J Mol Evol (1995) 40:56–63

JOURNAL OF MOLECULAR EVOLUTION

Synonymous and Nonsynonymous Substitutions in Mammalian Genes and the Nearly Neutral Theory

Tomoko Ohta

National Institute of Genetics, Mishima 411, Japan

Received: 20 May 1994 / Accepted: 22 August 1994

Abstract. The nearly neutral theory of molecular evolution predicts larger generation-time effects for synonymous than for nonsynonymous substitutions. This prediction is tested using the sequences of 49 single-copy genes by calculating the average and variance of synonymous and nonsynonymous substitutions in mammalian star phylogenies (rodentia, artiodactyla, and primates). The average pattern of the 49 genes supports the prediction of the nearly neutral theory, with some notable exceptions.

The nearly neutral theory also predicts that the variance of the evolutionary rate is larger than the value predicted by the completely neutral theory. This prediction is tested by examining the dispersion index (ratio of the variance to the mean), which is positively correlated with the average substitution number. After weighting by the lineage effects, this correlation almost disappears for nonsynonymous substitutions, but not quite so for synonymous substitutions. After weighting, the dispersion indices of both synonymous and nonsynonymous substitutions still exceed values expected under the simple Poisson process. The results indicate that both the systematic bias in evolutionary rate among the lineages and the episodic type of rate variation are contributing to the large variance. The former is more significant to synonymous substitutions than to nonsynonymous substitutions. Isochore evolution may be similar to synonymous substitutions. The rate and pattern found here are consistent with the nearly neutral theory, such that the relative contributions of drift and selection differ between the two types of substitutions. The results are also consistent with Gillespie's episodic selection theory.

Key words: Nearly neutral theory — Synonymous substitution — Nonsynonymous substitution — Mammalian genes

Introduction

One way to measure the effects of natural selection in molecular evolution is to estimate separately the rates of synonymous and nonsynonymous substitutions. Weak selection can affect both categories of substitutions, but obviously amino-acid-altering substitutions are more exposed to selection than synonymous ones. I predicted two decades ago that the rate of DNA evolution reflects generation number more strongly than does the rate of protein evolution, considering that the latter is more influenced by selection (Ohta 1973, 1974). At that time, most proteins examined were thought to have fixed functions preserved by long-term stabilizing selection that removed deleterious mutations. Thus the slightly deleterious mutation theory predicted a negative correlation between amino acid substitutions and the population size of the species. The population-size effect is diminished by the generation-time effect because large animals generally have small population sizes and *vice versa.* (For a recent review, see Chao and Carr 1993.)

Recently (Ohta 1993a), I tested the prediction that generation-time effect should be greater for synonymous than nonsynonymous substitutions using 17 genes in a mammalian star phylogeny (primates, artiodactyla, and rodentia). The prediction was verified. I now report the result of analysis of 49 genes. In addition, the variance of

evolutionary rate was calculated for the synonymous and the nonsynonymous substitutions. The variation caused by the lineage effect is quite large for the synonymous substitutions, but far less for the nonsynonymous substitutions.

Sequence Analysis

Nucleotide sequences were obtained from the genetic data bases maintained at the National Institute of Genetics (Japan), which include GenBank, DNA Data Bank of Japan, and EMBL. For acquisition and analysis of the data, I used the ODEN package of Ina (1992). The sequences used are listed in Table 1. They were chosen for large coding region to minimize functional change overtime. The last condition is important since we are interested in negative selection.

The method of Ina (1994) was used for estimating the numbers of synonymous and nonsynonymous substitutions. This method is based on Kimura's two-parameter model and is an improvement of that of Nei and Gojobori (1986). For two sequences compared, the nucleotide differences are divided into the synonymous and the nonsynonymous differences, and the numbers of synonymous sites and nonsynonymous sites are estimated. Then the numbers of synonymous and nonsynonymous substitutions per site are calculated by taking multiple-hits into account. Ina's large-scale simulation study indicates that this method gives the most satisfying estimates among the methods so far available for nuclear genes of mammals (Ina 1994). The new method of Li (1993) has also been used and gives similar estimates to those of Ina's method. Li's method is inapplicable in some cases, and the result is not included here.

Mean Numbers of Synonymous and Nonsynonymous Substitutions

A total of 49 gene loci have been analyzed. Figure 1 represents the result: the star phylogeny of the three orders with the numbers of synonymous and nonsynonymous substitutions per site beside each branch. The total number of sites compared is shown under the tree. The figures indicate that the generation-time effect is more conspicuous for synonymous substitutions than for nonsynonymous substitutions; note that rodentia have larger population-size and shorter generation-time than primates. The present analysis with a larger set of data confirms the previous conclusion.

