Genet. Res., Camb. (1987), **50**, *pp.* 181–185 *With* 1 *text-figure Printed in Great Britain*

Biochemical phylogeny of the eight species in the *Drosophila melanogaster* subgroup, including *D. sechellia* and *D. orena*

MARIE LOUISE CARIOU

Laboratoire de Biologie et Génétique Evolutives, C.N.R.S. 91190 – Gif-sur-Yvette, France

(Received 21 January 1987 and in revised form 27 May 1987)

Summary

The phylogenetic relationships of the eight species of the *Drosophila melanogaster* subgroup are examined on the basis of genetic variation at 33 putative enzyme loci. Values of Nei's genetic distance (*ds*) range from 0·28 to 1·74. *D. sechellia* appears closer to *D. simulans* than to *D. mauritiana*, the two former being the most closely related. *D. orena* is quite distantly related to *D. erecta* (*ds* = 1). Genetic differentiation supports the existence of three main lineages within the *melanogaster* subgroup and the *yakuba-teissieri* pair appears to be closer to the *melanogaster* lineage than to the *erecta-orena* one. Inferences of the times of species divergence from allozyme data are made and their agreement to other estimates is discussed.

1. Introduction

Since the pioneer phylogeny based on chromosomes (Lemeunier & Ashburner, 1976), several studies have focused on relationships within the *Drosophila melanogaster* subgroup. Dendrograms and networks have been established from different approaches, including the most recent molecular techniques, but the phylogenetic relationships of the species are not yet completely resolved because the different trees show some discrepancies. Thus a comparison between the greatest number of phylogenies is needed. A few enzyme polymorphism studies have been previously made but they failed to include the eight members of the *melanogaster* subgroup, more especially *D. orena* and *D. sechellia*.

This paper provides estimates of genetic divergence and phylogenetic relationships of *D. orena* and *D. sechellia* versus the other species. Time of divergences from allozymic data are also compared to the other estimates so far available.

2. Material and Methods

Strains. Samples of natural populations were investigated for five species. Isofemale lines were started from wild-caught females and three individuals were electrophoresed per locus for each line. *Drosophila melanogaster* (63 lines), *D. yakuba* (65 lines) and *D. teissieri* (75 lines) were collected in the Tai Forest (Ivory Coast). These populations are considered to be the best samples of the three species because West Africa represents their ancestral home range (Lachaise *et al.* 1987). *D. simulans* (20 lines) was captured on Mt. Kenya (Kenya) and *D. sechellia* (28 lines) in the Seychelles Islands.

Strains from the Laboratoire de Biologie et Génétique Evolutives collection (C.N.R.S., France) were used for the last three species: *D. mauritiana* (ref. 163·1, Mauritius Island), *D. erecta* (the two available strains from Ivory Coast, ref. 154·1, Lamto, and 220·5, Grand Bassam) and the single extant strain of *D. orena* founded from only one female (Mt. Lefo, Cameroon). For each locus and strain, at least 30 individuals were sampled. Electrophoretic techniques were the usual ones. The 32 loci analysed in starch gels are listed in Table 1. α-amylase was analysed using a 5% acrylamide gel as described in Daïnou *et al.* (1987).

Electrophoresis was performed on adult flies except for loci *Est11* and *2*, *Lap1* (third larval instar) and *Lap2* (pupa). The most common allele of *D. melanogaster* was numbered 100. The different alleles were then assigned a greater or lesser number according to their electrophoretic mobility relative to that of the *melanogaster* reference. The amylase alleles were nominated in agreement with previous usage and according to the recent nomenclature given in Daïnou *et al.* (1987).

Nei's standard genetic distance (*ds*) was used here with the UPGMA method for estimating branch lengths because of its superior performance in computer simulations (Nei *et al.* 1983). The *do* distance suggested by Gregorius (1984) was also used because it is metric and is bounded by 0 and 1.

