

# Note

## Genomic Background Predicts the Fate of Duplicated Genes: Evidence From the Yeast Genome

Ze Zhang<sup>\*,†,1</sup> and Hirohisa Kishino<sup>\*</sup>

<sup>\*</sup>Laboratory of Biometrics and Bioinformatics, Graduate School of Agriculture and Life Sciences, University of Tokyo, Tokyo 113-8657, Japan  
and <sup>†</sup>Institute for Bioinformatics Research and Development, Japan Science and Technology Agency, Tokyo 102-0081, Japan

Manuscript received September 24, 2003  
Accepted for publication December 31, 2003

### ABSTRACT

Gene duplication with subsequent divergence plays a central role in the acquisition of genes with novel function and complexity during the course of evolution. With reduced functional constraints or through positive selection, these duplicated genes may experience accelerated evolution. Under the model of subfunctionalization, loss of subfunctions leads to complementary acceleration at sites with two copies, and the difference in average rate between the sequences may not be obvious. On the other hand, the classical model of neofunctionalization predicts that the evolutionary rate in one of the two duplicates is accelerated. However, the classical model does not tell which of the duplicates experiences the acceleration in evolutionary rate. Here, we present evidence from the *Saccharomyces cerevisiae* genome that a duplicate located in a genomic region with a low-recombination rate is likely to evolve faster than a duplicate in an area of high recombination. This observation is consistent with population genetics theory that predicts that purifying selection is less effective in genomic regions of low recombination (Hill-Robertson effect). Together with previous studies, our results suggest the genomic background (*e.g.*, local recombination rate) as a potential force to drive the divergence between nontandemly duplicated genes. This implies the importance of structure and complexity of genomes in the diversification of organisms via gene duplications.

GENE duplication has long been thought to be one of the principal engines powering the evolution of protein function, allowing for increases in genomic complexity (OHNO 1970). Because a duplication event creates two fully functional overlapping copies, one of the two paralogs may evolve in a neutral manner right after duplication due to functional redundancy. In most cases, this duplicated gene accumulates deleterious loss-of-function mutations and becomes a pseudogene. There is a small chance that directional selection will lead it to acquire some novel function (OHNO 1970). However, this classical model predicts only an equal chance of evolutionary rate acceleration for each duplicate. Unless there is a selection pressure for overexpression of the genes, the two genes with identical functions cannot be stably retained in the genome (NOWAK *et al.* 1997). FORCE *et al.* (1999) explained the selective maintenance of both paralogs by subfunctionalization. The model suggests that each daughter gene adopts part of the function of the parental gene. That is, each developing member of the gene family acquires speci-

ficity in gene expression after duplication by the complementary loss of their *cis*-regulatory elements in each of the copies (FORCE *et al.* 1999). Subfunctionalization can also occur at the protein function level and can lead to functional specialization when one of the duplicate genes becomes better at performing one of the original functions of the parental gene (HUGHES 1999; ZHANG 2003).

Increasing evidence from genomic data indicates that asymmetric sequence divergence of duplicate genes is quite common. KONDRASHOV *et al.* (2002) analyzed 39 genomes from eubacteria, archaea, and eukaryotes and found asymmetric divergence in ~5% of the 101 duplicated gene pairs analyzed, whereas VAN DE PEER *et al.* (2001) found that half of the 26 duplicated gene pairs in zebrafish showed evidence of asymmetric divergence. More recently, a study by ZHANG *et al.* (2003) examined the evolutionary patterns for 250 duplicated gene pairs in the human genome by using the corresponding orthologs from the mouse genome as the outgroups and found that nearly 60% of duplicated pairs have evolved in an asymmetric divergent manner at the amino acid level. Moreover, CONANT and WAGNER (2003) found that 20–30% of duplicated gene pairs from four completely sequenced eukaryotic genomes showed asym-

<sup>1</sup>Corresponding author: School of Biosciences, University of Birmingham, Birmingham B15 2TT, United Kingdom.

metric evolution at the amino acid level. Although the classical model of neofunctionalization predicts acceleration in the evolutionary rate in one of the two duplicates, the model does not indicate which of the duplicates experiences the acceleration. In fact, it assumes implicitly an equal chance of acceleration between the two copies. Mounting evidence indicates that most duplicated genes are not redundant from the start because of selection for increased dosage (FORCE *et al.* 1999; GRAUR and LI 1999; KONDRASHOV *et al.* 2002). If this is a rule rather than an exception, both copies are subject to purifying selection after duplication (KONDRASHOV *et al.* 2002). As purifying selection intensity is stronger in a high-recombination environment (HILL and ROBERTSON 1966; CARVALHO and CLARK 1999; COMERON *et al.* 1999), we would expect that the rate of protein evolution should be lower in regions of high recombination.

