

## Research Article

# Is Speciation Accompanied by Rapid Evolution? Insights from Comparing Reproductive and Nonreproductive Transcriptomes in *Drosophila*

Santosh Jagadeeshan,<sup>1,2</sup> Wilfried Haerty,<sup>1</sup> and Rama S. Singh<sup>1</sup>

<sup>1</sup> Department of Biology, McMaster University, Hamilton, ON, Canada L8S 4K1

<sup>2</sup> Smithsonian Tropical Research Institute, P. O. Box 0834-03092, Balboa, Ancón, Panama

Correspondence should be addressed to Rama S. Singh, singh@mcmaster.ca

Received 10 January 2011; Revised 4 April 2011; Accepted 19 May 2011

Academic Editor: Jose M. Eirin-Lopez

Copyright © 2011 Santosh Jagadeeshan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The tempo and mode of evolutionary change during speciation have remained contentious until recently. While much of the evidence claiming speciation is an abrupt and rapid process comes from fossil data, recent molecular phylogenetics show that the background of gradual evolution is often broken by accelerated rates of molecular evolution during speciation. However, what kinds of genes affect or are affected by speciation remains unexplored. Our analysis of 4843 protein-coding genes in five species of the *Drosophila melanogaster* subgroup shows that while ~70% of genes follow clock-like evolution, between 17–19.67% of loci show signatures of accelerated rates of evolution in recently formed species. These genes show 2-3-fold higher rates of substitution in recently diverged species compared to older species. This fraction of loci affects a diverse range of functions. Only a small proportion of reproductive genes experience speciation-related accelerated changes but many sex- and -reproduction related genes show an interesting pattern of persistent rapid evolution suggesting that sex- and -reproduction related genes are under constant selective pressures. The identification of loci associated with accelerated evolution allows us to address the mechanisms of rapid evolution and speciation, which in our study appears to be a combination of both selection and rapid demographical changes.

## 1. Introduction

The tempo and mode of evolutionary change during speciation have remained a contentious issue for more than five decades. Evidence for abrupt and rapid changes during speciation came from fossil evidence but lacked mechanistic explanations for such a process [1–7]. Since the mid-1970s, molecular phylogenetic studies began to associate increased genetic changes with speciation events suggesting that rates of molecular evolution might be altered during speciation [8–11]. This trend has been recently confirmed by molecular phylogenies using large numbers of genes and a wide variety of taxa, explicitly showing that speciation is accompanied by accelerated rates of molecular evolution [12–15].

What remains unknown is the mode of change during speciation, that is, the kind of genetic changes associated with speciation. More specifically, the numbers and kinds of protein coding genes that change during speciation remain

unexplored. Accelerated evolution is only observed in a fraction of genes analyzed by Pagel et al. [13] as well as in other genome-wide estimates of molecular divergences [16–18]. A more systematic search to identify protein-coding genes that experience accelerated evolution during speciation will allow us to directly address the mode of evolutionary change during speciation, that is, divergences, in what kinds of genes, affect or are affected by speciation.

Taxa-wide evidence of sex- and reproduction-related genes evolving rapidly in sibling species [19, 20] has raised the possibility that sex-related genes might preferentially experience elevated rates of evolution during the speciation process, or may even drive speciation by causing reproductive isolation among diverging populations. This framework is supported by the fact that almost all candidate “speciation” genes identified, so far, are mainly sex related (with the exception of a few genes with other functions), evolve rapidly between closely related species, often show signatures of

adaptive evolution and have been invoked in the rapid evolution of hybrid sterility in different organisms [21–32]. After controlling for incomplete lineage sorting in the melanogaster subgroup, we put this framework to test by analyzing 4843 protein coding genes in the *Drosophila* transcriptome for signatures of speciation-related accelerated changes.

Divergence trends between recently formed species and old species provide a proxy to detecting signatures of speciation-related accelerated changes. Molecular divergences are expected to be proportional to the duration of a species' existence; newly formed species will have accumulated less molecular divergence compared to older species. Signatures of speciation-related accelerated evolution will be manifested as relatively higher rates of molecular evolution in newly formed species compared to species with longer postspeciation divergence times. Accordingly we asked the following questions: (1) are rates of molecular evolution in protein-coding genes affected by speciation events? That is, do genes show unexpectedly elevated rates of evolution in newly formed species relative to older species? (2) do sex-related genes preferentially exhibit accelerated speciation-related changes relative to non-sex related genes? And (3) do genes with accelerated rates of evolution show evidence of positive selection?

## 2. Methods

**2.1. Rationale.** If evolutionary rates of protein coding genes are not affected by speciation and evolve at a constant rate, we should be able to find a correlation between the length of time that species have diverged and the proportion of molecular divergence between these species. Species diverged for longer lengths of time would have accumulated proportionally higher amounts of genetic changes relative to species diverged for shorter periods of time. Conversely, if rates of molecular evolution were indeed affected by speciation, then this correlation will be broken and newer species might show relatively higher evolutionary rates compared to older species. Our analyses therefore primarily exploit the nature of molecular changes *since* divergence of species pairs (recent versus older) to infer speciation-related accelerated changes.

**2.2. The Phyletic System.** We used a phyletic system comprised of three pairs of closely related species from the melanogaster subgroup *D. simulans* and *D. sechellia* diverged about 0.3–0.6 Mya [33], *D. melanogaster* and *D. simulans* diverged about 4.3–6.5 Mya, and *D. yakuba* and *D. erecta* diverged about 8.1–12.7 Mya [34]. Divergence times for *D. simulans*-*D. melanogaster* and *D. yakuba*-*D. erecta* were recently reestimated by using over 2977 nuclear genes and by implementing a novel genomic-mutation distance approach correcting for codon bias [34] whereas the divergence times for the *D. simulans*-*D. sechellia* split were estimated using a small number of genes [33]. All of these divergence dates broadly concur with other independent estimates using mitochondrial loci (reviewed in Powell 1997), and there is also some general concordance with what little

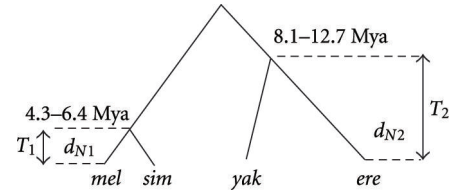


FIGURE 1: Schematic representation of rationale implemented in our methodology.  $d_{N1}$ : proportion of nonsynonymous substitutions in the more recently diverged species pair (diverged for time  $T_1$ ).  $d_{N2}$ : proportion of nonsynonymous substitutions in species that have diverged for longer periods of time ( $T_2$ ).

paleontological evidence is available for *Drosophila* (see [34]). We also used *D. pseudoobscura* and *D. persimilis* from the *D. obscura* group diverged for about 0.85 Mya [34]. These two species belong to the *Obscura* group that diverged 55 Mya from the melanogaster group [34], which brings about problems of saturation in  $d_S$  and gene expression differences [35]. Nevertheless, this species pair diverged for roughly the same amount of time as *D. simulans*-*D. sechellia*, a comparison of rates of protein evolution between the two species pairs will be useful to determine if acceleration in rates of evolution is commonly found among newly derived species.

