## Structure and Diversity of the Human T-Cell Receptor β-Chain Variable Region Genes

Jeffrey P. Tillinghast, Mark A. Behlke, Dennis Y. Loh

In order to characterize the variability of the expressed human T-cell receptor (TCR) β-chain repertoire and contrast this variability to the known murine β-chain repertoire, 15 independent complementary DNA (cDNA) clones containing TCR β-chain variable region  $(V_{\beta})$  genes were isolated from a human tonsil cDNA library. The nucleotide and derived amino acid sequences of these 15  $V_{\beta}$  genes were analyzed together with 7 previously defined sequences. Fifteen different human  $V_{\beta}$  genes could be identified from 22 independent sequences. By means of DNA hybridization and sequence homology comparisons, it was possible to group these 15 genes into ten distinct  $V_{\beta}$  subfamilies, each containing from one to seven members. Minimal polymorphism was noted between individuals, except in multimember subfamilies. The amino acid sequences of these genes contain conserved amino acids that are also shared by murine TCR  $V_{\beta}$  genes and immunoglobulins; no features were found that distinguish human  $V_{\beta}$  genes from their murine counterparts. Evaluation of secondary structure showed that maximum variability coincides with generally hydrophilic portions of the amino acid sequence, while specific hydrophobic regions were conserved in all  $V_{\beta}$  genes examined.

The T-cell receptor (TCR) is a heterodimer of disulfide-linked  $\alpha$ and  $\beta$  chains, each composed of an amino-terminal variable (V) region and a carboxyl-terminal constant (C) region (1-3). Antigen recognition by the TCR is presumed to occur in the V region (4). Unlike immunoglobulins, however, the TCR recognizes antigen only in the context of molecules encoded by the major histocompatibility complex (MHC), a phenomenon known as MHC restriction (5, 6). At present, the mechanisms responsible for MHC restriction are unknown. Murine TCR  $V_{\beta}$  genes are known to consist of a relatively small number of one- to threemember subfamilies, which show very little polymorphism among various strains of inbred mice (7-10). These V<sub> $\beta$ </sub> gene segments somatically rearrange to diversity (D) and joining (J) gene segments to create a functional variable region, a process that magnifies the germline-encoded variability.

The structure and diversity of murine TCR  $V_{\beta}$  genes have been well defined (7–10). To similarly characterize the structure and diversity of human TCR  $V_{\beta}$  genes, we examined 15  $V_{\beta}$ -containing complementary DNA's (cDNA's) and compared them with the murine (7–10) and human (11–15)  $V_{\beta}$  genes published to date. We find that human  $V_{\beta}$  genes show many similarities to their murine homologs in both amino acid sequence and predicted secondary structure. While the size and organization of the germline repertoire of human  $V_{\beta}$  genes is unknown, we estimate the expressed repertoire to be about 50 genes.

A cDNA library was constructed from a human tonsil, and a portion of the primary,

unamplified library was screened with a murine TCR  $C_{\beta}$  probe;  $C_{\beta}$ -positive cDNA clones were isolated, and their nucleotide sequences were determined. A total of 15 independent  $V_{\beta}$ -containing cDNA's were obtained; their nucleotide sequences are presented in Fig. 1 along with those of seven  $V_{\beta}$  genes previously defined by others. The sequences are grouped by homology so that duplicate isolates of a single  $V_{\beta}$  gene and related members of different subfamilies can be more easily recognized. It is not clear whether highly homologous isolates represent distinct members of the same subfamily or are actually allelic variants of the same gene isolated repeatedly from a polymorphic  $V_{\beta}$  gene pool. In this report, genes that are not identical will be presented as distinct  $V_{\beta}$ genes.