In this article, we compare evolutionary rates for the duplicated genes with different local recombination rates by the analysis of the *Saccharomyces cerevisiae* genome. It is well known that the arrayed structure of duplicated genes in the genome also affects their evolutionary pattern. Tandemly duplicated genes usually evolve in a concerted manner due mainly to unequal crossover and gene conversion mechanisms (LI 1997). In contrast, nontandemly duplicated genes are less affected by gene conversion (DROUIN 2002) and evolve in a divergent manner (ROONEY *et al.* 2002). Thus, this study into the evolution of nontandemly duplicated genes may provide us with some clues for understanding the evolution of new function through gene duplication and subsequent divergence. Our results examining nontandemly duplicated genes clearly show that the acceleration of evolutionary rate in one copy depends on the local recombination rate. This also implies that genomic background may drive the fate of the nontandemly duplicated genes.

The data set used in this article consists of the *S. cerevisiae* gene duplication database (<http://wolfe.gen.tcd.ie/>). The Smith-Waterman algorithm was used to search for paralogs. As most duplicate genes occurred a long time ago [ $\sim 100$  million years ago (MYA)], resulting from whole-genome duplication in *S. cerevisiae* (WOLFE and SHIELDS 1997), it is important to focus on duplicated genes with a high signal-to-noise ratio. To minimize the potential effect of different amino acid compositions, gene lengths, or selective constraints on the relation between recombination and evolutionary rate, a very stringent threshold with an expectation value of zero was chosen (PEARSON 1991; J. A. BIRDSELL, personal communication). Of the detected paralogs, we selected 47 sets of paralogous genes, 1 of which is in a high-recombination locus and the other in a low-recombination locus (GERTON *et al.* 2000; BIRDSELL 2002). Most genes had only one paralog; however, a few had multiple paralogs. When there was more than one high-

or low-recombination locus within a gene family, one was randomly selected for further analysis.

We then searched for the corresponding orthologous sequence for each of the selected paralogs from the *Candida albicans* genome data as a reference for the relative rate test between nontandem paralogs after gene duplication (Figure 1; WU and LI 1985). Because *S. cerevisiae* and *C. albicans* diverged  $\sim 140$ –330 MYA (SEOIGHE *et al.* 2000), it should be reasonable to use the ortholog from *C. albicans* as a reference for the relative rate test. The *C. albicans* genomic data (contig version 6) were downloaded from <http://www-sequence.stanford.edu/group/candida/>. We used two paralog sequences of each pair as queries to carry out the FASTA amino acid sequence similarity searches against *C. albicans* genomic data. The ortholog that the two copies had the same best hits and that showed  $>40\%$  amino acid identity to the two copies was selected as the outgroup, because outgroup with shorter branch lengths yields more trustworthy divergence estimates (CONANT and WAGNER 2003). For a few pairs, the two copies had different best hits. Then we looked at the InParanoid database that provides orthologous sequence cluster information between organism pairs of *S. cerevisiae* and *Caenorhabditis elegans*, *Drosophila melanogaster*, *Arabidopsis thaliana*, *Mus musculus*, and *Homo sapiens* (REMM *et al.* 2001) to find the appropriate outgroups. Finally, we found 30 duplicate pairs that have proper outgroups.

All the protein alignments were produced using CLUSTAL X (THOMPSON *et al.* 1994). Sequence alignments were edited with BioEdit (HALL 1999) and only unambiguously aligned regions were used for further analysis. Finally, the RRTree was applied to all the edited alignments that contain the appropriate outgroups for the relative rate tests (ROBINSON-RECHAVI and HUCHON 2000). All the alignments are available from the authors upon request.

