The EMBO Journal Vol.2 No.6 pp.927-934, 1983

Microdissection and cloning of the *white* locus and the 3B1-3C2 region of the *Drosophila* X chromosome

V.Pirrotta*, C.Hadfield[1] and G.H.J.Pretorius[2]

European Molecular Biology Laboratory, Postfach 10.2209, D-6900 Heidelberg, FRG

Communicated by V.Pirrotta
Received on 11 March 1983

Fragments from the 3B-3C region of the *Drosophila* X chromosomes were microdissected from salivary gland squashes and their DNA was cloned by the method developed by Scalenghe *et al.* (1981). These clones were used as starting points for a chromosome walk which covers 200 kb including bands 3B2 to 3C2. A number of deletion breakpoints were mapped on the cloned DNA allowing the localisation of several genes in the 3B region. The *white* locus in particular was isolated by microcloning the w^a insertion site. Two transcripts, of 2 and 2.4 kb, respectively, arise from the *white* region and its vicinity.

Key words: Drosophila/microcloning/repetitive sequences/*white* locus

Introduction

The chromosomal region between the *white* and *zeste* loci near the tip of the X chromosome of *Drosophila* has been intensively investigated and saturated with mutations both lethal and visible. It is therefore one of the best known regions of the *Drosophila* genome and contains a number of genes of interest for genetic and molecular analysis. Foremost among them are the *white* and *zeste* loci themselves. The *white* gene is responsible for the deposition and distribution of pigment in the eye. It is considered a complex locus because mutations which map in the proximal part of the locus appear to affect the control of its expression: they alter the spatial distribution of pigment, dosage compensation of the gene and its interaction with the *zeste* locus (reviewed by Judd, 1976). The *white* locus is often the target of chromosomal rearrangements. Many *w* mutations are unstable, generating deletions of flanking sequences and some have been shown to be due to the insertion of transposable elements (Collins and Rubin, 1982; Zachar and Bingham, 1982; Levis *et al.*, 1982b).

One of the more interesting features of the *white* locus is its interaction with *zeste*. The z^1 mutation decreases the expression of *white* if two copies of the proximal part of the *w* locus are present in close physical proximity, either on paired chromosomes or in tandem duplication (Jack and Judd, 1979). *Zeste* mutants have a similar effect on the transvection behaviour of the complex loci *bithorax* and *decapentaplegic* (Gelbart and Wu, 1982).

The *white* locus has been mapped cytogenetically at position 3C1-2 while *zeste* maps at 3A3-4. Between them are some 400 kb of genomic DNA containing a number of loci essential for normal development (Judd *et al.*, 1972), for fertility (Thierry-Mieg, 1982) and for circadian behavioural rhythm (Smith and Konopka, 1981).

Scalenghe *et al.* (1981) developed the microcloning technique and demonstrated its use by isolating recombinant clones from microdissected fragments of the 3A-3C region of the X chromosome. The technique consists of microdissecting small pieces of polytene chromosomes, extracting their DNA in nanoliter volumes and cloning restriction fragments in a phage vector. Recent improvements in the technique (Pirrotta *et al.*, 1983) have increased the resolution and yield of the overall cloning process. The fragment excised can be as small as 100-200 kb of genomic DNA, depending on the density of chromomeres in the region, and enough clones can be recovered to represent all the clonable restriction fragments. In the present work we have used clones obtained by microdissection to isolate a continuous array of overlapping molecular clones covering ~200 kb of genomic DNA including the *white* locus and most of the 3B region.

A complementary approach was used to clone a specific sequence to act as a genetic reference point. Many mutations, both spontaneous and induced, result from the insertion of a transposable element into a gene. In particular, the w^a mutation in the *white* locus was shown by Gehring and Paro (1980) and by Bingham and Judd (1981) to be due to the insertion of a *copia* element. We used *copia* DNA as a probe to identify the clone obtained by microdissection of the w^a locus, which contains *copia* sequences attached to flanking sequences of the insertion site.

Here we present the results of the chromosome walks in the 3B2-3C3 region and a preliminary analysis of the transcripts originating from the vicinity of the *white* locus as a first step in the study of the *zeste-white* interaction. A similar chromosome walk, starting from microdissection clones isolated from the vicinity of the *zeste* locus will be described elsewhere.

Results

Microdissection and cloning of the 3C region

Three fragments were microdissected from gt^{x11}/gt giant chromosome squashes, each containing most of the 3C region and part of the 3B region. The DNA of the fragments was extracted, digested with *Eco*RI and ligated to the λ641 vector in the oil chamber, yielding 256 recombinant clones. The size of the *Eco*RI insert contained in 30 of these recombinants, chosen at random, averaged 3.8 kb with a maximum insert of 10 kb and six inserts <0.5 kb.

The degree of redundancy of this minilibrary of clones specific for the microdissected region was tested with 28 of these clones by hybridising them back individually to the array of 256 recombinants. Eight of them were represented only once, 12 twice, five three times, two four times and one six times, giving, on average, a 2-fold representation. Taken as a rough measure, this indicates that there are ~125 clonable *Eco*RI fragments in the microdissected region. With an average size of 3.8 kb, we estimate that the region contains ~450 kb of genomic DNA. This estimate does not take into

[1]Present address: The Biocentre, Medical Sciences Building, Leicester University, University Road, Leicester LE1 7RH, UK.
[2]Present address: Department of Microbiology, Institute of Pathology, P.O. Box 2034, 0001 Pretoria, Republic of South Africa.
*To whom reprint requests should be sent.

