Volume 13 Number 6 1985 Nucleic Acids Research

Maternal inheritance of transcripts from three *Drosophila src*-related genes

Samuel C.Wadsworth*, Kornath Madhavan[1] and Diana Bilodeau-Wentworth

Cell Biology Group, Worcester Foundation for Experimental Biology, Shrewsbury, MA 01545, and [1]Biology Department, College of the Holy Cross, Worcester, MA 01610, USA

Received 7 December 1984; Revised and Accepted 15 February 1985

SUMMARY

The Drosophila genome contains three major sequences related to the v-src gene. Previously published molecular studies have confirmed the structural homology between v-src and two of the Drosophila sequences. We have sequenced a portion of the third v-src-related Drosophila gene and found that it also shares structural homology with vertebrate and Drosophila src-family genes. RNA sequences from each of the src genes are present in pre-blastoderm embryos indicating that they are of maternal origin. As embryogenesis proceeds, the levels of each of the src RNA sequences decline. The pre-blastoderm src gene transcripts contain poly(A) and are present on polyribosomes suggesting that they are functional mRNAs. Since the Drosophila src transcripts were maternally inherited, we also investigated their distribution in adult females. The majority of the src transcripts in adult females were contained in ovaries. Only low levels of the transcripts were detected in males. These results strongly suggest that an abundant supply of src protein is required during early embryogenesis, perhaps at the time of cellularization of the blastoderm nuclei.

INTRODUCTION

The genomes of lower eukaryotes such as Drosophila and yeast contain sequences related to vertebrate oncogenes (1,2,3,4). Thus it is possible to study cellular oncogene structure and function in different developmental and genetic contexts. V-src gene sequences, representing one of the most extensively studied oncogene families, are among those sequences that are conserved in Drosophila. Recombinant DNA clones representing three Drosophila genomic sequences complementary to the v-src gene have been reported (5,6). Nucleotide sequence analysis has confirmed that the v-src-complementarity of two of these genomic sequences is due to significant structural homology, (7) while the third has remained uncharacterized. The studies described in this report were initiated to fill two voids in the knowledge of Drosophila src-related sequences; the lack of nucleotide sequence data from the third src-related genomic segment and the lack of detailed knowledge of the expression of the three src-related genomic

segments at the RNA level.

We report here the nucleotide sequence of a portion of the third v-*src*-related *Drosophila* genomic segment. Analysis of the nucleotide sequence has revealed substantial homology with vertebrate *src*-family genes and with the two other *Drosophila* *src*-family genes. In addition, we show that each of the three *Drosophila* *src*-related sequences is expressed in the form of messenger RNA in preblastoderm embryos. The presence of RNA sequences from two of the *Drosophila* *src*-related genomic clones in embryos has recently been reported (8). The fact that these RNA sequences are present in preblastoderm embryos suggests that they are of maternal origin. In support of this hypothesis, we have also demonstrated that the *src* RNA sequences are abundant in ovaries relative to the levels in the remainder of the female and the male. As embryogenesis proceeds, the levels of these maternal RNA species decline suggesting that an abundant supply of the *Drosophila* *src* gene products is required during early development.

MATERIALS AND METHODS

Fly Culture

An Oregon R strain of *Drosophila melanogaster* was used as the source of RNA for all developmental stages analyzed. For the collection of newly fertilized eggs, well fed 5-6 day old adults were kept in a food bottle capped with a small agar dish smeared with yeast paste and acetic acid for 2 hr. After this precollection, eggs were collected for 1/2 hr (for 3/4 hr old embryos) or for an hr (for 2,4,16 and 20 hr old embryos) on a fresh, yeast-smeared agar dish. At the end of the collection, the eggs were washed from the agar dishes with distilled water, drained and stored at -80°C.

Screening the Bacteriophage Lambda Drosophila Genome Library

The *src* oncogene segment of Rous sarcoma virus that was used is contained within a *Pvu*II restriction site-terminated DNA fragment of approximately 800 base pairs (9). The v-*src* fragment was eluted from preparative agarose gels by electrophoresis and labeled by nick translation (10). A lambda phage library of *Drosophila* DNA (11) was screened by hybridization with the v-*src* probe using the technique of Benton and Davis (12). Specifically, filters were pre-hybridized for three hours at 60°C in 4 x SET [1X is 0.15 M NaCl, 0.03 M Tris-HCl (pH 8.0) and 2mM EDTA], 10 x Denhardt's solution (1 x is 0.02% each of Ficoll, polyvinylpyrrolidone and bovine albumin,13) , 0.1% sodium dodecylsulphate (SDS) and 0.1% sodium PPi and for an additional hour in the same solution containing denatured calf thymus DNA (Sigma) at 50 ug/ml. After overnight hybridization in the latter solution in the presence

of denatured ^{32}P-labeled v-*src* fragment probe, the filters were washed according to the following schedule: one hour at 60°C in the hybridization solution; three 45 minute intervals at 60°C in 3 x SET, 0.1% SDS and 0.1% sodium PPi; 45 minutes at 60°C in 0.3 x SET, 0.1% sodium PPi, 0.1% SDS and a final wash in 4 x SET at room temperature.

Extraction of RNA

RNA was extracted from from various stages of embryonic development by lysis in 8 M urea, 0.1 M Tris-HCl (pH 8.2), 0.3 M NaCl, 0.05 M EDTA and 4% Sarkosyl. One half g of CsCl was added per ml of lysate which was then layered over a 1.5 ml cushion of 5.7 M CsCl. The RNA was pelleted by centrifugation at 36K RPM in an SW50.1 rotor for 16 to 18 hours. The RNA pellet was dissolved in distilled water containing 0.1% SDS and then extracted once with an equal volume of 50% phenol-50% chloroform solution, twice with an equal volume of chloroform and ethanol precipitated. Oligo(dT)-cellulose chromatography was performed as suggested by the supplier (Collaborative Research). Polyribosomes were prepared by lysing embryos in 1% Triton X-100, 0.01 M $MgCl_2$, 0.01 M Tris-HCl (pH 7.6), 0.05 M KCl, 0.004 M dithiothreitol and 0.25 M sucrose (14). The triton extract was centrifuged through sucrose density gradients as described before (15).

