcagc	CCT	tgg	ACAG A	taactatgctgagc	+
gccage	-	tgggg	С	.aactatgctgagc	+
gccagcagct	CCA	ctgggg	AT	tatgetgage	+
gccagcagcta	CAC	ctgggggg	CC	atgctgagc	-
gccagcaget	т <u>с</u>	gggactggggg	-	taactatgctgagc	+
gccagcagct	Т	.ggactg	CCT	ctatgctgagc	+
gccagcagct	CCTT	gactggggg	-	ctatgctgagc	+
gccagcagct	CCCGGA	·····gggggg	-	ctatgctgagc	-
gccagcagcta	CTC	ctgggggg	-	ctatgctgagc	-
gccagcag	AAACA	·····ggg	- CT	.aactatgctgagc	
gccagcagc	CAAA ACGC	gggg		tgagc	
gccagcag	C CCC	gac	A A AGGGAG	taactatgetgage	+++
gccagcagc	A	gggactgg	AUGOAU	taactatgctgagc	+
gccagcagc gccagcag	GGA CC	gggac	ACCT	actatgetgage	+
gccagca	CCC	actggggg	ATT	tatgctgagc	+
gccagcagct	CG CC	ggga	AAGGGAG		+
gccage	CTCTCG C	ggg	AA	tatgctgagc	+
gccagcagct	CCCCCCA	gg	TAT	ctatgctgagc	-
gccagcag	ATCA	gact	AA	.aactatgctgagc	1.2
gccagcaget	CTA	gactggg	AGGG	tatgctgagc	+
gccagcaget	т	gactggg	A	atgctgage	+
gccagcagc	C CCC	gggactggg	TG	getgage	+
Core RAG2 (2c/2c)					
gccagcagct	CC	.ggactggg	т	tatgctgagc	+
gccagcaget	С	.ggactggggggg.		ctatgctgagc	+
gccagcagc	CCA	gggg	т	ctatgctgagc	+
gccagc	AGGT	.ggactg	TGG	atgctgagc	+
gccagcagct		.ggactg	TGGC	tgctgagc	+
gccagcageta	CA	·····ggg	Δ	taactatgctgagc	+
gccagcagc	CGC	.ggactg	CTC	.aactatgctgagc	+
gccagcagct	TAGCA	·····ggg	-	gctgagc	+
gccagcagc	-	.ggactgg	-	actatgctgage	+
gccagcagc	TCCA	ggg	т	ctatgctgagc	+
gccagcaget	С	.ggactgggggggc	-	tatgctgagc	+
gccagcagc	-	.ggact	CCT	ctatgctgagc	+
gccagcagc	CC	gggac	GGCG	ctatgetgage	+
gccagcagc	CCGCA C	ggggg	CTC TG	atgetgage	+
gccagcagct	C CCC	.ggactggg	AATA	tatgetgage	+
gccagcagc	CC CCC	gggac	AATA	atgctgagc atgctgagc	+
gccagcagc gccagcagct	CACA	gggact	GCT	tgctgagc	+
gccagcage	CG CCC	gggact	TA	atgetgage	+
gccagcagct	GCTCTTT	.ggac	AGAG	actatgetgage	+
gccagcagc	-	.ggact	CCT	ctatgetgage	+
gccagcagct	CCC	gggactggggggg.	-	tgctgage	+
gccagcagc	GGC	.ggactg	AAC	.aactatgetgage	+
gccagc	C	.ggact	CACGGGACA		+
gccagcagc	C	gggactggggg	CGA	getgage	+
gccagcagct	TG CCC	gggactggggggg	-	atgctgagc	-
	CCTCT		CCCA	.aactatgctgagc	+
gccagcagc	cerer	gactgg			
gccagcagc gccagcagct	CA		TC	ctatgctgagc	+

Figure 3. Sequences of CJs formed by V-D-J recombination. (A) CJs formed by V β 14-D β -J β 1.1 recombination in splenocytes were compared among WT (RAG1^{+/+} RAG2^{+/+}), core RAG1 (RAG1^{-/c} RAG2^{+/+}) and core RAG2 (RAG1^{+/+} RAG2^{c/c}) expressing mice. In V β 14-D β -J β 1.1 recombination, D β 1 within the D β 1-J β 1 gene cluster is expected to be utilized. Original coding sequences of VB14, DB1 and JB1.1 (boldface lowercase letters) and their flanking heptamer sequences (uppercase letters in parentheses) are indicated on the top. Nucleotide insertions are indicated by capital letters. Presumptive P nucleotides are underlined. All the clones contained different junctional sequences to each other. In-frame or out-of-frame rearrangement is denoted as '+' or '-' at the right end of each sequence. (B) CJ formed by V β 14-D β -J β 1.1 recombination in DP thymocytes. (C) CJ formed by V β 10-D β -J β 2.1 recombination in DP thymocytes. Original coding sequences of VB10, DB2 and J β 2.1 (boldface lowercase letters) and their flanking heptamer sequences (uppercase letters in parentheses) are indicated on the top. Although D β 1 can be used instead of D β 2 in at least one-fourth of all V β 10-D β -J β 2.1 CJ (69), but they could not be assigned due to similarity. See (A) and (B) for $D\beta 1$ coding sequence.