Table 1. *Electrophoretic mobilities (relative to* D. melanogaster*) of the most common allozymes at 33 loci in the eight species of the* melanogaster *subgroup*

	Species[a]								
Loci	mel	sim	mau	se	yak	teis	ore	ere 1	ere 2
Pgm	100	100	100	100	88[c]	93[c]	91	91	91
αGpdh[b]	100	100	100	100	100	100	98	98	98
Fum	100	100	100	100	102[c]	97[c]	100	100	100
G6pdh	100	100	100	100	99[c]	101[c]	106[d]	104[d]	104
Ald	100	100	100	100	100	100	112[d]	109[d]	109
Aldox	100	102	103	98	96	96	102[d]	105[d]	105
Odh	100	100	100	100	97	97	97	97	97
Adh[b]	100	94	102	94	88[c]	101[c]	104[d]	112[d]	112
Estc	100	107	112	107	97	100	101[d]	103[d]	103
Est6	100	100	100	102	100	102	100[d]	103[d]	103
Estp	100	100	100	100	105[c]	95[c]	105[d]	110[d]	110
Sod	100	100	100	100	98[c]	110[c]	100	100	100
Sdh	100	104	104	99	100	100	99[d]	106[d]	106
6Pgd	100	97	97	97	93[c]	95[c]	98[d]	96[d]	96
Acph	100	98	98	98	97	100	93	95	95
Xdh	100	101	100	100	102	102	102	102	101
Ca	100	100	100	100	103	106	103[d]	110[d]	110
Su	100	100	100	97	100	100	100	100	103
Hk1	100	100	100	100	97	97	100	100	100
Hk2	100	100	100	102	100	100	100[d]	102[d]	102
Hk3	100	100	100	100	100	100	100	100	100
Idh	100	104	104	104	104	104	112[d]	109[d]	109
Me	100	100	100	100	103	103	100[d]	103[d]	103
Mdh	100	100	100	100	103	103	100[d]	101[d]	101
Got	100	96	96	96	96	96	96	96	96
Pgi	100	100	103	100	103[c]	105[c]	95[d]	98[d]	98
Gdh	100	103	100	103	105	105	111[d]	107[d]	107
Gapdh	100	101	101	101	100	100	102	102	102
Lap1	100	100	100	100	100	100	100	100	100
Lap2	100	96	97	97	100	100	98[d]	105[d]	105
Es11	100	102	104	104	100	100	null[d]	105[d]	105
Es12	100	100	100	100	100	100	98[d]	103[d]	103
Amy	6	4·4	4·4	4·4	4·4[c]	3·4[c]	9[d]	8[d]	8

[a] mel, *D. melanogaster*; sim, *D. simulans*; mau, *D. mauritiana*; se, *D. sechellia*; yak, *D. yakuba*; teis, *D. teissieri*; ore, *D. orena*; ere 1 and 2, *D. erecta*.
[b] *αGpdh*[100] is the usual 'fast' allele; *Adh*[100] is the 'slow' allele.
[c] Loci diagnostic between *D. yakuba* and *D. teissieri*.
[d] Loci diagnostic between *orena* and *erecta*.

Table 2. *Estimates of Nei's* ds *genetic distance (above the diagonal) and* do *genetic distance (below the diagonal) (33 enzymatic loci)*

	mel	sim	mau	se	yak	teis	ore	ere 1	ere 2
melanogaster	—	0·545	0·503	0·623	0·935	1·008	1·142	1·533	1·737
simulans	0·357	—	0·296	0·281	0·998	1·239	1·013	1·507	1·490
mauritiana	0·345	0·215	—	0·317	0·882	1·244	1·069	1·506	1·664
sechellia	0·394	0·211	0·229	—	1·273	1·365	1·274	1·518	1·487
yakuba	0·492	0·513	0·438	0·577	—	0·392	1·118	1·256	1·546
teissieri	0·509	0·567	0·568	0·589	0·288	—	1·468	1·370	1·704
orena	0·553	0·510	0·525	0·571	0·540	0·605	—	0·938	1·130
erecta 1	0·625	0·619	0·619	0·619	0·570	0·588	0·489	—	0·073
erecta 2	0·655	0·616	0·643	0·612	0·624	0·643	0·540	0·073	—