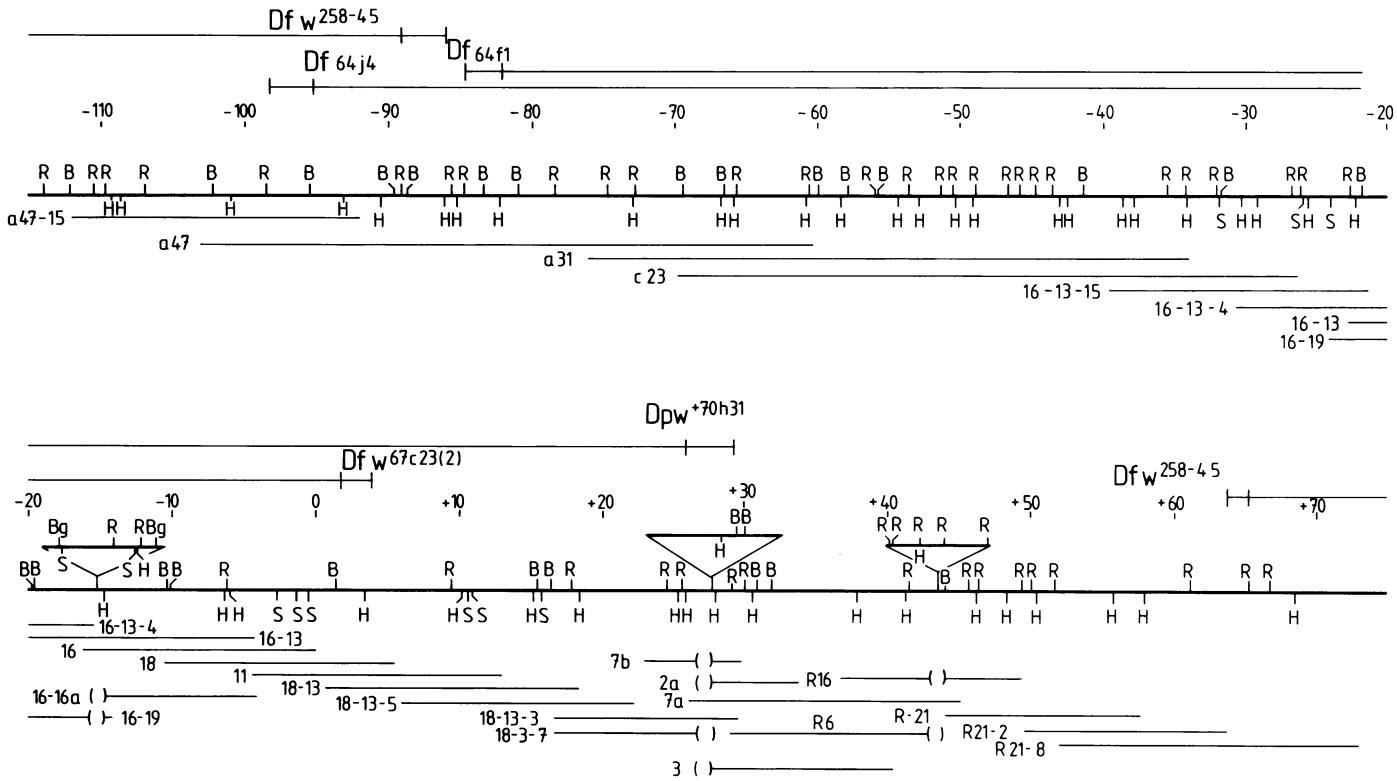

Fig. 1. Restriction map of the cloned region. The distance is measured in kb starting from the w^a insertion site. The positive direction is towards the centromere, negative towards the telomere. Only the most representative clones for each interval are indicated below the map. A parenthesis indicates that the clone enters the insertion sequence shown above. The restriction maps were checked in most cases using restriction fragments to hybridise back to Southern blots. Fragments <0.4 kb were generally not mapped. *Sal* sites were mapped only in the region −35 to +20. The breakpoints of deletions and duplications are indicated above the map by a solid line corresponding to DNA present. *Eco*RI = R, *Bam*HI = B, *Hind*III = H, *Sal*I = S.

account *Eco*RI fragments too large to fit in the λ641 cloning vector which accepts a maximum of 11 kb.

In situ hybridisation to the gt^{x11}/gt chromosomes showed that they contain a *copia* inserted at 3D-4A as well as a 412 element in the middle of the 3C complex. We, therefore, tested the minilibrary for the presence of repetitive sequences by hybridising the array of clones to total *Drosophila* DNA labelled by nick-translation. Sixteen percent or 40 of the clones gave detectable signals. Of these, 12 contained B104 sequences, one contained 412 sequences and three contained *copia*.

A second microcloning experiment was aimed at isolating the *white* locus. Eight chromosome fragments containing little more than the 3C region were microdissected from *gt* w^a chromosomes and the recombinant phage were screened directly for hybridisation to the *copia* element. From slightly over 1000 recombinants we obtained 31 signals of which about half were very strong and half considerably weaker. The more weakly hybridising clones were found to contain the 500-bp internal *Eco*RI fragment of the *copia* element. The stronger signals were all due to the same insert of ~8.5 kb containing one arm of *copia* attached to ~6 kb of flanking sequence. As shown by the restriction map determined later, the other *copia* arm and flanking sequence could not have been obtained in this experiment because the resulting *Eco*RI fragment is too large to fit in the λ641 vector.

We conclude that in the *gt* w^a strain the 3C region contains only one *copia* insertion which yields clonable, flanking *Eco*RI fragments. The deletion mapping experiments described below showed that this *copia* site corresponds to the *white* locus.

Chromosomal walk

Clones obtained from the microdissection experiments were used to screen genomic clone libraries to isolate a continuous set of overlapping clones. We used both phage and cosmid libraries constructed by inserting partial *Sau*3A digestion products into the λEMBL4 vector (Murray, 1983) or the cos4 tetracosmid vector (Reedy and Lehrach, unpublished data). In this article we will describe a set of clones which includes the *white* locus. We cut the w^a microdissection clone containing *copia* with *Eco*RI plus *Sac*I and isolated a non-repetitive flanking fragment with which to probe the genomic libraries. The clones obtained were used to walk further until the cloned region overlapped with clusters of clones isolated with other probes from the microdissection minilibrary.

The resulting clones, summarised in Figure 1, represent a continuous region of close to 200 kb extending ~110 kb to the left (distal) of the w^a insertion site and ~75 kb to the right (proximal) of it. Part of this region has also been cloned by Levis *et al.* (1982a) and by Goldberg *et al.* (1982). Our restriction map agrees for the most part with theirs with a few exceptions. Some of these are single restriction site heterogeneities which do not change the map relationships. They are mostly due to the fact that our clone libraries were constructed from an Oregon-R strain instead of a Canton-S strain. In other cases, however, the clones isolated from our genomic libraries represented arrangements different from those found in Canton-S. For example, clones isolated using probes from the −10 to −15 region contained a common sequence, usually from −6 to −15, but then branched off into as many as four different sequences. Figure 2 shows that

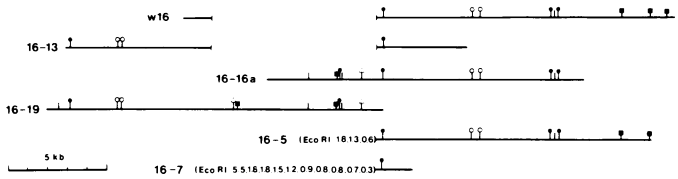

Fig. 2. Polymorphisms in the −15 region. The map shows the w16 clone and five clones obtained by screening the phage library with the left edge of w16. Clones 16-16a and 16-19 contain a repetitive element inserted in the site represented by the parentheses in clone w16 and 16-13. Clones 16-5 and 16-7 have a breakpoint at, apparently, the same site and continue with sequences unrelated to each other and with the preceding clones. The *Eco*RI fragments in the last two clones were not mapped. *Eco*RI: 1, *Bam*HI: ○, *Hind*III: ●, *Sal*: ■, *Bgl*II: △.

three of these clones, analysed by restriction mapping and cross-hybridisation, are clearly related. When suitably aligned, two of them, 16-19 and 16-16A, taken together correspond to the DNA of the third, 16-13, plus an insertion of 9−10 kb. Clone 16-13 hybridises to the part of 16-19 left of the insertion and to the part of 16-16A to the right of the insertion. Two other clones begin in the region covered by clone 16 but then diverge into sequences unrelated to each other or to those represented by the 16-19, 16-16A and 16-13 group. This polymorphism recalls the findings reported by Tschudi *et al.* (1982) for the sequences flanking the insertion site of a B104 transposable element. In that work, the insertion of B104 in the 5S RNA gene cluster gave rise to deletion variants in the population, some of which had lost the insertion alone while others showed deletions of the B104 plus flanking regions.

