*Genet. Res., Camb* (1995), **65**, *pp.* 95–103  

# Comparative population structuring of molecular and allozyme variation of *Drosophila melanogaster Adh* between Europe, West Africa and East Africa

VÉRONIQUE BÉNASSI AND MICHEL VEUILLE*
*Institut d'Ecologie, Université Paris VI, 7 quai Saint Bernard, 75252 Paris Cedex 5, France*

(*Received 5 September 1994 and in revised form 5 January 1995*)

**Summary**

Restriction enzyme molecular variation in *Drosophila melanogaster Adh* was compared between three natural populations from Europe, West Africa and East Africa. The frequency distribution of silent variation in the *slow* allele was compatible with the neutral model in all three samples. The number of haplotypes in East Africa was significantly higher than in the other two populations. The largest divergence, as measured by $F_{st}$, was between the East African population and a group made up from the West African, the European, and previously studied American populations. We suggest that a split first occurred within African populations at least 44000 years ago. European populations separated from West Africa more recently, between the last glacial maximum and the post-glacial optimum, 18000 to 8000 years ago. We suggest that this species was domesticated recently relative to human evolution, possibly with the advent of agriculture. Population differentiation with respect to the two allozymes, *fast* and *slow*, does not follow the geographical pattern of silent variation. It opposes European to both African populations, and probably results from selection for adaptation to alcohol in recent temperate populations.

## 1. Introduction

*Drosophila melanogaster* is a crucial model for the study of evolution because of our understanding of its genetics and of its recent evolutionary history. *D. melanogaster* is a species that became commensal to man. Although today considered a 'domestic' species (Dobzhansky, 1965), it was not artificially selected, and thus represents a case of evolution over a relatively short time. Two features of this evolution are well documented. First, this species is thought to have originated from Africa, where its closest relatives are present (Tsacas & Lachaise, 1974; Lachaise *et al.* 1988). Secondly, its domestication involved adaptation to ethanol-rich environments. Related African species feed on decaying fruit. *D. melanogaster* populations are found on similar substrates in houses, orchards and wine cellars (Parsons, 1975; David, 1988). In cold and temperate regions, an allele of the ADH (alcohol dehydrogenase) enzyme, ADH-*fast*, predominates and contributes to alcohol detoxication. This allele is more active than the other allele, ADH-*slow*, which is more abundant in tropical areas, including Africa (Vigue & Johnson, 1973; Oakeshott *et al.*, 1982; Singh, Hickey & David, 1982). The higher activity of this allele is due to a single amino acid change in the coding region and to two insertion/deletions, $\nabla_1$ and $\nabla_2$, in the adult first intron (Laurie *et al.* 1991; Matthews *et al.* 1992). Polymorphic modifiers are also present 5′ to the transcription unit (Jiang & Gibson, 1992*b*). The organization of the *Adh* locus is now characterized for a large part of the sequence in and around the coding region (Benyajati *et al.* 1983), and it has become a thoroughly studied gene–enzyme system (review in Chambers, 1988). This is a particularly suitable locus for population genetic studies, since selection on nucleotides can be related to functional organization. The sequences of a large number of naturally originating variants (Kreitman, 1983; Collet, 1988; Eisses *et al.* 1990; Gibson *et al.* 1990) show no more amino acid polymorphisms than was formerly suggested by protein electrophoresis. This is important, since the best technique for recording molecular variation in large samples, four-cutter restriction mapping (Kreitman & Aguadé, 1986*a*), does not reveal all variants, and thus requires that variation is already well characterized.

Four-cutter variation in Adh is known from North America (Kreitman & Aguadé, 1986*a*; Simmons *et al.* 1989) and West Africa (Bénassi *et al.* 1993). We

* Corresponding author.

extend this survey to a European and an East African sample. Our purpose is to compare genetical structuring in the allozyme polymorphism with genetical structuring in *slow* silent variation. This is possible because most silent variation at this locus occurs in *slow* haplotypes (Kreitman & Aguadé, 1986*b*; Hudson, Kreitman & Aguadé, 1987; Kreitman & Hudson, 1991). *Fast* and *slow* haplotypes share few polymorphic sites (Simmons *et al.* 1989; Bénassi *et al.* 1993), probably because *fast* appeared recently and exchanged a small number of nucleotides through crossover or gene conversion. Variation in *slow* is thus expected to be silent and to reflect historical relationships between populations, while variation in amino acids reflects selection.

## 2. Materials and methods

### (i) *Strains*

The random sample of European flies was collected in south-western France, between Cognac and Jarnac, in October 1991. Flies were collected on '*rafes*', the hard leftovers from the wine press, which are stored between vineyards until subsequent use as fertilizer. First-generation flies were mated to a $Cy^{o}/ap^{Xa}$ balancer stock for isolation of chromosome 2 haplotypes. Since most lines were *fast* (57 out of 70), an additional sample of 31 $Adh^{S}$ chromosomes was obtained from lines extracted by Jean David in 1991. These lines were obtained from $Adh^{S}$ flies collected in the INRA Grande Ferrade Zoological Station, in Bordeaux. They are not a random sample with respect to allozymes, but they are a random sample for molecular variation within the *slow* class. They differed little from Cognac *slow* lines, as will be shown below (see the section on $F_{st}$). The two samples were pooled for the purpose of comparing variation in *slow* haplotypes between geographical regions.

The sample of 65 East African lines was collected by Daniel Lachaise from wild areas of Malawi in 1991. It consists of 50 lines collected in the Kasungu National Park, 12 lines from Lilongwe, two from Zomba and one from Viphya. Wild caught males were mated to the $Cy^{o}/ap^{Xa}$ balancer stock for isolation of chromosome 2 haplotypes. Most lines were *slow* (63 out of 65).

We also used data from the Lamto Ecological Station in Ivory Coast (West Africa), which we described previously in a within-population study (Bénassi *et al.* 1993). They are here compared with data from the other populations.

All African lines (including the Ivory Coast sample) were collected in preserved wild areas, far from any human settlement. Vouidibio *et al.* (1989) showed that *D. melanogaster* populations from an African city, Brazzaville, were genetically closer to European samples than to populations from the surrounding country. This suggests that urban populations are contaminated by recent invasions from temperate areas.

We compared our samples to previously described American populations (Kreitman & Aguadé, 1986*a*, Simmons *et al.* 1989). We did not use the large body of data published by Berry & Kreitman (1993) from the North American East Coast, since this sample, being collected along a north–south cline, was not a random sample of populations.

