CAAGCTTGGAAATTGATGAG-3' and 5'-CTGG-TACCAACTGAAAGAAG-3'; J₈1, 5'-GGAATTCCT-TCTGCAAATACCTTG-3' and 5'-AGAGGGAAT-TCCTATGAGCT-3'; V₈1, 5'-AATAGGAATTCTAC-TGATGGTGG-3' and 5'-ATGGATCCAATGCTCT-GTTTTTAG-3'; J.1, 5'GTGGGATCCTTGTCCAAA-GAC-3' and 5'-GTGGATCCACAGTCACTTG-3'; J₈2, 5'-GCCGGATCCAAAAAACATCTG-3' and 5'-GGGATCCACAAAGAGCTC-3'; V_HS107, 5'-CTG-GAATTCGAAACAAAGCTAATG-3' and 5'-TCTG-GAGGAGGCTTGGTACA-3'; V_H81X, 5'-CCTGT-GAATCCAATGAATACG-3' and 5'-GTGGAGTCTG-GGGGAGGCTTA-3'; and J_H3, 5'-TGCAGAGAATC-TTGGTCCTG-3' and 5'-ACTTCAAGCTTCAGTTC-TGG-3'. Twenty-five to 30 cycles of amplification were done with a programmable thermal controller (MJ Research, Inc.) (1 min at 94°C, 1 min at 52°C to 60°C, and 30 s to 1 min at 72°C). A second round of amplification was done starting with 2 to 15 µl of the first-round mixture [M. A. Innis, D. H. Gelfand, J. J. Sninsky, T. J. White. Eds., PCR Protocols: A Guide to Methods and Applications (Academic Press, San Diego, CA, 1989), pp. 21_271

- Y. Ichihara, H. Havashida, S. Mivazawa, Y. Kuro-21. sawa, Eur. J. Immunol. 19, 1849 (1989).
- D. M. Asarnow et al., Cell 55, 837 (1988)
- 23. D. M. Asarnow, D. Cado, D. H. Raulet, Nature 362, 158 (1993).
- S. Itohara *et al.*, *Cell* **72**, 337 (1993).
 S. Itohara, N. Nakanishi, O. Kanagawa, R. Kubo, S. Tonegawa, *Proc. Natl. Acad. Sci. U.S.A.* **86**, 5094 (1989)
- 26. D. Raulet et al., Immunol. Rev. 120, 185 (1991); J. P. Allison and W. L. Havran, Annu. Rev. Immunol. 9. 679 (1991).
- 27. M. M. Davis and P. J. Biorkman. Nature 334, 395 (1988)
- 28. S. Gilfillan, A. Dierich, M. Lemeur, C. Benoist, D.

Mathis, Science 261, 1175 (1993).

- To construct the targeting vector, we cloned a 29. 0.7-kb genomic Eco RI-Hind III fragment 5' of the TdT gene into pBluescript SK (Stratagene) containing a 2.7-kb Eco RI-Hind III PGK-HSV-tk fragment in the Sst I site. A 1.6-kb Eco RI-Bal II PGK-neo fragment was blunt end-ligated in the opposite transcriptional orientation into the Hind III site. To obtain the final targeting vector, we inserted an 8-kb genomic Xba I*–Eco RI and 7-kb genomic Xho I–Kpn I* fragment respectively, into the Eco RI site and Xho I site of this construct (asterisk denotes sites in cloning vector). ES cells (CCE, 2 × 107) were transfected with Kpn I-linearized targeting vector (20 µg) and selected with G418 (0.4 mg/ml) and gancyclovir (1 µM) (16). Three hundred and sixty-two selected clones were screened by Southern (DNA) blot analysis of Eco R I-digested genomic DNA probed with the Sal I-Xba I fragment 3' of the region included in the targeting construct to identify 43 TdT+/clones. Four TdT+/- clones were selected in increased G418 concentration (1.6 to 4 mg/ml) to obtain $TdT^{-/-}$ clones (17). Subclones of single (TdT^{+/-}) or double (TdT^{-/-}) mutated ES cells were injected into blastocysts of RAG-2-deficient blastocysts and transferred into B6CBAF1/J Jackson Laboratory) females (18).
- 30. R. D. Garman, P. J. Doherty, D. H. Raulet, Cell 45, 733 (1986).
- 31 G. D. Yancopoulos et al., Nature 311, 727 (1984). We thank G. Rathbun and E. Oltz for critically 32. reading this manuscript, D. Mathis for sharing unpublished data, J. Chen for help with FACS analyses, and R. Sunshine for providing a murine TdT probe. Supported by the Howard Hughes Medical Institute and NIH grant Al20047.

24 May 1993; accepted 12 July 1993

Mice Lacking TdT: Mature Animals with an Immature Lymphocyte Repertoire

Susan Gilfillan, Andrée Dierich, Marianne Lemeur, Christophe Benoist, Diane Mathis

In adult animals, template-independent (or N) nucleotides are frequently added during the rearrangement of variable (V), diversity (D), and joining (J) segments of lymphocyte receptor genes, greatly enhancing junctional diversity. Receptor genes from adult mice carrying a mutation in the terminal deoxynucleotidyl transferase (TdT) gene have few N nucleotides, providing proof that this enzyme is essential for creating diversity. Unlike those from normal adults, receptor genes from adult mutant mice show extensive evidence of homology-directed recombination, suggesting that TdT blocks this process. Thus, switchon of the TdT gene during the first week after birth provokes an even greater expansion of lymphocyte receptor diversity than had previously been thought.

 ${f T}$ he repertoire of B and T cell antigen receptors expressed in adult animals is more diverse than that in perinates (1). One major difference is the amount of N region diversity at the junctions of rearranged immunoglobulin (Ig) and T cell receptor (TCR) gene segments. N nucleotides are rare in V(D)J junctions from fetal or newborn animals, but constitute a major component of the diversity of Igs and TCRs from adults (2-7). This dissim-

ilarity may be due to differential expression of TdT. Terminal deoxynucleotidyl transferase catalyzes template-independent addition of nucleotides in vitro (8), and the amount expressed in vivo correlates with the degree of N region diversity in antigen receptors (9, 10). Another difference between adult and perinatal repertoires lies in the diversity of V-J, V-D, and D-J junctional sequences. Examination of large sets of fetal and newborn Ig and $\gamma\delta$ TCR sequences revealed overrepresentation of some junctions, coincident with short stretches of homology between abutted gene segments (3-5, 11-

SCIENCE • VOL. 261 • 27 AUGUST 1993

13). In the case of $\gamma\delta$ TCRs, certain dominant junctions (termed "canonical") are functionally significant because they give rise to the quasi-monoclonal receptors in specific anatomical locations such as the skin. Overrepresented joints were not generally observed in adult sequences, only in some of those lacking N nucleotides. Initially, the presence of dominant junctions of $\gamma\delta$ TCRs was attributed to cellular selection (5, 14), but a preference for rearranging at short stretches of homology is more probable (15, 16). Why such homology-directed recombination is pronounced in perinates but rare in adults is an open question.

One approach to better understanding the adult-perinate dichotomy is to artificially produce mature animals with repertoires having immature features. Thus, we generated, through homologous recombination in embryonic stem cells, a strain of mice lacking TdT (17). The mutation of TdT we obtained was an insertion of the neomycin gene into exon 4, as illustrated in Fig. 1 and confirmed by extensive Southern (DNA) blot analysis. Given the predicted location of the TdT active site and its presumed globular nature (18), exons 4 to 7 are probably critical for TdT function. No mRNA corresponding to regions 3' of the neomycin insert was detected in thymus RNA from homozygous mutant mice after polymerase chain reaction (PCR) amplification, for which we used a primer pair on the 3' side of the insertion; nor was any revealed by in situ hybridization of the appropriate probe to thymic sections (19). Abrogation of protein expression was confirmed by staining of thymocytes with a polyclonal antiserum to TdT (19).

Homozygous mutant $TdT^{-/-}$ mice breed well and appear healthy in a conventional animal facility, are of normal size, and do not have increased susceptibility to infection, as is common for immunodeficient animals in our colony. The mutants show no marked abnormalities in the major T or B cell compartments and are capable of mounting T and B cell responses to complex antigens like keyhole limpet hemocyanin and ovalbumin (20).

