Variability levels, population size and structure of American and European *Drosophila montana* populations

JORGE VIEIRA*† & ANNELI HOIKKALA‡

†Institute of Cell, Animal and Population Biology, University of Edinburgh, Edinburgh EH9 3JT U.K.
and ‡Department of Biology, University of Oulu, P.O. Box 3000, FIN-90401 Oulu, Finland

The level and patterns of nucleotide diversity have been characterized for two X-linked loci, *fused* (*fu*; a region of 2362 bp) and *suppressor of sable* (*su(s)*; a region of 413 bp), in one European and one American *D. montana* population. Sequence variation at these loci shows that the two populations are divergent, although they may not be completely isolated. Data on the level of silent site variability at *su(s)* (1.1% and 0.5% for the European and American populations, respectively) suggest that the effective population sizes of the two populations may be similar. At the *fused* locus, one European sequence was highly divergent and may have resulted from gene conversion, and was excluded from the analysis. With this sequence removed, the level of silent site variability was significantly lower in the European population (0.28%) than in the American population (2.3%), which suggests a selective sweep at or near *fu* in the former population.

Keywords: DNA sequence variation, *Drosophila montana*, *fused*, population structure.

Introduction

Knowledge of the level and patterns of nucleotide polymorphisms within and between conspecific populations can give information about population size and structure as well as about the action of natural selection on DNA and protein sequences (Kimura, 1983; Ohta, 1992; Kreitman & Akashi, 1995). The efficiency of selection is dependent on the effective population size, which in turn is positively correlated with the level of intraspecific polymorphism at neutral sites (Kimura, 1983). Thus it can be possible to detect even very weak selection in very large populations (Kimura, 1983; McVean & Vieira, 1999).

Drosophila montana (a member of the *virilis* group of the subgenus *Drosophila*) is distributed around the northern hemisphere from the Atlantic coast of North America across Canada, Japan and northern Asia to Scandinavia (Throckmorton, 1982). Populations of this species exhibit variation in chromosome structure (Moorhead, 1954; Stone *et al.*, 1960) and morphological characters (Lakovaara & Hackman, 1973; Watabe &

Higuchi, 1979). Allozyme variability studies have been conducted so far only on North American *D. montana* populations (Baker, 1975, 1980). Thus it is not known whether there is a single world-wide *D. montana* population, and what is the level of polymorphism species-wide. This is despite European *D. montana* populations having been fairly well characterized for some ecological parameters and behavioural variability (see for instance Aspi & Lankinen, 1992; Suvanto *et al.*, 1994). In order to address this issue, we have characterized the level and patterns of nucleotide diversity for two X-linked genes, *fused* (*fu*) and *suppressor of sable* (*su(s)*) in one American and one European *D. montana* population.

Materials and methods

Two *D. montana* populations were analysed, one from Cache County, Utah (U.S.A.; 8 males), and another from Oulanka (Finland; 25 males). All males were collected in the field during summer 1999. We also used isofemale lines from both locations (three lines from Utah and two lines from Oulanka) to check possible changes in gene arrangement on the X chromosomes in the two populations.

Genomic DNA extraction was performed as described in Vieira & Charlesworth (1999). A 2.4-kb *fu* and

*Correspondence and present address: Departamento de Genética Molecular, Instituto de Biologia Molecular e Celular, Universidade do Porto, Rua do Campo Alegre 823, Porto 4150-180, Portugal. E-mail: jbvieira@ibmc.up.pt

469 bp *su(s)* PCR product was obtained from the genomic DNA of single males using the primers FUF and FU4IR (Vieira & Charlesworth, 2000) and SU(S)F and SU(S)R (Vieira & Charlesworth, 1999), respectively. Standard PCR amplification conditions were 30 cycles of denaturation at 94° for 30 s, primer annealing for 30 s at 52°, and primer extension at 72° for 2–3 min. Sequencing of both strands of these PCR products and analysis of DNA polymorphism was performed as in Vieira & Charlesworth (1999, 2000). GenBank accession numbers for *fu* are AY014455–AY014487 and for *su(s)* are AY014488–AY014503.

