Evolution of Single-Copy DNA and the ADH Gene in Seven Drosophilids

Laurence J. Zwiebel, Vivian H. Cohn, Donna R. Wright, and Gordon P. Moore

Division of Biological Sciences, Graduate Program in Cellular and Molecular Biology, The University of Michigan, Ann Arbor, Michigan 48109, USA

Summary. Single-copy DNA was isolated from *Drosophila melanogaster* and hybridized with total genomic DNA of *D. melanogaster, D. mauritiana, D. simulans, D. pseudoobscura, D. willistoni, D. hydei* and *D. virilis*. The duplexes were thermally eluted from hydroxyapatite and the data used to assess the relatedness of each species to *D. melanogaster*. The general pattern of relatedness was similar to that predicted by morphological methods but with some notable exceptions. The rate of nucleotide substitution was estimated to be greater than 0.66% of bases per million years. An unexpected, rapidly evolving component of *D. melanogaster* single-copy DNA was identified. The relatedness of these species was also studied with respect to the gene coding for alcohol dehydrogenase (ADH). The ADH gene, previously cloned from *D. melanogaster* (Goldberg 1980), was hybridized with Southern blots of genomic digests of the seven species. The intensity and position of the hybridizing bands suggest the amount of divergence of the gene. Divergence was quantitated by reassociation of a fragment of the cloned ADH gene with total DNA of the seven drosophilids and thermal elution of the resultant duplexes from hydroxyapatite. The ADH gene was isolated from genomic clone libraries of *D. melanogaster, D. simulans* and *D. mauritiana* and further studied by comparison of position of restriction sites. Species relationships deduced from comparison of total single-copy DNA and the ADH gene were consistent, demonstrating that a single gene can reflect divergence of the entire genome.

Key words: D. melanogaster — Genomic evolution — Thermal elution — Single-copy DNA — Alcohol dehydrogenase

Introduction

The idea that the course of evolution can be traced by analysis of biological macromolecules depends on the hypothesis that molecular divergence is linear with time (Wilson et al. 1977). As this concept has gained acceptance, phylogenetic analysis has moved to the molecular level. Interspecies molecular comparison have evolved from the study of biochemical properties of proteins (e.g. electrophoretic mobility, antigenicity, etc.) to their amino acid sequence and finally to the nucleotide sequence of DNA. Detailed interspecies comparison of individual genes has yielded important data regarding the processes of evolution (e.g. Perler et al. 1980; Efstratiadis et al. 1980). However, a disadvantage of fine-scale analysis is that only a minute fraction of the total genome can be studied using present methods. An alternative is analysis of DNA sequence relationships of a large fraction of the genome by hybridization/melting techniques. These procedures rely on the fact that mismatched DNA duplexes have a lowered melting point of about 1°C for each 1% of base mismatch (Hutton and Wetmur 1973). Single-copy DNA is prepared, labelled and hybridized with excess total DNA of the same (homologous) or different (heterologous) species (Britten and Kohne 1967; Kohne et al. 1971). The resultant duplexes are isolated and their thermal stability compared. The advantage of this method is that a huge number of DNA sequences are simultaneously compared. In the present study, we use interspecies comparisons at both the level of total single-copy DNA and a single cloned gene for an integrated view of evolution in drosophila.

Seven species of drosophila were chosen because of their position relative to *Drosophila melanogaster* in the evolutionary line of descent as deduced by morphological studies. As shown in Fig. 1 (adapted from Throck-
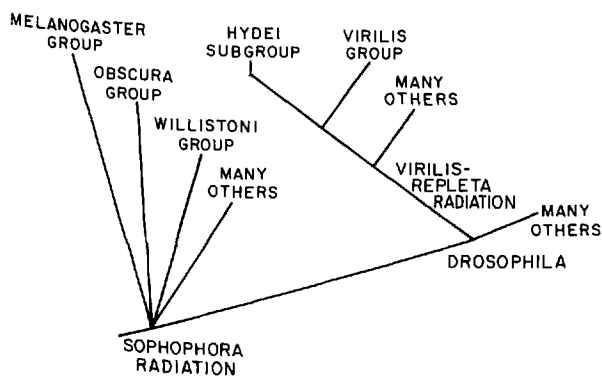
Fig. 1. Evolutionary line of descent of seven extant drosophilids (adapted from Throckmorton 1975)

morton 1975) the subgenus scaptodrosophila gave rise to the sophophora radiation, which includes the melanogaster group, as well as the drosophila radiation which itself gave rise to the virilis and repleta radiations. *D. simulans* and *D. mauritiana* were selected as members of the melanogaster group closely related to *D. melanogaster*. Indeed, *D. simulans* and *D. melanogaster* can be mated in the laboratory and probably form hybrids in nature as well (Sperlich 1962). *D. pseudoobscura* and *D. willistoni* were chosen to represent other groups in the sophophora radiation, while *D. virilis* and *D. hydei* represent the virilis-repleta radiation.

In addition to determination of the divergence of single-copy DNAs, evolution of the ADH gene has been studied. The ADH gene is particularly suitable for evolutionary comparison because it has been well characterized and cloned from *D. melanogaster*. Numerous ADH mutants exist and the gene has been mapped both genetically and cytologically. Electrophoretic variants of ADH protein have been identified, and their distribution in *D. melanogaster* populations studied. Moreover, the amino acid sequence of *D. melanogaster* ADH is known (Thatcher 1980) and the complete DNA sequence of the gene has been published (Benyajati et al. 1980, 1981). In the work reported here, genomic evolution in drosophila has been examined with respect to single-copy DNA relationships and sequence divergence of the ADH gene. We also describe isolation of clones containing the ADH gene from *D. simulans* and *D. mauritiana* and their preliminary characterization by restriction mapping.

Materials and Methods

Animal Procedures. The seven drosophilids were obtained from the Mid-Americas Drosophila Stock Center in Bowling Green, Ohio. All the species were maintained on standard *D. melanogaster* medium at 25°C except *D. pseudoobscura* which was maintained at 21°C.

