

Large scale screen for transposon insertions into cloned genes

(*Drosophila*/P element/reverse genetics/plasmid rescue)

BRUCE A. HAMILTON, MICHAEL J. PALAZZOLO*, JOO H. CHANG, K. VIJAYRAGHAVAN†, CAROL A. MAYEDA*, MICHAEL A. WHITNEY‡, AND ELLIOT M. MEYEROWITZ

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

Communicated by Norman H. Horowitz, December 14, 1990 (received for review August 31, 1990)

ABSTRACT We describe a method of screening for transposon insertions in or near *Drosophila* loci that correspond to cloned DNA sequences. We mobilize a modified P element transposon that carries a bacterial plasmid origin of replication and a drug-resistance marker. The genomic sequences flanking each transposon insertion site can then be rescued as a plasmid in *Escherichia coli*. Libraries of such plasmids, representing pools of transposon-mutagenized individuals, are used as hybridization probes against cloned sequences to determine whether a transposon has inserted next to a particular site in the genome. The number of loci that can be screened simultaneously by this procedure is quite large. We have screened an array of cDNA clones representing almost 700 distinct loci against libraries representing 760 mutagenized flies, and we obtained hybridization signals to 7 different cDNAs. Three of these events have been analyzed in detail and represent genuine insertions near genomic sequences that correspond to the cDNAs.

To understand the development and physiology of a nervous system in molecular detail, we have initiated a large scale analysis of RNAs expressed in the *Drosophila* brain. Our study began with the isolation of a large number of cDNA clones that represent messages expressed in the adult head but not in the preblastoderm embryo (1). Combining molecular and genetic analyses of these molecules should lead to testable models of their neuronal functions. An important requirement of this approach is the ability to identify mutations in the genetic loci that encode the cloned molecules. Such a genetic screen must meet several criteria: the screen must be simple and compatible with studying many loci simultaneously, it must have high sensitivity and low background, and it must not require any assumptions about mutant phenotypes. This paper describes a screening procedure that meets these criteria.

Although other methods to disrupt cloned *Drosophila* genes have been reported, none meets all of the criteria listed above. While classical genetic screens have been successful, they require a prediction of a mutant phenotype. Screening for loss of antibody binding (2, 3) requires unique reagents and assays for each locus. Methods based on PCR (4–8) require DNA sequence data and at least one unique oligodeoxynucleotide primer for each locus considered. Each locus must also be screened in a separate biochemical reaction. In addition, distinguishing true insertion events from spurious amplification products (ref. 7; B.A.H., M.A.W., C.A.M., and M.J.P., unpublished data) requires considerable effort in larger experiments.

We describe a strategy for reverse genetics that circumvents these problems. The method is based on plasmid rescue of genomic DNA (9, 10) that flanks the insertion sites of a modified P element transposon, PlacW (11). By using a

library of rescued sequences as a hybridization probe against an array of cDNA clones, we are able to assay for insertions in or near a large number of cloned loci in a single experiment. This requires no assumptions of expected mutant phenotypes and requires neither antisera nor DNA sequence data for PCR primers. Because PlacW carries a *w*⁺ minigene as a dominant genetic marker, it should be possible by standard mutagenesis methods to obtain deletion alleles of loci that are tagged but not functionally disrupted by the transposon. Enough probe could be synthesized from a single plasmid library to screen an array of cloned DNA representing the estimated sequence complexity of the *Drosophila* genome.

MATERIALS AND METHODS

Fly Strains and Crosses. We obtained new PlacW insertions by genetic mobilization. As a source of transposons we used C(1)RM, *y w* 4[PlacW], which carries four copies of PlacW on each homolog and was kindly provided by Dan Lindsley (University of California, San Diego). Transposase activity from the stable source P[ry⁺ Δ2-3](99B) (12) was supplied on a derivative of TM2, *Ubx ry*, that carries this insertion, kindly made available by John Merriam (University of California, Los Angeles). Other genetic elements have been described (13, 14). Flies were raised on standard medium (15) at 22°C. C(1)RM, *y w* 4[PlacW] virgins were mated to *shi*; TM2, *Ubx ry* P[ry⁺ Δ2-3](99B)/*Sb ry* P[ry⁺ Δ2-3](99B) males (12 females and 4 males per pint bottle) for 3–5 days and transferred or discarded. Each mating was transferred no more than twice. Dysgenic virgin females of the genotype C(1)RM, *y w* 4[PlacW]; TM2 *Ubx ry* P[ry⁺ Δ2-3](99B) were collected by heat treatment as described (16), and mated as described above to males of the genotype *w*; CyO; TM6, *Ubx/T(2;3)Xa* or *w*; CyO; TM6, *Hu/T(2;3)Xa*. Progeny males with pigmented eyes (indicating a PlacW transposition to an autosome) were individually mated to two or three females of an appropriate balancer strain to establish temporary lines. Males were removed for DNA preparations after 3–5 days.

