Simulation Study of the Reliability and Robustness of the Statistical Methods for Detecting Positive Selection at Single Amino Acid Sites

Yoshiyuki Suzuki and Masatoshi Nei

Institute of Molecular Evolutionary Genetics, Department of Biology, The Pennsylvania State University

Inferring positive selection at single amino acid sites is of biological and medical importance. Parsimony-based and likelihood-based methods have been developed for this purpose, but the reliabilities of these methods are not well understood. Because the evolutionary models assumed in these methods are only rough approximations to reality, it is desirable that the methods are not very sensitive to violation of the assumptions made. In this study we show by computer simulation that the likelihood-based method is sensitive to violation of the assumptions and produces many false-positive results under certain conditions, whereas the parsimony-based method tends to be conservative. These observations, together with those from previous studies, suggest that the positively selected sites inferred by the parsimony-based method are more reliable than those inferred by the likelihood-based method.

Introduction

Positive selection is an evolutionary event in which a mutant allele at a locus increases in frequency in a population because of a higher level of fitness compared with the wild-type allele. The occurrence of positive selection at the amino acid sequence level may be inferred from the observation that the rate of nonsynonymous nucleotide substitution (r_N) is higher than that of synonymous substitution (r_S) in protein-coding genes, i.e., $\omega > 1$, where $\omega = r_N/r_S$ (Hughes and Nei 1988, 1989).

Inferring positive selection at single amino acid sites is of biological and medical importance. By pinpointing the positively selected amino acid sites in a protein, one may infer the mechanism of positive selection, given that the biological function of these sites is well understood (e.g., antigen recognition sites in human leukocyte antigen [HLA] [Suzuki and Gojobori 1999]). In addition, neutralizing epitopes in viral proteins may be predicted as positively selected sites because such regions are expected to be positively selected (e.g., hepatitis C virus [Suzuki and Gojobori 2001]).

For inferring positive selection at single amino acid sites, Suzuki and Gojobori (1999) developed a parsimony-based (SG) method, whereas Yang et al. (2000) developed a likelihood-based (Yang) method. But the reliabilities of these methods are not well understood. The purpose of this article is to study the reliability and robustness of these methods by conducting computer simulation.

Materials and Methods

Methods

Both SG and Yang methods use a phylogenetic tree of protein-coding genes for detecting positive selection at single amino acid sites. In the SG method the following procedure is applied to each codon site. We first infer ancestral codons at all interior nodes of the phylogenetic tree by the maximum-parsimony (MP) method (Fitch 1971; Hartigan 1973). We then compute the total numbers of synonymous (c_S) and nonsynonymous (c_N) substitutions per codon site as well as the average numbers of synonymous (s_S) and nonsynonymous (s_N) sites per codon site for the entire phylogenetic tree. Selective neutrality is tested under the assumption that c_S and c_N are binomially distributed and that the probabilities of occurrence of synonymous and nonsynonymous substitutions are $s_S/(s_S + s_N)$ and $s_N/(s_S + s_N)$, respectively. If the neutrality is rejected and the relationship $c_N/s_N > c_S/s_S$ is observed, positive selection is inferred.

In the Yang method, ω is assumed to follow a certain probability distribution among codon sites in the sequence. Fourteen different distributions (M0–M13) were initially proposed to be used, but M3 and M8 have been recommended for the real data analysis and were actually used more often than others (Yang 2000; Yang et al. 2000). Model M3 is used in combination with M0. In M0, ω is assumed to be the same for all codon sites. In M3, codon sites are classified into three categories, 0, 1, and 2, for which ω is assumed to take values ω_0, ω_1, and ω_2, with proportions p_0, p_1, and p_2, respectively. In each model, free parameters are estimated by the maximum-likelihood (ML) method. If $\hat{\omega} > 1$ for any category in M3, we test whether M3 fits the data better than does M0 by the likelihood ratio test (LRT). Once the test is significant, we conclude that positively selected sites exist in the sequence. We then compute the posterior probability that a given codon site belongs to a category with $\hat{\omega} > 1$ in M3. If the probability is higher than a given confidence level, positive selection is inferred.

