Genes Genet. Syst. (2004) **79**, p. 19–26

Nucleotide variation of seven genes in *Drosophila kikkawai*

Hiroki Goto, Nobuyuki Inomata*, Alfred E. Szmidt and Tsuneyuki Yamazaki†

Department of Biology, Graduate School of Sciences, Kyushu University, Fukuoka 812-8581, Japan

(Received 29 October 2003, accepted 10 December, 2003)

We examined levels and patterns of nucleotide variation in 21 strains of *Drosophila kikkawai* from Miyako island, Japan for the partial regions of the following seven nuclear genes: *Adh*, *Ddc*, *esc*, *ksr*, *Pgi*, *su(f)*, and *Tpi*. The nucleotide variation at total sites (π_t) ranged from 0.0013 in the *ksr*, to 0.0173 in the *Adh*. The nucleotide divergence at total sites (K_t) between *D. kikkawai* and *D. lini* ranged from 0.0286 in the *Tpi* to 0.0687 in the *su(f)*. The levels of nucleotide polymorphism and divergence were heterogeneous among the investigated gene regions. The HKA test, which tests imbalance between the intra and interspecific nucleotide variation, showed that the intraspecific nucleotide variation in the *Pgi* region was much lower than the interspecific variation, while intraspecific variation in the *Tpi* region was only slightly lower than interspecific variation. The MK test showed an excess of low frequency replacement polymorphic changes in the *Adh* region, suggesting that most replacement mutations are deleterious. Fay and Wu's test detected an excess of newly arisen variants in the *Ddc* region. In total, four of the seven gene regions showed significant deviation from the neutrality.

Key words: *Drosophila kikkawai*, natural selection, nucleotide variation, polymorphism

INTRODUCTION

Understanding the maintenance mechanisms of nucleotide variation in natural populations is important in evolutionary genetics, because nucleotide variation within species is thought to be the raw material for its evolution. The levels and patterns of nucleotide variation in extant populations reflect evolutionary causes such as genetic drift, natural selection, population structure and history. Distinguishing effects of these evolutionary causes is often difficult because some of them can produce similar patterns of nucleotide variation. One way of overcoming this difficulty is to examine nucleotide variation at multiple loci in samples collected from discrete populations. The effects of selection can vary among genes and populations, while population structure and history should uniformly affect the patterns of nucleotide variation of different genes in samples taken from a single population. This approach, combined with recent advances in population genetics theory has considerably improved our ability to detect changes in nucleotide variation caused by various forms of natural selection (Otto, 2000). This is well illustrated by several recent studies on nucleotide variation in various organisms, which have provided much new evidence for different types of selection acting upon individual coding regions (*e.g.* Filatov and Charlesworth, 1999).

For the past two decades, *Drosophila melanogaster* is one of the most extensively studied organisms in terms of intraspecific nucleotide variation (see *e.g.* Moriyama and Powell, 1996). On the other hand, there are considerably fewer similar studies on other *Drosophila* species including *D. kikkawai* (*e.g.* Inomata and Yamazaki, 2002; Kovacevic and Schaeffer, 2000; Vieira and Charlesworth, 1999). In addition, most of these studies involved a very limited number of coding regions (*e.g.* Inomata and Yamazaki, 2002; Kovacevic and Schaeffer, 2000; Vieira and Charlesworth, 1999). Therefore, to build a more comprehensive picture of the evolutionary forces shaping the levels and patterns of nucleotide variation in *Drosophila*, it is desirable to accumulate data on nucleotide variation for additional genes in closely as well as distantly related species. To provide new information on this subject we studied nucleotide variation in partial regions of the following seven nuclear genes in one population of *D. kikkawai* from the Miyako island, Japan: *Alcohol dehydrogenase* (*Adh*), *Dopa decarboxylase* (*Ddc*), *extra sexcombs* (*esc*), *kinase suppressor of Ras* (*ksr*), *Phosphoglucose isomerase* (*Pgi*), *suppressor of forked* (*su(f)*)

Edited by Fumio Tajima

* Corresponding author: E-mail: ninomscb@mbox.nc.kyushu-u.ac.jp

† Present address: The Research Institute of Evolutionary Biology, 2-4-28, Kamiyoga, Setagaya-ku, Tokyo 150-0098.

Note: Nucleotide sequence data reported are available in the DDBJ Data Library under the accession nos. AB154542-AB154695.

and *Triose phosphate isomerase* (*Tpi*). With our sampling scheme, we hoped to avoid artificial patterns of nucleotide variation caused by combining samples collected from different populations. Taxonomically and phylogenetically, *D. kikkawai* is closer to *D. melanogaster* than to *D. pseudoobscura* but it belongs to a different species subgroup (*montium*) (*e.g.* Ashburner, 1989a; Inomata and Yamazaki, 2000; Zhang et al., 2003).