cDNA Gel Blots. cDNA inserts from λSWAJ3 clones (1, 17) were isolated by performing PCR directly on high-titer lysates of cloned bacteriophage with primers that abut the cloning site (SWAJ3.1, 5'-ATTTAGGTGACACTATAGAATACAC-3'; SWAJ3.2, 5'-CGGAAGCTTGGGCTGCAGGTCGACT-3'). λSWAJ3 contains no plasmid sequences that could hybridize to the plasmid probes in the screening procedure. DNA gel blots of PCR products were prepared by standard methods (18).

DNA Preparations. We prepared DNA from 10–20 flies by a standard method (19) with modifications. Flies were col-

*Present address: Department of Genetics, Washington University School of Medicine, St. Louis, MO 63110.

†Permanent address: Molecular Biology Unit, Tata Institute of Fundamental Research, Bombay 400005, India.

‡Present address: Department of Biology and Center for Molecular Medicine M-049, University of California, San Diego, La Jolla, CA 92023.

groups of 20 flies each. Half of the DNA from each preparation was pooled and aliquots were either partially digested with *EcoRI* or digested to completion with *Sac II*. Digested fly DNA was circularized in the presence of T4 DNA ligase. Transformation of *E. coli* produced 908 drug-resistant colonies from the *EcoRI*-digested DNA and 563 colonies from the *Sac II*-digested DNA. Plasmid DNA was prepared directly from colonies, yielding $\approx 8 \mu\text{g}$ per library. DNA from each library was labeled to high specific activity ($7\text{--}9 \times 10^8$ cpm/ μg) and equal amounts of radioactivity from each were combined as a hybridization probe against the filter array of cDNA targets. Autoradiography revealed hybridization to two cDNA clones: c3B7 and c4D12 (Fig. 2). Similar screens, representing *PlacW* insertions in 620 additional males, revealed hybridization signals to clones c1C2, c1F4, c4B9, c4E10, c6F9, and c8H9 (Table 1). Two additional clones in our collection gave a hybridization signal with every probe tested; these signals appear to be nonspecific and we have neither counted them as positive signals nor pursued them further.

Characterization of Rescued Plasmids. A rescued plasmid that corresponds to a *PlacW* insertion near a gene of interest should contain both a single contiguous piece of the fly genome that hybridizes to the cDNA and a *P* element terminal repeat sequence from the transposon. We have tested this for the first three presumed insertion events (c3B7, c4D12, and c8H9), as described below.

We isolated the relevant plasmids by colony filter hybridization to cDNA probes, using both the original plasmid-rescue libraries and primary transformant colonies representing either pools of 20 fly lines or single fly lines. A c3B7 cDNA probe allowed the isolation of nine independent plasmids from *EcoRI*-rescued libraries, but none from *Sac II*-rescued libraries. Each plasmid contained a 4-kilobase (kb) genomic fragment with no internal *EcoRI* sites. It is unlikely that these plasmids could have arisen by ligation of c3B7 homologous sequences to noncontiguous plasmid sequences, as this would generate an internal *EcoRI* site and a new rescued fragment size in each instance. Similarly, independent plasmids for c4D12 were isolated from *EcoRI* (partial digest), but not *Sac II*, libraries. Each of these contained either a 1.2-kb genomic fragment with no internal *EcoRI* sites or a 2-kb genomic fragment with one *EcoRI* site 1.2 kb from the *P* element end. Several independent c8H9 homologous plasmids were rescued by *Sac II* but not by *EcoRI*. Each of these

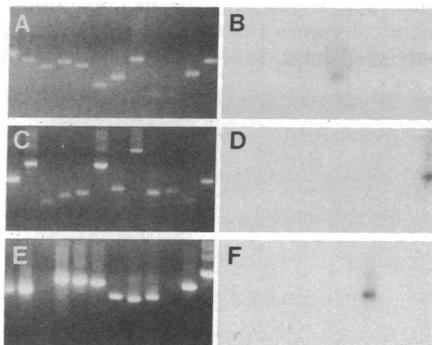


FIG. 2. Hybridization of rescued plasmids to cDNA gel blots. DNA gel blots of PCR-amplified cDNA inserts were hybridized to probes from rescued plasmids representing 140–160 mutagenized flies. (A) Agarose gel of some amplified cDNAs, including c3B7, stained with ethidium bromide and visualized by UV fluorescence. (B) Autoradiogram of the gel blot corresponding to A, showing hybridization to cDNA c3B7. (C) Gel including c4D12. (D) Autoradiogram of blot of C, showing hybridization to c4D12. (E) Gel including c8H9. (F) Autoradiogram corresponding to gel in E showing hybridization to c8H9.

