MOLECULAR AND CELLULAR BIOLOGY, June 1987, p. 2119–2127
0270-7306/87/060000-00$02.00/0

Vol. 7, No. 6

Primary Sequence and Developmental Expression of a Novel *Drosophila melanogaster src* Gene

RICHARD J. GREGORY, K. L. KAMMERMEYER, WALTER S. VINCENT III, AND SAMUEL G. WADSWORTH*
Worcester Foundation for Experimental Biology, Shrewsbury, Massachusetts 01545

Received 20 January 1987/Accepted 9 March 1987

We have sequenced a cDNA clone for the *Drosophila melanogaster* gene *Dsrc28C*, a homolog of the vertebrate gene c-*src*. The cDNA contains a single open reading frame encoding a protein of 66 kilodaltons which contains features highly conserved within the *src* family of tyrosine protein kinases. Novel structural features of the *Dsrc28C* protein include a basic pI and a polyglycine domain near the amino terminus. Cell-free translation of in vitro-transcribed RNA yielded a protein of the predicted size which could be immunoprecipitated by anti-v-*src* antisera. RNA blot hybridization revealed that the gene is expressed predominantly during embryogenesis, in imaginal disks of third-instar larvae, and in adult females. In situ hybridization showed that expression in adult females is largely confined to nurse cells and developing oocytes.

Members of the *src* tyrosine protein kinase family make up a major fraction of known cellular oncogenes (1). This gene family, characterized by extensive homology to the kinase domain of the transforming gene of Rous sarcoma virus (v-*src*), appears to have been highly conserved during evolution and contains representatives in the genomes of organisms as phylogenetically diverse as chickens and fruit flies. Although the function(s) of the intracellular members of the tyrosine protein kinase family are not known, the ability of retrovirally transduced members of the gene family to induce neoplastic transformation, the homology with the kinase domain of several growth factor receptors (9), and the high degree of evolutionary conservation within the gene family all suggest that tyrosine kinases play important roles in the control of cell growth.

The genome of the fruit fly, *Drosophila melanogaster*, contains three distinct genes with homology to v-*src* (10, 24). Hybridization and DNA sequence analyses have revealed that two of these genes, designated *Dsrc* and *Dash*, appear to be homologous to the vertebrate oncogenes *src* and *abl*, respectively (10, 23). Transcripts from these genes are maternally inherited and are present on polyribosomes in preblastoderm embryos, indicating that they encode gene products active during early embryonic development (13, 28). In addition, transcripts from *Dsrc* and *Dash* have been shown to be most abundant in embryos and pupae and to be present in lower amounts in larval and adult stages, except in adult ovarian tissue (23, 26, 28). This pattern of expression correlates well with the periods of maximal cell division during *Drosophila* development and suggests, by analogy with vertebrate *src* family genes, that these genes may play a role in the regulation of cell proliferation. However, examination of the tissue-specific expression of *Dsrc* showed that, after the first 8 h of development, transcripts of this gene accumulate predominantly in neural tissue and in differentiating smooth muscle. This result suggests that the *Dsrc* gene product may not participate in the regulation of cell division but may instead be involved in the differentiation of specific tissues (23).

The structure and expression of the third *D. melanogaster src* homolog, which we have designated *Dsrc28C* because of its location at position 28C in the polytene chromosome map, have only been partially characterized (28). *Dsrc28C* mRNA is also maternally inherited and is present on polysomes in developing embryos (28). In this study we present the complete primary structure and developmental expression of the *Dsrc28C* gene. Our results indicate that the protein encoded by the *Dsrc28C* gene bears extensive homology to the *src* and *abl* classes of vertebrate oncogenes. In addition, our sequence analyses have revealed that the *Dsrc28C* protein has structural features unique among the extensive *src* gene family. Hybridization results indicate that maximal expression of *Dsrc28C* occurs during embryogenesis and oogenesis while lower levels of expression occur in imaginal disks and during the first 24 h of metamorphosis.

MATERIALS AND METHODS

Isolation and sequencing of cDNA clones. cDNA clones were isolated from libraries and analyzed according to Maniatis et al. (15). Two separate bacteriophage libraries were screened for *Dsrc28C* cDNA. The first (obtained from Larry Kauvar and Tom Kornberg) contained 5×10^5 independent clones prepared from poly(A)-containing RNA isolated from 3- to 12-h-old embryos. The second library (obtained from Barry Yedvobnick) contained 3×10^5 independent clones and was prepared with poly(A)-containing RNA from adult females. Both libraries were constructed by using the λgt10 vector system. The genomic *Dsrc28C* probe used for screening the libraries was the 3.3-kilobase (kb) *Eco*RI fragment previously described (28; see Fig. 1). This clone was previously referred to as *Dsrc4*. Sequence analysis was performed according to Sanger et al. (21) on restriction fragments subcloned into M13mp series vectors (17). The sequence presented in Fig. 2 was determined from both strands, except for positions 2458 to 2755 in the noncoding 3′ extension.

RNA blot analysis. Extraction of RNA from various developmental stages was performed according to Wadsworth et al. (28). After a 1-h precollection, embryos were collected for 30 min or 1 h and staged from the midpoint of the collection. White prepupae were collected for staging of pupal development. Imaginal disks were supplied by James Fristrom. Total RNA (5 μg) was analyzed on 1% agarose–formaldehyde gels and transferred to nitrocellulose as described by Wadsworth et al. (28). Synthesis of the antisense RNA and hybridization to the RNA blot were

* Corresponding author.

performed according to Melton et al. (16). The probe was synthesized with phage SP6 RNA polymerase from a 0.6-kb *Xba*I-*Pst*I genomic restriction fragment (Fig. 1) cloned into plasmid pSP64 (Promega Biotec).

In vitro analysis of the *Dsrc28C* gene product. For transcription of the *Dsrc28C* cDNA, plasmid pWV411 was constructed by inserting approximately 2.2 kb of λ-6311, from the 5′ *Eco*RI site to the 3′ *Hin*cII site, into Bluescrybe M13+ (Stratagene). Truncation of this plasmid with *Hin*dIII allowed transcription of a 2.2-kb RNA from the T7 promoter. Transcription was carried out according to the protocol of the supplier, except that cap structure analog, 7meGpppG (Bethesda Research Laboratories), was added to a final concentration of 1 mM to the reaction mixture and the GTP concentration was reduced to 100 μM. From 5 μg of plasmid, about 25 μg of RNA was obtained. The presence or absence of cap structure analog had no effect on either RNA yield or fidelity of protein translation at high levels of RNA in the translation mix. The RNA was translated in 25-μl reactions of message-dependent rabbit reticulocyte lysate containing 1 mCi of [^{35}S]methionine (Amersham Corp.) per ml. Portions representing 10% of the reaction mixture were analyzed on 10% sodium dodecyl sulfate-polyacrylamide gels according to Laemmli (12).

