Volume 16 Number 12 1988 Nucleic Acids Research

Upstream and downstream transcriptional control signals in the yeast retrotransposon, TY

Alexandra M.Fulton, Peter D.Rathjen, Susan M.Kingsman and Alan J.Kingsman

Department of Biochemistry, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK

Received March 3, 1988; Revised and Accepted April 27, 1988

ABSTRACT

The yeast retrotransposon, Ty, shares many structural and functional features with retroviral proviruses. These include production of a terminally redundant major transcript. There are also two less abundant transcripts of 5.0kb and 2.2kb. Ty transcription is regulated by cell-type, that is it is reduced 5-20 fold in a/α diploids as compared to haploids. However control of expression of Ty is not well understood. By deletion analysis we have identified regions of the element which are involved in the activation and regulation of transcription. These signals are found both upstream and downstream of the mRNA start site. The downstream signals are within the region encoding the major Ty proteins. This organisation of transcriptional control signals is discussed with reference to the organisation of control signals in other yeast genes and in retroviral proviruses and other retro-elements.

INTRODUCTION

The yeast transposable element, Ty, is representative of a class of eukaryotic mobile genetic elements which includes the copia and copia-like elements of Drosophila (1), the TED element of Trichoplusia ni (2) and mouse intracisternal A-type particles (3). These elements or retrotransposons (4) are remarkably similar to avian and mammalian retroviral proviruses (5). They have the same overall structure of long terminal direct repeats (LTRs) flanking an internal region. They produce a terminally redundant major transcript and direct the assembly of virus-like particles (VLPs) with associated reverse transcriptase activity (6, 7, 8, 9, 10). Ty elements have also been shown to transpose to new genomic locations via an RNA intermediate (4).

The Ty element is 5.9kb in length with long terminal repeats (delta sequences) of ~335 base pairs flanking an internal or epsilon region (11). Three different sized transcripts have

been identified: 5.7 kilobases (kb), 5.Økb and 2.2kb (8, 12, 9). The 5.7kb RNA is the major transcript. It initiates around the Xho I site in the 5' delta and extends into the 3' delta such that it is terminally redundant for approximately 5Ø bases (Figure 1a) (8). It is this transcript which acts as the RNA intermediate in Ty transposition (4). The 5.7kb transcript is subject to cell-type specific control, that is transcription is repressed in a/α diploids compared to haploids and homozygous diploids (13).

Integration of a Ty element upstream of a gene such that the Ty element and the gene are divergently transcribed results in transcription of the gene being increased and placed under cell-type specific control, a reduction in transcription being observed in a/α diploids. These are known as Regulated Overproducing Alleles responding to Mating type and the effect is often called the ROAM effect (14). It is not known whether the cell-type specific control of Ty transcription is mediated via the same signals which exert the ROAM effect.

There is some confusion over the nature of the Ty 5.Økb transcript. It has been reported that it is 5' coterminal with the 5.7kb transcript, terminating within the epsilon region of the element (8). However, a 5.Økb transcript initiating within the epsilon region has also been described (12). This transcript is normally repressed but is seen at high levels in spt 3⁻ (suppressor of Ty) mutants (Winston et al, 1984). It is therefore unclear whether there are one or two 5.Økb transcripts.

No function has yet been ascribed to the minor 2.2kb transcript (9). A 2.Økb copia RNA has been identified (15) and in vitro this is translated to copia proteins (16). The 2.2kb Ty transcript encompasses the entire coding region of TYA and it is possible that it could be involved in the production of TYA proteins and the protease necessary for the cleavage of these proteins (17).

Despite the detailed knowledge which has been obtained regarding the structure and function of Ty elements, our understanding of the control of expression of Ty, or any other retrotransposon, is limited (18). Recently, Liao et al. (19) have analysed control of transcription in a class II Ty element. They define a TATA box and a putative UAS in the 5' delta of the element, both necessary for maximal expression of the element. In this paper we analyse control of transcription in a class I Ty

element, Ty1-15 (20). We demonstrate that the only transcriptional control signal in the 5' delta is an AT rich region (a potential TATA box). The 5' delta does not contain any transcriptional activators analagous to a yeast UAS (21) or a viral enhancer (22). We also investigate cell-type specific control of Ty transcription and show here that signals necessary for control of the Ty 5.7kb and 2.2kb transcripts are found within the internal, or epsilon, region of Ty.

MATERIALS AND METHODS

Bacterial and yeast strains and media

E.coli strain AKEC28 (C600 *thrC leuB6 thyA trpC1117 hsdRk hsdMk*) was used for plasmid manipulation and preparation. *Saccharomyces cerevisiae* strains: MD40/4c = α *ura2 trp1 leu2-3 leu2-112 his3-11 his3-15*; AH22 = a *leu2-3 leu2-112 his4-519 can*^{R}; SF30 = a/α *ura2/URA2 trp1/TRP1 leu2-3/leu2-3 leu2-112/leu2-112 his3-11,his3-15/HIS3 his4-519/HIS4 can*^{R}*/CAN*. *E.coli* media were prepared according to Miller, 1972 (23). Yeast media were prepared according to Hawthorne and Mortimer, 1960 (24).

Yeast transformation

Yeast was transformed according to the method of Hinnen *et al.*, 1978 (25).

DNA isolation

Plasmid DNA was isolated according to the method of Chinault and Carbon, 1979 (26) and minipreparations of plasmid DNA were prepared as described by Holmes and Quigley, 1981 (27).

Enzymes, fragment purification and nick translation

Enzymes were purchased from BRL or NBL and used according to the manufacturers' instructions. DNA fragments were purified from agarose gels by the method of Tabak and Flavell, 1978 (28). Nick translations were performed using a kit supplied by BRL and used according to the manufacturer's instructions. Probes were labelled to $4x10^7$-$1x10^8$ cpm ug^{-1}. Labelled dNTPs were supplied by Amersham plc.

DNA sequencing

DNA sequencing was carried out by the method of Sanger *et al.*, 1977 (29) using reagents provided by BRL.

