The EMBO Journal vol.6 no.2 pp.419–424, 1987

Developmental expression of *Drosophila melanogaster* retrovirus-like transposable elements

Susan M.Parkhurst[1] and Victor G.Corces

Department of Biology, The Johns Hopkins University, Baltimore, MD 21218, USA

[1]Present address: Imperial Cancer Research Fund, Developmental Biology Unit, Zoology Building, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

Communicated by G.Morata

We have determined the pattern of temporal expression of several *Drosophila* retrovirus-like transposable elements. Some of these elements can be grouped into classes whose members show a similar profile of developmental transcription. The members of the 412 class, which includes 412, mdg1, 17.6 and 3S18, are transcribed mainly in the early larval and pupal stages of development, with small differences among the various members. HMS Beagle and Springer constitute another class where RNA accumulation in the larval stages is higher than in pupae and the adult flies accumulate more RNA than any other stage of development. Finally, the transcription of other elements such as copia, 297 and B104 follows a specific and individual pattern distinct from those described above. These results suggest the existence of evolutionary relationships among different transposable elements in *Drosophila* and the involvement of different cellular genes in the control of their expression.

Key words: transposable elements/developmental expression/ *Drosophila melanogaster*

Introduction

Mobile genetic elements constitute up to 10% of the *Drosophila* genome and can be categorized depending on their structure and the mechanism of transposition they utilize (see Rubin, 1983; Finnegan and Fawcett, 1986 for reviews). Some transposable elements, such as P and hobo, have short inverted repeats at the ends; the internal part of the P element encodes an enzymatic activity that is involved in the process of transposition. Other elements, referred to as fold-back elements, contain long inverted repeats which are made up of short sequences repeated in tandem; the structure of these elements suggests that they do not use RNA as transposition intermediates. The members of a third type of transposable elements, the I and F families, lack terminal repeats and are similar to the Alu sequences in human DNA; these elements probably transpose by integration of double-stranded molecules made by reverse transcription of polyadenylated RNAs.

Most mobile elements in *Drosophila* belong to a fourth category of transposons, the copia-like or mdg family. These elements have structures similar to those of integrated retroviral proviruses, with two long terminal repeats (LTRs) 200–600 bp long which contain transcription initiation and termination signals. The central portion of these elements encodes proteins homologous to the endonuclease, protease and reverse transcriptase of vertebrate retroviruses. These elements probably excise from a particular place in the chromosome by homologous recombination between the two LTRs (Carbonare and Gehring, 1985; Mizrokhi *et al.*, 1985) and integrate in new places by a similar mechanism to that of vertebrate retroviruses (Flavell, 1984; Arkhipova *et al.*, 1986).

The transcription of *Drosophila* retroviral elements has been studied mainly in cultured cells and embryos (Flavell *et al.*, 1980; Scherer *et al.*, 1982; Schwartz *et al.*, 1982; Mossie *et al.*, 1985). These elements usually give rise to more than one RNA, the largest one corresponding to a transcript that initiates at the 5′ LTR and ends at the termination signals located at the 3′ LTR. The transcripts are polyadenylated and predominately nuclear. The transcription of three of these elements has been studied in whole flies. The copia element is transcribed into two different RNAs 5 and 2 kb long that accumulate at low levels during the embryonic stages of development. The levels of these transcripts increase during the larval stages and decrease again in adult flies (Flavell *et al.*, 1980; Schwartz *et al.*, 1982). The B104-encoded RNAs are present as maternal messages in unfertilized embryos and can also be detected in the early stages of embryogenesis but their expression has not been analyzed in later stages of development (Scherer *et al.*, 1982). The 412 element is transcribed late during embryonic development but is not present in larvae or adults (Schwartz *et al.*, 1982). These results suggest that different regulatory sequences are responsible for the expression of these elements, which must be under the control of different cellular genes. It is interesting to note in this respect that, as new mutant alleles of various genes are characterized, new and different retrotransposons [this term has been proposed by Boeke *et al.* (1985) to designate retrovirus-like transposable elements] are being identified which, with few exceptions, do not share sequence homology.