Our main objective was to test whether the investigated gene regions are evolving in the neutral fashion. We found that nucleotide variation was variable among the investigated gene regions. In addition, some statistical tests revealed significant deviation from the neutral expectation. We discuss the possible causes of these deviations.

MATERIALS AND METHODS

Fly samples. Twenty-one females of *Drosophila kikkawai* were collected on the Miyako island (24°47'N), Okinawa Prefecture, Japan. Thirteen of the 21 females were collected in October 2001. The other females were collected in March 2002. Twenty-one independent isofemale lines were established from the collected females, and kept at 25° on the standard cornmeal medium. There was no significant difference in the levels and patterns of nucleotide variation between the two collections (data not shown). Therefore, in the following analyses all 21 isofemale lines were treated as one sample from a single population. A single line of *Drosophila lini*, which is the sibling species of *D. kikkawai*, was obtained from the Tokyo Metropolitan University and used for interspecific comparisons of nucleotide variation.

Molecular analyses. Genomic DNA was extracted from single male flies using the standard method (Ashburner, 1989b). The conditions of PCR amplification using 30 cycles were as follows: denaturing at 94° for 30 seconds, annealing at 60° for 30 seconds and extension at 72° for 1 minute. Using gene specific primers, PCR amplification was done for partial regions of the following seven nuclear genes: *Alcohol dehydrogenase* (*Adh*), *Dopa decarboxylase* (*Ddc*), *extra sexcombs* (*esc*), *kinase suppressor of Ras* (*ksr*), *Phosphoglucose isomerase* (*Pgi*), *suppressor of forked* (*su(f)*), and *Triose phosphate isomerase* (*Tpi*). The PCR primers were as follows; 5'-TGATTTTCGTTGCCGGTCTGGGAGG-3' and 5'-CAGGGTGCCCAAGTCCAGTTTCCAGA-3' for the *Adh*, 5'-CCTACTTCCCCACGGCCAACTC-3' and 5'-GAGCAGTCAAAGTTCACCAGCA-3' for *Ddc*, 5'-CAACACGCTGCTGGGCAAGGATGA-3' and 5'-GATGTAGTTGTCCCGGTTTCCAGC-3' for the *esc*, 5'-TTGCCTGGAGTCGAAGCTGGTGCG-3' and 5'-TGGGTGGCGAGGGCGTAAGGGTTA-3' for the *ksr*, 5'-AAGGTCACAGAGTT-CGGCAT-3' and 5'-TGATGTCCCAAATGATGCCCTGAACGAAGA-3' for the Pgi, 5'-GACAAGGAATGGGGCATGGAGCG-3' and 5'-TGCTGTGATCTTTTGGTTCTCTG-3' for the *su(f)*, 5'-AAGATGAACGGCGACCAGAAGTC-3' and 5'-CTCCTTGGCGTTGGCGGCGGTCA-3' for the *Tpi*. Figure 1 shows the structures of the investigated gene regions. The length of the investigated gene regions were: 672, 580, 515, 494, 727, 395 and 545bp for *Adh*, *Ddc*, *esc*, *ksr*, *Pgi*, *su(f)* and *Tpi* respectively. Only a partial exon was amplified for the *Ddc* and *ksr* genes. For the other five genes, partial exons and introns were amplified. The structures of the investigated gene regions in *D. kikkaw*ai and *D. lini* were the same as those in *D. melanogaster*. Both strands of PCR products were directly sequenced, using the ABI PRISM 3100 Genetic Analyzer and DNA Sequencing Kit (BigDye termination Cycle-Sequencing Kit version 3.0, ABI). In the course of direct sequencing, heterozygous sites were found in some lines. Because some of these sites contained gaps, we could not obtain sequence information. To determine a haplotype sequence, heterozygous PCR products were subcloned into the pGEM-T Easy Vector (Promega). For heterozygous sequences without gaps, we sequenced only the part of the clone containing heterozygous sites. For heterozygous sequences with gaps, more than three independent clones were sequenced for both strands until two identical sequences were obtained. Only one haplotype sequence was determined for each fly. In total, 21 sequences were analyzed for each gene region. We found that an intron was lost in the *esc* gene region in one line of *D. kikkawai*. Analysis of nucleotide variation in the exons of the *esc* gene region revealed that there was little difference between data sets with and without intronless sequence (π_t = 0.0146 vs. 0.0142 respectively). Therefore, to examine nucleotide variation in the intron of the *esc* region we eliminated the intronless sequence from our

Figure 1. Structures of gene regions analyzed in this study. The black box indicates exon and the line indicates intron.

data set. Consequently, only 20 sequences were analyzed for *esc* gene regions. The sequences obtained in this study were deposited in the DNA Data Bank of Japan (DDBJ) under accession numbers: AB154542-AB154695.