Model M8 is used in combination with M7. In M7, ω is assumed to follow a beta distribution with $0 \leq \omega \leq 1$. In M8, codon sites are classified into two categories, 0 and 1, which exist with proportions p_0 and p_1, respectively. In category 0, ω is assumed to follow a beta distribution with $0 \leq \omega \leq 1$, and in category 1, ω takes a given value of ω_1. Positively selected amino acid sites are inferred in the same way as above, with M0 and M3 replaced by M7 and M8, respectively.

Computer Simulation

A random nucleotide sequence with 300 codon sites was allowed to evolve following a symmetrical

Key words: positive selection, parsimony, binomial probability, likelihood, posterior probability.

Address for correspondence and reprints: Yoshiyuki Suzuki, Institute of Molecular Evolutionary Genetics, Department of Biology, The Pennsylvania State University, 328 Mueller Laboratory, University Park, Pennsylvania 16802. E-mail: yis1@psu.edu.

Mol. Biol. Evol. 19(11):1865–1869. 2002

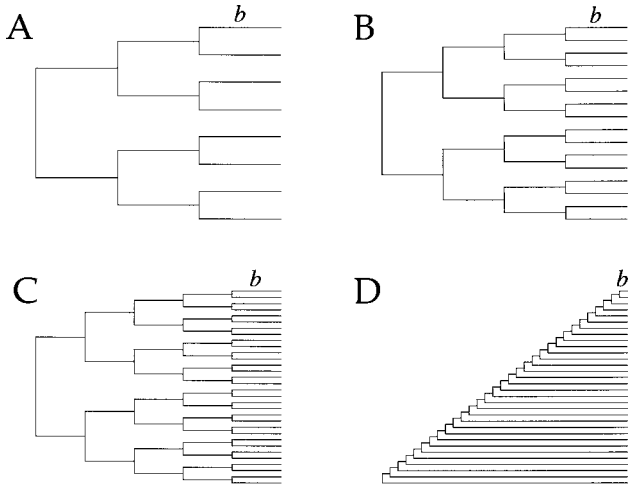

FIG. 1.—Phylogenetic trees used in the computer simulation. (*A*) Symmetrical phylogenetic tree with eight exterior nodes. (*B*) Symmetrical phylogenetic tree with 16 exterior nodes. (*C*) Symmetrical phylogenetic tree with 32 exterior nodes. (*D*) Asymmetrical phylogenetic tree with 32 exterior nodes. In (*A*), (*B*), and (*C*), all branch lengths (*b*) were set to be the same. In (*D*), the length of the shortest exterior branch (*b*) was set to be the same as those of all interior branches.

phylogenetic tree with 8 (fig. 1*A*), 16 (fig. 1*B*), or 32 (fig. 1*C*) exterior nodes. All branch lengths (*b*) were set such that the expected number of synonymous substitutions per synonymous site was 0.1. We assumed that $\omega = 0$ at half the codon sites and that $\omega = 1$ at the other half. The simulation scheme has been described in detail by Suzuki and Gojobori (1999).

The 8, 16, or 32 sequences generated above were analyzed by the SG and Yang methods. We used the computer program ADAPTSITE (Suzuki, Gojobori, and Nei 2001) for conducting the SG method and PAML (Yang 2000) for the Yang method. In both analyses we used the true (model) phylogenetic tree and 5% (95%) significance (confidence) level for inferring positive selection. The entire procedure was repeated 200 times for each parameter set such that the total number of codon sites examined was 60,000.

Results

The numbers of positively selected amino acid sites inferred by the SG and Yang methods in the simulation are presented in table 1. In the present study we mainly focus on the cases where nonselective sites ($\omega \leq 1$) are falsely identified as selective sites (false-positive results). In particular, we focus on the false-positive rate at the codon sites with $\omega = 1$ because we know the expected rate (type-I error) at these sites when a proper statistical method is used. The expected rate was 2.5% in the SG method because a two-tailed test with 5% significance level was conducted for testing the selective neutrality (the other 2.5% for negatively selected sites) and was 5% in the Yang method because the confidence level for inferring positive selection was 95%.