Isolation of DNA. DNA for the interspecies thermal melting experiments and for genomic blotting was isolated from adult flies frozen at -70°C. Five grams of flies were ground first in a

mortar, then in a Potter-Elvejen motor driven homogenizer in 0.05M Tris pH 7.6, 0.025M KCl, 0.05M Mg acetate, 0.35M sucrose. The homogenate was centrifuged at 4000 rpm in a Sorvall SS34 rotor for 15 min at 4°C. The pellet was suspended in 0.15M NaCl, 0.1M EDTA pH 8.0, then adjusted to 2% SDS and heated to 60°C for 10 min. The mixture was adjusted to 1M NaHClO$_4$, titrated to pH 8.0 and extracted once with redistilled phenol and three times with 24:1 chloroform/isoamyl alcohol. DNA was precipitated with ethanol, pelleted, and suspended in 0.1 X SSC. The mixture was incubated with 100 μg/ml preboiled RNAse A (Sigma) and 1500 units RNase Tl (Sigma) for 2 h at 37°C. 1000 units of α-amylase were added and incubation continued another hour. The DNA was again extracted with phenol and chloroform/isoamyl alcohol, precipitated and resuspended in 10 mM Tris pH 8.0, 1 mM EDTA. Yields were about 1 mg of DNA per 5 g flies.

Shearing and Sizing of Driver DNAs. DNAs used as driver in the thermal melting experiments were isolated as described above, and sheared by sonication in a Branson 185 cell disrupter. 200 μg of DNA were sonicated in 2 ml of 10 mM Tris pH 8.0, 1 mM EDTA for 40s on setting 3.2. The single-strand size of the sheared DNAs was estimated from their electrophoretic mobility relative to standards on alkaline-agarose gels (Hall et al. 1980). The gels were 1% agarose, 30 mM NaOH, 2 mM EDTA. Standards were ØX17RF DNA digested with Hae III and λ DNA digested with Hind III. The gels were stained in 1 μg/ml ethidium bromide and photographed with a yellow filter. As expected, the sheared DNAs had a broad length distribution. The mode of the length distribution was estimated to be approximately 1200 nucleotides for each driver DNA used.

Preparation of D. melanogaster Single-Copy Tracer. One hundred micrograms of total *D. melanogaster* (Oregon R) DNA were sonicated as described above then precipitated, resuspended in 1 ml of 0.12 M phosphate buffer pH 7.6 (PB), 1 mM EDTA, 0.05% SDS. The DNA was boiled 90 seconds in a sealed capillary tube then reassociated at 60°C to equivalent Cot 12 mol l^{-1}s. A water-jacketed glass column was packed with 0.4 g of hydroxyapatite (BioRad DNA grade HAP lot 19693) and the DNA applied at 50°C, 0.12M PB. DNA which was not bound by HAP (81% of input) was dialyzed to 1 mM EDTA, precipitated, resuspended in 0.1 mM EDTA then adjusted to 0.41 M PB in 0.2 ml. This was boiled 90 s, then reassociated at 60°C to equivalent Cot 1300 mol l^{-1} s. The reassociated DNA was dialyzed then labelled by the "gap-translation" method described by Galau et al. (1974) using ^{3}H-GTP (NEN) and *E. coli* polymerase I (BRL) but no deoxyribonuclease. The specific activity of the tracer was 5 × 10^5 cpm/μg. The tracer was precipitated, resuspended, adjusted to 0.12 M PB boiled, reassociated at 60°C to equivalent Cot 33 mol l^{-1}s and fractionated on HAP as above. Removal of double-stranded material at this step eliminated artifactual foldback DNA generated during the gap-translation procedure and any remaining repetitive DNA. Material unbound by HAP was passed over a Sephadex G-100 column then bound to HAP in 0.012 M PB in order to remove unincorporated nucleotides and short fragments of DNA.

The single-strand length of the single-copy tracer was determined by centrifugation through an isokinetic alkaline sucrose gradient (Noll 1967). V_{mix} = 9.84 ml, C_{res} = 43% sucrose (W/V), C_{flask} = 16% sucrose (W/V) in 0.1M NaOH. Gradients were centrifuged at 33,000 rpm for 30 h at 20°C in a Beckman SW41 rotor. The modal length of the tracer was determined in relation to internal standards previously sized in the electron microscope (Studier 1965). Sucrose gradients were fractionated and the absorbance at 260 nm measured. Each fraction was neutralized and counted in Aquasol 2 (NEN) in a liquid scintillation counter. Comparison of the migration of the ^{3}H-labelled

tracer to that of the optical standards indicated a modal length of 510 nucleotides.

Hybridization and Thermal Melting of Single-Copy DNA. Twenty-five micrograms of each sheared driver DNA were precipitated and resuspended in a final volume of 30 μl of 0.41M PB, 1 mM EDTA, 0.05% SDS. The drivers were sealed in capillary tubes with 10^4 cpm of D. melanogaster single-copy tracer, boiled 60 s then reassociated to equivalent Cot $\geq$ 2500 mol 1^{-1} s at 55°C. The ratio of driver DNA to tracer DNA was 1000. Samples were diluted to 1 ml of 0.12M PB, then applied to 0.4g HAP columns (three columns were run in series). Unbound material was washed away, then bound duplexes were eluted with 0.12 M PB by increasing temperature in 4°C increments as described by Moore et al. (1978). Eluate was collected directly into scintillation vials and counted.

Genomic Southern Blotting. The gene coding for D. melanogaster (Canton S) ADH was isolated, subcloned into the plasmid pBR322, named pSAC and graciously provided by Dr. David Goldberg. As shown in Fig. 5, the subclone consists of the smallest genomic Eco RI fragment which contains the entire structural gene as well as about 1 Kb of flanking DNA in both directions. pSAC DNA was prepared by the method of Clewell and Helinski (1970) and ^{32}P-labelled by nick-translation (Rigby et al. 1977) to specific activity 1-2 X 10^8 cpm/μg using all four α-^{32}P-dNTP (NEN). DNA of the seven drosophilids, prepared as above, was digested twice with a 3-fold excess of Eco RI and 2 μg were electrophoresed in 0.7% agarose gels for 15 h at 70 mA. DNA was transferred to nitrocellulose filters (BA 85 Schleicher and Schuell) using the Southern procedure (Southern 1975) except that denaturation was in 0.1M NaCl, 0.5 NaOH and neutralization was in 0.1M NaCl, 0.5M Tris. Hybridization was performed in 50% formamide (Fisher), 5 X SSC, 0.1% SDS, 10% dextran sulfate at 34°C. The filters were washed in 2 X SSC, 0.1% SDS, 0.1% Na pyrophosphate at 50°C. Autoradiograms were exposed at -70°C using Kodak XAR-5 X-ray film and Kodak C-2 film cassettes with Dupont Lightening Plus intensifiying screens.