Seven of the 22 nucleotide sequences in Fig. 1 are identical in their  $V_{\beta}$  coding region to previous  $V_{\beta}$  isolates (the sequences of these isolates are represented by dashes in Fig. 1). All duplicate isolates possess distinct junctional sequences and therefore represent independent, repeated usage of a single  $V_{\beta}$ gene. Interestingly, duplicate pairs of identical  $V_{\beta}$  genes were isolated both from the same source (ph 16 and 79, ph 7 and 34, and ph 26 and 29, from our tonsil cDNA library) and from different individuals (ATL 12-2 and ph 5, MT1-1 and Molt 4, YT35 and ph 11, and ATL 2-1 and ph 32). Our ph 7 and ph 34 genes are 99% homologous to the published Molt 4 and MT1-1 genes (Fig. 2) and may represent either allelic polymorphism or recently duplicated, yet distinct, genes; genomic Southern blot analysis indicates that ph 7 is part of a twomember  $V_{\beta}$  subfamily (Fig. 3). Similarly,

nine nucleotide differences were found between a germline human  $V_{\beta}$  gene and the YT35 cDNA (14). Our ph 8 and ph 11 cDNA's (Fig. 1) differ at the same nine nucleotides and correspond exactly to both of these published genes; genomic Southern blot analysis indicates that YT35 is part of a five-member  $V_{\beta}$  subfamily (14).

In summary, the 22 sequences presented in Fig. 1 include seven genes isolated in duplicate and eight genes isolated once, so that 15 distinct  $V_{\beta}$  genes have been defined. Thus, 32% (7/22) of  $V_{\beta}$  isolates to date involve repeated usage of previously defined  $V_{\beta}$  genes. If one assumes that all  $V_{\beta}$  genes are utilized equally, a statistical estimate can be made of the likely number of  $V_{\beta}$  genes that may be present in the human genome (9, 16, 17). The maximum likelihood estimate of the number of distinct  $V_{\beta}$  gene segments (n) is 25, which was obtained by determining the value of n that maximized the probability for 15 distinct gene segments and a total number of 22. A 95% onesided confidence bound of n = 52 was obtained by determining the smallest value of n for which the probability of observing at most 15 different segments among 22 segments examined was less than 0.05. If ph 34 and Molt 4 are allelic variants and not distinct genes, as discussed above, then the number of distinct gene segments is 14 and estimates of n will be slightly lower (18).

It is possible that our cDNA library is not representative of human  $V_{\beta}$  usage in general since it was derived from a single human tonsil, in which antigenic stimulation and clonal proliferation may have resulted in the relative overexpression of particular  $V_{\beta}$ genes. However, when our sequences are pooled with those previously identified, eight independent sources of  $V_{\beta}$  sequences are represented. Within the limitations of our sampling, we conclude that the expressed human  $V_{\beta}$  repertoire is less than 50 to 60 genes. This calculation will underestimate the actual number of human  $V_{\beta}$  genes if  $V_{\beta}$  usage is nonrandom in human peripheral lymphoid tissue; nonrandom  $V_{\beta}$  usage has already been demonstrated in the murine splenic T-cell population (10). Nevertheless, these data indicate that the human  $V_{\beta}$  repertoire is likely to be larger than the murine  $V_{\beta}$  repertoire of < 30 genes (7–10) but smaller than the immunoglobulin heavychain variable region  $(V_H)$  repertoire of 250 to 300 genes (19).

Clonal diversity in human  $\beta$ -chain variable regions is generated in a fashion identical to the murine V<sub> $\beta$ </sub> region, in which the

Howard Hughes Medical Institute, Departments of Medicine, Microbiology, and Immunology, Washington University School of Medicine, St. Louis, MO 63110.

diversity of a limited number of germline gene elements is enhanced by a flexible somatic recombination scheme. Toyonaga etal. (20) have shown that there are six J- region gene segments associated with  $C_{\beta}l$ and seven J regions associated with  $C_{\beta}2$ , with at least one diversity gene segment located 5' of each cluster of  $J_{\beta}$  elements (20). Analysis of the junctional regions of the cDNA sequences in Fig. 1 shows that diversity in the recombination event is apparently greater than would be expected



Fig. 1. Nucleotide sequences of 22 human  $V_{\beta}$  genes. A cDNA library in the  $\lambda$ gt10 cloning vector was constructed with polyadenylated RNA from a human tonsil (32, 33). The library was screened with a murine  $C_{\beta}$  probe; DNA from positive clones was subcloned into pUC 19 or M13mp 18. Nucleotide sequences were determined by either the method of Maxam and Gilbert (34) or the method of Sanger *et al.* (35), as modified by Biggin *et al.* (36). A synthetic primer complementary to  $C_{\beta}$  was used to sequence from  $C_{\beta}$  into the V region. Sequences begin within the leader; the J<sub>β</sub> elements are

aligned separately and are identified according to Toyonaga *et al.* (20). Gaps have been introduced to maximize homology. Genes that are identical to a previously defined sequence have their nucleotide sequences represented by dashes, indicating base identity with the nucleotide sequence immediately above the dash. Sources of  $V_{\beta}$  sequences are as follows: ph 5 to ph 79, this work; MT1-1, ATL 2-1, ATL 12-1, ATL 12-2 (11); YT35 (15); Molt 4 (13); and HPB-MLT (12).

