# ショウジョウバエにおいて推定されたホ乳類より高い塩基置換率


| | |
|---|---|
| 誌名 | The Japanese journal of genetics |
| ISSN | 0021504X |
| 著者 | 森山, 悦子 |
| 巻/号 | 62巻2号 |
| 掲載ページ | p. 139-147 |
| 発行年月 | 1987年4月 |

*Jpn. J. Genet.* (1987) **62**, *pp.* 139-147

# Higher rates of nucleotide substitution in Drosophila than in mammals

BY Etsuko N. MORIYAMA

*Department of Human Culture, Ochanomizu University, Bunkyo-ku, Tokyo*, 112
*and*
*National Institute of Genetics, Mishima, Shizuoka-ken*, 411

(*Received February* 5, 1987)

ABSTRACT

To examine whether the rate of nucleotide substitution is affected by generation time of the organism, I attempted to estimate an accurate rate of synonymous (silent) substitution in Drosophila lineages, using alcohol dehydrogenase (Adh) and heat shock protein 82 (hsp82) genes. The results obtained suggest that the rate of synonymous substitution in Drosophila lineages is roughly $10^{-8}$ per site per year. This rate is approximately two times higher than that of rodents and ten times greater than higher primates. The higher rate in Drosophila may be explained by the shorter generation times of the Drosophila species, though the possibility that the mutation mechanism in Drosophila may differ from that in mammals cannot be excluded.

## 1. INTRODUCTION

One of the most important concepts in molecular evolutionary studies is the "molecular evolutionary clock" or the rate-constancy hypothesis. This concept implies that for a given protein or part of DNA, the rate of amino acid or nucleotide substitution is roughly constant among diverse lineages as well as within lineages over a chronological time. Since this remarkable feature of molecular evolution was first suggested by Zuckerkandl and Pauling (1965), this concept has been useful particularly for constructing phylogenetic trees at the molecular level. However, the rate-constancy hypothesis has also been very controversial, since the reason that the rate of molecular evolution is constant *per year* rather than *per generation* is unclear. This controversy has been one of the points for dispute in the neutralist-selectionist argument. Since rate constancy *per year* has been supported by many comparative studies of protein sequences and immunological distance, it has been argued that such observations support the neutral theory of molecular evolution (Kimura, 1969, 1983). On the other hand, critics of this argument contend that if the neutral theory is true, the rate should be constant *per generation* rather than *per year* as mutation rate is sometimes believed to be constant *per generation* (or *replication*).

If mutation rates among organisms having widely different generation spans are roughly equal to each other when measured taking one generation as the unit, then the mutation rate *per year* for an organism having a shorter generation span should be much higher than that for one having a longer generation span. Since the neutral theory suggests that evolutionary rate is equal to the mutation rate for neutral alleles, the evolutionary rate *per year* of nearly neutral changes such as synonymous substitutions must be much higher in rodents (having a shorter generation span) than in higher primates (having a longer generation span). Indeed, recent works by Wu and Li (1985) and Kikuno *et al.* (1985) suggest that rodents have a rate of synonymous substitution higher than humans.

If an organism with a shorter generation span has a higher mutation rate than one with a longer generation span, the evolutionary rate of Drosophila DNA should be much higher than those of rodents and higher primates. This is because the generation time for Drosophila is much shorter than those of rodents and higher primates. To examine this point, I have collected the published DNA sequence data of Drosophila species and have estimated the rate of synonymous (silent) substitution. The results suggest that the rate of synonymous substitution for Drosophila species is roughly two times higher than those for mammals, implying that the generation time effect may be an important factor in determining evolutionary rate. Of course, it does not exclude the possibility that changes in the mutation mechanism may involve evolutionary rate, since the mutation mechanism in Drosophila is not fully understood.

## 2. MATERIALS AND METHODS

### *Genes in Drosophila*

In Drosophila species, only a small number of genes have been sequenced at present. Only Adh genes and heat shock protein genes are available for sequence comparisons. Thus, I used the nucleotide sequences of Adh genes (Bodmer and Ashburner, 1984; Fischer and Maniatis, 1985; Coyne and Kreitman, 1986) and hsp82 genes (Blackman and Meselson, 1986) for a total of eight species of Drosophila; *D. melanogaster*, *D. simulans*, *D. mauritiana*, *D. orena*, *D. sechellia*, *D. pseudoobscura*, *D. mulleri* and *D. virilis*. In particular, *D. mulleri* has two types of Adh genes, a larval type (Adh-1) and an adult type (Adh-2). For this reason, these two gene sequences for *D. mulleri* were also used in the present study.