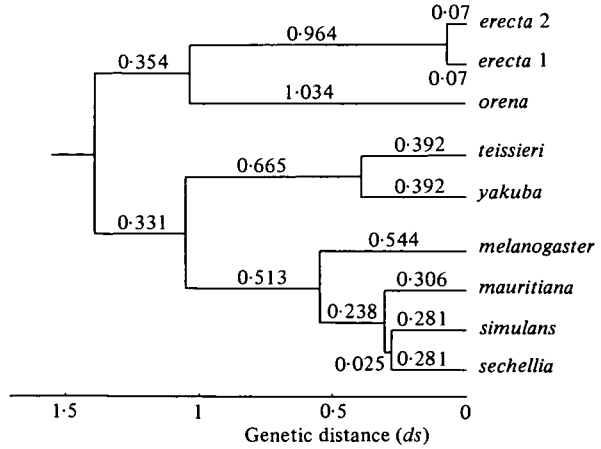

Fig. 1. Phylogenetic relationships between the eight species of the *Drosophila melanogaster* subgroup. The dendrogram was constructed by the UPGMA method from the Nei's (*ds*) genetic distance matrix.

3. Results

163 alleles were identified at 33 presumed genetic loci in the nine populations tested and the most common allozymes are presented in Table 1. Only two loci (*Hk3* and *Lap1*) were strictly monomorphic and identical in the eight species. Seven loci are diagnostic between *D. sechellia* and either *D. simulans* (*Aldox, Est6, Sdh, Xdh, Su, Lap2, Es11*) or *D. mauritiana* (*Aldox, Adh, Est6, Sdh, Su, Pgi, gdh*) and twenty between *D. orena* and *D. erecta* (see Table 1). It should be stressed that 9 out of the 33 loci are diagnostic between *D. yakuba* and *D. teissieri* (see Table 1) and the two species share very few alleles at polymorphic loci.

Nei's genetic distance (*ds*) and the (*do*) distance are presented in Table 2. Genetic differentiation between species is generally, high and both estimates, *ds* and *do*, are consistent. The lowest genetic distance (*ds*: 0·28, *do* = 0·21) is found between *D. simulans* and *D. sechellia*; the maximum (*ds* = 1·73, *do* = 0·65) involves *D. erecta* and *D. melanogaster*. Two phylogenetic trees were constructed from *ds* and *do* by using the UPGMA method. The trees produced are identical in topology whichever genetic distance is used yet they differ in branch lengths, especially for the most distant divisions. Only the *ds* tree is presented in Fig. 1.

4. Discussion

The fairly high genetic distance (*ds* = 1) between *D. erecta* and *D. orena* suggests a relatively old split. An ancient divergence of these two species, also supported by satellite (Strachan *et al.* 1982) and mtDNA data (Solignac *et al.* 1986) is more likely than a possible recent separation suggested by polytene and mitotic chromosome studies (Lemeunier & Ashburner, 1984).

D. sechellia, the last discovered species, appears genetically very close to the widespread *D. simulans* and its endemic relative, *D. mauritiana*. The distances are similar to the *D. simulans–D. mauritiana* distance found by Gonzalez *et al.* (1982). This is consistent with morphological hybridization (Lachaise *et al.* 1986; Coyne & Kreitman, 1986) and DNA data (Ashburner *et al.* 1984; Bodmer & Ashburner, 1984; Solignac *et al.* 1986).

D. yakuba and *D. teissieri* are identical for mtDNA (Solignac *et al.* 1986) while they are well differentiated by several other characters. Nei's distance between these species is 0·4, that is less than that previously found by Eisses *et al.* (1979) and Ohnishi *et al.* (1983). This value, half that between *D. erecta* and *D. orena*, indicates substantial genetic differentiation and a more recent split than that between the *erecta-orena* pair as also suggested by satellite DNA (Strachan *et al.* 1982) and ribosomal and histone gene families (Coen *et al.* 1982). The pattern of allozyme dif-

Table 3. *Divergence time estimates* (*Myr*) *of* Drosophila *species of the* melanogaster *subgroup. The two estimates based on allozymes are using Nei's formula* (*1975*) *with* $\alpha : 1 \times 10^{-7}$ *and Carson* (*1976*) *calibration* (*within parentheses*)

