
Site specific insertion of a type I rDNA element into a unique sequence in the *Drosophila melanogaster* genome

Michael J. Browne¹, Christopher A. Read², Heli Roiha³ and David M. Glover

Cancer Research Campaign, Eukaryotic Molecular Genetics Research Group, Department of Biochemistry, Imperial College of Science and Technology, London SW7 2AZ, UK

Received 11 October 1984; Revised and Accepted 8 November 1984

ABSTRACT

We describe a cloned segment of unique DNA from the Oregon R strain of *Drosophila melanogaster* that contains a short type I insertion of the kind principally found within rDNA. The predominant type I rDNA insertion is 5kb in length, but there are also a co-terminal sub-set of shorter type I elements that share a common right hand junction with the rDNA. The insertion that we now describe is another member of this sub-set. The right hand junction of the type I sequence with the unique DNA is identical to the right hand junction of the type I sequences with rDNA. There is no significant feature within the insertion sequence that could have determined the position of the left junction with the sequence into which it is inserted. Like the corresponding short type I insertions in rDNA, the insertion into the unique DNA is flanked on both sides by a duplicated sequence, which in this case is 10 base pairs long. The cloning of a sequence corresponding to the uninterrupted unique location was facilitated by the observation that the Karsnas strain of *D. melanogaster* contains only uninterrupted sequences of this kind. The duplicated sequence at the target site for the insertion is only present as a single copy in the uninterrupted DNA. The sequence of the target site for the insertion (ACTGTCT) in the unique segment shows a striking homology to the target in rDNA (ACTGTCCC).

INTRODUCTION

The rDNA of *Drosophila melanogaster* can contain two types of insertion in its 28S rRNA genes [1-5]. The two types of insertion occur at different sites within the rDNA [6]. They differ in their degree of conservation between members of the melanogaster species sub-group of *Drosophila*; the restriction endonuclease cleavage profile of the type II sequences being highly conserved whereas the type I sequences show a much more heterogenous pattern [7]. The two types of insertion also differ in their chromosomal distribution within a species. The type II sequences occur in about 15% of the rDNA units on both the X and Y chromosomes of *D. melanogaster* [8]. The type I insertions on the other hand occur in about 50% of the rDNA units of the X

chromosome and not in the rDNA of the Y chromosome [8,9].

There is a considerable amount of circumstantial evidence that indicates that the type I insertion is capable of a restricted degree of transposition. The majority of the type I elements are found either in the rDNA mainly as 5Kb insertions, or in the flanking heterochromatin where they occur as tandemly arranged 4-5kb units [10]. In one stretch of tandemly arranged heterochromatic type I elements, each unit was shown to be flanked on the left by a 20 nucleotide sequence and on the right by a 14 nucleotide sequence identical to the sequences flanking the insertion in rDNA [6]. This suggests that the sequences were once present in the nucleolus and have transposed into the flanking heterochromatin and subsequently become tandemly reiterated. Within the rDNA there exist a set of shorter type I insertions up to 1Kb in length. These correspond to the extreme right hand sequence of the 5Kb insertion and form a co-terminal sub-set. These smaller insertions all occur at the same site in the rDNA and are flanked on both sides by a segment of rDNA present only once in the uninterrupted unit. The length of the duplicated sequence varies from 7-15 nucleotides [11,12]. The 5Kb type I insertion is not flanked by a duplicated sequence but on the left hand side there is a deletion of 9 nucleotides of rDNA [6]. Such duplications and deletions are found in the sequences flanking a variety of transposable elements from both prokaryotic and eukaryotic organisms.

Previous analyses of the type I elements in non-rDNA sequences have focussed upon the organisation of the tandem arrays in the heterochromatin. In this study we examine the organisation of a type I element that is inserted into a unique sequence within the *D. melanogaster* genome. We find a strong homology between the target site of the insertion of the type I element into this unique sequence and the target site in rDNA.