	ATL 12-2	ph 22	ph 16	HPB- MLT	ph 34	Molt 4	ph 8	YT 35	ph 15	ph 21	ph 27	ATL 2-1	ph 29	ATL 12-1	ph 24
ATL 12-2		71	76	73	32	32	56	58	27	28	36	29	48	38	48
ph 22	82		65	66	31	31	53	54	30	33	40	36	45	42	45
ph 16	81	79		78	36	36	59	59	38	30	45	34	46	38	44
HPB-MLT	78	76	86		30	30	58	60	30	26	34	30	48	40	42
ph 34	49	51	49	52		98	31	32	31	27	32	35	29	28	28
Molt 4	49	51	49	52	99		31	31	31	27	32	35	29	28	28
ph 8	67	68	70	70	48	48		97	32	38	44	39	51	40	44
YT35	67	68	70	70	48	48	98		32	38	44	39	50	38	45
ph 15	52	51	50	45	46	46	54	53		53	57	5 <b>4</b>	33	27	29
ph 21	49	54	52	51	49	49	55	54	70	<u> </u>	55	57	32	32	34
ph 27	55	57	57	51	52	52	59	58	66	69	_	60	32	29	39
ATL 2-1	52	55	54	47	52	52	56	56	69	70	70		32	35	39
ph 29	56	59	57	56	50	50	60	59	52	52	56	50		47	40
ATL 12-1	55	62	56	59	49	49	57	57	44	52	53	50	62		41
ph 24	55	58	56	55	45	45	53	54	51	<b>4</b> 5	53	53	57	56	

Fig. 2. Homology matrix of 15 human  $V_{\beta}$  gene segments. Numbers above the diagonal indicate the percentage of homology between the sequences indicated on the x- and y-axes when compared at the amino acid level; numbers below the diagonal show percentage homology at the nucleotide level.  $V_{\beta}$  coding sequences, excluding leader segments, are compared.

solely from a flexible joining process. Putative D regions, defined as sequence not encoded by genomic J or identifiable V elements, vary significantly in length (Fig. 1). These junctional regions differ widely in the number of nucleotides originating from the two known germline  $D_{\beta}$  segment genes. The sequence of clone ph 16 has 15 putative "D-region" nucleotides, 13 of which appear to be encoded by  $D_{\beta}2.1$ , whereas ph 34 and YT35 have only 3 of 12 nucleotides and 3 of 15 nucleotides, respectively, of recognizable germline origin. We conclude that either more pronounced N-region diversification occurs during the human  $V_{\beta}$ - $D_{\beta}$ - $J_{\beta}$  recombination event than has generally been seen in murine  $\beta$ -chain cDNA's (7-9) or additional unidentified human germline  $D_{\beta}$  elements exist.

Comparison of germline and rearranged human or murine  $V_{\beta}$  genes suggests that somatic hypermutation is not significantly involved in the generation of TCR  $\beta$ -chain diversity (7–9, 11). T cells utilizing the same germline genes generally express identical  $V_{\beta}$  gene segments in sequenced cDNA's. Duplicate isolates of the same  $V_{\beta}$  genes are identical in seven cases examined here, supporting the conclusion that somatic hypermutation is a rare event.

Comparisons of the 15 different human  $V_{\beta}$  gene sequences to each other show eight genes that are greater than 75% homologous (Fig. 2). These genes should cross-hybridize on genomic Southern blot analysis and therefore define related  $V_{\beta}$  subfamilies.  $V_{\beta}$  genes ATL 12-2 (ph 5), ph 22, ph 16 (ph 79), and HPB MLT;  $V_{\beta}$  genes YT35 (ph 11) and ph 8; and  $V_{\beta}$  genes MT1-1 (Molt 4) and ph 34 (ph 7) define three distinct multimember subfamilies. The remaining genes are all <70% homologous to other isolates and define seven additional subfamilies.

lies. These seven subfamilies may also prove to have multiple members. In summary, 15 different  $V_{\beta}$  gene sequences examined define ten  $V_{\beta}$  subfamilies.

