J Mol Evol (1984) 20:251–264

Journal of Molecular Evolution

© Springer-Verlag 1984

Adjacent Chromosomal Regions Can Evolve at Very Different Rates: Evolution of the *Drosophila* 68C Glue Gene Cluster

Elliot M. Meyerowitz and Christopher H. Martin

Division of Biology 156-29, California Institute of Technology, Pasadena, California 91125, USA

Summary. The 68C puff is a highly transcribed region of the *Drosophila melanogaster* salivary gland polytene chromosomes. Three different classes of messenger RNA originate in a 5000-bp region in the puff; each class is translated to one of the salivary gland glue proteins sgs-3, sgs-7, or sgs-8. These messenger RNA classes are coordinately controlled, with each RNA appearing in the third larval instar and disappearing at the time of puparium formation. Their disappearance is initiated by the action of the steroid hormone ecdysterone. In the work reported here, we studied evolution of this hormone-regulated gene cluster in the *melanogaster* species subgroup of *Drosophila.* Genome blot hybridization experiments showed that five other species of this subgroup have DNA sequences that hybridize to *D. melanogaster* 68C sequences, and that these sequences are divided into a highly conserved region, which does not contain the glue genes, and an extraordinarily diverged region, which does. Molecular cloning of this DNA from *D. simulans, D. erecta, D. yakuba,* and *D. teissieri* confirmed the division of the region into a slowly and a rapidly evolving portion, and also showed that the rapidly evolving region of each species codes for third instar larval salivary gland RNAs homologous to the *D. melanogaster* glue mRNAs. The highly conserved region is at least 13,000 bp long, and is not known to code for any RNAs.

Key words: *Drosophila* — Genome evolution — 68C Glue gene cluster — *Drosophila melanogaster* species subgroup

Offprint requests to: E.M. Meyerowitz

Introduction

Puffs are regions of polytene chromosomes that are actively undergoing transcription (Pelling 1964). One of the largest puffs found in the *Drosophila melanogaster* salivary gland polytene chromosomes is the 68C puff, on the left arm of the third chromosome. This puff is present through much of the third larval instar, regressing several hours before the time of puparium formation (Ashburner 1967). The regression of the 68C puff is a direct result of an increase in the level of the steroid hormone ecdysterone in the larval hemolymph (Ashburner 1973, 1974; Ashburner and Richards 1976). Several lines of evidence indicate that 68C puff regression results from the binding of ecdysterone, itself bound to a protein receptor, to the DNA of the 68C region (Gronemeyer and Pongs 1980; Dworniczak et al. 1983).

Molecular cloning of 68C puff DNA followed by DNA, RNA, and protein sequence analysis has shown that the puff contains DNA sequences that are transcribed to produce three different polyadenylated messenger RNAs whose accumulation is under ecdysterone control (Meyerowitz and Hogness 1982; Crowley et al. 1983; Garfinkel et al. 1983; Crowley and Meyerowitz 1984). Each of these RNAs is translated in the salivary gland, and each of the resulting polypeptides is one of the salivary gland secretion proteins. This is a group of at least seven polypeptides that are produced in salivary gland cells during the third larval instar; they are secreted into the lumen of the salivary gland near the end of this developmental stage, and are expelled from the luman through the salivary gland duct to the larval substrate at the end of the third instar. The secretion then hardens to form a strong glue that binds the newly formed puparial case to a solid surface (Fraen-

kel and Brookes 1953). The three salivary gland secretion proteins coded for in the 68C puff are sgs-3, sgs-7, and sgs-8 (Crowley et al. 1983). These three proteins are related in their amino acid sequences and thus form a small, clustered gene family. The differences among the proteins have arisen both from single-nucleotide substitutions and from the appearance in one of the proteins (sgs-3) of a module of 234 amino acids not present in the other two (Garfinkel et al. 1983).

There are two major reasons for examining the DNA sequence topography of the 68C puff in species of *Drosophila* other than *D. melanogaster*. One is to understand the evolution of the three members of this diverged gene family, in particular to investigate the mechanism of modular evolution that gave rise to sgs-3. The other is to find those elements of the puff DNA that are conserved in evolution, in the expectation that the regions of sequence that are relatively conserved in related species of *Drosophila* will include those that interact with regulatory proteins coded elsewhere in the genome. Identification of conserved sequences would thus be a step toward understanding the relation of the 68C DNA sequences to the regulated expression of the 68C RNAs.

In the experiments reported here we have begun the analysis of the 68C glue gene cluster in the group of closely related *Drosophila* species making up the *melanogaster* species subgroup. By interspecies DNA hybridization and molecular cloning, we show that five of these species contain regions homologous to the 68C region of *D. melanogaster* that code for abundant salivary gland RNAs related to the *D. melanogaster* 68C glue messengers. We also discover that the 68C-homologous region is divided into adjacent blocks of sequence that evolve at very different rates, with the gene cluster found in a region that is evolving with extraordinary rapidity. Our results show that the rate of sequence evolution is a local property of chromosomal regions, and also serve as a first step toward understanding the relation of DNA sequence structure and its evolution to the regulated expression of the 68C glue gene cluster.

Materials and Methods

Materials. Restriction endonucleases were purchased from New England Biolabs. The large proteolytic fragment of *Escherichia coli* DNA polymerase I was from either New England Biolabs or New England Nuclear, and T4 DNA polymerase was purchased from New England Nuclear. T4 DNA ligase and *E. coli* DNA polymerase I were gifts of Dr. S. Scherer. ^{32}P-Labeled nucleoside triphosphates were from either Amersham or ICN. Avian Myeloblastosis Virus reverse transcriptase was a gift of G. Duyk. Nitrocellulose was purchased from Schleicher & Schuell.

The *Drosophila melanogaster* strains used were the homozygous third chromosome strain OR16f (Meyerowitz and Hogness 1982) and the Canton-S wild-type from the California Institute of Technology stock collection. The other fly species—*D. mauritiana, D. simulans* (*jv st pe*), *D. erecta, D. yakuba,* and *D. teissieri*—were from the California Institute of Technology *Drosophila* stock collection. Flies were cultured on standard food (Lewis 1960) at 18° or 22°C.

Nucleic Acid Preparations. Drosophila DNA was extracted from adult flies by freezing the flies in liquid N_2 and then powdering them in a mortar. The powder from 0.1–2 g of flies was added to 2.5 ml 0.2 M Tris-HCl, pH 8.0, 0.2 M ethylenediaminetetraacetate (EDTA), 1% sodium N-lauroyl sarcosine, 100 μg/ml proteinase K (Merck). This was incubated with gentle shaking at 48°C for 1 h, and then centrifuged at 10,000 rpm for 5 min in a Sorvall SS-34 rotor. The supernatant was brought to 4.0 ml with 10 mM Tris-HCl, pH 8.0, 1 mM EDTA; then 3.7 g CsCl and 0.4 ml 10 mg/ml ethidium bromide were added. This mixture was centrifuged at 53,000 rpm for 20 h in a Beckman VTi65 rotor, and then the ultraviolet-fluorescent band was removed with a syringe and gently butanol extracted four times to separate the DNA from the ethidium bromide. Following this, the DNA was precipitated by addition of 2 volumes of ethanol followed by gentle hand centrifugation. The DNA pellet was washed with 70% ethanol, air dried, and resuspended overnight at 4°C without agitation. The resulting DNA was pure and over 150,000 bp in fragment length.

Plasmid and bacteriophage DNA preparations were performed as described by Davis et al. (1980), with occasional modifications that did not affect the results. *Drosophila* salivary gland RNA was prepared from hand-dissected salivary glands by two phenol–chloroform extractions followed by two chloroform extractions and ethanol precipitation. Polyadenylated RNA was separated on an oligo(dT) cellulose (Collaborative Research) column as described by Maniatis et al. (1982).

Nucleic Acid Labeling. Nick translations followed the method of Rigby et al. (1977). 3′-End labeling of restriction fragments using T4 DNA polymerase was done as described by Maniatis et al. (1982). ^{32}P-Labeled cDNA was made from poly(A)$^+$ RNA hybridized to an oligo(dT) (Collaborative Research) primer in the presence of 100 μg/ml Actinomycin-D, in a modification of the reaction described by Lis et al. (1978).

Nuclease Digestions. Restriction endonuclease digestions of DNA were done as described by Davis et al. (1980).

Recombinant DNA. All *D. melanogaster* genomic libraries used are described by Meyerowitz and Hogness (1982). The genomic libraries from *D. simulans, D. erecta, D. yakuba,* and *D. teissieri* were produced by performing partial EcoRI digestion on high-molecular-weight adult DNA (see above) to give DNA with a mean size of 15,000–20,000 bp (15–20 kb). This DNA was subjected to sedimentation in a 10% to 40% sucrose gradient in 0.2 M sodium acetate, 10 mM Tris, 10 mM EDTA, pH 7.6, at 35,000 rpm in a SW41 swinging bucket rotor at 4°C for 15–20 h, and DNA in the size range 15–20 kb was isolated. This DNA was ligated to purified EcoRI arms of the vector λSep6 (Meyerowitz and Hogness 1982) using T4 DNA ligase (see Davis et al. 1980). The ligated DNA was treated with a λ in vitro packaging extract prepared using the *E. coli* strains NS428 and NS433 (Sternberg et al. 1977) and a modification of the procedure of Collins and Hohn (1978). The phage particles were plated on L agar in L soft agar with K802 cells (see Davis et al. 1980), without any amplification step, and screened as described below. Positive plaques were single-plaque purified twice more before proceeding. Restriction fragments of these λ clones were subcloned in plasmid or λ vectors by standard methods (Davis et al. 1980; Maniatis et al. 1982).