MATERIALS AND METHODS

Construction of Cloned DNAs

The recombinant λ MB8a was isolated from a partial library of size selected EcoRI fragments of *D. melanogaster* DNA cloned in λ gt.10WES. The construction of the library was described in Roiha and Glover [11]. The clones λ CR.H3 and λ CR.H6 were isolated from a library constructed by Dr. Kim Kaiser. This was made in λ EMBL4 using a size selected (approx. 15kb) partial Sau3A digest of DNA from the Karsnas strain of *D. melanogaster*.

DNA Preparation

DNA was extracted from flies as previously described by Roiha et al [7]. Plasmid DNA was prepared as previously described [6].

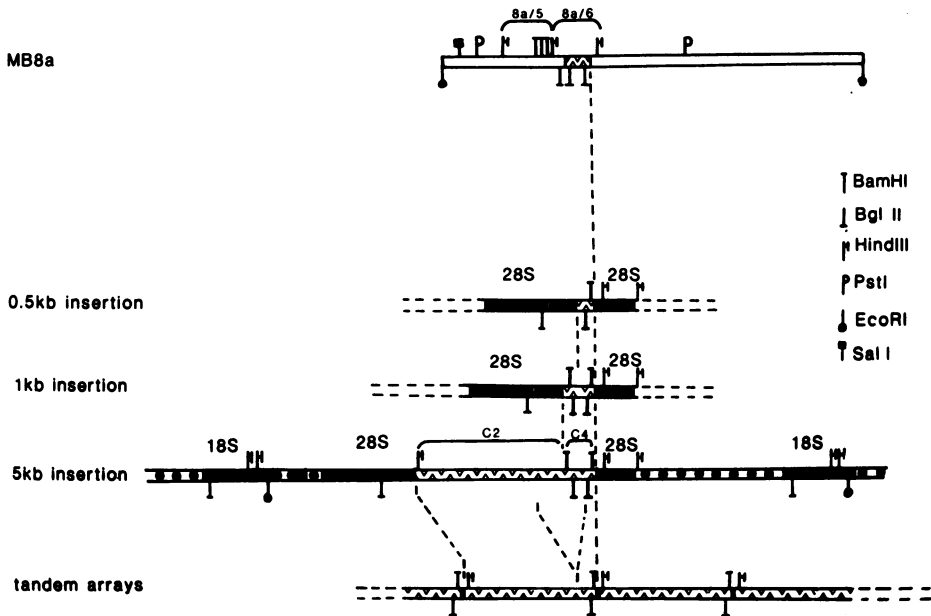


Figure 1. Physical Maps of rDNA and Non-rDNA Segments that Contain Type I Insertions. The uppermost map is of the cloned segment MB8a, described in this paper. The 12.5Kb EcoRI fragment was isolated from a partial library of *Drosophila melanogaster* EcoRI fragments, the construction of which was described in Roiha and Glover [11]. All cleavage sites for the enzymes SalI, PstI, BamHI and BglII are shown and were determined by standard double digestion techniques. The fragment has 8 cleavage sites for HindIII. Only those HindIII fragments containing BamHI sites (8a/5) or BglII sites (8a/6) have been mapped. The rDNA segments containing the 0.5Kb and 1Kb insertions correspond to the clones RI10 and MB27 described by Roiha and Glover [11]. The rDNA segment containing the 5Kb insertion is the 17Kb EcoRI fragment, Dm103 [1,13]. The sub-cloning of segments C2 and C4 from this clone have been described by Kidd and Glover [10] and Roiha et al [6]. The tandemly arranged type I sequences are a depiction of the clone Dm219 described by Kidd and Glover [10].