Nothing is known about the reason for the existence and maintenance of such a high number of different transposons in the *Drosophila* genome nor the origin and putative evolutionary relationships among these different elements. A plausible approach to solve this problem would be the analysis and comparison of DNA sequences from these various transposons, with the caveat that the sequences involved in control and regulation of their expression have not yet been identified. A second approach would be the analysis of the pattern of temporal expression of retroviral transposons on the basis that elements that have the same pattern of expression probably share common regulatory sequences and may be under the control of the same cellular genes, and, therefore, might have a common origin. In addition, evidence accumulated during the last few years strongly suggests a direct relationship between the mutagenic characteristics of retroviral elements and their transcriptional properties (Cullen *et al.*, 1984; Emerman and Temin, 1984; Winston *et al.*, 1984; Parkhurst and Corces, 1986b). A detailed knowledge of the temporal transcriptional specificity of these elements will thus also aid in the understanding of the mechanisms whereby they cause a mutant phenotype. Here we present an analysis of the expression during *Drosophila* development of most retrotransposons characterized to date in an attempt to approach these problems.

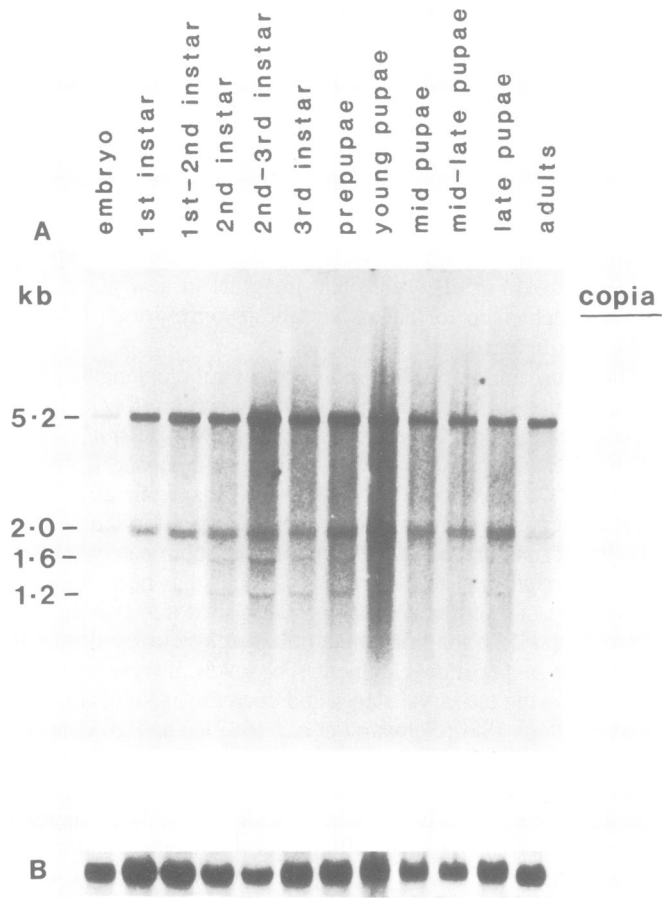

Fig. 1. Northern analysis of copia-encoded transcripts. **(A)** Poly(A)-containing RNA (25 μg) from different developmental stages of *Drosophila* was electrophoresed on a 0.8% formaldehyde−agarose gel and transferred to a BioTrans nylon filter. The membrane was then hybridized with a ^{32}P-labeled *Hpa*II DNA fragment containing exactly one copy of the copia element. **(B)** The same filter after hybridization with a cloned *Drosophila ras* gene to control for equal loading of the samples in the different lanes.

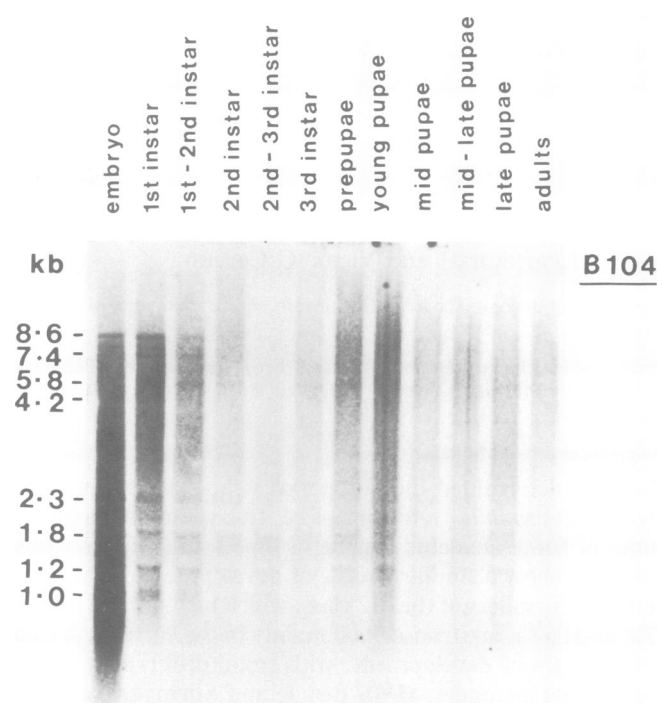

Fig. 2. Northern analysis of the developmental transcription of the B104 element. The same filter used in Figure 1 was dehybridized as described in Materials and methods and rehybridized with ^{32}P-labeled B104 DNA. The B104 probe was a 5.5-kb *Hin*dIII fragment that contains the 5′ LTR and approximately half of the internal sequences (Scherer *et al.*, 1982).