Table 1. Plasmid rescue and hybridization results

Probe	No. of lines	No. of colonies	cDNA clone(s)
1	140	1471	c3B7, c4D12
2	140	776	—
3	160	390	c8H9
4	160	500	c1C2, c1F4, c4B9
5	160	454	c1F4, c4E10

contained a 7-kb genomic fragment lacking internal *Sac II* sites.

A rescued plasmid and the original cDNA clone for each locus were used to probe duplicate gel blots of genomic DNA digested with a variety of restriction enzymes. For each locus, the rescued plasmid and the cDNA identify overlapping patterns of bands and each probe appears to be single copy in the genome. This further demonstrates that the recovered sequences represent discrete, tagged sites rather than ligations of disjoint sequences (or tagged dispersed repeat sequences).

A genuine rescue product should also contain one of the *P* element terminal repeat sequences from the *PlacW* transposon. We subjected one of the plasmids rescued from each locus to chain-termination sequencing (23) by using a synthetic primer that corresponds to the *P* element terminus (7). From each of these reactions, we obtained a unique sequence ladder, indicating that each of these plasmids contains a single *P* element terminus.

Isolation of Fly Lines Carrying the Identified Insertions. We have used two different methods to identify the single lines of flies that carry the desired *PlacW* insertions. The first method is based on PCR; the second is based on plasmid rescue.

We used sequence information from the rescued plasmids to design locus-specific PCR primers that lie adjacent to the *PlacW* terminal repeat; we used these primers in combination with the *P* element terminal repeat primer for PCR on fly genomic DNA templates. Amplification between the c3B7 and *P* element primers was used to detect a PCR product from DNA from pools of fly lines and then from individual lines. Only 1 of the original 20-fly DNA preparations in the pool that hybridized to c3B7 supports an amplification product of the size predicted from DNA sequence and illustrated by PCR on the rescued plasmid (≈ 220 base pairs; Fig. 3). Two of the 20 lines in this pool had failed to propagate. We prepared DNA from each of the remaining 18 lines for PCR. Nine of the lines gave amplification products of the appropriate size. This result was repeated with a second set of DNA preparations and has also been partially verified by plasmid rescue experiments and *in situ* hybridizations to polytene chromosomes (see below).

In addition to PCR sorting experiments, we assayed for the c3B7 and c4D12 insertions by plasmid rescue. PCR results indicated that a single pool contained both insertions. We

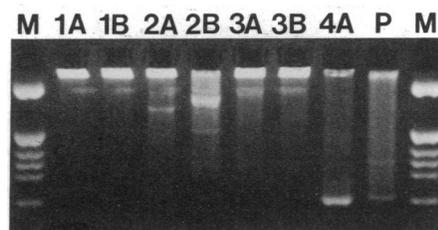


FIG. 3. PCR amplification between c3B7 and *P* element primers. Lanes M, size standards. Alphanumerically labeled lanes are amplification products from each of the pools of 20 flies used to make the first rescued-plasmid library. Lane P, control amplification from the rescued plasmid.

used *EcoRI* partial digests of new fly DNA preparations to rescue transposon-flanking sequences from each of 13 lines from this pool (2 lines were propagating poorly and so were not used to prepare DNA and 3 more lines did not yield drug-resistant colonies in this experiment). Minipreparations of plasmid DNA from 1–3 individual colonies for each line rescued were analyzed by DNA gel blot hybridization. A c3B7 probe hybridizes to plasmids from each of the lines implicated by PCR from which rescued plasmids were recovered: lines 4.01, 4.03, 4.08, 4.10, and 4.13. The filter was stripped of the c3B7 probe and hybridized to a c4D12 probe. The c4D12 probe identified one line, 4.09. Since the 13 plasmid-rescued lines were sufficient to identify one line bearing the c4D12 insertion and several bearing the c3B7 insertion, we have not tried further to rescue plasmids from any of the remaining 5 lines.