RESULTS

The Isolation of a Type I Element existing independently of rDNA and of other Type I Elements

The organisation of the most common type I insertions is shown in Figure 1. The major rDNA insertion is 5Kb length and is exemplified by the cloned segment Dm103 [13]. Tandem arrays of type I elements are roughly equally abundant and occur within the heterochromatin flanking the nucleolus [14,15]. Figure 1 shows

the arrangement of three tandemly repeating units which form part of the clone Dm209 [10]. The tandem arrays have diverged from the rDNA insertion and have deletions (as depicted in Figure 1) or insertions (not shown) as well as restriction site polymorphisms that probably reflect single base changes. Two shorter rDNA insertions of 1.0kb and 0.5kb are also depicted in this Figure. These are exemplified by the clones MB27 and R110 which represent a co-terminal sub-set of the 5Kb sequence [11]. Following our analysis of these shorter insertions we reasoned that they were probably derived from the 5Kb sequence and had become re-inserted into the rDNA at the same target site [11]. This resulted in a duplication of a 7 to 15 nucleotide rDNA segment flanking the insertion. We decided to search for similar elements derived from the right hand side of the 5Kb insertion in other regions of the genome. To this end we screened a partial library of *D. melanogaster* embryonic DNA for clones that would hybridise with the probe C4 (see Figure 1) but not with the probe C2 and not with an rDNA probe, pDm238 (a 12kb uninterrupted segment of rDNA). We isolated three clones which had this property and have characterised one of these, λ MB8a, in some detail.

A physical map of MB8a, the non-rDNA DNA segment containing a type I insertion is shown in Figure 1. It is a 12.5Kb EcoRI fragment that is cleaved at eight sites by HindIII. Two of the resulting HindIII fragments, 8a/5 and 8a/6 are shown on the physical map. The 12.5Kb EcoRI fragment has been re-cloned into pBR322 and Figure 2 shows digests of this recombinant plasmid, pMB8a, with a number of restriction endonucleases. The right hand panel of Figure 2 shows a Southern hybridisation of this gel probed with p225, which contains the sequences C2 and C4 (see Figure 1) cloned together in pBR322. The probe shows hybridisation to the vector sequences of pMB8a (marked 'v' in the Figure) and also to type I sequences present in a 0.45kb BglII fragment within the 1.35kb HindIII fragment, 8a/6.

The Isolation of the Corresponding 'Empty' Site

We have used segments of DNA from 8a/6 as probes in order to estimate their sequence complexity in the total genome, and secondly to isolate the corresponding DNA segments lacking the insertion. The fragment 8a/6 which contains type I sequences would be expected to hybridise to a complex pattern of fragments in total genomic DNA. We therefore carried out genomic Southern blots using the adjacent fragment 8a/5 as a probe. This sequence hybridises to a HindIII fragment of identical size within total genomic DNA. We quantitated the level of hybridisation of this sequence by comparing the autoradiographic response to that given by known amounts of plasmid DNA in adjacent tracks (data not