RNA and DNA Gel Electrophoresis, Blotting and Hybridizations

RNA was denatured at 55°C for 15 minutes in 50% formamide (Matheson 99%), 6% formaldehyde (Mallinkcrodt) with 0.01 M NaH_2PO_4/Na_2HPO_4 pH 7.0 and was electrophoresed through horizontal 1% agarose slab gels containing formaldehyde and phosphate buffer (16). Transfer of the RNA to nitrocellulose (Schleicher and Schuell) was performed immediately after electrophoresis using 20 X SSC (1 X is 0.15 M NaCl and 0.015 M NaCitrate) as the transfer fluid. After overnight transfer, the nitrocellulose sheets were baked at 80°C for 2 hours, prehybridized, hybridized, and washed as described (17) except that no carrier nucleic acid was used, dextran sulfate was omitted and 0.5% SDS was incorporated into the hybridization solution. Usually 5 x 10^6 cpm of ^{32}P labeled probe prepared by nick translation was hybridized to RNA blots for 12 to 18 hours. The temperature of hybridization was 67°C.

DNA was transferred to nitrocellulose (Schleicher and Schuell) by the method of Southern (18) using 6 x SSC . After overnight transfer, filters were baked at 80°C for 2 hours. The filters were hybridized and washed as described above for hybridization screening of phage libraries except that the temperature of hybridization was 67°C.

FIGURE 1. Restriction map of src4 genomic clone. A 3.3kb EcoRI segment from the src4 phage clone was inserted into the EcoRI site of the plasmid vector pUC8. Single, double or triple endonuclease digests were employed to generate the map shown. HindIII digests of lambda DNA run in parallel lanes were used to estimate the sizes of the src4 fragments. Restriction endonuclease cleavage sites that are not present within the src4 EcoRI segment include those for BglII, HindIII, SalI and XhoI. There are one SacI and two PvuII sites that have not been mapped. The regions of src4 that hybridized with the v-src probe are delimitted by the BamHI and HincII sites.

RESULTS

Isolation of v-src-Complementary Genomic Sequences

Drosophila genomic sequences complementary to the vertebrate src gene were isolated from a lambda phage library by hybridization with cloned v-src sequences from Rous sarcoma virus. Eight phage that hybridized strongly with v-src were plaque purified. The genomic segment represented by the clone src7 was isolated once while the genomic segments represented by the clones src4 and src1 were isolated three and four times respectively.

On the basis of restriction endonuclease mapping experiments and limited nucleotide sequencing, we concluded that the clones src1 and src7 corresponded respectively to Dsrc and Dash, two previously described v-src-related *Drosophila* clones (5,7). We refer to the src1 clone as Dsrc and the src7 clone as Dash throughout this report. The restriction map of src4 (Fig. 1) is distinct from that of Dsrc and Dash. By *in situ* hybridization of a src4 probe to polytene chromosomes, we determined that the src4 genomic segment is located on the left arm of chromosome 2 at 28C (data not shown). This chromosome location is in close agreement with the reported position of a third *Drosophila* genomic clone (designated S13) also isolated by hybridization with v-src (6). Thus we assume that the src4 and S13 clones represent the same genomic segment.

To compare the cloned DNA segments to genomic sequences, Southern blot hybridizations were carried out. To facilitate these comparisons, an 11.3kb BamHI segment from Dsrc, a 3.3 kb EcoRI segment from src4 and a 2.6 kb SalI segment from Dash, each contained within the *Drosophila* insert of the phage clones were subcloned into pUC8. Genomic DNA from Canton S adults and DNA from each of the plasmid subclones was cleaved with the appropriate

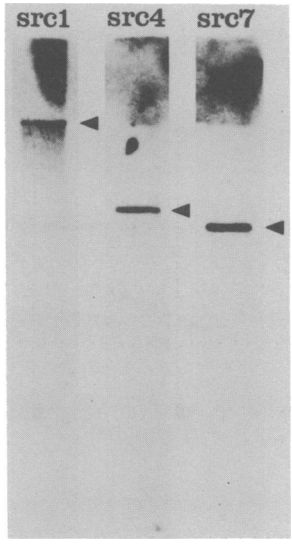
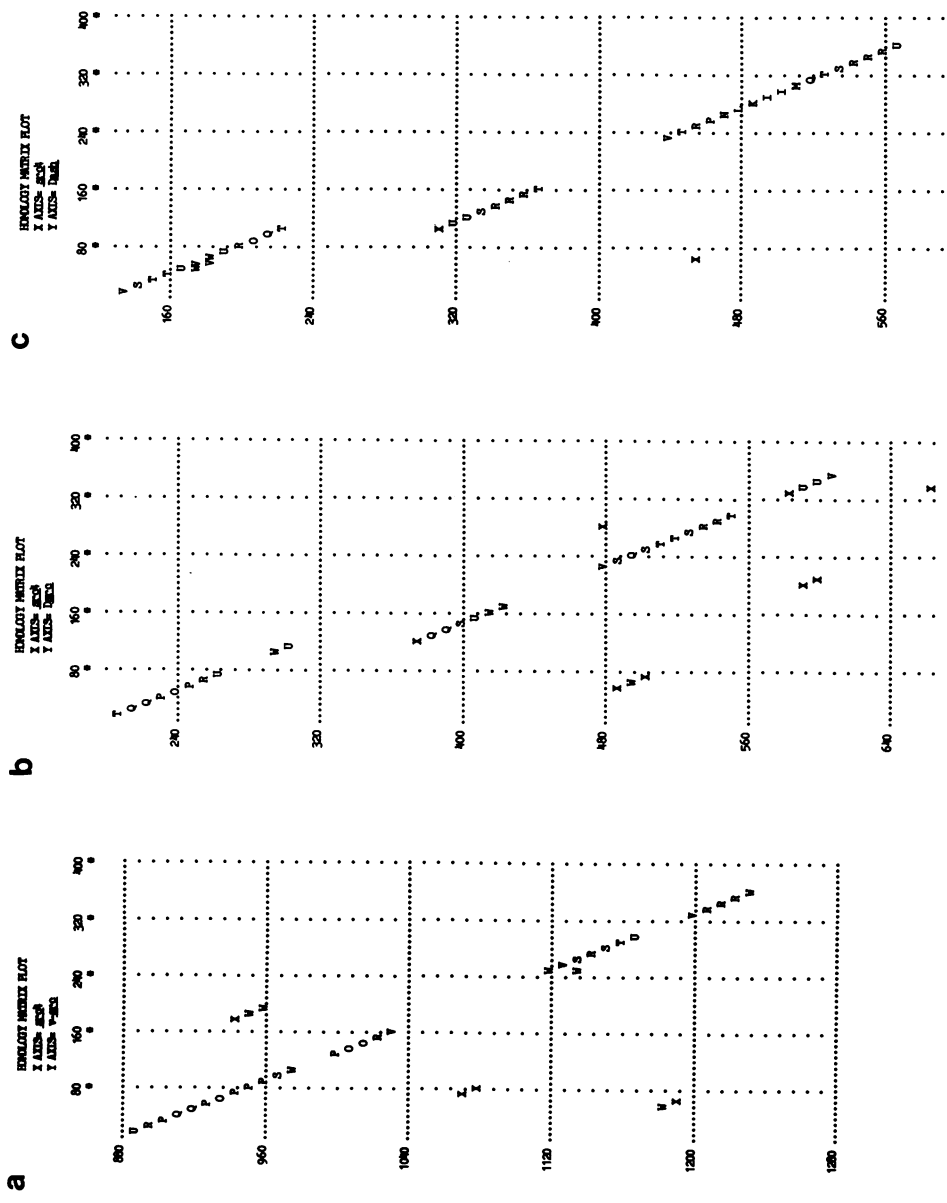