### *Divergence time in Drosophilids*

To estimate the rate of nucleotide substitution, it is necessary to know the divergence time between Drosophila species compared. The following three

estimates were obtained from paleontological and systematic data of Drosophilid.

(*1*) *Divergence between subgenera Sophophora and Drosophila:* Subfamily Drosophilinae consists of four genera; Neotanygastrella (fossil), Chymomyza, Scaptomyza and Drosophila. In genus Drosophila, there are two main subgenera, Sophophora and Drosophila. Since *D. melanogaster* and *D. virilis* belong to subgenera Sophophora and Drosophila respectively, the divergence between these two species should have occurred after the generic divergence. On the other hand, Neotanygastrella should have already existed about 30 million years (Myrs) before, because its fossil has been found in amber whose geological time was assigned to the mid-Oligocene (Wheeler, 1963). For this reason, the generic divergence within the subfamily Drosophilinae should have occurred more than 30 Myrs ago. Taking into account this fact and additional data on continental drift and geographic distribution of existent fruit flies, Throckmorton (1975) has proposed that the divergence between the two subgenera, Sophophora and Drosophila, could have occurred before the beginning of the Oligocene epoch, approximately 35 Myrs ago. It thus seems that the divergence time between *D. melanogaster* and *D. virilis* is more than 35 Myrs, possibly 35–40 Myrs. Similarly, the divergence time between *D. melanogaster* and *D. mulleri* may be 35–40 Myrs, because *D. mulleri* as well as *D. virilis* belongs to subgenus Sophophora. These divergence times are shorter than the estimates of Beverley and Wilson (1984).

(*2*) *Divergence of Sophophoran species groups:* Subgenus Sophophora consists of four species groups; obscura, melanogaster, willistoni and saltans. Throckmorton (1975) has suggested that the melanogaster and obscura species groups had separated during the period from mid-Oligocene to mid-Miocene (about 20–30 Myrs ago). Since *D. pseudoobscura* belongs to the obscura species group, it can be reasonably assumed that the divergence time between *D. pseudoobscura* and *D. melanogaster* is 20–30 Myrs.

(*3*) *Divergence of melanogaster species subgroups: D. simulans, D. mauritiana* and *D. sechellia* are sibling species of *D. melanogaster*. They are classified into the *melanogaster* complex. Similarly, *D. orena* is classified as part of the *yakuba* complex. For estimating divergence times among these species, the following approach has been taken. Genetic distances between *D. melanogaster* and *D. simulans* and between *D. simulans* and *D. mauritiana* have been estimated to be 0.450 and 0.179, respectively, using the electrophoretic mobilities of isozymes (Gonzalez *et al.* 1982). The genetic distance between *D. melanogaster* and *D. orena* also was estimated to be approximately 1.1 (Ashburner *et al.* 1984). Note that genetic distance (D) can be equated to

2ut, where t is the divergence time between the two species compared and u is the mutation rate per locus per year. Nei (1975) estimated the value of u to be $1\times10^{-7}$, assuming all mutations are neutral. Thus, divergence time (t) can be estimated by $D/(2\times10^{-7})$. In this way, the divergence times between *D. melanogaster* and *D. orena*, between *D. melanogaster* and *D. simulans*, and between *D. mauritiana* and *D. simulans* are estimated to be 5.5, 2.3, and 0.9 Myrs, respectively. The phylogenetic tree suggests that *D. simulans*, *D. mauritiana* and *D. sechellia* diverged from the ancestor of *D. melanogaster* (Ashburner *et al.* 1984; Coyne and Kreitman, 1986; Solignac *et al.* 1986). For this reason, the divergence times between *D. melanogaster* and *D. mauritiana*, and between *D. melanogaster* and *D. sechellia* are equal to that between *D. melanogaster* and *D. simulans*. Although the divergence times derived from electrophoretic data may not be as reliable as those derived from paleontological data, paleobiogeographical studies supports the estimate (2.3 Myrs) of the divergence time between *D. melanogaster* and its three sibling species (Lachaise *et al.*, personal communication). Thus, only the comparisons between *D. melanogaster* and its sibling species were used in the present study. These estimates of divergence times are summarized in Table 1.

