Genomic Blot Hybridization as a Tool of Phylogenetic Analysis: Evolutionary Divergence in the Genus *Drosophila**

Michael Loukas,¹ Christos Delidakis,² and Fotis C. Kafatos^{2,3}

¹ Department of Genetics, Agricultural College of Athens, Iera odos 75, Votanicos, Athens 11855, Greece

² Department of Cellular and Developmental Biology, The Biological Laboratories, Harvard University, 16 Divinity Avenue, Cambridge, MA 02138, USA

³ Institute of Molecular Biology and Biotechnology, Department of Biology, University of Crete, Heraklio Crete 71110, Greece

Summary. Comparative, quantitative Southern analysis of genomic DNA, using single-copy sequence probes, potentially is valuable for phylogenetic analysis. We have examined 27 *Drosophila* species, belonging to two subgenera, seven species groups, and ten subgroups, using a variety of cloned and characterized probes: twelve cloned sequences from *D. melanogaster,* two from *D. pseudoobscura,* and two from *D. grimshawi.* The data are generally congruent with accepted phylogenetic relationships in *Drosophila,* and confirm or clarify some previously uncertain relationships. The potential and limitations of the method are discussed.

Key words: DNA divergence $-$ Quantitative Southern analysis – Molecular phylogeny – *Drosophila* evolution

Introduction

Biological macromolecules provide a rich source of information on evolution, and are increasingly used to address important issues of phylogeny, population genetics and evolutionary theory. Various indices of molecular divergence have been used to infer phylogenetic relationships among different species, usually based on the hypothesis that **mo-** lecular divergence is approximately linear with time (Wilson et al. 1977; but see Britten 1986). These indices include antigenic properties (Collier and MacIntyre 1977; Beverley and Wilson 1984) or electrophoretic mobilities of proteins (Lakovaara et al. 1976; Gonzalez et al. 1982; Loukas et al. 1984), melting properties of genomic DNA hybrids retained on hydroxyapatite (Angerer et al. 1976; Zwiebel et al. 1982; Sibley and Ahlquist 1984), amino acid sequences in proteins (Thompson 1980), and nucleotide sequences of cloned genes (Bodmer and Ashburner 1984).

Melting properties of total DNA hybrids average out the divergence of many classes of sequences, which are known to evolve at different rates; sequence analyses are very labor-intensive and necessarily focus on a very limited number of genes. Between these two extremes, one can envisage two approaches for phylogenetic analysis based on DNA divergence of multiple genetic loci: comparative analysis of detailed restriction maps, and comparative, quantitative Southern analysis of genomic DNA, using single-copy sequence probes. The former approach has proved very useful for the analysis of intraspecies variation. In this report we explore the potential of the latter approach.

We have used 27 species of the genus *Drosophila,* including representatives of two major subgenera and seven species groups. For the 10 species of the *melanogaster* group, considerable phylogenetic information already exists (Lemeunier and Ashburner 1976; Eisses et al. 1979; Gonzalez et al. 1982; Oh-

Offprint requests to: Fotis C. Kafatos²

^{*} Presented at the FEBS Symposium on Genome Organization and Evolution, held in Crete, Greece, September 1-5, 1986

nishi et al. 1983; Bodmer and Ashburner 1984), Thus, analysis of that species group permitted us to Validate the method, as well as to verify some as-Pects of the phylogeny for which the supporting evidence was limited. The method was then applied to a Study of phylogeny in the *obscura* species group (es Lakovaara et al, 1976; Loukas et al. 1984; Steinemann et al. 1984; reviewed in Lakovaara and Saura 1982). Finally, we explored broad aspects of *Drosophila* phylogeny, by considering representatives of seven species groups.

Materials and Methods

DroSophila Species. Twenty seven species of the genus *Drosophila* ositions and abbreviations used hereafter are as follows:

The origin of the stocks was as follows:

- *1. Drosophila* Resource Center, Bowling Green: ma, si, au, tr, $ts, pu, se, ta, im, fo, hy, vi, ro, la$
- 2. Dept. of Genetics, Agricultural College of Athens: su, oh, am
- 3. A. Prevosti, Dept. of Genetics, Barcelona: md. gu
- 4. H. Burla, Zoologisches Institut u Museum der Univ., Zürich: tt, ss, bf
- 5. Museum of Comparative Zoology, Harvard: pe, gr
- 6. Biological Laboratories, Harvard: me, er, ps

All species were maintained at $23-25^{\circ}\text{C}$ on standard D. mel*anogaster* medium except for the *obscuta* group species which were maintained at 20°C and *D. grimshawi* which was mainrained at 18"C.

In addition, we used a stock of the mediterranean fruit fly , *Ceratitis capitata,* obtained *from* the NRC Democritus and maintained in Athens, Greece on a standard medfly diet.

DNA Preparation. Total nucleic acid was isolated from frozen male adult flies by the following procedure: After homogenization in 0.2 M sucrose, 50 mM EDTA, 10 mM Tris (pH 9.5) and 0.5% SDS and incubation at 68° C for 10 min, the homogenate was brought to a final concentration of 1,5 M potassium acetate and was incubated on ice for 30 min. Insolubles were removed by centrifugation at 15.000 rpm for 10 min in an Eppendorf microcentrifuge. DNA was precipitated with ethanol, pelleted and resuspended in TE solution (10 mM Tris pH 7.4 and 1 mM EDTA). The mixture was extracted three times with 24:24:1 phenol/chloroform/isoamyl alcohol, brought to a final concentration of 1.5 M ammonium acetate, precipitated with ethanol, pelleted and resuspended in TE solution. After one additional precipitation with ammonium acetate and ethanol, the pellet was washed in 70% ethanol and finally resuspended in TE solution. The solution was stored at -20° C. The concentration of DNA in that solution was calculated by utilizing the diphenylamine reaction as modified by Abraham et al. (1972). Samples were assayed in duplicate for each concentration tested. The optical densities were converted to DNA concentrations by using a standard curve made from known quantities of dNTPs.