It should be noted here that the large number of substitutions in the rodent lineage may be caused by its ancient branching. Indeed, Easteal (1990) and Easteal and Collet (1994) argued by comparing several gene sequences between marsupial and rodent or primate, that there is no generation-time effect for synonymous substitutions. However, the number of synonymous substitutions per site for such a distant species pair often exceeds one, and the estimated number is not very reliable. In addition, the genes examined by Easteal and Collet (1994) included members of multigene families: α_1-antitrypsin, α-globin, α-lactoalbumin, and β-casein, and some complications such as gene conversion and functional differentiation might have affected the substitution pattern (Ohta 1991). In fact, three (α_1-antitrypsin, α-globin, and β-casein) of the above four genes show larger divergence in primate than in rodent lineage (Table 1 of Easteal and Collet 1994). (For the analyses that show interaction of segmental gene conversion and selection on α_1-antitrypsin gene family in mammals, see Ohta 1994b.) If these three genes are omitted from the

Table 1. Sequences used in this study[a]

	Accession number		
	Primates	Artiodactyla	Rodentia
Acid phosphatase type 5	J04430	M98553	M76110
Alkaline phosphatase intestin	M15694	L07733	M61705
Alkaline phosphatase liver	X14174	M18443	Y00714
Aspartate aminotransferase cytosolic	M37400	X66020	J02623
Aspartate aminotransferase mitochondrial	M22632	Z25466	J02622
ATP synthase α	X59066	M22465	L01062
ATP synthase β	M27132	X05605	M19044
β-1, 4-galactosyl transferase	M22921	X14558	D00314
Carboxypeptidase	X51405	X04411	X61232
Connexin	M65188	J05535	M61896
Corticotropin releasing factor	V00571	M22853	X03036
Dopamine receptor D2	X51362	X51657	X55674
Glucose transporter	K03195	M60448	M22998
Hexokinase 1	M75126	M65140	J05277
Lactoferrin	X52941	X57084	J03298
Luteinizing hormone receptor	M63108	M29525	M81310
Myelin proteolipid protein	M54927	X03098	M15442
Neuroleukin	K03515	X07382	M14220
Neurophysin I	M11186	X00502	M88355
Neurophysin II	M11166	M25645	M88354
Opsin	K02281	M21606	X61084
Ornithine decarboxylase	X55362	M92441	M10624
Plasminogen activator inhibitor	X04744	X16383	M33960
Proopiomelanocortin	V01510	J00019	J00611
Protein disulfide isomerase	X05130	M17596	X06453
Terminal transferase	M11722	X04123	X04122
Thrombomodulin	D00210	M14657	X14432
Transforming growth factor β1	X02812	M36271	M13177
Transforming growth factor β2	Y00083	L08375	X57413
Transforming growth factor β3	X14149	X14150	M32745
Transforming growth factor β3 receptor	L07594	L07595	M80784
Urokinase-plasminogen activator	X02419	L03546	M17922

[a] The 17 gene loci used in Ohta (1993a) are not given here, but are included in the present analysis

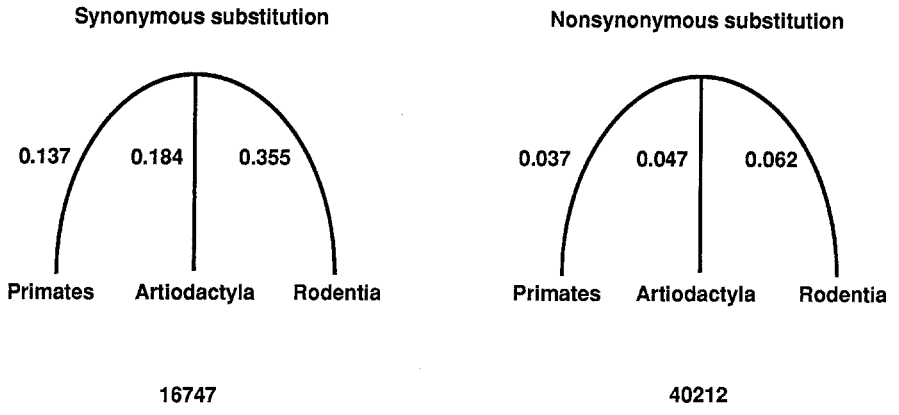

Fig. 1. Star phylogenies of 49 genes. Figures beside each branch are the estimated numbers of substitutions per site.

table, the tendency to rapid divergence of the rodent lineage can be observed. Of course, it is possible that the rodent line branched off before the primate-artiodactyl split (Li and Graur 1991). The difference in the branch lengths of Fig. 1 may be caused by both generation-time effect and ancient branching.

Let us examine the lineage effect of synonymous and nonsynonymous substitutions estimated by a weighting factor as in Gillespie (1991). The factor is the characteristic divergence of each lineage and the average is unity. The weight factors for synonymous and nonsynonymous substitutions of the 49 genes are given in Table 2. The values are slightly different from the previous estimates, but show a similar pattern.