Table 1. *Summary of divergence times (for details, see text).*

| Lineages compared | | | Time estimated (Myrs) | Time used in present study (Myrs) |
|---|---|---|---|---|
| *melanogaster*, *simulans*, *sechellia*, *mauritiana*, *orena*, *pseudoobscura* | — | *virilis*, *mulleri* | 35—40 | 40 |
| *melanogaster*, *simulans* | — | *pseudoobscura* | 20—30 | 30 |
| *melanogaster* | — | *orena* | 5.5 | |
| *melanogaster* | — | *simulans*, *mauritiana*, *sechellia* | 2.3 | 2.3 |
| *mauritiana* | — | *simulans* | 0.9 | |

## 3. RESULTS AND DISCUSSION

For the nucleotide sequences of the Adh genes and the hsp82 genes, the numbers of synonymous and nonsynonymous substitutions between Drosophila species compared were computed by the method of Nei and Gojobori (1986). Tables 2 and 3 show these numbers for the Adh and hsp82 genes, respectively.

Table 2. *The estimated numbers of synonymous (upper half) and nonsynonymous (lower half) substitutions between Adh genes for a given pair of Drosophila species, using the method of Nei and Gojobori (1986), based on the data of Bodmer and Ashburner (1984), Fisher and Maniatis (1985) and Coyne and Kreitman (1986).*

| | *melanogaster* | *simulans* | *mauritiana* | *orena* | *sechellia* | *mulleri*-2 | *mulleri*-1 |
|---|---|---|---|---|---|---|---|
| *melanogaster* | — | 0.038±0.014 | 0.038±0.014 | 0.129±0.028 | 0.060±0.018 | 0.821±0.109 | 0.884±0.119 |
| *simulans* | 0.004±0.002 | — | 0.022±0.011 | 0.124±0.027 | 0.032±0.133 | 0.820±0.110 | 0.883±0.119 |
| *mauritiana* | 0.009±0.004 | 0.005±0.003 | — | 0.135±0.029 | 0.043±0.015 | 0.790±0.105 | 0.850±0.114 |
| *orena* | 0.017±0.005 | 0.013±0.005 | 0.015±0.005 | — | 0.135±0.028 | 0.924±0.125 | 0.997±0.138 |
| *sechellia* | 0.003±0.002 | 0.0±0.0 | 0.005±0.003 | 0.013±0.005 | — | 0.772±0.102 | 0.835±0.111 |
| *mulleri*-2 | 0.114±0.015 | 0.111±0.015 | 0.114±0.015 | 0.116±0.015 | 0.112±0.015 | — | 0.120±0.027 |
| *mulleri*-1 | 0.120±0.015 | 0.121±0.015 | 0.124±0.016 | 0.126±0.016 | 0.121±0.015 | 0.022±0.006 | — |

Table 3. *The estimated numbers of synonymous (upper half) and nonsynonymous (lower half) substitutions between heat shock protein 82 genes for a given pair of Drosophila species, using the method of Nei and Gojobori (1986).*

| | *melanogaster* | *simulans* | *pseudoobscura* | *virilis* |
|---|---|---|---|---|
| *melanogaster* | — | 0.057±0.016 | 0.635±0.075 | 0.774±0.092 |
| *simulans* | 0.0±0.0 | — | 0.606±0.072 | 0.787±0.093 |
| *pseudoobscura* | 0.018±0.005 | 0.018±0.005 | — | 0.812±0.097 |
| *virilis* | 0.003±0.002 | 0.003±0.002 | 0.017±0.004 | — |

Figure 1 shows the relationship between the number of synonymous substitutions and the divergence time for various comparisons of Drosophila species. To examine whether or not the substitution rate for Drosophila species is higher than that for mammals, the rates of nucleotide substitution for Drosophila species were computed using the largest value of the estimated divergence times for each species comparison. This was done because a larger value for the divergence time gives a smaller value for the substitution rate. The estimated rate of synonymous substitution for Drosophila species ranges from $8.26\times10^{-9}$ to $13.0\times10^{-9}$/site/year, depending on species comparisons. However, it is clear from the figure that the linear relationship between these two quantities holds reasonably well. This implies that synonymous substitutions for Drosophila lineages occur at a roughly constant rate. The average rate of synonymous substitution for these Drosophila species is estimated to be $10.5\times10^{-9}$/site/year by regression analysis.

