The transcriptional control regions of the *copia* retrotransposon

Alan Sneddon and Andrew J.Flavell*

Department of Biochemistry, The University, Dundee DD1 4HN, UK

Received April 13, 1989; Revised and Accepted May 12, 1989

ABSTRACT

We have analysed the sequence elements that control expression of the *copia* retrotransposon. Expression of *copia* fusion constructs containing DNA sequence deletions and rearrangements was assayed by transient expression analysis. Progressive deletion and linker substitution identifies two regions on either side of the major transcriptional start sites in the *copia* long terminal repeat These regions are both required for high level expression in a cultured *Drosophila melanogaster* cell line but only the upstream region is required for *copia* expression in a *Drosophila hydei* cell line. A third control region lies downstream of the long terminal repeat in a region previously believed to contain no cis-acting regulatory sequences. We show by displacement and inversion of this region that it contains a transcriptional enhancer.

INTRODUCTION

The *Drosophila* transposable element *copia* is a member of the widespread family of retrotransposons which inhabits many, and perhaps all eukaryotic phyla. Close relatives to *copia*, as judged by structural features such as the relative gene order in the transposon, are found in *Drosophila* (1), yeast (2) and plants (3,4). More distant relatives include the vertebrate retroviruses and many other retrotransposons in both *Drosophila* and other organisms (5,6). All of these related transposable elements almost certainly transpose via RNA intermediates which are copied into DNA by reverse transcriptases encoded by the mobile genetic elements (5,7−9). *copia* RNA is therefore both an intermediate in transposition as well as the message for the proteins catalysing this process.

copia RNAs are probably the single most abundant mRNAs in cultured *Drosophila* cells (10). In the *Drosophila* organism *copia* RNAs, though less prevalent, are still major RNA species (11,12). *copia* expression is regulated in a temporal manner during the development of the insect (11−13) but knowledge of the particular tissues in which it is expressed is incomplete (14,15). The major *copia* RNAs share a staggered set of 5′ termini, the majority of which lie in the centre of one of the long terminal repeats (LTRs) which flank the mobile element (16). As a prelude to the determination of the regions of *copia* mobile element which control its developmental specificity, we have identified in detail the sequences which are required for *copia* expression in two *Drosophila* cell lines which support high levels of its gene expression.

Previous studies of the control regions which regulate *copia* gene expression have reached ambiguous conclusions (17−19). Sequence analysis of *copia* predicted the existence of a transcriptional enhancer in the 5′ untranslated region between the left long terminal repeat and the single long open reading frame of the retrotransposon (17). Transient expression analysis, however, has suggested that this region is inessential for high level

expression in cultured cells (18). Putative cis-acting control regions in the 5′ LTR have been localized by the latter technique to a region downstream from the transcriptional start sites and, surprisingly, no control regions were identified upstream of the transcribed region (19). However, these functional studies have been confused by the choice of a circularly-permuted *copia* based upon an extrachromosomal *copia* circular DNA (18−20) and the use of a cell line (21) derived from a *Drosophila* species which lacks *copias* (A. Flavell, unpublished).

We have created a comprehensive and systematic series of deletions and sequence replacements in the 5′ LTR of a genomic *copia* element to define accurately the regions which are necessary for its expression in two different *Drosophila* cell lines. Our data show here that three major regions are necessary for high levels of *copia* gene expression. One of these lies upstream of the major transcriptional start sites in the *copia* LTR. Another lies downstream of the start sites, but still inside the LTR. This latter region is only required for *copia* expression in a cultured *Drosophila melanogaster* cell line but is dispensible in a *Drosophila hydei* line. Lastly, we demonstrate that a third element lies downstream of the 5′ LTR and is a transcriptional enhancer.

MATERIALS AND METHODS

Construction of Recombinants

Recombinant DNA techniques were from Maniatis *et al* (22), except where indicated. The general structures of the recombinants are shown in Figure 1. Upstream deletions in the 5′ LTR (pKSCopCATΔ5′ series) were made as follows. Bal 31 exonuclease deletions (New England Biolabs) were made from an Msp 1 site (converted into a Sma 1 site beforehand) approximately 100bp upstream of the 5′ LTR of the *copia* element from the *copia* insertion in the *white* locus (17). The deletion endpoints were converted into Sma 1 sites and the resultant clones were grown up and the approximate extent of deletion determined by restriction digests. The truncated 5′ LTRs of selected deletants were excised as Sma 1-Apa 1 fragments and used to replace the undeleted corresponding Sma 1-Apa 1 5′ LTR fragment of pKSCopCAT (the unique Apa 1 site of *copia* is at position 283; reference 17). pKSCopCAT contains the CAT gene from pSVO-CAT (23) fused to a Pvu 11 site at position 820 of *copia* (structure shown in Figure 1) This replacement reconstructs genomic *copias* with progressive deletions into the 5′ LTR, followed by the normal *copia* sequence until the CAT gene at position 820. The sequence upstream of the 5′ LTR is the same for all these clones. A further set of Bal 31 deletions was made from the Sma 1 site of pKSCopCAT. All deletion breakpoints were identified by dideoxy DNA sequencing (23) using Sequenase™ (United States Biochemical, Cleveland).