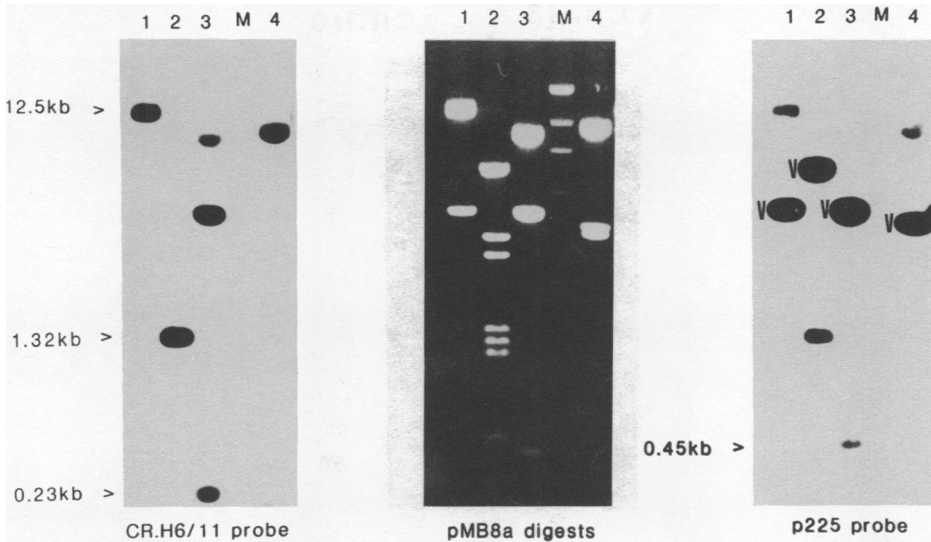


Figure 2. Identification of Restriction Fragments Containing the Type I Insertion in MB8a. The central panel shows the ethidium stained gel of pMB8a (MB8a subcloned from phage λ into pBR322). The plasmid has been digested with EcoRI, track 1; HindIII, track 2; EcoRI and BglII, track 3; EcoRI and BamHI, track 4. The gel was blotted onto nitrocellulose and probed with p225, segments C2 and C4 of Dml03 (Figure 1.) cloned together in pBR322 (right panel); or CR.H6/11, a HindIII fragment from a recombinant phage carrying a segment of *Drosophila* DNA carrying the 'empty' site (left panel). The CR.H6/11 probe is in M13mp8.

shown). This experiment indicated that the HindIII fragment 8a/5 was present as a single copy within the *D. melanogaster* genome. We therefore used this probe to examine the genomic DNA from a number of strains of *D. melanogaster*. The HindIII fragment 8a/6 that contains the type I insertion and the flanking HindIII fragment 8a/5 are both contained within a 6.3kb PstI fragment (see Figure 1.). We therefore expected that 8a/5 would hybridise to this 6.3kb PstI fragment in the genomic DNA of the Oregon R strain from which MB8a was isolated. If a type I sequence were not present at this site, then we expected to see a shorter fragment. We actually observed a more complex pattern of hybridisation which we believe is due to a repetitive sequence element in the vicinity of the right most PstI site. Nevertheless, we were able to identify a 6.3Kb Pst fragment, corresponding to the length of the fragment in MB8a, that was present in most strains of *D. melanogaster* but absent from the strain Karsnas. We therefore screened a library of DNA segments

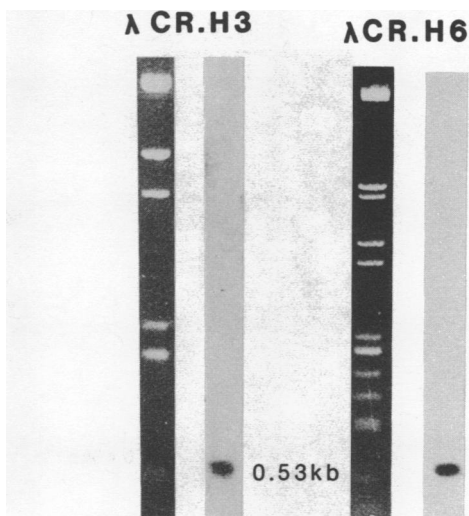


Figure 3. Two Cloned DNA Segments that Contain the 'Empty' Site. The recombinant phages λ CR.H3 and λ CR.H6 were isolated from a library of DNA from the Karsnas strain of *D. melanogaster* as described in the Materials and Methods. DNA from these recombinant phage were digested with several restriction endonucleases and blotted onto nitrocellulose for hybridisation with pMB8a/6, a sub-cloned HindIII fragment of MB8a that contains the type I insertion. The ethidium stained gel and Southern hybridisation pattern are shown for a double digest with HindIII and Sall. The 0.53kb fragment is also obtained with HindIII alone. It contains no sites for Sall or EcoRI, but has a single site for BglII.

from the Karsnas strain, and isolated clones that would hybridise with the fragment 8a/5 in the anticipation that these should contain the 'empty' site. The restriction endonuclease profiles of clones, λ CR.H3 and λ CR.H6, are shown in Figure 3. These two gels have been blotted and probed with the 1.32kb HindIII fragment 8a/6 that contains the type I insertion. It can be seen that in both cases digests that include HindIII generate a 0.53Kb fragment that hybridises with the probe.

We sub-cloned this 0.53kb HindIII fragment (CR.H6/11) from λ CR.H6, and used it to probe digests of MB8a (Figure 2.left panel). As predicted, it hybridised to the 1.32kb HindIII fragment that contains the insertion. It does not hybridise to the 0.45kb BglII fragment internal to the insertion, but gives a strong signal to the flanking 0.23kb BglII fragment.