For immunoprecipitation, portions (40%) of the reaction mixture were diluted 10-fold with Tris-buffered saline and reacted with 5 μl of tumor-bearing rabbit (TBR) serum (supplied by Joan Brugge) overnight at 4°C. Protein A-Sepharose was used to precipitate the antibody-antigen complex. The precipitated *Dsrc28C* protein was then eluted with 0.2 M glycine, pH 3.0, neutralized, and analyzed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis.

In situ hybridization. The template for the synthesis of antisense RNA probes was the same as that described above for RNA blot analysis. The cloned DNA was truncated at the polylinker *Sal*I site before in vitro transcription with SP6 RNA polymerase, according to the instructions provided by the manufacturer, to produce a 0.6-kb RNA probe. To construct the *Drosophila* epidermal growth factor receptor transcription template, a 3.5-kb *Eco*RI-*Hin*dIII genomic fragment containing virtually all of the amino acid coding sequences (29) was cloned into the Riboprobe pSP65 vector. The clone was truncated at the *Hin*dIII site before transcription to produce a 3.5-kb probe. The DNA templates were removed by brief DNase treatment, and the probes were then extracted with phenol and chloroform, ethanol precipitated, and suspended in distilled water. The probes were hydrolyzed to approximately 150 bases before hybridization (5). Ovaries were dissected from adult females and were transferred to O.C.T. embedding medium (Miles Laboratories) on a cryostat specimen holder, frozen by placing the specimen holder on dry ice, and stored at −70°C. Embedded tissue was sectioned in a cryostat at −20 to −24°C. Sections measuring 8 to 10 μm were collected on poly-L-lysine-coated slides (5) and placed on a slide warmer at 42°C for 25 to 45 min. The tissue sections were fixed to the slides and made permeable by the method of Hafen et al. (8).

For in situ hybridizations, the method of Cox et al. (5) was modified as follows. The RNA probe, carrier RNA (calf liver tRNA), and poly(A) were heated at 80°C for 3 min in distilled water before the other ingredients were added to yield the following final concentrations: 50% formamide, 0.6 M NaCl, 10 mM Tris hydrochloride (pH 7.5), 1 mM EDTA, 1× Denhardt solution (0.02% each of Ficoll, polyvinylpyrrolidone, and bovine serum albumin), 10% dextran sulfate, 1.5 mg of tRNA per ml, 0.5 mg of poly(A) per ml, and 0.2 to 0.3 mg of RNA probe per ml. Hybridizations were performed overnight at 48°C in a volume of 20 μl per slide, under baked cover slips sealed with rubber cement. Cover slips were removed, individual slides were treated with a solution containing RNase (20 μg/ml) in 0.6 M NaCl, 10 mM Tris hydrochloride (pH 7.5), and 1 mM EDTA at 37°C for 30 min, and then groups of slides were washed in the RNase buffer at 37°C for 30 min, in 2× SSC (1× SSC is 0.15 M NaCl plus 0.015 M sodium citrate) at room temperature for 30 min, and in 0.1× SSC at 48°C for 30 min, and dehydrated in a graded ethanol series. Autoradiography was as described by Pardue and Gall (18).

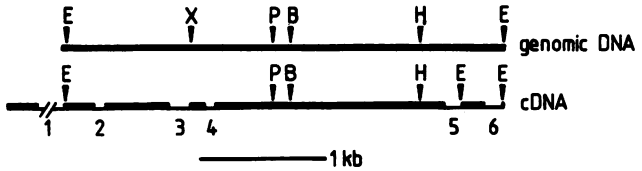

FIG. 1. Restriction maps of *Dsrc28C* genomic and cDNA clones. Colinear restriction maps of the genomic *Dsrc28C* restriction fragment isolated by Wadsworth et al. (28) and *Dsrc28C* cDNA clones described in this study. Introns are indicated by thin lines within the cDNA map. The actual size of intron 1 is not known but is greater than 5 kb, as determined by analysis of overlapping genomic DNA clones. Another intron(s) totaling approximately 100 base pairs also occurs between the *Bam*HI and *Hin*cII sites, as determined by comparative restriction mapping, but has not been precisely placed within the primary sequence. The *Xba*I site in the genomic DNA clone occurs within an intron. The extra *Eco*RI fragment within the cDNA sequence is created when intron 5 is spliced from the primary transcript. Abbreviations: E, *Eco*RI; X, *Xba*I; P, *Pst*I; B, *Bam*HI; H, *Hin*cII.

RESULTS

Isolation of *Dsrc28C* cDNA clones. *Drosophila Dsrc28C* cDNA clones were isolated from lambda phage libraries by hybridization with genomic *Dsrc28C* sequences. Two independent clones (λ-6311 and λ-6111) were recovered from the embryo library, and a single clone (λ-bud) was isolated from the adult female library. λ-6311 contained two *Eco*RI restriction fragments totaling 2.45 kb of insert cDNA, and λ-bud contained two *Eco*RI fragments totaling 0.65 kb of insert cDNA (restriction mapping indicated that clone 6111 was entirely contained within λ-6311). λ-bud was shown by DNA sequencing to overlap the 3′ end of λ-6311 by 150 nucleotides, so that the combined clones contained 2.95 kb of unique *Dsrc28C* cDNA. A comparative restriction map of the *Dsrc28C* genomic and cDNA clones is presented in Fig. 1. The positions of introns within the genomic sequence, deduced from sequence comparisons (see below), are also indicated.