RNA isolation and blotting

Total yeast RNA was isolated as described previously (30) from cultures grown to a density of 4-$6x10^6$ cells/ml, transferred to nitrocellulose filters and hybridised (31). Probe fragments

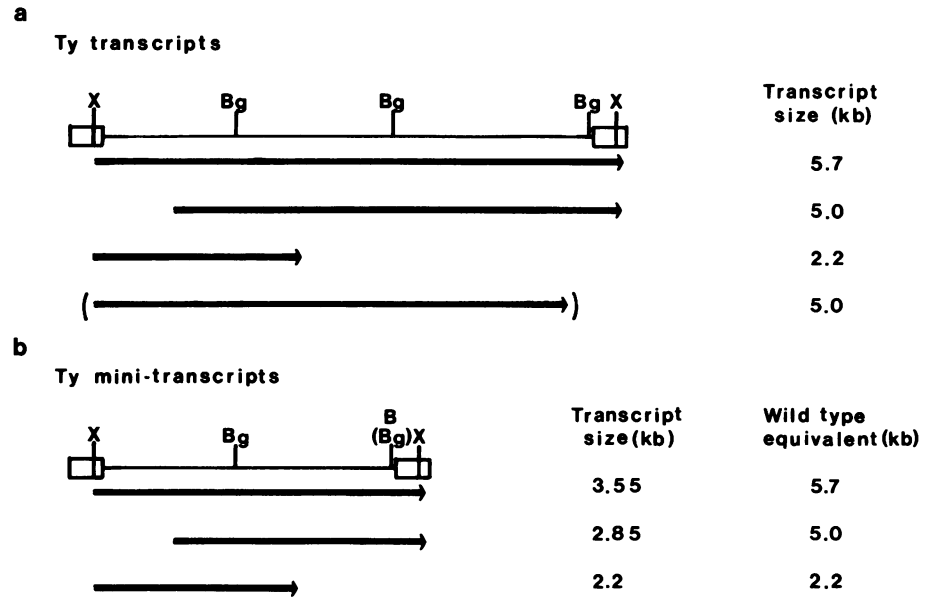


Figure 1 (a) Organisation of Ty1-15 showing identified transcripts as thick black arrows. (b) Structure of mini Ty elements showing expected transcripts as thick black arrows. B=BamH1; Bg=BglII; X=XhoI.

were labelled by nick translation (32). The probes used were the 1.25kb Pvu II/Bgl II fragment from Ty1-15 (Figure 1) for Figure 3b) and a human α2-interferon probe (33) for Figure 6b. A phosphoglycerate kinase (PGK) BamHI/BglII fragment from pMA778 (34) was used as a probe for a loading control in Figures 3b and 6b. For Figure 5b PGK and rRNA specific probes were purified from EcoRI digested plasmids pMA27 (35) and pYIRG12 (36). respectively.

S1 nuclease protection analysis

S1 nuclease protection analysis was carried out as described previously (37) except that DNA fragments were end-labelled using γ-ATP and polynucleotide kinase and the protected DNA fragments were separated on a denaturing 6% polyacrylamide gel. The probe used was a 240bp Sau 3A fragment extending from 90bp upstream of the XhoI site in the 5' delta into the epsilon region of the element (Figure 1a). The Sau3A sites are boxed in Figure 4. RNAs used in this analysis were aliquots of the same preparations used in the Northern blot analysis.

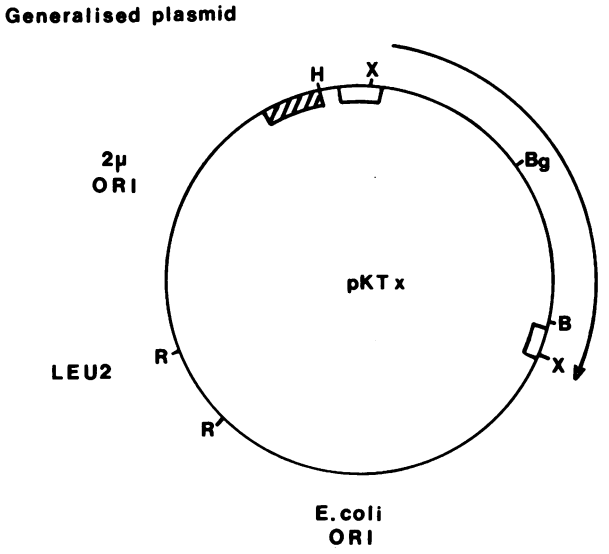


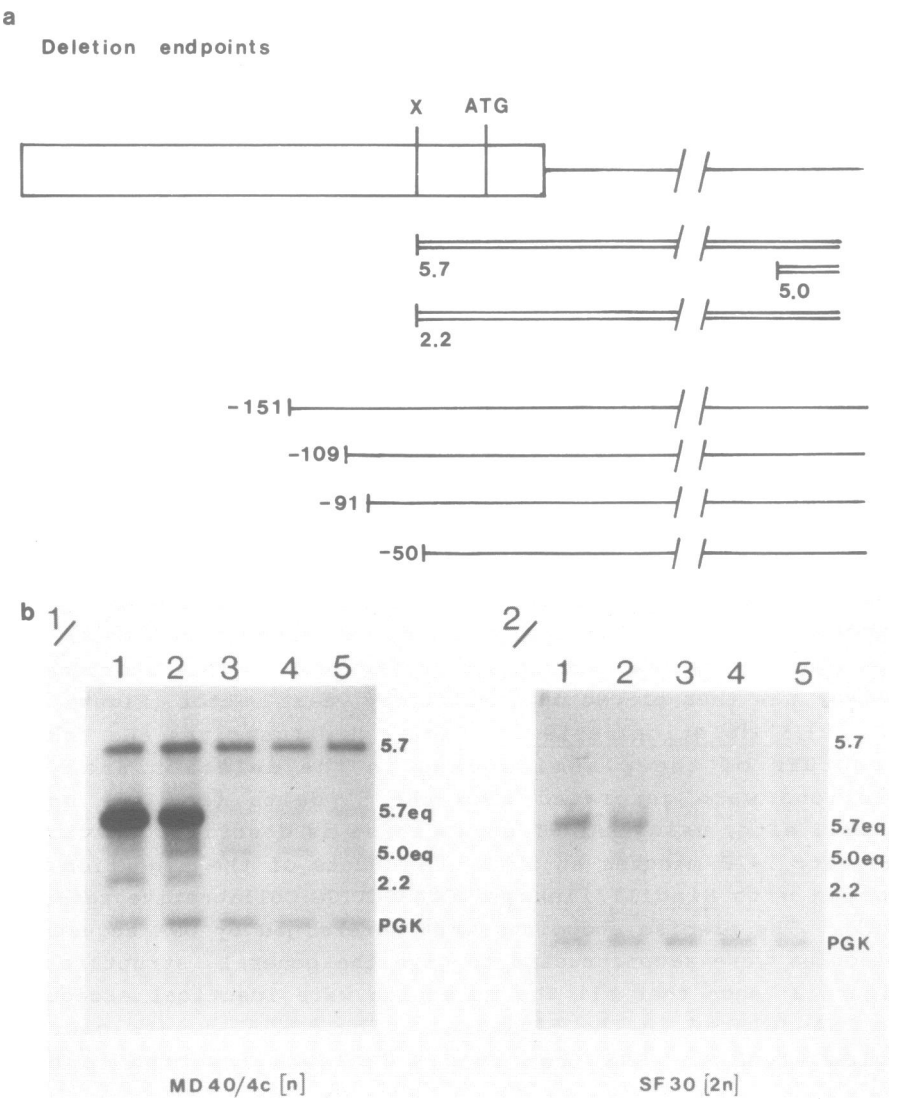
Figure 2: pKTx. Generalised plasmid illustrating reconstruction of the 5' deletion series molecules. ▨ PGK 5' untranscribed region, -1400 to -595. B=BamHl; Bg=BglII; H=HindIII; R=EcoRl; X=XhoI.

