and the amount of genetic variation in *D. melanogaster*[6].

To justify using evidence from DNA sequence variation to show that there has been a selective sweep involving *Sdic* and *Cdic*, it is necessary (if not sufficient) to demonstrate that there is less variation in these genes than would be expected from the background selection model for genes in this region of the X chromosome. This was not done by Nurminsky *et al.*[1], and indeed it is difficult to do. For example, although these genes are in a region where recombination is reduced in frequency compared with the middle of the X chromosome, the relation between the rate of recom-bination and the location is not known with any accuracy[2]. It is therefore not easy to predict the expected degree of variation under any model, in contrast to the tip of the X chromosome where the gradient of recombination frequency is better known[6].

The gene *Zw* is located at position D1 on the X chromosome, just distal to *Sdic* and *Cdic*, and has a nucleotide-site diversity (π) of 3.8×10^{-3} (ref. 7); the gene *su(f)* is located at 20E–F and has a π value of 0.5×10^{-3} (ref. 8). The π values for *Sdic* and *Cdic* (0.89×10^{-3} and 0.45×10^{-3}, respectively) are closer to those of *su(f)* than to that of Zw, but the large standard deviations of these estimates (0.73×10^{-3} and 0.50×10^{-3}, respectively) mean the true π values may lie between those for *Zw* and *su(f)*. Differences in selective constraints on different genes may also contribute to differences among loci; these can be accounted for by calibrating with respect to interspecific sequence comparisons[9], which was not done here[1].

The effects of selective sweeps can be detected from departures of frequencies of variants from neutral expectation, as measured by statistics such as Tajima's *D* (ref. 10). The small numbers of variant sites at *Sdic* and *Cdic* mean that such tests lack power in this case. Although Tajima's *D* values are negative for both loci (-0.085 and -0.104, respectively), indicating an excess of rare variants (as expected after a selective sweep[10]), their magnitudes are far below those needed for statistical significance.

We believe, therefore, that the data presented by Nurminsky *et al.* do not provide convincing evidence for a recent selective sweep. This does not imply that *Sdic* has not been fixed by selection. It is possible that the fixation event took place relatively soon after the divergence of *D. melanogaster* and *D. simulans*, and that variability in this region has since returned to that expected for its degree of recombination. Only much better characterization of the levels of genetic diversity and recombination frequencies in this region of the X chromosome can resolve this question.

Brian Charlesworth, Deborah Charlesworth
Institute for Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT, UK
e-mail: brian.charlesworth@ed.ac.uk

1. Nurminsky, D. I., Nurminskaya, M. V., De Aguiar, D. & Hartl, D. L. *Nature* **396,** 572–575 (1998).
2. Schalet, A. & Lefevre, G. in *The Genetics and Biology of* Drosophila (eds Ashburner, M. & Novitski, E.) 847–902 (Academic, London, 1976).
3. Begun, D. J. & Aquadro, C. F. *Nature* **356,** 519–520 (1992).
4. Charlesworth, B., Morgan, M. T. & Charlesworth, D. *Genetics* **134,** 1289–1303 (1993).
5. Gillespie, J. H. *Gene* **205,** 291–299 (1997).
6. Charlesworth, B. *Genet. Res.* **68,** 131–150 (1996).
7. Eanes, W. F., Kirchner, M. & Yoon, J. *Proc. Natl Acad. Sci. USA* **90,** 7475–7479 (1993).
8. Langley, C. H., MacDonald, J., Miyashita, N. & Aguadé, M. *Proc. Natl Acad. Sci. USA* **90,** 1800–1803 (1993).
9. Hudson, R. R., Kreitman, M. & Aguadé, M. *Genetics* **116,** 153–159 (1987).
10. Simonsen, K. L., Churchill, G. A. & Aquadro, C. F. *Genetics* **141,** 413–429 (1995).

Nurminsky and Hartl reply — One hallmark of persistent strong background selection is a severe diminution of codon usage bias[1–3]. An example is the *Drosophila* gene *rolled* (gene 1 in Fig. 1), which is located in the centromeric heterochromatin of chromosome 2, where recombination is severely restricted. Other genes shown in Fig. 1 are located near the base of the X chromosome. Genes 10 and 11 are *AnnX* and *Cdic*, respectively, which flank the *Sdic* gene[4]. As pointed out by Charlesworth and Charlesworth, gene 2, which is *su(f)*, has much less DNA sequence variation than gene 17 (*Zw*). This difference is consistent with the discordant levels of codon usage bias and suggests strong background selection at *su(f)* but not at *Zw*. The degree of codon usage bias shows an extremely sharp increase as the gene locations progress outwards from *su(f)*. The transition is near genes 5 (*sol*) and 6 (*slgA*), which are proximal to the *Cdic–AnnX* region.