The number of bands visualized on a Southern blot of genomic DNA hybridized with a  $V_{\beta}$  probe corresponds approximately to the number of related genes within a subfamily. When representative  $V_{\beta}$  probes from each subfamily are used in Southern blot analysis and the number of bands is counted, a more realistic estimate of the



Fig. 3. Genomic Southern blot of peripheralblood-leukocyte DNA from three individuals. DNA (10  $\mu$ g) (37) was digested with Eco RI, Pvu II, and Sac I as indicated, separated on 0.8% agarose gels, and transferred to nitrocellulose (38). DNA on the filters was hybridized to nicktranslated V<sub>g</sub>-specific probes (specific activity: 1 × 10<sup>8</sup> to 3 × 10<sup>8</sup> cpm/ $\mu$ g) at 68°C for >12 hours, under standard conditions (39). Filters were washed in 0.30M NaCl/0.03M sodium citrate at 68°C and autoradiographed. Positions of marker fragments from Hind III–digested  $\lambda$ DNA are indicated on the left.

human  $V_{\beta}$  repertoire can be made. At least seven bands are observed when a  $V_{\beta}$ ph 5 (ATL 12-2) probe is used (Fig. 3). Similarly, YT35 (our ph 11) shows five bands (14), while  $V_{\beta}ph$  7,  $V_{\beta}ph$  15, and  $V_{\beta}ph$  32 each show two bands (Fig. 3 and not shown).  $V_{\beta}$ ph 21,  $V_{\beta}$ ph 27, and  $V_{\beta}$ ph 29 each show a single band (not shown). Thus, representative  $V_{\beta}$  probes from eight separate subfamilies identify 21 bands on genomic Southern blots. Interestingly, 13 of the 21 putative  $V_{\beta}$  genes from these eight subfamilies have already been cloned and their sequences determined (Fig. 1). The human  $V_{\beta}$ gene family therefore appears to be composed of at least ten subfamilies, each containing one to seven members. When the composition of the human  $V_{\beta}$  gene family is compared to the murine  $V_{\beta}$  gene family (7-10), a similar small number of subfamilies is found. However, human  $V_{\beta}$  gene subfamilies are often composed of multiple members while murine  $V_{\beta}$  genes mostly belong to single member subfamilies. Therefore, while both the human and mouse species possess a similar number of  $V_{\beta}$  subfamilies, humans seem to have more than twice the number of actual  $V_{\beta}$  genes. It is not clear why human  $V_{\beta}$  subfamilies often expand to contain multiple members while their murine counterparts do not.

Little polymorphism has been seen in  $V_{\beta}$ gene subfamilies among various strains of inbred mice (7-10). It is not known to what extent polymorphism exists in the overall  $V_B$ gene family in the outbred human population. However, in a limited survey, we note the existence of both seemingly nonpolymorphic and highly polymorphic  $V_{\beta}$  gene subfamilies among the genes presented here. A  $V_{\beta}$  probe (Fig. 3) shows no restricfragment length polymorphism tion (RFLP) in a genomic Southern blot analysis of three individuals and thus represents a class of relatively nonpolymorphic  $V_{\beta}$  genes, similar to the pattern observed for most  $V_{\beta}$ genes among the inbred mouse population. In contrast, RFLP's are clearly seen among individuals in the  $V_{\beta}$ ph 5 (ATL 12-2) multimember subfamily (Fig. 3). Increased polymorphism in multimember subfamilies may result from gene exchanges between highly homologous but independent genes, as has been suggested to occur for immunoglobulins (21). Whether the RFLP's correlate with actual  $V_{\beta}$  sequence polymorphism is not known.