The system of recombinant clone nomenclature, which originated in the laboratory of D.S. Hogness, is as follows. All λ clones are prefixed with the letter λ, followed by a letter indicating the λ vector used: a for λ647 (Murray et al. 1977), b for λSep6, c for Charon 4 (Blattner et al. 1977), or e for λgt 10 (R. Davis, personal communication). Following this are two letters: Dm for *Drosophila melanogaster,* Ds for *D. simulans,* De for *D. erecta,* Dy for *D. yakuba,* or Dt for *D. teissieri.* Last is a number identifying the specific clone. Plasmid subclones begin with a single letter: a for pBR322 (Bolivar et al. 1977), f for pBR325 (Bolivar 1978), or q for DOA-1, a kanamycin-resistant, high-copy-number plasmid with multiple cloning sites (R.E. Pruitt and E.M. Meyerowitz, unpublished). The rest of the clone designation is as for the λ clones.

Gel Electrophoresis. Double-stranded DNA was subjected to electrophoresis in agarose gels cast and run in Tris–borate–EDTA buffer (Peacock and Dingman 1968). DNA was strand separated on gels as described by McDonell et al. (1977) and Meyerowitz and Hogness (1982). RNA was subjected to electrophoresis in horizontal agarose gels buffered with 40 mM sodium 3-(N-morpholino)propanesulfonate, pH 7.0, 5 mM sodium acetate, 1 mM EDTA, and containing 6% formaldehyde. Running buffer was the same as the gel buffer, but without formaldehyde. Prior to electrophoresis RNA was treated at 55°C for 15 min in 50% formamide, 6% formaldehyde, and RNA gel running buffer.

Size standards on Tris–borate–EDTA gels were restriction fragments of bacteriophage λ DNA. RNA gel size standards were single-stranded DNA prepared from HinfI-digested pBR322 and formaldehyde treated as was RNA, except that the temperature was 70°C.

Filter Binding and Hybridization of Nucleic Acids. DNA gels were denatured and neutralized as described by Southern (1975), and were transferred to nitrocellulose by the Southern procedure, using 20× SSPE (180 mM NaCl, 10 mM Na H_2PO_4, 8 mM NaOH, 1 mM Na_2 EDTA pH 7.0, Davis et al. 1980). *Drosophila* genome blots had 1.5 μg restricted DNA per lane. RNA gels were either rinsed in water; or rinsed and then treated for 45 min with 50 mM NaOH, 100 mM NaCl then neutralized in 100 mM Tris–HCl, pH 7.5, before blotting. The blot procedure was as for DNA gels. RNA extracted from three salivary gland lobes was used in each lane. Plaque and colony filters were prepared as described by Davis et al. (1980).

All hybridizations were in 50% formamide, 5× SSPE, 100 μg/ml sonicated and boiled salmon testis DNA (Sigma), 1× Denhardt's solution (0.02% Ficoll, 0.02% polyvinylpyrrolidone, 0.02% bovine serum albumin; Denhardt 1966), 0.1% sodium dodecyl sulfate (SDS), at 43° C. After hybridization, filters were washed in 1× SSPE, 0.1% SDS, at room temperature, unless otherwise indicated.

DNA Sequencing. DNA sequencing was performed by the modifications of the Maxam and Gilbert chemical method (1980) described by Garfinkel et al. (1983).

Results

Gel Blot Experiments

The first step in our analysis of evolution of the 68C puff was the demonstration that species of *Drosophila* other than *D. melanogaster* contain DNA sequences homologous to the 68C genes. The species chosen were those of the *melanogaster* species subgroup, which contains *D. melanogaster* and at least six other species: *D. simulans, D. mauritiana, D. yakuba, D. teissieri, D. erecta,* and *D. orena.* These are the *Drosophila* species most closely related to *D. melanogaster* in morphology, mitotic chromosome karyotype, and polytene chromosome banding pattern (Bock and Wheeler 1972; Lemeunier and Ashburner 1976; Lemeunier et al. 1978). DNAs from adults of all of these species except *D. orena* were purified and digested to completion with the restriction endonuclease EcoRI, and the resulting restriction fragments were separated by electrophoresis through an agarose gel. The gel pattern was transferred to a nitrocellulose filter and this genome blot filter was hybridized with a series of ^{32}P-labeled cloned 68C probes from *D. melanogaster.* Autoradiography revealed the presence and size of EcoRI fragments from each of the species that were homologous to the *D. melanogaster* probes. The intensity of hybridization and persistence of the signal through successively more stringent filter washes gave an estimate of the sequence divergence between the *D. melanogaster* probes and the homologous sequences from the other species.

The cloned *D. melanogaster* sequences used are depicted in Fig. 1. The first probe was λaDm1501-10, a genomic clone containing all or a substantial part of *D. melanogaster* EcoRI fragments 3.8, 4.7, 3.7, 2.6, and 7.0 kb long. The control *D. melanogaster* lane on the filter showed strong autoradiographic signals resulting from hybridization to bands of these sizes. The lanes containing DNAs from the other species showed hybridization to a smaller number of bands. For *D. simulans,* bands of 4.7, 3.8, and 3.7 kb showed strong signals, and weak hybridization was seen to EcoRI fragments of 2.6 and 1.4 kb. *D. mauritiana* also showed strong signals on bands of 4.7, 3.8, and 3.7 kb. *D. erecta* showed hybridization to bands of 4.3, 4.0, and 3.7 kb; *D. yakuba* to 8.4- and 4.2-kb fragments; and *D. teissieri* to bands at 4.7, 3.8, and 3.7 kb. Since the *D. melanogaster* DNA in λaDm1501-10 extends for almost 18 kb, it can be seen that only a fraction of the λaDm1501-10 insert hybridizes strongly to the genomic DNAs of the other species. This indicates that only part of the *D. melanogaster* probe has a high degree of similarity to sequences in the other species, with another part of the probe hybridizing weakly.

To determine which regions of the λaDm1501-10 insert were responsible for the strong cross-hybridization, several additional *D. melanogaster* clones were used as probes. λbDm2031 contains *D. melanogaster* DNA repesenting the leftward part of the DNA cloned into λaDm1501-10 (see Fig. 1). The *D. melanogaster* EcoRI fragments represented in λbDm2031 are 3.8, 4.7, and 3.7 kb long, the same size as the EcoRI fragments of *D. simulans, D.*

mauritiana, and *D. teissieri* that were strongly labeled with the λaDm1501-10 probe. A species genome blot filter hybridized with labeled λbDm2031 DNA gave the expected *D. melanogaster* pattern, and in the lanes containing DNA from the other species gave the same pattern of strongly labeled bands as that obtained using λaDm1501-10 as the probe. Thus, the *D. melanogaster* sequences with a high degree of similarity to sequences from the other species are those represented in λbDm2031, and not those that contain the salivary gland secretion protein genes.

This conclusion was tested by using λcDm2021, which includes all of the sequences found in λbDm2031 as well as several kilobases of additional DNA (Fig. 1). This clone hybridized strongly with the same bands as λbDm2031 did, and with smaller, additional bands that presumably were hybridized by those *D. melanogaster* sequences present in λcDm2021 and not in λbDm2031. Both the λbDm2031 hybridized filter and the λcDm2021 filter were washed at successively higher stringencies after the initial autoradiographic exposure. The first washes were in 0.01× SSPE at 47°C. On both filters, the hybridization patterns and intensities were unchanged by this treatment. After the first wash and autoradiographic exposure was a second wash of each filter in 0.01× SSPE at 52°C. This caused a uniform reduction in signal in all bands on both filters, with the signal change in the control *D. melanogaster* lanes paralleling that in the lanes representing the other species. Thus, the strongly hybridized regions of the 68C sequences in all of the species are not detectably diverged in DNA sequence, if one uses the melting temperature of filter-bound DNA duplexes as a crude divergence assay.

One further *D. melanogaster* probe from the conserved region of the 68C clones was used, aDm2003. This plasmid clone contains most of the rightmost fragment of the conserved region of *D. melanogaster,* the 3.7-kb EcoRI fragment, and much of a 2.6 kb EcoRI fragment, which is the leftmost fragment of the less-conserved RNA coding region. When hybridized to a species genome blot filter, the aDm2003 probe gave a strong signal at 3.7 kb and a weak signal at 2.6 kb in the *D. simulans* lane, a strong signal at 3.7 kb and a weak one at 4.0 kb in the *D. mauritiana* lane, strong hybridization to a 3.7-kb band and weak hybridization to a 4.2-kb fragment in the *D. erecta* track, and strong hybridization to a 3.7-kb *D. teissieri* band. Thus, not only is the 3.7-kb EcoRI band highly conserved by duplex melting criteria, but all of the species except *D. yakuba,* which was not tested in this experiment, appear to have the same EcoRI sites, spaced equally, surrounding this DNA. This indicates a very high degree of sequence conservation. That the remaining signal, from the 2.6-kb EcoRI fragment that includes one of the glue genes, was weak confirms that the boundary between conserved and less conserved sequences is approximately at the EcoRI site separating the 3.7-kb fragment from the adjacent EcoRI fragment to the right.

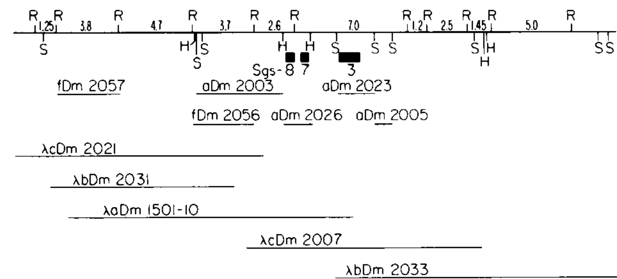

Fig. 1. *Drosophila melanogaster* cloned sequences used. The cloned *D. melanogaster* sequences used in this study are shown beneath a restriction endonuclease map of a portion of the 68C puff DNA of this species. The derivations of the map and clones are described in Meyerowitz and Hogness (1982) and in Garfinkel et al. (1983). The restriction endonuclease cleavage sites shown are those of EcoRI (R), HindIII (H), and SalI (S). The numbers on the restriction map are the sizes of each EcoRI fragment in kilobase pairs. The solid bars under the map show the positions and extents of the DNA coding for the glue proteins sgs-8, -7, and -3.