Data Analyses. DNA sequences were edited using SeqPup program version 0.6f (Gilbert, 1996). Multiple alignment of sequences was done using CLUSTAL X program (Thompson et al., 1997), and then corrected by hand. DNA variation, linkage disequilibrium, recombination parameter, Tajima's test (Tajima, 1989), HKA test (Hudson et al., 1987) and MK test (McDonald and Kreitman, 1991) were calculated using the DnaSP program version 3.14 (Rozas and Rozas, 1999). The multilocus HKA test was performed using the HKA program at Jody Hey's web site (http://lifesci.rutgers.edu/~heylab/). Fay and Wu's *H* test statistic (Fay and Wu, 2000) was calculated at the Fay's web site (http://crimp.lbl.gov/htest.html). To obtain critical values of this test statistic, 10,000 coalescent simulations were run for each gene. To test the heterogeneity in π/K ratio, where π is the nucleotide variation within species and K is nucleotide divergence between species, McDonald's tests (McDonald, 1998) were performed using the DNA slider program (McDonald, 1998). Statistical significance was estimated by computer simulations with 1,000 replicates.

RESULTS

Nucleotide variation. We examined nucleotide variation in the partial regions of the following seven genes in *Drosophila kikkawai*: *Adh*, *Ddc*, *esc*, *ksr*, *Pgi*, *su(f)* and *Tpi*. In total, there were 129 polymorphic sites and seven indels (insertions or deletions) in the 3969 bp length of the seven gene regions. All indels were found in introns of the *Adh*, *esc*, *su(f)* and *Tpi* gene regions. All seven replacement polymorphic sites were found in the *Adh* gene region. Three of these seven sites were singletons. There were no fixed replacement sites between *D. kikkawai* and *D. lini* in the *Adh* gene region. Nucleotide variation in the investigated gene regions is summarized in Table 1. The level of nucleotide variation varied among the investigated gene regions. The average number of nucleotide differences per site at total sites (π_t) for individual gene regions ranged from 0.0013 in the *ksr* to 0.0173 in the *Adh*, and the average number of nucleotide differences per site at the silent sites (π_s) ranged from 0.0050 in the *Pgi* to 0.0540 in the *Ddc*. The average values of π_t and π_s for the seven gene regions in *D. kikkawai* were 0.0105 and 0.0323, respectively. At the face value, the nucleotide variation observed in *D. kikkawai* was roughly two times higher than that in *D. melanogaster* (π_t: 0.0043, π_s: 0.0134; Moriyama and Powell, 1996). Nucleotide divergence at total sites (K_t) and silent sites (K_s) between *D. kikkawai* and *D. lini* was also variable among gene regions (see Table 2). It ranged from 0.0286 in the *Tpi* to 0.0687 in the *su(f)* for K_t, and from 0.0820 in the *Tpi* to 0.2077 in the *Ddc* for K_s. These K_s values were almost the same as those between *D. melanogaster* and *D. simulans* (see *e.g.* Ayala et al., 2002; Baines et al., 2002).

Table 1. Summary of nucleotide variation within *D. kikkawai*.

		Loci							
		Adh	*Ddc*	*esc*	*ksr*	*Pgi*	*su(f)*	*Tpi*	average
Total	Length	698	580	516	500	727	402	546	
	S[a]	32	29	25	4	3	6	30	
	π	0.0173	0.0134	0.0170	0.0013	0.0015	0.0058	0.0172	0.0105
	θ	0.0132	0.0149	0.0142	0.0023	0.0012	0.0042	0.0153	0.0093
Silent	Length	256.74	143.55	168.83	123.59	222.17	138.50	177.67	
	S[a]	25	29	25	4	3	6	30	
	π	0.0407	0.0540	0.0520	0.0052	0.0050	0.0166	0.0526	0.0323
	θ	0.0271	0.0600	0.0434	0.0091	0.0038	0.0121	0.0469	0.0289

[a]: the number of segregating sites.

Table 2. Summary of nucleotide divergence (K) between *D. kikkawai* and *D. lini*.