Using this substitution rate for Drosophila, divergence times between two species in the Drosophila lineages can be reestimated. In particular, the divergence time between *D. melanogaster* and *D. orena* and between *D. mauritiana* and *D. simulans* can be estimated to be 6.1 and 1.0 Myrs, respectively. These values are consistent with the values estimated from other lines of biological evidence already mentioned in Table 1. Thus, the estimated rate of synonymous substitution for Drosophila seems to be reasonable.

It is of particular interest to note that the rate of synonymous substitution for Drosophila is about two times higher than that for mammals, since the rate of synonymous substitutions for mammals is estimated to be, on the average, $5.0\times10^{-9}$/site/year (Miyata, 1982). The substitution rate for mammalian pseudogenes is also estimated at $5.0\times10^{-9}$/site/year (Li *et al.* 1981). Since the synonymous substitutions of functional genes and the substitutions of pseudogenes are free from functional constraints against amino acid changes, the rate for these substitutions may be almost the same as the mutation rate. Thus, the mutation rate for Drosophila could be more than

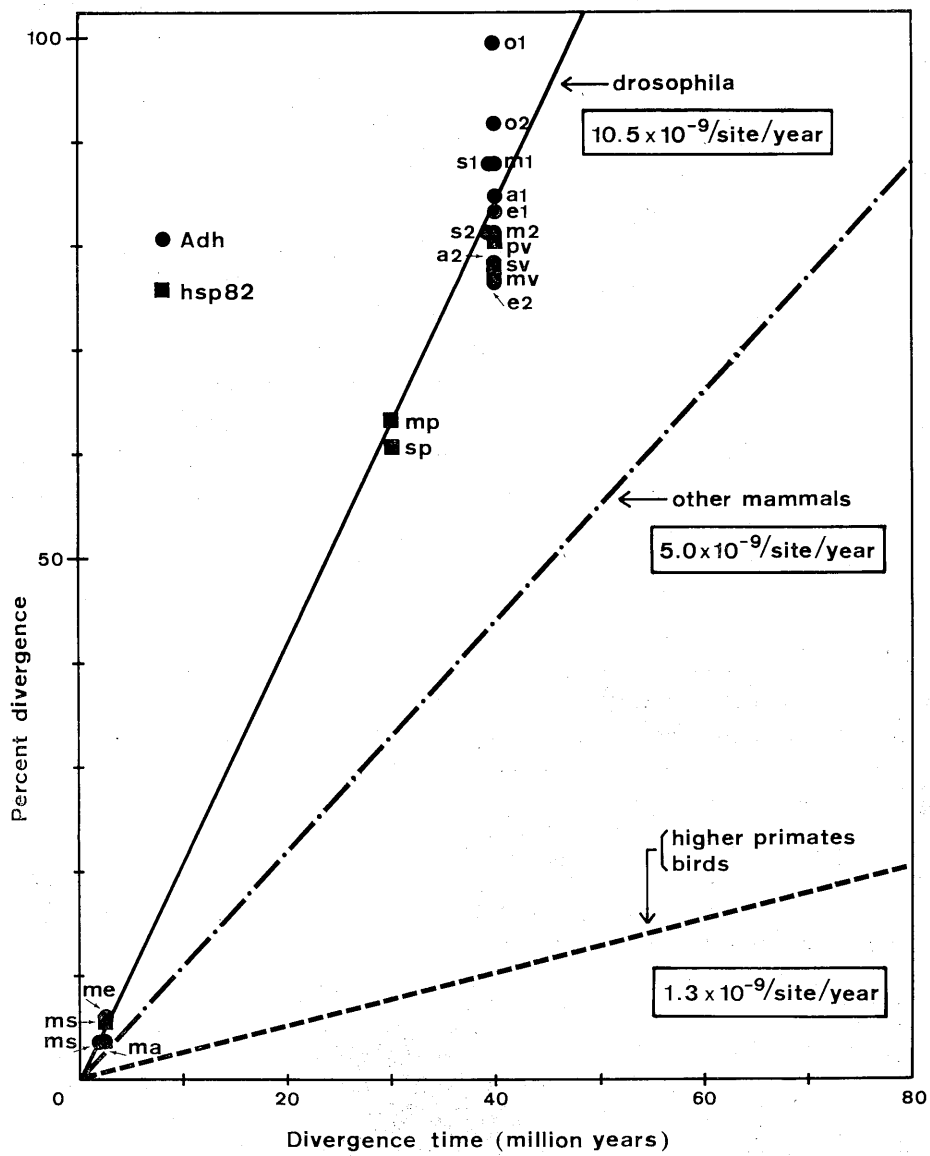


Fig. 1. The relationship between the number of synonymous substitutions and the divergence time for various comparisons of Drosophila species. Symbols "●" and "■" represent the data from Adh (Table 2) and hsp82 (Table 3) genes, respectively. A pair of letters with each symbol indicates species compared. For example, ms represents a comparison between *D. melanogaster* and *D. simulans*. The abbreviation for each species is as follows; m: *melangaster*, s: *simulans*, a: *mauritiana*, e: *sechellia*, o: *orena*, p: *pseudoobscura*, v: *virilis*, 1: *mulleri*-1, 2: *mulleri*-2. The rates of synonymous substitution for higher primates and birds (Britten, 1986) and other mammals (Miyata, 1982) are also shown in dashed lines.

two times higher than that for mammals when measured taking one year as the unit.

Britten (1986) has reported, mainly based on DNA-DNA hybridization data, that the rate of neutral DNA evolution for higher primates and birds is $1.3\times 10^{-9}$/site/year. Thus, the substitution rate in the Drosophila lineage is about eight times higher than his estimate for higher primates and birds. However, Britten (1986) also estimated that the rate of synonymous substitution for Drosophila as well as rodents and sea urchin is $6.6\times 10^{-9}$/site/year. Thus, the estimated rate for Drosophila in the present study is about 1.6 times higher than even Britten's estimate. It is probable that our estimate is more reasonable than Britten's, because he used only one set of data of DNA-DNA hybridization for a single comparison of Drosophila species.