We failed to identify an insertion line for c8H9 by either method. From the pool of 20 lines that produced the c8H9 *Sac* II rescue plasmid, 6 lines failed to propagate and could not be tested. This is anomalous: we usually recover offspring from >90% of single-male matings. It appears from the data that the initial c8H9 hybridization signal was due to a genuine insertion, but that the corresponding line was lost.

In Situ Hybridization to Polytene Chromosomes. As a final confirmation of the first two identified insertions, we tested whether the cDNAs for which we isolated insertions and a PlacW element in the identified line are associated with the same cytological location on polytene chromosomes (Fig. 4). Biotin-labeled probes prepared from cDNA clones in pEXLX (24) and from rescued genomic sequences were individually hybridized to larval salivary gland polytene chromosomes from a wild-type strain (Oregon-R). PlacW insertions were localized by hybridizing a biotinylated pBR322 probe to the plasmid sequences of the transposon in polytene chromosomes from identified insertion lines. Each probe identifies a single site in the genome and for each locus the cDNA, the rescued genomic sequence, and the PlacW element all map to the same cytological location: 45D for clone c4D12 and 90D for c3B7. In addition to the results shown in Fig. 4, we have tested the location of PlacW elements in lines 4.01 and 4.08; these also indicate single PlacW insertions at 90D. The rescued fragment identified by c8H9 hybridizes to a single site at 71F.

DISCUSSION

With the advent of modified transposons for enhancer traps (11, 25, 26), it has become possible to clone genes selected for either mutant phenotypes or expression patterns directed by

associated regulatory elements. We have used such modified transposons to take a different approach: isolating lesions in genes that correspond to cloned sequences.

Comparative Advantages of This Method. The screen described here offers several advantages not shared by previous methods. Any presumed insertion can be rapidly verified by characterizing the rescued plasmid before investing significant time in attempting to isolate and characterize a line of flies. Moreover, each of the presumed insertions we have analyzed appears to reflect a genuine correspondence between a contiguous rescued genomic fragment and a target cDNA, suggesting that the incidence of false positive signals should generally be low. Since the modified transposon carries a dominant visible marker (w^+), isolation of subsequent deletion alleles by imprecise excision of the transposon (27–29), radiation mutagenesis, or exposure to chemical mutagens known to cause deletions (30) should be straightforward. In addition, since very little extra effort is required to screen for insertions into additional loci and sufficient mass of rescued plasmid DNA is recovered from minipreparations, projects representing up to 10^4 unique target sequences (or more) should be feasible. This may prove useful for genome mapping and related studies.

Fidelity of the Screen. From probes representing 760 PlacW insertion lines, we obtained hybridization signals to eight clones in our array. We have characterized three of these insertions in detail to demonstrate that they are genuine, rather than artifacts of the screening procedure. Chain-termination sequencing shows that each rescued plasmid contains a single *P* element terminus, as predicted. Gel blot hybridization to restriction-digested genomic DNA shows that each rescued sequence represents a single continuous site in the genome, which encodes the corresponding cDNA. For c4D12 and c3B7, we have also shown that a PlacW element is present in the identified insertion line at the same cytological location as the rescued sequences and the cDNA. Similar criteria have now been used to verify and obtain lines for the c4B9 and c1C2 insertions (although the latter appears to be a transposon; ref. 31 and B.A.H. and J. Liao, unpublished data). Since all of the insertions analyzed appear genuine, we conclude that the screening procedure is reliable.

Screening Efficiency. We have used a collection of nearly 850 cDNA clones (all but the eye-specific clones in ref. 1) as hybridization targets in this screen. Cross-hybridization data suggest that these represent 682 discrete loci (1). However, preliminary DNA sequence data from a subset of almost 250 of these clones suggest that 2–5% of the collection comprise related or identical clones (M.A.W., C.A.M., B.A.H., and M.J.P., unpublished data). This is consistent with a report on