FIGURE 2. Whole genome Southern blot hybridizations. Two ug of DNA isolated from Canton S adults was electrophoresed through a 1% agarose gel after cleavage with either BamHI (lane 1), EcoRI (lane 2) or SalI (lane 3) and transferred to nitrocellulose. The individual lanes were cut apart and hybridized with probes prepared from the plasmid subclones of the Dsrc, src4 or Dash genomic clones. In each case, $4x10^5$ CPM of probe was applied per cm^2 of nitrocellulose. DNA from the Dsrc, src4 and Dash plasmid subclones was individually cleaved with the same restriction enzymes used to cleave the genomic DNA and run in parallel lanes. The arrow heads mark the positions of the ethidium bromide stained Drosophila DNA bands.

restriction endonuclease, electrophoresed and blotted in parallel (Fig. 2). In each case, the cloned segment corresponded in size to the genomic DNA segment detected by hybridization. Furthermore, a given probe did not cross-hybridize with the other src-family genomic sequences nor were repeated sequences were detected by any of the probes. In clone-to-clone hybridization experiments, we have detected weak hybridization between src4 and Dash sequences. No hybridization of Dsrc sequences to either src4 or Dash sequences could be detected under the conditions used (data not shown).

Sequence Homology Between src4 and Other src-Family Genes

Before conducting RNA expression studies, it was essential to determine whether the hybridization between v-src and src4 was due to significant nucleotide sequence homology. It was determined from Southern blot hybridization experiments that the central AccI site in src4 (Fig. 1) splits a region of v-src hybridization. Therefore, this AccI site and the nearby


a
b
c

BamHI site were end-labeled and the sequence determined by the chemical method of Maxam and Gilbert (19).

The src4 sequence was compared to other src-family sequences using the homology matrix computer program of Pustell and Kafatos (20) (Fig. 3a). Although there are short regions with homology lower than the cut-off value used in the analysis, the diagonal describing the homology between v-src and src4 is a straight line indicating that the two sequences are colinear. The overall homology between v-src and the src4 sequence is 55%. Short regions of higher homology also exist; 89% between nucleotides 997 and 1023 and 78% between nucleotides 1201 and 1227.

The src4 sequence was also compared to the previously published sequences from the two other Drosophila src-related genes (7) (Fig. 3b and 3c). In contrast to the colinearity of the v-src and src4 sequences, there are displacements in the diagonals describing the homology between src4 and Dsrc and Dash. The breaks in the homology diagonals seen in Figs. 3b and 3c correspond to the locations of putative introns within the latter two genes (7). From these nucleotide sequence comparisons, we tentatively conclude that there are no introns within the portion of src4 that has been sequenced. The presence of an open reading frame in the src4 sequence with homology to other src-family members supports this conclusion (shown below). The overall homology between src4 and Dsrc is 56%; the highest level is 81% between position 997 and 1023 in the v-src sequence. The overall homology between src4 and Dash is 59%; the highest level is 87% between positions 1141 and 1185 in the v-src sequence.

The nucleotide sequence of src4 and the deduced amino acid sequence are displayed in Fig. 4 along with the sequences from the analogous regions of v-src, v-abl, Dsrc and Dash. Only the amino acid-coding sequences are

