Letter:

Positive correlation between evolutionary rate and recombination rate in *Drosophila* genes with male-biased expression

Zhi Zhang and John Parsch

Department of Biology II, Section of Evolutionary Biology

University of Munich (LMU), Munich, Germany

**Corresponding author:**

John Parsch

Department of Biology II, Section of Evolutionary Biology

University of Munich (LMU)

Grosshaderner Str. 2, 82152 Munich, Germany

Email: parsch@zi.biologie.uni-muenchen.de

Phone: +49 89 2170 74 107, Fax: +49 89 2170 74 104

**Key words:** Adaptive evolution, Hill-Robertson interference, sex-biased genes

**Running head:** Evolutionary rate and recombination

Title length: 117 characters (including spaces)
Abstract length: 168 words
Total length of text (excluding Abstract): 10,294 characters (including spaces)
Total page requirement for all items: 3.75 pages
Number of references: 13
Abstract

Previous studies have shown that genes that are expressed predominantly or exclusively in males tend to evolve rapidly in comparison to other genes. In most cases, however, it is unknown whether this rapid evolution is the result of increased positive (or sexual) selection on male-expressed traits or if it is due to a relaxation of selective constraints. To distinguish between these two possibilities, we analyzed the relationship between the nonsynonymous substitution rate ($d_N$) and local recombination rate for 343 $Drosophila$ genes that were classified as male-, female-, or non-sex-biased in their expression. For the male-biased genes, a positive correlation between $d_N$ and recombination rate was observed. This can be explained by an increased rate of adaptive evolution in regions of higher recombination due to a reduction of Hill-Robertson interference. In contrast, the correlation between $d_N$ and recombination rate was negative for both female-, and non-sex-biased genes, suggesting that these genes are primarily subject to purifying selection, which is expected to be less effective in regions of reduced recombination.
Sexual dimorphism is common among higher eukaryotes and is thought to result from the differential action of natural (or sexual) selection on individuals of the two sexes. Darwin (1871) proposed that sexual selection, either through direct male-male competition or female mate choice, was responsible for the extravagant secondary sexual characteristics present in the males of many species. The rapid evolution of male reproductive traits also may play a role in the frequent occurrence of hybrid male sterility (Wu and Davis 1993). Modern molecular evolutionary studies suggest that sexual selection may affect a broad spectrum of sex-related genes (Civetta and Singh 1999; Singh and Kulathinal 2000; Swanson and Vacquier 2002). A recent study that used microarray data to identify \textit{Drosophila} genes with sex-biased expression and compared their rates of evolution between species found that genes with male-biased expression had significantly higher rates of nonsynonymous substitution ($d_N$) than genes with female- or non-sex-biased expression (Zhang, Hambuch, and Parsch 2004). The accelerated evolutionary rate of male-biased genes could have two different explanations. One possibility is that they are subject to less selective constraint than female- or non-sex-biased genes, allowing them to accumulate more neutral (or slightly deleterious) amino acid changes. Alternatively, male-biased genes could be frequent targets of positive (or sexual) selection, and thus accumulate more adaptive amino acid changes. An analysis of the available polymorphism data from \textit{D. melanogaster} supported the latter explanation: male-biased genes did not show an elevated level of nonsynonymous polymorphism as would be expected under the relaxed selective constraint hypothesis, but instead showed evidence of being subject to increased positive selection. However, there were several
limitations to this analysis that made this conclusion questionable. For example, the polymorphism comparison used a small number of sex-biased genes that were collected from a survey of the literature and, thus, included only a small fraction of the sex-biased genes in the genome. Furthermore, these genes likely were not a random sample because some were investigated with an a priori expectation of positive selection.

A further way to distinguish the evolutionary forces responsible for the increased $d_N$ in genes with male-biased expression is to examine the relationship between $d_N$ and local recombination rate. If most amino acid replacements are adaptive, then a positive correlation between $d_N$ and recombination rate is expected. This is because positive selection is more effective in regions of higher recombination due to a relaxation of Hill-Robertson interference among selected sites (Hill and Robertson 1966; Marais and Charlesworth 2003). In contrast, if most amino acid replacements are neutral (or slightly deleterious), then there should be no correlation (or a negative correlation) between $d_N$ and recombination rate. This is because purifying selection is less effective in regions of lower recombination for the same reason given above.

Betancourt and Presgraves (2002) investigated the relationship between $d_N$ and recombination rate for 255 orthologous genes of *D. melanogaster* and *D. simulans*. They found that genes in regions of high recombination had significantly higher $d_N$ than those in regions of low recombination. Their study focused primarily on a subset of male-expressed genes, namely those encoding accessory gland proteins (Acps), and suggested that most of the amino acid replacements in these genes were adaptive. When the Acp genes were removed from the analysis, there was no longer a significant difference in $d_N$ between genes in regions of high and low recombination. Recently, Marais et al. (2004)
investigated the relationship between $d_N$ and recombination rate for 630 genes compared between *D. melanogaster* and *D. yakuba*. For this data set, which was not enriched for male-expressed genes, a slightly negative correlation between $d_N$ and recombination rate was observed. The results of the two above studies suggest that the evolution of most genes is governed primarily by purifying selection, but that positive selection is responsible for the rapid evolution seen in male Acp genes. To investigate whether such adaptive evolution occurs in male-biased genes in general, we classified the genes used in Marais et al. (2004) as male-, female-, or non-sex-biased in their expression and examined the relationship between $d_N$ and recombination rate separately for each group.