Plasmid constructions

(a) Constructing a marked Ty element. To distinguish Tyl-15 from the 30-35 chromosomal copies of Ty the 2.15kb BglII fragment at the 3' end of the epsilon region of Tyl-15 was deleted and a BamHI linker was inserted (Figure 1). The expected transcripts from this shortened element and their wild type equivalents are illustrated in Figure 1. This shortened Ty element was then placed on a multicopy yeast vector (Figure 2).

(b) 5' deletion series. Figure 2 illustrates the general structure of the plasmids used in the deletion analysis. Deletions were generated into the 5' delta from the unique HindIII site, using BAL 31 exonuclease as described previously (38), for 2-10 minutes at 15°C. Products of the digestion were ligated with HindIII linkers (CCAAGCTTGG;Collabrative research, Inc.). The precise endpoints were determined by DNA sequencing. Molecules were reconstructed to give the general structure pKTx (Figure 2) such that all the molecules were identical except for the DNA deleted in the 5' delta. The PGK 5' non-transcribed region -1400 to -595 from pMA778 (34) was inserted to block

transcripts known to initiate in the LEU 2/2u region of the plasmid (J. Mellor, personal communication). This fragment is known to prevent readthrough transcription from plasmid sequences (34). Deletion endpoints are illustrated in Figure 3a and are numbered relative to the ATG of TYA (39). The undeleted control contains an intact 5' delta and flanking yeast chromosomal DNA for Ty1-15 (20).



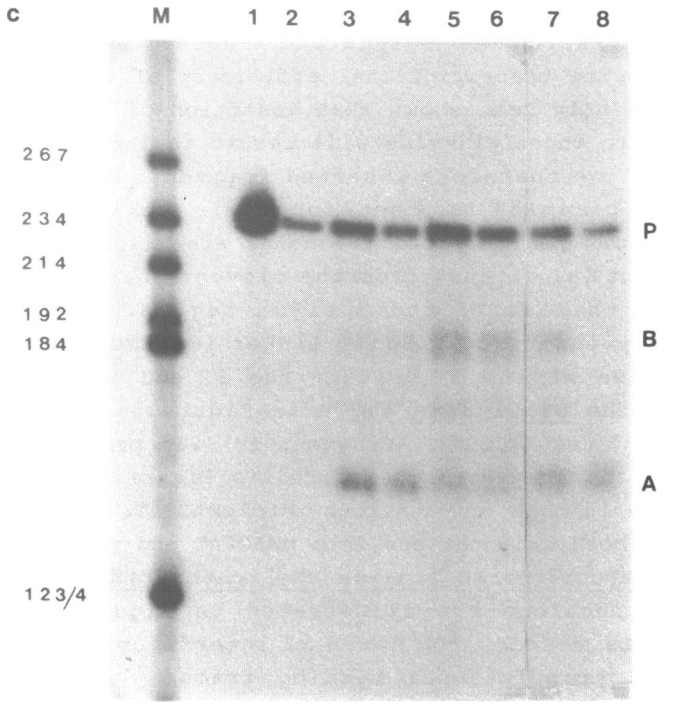


Figure 3: (a) Deletion endpoints through the 5' delta of Ty1-15. ATG: initiation codon of TYA. Endpoints are numbered with reference to the A of the ATG. X=XhoI. (b) Northern blot of RNAs from the 5' deletion series transformed into yeast strains haploid MD40/4c (panel 1) and diploid SF30 (panel 2). For each panel: Lane 1, undeleted control. Lane 2, -151 endpoint. Lane 3, -109 endpoint. Lane 4, -91 endpoint. Lane 5, -50 endpoint. c) S1 nuclease protection analysis of RNA from haploid cells. Lane M, size markers. pBR322 digested with HaeIII. Lane 1, untreated probe. Lane 2, S1 treated probe. Lane 3, undeleted control. Lane 4, -151 endpoint. Lane 5, -109 endpoint. Lane 6, -91 endpoint. Lane 7, -50 endpoint. Lane 8, control for levels of RNA from chromosomal copies of Ty. RNA used for this track was isolated from a yeast transformant containing plasmid pMA91-1 (Tuite et al, 1982). This plasmid has the same selectable marker as the other plasmids used in this experiment but does not contain any Ty sequences. P=undigested probe. A=146 base protected band from 5.7kb and 5.7kb equivalent transcripts. B=184-190 base protected bands from plasmid specific transcript through or initiating in the 3' delta.

(c) Constructions to test for transcriptional activation signals in the 5' delta. These constructions are illustrated in Figure 5a. The vectors used to assay for transcriptional

activators were pMA27AT and pMA763. These plasmids contain a copy of the yeast phosphoglycerate kinase gene (PGK) without a UAS, so reducing transcriptional efficiency of the gene (34, 40). It has previously been shown that insertion of a transcriptional activator into these plasmids will result in increased expression of PGK (41). We therefore inserted fragments of the Ty 5' delta into these plasmids to test for their ability to activate transcription. The inserted fragments are illustrated in Figure 5a. Fragment (a) extends from the eleventh base of the 5' delta to the eighth base of the epsilon region. The 5' end was generated by inserting a BglII linker (CAAAAGATCTTTTG) at the eleventh base of the 5' delta. The 3' end was generated by filling in the HinPI site shown in Figure 5a and inserting a BamHI linker (CCGGATCCGG). Fragment (b) was produced by cutting fragment (a) at the XhoI site marked in Figure 5a. Fragment (a) was inserted in both orientations into pMA27AT. Fragment (b) was inserted in both orientations into pMA27AT and pMA763.

(d)Constructions to analyse the role of the epsilon region. These constructions are illustrated in Figure 6a. The Ty transcript is marked with human α2 interferon cDNA inserted in the correct translational reading frame. 1 and 2 after a plasmid number indicate the interferon fragment inserted to maintain the correct reading frame (42). pKT231-1 has been described previously (42). In pKT202-2 the interferon cDNA is inserted at the PvuII site in Ty1-15 via insertion of a BglII linker (CAAAAGATCTTTTG) and in pKT48-2 the interferon cDNA is inserted at deletion endpoint 48 (43) which is 7 bases into the 5' delta (Figure 4).

