Genet. Res., Camb. (2003), **81**, *pp.* 89–90. © 2003 Cambridge University Press
DOI: 10.1017/S0016672302006067 Printed in the United Kingdom

Hill–Robertson interference in *Drosophila melanogaster*: reply to Marais, Mouchiroud and Duret

RICHARD M. KLIMAN[1] AND JODY HEY[2]*
[1] *Department of Biological Sciences, Cedar Crest College, 100 College Drive, Allentown, PA 18104, USA*
[2] *Department of Genetics, Rutgers University, 604 Allison Road, Piscataway, NJ 08865, USA*

Summary

The usage of preferred codons in *Drosophila melanogaster* is reduced in regions of lower recombination. This is consistent with population genetics theory, whereby the effectiveness of selection on multiple targets is limited by stochastic effects caused by linkage. However, because the selectively preferred codons in *D. melanogaster* end in C or G, it has been argued that base-composition-biasing effects of recombination can account for the observed relationship between preferred codon usage and recombination rate (Marais *et al.*, 2003). Here, we show that the correlation between base composition (of protein-coding and intron regions) and recombination rate holds only for lower values of the latter. This is consistent with a Hill–Robertson interference model and does not support a model whereby the entire effect of recombination on codon usage can be attributed to its potential role in generating compositional bias.

Marais *et al.* (2003) note that, because the preferred codons in *Drosophila melanogaster* end in either C or G, regional variation in mutation pressure will lead to regional variation in the usage of preferred codons. Because the process of crossing-over can increase the usage of C or G (Eyre-Walker, 1993; Birdsell, 2002), a positive correlation may arise between the recombination rate and the usage of preferred codons. However, because the effectiveness of natural selection is reduced, in principle, by insufficient opportunity for recombination among targets of selection (the Hill–Robertson effect) (Hill & Robertson, 1966; Felsenstein, 1974), the same correlation is predicted.

Marais *et al.* argue that the mutation pressure is sufficient to explain the positive correlation we observed between estimates of recombination rate and preferred codon usage in *D. melanogaster* (Hey & Kliman, 2002; see also Marais *et al.*, 2001); that is, we can not justify the argument that the correlation indicates Hill–Robertson interference. They suggest that a flawed measure of local compositional bias might be responsible, in part, for the results of our analyses. Specifically, we used two different measures of noncoding G+C content: (1) intron G+C for genes with introns; and (2) flanking G+C for genes lacking introns. Marais *et al.* point out that flanking G+C content is slightly higher than intron G+C content. However, there is no association between our measure of recombination rate (R) and the presence or absence of introns ($R_{\text{presence}} = 2{\cdot}431$; $R_{\text{absence}} = 2{\cdot}416$; $F_{2,12997} = 0{\cdot}220$; $p = 0{\cdot}639$). Thus, it is unlikely that the mixed measure of noncoding G+C content contributed to the positive correlation between codon bias (corrected for gene length and noncoding G+C content) and R (Hey & Kliman, 2002).

Regardless, we have performed new analyses using only those genes with average intron lengths of at least 100 bp ($N = 5900$). Following arcsine-root transformation of frequency data, third-codon-position G+C content (*GC3*) was regressed on gene length and intron G+C (*GCnc*). The correlation between the residuals and R is 0·0866 ($p < 0{\cdot}001$). When the genes are separated into those with $R > 2{\cdot}0$ and those with $R < 2{\cdot}0$, the results are striking. The correlation (r) for the low-recombination genes is 0·218 ($N = 2384$, $p < 0{\cdot}001$), whereas the correlation for high-recombination genes is slightly, but not significantly, negative ($r = -0{\cdot}027$, $N = 3516$, $p = 0{\cdot}105$). This begs the question of why, if the compositional bias effect of recombination rate is responsible for the variation in third-position G+C content (and, by extension, codon bias), there is no correlation at higher recombination rates. However, these results are easily compatible with a model in which Hill–Robertson effects