To evaluate the effect of a TdT deficiency on the lymphocyte repertoires of adult mice, we sequenced the V(D)J junctions of more than 300 rearranged Ig and TCR genes from adult animals (most from 6 to 8 weeks of age) (21). Representative sets of $V_{\gamma}3$ DNA sequences from total thymocytes (Fig. 2), V_H7183 DNA sequences from splenocytes (Fig. 3), and V₈8 RNA sequences from CD4⁺-CD8+CD3^{lo} thymocytes (Fig. 3) are shown. The enzyme TdT was responsible for the bulk of N region diversity because

Laboratoire de Génétique Moléculaire des Eucaryotes du CNRS et Unité 184 de Biologie Moléculaire de l'INSERM, Institut de Chimie Biologique, 11 rue Humann, 67085 Strasbourg Cédex, France.

the abundant N nucleotides found in rearranged genes of wild-type mice were absent in the mutant. Nonetheless, a few nucleotides not encoded in the germline do occur in these and our other data sets from $TdT^{-/-}$ mice; of the junctions that can be unequivocally assigned (the relevant V_{γ} , J_{γ} , D_{H} , J_{H} , V_{β} , D_{β} , and J_{β} genomic sequences being on record), about 3% had inserts of one to three nucleotides. Although PCR and sequencing errors may account for a few of these extra nucleotides, it is unlikely that such errors would be clustered so precisely at the joints. Similar inserts have been noted after rearrangement events involving other genes in nonlymphoid cell types. For example, about 10% of the junctions arising by circularization of transfected linear DNA have extra nucleotides (22). At least two mechanisms can be invoked to explain insertions in the absence of TdT: (i) incorporation of oligonucleotides during the joining process and (ii) nucleotide misincorporation by DNA polymerase at free ends during the fill-in and repair process that presumably occurs before joining. Some prokaryotic and eukaryotic DNA polymerases add one or two extra nucleotides when copying an oligonucleotide substrate (23). Despite the few template-independent nucleotides observed, we can now say with some certainty that TdT is the enzyme responsible for the difference in N region diversity characteristic of perinatal and adult antigen receptor repertoires, confirming earlier predictions (9, 10).

Analysis of the mutant mice also permitted us to evaluate the role of TdT in other aspects of the repertoire, in particular homology-directed recombination, as defined above. Rearranged V₂3-J₁ genes seem best suited for studying this aspect because three V-J junctions are significantly overrepresented in the fetal thymus. One junction, which comprises about 40% of the sequences, is in-frame and results in the invariant receptor found on essentially all intraepithelial lymphocytes in the skin [termed junction 1 by Itohora et al. (15)]; the two others, each comprising about 20% of the sequences, are out-of-frame (designated junctions 2 and 3). The dominance of these "canonical" junctions reflects a preference for recombining at short stretches of nucleotide homology (15, 16). Overrepresented joints are not a feature of the repertoire generated in adults. To determine whether TdT influences homology-directed recombination, we analyzed $V_{\gamma}3$ - $J_{\gamma}1$ sequences derived from pairs of TdT^{-7-} and TdT⁺ littermates. The sequences were obtained from either total thymocyte DNA or DNA from sorted populations of CD4⁻CD8⁻ and CD4⁺CD8⁺ thymocytes, enriched for cells committed to the $\gamma\delta$ or $\alpha\beta$ lineage, respectively. In all of the data sets from TdT^{-/-} animals (for example, Fig. 2), canonical joints were overused: junction 1 in 24 to 30% of the sequences, junction 2 in 20 to 23%, and junction 3 in 11 to 13% (Table 1). The data from sorted thymocyte populations

Fig. 1. Disruption of the TdT gene. A partial restriction map of the TdT gene, targeting vector, and mutant allele are shown. The positions of exons 5 and 6 are approximate (within 1.5 kb). For construction of the targeting vector, a 1.1-kb fragment of the neomycin resistance gene (pMC1neo-polyA, Stratagene) was cloned



into the Eco RV site in exon 4 of the 11-kb Xho I–Kpn I TdT genomic fragment, and the herpes simplex virus thymidine kinase gene was appended to the 3' end to allow double selection. RI, Eco RI; RV, Eco RV; K, Kpn I; X, Xho I; B, BgI II; and H, Hind III.

					10	nγ					
		TdT+						TdT			
V _v 3	Ρ	N	Ρ	J _v 1		V ₇ 3	Ρ	N	Ρ	J _y 1	
in frame				•		in'frame					
TGTGCCTGCTGGGAT				AGCTCAGGTTTT	#1	TGTGCCTGCTGGGAT				AGCTCAGGTTTT	#1
TGTGCCTGCTGGG <u>AT</u>				AGCTCAGGTTTT	#1	TGTGCCTGCTGGGAT				AGCTCAGGTTTT	#1
TGTGCCTGCTGG			т	ATAGCTCAGGTTTT		TGTGCCTGCTGGGAT				AGCTCAGGTTTT	#1
TGTGCCTGCTGG		TTA		AGCTCGGGTTTT		TGTGCCTGCTGGGAT				AGCTCAGGTTTT	#1
TGTGCCTGCTGGGATC		CCTC	AT	ATAGCTCAGGTTTT		TGTGCCTGCTGGGAT				AGCTCAGGTTTT	#1
out of frame						TGTGCCTGCTGGGAT				AGCTCAGGTTTT	#1
TGTGCCTGCTGGGAT			AT	ATAGCTCAGGTTTT		TGTGCCTGCTGGGAT				AGCTCAGGTTTT	#1
TGTGCCTGCTGGGAT			AT	ATAGCTCAGGTTTT		TGTGCCTGCT				CAGGTTTT	
TGTGCCTGCTGGGATC				TAGCTCAGGTTTT	#2	TGTGCCTGCTGGGATC				GCTCAGGTTTT	
TGTGCCTGCTGG			АT	ATAGCTCAGGTTTT		out of frame					
TGTGCCTGCTGGGA			АT	ATAGCTCAGGTTTT		TGTGCCTGCTGGGGTC				TAGCTCAGGTTTT	#2
TGTGCCTGCTGG				ATAGCTCAGGTTTT		TGTGCCTGCTGGGATC				TAGCTCAGGTTTT	#2
TGTGCCTGCTGGGA	(CCCCATCA	АТ	ATAGCTCAGGTTTT		TOTOCOTOCTOGATC				TAGCTCAGGTTTT	#2
TGTGCCTGCTGGGATCT		GGG		AGCTCAGGTTTT		TOTOCOTOCTOCCATC				TACCTCACCTTTT	#2
TGTGCCTGCTGGGATC		CGT	AT	ATAGCTCAGGTTTT		TGTGCCTGCTGGGATC				<u>magemence</u>	
TGTGCCTGCTGGGAT		т		TAGCTCAGGTTTT		TGTGCCTGCTGGGGATC				TAGCICAGGIIII	#2
TGTGCCTGCTGGG		GGG		AGCTCAGGTTTT		TGTGCCIGCIGGGATC				TAGCICAGGITTT	#2
TGTGCCTGCTGGG		G	AT	ATAGCTCAGGTTTT		TGTGCCTGCTGGGAT				ATAGCTCAGGTTTT	#3
TGTGCCTG		G	AT	ATAGCTCAGGTTTT		TGTGCCTGCTGGGAT				ATAGCTCAGGTTTT	#3
TGTGCCTGC		TC	т	ATAGCTCAGGTTTT		TGTGCCTGCTGGGAT				ATAGCTCAGGTTTT	#3
TGTGCCTG		GG	AT	ATAGCTCAGGTTTT		TGTGCCTGCTGGGAT				ATAGCTCAGGTTTT	#3
TGTGCCTGCTGGGATC		CGA	Ъ	ATAGCTCAGGTTTT		TGTGCCTGCTGGGAT				ATAGCTCAGGTTTT	#3
TGTGCCTGCTGGGATCT		GGGGG		GGTTTT		TGTGCCTGCTGG			AT	ATAGCTCAGGTTTT	
TGTGCCTGCTGGG		TC		ATAGCTCAGGTTTT		TGTGCCTGCTGG			AT	ATAGCTCAGGTTTT	
TGTGCCTGCTGGGAT		GAGGGG		TAGCTCAGGTTTT		TGTGCCTGCT				AGCTCAGGTTTT	
TGTGCCTGCTG		AAGGC		AGCTCAGGTTTT		TGTGCCTGCT				AGCTCAGGTTTT	
TGTGCCTGCTGGG		GGAA		ATAGCTCAGGTTTT		TGTGCCTGCTGG				ATAGCTCAGGTTTT	
TCTGCCTCCTGGG		GC		ATAGCTCAGGTTTT		TGTGCCTGC			AT	ATAGCTCAGGTTTT	
TGTGCCTGCTGGGA		с	AT	ATAGCTCAGGTTTT		TGTGCCTGCTGG				CTCAGGTTTT	
TGTGCCTGCTGGGA		AAN		ATAGCTCAGGTTTTT		TGTGCCTGCTGGGATC			AT	ATAGCTCAGGTTTT	
TGTGCCTGCTGGGATCT		TTGG		AGCTCAGGTTTT		TGTGCCTGCT				TAGCTCAGGTTTT	
TGTGCCTGCTG		AAGGC		AGCTCAGGTTTT		TGTGCCTGCT			АТ	ATAGCTCAGGTTTT	
TGTGCCTGCTGGG		GT	AT	ATAGCTCAGGTTTT		TGTGCCTGCTGG				CTCAGGTTTT	
TGTGCCTGCTGGG		GTC	т	ATAGCTCAGGTTTT		TGTGCCT			AT	ATAGCTCAGGTTTT	
TGTGCCTGCTGGG		CT		ATAGCTCAGGTTTT		TOTOCOTOCTOC		Ψ		CTCACCTTTT	

Fig. 2. Representative sequences of $V_{\gamma}3-J_{\gamma}1$ junctions from age-matched adult mice, either wild-type TdT⁺ (o/+ or +/+ genotypes) or mutant TdT^{-/-} (o/o genotype). The TCR_Y sequences were derived from total thymus DNA. The assignment of N nucleotides was based on the known sequences of the germline elements. Homologies of two or more nucleotides shared between joined segments are underlined. We did not score putative homologies of a single base because their significance is difficult to ascertain: in the absence of homology-induced joining, the probability of any junction containing a single-nucleotide "homology" is 38%. Homology was assigned assuming possible P inserts of up to five bases. P inserts are template-dependent nucleotides added to the ends of recombining gene segments; they are palindromic to the terminal few nucleotides of the unadulterated coding joint (*5*). Repeated junctions are shown only if they were obtained in independent amplifications from at least two mice. For $V_{\gamma}3$, the canonical junctions are numbered according to (*15*).