### (ii) *Restriction enzyme and allozyme polymorphism*

DNA extraction, PCR amplification of a 2·4 kb fragment encompassing the whole *Adh* gene, and restriction enzyme mapping were conducted as formerly described (Bénassi *et al.* 1993), except that restriction fragments were stained using ethidium bromide instead of being $^{35}$S-labelled. Restriction enzymes *Alu* I, *Ban* I, *Dde* I/*Bam*H I, *Hae* III, *Hha* I, *Msp* I, *Sau* 3AI, *Sau* 96I, and *Taq* I were used for comparison with previous studies. Comparisons with American populations did not involve the *Ban* I restriction enzyme, which was not used in one of these studies. *Fast/slow* allozyme polymorphism was recorded as previously described (Bénassi *et al.* 1993). No thermostable allele was detected.

## 3. Results

### (i) *Variation within populations*

A detailed account of the combinations of polymorphic sites found on each chromosome (haplotype) is given in Table 1. The frequency of the polymorphic sites in the three populations is shown in Table 2. We found 22 variable restriction sites. Eight were polymorphic in all three samples, four were polymorphic in two populations, and eleven were polymorphic in only one population: three in France, three in Ivory Coast, and five in Malawi. These eleven endemic polymorphisms were new, except for two from France (*Cfo* I 573 and *Hae* III 668) which had been previously found in North America (Simmons *et al.* 1989). All these endemic polymorphisms were found in *slow* haplotypes. One of them, *Dde* I 1875, from France, was found both in *slow* and *fast* haplotypes. The sites that are present in only one or two populations have a low frequency that never exceeded 13% for the rarest allele. All the other sites have been previously found in North America (Kreitman & Aguadé, 1986*a*; Simmons *et al.* 1989), and are therefore cosmopolitan polymorphisms. Of the eight polymorphisms that are present in all populations, five show a gradation in frequency from Malawi to Ivory Coast and then to France: *Sau3A* I 1354, *Dde* I 1518, *Dde* I 1527, *Alu* I 1596, and *Hae* III 1925. These sites are also the most heterozygous of all, and explain the population structuring that will be described below. Seven insertion/deletions (hereafter In-del) polymorphisms were found. Two of them, 'e' and 'g', were new and

https://doi.org/10.1017/S0016672300033115 Published online by Cambridge University Press

Table 1. *Sample of* Adh *haplotypes used in this study*

| IrrrIrrrIIIrrrrrrrrrIrrArrrrIrr† 11111 11 1112 22 a123b456cde78901234f56 7890g12 | L | M | C | B* | IrrrIrrrIIIrrrrrrrrrIrrArrrrIrr† 11111 11 1112 22 a123b456cde78901234f56 7890g12 | L | M | C | B* |
|---|---|---|---|---|---|---|---|---|---|
| .111.1.....111.111.111S.11.1.. | 1 | 1 | - | - | .11..1.....1111111..11S.111... | - | 1 | - | - |
| .11..1.....111.111.111S1.1.1.1 | 1 | - | - | - | ..1..1.....1111111.111S.11.... | - | 1 | - | - |
| .11..1.....111..11.111S1.1.1.1 | 3 | - | - | - | .11..1.....111.111.111S1.11... | - | 1 | - | - |
| .1...11....111..11.111S1.1.1.1 | 1 | - | - | - | .111.1....11111111.111S.11.... | - | 1 | - | - |
| .111.1.....111.111...1S.111..1 | 1 | - | - | - | .11..1.....11111.11.11S.111... | - | 1 | - | - |
| ..1.11.....111.111.111S.11.... | 4 | - | - | - | .111.1.....111.11111..S1.1.... | - | 1 | - | - |
| .11..1.....111.111.111S1.111.1 | 8 | - | - | - | .111.1.....1111111..11S1.111.. | - | 2 | - | - |
| .111.1.....111.111.1.1S.11...1 | 2 | - | - | - | .11..1.....111.111.111S1.11..1 | - | 1 | - | 1 |
| .11.11.....111.111.111S.111... | 1 | - | - | - | .11..1.....111.111.111S.11.... | - | 4 | - | 2 |
| .111.1.....1111111.111S1.111.1 | 1 | - | - | 2 | ..1..1.....1111111..11S.111... | - | 1 | - | - |
| .111.1.....111.111.111S1.1...1 | 1 | - | - | - | .11..1.....111.1111111S.111... | - | 1 | - | - |
| ..11.1.....111.111.111S1.111.1 | 4 | - | - | - | .11..1....11111111..11S.11.... | - | 1 | - | - |
| .111.1.....111.111...1S.11...1 | 10 | - | - | - | .11..1.....1111111..11S.11.... | - | 1 | - | - |
| ..1..1.....111.111.111S1.111.1 | 2 | - | - | 3 | .11..1.....1.1.111.111S.11.... | - | 1 | - | - |
| ..1.11.....111.111.111S.111... | 1 | - | - | - | .11..1....1111.111.111S.11.1.. | - | 1 | - | - |
| .111.1.....111.111.111S.11...1 | 6 | 2 | - | - | .11.11.....1111111..11S.1111.. | - | 1 | - | - |
| .111.1.....111.111.111S.11.1.1 | 1 | - | - | - | .11..1..1..111.111.111F1.11..1 | - | 1 | - | - |
| .111.......111.111...1S.111..1 | 1 | - | - | - | .11..1..1..111.111.1.1F1.11..1 | - | 1 | - | - |
| 111..1.....1111111.111S.11.... | 1 | - | - | - | .111.1....1111.111..11S1.11..1 | - | 1 | - | - |
| .111.1.....1111111.111S.11.1.1 | 1 | - | - | - | .111.1.....1111111..11S.11...1 | - | 1 | - | - |
| .111.1.....111.111...1S.11.1.1 | 1 | - | - | - | .11..1....1111.111..11S.111... | - | 1 | - | - |
| .111.......111.111...1S.1111.1 | 1 | - | - | - | .111.1....1111.111..11S.111... | - | 1 | - | - |
| .111.1.....111.111.111S1.111.. | 1 | - | - | - | .111.1....1111.111..11S.111..1 | - | 1 | - | - |
| .111.1.....111.111.111S.11.... | 4 | 1 | - | - | ..1..1.....1..111.111S.11.1.1 | - | - | - | 1 |
| .111.1.....111.111.111S1.111.1 | 12 | - | - | - | .11..1.....1..111.111S.1111.1 | - | - | - | 3 |
| .11..1..1..111.111.111F1.111.1 | 10 | - | - | - | .111.1.....111.111.111S1.1.1.1 | - | - | - | 3 |
| .11..1..1..1111111.111F.1111.1 | 1 | - | - | - | .111.......111.111....S.1111.1 | - | - | - | 2 |
| .11..1..11.1111111.111F1.111.1 | 3 | - | 2 | - | ..1..1.....111.111.111S.11...1 | - | - | - | 1 |
| .11..1..1..111.111.111F1.1.1.1 | 1 | - | 3 | - | .111.1.....111.111.111S.1111.1 | - | - | - | 2 |
| .111.1....1111.111.111S.11.... | - | 2 | - | - | ..11.1.....111.111.111S.1111.1 | - | - | - | 1 |
| .111.1.....1111111..11S.111... | - | 1 | - | - | .11..1.....111.111.111S.11.1.1 | - | - | - | 2 |
| .11..1.....111.111.111S.111..1 | - | 1 | - | - | .11..1.....111.111.111S.1111.1 | - | - | - | 1 |
| .111.1.....111.111..11S.11.... | - | 3 | - | - | .11..1.....1111111.111S1.11.11 | - | - | - | 1 |
| .111.1.....1111111.111S.11...1 | - | 1 | - | - | ..1..1.....111.111.111S1.11..1 | - | - | - | 2 |
| .111.1.....111.1111111S.11.... | - | 1 | - | - | .111.1.....1111111.111S1.11..1 | - | - | 3 | 3 |
| .111.1.....11111111.11S.111... | - | 2 | - | - | .111.1.....1111111.111S1.1...1 | - | - | 1 | 1 |
| .111.1.....111.111.111S1.11..1 | - | 5 | 1 | - | .111.1.....1111111.111F1.11... | - | - | 1 | - |
| .11..1....1111.111..11S.11...1 | - | 1 | - | - | .11..1.....1..111.111S.11.1.1 | - | - | 3 | - |
| ..1..1.....111.111.111S.1..... | - | 1 | - | - | .11..1.....1..111.111S.111..1 | - | - | 1 | - |
| ..1.11.....111.1111111S.11.... | - | 1 | - | - | .11..1..1..111.111.111F1.11..1 | - | - | 4 | - |
| ..11.1.....111.111..11S.11.... | - | 1 | - | - | .11..1.....1111111.111S1.11... | - | - | 1 | - |
| .11..1.....1111111.111S1.111.. | - | 1 | - | - | .11..1.....1..111.111S1.11..1 | - | - | 3 | - |
| .11..1.1...111111111111S.11.... | - | 1 | - | - | .11..1..11.1111111..11F1.1.1.1 | - | - | 7 | - |
| ..11.1....1111.111.111S.11.... | - | 1 | - | - | .11..1...1.111.111.11.F1.11..1 | - | - | 1 | - |
| ..11.1.....111.111.111S1.11..1 | - | 1 | - | - | .11..1..1..1111111.111F1.11..1 | - | - | 2 | - |
| .11..1.....111.111..11S.111..1 | - | 1 | - | - | .11..1..11.1111111.111F1.1.1.1 | - | - | 2 | - |
| .11..1....1111.111..11S.11.1.. | - | 2 | - | - | .11..1..1..1111111.111F1.11.11 | - | - | 1 | - |
| .111.1.....1111111.111S.11.... | - | 1 | - | - | .11..1..11.1111111.111F1.11..1 | - | - | 20 | - |
| .111.1.....111.111..11S.11...1 | - | 1 | - | - | .11..1..11.1111111.1.1F1.11..1 | - | - | 1 | - |
| .111.1....1111.1111.11S.11.... | - | 1 | - | - | .11..1..11.111.111.111F1.1...1 | - | - | 1 | - |
| .11..1.....111111.1.11S.111... | - | 1 | - | - | .11..1.....111.111.111F1.111.1 | - | - | 1 | - |
| .11..1.....11111111.11S.111... | - | 1 | - | - | .11..1...1.1111111.111F1.11..1 | - | - | 2 | - |
| .11..1.....111.111..11S.11.1.. | - | 1 | - | - | .11..1...1.1111111.11.F1.11..1 | - | - | 1 | - |
| .11..1....11111111..11S.11.1.. | - | 1 | - | - | .11..1.....111.111.111F1.11..1 | - | - | 8 | - |