Two other cases of heterogeneity in our Oregon-R library were found around position +27 and position +43 and corresponded to insertions of ~10 kb and 7 kb, respectively. In all three cases, the inserted DNA was highly repetitive as shown by hybridisation to a genomic Southern blot (Figure 3).

Because of the abundance of repetitive sequences encountered, screens of genomic libraries frequently turned up weakly hybridising clones which bore no resemblance to the adjoining clones either by restriction mapping or by homology with flanking sequences. These were usually due to short regions of homology (<1 kb) embedded in foreign sequences. To guard against such false walking steps, we used several independent criteria. Restriction fragments were tested by hybridising them with total nick-translated *Drosophila* DNA. Fragments which gave appreciable signals were considered repetitive and not used for further screening. At each walking step we examined a larger number of clones and demanded overlaps of more than one restriction fragment with previously secured clones. Selected clones were hybridised *in situ* to polytene chromosomes to check that we had not strayed out of the 3B-3C region. The correctness of the walk was confirmed in some cases by linking up with clones isolated independently by screening genomic libraries with different microdissection clones. In other cases the genomic clones were hybridised back to the microdissection mini-library to verify that they contained sequences coming from the microdissected region. Finally, hybridisation to genomic DNA containing genetically mapped rearrangements of the 3B-3C region confirmed the authenticity of the walk.

Deletion mapping

In(1)w^{m4} is an inversion with breakpoint at 3C1-2 which

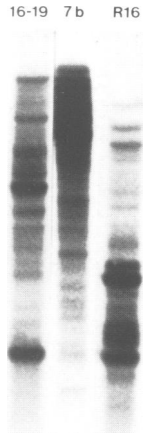

Fig. 3. Genomic Southern blots hybridised with clones corresponding to the three insertions.

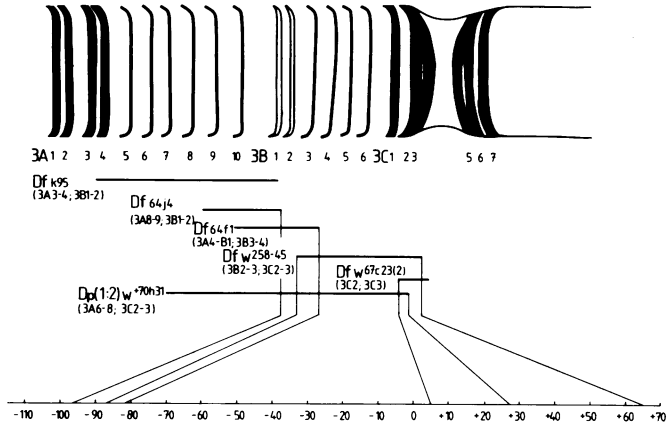

Fig. 4. Schematic summary of the chromosome region and the rearrangements examined. The bars indicate the approximate extent of the deletions or duplications with respect to the cytological map. The breakpoints identified are located in the map of the cloned region where 0 is the w^a insertion site, the positive direction is proximal and the negative direction distal.

places the *w* locus at the base of the X chromosome, near the *bb* locus. *In situ* hybridisation to polytene chromosomes of this strain showed that the cosmid clones in the region −30 to −105 lie more distal than 3C1-2 while the region 0 to +10 lies on the proximal side of the breakpoint. This orientation was confirmed and the localisation and extent of the chromosomal walk were determined by genomic Southern blots of DNAs carrying various deletions affecting the 3B-3C region. The results of this analysis are summarized in Figure 4.

Df(1)w$^{67c23(2)}$ is a small, male viable deletion which causes a white eye phenotype and, according to Lefevre and Green (1972), removes most but not all of the 3C2,3 doublet band. Hybridisation to the *w$^{67c23(2)}$* DNA with the w18 clone showed that all but the most proximal *Bam*HI fragment of this clone, at position +1.5 to +4.8, are still intact. Figure 5a shows the hybridisation of clone w11 to *w$^{67c23(2)}$* DNA cut with *Hind*III. Only one very large fragment hybridises, indicating that the deletion breakpoint is located between the *Bam* and *Hind*III sites at positions +1.5 and +3.5, respectively. Clones further to the right, including our most proximal clone, R21-8, give no hybridisation at all to this DNA, indicating that 75 kb to the right of the w^a insertion are still within the 3C3 band.

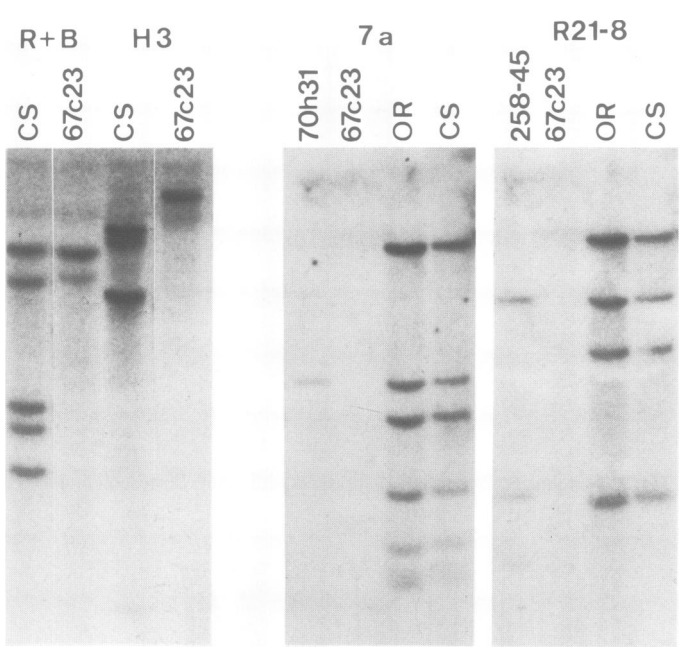

Fig. 5. Genomic Southern blots of strains carrying deletions. **(a)** DNA of Canton-S and *Df(1)w*$^{67c23(2)}$ flies digested with *Eco*RI + *Bam* or with *Hind*III, hybridised with labelled DNA of clone w11. **(b)** DNA of *Df(1)w*$^{258-45}$*/Dp (1;2) w*$^{+70h31}$, *Df(1)*$^{67c23(2)}$, Oregon-R and Canton-S flies cut with *Eco*RI + *Bam* and hybridised with labelled DNA of clone 7a. **(c)** The same DNAs cut with *Eco*RI and hybridised with labelled DNA of clone R21-8.