Several probes from the relatively unconserved glue RNA coding region of the 68C puff were also used in gel blot experiments. aDm2026 contains a 1.65-kb HindIII fragment from *D. melanogaster* that includes the *Sgs-7* and *Sgs-8* genes. When ^{32}P-labeled aDm2026 DNA was hybridized to a species genome blot filter, the hybridization to *D. melanogaster* sequences was much stronger than that to DNA of any of the other species. Each of the other species did show binding of the labeled probe to EcoRI fragments of various sizes. A series of washes of the filter in 0.01 × SSPE at temperatures of 43°, 49°, and 52.5°C, with autoradiography performed after each new wash, showed that the aDm2026-homologous sequences of *D. yakuba, D. teissieri,* and *D. erecta* lost all binding to aDm2026 DNA between 43° and 49°C, whereas the signals in the *D. mauritiana* and *D. simulans* lanes were considerably weakened after the 49°C wash, but not removed. A 52.5°C wash sufficed to remove almost all of the signal in the *D. mauritiana* and *D. simulans* lanes, while having little effect on *D. melanogaster* self-hybridization. Thus, the *D. melanogaster* sequences containing the *Sgs-7* and *Sgs-8* genes are more diverged from their homologous sequences in the other species than are the *D. melanogaster* sequences to the left of the glue genes, sequences that are not known to code for any *Drosophila* RNAs (Meyerowitz and Hogness 1982).

The next *D. melanogaster* DNA used as a labeled

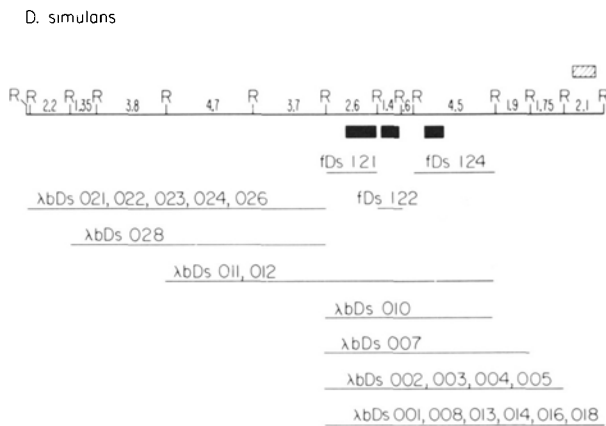

Fig. 2 Cloned sequences of *Drosophila simulans*. The original λ clones obtained from the *D. simulans* genomic library and several plasmid subclones used are shown in relation to a composite EcoRI restriction map of the cloned *D. simulans* DNA. The numbers on the map are distances between adjacent EcoRI sites in kilobase pairs. The hatched bar above the map indicates the maximum extent of the middle repetitive DNA element found in the 68C-homologous region in *D. simulans*. The solid bars below the map indicate the restriction fragments hybridized by cDNA made from abundant polyadenylated RNAs from the salivary glands of third instar larval *D. simulans*. All of the recombinant phage except λbDs007 are consistent with the composite map. λbDs007 has, in addition to the sequences shown, a 7.2-kb EcoRI fragment that is not present in any of the other clones and is not hybridized by *D. melanogaster* 68C sequences. It seems certain that this 68C-unrelated fragment was ligated to the 68C-homologous *D. simulans* DNA during the construction of the recombinant phage library, and that it derives from a random, noncontiguous region of the *D. simulans* genome

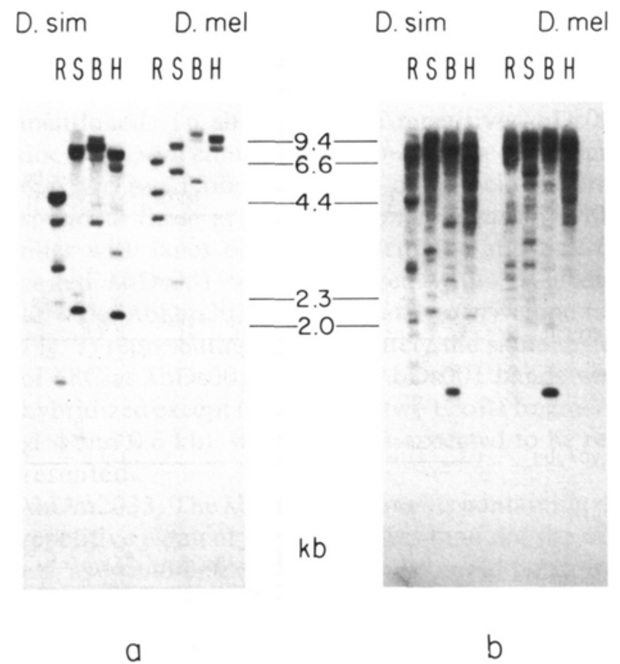

Fig. 3a,b. Hybridization of *Drosophila simulans* clones to *D. simulans* and *D. melanogaster* genomic DNA. *D. simulans* (D. sim) genomic DNA was digested with EcoRI (R), SalI (S), BamHI (B), or HindIII (H) in four separate reactions, each with 1.5 μg DNA. The digested samples were loaded in four adjacent lanes of a 0.9% agarose gel. The nearby group of lanes was loaded similarly with identical digests of *D. melanogaster* (D. mel) DNA. After electrophoresis the DNA in the gel was denatured and blotted to a nitrocellulose filter, and the filter was hybridized with a ^{32}P-labeled *D. simulans* λ clone probe and then autoradiographed. In **a** the probe was λbDs011. In **b** the signal was washed from the filter in a with boiling 0.01× SSPE, and the filter was rehybridized with λbDs001. The size standards are from λc-I857S7 DNA digested with HindIII

probe was aDm2023, a 2.4-kb SalI fragment containing the *Sgs-3* gene. Again, the *D. melanogaster* self-hybridization gave much stronger autoradiographic signals than did the hybridization of *D. melanogaster* sequences to the homologous sequences of other *Drosophila* species. A wash in 0.01× SSPE at 48°C reduced but did not eliminate the hybridization in the *D. yakuba, D. teissieri,* and *D. erecta* lanes. Thus, at least for some species, the RNA-coding region of the 68C puff is again less conserved in evolution than is the adjacent DNA. A final probing of a species genome blot was performed, using ^{32}P-labeled λcDm2007. The bands hybridized by this λ clone, which overlaps aDm2026 and aDm2023, included those hybridized by those two plasmid clones; the initial hybridization and subsequent washes at higher stringency confirmed the results obtained with those clones. Since λcDm2007 includes *D. melanogaster* sequences to the right of the RNA-coding region as well as the sequences containing the glue genes, there was hybridization to fragments not seen in the aDm2026 or aDm2023 experiments. These fragments were as intensely hybridized as *D. melanogaster* self-hybridized fragments of the same sizes, and after filter washes of either 48°C or 52.5°C in 0.01× SSPE these bands were still equal in intensity to the *D. melanogaster* bands. This indicates that the sequences to the right of the RNA-coding region are not highly diverged. No further experiments that analyzed this rightward region were performed.

Molecular Cloning

The general picture of 68C puff evolution gained from the genome blot experiments is of a highly diverged set of sequences containing the three glue genes, surrounded by highly conserved sequences that are not known to have any function in glue gene expression. To learn more about the evolution of this region and to establish a basis for DNA sequencing studies, the DNAs homologous to the *D. melanogaster* 68C sequences in *D. simulans, D. erecta, D. yakuba,* and *D. teissieri* were cloned. The first clones obtained were from *D. simulans,* the species most closely related to *D. melanogaster* (Sturtevant 1919). The strain of *D. simulans* used