Table 2. Comparison of D-J SJs

	WT	Core RAG1	Core RAG2
Dβ1-Jβ1.1			
Precise joints	43/50 (86%)	34/48 (71%)	35/47 (74%)
Imprecise joints	7/50 (14%)	14/48 (29%)	12/47 (26%)
N nucleotides	7/7 (100%)	13/14 (93%)	11/12 (92%)
Deletion	0/7 (0%)	3/14 (21%)	2/12 (17%)
Dβ2-Jβ2.1			
Precise joints	40/49 (82%)	9/46 (20%)	14/45 (31%)
Imprecise joints	9/49 (18%)	37/46** (80%) ^b	31/45** (69%) ^a
N nucleotides	9/9 (100%)	26/37 (70%)	28/31 (90%)
Coding sequences	0/9 (0%)	34/37 (92%)	13/31 (42%)
Deletion	0/9 (0%)	20/37 (54%)	7/31 (23%)
Dδ2–Jδ1			
Precise joints	41/48 (85%)	26/46 (57%)	36/46 (78%)
Imprecise joints	7/48 (15%)	20/46** (43%)	10/46 (22%)
N nucleotides	6/7 (86%)	20/20 (100%)	10/10 (100%)
Coding Sequences	1/7 (14%)	0/20 (0%)	0/10 (0%)
Deletion	0/7 (0%)	0/20 (0%)	0/10 (0%)

Number of each joint out of total number of analyzed was shown. The frequency of each joint is shown in parentheses. Significant differences versus WT mice, when compared using chi-squared test, are indicated by **(P < 0.01). Junctions were sequences from two independent animals, and the results were combined to calculate the number of each case.

^aDifferences from WT mice are statistically significant even when the data containing coding sequences are excluded, to exclude possible non-12/23 HJ from SJ (see Discussion).

^bDifference from WT mice is not significant when the data containing coding sequences are excluded.

associated with V(D)J recombination are normally repaired by NHEJ, alternative pathways must take part in completing the repair when conventional pathways are not available or not effective (40). When either RAG1 or RAG2 were truncated, we observed SJ that exhibit features consistent with repair via a pathway other than NHEJ.

Imprecise joints isolated from mice producing truncated RAG proteins frequently contained nucleotide deletions. A total of 90% of V δ 5-D δ 2 imprecise joints isolated from core RAG2 mice contained deletions, with an average loss of 17 nt, and the largest being a 74 bp deletion (Figure 1C). Thus, truncation of RAGs might directly influence the stability of the SEC in certain RSS pairs, allowing exonuclease attack of unprotected SE. The occurrence of deletions also suggests an end-joining defect, as end-joining deficiencies are known to

Table 3. Frequencies and aver	age length in	nucleotides	modification in CJ
-------------------------------	---------------	-------------	--------------------

	WT	Core RAG1	Core RAG2
Vβ14-Dβ-Jβ1.1			
Deletion within $V\beta$	24/40 (60%, -1.79)	21/40 (53%, -2.86**)	26/40 (65%, -3.34**)
P/N addition at VD junction	36/40 (90%, +2.72)	33/40 (83%, +2.72)	35/40 (88%, +3.87*)
Dβ ≤4 nt	17/40 (43%, -6.61)	18/40 (45%, -7.20)	17/40 (43%, -7.18)
P/N addition at DJ junction	20/40 (50%, +2.25)	34/40** (85%, +3.03)	18/40 (45%, +2.55)
Deletion within $J\beta$	23/40 (58%, -4.82)	27/40 (68%, -3.74)	30/40 (75%, -4.23)
Vβ10-Dβ-Jβ2.1			
Deletion within $V\beta$	30/30 (100%, -4.77)	30/30 (100%, -4.83)	30/30 (100%, -4.73)
P/N addition at VD junction	26/30 (87%, +3.19)	27/30 (90%, +3.74)	25/30 (83%, +3.20)
Dβ ≤4 nt	3/30 (10%, -6.03)	15/30** (50%, -8.83**)	6/30 (20%, -7.56)
P/N addition at DJ junction	20/30 (67%, +2.50)	25/30 (83%, +2.60)	24/30 (80%, +2.96)
Deletion within $J\beta$	25/30(83%, -2.60)	26/30(87%, -3.81)	28/30(93%, -3.93*)

Sequence data obtained from two independent genomic DNA samples were combined. Data obtained from spleen and DP thymocytes were combined in V β 14-D β -J β 1.1. Number of each event out of total number of analyzed were shown. The frequency and the average length of addition (+) and deletion (-) of each joining are shown in parentheses. Significant differences versus WT mice, when compared by chi-squared test (frequencies) or *t*-test (average length of addition and deletion), are indicated by * (P < 0.05) and ** (P < 0.01).

