

Letter to the Editors

Recombination Rates May Affect the Ratio of *X* to Autosomal Noncoding Polymorphism in African Populations of *Drosophila melanogaster*

Beatriz Vicoso and Brian Charlesworth¹

Institute of Evolutionary Biology, School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3JT, United Kingdom

Manuscript received October 31, 2008

IN African populations of *Drosophila melanogaster*, the level of silent variability on the *X* chromosome often exceeds three-fourths of the autosomal value (the ratio expected for neutral equilibrium), suggesting that demographic or selective effects may influence variability (ANDOLFATTO 2001; KAUER *et al.* 2002; MOUSSET and DEROME 2004; HUTTER *et al.* 2007; POOL and NIELSEN 2007; SINGH *et al.* 2007; POOL and NIELSEN 2008). Although these sites are not completely neutral (ANDOLFATTO 2005; HADDRILL *et al.* 2005; HALLIGAN and KEIGHTLEY 2006), they are less constrained than coding sites and are often used as a neutral proxy.

The level of variability of neutral mutations can be affected by selection at neighboring sites: neutral variants will be removed from the population if they are linked to deleterious mutations or swept to fixation if they are linked to beneficial variants (GORDO and CHARLESWORTH 2001). These effects are greatest when recombination rates are low, consistent with data on the relation between recombination and variability in *D. melanogaster* (BEGUN and AQUADRO 1992; SHAPIRO *et al.* 2007). Differences in recombination rates between the *X* chromosome and the autosomes therefore could affect the relative values of *X* chromosomal and autosomal diversities in *Drosophila*. The expected magnitude of this difference depends on the type of selective effect involved (selective sweeps *vs.* background selection), on the location of the genes being compared (high *vs.* low recombination regions), and on the selection coefficients and dominance coefficients of the variants under selection. This makes the expected magnitude hard to predict, but we note that CHARLESWORTH (1996, p. 139) showed that background selection could produce an effect on *X*-linked loci that yields a maximum diversity level for the highest recombining regions of the chromosome that was between 1.28 and 1.05 times that of an

autosomal arm, depending on which arm is used for the comparison.

A recent large-scale study of noncoding polymorphisms on the *X* and the autosomes (HUTTER *et al.* 2007) has confirmed that the levels of silent polymorphism were higher on the *X* chromosome than on the autosomes for African populations. To bypass the recombination problem, local recombination rates were estimated for *X*-linked and autosomal loci using the method of COMERON *et al.* (1999) and found to have very similar means; a difference in recombination was therefore excluded as the cause of this elevated π_X/π_A . However, this approach suffers from two problems that may underestimate the effect of recombination. First, recombination rates were estimated on the basis of comparisons of standard genetic maps with cytogenetic locations, which may lead to less precise estimates than recent estimates of physical positions, which are based on the whole-genome sequence. Second, since recombination does not occur in *D. melanogaster* males, these genetic maps were based on crossing over in females. It has been pointed out that for *X* chromosome–autosome comparisons recombination rates should be corrected to account for the fact that the *X* chromosome spends more time in females than do autosomes (LANGLEY *et al.* 1988; CONNALLON 2007). We have reanalyzed this data set using the genome-based recombination estimates described by SINGH *et al.* (2005) to test if π_X differs from $3\pi_A/4$ in the study of HUTTER *et al.* (2007), once recombination levels have been taken into account using these estimates. SINGH *et al.* (2005) followed a similar approach to COMERON *et al.* (1999) to estimate recombination rates but instead used the genome sequence to establish the physical locations of the genes used in the genetic maps, which should lead to more precise estimates. Their methods are explained on their website (<http://www.stanford.edu/~lipatov/recombination/methods.html>) and suggest that their estimates are reliable (the fit of their Marey maps overall was very good, with the exception of the extreme tip of the *X* chromosome). We multiplied the values for *X*-

¹Corresponding author: Institute of Evolutionary Biology, Ashworth Laboratories, King's Buildings, West Mains Rd., Edinburgh EH9 3JT, United Kingdom. E-mail: brian.charlesworth@ed.ac.uk