* L, Lamto (Ivory Coast); M, Malawi; C, Cognac (France); B, slow haplotypes from Bordeaux (France).

† I, insertion/deletion, numbered a through g; r, restriction sites numbered 1 through 22 (1: restriction; dot: no restriction); A, allozyme (F: fast; S: slow). Insertion/deletions c and d correspond to in-dels ∇1 and ∇2 of Kreitman (1983). Polymorphic sites are shown according to their order along the sequence. The enzyme and position corresponding to each restriction site is given below. For consistency with earlier studies, we follow Kreitman & Aguadé's (1986*a*) numbering of positions: r1: *Msp* I -345; r2: *Dde* I 9; r3: *Ban* I 102; r4: *Hae* III 287; r5: *Dde* I 321; r6: *Alu* I 550; r7: *Cfo* I 573; r8: *Msp* I 583; r9: *Hae* III 688; r10: *Hae* III816; r11: *Msp* I 818; r12: *Cfo* I; r13: *Hae* III 1196; r14: *Sau*3AI 1354; r15: *Hae* III; r16: *Sau*96 I 1425; r17: *Dde* I 1518; r18: *Dde* I 1527; r19: *Cfo* I 1570; r20: *Alu* I 1596; r21: *Dde* I 1881; 22: *Hae* III 1925.

https://doi.org/10.1017/S0016672300033115 Published online by Cambridge University Press

Table 2. *Frequency of polymorphic sites (%) in the total sample, and in the slow and fast subsamples*