RESULTS

Sequences upstream of the RNA initiation site are required for transcription of Ty.

To study Ty transcription it is necessary to analyse a single element and to distinguish its response from the ~35 endogenous Ty elements. This was achieved by constructing a mini Ty element (Figure 1b). To produce the mini element the 2.15kb Bgl II fragment at the 3' end of the epsilon region of Ty1-15 (Figure 1a) was deleted. The expected mini- transcripts and their wild type equivalents are illustrated in Figure 1b. We will discuss the mini-transcripts as wild type equivalents.

To investigate the role of sequences upstream of the RNA

initiation site in the control of Ty transcription we created a deletion series into the 5' delta as described in Materials and Methods. Deletion endpoints are illustrated in Figure 3a and are numbered relative to the first base of the ATG initiation codon of the major Ty protein, p1 (Figure 4) (30). Molecules were reconstructed to give the general structure pKTx (Figure 2) to ensure that all the plasmids were identical except for the DNA deleted in the 5' delta. To prevent readthrough transcription from other plasmid promoters a region of DNA that lacks any transcriptional control elements was inserted adjacent to the deletion endpoints at the unique Hind III site. The DNA fragment used was the PGK 5' non-transcribed region, -1400 to -595 (34).

Plasmids containing the deletions described in Figure 3a were transformed into the haploid yeast strain MD40/4c. Figure 3b, panel 1, is a Northern blot of RNA obtained from these transformants. The blot was probed with a Ty specific probe to assess the levels of Ty transcripts and also with a PGK specific probe to check for equal loading of all tracks (see Materials and Methods). Plasmid copy numbers were checked as described in Mellor et al, 1985 (35) and found to be consistently high and equivalent in each transformant (data not shown).

By comparison with the undeleted control (lane 1) deleting to within 151 bases of the ATG (lane 2) has no effect on the levels of the 5.7kb equivalent transcript. However in deletions to 109 bases from the ATG (lane 3) and beyond (lanes 4,5) this transcript is no longer detectable. Signals which are important for the production of the 5.7kb transcript are therefore located in the 42 bases deleted between the -151 and the -109 endpoints and sequences upstream of -151 are not required for normal levels of expression of this transcript. Only 99 bases upstream of the RNA initiation site are therefore required for normal expression levels. Densitometer scans of the autoradiographs support these conclusions (data not shown).

Analysis of the wild type Ty transcription pattern (Figure 3b, panel 1, lanes 1 and 2) shows the presence of the minor 2.2kb transcript (9). The response of this transcript to deletion of the 5' delta sequences is identical to that of the 5.7kb equivalent transcript. This suggests that the same promoter signals control the 5.7kb and the 2.2kb transcripts.

We have shown that the region between -151 and -109 contains a sequence for normal expression levels of the Ty 5.7kb and 2.2kb

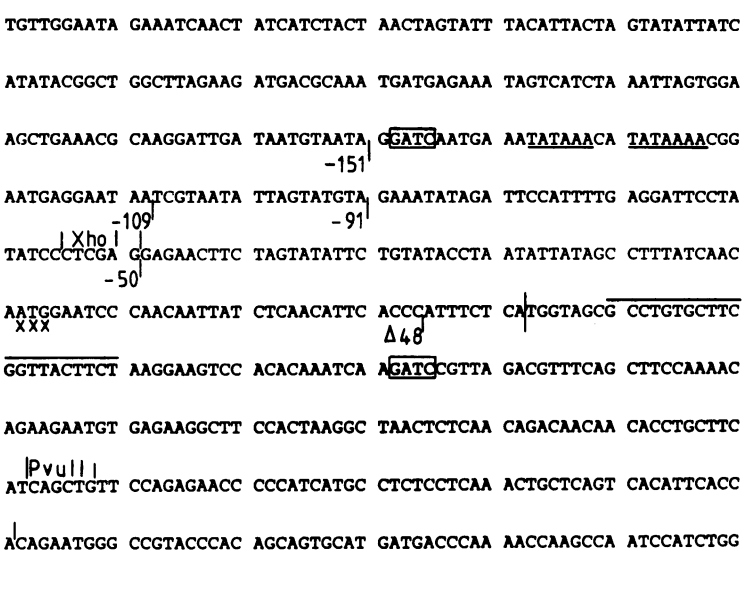


Figure 4: DNA sequence of the first 659 bases of Ty1-15. xxx, initiation codon of TYA. Underlined sequences, potential TATA boxes. Overlined sequences, homology to al:α2 consensus sequence (Miller et al, 1985, Siliciano and Tatchell, 1986), GATC, Sau3A sites defining probe used in S1 nuclease protection analysis. | delta/epsilon boundary. Numbers denote deletion endpoints.

transcripts. This region encompasses a very AT rich region (75% AT) of the 5' delta (Figure 4). This region could function as an upstream activator sequence as suggested for some other yeast genes e.g. HIS3 (44). Alternatively the region could provide TATA box function and two potential TATA boxes can be identified (Figure 4).

To assess the effect of the deletions on initiation of the 5.7kb transcript we carried out an S1 nuclease protection analysis (Figure 3c). The RNAs used in this analysis were aliquots of the same preparations run on the Northern gels. The probe used was a 240bp Sau 3A fragment (Figure 4). The protected band of approximately 146 bases (A) (Figure 3C, lane 3) confirmed that in the intact 5' delta the major 5.7kb transcript initiates the XhoI site (8) most probably at the A or G of the Xho I site (CTCGAG). Deleting to endpoint -151 (to within 99 bases of the RNA initiation site) does not change the RNA initiation site

(lane 4). In further deletions beyond endpoint -151 the protected band of 146 bases (Figure 3c, lane 5,6,7) is similar in intensity to the control track and is due to transcripts from the chromosomal copies of Ty (figure 3c, lane 8). In the deletions which lack the plasmid-encoded 5.7kb equivalent transcript (Figure 3c, lanes 5,6,7) new plasmid specific protected bands (B) of 184-190 bases are seen. These are consistent with a new transcript through, or initiating in, the 3' delta as the probe is homologous to 190 bases of the 3' delta.

In the Northern blot, Figure 3b, panel 1, a 5.0kb equivalent transcript can be detected at very low levels in all the deletions. This transcript is not clearly visible in the control track (lane 1). This may indicate that signals repressing the 5.0kb transcript are located between the 5' end of the delta and the -151 deletion endpoint but the low levels of transcript observed do not allow a firm conclusion to be drawn. The presence of the 5.0kb transcript from all the deleted molecules suggests that it is controlled independently of the 5.7kb equivalent and 2.2kb transcripts.