lead to large deletions. In fact, these joints are somewhat similar to SJ from artificial substrates introduced into XRCC4-deficient cells, as they tend to have large deletions with short stretches of homology at the junction (41). Either joining partner could have contributed these nucleotides, the presence of which is consistent with joining mediated by a homologous stretch of nucleotides between the two joining partners. Such MHEJ is thought to occur as a back-up to conventional NHEJ repair, since defects in NHEJ result in elevated MHEJ (38,39). In addition, CJ isolated from core RAG knock-in mice also appeared to have an increased frequency of nucleotide deletions, although the effect was less obvious than in SJ due to the imprecise nature of CJ.

Sequences homologous to the germline coding sequence were frequently observed in SJ isolated from core RAGexpressing mice. The presence of these coding sequences can be explained by a number of mechanisms. Oligonucleotide capture is a mechanism by which short DNA segments cleaved at both 5' and 3' ends can be incorporated into the junction (42,43). CJ with short oligonucleotides that could have been derived in this manner have been reported previously (44–46). The D δ 2 coding sequence trapped within a V δ 5-D δ 2 SJ of a core RAG2 sample (Figure 1C) could have resulted from this.

The frequent detection of long D β coding flank sequence attached directly to RSS cannot be explained by oligonucleotide capture, but is more likely the result of aberrant cleavage. For instance, RSS-like sequences found within coding gene segments may be mistakenly used as RAG substrates. In this regard, heptamer-like sequences containing CAC trinucleotide or nonamer-like sequences can serve as substrates for RAG proteins *in vitro* (47,48), and evidences indicating aberrant recombination using such sites have been reported not only at the breakpoint of chromosomal translocations associated with lymphoid neoplasia (33,49) but also in the TCR δ locus of normal mice (46).

Although RSS-like sequences can occur anywhere in the genome, we failed to find such a sequence in the coding sequences near the cleavage site, except for the bonafide RSS present on the opposite end of D β gene segments. The majority of D β 2-J β 2.1 SJ contained unmodified D β 2 coding flank directly attached to the 3'D β 2 RSS, while there were

frequent modifications observed between the end of DB2 coding sequence and the J β 2.1 RSS (Figure 2B). The structure of these aberrant joints suggests that the cleavage has occurred at the 5'D β 2 RSS instead of the 3'D β 2 RSS. The frequent deletion of D β coding sequence within V β 10-D β -J β 2.1 CJ isolated from core RAG1 samples (Figure 3C and Table 3) also supports the involvement of aberrant cleavage, since corresponding D_β2 coding sequences appeared in reciprocal SJ. Although the formation of a 12/12 deletional hybrid joints (HJs) between the CE of one element and the SE of another may sound unlikely, we speculate that short D β coding sequence flanked by 3' and 5' RSSs in the presence of core RAGs may result in such HJ. Indeed, HJ in apparent violation of the 12/23 rule has similarly been described in the TCR δ locus of WT mice, where $D\delta$ elements were involved (46). During V(D)J recombination, a RAG complex is first thought to form on a single RSS, followed by subsequent capture of a second RSS (50,51). Though the correct 12/23 pair may be favored by full-length RAGs, a complex formed with core RAGs may allow improper alignment with an RSS only 1-1.5 turns of the helix away. Or, the synaptic complex may initially form correctly with a 12/23 pair, but core RAGs could form an unstable complex that might lead to improper cleavage of a nearby heptamer. These possibilities may also be related to the finding that the RAG complex cleaves a non-B-DNA structure at the *Bcl-2* major breakpoint region (52). If RAG binding to the 3'end of a D gene segment can induce the formation of non-Bform DNA that extends to the 5' end of the D gene segment, this might provide a target for RAG nicking. Indeed, recent results showed that at least 12 bp of coding region, proximal to the heptamer, are protected in the synaptic complex formed by core RAG proteins (53). Thus the RSS flanking the other side of a D segment may be directly adjacent or partially contained within the synaptic complex, likely affecting the structure of the corresponding DNA.