Results

Transcription of the copia element

In order to study the pattern of temporal expression of various retrotransposons, we prepared poly(A)$^+$ RNA from different stages of development of a *Drosophila* stock carrying the mutations y^2 *cv v* f^1. Eggs were collected for 24 h and allowed to grow at 22°C and 70% relative humidity. Under these conditions, development from the embryonic to adult stage was completed in 12 days. Samples were collected every 24 h and RNA was prepared, electrophoresed and blotted as described in Materials and methods. The nylon filter was first probed with a *Drosophila ras* oncogene that is expressed at approximately the same levels throughout the development of the fly (Mozer *et al.*, 1985) in order to control for RNA degradation and equal loading of the different samples (Figure 1B). After dehybridization, the same blot was reprobed with labeled plasmid DNA containing the copia element. This element is 5146 bp long and contains 276-bp LTRs (Mount and Rubin, 1985). The transcription of copia has been previously studied by others (Flavell *et al.*, 1980; Schwartz *et al.*, 1982) and is presented here only because of the more detailed developmental anlaysis we have carried out and as a point of comparison with other transposable elements we have studied. Flavell *et al.* (1980) reported the presence of copia RNA in larvae but not in embryos or adults, whereas Schwartz *et al.* (1982) detected low levels of a 5.2-kb transcript in both embryos and adults whereas the levels of this RNA were higher in larval stages; an additional 2.1-kb RNA was present in Kc_0 cells but absent or in very low levels in whole animals. Figure 1A shows a detailed analysis of the developmental expression of copia. This transposable element encodes four different transcripts; the largest one is 5.2 kb and probably extends from the transcription initiation site located at the 5′ LTR, to the termination signals located in the 3′ LTR. The pattern of temporal expression of all four RNAs is identical, suggesting that the smaller RNAs are either transcribed from the same promoter or that they are the result of splicing of the 5.2-kb transcript. Alternatively, these RNAs could result from the transcription of internally deleted copia elements present in the genome. Copia transcripts are present in low amounts in embryos and their accumulation increases in the larval stages to reach a maximum in second−third instar larvae, then decreases slightly and is maintained at approximately the same levels until the adult stage.

Transcription of B104 and 297

The B104 element is 8.7 kb long with LTRs that extend 429 bp (Scherer *et al.*, 1982). Figure 2 shows the result obtained after dehybridization of the same blot used to study the transcription of copia and rehybridization with labeled plasmid DNA containing the B104 element. This transposon presents a complex pattern of RNA expression during *Drosophila* development. A transcript 8.6 kb long, probably corresponding to an RNA that extends between the two LTRs, can be observed in embryos and first instar larvae. Both of these stages, but especially embryos, accumulate large amounts of several RNAs, the most abundant