Our final data set consisted of 343 genes for which both the ratio of male to female expression (Parisi et al. 2003; Ranz et al. 2003) and an estimate of the local recombination rate were available (Marais et al. 2004). Because comparable estimates of gene expression and local recombination rates were not available for *D. yakuba*, all of our estimates are for *D. melanogaster*. The average $d_N$ of the male-biased genes was about four-fold greater than that of the female-biased genes and about two-fold greater than that of the non-sex-biased genes (table 1). Comparisons of $d_N$ for all pairwise combinations of the three groups were significant (Mann-Whitney test, $P \leq 0.01$), and these results held whether a 1.5-fold, two-fold, or three-fold expression difference between the sexes was used to define genes as sex-biased. Similarly, the synonymous substitution rate ($d_S$) was significantly higher for male-biased genes than for female- and non-sex-biased genes over all cutoffs ($P \leq 0.01$). These results cannot be explained solely by an increased mutation rate in male-biased genes, because $d_N/d_S$ for male-biased genes was also significantly higher than that of both female- and non-sex-biased genes ($P \leq 0.05$; with
the exception of male- vs. non-sex-biased at 1.5-fold where \( P = 0.18 \). The observed differences in evolutionary rate are consistent with those reported previously for a partially overlapping set of genes compared between *D. melanogaster* and *D. yakuba* (Zhang, Hambuch, and Parsch 2004). In total, 161 genes are common to the two studies.

To examine the type of natural selection acting on genes of the three sex-biased expression classes, we looked at the relationship between \( d_N \) and local recombination rate within each class of genes. For this, we used eight different estimators of local recombination rate (Marais et al. 2004). All eight estimators were positively correlated with \( d_N \) for male-biased genes, and in most cases this correlation was significant (table 2). In contrast, all eight estimators were negatively correlated with \( d_N \) for female- and non-sex-biased genes, and this correlation was significant for many of the estimators (table 2). Furthermore, the correlation coefficients of the male- and female-biased genes were significantly different \(( P < 0.05 \) for all recombination estimators, and those for the male- and non-sex-biased genes were significantly different \(( P < 0.05 \) for seven of the eight estimators. Differences between the correlation coefficients of female- and non-sex-biased genes were not significant for any recombination rate estimator. We also observed a positive correlation between \( d_S \) and recombination rate for the male-biased genes (table 2). A possible explanation for this is that there is an elevated rate of mutation in male-biased genes in regions of high recombination. However, mutational effects alone cannot explain our results, because \( d_N/d_S \) is also positively correlated with recombination rate (table 2; fig. 1). Our results also cannot be explained by differences between X-linked and autosomal genes. Although male-biased genes are significantly underrepresented on
the X chromosome, all of the above results are unchanged when only autosomal genes are considered (not shown).

Our analyses indicate that male-biased genes are subject to different selective forces than female- and non-sex-biased genes. The positive correlation between $d_N$ and recombination rate seen for male-biased genes suggests that they are often targets of positive selection, which is expected to be more effective in genomic regions with higher recombination rates due to a reduction of Hill-Robertson interference. In contrast, the negative correlation between $d_N$ and recombination rate seen for female- and non-sex-biased genes suggests that these genes are predominantly subject to purifying selection, which is expected to be less effective in regions of lower recombination allowing fixation of more slightly deleterious mutations. The reduced efficacy of purifying selection in regions of reduced recombination is also expected to affect male-biased genes and would counteract the positive correlation between $d_N$ and recombination rate. Thus a significantly positive correlation, as is seen in our data, is a conservative criterion for the inference of positive selection.

Marais et al. (2004) observed a slightly negative correlation between $d_N$ and recombination rate for 630 genes compared between D. melanogaster and D. yakuba. Their analysis, however, did not consider male-, female-, and non-sex-biased genes separately. Since the vast majority of genes are female- or non-sex-biased (table 1), the negative correlation in these genes would obscure any positive correlation present in the male genes.