Table 2. Weight factors for the synonymous and nonsynonymous substitutions obtained by the present data set

i	Primates	Artiodactyla	Rodentia
Synonymous $w_{s,i}$	0.607	0.817	1.576
Nonsynonymous $w_{n,i}$	0.751	0.965	1.284

Substitution Patterns at Individual Genes

Table 3 gives the estimated numbers of synonymous and nonsynonymous substitutions per site of the star phylogeny for each gene. The pattern similar to the average in Fig. 1 may be found in about 80% of genes. For the remaining 20% of genes, the lineage effect on synonymous and nonsynonymous branch lengths differs from the average pattern, but the difference is statistically insignificant in most cases. The following three cases compel our interest.

1. Each of three transforming growth factors (TGFs) shows peculiar patterns for nonsynonymous substitution. TGFβ1 and TGFβ2 are characterized by exceptionally rapid divergence in rodent lineage, whereas TGFβ3 shows rapid divergence in artiodactyl lineage. The three TGF genes are homologous and originated by duplication before mammalian divergence, and the peculiar pattern may be related to a system involving all TGFs.
2. Growth hormone shows unusually high divergence in the primate lineage. This may be caused by gene duplication in the primate lineage, since the human genome contains several growth-hormone-like genes in tandem duplication (Chen et al. 1989). They are expressed in the placenta and show a high amino-acid substitution rate (Ohta 1993b). The growth hormone gene retains its original function and is included in the present analysis. However, it has been argued that binding specificities for receptors of growth hormones from different species markedly differ and that most amino acid changes in rapidly evolving phase are caused by positive selection (Wallis 1994). If so, the inclusion of growth hormone in the present analysis may not be appropriate.
3. For interleukins 6 and 7, nonsynonymous substitution rate is higher than the synonymous substitution rate in the primate lineage. The excess of nonsynonymous substitution was noted by Wolfe and Sharp (1993) for the mouse vs rat comparison of interleukin 3 gene among the 363 genes examined. Although the excess of nonsynonymous substitution over synonymous substitution is statistically insignificant, interleukins may be subject to some positive selection.

Variances of Numbers of Synonymous and Nonsynonymous Substitutions

How variable is the evolutionary rate in general? It is now accepted that the variance of the evolutionary rate is often larger than the value expected under the simple Poisson process, but the magnitude and pattern of variation are still controversial. By using the star phylogenies of 49 gene loci, the variance of the substitution numbers has been estimated. The dispersion index, i.e., the ratio of the variance to the mean, is a useful measure. Gillespie

Table 3. Estimated numbers of synonymous and nonsynonymous substitutions per site of individual genes

Gene	Synonymous substitution				Nonsynonymous substitution			
	Primates	Artiodactyla	Rodentia	No. sites compared	Primates	Artiodactyla	Rodentia	No. sites compared
Acetylcholine receptor a	0.116	0.116	0.296	402	0.014	0.007	0.019	969
Acetylcholine receptor β	0.159	0.152	0.323	450	0.024	0.022	0.039	1050
Acid phosphatase type 5	0.270	0.112	0.420	304	0.031	0.058	0.030	665
Albumin	0.241	0.146	0.568	526	0.057	0.096	0.118	1294
Alkaline phosphatase intestin	0.212	0.233	0.344	431	0.047	0.097	0.091	1056
Alkaline phosphatase liver	0.137	0.175	0.309	459	0.031	0.030	0.046	1113
Aspartate aminotransferase cytosolic	0.133	0.195	0.328	367	0.017	0.027	0.030	873
Aspartate aminotransferase mitochondrial	0.143	0.179	0.289	388	0.014	0.023	0.011	902
ATP synthase a	0.111	0.167	0.364	525	0.008	0.005	0.008	1104
ATP synthase β	0.081	0.148	0.313	331	0.000	0.005	0.006	740
β-1, 4-galactosyl transferase	0.091	0.189	0.273	355	0.026	0.068	0.053	829
Carboxypeptidase	0.115	0.196	0.412	369	0.004	0.002	0.002	924
Connexin	0.188	0.100	0.351	339	0.006	0.005	0.008	807
Corticotropin-releasing factor	0.118	0.271	0.326	161	0.021	0.117	0.088	388
Dopamine receptor D2	0.116	0.107	0.209	397	0.010	0.008	0.011	933
Fibrinogen g	0.080	0.191	0.531	374	0.033	0.074	0.076	923
Glucose transporter	0.079	0.180	0.328	461	0.010	0.007	0.009	1015
Growth hormone	0.313	0.179	0.214	168	0.186	0.036	0.039	394
Growth hormone receptor	0.057	0.150	0.299	540	0.059	0.029	0.132	1357
Hexokinas I	0.115	0.193	0.402	783	0.021	0.031	0.022	1968
IGF binding protein 1	0.270	0.313	0.521	223	0.105	0.092	0.094	551
IGF binding protein 3	0.066	0.395	0.332	249	0.051	0.060	0.062	610
Insulin-like growth factor 1	0.067	0.200	0.386	101	0.005	0.008	0.023	241
Insulin-like growth factor 2	0.098	0.236	0.228	130	0.034	0.054	0.048	318
Interleukin 1a	0.132	0.179	0.339	215	0.084	0.069	0.158	556
Interleukin 1β	0.128	0.282	0.287	210	0.088	0.167	0.114	570
Interleukin 2	0.042	0.166	0.595	129	0.040	0.201	0.225	325
Interleukin 6	0.138	0.220	0.519	174	0.182	0.167	0.357	446
Interleukin 7	0.058	0.108	0.282	118	0.092	0.064	0.089	357
Lactate dehydrogenase A	0.113	0.127	0.512	287	0.020	0.017	0.017	709
Lactoferrin	0.146	0.303	0.384	580	0.074	0.137	0.130	1406
Luteinizing hormone receptor	0.118	0.124	0.344	638	0.041	0.031	0.047	1420
Myelin proteolipid protein	0.029	0.071	0.099	134	0.010	0.010	0.000	310
Neuroleukin	0.167	0.125	0.277	473	0.019	0.020	0.049	1201
Neurophysin I	0.072	0.088	0.291	143	0.032	0.033	0.072	343
Neurophysin II	0.092	0.221	0.242	97	0.044	0.084	0.015	254
Opsin	0.135	0.178	0.346	302	0.014	0.027	0.022	727
Ornithine decarboxylase	0.216	0.184	0.278	398	0.017	0.017	0.039	985
Plasminogen activator inhibitor	0.181	0.192	0.417	327	0.035	0.044	0.090	831
Prolactin	0.184	0.261	0.398	166	0.056	0.115	0.230	426
Proopiomelanocortin	0.091	0.204	0.251	179	0.021	0.022	0.075	454
Protein disulfide isomerase	0.179	0.248	0.370	454	0.018	0.014	0.029	1064
Terminal transferase	0.141	0.081	0.411	447	0.044	0.033	0.087	1073