Deletions from the 3′ side of the 5′ LTR (pKSCopCATΔ3′ series) were by Bal 31 digestion from the Apa 1 site of the *white-apricot copia* element from p3922a[15] (17). The deletion endpoints were converted into Sma 1 sites by linker addition, random deletants were cloned and the approximate extent of deletion determined by restriction mapping. Selected deleted 5′ LTRs were used (as Xba 1-Sma1 fragments) to replace the corresponding region of pKSCopCAT (whose Apa 1 site was converted into a Sma 1 site to accept the replacement). This resulted in a unidirectional deletion series into the 5′ LTR initiating from position 283, followed by the same *copia*-CAT fusion as was used in the upstream deletion series.

The linker scanning mutants were derived by combining appropriate pKSCopCATΔ5′ and pKSCopCATΔ3′ deletants by ligation of Sma 1-digested recombinants using the strategy

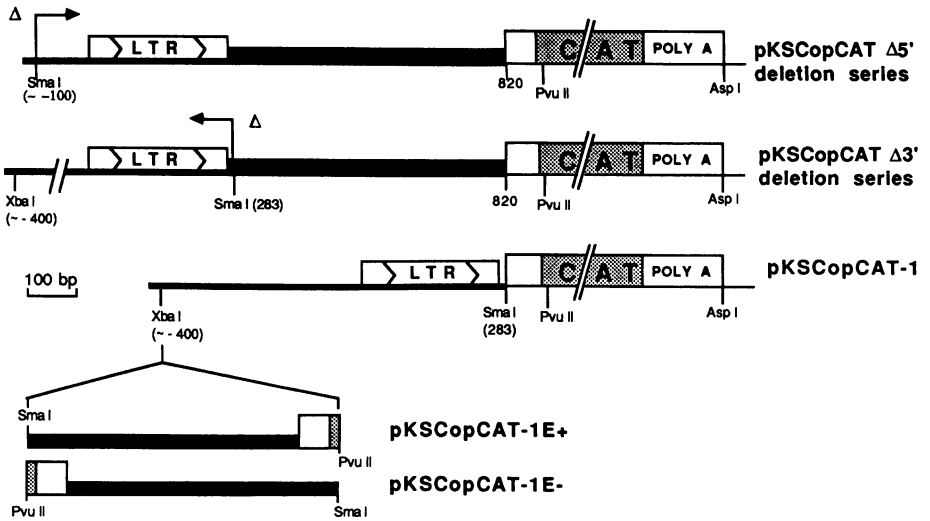

FIGURE 1. *Restriction maps of constructs used in this study.*
The nucleotide positions of the relevant sites on *copia* are shown. The enhancer-containing region is marked by a thick bar. The start points for the deletion series are arrowed. The translated region of the CAT gene (stippled box) and the polyadenylation region are not to scale

of McKnight (24). To ensure that the pLS42−66 mutant did not suffer an overall loss of nucleotides at the mutation site, a Sma 1-Hind 111-Sma 1 linker was used. The (36−71) duplication (pKSCopCATd1) was a fortuitous clone picked up during the screening for the pLS mutants. The (1−276) duplication is pCV2-CAT (a gift of John Sinclair). This construct is basically pCV2-gpt (18) with the CAT gene replacing the gpt gene. This construct is based upon a different vector (pAT 153) from our constructs and carries the 3′ region of *copia* upstream of the LTRs.

The translocated enhancer region constructs (pKSCopCAT-1E+ pKSCopCAT-1E− were derived from pKScopCAT-1 (Figure 1). This is basically a derivative of pKSCopCAT which lacks the region between the *copia* Apa 1 site (previously converted into a Sma 1 site, see above) and the Pvu 11 site at 820, where the CAT gene is joined to *copia*. The enhancer-containing fragment is a Sma 1-Pvu 11 fragment of pKSCopCAT spanning nucleotides 283−820 of *copia* plus 134 bp of the leader region to the CAT gene (25), inserted in either orientation into an Xba 1 site approximately 400bp upstream of the *copia* element of pKScopCAT-1.