| | Total | | | Slow | | | Fast | | |
|---|---|---|---|---|---|---|---|---|---|
| Population | M | L | F* | M | L | F | M | L | F |
| n | 65 | 85 | 70 | 63 | 70 | 43 | 2 | 25 | 57 |
| In-del a† | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Msp* I-345 | 89 | 87 | 100 | 89 | 84 | 81 | 100 | 100 | 100 |
| *Dde* I 9 | 100 | 99 | 100 | 100 | 99 | 100 | 100 | 100 | 100 |
| Ban I 102 | 51 | 56 | 9 | 52 | 69 | 44 | 0 | 0 | 2 |
| In-del b | 3 | 7 | 0 | 3 | 9 | 0 | 0 | 0 | 0 |
| *Hae* III 287 | 100 | 97 | 100 | 100 | 97 | 95 | 100 | 100 | 100 |
| *Dde* I 321 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | 0 | 0 |
| *Alu* I 550 | 2 | 0 | 0 | 2 | 0 | 0 | 0 | 0 | 0 |
| In-del c (∇1) | 3 | 18 | 61 | 0 | 0 | 0 | 0 | 0 | 75 |
| In-del d (∇2) | 0 | 4 | 53 | 0 | 0 | 0 | 0 | 20 | 65 |
| In-del e | 23 | 0 | 0 | 24 | 0 | 0 | 0 | 0 | 0 |
| *Cfo* I 573 | 100 | 100 | 90 | 100 | 100 | 74 | 100 | 100 | 100 |
| *Msp* I 583 | 98 | 100 | 100 | 98 | 100 | 100 | 100 | 100 | 100 |
| *Hae* III 688 | 100 | 100 | 90 | 100 | 100 | 74 | 100 | 100 | 100 |
| *hae* III 816 | 32 | 8 | 63 | 33 | 4 | 28 | 0 | 24 | 68 |
| *Msp* I 818 | 100 | 95 | 100 | 100 | 95 | 100 | 100 | 100 | 100 |
| *Cfo* I 864 | 98 | 100 | 100 | 98 | 100 | 100 | 100 | 100 | 100 |
| *Hae* III 1196 | 98 | 100 | 100 | 98 | 100 | 100 | 100 | 100 | 100 |
| In-del f | 17 | 0 | 0 | 17 | 0 | 0 | 0 | 0 | 0 |
| *Sau*3A I 1354 | 54 | 84 | 90 | 52 | 80 | 95 | 100 | 100 | 88 |
| *Hae III* 1425 | 97 | 81 | 99 | 98 | 77 | 95 | 50 | 100 | 98 |
| *Sau*96 I 1425 | 98 | 100 | 97 | 98 | 100 | 95 | 100 | 100 | 97 |
| *fast/slow* | 3 | 18 | 81 | 0 | 0 | 0 | 100 | 100 | 100 |
| *Dde* I 1518 | 23 | 56 | 94 | 21 | 49 | 60 | 100 | 96 | 100 |
| *Dde* I 1527 | 77 | 44 | 6 | 79 | 51 | 40 | 0 | 4 | 0 |
| *Cfo* I 1570 | 98 | 100 | 100 | 98 | 100 | 100 | 100 | 100 | 100 |
| *Alu* I 1596 | 45 | 55 | 76 | 43 | 47 | 72 | 100 | 96 | 77 |
| In-del g | 15 | 62 | 26 | 16 | 54 | 53 | 0 | 0 | 28 |
| *Dde* I 1881 | 0 | 0 | 1 | 0 | 0 | 2 | 0 | 0 | 2 |
| *Hae* III 1925 | 27 | 85 | 97 | 25 | 81 | 98 | 100 | 100 | 98 |

* M, Malawi; L, Lamto (Ivory Coast); F, France (Cognac); for France, the slow subsample includes the slow random sample from Bordeaux.
† Polymorphic sites are shown according to their order along the sequence; they are identified as on Table 1; frequencies are shown for the restriction alleles, the longest in-del alleles, and the fast allozyme.

endemic to Malawi. In-dels 'c' ($\nabla_1$) and 'd' ($\nabla_2$), which partly cause the higher enzyme activity of the *fast* allozyme, were polymorphic only in *fast* haplotypes.

Allozyme polymorphism was consistent with previous studies: *fast* predominates in France ($82{\cdot}8 \pm 11{\cdot}9\%$), and is rare in Africa. Its frequency is $17{\cdot}6 \pm 10{\cdot}8\%$ in Ivory Coast. Only two *fast* out of 65 chromosomes were found in Malawi. In previous studies, nucleotide diversity differed between *fast* and *slow* haplotypes. Since allozyme frequencies are very different between samples, an estimate is given for each class in Table 3. *Slow* haplotypes were more diverse, and therefore more informative. Two estimates of nucleotide site variation, $\pi$ (Nei, 1987) and $\theta$ (Watterson, 1975) were calculated. Molecular variation for *slow* haplotypes was remarkably constant, both between the two estimates, and between populations. For instance, $\theta$ varied between 0·0049 in Ivory Coast and 0·0060 in France, with a very similar value, 0·0059, in Malawi. As discussed by Charlesworth, Morgan & Charlesworth (1993), a difference between $\theta$ and $\pi$ would indicate a bias in frequency, due either to selection, population bottlenecks or founding events. No significant difference between the two estimators was found in any sample. A test of neutrality (Tajima, 1989) never rejected the hypothesis of a random distribution of allele frequencies within samples.

The Tajima test is conservative, since it is based on stochastic variance in a model with no recombination. Recombination occurs in *Adh*. From available data, Hudson (1987) estimated recombinational events to be roughly 1·6 times more frequent than nucleotide substitutions in *D. melanogaster Adh*. Haplotype diversity is therefore a complementary statistic for describing population variation. We found 109 haplotypes (including in-dels) in the three populations. Nine were shared by two samples and one was shared by the three samples. Haplotype diversity was different

https://doi.org/10.1017/S0016672300033115 Published online by Cambridge University Press

Table 3. *Molecular variation within populations*

| | N | Number of* haplotypes | Haplotype† diversity (%) | $\theta$‡ ($\times 10^3$) | $\pi$§ ($\times 10^3$) |
|---|---|---|---|---|---|
| Lamto | | | | | |
| Total | 85 | 29 | 94 | 5·06 (1·77) | 5·60 (2·00) |
| *Adh*$^F$ | 25 | 4 | 54 | 1·94 (3.22) | 1·00 (0.60) |
| *Adh*$^S$ | 70 | 25 | 93 | 4·94 (2·08) | 5·80 (2·00) |
| Malawi | | | | | |
| Total | 65 | 51 | 99 | 6·23 (2·50) | 5·80 (2·00) |
| *Adh*$^F$ | 2 | 2 | 50 | | |
| *Adh*$^S$ | 63 | 49 | 99 | 5·93 (2·50) | 5·70 (1·60) |
| France | | | | | |
| Total | 70 | 24 | 89 | 4·94 (2·08) | 3·20 (1·10) |
| *Adh*$^F$ | 57 | 17 | 83 | 2·76 (3·55) | 2·20 (0·80) |
| *Adh*$^S$ | 43 | 21 | 96 | 6·05 (1·86) | 6·30 (2·20) |

* The number of haplotypes includes restriction sites and in-dels.
† Haplotype diversity was calculated as $(n-1/n)(1-\Sigma p_i^2)$.
‡ $\theta$ was calculated after Watterson (1975) and standard error was calculated after Hudson (1982).
§ $\pi$ and standard error were calculated according to Nei & Miller (1990) using programme RESTSITE (Miller, 1991).