		Loci						
		Adh	*Ddc*	*esc*	*ksr*	*Pgi*	*su(f)*	*Tpi*
Total	Length	698	580	516	500	727	402	546
	K	0.0542	0.0531	0.0481	0.0309	0.0328	0.0687	0.0286
Silent	Length	256.77	143.55	168.8	121.89	222.18	137.53	177.67
	K	0.1367	0.2077	0.1407	0.1168	0.1029	0.1974	0.0820

Neutrality tests. We performed five neutrality tests to determine whether the observed levels and patterns of nucleotide variation fit the neutral model. The values of Tajima's *D* and Fay and Wu's *H* are shown in Table 3. Tajima's test statistics did not show significant deviation from the neutrality in the investigated gene regions. Fay and Wu's test detects an excess of newly arisen variants. We found significant deviation from the neutrality at 5% level in the *Ddc* region even after the Bonferroni correction (Table 3). The other gene regions did not show significant deviation from the neutrality.

The MK test tests the null hypothesis that the ratio of fixed replacement to synonymous substitutions between species is equal to that of replacement to synonymous polymorphic changes within species. When there was no fixed replacement substitution or replacement polymorphic change, we could not perform the *G*-test of independence. Therefore, we used the two-tailed Fisher's exact test to test the null hypothesis. The results are summarized in Table 4. Significance at 5% level was observed in the *Adh* gene, suggesting an excess of replacement polymorphic changes. When the MK test was performed for all seven gene regions together, we could not reject the null hypothesis (Table 4).

The multilocus HKA test showed significant deviation from the neutrality (x^2 = 15.7225, *d.f.* = 6, *P* = 0.0153). The HKA test was also conducted for all pairs of gene regions (Table 5). We found seven significant pairs of regions, although these pairs were no longer significant after the Bonferroni correction. The *Pgi* and *Tpi* regions showed imbalance between the intra and interspecific nucleotide variation (see Table 5). Compared with the other gene regions, the intraspecific nucleotide variation in the *Pgi* region was much lower than the interspecific variation (see Table 1 and 2). Howerver, for the *Tpi* gene region, intraspecific variation was only slightly lower

Table 3. Tests of the selective neutrality.

	Adh	*Ddc*	*esc*	*ksr*	*Pgi*	*su(f)*	*Tpi*
n	21	21	20	21	21	21	21
Tajima's *D*	1.1927	–0.3897	0.7654	–1.2142	0.8534	1.1640	0.4714
Fay and Wu's *H*	0.4048	–12.8095*	–3.2316	–1.3238	0.4095	0.0714	2.2905

n: number of sequences used, *: significant at 5% level after the Bonferroni correction for multiple comparisons.

Table 4. The McDonald and Kreitman tests.

		Synonymous		Replacement		Fisher's Exact Test
Gene	*n*	Fixed	Polymorphic	Fixed	Polymorphic	*P*-value
Adh	21	15	12	0	7	0.0106*
Ddc	21	19	31	1	0	0.3542
esc	20	9	20	1	0	0.3333
ksr	21	13	4	1	0	1.0000
Pgi	21	19	3	1	0	1.0000
su(f)	21	8	3	0	0	NC
Tpi	21	5	12	1	0	0.3333
Total		88	85	5	7	0.5451[a]

n: number of sequences used, *: significant at 5% level before the Bonferroni correction for multiple comparisons, NC: not calculated because of no fixed and polymorphic replacement substitution, [a]: G test of independence was applied.

Table 5. *P* values of the HKA test for all pairwise comparisons between gene regions.

	Adh	*Ddc*	*esc*	*ksr*	*Pgi*	*su(f)*
Ddc	0.9203					
esc	0.8449	0.9214				
ksr	0.1391	0.1232	0.0686			
Pgi	0.0247*	0.0209*	0.0183*	0.4388		
su(f)	0.0658	0.0560	0.0486*	0.8422	0.5342	
Tpi	0.3181	0.3797	0.4542	0.0288*	0.0033**	0.0089**

*: significant at 5% level before the Bonferroni correction for mutiple comparisons, **: significant at 1% level before the Bonferroni correction.