**2.3. Estimating Differences in Rates of  $d_N$  in relation to Speciation.** According to the divergence times recently reported by Tamura et al. [34], *D. erecta* and *D. yakuba* have diverged for an estimated length of time that is  $\sim 2$ -3 times greater than the divergence time between *D. melanogaster* and *D. simulans*. Given the neutrally expected linear relationship between genetic divergence and time, we should therefore expect  $\sim 2$ -3 times greater divergence in genes between *D. yakuba* and *D. erecta* relative to *D. melanogaster* and *D. simulans*. We calculated the expected rate of molecular divergence (rate of nonsynonymous mutations per nonsynonymous sites  $d_N$  and rate of synonymous mutations per synonymous sites  $d_S$ ) using the relationship given below. Since our interest ultimately was to determine differences in rates of protein evolution between the different species pairs we focused more on  $d_N$ . The ratio of nonsynonymous divergence ( $d_N$ ) between two-species pairs, a recently diverged species pair 1 and an older species pair 2, must be proportional to the ratio of their divergence times ( $T$ ), such that:

$$\log_2 \left( \frac{T_1}{T_2} \right) = \log_2 \left( \frac{d_{N1}}{d_{N2}} \right), \quad (1)$$

where, for a given species pair,  $T_1$  = divergence time of the more recently diverged pair and  $T_2$  = divergence time of the species with longer post-speciation divergence time. In our data, for example,  $T_1$  = 4.3–6.5 Mya (mel-sim) and  $T_2$  = 8.1–12.7 Mya (yak-ere), or,  $T_1$  = 0.4–0.6 Mya (sim-sec) and  $T_2$  = 8.1–12.7 Mya (yak-ere) (see Figure 1). We similarly calculated the ratio of synonymous divergence to species divergence time in all species pairs. Because of existing uncertainties in the divergence time estimates, particularly for the newer species [36, 37], we worked with

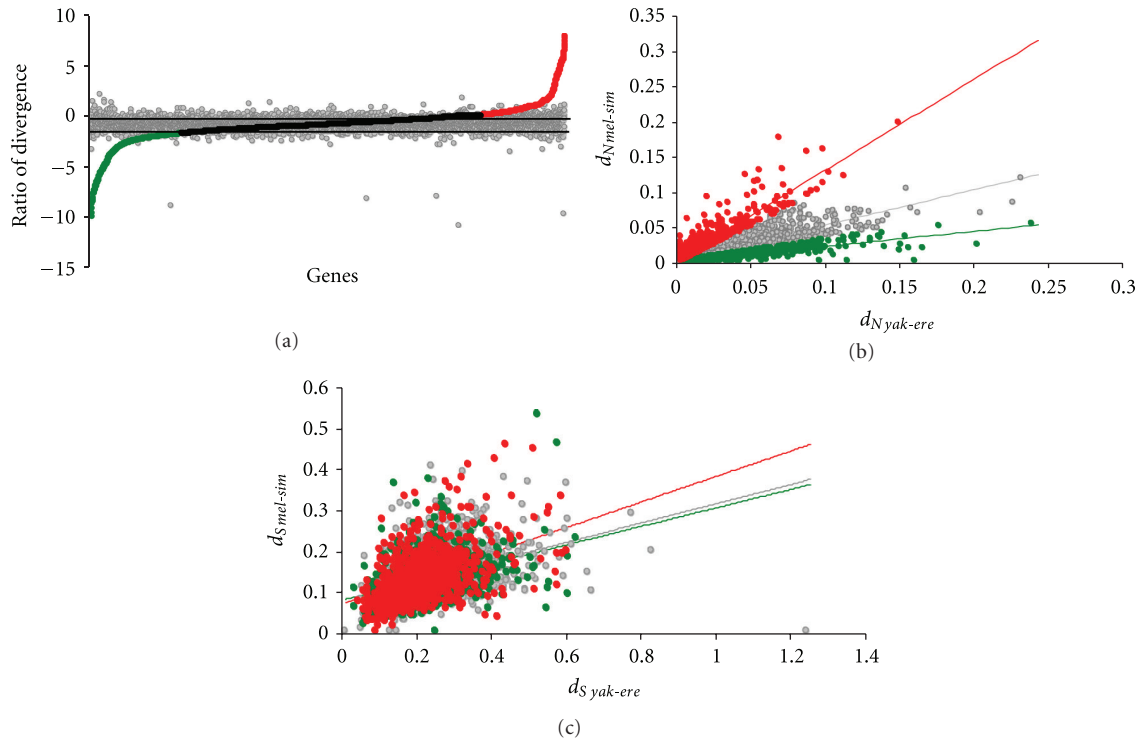


FIGURE 2: Distribution of ratio of divergence between *D. melanogaster*-*D. simulans* versus *D. yakuba*-*D. erecta*. (a) Distribution of molecular-divergence  $\log_2(d_{N1}/d_{N2})$  to divergence-time ratio  $\log_2(T_1/T_2)$ , between *D. melanogaster*-*D. simulans* and *D. yakuba*-*D. erecta*. Colored circles represent  $d_N$  ratios, and gray circles represent  $d_S$  ratios. The upper limit of the clockwise category is given by the  $\log_2$  ratio of divergence times (6.5/8.1) and the lower limit by the  $\log_2$  ratio 4.3/12.7. (b) Relationship of  $d_N$  between *D. melanogaster*-*D. simulans* and *D. yakuba*-*D. erecta*. The regression lines for each rate categories are also plotted. (c) Relationship of  $d_S$  between *D. melanogaster*-*D. simulans* versus *D. yakuba*-*D. erecta*; color codes: red: accelerated, black: clock-like, green: slow-rate categories, and grey:  $d_S$  ratios.

the estimated range of divergence times for each of the species pairs by incorporating the published upper and lower limits of divergence time estimates from Tamura et al. [34]. For instance, in comparisons between *D. melanogaster*-*D. simulans* versus *D. yakuba*-*D. erecta* pairs, the upper divergence-time limit represents a scenario where *D. melanogaster* and *D. simulans* diverged 4.3 Mya and *D. yakuba* and *D. erecta* diverged 12.7 Mya. Similarly, the lower divergence time limit represents a scenario where *D. melanogaster* and *D. simulans* diverged 6.5 Mya, and *D. yakuba* and *D. erecta* diverged 8.1 Mya. When  $\log_2(d_{N1}/d_{N2})$  is plotted against  $\log_2(T_1/T_2)$ , genes evolving in a clock-like manner (i.e.,  $d_{N1} = [d_{N2} * (T_1/T_2)]$ ) would fall within these divergence time-ranges. Genes that fall above the divergence time-ranges indicate accelerated evolution in the newer species and those falling below the lower limit are considered to evolve slowly in the newly formed species or to have much higher rates of evolution in the older species lineage.

Comparisons using the *D. simulans*-*D. sechellia* species pair grossly overestimated the number of genes in the accelerated rate category (94.89% of the genes fall under the accelerated category Figure 3(a)). This is an unlikely scenario given that the range of  $d_S$  falls well above the divergence times plotted in this graph. We believe that this is most likely due to a gross underestimation of the divergence time between *D. simulans* and *D. sechellia*. The

phylogenetic relationships in the *simulans* triad (*D. simulans*, *D. mauritiana*, and *D. sechellia*) have been contentious and unresolved [35] and the sole source of recent divergence times estimates using molecular data comes from Kliman et al.'s study [33], which used a small number of genes. We took a very generalized approach to reevaluate the divergence time for the *D. simulans*-*D. sechellia* split using  $d_S$  estimated from our dataset (Figure 2) in order to replot Figure 3(a). Relative to nonsynonymous divergence, we expect the ratio of synonymous divergence-to-species divergence times for most loci to fall within the clock-like category for most species comparisons (as it did in the previous comparison). We therefore applied the  $d_S$  boundaries conforming to the clock-like category from *D. melanogaster*-*D. simulans* versus *D. yakuba*-*D. erecta* comparisons (Figure 2) to the *D. simulans*-*D. sechellia* species pair. We were able to arrive at a generalized estimate of 1.16–3.05 Mya for the *D. simulans*-*D. sechellia* split which appears more likely (Kumar 2007, pers comm.). We therefore employed a slightly more stringent approach by applying 95% confidence intervals to the existing divergence boundaries (using  $d_S$  ratios), which consequently incorporated previous outliers (Table 1).

