Fifty-million-year-old polymorphism at an immunoglobulin variable region gene locus in the rabbit evolutionary lineage

CHEN SU AND MASATOSHI NEI*

Institute of Molecular Evolutionary Genetics and Department of Biology, Pennsylvania State University, University Park, PA 16802

ABSTRACT The Ig heavy chain variable region (V_H) genes encode the antigen-binding regions of antibodies. The rabbit genome contains more than 100 V_H genes, but only one (V_H1) is preferentially used in the VDJ gene rearrangement. Three highly divergent alleles occur at this V_H1 locus in most rabbit populations. These three V_H alleles are also present in snowshoe hare populations, indicating that the polymorphism of the V_H1 alleles is trans-specific. Here we report the results of a phylogenetic analysis of rabbit Ig germ-line V_H genes (alleles) together with V_H genes from humans and mice. We have found that all rabbit V_H genes belong to one mammalian V_H group (group C), which also includes various human and mouse V_H genes. Using the rate of nucleotide substitution obtained from human and mouse V_H sequences, we have estimated that the V_H1 polymorphism in the rabbit lineage has been maintained for about 50 million years. This extremely long persistence of V_H1 polymorphism is apparently caused by overdominant selection, though the real mechanism is unclear.

The Ig heavy chain variable region (V_H) genes of the rabbit *Oryctolagus cuniculus* are organized in the genome in a manner similar to those in humans and mice and contain multiple V_H, D (diversity), and J_H (joining) region segment genes. The number of V_H genes was estimated to be more than 100 (1). In the rabbit, however, only one V_H gene, i.e., the D-region-proximal V_H gene (V_H1), is preferentially used in the VDJ rearrangement (Fig. 1). Approximately 80% of the rabbit serum Ig molecules are composed of the product of this gene (2, 3).

There are two allotype groups of Ig molecules in rabbits: V_Ha and V_Ha^- (4). In the V_Ha group, there are three allotypes, a1, a2, and a3, and they are encoded by three alleles at the V_H1 locus, V_H1-a1, V_H1-a2, and V_H1-a3, respectively. In the V_Ha^- group, there are two major allotypes, x32 and y33, which are encoded by the V_Hx32 allele at the V_Hx locus and the V_Hy33 allele at the V_Hy locus, respectively. The occurrence of the V_Ha^- group in the rabbit serum Ig is much lower than that of the V_Ha group. Except for the V_H1, V_Hx, and V_Hy genes, most other V_H genes (including potentially functional genes and pseudogenes) are not used in the VDJ rearrangement. Instead, they are used as donor sequences in somatic gene conversion during B lymphocyte development. These genes play an important role in the diversification of the primary antibody repertoire in the rabbit gut-associated lymphoid tissue (5–8).

The alleles V_H1-a1, V_H1-a2, and V_H1-a3 at the V_H1 locus are found in all domestic and wild rabbit populations studied thus far (9, 10). Moreover, they are found even in an Alaskan population of the snowshoe hare *Lepus americanus* (Table 1 and ref. 11), which belongs to a genus different from that of the rabbit. Therefore, this polymorphism is apparently trans-specific and has persisted through several speciation events at
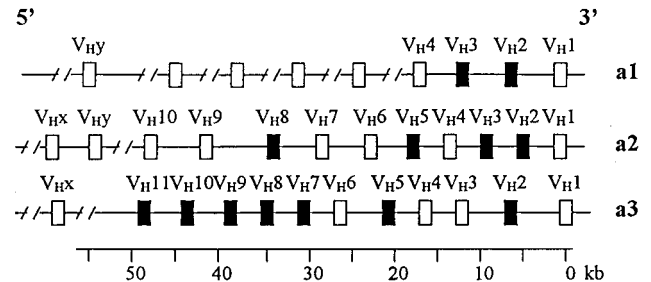
FIG. 1. The genomic organization of V_H genes in haplotypes V_H-a1, -a2, and -a3 (4). Each haplotype was estimated to have more than 100 V_H genes in a 750-kb region. Only genes with known genomic locations are shown. The putative pseudogenes are indicated with black boxes. V_Hx and V_Hy are at the 5′ side of the all other V_H genes shown here, but their locations are not well established. The genomic maps are approximate.

least from the time of divergence of rabbits and hares. In this aspect, the polymorphism at the rabbit V_H1 locus is similar to that at the MHC loci in mammals (12).

The polymorphism of the rabbit V_H1 gene is quite different from that of human and mouse V_H genes. The antigen-binding regions of human and mouse V_H genes have been shown to be subject to positive Darwinian selection (13), but the extent of genetic polymorphism is much lower than that of mammalian MHC loci (14, 15). For this reason, Nei and coworkers (16) speculated that the selection at human and mouse V_H genes is directional rather than of heterozygote advantage. However, the rabbit V_H1 locus has a high degree of polymorphism, and this polymorphism apparently has persisted in the population for a long evolutionary time. Note also that the V_H1 polymorphism exists as part of the haplotype polymorphism in association with the neighboring genes (Fig. 1), and this haplotype polymorphism is similar to that at the MHC-DRB loci in hominoids (54). It is therefore interesting to know the persistence time and to examine the possible mechanism of maintenance of this unusual polymorphism.