Probe Preparation. Twelve cloned probes derived from D. *melanogaster,* two from *D. pseudoobscura,* and two from *D. grimshawl* were used in this study. Table 1 lists the nature and origin of each probe, diagrams the location of gene(s) therein, indicates the restriction enzyme chosen to digest genomic DNA, and the expected size of the genomic fragment corresponding to the probe in the species of origin.

Each probe was excised from the corresponding ptasmid and recovered by electrophoresis in 0.9% low-melting agarose. The mixture of DNA-agarose was extracted at 65"C in a 1:1 suspension of phenol and 0.1 M Tris (pH 8.0), 0.5 M NaCl, 1 mM EDTA. After centrifugation the mixture was extracted twice again with phenol and once with 24:24:1 phenol/chloroform/isoamyl alcohol. After concentration with sec-butanol, the DNA was precipitated with sodium acetate and ethanol and resuspended in TE. For nick-translation, 0.2μ g of DNA and all four α -³⁹P-dNTPs were used in a standard reaction (Maniatis et al. 1982).

Southern Blotting. Genomic DNA samples (1 μ g, unless otherwise noted) were digested with the appropriate enzyme (Table 1), electrophoresed in 0.8% agarose and transferred to a nylon membrane (GeneScreen, NENuclear). Depurination was in 0.25 M HCl, denaturation in 0.5 M NaOH, 1 M NaCl $(2 \times 15 \text{ min})$ and neutralization in 0.5 M Tris (pH 7.4), 1.5 M *NaC1 (2 x 30* min). Prehybridization (15 min) and hybridization (36-48 h) was as recommended by Church and Gilbert (1984), at 65"C except for the indicated experiments which used 60°C. Filters were washed at 60°C as recommended by Church and Gilbert (1984), with 40

* All probes are from *D. melanogaster* except those ending in ps (*D. pseudoobscura*) and $gr(D.$ grimshawi)

t The probe also contains an unidentified gene downstream of chorion s16 (shown as double arrow in the diagram)

§ All probes are shown to scale, except those with asterisks (2 × reduced scale). Genes are diagrammed 5'-3'. Coding regions are indicated by wide and untranslated regions by narrow filled
boxes, introns by white boxes, a ** Enzyme used for digesting genomic DNA (abbreviated as in the diagram), and expected size of hybridizing fragment (in kb) are indicated

176

Table 1. The probes used

mM sodium phosphate (pH 7.2), 1 mM EDTA, in the presence 0f5% SDS (10 min, 30 min) and then in the presence of 1% SDS $(4 \times 15 \text{ min})$. Autoradiograms were exposed at -70° C with intensifier screens.

Quantitation of Band Intensities. Although the probe was in excess, the bands were not saturated under the conditions used, and their intensities could be used as a relative index of sequence divergence. No attempt was made to establish an absolute scale of intensities vs. percent sequence divergence. It should be noted, however, that probes hsp-82 and hsp-82ps have both been sequenced as have the corresponding genes in *D. pseudoobseura* and *D. virilis*, permitting evaluation of the relationship between hybridization intensity and degree of sequence divergence.

Because of the widely different band intensities, multiple ex-POSUres of variable duration were obtained from each autoradiogram, so that all intensities could be obtained from non-sat- .Urating exposures. Correction was made for the non-linearity introduced by the intensifying screens, by using a standard curve that was established by scanning bands corresponding to known amounts of DNA. All autoradiograms were scanned with a Helena Laboratories Quickscan RD densitometer under identical settings and the intensities were calculated by manual integration ^{of the} peaks. When multiple bands were present in the same sample, the intensity of each band was determined separately and then the intensities were summed. Intensities were always ex-Pressed relative to the self-hybridization standard, obtained by including in each blot an identical amount ofgenomic DNA from he Species of origin of the probe. Reproducibility was excellent. Using eight different probes, 27 samples were analyzed in du-Plicate, three in triplicate, and one in quadruplicate, with an OVerall mean deviation of 0.022 in relative intensity; in 10% of the cases the deviation was 0.00, in 13% it was 0.01, in 42% 0.02, m 19% 0.03, and in 16% 0.04. In general we consider differences of 0.03 or less as insignificant.

Results and Discussion

The Actin Genes in the Genus Drosophila

We have previously used DNA derived from the 5C actin locus of *D. melanogaster,* as probe for *in Situ* hybridizations with polytene chromosomes of Seven diverse *Drosophila* species (Loukas and Kafatos 1986a,b). Thus we determined that each of these species has six actin loci, which are widely dispersed, in a reasonably conservative pattern that reveals the homologies of chromosomal elements. Figure 1 shows the Southern patterns observed by hybridizing genomic DNA from 27 *Drosophila* Species and from the medfly, *Ceratitis capitata* (2 μ g), with an 1800 bp DNA probe derived from the transcribed region of the same 5C actin gene of D . *melanogaster* (Table 1). Multiple strong or very strong bands, typically 6 ± 1 , are characteristic of all *Drosophila* species. Six strong bands are observed in 14 species, five bands in 7, and seven bands in 4 Species. These results are consistent with six strongly COnserved actin genes throughout the genus, with due allowance for the possibility of chance coincidence of bands, of band splitting due to the presence ^{of} a restriction site within the region homologous

to the probe, or of DNA polymorphism within the stock. Only two species have more numerous strong bands (one each eight and ten bands), which might also be explained by internal restriction sites or by DNA polymorphism. Polymorphism is a likely explanation for the minor bands in approximately half of the species, including *D. melanogaster.*

For present purposes, two additional important observations can be derived from Fig. 1. First, the aggregate intensity of the bands is comparable (within a factor of two) in all *Drosophila* species, both distant and closely related. In fact, with the exception of the four very closely related species of the *melanogaster* subgroup *(melanogaster, mauritiana, sirnulans* and *erecta),* all 23 other *Drosophila* species show aggregate actin band intensities ranging between 0.45 and 0.70, relative to the *D. melanogaster* standard (Tables 2, 3, and 4). Presuming six highly conserved genes in all cases, that observation and the fact that we used equal amounts of DNA in all samples imply that the genome *size* is relatively constant throughout the genus. That will be important later, when individual band intensities are used as a measure of sequence divergence in different species. Surprisingly, only three weak bands are observed with genomic DNA of the medfly, suggesting either that the actin genes of this species are extensively diverged, or that its genome size is considerably larger than in *Drosophila.* Second, only a few actin bands of conserved size are observed, and only in very closely related *Drosophila* species *(melanogaster* and *mauritiana, mauritiana* and *simulans, auraria* and *triauraria).* Apparently, sequence evolution within the genus is so extensive, that coincidence of any bands encompassing more than coding region DNA might be expected, if at all, oniy in the case of sibling species.