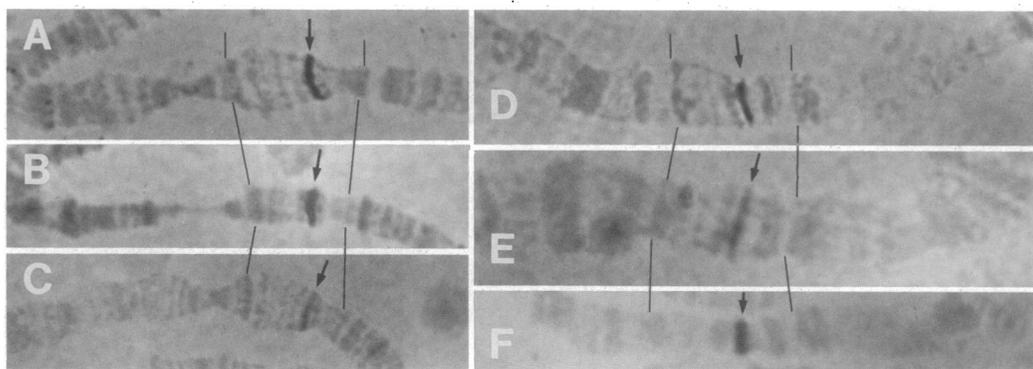


FIG. 4. *In situ* hybridization to polytene chromosomes. Arrows indicate the sites of hybridization signals (90D in A–C; 45D in D–F). Lines delimit the numbered division into which each signal falls. (A) Hybridization of cDNA clone c3B7-15 probe to polytene chromosomes from wild-type strain Oregon-R. (B) Rescued plasmid c3B7.R2 hybridized to Oregon-R chromosomes. (C) pBR322 probe hybridized to polytene chromosomes from the insertion line 4.03. Hybridization is to the plasmid sequences of the PlacW transposon. (D) cDNA clone c4D12-11, Oregon-R chromosomes. (E) Rescued plasmid 4D12R4.1, Oregon-R chromosomes. (F) pBR322 probe to insertion line 4.09.

eye-specific cDNAs from this collection (31). Thus, the number of distinct loci we have screened may be as low as 650.

We assayed 760 fly lines in this screen, predominantly carrying one PlacW element each. However, our dysgenesis scheme allows redundant isolations of single transposition events that occur in premeiotic germ cells or their progenitors; for example, 9 individual lines bear the c3B7 insertion. Such clustering of identical insertions is not surprising. While the transposase should be active quite early in development, the overall rate of transposition is low (15–25% of male progeny from the mass-mated dysgenic females show w^+ function) and any insertion that occurred very early in the germ lineage of one female would comprise a large fraction of the w^+ progeny in a given bottle (see *Materials and Methods*). Thus, the actual number of independent insertions represented by these flies is <760.

Modifications to the screening procedure should allow screening of the large numbers of mutagenized flies needed to find insertions near single target sites. Mating flies en masse and hybridizing DNA gel blots of rescued libraries to cloned probes could save considerable labor. Alternatively, rescued plasmid libraries could be screened by PCR using two locus-specific primers a known distance apart. Having rescued only genomic sequences that flank an insertion obviates the need for a transposon-specific primer and knowing the size of legitimate amplification products reduces background caused by spurious amplification products.

Sorting Lines. We described two methods for finding the appropriate insertion line among a pool of lines known to contain the insertion: PCR and flanking sequence plasmid rescue. PCR is simple, rapid, and extremely efficient; however, this high sensitivity makes it prone to contamination artifacts. Plasmid rescue assays are less sensitive to trace contamination, but they require more starting material.

The PCR sorting strategy we describe differs from those described by Ballinger and Benzer (7) and Kaiser and Goodwin (8) in two important respects. First, the veracity required of the PCR to determine which line carries the identified insertion is much lower than that required to ask whether any line carries any such insertion. Second, we can predict the size of the genuine amplification product because our gene-specific primer is derived by sequencing the rescued genomic fragment from the *P* element primer. This more closely resembles the method of Kim and Smithies (6) and ought to alleviate the signal/noise problems caused by spurious amplification products.

We thank D. L. Lindsley and J. Merriam for fly stocks; E. B. Lewis and S. E. Celniker for fly stocks, useful discussions, and comments on the manuscript; D. G. Ballinger for the *P* element oligodeoxynucleotide, fly stocks, and discussions; U. VijayRaghavan, S. Halsell, and T. Jack for practical advice; A. van der Blik for *shi*; and H. D. Lipshitz, C. H. Martin, and members of the Meyerowitz laboratory for comments on drafts of this paper. B.A.H. was supported in part by a U.S. Public Health Service Predoctoral National Research Service Award (T32 GM07616); M.J.P. is a Lucille P. Markey Scholar; K.V.R. was supported by a Rockefeller Foundation Biotechnology Career Fellowship. This work was supported by a scholar's grant from the Lucille P. Markey Charitable Trust (M.J.P.) and U.S. Public Health Service Program Grant GM40499 (E.M.M.).