Differences in substitution rates between taxonomic groups may be explained by difference in generation times, potential mutation rates or DNA repair mechanisms. The simplest explanation for the higher rate of synonymous substitution in Drosophila may be that a Drosophila species has a shorter generation time and therefore has higher mutation rates than mammals. In fact, the synonymous substitution rate is about 1.6 times higher than that for rodents, which is 1.3 times higher than other mammals such as bovines. An alternative explanation is that the higher rate of synonymous substitution for Drosophila may be due to changes in the mutation mechanism in Drosophila species. In particular, it is generally known that there are many transposable elements such as P elements, copia and copia-like elements in Drosophila species. These elements can cause mutations if they are inserted into the coding or controlling regions of genes. Several lines of data have shown that these elements raise the sensitivity for mutagens, involving the repair mechanism (Green *et al.* 1986; Nakamura *et al.* 1985). This implies that these elements can contribute to the elevation of spontaneous mutation rates. Therefore, the existence of these elements may support the explanation for a higher rate of synonymous substitution in Drosophila.

In the present study, it was speculated that differences in generation times can be a primary cause of a difference in evolutionary rates among different lineages. As mentioned before, however, the difference in the mutation (or repair) mechanisms may also be important. Britten (1986) maintains that changes in DNA repair mechanisms, as well as the differences in generation times or numbers of germline DNA replications per year, are the primary causes of differences in evolutionary rates. It is thus possible that both factors, generation time effect and changes of the mutation mechanism, affect differences in evolutionary rates for various lineages of organisms.

The author is grateful to Drs. J. Ohta and S. Ishiwa-Chigusa, Ochanomizu University and Drs. T. Maruyama, T. Ohta T. Watanabe and T. Gojobori, National Institute of Genetics at Mishima for their help and valuable comments. This study was supported by a Grant-in-Aid given to Dr.