SCIENCE • VOL. 261 • 27 AUGUST 1993

indicate that this is true of cells from the $\alpha\beta$ as well as the $\gamma\delta$ lineage. In contrast, these joints are rare in the rearranged $V_{\gamma}3$ - $J_{\gamma}1$ genes from TdT⁺ animals. According to current models of the rearrangement process (24), their use would result in genes lacking N region diversity because homology alignment would be the last step before ligation. The large major-

REPORTS

ity of sequences from TdT^+ animals carry N nucleotides, and less than half of those that lack N additions have a canonical junction. It might also be worth mentioning, should the currently preferred model of rearrangement prove incorrect, that the sequences with N nucleotides almost never have the same cutting points as those used to generate the canonical junctions.

Extensive use of homology-directed recombination has also been documented for the Ig heavy chain genes of perinatal mice (3, 4, 11, 12). Therefore, we analyzed V_H7183 -J_H4 and V_HJ558 -J_H2,3 sequences to see whether our observations on rearranged $V_{\gamma}3$ genes could be generalized (Fig. 3). The representative data set shows the frequent use of a few junctions—particularly certain D-J joints—in the sequences from TdT^{-/-} mice, all of which are situated at short stretches of nucleotide homology. Overused junctions are rare in sequences derived from TdT⁺ animals.

Analysis of rearranged TCRB genes

Fig. 3. Representative sequences of V_H7183-D-J_H4 and V_β8-D-J_β junctions from age-matched mice, either wild-type (TdT+) or mutant (TdT^{-/-}) as in Fig. 2. Immunoglobulin heavy chain sequences came from total spleen DNA and TCRB sequences from CD4+CD8+CD3lo thymocyte RNA. Assignment of N nucleotides was as in Fig. 2, except for the several members of the $V_{H}7183$ family for which the germline sequences are not known; in this instance, we relied on the fact that those V_H7183 family members germline whose sequences are known end with either . . . GCAAGACA or ... GCAAGAGA (12). Underlined bases are homologies of two or more Homologs nucleotides. were scored as in Fig. 2. The impression that the D segment is generally shorter in TdT-7- mice is an illusion caused by the routine placement of homologous nucleotides in the V or J seaments.

revealed a different situation (Fig. 3 and Table 1). Homology-directed recombination is not a dominant feature of the V_{β} repertoire in either TdT^{-/-} or TdT⁺ mice. In both types of mice, junctions are heterogenous: almost every sequence has a different V-D and D-J joint. That such heterogeneity is a feature common to V_{β} sequences, whether or not they show N region diversity, was confirmed by examination of published perinatal and adult V_{β} ,6,8 and 17a genes (6, 7, 10).

Two conclusions can be drawn from the data on homology-directed recombination. First, the presence of TdT somehow blocks this process. This could be a direct effect at the protein level—for example, the binding of TdT to the ends of segments might inhibit attachment of some protein needed to hold together short homologous stretches of nucleotides. Our data and those of others (4, 12) are consistent with this possibility: adult sequences devoid of N region diversity consistently have fewer dominant junctions

than do perinatal sequences lacking N nucleotides. The blocking of homologydirected recombination of TdT could also be an indirect effect; the addition of nucleotides to segment ends might set stretches of homology too far apart for the recombination machinery to join them effectively. This possibility has been suggested previously (12) and is supported by the preponderant use of homologies residing close to the ends in both Ig and TCR $\gamma\delta$ genes. A similar preference for homologous stretches near the extremities was observed after circularization of transfected linear plasmids in nonlymphoid cells (25).

The second point is that in animals whose receptor variety is severely restricted by the absence of N region diversity, some receptor gene families $(V_{\gamma}3-J_{\gamma}1 \text{ and IgH})$ use a strategy to further reduce variety, whereas the TCR genes do not. The D_{β} and J_{β} segments actually seem to be organized to avoid homology-directed recombination. Few stretches, even dinucle-

				TdT+				lg	H				TdT -				
V _H 7183	Р	N?	Р	D	Р	N	Р	J _H 4	V _H 7183	Р	N?	Р	D	Р	N	Р	J _H 4
GCAAG			GA	TCTATGATGGTTACTAC	G			ATGCTATG	gcaaga <u>ga</u>	T			GG				TTACTA TGCTATG
GCAAGAGA	TC	AGG		ATGATTACGAC		AGGAT		TTACTATGCTATG	GCAAGAGA	т			GG				TTACTA TGCTA TG
GCAAGAGA	TC	GG	A	TTTATTACTACGGTAGTAG				CTA TGCTATG	GCAAGACA	TG			ATGG				TTACTA TGCTATG
GCAAGAGA		GNG		TACTACGGTAGTAGCT		CCG		TATGCTATG	GCAAGACA	TG			GTA				ACTA TGCTATG
GCAAGA		TCCGC		AC		cc		TACTATGCTATG	GCAAGAGA				CTATAGTA				ACTA TGCTA TG
GCAAGACA		GGGT		TCTACTATGATTACGA		G		TACTATGCTATG	GCAAGAG				CCTACTATAGGTACG				ACTATGCTATG
GCAAGAGA		A		GGGA				ATTACTATGCTATG	GCAAGA				TGATTACG				ACTATGCTATG
gcaagaga	т	GG		AT <u>TACC</u> GTAGTAGCT	IC.			G	GCAAGAGA	т			GATTACG				ACTATGCTATG
GCAA		ACC		TCTACTATGATTACG		GTCG		TGCTATG	GAAAGAC	-			TTACTAAGGTAGTAG				CTATGCTATG
GCAAGACA	TG	ATG		ATGGTAACTAC	G	GGGG		CTATGCTATG	GCAAGACA				CTACGGTAGTAG				CTATCCTATC
GCAAG		GGGGCG		TIACTACGGTAATAG	~	ACC	-	ACTING	GCAAG		GG		ATGGTAAC				TACTATGCTATG
GCAAGAG		1000		TGATTACGAC	G	AGG		TOCTATO	GCAAGAGA	т			GATGGTTAC				TACTATGCTATG
CCANGAG		CCLCC		GOTACGAC	c	N		CTATCCTATC	GCAAGAGA	Ŧ			TACTACGG				TATG
CANG		CCCCCC		TTACT	G	0000		TATGCTATG	CONCA	-			TOOTTACTAC	G	c		TACTATCCTATC
GCAAGACA		GGGGTC		CGGCTAC		TTCC		CTATG	GCAAGA		TTC:		TGATTACG			АТ	ATTACTATCCTATC
GCAAGAC		GTCT	А	TTTATTACTACGGTAGT		TG		ATGCTATG	GCAAGAGA	т			TACTACGGTAG				TTACTATGCTATG
GCAAG		GGT	Ä	TTTATTACTACGGTAGTAG		NCCC	АΤ	ATTACTATGCTATG	GTAA	-							TGCTATG
GCAAGAGA		CGG		TTACTACGGTAGTAGC		GGG		TATGCTATG	GCAAGAGA	т			GGTTACTAC				GCTATG
GCAAG		с	A	TCTACTATGATTAC		CA		TACTATGCTATG	CONGICA	100			ATGG		A		TACTATCCTATG
GCAAGAGA	т			GGTAACTAC				GCTATG	CCARCACA			G	CCTACTATAGTAACTAC	G			CTATCCTATC
GCAAGAGA	т	GGG		CTATAGTAA		AGGGGG	2	TGCTATG	- unionan				comon no mo	5			
									•								