Table 3. Continued

Gene	Synonymous substitution				Nonsynonymous substitution			
	Primates	Artiodactyla	Rodentia	No. sites compared	Primates	Artiodactyla	Rodentia	No. sites compared
Thrombomodulin	0.169	0.237	0.570	300	0.112	0.125	0.125	724
Transforming growth factor β1	0.157	0.188	0.362	280	0.015	0.019	0.061	665
Transforming growth factor β2	0.098	0.132	0.324	385	0.004	0.002	0.020	854
Transforming growth factor β3	0.085	0.156	0.230	383	0.003	0.046	0.010	844
Transforming growth factor β3 receptor	0.138	0.311	0.344	746	0.038	0.054	0.063	1786
Urokinase-plasminogen activator	0.143	0.162	0.277	348	0.086	0.071	0.129	861

(1991) noted that the pattern of protein evolution and that of synonymous substitution are different. The silent substitution shows the lineage effect, and the dispersion index becomes small after removing the lineage effect by weighting. However, the nonsynonymous substitution pattern is more erratic. In the following, the dispersion index is examined in detail.

If the difference in substitution number among the three branches of the star phylogeny is mainly caused by the lineage effect, the dispersion index increases linearly with the mean substitution number per branch. Consider an extreme case where the branch length of the i-th lineage increases linearly with time, t, and let k_i be the lineage effect of the i-th lineage. Then the three branch lengths of the star phylogeny become k_1, k_2t, and k_3t. The mean is $\bar{k}t$, and the variance of the branch lengths is σ_k^2, where $\bar{k}$ and σ_k^2, are the mean and variance of k_i. Then the dispersion index is a linear function of t, $(\sigma_k^2/\bar{k})t$. On the other hand, if the variance of the branch lengths is mainly caused by the episodic clock (Gillespie 1987, 1991), the dispersion index should be independent of the mean substitution number. The relationship between the mean substitution number and the dispersion index is therefore important. The index has been calculated for the synonymous and for the nonsynonymous star phylogenies, both for the weighted and the unweighted branch lengths. In order to remove the lineage effect, the synonymous-weight factor is used for the synonymous values, and the nonsynonymous-weight factor is used for the nonsynonymous ones; $k_{s,i,w} = k_{s,i}/w_{s,i}$ and $k_{n,i,w} = k_{n,i}/w_{n,i}$, where $k_{s,i}$ and $k_{n,i}$ are the i-th branch lengths for synonymous and nonsynonymous substitutions, and another subscript, w, denotes the weighting.