T. Gojobori from the Ministry of Education, Science and Culture of Japan.

## REFERENCES

ASHBURNER, M., BODMER, M. and LEMEUNIER, F. (1984) On the evolutionary relationships of *Drosophila melanogaster*. *Devel. Genet.* **4**, 295–312.

BEVERLY, S. M. and WILSON, A. C. (1984) Molecular evolution in Drosophila and higher Diptera. II. A time scale for fly evolution. *J. Mol. Evol.* **21**, 1–13.

BLACKMAN, R. K. and MESELSON, M. (1986) Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the Drosophila hsp82 gene. *J. Mol. Biol.* **188**, 499–515.

BODMER, M. and ASHBURNER, M. (1984) Conservation and change in the DNA sequences coding for alcohol dehydrogenase in sibling species of Drosophila. *Nature* **309**, 425–430.

BRITTEN, R. J. (1986) Rates of DNA sequence evolution differ between taxonomic groups. *Science* **231**, 1393–1398.

COYNE, J. A. and KREITMAN, M. (1986) Evolutionary genetics of two sibling species, *Drosophila simulans* and *D. sechellia*. *Evolution* **40**(**4**), 673–691.

FISHER, J. A. and MANIATIS, T. (1985) Structure and transcription of the *Drosophila mulleri* alcohol dehydrogenase genes. *Nucl. Acids Res.* **13**, 6899–6917.

GONZALEZ, A. M., CABRERA, V. M., LARRUGA, J. M. and GULLON, A. (1982) Genetic distance in the sibling species *Drosophila melanogaster*, *Drosophila simulans* and *Drosophila mauritiana*. *Evolution* **36**, 517–522.

GREEN, M. M., TODO, T., RYO, H. and FUJIKAWA, K. (1986) Genetic-molecular basis for a simple *Drosophila melanogaster* somatic system that detects environmental mutagens. *Proc. Natl. Acad. Sci. USA* **83**, 6667–6671.

KIMURA, M. (1969) The rate of molecular evolution considered from the standpoint of population genetics. *Proc. Natl. Acad. Sci. USA* **63**, 1181–1188.

KIMURA, M. (1983) *The Neutral Theory of Molecular Evolution*. Cambridge University Press, London.

KIKUNO, R., HAYASHIDA, H. and MIYATA, T. (1985) Rapid rate of rodent evolution. *Proc. Japan. Acad.* **61**(**B**), 153–156.

LI, W.-H., GOJOBORI, T. and NEI, M. (1981) Pseudogenes as a paradigm of neutral evolution. *Nature* **292**, 237–239.

MIYATA, T. (1982) In: *Molecular Evolution, Protein Polymorphism and the Neutral Theory*. (ed. Kimura, M.) pp. 233–266. Japan Scientific Societies Press, Tokyo.

NAKAMURA, E., INOUE, T. and CHIGUSA, S. I. (1985) A study of the effect of EMS on male recombination in the P-M system of *D. melanogaster*. *Jpn. J. Genet.* **60**, 618.

NEI, M. (1975) *Molecular Population Genetics and Evolution*. North-Holland, Amsterdam.

NEI, M. and GOJOBORI, T. (1986) Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* **3**(**5**), 418–426.

SOLIGNAC, M., MONNEROT, M. and MOUNOLOU, J.-C. (1986) Mitochondrial DNA evolution in the *melanogaster* species subgroup of *Drosophila*. *J. Mol. Evol.* **23**, 31–40.

THROCKMORTON, L. H. (1975) The phylogeny, ecology, and geography of *Drosophila*, in: *Handbook of Genetics*. (ed. King, R. C.), **3**, pp. 421–469. Plenum Press, New York.

WHEELER, M. R. (1963) A note on some fossil Drosophilidae (Diptera) from the amber of Chiapas, Mexico. *J. Paleontol.* **37**, 123–124.

WU, C.-I. and LI, W.-H. (1985) Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. USA* **82**, 1741–1745.

ZUCKERKANDL, E. and PAULING, L. (1965) Evolutionary divergence and convergence in proteins, in: *Evolving Genes and Proteins*. (eds. Bryson, V. and Vogel, H. J.), pp. 97–116. Academic Press, New York.