2.4. Sequences and Rate Analyses. All sequences (*D. melanogaster*, *D. simulans*, *D. sechellia*, *D. yakuba*, *D. erecta*, *D. pseudoobscura*, and *D. persimilis*) were obtained from the

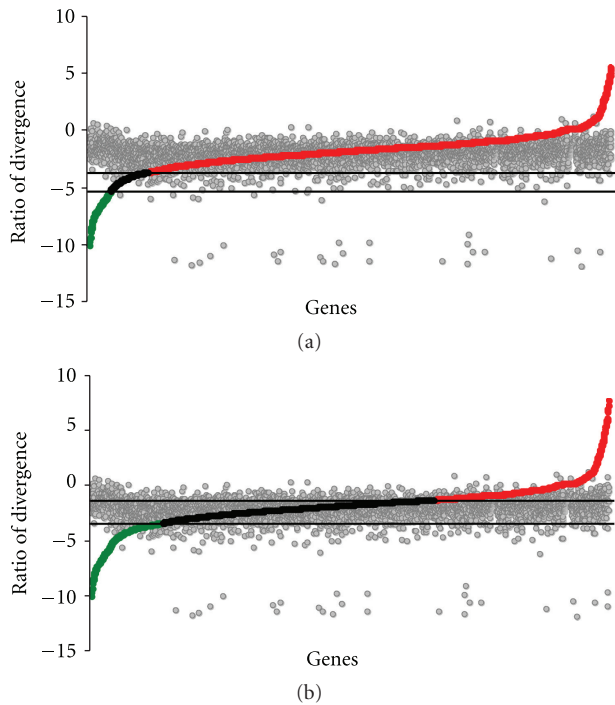


FIGURE 3: Distribution of ratio of divergence between *D. simulans*-*D. sechellia* and *D. yakuba*-*D. erecta*. (a) Distribution of ratio of amino acid replacement substitutions in separate clades  $\log_2(d_{N1}/d_{N2})$  to divergence time (according to Kliman et al. [33]) ratio  $\log_2(T_1/T_2)$ , *D. simulans*-*D. sechellia* and *D. yakuba*-*D. erecta*. (b) Re-evaluation of divergence times between these species.  $d_N$  ratios are plotted in colors, and  $d_S$  ratios are plotted in grey. Color codes: red circles accelerated evolution, black: clock-like, green: slow evolution categories, grey:  $d_S$  ratios.

recently sequenced genomes available on FlyBase (*Drosophila* 12 Genomes consortium 2007, <http://www.flybase.org/>). Sequences were aligned according to the corresponding protein alignment using CLUSTALW ver. 1.8 [38]. In order to remove any potential bias due to incomplete lineage sorting effects [39, 40], for each gene we compared the likelihood of trees differing in the placement of *D. yakuba* and *D. erecta* using PAML and restricted our analysis to genes for which the best tree involved *D. yakuba* and *D. erecta* as sister species. Pairwise estimates of  $d_N$  and  $d_S$  were determined using the program codeml in PAML [41]. Estimates of  $d_N$ ,  $d_S$ , and  $\omega$  along each lineage using branch site models and outputs of models M7 (neutral) versus M8 (positive selection) were also determined using PAML [42] and retrieved from a recent genome-wide analysis [43]. We were able to compute estimates of divergence to time ratio (see above) for 4843 orthologs between *D. melanogaster*-*D. simulans* versus and *D. yakuba*-*D. erecta*, and 4327 orthologs between *D. simulans*-*D. sechellia* versus *D. yakuba* - *D. erecta* comparisons as well as for 3988 genes from *D. pseudoobscura*-*D. persimilis* versus all melanogaster species pairs.

2.5. *Classification of Genes according to Site of Expression.* We classified genes according to their tissues of expression (testis, ovary, and head) by using the NCBI EST database (NCBI, <http://www.ncbi.nlm.nih.gov/UniGene/>). Genes that could not be classified into any tissue category were referred as unspecified. Based on EST data, we were able to classify tissue of expression for 3040 out of 4843 genes for comparisons involving *D. melanogaster*-*D. simulans* versus *D. yakuba*-*D. erecta*, 2686 genes for comparisons involving *D. simulans*-*D. sechellia* versus *D. yakuba*-*D. erecta* (Table S1).

### 3. Results

3.1. *Accelerated Rates of Molecular Evolution in Newly Formed Species.* Molecular divergence estimates, with respect to species divergence times, showed that protein coding genes fell into three distinct rate categories: (1) *accelerated evolution in younger species*: genes showing higher rates of molecular divergence in newly formed species relative to species diverged for longer lengths of time. (2) *clock-like evolution*: genes showing molecular divergence that corresponds to species divergence times, and (3) *slow evolution in younger species*: genes showing lower rates of molecular divergence in newly formed species in comparison to species with longer divergence times (Figure 2). Table 1 summarizes the fraction of genes that fall under each rate category for every species pair compared.

Most genes in our dataset (61–74%) fell into the clock-like rate category in all comparisons, indicating that evolutionary trajectories of most genes are unaffected by speciation events and their rates remained constant (Table 1). A small but discernable fraction of genes (17–19%) showed signatures of speciation-related accelerated evolution. Nonsynonymous divergences in these genes were 2-3-fold higher in newly formed species compared to older species (Table 1, Figures 2 and 3). A plot of  $d_N$  and  $d_S$  estimates between *D. melanogaster*-*D. simulans* versus *D. yakuba*-*D. erecta* shows that the distribution of  $d_N$  estimates is quite distinct between the accelerated, clock-like, and slow-rate categories while the distribution of  $d_S$  estimates is not (Figures 2(b) and 2(c)). This quite clearly indicates that elevated proportions of  $d_N$  is not always accompanied by correspondingly elevated proportions of  $d_S$  in genes with evidence of accelerated evolution, which is generally a sign of selection driven changes. We also obtained similar results using protein divergence estimates instead of the ratios of nucleotide divergence data. See files in Supplementary Material available online at doi: 10.4061/2011/595121. About 8–19% of protein coding genes that fell into the slow rate category showed extremely low levels of nonsynonymous divergence in newly formed species compared to older species (e.g., Avg  $d_{Nmel-sim} = 0.0057 \pm 0.007$ , Avg  $d_{Nyak-ere} = 0.2276 \pm 0.081$ , Figures 2, 4, and 5). This may be indicative of genes that have remained conserved in the evolution of the recently diverged species but that have diverged substantially with time in the older lineages.

These results quite unambiguously indicate two important trends: firstly, despite variances associated with divergence estimates, the 2-3-fold higher nonsynonymous divergence in newly formed species compared to species with

TABLE 1: Number and (percent) of genes falling under accelerated, clock-like, and slow-rate categories in recently diverged species when compared to older species pairs. (a) Analysis using the published divergence boundaries according to Kliman et al. [33] for *D. simulans* and *D. sechellia*, and Tamura et al. [34] for *D. melanogaster*-*D. simulans* and *D. yakuba*-*D. erecta*. (b) Analysis using the revised divergence time estimates for *D. simulans* and *D. sechellia*. (c) Analysis applying the 95% confidence interval boundaries on revised divergence estimates.

	Rate Categories		
	Slow	Clock-like	Accelerated
(a)			
<i>mel-sim</i> versus <i>yak-ere</i>			
$d_N$	1096 (22.63)	2251 (46.48)	1496 (30.89)
$d_S$	348 (7.19)	3628 (74.91)	867 (17.90)
<i>sim-sec</i> versus <i>yak-ere</i>			
$d_N$	173 (3.99)	314 (7.26)	3840 (88.75)
$d_S$	37 (0.86)	184 (4.25)	4106 (94.89)
(b)			
<i>sim-sec</i> versus <i>yak-ere</i>			
$d_N$	594 (13.73)	2275 (52.58)	1458 (33.69)
$d_S$	311 (7.19)	3241 (74.90)	775 (17.91)
(c)			
<i>mel-sim</i> versus <i>yak-ere</i>			
$d_N$	963 (19.89)	2996 (61.86)	884 (18.25)
$d_S$	242 (5.00)	4359 (90.00)	242 (5.00)
<i>sim-sec</i> versus <i>yak-ere</i>			
$d_N$	469 (10.84)	3007 (69.49)	851 (19.67)
$d_S$	216 (5.00)	3895 (90.00)	216 (5.00)
<i>pse-per</i> versus <i>sim-sec</i>			
$d_N$	327 (8.20)	2962 (74.27)	699 (17.53)
$d_S$	199 (5)	3590 (90)	199 (5)

much longer divergence times clearly represents increases in rates of protein evolution either during or immediately after speciation. Secondly, accelerated rates of molecular evolution are most apparent in nonsynonymous divergence and not in synonymous divergence (see Table 1, Figures 2 and 4), a broadly accepted sign of selection driven changes.