FIGURE 3. Comparison of src4 nucleotide sequence to v-src (3a), Dsrc (3b) and Dash (3c). The forward homology matrix program of Pustell and Kafatos (20) was used to compare the src4 sequence to the sequences listed above. Each individual character represents a point at which the program detected a level of sequence homology above the specified minimal value. A string of such characters describing a diagonal line indicates the length of the conserved sequences. The character ,X , indicates a weighted homology of 55%, the minimum value used in these comparisons. The higher alphabetical characters represent increasing levels of homology by 2% increments. The other parameters used to generate the matrices were a range of 20, scale factor of 0.95 and a compression of 10. The sequence of v-src was taken from Schwartz et al. (46). Nucleotide number one here is the first nucleotide in the amino acid-coding region and corresponds to position 7129 in the published v-src sequence. The sequences for Dsrc and Dash were taken from Hoffmann et al. (7). Their numbering system has been conserved.

```
src4                    GCG GTC AAG ATG ATG AAG GAA GGA ACC ATG TCC GAG GAC GAT TTC ATT GAG GAG GCC AAG
                        Ala Val Lys Met Met Lys Glu Gly Thr Met Ser Glu Asp Asp Phe Ile Glu Glu Ala Lys
v-src  (877-936)        GCC ATA AAG ACT CTG AAG CCC GGC ACC ATG TCC CCG GAG GCC TTC CTG CAG GAA GCC CAA
       (293-312)        Ala Ile Lys Thr Leu Lys Pro Gly Thr Met Ser Pro Glu Ala Phe Leu Gln Glu Ala Gln
Dsrc                    GCG GTC AAA ACG CTG CGC GCA GGC ACC ATG TCC ACG GCT GCT TTC CTT CAG GAG GCC GCG
                        Ala Val Lys Thr Leu Arg Ala Gly Thr Met Ser Thr Ala Ala Phe Leu Gln Glu Ala Ala
v-abl                   GCC GTG AAG ACC TTG AAG GAG GAC ACC ATG GAG GTG GAG GAG TTC CTG AAG GAA GCG GCG
                        Ala Val Lys Thr Leu Lys Glu Asp Thr Met Glu Val Glu Glu Phe Leu Lys Glu Ala Ala
Dash                    GCT GTT AAA ACG CTC AAG GAG GAC ACC ATG GCA CTG AAG GAC TTC CTC GAA GAG GCG GCC
                        Ala Val Lys Thr Leu Lys Glu Asp Thr Met Ala Leu Lys Asp Phe Leu Glu Glu Ala Ala

src4                    GTG ATG ACC AAG CTG CAG CAT CCA AAT CTT GTG CAG CTA TAT GGC GTC TGC ACC AAG CAC
                        Val Met Thr Lys Leu Gln His Pro Asn Leu Val Gln Leu Tyr Gly Val Cys Thr Lys His
v-src  (937-993)        GTG ATG AAG AAG CTC CGG CAT GAG AAG CTG GTT CAG CTG TAC GCA GTG GTG TCG    GAA
       (313-331)        Val Met Lys Lys Leu Arg His Glu Lys Leu Val Gln Leu Tyr Ala Val Val Ser    Glu
Dsrc                    ATT ATG AAG AAG TTC CGA CAC AAC CGC CTG GTG GCC CTC TAT GCC GTT TGC TCG CAG GAG
                        Ile Met Lys Lys Phe Arg His Asn Arg Leu Val Ala Leu Tyr Ala Val Cys Ser Gln Glu
v-abl                   GTG ATG AAG GAG ATC AAA CAC CCT AAC CTG GTG CAG CTG CTA GGG GTG TGT ACC CGG GAA
                        Val Met Lys Glu Ile Lys His Pro Asn Leu Val Gln Leu Leu Gly Val Cys Thr Arg Glu
Dash                    ATC ATG AAG GAA ATG AAG CAC CCT AAT CTG GTG CAG CTC ATT GGT GTT TGC ACC AGA GAA
                        Ile Met Lys Glu Met Lys His Pro Asn Leu Val Gln Leu Ile Gly Val Cys Thr Arg Glu

src4                    CGG CCC ATC TAC ATT GTG ACC GAG TAC ATG AAG CAC GGA TCC TTG TTG AAT TAC TTG CGA
                        Arg Pro Ile Tyr Ile Val Thr Glu Tyr Met Lys His Gly Ser Leu Leu Asn Tyr Leu Arg
v-src  (994-1053)       GAG CCC ATC TAC ATC GTC ATT GAG TAC ATG AGC AAG GGG AGC CTC CTG GAT TTC CTG AAG
       (332-351)        Glu Pro Ile Tyr Ile Val Ile Glu Tyr Met Ser Lys Gly Ser Leu Leu Asp Phe Leu Lys
Dsrc                    GAG CCC ATT TAC ATC GTG CAG GAG TAC ATG TCC AAG GGC AGT CTG CTG AAC TTC TTG CGC
                        Glu Pro Ile Tyr Ile Val Gln Glu Tyr Met Ser Lys Gly Ser Leu Leu Asn Phe Leu Arg
v-abl                   CCA CCA TTC TAC ATA ATC ACT GAG TTC ATG ACC TAT GGG AAC CTG CTG GAC TAC CTG AGG
                        Pro Pro Phe Tyr Ile Ile Thr Glu Phe Met Thr Tyr Gly Asn Leu Leu Asp Tyr Leu Arg
Dash                    CCA CCG TTC TAT ATC ATC ACC GAG TTT ATG TCG CAC GGC AAT CTG GTG GAC TTC CTG CGC
                        Pro Pro Phe Tyr Ile Ile Thr Glu Phe Met Ser His Gly Asn Leu Val Asp Phe Leu Arg

src4                       CGG CAT GAG AAG ACC CTG ATT GGT AAT ATG GGT CTA CTC CTT GAC ATG TGC ATA CAG
                           Arg His Glu Lys Thr Leu Ile Gly Asn Met Gly Leu Leu Leu Asp Met Cys Ile Gln
v-src  (1054-1107)      GGA GAG ATG GGC AAG TAC CTG          CGG CTG CCA CAG CTC GTC GAT ATG GCT GCT CAG
                        Gly Glu Met Gly Lys Tyr Leu          Arg Leu Pro Gln Leu Val Asp Met Ala Ala Gln
Dsrc                    GAG GGC GAT GAC CGT TAC TTG          CAC TTC GAA GAT CTC ATC TAC ATG CAC ACC CAG
                        Glu Gly Asp Asp Arg Tyr Leu          His Phe Glu Asp Leu Ile Tyr Met His Thr Gln
v-abl                   GAG TGT AAC CGG CAG GAG GTG          AGC GCC GTG GTA CTG CTC TAC ATG GCC ACA CAG
                        Glu Cys Asn Arg Gln Glu Val          Ser Ala Val Val Leu Leu Tyr Met Ala Thr Gln
Dash                    TCC GCC GGA CGC GAA ACG CTC          GAT GCA GTA GCG TTG CTG TAC ATG GCC ACT CAG
                        Ser Ala Gly Arg Glu Thr Leu          Asp Ala Val Ala Leu Leu Tyr Met Ala Thr Gln

src4                    GTT AGC AAG GGA ATG ACC TAC CTA GAG CGC CAT AAC TAC ATT CAC CGG GAT CTG GCT GCC
                        Val Ser Lys Gly Met Thr Tyr Leu Glu Arg His Asn Tyr Ile His Arg Asp Leu Ala Ala
v-src  (1108-1167)      ATT GCA TCC GGC ATG GCC TAT GTG GAG AGA ATG AAC TAC GTG CAC CGA GAC CTG CGG GCG
       (370-389)        Ile Ala Ser Gly Met Ala Tyr Val Glu Arg Met Asn Tyr Val His Arg Asp Leu Arg Ala
Dsrc                    GTG ACC ACC GGT ATG AAG TAT CTA GAG TCC AAG CAA GTC ATC CAC CGC GAT CTG ACG ACC
                        Val Thr Thr Gly Met Lys Tyr Leu Glu Ser Lys Gln Val Ile His Arg Asp Leu Thr Thr
v-abl                   ATC TCA TCA GCC ATG GAG TAC TTG GAG AAG AAG AAC TTC ATC CAC AGA GAC CTT GCT GCC
                        Ile Ser Ser Ala Met Glu Tyr Leu Glu Lys Lys Asn Phe Ile His Arg Asp Leu Ala Ala
Dash                    ATA GCG TCG GGA ATG AGC TAC CTG GAG TCG CGC AAC TAC ATT CAT CGC GAT CTC GCT GCC
                        Ile Ala Ser Gly Met Ser Tyr Leu Glu Ser Arg Asn Tyr Ile His Arg Asp Leu Ala Ala
```