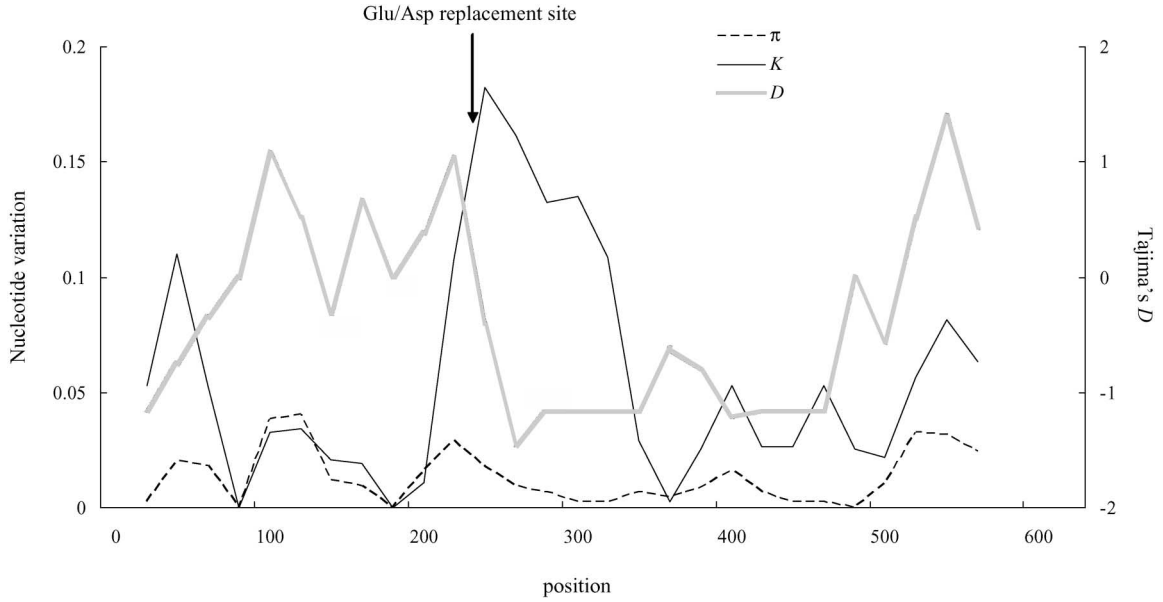

Figure 2. Sliding window plot analysis of nucleotide variation (π; dotted line) within *D. kikkawai*, nucleotide divergence (*K*; thin line) between *D. kikkkawai* and *D. lini* and Tajima's *D* (bold line) in the *Ddc*. Window size is 50bp and step size is 20bp. Arrow indicates the fixed replacement site between *D. kikkawai* and *D. lini*.

Table 6. The estimates of minimum number of recombination events per gene region.

	Adh	*Ddc*	*esc*	*ksr*	*Pgi*	*su (f)*	*Tpi*
Length	672	580	515	494	727	395	545
R_M	4	7	5	0	0	0	6

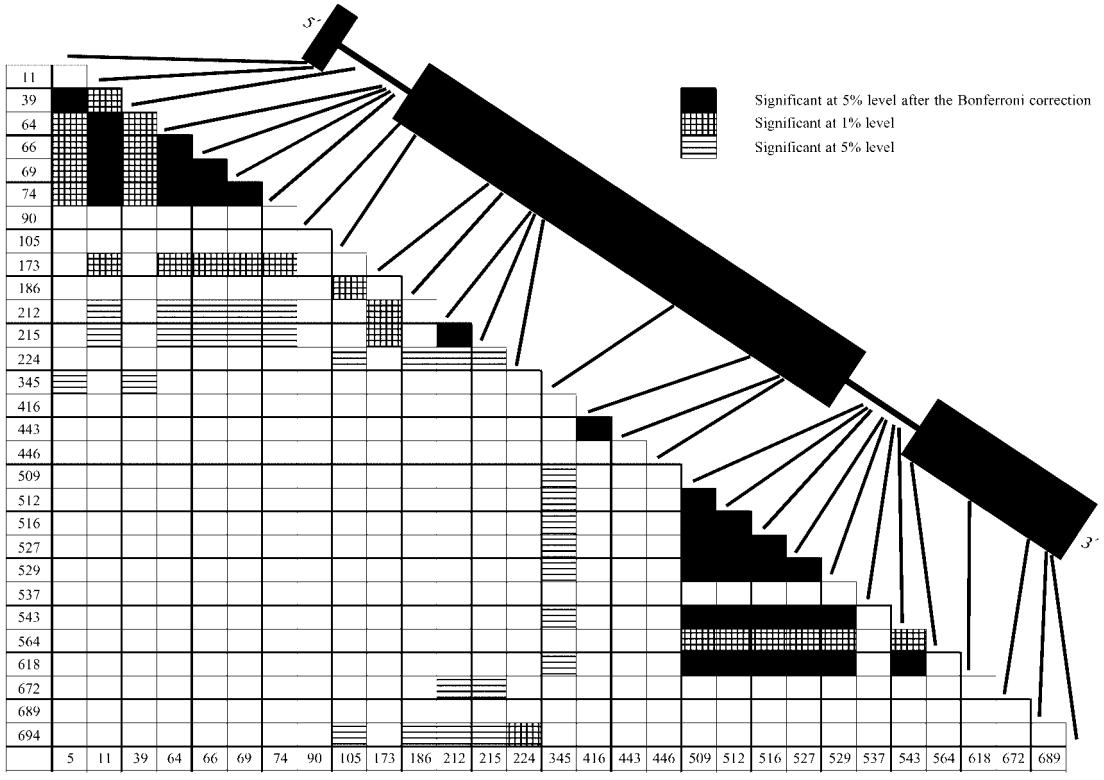

Figure 3. Linkage disequilibrium between polymorphic sites in the *Adh*. Only the informative sites were used for Fisher's exact test.

than interspecific variation (see Table 1 and 2).