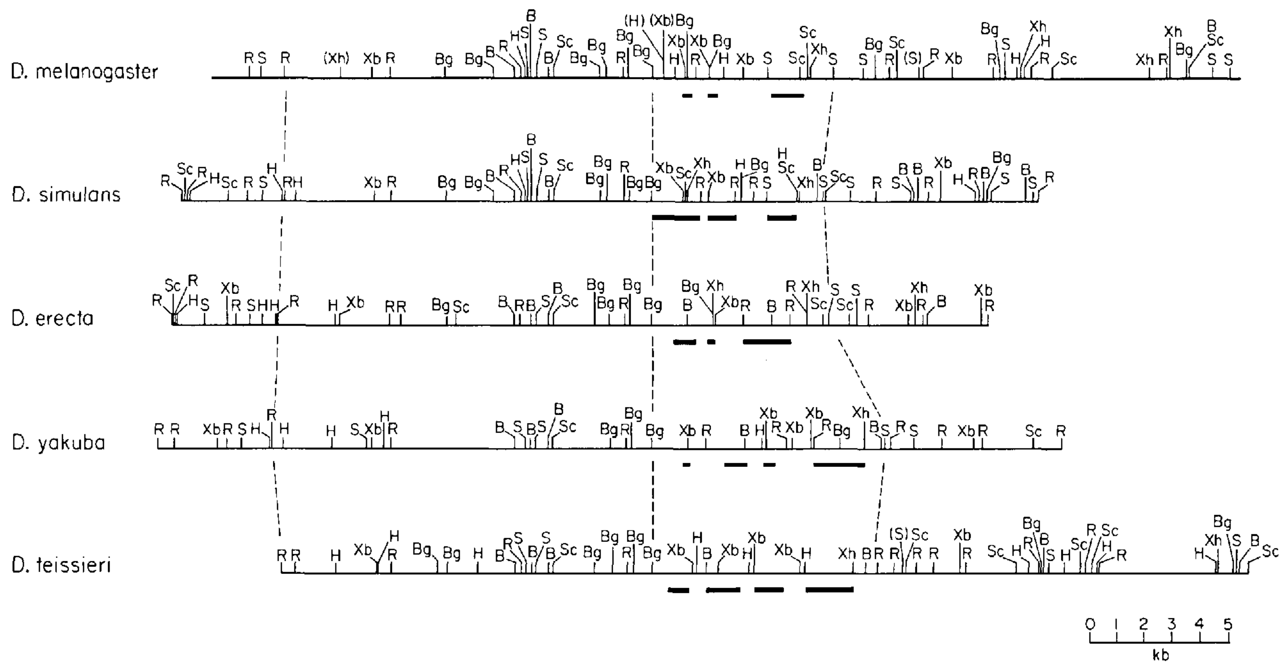

Fig. 4. Restriction maps of the cloned 68C-homologous sequences. All known BamHI (B), BglII (Bg), EcoRI (R), HindIII (H), SalI (S), SacI (Sc), XbaI (Xb), and XhoI (Xh) sites are depicted except for a single EcoRI site in *D. erecta,* which is omitted for the reason detailed in the caption to Fig. 5. Sites in parentheses on the *D. melanogaster* map are found in chromosomes of some wild-type strains, but not in others (Meyerowitz and Hogness 1982; Garfinkel et al. 1983). The parenthetical SalI site on the *D. teissieri* map is present in λbDt9100 but not in the overlapping λbDt9200. Beneath each map are solid bars showing the extents of the restriction fragments hybridized by [^{32}P]cDNA derived from third instar larval salivary gland polyadenylated RNA from each of the species (see Fig. 8 for more details). The maps are aligned by the series of common restriction endonuclease sites found to the left of the RNA-coding region. Three vertical dashed boundaries separate the maps into a leftward conserved region and a rightward RNA-coding region. The leftmost boundary is set at the EcoRI site that marks the left end of fDm2057 hybridization to the clones of each species, and is set at this point only because *D. teissieri* cloned sequences extend no farther to the left. Whether the conserved region continues beyond this point is not known, although comparison of the four species whose restriction maps do include DNA to the left of this boundary indicates that it probably does. The central boundary is the common BglII site just to the left of the RNA-coding region, this site separates conserved from diverged sequences by criteria of hybridization and restriction mapping. The rightward boundary marks the right end of the restriction fragments hybridized by aDm2023, where each species except *D. teissieri* has a SalI site. In *D. teissieri,* the boundary between hybridization of aDm2023 and aDm2005 is within the BamHI–EcoRI fragment that includes the RNA-coding-region boundary

was a homozygous third chromosome strain with the recessive third chromosome markers *jv, st,* and *pe.* DNA from adult flies was partially digested with EcoRI and 15- to 20-kb fragments were selected by sucrose gradient sedimentation and cloned into the EcoRI cloning vector λSep6 (see Materials and Methods). The resulting recombinant DNA library was not amplified, but was directly plated and screened by the plaque lift method, using the ^{32}P-labeled *D. melanogaster* clones aDm2026 and aDm2023 as probes. Fourteen independent clones were isolated that hybridized to both probes. These are λDs001, 002, 003, 004, 005, 007, 008, 010, 011, 012, 013, 014, 016, and 018. Figure 2 is a simple restriction endonuclease map of the DNA represented in these clones, showing the relation of the cloned segments. To obtain clones representing the highly conserved DNA adjacent to the RNA coding region, more recombinant phage were screened, using the *D. melanogaster* clone fDm2057 (see Fig. 1) as a labeled probe. The recombinant clones λbDs021, 022, 023, 024, 026, and 028 are all hybridized by fDm2057, and are also shown in Fig. 2.

That this collection of phage represents the sequence organization actually found in the *D. simulans* genome and does not result from the artificial joining of separate EcoRI fragments during the cloning procedure or from any other cloning artifact is shown by several facts. First, all regions of the composite restriction map, and all EcoRI junctions, were cloned more than once from a library of independent clones. In addition, the EcoRI fragments in the clones that hybridize to the *D. melanogaster* probes aDm2026 and aDm2023 are the same size as the EcoRI fragments of whole-genome *D. simulans* DNA hybridized by the same probes. Finally, two of the *D. simulans* λ clones were ^{32}P-labeled by nick translation and used as probes of genome blot filters containing *D. melanogaster* whole-genomic DNA digested with BamHI, EcoRI, HindIII, and SalI in different lanes and *D. simulans* genomic DNA similarly digested in a separate set of gel tracks. The

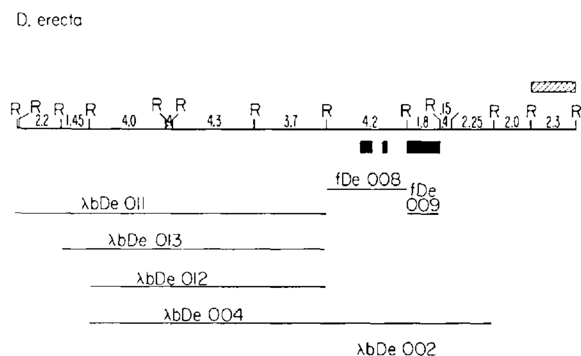

Fig. 5. Cloned sequences of *Drosophila erecta*. A composite restriction map of the cloned *D. erecta* sequences is depicted with the distances in kilobases between adjacent EcoRI (R) sites shown. The interval with the two numbers 0.15 and 0.4 contains two EcoRI fragments of these sizes; their order is unknown. Below the map are indicated the positions and extents of the λ clones and plasmid subclones used. Immediately below the map, solid bars show the location of the restriction fragments of the cloned DNA that hybridize to cDNA derived from abundant poly(A)$^+$ RNAs isolated from third instar salivary glands of *D. erecta*. Above the map, a hatched bar indicates the maximal extent of the middle repetitive DNA element found in the 68C-homologous region of *D. erecta*

first probe used was λbDs011. It hybridized to *D. melanogaster* fragments of the sizes expected from the known restriction map of the 68C region in this species. This indicates that no additional fragments from other genomic regions were incorporated into this phage during its construction, and that the *D. simulans* 68C region does not contain the breakpoints of any large inversions or translocations relative to the *D. melanogaster* sequence. λbDs011 also hybridized to the *D. simulans* restriction fragments expected from the restriction map of the phage, confirming that no large deletions or rearrangements occurred during the cloning of the *D. simulans* DNA (Fig. 3a). The second *D. simulans* λ probe used in hybridization to the genome blot filter was λbDs001. The autoradiogram resulting from this hybridization showed a multiplicity of labeled fragments in all lanes from both species, demonstrating that some part of λbDs001 contains a repetitive element present in numerous copies in both the *D. simulans* and the *D. melanogaster* genome (Fig. 3b). This element was localized to the 1.25-kb SalI–BamHI fragment internal to the 2.1-kb EcoRI fragment of λbDs001 by annealing ^{32}P-labeled, single-stranded *D. simulans* genomic DNA to a gel blot filter with lanes containing λbDs001 DNA digested with EcoRI, BamHI, and SalI. The autoradiogram of this filter showed strong labeling of the 2.1-kb EcoRI fragment, a 1.5-kb BamHI fragment, and a 1.65-kb SalI fragment; much weaker hybridization to the other *D. simulans* insert fragments; and no hybridization to phage vector DNA. Thus, the repetitive element is in the position shown in Fig. 2. Figure 4 shows a detailed restriction map of the 68C region of *D. simulans*, including the BamHI and SalI sites just mentioned. To show that the repetitive λbDs001 does represent contiguous *D. simulans* sequence and that the restriction fragments of this clone correspond to those at 68C in *D. melanogaster*, a blot filter with lanes of EcoRI-, BamHI-, and SalI-digested λbDs001 was hybridized with ^{32}P-labeled DNA of λbDm2033, a *D. melanogaster* clone (see Fig. 1) representing approximately the same region of 68C as λbDs001 does. All λbDs001 bands were hybridized except the leftmost two EcoRI fragments (1.4 and 0.6 kb), which are not expected to be represented in λbDm2033. The λbDs001 fragments containing the repetitive element hybridized less than did the others, as would be expected if a substantial portion of these DNA pieces contained sequences not present in the probe.

The next species whose 68C-homologous sequences were cloned was *D. erecta*. Production and screening of the *D. erecta* libraries was done as for *D. simulans*. The first screening used aDm2026 and aDm2023 as probes of duplicate plaque filters; two different clones hybridized by both probes were obtained. These are λbDe002 and λbDe004. A different set of clones from the *D. erecta* library was then probed with fDm2057, and three more positive plaques were obtained. These contained the phage λbDe011, λbDe012, and λDe013. A simple restriction map of the *D. erecta* sequences in these five clones and the relation of these clones to this composite map are shown in Fig. 5. A detailed restriction map of the *D. erecta* 68C-related region is in Fig. 4. Several results were obtained that show that the restriction map derived from the cloned *D. erecta* segments does correspond to the restriction map of the same sequences in the *D. erecta* genome. The first is that the restriction maps of all the overlapping regions of the λ clones are identical, thus eliminating rare cloning artifacts as a possibility in these regions. Also, hybridization of the ^{32}P-labeled *D. melanogaster* clones aDm2023 and aDm2026 to filter-bound EcoRI fragments of λbDe002 and λbDe004 showed that the sizes of the EcoRI fragments homologous to these probes are the same as the sizes of the *D. erecta* genomic EcoRI fragments hybridized by these *D. melanogaster* probes in the earlier species genome blot experiments. In addition, when λbDe004 DNA was ^{32}P-labeled and hybridized to a genome blot filter containing lanes with *D. erecta* genomic DNA digested with BamHI, EcoRI, HindIII, or SalI, the labeled bands on the filter correspond with the restriction fragment sizes expected from the clone restriction map. ^{32}P-Labeled λbDe004 DNA was also annealed to a genome blot

filter with separate lanes of *D. melanogaster* genomic DNA digested with BamHI, EcoRI, HindIII, or SalI. In each case the labeled restriction fragments were of the sizes expected from the known *D. melanogaster* 68C restriction map. Hybridizations to *D. erecta* and *D. melanogaster* genomic DNA digested with the same set of enzymes were done using ^{32}P-labeled λbDe002 as a probe as well. These showed, in addition to the expected bands, a distinct background smear in all lanes for both *D. erecta* and *D. melanogaster*. This smear, not seen in the λbDe004-hybridized genome blot filter, implies that λbDe002 contains a repetitive DNA element. Hybridization of ^{32}P-labeled *D. erecta* genomic DNA to a filter blotted from a gel containing lanes of λbDe002 DNA digested with EcoRI, XbaI, BglII, SacI, and a combination of EcoRI and each of the other enzymes showed that the repetitive element is entirely within the 2.1-kb EcoRI–XbaI fragment found in the 2.3-kb EcoRI fragment of λbDe002. Figure 5 shows the location of the repetitive element relative to the *D. erecta* restriction map.