Mating-type control of Ty transcripts

Although it is known that the levels of the Ty 5.7kb transcript are dependent on the mating-type status of the cell (13) the signals responsible for this control have not been identified. It is also not known whether the 2.2kb transcript is subject to this control. We therefore examined Ty transcription in haploid cells (MD40/4c) and a/α diploid cells (SF30) (Figure 3b) to identify the location of the signals which respond to the mating-type status of the cell. In the undeleted control it is clear that both the 5.7kb and 2.2kb transcripts exhibit cell-type specific control, that is they show a/α diploid repression. Comparison of the wild type and the -151 deletion in haploid cells (Figure 3b, panel 1) and a/α diploids (Figure 3b, panel 2) show that sequences upstream of deletion endpoint -151 are not necessary for a/α diploid repression of the 5.7kb equivalent and 2.2kb transcripts. Further deletion gives no information on the mating type control of the 5.7kb and 2.2kb transcripts as these transcripts are no longer detectable.

The mating-type control of the 5.0kb transcript is not addressed due to the low levels of expression of this RNA.

The 5' delta does not contain transcriptional activators.

As discussed above the AT rich region between -151 and -109

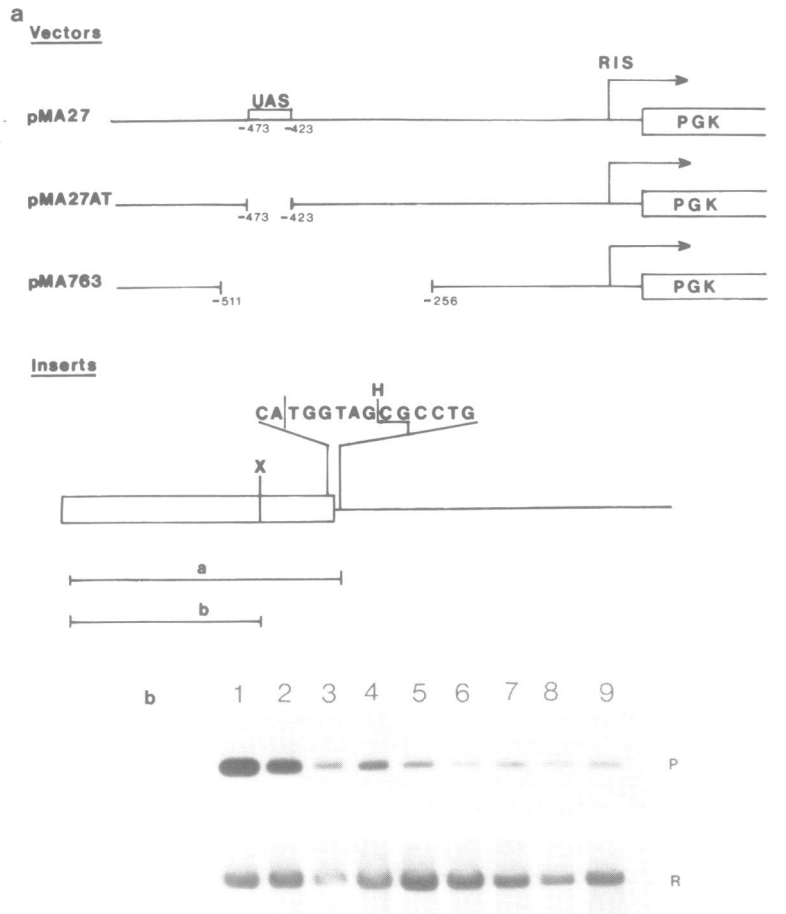


Figure 5: (a) Constructions used to screen for activation signals in the 5' delta. ☐ PGK UAS. RIS = RNA initiation site. ☐PGK☐ PGK coding sequences. H=HinPI; X=XhoI. (b) Northern blot of RNA from MD40/4c transformants containing the constructions illustrated in figure 5a. Lane 1, pMA27. Lane 2, pMA763. Lane 3, no plasmid. Lane 4, pMA763+b. Lane 5, pMA763+b. Lane 6, pMA27AT+b. Lane 7, pMA27AT+b. Lane 8, pMA27AT+a. Lane 9, pMA27AT+a. Arrows indicate the direction the inserted fragment would be transcribed in Ty. P=PGK transcript. R=ribosomal RNA loading control.

which is required for normal levels of Ty transcription (Figure 3b, panel 1) could act as a UAS or a TATA box. To assay for a UAS in the 5' delta sequence we therefore inserted the delta region and delta subfragments into pMA27AT. This plasmid

contains a copy of the phosphoglycerate kinase (PGK) gene without a UAS. Transcriptional efficiency of the gene is drastically reduced (34, 40). Transcription can be increased by inserting transcriptional activators in place of the PGK UAS (41). The fragments used in this analysis are illustrated in Figure 5a. Fragment **a** encompasses almost the entire 5' delta extending from the eleventh base of the 5' delta to a point seven bases past the end of the 5' delta (bases -290 to 45). This caused no activation of transcription (Figure 5b, lanes 8,9).

We then subdivided fragment **a** at the Xho I site which conveniently lies around the RNA start site. This gave fragment **b** (bases -290 to -56), a delta subfragment containing the sequences upstream of the RNA initation site (Figure 5a). This was also inserted into pMA27AT. No activation of transcription was observed (Figure 5b, lanes 6,7). We checked that this lack of activation was not dependent on distance from the PGK RNA initiation site by using a larger deletion of the PGK promoter, pMA763 (34). This placed fragment **b** closer to the RNA start site of PGK (Figure 5a), in a position similar to where it is located in Ty. Again no activation of transcription was observed (Figure 5b, lanes 4,5). A delta subfragment (bases -150 to 90) containing sequences downstream of the XhoI site has also been shown to lack activation signals (41). Ty does not therefore have a UAS or any activator in the 5' delta sequence as neither the delta nor the delta subfragment caused any activation of transcription in this heterologous assay system (Figure 5b) which is known to respond to a range of heterologous activators (41). We therefore propose that the region -151 to -109 is important because it contains a TATA element. This region lies within 57-99 bases of the Ty RNA initiation site, which places it in the normal functional range, for yeast genes, of a TATA box rather than a UAS (21, 45).

Epsilon sequences are necessary for cell-type specific control and transcriptional activation of Ty transcription.