Sequence analysis of *Dsrc* cDNA clones. The *Dsrc28C* cDNA sequence derived from λ-6311 and λ-bud is presented in Fig. 2. Although no poly(A) sequences were isolated with the cDNA clones, a potential polyadenylation site (AATAAA) is located approximately 25 nucleotides upstream from the 3′ end of the cDNA sequence. The *Eco*RI site at the 3′ end of the cDNA sequence corresponds to the *Eco*RI site at the 3′ end of the 3.3-kb genomic clone. This suggests that sequences 3′ to the end of the genomic clone may be transcribed in vivo and were not recovered in our cDNA clones. By comparing cDNA sequences with those from selected genomic DNA fragments, six intron positions were located, and their positions are indicated in Fig. 2. As

```
1    GAATTCCCCCCCCCCCCCCAAAAATAAATCGAGTTAAAAAACGCGTGTGCAACGCCAGCGTTTTTGGTTAGCCATCTCAGC
81   CGTTCCCAAATGATTCCCTGCGTGAGTTTGGAGAAACGAGCGTCATTGGCAACATGAAGGAGCGGGTCAAGGAAATGAAG
161  GTGTTCGGCTGCCGACTCAACTTCTGGAACCACATTGGTCACGAGCCTGACCAGTTCCAAAACCAAAGAAGGCAACGGCG
241  AGTTCTCCAGCCCAGAATTCAACGCGGCAGCAGTCAGTCCCAACAGCTCCACAACCAACAGTCAGTTCAGCTTGCAGCACA
321  ACAGCTCGGGGAGTCTCGGCGGAGGAGTGGGCGGCGGATTGGGAGGTGGAGGCAGCCTTGGCCTAGGCGGCGGAGGAGGA
401  GGTGGCGGCAGTTGCACGCCCACATCGCTACAGCCTCAGTCCTCGCTGACAACTTTCAAGCAGTCCCCAACTCTATTGAA
481  CGGCAACGGAAATCTATTGGATGCCAATATGCCTGGCGGTATACCCACTCCCGGAACTCCAAATTCCAAAGCCAAGGACA
561  ATTCCCACTTTGTCAAGCTGGTGGTGGCTCTCTATTTAGGCAAGGCCATCGAAGGCGGCGATTTATCCGTTGGGGAAAAG
641  AATGCCGAATACGAAGTGATAGATGACTCACAAGAGCACTGGTGGAAGGTCAAAGATGCCTTGGGGAATGTCGGCTATAT
721  ACCCAGCAACTATGTCCAAGCCGAAGCTCTTCTGGGCCTGGAGCGCTATGAATGGTATGTGGGCTACATGTCACGACAGA
801  GAGCCGAATCGCTGCTGAAGCAAGGTGATAAGGAGGGCTGCTTTGTGGTCAGGAAGTCATCGACCAAGGGTCTCTACACA
881  CTATCGCTGCATACCAAAGTTCCACAGTCGCATGTGAAGCACTACCACATCAAGCAAAATGCTCGCTGCGAGTATTATTT
961  GAGTGAGAAGCACTGCTGTGAAACAATTCCGGATCTGATTAACTACCATCGCCACAACTCTGGGGGTTTGGCTTGCCGAC
1041 TCAAATCCTCGCCCTGCGATCGCCCAGTTCCACCGACGGCGGGATTGTCCCACGACAAATGGGAGATCCATCCGATCCAA
1121 CTGATGCTGATGGAGGAGCTTGGATCGGGACAGTTTGGTGTTGTGCGGCGCGGCAAGTGGCGTGGGTCTATCGACACAGC
1201 GGTCAAGATGATGAAGGAAGGAACCATGTCCGAGGACGATTTCATTGAGGAGGCCAAGGTGATGACCAAGCTGCAGCATC
1281 CAAATCTTGTGCAGCTATATGGCGTCTGCACCAAGCACCGGCCCATCTACATTGTGACCGAGTACATGAAGCACGGATCC
1361 TTGTTGAATTACTTGCGACGGCATGAGAAGACCCTGATTGGTAATATGGGTCTACTCCTTGACATGTGCATACAGGTTAG
1441 CAAGGGAATGACCTACCTGGAGCGCCATAACTACATTCACCGGGATCTGGCTGCCCGCAACTGTCTCGTGGGTTCCGAGA
1521 ATGTCGTTAAAGTGGCCGACTTTGGATTGGCCCGATACGTTCTCGACGATCAGTATACCAGCTCTGGCGGAACCAAGTTC
1601 CCCATCAAGTGGGCCCCGCCCGAAGTGCTCAACTACACGCGCTTCTCCTCTAAGAGCGATGTGTGGGCATACGGTGTGCT
1681 GATGTGGGAGATCTTCACTTGCGGCAAGATGCCATACGGTCGCCTAAAGAATACCGAGGTTGTGGAGCGTGTGCAGCGCG
1761 GAATTATCCTAGAAAAACCAAAGTCGTGTGCCAAGGAGATTTATGATGTCATGAAGTTGTGCTGGTCACATGGACCCGAG
1841 GAGCGTCCCGCGTTCCGTGTGCTCATGGATCAGTTGGCTCTTGTGGCCCAAACGCTAACCGACTAAGGAGCGGACCAGAT
1921 TGCGGCACCCATCGCTGAAATTAACACGCACTGAACACACTGAAGCATCTTAAGATATTGAAGAACGTTAAAGGCATTAC
2001 ATTTAAGGTTACAAAAGCGGCGAAAGAAAGATATATTTAGATTTATACATATATGTATATGAAGCATAAGATTTAGTTAA
2081 GCGCGAGCCAGTTTACATTACGGCACAAGCCCCATATGATTATGGCATATTTGGATTATATGATAGAGCTAGAGCTGAGC
2161 AAATGATATCATTATGTTTATGAACGCGTTAACTGTTAGACGATGAGATACGAAGCGCAACAACTGACAGTCACATATAG
2241 TAACTATATTAAATGAGAAAACGGAAATGCATTAAAGTATAGAGAACTGTAGTTAATAGTCAAATTTTTGATACATTTCA
2321 CATACCGAAAGCGATATAAAAAAATTATAACGATAGATGCATATGATACGATATTTAAATGTGCTTTTTTAAATAGCGAG
2401 CCACTGTGCTAAGTAAAATATGATTGGATCTCTAAAGAAACAACAAACAAGAAAACGGTTATGCATTAAATATTAAATGT
2481 AAGCCTCAAGTCGAAAGCGAAATGTTTGCACAACTGATAAAGGATGAAGGAACCGAAGGATGATATGAAAAGAGTTAGCC
2561 AAGAAACATAAACGATTTTTAGACTAGATTTGTACGTGCGTAGTTTAAGTTTTGTGATTTATTGTTTAATTCTTGCAGAC
2641 GATTTTTATTTTTATTTAATTTTCATCTTGCATGGCGATCACCCACGCTCAAAGTATGAATCAAGTACGGTATAATTACG
2721 AGAAATTTATACGTTTATGCTTTATATTTTAACTGAATTCTTTCGTTTTTTTGCAATGTCTATGCGATATAAATCTATAT
2801 ATCTATACACATATTTTACCGTATTTTTATGAACGTTTTACTTAATTAATTTTAGCATTTGATTAAGCAACCAGAAATAC
2881 AAAAAAGAAAAGAAAAAACAGAAAATCAAATAAAATGTACGATAAACATTTAAAGAATTC
```

FIG. 2. Primary structure of *Dsrc28C* cDNA. The initiation and termination codons of the 590 codon open reading frame are boxed. A potential polyadenylation signal (AATAAA) is underlined. The positions of known introns in the genomic sequence are indicated with arrows.

the complete sequence of the genomic *Dsrc28C* clone was not determined, additional introns may be present within the genomic *Dsrc28C* DNA. Analysis of the nucleotide sequence revealed the presence of a 590-amino-acid open reading frame starting 140 nucleotides downstream from the 5′ end of the λ-6311 insert. This open reading frame is initiated by an ATG codon whose flanking sequences conform well to the *D. melanogaster* initiator consensus (D. Cavener, personal communication). A single in-frame stop codon is present 67 nucleotides upstream of the potential initiator.