Unlike recombination involving the D β 2-J β 2.1 cluster, imprecise SJ associated with the D β 1-J β 1.1 cluster contained no nucleotides with coding sequence homology (Figure 2A). We speculate that subtle differences in RSS pairs and/or length of D β segments (12 and 14 bp for D β 1 and D β 2, respectively) might significantly affect the DNA structure and/or stability of the synaptic complex, resulting in dramatic differences in how the synaptic complex can be formed and/ or resolved by core RAGs. Similarly, $D\beta1$ coding sequences detected within VB8.3 to DB1 rearrangements in core RAG2 samples could have arisen through deletional rearrangements between V β and 3'D β 1 RSS. Previous studies have shown that D-J rearrangement was significantly reduced when core RAG1 or core RAG2 was expressed, with nearly half of mature B cells containing germ line alleles (23,24). Since D β -to-J β rearrangement is not required for V β -D β joining (54), V β -to-D β rearrangement can occur in core RAGexpressing mice even before the completion of $D\beta$ -J β rearrangement. Thus in the presence of core RAG2, a 3'DB1 RSS might be mistakenly cleaved within a VD synaptic complex. Increased accumulation of HJ has been reported in cells expressing core RAG1 and/or RAG2 (15). The increased rate of HJ formation in core RAGs in combination with the structure of short D segments might account for HJ violating the12/23 rule. To confirm that the SJ results were not all due to the presence of HJ, we excluded data potentially arising from HJ by removing all data containing coding sequences from the calculation of frequencies (see legends for Tables 1 and 2). Consistent with most of the original analyses, we still found significant increases in imprecise SJ, except for data involving D β 2-J β 2.1 SJ of core RAG1 samples. This is likely due to an increase in the rate of D β 2-J β 2.1 associated HJ in the presence of core RAG1. Thus, we conclude that the expression of core RAG increases the frequency of imprecise SJ, in a manner highly influenced by specific RSS partners.

The coding sequences observed in V β 14-D β 1 SJ of core RAG2 samples could still involve a different mechanism, since they contained not only D β coding flank, but also V β coding flank sequences (Figure 1A). Similar aberrant joints have been reported previously in V β 14-D β 1 SJ isolated from transgenic mice expressing a T490A mutant of RAG2. T490 phosphorylation is known to promote cell-cycle-dependent RAG2 degradation (19). Since core RAG2 lacks this phosphorylation site due to C-terminal truncation, the protein is expected to behave like the T490A mutant, remaining active into S phase where HRR is the prevalent DSB repair pathway. Although homologous sequences observed within these SJ are shorter than those usually used in HRR, we cannot completely rule out the possibility that factors involved in HRR might participate in the repair of these. Importantly, T490A mutant expressed on a $RAG2^{-/-}$ background does not affect the frequency of precise VB14-DB1 SJ, thus, it is unlikely this point mutation affects complex stability or increases aberrant cleavage (19). On the other hand, core RAG2 expression reduced the frequency of precise joints associated with V β 14-D β 1 recombination from 90% in WT to 62% in core RAG2 (Table 1). Increased N additions and possible microhomology-mediated joining are likely responsible for the low fidelity of these joints. Microhomology was never observed in imprecise SJ isolated from T490A transgenic mice (19). This suggests that deletion of the C-terminal RAG2 region is responsible for lowering the frequency of precise joints, and not its aberrant expression in S phase, presumably due to an inability to provide a proper architectural role during the endjoining phase.

Importantly, the reduced accuracy of SJ formed via core RAG is sensitive to nucleotide sequences and/or the structure

Table 4. Comparison of RSS analyzed in this study

Position	7mer	12/23 spacer	9mer
3′Vβ14	CACA <u>C</u> TG	AGTAGGGTGGGGC	<u>G</u> CAAAAACC
2/3/00 2	CACACTO	AGACATCTGT	
3′Vβ8.3	CACAGTG	ATGTGTGGGCTTCCT CCCCTTTGC	ACA <u>G</u> AAA <u>GT</u>
3'Vβ10	CACAGT <u>T</u>	GTGCAGAGTCACTG TTTCCCTGT	<u>GCAC</u> AAACC
3'Vδ5	CAC <u>T</u> GTG	GTGCAGGTGCCCAG GGAGCCTGT	AC <u>CC</u> AAACC
5′Dβ1	CACAATG	TTACAGCTTTAT	ACAAAAAAG
3′Dβ1	CACGGTG	ATTCAATTCTATGG GAAGCCTTT	ACAAAAA
5′Dβ2	CACAATG	TTACATCGTGAT	ACAAAAAAG
3′Dβ2	CACAATG	ATTCAACTGGAAGA GGTGCTTTT	ACAAAAA <u>G</u> C
5′Dδ2	CACGGTG	CTACAGAGCTTT	GCAAAAACC
3'Dð2	CACAGTG	TTGCAAACCCCATA GGGACCTGT	ACAAAAACT
5′Jβ1.1	CACAGTG	CCATAGGATGAG	GAGAAAAAT
5′Jβ2.1	CACAGCA	GAAAAGGGCTAC	CAAGAATTC
5′Jδ1	CACAGCT	ACTGAGGCCATT	CCAAAAACC

Nucleotides different from the consensus sequences (heptamer, CACAGTG and nonamer, ACAAAAACC) are underlined.

of RSS. The RSS analyzed in this study are listed in Table 4. As RAGs undergo coupled RSS recognition and cleavage, both individual RSS sequences as well as their combinations are likely to contribute to the structure of the synaptic complex. SJ from different RSS combinations had quite distinct characteristics for both sides of SE. Although it has been reported that the quality of RSS affects accessibility to RAGs (54,55) and recombination frequencies (56,57), it may also affect the quality of recombination, especially when core RAGs are involved. The post-cleavage synaptic complex is known to be stable (3-5), however, certain RSS combinations such as V δ 5-D δ 2, may be inherently sensitive to structural change and become unstable when formed with truncated RAGs. Indeed, expression of mutant RAG1 proteins that form unstable post-cleavage complexes allow DNA ends to participate in both HRR and alternative NHEJ pathway (58). Thus, non-core regions of RAG might also have important roles for stabilizing the post-cleavage synaptic complex by minimizing the effects from RSS variations, since truncated RAGs result in higher frequency of imprecise SJ and modest increase in modification of CJ in vivo.