Figure 2A–D represents the results. The abscissa (x) is the mean substitution and the ordinate (y) is the dispersion index. For unweighted cases, the dispersion index increases as the mean number becomes larger. The correlation is significant for both synonymous and nonsynonymous substitutions, but it is higher for the former. When the lineage effect is removed by weighting, the correlation becomes smaller than the unweighted case, but it is still significant for the synonymous substitutions. The correlation almost disappears by weighting for the nonsynonymous substitutions. Bulmer (1989) found that the process of correcting for multiple hits increases the variance particularly when the distance is large. Therefore the significance of the remaining correlation of the weighted synonymous substitution may be spurious.

The mean dispersion index is: nonsynonymous, 8.46; weighted nonsynonymous, 5.60; synonymous, 25.01; weighted synonymous, 5.89. The unweighted values are considerably larger than the previous estimates (Kimura 1983; Gillespie 1991), but become smaller by weighting. It should be noted that these values may be overestimates because the estimated numbers of substitutions are used for calculation. According to Bulmer (1989), the bias is considerable when the branch length per site is 0.25 or more. Thus, the values of the synonymous substitution may be inflated. For nonsynonymous values, the bias is considered to be small. The value of the intercept at $x = 0$ would represent the magnitude of variation caused by the episodic type of substitution process. Unfortunately it is not statistically larger than unity in any case. Again the growth hormone gene shows a very deviant pattern. By weighting, the nonsynonymous dispersion index of this gene becomes larger and takes the largest value among the 49 genes. As pointed out before, the inclusion of this gene may not be appropriate.

In addition to the dispersion index, the coefficient of variation has been calculated. The coefficient of variation should stay constant if the variation is mostly caused by the lineage effect, but it decreases by increasing the branch length if it is due to the episodic process. Figure 3A–D represents the result. There is almost no correlation for the synonymous substitutions, but the coefficient of variation decreases as the branch length increases for the nonsynonymous substitutions. Thus, it is likely that nonsynonymous substitutions are episodic. In other words, protein evolution is characterized with bursts of substitutions separated by relatively quiescent periods.