Labelling and Strand-Separation of a Fragment of pSAC. A 610 nucleotide fragment of pSAC was prepared by digestion with Bam HI and Hind III as shown in Fig. 5. Digestion products were end-labelled with ^{32}P using modification of the procedure of Drouin and Symons (1979). Ten micrograms of pSAC were digested, precipitated and resuspended in a final reaction volume of 35 μl. Labelling was at 14°C for 1 h in 50 mM NaCl, 50mM Tris pH 7.4, 8mM MgCl$_2$ in the presence of 50 microcuries (1.85 x 10^6 becquerels) of each α-^{32}P-dNTP (NEN; 3000 mCi/mMol) and 2.5 units of the large fragment of *E. coli* DNA polymerase I (NEN) (Klenow and Hansen 1974). Labelled DNA was precipitated, then suspended in 25 μl of 30% (vol/vol) dimethylsulfoxide, 1 mM EDTA, 0.05% (wt/vol) xylene cyanol, 0.05% (wt/vol) bromphenol blue. Isolation of the 610 nucleotide fragment and separation of the DNA strands was performed as described by Maxam and Gilbert (1977) except that electrophoresis was in 4% (wt/vol) acrylamide crosslinked with 0.08% (wt/vol) methylene bis-acrylamide at 350 volts for 20 h. The gel was stained with ethidium bromide in the presence of 5mM ammonium acetate and the separated strands of the 610 nucleotide fragment were sliced out. DNA was eluted from the gel slices by diffusion into 10mM Tris, 1mM EDTA at 37°C overnight. The specific activity of the labelled, strand-separated tracer was estimated to be 3 x 10^6 cpm/μg. For the hybridizations discussed below, the slower migrating band was used. This was the nonsense strand labelled at the Bam site. The extent

of opposite strand contamination was measured as 7% by self-hybridization and binding to hydroxyapatite. The small amount of tracer-tracer duplex was normalized from the thermal elution curves of Fig. 4.

Hybridization and Thermal Elution of a Fragment of pSAC. Genomic DNA of seven drosophilids was prepared and sheared to a length of about 1000 nucleotides as described above. Aliquots of 150 μg were precipitated and suspended in a final volume of 100 μl of 0.41M PB, 0.05% SDS, 1mM EDTA and 10^4 cpm of the 610 nucleotide ADH probe whose preparation was described above. Each sample was sealed in a glass capillary tube, boiled for 90 s then incubated at 60°C for 36 h to an equivalent Cot of $\geq$ 2500. Samples were diluted to 0.12 M PB, bound to HAP and thermally eluted in 4°C increments from 60°C to 98°C as described above.

Isolation of ADH-Containing λ Clones. Genomic DNA of D. simulans and D. mauritiana was cloned in λ Charon 4A (Blattner et al. 1977) using methods described by Maniatis et al. (1978) and kindly provided by Dr. M. Messelson. pSAC was labelled with ^{32}P by nick-translation and used to screen λ libraries by modification of the method of Benton and Davis (1977) as described by Anderson et al. (1981). 60,000 recombinants of D. simulans were screened and four ADH-containing λ clones were isolated. Five ADH λ clones were isolated from 100,000 recombinants of D. mauritiana. Hybridization was done in 5 x SSC, 1% Sarkosyl at 58°C and washes were done in 3 x SSC, 0.5% Sarkosyl and 3 x SSC alone at 57–58°C.

Results

Evolutionary Divergence of Single-Copy DNA

Measurement of thermal stability of heterologous DNA duplexes provides an estimate of average sequence difference in DNA of related species. The conditions of reassociation dictate which sequences are, or are not, too divergent to form stable duplex structures. Using low stringency conditions (55°C, 0.41M phosphate buffer) which maximize hybridization of divergent sequences, a series of reassociations were performed between single-copy, ^{3}H DNA from *D. melanogaster* and excess, unlabelled DNA of seven drosophilids. Figure 2 presents representative elution profiles obtained when the hybridization products were bound to hydroxyapatite, then eluted by incremental increase in temperature. Data from the measurements of Fig. 2, as well as others, are summarized in Table 1.

The T_m 1/2 of the *D. melanogaster/D. melanogaster* duplex was 84.1°C (Fig. 2, Table 1). This is close to the theoretical value of 84.6°C, calculated using the equations of McConaughy et al. (1969) and Schildkraut and Lifson (1965), assuming 45% G-C content of *D. melanogaster* DNA (Laird and McCarthy 1968). The closeness of the measured value to the theoretical T_m 1/2 demonstrates that the single-copy tracer was not extensively nicked or contaminated with repeat DNA, since length reduction or repeat contamination would have lowered the T_m 1/2 significantly. The hydroxy-

Table 1. Divergence of single-copy DNAs measured by thermal elution from hydroxyapatite

Species	Number of determinations	Average T_m 1/2 (°C)[a]	ΔT_m 1/2 (°C)[b]	Average % reaction[c]	Corrected % reaction[d]
D. melanogaster	4	84.1	–	80.1	100
D. simulans	4	82.2	1.9	51.7	62.8
D. mauritiana	3	81.8	2.3	53.0	66.4
D. pseudoobscura	3	65.2	18.9	20.8	25.7
D. hydei	1	58.0	26.1	10.9	13.4
D. willistoni	3	57.6	26.5	13.4	16.5
D. virilis	3	57.9	2 62	15.1	18.6

[a] The T_m 1/2 values in multiple determinations were averaged. The variance in multiple determinations was $\leqslant 1°C$ for the closely related species (D. melanogaster, D. simulans and D. mauritiana). In more distant species, variation was $\leqslant 2.5°C$

[b] The ΔT_m 1/2 is the difference between the average T_m 1/2 observed for D. melanogaster/D. melanogaster duplexes and the average T_m 1/2 observed in heterologous measurements

[c] The percent reaction observed in multiple determinations was averaged; variation was $\leqslant 5\%$ of the total reactivity

[d] The percent reaction was corrected for the reactivity of tracer which was 80.1%

apatite method has some disadvantages relative to other melting techniques, notably the TEACl method discussed by (Britten et al. 1978). T_m 1/2 values measured by HAP elution are only reproducible to an accuracy of about ± 1°C. However, in cases where both methods have been employed the results have been similar (e.g. Hall et al. 1980).