Homology between members of different subfamilies ranges from 26 to 60% (Fig. 2). Comparison of the amino acid sequences of human  $V_{\beta}$  genes (Fig. 4) with the amino acid sequences of available murine  $V_{\beta}$  genes shows no difference in conserved residues; most of the residues conserved in  $V_{\beta}$  genes

					······································	/				
		10	20	30	40	50	<b>6</b> 0	70	80	90
ATL12-2 ph 22 ph 16 HPB-MLT ph 34 Moit 4 ph 8 YT35 ph 15 ph 21	MGTRLLCWVVLGFLGTDHT MGTSLLCWMALCLLGADQE MGTRLLFWVAFCLLGADHT MLLLLLLGPAGSGL MGSWTLCCVSLCILVAKHT MDSWTFCCVSLCILVAKHT MAQFGLLFSGAGLM MGPDLLGYVVLCILGAGPI	GAG VSOSPRYK VAKRO ISGVSONPRHK ITKRO GAG VSOSPSNK VTEKO GAVVSOHPSRVICKSO GAVVSOHPSWVICKSO DAGVIOSPRHE VTEMO EADIYOTPRYL VIGTO EADIYOTPRYL VIGTO	20 50DVALRCDPISG 50NVTFRCDPISG 5KDVELRCDPISG 5KDVELRCDPISG 5TSVKIECRSLDF 5TSVKIECRSLDF 50EVTLRCKPISG	30 HVSLFWY HNRLYWY HTALYWY HTALYWY QATTMFWY OATTMFWY HDYLFWY HNSLFWY HDKMYWY HFYMSWY	4U QQALGQGPEFI RQNPGQGPEFI RQSLGQGLEFI RQSLGQGPEFI RQFPKKSLMLN RQFPKKSLMLN RQFPKQSLMLN RQTMMRGLELL QQDPGMELH R QQDPGMELH R	50 : TY FONEAQL IY FOGNSAQ IY FOGNSAQ IY FOGNSAQ ATSNEGSKAT ATSNEGSKAT IY FNNNVPI IY FNNNVPI IY YSYGNVFY	60 : SGL PSDRF EKSGL LSDRT DKSGL PSDRF DDSGL PNDRF YEQG VEKDKF YEQG VEKDKF DDSGMPEDRF DDSGMPEDRF TEKGDLSSES TDKGDVPEGY	70 : FAQRPEGSVS SAERPKGSFS SAERTGGSVS FAVRPEGSVS LINHASLTLS LINHASLTLS SAKMPNASFS SAKMPNASFS SAKMPNASFS VSDFEPNE	80 : TLK IQRTOQE TLE IQRTEQG TLTIQRTOQE TLKIQRTERG TLTVTSAHPE TLVTSAHPE TLKIQPSEPR TLKIQPSEPR PLTLESARPS	90 E DSAVYLCAS DSAVYLCASS DSAVYLCASS DSSFYICSA DSSFYICSA DSSFYICSA DSAVYFCASS DSAVYFCASS DSAVYFCASS DSAVYFCASS
C ph 27 ATL2-1 C ATL2-1 C ph 29 C ATL12-1 C ph 24	OL VPF VLCGODTWM MÁSLLFFCGAFYLLGTGSM MDTRLLCCAVICLLGAGLS MCLRLLCCVAISFWGAST VLLCLLGAGPV	DAGITQSPRHKVTETG DAOVTOTPRNRIKTG NAGVMQNPRHLVRRRG DTKVTQRPRLLVKASE RAGVTQTPRHLIKTRG * *	TPVTLRCHQTEN KRIMLECSOTKG OEARLRCSPMKG OKAKMDCVPIKA OQVTLGCSPISG *	HRYMYWY HDRMYWY HSHVYWY HSYVYWY HRSVSWY	RQDPGLGLRQ RQDPGLGLR L RQLPEEGLKFN RKKLEEELKFL QQTLGQGLQFL	IH YSYGVKD IY YSFDVKD VY LQKENII VY FQNEELI .FEYF SETQ	TDKGEVSDGY TDKGEISDGY INKGEISDGY DESGMPKERF QKAEIINERF RNKGNFLGRF	SVSRKERNF SVSRSKTEDF SVSRQAQAKF SAEFPKEGPS LAQCSKNSSC SGRQFSNSRS	LLTLESPSPN LLTLESATSS SLSLESATPN ILRTOOVVRG TLETOSTESG EMNVSTLELG	QTSLIFCAS QTSVYFCAIS QTALYFCATS DSAAYFCASS DTALYFCASS DSALYLCAS * *



leader sequences are indicated and aligned separately. Asterisks identify invariant residues. Brackets group the 15 sequences into ten subfamilies according to homology (Fig. 2).