As no activation signals were identified in the 5' delta in our assay it appears that Ty transcription is either activator independent or that activators are located in the epsilon region. We had also failed to locate any cell-type specific controls in the 5' delta. We therefore decided to analyse the importance of the epsilon region in activation and cell-type specific control of Ty transcription. In this experiment the delta sequences

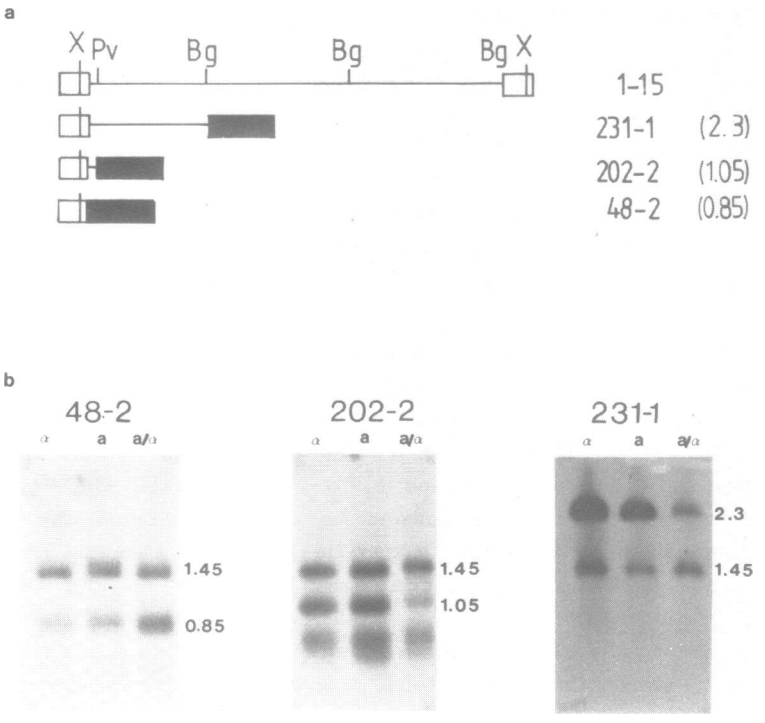


Figure 6: (a)Constructions used to analyse the role of the epsilon region in the control of Ty transcription. numbers in brackets = transcript size. Bg=BglII; Pv=PvuII; X=XhoII. (b) Northern blot of haploid MD40/4c, haploid AH22 and diploid SF30 transformants containing the construstions illustrated in figure 6a. Panel 1, pKT48-2. Panel 2, pKT202-2. Panel 3, pKT231-1. For each panel: Lane 1, AH22. Lane 2, MD40/4c. Lane 3, SF30. The 1.45 kb chromosomal PGK gene transcript which was used as an internal standard is indicated.

upstream of the RNA initiation site were retained and sequences were deleted from the epsilon region. Ty transcripts were marked with human α2 interferon cDNA, inserted in the appropriate Ty open reading frame. Interferon cDNA was inserted at three points in Ty1-15 (Figure 6a). These were at the first Bgl II site to produce pKT231-1 (42), at the first Pvu II site to produce pKT202-2 and a point 7 bases within the 3' end of the delta to produce pKT48-2 (Figure 4) (43). Plasmid copy number was consistently high in all transformants (data not shown).

Deleting sequences downstream of the first Pvu II site in Ty1-15 has no effect on the cell-type specific control of the

5.7kb transcript (pKT2Ø2-2 and pKT231-1, Figure 6b, panel 2 and 3). Control is lost however in the deletion which extends 7bp into the 5' delta (pKT48-2, Figure 6b, panel 1). These conclusions are supported by densitometer scans of the autoradiographs (data not shown). a/α diploid repression of the 5.7kb transcript therefore requires a sequence located in the 148 bases deleted between pKT2Ø2-2 and pKT48-2. Homozygous diploids were also examined. Transcript levels from all constructs were the same in these diploids as in haploid cells (data not shown).

Figure 6b, panel 1 also shows that the 5' delta will direct transcription of Ty, although at very low levels. These signals are however not sufficient to explain the normal expression levels of Ty transcripts. The major reduction in Ty transcription levels is seen when sequences between the first Pvu II site and the first Bgl II site are deleted (Figure 6b, panels 2 and 3, lane 1). This is consistent with there being, between these sites, transcriptional activation signals which elevate the basal levels of transcription to the normally observed levels of Ty transcription.

DISCUSSION

An array of signals both upstream and downstream of the RNA initiation site is involved in controlling Ty transcription. The first 15Ø nucleotides of the 5' delta are not required for normal levels of transcription of the Ty 5.7kb and 2.2kb RNAs but further deletion of 42 nucleotides encompassing an AT rich region of the 5' delta from -151 to -1Ø9 dramatically reduces the levels of Ty transcription. This AT rich region appears to function as a TATA box. There is no UAS in the 5' delta of this class of Ty element. This contrasts with the results of Liao et al., 1987 (19) who propose a UAS in the 5' delta of the class II element which they investigated. The 5' UAS activity in the class II element was however only detected in conjunction with the downstream elements. The sequence of the 5' deltas is identical in the area of the putative UAS (19; Figure 4) suggesting that the requirements of the downstream activators must be significantly different in the two classes of element. The Tyl-15 5' delta can, however, direct low basal levels of Ty transcription. It is interesting in this context that the 5' delta contains two regions of homology (>7Ø%) (see Figure 7) to the sequences recognised by the BAS2 gene product recently

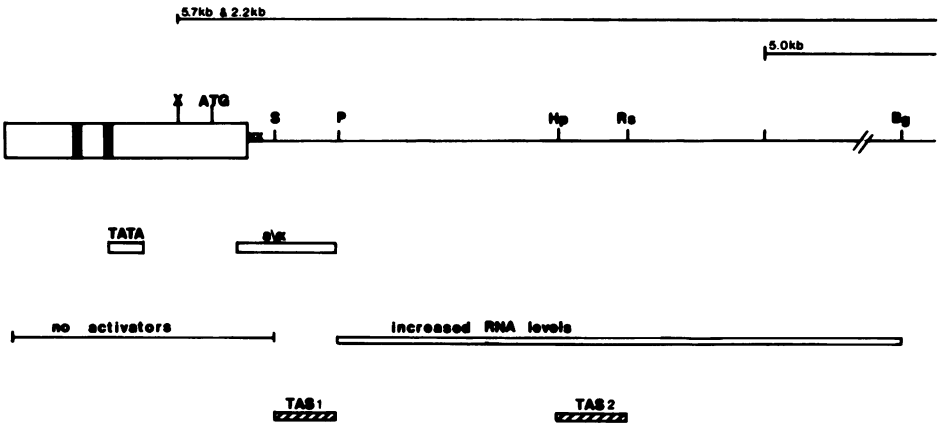


Figure 7: Summary of the location of signals involved in the control of transcription of Ty and neighbouring genes. ▭ signals affecting Ty transcription. ▨ Signals affecting the transcription of neighbouring genes. TAS = Ty activator sequence. ■ homology to the BAS2 consensus sequence (Arndt et al, 1987) -205 to -188 and -155 to -138. xx Homology to the α1:α2 consensus sequence (Miller et al, 1985, Siliciano and Tatchell, 1986). Bg=BglII; Hp=HpaI; P=PvuII; Rs=RsaII; S=Sau3A; X=XhoI.

identified upstream of HIS4 and PHO5 (46). In the case of HIS4 these sequences are required for the low basal expression levels of the gene.