TABLE 1

Mean noncoding polymorphism levels, number of autosomal and Xlinked loci used, and mean effective recombination rates in the HUTTER *et al.* (2007) data for African populations of *D. melanogaster*

| | Whole-sample comparison | |
|--------------------------------|--|-----------------|
| | Autosomes | X chromosome |
| Mean π (SE) | 0.0104 (0.0010) | 0.0117 (0.0004) |
| Mean Watterson's θ (SE) | 0.0114 (0.0011) | 0.0134 (0.0005) |
| No. of genes | 53 | 249 |
| Mean recombination rate (SE) | 1.37 (0.04) | 2.44 (0.02) |
| | Overlapping range of effective recombination rates | |
| | Autosomes | X chromosome |
| Mean π (SE) | 0.0115 (0.0014) | 0.0074 (0.0012) |
| Mean Watterson's θ (SE) | 0.0124 (0.0015) | 0.0088 (0.0011) |
| No. of genes | 32 | 14 |
| Mean recombination rate (SE) | 1.58 (0.01) | 1.55 (0.02) |

The recombination rates were obtained using the method of SINGH *et al.* (2005). Recombination estimates for Xlinked loci (centimorgans per megabase) were multiplied by two-thirds, autosomal ones by one-half. SE is the standard error. The overlapping range of recombination starts at the lowest recombination rate for an Xlinked locus (1.387 cM/Mb) and ends at the highest autosomal recombination rate (1.665 cM/Mb).

linked loci by two-thirds, and those of the autosomal loci by one-half, to obtain the effective recombination rate (LANGLEY *et al.* 1988; CONNALLON 2007).

We note, however, that methods based on standard map positions may be misleading, since these do not necessarily accurately reflect experimentally measured recombination frequencies (CHARLESWORTH 1996; BEGUN *et al.* 2007). Direct estimates of recombination based on high-density SNP maps would be preferable for this purpose, as are available for parts of the *D. pseudoobscura* genome (KULATHINAL *et al.* 2008). In addition, the presence of inversions (which are predominantly autosomal) may cause frequencies of crossing over on the X chromosome in the wild to be higher than laboratory measurements due to the interchromosomal effect of inversion heterozygosity on recombination (SCHULTZ and REDFIELD 1951; CHARLESWORTH 1996).

The results show that we cannot exclude differences in recombination rates as the cause of the higher polymorphism level for the X chromosome in their data (Table 1), as the ranges of recombination levels of X-linked and autosomal loci barely overlap, making the comparison between the two unreliable. On repeating the analysis using only loci that fall within overlapping ranges of effective recombination rate (1.39–1.67 cM/Mb), the mean noncoding diversity at X-linked loci is lower than three-fourths of the autosomal diversity ($\pi_X = 0.64\pi_A$; see Table 1). The results are unchanged if loci at the tip of the X chromosome are removed (supplemental material).

Because using different estimates of recombination can lead to rather different patterns, we reanalyzed the sample using the estimates of recombination described

in HEY and KLIMAN (2002; available at <http://lifesci.rutgers.edu/~heylab/>). Although the results varied depending on the estimates of recombination used (supplemental material), the main conclusion held for all of them: mean recombination rates for Xlinked loci in the sample were on average higher than for autosomal loci (the X:A ratio estimates ranged from 1.4 to 2.1), indicating the need to carefully control for rates of recombination before considering other hypotheses to account for π_X/π_A .

We thank David Begun, Chuck Langley, and two referees for their comments on the manuscript. This work was funded by a Portuguese Foundation for Science and Technology scholarship to B.V., and B.C. was supported by the Royal Society.