The McDonald's tests, which test the heterogeneity in π/K ratio within a gene region, were also performed, where π is nucleotide variation within *D. kikkawai* and K is nucleotide divergence between *D. kikkawai* and *D. lini*. The Runs statistics (K_R) value for the *Ddc* region showed the significant intragenic heterogeneity at 5% level (Figure 2).

Intragenic recombination and linkage disequilibrium. To determine whether recombination has affected the levels and patterns of nucleotide variation in the investigated gene regions we estimated the recombination parameter using methods proposed by Hudson and Kaplan (1985). The minimum number of recombination events per gene, R_M, is shown in Table 6. We could detect recombination in the *Adh*, *Ddc*, *esc* and *Tpi*, but not in the *ksr*, *Pgi* and *su(f)* regions.

Linkage disequilibrium between segregating sites was investigated for each of the seven partial gene regions with informative sites. Significant linkage disequilibria were found in the *Adh*, *esc* and *su(f)* gene regions. Especially, in the *Adh*, we found highly significant linkage disequilibria ($P < 0.05$ after the Bonferroni correction) between the adjacent segregating sites within each intron (Figure 3). There was a 70bp long cluster of significant linkage disequilibria in the first intron including six informative sites and another cluster (21bp) in the second intron including five informative sites. The pairs of sites with highly significant linkage disequilibrium were among G-A at 64, G-A at 66, C-T at 69 and T-C at 74 in the first intron, and G-C at 509, G-A at 512, T-A at 516, T-A at 527, A-T at 529 and G-A at 543 in the second intron. In the *esc*, significant linkage disequilibria were found in the part of 5' region (data not shown). Significant linkage disequilibria in the *su(f)* were found for all pairwise comparisons of four informative sites (data not shown). Unfortunately, we could not examine linkage disequilibria between individual genes because we could not infer the combination of haplotypes for individual gene regions.

DISCUSSION

DNA polymorphism. We examined nucleotide variation of *D. kikkawai* in the Miyako population. We found that the levels and patterns of nucleotide variation were variable among the seven gene regions included in this study. The multilocus HKA test showed significant deviation from the neutrality, which indicates the heterogeneity of π/K among the investigated gene regions. Moreover, other neutrality tests gave significant results for four out of the seven regions, although most of these results were not significant after the conservative Bonferroni correction. The effects of selection can vary among genes and populations, while population structure and history should uniformly affect the patterns of nucleotide variation of different genes in samples taken from a single population. We did not observe the uniform patterns of nucleotide variation among gene regions. This may imply that population structure and population history did not affect the observed patterns of nucleotide variation in the Miyako population of *D. kikkawai*. Miyako island is located in the Pacific Ocean on the southwest side of Japan. It is a small (about 160 km^2) flat island about 150 km away from the nearest island Ishigaki, and about 310 km away from the Okinawa main island. Taking into consideration the geographic location of the Miyako island migration from other populations is not likely to be frequent. Therefore, Miyako population of *D. kikkawai* may be a panmictic isolated population. To confirm this suggestion, information about nucleotide variation in other populations of *D. kikkawai* is necessary.

In this study, we examined partial gene regions. Therefore, our results are informative about the patterns of nucleotide variation, although they may not be sufficient for inferring whether these patterns are associated with the function of individual genes. In the following section, we discuss the possible causes of the observed deviations from the neutrality except for the demographic factors.

Possible causes of the deviations from neutrality. Using the HKA test, we found significance in seven of the twenty-one pairwise comparisons among the seven gene regions (see Table 5). The imbalance between intra and interspecific nucleotide variation was found in the *Pgi* and *Tpi* (see Table 5). Compared to the other gene regions investigated in our study, the intraspecific nucleotide variation in the *Pgi* region was much lower than the interspecific variation, while intraspecific variation in the *Tpi* region was only slightly lower than interspecific variation. These results suggest that both regions do not evolve in the neutral fashion. One possible explanation is that the observed patterns of nucleotide variation in the *Tpi* and *Pgi* regions could be due to a decrease and an increase of functional constraint in the *D. kikkawai* lineage, respectively. Other possible explanations for the *Pgi* region involve background selection (Charlesworth et al., 1993) or selective sweep (Hill and Robertson, 1966; Smith and Haigh, 1974), both of which could reduce intraspecific variation. Unfortunately, with our present results it is not possible to determine, which of these explanations is most reasonable. Further studies are necessary to clarify the cause of the observed imbalance between intra and interspecific variation in the *Pgi* and *Tpi* regions.