Because most duplicate gene pairs are very old, synonymous substitutions between the two copies are almost saturated. Therefore, we focused on the difference in the evolutionary rate of amino acid sequence between the two copies. Table 1 compares  $K_{DH}$ , the branch length of a copy at a high-recombination-rate region, with  $K_{DL}$ , that of another copy at a low-recombination-rate region (Figure 1). The average of the difference was  $-0.129$  and the  $t$  value was 2.325 (d.f. = 29), suggesting a significantly lower rate of the copies at higher-recombination-rate regions ( $P = 0.014$ ). We found that 14 of the 30 (46.6%) duplicate gene pairs showed asymmetry in the evolutionary rate of amino acid sequence (Table 1). Among the 14 duplicated gene pairs showing asymmetric evolution, 11 experienced an acceleration in the evolutionary rate of the copy in the region of low recombination (Table 1). This proportion is significantly higher than one-half ( $\chi^2 = 4.28$ ,  $P < 0.05$ ). This evidence that the paralog in a region of low recombination rate is more likely to evolve faster than the paralog in

TABLE 1

Comparison of the rates of evolution between the copies in high-recombination regions and those in low-recombination regions

High recombination <sup>a</sup>	Low recombination <sup>a</sup>	Outgroup	K		High vs. low recombination
			K <sub>DH</sub>	K <sub>DL</sub>	
YAL012w (1.989)	YGL184c (0.981)	ZK1127.10 ( <i>C. elegans</i> )	0.7016	1.08	**
YAL036c (1.849)	YGR173w (1.117)	T28D6.6 ( <i>C. elegans</i> )	0.4348	0.6535	**
YAL038w (3.139)	YOR347c (1.370)	Orf6.5754 ( <i>C. albicans</i> )	0.3289	0.4525	**
YAL061w (2.081)	YAL060w (1.430)	Orf6.3669 ( <i>C. albicans</i> )	0.7965	0.7906	NS
YAL062w (1.819)	YOR375c (0.936)	Orf6.3152 ( <i>C. albicans</i> )	0.3231	0.3077	NS
YBL016w (2.081)	YGR040w (1.067)	F43C1.2a ( <i>C. elegans</i> )	0.7505	0.6827	NS
YBR241c (2.27)	YGL104c (0.90)	Orf6.5168 ( <i>C. albicans</i> )	0.8161	0.9358	*
YCL040w (1.633)	YDR516c (1.254)	Orf6.6009 ( <i>C. albicans</i> )	0.6680	0.7076	NS
YDL155w (1.447)	YLR210w (0.981)	Orf6.8086 ( <i>C. albicans</i> )	0.8947	0.8414	NS
YDR188w (2.363)	YDL143w (1.284)	F01F1.8a ( <i>C. elegans</i> )	0.5948	1.4311	**
YER145c (1.809)	YER207w (1.174)	Orf6.8119 ( <i>C. albicans</i> )	0.5417	0.8731	**
YFR024c (1.977)	YHR016c (1.32)	Orf6.2516 ( <i>C. albicans</i> )	0.8334	0.7132	>
YHR096c (1.754)	YHR094c (1.09)	Orf6.2379 ( <i>C. albicans</i> )	0.6004	0.6391	NS
YIL053w (1.664)	YER062c (0.932)	Orf6.8673 ( <i>C. albicans</i> )	0.7316	0.7316	NS
YIL118w (1.516)	YPR165w (0.96)	Orf6.4937 ( <i>C. albicans</i> )	0.7424	0.2081	≥
YIL123w (1.62)	YNL066w (1.058)	Orf6.2071 ( <i>C. albicans</i> )	0.5533	0.5533	NS
YJL026w (2.001)	YGR180c (1.48)	Orf6.6635 ( <i>C. albicans</i> )	0.2544	0.6139	**
YJL052w (1.97)	YGR192c (1.497)	Orf6.8817 ( <i>C. albicans</i> )	0.2501	0.2580	NS
YJL116c (1.27)	YKR042w (1.25)	Orf6.7977 ( <i>C. albicans</i> )	0.5360	0.5565	NS
YJR049c (1.46)	YEL041w (1.01)	Orf6.8553 ( <i>C. albicans</i> )	0.8267	0.9153	NS
YJR077c (2.103)	YER053c (1.016)	Orf6.3051 ( <i>C. albicans</i> )	0.2497	0.9708	**
YKL035w (1.594)	YHL012w (1.064)	Orf6.1349 ( <i>C. albicans</i> )	0.2402	0.9356	**
YLR188w (1.719)	YPL270w (0.874)	Orf6.4270 ( <i>C. albicans</i> )	0.7841	0.9364	**
YLR350w (2.317)	YGR038w (1.221)	Orf6.2090 ( <i>C. albicans</i> )	0.7098	0.6312	NS
YLR378c (2.18)	YBR283c (1.039)	Orf6.5379 ( <i>C. albicans</i> )	0.3924	1.197	**
YMR056c (2.43)	YBL030c (1.374)	Orf6.8179 ( <i>C. albicans</i> )	0.2845	0.1746	≥
YOL059w (2.321)	YDL021w (1.031)	Orf6.4093 ( <i>C. albicans</i> )	0.5373	0.5058	NS
YOR173w (1.943)	YLR270w (1.173)	Orf6.7165 ( <i>C. albicans</i> )	0.7168	0.6688	NS
YOR374w (1.797)	YER073w (1.011)	Orf6.6640 ( <i>C. albicans</i> )	0.4880	0.4446	NS
YPL048w (1.683)	YKL081w (1.114)	Orf6.3352 ( <i>C. albicans</i> )	0.5763	0.6180	NS