Sequence Divergence and Phylogeny in the melanogaster *Group*

When eleven unique probes, derived from *D. melanogaster,* were hybridized with genomic DNA from 10 species of the *melanogaster* species group, rather coherent results were obtained which were consistent with the presumed phylogeny. The Southern patterns are presented in Figs. 2 and 3, and relative intensities of the bands are tabulated in Table 2.

As an example, let us consider the simple pattern obtained with act-1600, an actin probe 1.6 kb in length, largely encompassing the 3' untranslated and 3' flanking region of the 5C actin gene of *D. melanogaster.* In this case a single band is observed in each species, except for *D. simulans* which shows an additional minor band (presumably as a result of polymorphism, or of a new EcoRI site near the end of the probe-homologous region). In this pattern

M ps me ma si er au tr se ts pu ta

me sumd gu tt am bf ps pe ob ss

M

Fig. 1. Genomic blots using an actin probe from *D. melanogaster* (act-1800; see Table 1). Species from which the genomic DNA was derived are abbreviated as in Materials and Methods. *PandA* contains mostly species of the *melanogastergroup, Panel* B mostly species of the *obscura* group, and *Panel C* species of seven different groups, including the *melanogaster* group *(Mel),* the *obscura* group *(Obs),* the *immigrans* group *(Imm),* the *robusta* group (Rob), and the *hydei*, virilis and Hawaiian picture-winged groups (hy, vi, gr, respectively). All panels include a *D. melanogaster* standard (me). Panel C also includes medfly DNA (Cer*atitis capitata, ce).* Markers (M) were fragments of λ DNA digested with Hind III (23.0, 9.8, 6.6, 4.5, 2.5, and 2.2 kb from top to bottom)

(B)

the band intensities are high for members of the *melanogaster* subgroup *(D. melanogaster, D. mauritiana, D. simulans,* and *D. erecta),* moderate for the representative of the *takahashii* subgroup (D. *takahashii),* and uniformly low for the five representatives of the *montium* subgroup *(D. auraria, D. triauraria, D. serrata, D. tsacasi,* and *D. punjabiensis).* With the exceptions noted below, similar results, with respect to intensities, were obtained for all other probes tested.

Figure 4a schematizes the average distance of all species from *D. melanogaster,* measured in terms of the relative "deficit" of crosshybridization (one minus the average normalized intensity of crosshybridizing bands; \bar{x} in Table 2). Figure 4b presents a phylogenetic tree, which is consistent with our data as well as with previously available evidence.

The species of the *melanogaster* subgroup are thought to have diverged relatively recently, and thus their high band intensities were expected. Morphologically, these species are very similar, their male genitalia providing the only dependable distinguishing feature (ecologically, however, they are quite dissimilar: *D. melanogaster* and *D. sirnulanS* are cosmopolitan while the Afro-tropical species can be very restricted in distribution as in the case of *D. erecta* which feeds almost exclusively on the fallen fruits of the palm *Pandanus candelabrum [La*chaise and Tsacas 1974]). A gratifying agreement was observed in ranking according to DNA hybridization distance and according to presumed phyletic distance: the closest species to *D. melanogaster* is apparently *D. mauritiana,* closely followed by D. *simulans,* and then more distantly by *D. erecta. Le*meunier and Ashburner (1976), using polytene chromosome banding patterns, have classified the first three species in the *melanogaster* complex, and the last in the *yakuba* complex. The results are also

Table 2. Relative hybridizations in the *melanogaster* group*

* All values expressed relative to the species of origin of the probe. Deviant values are boxed (see text). Subgroups are separated by Spaees. The *obscura* group is indicated by *Obs.* Dash indicates hybridization not determined.

t Mean of all values except for act- 1800.

Table 3. Relative hybridization in the *obscura* group*

| | D. melanogaster probes | | | | | | | | D. pseudoobscura probes | | |
|------------------|------------------------|--------------------------|------|------|----------|----------|------|--------|-------------------------|--------------------|--|
| Species | act-1800 | $yp-1$ | adh | xdh | $cho-16$ | $hsp-82$ | Ñ. | xdh-ps | $hsp-82ps$ | $\bar{\mathbf{x}}$ | |
| | melanogaster group | | | | | | | | | | |
| me | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 0.14 | 0.08 | 0.11 | |
| obscura subgroup | | | | | | | | | | | |
| su | 0.61 | 0.14 | 0.03 | 0.06 | 0.17 | 0.19 | 0.12 | 0.12 | 0.20 | 0.16 | |
| md | 0.45 | 0.12 | 0.02 | 0.07 | 0.15 | 0.11 | 0.09 | 0.12 | 0.12 | 0.12 | |
| gu | 0.55 | 0.13 | 0.03 | 0.08 | 0.16 | 0.14 | 0.11 | 0.14 | 0.14 | 0.14 | |
| ob | 0.51 | — | 0.02 | 0.08 | 0.16 | 0.10 | 0.09 | 0.15 | 0.14 | 0.15 | |
| tt | 0.54 | $\overline{}$ | 0.02 | 0.08 | 0.15 | 0.15 | 0.10 | 0.15 | 0.17 | 0.16 | |
| am | 0.49 | | 0.02 | 0.08 | 0.14 | 0.10 | 0.09 | 0.11 | 0.14 | 0.13 | |
| Ss | 0.50 | – | 0.03 | 0.06 | 0.15 | 0.12 | 0.09 | 0.11 | 0.16 | 0.14 | |
| bf | 0.57 | | 0.04 | 0.09 | 0.15 | 0.16 | 0.11 | 0.11 | 0.19 | 0.15 | |
| pseudo, subgroup | | | | | | | | | | | |
| ps | 0.64 | | 0.06 | 0.08 | 0.15 | 0.16 | 0.11 | 1.00 | 1.00 | 1.00 | |
| pe | 0.63 | | 0.07 | 0.12 | 0.15 | 0.10 | 0.11 | 1.00 | 0.66 | 0.83 | |