Cell Transfection and Expression Analysis

DH 33 cells for transfection studies (21) were cultured in M3 medium (26), supplemented with 10% heat-inactivated foetal calf serum (Gibco-BRL). Kc cells (27) were grown in D20 medium (27) plus 2% heat-inactivated serum. CAT assays were as described by Gorman *et al*, (25). Transfections contained 10μg CAT fusion plasmid and 5μg pCβ3P2 as an internal control for transfection efficiency (these quantities being previously determined as non-saturating, data not shown). The latter plasmid contains a *copia-β* galactosidase fusion construct (Brierley and Flavell, in preparation). Quantitation of results was by scintillation counting of eluted ^{14}C spots for CAT assays and by spectrophotometric

TABLE 1. Deletion analysis of the *copia* long terminal repeat

5′ Deletions (pKSCopCATΔ5′ series) Deletion endpoint	*CAT activity* Kc cells			DH 33 cells		
0	100	–	–	100	–	–
15	nd	–	–	96	(2)	[1]
36	97	(2)	[3]	95	(2)	[2]
53	46	(4)	[1]	55	(3)	[2]
67	18	(3)	[2]	32	(3)	[1]
73	11	(3)	[1]	32	(3)	[1]
112	6	(2)	[3]	13	(2)	[1]
163	6	(3)	[1]	nd	–	–
283	11	(2)	[2]	8	(2)	[1]
3′ Deletions (pKSCopCATΔ3′ series)						
283	100	–	–	100	–	–
230	66	(2)	[18]	135	(2)	[5]
219	34	(2)	[9]	80	(2)	[7]
198	38	(7)	[9]	86	(4)	[3]
180	18	(2)	[3]	91	(2)	[6]
169	21	(2)	[1]	75	(2)	[18]
152	21	(4)	[4]	89	(2)	[7]
117	16	(3)	[4]	82	(4)	[1]
41	10	(2)	[1]	19	(2)	[6]
0	13	(2)	[1]	8	(2)	[4]
Replacement Mutants Substituted region						
42–52	49	(3)	[4]	67	(3)	[8]
42–66	22	(6)	[1]	34	(4)	[2]
Duplication Mutants Duplicated region						
2×(36–71)	90	(2)	[6]	109	(2)	[4]
2×(1–276)	102	(5)	[3]	84	(2)	[3]

Descriptions of the constructs are in Figure 1 and Materials and Methods. The nucleotide positions on the *copia* LTR are shown in the first column. The undeleted pKSCopCAT construct was used for the 100% reference in all cases. nd means the experiment was not done. The numbers in round brackets refer to the number of individual experiments which were performed with each construct and numbers in square brackets are the standard errors of each mean value.

absorbance at 420nm for β galactosidase assays (28). The results from each experiment were corrected for the individual transfection efficiencies, as determined by β galactosidase expression, then normalized between experiments to the CAT expression of pKSCopCAT. Each quoted CAT value is an average of at least two individual dishes, the actual numbers of experiments are shown in Table 1.

RNAs were prepared from transfected cells by the method of Savakis (29). DNA was removed from the RNA samples by digestion with RNAase-free DNAase (Boehringer) prior to electrophoresis in formaldehyde gels (22), blotting to Biodyne membranes (Pall corporation,) using the manufacturers protocol and probing with oligo-primed ^{32}P probe DNA (30). To ensure that contaminating DNA was not responsible for any of the observed

bands, the Northern blots were stripped of RNA in 0.2M NaOH, 20°, 30min then reprobed. Only parallel lanes containing control DNA gave any signal and these bands were of comparable intensity to those on the original probing, showing that the membrane was not harmed by this procedure (data not shown).

RESULTS

Mutation Analysis of the copia 5′ Long Terminal Repeat

We have mapped the control regions in the 5′ LTR of the *copia* element by generating two deletion series, beginning from each end of the LTR. This approach has two advantages. First, it circumvents the limitation to deletion analysis which arises when more than one important control region is present, namely that deletion of the first control region encountered in the series results in such a large drop in expression that the effect of removing any downstream regions is masked. Second, it enables the easy creation of small substitution mutations (or 'linker scanning mutations') which accurately pinpoint control regions (24).

To assay the expression of the *copia* mutant series, we fused each construct to the chloramphenicol acetyl transferase (CAT) gene (25) and the resulting fusion constructs were introduced into cultured *Drosophila* cells by calcium phosphate-mediated transfection. To relate our results to previous studies, we employed the DH 33 line of cultured *Drosophila hydei* cells as a recipient for our constructs (18−21). These cells support high level expression of transfected *copias*. However, *Drosophila hydei* does not contain any *copias*, so results obtained exclusively with this cell line are open to question. We therefore also introduced our constructs into the Kc line of *Drosophila melanogaster* cells (27). This line contains approximately 100−200 copies of endogenous *copias* (31) and supports high level expression of both these (16), and introduced *copia* elements.