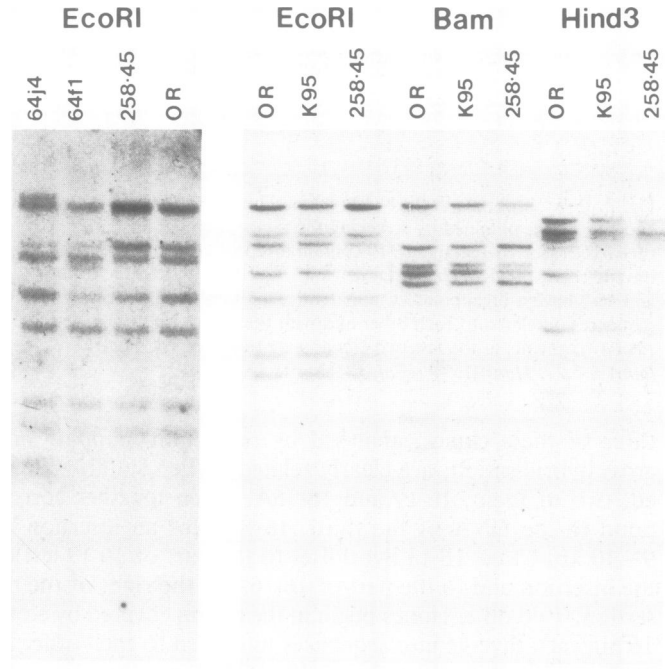

Fig. 6. Genomic Southern blots of strains carrying deletions. Genomic DNA from the different strains were cut with the enzymes indicated and hybridised with labelled DNA from clone a47. Bands which are included in a deletion appear relatively weaker than corresponding bands in the wild-type control (Oregon-R). For example, in the *Eco*RI digest of *Df(1)w*$^{258-45}$, all the bands except the top two are included in the deletion. The new band containing the breakpoint in this case is very small and weak and not visible in this picture (see, however, Figure 5c).

Two more breakpoints were localised in the region to the right of w^a. In the experiment shown in Figure 5C, DNA from *Df(1)258-45* (3B2-3; 3C2-3), cut with *Eco*RI + *Bam*, was hybridised with clone R21-8. Only the two most proximal fragments of that clone hybridise indicating that the proximal breakpoint of the w*258-45* deletion is near position +65. In the stock we used, the *w*$^{258-45}$ deletion was balanced by *Dp(1;2)w*^{70h31}, a duplication which covers approximately the same region. The genomic hybridisation shows, however, that the duplication falls somewhat short of the *w258-45* deletion, hence the missing bands in Figure 5c. The breakpoint of *Dp(1;2)w*^{70h31} was revealed by hybridising the same filter with clone 7a. Figure 5b shows that only the most distal of the *Eco*RI + *Bam* fragments of clone 7a is present in *Dp (1;2)w*^{70h31}, locating its breakpoint around position +28. The region between +28 and +65 is totally absent in this stock but, as shown also by the *w*$^{67c23(2)}$ deletion, this region is non-essential.

On the distal side we used a series of deletions with breakpoints in the 3B region (Judd *et al.*, 1972; Young and Judd, 1978). Since deletions in this region generally remove essential loci, they were balanced by the w^+Y duplication and the DNA was extracted from male flies. In genomic Southern blots of such DNAs, we expect that restriction fragments included in the deleted region would give relatively weaker hybridisation since they are contributed only by the balancing duplication. Fragments lying outside the deletion are contributed by both chromosomes and give a stronger signal. Fragments which cross the deletion breakpoint should give two weak hybridisation bands: a normal one contributed by the balancer, and one of unknown size corresponding to the new arrangement.

Figure 6 shows that the a47 cosmid clone contains several breakpoints in the 3B region. Hybridisation of a47 to the DNA of *Df(1)64j4* shows that the two most distal *Eco*RI fragments are weaker than expected and a new, weak band appears, corresponding to the breakpoint which must be located, therefore, within the second most distal *Eco*RI fragment, between −89 and −98. *Df(1)64f1* (3A9-B1; 3B3-4) has a more proximal breakpoint and decreases the hybridisation of the next three *Eco*RI fragments in the a47 map. *Df(1)w*$^{258-45}$ (3B2-3; 3C2-3) comes from the opposite direction. In this case, the two most distal *Eco*RI fragments are stronger and the proximal ones weaker. Using other restriction enzymes, the location of the breakpoints was mapped to shorter intervals. We did not detect any differences with DNA from *Df(1)K95* (3A3-4; 3B1-2) whose breakpoint presumably lies beyond our most distal clone. We have not yet mapped deletion breakpoints in the region between *white* and the a47 cosmid. This region contains repetitive sequences which prevent deletion mapping by this rapid method.

Repetitive and repetitious sequences

We originally used hybridisation with total labelled *Drosophila* DNA to Southern blots of the genomic clones as a diagnostic test to identify repetitive sequences. We found that a large proportion of the restriction fragments, including small ones <1 kb, gave a signal out of proportion to their size. Some, for example a 0.6-kb *Bam-Eco*RI fragment around position −22, give an extremely strong signal (Figure 8). We can distinguish at least two kinds of repetitive sequences detected by this test. One kind, such as dispersed gene families or transposable elements, has an extensive and conserved region of homology. Repetitive elements of this kind are found in clones from the region −60 to −20 which give a pattern of multiple bands when hybridised to a

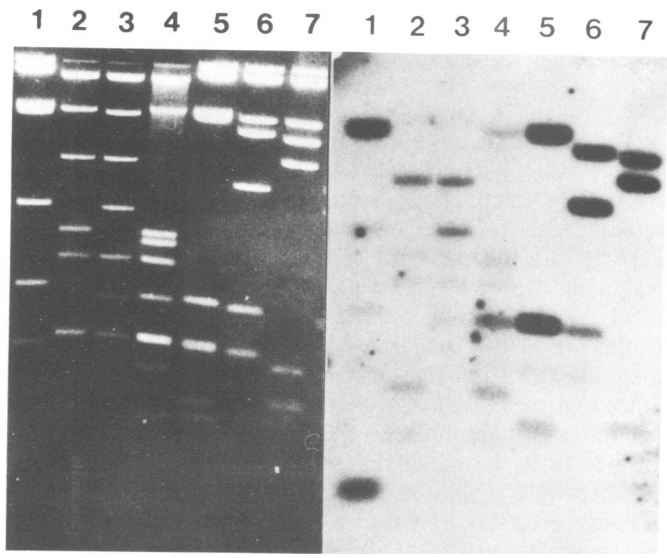

Fig. 7. Labelled *Drosophila* DNA hybridised to clones from the *white* region. The **left panel** shows the gel stained with ethidium bromide, the **right panel** shows the autoradiograph of the Southern blot. **(1)** clone 16-13 cut with *Eco*RI. **(2)** clone w16 cut with *Sal* + *Bam* + *Kpn*. **(3)** clone w16 cut with *Eco*RI + *Bam* + *Kpn*. **(4)** clone w18 cut with *Sal* + *Bam* + *Kpn*. **(5)** clone w11 cut with *Eco*RI + *Sal*. **(6)** clone 18-13 cut with *Eco*RI + *Sal*. **(7)** clone 18-13 cut with *Eco*RI + *Bam*.

genomic Southern blot. Figure 3 shows that the insertional polymorphisms in our Oregon-R clone library are also due to repetitive elements of this sort. Clone 7b, for example, which corresponds to an insertion at position +27, gives a highly repetitive pattern when hybridised to a genomic Southern blot while clone 7a, which corresponds to the empty site, hybridises like a unique sequence (Figure 5b).