In the SG method positive selection was inferred at 1 (0.003%), 35 (0.1%), and 140 (0.5%) sites when 8, 16, and 32 sequences were used, respectively, which was smaller than expected (table 1). In the Yang method we first analyzed eight sequences without conducting the LRT. Positive selection was inferred at 13,086 (44%) and 4,174 (14%) sites when models M3 and M8 were used, respectively, and at 3,493 (12%) sites when both models were used for the same data set. These false-positive rates were higher than expected. By conducting the LRT, however, the rate decreased to 3% in M8 but not in M3. Therefore, the number of positively selected sites inferred by both M3 and M8 was 738 (2%), which was smaller than expected. When 16 sequences were analyzed, however, positive selection was inferred at 3,685 (12%) sites by both M3 and M8, even after conducting the LRT. Similarly, 3,296 (11%) sites were inferred to be positively selected when 32 sequences were analyzed.

In the above analysis an increase in the number of sequences was accompanied by an increase in the number of nucleotide substitutions at each codon site in the entire phylogenetic tree. To examine which of these two factors, i.e., the number of sequences or the number of nucleotide substitutions, was responsible for the observed increase in the false-positive rate in the Yang method, we analyzed 16 sequences (fig. 1*B*) generated by the same scheme as above, but with $b = 0.05$ and eight sequences (fig. 1*A*) with $b = 0.2$. Note that the number of nucleotide substitutions for the former data set is similar to that for eight sequences with $b = 0.1$, whereas the number for the latter data set is similar to that for 16 sequences with $b = 0.1$. As shown in table 1, only the case of eight sequences with $b = 0.2$ produced as many false positives as the case of 16 sequences with $b = 0.1$, indicating that the increase in the number of nucleotide substitutions was mainly responsible for the increase in the false-positive rate.

To examine whether the high false-positive rate observed above is specific to the particular fractions of codon sites with $\omega = 0$ (50%) and $\omega = 1$ (50%), we considered the case where $\omega = 0$ at 90% of codon sites and $\omega = 1$ at the remainder with 32 sequences (fig. 1*C*). In the Yang method the false-positive rate (11%) was higher than expected, whereas in the SG method the rate (0.5%) was lower (table 1).

We also considered the case where $\omega = 0.5$ or 1.5 for 50% of codon sites and $\omega = 1$ for the remaining 50% to examine whether the high false-positive rate observed at the codon sites with $\omega = 1$ is specific to the particular ω value (0) at other sites. Again, the false-positive rate was high (12% for $\omega = 0.5$ and 7% for $\omega = 1.5$) in the Yang method but was low (1% for both) in the SG method (table 1).

Finally, to examine whether the high false-positive rate is specific to the symmetrical phylogenetic tree, we generated 32 sequences following an asymmetrical phylogenetic tree (fig. 1*D*), with $\omega = 0$ at half the codon sites and $\omega = 1$ at the other half. The false-positive rate was still high (10%) in the Yang method and was low (1%) in the SG method (table 1).

Note that in all the above simulations the SG method never inferred positive selection at the codon sites

Table 1
Numbers of Positively Selected Amino Acid Sites Inferred by the SG and Yang Methods in the Computer Simulation