src4		CGC AAC TGT CTC GTG GGT TTC GAG AAT GTC GTT AAA GTG GCC GAC TTT GGA TTG GCA CGA
		Arg Asn Cys Leu Val Gly Phe Glu Asn Val Val Lys Val Ala Asp Phe Gly Leu Ala Arg
v-src	(1168-1227)	GCC AAC ATC CTG GTG GGG GAG AAC CTG GTG TGC AAG GTG GCT GAC TTC GGG CTG GCA CGC
	(390-409)	Ala Asn Ile Leu Val Gly Glu Asn Leu Val Cys Lys Val Ala Asp Phe Gly Leu Ala Arg
Dsrc		CGT AAT GTG CTG ATC GGA GAG AAT AAT GTG GCG AAG ATT TGT GAT TTT GGA CTG GCG CGT
		Arg Asn Val Leu Ile Gly Glu Asn Asn Val Ala Lys Ile Cys Asp Phe Gly Leu Ala Arg
v-abl		CGG AAC TGC CTG GTA GGG GAA AAC CAC TTG GTG AAG GTG GCT GAT TTT GGC CTG AGC AGG
		Arg Asn Cys Leu Val Gly Glu Asn His Leu Val Lys Val Ala Asp Phe Gly Leu Ser Arg
Dash		CGC AAT TGC CTG GTG GGT GAT AAC AAG CTG GTC AAG GTG GCG GAT TTC GGC CTG GCA CGT
		Arg Asn Cys Leu Val Gly Asp Asn Lys Leu Val Lys Val Ala Asp Phe Gly Leu Ala Arg

src4		TAC GTT CTC GAC GAT CAA TAT ACC AGC TCG GCG GAA CCA AGT
		Tyr Val Leu Asp Asp Gln Tyr Thr Ser Ser Ala Glu Pro Ser
v-src	(1228-1269)	CTC ATC GAG GAC AAC GAG TAC ACA GCA CGG CAA GGT GCC AAG
	(410-423)	Leu Ile Glu Asp Asn Glu Tyr Thr Ala Arg Gln Gly Ala Lys
Dsrc		GTC ATC GCG GAT GAC GAG TAC CGC CCC AAG CAG GGA TCC CGG
		Val Ile Ala Asp Asp Glu Tyr Arg Pro Lys Gln Gly Ser Arg
v-abl		TTG ATG ACA GGG GAC ACC TAC ACG GCC CAT GCT GGA GCC AAA
		Leu Met Thr Gly Asp Thr Tyr Thr Ala His Ala Gly Ala Lys
Dash		TTG ATG CGG GAC GAC ACG TAT ACA GCA CAT GCC GGA GCC AAG
		Leu Met Arg Asp Asp Thr Tyr Thr Ala His Ala Gly Ala Lys

FIGURE 4. Nucleotide sequence of src4 and the deduced amino acid sequence; comparison to other src-family members. The nucleotide sequences and deduced amino acid sequences shown were aligned with the aid of manual and automatic sequence alignment computer programs (20). For the analysis of the data shown in this figure, the putative intron sequences have been removed from the Dsrc and Dash sequences of Hoffmann et al. (7). The src4 amino acid positions that match with any of the other four sequences are underlined. The actual homologies are noted in the text. The numbering system from the v-src sequence is used throughout. Amino acid number one represents the first amino acid in the v-src protein as deduced from the nucleotide sequence (46). Spaces in the sequence indicate that one or more of the other sequences have an amino acid codon in that position.

considered in these comparisons, the putative intron sequences in Dsrc and Dash have been removed. The comparisons reveal that the nucleotide sequence homology between src4 and other src-family genes shown in Fig. 3 is paralleled by substantial amino acid sequence homology. The src4 amino acid sequence matches with at least one of the other sequences in greater than 72% of the positions. When compared individually to the other src-family members, the src4 amino acid sequence homology is 50% with Dsrc, 53% with v-src, 54% with v-abl, and 58% with Dash. Within the sequenced region, we can discern no consistent pattern of amino acid sequence conservation. For example, between positions 319 and 330 in the v-src protein, eleven out of twelve positions are conserved among src4, v-abl, and Dash. Just three amino acids downstream, eight out of nine positions are conserved among src4, v-src, and Dsrc. Further downstream between positions 381 and 395, fourteen

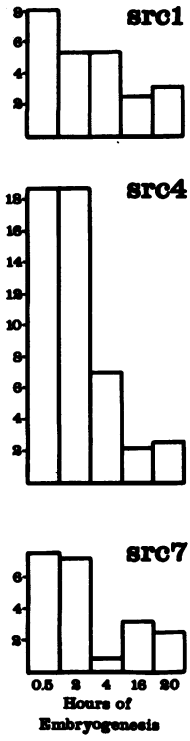


FIGURE 5. Expression of Drosophila src-family genes during embryogenesis. One ug aliquots of total RNA from the stages indicated were boiled in water and quenched on ice. Ten volumes of 20xSSC were added before filtering onto nitrocellulose. The nitrocellulose strips were hybridized with the indicated Drosophila src-family sequence probes as described in Materials and Methods. Subsequently the same filters were hybridized with a Drosophila actin 5C probe and the level of Dsrc probe hybridization was normalized to the level of actin hybridization. It has been shown that the actin mRNA levels during embryonic development detected by the actin 5C probe are essentially constant (47).

v-abl and fifteen Dash amino acids out of fifteen match with src4. More mismatches exist among src4, v-src, and Dsrc. Six amino acids further along, a short region (401 to 409) of 100% homology among src4, v-src and v-abl is seen, while there are mismatches among the three Drosophila genes. Thus, until more extensive sequence data is available, it would be premature to designate the src4 gene as a homolog of a specific src-family gene.

It should be noted that the lysine at position 295 in the v-src sequence is conserved in the src4 sequence. The conservaton of this particular amino acid may be relevant to the enzymatic activity of the protein encoded by the src4 gene since lysine 295 corresponds to lysine 71 in bovine cAMP-dependent kinase which binds ATP (21,22,23).