* For details, see Table 1 and the text

COnsistent with the electrophoretic analysis of 18 and 55 enzyme systems, respectively, by Eisses et **al. (1979) and Gonzalez et al. (1982), and with the** high resolution two-dimensional protein electro-Phoretic analysis ofOhnishi et al. (1983). The results are also consistent with data on mitochondrial, ribosomal, and satellite DNAs (Barnes et al. 1978; Cseko et al. 1979; Fauron and Wolstenholme 1980; Coen et al. 1982; Strachen et al. 1982). Finally, the results are consistent with the divergence of the *Adh* gene, as assayed by sequence analysis (Bodmer and Ashburner 1984). The only hybridization intensity values that are out of line, beyond the margin of error, are those of yp-3 in *D. mauritiana*, hsp-82 in *D. simulans,* and cho-s36 in *D. mauritiana* and D. *sirnulans;* duplicate experiments confirmed these deviations. For hsp-82 and cho-s36, the deviations (lower than expected intensities) can be explained by the high molecular weight of the bands (see below).

The five species of the *montium* subgroup studied showed quite uniform hybridization values with **D.** *melanogaster* **probes, averaging 0.13 to 0.15 for 10 different probes, and 0.60 to 0.70 for act-1800 (Ta-**

| | D. melanogaster probes | | | | | | D. pseud. | D. grimshawi probe | | |
|----------------|------------------------|--------------|--------------|--------------|--------------|--------------|---------------------|--------------------|------------------|--------------|
| Species | act- 1800 | $yp-1$ | $yp-3$ | cho- s36 | hsp- 82 | Ā | probe $hsp-82ps$ | cho- sl 8gr | cho- $s15+gr$ | Ā. |
| me tr | 1.00 0.59 | 1.00 0.19 | 1.00 0.21 | 1.00 0.17 | 1.00 0.22 | 1.00 0.20 | 0.08 0.09 | 0.01 0.03 | 0.12 0.11 | 0.07 0.07 |
| ps gu | 0.63 0.56 | 0.18 0.17 | 0.16 0.14 | 0.16 0.13 | 0.16 0.14 | 0.16 0.15 | 1.00 0.13 | 0.02 0.02 | 0.10 0.10 | 0.06 0.06 |
| la ro | 0.46 0.48 | 0.05 0.04 | 0.04 0.06 | 0.06 0.07 | 0.07 0.07 | 0.05 0.06 | 0.03 0.03 | 0.05 0.06 | 0.12 0.30 | 0.08 0.18 |
| gr | 0.47 | 0.04 | 0.05 | 0.07 | 0.10 | 0.07 | 0.04 | 1.00 | 1.00 | 1.00 |
| im fo | 0.57 0.54 | 0.09 0.09 | 0.07 0.06 | 0.15 0.14 | 0.16 0.16 | 0.12 0.11 | 0.06 0.06 | 0.08 0.07 | 0.20 0.21 | 0.14 0.14 |
| hy | 0.45 | 0.11 | 0.06 | 0.11 | 0.14 | 0.11 | 0.04 | 0.13 | 0.24 | 0.18 |
| vi | 0.46 | 0.03 | 0.06 | 0.11 | 0.12 | 0.08 | 0.03 | 0.09 | 0.13 | 0.11 |

Table 4. Relative hybridizations in diverse *Drosophila* species*

* For details, see Table 1 and the text. Heavy bars indicate species groups.

Fig. 2. Sequence divergence in the *melanogaster* group. Conventions as in Fig. 1. See also Table 1 and Materials and Methods. The probes were derived from *D. melanogaster:* act-1600 *(Panel A),* cutic *(Panel B),* adh *(Panel C)* and cho-sl9 *(Panel D).* Note the consistently high intensities corresponding to the *melanogaster* subgroup (with the *melanogaster* complex, me, ma, si slightly exceeding the *yakuba* complex, er). Note also the moderately high intensities in the *takahashii* subgroup (ta) and the low intensities in the *montium* group (au, tr, se, ts, pu). The intensities in the *obscura* group (ps) are even lower

ble 2). Interestingly, *D. takahashii* displayed a somewhat higher debree of hybridization (averaging 0.20 for 10 probes), in agreement with evidence from banding patterns that the *takahashii* group is closer to *D. melanogaster* than is the *montium* group (M. Ashburner, pers. commun.). The *D. takahashii* values were significantly higher than the entire range of the *montium* group values for seven probes, lower

l~i~. 3. Sequence divergence in the *melanogaster* group. Conventions as in Fig. 1. The probes were derived from *D. rnelanogaster: ~P-I (Panel A),* hsp-82 *(Panel B),* cho-s36 *(Panel C),* cho-sl6 *(Panel D),* xdh *(Panel* E), yp-3 *(Panel F).* Results comparable to Fig. 2. As a test of our procedures, 2μ g or 3μ g (rather than the usual 1 μ g) of *D. serrata* DNA was used in panels F and C, respectively; intensities reported in Table 2 have been corrected accordingly

for two probes, and equal to the average for one **.**

A representative of the *obscura* group, *D. pseucloobscura,* hybridized with the *D. melanogaster* probes slightly less intensely than the species of the