					LRT NOT CONDUCTED		LRT CONDUCTED	
SCHEME	TREE	*b*	METHOD	LRT+[a]	ω = 0, 0.5, or 1.5	ω = 1	ω = 0, 0.5 or 1.5	ω = 1
ω = 0/150	*A*	0.1	SG[b]				0 (0.00)	1 (0.00)
ω = 1/150[c] . . .			Yang (M3)	146 (0.73)	0 (0.00)[d]	13,086 (0.44)	0 (0.00)	13,086 (0.44)
			Yang (M8)	15 (0.08)	0 (0.00)	4,174 (0.14)	0 (0.00)	799 (0.03)
			Yang (M3 + M8)[e]	15 (0.08)	0 (0.00)	3,493 (0.12)	0 (0.00)	738 (0.02)
	B	0.1	SG				0 (0.00)	35 (0.00)
			Yang (M3)	143 (0.72)	5 (0.00)	15,655 (0.52)	5 (0.00)	15,655 (0.52)
			Yang (M8)	42 (0.21)	3 (0.00)	12,162 (0.41)	2 (0.00)	4,131 (0.14)
			Yang (M3 + M8)	40 (0.20)	3 (0.00)	10,217 (0.34)	2 (0.00)	3,685 (0.12)
	C	0.1	SG				0 (0.00)	140 (0.00)
			Yang (M3)	165 (0.83)	1 (0.00)	14,345 (0.48)	1 (0.00)	14,345 (0.48)
			Yang (M8)	70 (0.35)	3 (0.00)	8,484 (0.28)	2 (0.00)	4,659 (0.16)
			Yang (M3 + M8)	70 (0.35)	1 (0.00)	5,545 (0.18)	1 (0.00)	3,296 (0.11)
	B	0.05	SG				0 (0.00)	3 (0.00)
			Yang (M3)	143 (0.72)	2 (0.00)	13,265 (0.44)	2 (0.00)	13,265 (0.44)
			Yang (M8)	24 (0.12)	0 (0.00)	5,674 (0.19)	0 (0.00)	1,235 (0.04)
			Yang (M3 + M8)	24 (0.12)	0 (0.00)	5,071 (0.17)	0 (0.00)	1,171 (0.04)
	A	0.2	SG				0 (0.00)	4 (0.00)
			Yang (M3)	134 (0.67)	2 (0.00)	16,302 (0.54)	2 (0.00)	16,302 (0.54)
			Yang (M8)	39 (0.20)	1 (0.00)	13,857 (0.46)	0 (0.00)	4,543 (0.15)
			Yang (M3 + M8)	39 (0.20)	1 (0.00)	12,288 (0.41)	0 (0.00)	4,216 (0.14)
ω = 0/270	*C*	0.1	SG				0 (0.00)	27 (0.00)
ω = 1/30[f]			Yang (M3)	128 (0.64)	0 (0.00)	3,833 (0.64)	0 (0.00)	3,833 (0.64)
			Yang (M8)	21 (0.11)	2 (0.00)	3,240 (0.60)	0 (0.00)	630 (0.11)
			Yang (M3 + M8)	21 (0.11)	0 (0.00)	3,235 (0.60)	0 (0.00)	630 (0.11)
ω = 0.5/150 . .	*C*	0.1	SG				12 (0.00)	179 (0.01)
ω = 1/150[g] . . .			Yang (M3)	157 (0.79)	94 (0.00)	5,234 (0.17)	94 (0.00)	5,234 (0.17)
			Yang (M8)	139 (0.70)	55 (0.00)	4,699 (0.16)	55 (0.00)	4,699 (0.16)
			Yang (M3 + M8)	133 (0.67)	46 (0.00)	3,720 (0.12)	46 (0.00)	3,720 (0.12)
ω = 1.5/150 . .	*C*	0.1	SG				627 (0.02)	164 (0.01)
ω = 1/150[h] . . .			Yang (M3)	199 (1.00)	18,138 (0.60)	12,634 (0.42)	17,988 (0.60)	12,484 (0.42)
			Yang (M8)	200 (1.00)	8,527 (0.28)	4,594 (0.15)	8,547 (0.28)	4,594 (0.15)
			Yang (M3 + M8)	199 (1.00)	6,058 (0.20)	1,981 (0.07)	6,030 (0.20)	1,981 (0.07)
ω = 0/150	*D*	0.01	SG				0 (0.00)	168 (0.01)
ω = 1/150			Yang (M3)	157 (0.79)	0 (0.00)	11,300 (0.38)	0 (0.00)	11,300 (0.38)
			Yang (M8)	57 (0.29)	2 (0.00)	12,796 (0.43)	0 (0.00)	4,648 (0.15)
			Yang (M3 + M8)	57 (0.29)	0 (0.00)	7,635 (0.25)	0 (0.00)	2,895 (0.10)

[a] The number of replications in which the existence of positively selected sites was supported by the LRT in the Yang method. The proportion of the number out of a total of 200 replications is indicated in parentheses.

[b] The LRT is not relevant to the SG method.

[c] Assumed value of ω = 0 at 150 sites and ω = 1 at 150 sites.

[d] The proportion of positively selected sites.

[e] Positively selected sites inferred by both models M3 and M8 in the Yang method.

[f] Assumed value of ω = 0 at 270 sites and ω = 1 at 30 sites.

[g] Assumed value of ω = 0.5 at 150 sites and ω = 1 at 150 sites.

[h] Assumed value of ω = 1.5 at 150 sites and ω = 1 at 150 sites in a sequence.

with ω = 0, whereas the Yang method occasionally did (table 1).