Considering the close relationship between Ty and retroviruses it was surprising to find no major transcriptional activators within the 5' delta. Transcriptional control signals have, however, also been identified within retroviral transcriptional units e.g., in Bovine leukaemia virus (49) and in avian retroviruses (50). However, where these downstream activators have been found (47, 49, 50) there have also been major upstream activation signals. In the copia elements of Drosophila, transcriptional control signals may also be located 3' to the RNA start site (51). It therefore appears that the downstream location of transcriptional control signals may be a feature of retrotransposons.

Cell-type specific control of Ty transcription is clearly mediated via sequences located within the epsilon region of the element. This contrasts with the other yeast genes which exhibit a/α diploid repression. These include HO, RME, MAT α1,

STE 5 and KAR 1 (52, 53, 54, 55). Control of HO and MAT α1 requires interaction of a1:α2 with an upstream consensus sequence (56, 57) and no downstream sequences have been shown to be required. A search for homology to the a1:α2 upstream consensus sequence (T/CCA/GTGTnnA/TnAnnTACATCA) (56) in the region between the deletion endpoint in pKT48-2 and the PvuII site found GCCTGTGCTTCGGTTACTTCT (Figure 4) to be the best fit. With one insertion in the Ty string 15/20 nucleotides fit the consensus. This is weaker homology than that shown by the upstream regions of the a1:α2 controlled yeast genes (56, 57). These differences in mating type control of Ty and other a/α repressed genes suggests that the mechanism for control of Ty may be different, either using a1:α2 in an indirect way or independent of this control pathway.

Figure 7 shows the dispersed distribution of the Ty transcriptional control signals. Signals within Ty are also known to activate neighbouring genes and to place them under cell-type specific control (the ROAM effect) (41). These signals are called Ty activating sequences (TAS). There have been various other reports of epsilon sequences being required to mediate the ROAM effect. These sequences include an enhancer-like sequence 290 bases from the ATG with homology to the SV40 enhancer (58) in class II elements and two sequence blocks 350 and 520 bases from the ATG with homology to the SV40 enhancer and a1:α2 consensus sequences in class I elements (59, 60). However, the ROAM effect can be produced by the 5' Sau3A/PvuII fragment (bases 90 to 181, Figure 7) of Ty1-15 (TAS1) (41) which is also within the sequences necessary for mating type control of Ty transcription (bases 33-181,Figure 7). It may be that the mating-type control of Ty transcription and the ROAM effect are mediated via the same sequences. If this is the case then the weak homology to the a1:α2 sequence (bases 48-68) is unlikely to have any functional significance in diploid repression of Ty transcription as it lies outside the fragment required for the ROAM effect defined by Rathjen et al, 1987 (41).

'Host' control of transcription is not unique to Ty. For example copia elements in Drosophila show maximum levels of expression in larvae (16), the copia-like element 412 is expressed to the highest level in late embryos (61), the mouse intracisternal A-type particles (IAPs) show elevated expression in differentiated embryonal carcinoma cell lines (3) and the

mouse VL3Ø element has increased expression levels in proliferating cells (62). Cell-type specific control is also characteristic of some retroviruses e.g. Moloney murine sarcoma virus, MoMSV (63), bovine leukaemia virus, BLV (49) and mouse mammary tumour virus, MMTV (64) and sometimes involves signals within the transcriptional unit (BLV and MMTV, (49, 5Ø)). The widespread occurrence of cell-type specific control exerted on retro-elements by their `hosts' suggests that there may be a fine balance between retro-element expression and proliferation and cell viability.

ACKNOWLEDGEMENTS

We would like to thank Jane Mellor for helpful discussion and Mr. Ken Johnson for photography. This work was supported by an M.R.C. project grant to S.M.K. and A.J.K. P.D.R. is a Rhodes Scholar.

REFERENCES

1. Finnegan, D.J. (1985) Int. Rev. Cytol. **93**, 281-326.
2. Friesen, P.D., Rice, W.C., Miller, D.W. and Miller, L.K. (1986) Mol. Cell Biol. **6**, 1599-16Ø7.
3. Howe, C.C. and Overton, G.C. (1986) Mol. Cell Biol. **6**, 15Ø-157.
4. Boeke, J.D., Garfinkel, D.J., Styles, C.A. and Fink, G.R. (1985) Cell **4Ø**, 491-5ØØ.
5. Shapiro, J.A. (ed.) (1983) Mobile Genetic Elements New York: Academic Press.
6. Shiba, T. and Saigo, K. (1983) Nature **3Ø2**, 119-123.
7. Emori, Y., Shiba, T., Kanaya, S., Inouye, S., Yuki, S. and Saigo, K. (1985) Nature **315**, 773-776.
8. Elder, R.T, Loh, E.Y. and Davis, R.W. (1983) Proc. Natl. Acad. Sci. USA **8Ø**, 2432-2436.
9. Mellor, J., Malim, M., Gull, K., Tuite, M.F., McCready, S., Dibbayawan, T., Kingsman, S.M. and Kingsman, A.J. (1985) Nature **318**, 583-586.
1Ø. Garfinkel, D.J., Boeke, J.D. and Fink, G.R. (1985) Cell **42**, 5Ø7-517.
11. Cameron, J.R., Loh, E.Y. and Davis, R.W. (1979) Cell **16**, 739-751.
12. Winston, F., Durbin, K. and Fink, G.R. (1984) Cell **39**, 675-682.
13. Elder, R.T., St.John, T.P., Stinchcomb, D.T. and Davis, R.W. (198Ø) Cold Spring Harb. Symp. Quant. Biol. **45**, 581-591.
14. Errede, B., Cardillo, T.S., Wever, G. and Sherman, F. (198Ø) Cold Spring Harbor Symp. Quant. Biol. **45**, 593-6Ø7.
15. Carlson, M. and Brutlag, D. (1978) Cell **15**, 733-782.
16. Flavell, A.J., Ruby, S.W., Toole, J.J., Roberts, B.E. and Rubin, G.M. (198Ø) Proc. Natl.Acad. Sci. USA **77**, 71Ø7-7111.
17. Adams, S.E., Mellor, J., Gull, K., Sim, R.B., Tuite, M.F., Kingsman, S.M. and Kingsman, A.J. (1987) Cell **49**, 111-119.