The next DNA cloned was from *D. yakuba*. The first labeled *D. melanogaster* clone DNA used as a probe of the *D. yakuba* λ library was fDm2056; using this probe six positive plaques were obtained. These contained the *D. yakuba* genomic clones λbDy101, 102, 103, 104, 105, and 107. An additional clone, λbDy110, was subsequently obtained using aDm2023 as a probe of a separate portion of the *D. yakuba* library. Figure 6 shows the restriction map of the *D. yakuba* 68C-homologous region derived from the maps of these clones and the overlap of each λ clone insert with this map. Figure 4 shows a more detailed map of the *D. yakuba* sequences. As with the other species, the correspondence of the restriction map derived from the clones with that in the genomic DNA was shown in a variety of ways. As before, all regions of the phage clones that overlapped the same region of the composite restriction map showed identical restriction sites, eliminating the possibility of cloning artifacts in these areas. The sizes of EcoRI restriction fragments of *D. yakuba* genomic DNA hybridized by the *D. melanogaster* clone λcDm2021 in the earlier species genome blot experiments are all found in the *D. yakuba* clones that cover the left end of the composite restriction maps, and the sizes of the *D. yakuba* genomic EcoRI fragments hybridized by λcDm2007 in the species genome blots correspond to the sizes of the EcoRI fragments found in λbDy110. λbDy103, when ^{32}P-labeled and annealed to *D. yakuba* DNA that had been digested with BamHI, EcoRI, HindIII, or SalI, subjected to electrophoresis, and then transferred to nitrocellulose, showed hybridization to fragments of the same sizes as are present in the clones in all cases. When λbDy103 was used as a labeled probe

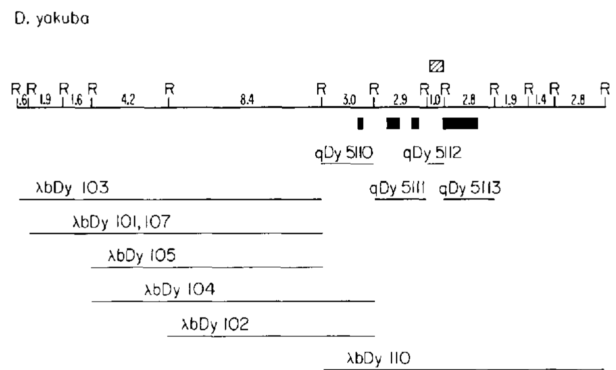

Fig. 6. Cloned sequences of *Drosophila yakuba*. A composite EcoRI (R) map of the cloned 68C-homologous DNA of *D. yakuba* is shown. Above the map, a hatched bar indicates the site of repetitive DNA; below the map, solid bars show the restriction fragments in the cloned DNA that hybridize to cDNA made from poly(A)$^+$ salivary gland RNA from *D. yakuba*. Below this, lines depict the extent of *D. yakuba* DNA in the plasmid and λ clones indicated

of *D. melanogaster* DNA digested with BamHI, EcoRI, HindIII, or SalI, the fragments hybridized were those that would be expected if the 68C regions of the two species are colinear. The results were different when λbDy110 was used as a labeled probe of genome blot filters containing restriction-endonuclease-digested *D. yakuba* and *D. melanogaster* genomic DNA samples. In this case, all *D. yakuba* lanes showed dark smears with a number of distinct bands superimposed. Thus, λbDy110 contains some DNA sequences related to sequences repeated many times in the *D. yakuba* genome. The *D. melanogaster* lanes do not show the dark smear that indicates hybridization of repetitive DNA; therefore the *D. yakuba* repetitive element is not highly repeated in the *D. melanogaster* genome. To localize the repetitive DNA in λbDy110, DNA of this λ clone was digested with EcoRI, XbaI, and a combination of both enzymes, and then subjected to electrophoresis in an agarose gel. The gel was blotted to nitrocellulose, and the resulting blot filter was hybridized with ^{32}P-labeled *D. yakuba* DNA. The λbDy110 fragments strongly hybridized were the 1.0-kb EcoRI piece and the 0.75-kb XbaI fragment wholly contained in this EcoRI fragment. The repetitive DNA was therefore shown to reside within this small XbaI fragment, in the position shown in Fig. 6.

The final λ library made used DNA from adult *D. teissieri*. A portion of these clones was screened on duplicate plaque lift filters, with both aDm2003 and aDm2023 as ^{32}P-labeled probes. One phage clone, λbDt9000, hybridized to aDm2003 and not aDm2023; one other clone, λbDt9008, hybridized to both probes. More clones were screened using aDm2005 as a labeled probe; λbDt9100 and λbDt9200 were thus obtained. A last set of λ clones

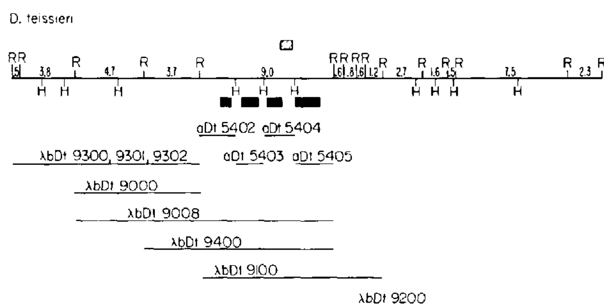

Fig. 7. Cloned sequences of *Drosophila teissieri*. A restriction endonuclease map of the cloned *D. teissieri* DNA is shown, with EcoRI (R) and HindIII (H) sites marked. The size of each EcoRI fragment in kilobase pairs is indicated. The hatched bar above the map shows the location of a repetitive DNA element; below the map are solid bars showing the location of the restriction fragments of the cloned DNA that hybridize to cDNA made from *D. teissieri* salivary gland poly(A)$^+$ RNA. Below this are lines showing the *D. teissieri* DNA represented in each of the clones used

was probed with fDm2056 and fDm2057 on duplicate plaque lift filters. λbDt9300, 9301, and 9302 were selected by both probes. The restriction maps of all of the clones were consistent with the map shown in Fig. 7, with three exceptions. One exception is the appearance of two EcoRI fragments (3.4 kb and 3.5 kb) at the right end of the insert of λbDt9000. These fragments are not hybridized by *D. melanogaster* 68C clones, and do not correlate with the fragments found in the analogous locations in λbDt9008, λbDt9400, and λbDt9100. We therefore conclude that these fragments were ligated to the 68C-homologous DNA in λbDt9000 during the λ clone construction, and do not represent the genomic DNA of the *D. teissieri* 68C-equivalent region. The second exception is the existence of an additional 0.7 kb of DNA, including a BglII site, in λbDt9100 and centered 1–2 kb to the left of the EcoRI site marking the right end of the 9.0-kb EcoRI fragment. This additional DNA is not present in λbDt9008 or λbDt9400, and may represent a polymorphism found in the population of *D. teissieri* flies from which the DNA was obtained. Finally, there is a single SalI site present in λbDt9100 but absent in the overlapping region of λbDt9200 (see Fig. 4). The usual battery of tests to determine if the restriction map derived from the clones was the same as that in the *D. teissieri* genome was applied: λbDt9200, λbDt9300, and λbDt9400 were each ^{32}P-labeled and used as hybridization probes of both *D. teissieri* and *D. melanogaster* genomic DNA that had been digested with the restriction endonucleases BamHI, EcoRI, HindIII, or SalI and then subjected to electrophoresis in an agarose gel and blotted to a nitrocellulose filter. λbDt9300 and λbDt9200 hybridized to fragments of the expected sizes in both *D. teissieri* and *D. melanogaster*, although with some enzymes both probes showed faint extra bands in the *D. teissieri* lanes. This is probably due to the presence of restriction-fragment-length polymorphism in the *D. teissieri* fly population from which the DNA was derived. λbDt9400 gave a highly repetitive signal (a dark smear with numerous discrete bands superimposed upon it) when hybridized to each of the four *D. teissieri* restriction digest lanes, indicating the presence of a repetitive DNA element in this clone. The *D. melanogaster* lanes did not show a repetitive pattern; rather, they showed fragments of the sizes found in the *D. melanogaster* 68C glue puff. The λbDt9400 repetitive element was localized to the position shown in Fig. 7 by hybridization of ^{32}P-labeled genomic DNA from *D. teissieri* to a blot filter with lanes of λbDt9400 DNA digested with both PvuII and XbaI. The only strongly labeled band was at the position of a 0.7-kb PvuII–XbaI fragment (see Fig. 8).

