A long, nontranslatable poly(A) RNA stored in the egg of the sea urchin *Strongylocentrotus purpuratus*

Frank J. Calzone, James J. Lee,[1] Nga Le,[2] Roy J. Britten, and Eric H. Davidson

Division of Biology, California Institute of Technology, Pasadena, California 91125 USA

Nontranslatable transcripts containing interspersed repetitive sequence elements constitute a major fraction of the poly(A) RNA stored in the cytoplasm of both the sea urchin egg and the amphibian oocyte. We report the first complete sequence of a representative interspersed maternal RNA transcript, called ISp1. The transcript is about 3.7 kb in length [including poly(A) tail]; and the 5′ half consists of a cluster of repetitive sequences, whereas the 3′ half is single copy. Other repetitive sequences occur in the 5′ and 3′ regions flanking the transcription unit. In several cloned alleles, the flanking repetitive and single-copy sequences differ, indicating a high degree of insertional and deletional rearrangement around, as well as within, the transcription unit. No significant open reading frames exist in any region of the ISp1 transcript, nor is it spliced to give rise to translatable mRNA in egg or embryo. A 620-nucleotide repetitive sequence element at the 5′ end of the ISp1 transcript is also represented in a large number of other long interspersed maternal poly(A) RNAs. In addition, this sequence appears in a prevalent set of small polyadenylated RNAs about 600-nucleotides in length, which disappear almost completely by the gastrula stage of development. The structural features of the ISp1 RNA uncovered in this work exclude several hypotheses of interspersed maternal poly(A) RNA origin and function.

Nontranslatable RNA transcripts containing single-copy genomic sequences interspersed with repetitive sequence elements constitute about 68% of the poly(A) RNA in the cytoplasm of the unfertilized egg of *Strongylocentrotus purpuratus* (Costantini et al. 1980) and a similar fraction of the oocyte RNA of *Xenopus laevis* (Anderson et al. 1982; Richter et al. 1984). This class of quantitatively important egg and oocyte transcript has been termed interspersed maternal RNA. The developmental function of interspersed maternal RNA remains unknown, although its general properties have been characterized extensively (reviewed by Calzone et al. 1985; Davidson 1986). The overall complexity of *S. purpuratus* egg RNA is 3.7×10^7 nucleotides (Galau et al. 1976), of which about 70% is included in the polysomal mRNAs of the early embryo (Hough-Evans et al. 1977). The remaining 30% is probably to be accounted for as nonmessage interspersed poly(A) RNA sequence. Several

prior studies indicate that the structure and role of interspersed maternal RNAs differ from those of mature maternal messages. Thus, interspersed maternal poly(A) transcripts isolated from *Xenopus* oocytes are not translatable in vitro and are not loaded on polysomes in vivo when injected into the oocyte (Richter et al. 1984). Furthermore, the size of the interspersed maternal transcripts of the *S. purpuratus* egg (3000–15000 nucleotides) is much greater, on average, than that observed for mature embryonic messages (Posakony et al. 1983). The representation of repetitive sequences in interspersed maternal RNA is reminiscent of that observed for nuclear RNAs. However, interspersed maternal RNA is not simply nuclear RNA that has been randomly transported to the cytoplasm of the egg during oogenesis or in the course of isolation. This follows from the fact that the complexity of both sea urchin and amphibian oocyte germinal vesicle RNAs is 5- to 10-fold greater than that of the interspersed maternal poly(A) RNAs stored in the cytoplasm of the respective eggs (reviewed in Davidson 1986). In any case, nuclear RNA leakage during isolation has been directly excluded for *Xenopus* oocytes by physical removal of the intact germinal vesical prior to extraction of RNA (Anderson et al. 1982).

[1]Current address: Department of Genetics and Development, College of Physicians and Surgeons of Columbia University, New York, New York 10032 USA.
[2]Current address: Washington University, St. Louis, Missouri 63110 USA.

In this paper we present the first complete sequence of a representative interspersed maternal RNA transcription unit and describe the structure of its RNA product. It seems clear from these results that this transcript per se cannot be utilized for translation in either egg or embryo. Nor are translatable spliced derivatives formed during embryogenesis. Thus, long poly(A) RNAs distinguished by interspersed repetitive sequence organization, such as that characterized in this work, apparently represent a new class of maternal cytoplasmic transcript that is functionally distinct from either message or message precursor.