**3.2. Higher Representation of Sex-Related Genes in the Accelerated and Clock-Like Rate Categories.** Identifying the range and, particularly, the kinds of loci in each rate category (accelerated, clock-like and slow) provides a starting point to broadly address the effects of demographic factors and selection during speciation. Demographic factors (drift, bottlenecks etc.) would affect a wide variety of loci whereas selection driven divergence would only affect specific functional classes of genes, such as sex-related genes which are expected to drive reproductive isolation in diverging populations [45, 46]. Due to lack of functional information for a large number of *Drosophila* genes, we used tissue of expression as a general and presumable indication of function (testis and ovary as reproductive tissues versus head, a presumably nonreproductive tissue). We tested the null hypothesis that genes expressed in each tissue-type are equally distributed within the accelerated, clock-like and slow rate categories. Because the expression data was determined using *D. melanogaster* data, only species from the *D. melanogaster* subgroup were analyzed. Given the

long divergence time between the *Obscura* group and the *melanogaster* group (55 Mya), it is likely that patterns of gene expression may be radically different in *D. pseudoobscura* and *D. persimilis*.

In the *D. melanogaster*-*D. simulans* versus *D. yakuba*-*D. erecta* comparison, a significantly higher proportion of testis-specific genes occupy the accelerated and clock-like rate categories compared to the slow-rate category ( $\chi^2$  testis specific = 8.08 and 6.78, df = 1, and  $P = 0.0134$  and  $0.0277$ , respectively, and a Bonferroni correction was applied, Table 2, [44]). Genes expressed in all three tissues—testis, ovary and head mostly fall in the accelerated category compared to both slow and clock-like rate categories ( $\chi^2$  testis, ovary-and-head, = 5.96 df = 1,  $P = 0.0439$ ,  $\chi^2 = 19.99$ , df = 1,  $P = 2.34 \times 10^{-5}$ , Bonferroni corrections were applied [44], Table 2). No significant differences were observed in the proportions of all other genes classes between rate categories.

In the *D. simulans*-*D. sechellia*/*D. yakuba*-*D. erecta* comparison, a significantly small proportion of testis-specific genes fall into the accelerated category compared to clock-like and slow rate categories ( $\chi^2 = 10.04$  and  $13.91$ , df = 1,  $P = 4.59 \times 10^{-3}$  and  $5.75 \times 10^{-4}$ , a Bonferroni correction was applied, Table 2).

**3.3. “Persistence” of Rapid Evolution in Testis-Specific Genes over Time.** That most testis-specific genes fall under the clock like rate category and only a small proportion fall in the

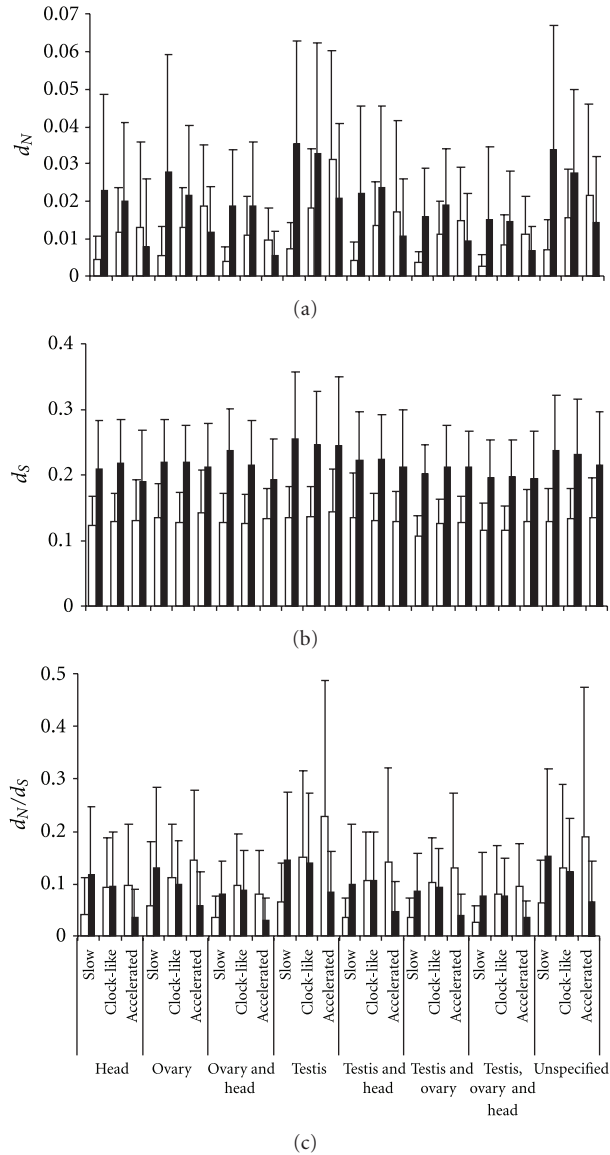


FIGURE 4: Relationship between rate category and tissue of expression. Distribution of (a)  $d_N$ , (b)  $d_S$ , and (c)  $d_N/d_S$  between rate categories (accelerated, neutral, and slow) and between tissue classifications. White bars: *D. melanogaster-D. simulans*, and black bars: *D. yakuba-D. erecta*. The error bars represent the standard deviation.

accelerated rate category would appear to contradict earlier evidence that testis specific genes are in general evolving rapidly [20, 47, 48]. We hypothesized that a “persistence” of rapid rates of evolution in testis specific genes (rapid rates of evolution over time) will explain why testis-specific genes largely occupy the clock-like rate category. Pagel et al. [13] also found such persisting elevated rates of evolution in many lineages. Persistence of rapid rates of evolution can be verified if testis specific genes in both younger and older lineages have high substitution rates in both the accelerated and clock-like rate categories compared to other genes. To test this, we first looked for a global difference in  $d_N$  and  $d_S$  between tissue

categories as well as between rate categories (Figure 5). We also performed more detailed analyses of differences in  $d_N$  and  $d_S$  between each tissue category using a Tuckey HSD test. Invariably, testis-specific genes show significantly higher  $d_N$  and  $d_N/d_S$  compared to genes in all other tissues in both younger and older species pairs ( $P < 0.05$  for both, Figures 4(a), 4(b), and 4(c)) (additional file 3). This implies that testis genes in the clock-like category are evolving faster than other genes but at a constant rate, whereas those in the accelerated rate category have experienced elevated rates of evolution in response to stronger selection during or after speciation. This is an important trend as it reveals a tempo of molecular evolution in the different classes of genes.

**3.4. Genes Evolving under Positive Selection in the Accelerated, Clock-Like, and Slow Rate Categories.** We compared the proportion of genes showing evidence of positive selection ( $\omega > 1.0$ ) among the different rate categories by implementing site models, branch-site and branch models in PAML [42]. Applying models M7 versus M8, none of the rate categories showed overrepresentation of genes evolving positively (slow versus accelerated,  $\chi^2 = 0.88$ ,  $df = 1$ ,  $P > 0.05$ , accelerated, slow versus clock-like  $\chi^2 = 0.93$  and  $2.46$ ,  $df = 1$ ,  $P > 0.05$  and  $0.35$  after Bonferroni corrections, Table 3). Using the branch-site model however, we observed a significant over-representation of genes showing positive selection in *D. simulans* in the accelerated rate category compared to the clock-like and slow rate categories ( $\chi^2 = 20.82$  and  $28.6$ ,  $df = 1$ ,  $P = 1.51 \times 10^{-5}$  and  $2.67 \times 10^{-7}$ , respectively after Bonferroni correction). Branch model tests also show a greater proportion of genes with foreground  $\omega >$  background  $\omega$  in the accelerated rate category in the *D. simulans* branch and the branch leading to the *D. melanogaster* clade (Table 3). A list of genes detected to be evolving under positive selection in each lineage can be found in additional file 4. Several genes involved in sensory stimuli, immune response, gametogenesis (spermatogenesis), transcription regulation, and hybrid incompatibilities (including *Hmr*, [21] are amongst the genes that show relatively large  $\omega$  estimates. Population genomic study of six *D. simulans* strains compared to *D. melanogaster* [49] found significant evidence of directional selection in genes affecting reproduction or spermatogenesis. Among the 1270 genes that show evidence of adaptive evolution from Singh [46], 505 are found in our comparison of *D. melanogaster-D. simulans/D. yakuba-D. erecta*. Among the 505 genes, 360 fall in the clock-like category, 108 in the accelerated category and 37 fall in the slow rate category. We detect a significant enrichment of genes under positive selection among the clock-like category  $\chi^2 = 7.25$ ,  $df = 1$ ,  $P = 7.09 \times 10^{-3}$ ) while we found a significant paucity among the slow rate category ( $\chi^2 = 40.05$ ,  $df = 1$ ,  $P = 2.48 \times 10^{-10}$ ). No significant effect is observed for the genes classified as rapidly evolving in the younger species ( $\chi^2 = 2.72$ ,  $df = 1$ ,  $P = 0.099$ ).