RNA-Homologous Regions of the Cloned Sequences

To find out if the 68C-equivalent regions in the species other than *D. melanogaster* contain DNA sequences that code for abundant polyadenylated third instar larval salivary gland RNAs, ^{32}P-labeled cDNA corresponding to the salivary gland poly(A)$^+$ RNA of each species was produced using oligo(dT) primers and reverse transcriptase. This labeled cDNA was than annealed to gel blot filters containing various restriction digests of cloned DNA from the same species. The DNA sequences hybridized by the cDNA were detected by autoradiography, and the sizes of the fragments hybridized indicated the DNA sequences that might code for RNA in each species. The restriction endonucleases used and the results are shown in Fig. 8. Since the reverse transcriptase reaction was performed in limiting amounts of one nucleotide (the labeled one, dCTP), it is unlikely that each RNA transcript was fully copied into cDNA. Rather, the 3′ ends are likely to be relatively overrepresented in the labeled cDNA; thus the sites indicated in the figure may not show the coding position of the 5′ end of each RNA. The results clearly show that the DNA homologous to the glue genes of *D. melanogaster* hybridizes to salivary gland cDNA in each species. It was also found that while *D. melanogaster, D. simulans,* and *D. erecta* each have three noncontiguous DNA regions hybridized by the cDNA, *D. yakuba* and *D. teissieri* each have four. It thus appears that the *D. yakuba* and *D. teissieri* 68C-homologous regions may contain an extra gene as compared with the other species, and therefore that the 68C gene family may have changed in size since the divergence of the species under study. This evidence alone does not exclude the possibility that the extra RNA-coding region in *D. yak-*

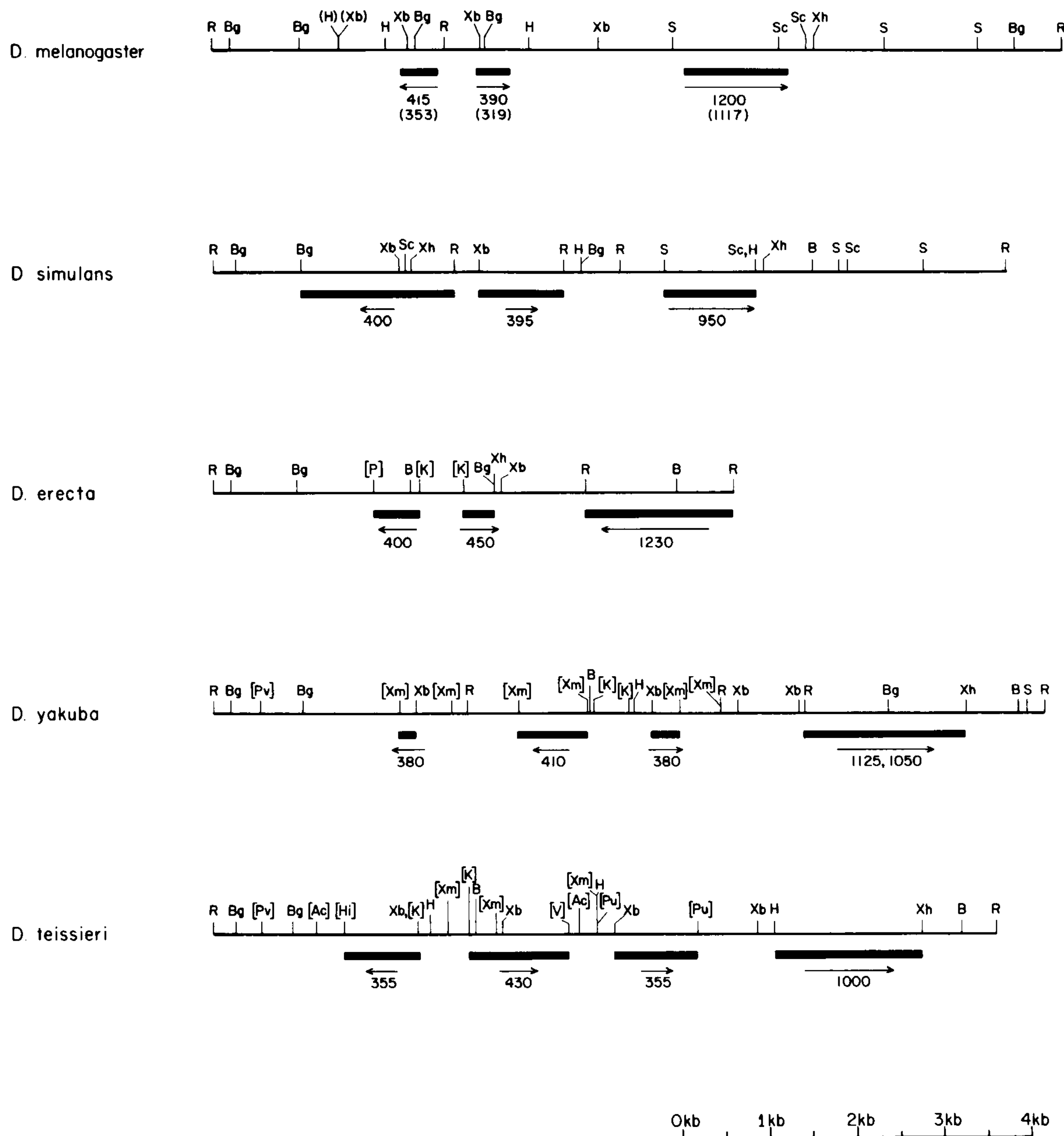

Fig. 8. Detailed restriction maps of 68C-homologous sequences coding for abundant third instar salivary gland RNAs. Restriction enzyme abbreviations are those used in Fig. 4, with the following additions: AccI (Ac), HincII (Hi), KpnI (K), PstI (P), PvuII (Pu), PvuI (Pv), EcoRV (V), and XmnI (Xm). Sites in parentheses are used as in Fig. 4. Sites in brackets indicate that only a subset of the sites recognized by the indicated enzyme are shown. The maps are aligned by the EcoRI site at the left edge. Filled bars below the maps indicate those restriction fragments that hybridize to ^{32}P-cDNA, as in Fig. 4. Arrows indicate the sizes of the RNAs hybridized by each of these regions and the direction of transcription of each of these RNAs. Below each arrow the size of each poly(A)$^+$ RNA is expressed in nucleotides. For *D. melanogaster* the sizes of the RNAs were determined using single-stranded DNA size standards. The extent of each *D. melanogaster* RNA was derived from DNA sequencing results (Garfinkel et al. 1983) and the predicted size of each mRNA [minus any poly(A) tail] is shown in parentheses. For the other species, the size of each transcript was determined using both single-stranded DNA and the *D. melanogaster* 68C RNAs as size standards. Two bands of approximately equal intensity were observed for the largest RNA of *D. yakuba*. This may be due to allelic variation in the *D. yakuba* stock used

uba and *D. teissieri* is due to entry of a new intervening sequence into a preexisting glue gene. This possibility would be excluded by finding that RNAs of different sizes are coded for by each of the four RNA regions in *D. yakuba* and *D. teissieri*, or by determining that all adjacent RNA-hybridized regions are transcribed in opposite directions.

The RNA sizes were determined by subjecting total RNA from third instar salivary glands of each species to electrophoresis in agarose–formaldehyde gels and then transferring the RNA to nitrocellulose filters by blotting. The resulting filters were hybridized with a ^{32}P-labeled restriction fragment from cloned DNA derived from the appropriate species,

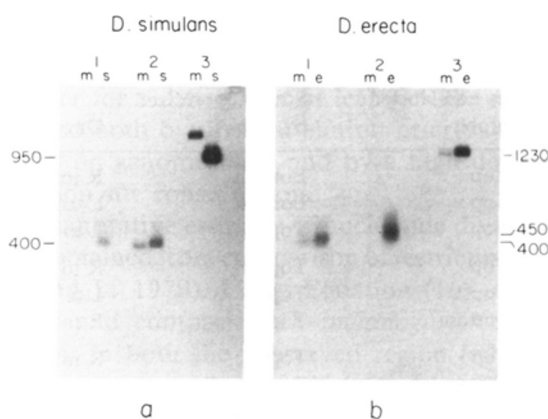

Fig. 9a,b. Hybridization of cloned probes to species RNAs. **a** *D. simulans* nick-translated probes hybridized to *D. melanogaster* (m) and *D. simulans* (s) total third instar salivary gland RNA. Probes used were (1) fDs121, (2) fDs122, and (3) fDs124. Both the fDs122 and fDs124 probes cross-hybridized to the expected *D. melanogaster* RNAs (sgs-7 and sgs-3, respectively). The lack of cross-hybridization by the fDs121 probe to any *D. melanogaster* RNA is probably due to the low level of the sgs-8 transcript produced by the OR16f *D. melanogaster* strain used (Crowley and Meyerowitz 1984). Numbers on the side indicate the lengths, in nucleotides, of the *D. simulans* RNAs. **b** *D. erecta* end-labeled probes hybridized to *D. melanogaster* (m) and *D. erecta* (e) total third instar salivary gland RNA. Probes used were (1) EcoRI insert of λeDe5020, (2) EcoRI insert of λeDe5021, and (3) EcoRI insert of fDe009. The probe containing the large RNA-coding region (fDe009) shows noticeable cross-hybridization to the sgs-3 transcript of *D. melanogaster*. The extent of cross-hybridization appears to be less than that in the similar *D. simulans* vs *D. melanogaster* experiment in **a**. The cross-hybridization observed using the λeDe5020 probe is to the sgs-7 transcript of *D. melanogaster* as determined by the size of the RNA, rather than to the expected sgs-8 RNA. It is not known if this indicates inversion of the region coding for the small RNAs of *D. erecta* in comparison with that of *D. melanogaster*, in addition to the *D. erecta* inversion that includes the gene coding for the large sgs-3 homologous RNA. Numbers on the side indicate the lengths, in nucleotides, of the *D. erecta* RNAs