Fig. 5. Variability and hydrophilicity analysis of human and murine  $V_{\beta}$  genes. Variability analysis was done according to Wu and Kabat (25). Relative hydrophilicity was determined from the algorithm of Hopp and Woods (29). Amino acid position is plotted on the *x*-axis. (a) Variability plot of the 15 human  $V_{\beta}$  genes translated in Fig. 3, including the D to J junctions. (b) Superimposed relative hydrophilicity plots of the following 12 human  $V_{\beta}$  genes: ph 5, ph 22, ph 16, ph 34, ph 8, ph 15, ph 21, ph 27, ph 32, ph 29,  $V_{\beta}ATL$  12-1, ph 24. (c) Superimposed relative hydrophilicity plots of the following 12 murine  $V_{\beta}$  genes [nomenclature according to Barth *et al.* (8) and Behlke *et al.* (10)]:  $V_{\beta}1$ ,  $V_{\beta}2$ ,  $V_{\beta}3$ ,  $V_{\beta}4$ ,  $V_{\beta}5.1$ ,  $V_{\beta}6$ ,  $V_{\beta}7$ ,  $V_{\beta}8.1$ ,  $V_{\beta}9$ ,  $V_{\beta}11$ ,  $V_{\beta}12$ , and  $V_{\beta}14$ . (d) Variability plot of 19 murine  $V_{\beta}$  genes; D to J junctions are not included.

are also present in TCR V<sub> $\alpha$ </sub> genes (22–24) and in the variable region genes of immunoglobulins (25). Despite a xenogeneic difference between their MHC-restriction elements, no common features could be defined that would distinguish human V<sub> $\beta$ </sub> genes from their murine counterparts. Variability among the 15 human V<sub> $\beta$ </sub> genes was analyzed in a Wu-Kabat variability plot (25) (Fig. 5a). In addition to variability encoded within the V<sub> $\beta$ </sub> gene segments, significant variability is introduced at the V<sub> $\beta$ </sub>-D<sub> $\beta$ </sub>-J<sub> $\beta$ </sub> junction. For comparison, a Wu-Kabat variability analysis of 19 murine V<sub> $\beta$ </sub> gene segments is included (Fig. 5d).

The hypervariable regions of immunoglobulins are believed to be involved in antigen interaction (26–28). By analogy, it would be predicted that MHC/antigen specificity should reside in the hypervariable regions of the TCR. Regions directly involved in ligand interaction should be in the hydrophilic regions of the molecule. Twelve human  $V_{\beta}$  genes and 12 murine  $V_{\beta}$  genes were analyzed with the hydrophilicity algorithm of Hopp and Woods (29). As shown



SCIENCE, VOL. 233

in Fig. 5, b and c, the areas of highest variability on the Wu-Kabat analysis correspond, in general, with hydrophilic regions of the molecule, consistent with these regions being available to interact with ligand. In addition, the hydrophilicity plot shows conservation of several hydrophobic regions, which are thought to be involved in the formation of a common tertiary structure shared among all TCR chains as well as members of the immunoglobulin supergene family (4, 30).

Pairwise comparison of the 15 human  $V_{\beta}$ genes in Figs. 1 and 4 with known murine  $V_{\mbox{\scriptsize B}}$  sequences does not present any clear picture of allelic evolutionary relationships. For example, human  $V_{\beta}$  genes ATL 12-2, ph 22, ph 16, HPB-MLT, ph 8, YT35, and ph 29 are all more homologous to the murine  $V_{\beta}$ 11 gene than any other known murine  $V_{\beta}$  gene (54 to 68% at the amino acid level, 64 to 78% at the nucleotide level). Some of these genes are actually more homologous to the murine gene than they are to each other. Thus seven different human  $V_{\beta}$  genes representing three different subfamilies all show significant homology to the same murine  $V_{\beta}$  gene. In other cases, significant homology is found in only a single murine-human pairwise combination. ATL 12-1 is most homologous to the murine  $\mathrm{V}_{\beta}3$  gene (54% amino acid, 66% nucleotide), ph 24 is most homologous to the murine  $V_{\beta}5.2$  gene (55% amino acid, 65% nucleotide), and ph 34 and Molt 4 are most homologous to the murine  $V_{\beta}$ 15 gene (47%) amino acid, 64% nucleotide). When pairwise comparisons between the most homologous murine and human  $V_{\beta}$  genes are aligned together, amino acid substitutions are seen to occur more often in hydrophilic regions of the molecule, and the frequency of these substitutions roughly parallels the plot of  $V_{\beta}$  variability in general. In these cases, therefore, evolution of  $V_{\beta}$  alleles between species is similar to the evolution of nonallelic  $V_{\beta}$  genes within a species, with changes being concentrated in the hydrophilic regions of the molecule.