Discussion

Because the actual pattern of evolutionary change of nucleotide sequences is quite complicated, it is desirable that the method for detecting positive selection at single amino acid sites is not very sensitive to violation of the assumptions made. Moreover, if the method is not unbiased, a conservative method is more preferable to a liberal method because we are mainly interested in whether positive selection operates or not, and the conclusion is usually drawn only from the statistical analysis without experiment (Nei and Kumar 2000, pp. 69–71). In the present study the SG method appeared to be conservative because the false-positive rate was lower than expected. These observations were probably due to the fact that the number of sequences was relatively small and the branch length in the phylogenetic tree was relatively large in the present study. Under such conditions, the SG method may underestimate the number of nucleotide substitutions because the MP method is used for the estimation, and it may become difficult to reject the selective neutrality (Suzuki and Gojobori 1999).

The Yang method, in contrast, appears to be liberal under the same conditions. When we did not conduct the LRT, the false-positive rate at the codon sites with ω = 1 was quite high in all cases examined. This probably occurred because the posterior probability is computed under the assumption that the estimate ($\hat{\omega}$) of ω is correct, although it is subject to the sampling error. The LRT appears to help reduce the cases where $\hat{\omega}$ happens to be greater than 1 by chance, but this test is

Table 2
Proportions of Negatively and Positively Selected Amino Acid Sites Inferred by the SG Method in the Computer Simulation[a]

Sequences	*b*	Selection	ω = 0.2	ω = 0.5	ω = 1	ω = 2	ω = 5
64	0.01	Negative	0.00 (0.08)	0.00 (0.05)	0.00 (0.02)	0.00 (0.00)	0.00 (0.00)
		Positive	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.01)	0.04 (0.21)
	0.02	Negative	0.01 (0.22)	0.00 (0.09)	0.00 (0.02)	0.00 (0.00)	0.00 (0.00)
		Positive	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.05)	0.15 (0.47)
	0.03	Negative	0.02 (0.33)	0.00 (0.12)	0.00 (0.02)	0.00 (0.00)	0.00 (0.00)
		Positive	0.00 (0.00)	0.00 (0.00)	0.00 (0.01)	0.01 (0.08)	0.22 (0.59)
128	0.01	Negative	0.01 (0.23)	0.00 (0.10)	0.00 (0.02)	0.00 (0.00)	0.00 (0.00)
		Positive	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.01 (0.06)	0.23 (0.57)
	0.02	Negative	0.06 (0.43)	0.01 (0.15)	0.00 (0.02)	0.00 (0.00)	0.00 (0.00)
		Positive	0.00 (0.00)	0.00 (0.00)	0.00 (0.01)	0.03 (0.15)	0.54 (0.84)
	0.03	Negative	0.14 (0.56)	0.02 (0.20)	0.00 (0.02)	0.00 (0.00)	0.00 (0.00)
		Positive	0.00 (0.00)	0.00 (0.00)	0.00 (0.01)	0.04 (0.21)	0.70 (0.92)

[a] The proportions of negatively and positively selected sites inferred by using the binomial probability are indicated inside the parentheses, whereas those inferred by using Fisher's exact test are indicated outside the parentheses. The results for the binomial probability are the same as those in table 3 of Suzuki and Gojobori (1999).

known to be sensitive to violation of the assumptions (Zhang 1999), as shown above. Anisimova, Bielawski, and Yang (2001) conducted a simulation using six sequences generated with $\omega = 0$ at half the codon sites and $\omega = 1$ at the other half and found that model M3 produced many false positives even after conducting the LRT but M8 did not do so (see also Anisimova, Bielawski, and Yang 2002). According to these results, they concluded that the Yang method is robust as long as multiple models are used. These observations are similar to our simulation results using eight sequences with $b = 0.1$. In our study with a larger number of sequences, however, both M3 and M8 produced many false positives regardless of the fractions of codon sites with $\omega = 0$ and $\omega = 1$, the ω values at the codon sites where $\omega \neq 1$, and the phylogenetic trees. In addition, an increase in the number of nucleotide substitutions appears to be mainly responsible for the increase in the false-positive rate. Because the number of nucleotide substitutions is equivalent to the sample size in the statistical test of positive selection, the Yang method appears to produce many false positives consistently under the above conditions. We have also shown that even amino acid sites with $\omega = 0$ may be identified as positively selected sites.

