Genomic Background Drives the Divergence of Duplicated *Amylase* Genes at Synonymous Sites in *Drosophila*

Ze Zhang*† and Hirohisa Kishino*

*Laboratory of Biometrics and Bioinformatics, Graduate School of Agriculture and Life Sciences,
University of Tokyo, Bunkyo-ku, Tokyo, Japan; †Institute for Bioinformatics Research and Development (BIRD),
Japan Science and Technology Agency (JST), Kawaguchi, Saitama, Japan

In some *Drosophila* species, there are two types of greatly diverged *amylase* (*Amy*) genes (*Amy* clusters 1 and 2), each encoding active *amylase* isozymes. Cluster 1 is located at the middle of its chromosomal arm, and the region has a normal local recombination rate. However, cluster 2 is near the centromere, and this region is known to have a reduced recombination rate. Although nonsynonymous substitutions follow a molecular clock, synonymous substitutions were accelerated in cluster 2 after gene duplications. This resulted in a higher GC content at the third codon position (GC3) and codon usage bias in cluster 1, and lower GC3 content and codon usage bias in the cluster 2. However, no systematic difference in GC content was observed in the first and second codon positions or the 3′-flanking regions. Therefore, differences in local recombination rate rather than mutation bias might explain the divergence at synonymous sites between the two *Amy* clusters within species (Hill-Robertson effect). Alternatively, the different patterns and levels of expression between the two clusters may imply that the reduced expression level in cluster 2 caused by chromatin potentiation decreased the codon bias. Both of these hypotheses imply the importance of the genomic background as a driving force of divergence between non-tandemly duplicated genes.

Introduction

Genomic sequencing has revealed that a high degree of sequence redundancy is very common in the genomes of most organisms (Rubin et al. 2000). Understanding the evolutionary mechanism of duplicated genes is therefore important for evolutionary genomics and systematic biology. The classical model of duplicated gene evolution holds that duplication creates two fully overlapping, redundant paralogous functional genes (Ohno 1970). Because of its functional redundancy, one paralog will tend to accumulate deleterious mutations, and these ultimately will be lost (Ohno 1970; Lynch and Conery 2000). Alternatively, but less likely, one of the duplicates might gain new functions, while the other paralog maintains its ancestral function. The classical model predicts the rapid loss of paralogs. Nevertheless, genomic data show that most duplicated genes have been preserved (Rubin et al. 2000). The model of subfunctionalization (Force et al. 1999) and the idea of non-neutrality for both duplicates (Kondrashov et al. 2002) explain the high level of preservation of duplicated genes in a genome. However, little is known about the evolutionary forces that drive their divergence.

Eukaryotic genomes are not uniform in recombination and mutation rates (Wolfe, Sharp, and Li 1989; Hey and Kliman 2002), and these affect the evolutionary rates and patterns of genes (Stephan and Langley 1989; Takano-Shimizu 1999, 2001; Munte, Aguade, and Segarra 2001). Furthermore, local recombination and mutation rates may or may not vary between the two duplicated gene copies. At one extreme, tandemly repeated genes may have a similar genomic background and be likely to evolve via concerted activity. At the other extreme, duplicated genes

located far apart may have very different genomic backgrounds and experience very different evolutionary processes. Therefore, genomic background factors such as local recombination and mutation rates may predict the fates of the latter kind of recently duplicated genes.

Thornton and Long (2002) found that the average ratio of nonsynonymous to synonymous substitutions between duplicated genes on the X chromosome is significantly higher than the genome average in *Drosophila melanogaster*, implying that genomic locations affect the divergence between duplicated genes. Based on their survey for new retrogenes and the functionality and evolution of those genes, they found that there is a significant excess of retrogenes from the X chromosome that retropose to autosomes. Moreover, most X-derived autosomal retrogenes have evolved a testicular expression pattern (Betran, Thornton, and Long 2002). These observations may be explained by natural selection favoring those new retrogenes that moved to autosomes and thus avoided X inactivation; they also suggest the importance of genome position for the origin of new genes.