Table 1. λgt10 subclones used for transcription direction mappings

Clone	Source
D. erecta	
λeDe5020	2.4-kb *Eco*RI–*Kpn*I fragment of fDe008
λeDe5021	1.4-kb *Kpn*I–*Eco*RI fragment of fDe008
D. yakuba	
λeDy5120	2.9-kb *Eco*RI insert of qDy5110
λeDy5121	1.5-kb *Eco*RI–*Kpn*I fragment of qDy5111
λeDy5122	1.1-kb *Kpn*I–*Eco*RI fragment of qDy5111
λeDy5123	2.7-kb *Eco*RI insert of qDy5113
D. teissieri	
λeDt5420	2.5-kb *Eco*RI–*Hin*dIII insert of aDt5402
λeDt5421	1.9-kb *Hin*dIII insert of aDt5403
λeDt5422	2.0-kb *Hin*dIII insert of aDt5404
λeDt5423	2.6-kb *Hin*dIII–*Eco*RI insert of aDt5405

Where necessary, *Eco*RI linkers were ligated onto the blunt-ended fragment. The fragments were all cloned into the *Eco*RI site of λgt10

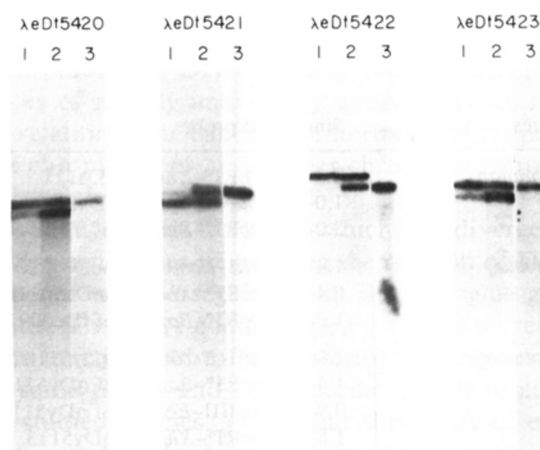

Fig. 10. Transcription direction mapping of *D. teissieri* RNAs. Three hundred nanograms of each of the indicated λgt10 subclones (see Table 1) was denatured with 0.1 N NaOH at 37°C for 10 min, loaded onto a wide lane of a 0.4% agarose gel [2 mM EDTA, 40 mM Tris-acetate (pH 8.3)], and separated by electrophoresis at 0.75 V cm^{-1} for approximately 24 h. Since each of the λgt10 subclones used contains only a single RNA-coding block, the direction of transcription of each of the genes can be determined by observing which of the two separated strands hybridizes to a single-stranded cDNA probe synthesized from poly(A)$^+$ third instar salivary gland RNA. It is also necessary to determine the 5′-to-3′ orientation of the insert DNA of each of the separated strands relative to the map shown in Fig. 8. This was done utilizing a restriction fragment of DNA homologous to the λgt10 insert DNA labeled at only one of its 3′ ends. This probe hybridizes to the strand that has a 5′-to-3′ direction opposite to that of the end-labeled strand of the probe. After electrophoresis, the gel was denatured in 1.5 M NaCl, 0.5 M NaOH for 1 h and then neutralized in 1 M Tris–HCl (pH 8.0), 1.5 M NaCl for 1 h. The gel was then blotted to a nitrocellulose filter. Each lane of the baked filter was then cut into three strips. The leftmost strip (strip 1) was hybridized with a probe labeled at a single 3′ end. The middle strip (strip 2) was hybridized with a probe labeled at both of its 3′ ends and was used to register the location of both of the separated strands. The rightmost strip (strip 3) was hybridized with the cDNA probe. The probes used and the transcription directions obtained are shown in Table 2 and in Fig. 8

from each of the RNA-coding blocks. Adjacent to each RNA lane on each blot filter was *D. melanogaster* salivary gland RNA included to detect cross-hybridization between the DNA of each species and the specific glue RNAs of *D. melanogaster*. Single-stranded DNA size standards were also included in each gel. Autoradiograms from the *D. simulans* and *D. erecta* experiments are shown in Fig. 9. The sizes of the RNAs are shown in Fig. 8. The cloned DNA probes used in these RNA blot experiments often hybridized weakly to the *D. melanogaster* 68C glue RNAs that were in the lanes adjacent to the strongly labeled RNAs from the same species as the probe (Fig. 9). Thus, the *D. melanogaster* glue RNAs appear to be homologous to, though quite diverged from, the similar RNAs of each of the other species.

The transcription direction of each of the RNAs

Table 2. Summary of transcription direction mapping experiments

Species	Single-end probe[a]	Strand hybridized by end-labeled probe	Strand hybridized by cDNA	Direction[b]
D. simulans	2.2-kb *Eco*RI*–*Xho*I of fDs121	Top	Bottom	R to L
	1.0-kb *Xba*I–*Eco*RI* of fDs122	Top	Bottom	L to R
	2.2-kb *Eco*RI*–*Bam*HI of fDs124	Top	Top	L to R
D. erecta	1.4-kb *Bgl*II–*Eco*RI* of λeDe5020[c]	Top	Top	R to L
	0.4-kb *Eco*RI*–*Xba*I of λeDe5021[c]	Top	Top	L to R
	1.0-kb *Eco*RI*–*Bam*HI of fDe009	Bottom	Top	R to L
D. yakuba	1.9-kb *Bgl*II–*Eco*RI* of qDy5110	Top	Top	R to L
	1.4-kb *Eco*RI*–*Bam*HI of qDy5111	Bottom	Top	R to L
	0.9-kb *Hin*dIII–*Eco*RI* of qDy5111	Top	Bottom	L to R
	1.8-kb *Eco*RI*–*Xho*I of qDy5113	Top	Top	L to R
D. teissieri	1.6-kb *Bgl*II–*Hin*dIII* of aDt5402	Top	Top	R to L
	1.5-kb *Bam*HI–*Hin*dIII* of aDt5403	Bottom	Top	L to R
	0.9-kb *Pvu*II–*Hin*dIII* of aDt5404	Top	Bottom	L to R
	1.7-kb *Hin*dIII*–*Xho*I of aDt5405	Top	Top	L to R

[a] Asterisk indicates 3′ end-labeled site
[b] "R to L" indicates right to left; "L to R" indicates left to right. See Fig. 8
[c] *Eco*RI ends of λgt10 subclones were generated during cloning.

was also determined. The strategy used is shown in Fig. 10. The DNA fragments and probes used, and the results obtained, are listed in Tables 1 and 2 and depicted in Fig. 8. The transcription direction of the rightmost *D. erecta* RNA was also determined by DNA sequencing to confirm that it is indeed inverted relative to the orientation of transcription of the similar RNA from all of the other species. The EcoRI fragment coding for this RNA was labeled at the 3′ ends of both strands, using the large proteolytic fragment of *E. coli* DNA polymerase I to add ^{32}P-labeled residues. After digestion with BamHI, the larger of the two resulting fragments (see Figs. 4 and 8), now labeled at only one end, was sequenced using the chemical sequencing method of Maxam and Gilbert (1980). The sequence obtained showed clearly that the DNA adjacent to the EcoRI site is homologous to the 3′ end of the *D. melanogaster Sgs-3* gene sequence and that it is indeed inverted relative to the *D. melanogaster* orientation. This sequence will be presented at a later date (C. Martin and E. Meyerowitz, work in progress). The results from all of the hybridizations show that adjacent coding regions in all of the species code for RNAs of different sizes or of opposite transcription directions. This eliminates the possibility that the enlargement of the RNA coding region of *D. yakuba* and *D. teissieri* is due solely to addition of one or more new intervening sequences.

Discussion

Several conclusions are possible from the results presented. The first is that the five closely related *Drosophila* species studied all do contain DNA sequences hybridized by the 68C glue gene cluster of *D. melanogaster*, and that in all of the species these sequences contain DNA that is transcribed to give several abundant polyadenylated RNA species in third instar larval salivary glands. This is consistent with previous work showing that *D. simulans, D. yakuba*, and *D. teissieri* all have puffs similar to that found at 68C in *D. melanogaster*, at the analogous chromosomal position (Ashburner and Lemeunier 1972; Ashburner and Berendes 1978). Although all of the species have a 68C-homologous gene cluster, it is clear that the 68C gene family has evolved since the divergence of the species studied. That the DNA sequence of the genes has changed is evidenced by the difference in restriction endonuclease sites within the individual genes, by the different sizes of the RNAs in the different species, and by the weak cross-hybridization between the genes of the various other species to the *D. melanogaster* glue RNAs. Divergence is also shown by the difference in the number of genes (some of which may in fact be pseudogenes) in the species, and by the inversions of certain of the genes relative to the others. This divergence includes more than just the RNA-coding DNA; the entire region of chromosomal DNA that includes the gene family and all of the flanking sequences is remarkably different in the *Drosophila* species studied. This is shown both by the thermal elution of labeled *D. melanogaster* DNA probes from the DNA representing this region in genome blot experiments and by the virtual absence of any conserved restriction endonuclease sites in the entire region, as shown in Figs. 4 and 8. In striking contrast is the adjacent DNA to the left (toward the telomere in *D. mela-*

nogaster). This set of sequences shows an extraordinary degree of conservation from one species to the other for a distance of at least 13 kb, as demonstrated both by thermal elution of cross-species hybrids on genome blots and by a high degree of restriction site conservation.

A quantitative estimate of nucleotide divergence can be obtained from comparison of restriction maps (Nei and Li 1979). Using Equation (16) of these authors and comparing *D. melanogaster* and *D. simulans* in both the conserved region (where *D. melanogaster* has 18 restriction sites and *D. simulans* 19, with 18 shared) and the RNA-coding region (where *D. melanogaster* has 13 sites and *D. simulans*, 15, with 6 apparently shared) it can be estimated that the sequences in the conserved regions of these species are diverged by less than 0.5%, whereas in the RNA-coding region the mean frequency of nucleotide substitution per position is about 18%. Thus, at least at the 68C glue locus, evolutionary divergence occurs at very different rates in adjacent blocks of chromosomal DNA sequence.