Homology of *Dsrc28C* with *src* family tyrosine protein kinases. In Fig. 3, we present a conceptual translation of the *Dsrc28C* open reading frame and compare this protein with the prototypical member of this gene family, $p60^{c\text{-}src}$. Several gaps have been introduced into the sequences to maximize homologies. This comparison reveals that *Dsrc28C* bears considerable homology to c-*src* throughout 70% of its length; only the N-terminal 140 amino acids of *Dsrc28C* appear completely unrelated. Of these 140 amino acids, 58 make up an N-terminal extension compared with c-*src*, resulting in a protein with a molecular mass 6 kilodaltons larger than that of the c-*src* gene product. Within the kinase domains of these two proteins, defined here as extending from amino acid 270 to 470 of $p60^{c\text{-}src}$, identical amino acid residues are present at 54% of the positions. Several features found to be highly conserved in *src* family tyrosine kinases (11, 27, 30), including lysine-295, tyrosine-416, and leucine-516, are present in *Dsrc28C*. The kinase domain segment of the *Dsrc28C* open reading frame was also compared with other members of the *src* kinase family. Although *Dsrc28C* did appear to be more closely related to the *src* and *abl* class of kinases, this comparison revealed no clear evolutionary relationship of *Dsrc28C* with a specific member of this gene family (Table 1).

Cell-free translation of *Dsrc28C* RNA sequences. A segment of *Dsrc28C* cDNA containing the complete open reading frame was transferred to a plasmid vector downstream from a bacteriophage T7 promoter in the translational sense orientation. Increasing amounts of in vitro-transcribed RNA from this construct (pWV411) were translated in a rabbit reticulocyte message-dependent lysate (19), and the products were displayed on sodium dodecyl sulfate-polyacrylamide gels (Fig. 4). At the lower *Dsrc28C* RNA levels, the predominant translation product corresponded to the molecular weight predicted from the cDNA sequence shown in Fig. 3. With increasing amounts of RNA added, additional translation products were observed. The apparent molecular weights of the less abundant translation products were consistent with translational initiation at internal methionine residues in the *Dsrc28C* cDNA sequence. The appearance of these translation products was dependent on the addition of *Dsrc28C* RNA. Homology of the *Dsrc28C* protein to the *src* gene product is further substantiated by the fact that the cell-free translation products were recognized by two independent antisera (TBR sera) from rabbits bearing Rous sarcoma virus-induced tumors (Fig. 4, lanes E and F). Four additional TBR sera tested did not recognize the *Dsrc28C* polypeptides (data not shown). The ability of the two antisera to recognize the smaller *Dsrc28C* translation products resulting from internal initiation, as well as the full-length protein, indicates that the antisera are recognizing epitopes within the carboxyl half of the intact protein, the segment containing the highly conserved kinase domain.

Developmental expression of *Dsrc28C* mRNA. The temporal pattern of *Dsrc28C* RNA expression was determined by hybridization of a cRNA probe with total RNA from various developmental stages (Fig. 5). This probe hybridizes to the 3.5-kb RNA species previously identified as *Dsrc28C* RNA (28). From each developmental stage, 5 μg of total RNA was electrophoresed in denaturing formaldehyde gels, transferred to nitrocellulose, and hybridized with the ^{32}P-labeled probe. The results of these experiments revealed that *Dsrc28C* RNA levels are highest during early to middle embryogenesis and decline thereafter. *Dsrc28C* RNA is not detectable in total larval RNA (larval RNA lanes not shown). However, imaginal disks, which represent a minor proportion of cells within third-instar larvae, preferentially express *Dsrc28C* RNA. After a minor peak of expression during the first 24 h of metamorphosis, *Dsrc28C* RNA levels decline to undetectable levels. As previously determined (28), expression in adult males is nearly undetectable and in females is primarily confined to ovarian tissue (see below), indicating that the *Dsrc28C* RNA expressed in early embryos is of maternal origin. However, the presence of comparable levels of *Dsrc28C* RNA per microgram of total RNA during early and midembryogenesis indicates that transcription of *Dsrc28C* also takes place in the developing embryo. Control experiments using an RNA probe complementary to the protein-coding region of the actin 5C gene yielded expression patterns comparable to published studies (data not shown; 7). We have not determined the origin of the bands seen in addition to the major 3.5-kb transcript in Fig. 5. These bands were not consistently observed and may represent either artifacts of the electrophoresis conditions or low levels of transcripts arising from alternative processing such as takes place with transcripts from the *Dash* and *Drosophila* epidermal growth factor receptor genes (22, 26).

Spatial distribution of *Dsrc28C* transcripts. To learn if the *Dsrc28C* RNA species were preferentially expressed in certain tissues, in situ hybridizations were performed. Even though the highest levels of *Dsrc28C* RNA are expressed during embryogenesis, we have not observed preferential expression in any region of the embryo. *Dsrc28C* transcripts appear to be uniformly distributed throughout the cells of embryo regardless of developmental stage (data not shown). Likewise, in situ hybridizations carried out on 24-h pupae have not revealed significant concentrations of the *Dsrc28C* transcripts in any tissue.

After early metamorphosis, *Dsrc28C* RNA levels fall to undetectable levels until adulthood, when ovarian tissue preferentially expresses *Dsrc28C* transcripts. Results of in situ hybridizations on ovaries from adult females revealed that *Dsrc28C* RNA sequences are localized within the cytoplasm of nurse cells and developing oocytes and are absent from the adjacent layer of follicle cells (Fig. 6a through d). As a control for nonspecific binding of probe, in situ hybridizations were performed with antisense RNA transcripts from the *Drosophila* epidermal growth factor receptor homolog (14, 29). This probe hybridizes to ovarian tissue in a pattern complementary to that observed with probes specific for *Dsrc28C* (Fig. 6e and f; K. L. Kammermeyer and S. C. Wadsworth, Development, in press). These results suggest that within the ovary there is cell lineage-specific expression of the *Dsrc28C* gene.