NS, not significant. \* and \*\* indicate that the duplicates with low-recombination rates evolved significantly faster than the corresponding paralogs with high-recombination rates at 5 and 1% probabilities, respectively. > and ≥ indicate that the duplicates with high-recombination rates evolved significantly faster than the corresponding paralogs with low-recombination rates at 5 and 1% probabilities, respectively.

<sup>a</sup> The numbers in parentheses are the recombination rates from GERTON *et al.* (2000).

an area of high-local-recombination rate is consistent with population genetics theory, which predicts that purifying selection is less effective in genomic regions of low recombination (HILL and ROBERTSON 1966; CARVALHO and CLARK 1999; COMERON *et al.* 1999). As most amino acid changes tend to be deleterious (LI 1997), relaxation of the purifying selection to favor the copy in areas with a low-local-recombination rate allows for their accumulation over those located in a region of high-local-recombination rate.

We did not detect a significant difference in the evolutionary rate between the two copies for 16 of the 30 (53.4%) duplicate gene pairs. The averages of  $K_{DH}$  for the groups for which two copies show evolutionary asymmetry and symmetry after duplication are 0.522 and 0.622, respectively (Figure 1 and Table 1). Although there is no significant difference in average evolutionary

rate between the two groups ( $P = 0.18$  for  $t$ -test), the group showing evolutionary symmetry has a higher average of  $K_{DH}$  than the group with the pattern of evolutionary asymmetry. The GC3 contents of copies with higher recombination rates were examined for the two groups. They had very similar average GC3 contents (0.461 and 0.465, respectively). Therefore, different evolutionary patterns of two groups do not appear to be due to base composition. Probably this is because the duplication events and divergence between *S. cerevisiae* and *C. albicans* are too old to guarantee sufficient power of the relative rate test. Alternatively, some of them may have been in the process of subfunctionalization, during which the sites of the two copies experienced complementary acceleration. As a result, the different rates were cancelled out over the sites.

WOLFE and SHIELDS (1997) found ~55 duplicate

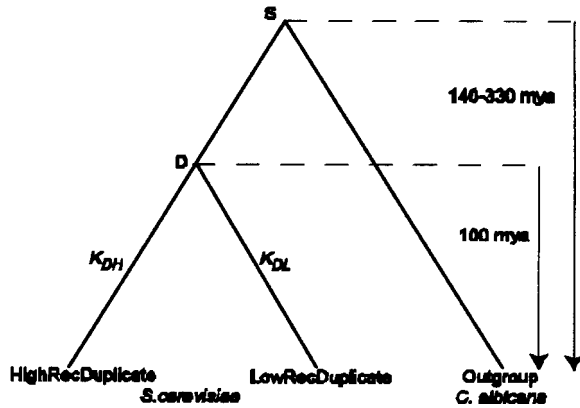


FIGURE 1.—A tree for the relative rate test to detect the difference in evolutionary rate between the high- and low-recombination duplicates in the *S. cerevisiae* genome using the ortholog from *C. albicans* as a reference. The approximate duplication time, 100 MYA, is from WOLFE and SHIELDS (1997), and the splitting time of *S. cerevisiae* and *C. albicans*, 140–330 MYA, is from SEOIGHE *et al.* (2000).

blocks in the *S. cerevisiae* genome. To contrast the above result with a negative control, we randomly selected 40 duplicate gene pairs of which two copies of each pair have similar recombination rates and come from different paralogous blocks in the *S. cerevisiae* gene duplication database (<http://wolfe.gen.tcd.ie/>). Of 40 duplicate pairs, 8 were found to show asymmetric evolution (results not shown). This proportion (20%) is much lower than that (46.6%) of duplicate pairs, which have very different recombination rates. This suggests that recombination rate may be a potential force to drive the divergence of duplicate genes.