The results presented in Table 1 are generally consistent with the phylogeny shown in Fig. 1. The divergence between the sibling species D. melanogaster and D simulans or D. mauritiana is much less (ΔT_m 1/2 of 1.9°C and 2.3°C respecitvely) than any of the other species compared. The divergence between D. melanogaster and D. pseudoobscura leads to a ΔT_m 1/2 of 18.9°C. According to the data listed in Table 1, these species, both in the sophophora radiation, are more closely related than is D. melanogaster with D. virilis or D. hydei. The data in Table 1 do not distinguish between the divergence of D. melanogaster and D. willistoni, D. hydei or D. virilis, all of whose genomes have little homology with D. melanogaster.

In addition to providing quantitation of sequence divergence, the data of Table 1 also show what fraction of single-copy DNA is capable of forming homologous and heterologous duplexes. This parameter is generally proportional to the phylogenetic relationships indicated both by Fig. 1 and by the melting data. However, measurement of the fraction of D. melanogaster single-copy DNA capable of hybridization with total DNA of D. simulans or D. mauritiana yielded an unexpected result. Only about 65% of the D. melanogaster tracer hybridized with either of the other genomes. This result indicates that $\sim 35\%$ of D. melanogaster single-copy sequences are totally absent in D. simulans and D. mauritiana or so modified that they fail to hybridize. This is true although 1) the ΔT_m 1/2 of the duplexes which do form is very low, 2) D. melanogaster and D. simulans are

extremely similar morphologically and can cross-mate and 3) the polytene chromosome banding patterns of these species are similar (Patterson and Stone 1952).

Evolutionary Divergence of the ADH Gene Studied by Genomic Blotting

The gene coding for ADH was studied in the same seven drosophilids. A plasmid containing the D. melanogaster ADH gene (pSAC, see Fig. 5) was labelled and used to probe Southern transfers of total genomic DNA; Fig. 3 shows the results. The intensity and position of the bands gives qualitative indication of how closely the ADH gene is related in each species to that of D. melanogaster. Analysis of several exposures of the blot shown in Fig. 3, as well as a number of other blots done under varying conditions lead to the synthesis of data shown in Table 2. No single set of conditions is appropriate to show clearly all the bands because the range of divergence in the seven species is too great. Each of the species appears to contain the ADH gene and the genes are similar enough to those of D. melanogaster to hybridize under the conditions used. The relationships of D. melanogaster to the other species as deduced from the ADH band intensities listed in Table 2 are consistent with the relationships indicated by comparison of total single-copy DNA shown in Table 1. Thus divergence of DNA coding for a single gene appears to reflect the vastly larger data base represented by total single-copy DNA. In the case of D. melanogaster, D. simulans and D. mauritiana not only are the bands of high intensity, the similarity in band size indicates conservation of restriction sites. Hybridization of another single-copy probe with genomic DNA of the same species yielded an identical ordering of species relatedness (L. Zwiebel unpublished observation).
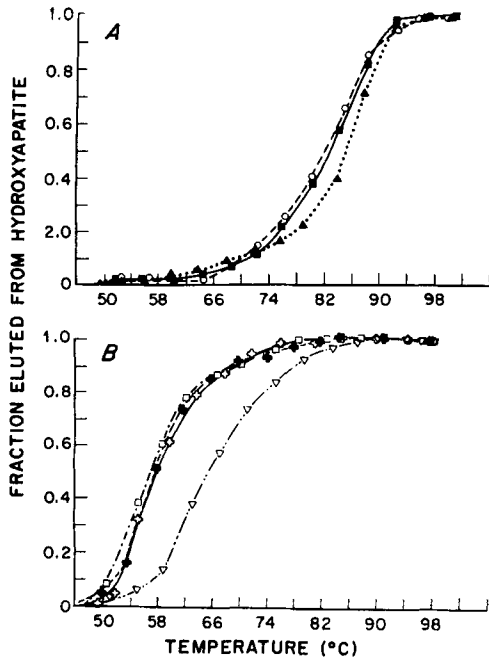
Fig. 2 A and B. Thermal elution from hydroxyapatite of duplexes formed between H³-labelled *D. melanogaster* single-copy DNA and excess total DNA of seven drosophilids. Reassociations were in 0.41M PB, 0.05% SDS, 1 mM EDTA at 55°C to equivalent Cot $\geqslant$ 2500 mol l⁻¹s. After hybridization, samples were diluted to 0.12M PB and applied to a series of 0.4g hydroxyapatite columns at 46°C. Temperature was increased in approximately 4°C increments and the DNA fragments rendered single-stranded at each increment were eluted with 0.12M PB. Driver DNA's were: A (▲) *D. melanogaster*, (○) *D. mauritiana*, (■) *D. simulans*; B (▽) *D. pseudoobscura*, (◊) *D. hydei*, (✚) *D. willistoni*, (□) *D. virilis*

Table 2. Order of band intensity in blots of Eco R1 digested genomic DNA probed with pSAC

Species	Band length (Kb)[a]	Ordered band Intensity[b]
D. melanogaster	4.7[c]	1
D. simulans	4.7	2[d]
D. mauritiana	4.7	2[d]
D. pseudoobscura	21, 10.2, 5.2	4
D. hydei	6.8	5
D. willistoni	2.9, 2.2	6
D. virilis	e	7

a Determined relative to the electrophoretic mobility of Hind III digested λ DNA

b The intensity of the faint bands in Fig. 3 was photographically enhanced to show their positions more clearly. The order of band intensity was qualitatively determined from many blots carried out under a wide variety of hybridization conditions. No one set of hybridization conditions was appropriate for all of the species since the range of sequence divergence was great

c The faint band at 6.8 Kb seen in Fig. 3 was the product of incomplete digestion of *D. melanogaster* genomic DNA. The Eco RI site 5' to the ADH gene was incompletely digested leading to a larger fragment as shown in the restriction map of Fig. 5

d The intensity of the ADH band in *D.mauritiana* and *D. simulans* genomic DNA was not reproducibly different

e The signal was so weak that band position could not be reproducibly determined
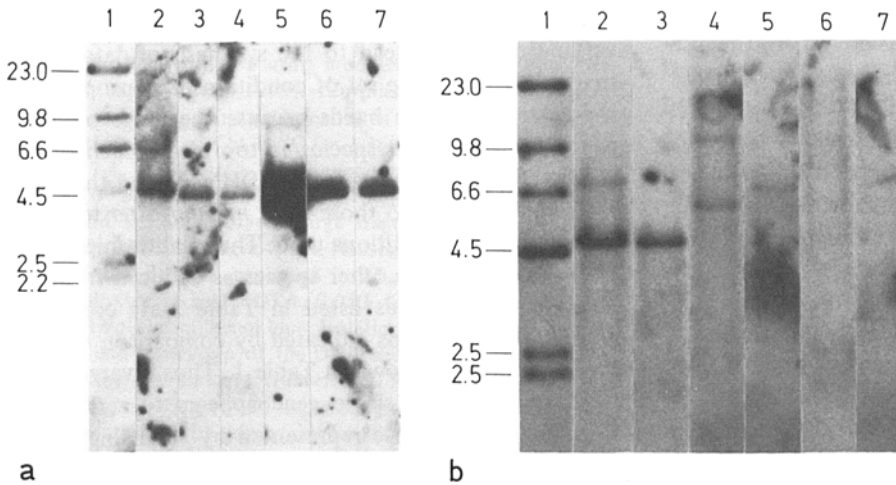