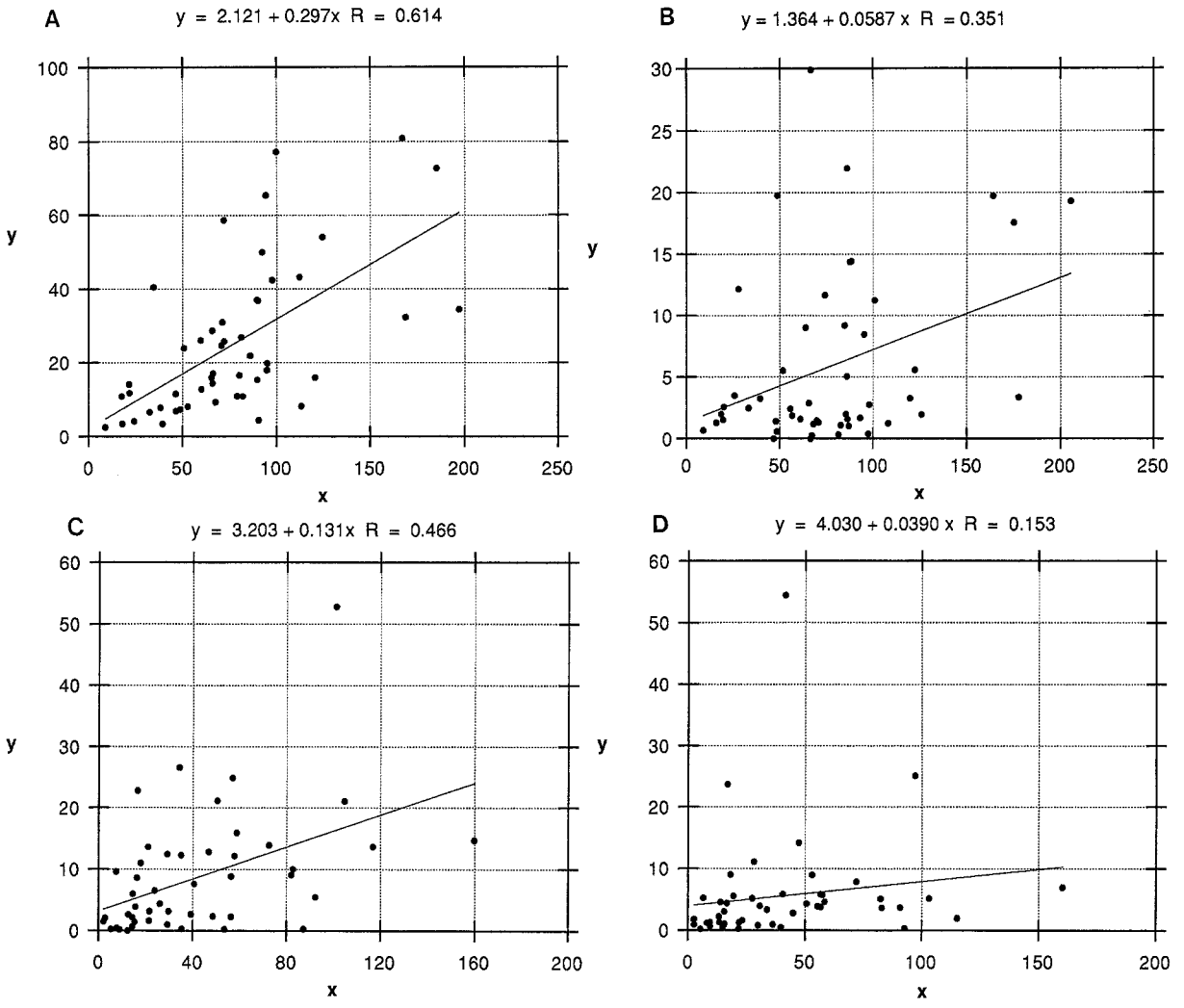

Fig. 2. The relationship between the mean substitution number per branch (x) and the dispersion index (y) for **(A)** the synonymous substitution, **(B)** the weighted synonymous substitution; **(C)** the nonsynonymous substitution; **(D)** the weighted nonsynonymous substitution.

Under the selection model of Gillespie (1991), the episodic process is caused by environmental changes—i.e., when environment changes, more than one amino acid replacement is needed to modify the function. Under the nearly neutral model (Ohta 1992), it is caused by the interplay between fluctuation of population size and weak selection—i.e., while the population size is small, slightly deleterious amino acid substitutions, that very slightly disturb the protein structure and function, may occur. The compensatory substitutions coming from the higher-order structure of proteins follow, resulting in the episodic pattern of substitution. The two hypotheses are indistinguishable here.

Discussion

The patterns of synonymous and nonsynonymous substitutions of 49 gene loci confirm the previous conclusion that the number of nonsynonymous substitutions relative to synonymous is larger in the primate lineage than in the rodent lineage—i.e., this relative number is negatively correlated with the species population size. This result strongly indicates that negative selection has a major effect on amino acid substitutions and that both random genetic drift and selection are important—i.e., slightly deleterious mutation theory.

One has to note that this theory never predicts that the gene system constantly deteriorates. From the moment the theory was proposed, the following features were emphasized (Ohta 1973): (1) The species population size fluctuates and the effectiveness of selection accordingly varies. Once in a while, such as at the time of speciation, the population size becomes small and drift predominates, and in the other period, the population gets large and selection becomes effective. (2) Compensatory mutations may be common at the molecular level if one considers higher-order structures of proteins and nucleic acids. In terms of the above features, the genetic systems are thought to be fluctuating.