MATERIALS AND METHODS

V_H1 Polymorphism: Minimum Persistence Time. Because the V_H1 polymorphism is shared by rabbits and snowshoe hares, this polymorphism must have existed in the ancestral population of these two species. Therefore, finding the time of divergence between these two species will give us a minimum estimate of the time of persistence of V_H1 gene polymorphism. Unfortunately, no fossil record is available that can be used for this purpose. We therefore took the molecular approach to

Table 1.   Frequencies of Ig V_H1 alleles in rabbit and snowshoe hare populations

Allele	Domestic rabbits* ($n = 112$)	Wild rabbits				Snowshoe hares Yukon* ($n = 684$)
		France* ($n = 157$)	Australia* ($n = 439$)	Kerguelen* ($n = 1,338$)	Europe *et al.*.* ($n = 1,601$)	
V_H1-a1	0.196	0.484	0.515	0.172	0.479	0.335
V_H1-a2	0.191	0.251	0.157	0.136	0.133	0.528
V_H1-a3	0.61	0.222	0.328	0.691	0.388	0.137

n, Number of individuals studied in each population.
*Original data from refs. 9–11.

estimate this time, using information on the divergence time between humans and rabbits for calibrating the time scale. We searched for DNA and protein sequences available for rabbits, snowshoe hares, and humans, examining the databases of GenBank, HOVERGEN (17), and TREEBASE (http://phylogeny.harvard.edu/treebase) and found four genes that can be used for our purpose: mitochondrial cytochrome *b* gene (cytb), polychlorinated biphenyl (PCB) binding protein gene, 12S rRNA gene, and Ig κ chain constant region (C_κ) gene. Previously, for a similar purpose, Bouton and van der Loo (18) used the number of synonymous nucleotide substitutions for the C_κ gene to eliminate the effect of positive Darwinian selection. We followed their approach with this gene and used synonymous distances obtained by the method of Nei and Gojobori (19). For the 12S rRNA gene, we used the Jukes–Cantor distance (20). For the other two genes, we used Poisson correction distances for amino acid sequences, but other distance measures (e.g., gamma distances) gave essentially the same results.

Let d_{12}, d_{13}, and d_{23} be the distances between rabbits and hares, rabbits and humans, and hares and humans, respectively. Under the assumption of a molecular clock, we can estimate the time of divergence (t) between rabbits and hares by

$$\hat{t} = \frac{2d_{12}}{d_{13} + d_{23}} T, \qquad [1]$$

where T is the time of divergence between humans and rabbits and was assumed to be 90 million years (MY) (21). The SE of $\hat{t}$ was obtained by the bootstrap method with 1,000 replications for each gene. The hypothesis of the molecular clock was examined by using Tajima's (22) relative rate test.

V_H1 Polymorphism: More Appropriate Persistence Time. The above method, however, may give a serious underestimate of the persistence time of the V_H1 polymorphism, because the polymorphism may have existed a long time before the divergence of rabbits and snowshoe hares. A more appropriate estimate of the persistence time is obtained by comparing the extent of sequence divergence between the V_H1-a1, V_H1-a2, and V_H1-a3 alleles (haplotypes) with that between human and mouse V_H gene sequences. Because we have some idea about the time of divergence between humans and mice, it is possible to estimate the time of divergence between the rabbit V_H1-a1, V_H1-a2, and V_H1-a3 alleles, assuming that the most closely related human and mouse V_H genes diverged at the time of speciation. This assumption is unlikely to be correct, because even the most closely related human and mouse V_H genes would have diverged before the divergence between the human and mouse lineages. However, this method is still useful for obtaining a conservative estimate of the time of divergence of V_H1 polymorphic alleles; we therefore used this approach. For this purpose, we used the method of the linearized tree devised by Takezaki *et al.* (23) to estimate the persistence time of the polymorphic alleles.

V_H Sequences Used in the Analysis. We used the following rabbit V_H genes in the analysis (the numbers in parentheses are the GenBank accession numbers):

V_H-a1 haplotype: V_H1-a1 (M93171), V_H3-a1ψ (M93177), V_H4-a1 (M93181).

V_H-a2 haplotype: V_H1-a2 (M93172), V_H2-a2ψ (M93175), V_H3-a2ψ (M93178), V_H4-a2 (M93182), V_H5-a2 (U51025), V_H6-a2 (U51026), V_H7-a2 (U51027), V_H8-a2 (U51028), V_H9-a2 (U51029), V_Hx32 (L03846), V_Hy33 (L03890).

V_H-a3 haplotype: V_H1-a3 (M93173), V_H2-a3ψ (M93176), V_H3-a3 (M93179), V_H4-a3 (M93183), V_H5-a3ψ (M93184), V_H6-a3 (M93185), V_H7-a3ψ (M93186), V_H8-a3ψ (L27311), V_H9-a3 (L27312), V_H10-a3ψ (L27313), V_H11-a3ψ (L27314).