The Yang method appears to have another problem when the number of sequences used is large. When Suzuki and Nei (2001) analyzed 218 sequences of the HLA gene by the Yang method, they often failed to find the ML estimates of free parameters because of the existence of multiple local optima on the likelihood surface. Su, Nguyen, and Nei (2002) also obtained different results depending on the initial ω values when they analyzed 70 dimeric immunoglobulin (Ig) variable region (V_HH) genes and 96 conventional Ig heavy chain (V_H) genes. Note that even when putative ML estimates were found, in both studies, the overall reliability of the Yang method appeared to be similar to or lower than that of the SG method. In the present study we also compared the overall efficiency of detecting positively selected sites between the SG and Yang methods, using the case with $\omega = 1.5$ at half the codon sites and $\omega = 1$ at the other half. When we computed the odds ratio which is defined by $n_{1.5}(30{,}000 - n_1)/(30{,}000 - n_{1.5})n_1$, where $n_{1.5}$ is the number of positively selected sites inferred at the codon sites with $\omega = 1.5$, and n_1 is the number at the codon sites with $\omega = 1$; the ratios were 3.88 for the SG method and 3.56 for the Yang method, suggesting that the former method had a slightly higher efficiency than did the latter method. In conclusion, both our simulation studies and real data analyses suggest that the positively selected sites inferred by the parsimony-based method are more reliable than those inferred by the likelihood-based method.

Several other methods similar to the SG method have been proposed for detecting positive selection at single amino acid sites. Fitch et al. (1997) proposed to use $C_S/(C_S + C_N)$ and $C_N/(C_S + C_N)$ as the probabilities of occurrence of synonymous and nonsynonymous substitutions at every codon site, respectively, for computing the binomial probability, where C_S and C_N are the sums of c_S and c_N over all codon sites in the sequence, respectively. This method effectively assumes that the average of ω over all codon sites represents the neutrality level. It is, therefore, possible that the results include the false positives or negatives according to the sequence analyzed. For example, if $\omega = 0.1$ at a certain fraction of codon sites and $\omega = 0.2$ at others in the sequence, the latter sites may be inferred as positively selected, although they are actually negatively selected, because the neutrality level is assumed to be somewhere between $\omega = 0.1$ and $\omega = 0.2$.

Yamaguchi-Kabata and Gojobori (2000) proposed to use Fisher's exact test instead of the binomial probability for testing the selective neutrality in the SG method. They computed the expected numbers of synonymous ($E(c_S)$) and nonsynonymous ($E(c_N)$) substitutions at a given codon site approximately by $E(c_S) = (c_S + c_N)s_S/(s_S + s_N)$ and $E(c_N) = (c_S + c_N)s_N/(s_S + s_N)$, respectively. They then conducted Fisher's exact test for the 2×2 contingency table, which consisted of c_S and c_N in the first and second columns in the first row, respectively, and $E(c_S)$ and $E(c_N)$ in the first and second

columns in the second row, respectively. Strictly speaking, however, Fisher's exact test was originally designed for the cases where the marginal totals (in the present example, $c_S + c_N$, $E(c_S) + E(c_N)$, $c_S + E(c_S)$, and $c_N + E(c_N)$) in the table were fixed, which is not applicable to the present case (Sokal and Rohlf 1995, pp. 724–743). To examine the reliability of this method, we analyzed 64 or 128 sequences which were generated following symmetrical phylogenetic trees under the assumption that all codon sites in a sequence (300 codon sites) had the same ω value, i.e., ω = 0.2, 0.5, 1, 2, or 5, and b = 0.01, 0.02, or 0.03 (Suzuki and Gojobori 1999). (Note that it is not very meaningful to include the Yang method in this analysis because the efficiency of detecting positively selected sites in the Yang method depends on the distribution of ω values over all codon sites [Yang et al. 2000], whereas positive selection is detected at each codon site independently in the SG method.) Similar to the case where binomial probability was used, the false-positive rate of the SG method using Fisher's exact test was low under various conditions (table 2). But the probability of correctly identifying the positively and negatively selected sites (true-positive rate) was significantly higher in all cases examined when the binomial probability was used than when Fisher's exact test was used. It, therefore, appears to be better to use the binomial probability than Fisher's exact test in the SG method.