It has been suggested that in the *Adh* of *D. melanogaster*, the Fast/Slow (Lys/Thr) isozyme polymorphism is

maintained by balancing selection (Kreitman, 1983; Kreitman and Aguade, 1986; Kreitman and Hudson, 1991). An excess of fixed replacement substitutions was observed between *D. melanogaster* and its sibling species *D. simulans*, suggesting adaptive evolution after their speciation (McDonald and Kreitman, 1991). We found seven replacement polymorphic changes in *D. kikkawai* and no fixed replacement substitution between *D. kikkawai* and *D. lini*. The number of newly arisen variants at these seven replacement sites were low; one at three sites, two at three sites and three at one site. Taking into account the absence of fixed replacement substitution between the two species, these observations suggest that most replacement mutations are deleterious in the *Adh* gene region of *D. kikkawai*.

We observed a cluster of linkage disequilibria between the adjacent sites in the *Adh* introns. One explanation for our observation is that recombination is not sufficient to break down linkage between adjacent segregating sites. However, we did not find significant linkage disequilibrium between other adjacent segregating sites in this gene region. Another explanation is that the observed linkage disequilibrium was caused by an admixture of subdivided populations. This is also not likely because we did not detect similar clusters of linkage disequilibria in any of the other investigated gene regions of *D. kikkawai*. Therefore, the most plausible explanation is that natural selection acts on the *Adh* gene. Similar cluster of linkage disequilibria was found in the intron of *D. pseudoobscura* (Schaeffer and Miller, 1993). In *D. pseudoobscura*, epistatic selection seems to act on the intron of the *Adh* to maintain the pre-mRNA secondary structure (Schaeffer and Miller, 1993; Kirby et al., 1995). Further theoretical and experimental studies are necessary to confirm whether epistatic selection acts on the intron of the *Adh* in *D. kikkawai*.

Fay and Wu's test showed that the *H* statistics was highly significant in the *Ddc* gene region (Table 3), and the McDonald's test was also significant. Close examination of Tajima's *D* using the window plot analysis showed that the *D* value was –1.45 around 260th nucleotide site (Figure 2). Around this site, we found that interspecific nucleotide variation was high and intraspecific nucleotide variation was low compared to the other sites of this gene region (Figure 2). That is, the difference between intra and interspecific variation was larger. At the 231st position, we found G/T replacement substitution, which results in Glu/Asp amino acid substitution between *D. kikkawai* and *D. lini*. One possible scenario is as follows; adaptive replacement substitution has occurred and was fixed in the *Ddc* of *D. kikkawai*, and then variation of the linked sites was reduced by selective sweep (Hill and Robertson, 1966; Smith and Haigh, 1974). In *Drosophila*, the *Ddc* catalyzes the decarboxylation of dopa into dopamine, and functions in the epidermis and in the central neural system (Lunan and Mitchell, 1969; Wright et al., 1976). However, the relationship between the replacement substitution at 231st nucleotide position and the observed imbalance between intra and interspecific variation in the *Ddc* gene region remains unclear. Studies of additional populations would help to clarify this issue.

We thank Mr. T. Kado for useful discussion. We thank Dr. Y. Fuyama for *Drosophila lini*, Dr. M. Watada for helpful suggestions on *Drosophila* collection. This work was supported by research grants to N. I, A. E. S. and T. Y. from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

REFERENCES

Ashburner, M. (1989a) *Drosophila*, a laboratory handbook. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Ashburner, M. (1989b) *Drosophila*, a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Ayala, F. J., Balakirev, E. S., and Saez, A. G. (2002) Genetic polymorphism at two linked loci, *Sod* and *Est-6*, in *Drosophila melanogaster*. Gene **300**, 19–29.

Baines, J. F., Chen, Y., Das, A., and Stephan, W. (2002) DNA sequence variation at a duplicated gene: Excess of replacement polymorphism and extensive haplotype structure in the *Drosophila melanogaster bicoid* region. Mol. Biol. Evol. **19**, 989–998.