The *Amy* genes of *Drosophila* constitute a relatively small multigene family with two to seven members in different species. Inomata and Yamazaki (2000) first found that in *D. kikkawai* and its sibling species there are two divergent *Amy* gene clusters, each encoding for active isozymes. Cluster 1 is in the middle of the B arm of chromosome 2, thought to be a region with a normal recombination rate, and cluster 2 is near the centromere, a region with reduced recombination (Ashburner 1989). The two clusters exhibit significant divergence at synonymous sites and different expression levels and patterns (Inomata and Yamazaki 2000). Similar observations were reported in *D. ananassae* (Da Lage, Maczkowoak, and Carious 2000). Zhang et al. (2002) showed that the difference in GC3 content at synonymous sites between clusters 1 and 2 was caused primarily by the changes in selection intensity immediately after gene duplication in the *montium* subgroup. Based on analyses of the coding

and 3′-flanking regions of the extended *Amy* gene sequences, we show here that the difference in local recombination rate rather than mutation bias has contributed significantly to the divergence at synonymous sites between clusters 1 and 2 in *Drosophila* species. Alternatively, the different patterns and levels of expression between the two clusters might be caused by chromatin potentiation, and this may explain the decreased codon bias in clusters 2. Both of these hypotheses suggest that genomic background has had a significant effect on the divergence of non-tandemly duplicated genes.

Materials and Methods

Twenty-eight complete *Amy* gene sequences found in the two clusters within each species for the *melanogaster* group, and four from the *obscura* group were retrieved from GenBank according to previous studies (Da Lage, Maczkowoak, and Carious 2000; Inomata and Yamazaki 2000; Zhang et al. 2002). To infer the evolutionary forces of divergence of duplicated genes, sequences without apparent expression information were ignored in this study. The *Amy* genes code for 494 amino acids (1,482 nucleotides). Only the *Amy*4N and *Amy*i5 genes in *D. ananassae* have an additional amino acid (Arg) in the signal peptide that encompasses the first 18 amino acids (Da Lage, Maczkowoak, and Carious 2000). After this additional amino acid had been removed, the length of sequences analyzed in this study was 494 amino acids.

As there is a great difference in GC3 content between the two types of *Amy* genes (Da Lage, Maczkowoak, and Carious 2000; Inomata and Yamazaki 2000; Zhang et al. 2002), the sequences of the first and second codon positions were used for phylogenetic analysis to reduce the effects of compositional bias on phylogenetic reconstruction. Neighbor-Joining (NJ), maximum parsimony (MP), and maximum likelihood (ML) methods, implemented in PAUP* 4.0 (Swofford 2001), were used for phylogenetic analysis. NJ analyses were carried out using the JC69, K80, and TN93 distances to examine their effects on topological stability. A heuristic tree search under parsimony was performed using the tree-bisection-recombination (TBR) swapping algorithm. Maximum likelihood trees were generated under the general time-reversible (GTR) model of evolution with a discrete gamma model (dγ) allowing for four categories of rate variation among sites (Swofford 2001). Heuristic searches under the ML optimality criterion were conducted using an MP starting tree and an NNI branch-swapping algorithm. The accuracy of the tree topology was assessed by bootstrap analysis, with 1,000 resampling replicates for the MP and NJ methods and 100 replicates for the ML method.

To test for differences in evolutionary rate between the two types of the *Amy* genes, the distance-based method of Li and Bousquet (1992), implemented in RRTree (Robinson-Rechavi and Huchon 2000), and a likelihood ratio test method of Muse and Gaut (1994), implemented in Hy-Phy (Muse and Pond 2000), were used for relative rate tests. The tests were applied to synonymous and nonsynonymous substitution rates separately.

Results and Discussion

Because the topologies of NJ, MP, and ML trees reconstructed by using the first and second codon positions were similar in overall structure, we show only the ML tree (fig. 1). *D. ananassae* and the *montium* subgroups belong to the *melanogaster* group, which is a sister of the *obscura* group. The phylogeny in figure 1 is consistent with the classical classification and previous results (Inomata, Tachida, and Yamazaki 1997). There are clearly two paralogous *Amy* clusters both in *D. ananassae* and in the *montium* subgroup, as indicated previously (Da Lage, Maczkowoak, and Carious 2000; Inomata and Yamazaki 2000; Zhang et al. 2002). The two *Amy* clusters within species have diverged in sequence, especially at synonymous sites, and have different GC3 contents. Clusters 1 and 2 differ in GC3 content by about 10% in *D. ananassae* and by 18% in the *montium* subgroup. The scaled chi-squared value of codon usage bias in cluster 1 is about twice that in cluster 2 within *D. ananassae* and within the *montium* subgroup (table 1). However, the average values of GC content at the first and second codon positions (GC12) are almost the same for all gene clusters (table 1), suggesting the divergence of base composition only at the third codon positions.