The number immediately following "V_H" in the above gene notation indicates the genomic order beginning from the 3′ end of the V_H gene region (see Fig. 1). The notation a1, a2, or a3 refers to the haplotype to which each gene belongs. For example, V_H2-a3 refers to the second V_H gene from the 3′ end in haplotype a3. We included all available rabbit germ-line V_H genes with known genomic locations except V_H2-a1, for which only a partial sequence was available. The pseudogenes are indicated by "ψ" at the end of the gene notation. (For details of the sequences, see refs. 2, 3, and 24.) We included rabbit V_H pseudogenes in this analysis because they are used as donor sequences in somatic gene conversion and are not really "dead" genes as in the case of chicken pseudogenes (25).

A large number of V_H genes are present in the human and mouse genomes, and human and mouse V_H genes have been classified into 7 and 15 families, respectively (26, 27). To choose the appropriate human and mouse sequences for our analysis, we first constructed a phylogenetic tree with all of the human and mouse genes available from the GenBank and VBASE database (26). This tree showed that the genes belonging to a V_H family of a species always form a reliable cluster. We therefore used V_H sequences in the subsequent analysis that are representatives from each of the human and mouse V_H families, excluding pseudogenes.

The rabbit, human, and mouse V_H gene sequences were aligned by the CLUSTAL W program (28), followed by visual adjustments. In this study, only the framework regions (FR1, FR2, and FR3) of V_H genes were used, because the complementarity-determining regions (CDRs) were so divergent that they were unalignable (29). The total number of nucleotides used for a V_H sequence was 225 bp, but some rabbit sequences contained a few codon deletions. By contrast, the rabbit V_Hy33 gene had a unique 4-codon insertion, and this insertion was eliminated from the analysis.

Phylogenetic Analysis. To construct a phylogenetic tree for the rabbit, human, and mouse V_H genes, we used the neighbor-joining method (30) with the horned shark gene V_H101 and the little skate gene V_H278 as outgroup sequences. These outgroup sequences were previously used by Nei and coworkers (16) when rooting a tree for V_H genes from higher vertebrates. We used Kimura's two-parameter distance (31) and the Kimura gamma distance (32) with a shape parameter of $a = 2.64$ to measure the evolutionary distance for each pair of sequences. However, because these two distance measures gave essentially the same results, we present only the results from Kimura's two-parameter distance. All of these computations, except the estimation of gamma parameter a, were conducted by using the computer program MEGA (33). The

gamma parameter *a* was estimated by the program PAML (34). Because there were only a few gaps in the sequence alignment, we used the pairwise deletion option in MEGA to compute evolutionary distances, but essentially the same results were obtained when all alignment gaps were eliminated (the complete deletion option). In the construction of the phylogenetic tree, we also used the maximum parsimony method (35), but the tree obtained by this method was similar to the neighbor-joining tree. Therefore, we present only the neighbor-joining tree in this paper.

RESULTS

Time of Divergence Between Rabbits and Snowshoe Hares. Tajima's test (22) showed that the molecular clock hypothesis could not be rejected for any of the four genes used. Although this test examines only the equality of the branch lengths leading to the rabbit and the hare lineages, we assumed that the molecular clock approximately applies to all three sequences, including the human lineage. Estimates of the time of divergence between rabbits and hares obtained under this assumption (Eq. **1**) are presented in Table 2. The estimate varies from 13.1 to 29.0 MY ago, the average being 21.9 ± 3.8 MY ago. Recently, Halanych and Robinson (36) used 12s rRNA data to estimate the time of generic radiation within the Leporidae (hares and rabbits), which corresponds roughly to the time of divergence between the snowshoe hares and European rabbits in which we are interested (36, 37). By assuming that the time of divergence between ochotonids (pikas) and leporids was 30–40 MY ago (37), they estimated that the time of generic radiation in the Leporidae was 12.2 to 16.3 MY ago. Because they used a lower bound of the divergence time between pikas and leporids, their estimate is probably too conservative. Therefore, their estimate is not far from ours (20 MY ago).

Phylogenetic Tree for Human, Mouse, and Rabbit V_H Genes. The first step in constructing a linearized tree is to produce a phylogenetic tree without the assumption of a molecular clock and then to eliminate sequences that do not follow the clock (23). The phylogenetic tree is presented in Fig. 2. In this figure the mouse genes are denoted by the V_H family numbers to which they belong, whereas the human genes are designated by their V_H family numbers followed by the locus numbers as indicated by Tomlinson *et al.* (26).

Ota and Nei (29) have shown that the human and mouse V_H genes can be divided into three major V_H groups, A, B, and C (see Fig. 2). By contrast, all rabbit V_H genes are closely related to one another and form one monophyletic group which is included in group C. As far as the sequenced genes are concerned, there are no other group genes in the rabbit genome. This clustering pattern is similar to that of the chicken V_H genes, which also belong to group C.