The results of this analysis are shown in Table1. Deletion of up to 36 base pairs (bp) from the left (or 5′) end of the LTR has no effect upon expression in either cell type. Removal of the next 17bp results in a two-fold drop in *copia* expression and loss of 14 bp more (a total deletion of 66bp) yields a further drop of approximately two-fold to a level of 18% and 32% of the normal levels for *D. melanogaster* and *D. hydei* cells respectively. Substitution of the sequence between nucleotides 42 and 52 with an oligonucleotide linker removes roughly one half of the CAT expression, an equivalent drop to that caused by the 1 to 52 deletion suggesting that the nucleotides between 1 and 41 are dispensible. Further replacement of bases 53 to 66 drops the level of expression by another factor of two-fold. Taken together, these data identify a region lying between nucleotides 42 and 66 which stimulates *copia* gene expression by a factor of three to five-fold, depending upon the cell type. This region lies upstream of the *copia* transcriptional start sites (16) in a region previously believed to contain no transcriptional regulatory sequences (19). Furthermore, it probably extends further in the 3′ direction because removal of six more bp (to position 73) drops the activity of *copia* constructs in Kc cells by about 40% further and an extra 45bp deletion (up to nucleotide 112) results in a drop of CAT expression in DH 33 cells to a level roughly ten-fold below an unmutated *copia*.

Further deletion from the 5′ end of the *copia* 5′ LTR has little or no effect on the residual level of *copia* gene expression, indeed the removal of the entire 5′ LTR still leaves roughly 10% of the normal *copia* activity in both cell types. This residual level of expression is dependent upon the presence of nucleotides 283−820 of *copia* (the region downstream of the 5′ LTR) because removal of these bases yields no CAT activity above the untransfected cell background (2% or less for pSVO-CAT, see Table 2).

TABLE 2. *Copia* contains a transcriptional enhancer

Construct	CAT Activity					
	Kc Cells			DH 33 cells		
pKSCopCAT	100	–	–	100	–	–
pKSCopCAT-1	24	(3)	[2]	4	(9)	[1]
pKSCopCAT-1E+	nd	–	–	86	(2)	[3]
pKSCopCAT-1E−	nd	–	–	111	(3)	[2]
pSVO-CAT	2	(2)	[0]	0	(4)	[0]

Descriptions of the constructs are in Figure 1 and Materials and Methods. nd means the experiment was not done. The numbers in round brackets refer to the number of individual experiments which were performed with each construct and numbers in square brackets are the standard errors of each mean value.

Deletion from the 3′ end of the LTR reveals a major difference in the regions which are important for *D. melanogaster* and *D. hydei* expression (Table 1). Removal of sequences between nucleotides 180 and 283 has no significant effect upon the levels of CAT activity in *D. hydei* cells yet the same deletion abolishes approximately 80% of activity in *D. melanogaster* cells. Further deletion to nucleotide 117 has little effect, demonstrating the lack of control regions in this area. Deletions which extend further 5′ are consistent with the results of the 5′ deletion series but generate no extra information.

We also assayed two mutants which contain duplications of LTR sequence (Table 1). pKSCopCATd1 contain two tandem copies of nucleotides 36−71 and pCV2-CAT contains a tandem duplication of the complete LTR (with a single base between). Neither duplication gave significantly elevated *copia*-CAT expression.

The broad conclusions from these data are shown in Figure 2. There are two major areas which affect the efficiency of *copia* gene expression in *D. melanogaster* cells. These straddle the major transcriptional start sites (16), which themselves are dispensible. The only other portion of the LTR which is not required for high levels of *copia* expression is the first 36 bp. This region is presumably required as a target site for the integrase enzyme responsible for the insertion of extrachromosomal *copia* DNA into the *Drosophila* chromosome (17). In all these studies it is impossible to estimate from such data the exact extent of the control regions unless incremental steps of one base are removed.

In *D. hydei* cells the transcriptional control regions are significantly different. The region upstream of the transcriptional start site is the only important sequence for *copia* expression in these cells. It extends further 3′ than is the case for *D. melanogaster* and is comprised of at least two elements. Thus *copia* requires markedly different regulatory regions for optimal gene expression. These effects are presumably mediated by transcriptional rgulatory factors which recognise specific nucleotide sequences and are present in differing amounts in the two cell types.