Fig. 3 A and B. Blot hybridization of a cloned ADH gene with genomic DNA of seven drosophilids. 2 µg of each genomic DNA were digested with Eco RI and electrophoresed through a 0.7% agarose gel. The DNA was transferred to a nitrocellulose filter, hybridized with the probe (³²P-labelled pSAC) DNA and autoradiographed, as described in the Methods section. The DNAs were A 1) Hind III digested λ DNA, 2) *D. melanogaster*, 3) *D. simulans*, 4) *D. mauritiana*, 5) pSAC, 6) isolated ADH-containing λ clone of *D. simulans*, 7) isolated ADH-containing λ clone of *D. mauritiana*; B 1) Hind III digested λ DNA, 2) *D. melanogaster*, 3) *D. simulans*, 4) *D. pseudoobscura*, 5) *D. hydei*, 6) *D. willistoni*, 7) *D. virilis*

Some of the lanes of the blot shown in Fig. 3 contain more than one hybridizing band. In the *D. melanogaster* lane, the faint band at 6.8 Kb results from incomplete digestion of the Eco R1 site 5' to the ADH gene (see Fig. 5). Interestingly, the same site has been resistant to digestion in genomic clones as well. It seems that the nucleotide sequence which makes up this particular Eco R1 site, or its immediate sequence environment, renders it a less suitable substrate for the enzyme than other Eco R1 sites. However, more exhaustive digestion results in complete cleavage. The extra bands seen in digests of *D. pseudoobscura* and *D. hydei* are probably not the result of partial digestion since hybridization with a different single-copy probe resulted in only one band (data not shown). Aside from incomplete digestion, several explanations are possible. These include multiple alleles in the population, hybridization of the ADH probe with a related sequence in the genome, or multiple copies of the ADH gene itself. There is genetic evidence for the presence of more than one ADH gene in some species of drosophila (D. Sullivan personal communication).

Evolutionary Divergence of the ADH Gene Studied Using Thermal Elution

In order to quantitate interspecies divergence of the ADH gene, a 610 nucleotide fragment of pSAC was end-labelled, strand-separated, and hybridized with excess genomic DNA of the seven species. The resultant duplexes were bound to hydroxyapatite and thermally eluted. As shown in Fig. 5, the 610 nucleotide fragment contains 125 nucleotides of 5' flanking sequence, 420 nucleotides of protein coding DNA and the 5' intron of the ADH gene which is 65 nucleotides in length. Although the 610 nucleotide probe contains some DNA flanking the ADH gene as well as one of the introns, the temperature at which DNA duplex elutes from hydroxyapatite is dependent upon the most homologous sequence elements. Thus this probe is suitable to measure sequence divergence of the ADH structural gene.

Fig. 4 shows the thermal elution profiles; the data are listed in Table 3. The order of species relatedness deduced from quantitation of divergence of the 610 nucleotide fragment is consistent with that generated from blotting (Table 2) and the single-copy melts (Table 1). Comparisons of Tables 1 and 3 show that this fragment is more highly conserved during evolution than total single-copy DNA. Thus comparison of its divergence is better able to resolve patterns of relatedness in the more distant species. While *D. willistoni, D. hydei* and *D. virilis* appeared equally distant from *D. melanogaster* in Table 1, the order of their relatedness (with respect to the ADH gene) is resolved in Table 3.

As a control, the separated strands of the 610 nucleotide fragment were mixed, reassociated and thermally
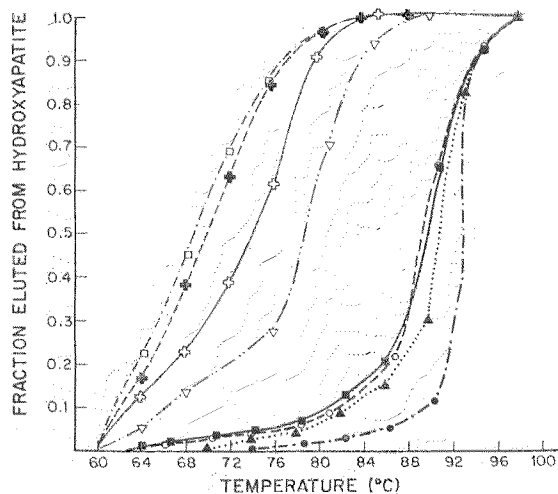

Fig. 4. Thermal elution from hydroxyapatite of duplexes formed between a 610 nucleotide fragment of pSAC and genomic DNA of seven drosophilids. The 610 nucleotide fragment (see Fig. 5) was end-labelled and strand-separated as described in the Methods section. 150 μg of driver and 5000 cpm of tracer were incubated in each reaction in 0.41 PB, 0.05% SDS, 1 mM EDTA at 65°C to Cot ≥ 2000. Thermal elution was as in Fig. 4 except that the sample was applied to the column at 60°C. Driver DNAs were (•) tracer-tracer self reaction, (▲) *D. melanogaster*, (○) *D. simulans*, (■) *D. mauritiana*, (▽) *D. pseudoobscura*, (◆) *D. hydei*, (✳) *D. willistoni*, (□) *D. virilis*

eluted from hydroxyapatite. The T_m 1/2 of the self-reassociated tracer was 92.8°C. When individual strands of the tracer were reassociated with *D. melanogaster* DNA, the T_m 1/2 of the resultant duplex was 90.8°C. The cloned ADH gene was derived from Canton S *D. melanogaster* while the driver DNA was Oregon R strain. Thus the 2°C reduction in T_m 1/2 of the driven duplex relative to the clone self-reassociation probably reflects divergence in the ADH gene between these two strains of *D. melanogaster*. The sibling species of *D. melanogaster* show a further reduction in T_m 1/2 of only 0.8°C. *D. simulans* and *D. mauritiana* are indistinguishable. The *D. pseudoobscura* ADH fragment is the next most related to *D. melanogaster* having a ΔT_m 1/2 of 11.9°C. This is consistent with the single-copy melting data of Fig. 2 and the genomic blots shown in Fig. 3. Again comparisons at the DNA level show that *D. willistoni* is more distantly related to *D. melanogaster* than *D. pseudoobscura* or even *D. hydei*. *D. virilis* is once more shown to be the most distantly related to *D. melanogaster* of the seven.