Distribution of the 'Target' HindIII Fragment Between Strains of *D. melanogaster*

We have used the sub-cloned 0.53Kb HindIII fragment

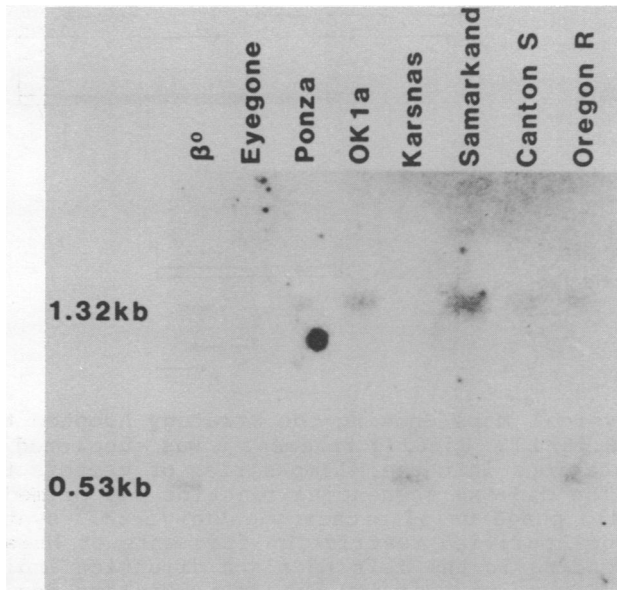


Figure 4. Hybridisation of the HindIII Fragment Containing the Empty Site of λ CR.H6 to Different Strains of *D. melanogaster*. DNA was extracted from the indicated strains of *D. melanogaster* using the procedure described by de Cicco and Glover [15]. The strain β^0 is a fly carrying null allele for the β subunit of larval serum protein 1 and was isolated by Dr David Roberts, Oxford. Ponza is a wild-type stock and Eyegone, a stock carrying the marker *eyg* at map position 3. 35.5. These two stocks also carry null-alleles for the gamma sub-unit of LSP1 (Roberts, personal communication). The Oklahoma (OK1a) line was established in 1970 from flies collected in the wild [16], and was obtained from the University of Cambridge. The DNA was digested with HindIII fractionated by electrophoresis and transferred from the agarose gel onto nitrocellulose for blotting with CR.H6/11.

(CR.H6/11) to examine the distribution of homologous HindIII fragments in the genomes of eight strains of *D. melanogaster*. The results of this experiment are shown in Figure 4. We see two bands of hybridisation at 1.32Kb and 0.53Kb which correspond to interrupted and uninterrupted HindIII fragments, respectively. Both bands are present in the β^0 and Oregon R strains indicating chromosomes with and without the insertion. The Ponza, OK1a, Samarkand and Canton S strains have only the 1.32Kb band indicative of the interrupted sequence, whereas the Karsnas strain has only the 0.53Kb band indicating the absence of the insertion at this site.