Copia Contains a Transcriptional Enhancer

The deletion studies described above show that both *D. melanogaster* and *D. hydei* cells support a significant level of *copia* expression even in the absence of the entire 5′ LTR. The region responsible for this residual activity lies in *copia*, because removal of all of the *copia* sequences from the CAT fusion construct (pSVO-CAT) yields no measurable CAT activity in both cell types (Table 2). To further demonstrate the importance of this region, we specifically removed it from pKSCopCAT. This resulted in a five-fold drop in expression in Kc *D. melanogaster* cells and a twenty five-fold drop in *D. hydei* cells

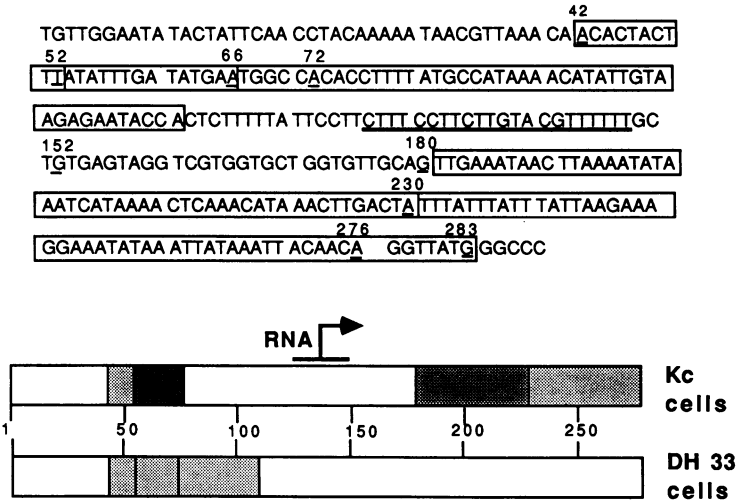

FIGURE 2. *The Locations of Control Regions in the copia* LTR
The nucleotide positions of breakpoints mentioned in the text on the *copia* LTR are shown by numbers and light underlining. Boxed parts of the sequence contain control regions mentioned in the text. The position of the multiple transcriptional start sites (16) on the sequence is shown by heavy underlining. The approximate locations of the control regions for the two cell types used in this study are shown schematically below the sequence. The estimated importance of each boxed region to *copia* expression is indicated by the density of shading, deletion of a lightly shaded region results in an approximate twofold drop in expression and deletion of the most densely shaded region yields roughly a fourfold reduction in *copia*-CAT expression.

(Table 2). This region contains a sequence repeat which resembles a sequence found in the mammalian SV40 viral enhancer (17,32). We therefore tested the possibility that this region contains a transcriptional enhancer by removing it from its normal position and replacing it in either orientation 5′ to the *copia* LTR (pKSCopCAT-1E+ and pKSCopCAT-1E−). These alterations gave high levels of CAT activity in DH 33 cells (Table 2), demonstrating that the region between nucleotides 283 and 820 does indeed contain a transcriptional enhancer.

We next measured the RNA levels in these cells to determine whether the CAT activities accurately reflect the RNA concentrations and eliminate the possibility that translational effects are causing the reductions in gene expression when the deleted regions lie in the transcribed parts of *copia*. We chose recombinants which each lack one of the three control regions identified by these studies and assayed cell cultures transfected with these, plus control pKSCopCAT, for RNA levels (Figure 3). The levels of *copia*-CAT RNAs (Figure 3B) are approximately equivalent to the respective CAT activities (Figure 3A) for these constructs. Probing the stripped blot for β galactosidase-specific RNA demonstrates that equivalent expression levels of this control sequence were obtained with each of the transfected samples (Figure 3C).

DISCUSSION

The data presented here support three major conclusions. First, *copia* contains a region upstream of the transcriptional start site in the 5′ LTR which is required for high level

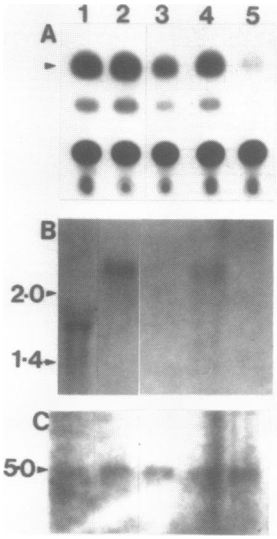

FIGURE 3. *CAT activities are approximately proportional to RNA levels*
Lane 1 pKSCopCAT-1E−; Lane 2 pKSCopCAT; Lane 3 pKSCopCATLS(42−66); Lane 4 pKSCopCATΔ3′-180; Lane 5 pKSCopCAT-1. DH33 cells were transfected with the above plasmids plus pCβ3P2 internal control. Each cell sample was split into three portions for CAT and β galactosidase activity determinations and RNA preparation respectively. CAT and Northern blot assays were performed using quantities normalized for the relative β galactosidase activities of the sample preparations. A, CAT activities-monoacetylated chloramphenicol is arrowed. B, Northern blot analysis with CAT probe. RNAs from pKSCopCAT-1E− and pKSCopCAT-1 (Lanes 1 and 5 respectively) are 540bp shorter than the others and their mobility is somewhat distorted by rRNA. The mobilities of marker DNAs are shown in kb. C, The filter was stripped and reprobed with β galactosidase probe. The mobility of a marker 5.0kb DNA is shown.

expression in cultured *Drosophila* cells. Second, a further region downstream of the RNA start sites in the 5′ LTR is needed for optimal activity in *D. melanogaster* cells but is unnecessary in *D. hydei* cells. Third, another control region lies 3′ to the *copia* LTR and this is a transcriptional enhancer.