Results

X chromosome nucleotide variability levels

Figure 1 shows the haplotype structure of a *D. montana* population from Oulanka (O1–O25 sequences) and Utah, U.S.A. (U2–U4, U6–U9 and U11 sequences). The haplotypes are based on a 2.4-kb *fu* genomic DNA fragment, which includes most of the coding region of this gene, the four introns, and a small part of the 5' flanking region. Visual inspection of the sequence data

in Fig. 1 reveals that in the Oulanka population, the individual O25 alone defines 12 new polymorphic sites relative to the remaining 24 sequences, 11 of these sites being confined to a small 145 bp region (Fig. 1, sites 1573–1718). This pattern is suggestive of gene conversion between two divergent sequences. Since nine out of the 11 polymorphic sites confined to region 1573–1718, are present in the Utah population, it is possible that we have detected a gene conversion event between an American-like and a European-like *fu* sequence. That O25 is a sequence of a *D. montana* male has previously been confirmed by recording the male's courtship song and by testing whether it mates with *D. montana* females (all Oulanka males were run through these tests in another experiment). The highly divergent sequence O25 was excluded from the polymorphism analyses. If included it would greatly inflate the value of Watterson's θ estimator based on the number of segregating sites and the assumption of equilibrium (Watterson, 1975). The value of π (Nei, 1987) is, however, only slightly affected.

The level of silent site (synonymous, intron and 5' flanking sites) variability in the Oulanka population was about 10% of that found for the Utah population (Table 1). As this difference could be due to a different

[illegible]

Fig. 1 *Drosophila montana* fu haplotypes in one American and one European population. Dots represent the same nucleotide as in the first sequence. Code is f for 5' flanking region sites, i for intron sites, s for synonymous sites and r for replacement sites. O is for Oulanka and U for Utah sequences.

Table 1 DNA sequence variation summary at *fu* in *Drosophila montana*

Sample		All (2362)	5'fl (47)	nsyn (1590)	syn (482)	int (243)	sil (772)
Oulanka	<i>S</i>	10	2	2	5	1	8
Utah		47	5	1	28	13	46
Oulanka	π	0.0011 $\pm$ 0.0007	0.0066 $\pm$ 0.0106	0.0003 $\pm$ 0.0003	0.0031 $\pm$ 0.0019	0.0019 $\pm$ 0.0013	0.0029 $\pm$ 0.0018
Utah		0.0075 $\pm$ 0.0042	0.0327 $\pm$ 0.0281	0.0002 $\pm$ 0.0003	0.0236 $\pm$ 0.0128	0.0190 $\pm$ 0.0125	0.0227 $\pm$ 0.0129
Oulanka	θ	0.0011 $\pm$ 0.0005	0.0114 $\pm$ 0.0085	0.0003 $\pm$ 0.0003	0.0028 $\pm$ 0.0015	0.0011 $\pm$ 0.0011	0.0028 $\pm$ 0.0013
Utah		0.0077 $\pm$ 0.0034	0.0410 $\pm$ 0.0242	0.0002 $\pm$ 0.0002	0.0224 $\pm$ 0.0103	0.0206 $\pm$ 0.0102	0.0230 $\pm$ 0.0103

Sample size is 24 for the Oulanka population and eight for the Utah population. π (Nei, 1987) is the average number of differences per base pair, and θ is Watterson's estimator based on the number of segregating sites (Watterson, 1975), at nonsynonymous sites (nsyn), at synonymous sites (syn), at intron sites (int), at 5' noncoding flanking sites (5'fl) or at silent sites (sil; 5' fl + int + syn). The standard deviations of π and θ due to stochastic factors, including sampling variance, were calculated according to Nei (1987; pp. 254–258) and Tajima (1993; pp 37–59). In order to obtain a conservative estimate, these standard deviations were calculated under the assumption of no recombination (Nei, 1987).

location of the *fu* locus on the X chromosome (see Discussion), we studied the salivary gland chromosomes of F₁ females from crosses between two isofemale lines from Oulanka and three isofemale lines from Utah. This study did not reveal any gross cytological rearrangements on the X chromosome, which suggests that *fu* is located in the same chromosome region in both populations.