DISCUSSION

The *Dsrc28C* cDNA clones isolated in this study comprise approximately 3 kb of unique sequence. As the mRNA for *Dsrc28C* is approximately 3.5 kb long, we have not isolated the complete cDNA sequence. However, since the sequence presented here contains a complete 590-amino-acid reading

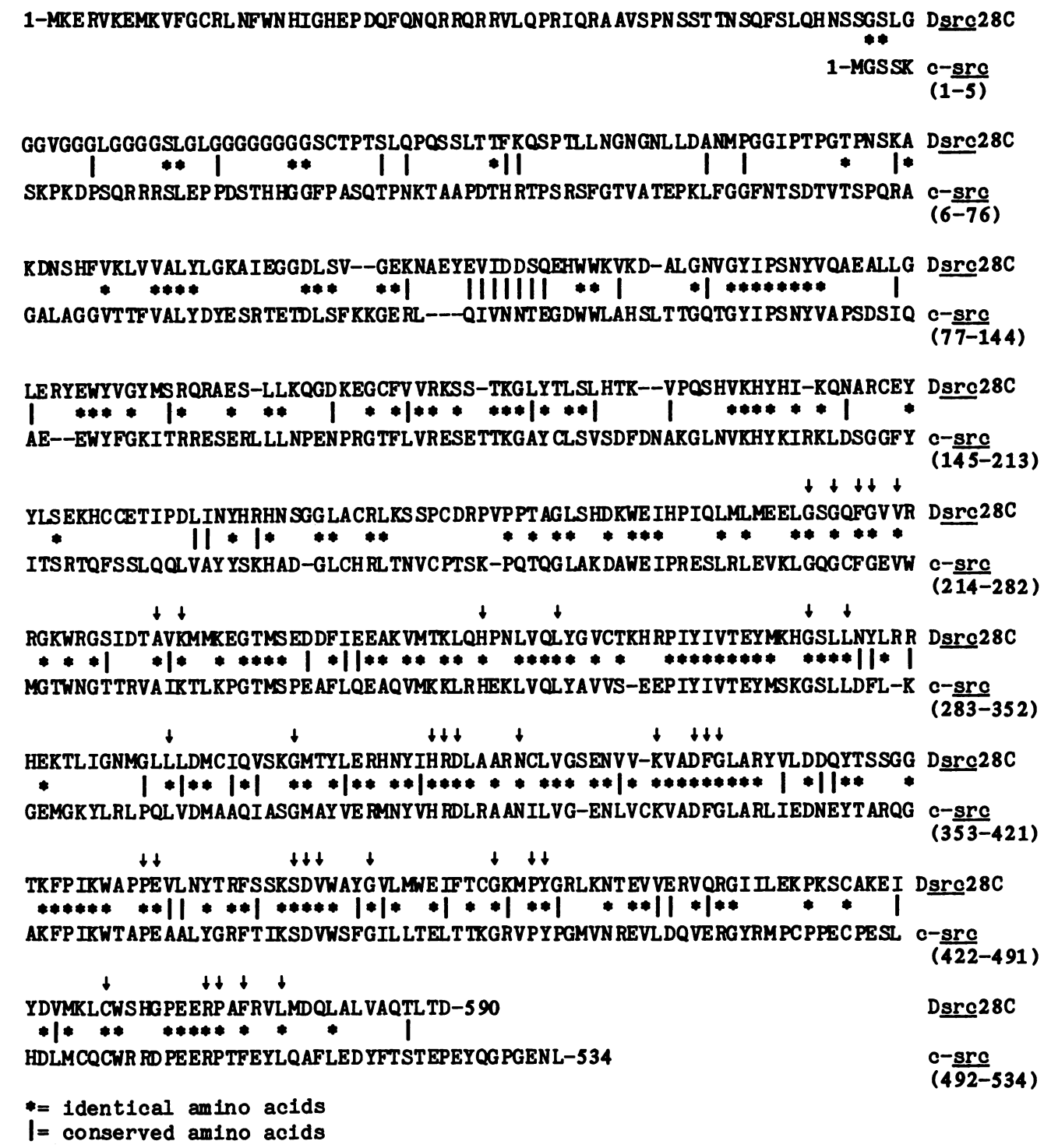

FIG. 3. Comparison of the predicted product of the *Dsrc28C* gene with the sequence of pp60$^{c\text{-}src}$. The conceptual translation of the *Dsrc28C* gene product is compared with the amino acid sequence of the c-*src* gene product as described by Takeya and Hanafusa (25). Conserved amino acids have been assigned according to Chambon et al. (3). Highly conserved features of tyrosine-protein kinases, noted by Van Beveren and Verma (27) and Yaciuk and Shalloway (30), are indicated with arrows.

frame, the bulk of the missing sequence is apparently derived from the 5′ and 3′ untranslated regions of the mRNA. We have determined by comparison with limited sequence data from genomic clones that the primary transcript of the *Dsrc28C* gene contains at least six introns of various sizes. The intron(s) nearest the 5′ end of the gene (Fig. 1) appears to be at least 5 kb long, as we were unable to locate sequences from the 5′ end of the cDNA within genomic lambda clones extending over 5 kb upstream. Since the entire protein-coding sequence of the *Dsrc28C* gene was

TABLE 1. Homology comparison of the kinase domain of *Dsrc28C* with that of other *src* family genes[a]

Gene	% Homology
c-*src*	54
Dash	54
Dsrc	51
v-*fgr*	51
v-*abl*	50
v-*fps*	44
v-*fes*	43
v-*fms*	39
c-*raf*	28

[a] The kinase domain is defined here as corresponding to amino acids 270 to 470 of pp60$^{c\text{-}src}$. The percent homology was calculated by dividing the number of exact matches by 200 (the length of the compared segments).

obtained from the cDNA clones, no further attempt was made to locate 5′ noncoding sequences within the genomic DNA. However, only 33 amino acids of coding sequence from the amino terminus were not verified by comparison with the genomic DNA sequence. As only one cDNA clone having the extreme 5′ DNA sequences was isolated, and this segment was not analyzed in the genomic DNA, the formal possibility exists that the 33 amino-terminal amino acids of the open reading frame do not represent the structure of the primary *Dsrc28C* gene product in vivo.