Expression of Drosophila src RNA During Embryogenesis

To learn if the Drosophila src-family genes were expressed at the RNA level, we initially tested RNA from a wide variety of developmental stages by hybridization with Dsrc, src4, and Dash probes. In general, the levels of transcripts complementary to each of the probes were highest during embryogenesis and metamorphosis and lowest during the larval and adult

Table 1. MATERNAL DROSOPHILA src RNA SPECIES ARE PRESENT ON POLYRIBOSOMES

	% Total 1.5 hour embryo RNA on polyribosomes	% Total RNA on EDTA disrupted polyribosomes
D*src*	61	5
*src*4	38	1
D*ash*	54	2

RNA was purified from all portions of the sucrose gradients by phenol extraction and ethanol precipitation. Aliquots of the RNA were dotted onto nitrocellulose and hybridized as described in Experimental Procedures. The values shown represent the fraction of the total hybridizable RNA present in the polyribosome region of the gradient.

stages. In this report, we will focus on RNA expression during oogenesis and embryogenesis. Shown in Fig. 5 are the results of dot hybridizations using RNA extracted from various stages of embryonic development. The RNA sequences complementary to each of the *src*-related genes are present at their highest levels at 0.5 hours of development and decline in abundance as embryogenesis proceeds. The presence of these RNA sequences prior to the onset of zygotic transcription, which occurs at approximately 2.5 hours of development (24,25,26), is strong evidence that the *src* transcripts are of maternal origin. The maternal nature of the D*src* and D*ash* transcripts has been reported (8).

To define the nature of these maternal RNA sequences, two types of experiments were carried out. First, the presence of the maternal *src* transcripts on polyribosomes was established. Embryos at ninety minutes of development were homogenized and fractionated on sucrose gradients either in the presence or absence of EDTA. By performing dot hybridizations with D*src*, *src*4, or D*ash* probes, the fraction of the total hybridizable RNA detected by each probe that was present in the polyribosome region of the gradient was determined (Table 1). Between 40 and 60% of the *src*-family RNA sequences were present on polyribosomes; very similar to the fraction of total poly(A)-containing RNA present on *Drosophila* embryo polyribosomes (27). Because virtually all of these RNA sequences were displaced from polyribosomes when centrifuged in the presence of EDTA, we concluded that a substantial fraction of the maternal *Drosophila src* RNA sequences are in the form of messenger RNA.

To further analyze the form of the maternal *src* RNA species, we performed RNA gel blot hybridizations. Poly(A)-containing RNA from whole eggs 0.5 hour after fertilization and from polyribosomes isolated 1.5 hour after fertilization was fractionated on formaldehyde-agarose gels, blotted onto

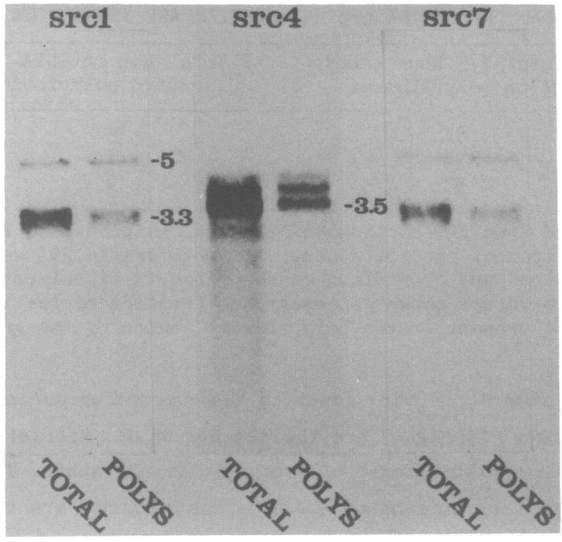


FIGURE 6. RNA gel transfer hybridization analysis of poly(A)-containing RNA from total embryos and from polyribosomes. Poly(A)-containing RNA from 0.5 hr embryos or from polyribosomes from 1.5 hr embryos was electrophoresed in the presence of formaldehyde as described in Experimental Procedures. Each lane represents poly(A)-containing RNA from 20ug of total RNA. The pairs of lanes were hybridized with the indicated probes. The central panel of the figure was printed from a shorter autoradiographic exposure so that the src4 panel would not be overexposed. The numbers adjacent to the lanes indicate the approximate sizes of the transcripts as determined in other experiments by co-electrophoresis with mouse rRNA.

nitrocellulose and hybridized with Dsrc, src4, and Dash probes (Fig. 6). In each case, the RNA species present on polyribosomes are identical to those seen in total embryo RNA. The Dsrc-complementary transcripts are approximately 5kb and 3.3kb. The slower migrating species actually consists of two RNA bands that are not resolved on the gel shown in Fig. 6. The 3.3kb species is always a diffuse band. The major src4-complementary transcript is approximately 3.5kb. In most other experiments, a single RNA band of 3.5kb was seen. The Dash-complementary transcripts were indistinguishable in size from the Dsrc-complementary transcripts. It is important to reiterate here, that, as shown in Fig 2, cloned Dsrc and Dash sequences do not hybridize with each other or any other genomic sequences under the conditions used. Thus, the similarity in the sizes of the RNA bands in the Dsrc and Dash lanes is not due to cross-hybridization.

Since both the fast and slow migrating forms of the Dsrc and Dash

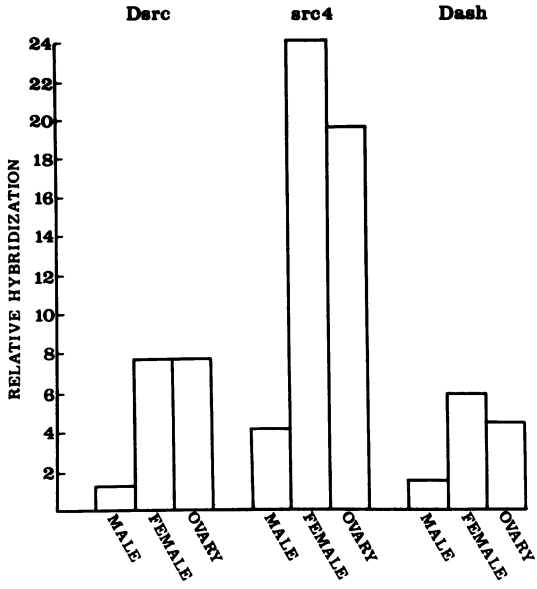


FIGURE 7. Dot hybridization analysis of Drosophila src-family RNA sequences in males and females. Ovaries were hand-dissected from 5 day-old females and RNA was isolated as described in Materials and Methods. Denaturation, application to nitrocellulose and hybridization were as described in the legend to Figure 5. The hybridization levels with src probes were normalized to the hybridization levels detected by the actin 5C probe. The hybridization data is plotted on a per animal basis.

transcripts are found on polyribosomes, it is unlikely that the slow forms are obligate precursors to the faster migrating forms. These data give rise to the speculation that the fast and slow forms represent messenger RNA species resulting from alternate synthetic pathways. Since we have quantitated the expression of these RNA sequences by dot hybridizations, (Fig 5) we do not know if the ratios of these RNA species remain constant throughout embryogenesis. Additional experiments will be required to establish the structure of these mRNA species.