18. Fulton, A.M., Adams, S.E., Mellor, J., Kingsman, S.M., and Kingsman, A.J. (1987) Microbiol. Sci. **4**, 180-185.
19. Liao, X.-B., Clare, J. and Farabaugh, P.J. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 8520-8524.
20. Kingsman, A.J., Gimlich, R.L., Clarke, L., Chinault, A.C. and Carbon, J. (1981) J. Mol. Biol. **145**, 619-632.
21. Guarente, L. (1984) Cell **36**, 799-800.
22. Serfling, E., Jasin, M. and Schaffner, W. (1985) Trends in Genetics **1**, 224-230.
23. Miller, J.H. (1972) New York. Cold Spring Harbor Laboratory.
24. Hawthorne, D.C. and Mortimer, R.K. (1960) Genetics **45**, 1085-1110.
25. Hinnen, A., Hicks, J.B. and Fink, G.R. (1978) Proc. Natl. Acad. Sci. U.S.A. **75**, 1929-1933.
26. Chinault, A.C. and Carbon, J.A. (1979) Gene **5**, 111-126.
27. Holmes, D.S. and Quigley, M. (1981) Anal. Biochem. **114**, 193-197.
28. Tabak, H.F. and Flavell, R.A. (1978) Nucleic Acids Res. **5**, 2321-2332.
29. Sanger, F., Niklen, S. and Coulson, A.R. (1977) Proc. Natl. Acad. Sci. U.S.A. **74**, 5463-5457.
30. Dobson, M.J., Mellor, J., Fulton, A.M., Roberts, N.A., Bowen, B.A., Kingsman, S.M. and Kingsman, A.J. (1984) EMBO J. **3**, 1115-1119.
31. Thomas, P. (1980) Proc. Natl. Acad. Sci. U.S.A. **77**, 5201-5204.
32. Rigby, P.W.J., Dieckman, M., Rhodes, C. and Berg, P.J. (1977) J. Mol. Biol. **113**, 237-251.
33. Tuite, M.F., Dobson, M.J., Roberts, N.A., King, R.M., Burke, D.C., Kingsman, S.M. and Kingsman, A.J. (1982) EMBO J. **1**, 603-608.
34. Ogden, J.E., Stanway, C.A., Kim, S., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986) Mol. Cell Biol. **6**, 4335-4343.
35. Mellor, J., Dobson, M.J., Roberts, N.A., Kingsman, A.J. and Kingsman, S.M. (1985) Gene **33**, 215-226.
36. Petes, T.D., Hereford, L.M. and Skyrabin, K.G. (1978) J.Bacteriol. **134**, 295-305.
37. Wilson, W., Malim, M.H., Mellor, J., Kingsman, A.J. and Kingsman, S.M. (1986) Nucleic Acids Res. **14**, 7001-7016.
38. Dobson, M.J., Tuite, M.F., Roberts, N.A., Kingsman, A.J. and Kingsman, S.M. (1982) Nucleic Acids Res. **10**, 2625-2637.
39. Fulton, A.M., Mellor, J., Dobson, M.J., Chester, J., Warmington, J.R., Indge, K.J., Oliver, S.G., de la Paz, P., Wilson, W., Kingsman, A.J. and Kingsman, S.M. (1985) Nucleic Acids Res. **13**, 4097-4112.
40. Stanway, C., Mellor, J., Ogden, J.E., Kingsman, A.J. and Kingsman, S.M. (1987) Nucleic Acids Res. **15**, 6855-6873.
41. Rathjen, P.D., Kingsman, A.J. and Kingsman, S.M. (1987) Nucleic Acid Res. **15**, 7309-7324.
42. Mellor, J., Fulton, S.M., Dobson, M.J., Wilson, W., Kingsman, S.M. and Kingsman, A.J. (1985) Nature **313**, 243-246.
43. Bowen, B.A., Fulton, A.M., Tuite, M.F., Kingsman, S.M. and Kingsman, A.J. (1984) Nucleic Acids Res. **12**, 1627-1640.
44. Struhl, K. (1985) Proc. Natl. Acad. Sci. USA **82**, 8419-8423.
45. Stanway, C.A., Kingsman, A.J. and Kingsman, S.M. (1987) Bioessays **7**, 62-68.

46. Arndt, K.T., Styles, C. and Fink, G.R. (1987) Science **237**, 874-880.
47. Varmus, H. and Swanstrom, R. (1982) In "RNA Tumour Viruses" Weiss, R., Teich, N., Varmus, H. and Coffin, J. (eds.) pp.369-512 New York: Cold Spring Harbor Laboratory.
48. Varmus, H.E. (1983) In "Mobile Genetic Elements" Shapiro,J.A. (ed.) pp. 411-503 New York: Academic Press.
49. Derse, D., Cardonna, S.J. and Casey, J.W. (1985) Science **227**, 318-320.
50. Arrigo, S., Yun, M. and Beeman, K. (1987) Mol.Cell Biol. **7**, 388-397.
51. Sinclair, J.H., Burke, J.F., Ish-Horowicz, D. and Sang, J.H. (1986) EMBO J. **5**, 2349-2354.
52. Klar, A.J.S., Strathern, J.N., Broach, J.R. and Hicks, J.B. (1981) Nature **289**, 239-244.
53. Nasmyth, K.A., Tatchell. K., Hall, B.D., Astell, C. and Smith, M. (1981) Nature **289**, 244-250.
54. Jensen, R., Sprague, G.F. Jr. and Herskowitz, I. (1983) Proc. Natl. Acad. Sci. U.S.A. **80**, 3035-3039.
55. Mitchell, A.P. and Herskowitz, I. (1986) Nature **319**, 738-742.
56. Miller, A.M., MacKay, V.L. and Nasmyth, K.A. (1985) Nature **314**, 598-603.
57. Siliciano, P.G. and Tatchell, K. (1986) Proc. Natl. Acad. Sci. USA **83**, 2320-2324.
58. Roeder, G.S., Rose, A.B. and Pearlman, R.E. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 5428-5432.
59. Errede, B., Company, M., Ferchak, J.D., Hutchinson III, C.A. and Yarnell, W.S. (1985) Proc. Natl. Acad. Sci. U.S.A. **82**, 5423-5427.
60. Errede, B., Company, M. and Hutchinson III, C.A. (1987) Mol. Cell Biol. **7**, 258-265.
61. Scherer, G., Telford, J., Baldari, C. and Pirotta, V. (1981) Dev. Biol. **86**, 438-447.
62. Dragani, T.A., Manenti, G., Della-Porta, G., Gattoni-Celli, S. and Weinstein, I.B. (1986) Cancer Res. **46**, 581-584.
63. Miksicek, R., Heber, A., Schmid, W., Danesch, U., Posseckert, G., Beato, M. and Schutz, G. (1986) Cell **46**, 283-290.
64. Payvar, F., DeFranco, D., Firestone, G.L., Edgar, B., Wrange, O., Okret, S., Gustafsson, J.A. and Yamamoto, K.R. (1983) Cell **35**, 381-392.