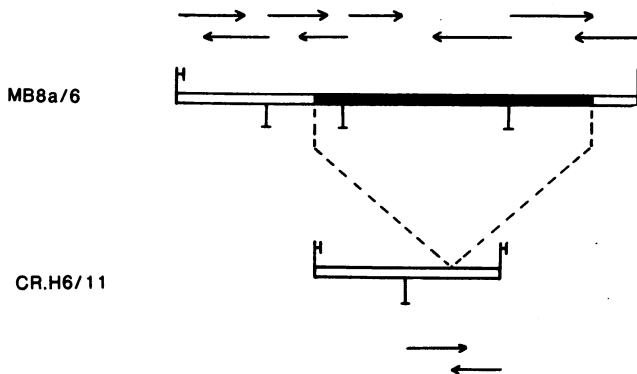


Figure 5. Physical Maps Showing the Strategy Adopted to Sequence MB8a/6 and CR.H6/11. HindIII fragment 6 was subcloned from pMB8a in both orientations into the M13mp series of vectors for sequencing. The dideoxy sequencing reaction was primed from the recombinant M13 phage using either the 'universal' synthetic primer using gel purified restriction fragments of MB8a/6 as primer. Sequencing in the left to right direction indicated by arrows in the Figure was carried out on a template designated MB8a/6.W5 and sequencing from right to left on a template designated MB8a/6.W7 as the template DNA. The use of restriction fragments as primers means that there are short gaps in the sequence of the insertion which we have not completed but the sequence at the junctions between the type I sequence and its flanking DNA is complete and has been read from both strands. The HindIII fragment MB8a/6 is 1.32Kb in length and corresponds to the upper band of hybridisation in Figure 4. Whereas the HindIII fragment CR.H6/11 is 0.53Kb in length and corresponds to the lower band of this length in Figure 4.

DNA Sequence Analysis at the Insertion Site

The physical maps of the interrupted and uninterrupted HindIII fragments are shown in Figure 5, together with the strategy that was adopted for their sequencing. The sequence of CR.H6/11 around the site of the insertion is given in Figure 6a. The numbering in this Figure is from the point of the insertion. The sequence of the left and right junctions of the type I insertion with this non-rDNA segment is shown in Figure 6b and is derived from the sequencing of MB8a/6. It can be seen that the type I insertion is flanked by a duplication of 10 bases at the target site for insertion. A comparison is shown in Figure 6c with the type I insertion in the rDNA clone MB27 [11]. In this case insertion of the type I sequence has resulted in the duplication of a 7 base pair sequence at the target site (although in other clones the length of duplication varies [11]). It can be seen that the right hand junction of the element with



Figure 6. Sequence at the Site of the Type I Insertion in MB8A and of the Corresponding Empty Site in CR.H6.
A. The sequence flanking the type I insertion site in CR.H6/11. The sequence flanking the type I insertion in MB8a is identical except that the 10 base pair element indicated by asterisks is duplicated on either side of the type I element.
B. Sequences at the junctions of the type I sequence with the non-rDNA flanking sequences. The sequence of the 765 base pair insertion is not given. We observed 5% sequence divergence between this sequence and the previously published sequence of the insertion of Dm103 [6,11].
C. Sequence at the junctions at the 1Kb type I insertion in the 28S rDNA of the clone MB27 [11].

the rDNA is identical to the junction between the type I sequence and the non-rDNA element in MB8a. The insertion in MB8a therefore forms one other member of the co-terminal subset of type I insertion sequences. There is a striking similarity between the sequence at the site of the insertion in the rDNA and in MB8a. In each case the duplicated segment begins with TGTPyCPy and is preceded by AC in both the uninterrupted rDNA and unique sequence. This reinforces our earlier observations regarding the strong specificity of this transposon [11]. The

insertion in MB8a is 765 base pairs in length and corresponds to the extreme sequences from the right hand side of the rDNA insertions. We can detect 5% sequence divergence between the rDNA insertion of Dml03 and the MB8A insertion from these experiments.

DISCUSSION

We have isolated a unique sequence from the genome of *D. melanogaster* that is interrupted by an insertion element previously characterised as occurring with the gene for 28S ribosomal RNA. This clone differs from other clones of non-rDNA type I insertion sequences that have been previously characterised and which consist predominantly of tandemly arranged type I elements [10]. The insertion of a type I element into a unique non-rDNA segment is another piece of evidence supporting the thesis that type I sequences are capable of transposition. The non-repetitive nature of the target site for the insertion has allowed us to look in other strains of *D. melanogaster* for the corresponding 'empty' site. We analysed the genomic DNA from 8 strains of *Drosophila melanogaster* for the restriction profile of unique sequences immediately flanking the type I insertion and were able to find examples of both the interrupted and non-interrupted unique sequence.