During the past decade, much attention has been focused on the effect of recombination on evolutionary rates and patterns of genes (STEPHAN and LANGLEY 1989; TAKANO-SHIMIZU 1999, 2001; MUNTE *et al.* 2001; BIRDSELL 2002). However, we are still in a state of relative ignorance about how this affects the divergence between the nontandemly duplicated genes. Recently, BIRDSELL (2002) found a significant effect of recombination on divergence of GC (GC3) content at third codon positions between the high and low recombinational duplicates in the *S. cerevisiae* genome. He proposed a “constraint hypothesis” (a modified biased gene conversion hypothesis) that explains the observations. Selection is not directly invoked in this hypothesis.

Recently, THORNTON and LONG (2002) found that the average ratio of nonsynonymous-to-synonymous substitutions between the duplicated genes on the X chromosome is significantly higher than the genome average in *D. melanogaster*. On the basis of the survey for new retrogenes and their functionality and evolution, they further found a significant excess of retrogenes from the X chromosome that retrotranspose to autosomes. Moreover, most X-derived autosomal retrogenes have evolved a testicular expression pattern (BETRAN *et al.* 2002). These

observations may be explained by natural selection favoring those new retrogenes that moved to autosomes and thus avoided X inactivation in spermatocytes and suggest the importance of genome position for the origin of new genes. More recently, ZHANG and KISHINO (2004) presented an example in which changes in local recombination rate drove the divergence at synonymous sites between the duplicated *amylase* genes in *Drosophila*.

In this study, we found a significant effect of recombination on divergence in protein evolution between the two duplicates in different recombination rate regions. Differences in purifying selection intensity triggered by differences in local recombination rate can explain these observations. When genome data from many closely related species become available, it will be interesting to examine the site-specific asymmetry of duplicate genes (GU 1999; KNUDSEN and MIYAMOTO 2001; WANG and GU 2001; KNUDSEN *et al.* 2003). Together with previous studies, the results suggest that genomic background (*e.g.*, local recombination rate) has a potential to drive divergence between nontandemly duplicated genes, although it may act through different mechanisms. This also implies that the genomic background should be taken into account to better understand the evolution of duplicated genes.

We thank J. A. Birdsell for kindly sending us sequence data of *S. cerevisiae* duplicated genes and for helpful discussions. This study has been supported by the Institute for Bioinformatics Research and Development, Japan Science and Technology Agency.

#### LITERATURE CITED

- BETRAN, E., K. THORNTON and M. LONG, 2002 Retroposed new genes out of the X in *Drosophila*. *Genome Res.* **12**: 1854–1859.
- BIRDSELL, J. A., 2002 Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Mol. Biol. Evol.* **19**: 1181–1197.
- CARVALHO, A. B., and A. G. CLARK, 1999 Intron size and natural selection. *Nature* **401**: 344.
- COMERON, J. M., M. KREITMAN and M. AGUADE, 1999 Natural selection on synonymous site is correlated with gene length and recombination in *Drosophila*. *Genetics* **151**: 239–249.
- CONANT, G. C., and A. WAGNER, 2003 Asymmetric sequences divergence of duplicate genes. *Genome Res.* **13**: 2052–2058.
- DROUIN, G., 2002 Characterization of the gene conversions between the multigene family members of the yeast genome. *J. Mol. Evol.* **55**: 14–23.
- FORCE, A., M. LYNCH, F. B. PICKETT, A. AMORES, Y. YAN *et al.*, 1999 Preservation of duplicate genes by complementary, degenerative mutations. *Genetics* **151**: 1531–1545.
- GERTON, J. L., J. DERISI, R. SHROFF, M. LICHTEN, P. O. BROWN *et al.*, 2000 Global mapping of meiotic recombination hotspots and coldspots in the yeast *Saccharomyces cerevisiae*. *Proc. Natl. Acad. Sci. USA* **97**: 11383–11390.
- GRAUR, D., and W.-H. LI, 1999 *Fundamentals of Molecular Evolution*, Ed. 2. Sinauer Associates, Sunderland, MA.
- GU, X., 1999 Statistical methods for testing functional divergence after gene duplication. *Mol. Biol. Evol.* **16**: 1664–1674.
- HALL, T. A., 1999 BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**: 46–49.
- HILL, W. G., and A. ROBERTSON, 1966 The effect of linkage on limits to artificial selection. *Genet. Res.* **8**: 269–294.