Molecular studies have indicated that there is one *Amy* gene cluster with three copies in the *obscura* species (Brown, Aquadro, and Anderson 1990; Steinemann and Steinemann 1999). Therefore, four *Amy* sequences from the *obscura* group species were used as the outgroups for the relative rate test to examine significant differences in evolutionary rate between the two clusters within species. The results of the relative rate test implemented in RRTree (Robinson-Rechavi and Huchon 2000) indicated that, after gene duplication, synonymous substitutions in clusters 2—with lower GC3 contents—of both *D. ananassae* and the *montium* subgroup species were significantly accelerated in comparison with the corresponding clusters 1—with higher GC3 contents. However, there are no significant differences in the substitution rate of amino acids between the two clusters in *D. ananassae* or in the *montium* subgroup (table 2). These results were substantiated by a likelihood ratio test of Muse and Gaut (1994), which tests rate constancy between the two sequences with a third outgroup sequence. Using the *miranda Amy* 1 as the outgroup, significant differences in the rates of synonymous substitutions were detected in all the pairs between the two sequences from the two gene clusters within species. The smallest log-likelihood difference in all the comparisons was 4.1, with $P <$ 0.05 (1 df). However, no significant differences in the rate of nonsynonymous substitutions were found between any pair.

The classical model predicts that one of two gene copies will evolve neutrally or under less functional constraint immediately after gene duplication (Ohno 1970). This relaxation of selection or functional constraint should have the same effect on both synonymous sites and nonsynonymous sites. Accordingly, there should be a corresponding acceleration of nonsynonymous substitutions. However, we did not observe such acceleration in either of the two clusters (table 2). This is contrary to the prediction from the classical model.
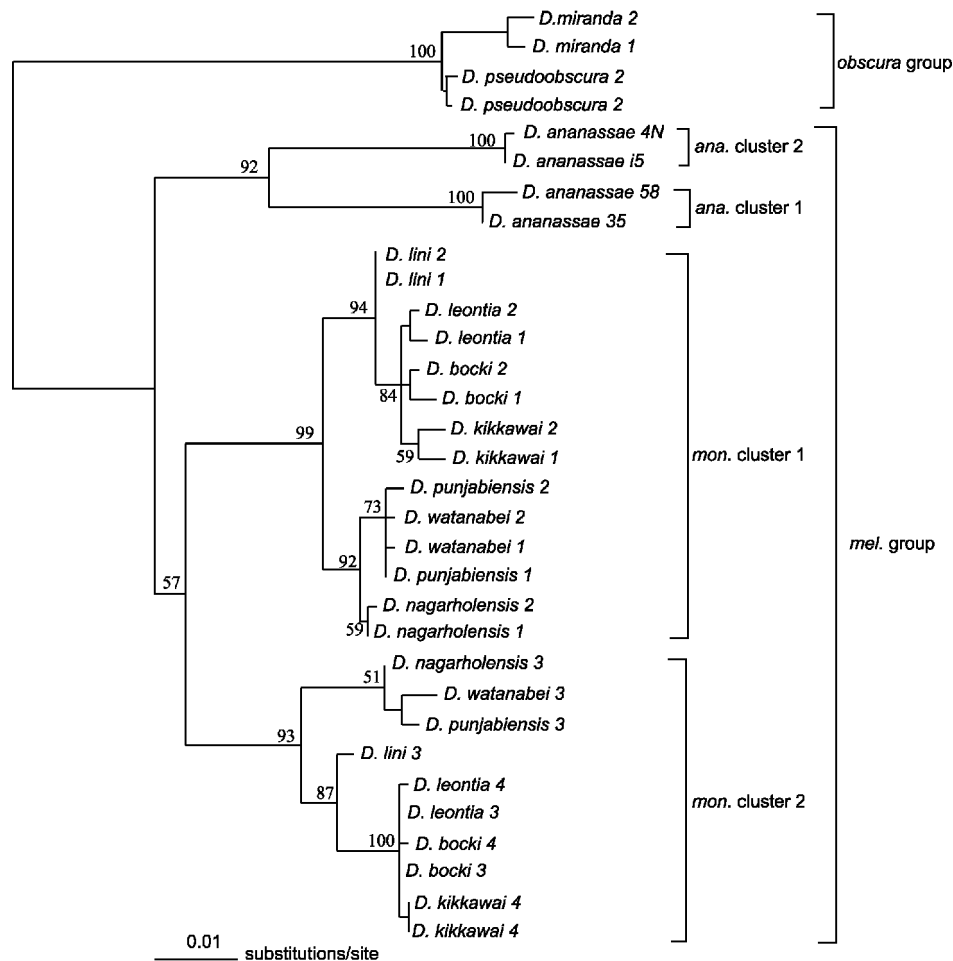
FIG. 1.—Gene tree representing relationships among the *Amy* genes studied in *Drosophila*. The maximum-likelihood tree (−lnL = 2290.71) reconstructed by the first and second codon positions is presented; branch lengths were optimized with likelihood using the GTR + dγ model of evolution. The numbers at internal nodes are bootstrap probabilities. Values below 50% are not shown. The accession numbers are AB035055–035069 and AB078765–078773 for 24 *Amy* gene sequences in the *montium* subgroup, U53698 for *Amy* 58 and *Amy* 38, and U53477–53478 for *Amy* 4N and *Amy* i5 in *D. ananassae*; Y15603–15604 for *D. miranda Amy* 1 and 2, and X76240–76241 for *D. pseudoobscura Amy* 1 and 2.