Drosophila src Transcripts are Most Abundant in Ovaries

Maternally inherited Drosophila RNA species are synthesized in the 15 interconnected nurse cells within the ovary and are deposited in the developing egg. We hybridized the Dsrc, src4, and Dash probes with RNA isolated from whole males, whole females and from ovaries (Fig. 7). The levels of RNA complementary to each Drosophila src gene were lowest in male and highest in whole female and ovary. Adult Drosophila females contain

three to four times the amount of RNA as do males (28). In agreement with this fact, we find that approximately 75% of the total female RNA is present in ovaries (unpublished observation). Thus the near equivalence of the RNA levels in whole female and in ovary strongly suggests that the major site of expression of Drosophila src-family genes in adults is in the ovary. We have detected the src4 3.5kb transcript at low adundance in male RNA (data not shown), suggesting that expression of at least this gene is not restricted to females. We have not investigated the presence of the Dsrc and Dash transcripts in males .

DISCUSSION

The earliest stages of embryonic development are directed by maternal components deposited in the egg during maturation. Genetic and molecular approaches have been applied to identify maternal components active during and perhaps controlling early embryonic development. At the genetic level, numerous mutations identifying Drosophila genes active during oogenesis or preblastoderm development have been characterized (reviewed in reference 29). Such mutations are broadly referred to as maternal effect mutations, since expression of the mutant phenotype is under control of the maternal genome. Two classes of maternal effect mutations are those in which no eggs are produced or eggs with grossly altered morphology are produced. The third major class of maternal effect mutations is that in which eggs of normal morphology are laid, but which fail to develop normally. Because the normal period of activity of some of the latter genes is before the onset of zygotic transcription, it is apparent that these mutations affect the quality or quantity of maternal components essential for normal development. In Caenorhabditis as well, a number of maternal effect mutations have been characterized (30,31). Temperature shift experiments have shown that a subset of these mutations define genes whose activity is essential to the proper execution of the earliest stages of embryonic development. Thus the genetic experiments have clearly established the existence of maternally inherited components that function during early embryonic development.

The pattern of expression of Drosophila src mRNA sequences that we have documented here would seem to model closely the maternally inherited determinants defined by genetic means. Src RNA species are present in embryos before the onset of zygotic transcription, strongly suggesting their maternal origin. We have shown that the src RNA sequences are particularly abundant in ovary, the site of synthesis of maternal RNA species, as compared to whole male and female non-ovarian tissue. It is important to note here

that maternal gene expression does not exclude other modes of expression of the same sequences at other stages of development. Thus our results do not suggest a lack of expression of Drosophila src-family sequences in male and female non-ovarian tissue; rather the high levels of src transcripts in ovary highlight their participation in early embryonic development.

Hybridization studies in a variety of systems have identified complex classes of egg RNA sequences (Reviewed in reference 32,33,34,35,36). The levels of the majority of these egg RNA sequences are either maintained or increase throughout embryogenesis, presumably through the replacement of maternal RNA sequences by newly transcribed RNA sequences. These studies have also identified a minor class of egg RNA sequences that are lost as development proceeds. High resolution two dimensional protein gel analyses and the cloning of specific sequences have afforded a more detailed view of the pattern of regulation of individual early embryonic mRNA sequences (37,38,39,40,41,42,43,44). In general, these analyses have supported the conclusions derived from hybridization experiments; examples of mRNA sequences maintained or increasing throughout embryogenesis are in the majority while few instances of sequences declining in abundance were found. Of the cloned maternally expressed RNA sequences, only a handful have been correlated with known genes. Thus the Drosophila src-family sequences represent especially interesting examples of maternally expressed genes since they represent the minority case of RNA sequences that decrease in abundance as embryogenesis proceeds (Fig. 5). This is especially prominant for src4 and Dash sequences where the decline in RNA levels is coincident with the time of onset of zygotic transcription. These results, in addition to those discussed above, tend to emphasive the participation of the Drosophila src-family gene products in early developmental events.

A commonly held view is that the function of maternal mRNA species is to provide a means of production of the encoded proteins at a time when the embryonic genes cannot be transcribed. In the present case, this implies that the functions of the Drosophila src-family proteins are required during pre-blastoderm development. Although the available data do not directly shed light on the functional roles of the src proteins, the hybridization pattern of src-family RNA species that we have described suggests an hypothesis to direct future experiments. It has been demonstrated that vertebrate src-family proteins are intracellular, membrane-associated proteins. In preblastoderm Drosophila embryos there are no cells (with the exception of the precociously segregating pole cells) and the only fully formed plasma

membrane encloses the entire egg lying just beneath the vitelline membrane. Up until the blastoderm stage, a series of syncytial nuclear divisions take place resulting in the distribution of the the vast majority of the nuclei around the periphery of the egg. At the cellular blastoderm stage, these nuclei (which number between 3500 and 5000) are synchronously enclosed by membranes in a two-stage process described by Fullilove and Jacobson (45). The source of approximately half of the required membrane appears to be the abundant villi in the plasma membrane surrounding the egg. However, the remainder of the required membrane must be formed from pre-existing vesicles or synthesized de novo. The maternally inherited *src* mRNAs could be directing the synthesis of proteins for incorporation into the blastoderm cells as they form. Members of the vertebrate *src*-family have been shown to be centrally involved in the control of cell growth. We reason that because of the high degree of amino acid conservation among the *Drosophila* and vertebrate *src*-family sequences, the *Drosophila* proteins are likely to have analogous functions. Thus the incorporation of preformed *src* proteins into blastoderm cells might well be crucial in maintaining the rapid pace of cell growth and development characteristic of embryogenesis.

Portions of the three major *src*-related *Drosophila* genes have now been sequenced and mRNA species complementary to each have been demonstrated. Thus it is a reasonable hypothesis that the chromosomal loci identified by *in situ* hybridization contain functional genes in which mutations might be isolated. The fact that the *src*-family RNA sequences are maternally inherited clearly suggests that the search for such mutations should include strategies to recover mutations with maternal-effect phenotypes.

ACKNOWLEDGEMENTS

We thank W. Bender for providing the *Drosophila* lambda phage library; J. M. Bishop and H. E. Varmus for the use of the recombinant plasmid containing v-*src* sequences; H. Robinson, D. Steffen and M.M. Madhavan for useful discussions;and D. Steffen for help with the nucleotide sequence analysis.