Another kind of sequence might be called repetitious rather than repetitive. Such sequences are either weakly homologous or short, but highly abundant and dispersed in the genome. To this class belong sequences in the region −60 to −110 or −20 to +20. Clones containing these sequences give only the expected bands when hybridised to genomic Southern blots since the less homologous sites contribute only a weak background. However, when total labelled *Drosophila* DNA is hybridised to the cloned sequences, the additive contribution from many genomic sites is stable and resistant to stringent washing conditions (several hours at 67°C, 0.1 x SSC).

We analysed in more detail the distribution of repetitious sequences in the region surrounding the w^a insertion site. Figure 7 shows that many of the fragments in this region give detectable signals summarized in the map in Figure 9. Notably free of repetitious sequences are the *Sal-Sal* fragment around −2 and the *Eco*RI-*Bam* region at −7.5 to −10. In the region to the right of w^a we find not only an abundance of repetitive sequences but also a degree of internal homology. Part of clone 18-13-3, for example, in the region +26 to +30, cross-hybridises weakly with fragments of clones R16 and R21 in the region +43 to +49.

Transcription of the 3B region

The 3C2-3B region contains a number of genes necessary for normal development (Judd *et al.*, 1972) as well as loci for affecting fertility (Thierry-Mieg, 1982), periodicity (Smith and Konopka, 1981) and the *white* gene itself. We looked at the transcriptional activity of this region by hybridising Southern blots of the clones to cDNA from early pupae.

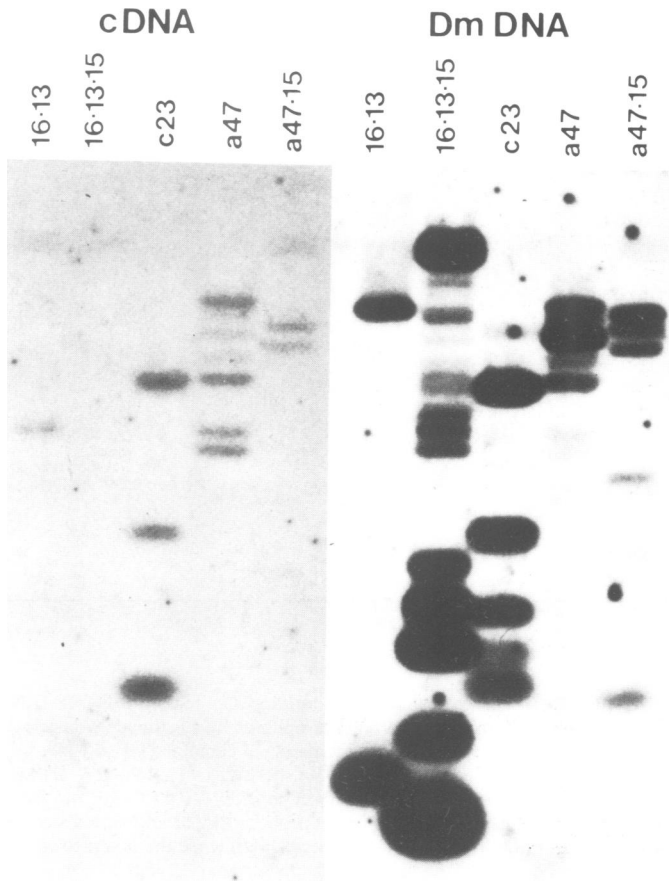

Fig. 8. Transcription and repetitivity in the region to the left of *white*. Clones representing the region left of *white* were cut with *Eco*RI, run on duplicate gels and blotted. One filter **(left)** was hybridised with cDNA made from early pupae. The other filter **(right)** was hybridised with labelled *Drosophila* DNA to display repetitive sequences. Note the very strongly hybridising band of ~300 bp in clone 16-13-15 DNA. This sequence hybridises so strongly that partial digestion products not visible in the stained gel are responsible for several of the other bands visible in this lane.

Nearly every *Eco*RI fragment in the region −70 to −110 hybridises to the cDNA (Figure 8). Although some of these fragments contain repetitious sequences by the criteria discussed in the previous section, they are individually too weakly homologous to be detectable in genomic blots. We suppose, therefore, that the hybridisation to cDNA may reflect real transcription of the *zw3, zw6* and *zw12* loci which, according to the genetic data of Judd *et al.* (1972), map in this region.

Three *Eco*RI fragments from cosmid c23 hybridise to the cDNA while the region −40 to −10 is silent despite the presence of very highly repetitive sequences.

Transcription of the white region

The w^a insertion roughly divides the *white* locus into proximal and distal regions. The proximal region is believed to be the control region (Judd, 1976; Zachar and Bingham, 1982) while the distal region is the candidate for the structural part of the gene. Rearrangements that affect the function of the *white* locus have been mapped by Levis *et al.* (1982b) between positions +10.6 and −3.2. We first looked for transcripts from the region from −15 to +15 by hybridising blots of the various clones with cDNA made from larvae or pupae. Signals were obtained with the *Bam-Kpn* fragment from −7 to −10 and the *Sal-Sal* and *Sal-Bam* fragments in the region

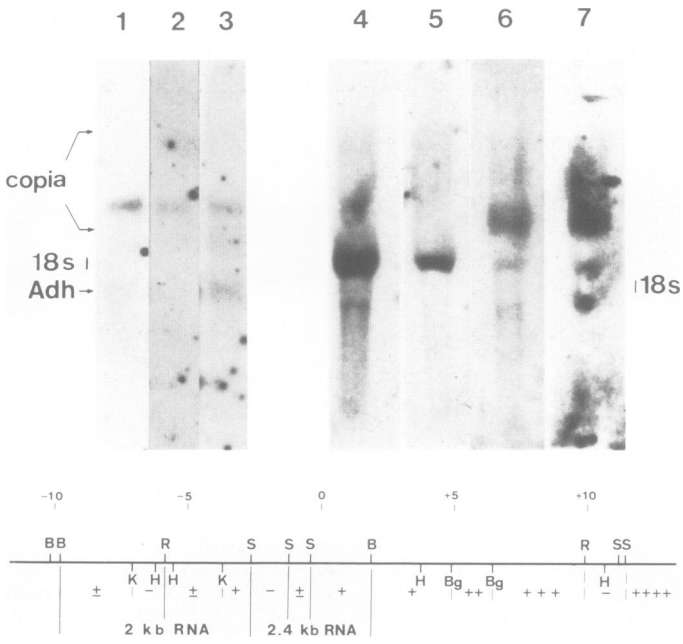

Fig. 9. Northern blots of transcripts from the *white* region. Poly(A)$^+$ RNA (still containing up to 50% rRNA) electrophoresed in formaldehyde gels was blotted on nitrocellulose filters. **Lanes 1–3** contained larval RNA hybridised with the fragments: **(1)** *Sal-Sal*, position −1.2 to −2.6. **(2)** *Sal-Bam*, position +1.9 to −1.4. **(3)** *Sal-Sal*, position −0.4 to −1.2. In **lanes 1** and **3**, *Adh* sequences included in the probe provided an internal size marker (*Adh* mRNA = 1.1 kb; Goldberg, 1980) while the two arrows point to the position of the major *copia* transcripts (~2 and 5 kb), detected in a parallel lane. **Lanes 4–7** contained RNA from adult flies hybridised with the following fragments: **lane 4**, *Bam-Eco*RI, position −9.8 to −6.9, **lane 5**, *Eco*RI-*Sal*, position −6.9 to −2.6, **lane 6**, *Sal-Sal*, position −2.6 to −1.2, **lane 7**, *Sal-Bam*, position −1.2 to +1.9. The map below summarizes the results of the Northern blots and of the hybridisation with labelled *Drosophila* DNA such as shown in Figure 7. Intervals marked ± gave weak but detectable hybridisation.