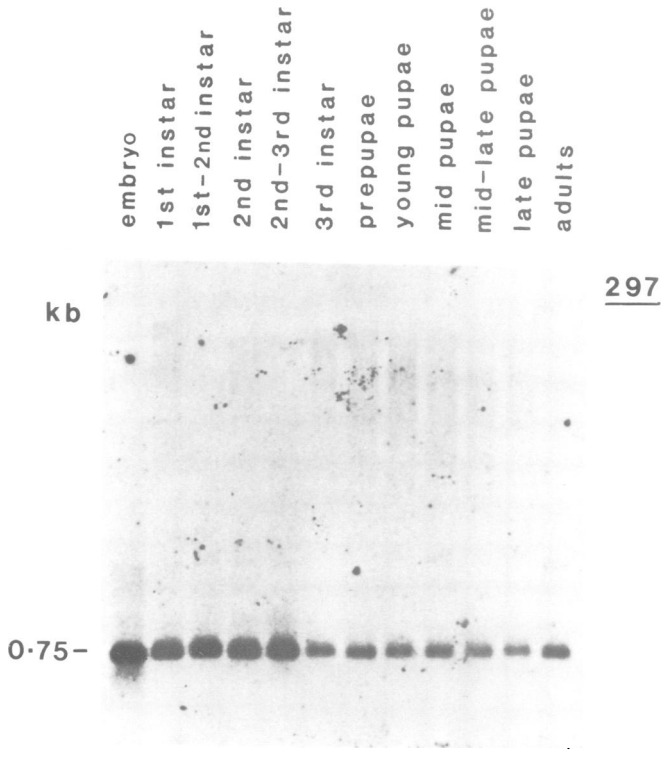

Fig. 3. Developmental analysis of 297 transcription. The same membrane utilized in Figure 1 was dehybridized and rehybridized with ^{32}P-labeled 297 DNA. The 297 probe used was a 3.5-kb *Xba*I−*Xho*I fragment containing an internal segment of the element. The same result was obtained using cDM4006, a clone containing the entire 297 element with 5 kb of flanking genomic sequences (Mossie *et al.*, 1985).

ones ~1 kb long. It is interesting to note that some of the smaller transcripts are present preferentially in late larval and pupal stages, at times when the larger RNAs are not expressed; these RNAs could appear as a consequence of alternative splice patterns of the high mol. wt RNAs, preferential stabilization of particular messages, or the use of different promoters at various development stages.

The pattern of transcription of the 297 element, analyzed on the same blot, is shown in Figure 3. This element is 7.0 kb long and contains two 415-bp direct terminal repeats (Ikenaga and Saigo, 1982). The LTRs of 297 show considerable homology with those of the 17.6 element (Kugimiya *et al.*, 1983). In contrast to the rest of the elements analyzed, we have been unable to detect a full-size transcript corresponding to an RNA that extends between both LTRs. This element encodes a single RNA 0.75 kb long that is expressed at high levels in embryos and during the first and second instars of larval development. The accumulation of this transcript decreases in third instar larvae and is maintained at this reduced level during pupal and adult states.

Transcription of the 412 class

The transposable elements we have examined so far present a unique pattern of developmental transcription. Other elements, showing similar patterns of temporal expression, can be placed into 'transcription classes'. One such class includes the 412, mdg1, 17.6 and 3S18 elements. Northern analysis of the developmental expression of these four elements is shown in Figure 4. The 412 element is 7.6 kb long with two 481-bp LTRs (Will *et al.*, 1981), and is transcribed into a 7.3-kb RNA that accumulates in embryos and first instar larvae (Figure 4A) and is not detectable again until the pupal stages of development. Adult animals express low but detectable levels of this transcript. In addition, two small RNAs 1.4 and 1.2 kb long accumulate in the embryonic and early larval stages, but are barely detectable at other times during development. The transcription of mdg1 follows a pattern very similar to that of 412. The mdg1 element is 7.3 kb, it contains two LTRs 442 bp long (Kulguskin *et al.*, 1981) and it encodes a major 7.3-kb RNA that accumulates in embryos, first instar larvae and pupae (Figure 4B), the same stages of development when 412 transcription is observed. Two small RNAs are also encoded by this element.

The next member of this class, the 17.6 element, is 7.4 kb long wih two direct repeats that extend 512 bp (Ikenaga and Saigo, 1982). Northern analysis of the transcription of this element shows the presence of a 7.3-kb transcript that is present at low levels in embryos; its accumulation increases in the first−second instar larvae then decreases in later stages of larval development (Figure 4C). This RNA is present at low levels in pupae and is scarcely detectable in adult flies. Two additional transcripts 1.4 and 1.2 kb long are also encoded by 17.6 but their temporal expression is slightly different from that of the corresponding size RNAs encoded by 412 and mdg1 in that the 17.6-encoded small RNAs accumulate at high levels during the pupal stages of development. It is interesting to note that, in spite of the high degree of sequence homology between their LTRs (Kugimiya *et al.*, 1983), the transcription of 17.6 and 297 follows very different developmental profiles, suggesting that the DNA sequences responsible for the temporal expression of these elements reside outside the conserved regions of the LTRs.

The last *Drosophila* retrotransposon that we have grouped in the 412 class is 3S18. This element is 6.5 kb long with two LTRs ~500 bp in length (Bell *et al.*, 1985). The pattern of temporal transcription of this element differs from that of other members of the same class in that 3S18 encodes two large transcripts 6.3 and 5.2 kb long; the larger of these two transcripts accumulates only in embryos and first instar larvae (Figure 4D), whereas the 5.2-kb RNA accumulates with a similar developmental profile as 412, mdg1 and 17.6, though the highest levels of expression during pupal development correspond to the pre-pupae stage. No small transcripts encoded by 3S18 could be detected, possibly due to the use of an internal fragment of the element as a hybridization probe.