Increased frequency in aberrant joints associated with the expression of core RAGs have never been reported in studies using extrachromosomal substrates (8-12,28,29,31,59-61). There are several reasons why these effects have never been found in transfection studies. First, RSS tested for joining in plasmid substrates tended to be limited to so-called 'consensus' RSS. Our results demonstrate that not all combinations of RSS result in extensive changes to CJ or SJ. Although endogenous RSS have been tested for ex vivo recombination and in vitro binding/cleavage assays (56,57,62,63), specific features of these joints were not addressed in these studies. Thus, it will be important to compare core and full-length RAGs for quality of end-joining between different RSS pairs formed ex vivo. In addition, LM-PCR should be performed with and without T4 DNA polymerase treatment, since we cannot completely rule out the possibility that

truncated RAGs might produce aberrant SE with certain pairs of RSS. Second, artificial substrates are composed of limited coding flank sequences. Since an importance of coding flank sequence to recombination frequency has been shown in the earlier study (64), it has shown to play a critical role in binding (53,65,66), cleavage (47,48,67) and accessibility control by RAGs (68). Thus, coding flank sequences might also affect the joining reaction. Third, the structure of short D gene segments with RSS at both 3' and 5' ends might account for the aberrant cleavage or misalignment of core RAGs. From this point of view. D segments attached to 3' and 5'RSS should be tested in artificial substrates to compare the end-processing of fulllength RAGs versus core RAGs. Fourth, chromatin structure in endogenous loci might make core RAGs more sensitive to subtle difference in RSS. To further address these issues, the structure of cleaved ends, junctions and stability of the SEC will need to be compared in full-length versus core RAGs, and with different combinations of RSS, including natural D coding segments, in naked or chromatinized substrates. Our results indicate that the non-core regions of RAGs are critical for tolerating variations between RSS and to suppress aberrant joints in endogenous loci. In combination with our findings, future analyses focusing on the above aspects in vitro should reveal the complete role of RAG proteins in V(D)J joining.

ACKNOWLEDGEMENTS

The authors thank Craig H. Bassing and Yan Dai for critical reading of the manuscript, Hiroshi Nakase for stimulating discussions and Kaori Gotoh for help with our initial work. This study was supported by grants from the Ministry of Education, Science, Sports and Culture of Japan (Y.T. and Y.A.) and NIH grants (F.W.A.). F.W.A. is the investigator of the HHMI.

REFERENCES

- Bassing,C.H., Swat,W. and Alt,F.W. (2002) The mechanism and regulation of chromosomal V(D)J recombination. *Cell*, **109** (Suppl.), S45–S55.
- Gellert,M. (2002) V(D)J recombination: RAG proteins, repair factors, and regulation. Annu. Rev. Biochem., 71, 101–132.
- Agrawal,A. and Schatz,D.G. (1997) RAG1 and RAG2 form a stable postcleavage synaptic complex with DNA containing signal ends in V(D)J recombination. *Cell*, 89, 43–53.
- Hiom,K. and Gellert,M. (1998) Assembly of a 12/23 paired signal complex: a critical control point in V(D)J recombination. *Mol. Cell*, 1, 1011–1019.
- Jones, J.M. and Gellert, M. (2001) Intermediates in V(D)J recombination: a stable RAG1/2 complex sequesters cleaved RSS ends. *Proc. Natl Acad. Sci. USA*, 98, 12926–12931.
- Brandt,V.L. and Roth,D.B. (2002) A recombinase diversified: new functions of the RAG proteins. *Curr. Opin. Immunol.*, 14, 224–229.
- Dai, Y., Kysela, B., Hanakahi, L.A., Manolis, K., Riballo, E., Stumm, M., Harville, T.O., West, S.C., Oettinger, M.A. and Jeggo, P.A. (2003) Nonhomologous end joining and V(D)J recombination require an additional factor. *Proc. Natl Acad. Sci. USA*, **100**, 2462–2467.
- Cuomo,C.A. and Oettinger,M.A. (1994) Analysis of regions of RAG-2 important for V(D)J recombination. *Nucleic Acids Res.*, 22, 1810–1814.
- Kirch,S.A., Sudarsanam,P. and Oettinger,M.A. (1996) Regions of RAG1 protein critical for V(D)J recombination. *Eur. J. Immunol.*, 26, 886–891.
- Sadofsky, M.J., Hesse, J.E., McBlane, J.F. and Gellert, M. (1993) Expression and V(D)J recombination activity of mutated RAG-1 proteins. *Nucleic Acids Res.*, 21, 5644–5650.