TCR β

NORMATIA A CAGG TUTN CAMCHORAN L MODINTIDA A GAMACTICORA CAGG AACACHORAN ACACHORAN	V _B 8.1	Р	Ní	Р	D	Р	Ν	Р		J	V _B 8.1	Р	Ν	Р	D	Р	Ν	Р		J
MACHATINA A CAGG TUTN CAAACTCCCAR 1.2 ACCAT	AGCAGTGA				GACAGGGG	•		•	AAACACAGAA	1.1	GGCAGTGAT		••		0			•	AACACAGAA	1.1
ACCONTING A C CALAGES ACCASTA COCOST C CALAGES ACCASTA C COCAST C COCAST C COCAST C COCAST C COCAST C COCAST C CCAST A TUCTBOANT C COCAST A TUCTBOANT C COCASTA C COCASTA C COCASTA C COCASTA C COCASTA C <	AGCAGTGA				CAGG		TCTN		CANACTCCGAC	1.2	AGCAGT				ACAGG				AAACACAGAA	
ACCM ACT CHOG A CONSTANT CACAGG ACCMG CACAGG ACTCCCAC ACCMG C C CGACAGG CATT TOTSMARAT ACGOGTANT CACGOGTANT CACGOGTANT TOTSMARAT TOTSMARAT ACGOGTANT TOTSMARAT <td>ACCACTGA</td> <td></td> <td></td> <td></td> <td>Α</td> <td></td> <td></td> <td>G</td> <td>CAAACTCCGAC</td> <td></td> <td>AGCAGTGA</td> <td></td> <td></td> <td></td> <td>GGGG</td> <td></td> <td></td> <td>G</td> <td>CAAACTCCGAC</td> <td>1.2</td>	ACCACTGA				Α			G	CAAACTCCGAC		AGCAGTGA				GGGG			G	CAAACTCCGAC	1.2
MACCART CC GGARCA COLV A THACTMONAT TICCMINAG GGARC A THACTMONAT TICTORGANT ACCART TICTORGANT CL ACCART TICTORGANT CL ACCART TICTORGANT CL ACCART TICTORGANT CL ACCART	AGCAG		ACT		CAGG		A		CTGGAAAT	1.3	AGCAG				GGACAGGG				ACTCCGAC	
LOCAD CACAGOGO A THACTATOR 2.1 ACCAGTRATO G TTOT 2.1. AGCORTANTO CA. G CAGOG CAGOGO ACTACTANA 2.3 ACCAGT CACAGO CALAGOG CALAGOG CALAGOGO CALAGOGO CALAGOGO CALAGOGO CALAGOGO CALAGOGO CALAGOGO CALAGOGO CALAGAC 2.5 ACCAGTRATO CACAGTRATO CACAGTRATO CACAGTRATO CACAGTRATO ACCAGTRATO CACAGTRATO CACACACA <td>AGCAGT</td> <td></td> <td>c</td> <td>cc</td> <td>GGGACA</td> <td></td> <td>CGAT</td> <td>А</td> <td>TTCTGGAAAT</td> <td></td> <td>TCCANCAG</td> <td></td> <td></td> <td></td> <td>GGGAC</td> <td></td> <td></td> <td>λ</td> <td>TTCTGGAAAT</td> <td>1.3</td>	AGCAGT		c	cc	GGGACA		CGAT	А	TTCTGGAAAT		TCCANCAG				GGGAC			λ	TTCTGGAAAT	1.3
Inconstruction CG Choose CA ACTUCAGAN 2.3 ACCAG GEACAGGE CLAGGE CLAGGE <th< td=""><td>AGCAG</td><td></td><td>-</td><td></td><td>GACAGGGG</td><td></td><td></td><td>A</td><td>TAACTATGCT</td><td>2.1</td><td>AGCAGTGATG</td><td></td><td></td><td></td><td>G</td><td></td><td></td><td></td><td>TGCT</td><td>2.1</td></th<>	AGCAG		-		GACAGGGG			A	TAACTATGCT	2.1	AGCAGTGATG				G				TGCT	2.1
ACCARTMANC CA G CTOGGG ACCARGAC 2.5 ACCARTAS CAGGOTA B.2 ACCONTANC G GGACAGGG TTCTGAMAT 1.3 ACCONTANC CAGGOTGA CAGGOTGA CAGGOTGA CAGGOTGA CAGGOTGA CAGGOTGA CAGGOTGA CAGGOTGA C ACCONTANTO CACAGGOG CAGAN A ACCGOTGANC C TOGG CC CTUTGGAAN A TTUTGGAAN ACCONTANTO CACAGGOGG CAGAN A ACCGOTGANC C GACATTOGGA AC TTUTGGAAN A CTUTGGAANT ACCGOTGAN C ACCGOTANT C ACCGOTANT C ACAGOTANT ACCGOTANT C ACCGOTANT ACCGOTANT C GGAC ACAGOTANT ACCGOTANT ACCAGTANT ACCAGT	AGCGGTGATG	CA	CG		CAGGG		CA		AGTGCAGAA	2.3	AGCAG				GGACAGGG				CAGAA	2.3
8.2 ACCOSTGATO GGGACAGGG TTCTGGANAT 1.3 ACCOSTGAT 8.2 CAGGAGGG CAGAN 1.4 ACCOSTGAT ACCOSTGAT ACCOSTGAT CC CAGAN 1.4 ACCOSTGAT CAGAN 1.4 ACCOSTGAT CC TGGG CC CTMGCT ACCOSTGAT GACACA ACTOTGANAT CAGAN 1.4 ACCOSTGAT C GACACA ACCOSTGAT GGACA TTGTGANAT ACCOSTGAT GGACA TTGTGANAT ACCOSTGAT GGACA TTGTGANAT ACCOSTGAT GGACA TTGTGGANAT ACCOSTGAT GGACA TTGTGGANAT ACCOSTGAT GGACA TTGTGANAT ACCOSTGAT GGACA TTGTGGANAT ACCOSTGAT GGACA TTGTGGANAT ACCOSTGAT GGACA TTGTGGANAT ACCOSTGAT GGACA TTGTGGANAT ACCOSTGAT GGACA TTTTGGANAT ACCOSTGAT GGACA TTTTGGANAT ACCOSTGAT GGACA TTTTTGANAT ACCOSTGAT GGACA ACCOSTGAT ACCOSTGAT GGACA ACCAGA TT ACCAGAT ACCAGAT ATT ACCAGTGA ACCA	AGCAGTGATG	CA	G		CTGGGG				ACCAAGAC	2.5	AGCAGTGATG								GTCAAAAC	2.4
ACCONTANTS G GGACAGGG TTCTERLANT 1.3 ACCONTANT 8.2 ACCONTANTS TGGG C AACGATANT C AACCANTANT AACCANTANTANT AACCANTANT AACCANTANT AACCANTANT AACCANTANT AACCANTANT AACCANTANT AACCANTANT AACCANTANT AACCANTANTANT AACCANTANT AACCANTANTANT AACCANTANTANT AACCANTANTANTANTANTANTANTANT AACCANTANTANTANTANTANTANTANTANTANTANTANA	8.2										AGCAGT <u>GA</u>				CAGGGGG				CAAGAC	2.5
ACCONTRAT C GACAGOGG C C CTATOCT 1.4 ACCONTA C GACAGOGG G CC CAGAA 1.1 ACCONTA C C TOGG C C CTATOCT 1.4 ACCONTA C GACAGOGG C TATOCT A ACCONTA C GACTOGGG AAC CTATACT A ACCONTA C GACAGOGG C C CAGAA 1.1 ACCONTA C C GACAGOGG C C CAGAA 2.5 ACCONTA C C GACAGOGG ACA CAGAGOG C C CACAGOGG C C ACCONTA C C GACAGOGG ACA GAAA 1.1 ACCANTA C C GACAGOGG ACA GAAA A A ACCANTA C C GACAGOGG ACA GAAAA 1.1 ACCANTA C C GACAGOGG ACA GAAAAA 1.1 ACCANTA C C GACAGOGG ACA GAAAAAAAAAAAAAAAAAAAA	ACCOUNTS		0		CCCACACGC				TTOTOCAAAT	1.3	8.2									
ACCORTIGATIO C. TOGG C. CINNECCE 1.1 ACCORTIGA C. C. CAA A. CI CINNECCE 1.2.1 ACCORTIGA C. GAACA TTUGAAAT 1.3 ACCORTIGA C. GAACA A. CI CINNECCE 1.2.1 ACCORTIGA C. GAACA TTUGAAAT 1.3 ACCORTIGAT C. GAACA A. CI CINNECCE 1.2.1 ACCORTIGAT C. GAACA C. CAACA A. CI CINNECCE 1.2.1 ACCORTIGAT C. GAACA GAACA A. CI CINNECCE 1.2.1 ACCORTIGAT C. CI CAGGGG A.A. CI CINNECCE 1.2.1 ACCORTIGAT C. CI CAGGGG A.A A. A CAAA TI A.ACCAACACAC 2.5 ACCORTIGAT C. CI GAACAGGGG A. CIGAAAT 1.3 ACCORTIGAT C. GAACAGGGG A. CIGAAAT 1.3 ACCORTIGAT A. CCCCT G. GAACAGGGG C. CIGAAAT 1.3 ACCORTIGAT A. CCCCT G. GAACAGGG C. CIGAAAT 1.3 ACCORTIGAT A. CCCCT G. GAACAGGG C. CIGAAAT 1.4 ACCORTIGAT A. CCCCT G. GAACAGGG C. CIGAAAT 1.4 ACCORTIGAT A. CCCCT G. GAACAGGG C. CIGAAAT 1.4 ACCORTIGAT G. GAACAGGGG T. TICICINTAT 1.6 ACCORTIGAT G. GAACAGGGG C. CIGAAAT 1.4 ACCORTIGAT G. GAACAGGGG C. CIGAAAT 1.4 ACCORTIGAT G. GAACAGGGG C. CIGAAAT 1.4 ACCORTIGAT G. GAACAGGGGG T. CICAACACA ACCAGTGAT G. GAACAGGGGG T. CICAAACACACACACACACACACACACACACACACACAC	AGCGGTGATG		G		GOGACAGGG				ALCON	1 4	ACCOUNTS				Changed				CAGAA	1 1
ACCONTRA C C C C C C C C C C C C C C C C C C C	ACCORTINE	~			TOGO		~		CTATGCT	2.1	AGCOGIGAIG				C				AACTICCGAC	1.2
Naccorrown C	AGCGGTG	C	000		CA				TATGCT		AGCGGTG				CCACA				TOTOGAAAT	1.3
Maccorrow A CT ADTCALGUAS A CT ADTCALGUAS	AGCGGTGA		c		GACTGGGG		AAC		CTATGCT		AGCGGTGATG	c			GGG				TAAT	1.6
Inscissment C, A, G GGGCA COLAGIC 2: Inscissment C, A GGACAGGGGC THEORY CACCORD AGCG GGGCA GGGCA ACA TT TGAA ACCONTRANC CACCORD ACCONTRANC CACCORD ACA TT MCCAGTGA ACA TT MCCAGTGA ACC ACCA TT MCCAGTGA ACA TT MCCAGTGA ACCA TT MCCAGTGA ACCAGTGA ACA TT MCCAGTGA ACCAGTGA ACCAGTGA <td< td=""><td>AGCGGTGAT</td><td></td><td>T</td><td></td><td>GGACA</td><td></td><td>A</td><td>CT</td><td>AGTGCAGAA</td><td>2.3</td><td>AGCGGTGAT</td><td>č</td><td></td><td></td><td>000</td><td></td><td></td><td></td><td>AAT</td><td></td></td<>	AGCGGTGAT		T		GGACA		A	CT	AGTGCAGAA	2.3	AGCGGTGAT	č			000				AAT	