- Palazzolo, M. J., Hyde, D., Raghavan, K. V., Mecklenberg, K., Benzer, S. & Meyerowitz, E. M. (1989) *Neuron* **3**, 527–539.
- Van Vactor, D., Jr., Krantz, D. E., Reinke, R. & Zipursky, S. L. (1988) *Cell* **52**, 281–290.
- Katz, R., Moats, W. & Jan, Y. N. (1988) *EMBO J.* **7**, 3471–3477.
- Saiki, R. K., Scharf, S., Faloona, F., Mullis, K. B., Horn, G. T., Erlich, H. A. & Arnheim, N. (1985) *Science* **230**, 1350–1354.
- Saiki, R. K., Gelfand, D. H., Stoffel, B., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–491.
- Kim, H.-S. & Smithies, O. (1988) *Nucleic Acids Res.* **16**, 8887–8903.
- Ballinger, D. G. & Benzer, S. (1989) *Proc. Natl. Acad. Sci. USA* **86**, 9402–9406.
- Kaiser, K. & Goodwin, S. F. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1686–1690.
- Perucho, M., Hanahan, D. & Wigler, M. (1980) *Cell* **22**, 309–317.
- Steller, H. & Pirrotta, V. (1986) *Mol. Cell. Biol.* **6**, 1640–1649.
- Bier, E., Vaessin, H., Shepherd, S., Lee, K., McCall, K., Barbel, S., Ackerman, L., Carretto, R., Uemura, T., Grell, E., Jan, L. Y. & Jan, Y. N. (1989) *Genes Dev.* **3**, 1273–1287.
- Robertson, H. M., Preston, C. R., Phillis, R. W., Johnson-Schlitz, D. M., Benz, W. K. & Engels, W. R. (1988) *Genetics* **118**, 461–470.
- Lindsley, D. L. & Grell, E. H. (1968) *Genetic Variations of Drosophila melanogaster* (Carnegie Institute of Washington, Washington, DC).
- Lindsley, D. L. & Zimm, G. (1990) *Drosoph. Inf. Serv.* **68**, 1–382.
- Lewis, E. B., (1960) *Drosoph. Inf. Serv.* **34**, 117–118.
- Grigliatti, T. (1986) in *Drosophila: A Practical Approach*, ed. Roberts, D. B. (IRL, Oxford, England), pp. 39–58.
- Palazzolo, M. J. & Meyerowitz, E. M. (1987) *Gene* **52**, 197–206.
- Sambrook, J., Fritsch, E. F. & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Lab., Cold Spring Harbor, NY), 2nd Ed.
- Bender, W., Speierer, P. & Hogness (1983) *J. Mol. Biol.* **168**, 17–33.
- Dower, W. J., Miller, J. F. & Ragsdale, C. W. (1988) *Nucleic Acids Res.* **16**, 6127–6145.
- Holmes, D. S. & Quigley, M. (1981) *Anal. Biochem.* **114**, 193–197.
- Feinberg, A. P. & Vogelstein, B. (1984) *Anal. Biochem.* **137**, 266–267.
- Sanger, F., Nicklen, S. & Coulson, A. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 5463–5467.
- Palazzolo, M. J., Hamilton, B. A., Ding, D., Martin, C. M., Mead, D. A., Mierendorf, R. C., Raghavan, K. V., Meyerowitz, E. M. & Lipshitz, H. D. (1990) *Gene* **88**, 25–36.
- O'Kane, C. J. & Gehring, W. J. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 9123–9127.
- Wilson, C., Pearson, R. K., Bellen, H. J., O'Kane, C. J., Grossniklaus, U. & Gehring, W. J. (1989) *Genes. Dev.* **3**, 1301–1313.
- Voelker, R. A., Greenleaf, A. L., Gyurkovics, H., Wisely, G. B., Huang, S.-M. & Searles, L. L. (1984) *Genetics* **107**, 279–294.
- Daniels, S. B., McCarron, M., Love, C. & Chovnick, A. (1985) *Genetics* **109**, 95–117.
- Salz, H. K., Cline, T. W. & Schedl, P. (1987) *Genetics* **117**, 221–231.
- Ashburner, M. (1989) *Drosophila: A Laboratory Handbook* (Cold Spring Harbor Lab., Cold Spring Harbor, NY).
- Hyde, D. R., Mecklenburg, K. L., Pollock, J. A., Vihtelic, T. S. & Benzer, S. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 1008–1012.