−3 to +1.6 but not (or more weakly) to the *Kpn-Kpn* and *Kpn-Sal* fragments (see map in Figure 9). Very faint hybridisation to the cDNA could also be detected with the region +10 to +15 and again to more proximal fragments, but no known genes map in this interval and, because of the presence of repetitious sequences, we cannot be sure that these transcripts arise from this region.

To analyse the transcripts from the *white* region, we hybridised Northern blots of larval, pupal and adult fly RNA to different probes from the +1.5 to −10 region. The RNA which comes from this region is clearly a very rare one and could not be detected with total RNA even when 40 μg were loaded per track. However, using 10–25 μg of poly(A)$^+$ RNA, we found a transcript of ~2.4 kb which hybridises to the *Bam-Sal* 2-kb fragment as well as to the 0.8-kb and the 1.4-kb *Sal-Sal* fragments (Figure 9). Roughly equal signals were obtained with early larval, late larval, pupal and fly RNAs. A different RNA species of ~2 kb hybridises with the fragments immediately to the left. Both the *Eco*RI-*Sal* fragment at position −3 to −6 and the *Bam-Eco*RI fragment from −6 to −10 reveal this second transcript which is at least five times more abundant than the first. Of the two, we suppose the −3 to +1.5 region to be the more likely candidate for the *white* transcript since it is closest to the w^a insertion, to the $w^{67c23(2)}$ deletion which presumably removes the regulatory part of the gene, and it includes the majority of the *white* mutations mapped by Rubin *et al.* (1982), Goldberg *et al.* (1982) and Zachar and Bingham (1982).

Discussion

The microcloning technique yields a minilibrary of clones which, within the limits imposed by the cloning vector, gives a reasonably complete representation of the microdissected region. With just three dissected fragments, we obtained, on average, a 2-fold representation of each clonable restriction fragment. An even higher efficiency was obtained in the second microcloning experiment in which eight chromosome fragments yielded the desired sequence 15 times. This experiment illustrates one limitation of the microcloning approach. Since at present high efficiency can be obtained only by cloning restriction fragments, regions containing fragments too large to fit in the vector remain unrepresented in a single experiment and would have to be looked for using different cloning vectors and restriction enzymes.

The same experiment illustrates the power of the microcloning approach for the isolation of loci which can be molecularly tagged by the insertion of a known transposable element. Rubin *et al.* (1982) have recently identified a transposable element called P which, in appropriate 'dysgenic' crosses, can be mobilised, generating insertional mutations. The desired sequence can in principle be isolated from a genomic library made from the dysgenic mutant by isolating all the P-containing clones and screening by *in situ* hybridisation until the clone coming from the desired locus is identified. On the other hand, microcloning from an insertional mutant allows direct isolation of the clone derived from the desired locus. The minilibrary of microdissection clones serves as a convenient source of multiple entry points for chromosomal walks and as an independent reference to check that the walk has not strayed out of the region of interest and into other chromosomal sites. This is a serious danger due to the frequency with which repetitive or repetitious sequences are encountered.

The region we have cloned contains much repetitive material. Some of these sequences are large and conserved in the genome and give a discrete pattern of bands when hybridised to a genomic Southern blot. Sequences of this sort were associated with the three cases of insertional polymorphisms encountered in the Oregon-R population from which our genomic libraries were constructed. Variants corresponding to the presence or absence of the repetitive element indicate that these sequences are highly mobile.

More puzzling are the repetitious sequences detected by hybridisation to total labelled *Drosophila* DNA. They could be due to trivial homologies, such as local clusters of very high GC content, or they could be due to short but widely spread conserved sequences such as those reported by Wensink *et al.* (1979). One interesting class of short repetitive sequences one might expect to find is that of control elements responsible for dosage compensation of X chromosome genes.

Our chromosomal walk can be roughly divided into two halves. The proximal half starts in a region of very dense bands (3C1-3) where progress from one band to the next requires crossing 50–100 kb of genomic DNA. On the distal side, however, we move into the 3B region which is much less condensed and contains a number of very faint bands. In this region, as shown by our deletion mapping, bands contain 10 kb or less of genomic DNA.

Sorsa *et al.* (1973) have positioned the *w* locus at the distal edge of band 3C2, perhaps in a very faint band situated between 3C1 and 3C2. The region proximal to *w* contains no known genes until the *roughest* and *verticals* loci which are usually mapped in 3C5-6. The region in between, 3C3-3C5, has a number of unusual characteristics because of which it was sometimes said to contain intercalary heterochromatin (reviewed by Zhimulev *et al.*, 1981). It corresponds to a chromosome constriction which is frequently involved in ectopic pairing and chromosome breaks, its DNA is underreplicated in polytene chromosomes and contains highly repetitious sequences. This region and a few others which share its properties might play a role in initiating chromosome pairing. It is possible that our most proximal clones might be entering this region.

On the distal side of the *white* gene, a number of genetic loci can be assigned positions in our map. Judd *et al.* (1972) have identified four complementation groups defining loci *zw9, zw11, zw5* and *zw7* which lie between *w* and the *64f1* deletion or between position −5 and −82 in our map. Mutations in these loci are generally recessive lethal except for *zw9* and one allele in *zw5* which are semilethal with defects in wings or appendages and sterility. The right borders of *Df(1)64f1* and of *Df(1)w*$^{258\text{-}45}$ define an interval (positions −82 to −89) within which map the loci *zw12* and *par*. *zw12* mutations are recessive lethal while the *par* locus is defined by a single mutation with a bewildering range of pleiotropic effects (Thierry-Mieg, 1982). Homozygous *par* flies are female sterile at 29°C while at lower temperatures they lack germ cells and have defective cuticular patterns. The mutation interacts with other loci in the 3A-3C region including the *zeste* locus.

The interval between the *258-45* and *64j4* breakpoints, corresponding to positions −85 to −99, contains *zw6* whose alleles are recessive lethal. Beyond, but in the close vicinity of the *64j4* breakpoint, are the *zw3* and the *per* locus (Young and Judd, 1978; Smith and Konopka, 1981). Mutations in the latter locus alter the period of circadian activity in the fly or cause arrhythmic activity.