Note Added in Proof

After this paper was accepted for publication, we learned that the PAML version 3.0a used in our study may produce less accurate results than the latest version, 3.12, when the number of sequences used is very large. We therefore repeated our computer simulation using this new version for the case of 32 sequences that were generated according to the model tree in Fig. 1*C*. The computational speed of version 3.12 is very slow, and we examined only the first 50 replications of this case with $\omega = 0$ for the first half of the codon sites and $\omega = 1$ for the remaining codon sites. However, the results obtained were essentially the same as those presented in Table 1. That is, positive selection was inferred at 4116 (55%) and 1051 (14%) sites, respectively, when models M3 and M8 were used with the LRT and at 901 (12%) sites when both models were used. Therefore, our conclusion appears unaffected by the use of the latest version of PAML.

Acknowledgments

The authors thank Jianzhi Zhang and two anonymous reviewers for their valuable comments. This work was supported by a grant from the National Institutes of Health (GM20293) to M.N. Y.S. was supported by the JSPS Research Fellowships for Young Scientists.

LITERATURE CITED

ANISIMOVA, M., J. P. BIELAWSKI, and Z. YANG. 2001. Accuracy and power of the likelihood ratio test in detecting adaptive molecular evolution. Mol. Biol. Evol. **18**:1585–1592.

———. 2002. Accuracy and power of Bayes prediction of amino acid sites under positive selection. Mol. Biol. Evol. **19**:950–958.

FITCH, W. M. 1971. Toward defining the course of evolution: minimum change for a specific tree topology. Syst. Zool. **20**:406–416.

FITCH, W. M., R. M. BUSH, C. A. BENDER, and N. J. COX. 1997. Long-term trends in the evolution of H(3) HA1 human influenza type A. Proc. Natl. Acad. Sci. USA **94**:7712–7718.

HARTIGAN, J. A. 1973. Minimum mutation fits to a given tree. Biometrics **29**:53–65.

HUGHES, A. L., and M. NEI. 1988. Pattern of nucleotide substitution at major histocompatibility complex class I loci reveals overdominant selection. Nature **335**:167–170.

———. 1989. Nucleotide substitution at major histocompatibility complex class II loci: evidence for overdominant selection. Proc. Natl. Acad. Sci. USA **86**:958–962.

NEI, M., and S. KUMAR. 2000. Molecular evolution and phylogenetics. Oxford University Press, Oxford, N.Y.

SOKAL, R. R., and F. J. ROHLF. 1995. Biometry. 3rd edition. Freeman, New York.

SU, C., V. K. NGUYEN, and M. NEI. 2002. Adaptive evolution of variable region genes encoding an unusual type of immunoglobulins in camelids. Mol. Biol. Evol. **19**:205–215.

SUZUKI, Y., and T. GOJOBORI. 1999. A method for detecting positive selection at single amino acid sites. Mol. Biol. Evol. **16**:1315–1328.

———. 2001. Positively selected amino acid sites in the entire coding region of hepatitis C virus subtype 1b. Gene **276**: 83–87.

SUZUKI, Y., T. GOJOBORI, and M. NEI. 2001. ADAPTSITE: detecting natural selection at single amino acid sites. Bioinformatics **17**:660–661.

SUZUKI, Y., and M. NEI. 2001. Reliabilities of parsimony-based and likelihood-based methods for detecting positive selection at single amino acid sites. Mol. Biol. Evol. **18**:2179–2185.

YAMAGUCHI-KABATA, Y., and T. GOJOBORI. 2000. Reevaluation of amino acid variability of the human immunodeficiency virus type 1 gp120 envelope glycoprotein and prediction of new discontinuous epitopes. J. Virol. **74**:4335–4350.

YANG, Z. 2000. PAML (phylogenetic analysis by maximum likelihood). Version 3.0. University College London, London.

YANG, Z., R. NIELSEN, N. GOLDMAN, and A.-M. K. PEDERSEN. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics **155**:431–449.

ZHANG, J. 1999. Performance of likelihood ratio tests of evolutionary hypotheses under inadequate substitution models. Mol. Biol. Evol. **16**:868–875.

NARUYA SAITOU, reviewing editor

Accepted June 7, 2002