Table 4. *Comparison of $F_{st}$ (%) for ADH allozymes and for restriction polymorphisms among slow haplotypes*

| | | | Restriction polymorphism in slow haplotypes † | | |
|---|---|---|---|---|---|
| | | Allozymes* | Total (2·4 kb) | 3′ half (0·9 kb) | Excluding Hae III 1925 |
| France | Malawi | 77 ($P < 10^{-4}$) | 27 (9) | 38 (13) | 20 (5) |
| Lamto | Malawi | 9 ($P < 5 \times 10^{-3}$) | 17 (6) | 20 (10) | 11 (3) |
| USA | Malawi | 17 ($P < 10^{-4}$) | 19 (8) | 28 (12) | 12 (2) |
| France | Lamto | 59 ($P < 10^{-4}$) | 9 (3) | 7 (3) | 9 (3) |
| USA | Lamto | 4 ($P < 2 \times 10^{-3}$) | 4 (2) | 3 (3) | 5 (3) |
| France | USA | 39 ($P < 10^{-4}$) | 3 (1) | 2 (1) | 3 (1) |

* $F_{st}$ probability was calculated by resampling individuals from the pooled population.
† Standard error was estimated through jacknifing over restriction sites.

in the three populations. Among *slow* haplotypes, we found 25 haplotypes out of 85 gametes in Lamto, 24 out of 70 in France (Bordeaux and Cognac), and 51 out of 65 in Malawi. There were twice as many haplotypes in Malawi as in the other two populations, though the sample size was slightly smaller. Since sample sizes were not too different, we compared these numbers using a Fisher exact test. There was no significant difference between France and the Ivory Coast, while Malawi had significantly more haplotypes than Ivory Coast ($P < 10^{-7}$) and France ($P < 10^{-6}$). A larger number of haplotypes implies a larger population effective size.

### (ii) *Genetical structuring between populations*

The amount of divergence between populations is classically measured by Wright's (1951) $F_{st}$. Formulae are provided by Weir & Cockerham (1984) for independent isolated loci. For linked nucleotide sites, a formula is provided by Hudson, Slatkin & Maddison (1992). Although these methods apply to different kinds of variation, they follow the same rationale, and the two series of estimates will follow the same trends in the same set of populations. The requisite is that the variation is neutral. We compared the $F_{st}$ for *fast/slow* allozyme frequencies (using Weir & Cockerham's formula) to the $F_{st}$ for *slow* silent nucleotide variation (using Hudson, Slatkin & Maddison's formula). Results are shown in Table 4. They include values from randomly sampled American populations (Kreitman & Aguadé, 1986*a*; Simmons *et al.* 1989). An important contrast appears between allozymic and silent variation. The allozyme polymorphism broadly differentiates between North and South (Table 4, first column). A very high $F_{st}$ was found between

https://doi.org/10.1017/S0016672300033115 Published online by Cambridge University Press

France and Africa (0·58 for Ivory Coast, 0·77 for Malawi). North American populations were intermediate. For *slow* silent variation (Table 4, second column), populations from France, North America and Ivory Coast were very close to each other. Their value (range 0·01–0·09) was similar to that between two American populations (average: 0·08, range 0·01–0·23), between the two French sub-samples (Cognac and Bordeaux, $F_{st} = 0·06$), and between the two Malawi sub-samples (Kasungu and Lilongwe, $F_{st} = 0·02$). All these values are smaller than any of the values including Malawi (range: 0·23–0·37). In other words, *slow* silent variation opposes two groups: East African and the other three populations. Ivory Coast is closer to Malawi than the other samples. We wanted to know if this result was biased by the existence of the nearby amino acid polymorphism. Frequency changes between *slow* and *fast* will change the frequency of silent alleles through hitch-hiking, not only between, but also within allozyme classes: recombination, though low, does occur, and will tend to increase the divergence between populations for polymorphic sites located far from the allozyme site. $F_{st}$ calculated on the half of the gene that surrounds the *fast/slow* site (covering polymorphic sites *Hae* III 1196 through *Hae* III 1925), yields very similar values to $F_{st}$ values obtained for the whole gene (Table 4, third column). The five most heterozygous sites were located between 1354 and 1925, around the F/S site (1490). This explains why removing more distant sites does not affect the geographical pattern. A site located in the intergenic region between *Adh* and *Adh-dup*, *Hae* III 1925, substantially increases population differentiation. When bootstrapping $F_{st}$ over sites, its value was not affected by the removal of any site (data not shown), except for this one (Table 4, fourth column). Ivory Coast then becomes intermediate between Malawi ($F_{st} = 0·11$) and France ($F_{st} = 0·09$), while North America becomes very close to Ivory Coast ($F_{st} = 0·05$).

## 4. Discussion

The population structure of *Adh* presents a strong contrast between silent variation and amino acid variation. Silent variation in *slow* opposes the East African sample to a group comprised of the West African, the European and the North American samples. Population structure in allozymes shows no relation to this and opposes Europe to the two African populations.

Two conclusions flow from these results. First. since *slow* silent variation is probably not affected by selection, the contrast between East Africa and the other populations suggests that a major split occurred during *D. melanogaster* evolution. Secondly, the contrast found in the same locus between allozyme and silent variations indicates that allozymes are subject to selection.

### (i) *History of* D. melanogaster *as suggested by silent variation in* Adh

Several studies have shown that *D. melanogaster* originated in Africa. This hypothesis was put forward by Tsacas & Lachaise (1978) from a biogeographical study of the *melanogaster* subgroup. Molecular studies showed that populations from North America and south-eastern Africa were very divergent from each other (Eanes *et al.* 1989; Begun & Aquadro, 1993, 1994*a*). Since no intervening population was involved in these studies, several explanations were possible. Begun & Aquadro (1993) noted that allozyme polymorphism was more similar for Benin and the USA than molecular polymorphism for Zimbabwe and the USA. Possible technical explanations of these differences have been discussed (Begun & Aquadro, 1994*b*; Lemeunier *et al.* 1994; Singh & Hale, 1994). Since Benin is close to Ivory Coast and thus belongs to West Africa, we consider that these observations confirm our hypothesis of an important differentiation of this species between West and East Africa.

Lemeunier & Aulard (1992) reviewed studies carried out on chromosomal inversions in Africa. They compared five populations from West Africa (Ivory Coast, Benin, Cameroon, Gabon and Congo) with five populations from East Africa (Botswana, Zimbabwe, Zambia, Tanzania and Kenya). Recurrent polymorphic inversions (that is, inversions that are found only in the African continent, and in at least two populations) are different between the two areas. Inversions *In(1)* 16D18D, *In(2L)* 22A; 26B, and *In(2R)* 48B; 56A are found in western populations, while inversions *In(1)* 12A; 18D and *In(3L)* 62D; 68A are found in eastern populations. Lemeunier & Aulard did not interpret this map in terms of history. Their observation is, however, consistent with our conclusion of a split between West and East Africa.