We have not attempted to characterise in any detail the transcription of the genes contained in the cloned region except for the *white* locus and its vicinity. The *white* gene is active throughout the lifetime at least from early larvae to adult flies. It is transcribed into a very rare poly(A)$^+$ RNA of ~2.4 kb, most or all of which comes from the region between positions +1.5 and −3. Although the entire transcript could fit in this region, we cannot exclude the existence of introns. Transcription might, for example, initiate in the *white* proximal region. This would account for the existence of regulatory mutations in the region +3 to +10 and for the effect of the *w*$^{67c23(2)}$ deletion which inactivates the gene by deleting the *white* proximal region. It is interesting to note that, in the *w*^{a} mutant, a *copia* element is inserted in the *Bam-Sal* fragment (positions +1.5 to −0.5) within 0.3−0.4 kb of the *Sal* site. Preliminary S1 mapping experiments indicate that the insertion is within the transcribed region and that the direction of transcription is proximal to distal in accordance with the genetic data.

The identity of the second transcript is still not clear. It might conceivably still belong to the *w* locus: one *w* allele, w^{Bwx} has been mapped considerably more distal than the others. However, no other *w* mutations have been mapped more distal than position −3.0 and duplications of the *w* locus, apparently able to confer full activity, have a breakpoint around the same position (Levis *et al.*, 1982a; Zachar and Bingham, 1982). A more detailed analysis of these transcripts, of their relationship and of the effect of *zeste* mutations on them is under way. The region surrounding the *zeste* locus has also been cloned (Mariani and Pirrotta, unpublished data) and should enable us to study the molecular basis for the *zeste* effect.

Materials and methods

Drosophila strains

Oregon-R Heidelberg. Our stock of a standard wild-type strain. It was used to prepare the genomic libraries.

Canton-S. A wild-type strain which has a number of restriction site differences in the *white* region, with respect to Oregon-R.

y sc gt^{X11}/FM6 and *gt*: used to prepare *gt*^{X11}*gt* giant larvae.

gt^{13z} *gt w*^{a}: used to microdissect the *w*^{a} locus.

In(1)w^{m4}: carries an inversion with breakpoints 3C1-2; 20A which places the *w* locus near *bb* and gives position effect variegation for *w* (Cooper, 1959).

Df(1) w$^{67c23(2)}$ *sn*^{3}: carries a small homozygous viable deletion within the 3C2-3 doublet and has a white eye phenotype (Lefevre and Green, 1972).

C(1) DX ywf(Df(1) w$^{258\text{-}45}$*sn*^{3}*; Dp(1;2)w*$^{+70h31}$: carries the *w*$^{258\text{-}45}$ deletion 3B2-3;3C2-3, balanced in the male by the *w*$^{+70h31}$ duplication (3A6-8; 3C2-3;31) (Lefevre and Wilkins, 1966; Judd *et al.*, 1972). Only males were used to isolate DNA from this and the following three strains.

C(1) DX yf/Df(1)64j4, y/w$^+$*Y*: carries the *64j4* deletion 3A9-3B1; 3B3-4 (Judd *et al.*, 1972).

C(1) DX yf/Df(1)64f1/w$^+$*Y*: carries the *64f1* deletion 3A9-3B1;3B3-4 (Judd *et al.*, 1972).

C(1) DX yf/Df(1) K95,y/w$^+$*Y*: carries the *K95* deletion 3A3-4; 3B1-2 (Judd *et al.*, 1972).

Microdissection and microcloning

The details of the procedure are described by Pirrotta *et al.* (1983). Salivary glands, exposed to 45% acetic acid for a maximum of 1.5 min, were squashed on 24 x 60 mm unsubbed coverslips with a 20 x 20 mm siliconised coverslip, frozen in liquid nitrogen and the siliconised coverslip flicked off with a razor blade. The squash was washed in absolute ethanol and air dried before placing in the oil chamber for dissection. Chromosome fragments were dissected using a de Fonbrune micromanipulator and a glass needle with the tip bent at a 40° angle. The fragments were deposited in a 0.5 nl aqueous droplet containing 0.5 mg/ml proteinase K and 0.1% SDS. The droplet was extracted three times with four volumes of water saturated phenol and then with $CHCl_3$. *Eco*RI (0.2 nl at 120 units/μl) containing 3 x *Eco*RI buffer (50 mM Tris pH 7.5, 50 mM NaCl, 10 mM $MgCl_2$ and 5 mM dithiothreitol) was added and the entire oil chamber incubated at 37°C for 2 h. The enzyme was inactivated by placing the oil chamber in a moist Petri dish at 70°C for 20 min. After cooling, 0.5 nl of *Eco*RI cut λ641 DNA (100 μg/ml DNA, 1 mM ATP in *Eco*RI buffer) were added, followed by 0.2 μl of T4 ligase (1−4 units/μl). After ligation overnight at 5°C, the products were recovered, diluted to 2 μl and packaged *in vitro* as described by Scherer *et al.* (1981).

Genomic libraries

These were made from partial *Sau*3A digests of embryonic Oregon-R DNA. The digests were fractionated on a NaCl gradient and fractions corresponding to 40−50 kb were used for cosmid cloning while the 18−24 kb range was used for λ cloning. The λ vector used was λEMBL4 (Frischauf *et al.*, in preparation; described by Murray, 1983), an extremely convenient cloning vector with excellent growing ability which allows *spi* selection of recombinants. This vector accepts 7−22 kb inserts which can then be excised using the *Eco*RI sites which flank the *Bam* sites used for insertion. The λ vector was cut with *Bam*, heated to 70°C for 10 min to inactivate the enzyme and ligated to a 5-fold smaller weight of size-selected *Sau*3A partial products. The ligation mixture was packaged *in vitro* (Scherer *et al.*, 1982) and plated on Q359, a strain lysogenic for P2 which selects the recombinant phages. About 3 x 10^5 recombinants were obtained per μg of λEMBL4 DNA.

The cosmid vector cos4, constructed by Reedy and Lehrach, contains two sets of double *cos* sites in tandem orientation. A *Pvu*II site separates the two sets and allows linearisation of the vector. This vector allows a more convenient and efficient version of the phosphatase procedure of Ish-Horowicz and Burke (1981). The *Pvu*II ends were treated with phosphatase to prevent religation and the vector was cut with *Bam* to provide an insertion site for partial *Sau*3A fragments.

Phage libraries were screened by the method of Benton and Davis (1977).

Cosmid libraries were screened by the method of Hanahan and Meselson (1980) with some modifications.