This work was supported by grants to S.C.W. from the National Cancer Institute (CA-12708), the March of Dimes Birth Defects Foundation (Basil O'Connor Starter Research Grant No. 5-345) and the American Cancer Society, Massachusetts Division Inc. and by a grant to K.M. from Monsanto Company. Central research facilities at the WFEB were supported in part by The Mimi Aaron Greenberg Memorial Cancer Research Fund.

*To whom correspondence should be addressed

REFERENCES

1. Shilo, B.-Z. and R.A. Weinberg (1981). Proc. Natl. Acad. Sci.USA 78, 6789-6789.
2. DeFeo-Jones, D., Scolnick,E., Koller,R., and Dhar, R. (1983). Nature 306, 707-709.
3. Gallwitz, D., Dorath, C., and Sander, C. (1983). Nature 306, 704-707.
4. Powers, S., Kataoka, T., Fasano, O., Goldfarb, M., Strathern, J., Broach, J., and Wigler, M. (1984). Cell 36, 607-612.
5. Hoffman-Falk, H., P. Einat, B.-Z. Shilo and F.M. Hoffman (1983). Cell 32, 589-598.
6. Simon, M. A., Kornberg, T. B., and Bishop, J. M. (1983). Nature 302, 837-839.
7. Hoffmann, F. M., Fresco, L. D., Hoffman-Falk, H., and Shilo, B.-Z. (1983). Cell 35, 393-401.
8. Lev, Z., Leibovitz, N., Segev, O., and Shilo, B.-Z. (1984). Mol. Cell. Biol. 4, 982-984.
9. Czernilofsky, A.P., Levinson, A.D., Varmus, H.E., Bishop, J.M., Tischer, E. and Goodman, H.M. (1980). Nature 287, 198-203.
10. Maniatis, T., A. Jeffrey, and D.G. Kleid (1975). Proc. Natl. Acad. Sci. USA. 72, 1184-1188.
11. Maniatis, T., R.C. Hardison, E. Lacy, J. Lauer, C. O'Connell and D. Quon (1978). Cell 15, 687-701.
12. Benton, W.D. and R.W. Davis (1977). Science 196, 180-182.
13. Denhardt, D.T. (1966). Biochem. Biophys. Res. Comm. 23, 641-646.
14. Boshes, R.A. (1970). J. Cell Biol. 46, 477-490.
15. Craig, E.A., B.J. McCarthy and S.C. Wadsworth (1979). Cell 16, 575-588.
16 Lehrach, H., D. Diamond, J.M. Wozney, and H. Boedtker (1977). Biochemistry 16, 4743-4751.
17. Thomas, P.S. (1980). Proc. Natl. Acad. Sci. USA 77, 5201-5205.
18. Southern, E.M. (1975). J. Mol. Biol. 98,503-517.
19. Maxam, A. M., and Gilbert, W. (1980). Meth. Enzymol. 65, 499-560.
20. Pustell, J., and Kafatos, F.C. (1982). Nuc. Acids Res. 10, 4765-4782.
21. Barker, W. C., and Dayhoff, M. O. (1982). Proc. Nat. Acad. Sci. USA 79, 2836-2839.
22. Shozo, S., Parmalee, D.C., Wade, R.D., Kumar, S., Ericsson, R.H., Walsh, K.A., Neurath, H., Long, G.L., Demaille, J.G., Fischer, E.H., and Titani, K. (1981). Proc. Nat. Acad. Sci. USA 78, 848-851.
23. Zoller, M.J. and Taylor, S.S. (1979). J. Biol. Chem. 254, 8363-8368.
24. Lamb, M.M. and C.D. Laird (1979). Devel. Biol. 52, 31-42.
25. McKnight, S.L. and O.L. Miller, Jr. (1976). Cell 8, 305-319.
26. Zalokar, M. (1976). Devel. Biol. 49, 425-437.
27. Lovett, J.A., and Goldstein, E.S. (1977). Devel. Biol. 61, 70-78.
28. Church, R.B. and F.W. Robertson, (1966). J. Exp. Zool. 162, 337-353.
29. Mahowald, A.P. (1983). Genetic analysis of oogenesis and determination. In Time, space, and pattern in embryonic development. (Alan R. Liss, Inc.,New York). pp. 349-363.
30. Wood, W.B., Hecht, R., Carr, S., Vanderslice, R., Wolf, N., and Hirsh, D. (1980). Devel. Biol. 74, 446-469.
31. Isnenghi, E., Cassada, R., Smith, K., Denich, K., Radnia, K., and von Ehrenstein, G. (1983). Devel. Biol. 98, 465-480.
32. Davidson, E.H. (1976). Gene activity in early development. (New York; Academic Press).

33. Hough-Evans, B.R., Wold, B.J., Ernst, S.G., Britten, R.J., and Davidson, E.H. (1977). Devel. Biol. 60, 258-277.
34. Perlman, S., and Rosbash, M. (1978). Devel. Biol. 63, 197-212.
35. Arthur, C.G., Weide, C.M., Vincent, W.S., and Goldstein, E.S. (1979). Exper. Cell Res. 121, 87-94.
36. Imre, K., and Smith, M.J. (1982). Devel. Biol. 89, 56-63.
37. Dworkin, M.B. and Dawid, I.B. (1980). Devel. Biol. 76, 449-464.
38. Lasky, L.A., Lev, Z., Xin, J.-H., Britten, R.J. and Davidson, E.H. (1980). Proc. Nat. Acad. Sci. USA 77, 5317-5321.
39. Colot, H.V. and Rosbash, M. (1982). Devel. Biol. 94, 79-86.
40. Bedard, P.-A. and Brandhorst, B.P. (1983). Devel. Biol. 96, 74-83.
41. Trumbly, R.J. and Jarry, B. (1983). Embo J. 2,1281-1290
42. Zimmerman, J.L., W. Petri and M. Meselson (1983). Cell 32, 1161-1170.
43. Vincent, A., O'Connell, P., Gray, M. and Rosbash, M. (1984). Embo J. 3, 1003-1013.
44. Madhavan, K., Bilodeau-Wentworth, D. and Wadsworth, S.C. (1984). In Press
45. Fullilove, S.L., and Jacobson, A.G. (1971). Devel. Biol. 26, 560-577.
46. Schwartz, D.E., Tizard, R. and Gilbert, W. (1983). Cell 32, 853-869.
47. Fyrberg, E.A., Mahaffey, J.W., Bond, B.J. and Davidson, N. (1983). Cell 33, 115-123.