montium subgroup. Conversely, when *D. pseudoobscura* probes (xdh-ps and hsp-82ps) were used, all species of the *melanogaster* group (except *D. takahashii)* hybridized equally weakly (Table 2). Both of these results were expected, and confirm the dis-

me me sl er *tO* au tr (ts, pu, se) *obscura* oroup

Fig. 4. (a). Sequence divergence, relative to *D. melanogaster,* in the *rnelanogaster* group (subgroup and complex *melanogaster,* me, ma, si; subgroup *melanogaster,* complex *yakuba,* er; subgroup *takahashii, ta; montium subgroup), and in the <i>obscura group* (ps). Relative divergences from *D. melanogaster* are represented by concentric circles, with the radius (arrow) accompanied by the average relative intensity of hybridization with 11 probes (see Table 2). (b). Phylogenetic relationships, beased on the data summarized in (a), and on previous evidence (see text). The close relationship of au and tr is suggested by the frequent coincidence of hybridizing bands. Branch lengths in the dendrogram are only approximate

tinctness of the *melanogaster* group of species relative to the *obscura* group. The abnormally low values for *D. takahashii* using *D. pseudoobscura* probes might suggest a somewhat higher genome size, or faster rate of sequence divergence (cf. Britten 1986) in that species; either interpretation, in conjunction with the results using *D. melanogaster* probes, reinforces the evidence that *D. takahashii* is closer to *D. melanogaster* than is the *montium* group.

In qualitative terms, similar fragment lengths were observed with several probes (act- 1600, yp- 1, adh, cho-sl9, cho-sl6, and cho-sl5, s18) for *D. melanogaster, D. mauritiana* and/or *D. simulans.* This reinforces the already summarized evidence that these three species are very closely related. Among the species of the *montium* subgroup, only *D. auraria* and *D. triauraria* frequently displayed restriction fragments of the same length (probes cutic, adh, hsp-82, cho-s16). This agrees with the known phylogenetic proximity of these two species (Ohnishi et al. 1983), which could not be inferred from hybridization intensities, since the use of exclusively D . *melanogaster* probes resulted in comparisons of each species to *D. melanogaster,* rather than to each other.

Sequence Divergence and Phylogeny in the obscura *Group*

Similar experiments were performed on 10 species of the *obscura* group, using five probes derived from *D. melanogaster* and two from *D. pseudoobscura* (a member of the *obscura* group). Results are presented in Fig. 5, and intensity values are tabulated in Table 3.

In general, when the probes were from *D. melanogaster,* the *obscura* group species showed uniformly weak bands (averages for four probes relative to *D. melanogaster* ranging from $0.09-0.12$ in 10 species). In conjunction with the findings discussed in the previous section, these results confirm that the *obscura* group is a distinct taxon, all species of which are equally distant from *D. melanogaster.* The small variations in intensities with each probe presumably resulted from random differences in sequence divergence and not from differences in phyletic age or genome size, since they were not consistent for different probes.

When probes derived from *D. pseudoobscura* were used, the only other species that showed intense bands was *D. persimilis,* the well-known sibling species of *D. pseudoobscura* (both North American, in the subgroup *pseudoobscura*; Lakovaara and Saura *1982).* The remaining eight, palearctic species, which are classified in the *obscura* subgroup (Lakovaara and Saura 1982), showed uniformly weak bands, comparable or only slightly stronger than those of *D. melanogaster.* The latter observation indicates that the phyletic separation of the *obscura* and *pseudoobscura* subgroups is not much more recent than the separation of the *melanogaster* and *obscura* groups. A similar conclusion applies to the sepa" ration of the *melanogaster* and *montium* subgroups (see Table 2 and Fig. 4b; Eisses et al. 1979).

It should be noted that the patterns obtained with *D. melanogaster* and *D. pseudoobscura* probes from the equivalent gene were not always qualitatively identical, presumably because the probes were overlapping rather than strictly comparable (see Table 1). In the case of hsp-82, an additional possibility

Fig. 5, Sequence divergence in the *obscura* group. Conventions as in Fig. 1. Four probes were derived from *D. melanogaster* and gave comparable low intensities with all species of the *obscura group.* These probes were: hsp-82 *(Panel A),* eho-sl6 *(Panel B),* adh *(Panel D)* and xdh *(Panel E).* Two probes were derived from *D. pseudoobscura:* hsp-82ps *(Panel C)* and xdh-ps *(Panel 10.* Note the close relationship of ps and pe (most evident by similar intensities in Panels C and F), and the close relationship of su, md, gu (suggested by band coincidences in Panels B, E and F)

for the observation of weak extra bands is random Convergence to sequences of other members of the heat-shock gene family.

Some coincidence of restriction fragments was observed between *D. pseudoobscura* and *D. persimilis,* as expected from their status as sibling species (hsp-82ps; cho-sl6). In addition, the three known sibling species, *D. subobscura, D. madeirensis,* and *D. guanche,* and especially the first two, frequently displayed coincidence of restriction fragments (xdh, cho-sl6, yp-1; for yp-1 data not presented, with a fragment of 2.4 kb showing sequence conservation). In addition, coincidences of restriction fragments were observed between *D. obscura* and *D. ambigua*

Fig. 6. Phylogenetic relationships in the *obscura* group. The close relationships within the *pseudoobscura* subgroup (ps, pc) and in the su, md, gu triad of the *subobscura* group are supported by Fig. 5 as well as previous evidence (see text). The relationships within the rest of the *obscrua* group (bf, ss, am, ob, tt) are only based on evidence summarized by Lakovaara and Saura (1982). Branch lengths are approximate

(cho-s 16); *D. bifasciata* and *D. ambigua* (xdh); and surprisingly *D. bifasciata* and *D. pseudoobscura* (hsp-82).

These findings are in accordance with the phylogenies proposed for the *obscura* group by Lakovaara et al. (1976) and Loukas et al. (1984) (see also review by Lakovaara and Saura 1982). In agreement with electrophoretic data (Loukas et al. 1984) and with cytological data and results of interspecies crosses (Krimbas and Loukas 1984), our limited data confirm that the triad *subobscura-madeirensisguanche* is closely related, and probably that *D. subobscura* is more closely related to *D. madeirensis* than to *D. guanche.* Thus, the phylogeny diagrammed in Fig. 6 is consistent with all data to date.