The study of other loci and of other populations will be necessary to confirm this conclusion. For instance, our study did not involve Asian populations. Geographic variation in *D. melanogaster* was first observed in a biometrical comparison of Japanese and French populations (Teissier, 1957). Restriction polymorphism studies, however, showed little change between Chinese and Australian populations (Jiang & Gibson, 1992*a*). Likewise, mitochondrial DNA shows little differentiation between Asian and American populations (Hale & Singh, 1991). Thus, Asian populations are probably closer to derived than to ancestral populations of *D. melanogaster*.

Variation in $Adh^S$ suggests that the first divergence occurred between East and West African populations. The Ivory Coast population is probably close to the ancestral population which led to European and North American populations. We formerly showed (Bénassi *et al.* 1993) that 12% of Ivory Coast chromosomes involve a set of endemic haplotypes,

https://doi.org/10.1017/S0016672300033115 Published online by Cambridge University Press

which we called 'Washington-*slow*-intermediates'. They connect two groups of *slow* haplotypes. One is comprised of most of the *slow* haplotypes. The other is a group called 'Washington-slow-like' ('*Wa-s*'). In North American samples, such intermediates have not been found (Simmons *et al.* 1989; Berry & Kreitman, 1993). Several changes, including two restriction sites, separate the two groups, although many chromosomes from each group are found in the USA. The intermediates are also lacking in France (this study) and in Spain (Aguadé, 1988). This pattern may reflect colonization pathways. Haplotypes that were present in a West African ancestor may have been lost in the colonization process. Since levels of variation are, however, very similar between West Africa and France, migration can only have involved a few founding effects.

The level of divergence between populations is given by $F_{st}$. Previous estimates of $F_{st}$ between East Africa and North America extend over a wide range, from 0·25 to 0·60, for X-linked loci (Begun & Aquadro, 1993) that are subject to selection. Some of them (*y*, *ac*, *su(f)*) may be driven by selective hitch-hiking (Begun & Aquadro, 1991), whilst two others (*G6pd* and *Pgd*) are probably directly subjected to selection (Eanes *et al.* 1993; Begun & Aquadro, 1994*a*). Our estimate of $F_{st}$ between East Africa and North America is 0·26. This provides an accurate estimate of the divergence, since it was calculated for silent variation in a locus located in a freely recombining region of an autosome (Ashburner, 1989).

The nucleotide divergence between populations can be calculated using Nei's (1987) $D_A$. Its value for $Adh^S$ silent variation is $0{\cdot}00222 \pm 0{\cdot}00113$ per nucleotide between Ivory Coast and Malawi, and $0{\cdot}00054 \pm 0{\cdot}00026$ between France and Ivory Coast. In comparison, the average number of differences in the *Adh* region delimited by our primers was 0·021 between *D. simulans* and *D. mauritania*. Hey & Kliman (1993) estimated these species to have diverged 770 kyr (thousand years) ago, with a range of 580–860 kyr. A ratio of these estimates gives an order of magnitude for the time elapsed since the divergence between populations. This time is 81 kyr for France and Malawi, 44 kyr for Ivory Coast and Malawi, and 19 kyr for France and Ivory Coast. These values are rough calculations. The estimated divergence between species is based on assumptions on substitution rates and highly depends on evolutionary models (for a comparison between several models, see Ashburner, Bodmer & Lemeunier, 1984). They also imply a complete isolation model. This is an unlikely hypothesis for the divergence between African populations since they coexist on the same continent. An isolation model is a more reasonable hypothesis for the divergence between Africa and Europe. A potentially efficient climatic barrier separated Europe from Africa 18 kyr ago during the last glacial maximum (see map fig. 1, in Adams *et al.* 1990). A continuous and wide tropical desert extended from the Atlantic Coast of Africa (Mauritania) to Far-Eastern Asia (Eastern China). On the other hand, this barrier was open for a short time during the post-glacial climatic optimum, 8·0 kyr ago (Branchu *et al.* 1993). The Sahara area then decreased and was partly replaced by grassland. The whole period extending from before the last glacial maximum to the postglacial optimum is compatible with our data, when accounting for the margin of error in our estimates.

An alternative hypothesis would be that European and African *D. melanogaster* populations separated at the same time as early *Homo sapiens* populations. According to Cavalli-Sforza *et al.* (1988), African and non-African human populations separated at least 92 kyr ago. Strong disagreement persists as to whether mitochondrial DNA variation supports this African origin hypothesis (Cann, Stoneking & Wilson, 1987; Vigilant *et al.* 1991; Tamura & Nei, 1993) or not (Ruvolo *et al.* 1993). It nonetheless provides an order of magnitude for the earliest human divergence. This data of 92 kyr is compatible with an 81 kyr estimate for the divergence between European and East African fruitflies, but it is not compatible with a 19 kyr estimate for the divergence between European and West African fruitflies.

Fruitflies are closely associated with mankind and probably originated in Africa, but their current geographic pattern suggests that *Drosophila melanogaster* came to Europe long after humans. European fruitflies are truly 'domestic' in the sense that they are found only in or near houses (Parsons, 1974; David, 1988). We suggest that their domestication occurred at some time between 18000 and 8000 years ago. This period corresponds to landmarks in human cultural history that may have been important in *Drosophila* domestic ecology: the beginnings of agriculture in pre-pottery neolithic villages of the Middle East around 10–11000 years ago and the domestication of fruit-trees in the Chalcolithic 6000 years ago (Zohary & Hopf, 1988).

### (ii) *Comparing allozyme and silent variation*

The second outcome of this study is that the geographical structuring with respect to *fast/slow* alleles is not related to that for silent variation. Many ecological, molecular, physiological, and population studies have suggested that selection acts on this allozyme polymorphism. Population studies have found north–south clines in ADH allozyme frequency on continents that *D. melanogaster* probably colonized only recently, such as North America and Australia. We have compared amino acid and silent variation in the same sample and showed that they have different geographical patterns. Since history is the best way of explaining structuring in silent variation, changes in *fast/slow* frequencies cannot be accounted for by

history alone. Berry & Kreitman (1993) carried out a similar four-cutter study between populations sampled along the ADH cline of the North American East Coast. They showed that a north–south cline existed for *fast/slow* and for the $\nabla_1$ in-del, but not for $Adh^S$ molecular variation. Berry & Kreitman's statistical treatment provides good evidence that selection on the $\nabla_1$ in-del polymorphism was an important factor in *fast/slow* evolution. The conclusion of both studies is that selection acts on allozyme frequencies. In both cases, *fast* predominates in the north, and *slow* in the south. A possibility is that a north–south differentiation was established independently in the Old World and in the New World. In agreement with this, substantial changes in ADH allozyme frequencies are observed in semi-natural experimental devices (van Delden & Kamping, 1989). Another possibility is that the differentiation was first established between Europe and West Africa at a critical time of *D. melanogaster* evolution. Populations from different origins would then have colonized the New World, resulting in the secondary formation of a cline through migration. Both interpretations are based on an initial change in frequency that was faster for *fast/slow* than for any silent site. This agrees with the prevalent idea that the domestication of *D. melanogaster* involved adaptation to alcohol through processing of environmental ethanol by ADH.