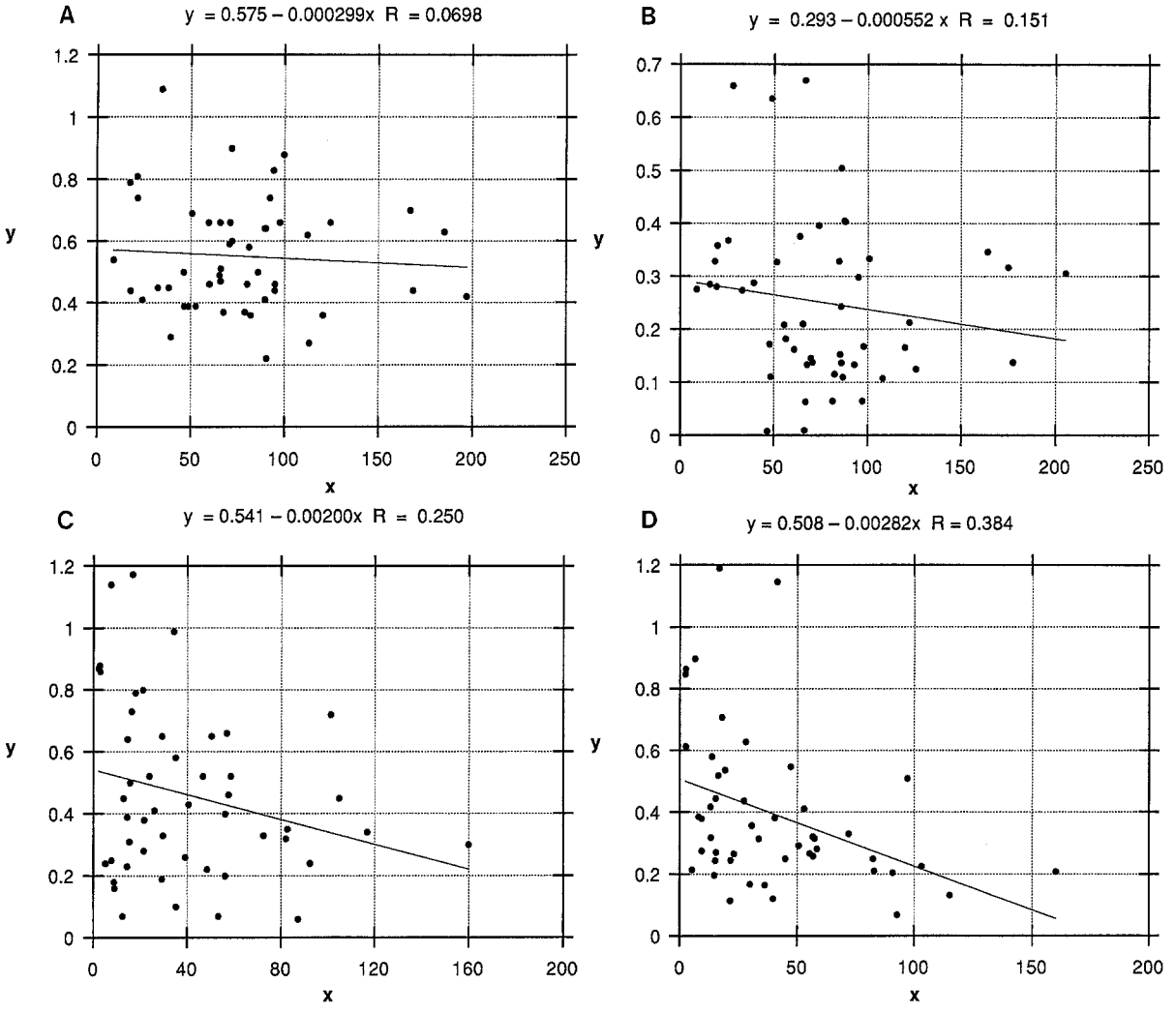

Fig. 3. The relationship between the mean substitution number per branch (x) and the coefficient of variation for **(A)** the synonymous substitution; **(B)** the weighted substitution; **(C)** the nonsynonymous substitution; **(D)** the weighted nonsynonymous substitution.

This aspect seems to be different from the selectionists', e.g., Gillespie's view (Gillespie 1991). He thinks that the species population size is large enough and therefore the gene system moves only through selection caused by environmental change. The difference seems to be an emphasis not on selection or drift, but on relative importance of the two processes.

From the dispersion index of synonymous substitutions, the lineage effect is thought to be the major factor of variation, in accord with the result of Gillespie (1991). The result implies that evolutionary force is systematic. Variation of nonsynonymous substitutions seems to be more complicated, and both the lineage and the episodic effects are present. For the moment it is difficult to clarify whether the lineage effect or the episodic process is more important for inflating the variance.

How is the present result connected to the isochore structure of mammalian chromosomes? The isochore structure would be beneficial to the organisms in relation to chromosome mechanics and gene expression. The issue of drift vs selection for the isochore evolution (Bernardi et al. 1985; Bernardi 1989; Holmquist 1992; Wolfe et al. 1989) cannot be settled unless both drift and selection are properly taken into account. The difference between the pattern of synonymous substitutions and that of nonsynonymous substitutions found is consistent with the nearly neutral theory, provided that relative contributions of drift vs selection differ between the two. Of course, selection pressure is greater in nonsynonymous substitutions than in synonymous substitutions. Also, selection acts differently between the two; for synonymous substitutions, selection would approach an optimum state with respect to codon usage and GC content (e.g., Kimura 1981), and for nonsynonymous substitutions, selection would optimize protein structure and function.

Positive selection may be detected when the patterns of synonymous and nonsynonymous substitutions deviate from the general pattern. Already there are several examples in which amino acid substitutions are accelerated in connection with functional differentiation (Ohta

1991). For example, the gene of fetal γ hemoglobin of higher primates shows the accelerated amino acid substitution following gene duplication (Fitch et al. 1991), and the stomach lysozyme of ruminants is also characterized by the elevated amino acid substitution via duplication (Irwin and Wilson 1990). Other examples include genes for visual pigment (Yokoyama and Yokoyama 1990), α_1-antitrypsin (Borriello and Krauter 1991; Rheaume et al. 1994), growth-hormone-like protein (Ohta 1993b), and troponin C fast and slow types (Parmacek et al. 1990; Ohta 1994a).

In any case, the substitution process is complex, and separation of drift and selection is impossible. Data analysis reveals various interesting patterns, particularly when the numbers of synonymous and nonsynonymous substitutions or the functional and nonfunctional substitutions can be separately estimated.

Acknowledgments. I thank Dr. William B. Provine, Dr. Fumio Tajima, and Dr. Yasuo Ina and the two anonymous referees for their many useful comments on the manuscript. This work was supported by a Grant-in-Aid from the Ministry of Education, Science and Culture of Japan. This is contribution No. 2005 from the National Institute of Genetics, Mishima, 411 Japan.