Results

Genomic sequence organization

Figure 1a provides a summary of the genomic sequence organization of the representative interspersed maternal RNA transcription unit that we cloned and characterized. In the following text, we refer to this as the ISp1 transcription unit. Five different repetitive sequence elements (A–E) were identified in the 5'-flanking and the transcribed sequences of the particular ISp1 allele (B1) diagramed at the top of Figure 1a. Each repetitive sequence element was characterized using restriction fragments of this allele as probes for genome blot hybridizations, as shown in Figure 2, and for hybridizations with the various other cloned alleles that are diagramed in Figure 1b (data not shown).

The length of repeat A is at least 2.8 kb. The genome blot hybridization patterns observed with probes for a portion of repeat A among different S. purpuratus individuals suggest a genomic repetition frequency of only two to four copies per haploid genome (Fig. 2b). However, certain regions of repeat A, at least as it occurs in allele B1, contain a higher prevalence sequence element. When included in a hybridization probe, this sequence reacts weakly with a much larger number of genomic restriction fragments (Fig. 2a). The genome blot hybridization patterns obtained with probes containing sequences of repeats B, C, D, and E (Fig. 2c,f,g,h) are typical for moderately repetitive, interspersed sequence elements that are represented in a few hundred to several thousand copies per genome (reviewed by Thomas et al. 1981; Calzone et al. 1985). The allelic distribution of genomic restriction fragments that react with probes derived from the region to the right of repeat E (see Fig. 1a)

demonstrates that this region is single copy (Fig. 2i; cf. single copy genome blots reported by Thomas et al. 1981; Posakony et al. 1983; Lee et al. 1984). The region separating repeats B and C is also single copy, although, as discussed below, some of the single-copy sequence present in allele B1 is absent from the homologous regions of other alleles.

Repeats C, D, and E are located within the ISp1 transcription unit, as indicated in Figure 1a. Transcript mapping experiments presented below show that these three contiguous repetitive elements are positioned in the 5' half of the ISp1 transcript and that single-copy sequences are confined to the 3' portion of the transcript. It is unlikely that repeats C, D, and E are always contiguous where they occur in other genomic locations. Thus, comparison of the number of restriction fragments reacting in the genome blot hybridizations with probes representing each repetitive element suggests that there are severalfold fewer copies per genome of repeats D and E than of repeat C. Therefore, repeat C must be present at many other genomic locations without repeats D and E.

Frequent rearrangements within the ISp1 transcription unit and its flanking sequences

Figure 2d shows that three of the four individual diploid genomes sampled lack the entire single-copy probe sequence of allele B1 utilized for this blot hybridization (i.e., probe 70; see Fig. 1a). Similarly, it can be seen that in Figure 2e one of the genomes of individual F and both of the genomes of individual M lack the adjacent single-copy sequence element (probe 40). Both alleles of individual F have been cloned, as shown in Figure 1a. In an extensive series of clone blot hybridizations (not shown), it was found that although all three alleles analyzed (i.e., B1, F1, and F2) share approximately 400 nucleotides of single-copy sequence immediately adjacent to the start of transcription, all homology between allele F2 and B1 ends at this point, whereas homology between alleles F1 and B1 extends about 200 nucleotides upstream. In each case, the homologous region is terminated by a different repetitive sequence, namely repeat B in allele B1, repeat G in allele F1, and repeat H in allele F2 (see Fig. 1a,b; Fig. 2k,l).