The phylogenetic tree in Fig. 2 suggests that the human and rabbit genes evolved at nearly the same rate, except for a few group B human genes. By contrast, most mouse genes appear to have evolved considerably faster than human and rabbit genes. We therefore examined the heterogeneity of evolutionary rate among different lineages by using the two-cluster test
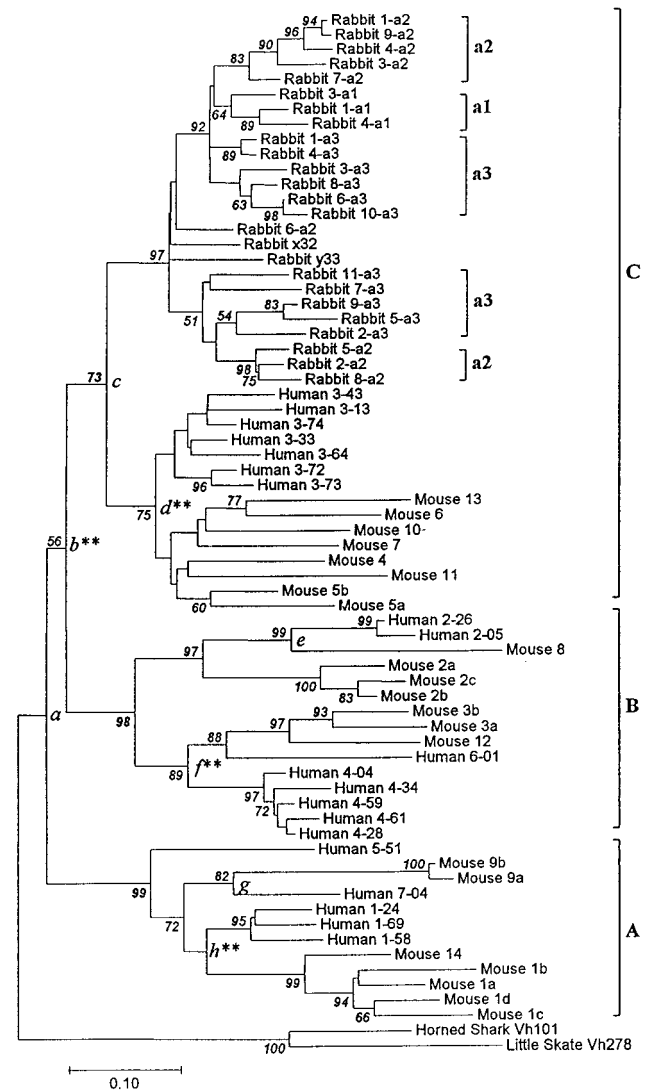

FIG. 2. Neighbor-joining tree for rabbit, human, and mouse V_H genes. The root of the tree is determined by V_H genes from cartilaginous fish. Numbers given to interior branches are the bootstrap values (P_B) when 1000 replicate resamplings were done. The P_B values lower than 50% are not shown. The branch lengths (see scale at bottom) are measured in terms of the number of nucleotide substitutions per site in the framework regions (225 bp). Three mammalian V_H groups, i.e., groups A, B, and C, are indicated by the brackets. The nodes (*d*, *e*, *f*, *g*, and *h*) are indicated for comparison of human and mouse gene clusters; ∗∗ indicates significance at the 1% level.

(23) of Takezaki and coworkers. This test examines the null hypothesis that the average evolutionary rates of two clusters separated by a node of a tree are the same.

Application of this test to the two clusters separated by node *a* (see Fig. 2) shows that the difference between the average

Table 2. Estimates of the divergence time (*T*) between rabbits and snowshoe hares

Gene	Distance, nucleotide substitutions per 100 sites			*T*, MY
	Human–rabbit	Human–hare	Rabbit–hare	
PCB ($n = 91$)	48.6 ± 8.3	46.8 ± 8.1	6.8 ± 2.8	13.1 ± 5.3
Cytb ($n = 240$)	21.3 ± 3.3	20.7 ± 3.2	5.6 ± 1.6	20.3 ± 5.8
12S rRNA ($n = 714$)	25.0 ± 2.2	26.0 ± 2.3	7.1 ± 1.1	25.1 ± 3.9
Ig C$_\kappa$ ($n = 65$)	60.5 ± 15.9	46.0 ± 12.7	13.2 ± 5.6	29.0 ± 12.2
Average				21.9 ± 3.8

n, Number of codons after all gaps were eliminated except for the 12S rRNA gene, for which *n* refers to number of nucleotide sites after all gaps were eliminated.

evolutionary rate of group A genes and of group B and C genes is not statistically significant. The two clusters connected by node *c* also do not show a significant rate difference. However, the average rate of group B genes is significantly higher than that of group C genes (see node *b*). There are five comparisons of human and mouse gene clusters (node *d, e, f, g,* and *h*), and three of them are significant at the 1% level. These tests support the idea that the mouse genes generally evolved faster than the human and rabbit genes. By contrast, the human and rabbit genes seem to have evolved nearly at the same rate.