The protein encoded by the *Dsrc28C* open reading frame has several interesting properties. Among these is a stretch of 27 amino acids in the N-terminal third of the protein which is 74% glycine, including 8 consecutive glycine residues. This segment of the protein should be relatively unstructured, allowing the amino-terminal region considerable freedom of movement with respect to the rest of the protein. The protein encoded by the open reading frame is quite basic (pI of 10.2). The amino-terminal segment of the protein preceding the glycine-rich region contains an unusual number of basic amino acids, resulting in a pI of 12.2 for this portion of the protein. By comparison, the pIs of c-*src* and *Dsrc* are 8.7 and 6.8, respectively. In addition, the glycine at position 2 of the avian *src* protein, which serves as the myristylation site responsible for anchoring p60^{src} to the cell membrane (6), is not present adjacent to the potential initiator methionine of *Dsrc28C*. The conceptual translation also reveals no N-terminal hydrophobic domain capable of anchoring the protein to a cell membrane. Thus, *Dsrc28C* may not be membrane associated, or it may instead interact with another membrane-bound component. In contrast, the *Dsrc* gene product is predicted to have glycine in position 2 and thus may be anchored to the membrane by myristylation of this amino acid (23). Although in vitro translation initiates primarily at methionine codon 1 of the open reading frame, it is possible that many of the unusual features of the amino terminus of *Dsrc28C* might not be expressed in vivo if the protein initiates at methionine-115 of the open reading frame. The lack of homology with other tyrosine kinases upstream of this point indicates that initiation from this methionine is a possibility, although the sequences surrounding this methionine codon do not match the *D. melanogaster* initiator consensus as well as the sequences surrounding methionine-1 of the open reading frame do (D. Cavener, personal communication). Because of a slightly shorter carboxyl terminus, *Dsrc28C* also lacks the amino acid equivalent to tyrosine-527 of c-*src*, an amino acid conserved in *Dsrc*. Phosphorylation of this amino acid has been implicated as a mechanism for negatively regulating p60$^{c\text{-}src}$ kinase activity (4). While the close correlation between the predicted size of the *Dsrc28C* protein and the cell-free translation product derived from the cDNA indicate that these unique structural features of the protein are present in at least the primary in vitro translation product, immunological analyses will be necessary to verify the expression of these features in vivo.

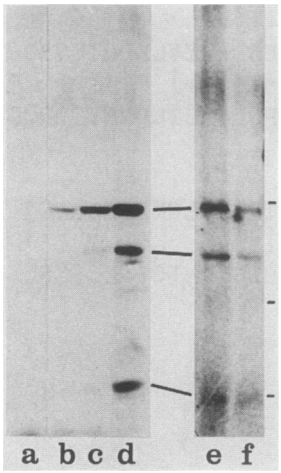

FIG. 4. In vitro translation of *Dsrc28C* RNA and immunoprecipitation of the translation products. RNA transcripts containing the complete *Dsrc28C* open reading frame were synthesized in vitro and translated in a rabbit reticulocyte message-dependent lysate containing [^{35}S]methionine. Samples were either directly analyzed by electrophoresis through denaturing sodium dodecyl sulfate-polyacrylamide gels (lanes a to d) or immunoprecipitated with anti-*src* sera from TBR before electrophoresis (lanes e and f). Lane a, No RNA added to message-dependent lysate; lanes b through d, ~0.2, 0.8, or 2 μg of *Dsrc28C* RNA added to each reaction; lanes e and f, proteins precipitated from reactions of *Dsrc28C* RNA and message-dependent lysate with two independent TBR sera. Lanes a to d were exposed for 1 day, and lanes e and f were exposed for 30 days. Molecular weight markers of 66,000, 45,000, and 29,000 are indicated on the right side of the figure.

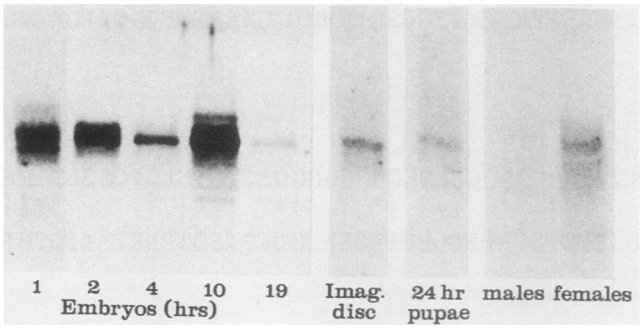

FIG. 5. Developmental expression of *Dsrc28C* mRNA. Total RNA (5 μg) was prepared from animals at progressive stages of development, fractionated on formaldehyde-agarose gels, transferred to nitrocellulose, and hybridized with an antisense RNA probe transcribed from a 0.6-kb *Xba*I-*Pst*I segment of *Dsrc28C* genomic DNA (Fig. 1). This probe hybridizes predominantly to a 3.5-kb RNA species previously identified as *Dsrc28C* mRNA (28). No hybridizing RNA was detected in whole first-, second-, or third-instar larvae. The exposure times are equivalent for all lanes. The embryonic time points represent the number of hours after egg laying. Other time points represented are imaginal disks from third instar larvae (Imag. disc, 24 h postpupariation (24 h pupae), adult males, and adult females.