Comparison of the ADH Gene in D. melanogaster, D. simulans and D. mauritiana by Restriction Mapping

In order to initiate a more detailed analysis of the ADH gene in closely and distantly related drosophilids, pSAC was used to screen genomic λ clone libraries. The first result of these screens has been isolation of the ADH

Table 3. Evolutionary divergence of a fragment of pSAC measured by thermal elution from hydroxyapatite

Species	T_m 1/2 (°C)[a]	Interspecies[b] ΔT_m 1/2 (°C)[b]	Polymorphic ΔT_m 1/2 (°C)[c]	ΔT_m 1/2 of Single-copy DNA (C°)[d]
Tracer self-reassociation[e]	92.8°	–	–	–
D. melanogaster	90.8°	–	2.0°	–
D. simulans	90.0°	0.8°	–	1.9°
D. mauritiana	90.0°	0.8°	–	2.3°
D. pseudoobscura	79.0°	11.8°	–	18.9°
D. hydei	74.0°	16.8°	–	26.1°
D. willistoni	69.7°	21.1°	–	26.5°
D. virilis	68.7°	22.1°	–	26.2°

a The average T_m 1/2 value in double determinations; variation was $\leqslant 1°C$

b The interspecies ΔT_m 1/2 is the difference between the average T_m 1/2 observed for D. melanogaster/D. melanogaster duplex and the average T_m 1/2 observed in heterologous measurements

c The polymorphic ΔT_m 1/2 is the difference between the average T_m 1/2 of the self-reassociated tracer and the average T_m 1/2 observed when the tracer was reassociated with D. melanogaster DNA. The clone was derived from Canton S while the driver was Oregon R strain DNA

d From Table 1

e The two separated strands of the D. melanogaster 610 nucleotide fragment were reassociated with each other and the duplex thermally eluted from hydroxyapatite
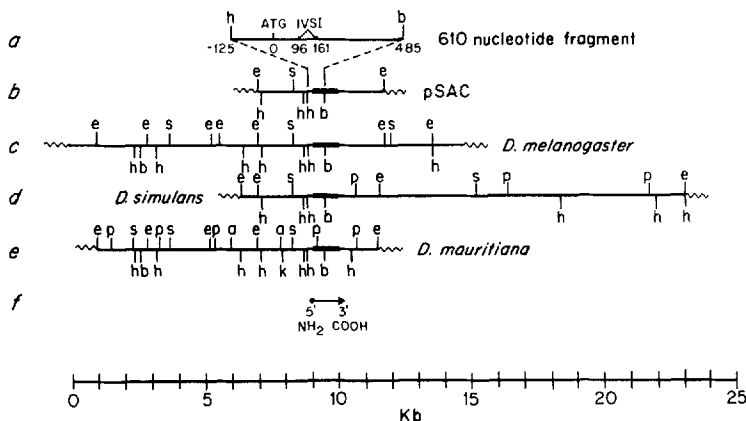

Fig. 5. Restriction maps of ADH-containing clones. *a* The 610 nucleotide fragment used in Fig. 4 and Table 3, *b* pSAC, an ADH-containing subclone from an isolated ADH-containing λ clone from D. melanogaster (Goldberg 1980), *c* an isolated ADH-containing λ clone from D. melanogaster (Goldberg 1980), *d* an isolated ADH-containing λ clone from D. simulans, *e* an isolated ADH-containing λ clone from D. mauritiana, *f* The location and direction of transcription of the ADH gene. The symbols represent (—) drosophila DNA, (∿) vector DNA, (E) Eco R1; (H) Hind III; (B) Bam HI: (S) Sal I; (P) Pst I; (A) Ava I; (K) Kpn I. Pst I, Kpn I, and Ava I sites were not available for D. melanogaster and Kpn I and Ava I sites were not determined for D. simulans

gene from D. simulans and D. mauritiana. In these species, as well as in D. melanogaster, the ADH gene resides in a 4.7 Kb Eco R1 fragment; this is shown in Fig. 3. Restriction maps of each clone were established by standard techniques and are shown in Fig. 5. The restriction map of the D. melanogaster ADH containing λ clone has been published previously (Goldberg 1980). The direction of transcription of the gene has been established by sequencing (Benyajati et al. 1981). The D. simulans and D. mauritiana maps show a number of sites in common with each other and with D. melanogaster. Comparison of these sites allows orientation of the gene whose 5' end resides very close to a Hind III site common to all three species. In addition to the common Hind III site, D. melanogaster, D. simulans and D. mauritiana also share two Eco R1 sites, a Sal I site and two additional Hind III sites flanking the ADH gene. There is also one conserved Bam site within the gene. Pst I sites have not been mapped in D. melanogaster but there is one site shared between D. simulans and D.

mauritiana as well as one which is not shared. There is a high degree of restriction site sharing within and directly flanking the ADH gene in the three sibling species. Although the number of sites compared does not allow quantitation of sequence divergence, this provides independent evidence of the similarity of these genomes, and suggests that this gene and its immediate sequence environment may represent a region of conserved DNA sequence.

Discussion

Use of Interspecies Comparison at the DNA Level to Assess Species Relatedness

Four different techniques have been used to assess the relatedness of seven drosophilids at the DNA level. These were thermal elution of interspecies single-copy DNA hybrids, genomic blotting using the cloned ADH gene

as a probe, thermal elution of interspecies hybrids of ADH, and in a limited set of measurements, restriction mapping of isolated ADH-containing genomic clones. Phylogenetic conclusions based on divergence of a single gene are subject to uncertainty. However, based on the measurements reported here, a consistent ordering of the relatedness of these species to *D. melanogaster* was obtained and can be compared to that predicted by the phylogeny in Fig. 1. An unexpected finding was that *D. willistoni*, thought to have diverged from the line leading to *D. melanogaster* at about the same time as *D. pseudoobscura* (Throckmorton 1975), was found by several lines of evidence to be much more distantly related to *D. melanogaster*. Beverley and Wilson (1982) have recently come to a similar conclusion based on immunological evidence. With respect to ADH, Table 3 shows *D. willistoni* to be more distant even than *D. hydei* and this is supported by the blot of Fig. 3. There are two possible interpretations of this result. One is that the phylogeny is simply incorrect regarding placement of *D. willistoni*. A second is that after divergence from the line leading to *D. melanogaster*, the *D. willistoni* line diverged at a much more rapid rate than the line leading to *D. pseudoobscura*. While less striking than the divergence of *D. willistoni*, the phylogeny shown in Fig. 1 did not predict that *D. hydei* would appear to be more closely related to *D. melanogaster* than would *D. virilis*.