Drosophila RNA

RNA was isolated from larvae, pupae or adult flies. Several methods were tried but the best results were obtained by homogenising the material, still frozen at −70°C, with a Polytron homogeniser in 100 mM sodium cacodylate, pH 7.5, 0.5 M NaCl, 2% diethyl pyrocarbonate and 2% SDS. The homogenate was centrifuged for 10 min at 10 000 r.p.m. then extracted three times with phenol:chloroform:isoamyl alcohol (25:25:1) and finally precipitated with 3 volumes ethanol. Poly(A)$^+$ RNA was prepared by passing the RNA through a poly(U)-Sephadex column in buffer containing 10 mM Tris pH 7.5, 10 mM EDTA, 0.2% SDS and 0.2 M NaCl. The column was washed extensively with the same buffer minus salt and eluted with 10 mM Tris pH 7.5, 10 mM EDTA and 50% formamide (recrystallised). After one such passage, the RNA frequently still contained up to 50% rRNA.

Drosophila DNA

DNA for genomic analysis was extracted from flies by gentle homogenisation with a few pestle strokes in a Dounce homogeniser in buffer containing 0.1 M Tris, 0.1 M EDTA, 1% SDS, 0.2% diethyl pyrocarbonate, pH 9. The extract was incubated at 65°C for 30 min, then made 1 M potassium acetate and chilled in ice for 30 min. The precipitate was pelleted and the supernatant was precipitated with 0.5 volumes isopropanol. The DNA pellet was washed twice with 70% ethanol, dried briefly and resuspended in 10 mM Tris, pH 7.5, 1 mM EDTA.

Restriction fragments

Labelled fragments for screening genomic libraries or for restriction mapping were prepared by cutting cloned DNA with a suitable restriction enzyme and labelling the ends by filling in with DNA polymerase. The fragments were separated on an agarose gel and the desired bands were cut out, boiled to melt the agarose and denature the DNA and added to the hybridisation mixture.

For nick-translation or for subcloning, fragments were purified from a low melting agarose gel by the method of Tautz and Renz (1983).

Southern and Northern hybridisation

Southern blots were made from agarose gels according to Southern (1975). After baking, the nitrocellulose filters were prehybridised with 2 x SSC, 10 x Denhardt's solution (Denhardt, 1966) at 67°C and then hybridised with labelled probes under the same conditions. For blots of genomic DNA the probe generally contained 10% dextran sulfate and 100 μg/ml denatured salmon sperm DNA. Hybridisation of clones with total genomic DNA, labelled by nick-translation was carried out under the same conditions in the presence of 10% dextran sulfate and salmon sperm DNA.

For Northern blots, RNA was denatured in 2 M formaldehyde, 50% deionised formamide at 65°C for 5 min before loading on 1% agarose gels containing 2 M formaldehyde 0.2 M morpholinopropanesulfonic acid, 50 mM Na acetate. After electrophoresis, the gel was soaked in 20 x SSC for 30 min with agitation before blotting on nitrocellulose filters. Hybridisation was carried out in 50% formamide with or without 10% dextran sulfate, 10 x Denhardt's solution, 100 μg/ml denatured salmon sperm DNA in 2 x SSC at 43°C.

Acknowledgements

We are grateful to B.H.Judd, M.M.Green and M.Gans for providing fly stocks, to N.Murray and H.Lehrach for the λ EMBL4 and cos4 vectors, to Christa Bröckl for excellent technical assistance and to Wendy Moses for typing the manuscript. C.Hadfield was supported by EMBO and EMBL fellowships and G.Pretorius by a stipend from the Medical Research Council of the Union of S.Africa.

References

Benton,W.D. and Davis,R.W. (1977) *Science (Wash.),* **196**, 180-192.
Bingham,P.M. and Judd,B.H. (1981) *Cell,* **25**, 705-711.
Collins,M. and Rubin,G.M. (1982) *Cell,* **30**, 71-79.
Cooper,K.W. (1959) *Chromosoma,* **10**, 535-588.
Denhardt,D.T. (1966) *Biochem. Biophys. Res. Commun.,* **23**, 641-646.
Gehring,W.J. and Paro,R. (1980) *Cell,* **19**, 897-904.
Gelbart,W.M. and Wu,C. (1982) *Genetics,* **102**, 179-189.
Goldberg,D.A. (1980) *Proc. Natl. Acad. Sci. USA,* **77**, 5794-5798.
Goldberg,M.L., Paro,R. and Gehring,W.J. (1982) *EMBO J.,* **1**, 93-98.
Hanahan,D. and Meselson,M. (1980) *Gene,* **10**, 63-67.
Ish-Horowicz,D. and Burke,J.F. (1981) *Nucleic Acids Res.,* **9**, 2989-2998.
Jack,J.W. and Judd,B.H. (1979) *Proc. Natl. Acad. Sci. USA,* **76**, 1368-1372.
Judd,B.H. (1976) in Ashburner,M. and Novitski,E. (eds.), *The Genetics and Biology of Drosophila,* Academic Press, NY, pp. 767-799.
Judd,B.H., Shen,M.W. and Kaufman,T.C. (1972) *Genetics,* **71**, 139-156.
Lefevre,G. and Green,M.M. (1972) *Chromosoma,* **36**, 391-412.
Lefevre,G. and Wilkins,M.D. (1966) *Genetics,* **53**, 175-187.
Levis,R., Bingham,P.M. and Rubin,G.M. (1982a) *Proc. Natl. Acad. Sci. USA,* **79**, 564-568.
Levis,R., Collins,M. and Rubin,G.M. (1982b) *Cell,* **30**, 551-565.
Murray,N.E. (1983) in Hendrix,R.W., Roberts,J.W., Stahl,F. and Weisberg, R.A. (eds.), *Lambda II,* Cold Spring Harbor Laboratory Press, NY, in press.
Pirrotta,V., Jäckle,H. and Edström,J.E. (1983) in Hollaender,A. and Setlow, J.K. (eds.) *Genetic Engineering, Principles and Methods,* vol. **5**, Plenum Press, in press.
Rubin,G.M., Kidwell,M.G. and Bingham,P.M. (1982) *Cell,* **29**, 987-994.
Scalenghe,F., Turco,E., Edström,J.E., Pirrotta,V. and Melli,M. (1981) *Chromosoma,* **82**, 205-216.
Scherer,G., Telford,J., Baldari,C. and Pirrotta,V. (1981) *Dev. Biol.,* **86**, 438-447.
Smith,R.F. and Konopka,R.J. (1981) *Mol. Gen. Genet.,* **183**, 243-251.
Sorsa,V., Green,M.M. and Beermann,W. (1973) *Nature New Biol.,* **245**, 34-37.
Southern,E.M. (1975) *J. Mol. Biol.,* **98**, 503-517.
Tautz,D. and Renz,M. (1983) *Anal. Biochem.,* in press.
Thierry-Mieg,D. (1982) *Genetics,* **100**, 209-237.
Tschudi,C., Pirrotta,V. and Junakovic,N. (1982) *EMBO J.,* **1**, 977-985.
Wensink,P.C., Tabata,S. and Pachl,S. (1979) *Cell,* **18**, 1231-1246.
Young,M.W. and Judd,B.H. (1978) *Genetics,* **88**, 723-742.
Zachar,Z. and Bingham,P.M. (1982) *Cell,* **30**, 529-541.
Zhimulev,I.F., Belyaeva,E.S. and Semeshin,V.F. (1981) *CRC Crit. Rev. Biochem.,* **11**, 303-340.