Sequence Divergence and Phylogeny of Distantly Related Species Groups

Similar experiments were performed on a selection of 11 species, representing seven species groups: the already considered *melanogaster* and *obscura* groups of the subgenus *Sophophora* (two species each), and five species groups of the subgenus *Drosophila,* namely the *robusta* and *immigrans* groups (two species each) and the *hydei, virilis,* and Hawaiian picture-winged groups (one species each). The probes used were derived from *D. melanogaster, D. pseu-* *doobscura,* or the picture-winged fly *D. grimshawi.* To facilitate detection of distantly related sequences, hybridization was performed at 60°C rather than the 65° C used in the previous experiments. Results are presented in Fig. 7, and band intensities are tabulated in Table 4.

Within the accuracy of these experiments, the main conclusion supports the classification of the species groups into two distinct subgenera. Besides the species of origin, the most intense bands with either *D. melanogaster* or *D. pseudoobscura* probes were observed almost invariably in the *melanogaster* and *obscura* groups of the subgenus *Sophophora,* while the bands of the other species (subgenus *Drosophila)* were slightly but distinctly weaker. Conversely, with the *D. grimshawi* probes, the bands of the *melanogaster* and *obscura* groups were slightly less intense than those of species classified in the subgenus *Drosophila.* In this context, it should be noted that the chromosomal locations of actin genes in *D. grimshawi,* measured in relative distances from the proximal end of each chromosomal element, are very similar to those in *D. melanogaster* and D. *madeirensis,* but are rather different in *D. virilis, D. robusta,* or *D. hydei* (Loukas and Kafatos 1986b). Nevertheless, the hybridization results classify D. *grimshawi* with the latter three species in agreement with the taxonomic evidence. The degree of chromosomal rearrangement may not be related to sequence divergence or phyletic distance among drosophilids.

The data are not adequate to distinguish among the species groups of the subgenus *Drosophila,* in terms of evolutionary distance. Carson and Stalker (1969) have proposed that the ancestor of the Hawaiian *Drosophila* might have been a member of the *robusta* group, although an origin from the *immigrans* group has also been considered (Throckmorton 1975). Although *D. grimshawi* more closely resembles the *robusta* group in terms of hybridizing with *D. melanogaster* and *D. pseudoobscura* probes less intensely than do the members of the *immigrans* group, the *robusta* and *immigrans* groups appear equidistant from *D. grimshawi.* Coincidence of restriction fragments also offers no evidence on this matter although it agrees with taxonomic assignments to species groups. The only coincidences observed are for cho-s16, s19 and yp-3 bands of D. *immigrans* and *D. formosana* (both of the *immi-*

Fig. 7. Sequence divergence among seven species groups of *Drosophila.* Conventions as in Fig. 1. Two probes were derived from *D. grimshawi* (subgenus *Drosophila):* cho-s 15 +gr *(Panel A)* and cho-sl 8gr *(Panel B).* Note that high intensities are only seen in D. *grimshawi,* and moderately high intensities only in the other four species groups of the subgenus *Drosophila (Rob, lmm,* hy, vi); intensities are almost invariably lowest in the subgenus *Sophophora (Mel, Obs).* The other four probes were derived from the subgenus *Sophophora.* One, hsp-82ps *(Panel D),* was from *D. pseudoobscura,* and three were from *D. melanogaster:* cho-s36 *(Panel C),* yp-3 *(Panel E)* and yp-1 *(Panel F).* Except for Panel C, where some differences were marginal, the probes derived from the subgenus *Sophophora* dearly hybridized with the species of that subgenus more intensely than with the species of the subgenus *Drosophila*

I^rig. 7. Continued. In Panel C, 2 μg (rather than the usual 1 μg) of *D. grimshawi* DNA was used; the intensity reported in Table 4 has been corrected accordingly

grans group), and for the hsp-82 band ofD. *lacertosa* and *D. robusta* (both of the *robusta* group). Further work is necessary for exploring the phylogeny of the subgenus *Drosophila.*

Recapitulation and Conclusions: The Potential of the Method and Its Limitations

Considered as a whole, the data presented validate Southern hybridization as a useful tool for phylogenetic analysis. With few exceptions, our data for individual probes were congruent with the commonly accepted phylogenies of *Drosophila* species belonging to various species complexes, subgroups, groups, and subgenera. When all the probes were considered, the agreement was excellent.

Under the conditions used, discrimination was particularly good for closely related species. Thus, sibling species were easily discerned (and tentatively ordered in order of similarity); the *melanogaster* and *yakuba* species complexes were clearly separated and subgroups were defined unambiguously. The method showed that, in terms of affinity to *D. melanogaster, D. takahashii* is closer than .the *montium* subgroup, in accordance with unpublished evidence ofM. Ashburner (pers. commun.). Distantly related subgroups within either the *melanogaster* or the *obscura* group appeared almost as disparate as these two groups.

For higher taxa, the data were sufficient to classify the species groups into two subgenera, as commonly accepted. With only two probes from the subgenus *Drosophila,* it was not possible to determine the relative affinities of the groups comprising that subgenus.

One limitation of the present data was the small number of species from which the probes were derived. In the nature of things, this type of experiment evaluates the distance of various species from the species of origin of the probe, but not from each other (Fig. 4a). Establishment of a dendrogram requires either independent taxonomic evidence or the use of probes from several species so that the affinity of any two species can be directly evaluated.

A less important limitation was the lack of direct estimates for genome size; the data suggest that it does not vary widely within the genus but actual measurements would be needed for greater accuracy. A third limitation was that the stringency of hybridization was not varied widely. Undoubtedly, parallel experiments at progressively more permissive criteria would have permitted better evaluation of the affinities among the more distantly related taxa.