* Corresponding author. e-mail: hey@biology.rutgers.edu

Table 1. *Pearson's correlations for recombination rate and G + C content*

Genes	*R* vs *GCnc*	*R* vs *GC3*	*GCnc* vs *GC3*
All ($N = 5900$)	$r = 0{\cdot}0299$ ($p = 0{\cdot}022$)	$r = 0{\cdot}0820$ ($p < 0{\cdot}001$)	$r = 0{\cdot}2336$ ($p < 0{\cdot}001$)
$R < 2{\cdot}0$ ($N = 2384$)	$r = 0{\cdot}0791$ ($p < 0{\cdot}001$)	$r = 0{\cdot}2264$ ($p < 0{\cdot}001$)	$r = 0{\cdot}2760$ ($p < 0{\cdot}001$)
$R > 2{\cdot}0$ ($N = 3516$)	$r = -0{\cdot}0289$ ($p = 0{\cdot}087$)	$r = -0{\cdot}0379$ ($p = 0{\cdot}025$)	$r = 0{\cdot}1789$ ($p < 0{\cdot}001$)

are essentially mitigated once *R* exceeds some threshold value.

Marais *et al*. also argue that *GCnc* is not a good measure for mutation pressure. However, at present, we really have no direct measure of mutation pressure and must rely on a correlated proxy. If variance in mutation pressure is due largely to variance in recombination rate, then both *GCnc* and *GC3* should covary positively with *R*. These correlations are, in fact, positive for the 5900 genes analysed above. However, *GC3* and *GCnc* correlate with each other more strongly than either correlates with *R* (Table 1), suggesting that compositional influences other than recombination rate affect coding and noncoding regions alike. When we consider only the low-recombination genes, the correlation between the two G + C contents is greater than the correlation between *R* and either G + C content. In the high-recombination genes, the two G + C contents are positively correlated, but both are slightly negatively correlated to *R* (Table 1). Thus, it is difficult to argue that *R* serves as a better proxy for mutation pressure on *GC3* (and codon usage) than does *GCnc*.

Finally, we question the 'simple test' of the mutation-pressure model vs the Hill–Robertson model. Marais *et al*. argue that the former predicts that both *GCnc* and *GC3* should be affected by recombination rate, whereas the latter predicts only an effect on *GC3*. The underlying assumption is that only *GC3* is subject to selection. However, there appears to be a fairly strong overall mutation pressure in *D. melanogaster*, as in many organisms, towards A and T. Selection on functional regions of noncoding DNA would help to maintain G and C at some constrained positions. Thus, even *GCnc* could be influenced by Hill–Robertson effects.

We agree with Marais *et al*. that regional variation in mutation pressure probably contributes to among-gene variation in preferred codon usage. However, the analyses presented here, in Hey & Kliman (2002) and recently by Marais & Piganeau (2002) support the hypothesis that Hill–Robertson effects contribute to variance in codon bias in *D. melanogaster*.

This research was supported by National Institutes of Health grants R01GM54684 to J.H. and R15HG02456 to R.M.K.

References

Birdsell, J. A. (2002). Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Molecular Biology and Evolution* **19**, 1181–1197.

Eyre-Walker, A. (1993). Recombination and mammalian genome evolution. *Proceedings of the Royal Society of London Series B* **252**, 237–243.

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics* **78**, 737–756.

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

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

Marais, G., Mouchiroud, D. & Duret, L. (2001). Does recombination improve selection on codon usage? Lessons from nematode and fly complete genomes. *Proceedings of the National Academy of Sciences of the USA* **98**, 5688–5692.

Marais, G. & Piganeau, G. (2002). Hill–Robertson interference is a minor determinant of variations in codon bias across *Drosophila melanogaster* and *Caenorhabditis elegans* genome. *Molecular Biology and Evolution* **19**, 1399–1406.

Marais, G., Mouchiroud, D. & Duret, L. (2003). Neutral effect of recombination on base composition in *Drosophila*. *Genetical Research* **81**, 79–87.