The *fu* gene sequences of Oulanka population had only one replacement site that was not a singleton (i.e. a site that appeared only once in our sample). This was a replacement of a serine (TCG) by a leucine (TTG) at position 1402, which was present at a frequency of 20% (5 out of 25 sequences analysed have leucine at this position). The Utah population had no non-singleton replacement polymorphisms in the analysed *fu* region.

The level of variability was also estimated for another X-linked gene, *su(s)* (Table 2). Figure 2 shows the haplotype structure of the European and American *D. montana* populations for this locus. In the Oulanka population variability at silent sites (synonymous and intron sites) was about 1.1%, whereas in the Utah population it was about 0.5%. The analysed *su(s)* region included a short run of glutamines that is polymorphic in size (Table 3). The size of this section varied between eight and 10 repeats.

Linkage disequilibrium and recombination parameters

The Utah *fu* sequences showed evidence for a minimum of four recombination events in the history of the sample (Hudson & Kaplan, 1985). All four possible gametic types were found in 167 out of 1081 pairwise comparisons involving the 47 polymorphic sites (see Table 1). No significant linkage disequilibrium (by the χ^2 method with sequential Bonferroni correction; Rice, 1989) was detected in 300 pairwise comparisons involving the 25 informative polymorphic sites. Two different estimators of the level of recombination between adjacent sites (*C*) were 0.012 (Hey & Wakeley, 1997) and 0.119 (Hudson, 1987). For an X-linked locus $C = 3N_e c$, where *c* is the population average recombination frequency per nucleotide site and *N_e* is the effective population size. Linkage disequilibrium could not be studied in *su(s)* in the Utah population since this locus, in the region analysed, had only one segregating site that was not a singleton.

There was no evidence for recombination in the *fu* locus in the Oulanka population. In this population *D'*, a scaled measure of the degree of association between sites (Lewontin, 1964), was always 1. The level of linkage disequilibrium was, however, statistically significant in

Table 2 DNA sequence variation summary at *su(s)* in *Drosophila montana*

		All (413)	nsyn (223)	syn (56)	int (134)	sil (190)
Oulanka	<i>S</i>	8	2	0	6	6
Utah		2	0	0	2	2
Oulanka	π	0.0061 $\pm$ 0.0043	0.0018 $\pm$ 0.0027	0	0.0159 $\pm$ 0.0104	0.0112 $\pm$ 0.0074
Utah		0.0021 $\pm$ 0.0016	0	0	0.0064 $\pm$ 0.0051	0.0045 $\pm$ 0.0036
Oulanka	θ	0.0068 $\pm$ 0.0035	0.0032 $\pm$ 0.0024	0	0.0158 $\pm$ 0.0087	0.0112 $\pm$ 0.0061
Utah		0.0021 $\pm$ 0.0016	0	0	0.0065 $\pm$ 0.0051	0.0046 $\pm$ 0.0036

Sample size is 10 for Oulanka and six for Utah; definitions as in Table 1.

	1124
	12278990362
	63667066075
Code	iiiiiiiirr
O13	CGGTCCACGG
O2	
O14	A.....
O4	A.....
O16	AA.....
O3	AA...T....C
O7	AA...T.....
O5	.A.....
O8	.A.C.....A.
O9	..A.....
U11	AA....TG...
U2	AA....TG...
U6	AA....TG...
U8	AA....TG...
U3	AA....T.A..
U9	AA....TGA..