LACCOGTINATION GOGGCA GOGGCA A CTCCTATGAA 2.6 ACCOGTINATIO TA ACA TT TGAA ACA TT MCCCGTONATION TGGGG ACA TT MCCCGTONATION TT MCCCGTONATION TC CCMAGAAC ACA MCCCGTONATION TC CCMAGAACACC MCCCGTONATION ACA TT MCCCGTONATION TC CCMAGACACC MCCCGTONATION ACA TT MCCCGTONATION TC CCMAGACACC MCCCGTONATION ACA TT MCCCGTONATION ACA ACCAGTONATION ACA TT MCCCGTONATION ACA ACCAGTONATION ACA AA ACA AA ACA ACA ACA ACA AA ACA AA ACA ACA AA ACA AA ACA AA ACA AA ACA ACA AA ACA ACA ACA AA ACA AA ACA AA ACA ACA AA ACA AA ACA ACA AA ACA ACA AA ACA ACA AA ACA ACA	AGCGGTGATG	CA	Ğ		GGGAC				CAAGAC	2.5	AGCAGTG				GGACAGGGGGC				TATGCT	2.1
ACCOGTENTIG TA ACA TT TGAA ACC ACA TT AACA TT ACA AC	AGCGG		GGGCA		GGGACA		λ		CTCCTATGAA	2.6	AGCGGTGATG				GGGG				CACCGGG	2.2
8.3 ACCONTINAT TGC AACCARGAC AGCASTEAN CCT CAGAGGG ACCARGA 1.1 MCCOSTEANT GGACTOGG T CCLAGAGA AGCASTEAN C GACAG AA AGAA MCGASTEANT CAGAGGGG T CCLAGGGG MCCARGEGG MCCARGEGGG MCCARGEGGG MCCARGEGGG MCCARGEGGGG MCCARGEGGGGG MCCARGEGGGGG MCCARGEGGGGG MCCARGEGGGGG MCCARGEGGGGG MCCARGEGGGGG MA MCCARGEGGGG MCCARGEGGGGG MCCARGEGGGGG MCCARGEGGGGG MCCARGEGGGGG MCCARGEGGGGG MCCARGEGGGG MCCARGEGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGGG MCCARGEGGGGGGG MCCARGEGGGGGGG MCCARGEGGGGGGG MCCARGEGGGGGGG MCCARGEGGGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGG MCCARGEGGGGGGG <	AGCGGTGATG		TA		ACA		TT		TGAA		AGCG				АСА			TT	AACCAAGAC	2.5
ACCARTGAT CCT CAGGGG ACIGAT ACCARTGA CGACTGGG T CCLARACACC ACCARTGA GT GGACA AA AGAA AGAA AGAA AGAA CTGGGGGG T TGGGGGG T TAA 2.6 AGCARTGA C GACAGGGG AGA GGALATI 1.3 MCCGGTGATG AGCAGTGA AA AGCARTGA C GGACAGGG CC CACAGGGG AGAA AA AGCARTG CC GGGACAGGG CC CACAGGGT 1.5 AGCARTGA AGCAGTGA GGGGGC G AA AGCARTGA ACCARTGA T T TCCTAGGCT AGCAGTGA GGGGGC G CCAA AGCARTGA AGCARTGA T T TCCTAGGCT AGCAGTGA GGGGGC G CAA AGCARTGA CC GGGGACAGGGG T TOCT 2.1 AGCAGTGA GGGGACAGGGC AGCAGTGA AGCARTGA GGGGGGGGG A TAACTTRGCT AGCAGTGA GGGGACAGGGG AGCAGTGA GGGGGACGG CCAA AGCAGTGA GCCCC GGGGGGGG ATATCTRTGCT AGCAGTGA GGGGACAGGGG T C AGCAGTGA GCCCC GGGGGGGG ACCAGTGA GGGACAGGGGC T C <td< td=""><td>8.3</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>AGCGGTGAT</td><td></td><td></td><td></td><td>TGG</td><td></td><td></td><td></td><td>AACCAAGAC</td><td></td></td<>	8.3										AGCGGTGAT				TGG				AACCAAGAC	
ACCARTGA GCL GARACA NA ACCARTGA CTGGGGGG TCAA 2.6 ACCARTGA C GARACA ACA GARATA I.A ACCORTGATG CTGGGGGG TCAA 2.6 ACCARTGA C GARACAGAGA ACA GARATA I.A ACCORTGATG ACCOGTGATG ACCOGTGATG ACCARTGA GGGGCC G CCANACAA ACCARTGA GGGGCC G TT ACCARTGA GGGACAGGGGC G TT ACCARTGA GGGACAGGGGC G TT ACCARTGA GGGACAGGGGC G TT ACCARTGA GGGACAGGGGC G TT ACCARTGA ACCARTGA GGGACAGGGGC G TT ACCARTGA ACCARTG	ACCACTERAT		COT		CAGGGG				ACAGAA	1.1	AGCGGTGA				GGACTGGG		т		CCAAGACACC	
AGCARTGA C GACAGGGGG AGA GGUAT 1.3 AGCOGTQATG GGGGC AGCAGTG AGCOGTQATG AGCAGTG AG	ACCAGIGAT		CT.		GGACA		**		AGAA		AGCGGTGATG	CA			CTGGGGGG				TGAA	2.6
AGCAGTG C GGGCAGGG C C CAACCAGGT 1.4 AGCAGTG C C GGGCAGGG C C CAACCAGGT 1.4 AGCAGTGAT ACCCT C C GGGCAGGG T T TTCCTATAAT 1.6 AGCAGTGAT C G GGGCAGGG C T T TTCCTATAAT 1.6 AGCAGTGAT C G GGGCAGGG C T T TTCCTATAAT 1.6 AGCAGTGA G GGGCAGGG C T T TTCCTATAAT 1.6 AGCAGTGA G GGGCAGGG C A TAACTATGCT A AGCAGTGA C G GGGCAGGG C A TAACTATGCT A AGCAGTGAT A GGGGGGG G G AAC AGCAGTGAT C GGGGGGG G G AAC AGCAGTGAT C GGGGGGG A TTACTATGAT 2.6 AGCAGTGA G GGGCAGGG C T C CAAACC 2.5 AGCAGTGA A GGGCAGGG C TATGAA 2.6 AGCAGTGA A GGGCAGGG C TATGAA 2.6 AGCAGTGA A GGGCAGGG C TATGAA 2.6 AGCAGTGA A GGGCAGGGG T TT C CTCTATGAA A AGCAGTGA A GGGGGGG C TTT C CTCTATGAA A AGCAGTGA C C GGGGGGG C TATGAA 2.6 AGCAGTGA C C GGGCAGGGG C TATGAA 2.6 AGCAGTGA C C GGGCAGGGG C TATGAA 2.6 AGCAGTGA C C GGGCAGGGG C TATGAA 2.6 AGCAGTGA G C CC GGGCAGGGG C TATGAA 2.6 AGCAGTGA C C GGGGGGGG C CTCT CTCTCTATGAA ACCAGTGA GGGGGGG C CAAGAC 2.5 AGCAGTGA C C GGGGGGG C TATGAA 2.6 AGCAGTGA C C GGGGGGG C CTCT CTCTCTATGAA A AGCAGTGA C C GGGGGGG C CTCT CTCTCTATGAA ACCAGTGA GGGGGGG C CAAGAC 2.5	AGCAGIGA		C C		GACAGGGGG		AGA		GGAAAT	1.3	AGCGGTGATG								AA	
AGCARTGA CC GGGACAGGG CC CAACCAGGE 1.5 AGCARTGA ACCARTG T CCAGGET ACCARTGA AGCARTGA TG T TCCTYATAT 1.6 AGCARTGA AGCARTGA GGGGC G CCAACGAA (A AGCARTGA TG T TCCTYATAT 1.6 AGCARTGA GGGACAGGGG T TGCT 2.1 AGCARTGA GGGACAGGGG ACTARTGCT AGCARTGA GGGACAGGGG ACTARTGCT AGCARTGA GGGGGGGG A AGCARTGA GGGGGGGG C AGCARTGA GGGGGGGG ACCARTGA GCCARCGA GGCACTGGGGG ACCARTGA AGCARTGA GGGGGGGG ACCARTGA AGCARTGA GGGGGGGG ACCARTGA AGCARTGA GGGGGGGG ACCARTGA AGCARTGA GGGGGGGGG ACCARTGA AGCARTGA GGGGGGGGG TT AGCARTGA GGGGGGGGG TT AGCARTGA GGGGGGGGG TT AGCARTGA GGGGGGGGGG TT AGCARTGA GGGGGGGGG TT AGCARTGA GGGGGGGGG TT AGCARTGA GGGGGGGGG TT AGCARTGA GGGGGGGG	AGCAGIGA		C		00000				CGAA	1.4	AGCGGTGATG								AA	
ACCARTENT CCCT CA T CCANGES C CALL ACCARTENT CACCT CA T T CCANCENTANT I.6 ACCARTENTS CA GACARTENTS I.6 ACCARTENTS CACARTENTS CACART	AGCAGIG		~		00000		œ		CARCEAGOOT	1.5	8.3									
ACCARTGATG ACCARTGATG C TG GGGACAGGGG T TTCCTNTNAT 1.6 ACCARTGATG C G GGGACAGGG T TCCTCTNTNAT 1.6 ACCARTGATG C G GGGACAGGGGC T TCCTCTNTNAT 1.6 ACCARTGATG G GGGACAGGGGGC C ACCARGA ACCARTGATG G GGGACAGGGGGC C ACCARGA ACCARTGAT A GGGGGGG C G G AAC ACCARTGATG C GGGACAGGGGGC C TC CANAAC 2.4 ACCARTGATG GGGACAGGGGGC G G AAC ACCARTGATG C GGGGGGGG C TC CANAAC 2.5 ACCARTGATG GGGACAGGGGGC C TC ACCARTGATG GGGACAGGGGGC C TC ACCARTGATG GGGACAGGGGGC C TC ACCARTGATG GGGACAGGGGGG C TC ACCARTGATG GGGACAGGGGGG C C ACCARTGATG C GGGACAGGGGGG C C ACCARTGATG GGGACAGGGGGG C C ACCARTGATG GGGACAGGGGGG C C ACCARTGATG GGGACAGGGG C C ACCARTGA ACCARTGATG GGGACAGGGG C C ACCARTGA ACCARTGA A C ACCARTGA A C ACCARTGA A C ACCARTGA A C ACCARTGA A C ACCARTGA C C ACCARTGA C	AGCAGIG		ACCOR		Ch		T		CONCECT	1.5	10000				00000	~			COL NOCK N	1.4
ACCAGTONITS CA. G. CLANCIAL G. CLANCIAL ACCAGTO GA CLANCIAL GA CLANCIAL ACCAGTO GA CLANCIAL GA CLANCIAL ACCAGTO GA CLANCIAL ACCAGTO GA CLANCIAL ACCAGTO GGALACAGGGGGC ACCAGTO ACCAGTO GGALACAGGGGGC ACCAGTO ACCAGTO GGALACAGGGGG ACCAGTO ACCAGTO GGALACAGGGGGC 1.6 ACCAGTOR GGALACAGGGG T ACCAGTOR GGALACAGGGG T. ACCAGTOR GGALACAGGGGG C ACCAGTOR GGALTGGGGGGGC T ACCAGTOR GGALTGGGGGG ACCAGTOR GALTGGGGGG T T ACCAGTOR GGALTGGGGGGC ACCAGTOR GALTGGGGGG T T ACCAGTOR GGALTGGGGGGC ACCAGTOR GALTGGGGG T T ACCAGTOR GGALTGGGGGGC T ACCAGTOR GALTGGGGG CCCCC CAGGGGG CCCCC CAGGGGG CCCCC CAGGGGG CCCCCC CAGGGGGG CCCCCC CAGGGGGG CCCCCC CACAGTOR CACAGTOR CACAGTOR CACAGTOR CACAGTOR CACAGTOR CACAGTOR CACAGTOR CACAGTOR CACAGTOR <td>AGCAGTGATG</td> <td></td> <td>ACCCI</td> <td></td> <td>TG</td> <td></td> <td>- </td> <td></td> <td>TTCCTATAAT</td> <td>1.6</td> <td>AGCAGIG</td> <td></td> <td></td> <td></td> <td>GGGGC</td> <td>G</td> <td></td> <td></td> <td>CCAACGAA</td> <td>1.4</td>	AGCAGTGATG		ACCCI		TG		- 		TTCCTATAAT	1.6	AGCAGIG				GGGGC	G			CCAACGAA	1.4