Note added in proof: Similar results have been obtained by Kimura et al. (31).

**REFERENCES AND NOTES** 

- O. Acuto et al., Cell 34, 717 (1983).
   J. Kappler, R. Kubo, K. Haskins, J. White, P. Marrack, *ibid.*, p. 727.
   B. W. McIntyre and J. P. Allison, *ibid.*, p. 739.
   J. Novotny, S. Tonegawa, H. Saito, D. Kranz, H. Eisen, Proc. Natl. Acad. Sci. U.S.A. 83, 742 (1986).
   R. M. Zinkernagel and P. C. Doherty, Nature (London) 248, 701 (1974).
   G. M. Shearer, Eur. J. Immunol. 4, 527 (1974).
   Pattern et al. Nature (London) 312, 40 (1984).

- O. N. Shcate, Lut. J. Humanic. 7, 227 (1974).
   P. Patten et al., Nature (London) 312, 40 (1984).
   R. Barth et al., ibid. 316, 517 (1985).
   M. Behlke et al., Science 229, 566 (1985).
   M. Behlke, H. S. Chou, K. Huppi, D. Y. Loh, Proc. Natl. Acad. Sci. U.S.A. 83, 767 (1986).

- 11. K. Ikuta, T. Ogura, A. Shimizu, T. Honjo, *ibid.* 82, 7701 (1985).
- N. Jones et al., Science 227, 311 (1985).
   A. Tunnacliffe, R. Kefford, C. Milstein, A. Forster, T. H. Rabbitts, Proc. Natl. Acad. Sci. U.S.A. 82,
- 5068 (1985).
- G. Siu et al., Cell 37, 395 (1984).
   Y. Yanagi et al., Nature (London) 308, 145 (1984).
   M. Weigert, R. Riblet, Cold Spring Harbor Symp.
- Quant. Biol. 41, 839 (1977).
   M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions (National Bureau of Standards Applied Mathematics Series No. 55, Govern-ment Printing Office, Washington, DC, 1970). 18. A maximum likelihood estimate of n = 21 was
- obtained by determining the value of n that maximized P(d) for d and r fixed at d = 14 and r = 22. A 95% one-sided confidence bound of n = 40 was obtained by determining the smallest value of n for which the probability of observing at most 14 different segments among 22 segments examined was less than 0.05.
- S. Tonegawa, Nature (London) 302, 575 (1983).
   B. Toyonaga, Y. Yoshikai, V. Vadasz, B. Chin, T. W. Mak, Proc. Natl. Acad. Sci. U.S.A. 82, 8624
- (1985)

- L. Hood, J. H. Campbell, S. C. R. Elgin, Annu. Rev. Genet. 9, 305 (1975).
   D. Becker et al., Nature (London) 317, 430 (1985).
   Y. Chien et al., ibid. 312, 31 (1984).
   H. Saito et al., ibid., p. 36.
   E. A. Kabat, T. T. Wu, H. Bilofsky, M. Reid-Miller, H. Perry, Sequences of Proteins of Immunological Inter-ret (National Interiments of Lashkeit Derberdet MPest (National Institutes of Health, Bethesda, MD, 1983).