This arrangement of regulatory regions is reminiscent of other retrotransposons. The Ty element of yeast is a structural relative of *copia* (2), sharing the same gene order and a similar primer tRNA for initiation of reverse transcription (although there is little DNA sequence homology between the two retrotransposons). Ty contains regions internal to its LTRs which can stimulate the transcription of adjacent genes and analysis of revertants has localized an element resembling the SV40 enhancer core sequence, which is required for this effect (33−34). Ty also possesses control sequences, in addition to the enhancer-like element, both upstream and downstream of its transcriptional start site (see above references). The *gypsy* retrotransposon of *Drosophila* also contains enhancer-like elements in its 5′ untranslated leader region (35). These are required for the mutagenic effect of the mobile element on adjacent genes (36−38; Flavell *et al*, unpublished). The presence of enhancers therefore appears to be a common property of retrotransposons and the resulting interference with adjacent gene expression, with its consequent deleterious effects on the host, may be the price paid by the mobile elements for high transcription levels. There is as yet no definite example of *copia* altering the expression of an adjacent gene by the action of its enhancer element, although the effects on the *white* locus of this element in the *white-apricot* allele could be explained at least partly in this way (39).

The actual nature of the control sequences is at present obscure but can be guessed at from their sequences. The upstream region contains two adjoining motifs (between bases 68 and 86) reminiscent of the MLTF/USF and CREB/ATF binding sites (GGCCACACCTTT; TATGCCA for *copia* and GGCCACGTGACC; [T/G][T/A]CGTCA for MLTF/USF and CREB/ATF respectively (40−42). The enhancer has not been accurately mapped by us but a repeat, which was postulated by Mount and Rubin to be

a transcriptional regulatory element, lies in the fragment which we have shown to contain the enhancer (17). This repeat contains a region displaying dyad symmetry (ttTTTCACAttctTGTGAAAt[a/t]) and a part of this is similar to the GTll-1B motif (CTGTGGAATG) which is recognised by the AP5 transcription factor (43) and found originally in the SV40 enhancer (32).

The nature of the downstream *D. melanogaster*-specific control sequence is difficult to assess because we have not localized it accurately, indeed, our data do not distinguish whether this element exerts its effects at the transcriptional level at all. Rather, this region may be required for the stability of *copia* RNAs. However, we think that this latter possibility is less likely because tissue-specific transcriptional effects are well known but cell line specific destabilization of mRNAs has not been as well established.

The results described here have revealed major differences in the regions which are necessary for high *copia* activity in different *Drosophila* cells. The transcribed region of the LTR is inessential in *D. hydei* cells, yet its removal results in an 80% loss of activity in *D. melanogaster* cells. Conversely, the absence of the enhancer is five-fold more deleterious in a *D. hydei* cell than a *D. melanogaster* cell. Similar cell line-specific, differential effects of mutations upon gene expression have been observed in studies of the transcriptional regulatory regions of the mammalian tumour virus SV40 (44). These differences correlate with the differences in abundance in these cell types of the protein factors which interact with these control regions. We think it likely that *copia*, like SV40 has evolved a complex set of regulatory regions in order to support its expression in a variety of different cell types. Both *Drosophila* lines are derived from embryonic tissue but it is unclear which cell lineages eventually formed the immortal culture, as during the several months required for the establishment of *Drosophila* cell lines many different cell types, including those obviously adult-specific, emerge then perish. *copia* RNA is found throughout development of the organism (11−13) and while little is known of the tissues in which it is expressed, it is apparent that its transcription extends beyond those tissues required simply for its inheritance, that is the germ cells (14−15). We predict that the deletions affecting specific control sequences described here will lead to tissue-specific alterations in the expression pattern of *copia*. Reintroduction of such mutagenized constructs into the fly should therefore prove useful in probing the tissue specificity of the trans-acting factors which recognise these regions in both *copia* and other *Drosophila* genes.