It is instructive to consider the relationship between relative hybridization intensities and degree of sequence identity for the *hsp82* gene, which has been sequenced in four species (Blackman and Meselson 1986), and for which probes were derived from two species (Table 5). Clearly, the conditions used were stringent: 89.1-90.3% sequence identity corresponded to 0.07-0.16 hybridization, relative to the homospecific standard. Closer examination of Table 5 in conjunction with Figs. 3, 5, and 7, reveals a fourth limitation of the data: when the hybridizing band was of very high molecular weight (as for hsp-82 in *D. simulans* and *D. virilis),* the relative intensity was somewhat lower than expected on the basis of sequence conservation. The effect is due to partial degradation of bands that approach the average size of the genomic DNA being analyzed. This limitation can be counteracted by the choice of enzymes for restriction of genomic DNA, by the use of multiple probes that hybridize to bands of widely different size, or by the use of very high molecular weight genomic DNA. The limitation would be circumvented altogether by the use of dotblots or slot-blots, rather than Southern blots. That would eliminate the information on band coincidence between closely related species, but would also result in considerable savings in effort, and would eliminate the annoyance of polymorphisms. On the other hand, under permissive conditions spurious crosshybridization to distantly similar sequences might create problems which could be avoided by examination of Southern blots.

Table 5 also shows that at the sequence level this part of the *hsp82* gene is unexpectedly similar between *D. virilis* and *D. melanogaster* (more than between *D. pseudoobscura* and either *D. melanogaster* or *D. virilis).* Such random deviations are to be expected, and presumably account for some irregularities in Tables 1-4. To counteract them, it is necessary to base the analysis on multiple probes.

A recommended strategy for further development of this method would be as follows:

1. Clone selected unique sequences from several key species (covering the range of expected phylogenetic affinities). The sequences need not be extensively characterized, although it would be preferable that some of the probes from different species be the same, and that they include some known genes, diverging at different rates. We recommend that at least three different sequences be used from each key species.

2. Determine genome sizes, if there is likelihood of major variations.

3. Prepare many replicate slot-blots, using commercially available equipment. They are produced more easily than Southern blots, and are easy to scan.

4. Hybridize the blots with the various probes under various conditions of stringency (recommended: $5-40^{\circ}$ below Tm, depending on the phylogenetic distance of the species). Scan the blots.

Table 5. Percent sequence identities in hsp-82 compared to relative hybridization intensities*

| probe Genomic DNA | | D. melano- gaster 796 bp | D. pseudo- obscura 1035 bp | | |
|--------------------------------|-----------|----------------------------------|----------------------------------|--|--|
| D. melanogaster | kb 7.5 | 100% (1.0) | (0.08) 89.5% (0.10) (0.08) | | |
| D. simulans | 19 | 98.5% (0.53) | 89.8% (0.07) | | |
| D. pseudoobscura | 3.0 | (0.16) 89.1% (0.16) (0.16) | (1.0) 100% | | |
| $D.$ virilis | 18 | 90.3% (0.12) | 87.4% (0.03) | | |

* Percent identities for the region hybridizing with the probe were calculated from Blackman and Meselson (1986). Relative hybridization intensities are from Tables 2-4. See text for details.

5. After the results are analyzed, cross-check uncertainties by performing selected Southern blots.

We view this approach as complementary rather than necessarily preferable to existing methods. It ^{is} best suited to evaluating relative affinities to key Species, rather than working out *ab initio* the full topology, or accurately estimating phylogenetic distances. Thus, the method should be particularly useful for situations in which one wants to fit uncertain taxa into an otherwise well-established phylogenetic framework (cf. the conclusions concerning the *takahashii* group). The method is clearly more time-COnsuming and less quantitative than melting anal-Ysis of total DNA hybrids (Sibley and Ahlquist 1984; Britten 1986). However, a significant advantage is the fact that multiple discrete probes are used, and that one may choose as probes known genes, rather than the ill-defined extragenic sequences that inevitably dominate hybridizations of total *DNA.* Finally, since the probes are cloned, one always has the option of resorting to the ultimate arbiter of Uncertainties: determination of nucleotide sequences.

Acknowledgments. We thank R. Blackman, E. Fyrberg, J.C. Martinez-Cruzado, W. Orr, M. Riley, M. Snyder, N. Spoerel and P. Wensink for gifts of cloned sequences; P. Hatzopoulos and S.A. Mitsialis for advice; B. Klumpar for technical help and photography; M. Youk-See for graphics; and A. Houlaki and E. Fenerjian for secretarial assistance. The work was supported by American Cancer Society and NIH grants to F.C. Kafatos; M. LOukas was recipient of a long-term EMBO fellowship.