The rabbit genes can be divided into five clusters and three isolated sequences (6-a2, x32, and y33), but none of the cluster pairs showed a significant difference in evolutionary rate. We also applied the branch length test of Takezaki and coworkers (23), using their *U* statistic. This test also did not reject the molecular clock hypothesis for rabbit sequences. Therefore, we can conclude that all of the rabbit genes evolved at nearly the same rate. The tree in Fig. 2 shows that the genes in the same haplotype (a1, a2, and a3) do not necessarily form a mono-phyletic group, suggesting that different clusters of rabbit genes existed before the emergence of these three haplotypes.

In the present paper, we are primarily interested in the polymorphism at the V_H1 locus. Therefore, we consider the persistence time of the top three clusters of haplotypes, a1, a2, and a3 in the tree of Fig. 2. In the standard way of constructing a linearized tree, all of the sequences that evolved too fast or too slow in comparison with the average rate for the entire set of sequences or the average rate for a specific group of sequences (e.g., rabbit-specific sequences) are eliminated. In the present case, however, if we eliminate all the mouse sequences that evolved faster, we cannot estimate the time of divergence between the rabbit polymorphic alleles (allelic lineages), because it is difficult to calibrate the time scale of evolution without information on the divergence time between the human and mouse lineages.

One solution to this dilemma is to use the mouse genes despite their faster evolution, with the understanding that this approach would give an underestimate of the persistence time of rabbit V_H1 gene polymorphism. This approach is acceptable for our purpose because we are interested in a lower bound of the persistence time. Previously we noted that the duplicate genes that were generated before the divergence between humans and mice would also give an underestimate of the persistence time of rabbit V_H1 polymorphism. Therefore, our estimate would be a lower bound in any case.

With these caveats, we constructed a linearized tree (Fig. 3). Note that this tree was constructed only for obtaining a lower bound estimate of the persistence time of V_H1 gene polymor-phism (indicated by an asterisk in Fig. 3); this tree should not be used for estimating the time of divergence for other clusters of genes (e.g., V_H gene clusters A and B). Fig. 3 shows that there are five bifurcating or multifurcating nodes that repre-sent the divergence between the human and mouse genes (indicated by five arrows in Fig. 3). The average of the branch lengths between these five nodes and the tips of the branches is 0.14 ± 0.01.

At the present time, there is no good geological estimate of the time of divergence between humans and mice. However, the recent finding of the fossils of ungulates, which were dated 85 MY old, suggests that humans and mice might have diverged more than 85 MY ago (38). In fact, molecular data suggest that they diverged about 110 MY ago (21). In this paper, we use 100 MY, as suggested by Novacek (39). We then obtain the rate of nucleotide substitution for V_H genes as 1.4×10^{-9} per site per year per lineage. The average branch length between the rabbit V_H1 polymorphic alleles and their ancestral node is 0.07. Therefore, dividing 0.07 by 1.4×10^{-9}, we obtain 50 MY as an estimate of the persistence time of rabbit V_H1 polymorphism. The SE of this estimate is 7 MY, thus the 95% confidence interval of the persistence time is 36–64 MY. This estimate is
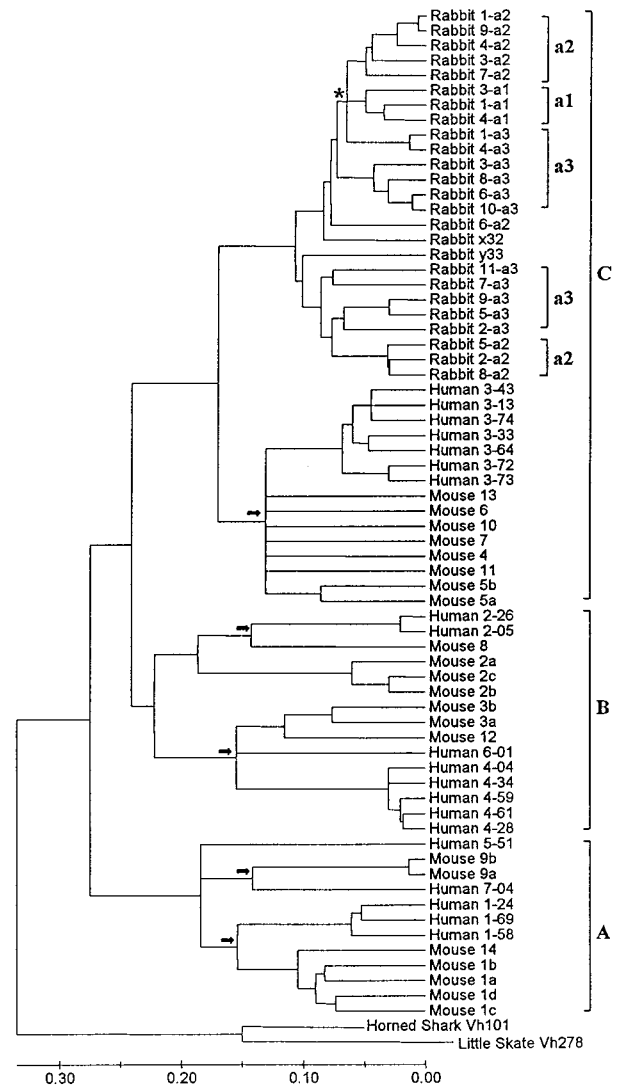