The *Amy* genes with lower GC3 contents are located in the regions near the centromeres of chromosomes 3 and 2 in *D. ananassae* (Da Lage, Maczkowoak, and Carious. 2000) and in *D. kikkawai* (Inomata and Yamazaki 2000). These regions have reduced local recombination rates (Ashburner 1989). In contrast, *Amy* genes with higher GC3 contents are located on the center of the arm of chromosome 2, and they have a normal local recombination rate (Ashburner 1989). This suggests that the local recombination rate affected nucleotide divergence between the two *Amy* gene clusters within species. Based on the pattern of polymorphism and divergence at synonymous sites, synonymous substitutions have been found to be subject to weak selection against major and non-major codons (Akashi 1995). Increasing evidence suggests that natural selection acts on synonymous sites in genes of *Drosophila* (Takano-Shimizu 1999; Munte, Aguade, and Segarra 2001). The results in table 2 indicate that the lower local recombination rate relaxed the selection constraint on synonymous substitutions in cluster 2 in the different subgroups because of the Hill–Robertson effect (Hill and Robertson 1966).

It should be pointed out that the differences in recombination rate between clusters 1 and 2 were qualitatively inferred based on their locations on particular chromosomes. Local recombination rates of orthologous regions may vary among the genomes of related species. Takano-Shimizu (1999) observed differences in the GC3 contents of the *yellow* gene between closely related species of *Drosophila* and experimentally suggested a difference in local recombination rates as a potential cause. In the present analysis, however, GC3 contents were very similar within clusters 1 and 2 (table 1). Thus, we have no reason to expect large variations in local recombination rates within the clusters.

Local mutation bias may explain the divergence between the two *Amy* clusters within species. However, Zhang et al. (2002) examined the GC contents of the introns and the 5′-flanking nucleotide sequences, and they found no difference in mutation bias between clusters 1 and 2 in the *montium* subgroup. It may be argued that the intron and 5′-flanking sequences are under some selection constraint. The unique short intron has ~50% of the sequence corresponding to elements for the splicing

Table 1
Average GC Contents and Codon Bias for Different *Amy* Gene Clusters

Gene Cluster[a]	GC12 (%)[b]	GC3 (%)[b]	3′-Flanking GC (%)[c]	Scaled Chi-Square
The *montium Amy* cluster 1 (14)	50.21 (0.25)	88.73 (1.89)	28.83 (9)	0.959
The *montium Amy* cluster 2 (10)	49.38 (0.20)	69.83 (1.88)	32.81 (8)	0.467
Ananassae Amy cluster 1 (2)	49.19 (−)	73.90 (−)	36.75 (2)	0.587
Ananassae Amy cluster 2 (2)	49.24 (−)	63.10 (−)	32.51 (2)	0.330
The *obscura Amy* cluster (4)	50.05 (0.06)	87.60 (0.25)	46.85 (3)	1.036

NOTE.—Scaled chi-square defined by Shields et al. (1988) was used to measure codon bias; GC12: GC content at the first and second codon positions; GC3: GC content at (synonymous) third codon positions.