References

- Abraham G, Scaletta C, Vaughan J (1972) Modified diphenylamine reaction for increased sensitivity. Analytical Biochemistry 49:547-549
- Angerer RC, Davidson EH, Britten RJ (1976) Single copy DNA and structural gene sequence relationships among four sea urchin species. Chromosoma 56:213-226
- Barnes SR, Webb DA, Dover G (1978) The distribution of satellite and main-band DNA components in the *melanogaster* species subgroup *of Drosophila.* Chromosoma 67:341- 363
- Barnett T, Pachl C, Gergen J, Wensink P (1980) The isolation and characterization of Drosophila yolk protein genes. Cell 21:729-738
- Beverley SM, Wilson AC (1984) Molecular evolution in *Drosophila* and the higher diptera. II. A time scale for fly evolution. J Mol Evol 21:1-13
- Blackman RK, Meselson M (1986) Interspecific nucleotide sequence comparisons used to identify regulatory and structural features of the *Drosophila hsp82* gene. J Mol Biol 188:499- 515
- Bodmer M, Ashburner M (1984) Conservation and change in the DNA sequences coding for alcohol dehydrogenase in sibling species of *Drosophila.* Nature 309:425-430
- Bond BJ, Davidson N (1986) The *Drosophila melanogaster* actin 5C gene uses two transcription initiation sites and three polyadenlyation sites to express multiple mRNA species. Mol Cel Biol 6:2080-2088
- Britten RJ (1986) Rates of DNA sequence evolution differ between taxonomic groups. Science 231:1392-1398
- Carson HL, Stalker HD (1969) Polytene chromosome relationships in Hawaiian species *of Drosophila.* IV. The *D. primaeva* subgroup. Univ Texas Publ 6918:85-94
- Church GM, Gilbert W (1984) Genomic sequencing. Proc Nail Acad Sci USA 81:1991-1995
- 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.* J Mol Biol 158: 17-35
- Collier CE, Maclntyre RJ (1977) Micro-complement fixation studies on the evolution of alpha-glycerophosphate dehydrogenase within the genus *Drosophila.* Proc Natl. Acad Sci USA 74:684-688
- Cot6 B, Bender W, Curtis D, Chovnick A (1986) Molecular mapping of the *rosy* locus in *Drosophila melanogaster.* Genetics 112:769-783
- Cseko YMT, Dower NA, Minoo P, Lowenstein L, Smith GR, Stone J, Sederoff R (1979) Evolution of polypyrimidines in *Drosophila.* Genetics 92:459-484
- Eisses K, vanDijk H, vanDelten W (1979) Genetic differentiation within the *melanogaster* species group of the genus *Drosophila (Sophophora).* Evolution 33:1063-1068
- Fauron CM-R, Wolstenholme D (1980) Extensive diversity among *Drosophila* species with respect to nucleotide sequences within the adenine + thymine-rich region of mitochondrial DNA molecules. Nucl Acids Res 8:2439-2452
- Fyrberg E, Kindle K, Davidson N (1980) The actin genes of Drosophila: a dispersed multigene family. Cell 19:365-378
- Goldberg DA (1980) Isolation and partial characterization of the *Drosophila* alcohol dehydrogenase gene. Proc Natl Acad Sci 77:5794-5798
- Gonzalez A, Cabrera V, Larruga J, Gullon A (1982) Genetic distance in the sibling species *Drosophila melanogaster, Drosophila simulans* and *Drosophila rnauritiana.* Evolution 36: 517-522
- Gritfin-Shea R, Thireos GC, Kafatos FC (1982) Organization of a cluster of four chorion genes in *Drosophila* and its relationship to developmental expression and amplification. Devel Biol 91:325-336
- Hung M-C, Wensink PC (1981) The sequence of the *Drosophila melanogaster* gene for yolk protein- 1. Nucl Acids Res 9:6407- 6419
- Krimbas CB, Loukas M (1984) Evolution of the obscura group *Drosophila* species. I. Salivary chromosomes and quantitative characters in *D. subobscura* and two closely related species. Heredity 53:469-482
- Lachaise D, Tsacas L (1974) Les *Drosophilidae* des savanes préforestières de la région tropicale de Lamto (Côte d'Ivoire). II. Le peuplement des fruits de *Pandanus candelabrum (Pan*danacées). Ann Univ Abidjan E 7:153-192
- Lakovaara S, Saura A, Lankinen P, Pohjola L, Lokki J (1976) The use of isoenzymes in tracing evolution and in classifying Drosophilidae. Zoologica Scripta 5:173-179
- Lakovaara S, Saura A (1982) Evolution and speciation in the *Drosophila obscura* group. In: Ashburuer M, Carson HL, Thompson JN (eds) The genetics and biology of Drosophila, vol. 3b. Academic Press, London, NY, pp 1-59
- Lemeunier F, Ashburner M (1976) Relationships within the *melanogaster* species subgroup of the genus *Drosophila (Sophophoraj.* II. Phylogenetic relationships between six species based upon polytene chromosome banding sequences. Proc R Soe Lond B 193:275-294
- Loukas M, Krimbas CB, Vergini Y (1984) Evolution of the obscura group *Drosophila* species. II. Phylogeny of ten species based on electrophoretic data. Heredity 53:483-493
- Loukas M, Kafatos FC (1986a) Chromosomal locations of actin genes are conserved between the *rnelanogaster* and *obscura* groups of *Drosophila.* Genetica (in press)
- Loukas M, Kafatos FC (1986b) The actin loci in the genus *Drosophila:* Establishment of chromosomal homologies among distantly related species by in situ hybridization. Chromosoma (in press)
- Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. A laboratory manual. Cold Spring Harbor Laboratory

Ohnishi S, Kaanishi M, Watanabe TK (1983) Biochemical phy-

logenies of *Drosophila:* protein differences detected by twodimensional electrophoresis. Genetica 61:55-63

- Sibley CG, Ahlquist JE (1984) The phylogeny of the hominoid primates as indicated by DNA-DNA hybridization. J Mol Evol 20:2-15
- Snyder M, Hirsh J, Davidson N (1981) The cuticle genes of *Drosophila:* A developmentally regulated gene cluster. Cell 25:165-177
- Spradling AC (1981) The organization and amplification of two chromosomal domains containing *Drosophila* chorion genes. Cell 27:193-201
- Steinemann M, Pinsker W, Sperlich D (1984) Chromosome homologies within the *Drosophila obscura* group probed by in situ hybridization. Chromosoma 91:46-53
- Strachen T, Coen E, Webb D, Dover G (1982) Modes and rates of change of complex DNA families of *Drosophila.* J Mol Biol 158:37-54
- Thompson EOP (1980) Amino acid sequences of globin chains and their use in phylogenetic divergence point estimations. In: Sigman DS, Brazier MAB (eds) The evolution of protein structure and function. Academic Press, NY, pp 267-298
- Throckmorton LH (1975) The phylogeny, ecology and geography *of Drosophila.* In: King RC (ed) Handbook of genetics, vol. 3. Plenum Press, NY, pp. 421-469
- Wilson AC, Carlson S, White T (1977) Biochemical evolution. Ann Rev Biochem 46:573-639
- Zwiebel LJ, Cohn VH, Wright DR, Moore GP (1982) Evolution of single-copy DNA and the ADH gene in seven drosophilids. J Mol Evol 19:62-71