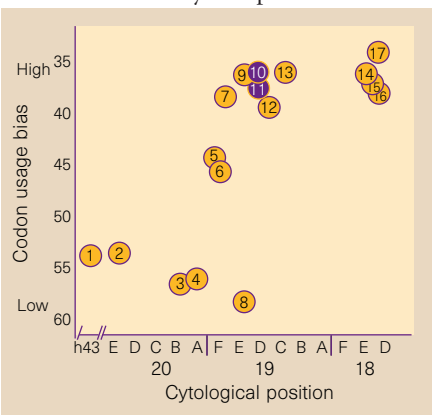

Figure 1 Codon usage bias is scaled according to the effective number of codons[7]. The vertical axis is inverted because a smaller effective number of codons corresponds to a greater bias in codon usage. A similar pattern is seen with other measures of codon bias, such as the χ^2/L statistic[8] (data not shown). The genes and their accession numbers are: 1, *rl* (M95124); 2, *su(f)* (X62679); 3, *S6kII* (L28945); 4, *fog* (U03717); 5, *sol* (M64084); 6, *slgA* (L07330); 7, *dod* (U35140); 8, *shakB* (U17330); 9, *run* (X56432); 10, *AnnX* (M34069); 11, *Cdic* (AF070699); 12, *Pbprp2* (U05981); 13, *Pp4-19C* (Y14213); 14, *Mer* (U49724); 15, *Cdc42* (U11824); 16, *Bap* (X75910); and 17, *Zw* (M26673, M26674).

The result is that *AnnX* and *Cdic* are located in a region of codon usage bias similar to that of *Zw*. The cytological region 19DE might therefore support a level of DNA sequence variation of *Sdic* and *Cdic* comparable to that of *Zw*. However, based on the amount of polymorphism observed for *Zw*, the probability of obtaining a value as low or lower than that for *Sdic* and *Cdic* is about 0.043 and 0.008, respectively. These estimates are based on 10,000 simulations using Watterson's formula[5] for pairwise mismatches in the infinite-alleles model with no recombination, so they should be conservative. There seems to be a statistically significant difference between *Zw* and the other two genes.

The evolution of *Sdic* required an initial duplication and gene fusion accompanied or followed by three deletions, two more insertions/deletions (including one that created a new splice junction), 11 nucleotide substitutions (including reversal of a chain-terminating codon), and a ten-fold tandem reiteration of the *Sdic* coding sequence. Although all these changes may have occurred shortly after the divergence of *D. melanogaster* and *D. simulans*, the similarity in degree of codon usage bias of *Cdic* and *Zw*, contrasted with the significant discrepancy in their levels of nucleotide diversity, provides independent evidence in support of our original inference of at least one relatively recent selective sweep. The negative values of Tajima's *D* statistic[6] for *Sdic* and *Cdic* also support this idea, notwithstanding their lack of statistical significance.

Charlesworth and Charlesworth are correct in pointing out that all the genes in the region 19DE might have limited DNA sequence variation as a result of background selection, despite the higher than predicted level of codon usage bias indicated in Fig. 1. We agree that a much more complete characterization of the levels of genetic diversity and recombination in this region would be informative.

Dmitry I. Nurminsky, Daniel L. Hartl
Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA

1. Kliman, R. M. & Hey, J. *Mol. Biol. Evol.* **10,** 1239–1258 (1994).
2. Munte, A., Aguadé, M. & Segarra, C. *Genetics* **147,** 165–175 (1997).
3. Kreitman, M. E. & Antezana, M. in *Evolutionary Genetics from Molecules to Morphology* (eds Singh, R. S. & Krimbas, C.) (Cambridge Univ. Press, New York, in the press).
4. Nurminsky, D. I., Nurminskaya, M. V., De Aguiar, D. & Hartl, D. L. *Nature* **396,** 572–575 (1998).
5. Watterson, G. A. *Theor. Popul. Biol.* **7,** 256–276 (1975).
6. Tajima, F. *Genetics* **123,** 585–595 (1989).
7. Wright, F. *Gene* **87,** 23–29 (1990).
8. Shields, D. C., Sharp, P. M., Higgins, D. G. & Wright, F. *Mol. Biol. Evol.* **5,** 704–716 (1988).