- HUGHES, A. L., 1999 *Adaptive Evolution of Genes and Genomes*. Oxford University Press, Oxford.
- KNUDSEN, B., and M. M. MIYAMOTO, 2001 A likelihood ratio test for evolutionary rate shifts and functional divergence among proteins. *Proc. Natl. Acad. Sci. USA* **98**: 14512–14517.
- KNUDSEN, B., M. M. MIYAMOTO, P. J. LAIPIS and D. N. SILVERMAN, 2003 Using evolutionary rates to investigate protein functional divergence and conservation: a case study of the carbonic anhydrases. *Genetics* **164**: 1261–1269.
- KONDRASHOV, F. A., I. B. ROGOZIN, Y. I. WOLF and E. V. KOONIN, 2002 Selection in the evolution of gene duplications. *Genome Biol.* **3**: RESEARCH0008.1–0008.9.
- LI, W.-H., 1997 *Molecular Evolution*. Sinauer Associates, Sunderland, MA.
- 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.
- NOWAK, M. A., M. C. BOERLIJST, J. COOKE and J. M. SMITH, 1997 Evolution of genetic redundancy. *Nature* **388**: 167–170.
- OHNO, S., 1970 *Evolution by Gene Duplication*. Springer-Verlag, Heidelberg, Germany.
- PEARSON, W. R., 1991 Searching protein sequence libraries: comparison of the sensitivity and selectivity of the Smith-Waterman and FASTA algorithms. *Genomics* **11**: 635–650.
- REMM, M., C. E. V. STORM and E. L. SONNHAMMER, 2001 Automatic clustering of orthologs and in-paralogs from pairwise species comparisons. *J. Mol. Biol.* **314**: 1041–1052.
- ROBINSON-RECHAVI, M., and D. HUCHON, 2000 RRTree: relative-rate tests between groups of sequences on a phylogenetic tree. *Bioinformatics* **16**: 296–297.
- ROONEY, A. P., H. PIONTKIVSKA and M. NEI, 2002 Molecular evolution of the histone 3 multigene family. *Mol. Biol. Evol.* **19**: 68–75.
- SEOIGHE, C., N. FEDERSPIEL, T. JONES, N. HANSEN, V. BIVOLAROVIC *et al.*, 2000 Prevalence of small inversions in yeast gene order evolution. *Proc. Natl. Acad. Sci. USA* **97**: 14433–14437.
- STEPHAN, W., and C. H. LANGLEY, 1989 Molecular genetic variation in the centromeric region of the X chromosome in three *Drosophila ananassae* populations. I. Contrasts between the *vermilion* and *forked* loci. *Genetics* **121**: 89–99.
- 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.
- THOMPSON, J. D., D. G. HIGGINS and T. J. GIBSON, 1994 CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**: 4673–4680.
- THORNTON, K., and M. LONG, 2002 Rapid divergence of gene duplicates on the *Drosophila melanogaster* X chromosome. *Mol. Biol. Evol.* **19**: 918–925.
- VAN DE PEER, Y., J. S. TAYLOR, I. BRAASCH and A. MEYER, 2001 The ghost of selection past: rates of evolution and functional divergence of anciently duplicated genes. *J. Mol. Evol.* **53**: 436–446.
- WANG, Y., and X. GU, 2001 Predicting functional divergence of caspase gene family. *Genetics* **158**: 1311–1320.
- WOLFE, K. H., and D. S. SHIELDS, 1997 Molecular evidence for an ancient duplication of the entire yeast genome. *Nature* **387**: 708–713.
- WU, C.-I., and W.-H. LI, 1985 Evidence for higher rates of nucleotide substitution in rodents than in man. *Proc. Natl. Acad. Sci. USA* **82**: 1741–1745.
- ZHANG, J., 2003 Evolution by gene duplication: an update. *Trends Ecol. Evol.* **18**: 292–298.
- ZHANG, P., Z. GU and W.-H. LI, 2003 Different evolutionary patterns between young duplicate genes in the human genome. *Genome Biol.* **4**: R56.
- ZHANG, Z., and H. KISHINO, 2004 Genomic background drives the divergence of duplicated *amylase* genes at synonymous sites in *Drosophila*. *Mol. Biol. Evol.* **21**: 222–227.

Communicating editor: S. YOKOYAMA