Other lines of evidence also indicate that the type I elements are capable of transposition. The unit type I elements within tandem arrays are each flanked by tiny rDNA segments indicating that these sequences were once within the rDNA and have subsequently undergone transposition [6]. The major 5Kb rDNA insertion itself is flanked on the left by a deletion of 9 nucleotides of rDNA indicative of some ancestral sequence rearrangement at this site [6]. There are also rDNA units that contain smaller insertions which appear to be derived from the larger 5Kb insertion. These small insertions comprise a co-terminal subset of sequences which have identical junctions at their right hand side with the rDNA, but have their left hand junctions at an apparently random site up to 1Kb from the right hand junction. The rDNA units into which these shorter type I elements are inserted do not have deletions at the insertion site but rather the insertion is flanked by a duplication of 7-15 base pairs at the target site [11,12]. The type I elements are, however, unusual compared to other transposable elements in that they have neither inverted or directly repetitious terminal sequences, nor do they end in a stretch of poly A.

The organisation of the type I sequences within the unique DNA segment that we have characterised has many features in

common with the organisation of the short type I elements in rDNA. There is an identical right hand junction between the type I insertions and the flanking DNA at these two locations. The type I insertion at this unique site is therefore another member of the co-terminal subset of short type I elements. It is flanked by a duplication of the target sequence which in this case is 10 base pairs long. There is also striking homology at the target sites of the non-rDNA and the rDNA. In both cases the insertion occurs within a sequence ACTGTPyCPy and there is duplication of the sequence TGTPyCPyN where N represents the subsequent nucleotides and can be a sequence between 1 and 9 nucleotides in length in the clones that we have examined.

The ability of the right hand sequences from the type I element to exist independently of the left hand sequences is seen not only in *D. melanogaster* but in the analogous elements within the sibling species [7]. This leads to complex highly polymorphic patterns of the type I sequences between the species. As a consequence of this complexity it is difficult to assess the rate of transposition of the rDNA element. The finding of a unique sequence showing polymorphism for the presence of a type I insertion may now provide one means of addressing this problem.

ACKNOWLEDGEMENTS

This work was supported by grants from the Medical Research Council and the Cancer Research Campaign (C.R.C). D.M.G. holds a Career Development Award from the C.R.C. M.J.B. was a visiting worker from Beechams Pharmaceuticals during 1981.

¹Permanent address: Beecham Pharmaceuticals Research Division, Biosciences Research Center, Great Burgh, Yew Tree Bottom Road, Epsom, Surrey KT18 5XQ, UK

²Present address: Amersham International plc, Amersham, Buckinghamshire HP7 9LL, UK

³Present address: Department of Biochemistry, University of California, Berkeley, CA 94720, USA

REFERENCES

1. Glover, D.M. and Hogness, D.S. (1977). *Cell* 10, 167-176.
2. White, R.L. and Hogness, D.S. (1977). *Cell* 10, 177-192.
3. Wellauer, P.K. and Dawid, I.B. (1977). *Cell* 10, 193-212.
4. Pellegrini, M., Manning, J. and Davidson, N. (1977). *Cell* 10, 213-224.
5. Roiha, H. and Glover, D.M. (1980). *J. Mol. Biol.* 140, 341-355.
6. Roiha, H., Miller, J.R., Woods, L.C. and Glover, D.M. (1981). *Nature* 290, 749-753.
7. Roiha, H., Read, C.A., Browne, M.J. and Glover, D.M. (1983). *E.M.B.O. Journal* 2, 721-726
8. Wellauer, P.K., Dawid, I.B. and Tartof, K.D. (1978). *Cell* 14, 269-278.

9. Tartof, K.D. and Dawid, I.B. (1976). *Nature* 263, 27-30
10. Kidd, S.J. and Glover, D.M. (1980). *Cell* 19, 103-119.
11. Roiha, H. and Glover, D.M. (1981). *Nucleic Acids Res.* 9, 5521-5532.
12. Dawid, I.B. and Rebbert, M.L. (1981). *Nucleic Acids Res.* 9, 5011-5020.
13. Glover, D.M., White, R.L., Finnegan, D.J. and Hogness, D.S. (1975). *Cell* 5, 149-157.
14. Appels R. and Hilliker A.J. (1982) *Genet. Res.* 39 149-156
15. de Cicco, D.V. and Glover, D.M. (1983). *Cell* 32, 1217-1225
16. Woodruff, R.C. and Thompson, J.N. (1977). *Heredity* 38, 291-307