Fig. 2 *Drosophila montana* *su(s)* haplotypes in one American and one European population. Dots represent the same nucleotide as in the first sequence. Code is i for intron sites, s for synonymous sites and r for replacement sites. O is for Oulanka and U for Utah sequences.

Table 3 *Drosophila montana* *su(s)* homopolymer length variation

Population	<i>N</i>	<i>Q</i> ₈	<i>Q</i> ₉	<i>Q</i> ₈ LQ
Oulanka	10	7	3	0
Utah	6	2	3	1

N, number of individuals analysed.

only four out of 28 pairwise comparisons (χ^2 -test with sequential Bonferroni correction; Rice, 1989). *su(s)* also showed high linkage disequilibrium (D' always 1) in the Oulanka population. However, significant linkage disequilibrium (by the χ^2 -test with sequential Bonferroni correction; Rice, 1989) was not detected in any of the six pairwise comparisons. There was no clear association between *fu* and *su(s)* haplotypes.

Divergence between *D. montana* populations at *fu* and *su(s)*

The Utah and Oulanka populations had two fixed differences in *fu* (one synonymous at site 1904 and one intron substitution at site 1063; Fig. 1) and one fixed difference in *su(s)* (an intron substitution at position 96; Fig. 2). The average level of divergence between the two

Table 4 Average level of divergence between the *Drosophila montana* Oulanka and Utah populations

Locus	All	5'fl	nsyn	syn	int	sil
<i>fu</i>	0.0074	0.0225	0.0002	0.0219	0.0228	0.0223
<i>su(s)</i>	0.0100	—	0.0009	0	0.0300	0.0209

Definitions as in Table 1.

populations at silent sites was about 2.2% and 2.1% per site for *fu* and *su(s)*, respectively (Table 4). The value of F_{ST} (Hudson *et al.*, 1992a) was 0.39 (0.41 if sequence O25 is excluded) and 0.59 for *fu* and *su(s)*, respectively. The permutation test of Hudson *et al.* (1992b) revealed significant isolation ($P < 0.001$) between the two populations for both *fu* and *su(s)* data. It therefore seems that the two populations are differentiated. They may, however, not be completely isolated, which is best illustrated by looking at sequence O25. This sequence is likely to be the result of a gene conversion event between an American-like and a European-like *fu* sequence.

Neutrality tests

When test statistics that assume no recombination are used for detecting selection in regions of normal recombination, their power is usually low (Wall, 1999). Although the true rate of recombination to which the *fu* gene is exposed is unclear, there was evidence for recombination among the sequences from the Utah population (see above). Therefore, we have included recombination in the five statistical tests of departure from neutrality (Tajima, 1989; Fu & Li, 1993; Kelly, 1997; Wall, 1999), using the methods described in detail in Filatov & Charlesworth (1999). The *B* and *Q*-tests, based on the proportion of pairs of adjacent segregating sites that are congruent (i.e. have consistent genealogies; Wall, 1999), were significant ($P < 0.05$) for the eight *fu* sequences from the Utah population, suggesting a deviation from neutrality, although the significant deviation could be due to the violation of other assumptions of these tests. In these tests *C* was assumed to be higher than 0.010, which is a smaller value than the smallest estimate of *C* (0.012) for the Utah population. The other neutrality tests performed for the data were not significant. These tests include Kelly's test, which examines regions for the excess of linkage disequilibrium compared with that expected under neutrality (Kelly, 1997), the Fu and Li D^* statistical test (Fu & Li, 1993) and Tajima's *D*-test (Tajima, 1989). Therefore the conclusion of significant deviations from neutrality in the Utah population is only tentative.