We thank Daniel Lachaise for collecting population samples from the Ivory Coast and from Malawi (a project supported by CNRS and Ecotrop), Jean David for providing isogenic $Adh^S$ lines from Bordeaux, Matthew Cobb and Yannis Michalakis for comments on the manuscript, and Daniel Zohary for a helpful and cheerful discussion. Monty Slatkin drew our attention to possible biases due to recombination. Comparing polymorphism and species divergence was suggested to us by Brian Charlesworth.

## References

Adams, J. M., Faure, H., Faure-Denard, L., McGlade, J. M. & Woodward, F. I. (1990). Increase in terrestrial carbon storage from the last glacial maximum to the present. *Nature* **348**, 711–714.

Aguadé, M. (1988). Restriction map variation at the *Adh* locus of *D. melanogaster* in inverted and non-inverted chromosomes. *Genetics* **119**, 135–140.

Ashburner, M. (1989). *Drosophila, a Laboratory Handbook*. Cold Spring Harbor: Cold Spring Harbor Laboratory Press.

Ashburner, M., Bodmer, M. & Lemeunier, F. (1984). On the evolutionary relationships of *Drosophila melanogaster*. *Developmental Genetics* **4**, 295–312.

Begun, D. J. & Aquadro, C. F. (1991). Molecular population genetics of the distal portion of the X chromosome in *Drosophila*: evidence for genetic hitchhiking for the *yellow-achaete* region. *Genetics* **129**, 1147–1158.

Begun, D. J. & Aquadro, C. F. (1993). African and North American populations of *Drosophila melanogaster* are very different at the DNA level. *Nature* **365**, 548–550.

Begun, D. J. & Aquadro, C. F. (1994*a*). Evolutionary inferences from DNA variation at the *6-phosphogluconate dehydrogenase* locus in natural populations of *Drosophila*: selection and geographic differentiation. *Genetics* **136**, 155–171.

Begun, D. J. & Aquadro, C. F. (1994*b*). Regional variation in fruitflies. *Nature* **369**, 449.

Bénassi, V., Aulard, S., Mazeau, S. & Veuille, M. (1993). Molecular variation of *Adh* and *P6* genes in African populations of *Drosophila melanogaster* and its relation to chromosomal inversions. *Genetics* **134**, 789–799.

Benyajati, C., Spoerel, N., Haymerle, H. & Ashburner, M. (1983). The messenger RNA for alcohol dehydrogenase in *Drosophila melanogaster* differs in its 5′ end in different developmental stages. *Cell* **33**, 125–133.

Berry, A. & Kreitman, M. (1993). Molecular analysis of an allozyme cline: alcohol dehydrogenase in *Drosophila melanogaster* on the East coast of North America. *Genetics* **134**, 869–893.

Branchu, P., Haure, H., Ambrosi, J.-P., van Zinderen Bakker, E. M. & Faure-Denard, L. (1993). Africa as source and sink for atmospheric carbon dioxide. *Global and Planetary Change* **7**, 41–49.

Can, R. L., Stoneking, M. & Wilson, A. C. (1987). Mitochondrial DNA and human evolution. *Nature* **325**, 31–36.

Cavalli-Sforza, L. L., Pizza, A., Menozzi, P. & Mountain, J. (1988). Reconstruction of human evolution: bringing together genetic, archeological, and linguistic data. *Proceedings of the National Academy of Sciences, USA* **85**, 6002–6006.

Chambers, G. K. (1988). The *Drosophila alcohol dehydrogenase* gene–enzyme system. *Advances in Genetics* **25**, 39–107.

Charlesworth, B., Morgan, M. T. & Charlesworth, D. (1993). The effect of deleterious mutations on neutral molecular variation. *Genetics* **134**, 1289–1303.

Collet, C. (1988). Recent origin of a thermostable alcohol dehydrogenase allele of *Drosophila melanogaster*. *Journal of Molecular Evolution* **27**, 142–146.

David, J. (1988). Ethanol adaptation and alcohol dehydrogenase polymorphism in *Drosophila*: from phenotypic functions to genetic structures. In *Population Genetics and Evolution* (ed. G. de Jong), pp. 163–172. Berlin, Heidelberg: Springer-Verlag.

Dobzhansky, Th. (1965). 'Wild' and 'domestic' species of *Drosophila*. *In The Genetics of Colonizing Species* (ed. H. G. Baker and G. Ledyard Stebbins), pp. 533–546. New York and London: Academic Press

Eanes, W. F., Ajioka, J. W., Hey, J. & Wesley, C. (1989). Restriction map variation associated with the *G6PD* polymorphism in natural populations of *Drosophila melanogaster*. *Molecular Biology and Evolution* **6**, 384–397.

Eanes, W. F., Kirchner, M. & Yoon, J. (1993). Evidence for adaptive evolution of the *G6pd* gene in *Drosophila melanogaster* and *Drosophila simulans* lineages. *Proceedings of the National Academy of Sciences, USA* **90**, 7475–7479.

Eisses, K. T., Andriesse, A. J., Douwe de Beer, A., Thörig, G. E. W. & Weisbeek, P. J. (1990). Analysis of the gene encoding the multifunctional alcohol dehydrogenase alozyme ADH-*71K* of *Drosophila melanogaster*. *Molecular Biology and Evolution* **7**, 459–469.

Gibson, J. G., Reed, D., Wilks, A. V. & Jiang, C. (1990). Molecular similarity of *Drosophila melanogaster* thermostable alleles from populations on different continents. *Heredity* **65**, 1–6.

Hale, L. R. & Singh, R. S. (1991). A comprehensive study of genic variations in natural populations of *Drosophila melanogaster*. IV. Mitochondrial DNA variation and the role of history *vs.* selection in the genetic structure of geographic populations. *Genetics* **129**, 103–117.