LITERATURE CITED

- ANDOLFATTO, P., 2001 Contrasting patterns of X-linked and autosomal nucleotide variation in *Drosophila melanogaster* and *Drosophila simulans*. *Mol. Evol. Biol.* **18**: 279–290.
- ANDOLFATTO, P., 2005 Adaptive evolution of non-coding DNA in *Drosophila*. *Nature* **437**: 1149–1152.
- BEGUN, D. J., and C. F. AQUADRO, 1992 Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature* **356**: 519–520.
- BEGUN, D. J., A. K. HOLLOWAY, K. STEVENS, L. W. HILLIER, Y. P. POH *et al.*, 2007 Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. *PLoS Biol.* **5**: e310.
- CHARLESWORTH, B., 1996 Background selection and patterns of genetic diversity in *Drosophila melanogaster*. *Genet. Res.* **68**: 131–149.
- COMERON, J. M., M. KREITMAN and M. AGUADÉ, 1999 Natural selection on synonymous sites is correlated with gene length and recombination in *Drosophila*. *Genetics* **151**: 239–249.
- CONNALLON, T., 2007 Adaptive protein evolution of X-linked and autosomal genes in *Drosophila*: implications for faster-X hypotheses. *Mol. Evol. Biol.* **24**: 2566–2572.
- GORDO, I., and B. CHARLESWORTH, 2001 Genetic linkage and molecular evolution. *Curr. Biol.* **11**: R684–R686.

- HADDRILL, P. R., B. CHARLESWORTH, D. L. HALLIGAN and P. ANDOLFATTO, 2005 Patterns of intron sequence evolution in *Drosophila* are dependent upon length and GC content. *Genome Biol.* **6**: R67.
- HALLIGAN, D. L., and P. D. KEIGHTLEY, 2006 Ubiquitous selective constraints in the *Drosophila* genome revealed by a genome-wide interspecies comparison. *Genome Res.* **16**: 875–884.
- HEY, J., and R. KLIMAN, 2002 Interactions between natural selection, recombination and gene density in the genes of *Drosophila*. *Genetics* **160**: 595–608.
- HUTTER, S., H. LI, S. BEISSWANGER, D. DE LORENZO and W. STEPHAN, 2007 Distinctly different sex ratios in African and European populations of *Drosophila melanogaster* inferred from chromosome-wide single nucleotide polymorphism data. *Genetics* **177**: 469–480.
- KAUER, M., B. ZANGERL, D. DIERINGER and C. SCHLOTTERER, 2002 Chromosomal patterns of microsatellite variability contrast sharply in African and non-African populations of *Drosophila melanogaster*. *Genetics* **160**: 247–256.
- KULATHINAL, R. J., S. M. BENNETT, C. L. FITZPATRICK and M. A. F. NOOR, 2008 Fine-scale mapping of recombination rate in *Drosophila* refines its correlation to diversity and divergence. *Proc. Natl. Acad. Sci. USA* **105**: 10051–10056.
- LANGLEY, C. H., E. A. MONTGOMERY, R. H. HUDSON, N. L. KAPLAN and B. CHARLESWORTH, 1988 On the role of unequal exchange in the containment of transposable element copy number. *Genet. Res.* **52**: 223–235.
- MOUSSET, S., and N. DEROME, 2004 Molecular polymorphism in *Drosophila melanogaster* and *D. simulans*. What have we learned from recent studies? *Genetica* **120**: 79–86.
- POOL, J., and R. NIELSEN, 2007 Population size changes reshape genomic patterns of diversity. *Evolution* **61**: 3001–3006.
- POOL, J. E., and R. NIELSEN, 2008 The impact of founder events on chromosomal variability in multiply mating species. *Mol. Evol. Biol.* **25**: 1728–1736.
- SCHULTZ, J., and H. REDFIELD, 1951 Interchromosomal effects on crossing over in *Drosophila*. *Cold Spring Harb. Symp. Quant. Biol.* **16**: 175–197.
- SHAPIRO, J. A., W. HUANG, C. ZHANG, M. J. HUBISZ, J. LU *et al.*, 2007 Adaptive genic evolution in the *Drosophila* genomes. *Proc. Natl. Acad. Sci. USA* **104**: 2271–2276.
- SINGH, N. D., J. C. DAVIS and D. A. PETROV, 2005 X-linked genes evolve higher codon bias in *Drosophila* and *Caenorhabditis*. *Genetics* **171**: 145–155.
- SINGH, N., J. MACPHERSON, J. JENSEN and D. PETROV, 2007 Similar levels of X-linked and autosomal nucleotide variation in African and non-African populations of *Drosophila melanogaster*. *BMC Evol. Biol.* **7**: 202.

Communicating editor: D. BEGUN