FIG. 3. Linearized tree constructed under the assumption of a molecular clock for the same V_H genes shown in Fig. 2. The arrows indicate the time of divergence of the putatively orthologous human and mouse V_H lineages. The asterisk (*) near the top indicates the node corresponding to the divergence of three rabbit V_H1 alleles. The scale at the bottom is nucleotide substitutions per site.

much greater than the minimum estimate (20 MY) obtained from the comparison of rabbit and hare gene sequences.

DISCUSSION

Persistence Time of Rabbit V_H1 Polymorphism. The above statistical analysis suggests that the V_H1 polymorphism in the rabbit lineage has persisted for about 50 MY. This persistence time is extremely long and various factors affecting the esti-mate suggest that the estimate is a lower-bound, rather than an upper-bound estimate. We have already mentioned that the gene duplication preceding the separation between the human and the mouse lineages and the faster rates of mouse V_H genes both tend to give an underestimate. One might think that the use of group C genes only, excluding group A and B genes, would give a better estimate. However, this approach actually gives a higher estimate (70 MY), because the divergence between the human and mouse sequences is shorter than that for group B and C genes (Fig. 3).

The results presented in Fig. 3 are based on the Kimura distance, without considering the variation in substitution rate

Table 3. Average nucleotide substitutions per synonymous site ($\bar{d}_S$) and per nonsynonymous site ($\bar{d}_N$) at the rabbit Ig V_H loci

Genes	No. of sequences	CDRs			FRs		
		$\bar{d}_S$	$\bar{d}_N$	$\bar{d}_N/\bar{d}_S$	$\bar{d}_S$	$\bar{d}_N$	$\bar{d}_N/\bar{d}_S$
V_H1 alleles	3	0.17 ± 0.11	0.16 ± 0.05	0.94	0.11 ± 0.04	0.14 ± 0.03	1.3
Functional genes	16	0.23 ± 0.09	0.21 ± 0.04	0.91	0.13 ± 0.03	0.19 ± 0.02	1.4
Pseudogenes	9	0.23 ± 0.09	0.23 ± 0.04	1.0	0.13 ± 0.03	0.20 ± 0.02	1.5*

$\bar{d}_N$ and $\bar{d}_S$ were estimated by the method of Nei and Gojobori (19), and the standard errors were computed by the method of Nei and Jin (42). All computations were done with the sequence data after complete deletion, which excluded all sites having one or more alignment gaps.
*$\bar{d}_N$ is significantly higher than $\bar{d}_S$ at the 5% level ($P < 0.05$).

among different sites, and this might have given an overestimate of persistence time. However, when we used the Kimura gamma distance, the persistence time decreased only slightly (46 MY). The use of amino acid sequence data with a gamma parameter ($a = 2.00$) gave a much higher estimate of divergence time, probably because synonymous substitutions are ignored in the latter analysis.

In this study, we assumed that the human and mouse lineages diverged about 100 MY ago, but this could be an overestimate. If we use the available fossil record, which appears to be at most 85 MY, the estimate of persistence time of V_H1 polymorphism becomes 43 ± 6 MY. Even this length of time is very long.

With regard to the level of allelic polymorphism, only a few MHC class I and class II loci are comparable to the rabbit V_H1 locus. Some of the polymorphic MHC allelic lineages have apparently persisted for more than 25 MY, predating the separation between the human lineage and the Old World monkey lineage (43–46). However, polymorphic alleles of MHC class I and class II loci are not shared by hominoids and New World monkeys, suggesting that MHC polymorphism is less than 40 MY old (47–49, 52). Some authors claimed older polymorphism (e.g., ref. 50), but these claims are not substantiated by the critical test proposed by Nei and Rzhetsky (51).

Mechanism of Maintenance of V_H1 Polymorphism. What is the mechanism of maintenance of this polymorphism for such a long time? The most likely explanation is that the three polymorphic alleles or allelic lineages have been maintained by overdominant selection, as in the case of MHC loci (40, 41). However, the V_H1 locus has only three polymorphic alleles, and they are very divergent at the nucleotide level. MHC loci usually contain a large number of polymorphic alleles. For example, the class II MHC locus *DRB1* in humans contains more than 100 alleles (53), indicating that the polymorphic patterns of MHC alleles and V_H1 alleles are different and suggesting that the three V_H1 alleles have had specific adaptive significance. It seems that these three alleles have adapted to cope with three different sets of pathogens and have become more or less irreplaceable.