[a] The numbers in parentheses are the numbers of sequences analyzed in each gene cluster.

[b] The numbers in parentheses are the standard deviations for each gene cluster.

[c] The numbers in parentheses are the numbers of sequences analyzed in each gene cluster when 3′-flanking nucleotide sequences are available. The lengths of the sequences range from 121 to 400 bp.

reaction (Mount et al. 1992) and the 5′-flanking region harbors the promoter and other regulatory elements of gene expression. We therefore examined the 24 3′-flanking nucleotide sequences available from GenBank. The sequence lengths varied from 121 bp to 400 bp. Although the alignment columns without gaps are too short to estimate the detailed phylogenetic relations, the clusters formed groups consistent with the coding regions (results not shown). In contrast to the coding regions, the average GC content in the 3′-flanking region of the *Amy* gene cluster 1 is relatively smaller than that of the *Amy* gene cluster 2 in the *montium* subgroup (table 1). In *D. ananassae*, the GC contents of the 3′-flanking regions are 45.75 and 27.75 for *Amy35* and *Amy58* of cluster 1, respectively, and 28.26 and 36.75 for *Amy4N* and *Amyi5* of cluster 2, respectively. The four *D. ananassae Amy* genes have 3′-flanking regions with a shared length of 400 bp. There is large GC content variation in these regions compared with the coding regions. There is no evidence that mutation bias has shaped the composition patterns of the two *Amy* clusters within species. Therefore, in the case of the *Amy* gene family in *Drosophila*, local recombination rate may be an important factor in the genomic background. By comparing orthologous sequences in *Drosophila* species, Takano-Shimizu (2001) observed positive correlation in GC content between coding and noncoding regions. However, we did not find such a correlation in the comparison of paralogous genes. Mutation bias explains species-specific GC contents, but location effect via local recombination rate makes a large contribution to the divergence between duplicated genes.

Local recombination rates affect natural selection through changes in the effective population size. Thus, changes in the effective population size should affect all types of substitutions in genes. However, changes in local recombination rate do not seem to have affected the corresponding divergence at amino acid level between the two *Amy* clusters within species (table 2). Alpha-amylase plays a major role in the digestive processes of carbohydrates by hydrolyzing starch from food substrates into smaller sugars, such as maltose and glucose. Both *Amy* clusters are active and expressed (Da Lage, Maczkowoak, and Carious 2000; Inomata and Yamazaki 2000; Zhang et al. 2002). On the one hand, strongly purifying selection might prevent changes to amino acids. On the other hand, this implies that most possible replacement substitutions are deleterious, and suggests that rates of amino acid replacement are insensitive to differences in the effective population size of the *Amy* gene region. Similar results have been observed in a study on the *y* gene in *Drosophila* (Munte, Aguade, and Segarra 2001). In addition, Zeng et al. (1998) inferred that the rates of amino acid replacement in *Drosophila* were not overdispersed. One possibility is that effective population sizes in *Drosophila* are large enough for most non-synonymous mutations to be effectively deleterious, and therefore they do not become fixed (Zeng et al. 1998).

It is also of interest that the codon usage bias of one gene is positively related to its expression level (Shields et al. 1988). As indicated in previous studies (Inomata and Yamazaki 2000; Da Lage, Maczkowoak, and Carious 2000; Zhang et al. 2002), clusters 1 and 2 have different patterns and levels of expression, cluster 2 being expressed less than cluster 1. Changes in codon usage bias might therefore be caused by changes in expression level. In cluster 2, a decrease of expression level has led to