In the Oulanka population, the *fu* sequence O25 alone defines 12 new polymorphic sites relative to the remaining 24 sequences. Such an event violates the assumption

of equilibrium, which is an important assumption of the statistical tests of deviations from neutrality. Consequently, sequence O25 has not been incorporated in any of the statistical analyses for detecting deviations from neutrality. Since there was no evidence for recombination at the *fu* or *su(s)* locus in the Oulanka population, we did not include recombination in the statistical tests. No significant deviations from neutrality were detected using the same five statistical tests as above (Tajima, 1989; Fu & Li, 1993; Kelly, 1997; Wall, 1999). Deviations from neutrality in the Oulanka population are not detected even if *C* is assumed to be as high as the smallest estimate of *C* (0.012) for the Utah population. The power of these tests is, however, limited with a small ($S < 11$; see Tables 1 and 2) number of segregating sites.

Discussion

Excluding the O25 sequence (the highly divergent European sequence that may have resulted from a gene conversion event; see Results), the estimated level of silent site nucleotide variability at *fu* was found to be about eight times smaller in European than in American *D. montana* population. In *su(s)*, the situation was the converse: here the estimated level of silent site nucleotide variability was about 2.5 times higher in European than in American population. Coalescent simulations of gene trees performed as described in Yi & Charlesworth (2000), showed that the level of silent DNA polymorphism at *fu* in the Utah population differs significantly from that in the Oulanka population (Table 1; $P < 0.05$; two tailed test; 10 000 runs). The same simulations for *su(s)* showed no significant difference between populations (Table 2; $P > 0.05$; two tailed test; 10 000 runs). The data on the level of silent site variability at *su(s)* therefore suggests a similar effective population size (on the order of $0.5\text{--}1 \times 10^6$ individuals) for the American and European *D. montana* populations, assuming a neutral mutation rate of 10^{-9} per site per generation (Aquadro *et al.*, 1994; Vieira & Charlesworth, 1999).

Both *fu* and *su(s)* support the hypothesis that American and European populations are differentiated, although the detection of one American-like *fu* sequence region in the latter population suggests, however, that the populations are not completely isolated (see Results). Therefore the apparent differentiation between the two populations at *fu* cannot be accounted for as the result of the putative selective sweep at or near this locus in the Oulanka population.

Genes located in regions of low recombination are usually less variable than genes located in regions of high levels of recombination (Aguadé *et al.*, 1989; Begun & Aquadro, 1992; Aquadro *et al.*, 1994). Therefore the *fu* data could reflect the presence of different

X-chromosome arrangements in the two populations. All individuals sequenced here were wild-caught males and therefore it was not possible to study their gene arrangements. X-chromosomal inversion polymorphism seems, however, an unlikely explanation for the different variability levels found at *fu* in the two populations, since we did not detect any gross chromosomal changes in the salivary gland chromosomes of F_1 females.

It could be that selection acting at or near *fu* in the Finnish Oulanka population has distorted the level of polymorphism at this locus (Maynard Smith & Haigh, 1974; Kaplan *et al.*, 1989; Stephan, 1995; Barton, 1998). Significant departures from neutrality are, however, difficult to show, since the power of statistical tests is limited with a small number of segregating sites. Since sequences for *fu* and *su(s)* are available for *D. virilis* (Vieira & Charlesworth, 1999), we could, in principle, address this issue by contrasting the levels of divergence and polymorphism in the European population for the two genes. The HKA test (Hudson *et al.*, 1987) could not, however, be performed as the intron sequences of *D. montana su(s)* gene were too dissimilar to permit unambiguous alignment with *D. virilis* sequences.

At *fu*, in the Utah population a significant deviation from neutrality was detected when two neutrality test statistics based on levels of linkage disequilibrium compared with that expected under neutrality (the B and Q-tests; Wall, 1999; see Results) were performed. However, three other neutrality test statistics, including Kelly's test (Kelly, 1997), which also examines regions for the excess of linkage disequilibrium compared with that expected under neutrality, did not detect any significant deviations from neutrality. The conclusion of positive selection on *fu* in the Utah population is therefore only very tentative. Additional data on *fu* and other linked and unlinked X-chromosomal genes for American and European populations is needed in order to fully address these issues. Although the linkage disequilibrium data obtained in the European population are compatible with neutrality, this aspect also deserves further research.