Our conclusions of both the number and locations of *copia* control sequences are in contrast to those obtained previously (18−19). Sinclair *et al* concluded that the major regulatory region for *copia* gene expression in *D. hydei* cells lies in the LTR downstream of the transcriptional initiation site with no control element present in the region downstream of this. These former studies were based upon a circularly-permuted extrachromosomal *copia* circle which juxtaposed the 3′ LTR with the 5′ LTR (20). The 3′ LTR is identical to the 5′ LTR and these constructs therefore carried two sets of the upstream transcriptional regulatory regions. Under such circumstances the deletion of the *copia* enhancer might have no measurable effect. Studies on the SV40 transcriptional enhancer have shown that defects resulting from deletion of part of a transcriptional regulatory region can be compensated for by duplication of other control regions (45). We therefore suggest that the failure to detect the presence of the *copia* enhancer in these former studies was due to extra positive control regions in the 3′ LTR of the *copia* circular molecule compensating for the loss of the enhancer. Another discrepancy between our two studies can be resolved if this explanation is correct. Our two groups support different conclusions concerning

the importance of the transcribed region of the 5′ LTR to *copia*. gene expression. The studies of Sinclair *et al* suggested that deletion of the last 73bp of the LTR leads to a ten-fold decrease in activity in *D. hydei* cells. However, our data show that this region can be deleted with almost no effect upon expression in such cells. However, the deletant in the previous study had lost not only the last 73bp of one LTR but also lacked the extra LTR and the enhancer. We propose that it was the loss of both of these two control elements, not the 73bp of the 5′ LTR, that led to the decrease in expression.

Put together, our two sets of data are consistent and support a model that optimal expression of *copia* in *D. hydei* cells can be achieved if the enhancer and the upstream region of an LTR are present and that the enhancer is dispensible only if an extra copy of the LTR upstream control region is provided in the special case of an extrachromosomal 2 LTR *copia* circle. It is important to note that addition of an extra LTR can only compensate for loss of the enhancer; it is unable to boost expression above the levels of an unmutated *copia* (compare pCV2-CAT with pKSCopCAT in Table 1 and pCV1*gpt* with pCV2*gpt* in reference 19).

Another difference between our results and those of Sinclair *et al* cannot be due to the presence of an extraneous LTR. Their studies indicated that deletion of the first 69 bp of the copia 5′ LTR had negligible effect upon expression of a *copia-gpt* (xanthine guanine phosphoribosyl transferase) fusion construct which contained the same *copia* regions as our constructs. However, the 69 bp deletion did result in a two-fold drop in *copia* expression when combined with an enhancer deletion. Our results show that a smaller deletion (bases1 to 67) results in a threefold drop in CAT expression in the same cell type, even when the enhancer is present. We do not know the source of this discrepancy but it may be due to the differences in the assay methods or the sequences of the particular *copia* elements employed by our two laboratories.

ACKNOWLEDGEMENTS

We thank Jane Forrester for technical assistance. We are also grateful to John Sinclair for pCV2CAT plasmid and helpful discussions and to Cathy Brierley for pCβ3P2 plasmid. Both Cathy Brierley and Luke Alphey also provided useful advice and discussion. This work was supported by the MRC and the SERC.

*To whom correspondence should be addressed

REFERENCES

1. Fourcade-Peronnet, F., d'Auriol, L., Becker, J., Galibert,F. and Best-Belpomme, M. (1988) Nucleic Acids Res. **16**, 6113–6125.
2. Clare, J. and Farabaugh, P.J. (1985) Proc. Natl. Acad. Sci. USA **82**: 2828–2833.
3. Voytas, D.F. and Ausubel,F.M.. (1988) Nature **336**, 242–244.
4. Grandbastein, M-A., Spielman, A. and Caboche, M. (1989) Nature **337**, 376–380.
5. Varmus, H.E. (1983) p411–503. *In* J.A.Shapiro (ed) Mobile Genetic Elements. Academic Press Inc., Orlando, Fla USA.
6. Finnegan, D.J. (1985) International Review of Cytology, **93**, 281–326.
7. Boeke, J.D., Garfinkel, D.J., Styles, C.A. and Fink, G.R.. (1985) Cell **40**, 491–500.
8. Flavell, A.J. and Ish-Horowicz, D. (1981) Nature **292**, 591–595.
9. Arkhipova, I.R., Mazo, A.M., Cherkasova, V.A., Gorelova, T.V., Schuppe, N.G. and Ilyin, Y.V. (1986) Cell **44**, 555–563.
10. Carlson, M. and Brutlag,D. (1979) Proc. Natl. Acad Sci. **75**, 5898–5902.
11. Schwartz, H.E., Lockett, T.J. and Young, M.W. (1982) J. Mol. Biol. **157**, 49–58.