Rate of Single-Copy DNA Divergence

Because of the lack of fossil evidence, it is extremely difficult to estimate divergence times for drosophilids. The best available estimate of divergence time between *D. melanogaster* and *D. virilis* is about forty million years (Throckmorton, personal communication). Referring to data of Table 1, and assuming 1°C change in T_m 1/2 corresponds to 1% sequence mismatch (Britten et al.

1974) we can calculate a change of 26.2/40 = 0.66% of bases per million years for the 18.6% of single-copy DNA which is still similar enough to form interspecific duplex. Since 81.4% of the *D. virilis* genome is too divergent to hybridize with *D. melanogaster* at all, the estimate of 0.66% per million years is clearly a lower limit. Calculation of this average rate does not necessarily imply a constant rate over time. The minimum rate of divergence of 0.66% per million years for single-copy DNA in drosophila is similar to the rate calculated for sea urchin DNA of 0.5% per million years (Hall et al. 1980). In Hawaiian drosophila, a rate of 0.4% per million years has been reported (Hunt et al. 1981). A slower rate, 0.1% per million years, has been reported for primate DNA's (Britten and Davidson 1976).

The Nature of Single-Copy DNA Divergence

There are some clear differences between the nature of single-copy DNA divergence observed here and that reported in other interspecies comparisons. For example, Angerer et al. (1976) compared two distantly related species of sea urchin i.e. *S. purpuratus* and *L. pictus*. Only about 12% of the single-copy DNA was cross-hybridizable. However, the T_m 1/2 of that 12% was relatively high, demonstrating the existence of a conserved class of single-copy DNA. In contrast, the distantly related drosophilids compared here do not contain an analogous conserved component in their genomes. An interesting observation of this study is that the genome of *D. melanogaster* appears to contain a large, rapidly evolving, single-copy component. This follows from the finding that about 35% of *D. melanogaster* single-copy fails to hybridize with *D. simulans* or *D. mauritiana* total DNA (Table 1). A similar observation has recently been reported in comparisons of closely related species of Hawaiian drosophila (Hunt et al. 1981). That these observations in drosophila are unusual can be seen from

Table. 4. Melting temperature and percent duplex formation in selected interspecies comparisons

Species compared	Single-copy ΔT_m 1/2 (°C)	%Reaction	Source
S. purpuratus/S. franciscanus	14°	90%	Hall et al. (1980)
Mouse/Rat	18°	80–85%	Kohne et al. (1972) Rosbash et al. (1975)
Man/Capuchin	10.5°	90–95%	Kohne et al. (1972)
Cow/Pig	14°	80%	Kohne et al. (1972)
Individual *S. purpuratus/* Second individual *S. purpuratus*	4°	95%	Britten et al. (1978)
D. picticornis/ or *D. heteroneura* *D. silvestris*	2.1°	62%	Hunt et al. (1981)
D. melanogaster/ or *D. simulans* *D. mauritiana*	2.1°	65%	Average of data from Table 1 of this paper

data listed in Table 4. The first four comparisons show instances where evolutionary sequence divergence has lowered the melting point of interspecific duplexes by a large amount i.e. 10–18°C. However, most of the single-copy DNA of these genomes remains cross-hybridizable (80–95%). In fact, for a large number of published comparisons (of which those in Table 4 are only a fraction) it is generally true that the extent of inter-species hybridization remains relatively high even when sequence divergence reaches a level of 10–15%. In very distantly related species, of course, cross-hybridization is low. In very closely related organisms, such as the comparison of two individual *S. purpuratus* cited in Table 4, cross-hybridization is essentially complete and ΔT_m 1/2 is low.

The last two citations of Table 4 show comparisons of drosophilids known by several criteria to be closely related. As expected, the ΔT_m 1/2 is small. Surprisingly however, a significant fraction of the single-copy DNA fails to cross-hybridize at all. It appears that drosophilids are undergoing two completely different evolutionary processes. One is the normal accumulation of mutations common to DNA of all species. A second is a rapid process which results in lack of hybridization between a large fraction of the genome of closely related species. This second process is not seen in the other species listed in Table 4 and may be unique to drosophila. The recent discovery of mobile elements in drosophila genomes (e.g. Strobel et al. 1979) suggests that processes of insertion and deletion of such elements might explain the results obtained here. Such mobile elements have not been previously reported in single-copy DNA.

D. melanogaster and *D. simulans* have similar polytene chromosome banding patterns and interspecific hybrids exhibit almost normal chromosomal pairing (Patterson and Stone 1952). Thus it is unlikely that rapidly evolving DNA is all located in one or a few regions of the genome. If it were, these regions would be evident as areas where homologous chromosomes could not pair in interspecific hybrids. Therefore it seems likely that rapidly evolving DNA is interspersed throughout the genome with DNA undergoing the more usual processes of evolution. On the other hand, rapidly evolving DNA probably is not interspersed with a very short periodicity. The tracer used in the measurements listed in Table 1 had a length of about 500 nucleotides. If interspersion periodicity were much shorter than 500 nucleotides, most tracer molecules would contain both rapidly evolving and "normal" DNA. These molecules would form duplexes in the interspecific hybridization experiments and a larger cross-hybridization would have been observed. Thus the rapidly evolving DNA is neither localized in one or a few positions, nor interspersed at short periodicity. It must, therefore, be interspersed in many genomic locations but with an average spacing considerably greater than 500 nucleotides. Further

characterization of rapidly evolving DNA must await cloning and detailed study of individual sequences.