Transcription of the HMS Beagle class

We have placed the HMS Beagle and Springer elements in a separate class from the elements described above. HMS Beagle is 7.3 kb long and contains 266-bp LTRs (Snyder *et al.*, 1982). It encodes a major 7.0-kb RNA that is transcribed at low levels in embryos, it accumulates in higher amounts in first instar larvae and is not detectable in third instar stages (Figure 5A). The expression of this RNA resumes in pre-pupae and is maintained at low levels for the remainder of the pupal stages of development. The amount of this RNA present in adult flies is higher than in any other stage of development, this being a major difference between these elements and members of the 412 class. Two additional RNAs 1.7 and 1.2 kb long are also detectable in some stages of development and are transcribed with a different developmental specificity (Figure 5A).

The transcription of the Springer element follows a very similar temporal profile to that of HMS Beagle. Springer is 8.8 kb long and contains two 405-bp LTRs (Karlik and Fyrberg, 1985). It gives rise to a 7.4-kb RNA that accumulates at the highest levels

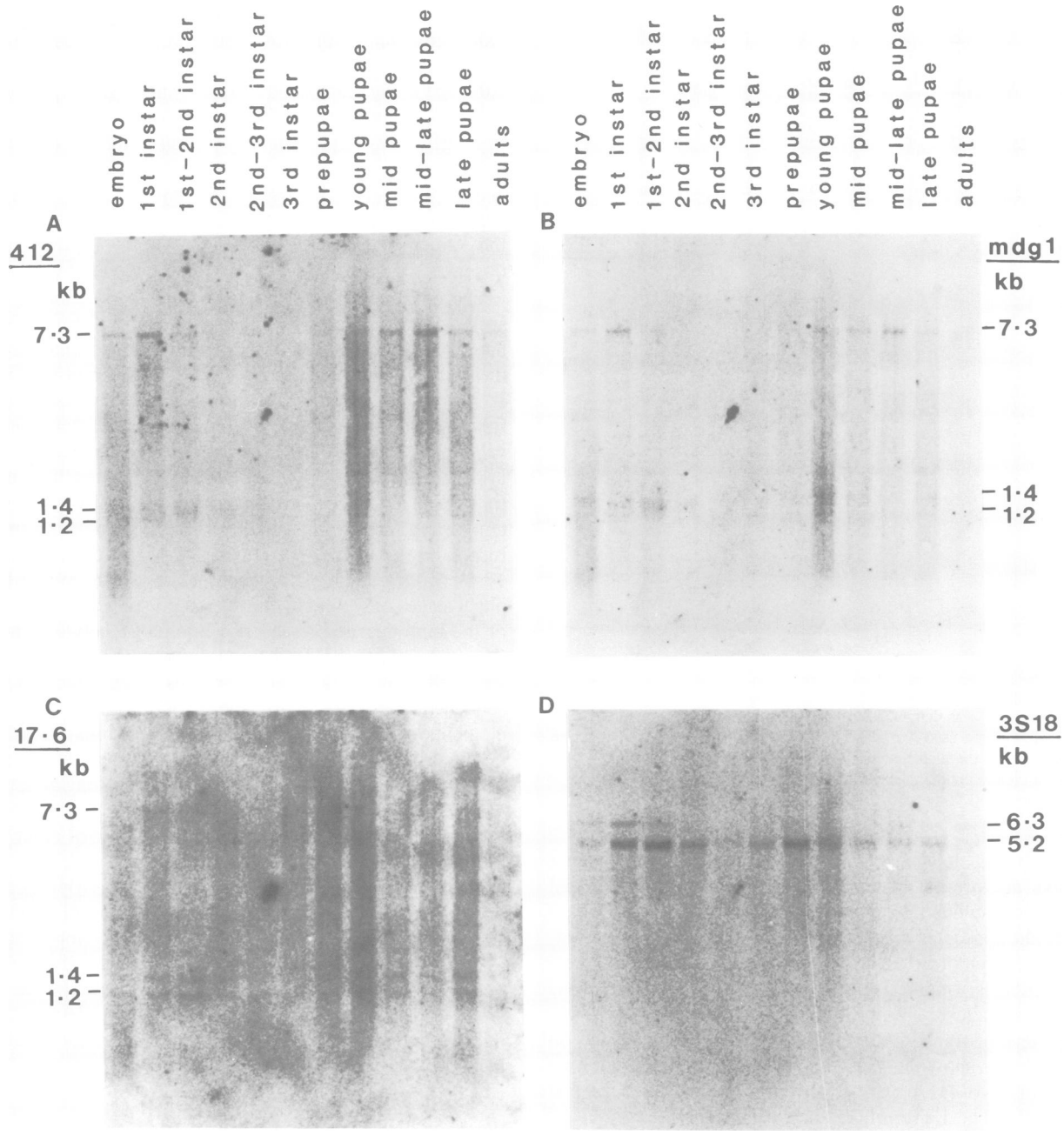

Fig. 4. Transcriptional analysis of the 412 family of transposable elements. The same filter used in Figure 1 was hybridized with cloned DNA corresponding to the 412 **(A)**, mdg1 **(B)**, 17.6 **(C)**, and 3S18 **(D)** elements. The 412 probe was an 8.5-kb *Bam*HI fragment containing the entire 412 element with 2 kb of flanking sequences. The mdg1 probe was an 8.3-kb *Eco*RI fragment containing the complete element with 1 kb of flanking genomic sequences. The 17.6 probe was a 7-kb *Cla*I fragment containing one copy of the LTR and all the internal sequences (Ikenaga and Saigo, 1982). The 3S18 probe was a 4-kb *Bam*HI fragment containing sequences internal to the LTRs only (Bell *et al.*, 1985).