12. Flavell, A.J., Ruby,S.W., Toole, J.J., Roberts, B.E. and Rubin., G.M. (1980) Proc. Natl. Acad. Sci. (USA) **77**, 7107–7111.
13. Parkhurst, S.M. and Corces, V.G. (1987) EMBO J. **6**, 419–424.
14. Steller, H. and Pirrotta, V. (1984) EMBO J. **3**, 165–173.
15. McDonald, J.F., Strand, D., Brown, M.R., Paskewitz, S.M., Csink A.M. and Voss, S.H.. (1988) *In* M.E. Lambert, J.F. McDonald and I.B.Weinstein (eds) Eukaryotic transposable elements as mutagenic agents: Banbury Report 30, Cold Spring Harbor Laboratory, New York.
16. Flavell, A.J., Levis,R., Simon.M.A. and Rubin., G.M. (1981) Nucleic Acids Res. **9**, 6279–6291.
17. Mount, S.M. and Rubin, G.M.. 1985. Mol. Cell. Biol. **5**, 1630–1638.
18. Burke, J.F., Sinclair, J.H., Sang, J.H. and Ish-Horowicz, D. (1984) EMBO J. **3**, 2549–2554.
19. Sinclair, J.H., Burke,J.F., Ish-Horowicz, D. and Sang, J.H.. (1986) EMBO J. **5**, 2349–2354 .
20. Sinclair, J.H., Burke, J.F., Ish-Horowicz, D. and Sang,J.H..(1983) Nature **306**, 198–200.
21. Sondermeijer, P.J.A., Derksen, J.W.M. and Lubsen, N.H. (1980) In Vitro **16**, 913–914
22. *In* Maniatis, T., Fritsch, E.F. and Sambrook, J. (eds) Molecular cloning–a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. published 1982.
23. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. USA **74**, 5463–5467.
24. McKnight, S.L. and Kingsbury, R.. (1982) Science **217**, 316–324.
25. Gorman, C.M., Moffat, L.F. and Howard., B.H. (1982) Mol. Cell. Biol. **2**, 1044–1051
26. Shields, G. and Sang J.H. (1977) *Drosophila*. Information Service. **51**, 161.
27. Echalier, G. and Ohanessian, A. (1970) *In vitro* **6**, 162–172.
28. Miller,J.H. (1972) p352–355 In Experiments in Molecular Genetics. Cold Spring Harbor Laboratories, Cold Spring Harbor, New York.
29. Savakis, C., Ashburner, M. and Willis, J.H. (1986) Devel. Biol. **114**, 194–207.
30. Feinberg, A.P. and Vogelstein, B. (1984) Analytical Biochem. **137**, 266–7.
31. Potter, S.S., Brorein, W.J., Dunsmuir, P. and Rubin, G.M.. (1979) Cell **17**, 415–427.
32. Weiher, H., Konig, M. and Gruss, P. (1983) Science **219**, 626–631.
33. Roeder,G.S., Rose, A.B. and Pearlman, R.E. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 5428–5432.
34. Fulton,A.M., Rathjen, P.D., Kingsman, S.M. and Kingsman, A.J. (1988) **16**, 5439–5458.
35. Marlor, R.L., Parkhurst, S.M. and Corces, V.G. (1986) Mol. Cell. Biol. **6**, 1129–1134.
36. Geyer, P.K., Green, M.M. and Corces, V.G. (1988) Proc. Natl. Acad. Sci. USA **85**, 8593–8597.
37. Peifer, M. and Bender, W.. (1988) Proc. Natl. Acad. Sci. USA **85**, 9650–9654.
38. Mazo, A.M., Mizrokhi, L.V., Karavanov, A.A., Sedkov, Y.A., Krichevskaja, A.A. and Ilyin, Y.V. (1989) EMBO J. **8**, 903–911.
39. Mount, S.M., Green, M.M. and Rubin, G.M. (1988) Genetics **118**, 221–234.
40. Carthew, R.W., Chodosh, L.A. and Sharp, P.A. (1985) Cell **43**, 439–448.
41. Sawadago, M. and Roeder, R.G. (1985) Cell **43**, 165–175.
42. Montminy, M.R., Sevarino, K.A., Wagner, J.A., Mondel, G. and Goodman, R.H. (1986) Proc. Natl. Acad. Sci. USA **86**, 6682–6686.
43. Chiu, R., Imagawa, M., Imbra, R.J., Bockoven, J.R. and Karin, M. (1987) Nature **329**, 648–651.
44. Nomiyama, H., Fromenthal, C., Xiao, J.H. and Chambon, P. (1987) Proc. Natl. Acad. Sci. USA. **84**, 7881–7885.
45. Herr, W. and Clarke. J. (1986) Cell **45**, 461–470.

This article, submitted on disc, has been automatically converted into this typeset format by the publisher.