- Sadofsky,M.J., Hesse,J.E. and Gellert,M. (1994) Definition of a core region of RAG-2 that is functional in V(D)J recombination. *Nucleic Acids Res.*, 22, 1805–1809.
- Silver,D.P., Spanopoulou,E., Mulligan,R.C. and Baltimore,D. (1993) Dispensable sequence motifs in the RAG-1 and RAG-2 genes for plasmid V(D)J recombination. *Proc. Natl Acad. Sci. USA*, **90**, 6100–6104.
- Elkin,S.K., Matthews,A.G. and Oettinger,M.A. (2003) The C-terminal portion of RAG2 protects against transposition *in vitro*. *EMBO J.*, 22, 1931–1938.
- Melek, M. and Gellert, M. (2000) RAG1/2-mediated resolution of transposition intermediates: two pathways and possible consequences. *Cell*, **101**, 625–633.
- Sekiguchi, J.A., Whitlow, S. and Alt, F.W. (2001) Increased accumulation of hybrid V(D)J joins in cells expressing truncated versus full-length RAGs. *Mol. Cell*, 8, 1383–1390.
- Swanson, P.C., Volkmer, D. and Wang, L. (2004) Full-length RAG-2, and not full-length RAG1, specifically suppresses RAG-mediated transposition, but not hybrid joint formation or disintegration. *J. Biol. Chem.*, 279, 4034–4044.
- Tsai, C.L. and Schatz, D.G. (2003) Regulation of RAG1/RAG2-mediated transposition by GTP and the C-terminal region of RAG2. *EMBO J.*, 22, 1922–1930.
- Jones, J.M. and Gellert, M. (2003) Autoubiquitylation of the V(D)J recombinase protein RAG1. *Proc. Natl Acad. Sci. USA*, **100**, 15446–15451.
- Lee, J. and Desiderio, S. (1999) Cyclin A/CDK2 regulates V(D)J recombination by coordinating RAG-2 accumulation and DNA repair. *Immunity*, **11**, 771–781.
- Yurchenko, V., Xue, Z. and Sadofsky, M. (2003) The RAG1 N-terminal domain is an E3 ubiquitin ligase. *Genes Dev.*, 17, 581–585.
- Corneo, B., Benmeral, A. and Villartay, J.P. (2002) A short peptide at the C-terminus is responsible for the nuclear localization of RAG2. *Eur. J. Immunol.*, **32**, 2068–2073.
- Spanopoulou, E., Cortes, P., Shih, C., Huang, C.M., Silver, D.P., Svec, P. and Baltimore, D. (1995) Localization, interaction, and RNA binding properties of the V(D)J recombination-activating proteins RAG1 and RAG2. *Immunity*, 3, 715–726.
- Akamatsu,Y., Monroe,R., Dudley,D.D., Elkin,S.K., Gartner,F., Talukder,S.R., Takahama,Y., Alt,F.W., Bassing,C.H. and Oettinger,M.A. (2003) Deletion of the RAG2 C-terminus leads to impaired lymphoid development in mice. *Proc. Natl Acad. Sci. USA*, 100, 1209–1214.
- Dudley, D.D., Sekiguchi, J., Zhu, C., Sadofsky, M.J., Whitlow, S., DeVido, J., Monroe, R.J., Bassing, C.H. and Alt, F.W. (2003) Impaired V(D)J recombination and lymphocyte development in core RAG1-expressing mice. *J. Exp. Med.*, **198**, 1439–1450.
- Liang,H.E., Hsu,L.Y., Cado,D., Cowell,L.G., Kelsoe,G. and Schlissel,M.S. (2002) The 'dispensable' portion of RAG2 is necessary for efficient V-to-DJ rearrangement during B and T cell development. *Immunity*, 17, 639–651.
- 26. Villa,A., Sobacchi,C., Notarangelo,L.D., Bozzi,F., Abinun,M., Abrahamsen,T.G., Arkwright,P.D., Baniyash,M., Brooks,E.G., Conley,M.E. *et al.* (2001) V(D)J recombination defects in lymphocytes due to RAG mutations: severe immunodeficiency with a spectrum of clinical presentations. *Blood*, **97**, 81–88.
- Steen,S.B., Han,J.O., Mundy,C., Oettinger,M.A. and Roth,D.B. (1999) Roles of the 'dispensable' portions of RAG-1 and RAG-2 in V(D)J recombination. *Mol. Cell Biol.*, **19**, 3010–3017.
- Qiu,J.X., Kale,S.B., Yarnell Schultz,H. and Roth,D.B. (2001) Separation-of-function mutants reveal critical roles for RAG2 in both the cleavage and joining steps of V(D)J recombination. *Mol. Cell*, 7, 77–87.
- Yarnell Schultz, H., Landree, M.A., Qiu, J.X., Kale, S.B. and Roth, D.B. (2001) Joining-deficient RAG1 mutants block V(D)J recombination *in vivo* and hairpin opening *in vitro*. *Mol. Cell*, 7, 65–75.
- Tsai,C.L., Drejer,A.H. and Schatz,D.G. (2002) Evidence of a critical architectural function for the RAG proteins in end processing, protection, and joining in V(D)J recombination. *Genes Dev.*, 16, 1934–1949.
- Huye,L.E., Purugganan,M.M., Jiang,M.M. and Roth,D.B. (2002) Mutational analysis of all conserved basic amino acids in RAG-1 reveals catalytic, step arrest, and joining-deficient mutants in the V(D)J recombinase. *Mol. Cell Biol.*, 22, 3460–3473.