AGCAGTGA GGACTGGGGG ANCTATOCT AGC GGGACAGGGGGC AUCAGGCT 1.5 GCCAGCAG GCCC GGGGGG A TANCTATOCT AGC GGGACAGGGGGC GGACAGGGGGC T.C.C AGCAGTGA A GGGGGGG C CALAAC 2.4 AGCAGTGA GGGACAGGGGC TOCT 2.1 AGCAGTGATG C GGGGGGGG G AAC C GGGACAGGGGC TOCT 2.1 AGCAGTGATG GGGGGGGG G AAC C C GGACAGGGGC TOCT 2.1 AGCAGTGATG GGGGGGGG AACAAC C AGCAGTGA GGACTGGGGG TOCT T AGCAGTGATGA GGGGGGGG ACCAAGTGA GGGACTGGGG TATOTA AGCAGTGA GGACTGGGG ACCACGTC 2.2 AGCAGTGA A GGACTGGGGG TATGAA 2.6 AGCAGTGA GGGACTGGGG TATOTA AGCAGTGA A GGACTGGGGGGC TATOTA AGCAGTGA GGGGGGG GTCCGGAA AGCAGTGAT C GGGGGGGG TATOTA AGCAGTGA GGGGGGG GTCCGGAA AGCAGTGAT C C CAGTGAGGA C C CAGTGAA AGCAGTGAT C CAGGGGGG <t< td=""><td>AGCAGTGATG</td><td>CA</td><td>G</td><td></td><td>GGGACAGGG</td><td></td><td>т</td><td></td><td>TGCT</td><td>2.1</td><td>AGCGGTGATG</td><td></td><td></td><td></td><td>G (1)</td><td></td><td></td><td></td><td>CARCEAR</td><td></td></t<>	AGCAGTGATG	CA	G		GGGACAGGG		т		TGCT	2.1	AGCGGTGATG				G (1)				CARCEAR	
COCKEGNAG CCCC CGGGGGG A TANCTATOCT ACCATGA CGGGGGGGC TANCTATOCT ACCATGA CGGGGGGG C CANAC ACCATGA CGGACAGGGGC TOCT 1.6 ACCATGA CGGGGGGGC GC CANAC ACCATGA CGGACAGGGGC TOCT 1.6 ACCATGA CGGGGGGGC GC CANC ACCATGA CGGACAGGGGC TOCT T ACCATGATG CGGGGGGGC GGGGGGGG ACCATGATG CGACATGGGGG T T ACCATGATG CGGGGGGGC ACCATGATG CACATGATGG ACCATGATG CACATGAGGG T ACCATGA ACCATGATG CACATGAGGG ACCATGATG CACATGAGGG TATGCA ACCATGA CACATGAGGG CACATGAGGG CACATGA CACATGA CACATGAGGG CACATGAGGG CACATGA CACATGA CACATGAGGG CACATGAGGG CACATGAGGG CACATGAG ACCATGA CACATGAGGG CACATGAGGG CACATGAGGG CACATGAGGG CACATGAGGG CACATGAGGG CACATGAGGGG CACATGAGGGG CACATGAGGGG CACATGAGGGG CACATGAGGG CACATGAGGGG CACATGAGGGG CACATGAGGGG CACATGAGGGG CACATGAGGGG CACATGAGGGG CACATGAGGGG CACATGAGGGG CACATGAGGGGG CACATGAGGGG CACATGAGGGGG <t< td=""><td>ACCAGTGA</td><td></td><td>-</td><td></td><td>GGACTGGGGG</td><td></td><td>-</td><td></td><td>AACTATGCT</td><td></td><td>AGCAGIG</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td><td>NACONCOLUM</td><td></td></t<>	ACCAGTGA		-		GGACTGGGGG		-		AACTATGCT		AGCAGIG								NACONCOLUM	
MCMGT A GGGGGG TC CANANC 2.4 ACCASTA GCASTA GGGGGGC T T ACCASTATAS C GGGGGGG G AAC GCASTA GGACTGGGGGGCC T ACCASTATAG GGGGGGG GCASTA ACCASTA GGACTGGGGGCC T AGCASTATAG GGGGGGG AACASTA ACCASTA GGACTGGGGG TATASA AGCASTATAG GGGACTGGGGG ACCASTA GGACTGGGGG ACCASTA AGCASTAGA ACCASTA GGACTGGGGG ACCASTA AGCASTAGA A GCASTGGGGGC TATASA ACCASTA GGACTGGGG GACASGGG AGCASTA A GCASTGGGGC TATASA ACCASTGA GG GTUCAGAA AGCASTA C C GGGACTGGGGC TATASA ACCASTGA GGGGGGG CAAAAC AGCASTAT C AGGGGGG TCT CTCTTATAA ACCASTGA C AGCASTGA AGCASTAT C CAGGGG TCT CTCTTATAA ACCASTGA C AGCASTGA AGCASTATA C C CAGGGG TCT CTCTTATAA ACCASTGA C AGCASTGA AGCASTATA C C CAGGGG TCT CTCTTATAA ACCASTGA C AGCASTGA AGCASTATAT C C CAGGGGG A	GCCAGCAG		9000		GGGGGG			A	TAACTATGCT		AGC				GGGACAGGGGGG				AACCAGGCI	1.4
AGCAGTIANTG C COGGGGGG G CAC ACCAGTAGATG C COUNT OF	AGCAGT		A		GGGGGG		TC		CAAAAC	2.4	AGCAGIGA				CCACECCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCCC					2 1
NGCNGTGATG GGGGGG NAGIC 2.5 NGCA NGCNGTGATG GACTGGGG TATGCT NGCA TCAN GACTGGGG ACCANGAC AGCAGTG GGCACTGGGG TATGCT NGCAG CCCC CAGGGG TATGAA 2.6 AGCAGTG GGCACTGGGG ACCAGTG GGCACTGGG ACCAGTG GGCACTGGGG CCCC CAGGGG CCCCC CAGGGGC TATGAA AGCAGTG GGCAGTGGA GG GTCCAGAA 2.3 AGCAGTG GC C GGGACTGGGGGG TATGAA AGCAGTGA GGGGGGG GGGGGGG CCCC GGGGGGG CCAAGAC AGCAGTGA C AGCAGTGA AGCAGTGA AGCAGTGA C AGCAGTGA AGCAGTGA AGCAGTGA C AGCAGTGA AGCAGTGA AGCAGTGA AGCAGTGA AGCAGTGA AGCAGTGA AGCAGTGA C AAGCAG AGCAGTGA AG	AGCAGTGATG	с			GGGGGGGC	G	G		AAC		AGCAGIG	c			GGAC10000000				1001 T	
AGC TCAA GACTOGOGG ACCAGG ACCAGG GACLAGU GACLAGU GACLAGU ACCAGG ACCAGG ACCAGG ACCAGG ACCAGG CACUAGU GGACTOGOG ACCAGGG CACUAGU GGACTOGOG CACUAGU GGACTOGOG GGACTOGOGOGO CACUAGU GGACTOGOGOGO GGACTOGOGOGO CACUAGU GGACTOGOGO CACUAGU CAU CAU </td <td>AGCAGTGATG</td> <td></td> <td></td> <td></td> <td>GGGGGG</td> <td></td> <td></td> <td></td> <td>AAGAC</td> <td>2.5</td> <td>AGCAGIG</td> <td>74</td> <td></td> <td></td> <td>ChOTCOCC</td> <td></td> <td></td> <td></td> <td>TATCOT</td> <td></td>	AGCAGTGATG				GGGGGG				AAGAC	2.5	AGCAGIG	74			ChOTCOCC				TATCOT	
AGCAG CCCC CAGGGG TATGAA 2.6 AGCAGTG GACAGGG CACAGGG GACAGGG CACAGGG CACAGGG CACAGGGG CACAGGG CACAGGGG CACAGGG CACAGGG CACAGGG CACAGG CACAGGG CACAGG CACAGGC CACAGGC CACAGGC CACAGGC CACAGGC CACAGGG CACAGGC CACAGGC </td <td>AGC</td> <td></td> <td>TCAA</td> <td></td> <td>GACTGGGGG</td> <td></td> <td></td> <td></td> <td>ACCAAGAC</td> <td></td> <td>AGCAGIG</td> <td></td> <td></td> <td></td> <td>CCCACTCCC</td> <td></td> <td></td> <td></td> <td>ACACCEGT</td> <td>2.2</td>	AGC		TCAA		GACTGGGGG				ACCAAGAC		AGCAGIG				CCCACTCCC				ACACCEGT	2.2
AGCAGTGA A GGACTGGGGGGC TATGAA AGCAGTGA GGACTGGGGGG GGCGGGG CAAGAC 2.5 AGCAGTGA GGGGGGGG TCT CTCTGTATGAA AGCAGTGA GGGGGGG CAAGAC 2.5 AGCAGTGAT C C CAGGGGG TCT CTCTGTATGAA AGCAGTGA GGGGGGGG CAGAC 2.5	AGCAG		cccc		CAGGGG				TATGAA	2.6	AGCAGIG				CACAGGGG				CCAGAA	2 7
AGCAG GC CC GGGACTGGGGGC TATGAA ACC TGGGGG CAAAAA 2.4 AGCAGTGAT CAGGGGG TCT CTCTATGAA AGC TGGGGG CAAAAA 2.4 AGCAGTGAT CAGGGGG ACTATTATAA AGCAGTGAA C AGCAGTGAC 2.4 AGCAGTGAT CAGGGGG A CTATTGAA AGCAGTGAA C AGCAGTGAC AGCAGTGAA C AGCAGTGAA AGCAGTGAA GGGGGGGG C AAGCAGTGAA C AGCAGTGAA AGCAGTGAA GGGGGGGG C AAGCAGTGAA AGCAGTGAA	AGCAGTGA		λ		GGACTGGGGGGC				TATGAA		AGCAGIG				GACA0000				CTCCACAA	
AGCAGTGAT CAGGGG TCT CTCCTATGAA AGCAGTGA CGGGGGG CAAGAC 2. AGCAGTGA C CAGGGG A CTATGCA AGCAGTGA GGGGGGG CAAGAC 2. AGCAGTGATG CA GGGGG TCT CTCCTATGAA AGCAGTGA GGGGGGG CAAGAC 2.5	AGCAG		GC	cc	GGGACTGGGGGC				TATGAA		AGCAGIGA				TOGOOG				CRABAC	2.4
AGCAGT C CAGGGG A CTATGAA ACCAGAGA GGGGGGG CAAGAC 2.5 AGCAGTGATG CA GGGGG TCT CTCCTATGAA ACCAGTGA GGGGGGG CAAGAC 2.5	AGCAGTGAT				CAGGGGG		TCT		CTCCTATGAA		ACCACTICA				100000				AGTCAAAAC	
AGCAGTGATG CA GGGGG TCT CTCCTATGAA AGCAGTGATG CG AG CTCCTATGAA A	AGCAGT		с		CAGGGG		A		CTATGAA		ACCACITCA				0000000				CAACAC	2.5
	AGCAGTGATG		CA		GGGGG		TCT		CTCCTATGAA		AGCAGIGA				0000000			NG.	CTCCTATCAA	2.6