Hughes and Nei (40, 41) detected positive selection by comparing the number of synonymous substitutions per synonymous site (d_S) and the number of nonsynonymous substitutions per nonsynonymous site (d_N). We computed the average values ($\bar{d}_S$ and $\bar{d}_N$) of d_S and d_N for the comparisons of the three alleles of the V_H1 locus, 16 seemingly functional V_H genes, and 9 apparently pseudogenes, separately. (Note that these pseudogenes are not really dead genes, because they are used for somatic gene conversion.) The results for the CDRs and the FRs are presented separately in Table 3. In the CDRs, $\bar{d}_N$ is nearly equal to $\bar{d}_S$, and there is no significant difference between them for any gene group. These results are different from those obtained in the human, mouse, and chicken genes (13, 25). In the latter genes $\bar{d}_N$ was significantly higher than $\bar{d}_S$, suggesting that they are subject to positive selection.

However, the equality of $\bar{d}_N$ and $\bar{d}_S$ does not mean that no positive selection is operating at the rabbit V_H loci, because a high rate of nonsynonymous substitution may occur at limited nucleotide sites of the CDRs. It is also possible that once a given set of motif amino acids is established for each allelic lineage to cope with a given set of foreign pathogens, there is no need to have further amino acid substitutions. In other words, without any further accelerated nonsynonymous substitution overdominant selection can be maintained. Therefore, the equality of $\bar{d}_N$ and $\bar{d}_S$ is not inconsistent with the hypothesis of overdominant selection.

Table 3 shows that in the FRs $\bar{d}_N$ is generally higher than $\bar{d}_S$ and the difference is statistically significant for the pseudogene group. These results are different from those of human and mouse V_H genes, where $\bar{d}_N$ was lower than $\bar{d}_S$. However, chicken V_H pseudogenes, which are also used for somatic gene conversions, show somewhat similar results, $\bar{d}_N$ being nearly equal to $\bar{d}_S$ (25). Therefore, the relatively high value of $\bar{d}_N$ in comparison to $\bar{d}_S$ seems to be a common feature for V_H pseudogenes that are used as donor sequences for somatic gene conversion.

In conclusion, the rabbit V_H1 polymorphism appears to have been maintained for an extremely long time. It is trans-specific, and the persistence time of the polymorphism appears to be much longer than that of MHC polymorphism, which is also trans-specific. However, the mechanism of maintenance of the rabbit V_H1 polymorphism is yet to be investigated.

1. Gallarda, J. L., Gleason, K. S. & Knight, K. L. (1985) *J. Immunol.* **135,** 4222–4228.
2. Knight, K. L. & Becker, R. S. (1990) *Cell* **60,** 963–970.
3. Raman, C., Spieker-Polet, H., Yam, P. C. & Knight, K. L. (1994) *J. Immunol.* **152,** 3935–3945.
4. Knight, K. L. & Tunyaplin, C. (1995) in *Immunoglobulin Genes*, eds. Honjo, T. & Alt, F. W. (Academic, San Diego), pp. 289–314.
5. Fitts, M. G. & Metzger, D. W. (1990) *J. Immunol.* **145,** 2713–2717.
6. Knight, K. L., Becker, R. S. & DiPietro, L. A. (1991) *Adv. Exp. Med. Biol.* **292,** 235–244.
7. Roux, K. H., Dhanarajan, P., Gottschalk, V., McCormack, W. T. & Renshaw, R. W. (1991) *J. Immunol.* **146,** 2027–2036.
8. Knight, K. L. & Winstead, C. R. (1997) *Int. Rev. Immunol.* **15,** 129–163.
9. Cazenave, P. A., Benammar, A., Sogn, J. A. & Kindt, T. J. (1987) in *The Rabbit in Contemporary Immunological Research*, ed. Dubiski, S. (Longman, Harlow, UK), pp. 148–163.
10. van der Loo, W., Bousses, P., Arthur, C. P. & Chapuis, J. L. (1996) *Genetics* **144,** 1181–1194.
11. van der Loo, W. (1987) in *The Rabbit in Contemporary Immunological Research*, ed. Dubiski, S. (Longman, Harlow, UK), pp. 164–190.
12. Klein, J. & Figueroa, F. (1986) *Crit. Rev. Immunol.* **6,** 295–386.
13. Tanaka, T. & Nei, M. (1989) *Mol. Biol. Evol.* **6,** 447–459.
14. Li, H. & Hood, L. (1995) *Genomics* **26,** 199–206.
15. Sasso, E. H., Buckner, J. H. & Suzuki, L. A. (1995) *J. Clin. Invest.* **96,** 1591–1600.
16. Nei, M., Gu, X. & Sitnikova, T. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 7799–7809.
17. Duret, L., Mouchiroud, D. & Gouy, M. (1994) *Nucleic Acids Res.* **22,** 2360–2365.