- Miller,S.A., Dykes,D.D. and Polesky,H.F. (1988) A simple salting out procedure for extracting DNA from human nucleated cells. *Nucleic Acids Res.*, 16, 1215.
- Lewis,S.M. (1994) The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. *Adv. Immunol.*, 56, 27–150.
- Starr, T.K., Jameson, S.C. and Hogquist, K.A. (2003) Positive and negative selection of T cells. *Annu. Rev. Immunol.*, 21, 139–176.
- Clatworthy, A.E., Valencia, M.A., Haber, J.E. and Oettinger, M.A. (2003) V(D)J recombination and RAG-mediated transposition in yeast. *Mol. Cell*, 12, 489–499.
- Kanaar,R., Hoeijmakers,J.H.J. and van Gent,D.C. (1998) Molecular mechanisms of DNA double-strand break repair. *Trends Cell Biol.*, 8, 483–489.
- Baumann,P. and West,S.C. (1998) DNA end-joining catalyzed by human cell-free extracts. *Proc. Natl Acad. Sci. USA*, 95, 14066–14070.
- Kabotyanski,E.B., Gomelsky,L., Han,J.O., Stamato,T.D. and Roth,D.B. (1998) Double-strand break repair in Ku86- and XRCC4-deficient cells. *Nucleic Acids Res.*, 26, 5333–5342.
- Verkaik,N.S., Esveldt-van Lange,R.E., van Heemst,D., Bruggenwirth,H.T., Hoeijmakers,J.H., Zdzienicka,M.Z. and van Gent,D.C. (2002) Different types of V(D)J recombination and end-joining defects in DNA double-strand break repair mutant mammalian cells. *Eur. J. Immunol.*, **32**, 701–709.
- Roth,D.B. (2003) Restraining the V(D)J recombinase. *Nature Rev. Immunol.*, 3, 656–666.
- Li,Z., Otevrel,T., Gao,Y., Cheng,H.L., Seed,B., Stamato,T.D., Taccioli,G.E. and Alt,F.W. (1995) The XRCC4 gene encodes a novel protein involved in DNA double-strand break repair and V(D)J recombination. *Cell*, 83, 1079–1089.
- Roth,D.B., Proctor,G.N., Stewart,L.K. and Wilson,J.H. (1991) Oligonucleotide capture during end joining in mammalian cells. *Nucleic Acids Res.*, 19, 7201–7205.
- Lieber, M.R. (1992) The mechanism of V(D)J recombination: a balance of diversity, specificity, and stability. *Cell*, **70**, 873–876.
- Lieber, M.R., Hesse, J.E., Lewis, S., Bosma, G.C., Rosenberg, N., Mizuuchi, K., Bosma, M.J. and Gellert, M. (1988) The defect in murine severe combined immune deficiency: joining of signal sequences but not coding segments in V(D)J recombination. *Cell*, 55, 7–16.
- Schuler, W., Ruetsch, N.R., Amsler, M. and Bosma, M.J. (1991) Coding joint formation of endogenous T cell receptor genes in lymphoid cells from scid mice: unusual P-nucleotide additions in VJ-coding joints. *Eur. J. Immunol.*, 21, 589–596.
- Carroll,A.M., Slack,J.K. and Mu,X. (1993) V(D)J recombination generates a high frequency of nonstandard TCR D delta-associated rearrangements in thymocytes. *J. Immunol.*, 150, 2222–2230.
- Ramsden, D.A., McBlane, J.F., van Gent, D.C. and Gellert, M. (1996) Distinct DNA sequence and structure requirements for the two steps of V(D)J recombination signal cleavage. *EMBO J.*, 15, 3197–3206.
- Cuomo,C.A., Mundy,C.L. and Oettinger,M.A. (1996) DNA sequence and structure requirements for cleavage of V(D)J recombination signal sequences. *Mol. Cell Biol.*, 16, 5683–5690.
- Tycko,B. and Sklar,J. (1990) Chromosomal translocations in lymphoid neoplasia: a reappraisal of the recombinase model. *Cancer Cells*, 2, 1–8.
- Jones, J.M. and Gellert, M. (2002) Ordered assembly of the V(D)J synaptic complex ensures accurate recombination. *EMBO J.*, 21, 4162–4171.
- Mundy,C.L., Patenge,N., Matthews,A.G. and Oettinger,M.A. (2002) Assembly of the RAG1/RAG2 synaptic complex. *Mol. Cell Biol.*, 22, 69–77.