**3.5. Effect of Local Recombination Rates.** Gene evolution may be influenced by their chromosomal location [50, 51] as well

TABLE 2: Number of genes present in each rate category when comparing *D. melanogaster*-*D. simulans* to *D. yakuba*-*D. erecta* and *D. simulans*-*D. sechellia* and to *D. yakuba*-*D. erecta*. Under brackets are the proportion of genes within tissue and the proportion within the rate category. Enrichment within a category was tested using  $\chi^2$  test and the Bonferroni corrections were applied [44].

Tissue	Accelerated	Clock-like	Slow	Total
<i>D. melanogaster</i> - <i>D. simulans</i> versus <i>D. yakuba</i> - <i>D. erecta</i>				
H	129 (17.18); (14.59)	445 (60.59); (14.85)	177 (22.63); (18.38)	<b>751 (15.51)</b>
O	47 (20.26); (5.32)	147 (63.79); (4.91)	38 (15.95); (9.95)	<b>232 (4.79)</b>
OH	35 (15.70); (3.96)	138 (63.68); (4.61)	50 (20.62); (5.19)	<b>223 (4.6)</b>
T	142 (20.67); (16.06)	440 (65.50); (14.69)	105 (14.70); (10.90)	<b>687 (14.19)</b>
TH	114 (20.36); (12.90)	357 (64.46); (11.92)	89 (15.72); (9.24)	<b>560 (11.56)</b>
TO	32 (20.65); (3.62)	99 (64.52); (3.30);	24 (14.84); (2.49)	<b>155 (3.2)</b>
TOH	116 (21.17); (13.12)	231 (55.97); (7.71)	80 (17.57); (8.31)	<b>427 (8.82)</b>
un	269 (14.88); (30.43)	1139 (64.66); (38.02)	400 (20.79); (41.54)	<b>1808 (37.33)</b>
Total	<b>884 (18.25)</b>	<b>2996 (61.86)</b>	<b>963 (19.89)</b>	<b>4843</b>
<i>D. simulans</i> - <i>D. sechellia</i> versus <i>D. yakuba</i> - <i>D. erecta</i>				
H	149 (21.94); (17.51)	458 (67.45); (15.23)	72 (10.6); (15.35)	<b>679 (15.69)</b>
O	32 (16.24); (3.76)	146 (74.11); (4.86)	19 (9.64); (4.05)	<b>197 (4.55)</b>
OH	48 (32.65); (5.64)	130 (65.99); (4.32)	19 (9.64); (4.05)	<b>197 (4.55)</b>
T	101 (16.13); (11.87)	427 (68.21); (14.20)	98 (15.65); (20.90)	<b>626 (14.47)</b>
TH	101 (19.65); (11.87)	363 (70.62); (12.07)	50 (9.73); (10.67)	<b>514 (11.88)</b>
TO	26 (19.55); (3.06)	99 (74.44); (3.29)	8 (6.02); (1.71)	<b>133 (3.07)</b>
TOH	77 (22.65); (9.05)	217 (63.82); (7.22)	46 (13.53); (9.81)	<b>340 (7.86)</b>
un	317 (19.32); (37.25)	1167 (71.11); (38.81)	157 (9.77); (33.48)	<b>1641 (37.92)</b>
Total	<b>851 (19.67)</b>	<b>3007 (69.49)</b>	<b>469 (10.84)</b>	<b>4327</b>

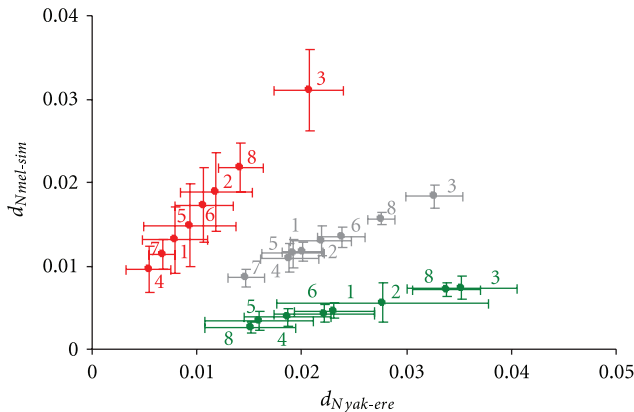


FIGURE 5: Relationship between evolutionary rates and tissue of expression. Relationship between  $d_N$  *D. melanogaster*-*D. simulans* and  $d_N$  *D. yakuba*-*D. erecta* for gene classifications in each rate category: red: accelerated, grey: clock-like, green: slow. Tissue classifications: (1) head-specific; (2) ovary-specific; (3) testis-specific; (4) ovary and head; (5) ovary and testis; (6) testis and head; (7) Testis; ovary and head; (8) unknown. Error bars represent 95% confidence intervals.

as by local recombination rates [43]. Using recombination rates in *D. melanogaster* computed by Hey and Kliman [52], we found significantly higher average recombination rates for genes in the accelerated rate category compared to the clock-like and slow categories in the *D. melanogaster*-*D. simulans*/*D. yakuba*-*D. erecta* comparison (Kruskal Wallis

rank sum test,  $P < 2.2 \times 10^{-16}$  and  $P = 0.0234$ , respectively, a Bonferroni correction was applied). There was no significant difference between the clock-like and slow rate categories (Kruskal Wallis rank sum test,  $P = 0.323$ , a Bonferroni correction was applied). The *D. simulans*-*D. sechellia*/*D. yakuba*-*D. erecta* comparison showed no such effect of recombination in all rate categories (Kruskal Wallis rank sum test,  $P = 0.67$ ,  $0.30$  and for clock-like versus accelerated, accelerated versus slow, clock-like versus slow, respectively after Bonferroni correction).

## 4. Discussion

**4.1. What Kinds of Genes Change during Speciation?.** Despite recent evidence linking accelerated rates of molecular evolution to speciation events [12, 13], the kinds of protein coding genes that might experience accelerated rates of evolution during or immediately after speciation require investigation [15]. Are rates of evolution in all genes likely to be altered during or even after speciation? If not, what genes do show speciation associated changes? Answers to these questions will be crucial to understanding the mechanism(s) driving speciation and the molecular evolutionary consequences of speciation.

This study analyzes the ratio of molecular evolutionary rates to the ratio of species divergence times, therefore clock-like evolution should not be taken as a sign of gradual evolution; it only implies a constant rate of evolution. This is exemplified by the large representation of testis-specific genes in the clock-like rate category, which are in fact

evolving much faster than other genes in the same category. Furthermore, *Drosophila* is one of those groups that lack detailed fossil records and incomplete taxon sampling is an obvious but unavoidable problem in this, or any study of this nature. Apart from extinct taxa, undiscovered and undescribed taxa are an additional problem. We are also lacking the genome sequence of *D. mauritiana*, a member of the *D. simulans*-*D. sechellia*-*D. mauritiana* triad as well as the genome of *D. santomea* which recently split from *D. yakuba* [53]. These factors are bound to affect true assessments of speciation-related changes in rates of molecular evolution. However, within the currently accepted phylogenetic network of the melanogaster subgroup, our study provides an approach to identify genes that change during speciation and out results report the fraction and identity of genes in the *Drosophila* transcriptome that show signatures of accelerated speciation-related changes that can be further investigated.

**4.2. Evidence for Accelerated Evolution.** In our study, evidence of speciation-related accelerated evolution is detected in a small but discernable fraction of protein coding genes. This corresponds to what Pagel et al. [13] found across taxa. But what is striking is that the acceleration is most apparent in nonsynonymous divergence. That  $d_S$  in most genes fall under the clock-like category as opposed to  $d_N$  is not entirely surprising, as we expect  $d_N$  changes to be more sensitive to selection and may therefore show higher variance. Nevertheless, the 2-3 fold higher estimates of  $d_N$  in the recently diverged species compared to species with much longer divergence times is a strict deviation from clock-like evolution despite the variance that may be involved. More importantly, our survey reveals that the more recently diverged species (*D. simulans*-*D. sechellia*) have a slightly greater proportion of genes that show accelerated rates of evolution (Table 1), further strengthening the case for a causal link between acceleration in rates of molecular evolution and speciation.