The sequence contents of the transcribed regions of the three alleles cloned are homologous, except that al-

Figure 1. Genomic sequence organization of ISp1 clones. (a) The structures of three allelic isolates of ISp1 are represented schematically. Boxes indicate different repetitive sequence elements identified by capital letters (A–H). Solid lines indicate single-copy genomic sequences. The black box in repeat A indicates a high frequency domain of this sequence. The black box in repeat E locates the 2109A-10 repeat E probe of Posakony et al. (1983), and the bracket in the adjacent single-copy segment the 10HB single copy probe utilized by the same investigators in their initial identification of the ISp1 transcript. cDNA clone 2109A-10 also includes the 10HB single-copy sequence probe. The lengths of repeats G and H may be shorter than indicated. The orientation and length of the primary transcript of ISp1 is shown directly above allele B1. Key restriction sites are mapped in ISp1 allele B1 for reference below. The probes used for genome blot, RNA gel blot, and nuclease mapping experiments are mapped below allele B1. (See Figs. 2–4 for details.) (b) Restriction maps of genomic and cDNA isolates of ISp1. A restriction map of the genomic inserts in λ clones representing each ISp1 allele in a is shown. Note that the scales in a and b are different. Restriction sites are as follows: AccI (A), AluI (B), AhaI (C), AvaII (D), BamHI (E), BglII (F), DdeI (G), EcoRI (H), HaeIII (I), HincIII (J), HindIII (K), HpaII (L), PstI (M), RsaI (N), SalI (O), Sau3AI (P), TaqI (Q), XbaI (R), XhoI (S), XmaI (T). The lines connecting the restriction maps of different genomic inserts indicate shared restriction sites.
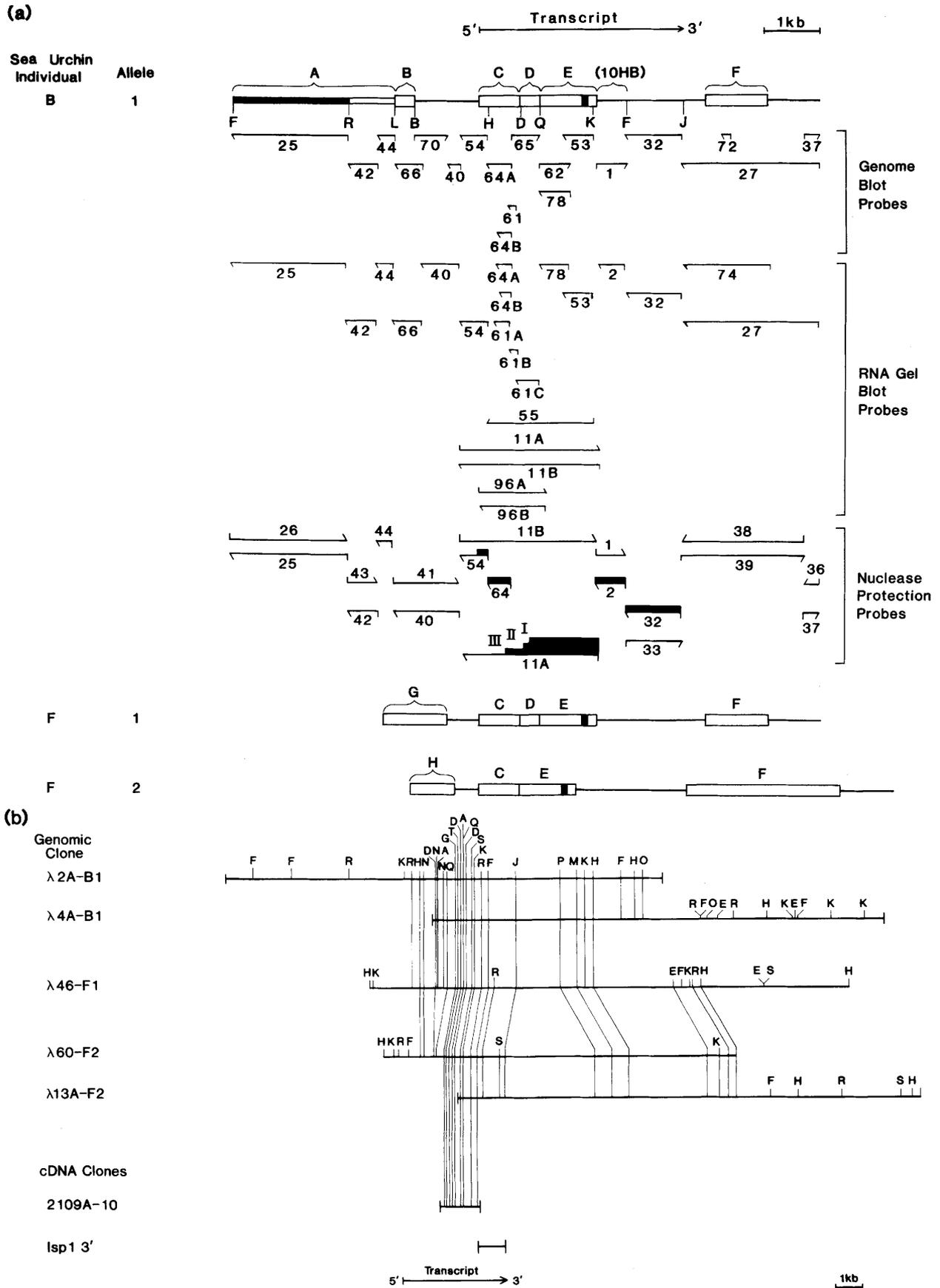
Figure 1. (*See facing page for legend.*)
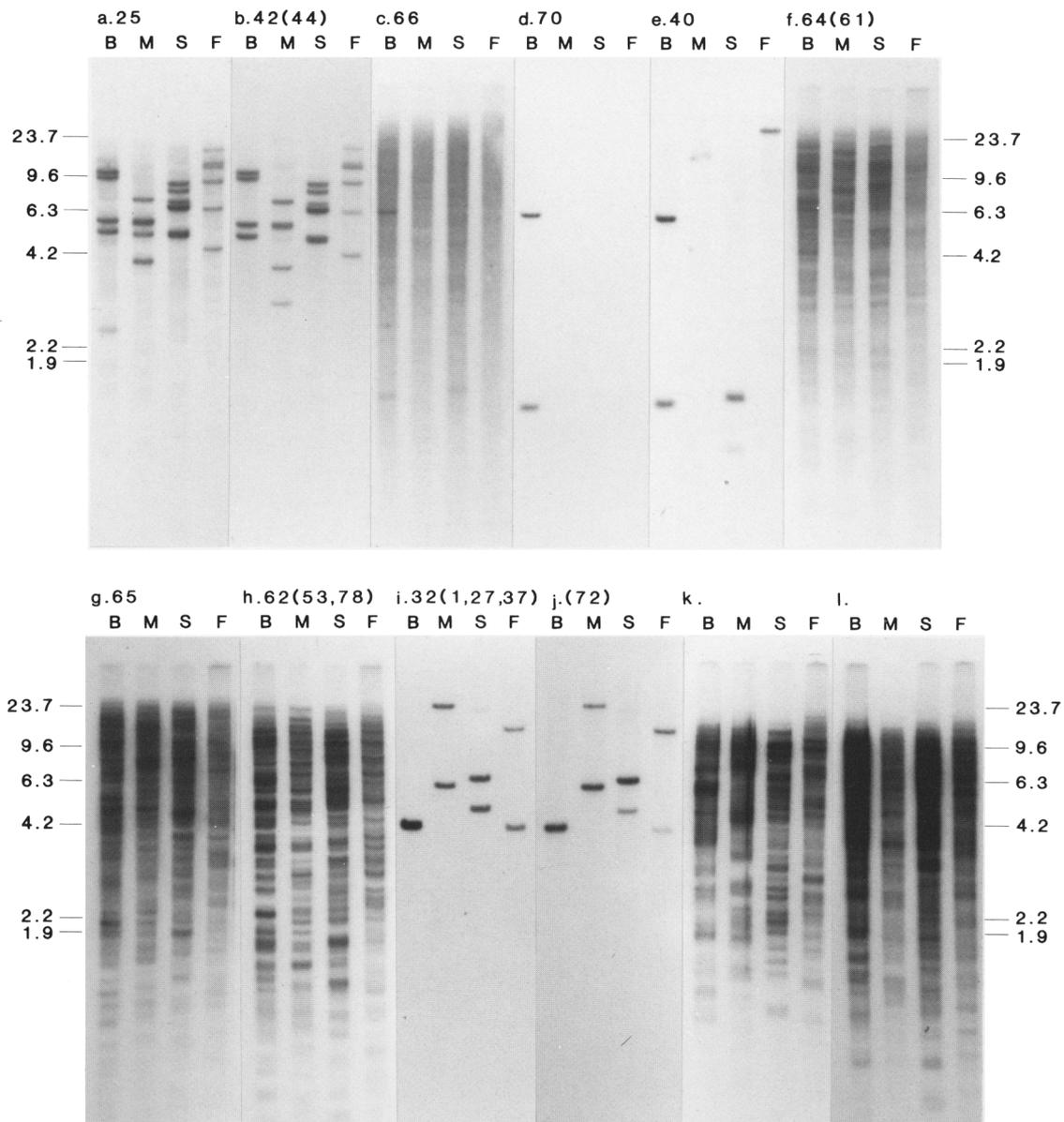
Figure 2. Sequence organization of ISp1 region established by genome blot hybridizations. DNA gel blots containing *Hind*III digests of sperm DNA of *S. purpuratus* individuals B, M, S, and F were reacted with a series of single-stranded DNA or RNA probes. [See Fig. 1a (genome blot probes) for the location in the map of individual probe sequences.] (*a–l*) Genome blots. In each case, the autoradiograph shown was obtained with the initial probe designated, and additional probes listed in the parentheses gave identical results to those shown. The probe used in *k* consisted of a DNA fragment representing repeat G of allele F1, and that used in *l* consisted of a fragment having the sequence of repeat H of allele F2.