Acknowledgements

We thank Jim Fry for collecting and sending us the flies from Utah population and C.P. Vieira and B. Charlesworth for helpful comments on the work. J.V. wishes to thank B. Charlesworth for help in running the coalescent simulations. We thank John Brookfield and the anonymous reviewers for several suggestions that greatly improved this work. J.V. is supported by the Fundacao para a Ciencia e Tecnologia (PRAXIS XXI/BPD/14120/97) and A.H. by the Academy of Finland (36166).

References

- AGUADÉ, M., MIYASHITA, N. AND LANGLEY, C. H. 1989. Restriction-map variation at the *zeste-tko* region in natural populations of *Drosophila melanogaster*. *Mol. Biol. Evol.*, **6**, 123–130.
- AQUADRO, C. F., BEGUN, D. J. AND KINDAHL, E. C. 1994. Selection, recombination and DNA polymorphism in *Drosophila*. In: Golding, B. (ed.) *Non-Neutral Evolution: Theories and Molecular Data*, pp. 46–56. London: Chapman & Hall.
- ASPI, J. AND LANKINEN, P. 1992. Frequency of multiple insemination in a natural population of *Drosophila montana*. *Hereditas*, **117**, 169–177.
- BAKER, W. K. 1975. Linkage disequilibrium over space and time in natural populations of *Drosophila montana*. *Proc. Natl. Acad. Sci. U.S.A.*, **72**, 4095–4099.
- BAKER, W. K. 1980. Evolution of the alpha-esterase duplication within the montana subphylad of the virilis species group of *Drosophila*. *Genetics*, **94**, 733–748.
- BARTON, N. H. 1998. The effect of hitch-hiking on neutral genealogies. *Genet. Res.*, **72**, 123–133.
- BEGUN, D. J. AND AQUADRO, C. F. 1992. Levels of naturally occurring DNA polymorphism correlate with recombination rates in *D. melanogaster*. *Nature*, **356**, 519–520.
- FILATOV, D. A. AND CHARLESWORTH, D. 1999. DNA polymorphism, haplotype structure and balancing selection in the *Leavenworthia Pgi* locus. *Genetics*, **153**, 1423–1434.
- FU, Y. X. AND LI, W. H. 1993. Statistical tests of neutrality of mutations. *Genetics*, **133**, 693–709.
- HEY, J. AND WAKELEY, J. 1997. A coalescent estimator of the population recombination rate. *Genetics*, **145**, 833–846.
- HUDSON, R. R. 1987. Estimating the recombination parameter of a finite population model without selection. *Genet. Res.*, **50**, 245–250.
- HUDSON, R. R. AND KAPLAN, N. L. 1985. Statistical properties of the number of recombination events in the history of a sample of DNA sequences. *Genetics*, **111**, 147–164.
- HUDSON, R. R., KREITMAN, M. AND AGUADÉ, M. 1987. A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153–159.
- HUDSON, R. R., SLATKIN, M. AND MADDISON, W. P. 1992a. Estimation of levels of gene flow from DNA sequence data. *Genetics*, **132**, 583–589.
- HUDSON, R. R., BOOS, D. D. AND KAPLAN, N. L. 1992b. A statistical test for detecting geographic subdivision. *Mol. Biol. Evol.*, **9**, 138–151.
- KAPLAN, N. L., HUDSON, R. R. AND LANGLEY, C. H. 1989. The 'hitch-hiking' effect revisited. *Genetics*, **123**, 887–899.
- KELLY, J. K. 1997. A test of neutrality based on interlocus associations. *Genetics*, **146**, 1197–1206.
- KIMURA, M. 1983. *The Neutral Theory of Molecular Evolution*. Cambridge, Cambridge University Press.