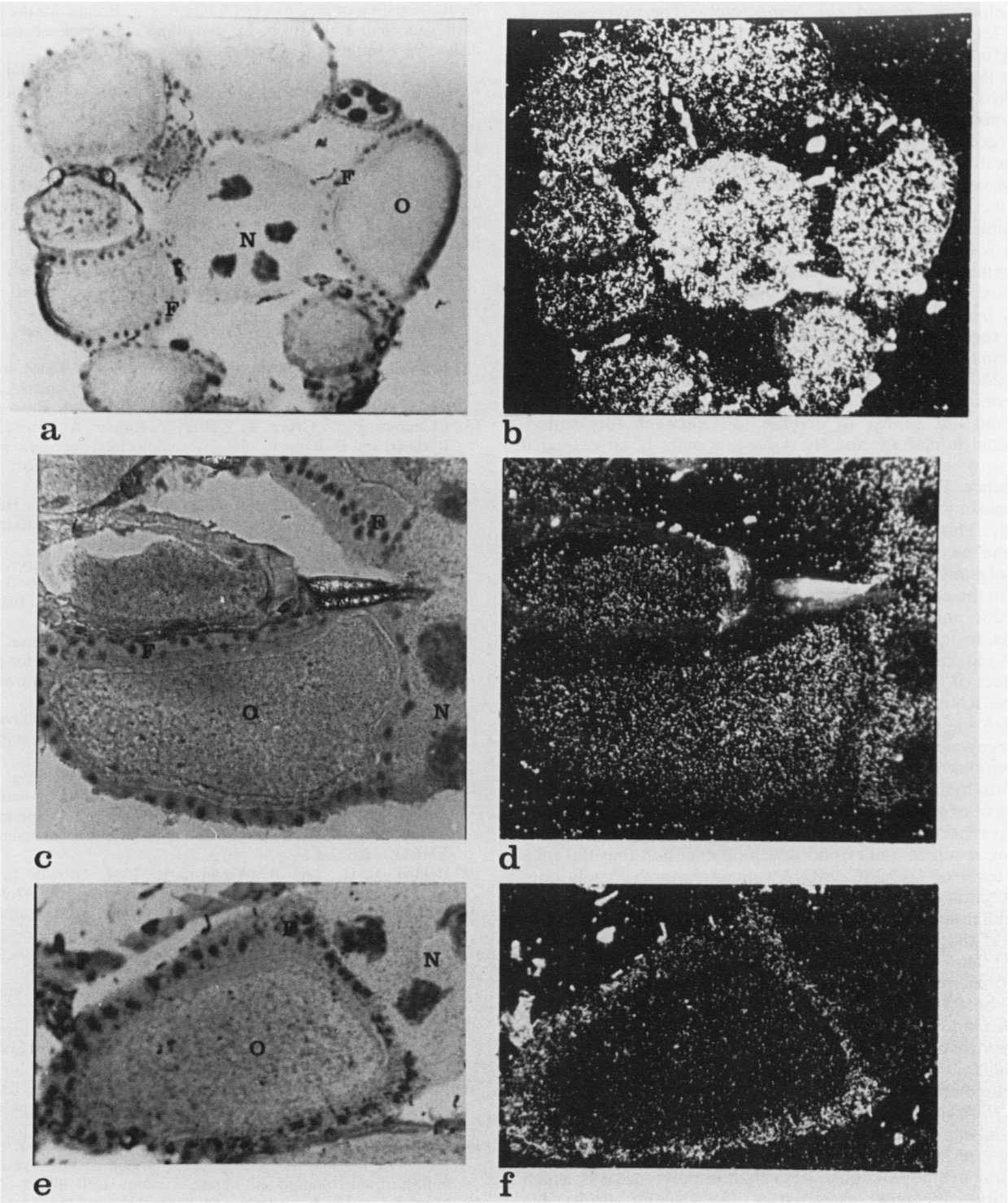

FIG. 6. In situ hybridization of *Dsrc28C* and DER antisense probes to sections of ovary. (a and c) Giemsa-stained sections hybridized with antisense *Dsrc28C* probe; (e) Giemsa-stained section hybridized with antisense *Drosophila* epidermal growth factor receptor probe; (b, d, and f) dark-field illumination of the same sections shown in a, c, and e, respectively. Autoradiographic exposure was 5 to 7 weeks for both probes. Abbreviations: O, oocyte; N, nurse cells; F, follicle cells.

Homology comparisons of the *Dsrc28C* kinase domain with other members of the tyrosine protein kinase family revealed that the protein is most closely related to the *src* and *abl* kinases. Although *Dsrc28C* sequences had previously been shown to hybridize to a probe from the v-*fps* oncogene (24), the sequence comparisons shown in Table 1 clearly indicate that this is a more distant relationship. Interestingly, the degree of homology of *Dsrc28C* with chicken c-*src* is as high or higher than its homology with either *Dsrc* or *Dash* (Table 1), the *D. melanogaster*

homologs of *src* and *abl*, respectively. The divergence of *Dsrc28C* from the other *D. melanogaster src* family genes is therefore probably an evolutionarily distant event and suggests that these kinases have evolved to fulfill different roles in *Drosophila* development. The amino-terminal regions of the various members of the tyrosine kinase gene family have diverged significantly compared with the kinase domain, suggesting that each protein has a specialized function. However, the homology of *Dsrc28C* with the *src* and *abl* members of the kinase gene family extends into the N-terminal regions of these proteins, although at a lower level than within the kinase domain. This homology includes a sequence of eight amino acids (GYIPSNYV) that is identical in c-*src* and *Dsrc28C* but is only partially conserved in *Dsrc*. Only *src*, *yes*, and *abl* have been shown to have homology in this region (2), pointing to a common origin for these proteins and *Dsrc28C*. In addition, the point at which the *Dsrc28C* and c-*src* sequences appear to diverge within the N-terminal region is also the point at which the sequences of *src* and *abl* appear to diverge (2). Between this amino-terminal homology and the kinase domain is an extended region of approximately 90 amino acids centered on the sequence FVVRKSSTKGLYTLS with homology to the cytoplasmic members of the tyrosine-protein kinase gene family. This domain, absent in tyrosine kinases that span the plasma membrane, may play a role in the interaction of cytoplasmic tyrosine kinases with specific effector molecules within the cell (20). A further indication of the relatedness of *Dsrc28C* and vertebrate *src* family genes is the ability of two of six tested antisera from rabbits bearing Rous sarcoma virus-induced tumors to recognize the in vitro translation product of the *Dsrc28C* gene.

The RNA blots shown in Fig. 5 illustrate that *Dsrc28C* mRNA is expressed predominantly in early to middle embryogenesis. Wadsworth et al. (28) have previously shown that *Dsrc28C* mRNA is maternally inherited, and the in situ hybridization results presented here confirm this mode of expression of the *Dsrc28C* gene. The data presented in this report indicate that transcription of the gene continues during much of embryonic development such that the relative level of *Dsrc28C* mRNA remains approximately constant until late embryogenesis. However, because in situ hybridization experiments did not reveal preferential expression of the gene in any cell lineage or tissue, no inference about the functional role of the gene product during embryogenesis can be drawn from these studies. Nevertheless, because of the abundant maternal expression of the RNA and because of its continual expression during embryogenesis, we hypothesize that the primary functional period for the *Dsrc28C* protein is during early embryonic development.

In third-instar larvae, we were able to demonstrate preferential expression of *Dsrc28C* transcripts in imaginal disks, suggesting a role for the *Dsrc28C* protein in these mitotically dividing cell populations rather than in polytene cells which constitute the bulk of the larva. The *Dsrc28C* transcripts seen in 24-h pupae could represent conserved imaginal disk transcripts or new gene transcription. As in developing embryos, in situ hybridizations failed to reveal preferential distribution of the *Dsrc28C* RNA in any tissue, suggesting that the pupal transcripts are present in many cell types at low levels.

ACKNOWLEDGMENTS

We thank Matthew B. Potter for skilled technical assistance. Thanks also are due to B. Yedvobnick, L. Kauvar, and T. Kornberg for the donation of lambda cDNA libraries, J. Brugge for the gift of TBR sera, and J. Fristrom for providing mass isolated imaginal disks. We also thank G. Gregory for preparing Fig. 1.