- 26. L. M. Amzel and R. J. Poljak, Annu. Rev. Biochem. 48, 961 (1979)
- D. R. Davies and H. Metzger, Annu. Rev. Immunol. 1, 87 (1983). 27.
- T. Wu and E. Kabat, J. Exp. Med. 132, 211 (1970).
   T. P. Hopp and K. R. Woods, Proc. Natl. Acad. Sci. U.S.A. 78, 3824 (1981). 29.
- J. Novotny and E. Haber, *ibid.* 82, 4592 (1985). N. Kimura, B. Toyonaga, Y. Yoshikai, T. Mak, J.
- 31. 32.
- N. Kimura, B. Toyonaga, Y. Yoshikai, T. Mak, J. Exp. Med., in press.
  T. V. Huynh, R. A. Young, R. N. Davis, in DNA Cloning: A Practical Approach, D. Glover, Ed. (IRL Press, Oxford, 1984), pp. 49–78.
  R. Epstein et al., J. Exp. Med. 161, 1219 (1985).
  A. Maxam and W. Gilbert, Proc. Natl. Acad. Sci. U.S.A. 74, 560 (1977).
  F. Sanger, S. Nicklen, A. Coulsen, *ibid.*, p. 5463.
  M. D. Biggin, T. J. Gibson, G. F. Hong, *ibid.* 80, 3963 (1983).
- 33. 34.
- 35. 36.
- 37
- 38
- 3963 (1983).
  G. I. Bell et al., ibid. 78, 5759 (1981).
  E. M. Southern, J. Mol. Biol. 98, 503 (1975).
  T. Maniatis, E. Fritsch, J. Sambrook, Molecular Cloning: A Laboratory Manual (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1982).
  We thank Dr. Daniel Hartl for assistance with the tetristical analysis and Medline Baselment for manual for the second sec 39
- 40. statistical analysis and Madline Pearlman for manuscript preparation. Supported by funds from the Howard Hughes Medical Institute (D.Y.L.), NIH training grant 2-T32-A100112 (J.P.T.), and Nation-al Institute of Health Research Service Award, GM 07200, Medical Scientist Training Program, from the National Institute of General Medical Sciences (M.A.B.).

3 March 1986; accepted 6 June 1986

## Replication of the B19 Parvovirus in Human Bone Marrow Cell Cultures

Keiya Ozawa,\* Gary Kurtzman, Neal Young

The B19 parvovirus is responsible for at least three human diseases. The virus was successfully propagated in suspension cultures of human erythroid bone marrow from patients with hemolytic anemias; release of newly synthesized virus into the supernatants of infected cultures was observed. This culture system allowed study at a molecular level of events associated with the B19 life cycle. The B19 parvovirus replicated through high molecular weight intermediate forms, linked through a terminal hairpin structure. B19 replication in vitro was highly dependent on the erythropoietic content of cultures and on addition of the hormone erythropoietin.

HE B19 PARVOVIRUS WAS DISCOVered in the sera of normal blood donors in 1975 (1) and subsequently was identified as the causative agent of (i) transient aplastic crisis of hemolytic disease (2), (ii) the common childhood exanthem called fifth disease (3), and (iii) a polyarthralgia syndrome in normal adults (4). In addition, B19 parvovirus may be responsible for some cases of hydrops fetalis (5). The genome of the B19 virus consists of about 5.4 kb of single-stranded DNA (6-8), which encodes at least two capsid proteins of about 84 and 58 kD (9, 10). Genetic studies have been hampered by the inability to propagate the B19 parvovirus in vitro. Even though B19-containing sera are potent inhibitors of colony formation by bone marrow erythroid progenitor cells (11), clonogenic assays are inadequate for either maintenance or molecular study of the B19 virus. We report here successful propagation of B19 parvovirus in a tissue culture system by means of fresh human bone marrow cells.

Serum containing B19 parvovirus (Minor II, 60 µg of B19 DNA per milliliter) was obtained from a patient with sickle cell disease and transient aplastic crisis during a recent B19 epidemic (12). It was adsorbed to erythroid bone marrow mononuclear cells that had been obtained with informed consent from patients with hemolytic anemias; the infected cells were then grown in standard growth medium in the presence of the erythroid cell-specific hormone erythropoietin. Virus addition only slightly de-

Cell Biology Section, Clinical Hematology Branch, Na-tional Heart, Lung, and Blood Institute, Bethesda, MD 20892.

ACRF 7C108, National Institutes of Health, Bethesda, MD 20892. \*To whom communications should be addressed at