in first instar larvae and adults, but is also present at low levels in other stages of development (Figure 5B). Three additional transcripts 2.0, 1.6 and 1.2 kb are also encoded by this transposable element and accumulate with the same developmental specificity as the full-length transcript.

Discussion

We have examined the pattern of developmental expression of various *Drosophila* retroviral transposable elements. The flies utilized in these studies belong to a y^2 *cv v* f^1 stock and nothing is known about the number or chromosomal locations of the different transposable elements in this stock. Several of the elements tested, such as BS, Sancho I, Sancho II, and mdg3 gave no hybridization in the conditions used in this study (data not shown). This could be due to the low rates of accumulation of the RNAs encoded by these elements or to their low copy number in the genomes of the flies tested. The elements copia, B104, 297, 412, mdg1, 17.6, 3S18, HMS Beagle and Springer are extensively transcribed in this stock, allowing a detailed study of their temporal expression. The transcription of some of these elements has also been studied in other strains of *Drosophila melanogaster*

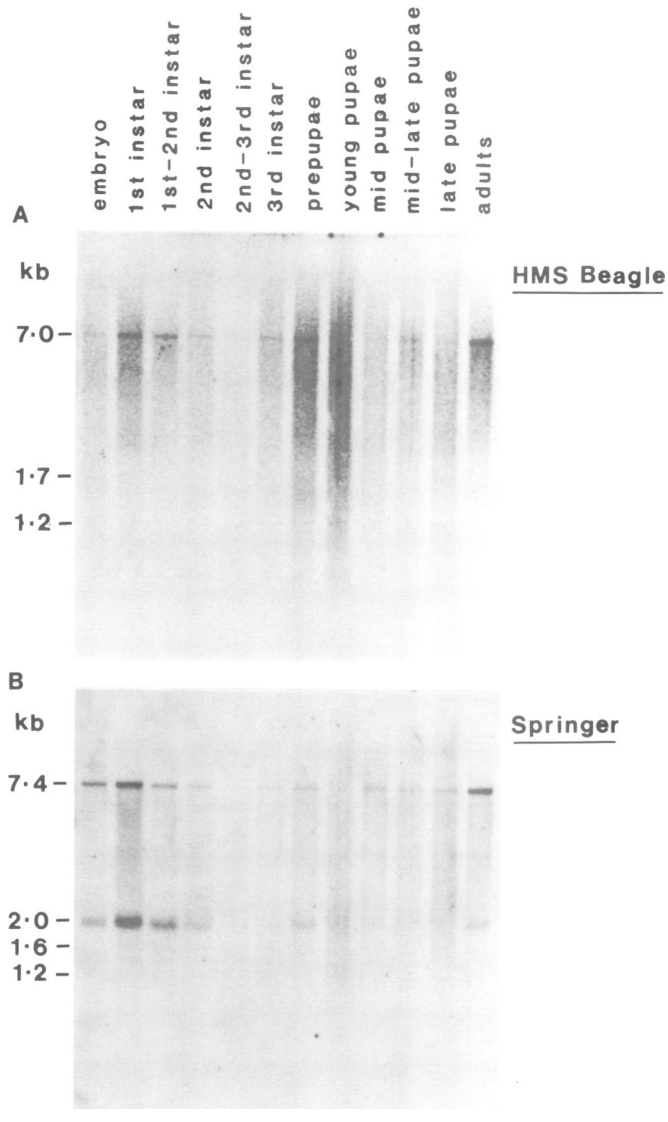

Fig. 5. Northern analysis of HMS Beagle and Springer temporal expression. The filter used in the experiments described in Figure 1 was hybridized with ^{32}P-labeled HMS Beagle **(A)** and Springer **(B)** DNAs. The HMS Beagle probe was an internal 5-kb *Xho*I fragment containing no LTR sequences (Snyder *et al.*, 1982). The Springer probe was an internal 4.5-kb *Sal*I fragment (Karlik and Fyrberg, 1985).