SCIENCE • VOL. 261 • 27 AUGUST 1993

Table 1. Use of homology in TdT-/- and TdT+ V(D)J junctions. Most data sets were compiled from independent amplifications from at least two mice. Only homologies of two nucleotides or more were counted, and as in the figures we considered the possibility of P inserts of up to five nucleotides. The V₃3 canonical junctions are numbered as in (15). DP, CD4+CD8+; DN, CD4-CD8- thymocytes.

		V _y 3	_	V ₈ 8 DP	V _H 783/J558			
	Total	DP	DN	CD3 lo	(D-J only)			
		TdT-						
No. of junctions	64	54	46	122	39			
No. of junctions with N	3	0	3	3	2			
No. of homologies > 1 bp (% in N ⁻ junctions)	40 (66)	35 (65)	29 (67)	22 (18)	20 (54)			
No. of canonical junctions								
(% IOIal junctions)	17 (97)	16 (20)	11 (24)					
Junction 2	17 (27)	11 (20)	9 (20)					
Junction 3	7 (11)	7 (13)	5 (20)					
Total	39 (61)	34 (63)	25 (54)					
	00 (01)	TdT+	20 (0 !)					
No. of junctions	55	54	30	82	23			
No. of junctions with N	42	38	12	50	17			
No. of homologies > 1 bp (% in N ⁻ junctions)	6 (46)	7 (44)	8 (44)	2 (6)	2 (33)			
No. of canonical junctions (% total junctions)								
Junction 1	2 (4)	1 (2)	0					
Junction 2	2 (4)	1 (2)	5 (17)					
Junction 3	2 (4)	5 (9)	3 (10)					
Total	6 (11)	7 (13)	8 (27)					

otides, are shared between the two: the D segments are essentially G nucleotide stretches, whereas few G nucleotides are found in the J segments. Consequently, TCR β genes will be very diverse even in perinatal animals.