References

Anderson DM, Sheller RH, Posakony JW, McAllister LB, Trabert SG, Beall C, Britten RJ, Davidson EH (1981) Repetitive sequences in the sea urchin genome. I. Distribution of members of specific repetitive families. J Mol Biol 145:5–28

Angerer RC, Davidson EH, Britten RJ (1976) Single copy DNA and structural gene sequence relationships among four sea urchin species. Chromosoma 56:213–226

Benton WD, Davis RW (1977) Screening λ recombinant clones by hybridization to single plaques in situ. Science 196: 180–182

Benyajati C, Wang N, Reddy A, Weinberg E, Sofer W (1980) Alcohol dehydrogenase in *Drosophila*: Isolation and characterization of messenger RNA and cDNA clone. Nucleic Acids Res 8:5649–5665

Benyajati C, Place AR, Powers DA, Sofer W (1981) Alcohol dehydrogenase gene of *Drosophila melanogaster*: Relationship of intervening sequences to functional domains in the protein. Proc Natl Acad Sci USA 78:2717–2721

Beverley SM, Wilson AC (1982) Molecular evolution in *Drosophila* and higher diptera I: Micro-complement fixation studies of a larval hemolymph protein. J Mol Evol 18: 251–264

Blattner FR, Williams BG, Blechl AE, Denniston-Thompson K, Faber HE, Furlong LA, Grunwald DJ, Keifer DO, Moore DD, Schumm JW, Sheldon EL, Smithies O (1977) Charon phages: Safer derivatiues of bacteriophage lambda for DNA cloning. Science 196:161–169

Britten RJ, Kohne DE (1967) Nucleotide sequence repetition in DNA. Carnegie Inst Wash Yearbook 65:98–117

Britten RJ, Graham DE, Neufeld BR (1974) Analysis of repeating DNA sequences by reassociation. In: Grossman L, Moldave K (eds) Methods in Enzymology 29 E, Academic Press, New York, p 363

Britten RJ, Davidson EH (1976) DNA sequence arangement and preliminary evidence on its evolution. Fed Proc 35:2151–2157

Britten RJ, Cetta A, Davidson EH (1978) The single-copy DNA sequence polymorphism of the sea urchin *Strongylocentrotus purpuratus*. Cell 15:1175–1186

Clewell DB, Helinski DR (1972) Effect of growth conditions on the formation of the relaxation complex of supercoiled Col E1 desoxyribonucleic acid and protein in *Escherichia coli.* J Bacteriol 110:1135

Drouin J, Symons RJ (1979) In: Cummings D, Borst P, David I, Weisman S, Fox G (eds) Extra Chromosomal DNA, ICN-UCLA Symposia vol 15 Academic Press, New York, pp 471–484

Efstratiadis A, Posakony JW, Maniatis T, Lawn RM, O'Connell C, Spritz RA, DeRiel JK, Forget BG, Weissman SM, Slighton JL, Blechel AF, Smithies O, Baralle FE, Shoulders CC, Proudfoot NJ (1980) The structure and evolution of the human β-globin gene family. Cell 21:653–668

Galau GA, Britten RJ, Davidson EH (1974) A measurement of the sequence complexity of polysomal messenger RNA in sea urchin embryos. Cell 2:9–21

Goldberg DA (1980) Isolation and partial characterization of the *Drosophila* alcohol dehydrogenase gene. Proc Natl Acad Sci USA 77:5794–5798

Hall TJ, Grula JW, Davidson EH, Britten RJ (1980) Evolution of the sea urchin non-repetitive DNA. J Mol Evol 16:95–110

Hunt JA, Hall TJ, Britten RJ (1981) Evolutionary distances in Hawaiian *Drosophila* measured by DNA reassociation. J Mol Evol 361–367

Hutton JR, Wetmur JG (1973) Effect of chemical modification on the rate of resaturation of deoxyribonucleic acid: Deaminated and glyoxalated deoxyribonucleic acid. Biochemistry 12:558–563

Klenow H, Overgaard-Hansen OK (1976) The N-terminal amino acid sequences of DNA polymerase I from *Escherichia coli* and of the large and small fragments obtained by a limited proteolysis. Eur J Biochem 45:623–627

Kohne DE, Chiscon JA, Hoyer BH (1971) Carnegie Inst Wash Yearbook 69:488–502

Laird CD, McCarthy BJ (1968) Magnitude of interspecific nucleotide sequence variablity in *Drosophila*. Genetics 60:303–322

Maniatis T, Hardison RC, Lacy E, Laver J, O'Connell C, Quon D, Sim GK, Efstratiadis A (1978) The isolation of structural genes from libraries of eucaryotic DNA. Cell 15:687–701

Maniatis T, Fritsch EF, Sambrock J (1981) In: Molecular Cloning Cold Spring Harbor publishers.

Maxam A, Gilbert W (1977) A new method for sequencing DNA. Proc Natl Acad Sci USA 77:5791–5798

McConaughy BL, Laird CD, McCarthy BJ (1969) Nucleic acid reassociation in formamide. Biochemistry 8:3289–3295

Moore GP, Scheller RH, Davidson EH, Britten RJ (1978) Evolutionary change in the repetition frequency of sea uchin DNA sequences. Cell 15:649–660

Noll H (1967) Characterization of macromolecules by constant velocity sedimentation. Nature 215:360–363

Patterson JT, Stone WS (1952) In: Evolution in the Genus *Drosophila*. McMillan, New York

Perler F, Efstratiadis A, Lomedico P, Gilbert W, Kolodner R, Dodgson J (1980) The evolution of genes: the chicken preproinsulin gene. Cell 20:555–556

Rigby PW, Dieckman M, Rhodes C, Berg P (1977) Labeling deoxyribonucleic acid to high specific activity in vitro by nick translation with DNA polymerase I. J Mol Biol 113:237–251

Rosbash M, Campo MS, Gummerson KS (1975) Conservation of cytoplasmic poly(A)-containing RNA in mouse and rat. Nature 258:682–686

Schidkraut C, Lifson N (1965) Dependence of the melting temperature of DNA on salt concentration. Biopolymeres 3:195–208

Southern EM (1975) Detection of specific sequences among DNA fragments seperated by gel electrophoresis. J Mol Biol 98:503–517

Sperlick D (1962) Dros Info Ser 36:118

Strobel E, Dunsmuir P, Rubin G (1979) Polymorphisms in the chromosomal locations of elements of the 412, Copia and 297 dispersed repeated gene family in Drosophila. Cell 17:429–439

Studier FW (1965) Sedimentation studies of the size and shape of DNA. J Mol Biol 11:373–390

Thatcher D (1980) The complete amino acid sequence of 3 alcohol dehydrogenase alloenzymes. Biochem J 187:875–886

Throckmorton LH (1975) In: King RC (ed) Hanbook of Genetics Volume 7 Plenum Press, New York, p 421–469

Wilson AC, Carlson SS, White TJ (1977) Biochemical evolution. Annu Rev Biochem 46:573–639