References

Bernardi G (1989) The isochore organization of the human genome. Ann Rev Genet 23:637–661

Bernardi G, Olofsson B, Filipski J, Zerial M, Salinas J, Cuny G, Meunier-Rotival M, Rodier F (1985) The mosaic genome of warm-blooded vertebrates. Science 228:953–958

Borriello F, Krauter KS (1991) Multiple murine α_1-protease inhibitor genes show unusual evolutionary divergence. Proc Natl Acad Sci USA 88:9417–9421

Bulmer M (1989) Estimating the variability of substitution rates. Genetics 123:615–619

Chao L, Carr DE (1993) The molecular clock and the relationship between population size and generation time. Evolution 47:688–690

Chen EY, Liao YC, Smith DH, Barrera-Saldana HA, Gelinas RE, Seeburg PH (1989) The human growth hormone locus: nucleotide sequence, biology and evolution. Genomics 4:479–497

Easteal S (1990) The pattern of mammalian evolution and the relative rate of molecular evolution. Genetics 124:165–173

Easteal S, Collet C (1994) Consistent variation in amino-acid substitution rate, despite uniformity of mutation rate: protein evolution in mammals is not neutral. Mol Biol Evol 11:643–647

Fitch DHA, Bailey WJ, Tagle DA, Goodman M, Sieu L, Slightom JL (1991) Duplication of the γ-globin gene mediated by repetitive L1 LINE sequences in an early ancestor of simian primates. Proc Natl Acad Sci USA 88:7396–7400

Gillespie JH (1987) Molecular evolution and the neutral allele theory. Oxf Surv Evol Biol 4:10–37

Gillespie JH (1991) *The causes of molecular evolution.* Oxford University Press, Oxford

Holmquist G (1992) Review article: chromosome bands, their chromatin flavors and their functional features. Am J Hum Genet 51:17–37

Ina Y (1992) ODEN. National Institute of Genetics, Mishima, Japan

Ina Y (1994) New methods for estimating the numbers of synonymous and nonsynonymous substitutions. J Mol Evol (in press)

Irwin DM, Wilson AC (1990) Concerted evolution of ruminant stomach lysozymes. J Biol Chem 265:4944–4952

Kimura M (1981) Possibility of extensive neutral evolution under stabilizing selection with special reference to nonrandom usage of synonymous codons. Proc Natl Acad Sci USA 78:5773–5777

Kimura M (1983) *The neutral theory of molecular evolution.* Cambridge University Press, Cambridge

Li W-H (1993) Unbiased estimation of the rates of synonymous and nonsynonymous substitution. J Mol Evol 36:96–99

Li W-H, Graur D (1991) *Fundamentals of molecular evolution.* Sinauer, Sunderland

Nei M, Gojobori T (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. Mol Biol Evol 3:418–426

Ohta T (1973) Slightly deleterious mutant substitutions in evolution. Nature 246:96–98

Ohta T (1974) Mutational pressure as the main cause of molecular evolution and polymorphisms. Nature 252:351–354

Ohta T (1991) Multigene families and the evolution of complexity. J Mol Evol 33:34–41

Ohta T (1992) The nearly neutral theory of molecular evolution. Ann Rev Syst Ecol 23:263–286

Ohta T (1993a) An examination of the generation-time effect on molecular evolution. Proc Natl Acad Sci USA 90:10676-10680

Ohta T (1993b) Pattern of nucleotide substitutions in growth hormone-prolactin gene family: a paradigm for evolution by gene duplication. Genetics 134:1271–1276

Ohta T (1994a) Further examples of evolution by gene duplication revealed through DNA sequence comparisons. Genetics (in press)

Ohta T (1994b) On hypervariability at the reactive centre of proteolytic enzymes and their inhibitors. J Mol Evol (in press)

Parmacek MS, Bengur AR, Vara AJ, Leiden JM (1990) The structure and regulation of expression of the murine fast skeletal troponin C gene. J Biol Chem 265:15970–15976

Rheaume C, Goodwin RL, Latimer JJ, Baumann H, Berger FG (1994) Evolution of murine α_1-proteinase inhibitors: gene amplification and reactive center divergence. J Mol Evol 38:121–131

Wallis M (1994) Variable evolutionary rates in the molecular evolution of mammalian growth hormone. J Mol Evol 38:619–627

Wolfe KH, Sharp PM (1993) Mammalian gene evolution: nucleotide sequence divergence between mouse and rat. J Mol Evol 37:441–456

Wolfe KH, Sharp PM, Li W-H (1989) Mutation rates differ among regions of the mammalian genome. Nature 337:283–285

Yokoyama S, Yokoyama R (1990) Molecular evolution of visual pigment genes and other G-protein-coupled genes. In: Takahata, N, Crow JF (eds) Population biology of genes and molecules. Baifukan, Tokyo