**4.3. Evidence for Persistent Rapid Evolution in Sex Genes.** Sexually reproducing organisms are influenced both by natural and sexual selection. Widespread rapid evolution of sex-related genes in *Drosophila* genomes [18] imply that regardless of how speciation occurred, sex and reproduction related genes that are causally involved in establishing reproductive isolation would be constantly under strong selection [46]. Therefore we would not expect to find an over-representation of sex-related genes in the accelerated rate category alone. This is specifically illustrated by the persistence of rapid rates of evolution in testis specific genes in our study. In addition, testis-specific genes and genes expressed in all three tissues are over-represented in the accelerated and clock-like rate category in the *D. melanogaster*-*D. simulans* pair (Table 2). But in the more recently diverged *D. simulans*-*D. sechellia* pair, only testis-specific genes are overrepresented in the clock-like and slow rate categories (Table 2). These results support a scenario where sex-related genes are under constant but higher selective pressure.

TABLE 3: Comparison of the proportion of genes showing evidence of positive selection under the site model, branch-site model, and branch model between slow ( $n = 963$ ), clock-like ( $n = 2996$ ), and accelerated ( $n = 884$ ) categories.  $P$  values for the different tests were corrected using a Bonferroni correction prior to the comparisons. *mel*: *D. melanogaster*, *sim*: *D. simulans*, *sec*: *D. sechellia*, *simsecmel*: branch leading to the *D. melanogaster* clade, *yak*: *D. yakuba*, *ere*: *D. erecta*, and *yakere*: branch leading to the *D. yakuba* clade.

Model	Slow	Clock-like	Accelerated
Site model			
<i>M7/M8</i>	6	40	9
Branch-site model			
<i>mel</i>	0	8	0
<i>sim</i>	5	51	39
<i>sec</i>	10	29	9
<i>simsecmel</i>	3	6	3
<i>yakere</i>	3	4	0
Branch model			
<i>mel</i>	0	6	6
<i>sim</i>	0	3	24
<i>sec</i>	1	13	3
<i>simsecmel</i>	1	59	80
<i>yak</i>	20	20	0
<i>ere</i>	36	26	0
<i>yakere</i>	68	83	0

**4.4. Factors Driving Accelerated Evolution during Speciation.** We find no evidence of widespread positive selection in the accelerated rate category. The higher proportion of nonsynonymous divergence we observe in the accelerated rate category cannot be ruled out due to relaxed constraint or accumulation of slightly deleterious mutations [16, 54, 55]. The only tenable links to acceleration that we find appear to be influenced by local recombination rates (higher in the *D. melanogaster*-*D. simulans* comparison but lower in the *D. simulans*-*D. sechellia* comparison) and a relatively small proportion of genes with lineage specific positive selection (Table 3). Variation in recombination rates between species can be an important force driving divergence, as new allele combinations can be produced at different rates within species [56]. Rate of recombination also appears to be positively correlated with levels of polymorphism and high polymorphism would be expected to correlate with levels of divergence [57]. Nonetheless, sex-related genes involved in gametogenesis, hybrid incompatibilities as well as genes involved in metabolism and sensory functions do show evidence of positive selection. This implies a loci-specific role of sex genes and sexual selection in speciation along with some ecological specialization. Our data therefore indicates that acceleration in rates of evolution is not purely a consequence of strong selection but most likely a combination of factors, such as demographic processes as well as some form of selection (a few genes in the accelerated category are evolving positively, Table 3). These data are interesting particularly in *D. sechellia*, which is a relatively young species and one that is likely to have undergone founder speciation (see [33]).



According to the founder effect model of speciation [3], speciation occurs as a consequence of major demographic processes in which bottlenecks play an important role. In such a case, a large number of loci involved in a wide range of functions would be affected (but see [58] for a more recent critical analysis of bottlenecks and speciation). Our results as well as those of Pagel et al. [13] do not support the notion of a genetic revolution as a consequence of speciation; the range of loci affected during speciation is rather limited. However, loci affecting a wide range of functions show accelerated rates of evolution (Figures 1, 2, and 3, Tables 1 and 2). This supports a scenario where, as founding populations adapt to new ecological niches the evolution of modified or new behaviors affecting sex and reproduction as well as other ecological adaptations can occur rapidly [59–62]. *D. sechellia* has evolved an intricate ecological relationship with its plant host *Morinda citrifolia*, on which *D. sechellia* females oviposit their eggs [63]. Compounds in the pulp of the immature *M. citrifolia* fruit are toxic to other species of the melanogaster subgroup but not to *D. sechellia* larvae that feed and grow on it [63]. A large proportion of head-expressed genes show accelerated evolution in the *D. simulans*-*D. sechellia* pair, which might reflect rapid behavioral and sensory modifications that occurred during the host-plant specialization between *D. sechellia* and *M. citrifolia*. Genes involved in sensory perception (*Gr2a*, *Gr21a*, *Gr43a*, *Obp50a*, *Or67d*, *king-tubby*, *CG32683*), sensory organ development (*amos*, *Brd*, *mib1*, *Oseg1*, *Poxn*), detoxification (*kraken*), metabolism, and oogenesis (*cup*, *retn*, *kel*, *spir*, *Tm1*) found to be involved in host specialization by Dworkin and Jones [64] are among those that show evidence of accelerated speciation-related rates of evolution in our study.

The *D. sechellia* data demonstrates the importance of founder-effect and subsequent ecological divergence in driving speciation and as a consequence, causing elevated rates of evolution of relevant protein coding genes associated with the speciation event. Such cases of speciation driven by founder-effect and subsequent resource specialization are not uncommon in *Drosophila* [45, 59, 65]. *D. sechellia*, *D. mauritiana* and *D. santomea* (a recently described sibling species to *D. yakuba*) are all insular species within the melanogaster subgroup. Among these, *D. sechellia* and *D. santomea* show resource specialization having evolved special ecological relationships of utilizing specific host plants either for food or oviposition [53, 63–65], while their parent lineages do not. Such species-specific ecological adaptations are likely to influence accelerated evolution in the newly formed species but not necessarily in the parent lineages. Our study supports this reasoning; we find ~400 genes in the accelerated rate category in the *D. simulans* versus *D. sechellia* comparison; all of which fall in the clock-like rate category in the *D. melanogaster* versus *D. simulans* comparison (see additional file 5). Many of these genes, as mentioned above, have been found to be involved in host-plant specialization [64]. Further detailed study of such genes between *D. sechellia* and *D. simulans*, as well as the inclusion of *D. santomea* genome (when completed) in a similar study should shed some light on our claim. However, it should be noted that while *D. sechellia* is isolated from *D.*

*simulans*, *D. santomea* and its sibling species *D. yakuba* are sympatric on the Island of San Tome and there is substantial divergence of sexual traits between the two species [53, 66–69]. We might therefore expect to see higher proportions of sex-related genes as well as genes involved in ecological specializations to show signatures of accelerated evolution between *D. santomea* and *D. yakuba*, much like what we observed between the sympatric *D. melanogaster* and *D. simulans* in our study.

## 5. Conclusion

Signatures of speciation-related accelerated rates of evolution are detected in all newly evolved species analyzed in this study. This is not observed as a widespread genomic signature but is restricted to a fraction of the genome that affects widely different functions. The range and kinds of loci in this rapidly evolving fraction identified in this study complements the recently reported punctuational effect observed across a variety of taxa [13, 15] and improves our focus of studying the genetics of speciation. Population genetic studies of these ‘candidate’ genes can establish the nature of selective forces driving their elevated rates of evolution. In our case, we find a demography-selection driven effect. Testis specific genes and sex-related genes show persistently high rates of evolution, indicating that sexual selection is a constant pressure in diverging and established populations. In the *D. simulans*-*D. sechellia* pair, our data reinforce previous evidence that founder effect and ecological specialization played important roles in their speciation process. Our results support the growing appreciation that speciation is often driven by a combination of demographic fluctuations, ecological adaptation, and sexual selection and can seldom be attributed to one single important factor [45, 70].