	Species compared			
	sim/mau[a]	mel/sim mel/sim-mau-se	mel/yak mel/yak-tei	mel/ere mel/ere-ore
Allozymes distance				
Present study	1·5 (0·59)	2·7 (1·09)	5·3 (2·1)	6·9 (2·8)
Immunological distances				
Beverley & Wilson (1982, 1984)	—	0·5	17	—
Nucleotide sequences of ADH				
Ashburner *et al.* (1984)	3·0–3·5	3·8–4·0	—	15–37
Bodmer & Ashburner (1984)	2·9	3·9	—	13
Cohn *et al.* (1984)	2·7	4·7	—	—
Eastel & Oakeshott (1985)	2·0–7·7	3·0–9·4	—	9·9–30·2
Stephens & Nei (1985)	0·86–1·45	2·0–3·5	—	—
Paleo-biogeographic arguments				
Lachaise *et al.* (1987)	—	2·5–3·5	—	6–15

[a] Abbreviations as note *a* in Table 1.

https://doi.org/10.1017/S0016672300023673 Published online by Cambridge University Press

ferentiation between the two species supports either an ancient speciation, as also suggested by paleobiogeographic arguments (Lachaise *et al.* 1987), or a quantum speciation (Solignac *et al.* 1986), but excludes an introgression process (Solignac *et al.* loc. cit.) because heterozygotes between specific alleles are missing.

Three main clusters are defined and indicate the existence of three main evolutionary lineages within the subgroup. The topology given in Fig. 1 is consistent with those more recently inferred from 2D electrophoresis (Ohnishi *et al.* 1983), amylase polymorphism (Daïnou *et al.* 1987) and mitochondrial DNA (Solignac *et al.* 1986) in clustering the *yakuba-teissieri* species to the *melanogaster* complex, rather than to the *erecta-orena* pair. *D. sechellia* is found somewhat closer to *D. simulans* than *D. mauritiana* is, but the differences between the various distances are trivially small and the tree still remains ambiguous for the respective positions of these branching points. Therefore, the chronology of the speciation events remains unresolved for these three species.

Several calibrations of the molecular (electrophoretic) clock have been provided (see Thorpe, 1982, for a review). Considering their diversity and the evidence that protein loci evolve at different rates in different groups, the two following values have been considered: Nei originally suggested a calibration of $1\ D = 5$ Myr using a mean mutation rate of 1×10^{-7} per locus per year, all mutations being neutral. Carson (1976) postulated an increase in D of 0·01 every 20000 years ($1\ D = 2$ Myr) for Hawaian species of *Drosophila*.

Some estimates of divergence time obtained by several investigators using various techniques are shown in Table 3. Large discrepancies exist between the different values, depending on different assumptions. Such inferences of divergence time are highly speculative, but it should be stressed that allozyme estimates (using Nei values), the lowest values established by *Adh* DNA sequencing (Stephens & Nei, 1984) and Paleo-biogeographic arguments (Lachaise *et al.* 1987) are congruent and thus may correspond to the true divergence times.

This manuscript has benefited from discussions and comments from J. R. David, D. Lachaise and F. Lemeunier. I thank Miss N. Groseille for excellent technical assistance, J. P. Gauthier for computer assistance and Mrs E. Simonneau for typing the manuscript. This work was partly supported by a grant from the Centre National de la Recherche Scientifique, ATP 'Biologie des Populations'.

References

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

Beverley, S. M. & Wilson, A. C. (1982). Molecular evolution in *Drosophila* and higher Diptera. I. Micro-complement fixation studies of a larval hemolymph protein. *Journal of Molecular Evolution* **18**, 251.

Beverley, S. M. & Wilson, A. C. (1984). Molecular evolution in *Drosophila* and higher Diptera. II. A time scale for fly evolution. *Journal of Molecular Evolution* **21**, 1.