 Raghavan,S.C., Swanson,P.C., Wu,X., Hsieh,C.L. and Lieber,M.R. (2004) A non-B-DNA structure at the Bcl-2 major breakpoint region is cleaved by the RAG complex. *Nature*, 428, 88–93.

Nucleic Acids Research, 2004, Vol. 32, No. 15 4549

- Nagawa, F., Hirose, S., Nishizumi, H., Nishihara, T. and Sakano, H. (2004) Joining mutants of RAG1 and RAG2 demonstrate impaired interactions with the coding-end DNA. *J. Biol. Chem.* [Epub ahead of print].
- Sleckman,B.P., Bassing,C.H., Hughes,M.M., Okada,A., D'Auteuil,M., Wehrly,T.D., Woodman,B.B., Davidson,L., Chen,J. and Alt,F.W. (2000) Mechanisms that direct ordered assembly of T cell receptor beta locus V, D, and J gene segments. *Proc. Natl Acad. Sci. USA*, 97, 7975–7980.
- 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*, 405, 583–586.
- Montalbano,A., Ogwaro,K.M., Tang,A., Matthews,A.G., Larijani,M., Oettinger,M.A. and Feeney,A.J. (2003) V(D)J recombination frequencies can be profoundly affected by changes in the spacer sequence. *J. Immunol.*, **171**, 5296–5304.
- Lee,A.I., Fugmann,S.D., Cowell,L.G., Ptaszek,L.M., Kelsoe,G. and Schatz,D.G. (2003) A functional analysis of the spacer of V(D)J recombination signal sequences. *PLoS Biol.*, 1, 56–69.
- Lee,G.S., Neiditch,M.B., Salus,S.S. and Roth,D.B. (2004) RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell*, **117**, 171–184.
- Fugmann,S.D., Villey,I.J., Ptaszek,L.M. and Schatz,D.G. (2000) Identification of two catalytic residues in RAG1 that define a single active site within the RAG1/RAG2 protein complex. *Mol. Cell*, 5, 97–107.
- Fugmann,S.D. and Schatz,D.G. (2001) Identification of basic residues in RAG2 critical for DNA binding by the RAG1–RAG2 complex. *Mol. Cell*, 8, 899–910.
- Purugganan, M.M., Shah, S., Kearney, J.F. and Roth, D.B. (2001) Ku80 is required for addition of N nucleotides to V(D)J recombination junctions by terminal deoxynucleotidyl transferase. *Nucleic Acids Res.*, 29, 1638–1646.
- Yu,K., Taghva,A. and Lieber,M.R. (2002) The cleavage efficiency of the human immunoglobulin heavy chain VH elements by the RAG complex. *J. Biol. Chem.*, 277, 5040–5046.
- 63. Olaru, A., Patterson, D.N., Villey, I. and Livak, F. (2003) DNA–Rag protein interaction in the control of selective D gene utilization in the TCR beta locus. *J. Immunol.*, **171**, 3605–3611.
- Gerstein, R.M. and Lieber, M.R. (1993) Coding end sequence can markedly affect the initiation of V(D)J recombination. *Genes Dev.*, 7, 1459–1469.
- Eastman,Q.M., Villey,I.J. and Schatz,D.G. (1999) Detection of RAG protein–V(D)J recombination signal interactions near the site of DNA cleavage by UV cross-linking. *Mol. Cell Biol.*, **19**, 3788–3797.
- Mo,X., Bailin,T. and Sadofsky,M.J. (2001) A C-terminal region of RAG1 contacts the coding DNA during V(D)J recombination. *Mol. Cell Biol.*, 21, 2038–2047.
- 67. Yu,K. and Lieber,M.R. (1999) Mechanistic basis for coding end sequence effects in the initiation of V(D)J recombination. *Mol. Cell Biol.*, **19**, 8094–8102.
- Jung, D., Bassing, C.H., Fugmann, S.D., Cheng, H.L., Schatz, D.G. and Alt, F.W. (2003) Extrachromosomal recombination substrates recapitulate beyond 12/23 restricted VDJ recombination in nonlymphoid cells. *Immunity*, 18, 65–74.
- Livak, F., Burtrum, D.B., Rowen, L., Schatz, D.G. and Petrie, H.T. (2000) Genetic modulation of T cell receptor gene segment usage during somatic recombination. J. Exp. Med., 192, 1191–1196.