## Acknowledgments

The authors are grateful to the AAA group for making the genomes data available to them. Richard Morton, Alberto Civetta, Sudhir Kumar, Harilaos Lessios, and two anonymous reviewers provided useful and critical comments that greatly improved this manuscript. The authors would also like to thank Ben Evans, Jonathan Stone, and Osamu Miura for very useful discussions, comments, and criticisms on previous versions of this manuscript. This study was supported by a National Science and Engineering Research Grant to R.S.S. and a Smithsonian Institution Molecular Evolution Postdoctoral Fellowship to S. Jagadeeshan. S. Jagadeeshan and W. Haerty contributed equally to this work.

## References

- [1] N. Eldredge and S. J. Gould, “Punctuated equilibria: an alternative to phyletic gradualism,” in *Models in Paleobiology*, Freeman, San Francisco, Calif, USA, 1972.
- [2] S. J. Gould and N. Eldredge, “Punctuated equilibrium comes of age,” *Nature*, vol. 366, no. 6452, pp. 223–227, 1993.

- [3] T. G. Barraclough, A. P. Vogler, and P. Harvey, "Revealing the factors that promote speciation," *Philosophical Transactions of the Royal Society B*, vol. 353, no. 1366, pp. 241–249, 1998.
- [4] J. B. C. Jackson and A. H. Cheetham, "Tempo and mode of speciation in the sea," *Trends in Ecology and Evolution*, vol. 14, no. 2, pp. 72–77, 1999.
- [5] T. G. Barraclough and V. Savolainen, "Evolutionary rates and species diversity in flowering plants," *Evolution*, vol. 55, no. 4, pp. 677–683, 2001.
- [6] M. J. Benton and P. N. Pearson, "Speciation in the fossil record," *Trends in Ecology and Evolution*, vol. 16, no. 7, pp. 405–411, 2001.
- [7] G. Hunt, "The relative importance of directional change, random walks, and stasis in the evolution of fossil lineages," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 47, pp. 18404–18408, 2007.
- [8] J. C. Avise and F. J. Ayala, "Genetic change and rates of cladogenesis," *Genetics*, vol. 81, no. 4, pp. 757–773, 1975.
- [9] J. C. Avise and F. J. Ayala, "Genetic differentiation in speciose versus depauperate phylads: evidence from the California minnows," *Evolution*, vol. 30, no. 1, pp. 46–58, 1976.
- [10] J. C. Avise, "Is evolution gradual or rectangular? Evidence from living fishes," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 74, no. 11, pp. 5083–5087, 1977.
- [11] D. P. Mindell, J. S. Sites, and D. Graur, "Speciation evolution: a phylogenetic test with allozymes in *Sceloporus* (Reptilia)," *Cladistics*, vol. 5, no. 1, pp. 49–61, 1989.
- [12] A. J. Webster, R. J. H. Payne, and M. Pagel, "Molecular phylogenies link rates of evolution and speciation," *Science*, vol. 301, no. 5632, p. 478, 2003.
- [13] M. Pagel, C. Venditti, and A. Meade, "Large punctuational contribution of speciation to evolutionary divergence at the molecular level," *Science*, vol. 314, no. 5796, pp. 119–121, 2006.
- [14] I. J. Leitch, J. M. Beaulieu, K. Cheung, L. Hanson, M. A. Lysak, and M. F. Fay, "Punctuated genome size evolution in Liliaceae," *Journal of Evolutionary Biology*, vol. 20, no. 6, pp. 2296–2308, 2007.
- [15] C. Venditti and M. Pagel, "Speciation as an active force in promoting genetic evolution," *Trends in Ecology and Evolution*, vol. 25, no. 1, pp. 14–20, 2010.
- [16] L. Arbiza, J. Dopazo, and H. Dopazo, "Positive selection, relaxation, and acceleration in the evolution of the human and chimp genome," *PLoS Computational Biology*, vol. 2, no. 4, Article ID e38, 2006.
- [17] J. Savard, D. Tautz, and M. J. Lercher, "Genome-wide acceleration of protein evolution in flies (Diptera)," *BMC Evolutionary Biology*, vol. 6, Article ID 7, 2006.
- [18] A. G. Clark, M. B. Eisen, D. R. Smith et al., "Evolution of genes and genomes on the *Drosophila* phylogeny," *Nature*, vol. 450, no. 7167, pp. 203–218, 2007.
- [19] T. M. Panhuis, N. L. Clark, and W. J. Swanson, "Rapid evolution of reproductive proteins in abalone and *Drosophila*," *Philosophical Transactions of the Royal Society B*, vol. 361, no. 1466, pp. 261–268, 2006.
- [20] W. Haerty, S. Jagadeeshan, R. J. Kulathinal et al., "Evolution in the fast lane: rapidly evolving sex-related genes in *Drosophila*," *Genetics*, vol. 177, no. 3, pp. 1321–1335, 2007.
- [21] D. A. Barbash, P. Awadalla, and A. M. Tarone, "Functional divergence caused by ancient positive selection of a *Drosophila* hybrid incompatibility locus," *PLoS Biology*, vol. 2, no. 6, Article ID E142, 2004.
- [22] N. J. Brideau, H. A. Flores, J. Wang, S. Maheshwari, X. Wang, and D. A. Barbash, "Two Dobzhansky-Muller Genes interact to cause hybrid lethality in *Drosophila*," *Science*, vol. 314, no. 5803, pp. 1292–1295, 2006.
- [23] J. P. Masly, C. D. Jones, M. A. F. Noor, J. Locke, and H. A. Orr, "Gene transposition as a cause of hybrid sterility in *Drosophila*," *Science*, vol. 313, no. 5792, pp. 1448–1450, 2006.
- [24] D. C. Presgraves, L. Balagopalan, S. M. Abmayr, and H. A. Orr, "Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*," *Nature*, vol. 423, no. 6941, pp. 715–719, 2003.
- [25] C. T. Ting, S. C. Tsaur, M. L. Wu, and C. I. Wu, "A rapidly evolving homeobox at the site of a hybrid sterility gene," *Science*, vol. 282, no. 5393, pp. 1501–1504, 1998.
- [26] J. Wittbrodt, D. Adam, B. Malitschek et al., "Novel putative receptor tyrosine kinase encoded by the melanoma-inducing Tu locus in *Xiphophorus*," *Nature*, vol. 341, no. 6241, pp. 415–421, 1989.
- [27] K. S. Zigler, M. A. McCartney, D. R. Levitan, and H. A. Lessios, "Sea urchin bindin divergence predicts gamete compatibility," *Evolution*, vol. 59, no. 11, pp. 2399–2404, 2005.
- [28] H. A. Orr, "The genetic basis of reproductive isolation: insights from *Drosophila*," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, supplement 1, pp. 6522–6526, 2005.
- [29] W. J. Swanson and V. D. Vacquier, "The rapid evolution of reproductive proteins," *Nature Reviews Genetics*, vol. 3, no. 2, pp. 137–144, 2002.
- [30] E. C. Metz and S. R. Palumbi, "Positive selection and sequence rearrangements generate extensive polymorphism in the gamete recognition protein bindin," *Molecular Biology and Evolution*, vol. 13, no. 2, pp. 397–406, 1996.
- [31] K. S. Zigler and H. A. Lessios, "Evolution of bindin in the pantropical sea urchin Tripneustes: comparisons to bindin of other genera," *Molecular Biology and Evolution*, vol. 20, no. 2, pp. 220–231, 2003.
- [32] C. G. Artieri, W. Haerty, and R. S. Singh, "Association between levels of coding sequence divergence and gene misregulation in *Drosophila* male hybrids," *Journal of Molecular Evolution*, vol. 65, no. 6, pp. 697–704, 2007.
- [33] R. M. Kliman, P. Andolfatto, J. A. Coyne et al., "The population genetics of the origin and divergence of the *Drosophila simulans* complex species," *Genetics*, vol. 156, no. 4, pp. 1913–1931, 2000.
- [34] K. Tamura, S. Subramanian, and S. Kumar, "Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks," *Molecular Biology and Evolution*, vol. 21, no. 1, pp. 36–44, 2004.
- [35] Y. Zhang, D. Sturgill, M. Parisi, S. Kumar, and B. Oliver, "Constraint and turnover in sex-biased gene expression in the genus *Drosophila*," *Nature*, vol. 450, no. 7167, pp. 233–237, 2007.
- [36] S. Kumar, A. Filipowski, V. Swarna, A. Walker, and S. B. Hedges, "Placing confidence limits on the molecular age of the human-chimpanzee divergence," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 52, pp. 18842–18847, 2005.
- [37] F. Rodríguez-Trelles, R. Tarrío, and F. J. Ayala, "A methodological bias toward overestimation of molecular evolutionary time scales," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 12, pp. 8112–8115, 2002.
- [38] J. D. Thompson, D. G. Higgins, and T. J. Gibson, "CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap

- penalties and weight matrix choice," *Nucleic Acids Research*, vol. 22, no. 22, pp. 4673–4680, 1994.
- [39] D. A. Pollard, V. N. Iyer, A. M. Moses, and M. B. Eisen, "Widespread discordance of gene trees with species tree in *Drosophila*: evidence for incomplete lineage sorting," *PLoS Genetics*, vol. 2, no. 10, p. e173, 2006.
- [40] A. Wong, J. D. Jensen, J. E. Pool, and C. F. Aquadro, "Phylogenetic incongruence in the *Drosophila melanogaster* species group," *Molecular Phylogenetics and Evolution*, vol. 43, no. 3, pp. 1138–1150, 2007.
- [41] Z. Yang, "PAML: a program package for phylogenetic analysis by maximum likelihood," *Computer Applications in the Biosciences*, vol. 13, no. 5, pp. 555–556, 1997.
- [42] Z. Yang and R. Nielsen, "Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages," *Molecular Biology and Evolution*, vol. 19, no. 6, pp. 908–917, 2002.
- [43] A. M. Larracuente, T. B. Sackton, A. J. Greenberg et al., "Evolution of protein-coding genes in *Drosophila*," *Trends in Genetics*, vol. 24, no. 3, pp. 114–123, 2008.
- [44] S. Holm, "A simple sequentially rejective multiple test procedure," *Scandinavian Journal of Statistics*, vol. 6, no. 2, pp. 65–70, 1969.
- [45] J. A. Coyne and A. H. Orr, *Speciation*, Sinauer Associates, Sunderland, Mass, USA, 2004.
- [46] R. S. Singh, "Toward a unified theory of speciation," in *Evolutionary Genetics from Molecules to Morphology*, R. S. Singh and C. Krimbas, Eds., pp. 570–608, 2000.
- [47] A. Civetta and R. S. Singh, "High divergence of reproductive tract proteins and their association with postzygotic reproductive isolation in *Drosophila melanogaster* and *Drosophila virilis* group species," *Journal of Molecular Evolution*, vol. 41, no. 6, pp. 1085–1095, 1995.
- [48] S. Jagadeeshan and R. S. Singh, "Rapidly evolving genes of *Drosophila*: differing levels of selective pressure in testis, ovary, and head tissues between sibling species," *Molecular Biology and Evolution*, vol. 22, no. 9, pp. 1793–1801, 2005.
- [49] D. J. Begun, A. K. Holloway, K. Stevens et al., "Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*," *PLoS Biology*, vol. 5, no. 11, p. e310, 2007.
- [50] J. F. Baines, S. A. Sawyer, D. L. Hartl, and J. Parsch, "Effects of X-linkage and sex-biased gene expression on the rate of adaptive protein evolution in *Drosophila*," *Molecular Biology and Evolution*, vol. 25, no. 8, pp. 1639–1650, 2008.
- [51] N. D. Singh, A. M. Larracuente, and A. G. Clark, "Contrasting the efficacy of selection on the X and autosomes in *Drosophila*," *Molecular Biology and Evolution*, vol. 25, no. 2, pp. 454–467, 2008.
- [52] J. Hey and R. M. Kliman, "Interactions between natural selection, recombination and gene density in the genes of *Drosophila*," *Genetics*, vol. 160, no. 2, pp. 595–608, 2002.
- [53] D. Lachaise, M. Harry, M. Solignac, F. Lemeunier, V. Benassi, and M. L. Cariou, "Evolutionary novelties in islands: *Drosophila santomea*, a new *Drosophila melanogaster* sister species from São Tomé," *Proceedings of the Royal Society B*, vol. 267, no. 1452, pp. 1487–1495, 2000.
- [54] J. Charlesworth and A. Eyre-Walker, "The other side of the nearly neutral theory, evidence of slightly advantageous back-mutations," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 104, no. 43, pp. 16992–16997, 2007.
- [55] S. Y. W. Ho, M. J. Phillips, A. Cooper, and A. J. Drummond, "Time dependency of molecular rate estimates and systematic overestimation of recent divergence times," *Molecular Biology and Evolution*, vol. 22, no. 7, pp. 1561–1568, 2005.
- [56] R. K. Butlin, "Recombination and speciation," *Molecular Ecology*, vol. 14, no. 9, pp. 2621–2635, 2005.
- [57] R. J. Kulathinal, S. M. Bennett, C. L. Fitzpatrick, and M. A. F. Noor, "Fine-scale mapping of recombination rate in *Drosophila* refines its correlation to diversity and divergence," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 105, no. 29, pp. 10051–10056, 2008.
- [58] M. Nei, "Bottlenecks, genetic polymorphism and speciation," *Genetics*, vol. 170, no. 1, pp. 1–4, 2005.
- [59] H. L. Carson, "The genetic system, the deme, and the origin of species," *Annual Review of Genetics*, vol. 21, pp. 405–423, 1987.
- [60] J. B. Losos, K. I. Warheit, and T. W. Schoener, "Adaptive differentiation following experimental island colonization in *Anolis* lizards," *Nature*, vol. 387, no. 6628, pp. 70–73, 1997.
- [61] J. B. Losos, T. W. Schoener, R. B. Langerhans, and D. A. Spiller, "Rapid temporal reversal in predator-driven natural selection," *Science*, vol. 314, no. 5802, p. 1111, 2006.
- [62] J. N. Thompson, "Rapid evolution as an ecological process," *Trends in Ecology and Evolution*, vol. 13, no. 8, pp. 327–329, 1998.
- [63] S. R'Kha, P. Capi, and J. R. David, "Host-plant specialization in the *Drosophila melanogaster* species complex: a physiological, behavioral, and genetical analysis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 88, no. 5, pp. 1835–1839, 1991.
- [64] I. Dworkin and C. D. Jones, "Genetic changes accompanying the evolution of host specialization in *Drosophila sechellia*," *Genetics*, vol. 181, no. 2, pp. 721–736, 2009.
- [65] M. L. Cariou, J. F. Silvain, V. Daubin, J. L. Da Lage, and D. Lachaise, "Divergence between *Drosophila santomea* and allopatric or sympatric populations of *D. yakuba* using paralogous amylase genes and migration scenarios along the Cameroon volcanic line," *Molecular Ecology*, vol. 10, no. 3, pp. 649–660, 2001.
- [66] A. J. Moehring, A. Llopart, S. Elwyn, J. A. Coyne, and T. F. C. Mackay, "The genetic basis of prezygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to mating preference," *Genetics*, vol. 173, no. 1, pp. 215–223, 2006.
- [67] A. J. Moehring, A. Llopart, S. Elwyn, J. A. Coyne, and T. F. C. Mackay, "The genetic basis of postzygotic reproductive isolation between *Drosophila santomea* and *D. yakuba* due to hybrid male sterility," *Genetics*, vol. 173, no. 1, pp. 225–233, 2006.
- [68] J. A. Coyne, S. Y. Kim, A. S. Chang, D. Lachaise, and S. Elwyn, "Sexual isolation between two sibling species with overlapping ranges: *Drosophila santomea* and *Drosophila yakuba*," *Evolution*, vol. 56, no. 12, pp. 2424–2434, 2002.
- [69] J. E. Blyth, D. Lachaise, and M. G. Ritchie, "Divergence in multiple courtship song traits between *Drosophila santomea* and *D. yakuba*," *Ethology*, vol. 114, no. 7, pp. 728–736, 2008.
- [70] M. G. Ritchie, "Evolution: feathers, females, and fathers," *Science*, vol. 318, no. 5847, pp. 54–55, 2007.