Charlesworth, B., Morgan, M. T., and Charlesworth, D. (1993) The effect of deleterious mutations on neutral molecular variation. Genetics **134**, 1289–1303.

Fay, J. C., and Wu, C. I. (2000) Hitchhiking under positive Darwinian selection. Genetics **155**, 1405–1413.

Filatov, D. A., and Charlesworth, D. (1999) DNA polymorphism, haplotype structure and balancing selection in the Leavenworthia PgiC locus. Genetics **153**, 1423–1434.

Gilbert, D. G. (1996) SeqPup, biosequence editor and analysis platform. Bionet. Software.

Hill, W. G., and Robertson, A. (1966) The effect of linkage on limits to artificial selection. Genet. Res. **8**, 269–294.

Hudson, R. R., and Kaplan, N. L. (1985) Statistical properties of the number of recombination events in the history of a sample of DNA sequences. Genetics **111**, 147–164.

Hudson, R. R., Kreitman, M., and Aguade, M. (1987) A test of neutral molecular evolution based on nucleotide data. Genetics **116**, 153–159.

Inomata, N., and Yamazaki, T. (2000) Evolution of nucleotide substitutions and gene regulation in the *Amylase* multigenes in *Drosophila kikkawai* and its sibling species. Mol. Biol. Evol. **17**, 601–615.

Inomata, N., and Yamazaki, T. (2002) Nucleotide variation of the duplicated *Amylase* genes in *Drosophila kikkawai*. Mol. Biol. Evol. **19**, 678–688.

Kirby, D. A., Muse, S. V., and Stephan, W. (1995) Maintenance of pre-mRNA secondary structure by epistatic selection. Proc. Natl. Acad. Sci. USA **92**, 9047–9051.

Kovacevic, M., and Schaeffer, S. W. (2000) Molecular population genetics of X-linked genes in *Drosophila pseudoobscura*. Genetics **156**, 155–172.

Kreitman, M. (1983) Nucleotide polymorphism at the *Alcohol dehydrogenase* locus of *Drosophila melanogaster*. Nature **304**, 412–417.

Kreitman, M., and Hudson, R. R. (1991) Inferring the evolution-

ary histories of the *Adh* and *Adh-dup* loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. Genetics **127**, 565–582.

Kreitman, M. E., and Aguade, M. (1986) Excess polymorphism at the *Adh* locus in Drosophila melanogaster. Genetics **114**, 93–110.

Lunan, K. D., and Mitchell, H. K. (1969) The metabolism of tyrosine-o-phosphate in *Drosophila*. Arch. Biochem. Biophys. **132**, 450–456.

McDonald, J. H. (1998) Improved tests for heterogeneity across a region of DNA sequence in the ratio of polymorphism to divergence. Mol. Biol. Evol. **15**, 377–384.

McDonald, J. H., and Kreitman, M. (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. Nature **351**, 652–654.

Moriyama, E. N., and Powell, J. R. (1996) Intraspecific nuclear DNA variation in *Drosophila*. Mol. Biol. Evol. **13**, 261–277.

Otto, S. P. (2000) Detecting the form of selection from DNA sequence data. Trends. Genet. **16**, 526–529.

Rozas, J., and Rozas, R. (1999) DnaSP version 3: An integrated program for molecular population genetics and molecular evolution analysis. Bioinformatics **15**, 174–175.

Schaeffer, S. W., and Miller, E. L. (1993) Estimates of linkage disequilibrium and the recombination parameter determined from segregating nucleotide sites in the *Alcohol dehydrogenase* region of *Drosophila pseudoobscura*. Genetics **135**, 541–552.

Smith, J. M., and Haigh, J. (1974) The hitch-hiking effect of a favourable gene. Genet. Res. **23**, 23–35.

Tajima, F. (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. Genetics **123**, 585–595.

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., and Higgins, D. G. (1997) The Clustal_X windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. **25**, 4876–4882.

Vieira, J., and Charlesworth, B. (1999) X chromosome DNA variation in *Drosophila virilis*. Proc. R. Soc. Lond. B. Biol. Sci. **266**, 1905–1912.

Wright, T. R., Bewley, G. C., and Sherald, A. F. (1976) The genetics of dopa decarboxylase in *Drosophila melanogaster*. II. Isolation and characterization of dopa-decarboxylase-deficient mutants and their relationship to the alpha-methyl-dopa-hypersensitive mutants. Genetics **84**, 287–310.

Zhang, Z., Inomata, N., Cariou, M. L., Da Lage, J. L., and Yamazaki, T. (2003) Phylogeny and the evolution of the *Amylase* multigenes in the *Drosophila montium* species subgroup. J. Mol. Evol. **56**, 121–130.