Table 2
Results for Relative Rate Tests

Gene Cluster 1	Gene Cluster 2	K_{O1}	K_{O2}	Z	Comparison for Evolutionary Rates
Synonymous substitutions					
D. ananassae Amy cluster 1	*D. ananassae Amy* cluster 2	0.5708	0.8640	3.2513*	O2 > O1
D. montium subgroup *Amy* cluster 1	*D. montium* subgroup *Amy* cluster 2	0.4645	0.6326	2.7282*	O2 > O1
Nonsynonymous substitutions					
D. ananassae Amy cluster 1	*D. ananassae Amy* cluster 2	0.0740	0.0762	0.3478	O2 = O1
D. montium subgroup *Amy* cluster 1	*D. montium* subgroup *Amy* cluster 2	0.0714	0.0706	0.1491	O2 = O1

NOTE.—The *Amy* gene cluster of the *obscura* group in figure 1 was used as an outgroup.

* Significance of Z-test at 1% probability level.

a decrease in selection on codon usage bias. Consequently, codon usage bias has gone down and synonymous substitution rate has gone up (table 1). Two factors might explain this variation. One possibility is changes to the regulatory elements such as *cis*-sequences or *trans*-acting elements. In fact, cluster 2 in general has lost some *cis*-regulatory elements (Da Lage, Maczkowoak, and Carious 2000; Inomata and Yamazaki 2000; Zhang et al. 2002), which might be caused by relaxation of purifying selection in a lower recombination rate region. The other possibility lies in chromosomal domains of expression, such as chromatin potentiation in higher eukaryotes (Kramer et al. 1998; Boutanaev et al. 2002), because cluster 2 is near the centromere and cluster 1 is in the middle of its chromosomal arm; they therefore probably have different chromatin structures. It is not clear which of these scenarios is true. However, all the evidence suggests that genomic background drives the divergence between the two *Amy* clusters within *Drosophila* species, although it may act through different mechanisms.

Thus, the change of selection intensity triggered by genomic background along a genome seems to be the most general model that can account for the divergence at synonymous sites but not at amino acid level between the two *Amy* gene clusters within *Drosophila* species. With these clues from the evolution of the *Amy* gene family in *Drosophila*, together with emerging evidence from genomic data (see review in Kondrashov et al. 2002), it is the most likely that duplicated genes are not redundant from the start because of selection for increased dosage (Grauer and Li 1999; Force et al. 1999). Thus, after duplication, either copy should be subject to purifying selection (Kondrashov et al. 2002). Under certain selection pressures, the fates of duplicate genes will depend on their genomic backgrounds. Concerted evolution between tandemly repeated genes is also consistent with the genomic background hypothesis, because they share a similar genomic background (Grauer and Li 1999). In contrast, if two duplicate genes are located in different regions, the difference in genomic backgrounds will be a driving force for divergence. Because eukaryotic genomes are heterogeneous in recombination and mutation rates and in chromatin potentiation, this heterogeneity will accelerate the divergence of non-tandemly duplicated genes.

Acknowledgments

Literature Cited

Akashi H. 1995. Inferring weak selection from patterns of polymorphism and divergence at ''silent'' sites in Drosophila DNA. Genetics **139**:1067–1076.

Ashburner, M. D. 1989. Drosophila. A laboratory handbook. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Betran, E., K. Thornton, and M. Long. 2002. Retroposed new genes out of the X in Drosophila. Genome Res. **12**: 1854–1859.

Boutanaev, A. M., A. I. Kalmykova, Y. Y. Sheveiyov, and D. I. Nurminsky. 2002. Large clusters of co-expressed genes in the Drosophila genome. Nature **420**:666–669.

Brown, C. J., C. F. Aquadro, and W. W. Anderson. 1990. DNA sequence evolution of the amylase multigene family in *Drosophila pseudoobscura*. Genetics **126**:131–138.

Da Lage, J.-L., F. Maczkowoak, and M.-L. Carious. 2000. Molecular characterization and evolution of the amylase multigene family of *Drosophila ananassae*. J. Mol. Evol. **51**:391–403.

Force, A., M. Lynch, B. Pickett, A. Amores, Y.-L. Yan et al. 1999. Preservation of duplicate genes by complementary, degenerative mutations. Genetics **151**:1531–1545.

Grauer, D., and W.-H. Li. 1999. Fundamentals of molecular evolution, 2nd Edition. Sinauer Associates, Sunderland, Mass.