Our results are similar to those seen for male-specific Acp genes. However, the genes analyzed here came from a random EST survey and were classified only by their degree
of sex-biased expression. They are not enriched for genes of a particular functional class. In fact, none of the male-biased genes used in our analysis match annotated Acp genes or putative Acp genes identified in an accessory gland-specific EST screen (Swanson et al. 2001). Thus it appears that rapid evolution due to positive selection is a general feature of male-biased genes, and is not limited to a relatively small set of Acp genes.
Literature Cited


Table 1

Evolutionary Rates of Genes with Male-, Female-, and Non-sex-biased Expression

<table>
<thead>
<tr>
<th>Cutoff(^a)</th>
<th>Bias</th>
<th>(n^b)</th>
<th>(d_{s'})</th>
<th>(d_s)</th>
<th>(d_{s'}/d_s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5-fold</td>
<td>Male</td>
<td>68</td>
<td>0.037</td>
<td>0.381</td>
<td>0.094</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>159</td>
<td>0.012</td>
<td>0.248</td>
<td>0.048</td>
</tr>
<tr>
<td></td>
<td>Non-sex</td>
<td>115</td>
<td>0.026</td>
<td>0.284</td>
<td>0.071</td>
</tr>
<tr>
<td>2-fold</td>
<td>Male</td>
<td>32</td>
<td>0.043</td>
<td>0.381</td>
<td>0.114</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>119</td>
<td>0.012</td>
<td>0.233</td>
<td>0.046</td>
</tr>
<tr>
<td></td>
<td>Non-sex</td>
<td>192</td>
<td>0.021</td>
<td>0.305</td>
<td>0.069</td>
</tr>
<tr>
<td>3-fold</td>
<td>Male</td>
<td>17</td>
<td>0.045</td>
<td>0.393</td>
<td>0.116</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>75</td>
<td>0.010</td>
<td>0.208</td>
<td>0.042</td>
</tr>
<tr>
<td></td>
<td>Non-sex</td>
<td>252</td>
<td>0.021</td>
<td>0.303</td>
<td>0.069</td>
</tr>
</tbody>
</table>

\(^a\)Fold cutoff used to define genes as sex-biased.

\(^b\)Number of genes in each class.
Table 2

Correlation Between Substitution Rate and Recombination Rate for Genes with Male-, Female-, and Non-sex-biased Expression

<table>
<thead>
<tr>
<th>Bias</th>
<th>ACE</th>
<th>RTE</th>
<th>HKw</th>
<th>HKp</th>
<th>CC99</th>
<th>KH93</th>
<th>CK00</th>
<th>MMD01</th>
</tr>
</thead>
<tbody>
<tr>
<td>$d_N$</td>
<td>Male</td>
<td>0.24*</td>
<td>0.26*</td>
<td>0.37**</td>
<td>0.32**</td>
<td>0.21</td>
<td>0.21</td>
<td>0.25*</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>-0.14</td>
<td>-0.15</td>
<td>-0.18*</td>
<td>-0.18*</td>
<td>-0.12</td>
<td>-0.17*</td>
<td>-0.10</td>
</tr>
<tr>
<td></td>
<td>Non-sex</td>
<td>-0.04</td>
<td>-0.22*</td>
<td>-0.22*</td>
<td>-0.24**</td>
<td>-0.17</td>
<td>-0.23*</td>
<td>-0.14</td>
</tr>
<tr>
<td>$d_S$</td>
<td>Male</td>
<td>0.17</td>
<td>0.18</td>
<td>0.27*</td>
<td>0.26*</td>
<td>0.10</td>
<td>0.08</td>
<td>0.05</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>-0.13</td>
<td>-0.11</td>
<td>-0.11</td>
<td>-0.08</td>
<td>-0.12</td>
<td>-0.05</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>Non-sex</td>
<td>-0.10</td>
<td>-0.04</td>
<td>-0.04</td>
<td>-0.01</td>
<td>-0.06</td>
<td>-0.04</td>
<td>-0.05</td>
</tr>
<tr>
<td>$d_N/d_S$</td>
<td>Male</td>
<td>0.19</td>
<td>0.21</td>
<td>0.30*</td>
<td>0.28*</td>
<td>0.14</td>
<td>0.18</td>
<td>0.25*</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>-0.15</td>
<td>-0.14</td>
<td>-0.19*</td>
<td>-0.16*</td>
<td>-0.11</td>
<td>-0.16*</td>
<td>-0.13</td>
</tr>
<tr>
<td></td>
<td>Non-sex</td>
<td>-0.07</td>
<td>-0.19*</td>
<td>-0.19*</td>
<td>-0.21*</td>
<td>-0.10</td>
<td>-0.14*</td>
<td>-0.09</td>
</tr>
</tbody>
</table>

NOTE.—The eight estimators of local recombination rate correspond to those from Marais et al. (2004); Numbers indicate Pearson’s correlation coefficient. Genes were classified using the 1.5-fold cutoff (see table 1).

* $P < 0.05$; ** $P < 0.01$. 
Figure legend

FIG. 1.–Correlation between $d_{s}/d_{S}$ and local recombination rate (estimated by HKw [table 2] and using the 1.5-fold cutoff [table 1]) for genes with (A) male-biased ($R = 0.30, P = 0.02$), (B) female-biased ($R = -0.19, P = 0.02$), and (C) non-sex-biased expression ($R = -0.19, P = 0.04$). Least-squares linear regression lines are shown.
Fig. 1

Graphs A, B, and C show the relationship between recombination rate and dN/dS. Graph A displays a positive correlation, while Graph B shows a negative correlation. Graph C presents a scatter plot without a clear trend.