This work was supported by Public Health Service grant NCI RO1 CA-35911 from the National Cancer Institute and a grant from the March of Dimes Birth Defects Foundation to S.C.W. Central research facilities at the Worcester Foundation for Experimental Biology were supported in part by Public Health Service grant CA-12708 from the National Cancer Institute and the Mimi Aaron Greenberg Memorial Cancer Research Fund. K.L.K. was the recipient of a postdoctoral traineeship (T32 HDO7312). W.S.V. III was the recipient of a National Research Service Award (5F32 CA8011) from the National Cancer Institute.

LITERATURE CITED

1. **Bishop, J. M.** 1983. Cellular oncogenes and retroviruses. Annu. Rev. Biochem. **52**:301–354.
2. **Ben-Neriah, Y., A. Bernards, M. Paskind, G. Q. Daley, and D. Baltimore.** 1986. Alternative 5′ exons in c-*abl* mRNA. Cell **44**:577–586.
3. **Chambon, P., S. Green, P. Walter, V. Kumar, A. Krust, J. M. Bornert, and D. Banner.** 1986. Steroid hormone receptor genes: cloning, organization and expression. ICSU Short Rep. **4**:246–249.
4. **Cooper, J. A., K. L. Gould, C. A. Cartwright, and T. Hunter.** 1986. Tyr527 is phosphorylated in pp60c-*src*: implications for regulation. Science **231**:1431–1434.
5. **Cox, K., D. V. DeLeon, L. M. Angerer, and R. C. Angerer.** 1984. Detection of mRNAs in sea urchin embryos by *in situ* hybridization using asymmetric RNA probes. Dev. Biol. **101**:485–502.
6. **Cross, F. R., E. A. Garber, D. Pellman, and H. Hanafusa.** 1984. A short sequence in the $p60^{src}$ N terminus is required for $p60^{src}$ myristylation and membrane association and for cell transformation. Mol. Cell. Biol. **4**:1834–1842.
7. **Fyrberg, E. A., J. W. Mahaffey, B. J. Bond, and N. Davidson.** 1983. Transcripts of the six *Drosophila* actin genes accumulate in a stage- and tissue-specific manner. Cell **33**:115–123.
8. **Hafen, E., M. Levine, R. L. Garber, and W. L. Gehring.** 1983. An improved *in situ* hybridization method for the detection of cellular RNAs in *Drosophila* tissue sections and its application for localizing transcripts of the homeotic antennapedia complex. EMBO J. **2**:617–623.
9. **Helden, C.-H., and B. Westermark.** 1984. Growth factors: mechanism of action and relation to oncogenes. Cell **37**:9–20.
10. **Hoffmann-Falk, H., P. Einat, B.-Z. Shilo, and F. M. Hoffmann.** 1983. *Drosophila* melanogaster DNA clones homologous to the vertebrate oncogenes: evidence for a common ancestor to the *src* and *abl* cellular genes. Cell **32**:589–598.
11. **Hunter, T., and J. A. Cooper.** 1985. Protein-tyrosine kinases. Annu. Rev. Biochem. **54**:897–930.
12. **Laemmli, U. K.** 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature (London) **227**:680–685.
13. **Lev, Z., N. Leibovitz, O. Segev, and B.-Z. Shilo.** 1984. Expression of the *src* and *abl* cellular oncogenes during development of *Drosophila melanogaster*. Mol. Cell. Biol. **4**:982–984.
14. **Livneh, E., L. Glazer, D. Segal, J. Schlessinger, and B.-Z. Shilo.** 1985. The *Drosophila* EGF receptor homolog: conservation of both hormone binding and kinase domains. Cell **40**:599–607.
15. **Maniatis, T., E. F. Fritsch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Melton, D. A., P. A. Krieg, M. R. Rebagliati, T. Maniatis, K. Zinn, and M. R. Green.** 1984. Efficient *in vitro* synthesis of biologically active RNA and RNA hybridization probes from plasmids containing a bacteriophage SP6 promoter. Nucleic Acids Res. **12**:7035–7056.
17. **Messing, J., and J. Viera.** 1982. A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene **19**:269–276.
18. **Pardue, M. L., and J. G. Gall.** 1975. Nucleic acid hybridization

to the DNA of cytological preparations. Methods Cell. Biol. **10**:1–16.

19. **Pelham, H. R. B., and R. J. Jackson.** 1976. An efficient mRNA-dependent translation system from reticulocyte lysates. Eur. J. Biochem. **67**:247–256.
20. **Sadowski, I., J. C. Stone, and T. Pawson.** A noncatalytic domain conserved among cytoplasmic protein-tyrosine kinases modifies the kinase function and transforming activity of Fujinami sarcoma virus P130$^{gag\text{-}fps}$. Mol. Cell. Biol. **6**:4396–4408.
21. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
22. **Schejter, E. D., D. Segal, L. Glazer, and B.-Z. Shilo.** 1986. Alternative 5′ exons and tissue-specific expression of the *Drosophila* EGF receptor homolog transcripts. Cell **46**:1091–1101.
23. **Simon, M. A., B. Drees, T. B. Kornberg, and J. M. Bishop.** 1985. The nucleotide sequence and the tissue-specific expression of *Drosophila* c-*src*. Cell **42**:831–840.
24. **Simon, M. A., T. B. Kornberg, and J. M. Bishop.** 1983. Three loci related to the *src* oncogene and tyrosine-specific protein kinase activity in *Drosophila*. Nature (London) **302**:837–839.
25. **Takeya, T., and H. Hanafusa.** 1983. Structure and sequence of the cellular gene homologous to the RSV *src* gene and the mechanism for generating the transforming virus. Cell **32**: 881–890.
26. **Telford, J., J. Burckhardt, B. Butler, and V. Pirrotta.** 1985. Alternative processing and developmental control of the transcripts of the *Drosophila abl* oncogene homologue. EMBO J. **4**:2609–2615.
27. **Van Beveren, C., and I. M. Verma.** 1986. Homology among oncogenes. Curr. Top. Microbiol. Immunol. **123**:91–98.
28. **Wadsworth, S. C., K. Madahavan, and D. Bilodeau-Wentworth.** 1985. Maternal inheritance of transcripts from three *Drosophila src*-related genes. Nucleic Acids Res. **13**:2153–2170.
29. **Wadsworth, S. C., W. S. Vincent III, and D. Bilodeau-Wentworth.** 1985. A *Drosophila* genomic sequence with homology to human epidermal growth factor receptor. Nature (London) **314**:178–180.
30. **Yaciuk, P., and D. Shalloway.** 1986. Features of the pp60$^{v\text{-}src}$ carboxyl terminus that are required for transformation. Mol. Cell. Biol. **6**:2807–2819.