Thus, the enzyme TdT catalyzes the bulk of N nucleotide addition to V. D. and J segment ends and blocks homologydirected recombination. These two processes greatly magnify the number of antigen receptor specificities expressed by adults as compared with perinates. It remains to be determined which is more critical for the development and wellbeing of the animal, the relatively restricted perinatal or the highly diverse adult repertoire.

The accompanying paper by Komori et al. arrived at similar conclusions about mice carrying a different TdT mutation (26).

REFERENCES AND NOTES

- 1. C. Benoist and D. Mathis, Curr. Opin. Immunol. 4, 2 (1992).
- L. Carlsson and D. Holmberg, Int. Immunol. 2, 2. 639 (1990); J. F. Elliot, E. P. Rock, P. A. Patten, M. M. Davis, Y.-H. Chien, *Nature* **331**, 627 (1988); W. T. McCormack, L. W. Tjoelker, G. Stella, C. E. Postema, C. B. Thompson, *Proc. Natl. Acad. Sci. U.S.A.* 88, 7699 (1991); J. F. George, Jr., and H. W. Schroeder, Jr., J. Immunol. 148, 1230 (1992).
- 3. H. Gu, I. Förster, K. Rajewsky, EMBO J. 9, 2133 (1990).
- A. J. Feeney, J. Exp. Med. 172, 1377 (1990).
- 5. J. J. Lafaille, A. DeCloux, M. Bonneville, Y. Taka-

gaki, S. Tongegawa, *Cell* **59**, 859 (1989).

- 6. M. Bogue, S. Candéias, C. Benoist, D. Mathis, EMBO J. 10, 3647 (1991).
- A. J. Feeney, J. Exp. Med. 174, 115 (1991).
- F. J. Bollum, in The Enzymes, P. D. Boyer, Ed. (Academic Press, New York, 1974), pp. 145-171.
- 9 S. V. Desiderio et al., Nature 311, 752 (1984); G. D. Yancopoulos, T. K. Blackwell, H. Suh, L. Hood, F. W. Alt, *Cell* 44, 251 (1986); N. R. Landau, D. G. Schatz, M. Rosa, D. Baltimore, Mol. Cell. Biol. 7, 3237 (1987); S. Kallenbach, N. Doyen, M. Fanton d'Andon, F. Rougeon, Proc. Natl. Acad. Sci.
- *U.S.A.* **89**, 2799 (1992). 10. M. Bogue, S. Gilfillan, C. Benoist, D. Mathis, *Proc.* Natl. Ācad. Sci. U.S.A. 89, 11011 (1992).
- 11. Y. Ichihara, H. Hayashida, S. Miyazawa, Y. Kuro-
- sawa, Eur. J. Immunol. 19, 1849 (1989).
 12. A. J. Feeney, J. Immunol. 147, 4343 (1991); ibid. 149, 222 (1992).
- 13. D. M. Asarnow, T. Goodman, L. LeFrancois, J. P. Allison, Nature 341, 60 (1989); L. K. Aguilar and J.
- W. Belmont, J. Immunol. 146, 1348 (1991).
 J. J. Lafaille, W. Haas, A. Coutinho, S. Tonegawa, Immunol. Today 11, 75 (1990); S. Itohara and S. Tonegawa, Proc. Natl. Acad. Sci. U.S.A. 87, 7935 (1990)
- 15. S. Itohara et al., Cell 72, 337 (1993).
- D. M. Asarnow, D. Cado, D. H. Raulet, Nature 362, 16. 158 (1993).
- 17. To produce $TdT^{-/-}$ mice, we constructed a targeting vector as illustrated in Fig. 1. It was linearized and electroporated into D3 embryonic stem cells as described (27). Of 151 Gancyclovir/G418-resistant D3 clones analyzed, one carrying the predicted integration was identified, and this was confirmed by extensive Southern blot analysis. It was expanded and injected into blastocyts, which were reimplanted into pseudopregnant females. Chimeras were crossed to C57BI/6 mice, and offspring carrying the mutation were intercrossed to produce homozygous $TdT^{-/-}$ mice. The presence of the mutation was confirmed by Southern blot analysis of Eco RI–, Eco RV–, or Bgl II–digested tail DNA with the use of a PCR-generated cDNA probe spanning exons 4 to 6. Mice were housed in a conventional animal facility and handled

according to EEC guidelines. 18. J. K. Farrar, R. K. Evans, C. M. Beach, S. Cole-

- man, Biochemistry 30, 3075 (1991). S. Gilfillan, unpublished data.
- 20. The immune system of $TdT^{-/-}$ mice was char-
- acterized (S. Gilfillan et al., in preparation).
- 21 DNA and RNA were isolated from sorted cells and tissues with the use of standard techniques. Briefly, thymocyte suspensions from individual mice were stained with appropriate combinations of monoclonal antibodies to CD4, CD8 (Caltag), and CD3 (KT3) (28) and electronically sorted (29). RNA was prepared by NP-40 lysis from 1×10^5 to 2×10^5 sorted cells to which 1×10^6 HeLa cells were added as carrier; DNA was isolated directly from 1 \times 10⁵ to 5 \times 10⁶ sorted cells. RNA was isolated from spleen tissue and cDNA synthesized as described (29). DNA was prepared by proteinase K digestion followed by phenol-chloroform extraction. PCR was carried out in a DNA Thermal Cyler (Perkin-Elmer/Cetus); products were digested with appropriate restriction enzymes, cloned into M13mp19, and sequenced by either the dideoxy method or by automated se-quencing (Applied Biosystems). The conditions and primers (restriction sites are underlined) used and phillers (restriction sites are in definited) used were as follows: V_{g8} , QN199 (GAGGAAAGGTGA-CATTGAGC) (7) and QN200 (CGAC<u>GCATGC-</u> TGGTATCGGCAGGAC) (7); C_g, MQ284 (AGCA-CACGAGGGTAGCCTT) (29) and MS175 (GACA-GAACTTT<u>GAATTCCTCTGCTTTTGATGG</u>) (29); V.3, QR239 (CACCATTCTGCAGTGGTACCAAC TG) (5) and OU47 (GTCCTGACCTCCTGCAGC CGCTTGGAAATT) (30); Jy1, QR238 (CAGAGG GAATTCCTATGAGCTTAGT) (5); V_H7183, QR194 (CGCGCTGCAGCGTGGAGTCTGGGGGAGGCT-TA) (31) and QU234 (GTGCAGC<u>CTGCAG</u>AGTC-CCTGAAACTCTCC); J_H^4 , QR240 (CCTGGAGA-GGGAATTCTTACCTGAGGGAG) and QQ4 (GCGC-TCGAGAATTCGGTGACTGAGGTT) (31); V_HJ558 QQ3 (GTCCAACTGCAGCAGCCTGGGGGCTG-AG); and C_µ, QQ2 (GGGGAATTCATTTGGGAA-GGACTGACT) (4). For V_p8, cDNA was amplified as described (7). We amplified V_y3 and V_H7183 junctions from 1 to 2 μ g of Eco RI–digested DNA using two rounds of PCR. For the primary amplification, samples were denatured for 6 min at 94°C and then subjected to 25 amplification cycles-each consisting of 30 s at 94°C, 1 min at 55°C, and 1 min at 72°C-followed by a 10-min elongation at 72°C. One microliter of the primary reaction was used for the secondary PCR in which the annealing step was changed to 30 s at 60°C. We amplified V_HJ558 junctions from cDNA using a single set of $V_{\rm H}$ and C_{μ} primers as follows: 6 min denaturation at 94°C; 25 cycles of 45 s at 94°C, 1 min at 60°C, and 2 min at 72°C; and a final elongation of 10 min at 72°C
- 22 D. B. Roth, X.-B. Chang, J. H. Wilson, Mol. Cell. Biol. 9, 3049 (1989).
- 23 J. M. Clark, Nucleic Acids Res. 16, 9677 (1988).
- M. R. Lieber, Cell 70, 873 (1992) 24.
- 25. D. B. Roth and J. H. Wilson, Mol. Cell. Biol. 6, 4295
- (1986).
 - 26. T. Komori, A. Okada, V. Stewart, F. Alt, Science 261, 1171 (1993).
- D. Cosgrove et al., Cell 66, 1051 (1991).
- 28. K. Tomonari, Immunogenetics 28, 455 (1988)
- 29. S. Candéias, C. Waltzinger, C. Benoist, D. Mathis, J. Exp. Med. 174, 989 (1991).
- C. Schleussner et al., Thymus 20, 195 (1992). L. Carlsson, C. Overmo, D. Holmberg, Int. Immunol. 4, 549 (1991).
- We thank F. Alt for discussing unpublished data, P. 32 Marchal-Bohn and B. Laborde for help with the sequencing, S. Vicaire for automated sequencing, C. Waltzinger and C. Ebel for cell sorting, C. Schleussner for oligonucleotides and for advice on V.3 PCR, and P. Charles, C. Repis, P. Michel, N. Zinck, and S. Metz for maintaining the mouse colony. Supported by the Institut National de la Santé et de la Recherche Médicale and the Centre National de la Recherche Scientifique. S.G. received a fellowship from the American Cancer Society.

25 May 1993; accepted 12 July 1993

1178

SCIENCE • VOL. 261 • 27 AUGUST 1993