Hey, J. & Kliman, R. M. (1993). Population genetics and phylogenetics of DNA sequence variation at multiple loci

within the *Drosophila melanogaster* species complex. *Molecular Biology and Evolution* **10**, 804–822.
Hudson, R. R. (1982). Estimating genetic variability with restriction endonucleases. *Genetics* **100**, 711–719.
Hudson, R. R. (1987). Estimating the recombination parameter of a finite population model without selection. *Genetical Research* **50**, 245–250.
Hudson, R. R., Kreitman, M. & Aguadé, M. (1987). A test of neutral molecular evolution based on nucleotide data. *Genetics* **116**, 153–159.
Hudson, R. R., Slatkin, M. & Maddison, W. P. (1992). Estimation of levels of gene flow from DNA sequence data. *Genetics* **132**, 583–589.
Jiang, J. C. & Gibson, J. B. (1992*a*). The alcohol dehydrogenase polymorphism in natural populations of *Drosophila melanogaster*: restriction map variation in the region of the *Adh* locus in populations from two hemispheres. *Heredity* **68**, 1–14.
Jiang, J. C. & Gibson, J. B. (1992*b*). The alcohol dehydrogenase polymorphism in natural populations of *Drosophila melanogaster*: ADH activity variation, restriction site polymorphism and the *Adh* cline. *Heredity* **68**, 337–344.
Kreitman, M. (1983). Nucleotide polymorphism at the *alcohol dehydrogenase* gene region of *Drosophila melanogaster*. *Nature* **304**, 412–417.
Kreitman, M. & Aguadé, M. (1986*a*). Genetic uniformity of two populations of *Drosophila melanogaster* as revealed by filter hybridization of four-nucleotide restriction enzyme digests. Proceedings of the National Academy of Sciences, USA **83**, 3562–3566.
Kreitman, M. & Aguadé, M. (1986*b*). Excess polymorphism at the *Adh* locus in *Drosophila melanogaster*. *Genetics* **114**, 93–110.
Kreitman, M. & Hudson, R. (1991). Inferring the evolutionary histories of the *Adh* and *Adh-dup* loci in *Drosophila melanogaster* from patterns of polymorphism and divergence. *Genetics* **127**, 565–582.
Lachaise, D., Cariou, M.-L., David, J. R., Lemeunier, F. & Tsacas, L. (1988). Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evolutionary Biology* **22**, 159–225.
Laurie, C. C., Bridgham, J. T. & Choudary, M. (1991). Association between DNA sequence variation and variation in expression of the *Adh* gene in natural populations of *Drosophila melanogaster*. *Genetics* **129**, 489–499.
Lemeunier, F. & Aulard, S. (1992). Inversion polymorphism in *Drosophila melanogaster*. In *Drosophila Inversion Polymorphism*, (ed. C. B. Krimbas and J. R. Powell), Cleveland: CRC Press.
Lemeunier, F., Aulard, S., Bénassi, V. & Veuille, M. (1994). Fruitfly origins. *Nature* **371**, 25.
Matthews, P., Agrostis, A., Taylor, A. C. & McKechnie, W. (1992). An association between ADH protein levels and polymorphic nucleotide variation in the *Adh* gene of *Drosophila melanogaster*. *Molecular Biology and Evolution* **9**, 526–536.
Miller, J. C. (1991). Restsite: a phylogenetic program that sorts raw restriction data. *Journal of Heredity* **82**, 262–263.
Nei, M. (1987). *Molecular Evolutionary Genetics*. New York: Columbia University Press.
Nei, M. & Miller, J. C. (1990). A simple method for estimating the average number of nucleotide substitutions within and between populations from restriction data. *Genetics* **125**, 873–879.
Oakeshott, J. G., Gibson, J. B., Anderson, P. R., Knibb, W. R., Anderson, D. G. & Chambers, G. K. (1982). Alcohol dehydrogenase and Glycerol-3-phosphate dehydrogenase clines in *Drosophila melanogaster* on different continents. *Evolution* **36**, 86–96.
Parsons, P. A. (1975). The comparative evolutionary biology of the sibling species, *Drosophila melanogaster* and *D. simulans*. *Quarterly Review of Biology* **50**, 151–169.
Ruvolo, M., Zehr, S., von Dornum, M., Pan, D., Chang, B. & Lin, J. (1993). Mitochondrial *COII* sequences and modern human origins. *Molecular Biology and Evolution* **10**, 1115–1135.
Simmons, G. M., Kreitman, M., Quattelbaum, W. F. & Miyashita, N. (1989). Molecular analysis of the alleles of *Alcohol dehydrogenase* along a cline in *D. melanogaster*. *Evolution* **43**, 393–409.
Singh, R. S. & Hale, L. R. (1994). Regional variation in fruitflies. *Nature* **369**, 449.
Singh, R. S., Hickey, D. A. & David, J. (1982). Genetic differentiation between geographically distant populations of *Drosophila melanogaster*. *Genetics* **101**, 235–256.
Tajima, F. (1989). Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics* **123**, 585–595.
Tamura, K. & Nei, M. (1993). Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biology and Evolution* **10**, 512–526.
Teissier, G. (1957). Discriminative biometrical characters in French and Japanese *Drosophila melanogaster*. *Cytologia suppl. Proceedings of the International Genetics Symposia*, 1956: 502–505.
Tsacas, L. & Lachaise, D. (1974). Quatre nouvelles espèces de la Côte d'Ivoire du genre *Drosophila*, groupe *melanogaster* et discussion de l'origine du sous-groupe *melanogaster* (Diptera: Drosophiliidae). *annales de l'Université d'Abidjan* **E7, 193–211.**
van Delden, W. & Kamping, A. (1989). The association between the polymorphisms at the *Adh* and *alpha-Gpdh* loci and the *In(2L)t* inversion in *Drosophila melanogaster* in relation to temperature. *Evolution* **43**, 775–793.
Vigilant, L., Stoneking, M., Harpending, H., Hawkes, K. & Wilson, A. C. (1991). African populations and the evolution of human mitochondrial DNA. *Science* **253**, 1503–1507.
Vigue, C. L. & Johnson, F. M. (1973). isozyme variability in species of the genus *Drosophila*. VI. Frequency–property–environment relationships of allelic alcohol dehydrogenases in *D. melanogaster*. *Biochemical Genetics* **9**, 213–227.
Vouidibio, J., Capy, P., Defaye, D., Pla, E., Sandrin, J., Csink, A. & David, J. (1989). Short-range genetic structure of *Drosophila melanogaster* populations in an Afrotropical urban area and its significance. *Proceedings of the National Academy of Sciences, USA* **86**, 8442–8446.
Watterson, G. A. (1975). On the number of segregating sites in genetical models without recombination. *Theoretical Population Biology* **7**, 256–276.
Weir, B. S. & Cockerham, C. C. (1984). Estimating *F*-statistics for the analysis of population structure. *Evolution* **38**, 1358–1370.
Wright, S. (1951). The genetical structure of populations. *Annals of Eugenics* **15**, 323–354.
Zohary, D. & Hopf, M. (1988). *Domestication of Plants in the Old World*. Oxford: Clarendon Press.