It also appears that the processes leading to the observed divergence are different in the adjacent regions. In all five species the spacing between shared restriction sites is the same in most of the conserved region, with different unshared sites appearing or disappearing against an otherwise constant background. This implies that the primary process in divergence is single-nucleotide substitution, although of course tiny deletions or substitutions of small blocks of nucleotides would not have been detected in our experiments. In contrast, the rapidly evolving RNA-coding region is subject to insertions, deletions, inversions, and an extraordinary number of apparent single-site changes. In fact, these processes are so evident that it is meaningless to try to estimate the levels of nucleotide divergence in this region between any of the species except for the closely related siblings *D. melanogaster* and *D. simulans*, since the available methods for making such calculations (Nei and Li 1979; Engels 1981) assume that only single-nucleotide substitutions have occurred.

That *Drosophila* genomes contain large interspersed blocks of rapidly evolving and more slowly evolving DNA has been predicted from the results of thermal elution studies on interspecies hybrids of single-copy genomic DNA (Zwiebel et al. 1982). The work reported here confirms this in a specific instance, and also shows that the rapidly evolving DNA can code for messenger RNAs. The possibility that mammalian genomes may also contain interspersed blocks of DNA with different evolutionary rates has been shown by analysis of the mouse major histocompatibility complex (Hood et al. 1983). The evolutionary mechanisms that result in disparate rates of sequence divergence in adjacent domains of chromosomal DNA are unknown. Whether the blocks of rapidly and slowly evolving DNA bear any relation to the bands and interbands of the polytene chromosomes, or to other chromatin features, is also unknown.

What is clear is the use to which this divergence pattern can be put in studying the relation of DNA sequence to gene regulation in the 68C glue gene cluster. The RNA-coding region seems to serve the same function and to be transcribed in response to the same tissue- and time-specific signals in all of the studied species. Proof that the DNA of each species does indeed respond to identical intracellular signals will be sought in interspecies P-factor-mediated transformation experiments. It is already known that the cloned *Sgs-3* gene of *D. melanogaster* (including none of the conserved region sequences) is expressed normally after P-factor-mediated reintroduction to the *D. melanogaster* genome (Crosby 1983; M. Crosby and E. Meyerowitz, work in progress). If the genes of the other species are expressed normally when integrated into the *D. melanogaster* genome, DNA sequencing studies should reveal which, if any, regions of the genic and intergenic sequences have been conserved in evolution and thus may be functionally significant, since conserved islands of sequence should be evident against the remarkably diverged background of the surrounding DNA. These studies will also show if all of the genes are capable of coding for proteins, and if the proteins coded for are similar to the 68C glue polypeptides of *D. melanogaster*. Sequencing studies should also point to the nature of the events that result in extremely rapid divergence in defined chromosomal segments, and may show any special features of the DNA sequences found at the sharp boundary between the rapidly and slowly evolving regions.

Acknowledgments. We would like to acknowledge Mark Garfinkel's construction and screening of a portion of the *D. teissieri* library, and to thank Lynn Crosby, Tom Crowley, Mark Garfinkel, Pete Mathers, and Bob Pruitt for their comments on the manuscript. This work was supported by NIH grant GM28075 to E.M.M. C.H.M. was supported by a National Science Foundation predoctoral fellowship.

References

Ashburner M (1967) Patterns of puffing activity in the salivary gland chromosomes of *Drosophila*. I. Autosomal puffing patterns in laboratory stock of *Drosophila melanogaster*. Chromosoma 21:398–428

Ashburner M (1973) Sequential gene activation by ecdysone in polytene chromosomes of *Drosophila melanogaster*. I. Dependence upon ecdysone concentration. Dev Biol 35:47–61

Ashburner M (1974) Sequential gene activation by ecdysone in polytene chromosomes of *Drosophila melanogaster*. II. The

effects of inhibitors of protein synthesis. Dev Biol 39:141–157

Ashburner M, Berendes HD (1978) Puffing of polytene chromosomes. In: Ashburner M, Wright TRF (eds) The genetics and biology of *Drosophila*. vol 2b. Academic Press, London pp 315–395

Ashburner M, Lemeunier F (1972) Patterns of puffing activity in the salivary gland chromosomes of *Drosophila*: VII. Homology of puffing patterns on chromosome arm 3L in *D. melanogaster* and *D. yakuba*, with notes on puffing in *D. teissieri*. Chromosoma 38:283–295

Ashburner M, Richards G (1976) Sequential gene activation by ecdysone in polytene chromosomes of *Drosophila melanogaster*. III. Consequences of ecdysone withdrawal. Dev Biol 54:241–255

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

Bock IR, Wheeler MR (1972) The *Drosophila melanogaster* species group. University of Texas Publication 7213. University of Texas, Austin

Bolivar F (1978) Construction and characterization of new cloning vehicles. III. Derivatives of plasmid pBR322 carrying unique EcoRI sites for selection of EcoRI generated recombinant DNA molecules. Gene 4:121–136

Bolivar F, Rodriguez RL, Greene PJ, Betlach MC, Leyneker HL, Boyer HW, Crossa JH, Falkow S (1977) Construction and characterization of new cloning vehicles. II. A multipurpose cloning system. Gene 2:95–113

Collins J, Hohn B (1978) Cosmids: a type of plasmid gene-cloning vector that is packageable in vitro in bacteriophage λ heads. Proc Natl Acad Sci USA 75:4242–4246

Crosby MA (1983) Determination of sequences necessary for regulated expression of the Sgs-3 gene. In: Caltech Biology Annual Report. California Institute of Technology, Pasadena, pp 58–59

Crowley TE, Meyerowitz EM (1984) Steroid regulation of RNAs transcribed from the *Drosophila* 68C polytene chromosome puff. Dev Biol 102:110–121

Crowley TE, Bond MW, Meyerowitz EM (1983) The structural genes for three *Drosophila* glue proteins reside at a single polytene chromosome puff locus. Mol Cell Biol 3:623–634

Davis RW, Botstein D, Roth JR (1980) Advanced bacterial genetics. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Denhardt DT (1966) A membrane-filter technique for the detection of complementary DNA. Biochem Biophys Res Commun 23:641–646

Dworniczak B, Kobus S, Schaltmann-Eiteljorge K, Pongs O (1983) Ecdysterone, ecdysterone receptor, and chromosome puffs. In: Roy AK, Clark JH (eds) Gene regulation by steroid hormones II. Springer-Verlag, New York, pp 79–91

Engels WR (1981) Estimating genetic divergence and genetic variability with restriction endonucleases. Proc Natl Acad Sci USA 78:6329–6333

Fraenkel G, Brookes VJ (1953) The process by which the puparia of many species of flies become fixed to a substrate. Biol Bull 105:442–449

Garfinkel MD, Pruitt RE, Meyerowitz EM (1983) DNA sequences, gene regulation and modular protein evolution in the *Drosophila* 68C glue gene cluster. J Mol Biol 168:765–789

Gronemeyer H, Pongs O (1980) Localization of ecdysterone on polytene chromosomes of *Drosophila melanogaster*. Proc Natl Acad Sci USA 77:2108–2112

Hood L, Steinmetz M, Malissen B (1983) Genes of the major histocompatibility complex of the mouse. Annu Rev Immunol 1:529–568

Lemeunier F, Ashburner M (1976) Relationships within the *melanogaster* species subgroup of the genus *Drosophila* (*sophophora*). II. Phylogenetic relationships between six species based upon polytene chromosome banding sequences. Proc R Soc Lond [Biol] 193:275–294

Lemeunier F, Dutrillaux B, Ashburner M (1978) Relationships within the *melanogaster* subgroup species of the genus *Drosophila* (*sophophora*). Chromosoma 69:349–361

Lewis EB (1960) A new standard food medium. Dros. Inf. Ser. 34:117–118

Lis JT, Prestidge L, Hogness DS (1978) A novel arrangement of tandemly repeated genes of a major heat shock site in *D. melanogaster*. Cell 4:901–919

Maniatis T, Fritsch EF, Sambrook J (1982) Molecular cloning. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York

Maxam AM, Gilbert W (1980) Sequencing end-labeled DNA with base-specific chemical cleavages. Methods Enzymol 65: 499–560

McDonell MW, Simon MN, Studier FW (1977) Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J Mol Biol 110:119–146

Meyerowitz EM, Hogness DS (1982) Molecular organization of a *Drosophila* puff site that responds to ecdysone. Cell 28: 165–176

Murray NE, Brammar WJ, Murray K (1977) Lambdoid phages that simplify the recovery of in vitro recombinants. Mol Gen Genet 150:53–61

Nei M, Li W-H (1979) Mathematical model for studying genetic variation in terms of restriction endonucleases. Proc Natl Acad Sci USA 76:5269–5273

Peacock C, Dingman CW (1968) Molecular weight estimation and separation of ribonucleic acid by electrophoresis in agarose–acrylamide composite gels. Biochemistry 7:668–674

Pelling C (1964) Ribonukleinsäure-Synthese der Riesenchromosomen. Autoradiographische Untersuchungen an *Chironomus tentans*. Chromosoma 15:71–122

Rigby PWJ, Dieckmann 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

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

Sternberg N, Tiemeier D, Enquist L (1977) In vitro packaging of a λ Dam vector containing EcoRI fragments of *Escherichia coli* and phage P1. Gene 1:255

Sturtevant AH (1919) A new species closely resembling *Drosophila melanogaster*. Psyche (Stuttg) 26:153–155

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