Bodmer, M. & Ashburner, M. (1984). Conservation and change in the DNA sequences coding for alcohol dehydrogenase in sibling species of *Drosophila*. *Nature* **309**, 425.

Carson, H. L. (1976). Inference of the time of origin of some *Drosophila* species. *Nature* **259**, 395.

Coen, E., Strachan, T. & Dover, G. (1982). Dynamics of concerted evolution of ribosomal DNA and histone gene families in the *melanogaster* species subgroup of *Drosophila*. *Journal of Molecular Biology* **158**, 17.

Cohn, V. H., Thompson, M. A. & Moore, G. P. (1984). Nucleotide sequence comparison of the *Adh* gene in three Drosophilids. *Journal of Molecular Evolution* **20**, 31.

Coyne, J. A. & Kreitman, M. (1986). Evolutionary genetics of two sibling species of *Drosophila*. *Evolution* **40**, 673.

Daïnou, O., Cariou, M. L., David, J. R. & Hickey, D. (1987). Amylase gene duplication: an ancestral trait in the *Drosophila melanogaster* species subgroup. *Heredity* (In the Press).

Easteal, S. & Oakeshott, J. G. (1985). Estimating divergence times of *Drosophila* species from DNA sequence comparisons. *Molecular Biology and Evolution* **2**, 87.

Eisses, K. T., Van Dijk, H. & Van Delden, W. (1979). Genetic differentiation within *melanogaster* species group of the genus *Drosophila* (*Sophophora*). *Evolution* **33**, 1063.

Gonzalez, A. M., Cabrera, V. M., Larruga, M. M. & Gullon, A. (1982). Genetic distance in the sibling species *Drosophila melanogaster*, *Drosophila simulans* and *Drosophila mauritiana*. *Evolution* **36**, 517.

Gregorius, H. R. (1984). A unique genetic distance. *Biometrical Journal* **26**, 13.

Lachaise, D., Cariou, M. L., David, J. R., Lemeunier, F., Tsacas, L. & Ashburner, M. (1987). Historical biogeography of the *Drosophila melanogaster* species subgroup. *Evolutionary Biology* **22**, (In the Press).

Lachaise, D., David, J. R., Lemeunier, F., Tsacas, L. & Ashburner, M. (1986). The reproductive relationships of *Drosophila sechellia* with *D. mauritiana*, *D. simulans*, and *D. melanogaster* from the Afrotropical region. *Evolution* **40**, 262.

Lemeunier, F. & Ashburner, M. (1976). Relationships within the *melanogaster* species subgroup of the genus *Drosophila* (*Sophophora*). II. Phylogenetic relationships between six species based upon chromosome banding sequences. *Proceedings of the Royal Society of London* **193**, 275.

Lemeunier, F. & Ashburner, M. (1984). Relationships within the *melanogaster* species subgroup of the genus *Drosophila* (*Sophophora*). IV. The chromosomes of two new species. *Chromosoma* **89**, 343.

Nei, M., Tajima, F. & Tateno, Y. (1983). Accuracy of estimated phylogenetic trees from molecular data. *Journal of Molecular Evolution* **19**, 153.

Ohnishi, S., Kawanishi, M. & Watanabe, T. K. (1983). Biochemical phylogenies of *Drosophila*: protein differences detected by two-dimensional electrophoresis. *Genetica* **6**, 55.

Solignac, M., Monnerot, M. & Mounolou, J.-C. (1986). Mitochondrial DNA evolution in the *melanogaster* species subgroup of *Drosophila*. *Journal of Molecular Evolution* **23**, 31.

Stephens, J. C. & Nei, M. (1985). Phylogenetic analysis of polymorphic DNA sequences at the Adh locus in

Drosophila melanogaster and its sibling species. *Journal of Molecular Evolution* **22**, 289.

Strachan, T., Coen, E., Webb, D. A. & Dover, G. (1981). Modes and rates of change of complex DNA families of *Drosophila*. *Journal of Molecular Biology* **157**, 37.

Thorpe, J. P. (1982). The molecular clock hypothesis: biochemical evolution, genetic differentiation and systematics. *Annual Revue of Ecology and Systematics* **13**, 139.