in which the chromosomal location of the retrotransposons is presumably different and the pattern of their developmental transcription was found to be the same as that reported above (data not shown). This suggests that the temporal expression of these elements is in general not significantly affected by their chromosomal location.

Our main conclusion is the existence of widely diverse patterns of expression among some retrotransposons whereas other elements show very similar profiles of temporal transcription. For example, the copia element is transcribed in every stage of *Drosophila* development, with the highest levels of accumulation taking place in second–third instar larvae. Other elements such as 412 and mdg1, are expressed preferentially in embryos and pupae, but their expression in adult flies is very low or non-detectable. The elements HMS Beagle and Springer encode RNAs that accumulate mostly in early larval stages and adults. In addition, the gypsy element, whose transcription and DNA structure we have previously described (Marlor *et al.*, 1986; Parkhurst and Corces, 1986a), is expressed preferentially in the pupal stages of development. These results, taken together, suggest the possibility of evolutionary relationships among the various members of a class which present similar transcription patterns, as well as the existence of evolutionary divergencies among elements with widely different profiles of developmental expression.

The differences and similarities in the patterns of developmental expression of these various elements can also be interpreted in the context of the molecular basis of their transcriptional regulation. The fact that groups of elements share the same profile of temporal expression suggests the existence of common regulatory sequences among these elements as well as the possibility that the same *trans*-acting cellular genes are involved in their control. The elements 297 and 17.6 show extensive nucleotide homology in the DNA sequence of the LTRs (Kugimiya *et al.*, 1983) while the patterns of developmental transcription are very different. This might suggest the presence of control elements outside of the LTRs. It is interesting to note that 412 and mdg1, which have very similar profiles of developmental transcription, also share a 27-bp region of perfect homology adjacent to the left-hand LTR (Will *et al.*, 1981). Thus, the information on the temporal expression of retrotransposons, together with structural analyses on their DNA sequence, could afford their evolutionary classification as well as the identification of the DNA sequences involved in their transcriptional regulation.

Materials and methods

Nucleic acids

Isolation and labeling of plasmid DNA was carried out by standard procedures (Maniatis *et al.*, 1982). Total RNA from different stages of *Drosophila* development was isolated by lysing the tissues in a Dounce homogenizer in 4 M guanidine isothiocyanate, 0.2% *N*-lauroyl sarcosine, 150 mM mercaptoethanol, 12.5 mM EDTA, and 50 mM Tris hydrochloride pH 7.5. After the addition of one volume of 100 mM sodium acetate pH 5.0 at 65°C, the RNA was extracted with phenol–chloroform at the same temperature and then kept on ice for 10 min. The cycle of phenol extraction–ice treatment was repeated three times and the RNA was then precipitated with two volumes of ethanol (Parkhurst and Corces, 1985). The poly(A)$^+$ fraction was isolated by chromatography of the total RNA on oligo(dT)cellulose.

Gel electrophoresis and filter hybridizations

Twenty-five micrograms of poly(A)$^+$ RNA obtained from different stages of *Drosophila* development were electrophoresed on a 0.8% agarose–formaldehyde gel and then transferred to a BioTrans nylon filter (ICN Pharmaceuticals Inc., Irvine, CA) in 20 × SSC. The filter was hybridized with ^{32}P-labeled DNA in a solution containing 50% formamide, 5 × SSC, 10 mM phosphate pH 6.7, 10% dextran sulfate, 0.1% Ficoll, 0.1% polyvinylpyrrolidone, 0.1% bovine serum albumin and 0.1% sodium dodecyl sulfate at 42°C, and then washed twice in 2 × SSC, 0.1% sodium dodecyl sulfate at room temperature and twice in 0.1% SSC, 0.1% sodium dodecyl sulfate at 51°C. After exposure of the film, the filter was dehybridized in 50% formamide, 10 mM phosphate pH 6.7 at 65°C and rehybridized with a different ^{32}P-labeled DNA. Under these conditions, a nylon filter could be hybridized and rehybridized an average of 20 times without noticeable loss of RNA. In those cases in which homology exists between the LTRs of different transposable elements, internal fragments were used as probes to avoid cross-hybridization artefacts. The exact nature of the DNA sequences used for each hybridization probe is given in the figure legends.