18. Bouton, C. & van der Loo, W. (1997) *Immunogenetics* **45,** 444–446.
19. Nei, M. & Gojobori, T. (1986) *Mol. Biol. Evol.* **3,** 418–426.
20. Jukes, T. H. & Cantor, C. R. (1969) in *Mammalian Protein Metabolism*, ed. Munro, H. N. (Academic, New York), pp. 21–132.
21. Kumar, S. & Hedges, S. B. (1998) *Nature (London)* **392,** 917–920.
22. Tajima, F. (1993) *Genetics* **135,** 599–607.
23. Takezaki, N., Rzhetsky, A. & Nei, M. (1995) *Mol. Biol. Evol.* **12,** 823–833.
24. Crane, M. A., Kingzette, M. & Knight, K. L. (1996) *J. Exp. Med.* **183,** 2119–2121.
25. Ota, T. & Nei, M. (1995) *Mol. Biol. Evol.* **12,** 94–102.
26. Tomlinson, I. M., Williams, S. C., Ignatovich, O., Corbett, S. J. & Winter, G. (1996) *VBASE Sequence Directory* (Medical Research Council Centre for Protein Engineering, Cambridge, UK).
27. Kofler, R., Geley S., Kofler, H. & Helmberg, A. (1992) *Immunol. Rev.* **128,** 5–21.
28. Thompson, J. D., Higgins, D. G. & Gibson, T. J. (1994) *Nucleic Acids Res.* **22,** 4673–4680.
29. Ota, T. & Nei, M. (1994) *Mol. Biol. Evol.* **11,** 469–482.
30. Saitou, N. & Nei, M. (1987) *Mol. Biol. Evol.* **4,** 406–425.
31. Kimura, M. (1980) *J. Mol. Evol.* **16,** 111–120.
32. Jin, L. & Nei, M. (1990) *Mol. Biol. Evol.* **7,** 82–102.
33. Kumar, S., Tamura, K. & Nei, M. (1993) MEGA, Molecular Evolutionary Genetics Analysis (Pennsylvania State Univ., University Park).
34. Yang, Z. (1997) *Phylogenetic Analysis by Maximum Likelihood* (PAML) (Dept. of Integrative Biology, Univ. of California, Berkeley).
35. Swofford, D. L. (1999) PAUP, Phylogenetic Analysis Using Parsimony (Sinauer, Sunderland, MA), Version 4.
36. Halanych, K. M. & Robinson, T. J. (1999) *J. Mol. Evol.* **48,** 369–379.
37. Dawson, M. R. (1981) in *Proceedings of the World Lagomorph Conference*, eds. Myers, K. & MacInnes, C. D. (Univ. of Guelph Press, Guelph, ON), pp. 1–8.
38. Archibald, J. D. (1996) *Science* **272,** 1150–1153.
39. Novacek, M. J. (1992) *Nature (London)* **356,** 121–125.
40. Hughes, A. L. & Nei, M. (1988) *Nature (London)* **335,** 167–170.
41. Hughes, A. L. & Nei, M. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 958–962.
42. Nei, M. & Jin, L. (1989) *Mol. Biol. Evol.* **6,** 290–300.
43. Mayer, W. E., Jonker, M., Klein, D., Ivanyi, P., van Seventer, G. & Klein, J. (1988) *EMBO J.* **7,** 2765–2774.
44. Fan, W. M., Kasahara, M., Gutknecht, J., Klein, D., Mayer, W. E., Jonker, M. & Klein, J. (1989) *Hum. Immunol.* **26,** 107–121.
45. Klein, J., Kasahara, M., Gutknecht, J. & Figueroa, F. (1990) *Chem. Immunol.* **49,** 35–50.
46. Lawlor, D. A., Ward, F. E., Ennis, P. D., Jackson, A. P. & Parham, P. (1988) *Nature (London)* **335,** 268–271.
47. Trtkova, K., Mayer, W. E., O'hUigin, C. & Klein, J. (1995) *Mol. Phylogenet. Evol.* **4,** 408–419.
48. Antunes, S. G., de Groot, N. G., Brok, H., Doxiadis, G., Menezes, A. A., Otting, N. & Bontrop, R. E. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 11745–11750.
49. Gyllensten, U., Bergstrom, T., Josefsson, A., Sundvall, M., Savage, A., Blumer, E. S., Giraldo, L. H., Soto, L. H. & Watkins, D. I. (1994) *Immunogenetics* **40,** 167–176.
50. Figueroa, F., O'hUigin, C., Tichy, H. & Klein, J. (1994) *J. Immunol.* **152,** 4455–4465.
51. Nei, M. & Rzhetsky, A. (1991) in *Evolution of MHC Genes*, eds. Klein, J. & Klein, D. (Springer, Heidelberg), pp. 13–27.
52. Watkins, D. I. (1995) *Crit. Rev. Immunol.* **15,** 1–29.
53. Lefranc, M. P., Giudicelli, V., Ginestoux, C., Bodmer, J., Muller, W., Bontrop, R., Lemaitre, M., Malik, A., Barbie, V. & Chaume, D. (1999) *Nucleic Acids Res.* **27,** 209–212.
54. Klein, J., O'hUigin, C., Figueroa, F., Mayer, W. E. & Klein, D. (1993) *Mol. Biol. Evol.* **10,** 48–59.