- KREITMAN, M. AND AKASHI, H. 1995. Molecular evidence for natural selection. *Ann. Rev. Ecol. Syst.*, **26**, 403–422.
- LAKOVAARA, L. AND HACKMAN, W. 1973. *Drosophila ovivororum* sp. n., a new species of the *Drosophila virilis* group (Diptera, Drosophilidae). *Notulae Entomologicae*, **LIII**, 167–169.
- LEWONTIN, R. C. 1964. The interaction of selection and linkage. I. General considerations; heterotic models. *Genetics*, **49**, 49–67.
- MAYNARD SMITH, J. AND HAIGH, J. 1974. The hitch-hiking effect of a favourable gene. *Genet. Res.*, **23**, 23–35.
- MCVEAN, G. A. T. AND VIEIRA, J. 1999. The evolution of codon preferences in *Drosophila*: a maximum likelihood approach to parameter estimation and hypothesis testing. *J. Mol. Evol.*, **49**, 63–75.
- MOORHEAD, P. S. 1954. Chromosome variation in giant forms of *Drosophila montana*. *University of Texas Publ.*, **5422**, 106–129.
- NEI, M. 1987. *Molecular Evolutionary Genetics*. New York: Columbia University Press.
- OHTA, T. 1992. The nearly neutral theory of molecular evolution. *Ann. Rev. Ecol. Syst.*, **23**, 263–286.
- RICE, W. R. 1989. Analyzing tables of statistical tests. *Evolution*, **43**, 223–225.
- STEPHAN, W. 1995. An improved method for estimating the rate of fixation of favorable mutations based on DNA polymorphism data. *Mol. Biol. Evol.*, **12**, 959–962.
- STONE, W. S., GUEST, W. C. AND WILSON, F. D. 1960. The evolutionary implications of the cytological polymorphism and phylogeny of the *virilis* group of *Drosophila*. *Proc. Natl. Acad. Sci. U.S.A.*, **46**, 350–361.
- SUVANTO, L., HOIKKALA, A. AND LIIMATAINEN, J. O. 1994. Secondary courtship songs and inhibitory songs of *Drosophila virilis*-group males. *Behav. Genet.*, **24**, 85–94.
- TAJIMA, F. 1989. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.
- TAJIMA, F. 1993. Measurement of DNA polymorphism In: Takahata, N. and Clark, A. G. (eds) *Mechanisms of Molecular Evolution*, Sinauer Associates, Sunderland, MA.
- THROCKMORTON, L. H. 1982. The *virilis* species group. In: Ashburner, M., Carson, H. L. and Thompson, J. N. Jr (eds) *The Genetics and Biology of Drosophila*, vol. 3b. Academic Press, New York, pp. 227–296.
- VIEIRA, J. AND CHARLESWORTH, B. 1999. X chromosome DNA variation in *Drosophila virilis*. *Proc. R. Soc. B*, **266**, 1905–1912.
- VIEIRA, J. AND CHARLESWORTH, B. 2000. Evidence for selection at the *fused* locus of *Drosophila virilis*. *Genetics*, **155**, 1701–1709.
- WALL, J. D. 1999. Recombination and the power of statistical tests of neutrality. *Genet. Res.*, **74**, 65–79.
- WATABE, H.-A. AND HIGUCHI, C. 1979. On a new species of the *virilis* group of the genus *Drosophila* (Diptera, Drosophilidae), with revision of the geographical distribution of the group. *Ann. Zool. Jap.*, **52**, 203–211.
- WATTERSON, G. A. 1975. On the number of segregating sites in genetical models without recombination. *Theor. Pop. Biol.*, **7**, 256–275.
- YI, S. AND CHARLESWORTH, B. 2000. Contrasting patterns of molecular evolution of the genes on the new and old chromosomes of *Drosophila miranda*. *Mol. Biol. Evol.*, **17**, 703–717.