Acknowledgements

We would like to thank Drs W.Bender, A.Bogardus, E.Fyrberg, G.Georgiev, R.Levis, P.Mathers, E.Meyerowitz, L.Mizrokhi, J.Modolell, K.Mossie, M.Pellegrini, G.Rubin, K.Saigo, P.Schedl, M.Snyder, A.Spradling, H.Varmus and W.Zerges for gifts of cloned transposable elements. This work was supported by Public Health Service grants GM32036 and GM35463 from the National Institutes of Health.

References

Arkhipova,I.R., Mazo,A.M., Cherkasova,V.A., Gorelova,T.V., Schuppe,N.G. and Ilyin,Y.V. (1986) *Cell*, **44**, 555–563.

Bell,J.R., Bogardus,A.M., Schmidt,T. and Pellegrini,M. (1985) *Nucleic Acids Res.*, **13**, 3861–3871.

Boeke,J.D., Garfinkel,D.J., Styles,C.A. and Fink,G.R. (1985) *Cell*, **40**, 491–500.

Carbonare,B.D. and Gehring,W.J. (1985) *Mol. Gen. Genet.*, **199**, 1–6.

Cullen,B.R., Lomedico,T. and Ju,G. (1984) *Nature*, **307**, 241–245.

Emerman,M. and Temin,H.M. (1984) *Cell*, **39**, 459–467.

Finnegan,D.J. and Fawcett,D.H. (1986) In Maclean, N (ed.), *Oxford Surveys in Eukaryotic Genes*. Oxford University Press, Vol. 3, pp. 1–62.

Flavell,A.J. (1984) *Nature*, **310**, 514–516.

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.

Ikenaga,H. and Saigo,K. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 4143–4147.

Karlik,C.C. and Fyrberg,E.A. (1985) *Cell*, **41**, 57–66.

Kugimiya,W., Ikenaga,H. and Saigo,K. (1983) *Proc. Natl. Acad. Sci. USA*, **80**, 3193–3197.

Kulguskin,V.V., Ilyin,Y.V. and Georgiev,G.P. (1981) *Nucleic Acids Res.*, **9**, 3451–3463.

Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Marlor,R.L., Parkhurst,S.M. and Corces,V.G. (1986) *Mol. Cell. Biol.*, **6**, 1129–1134.

Mizrokhi,L.J., Obolenkova,L.A., Priimagi,A.F., Ilyin,Y.V., Gerasimova,T.I. and Georgiev,G.P. (1985) *EMBO J.*, **4**, 3781–3787.

Mossie,K.G., Young,M.W. and Varmus,H.E. (1985) *J. Mol. Biol.*, **182**, 31–43.

Mount,S.M. and Rubin,G.M. (1985) *Mol. Cell. Biol.*, **5**, 1630–1638.

Mozer,B., Marlor,R., Parkhurst,S. and Corces,V. (1985) *Mol. Cell. Biol.*, **5**, 885–889.

Parkhurst,S.M. and Corces,V.G. (1985) *Cell*, **41**, 429–437.

Parkhurst,S.M. and Corces,V.G. (1986a) *Mol. Cell. Biol.*, **6**, 47–53.

Parkhurst,S.M. and Corces,V.G. (1986b) *BioEssays*, **5**, 52–57.

Rubin,G.M. (1983) In Shapiro,J.A. (ed.), *Mobile Genetic Elements*. Academic Press, Orlando, FL, pp. 329–361.

Scherer,G., Tschudi,C., Perera,J., Delius,H. and Pirrotta,V. (1982) *J. Mol. Biol.*, **157**, 435–452.

Schwartz,H.E., Lockett,T.J. and Young,M.W. (1982) *J. Mol. Biol.*, **157**, 49–58.

Snyder,M.P., Kimbrell,D., Hunkapiller,M., Hill,R., Fristrom,J. and Davidson,N. (1982) *Proc. Natl. Acad. Sci. USA*, **79**, 7430–7434.

Will,B.M., Bayev,A.A. and Finnegan,D.J. (1981) *J. Mol. Biol.*, **153**, 897–915.

Winston,F., Durbin,K.J. and Fink,G.R. (1983) *Cell*, **39**, 675–682.

Received on 17 October 1986; revised on 12 November 1986