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

Hey, J., and R. M. Kliman. 2002. Interactions between natural selection, recombination and gene density in the genes of Drosophila. Genetics **160**:595–608.

Inomata, N., H. Tachida, and T. Yamazaki. 1997. Molecular evolution of the Amy multigenes in the subgenus *Sophophora* of *Drosophila*. Mol. Biol. Evol. **14**:942–950.

Inomata, N., and T. Yamazaki. 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.

Kondrashov, F. A., I. B. Rogozin, Y. I. Wolf, and E. V. Koonin. 2002. Selection in the evolution of gene duplications. Genome Biol. **3**:research 0008.1–0008.9.

Kramer, J. A., J. R. McCarrey, D. Djakiew, and S. A. Krawetz. 1998. Differentiation: the selective potentiation of chromatin domains. Development **125**:4749–4755.

Li, P., and J. Bousquet. 1992. Relative-rate test for nucleotide substitutions between two lineages. Mol. Biol. Evol. **9**:1185–1189.

Lynch, M., and J. S. Conery. 2000. The evolutionary fate and consequences of duplicate genes. Science **290**:1151–1155.

Mount, S. M., C. Burks, G. Hertz, G. D. Stormo, O. White, and C. Fields. 1992. Splicing signals in Drosophila: intron size, information content, and consensus sequences. Nucleic Acids Res. **20**:4255–4262.

Munte, A., M. Aguade, and C. Segarra. 2001. Changes in the recombinational environment affect divergence in the yellow gene of Drosophila. Mol. Biol. Evol. **18**:1045–1056.

Muse, S. V., and B. S. Gaut. 1994. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates with application to the chloroplast genome. Mol. Biol. Evol. **11**:715–724.

Muse, S. V., and S. K. Pood. 2000. Hy-Phy user manual. North Carolina State University, Raleigh; University of Arizona, Tucson.

Ohno, S. 1970. Evolution by Gene Duplication. Springer-Verlag, Heidelberg.

Robinson-Rechavi, M., and D. Huchon. 2000. RRTree: relative-rate tests between groups of sequences on a phylogenetic tree. Bioinformatics **16**:296–297.

Rubin, G. M., M. D. Yandell, J. R. Wortman et al. (50 co-authors). 2000. Comparative genomics of the eukaryotes. Science **287**:2204–2215.

Shields, D. C., P. M. Sharp, D. G. Higgins, and F. Wright. 1988. ''Silent'' sites in Drosophila genes are not neutral: evidence of

selection among synonymous codons. Mol. Biol. Evol. **5**:704–716.

Steinemann, S., and M. Steinemann. 1999. The amylase gene cluster on the evolving sex chromosomes of *Drosophila miranda*. Genetics **151**:151–161.

Stephan, W., and C. H. Langley. 1989. Molecular genetic variation in the centromeric region of the X chromosome in three *Drosophila ananassae* populations. 1. Contrasts between the *vermilion* and *forked* loci. Genetics **121**:89–99.

Swofford, D. L. 2001. PAUP*: phylogenetic analysis using parsimony (*and other methods), Version 4. Sinauer Associates, Sunderland, Mass.

Takano-Shimizu, T. 1999. Local recombination and mutation effects on molecular evolution. Genetics **153**:1285–1296.

Takano-Shimizu, T. 2001. Local changes in GC/AT substitution biases and in crossover frequencies on Drosophila chromosomes. Mol. Biol. Evol. **18**:606–619.

Thornton, K., and M. Long. 2002. Rapid divergence of gene duplicates on the *Drosophila melanogaster* X chromosome. Mol. Biol. Evol. **19**:918–925.

Wolfe, K. H., P. M. Sharp, and W.-H. Li. 1989. Mutation rates differ among regions of the mammalian genome. Nature **337**:283–285.

Zeng, L.-W., J. M. Comeron, B. Chen, and M. Kreitman. 1998. The molecular clock revisited: the rate of synonymous vs. replacement change in Drosophila. Genetica **102/103**:369–382.

Zhang, Z., N. Inomata, T. Ohba, M.-L. Cariou, and T. Yamazaki. 2002. Codon bias differentiates between the duplicated *amylase* Loci following gene duplication in *Drosophila*. Genetics **161**:1187–1196.