lele F2 lacks completely the region defined as repeat D in allele B1. This is shown diagramatically in Figure 1a and can be perceived in more detail in the restriction maps of Figure 1b. Deletion of repeat D was not a cloning artifact, as the sizes of the relevant *Hind*III restriction fragments observed in the cloned isolates (i.e., λ46 and λ60; Fig. 1b) are the same as those displayed in the DNA of individual F by the mainly single-copy probe 54 utilized for genome blots (not shown).

A dominant feature of the 3′-flanking region of the ISp1 transcription unit is the repetitive sequence labeled F in Figure 1a. In allele B1, this consists of a tandem array of 40, more or less similar 25 nucleotide monomers, as shown explicitly in the primary sequence given below. Unlike the other repetitive sequence elements considered, repeat F is not interspersed in the genome but occurs in a single location, as demonstrated in Figure 2j. The number of monomeric elements included in repeat F, as it occurs in different genomes, varies considerably. Thus, although this number is similar in alleles F1 and B1, the greater length of repeat F in allele F2 indicates approximately 68 more monomeric

units. Similarly, the genome blots in Figure 2j show that in five of the eight alleles tested, the relative intensity of the restriction fragments detected with a short probe consisting of five copies of the 25 nucleotide element is not constant (compare the equivalent reactions of the single-copy probes represented in the same restriction fragments in the experiments of Figure 2i).

In summary, these data demonstrate the following: repetitive sequence insertions (or single-copy sequence deletions) in the 5′-flanking region of the ISp1 transcription unit beyond about 400 nucleotides from the start site; the deletion of an internal repeat sometimes included in the transcription unit; and alterations in the length of a satellitelike repeat near the 3′ terminus of the transcript. The polymorphism in restriction sites farther downstream, illustrated in Figure 1b, suggests that there may also exist many additional alternatives in sequence organization in the more distal 3′ flanking sequences of this highly variable genomic region as well.

The 3.7-kb ISp1 transcript is not spliced or processed in egg or embryo

Mapping the ISp1 RNA presented special problems because the entire 5′ region of this transcript consists of repetitive sequences also included in other maternal poly(A) RNAs. This can be seen in the RNA gel blot hybridizations shown in Figure 3a, which were carried out under relatively stringent criterion conditions (see legend), using a single-stranded repeat C probe. A heterogenous collection of large transcripts ranging in size from 2 to 15 kb is displayed by this probe in RNAs of egg-, blastula-, gastrula-, and pluteus-stage embryos. This is typical of the interspersed maternal poly(A) RNA sets observed with other repetitive sequence probes in earlier studies (Thomas et al. 1981; Posakony et al.

1983). As also observed previously, both complements of repeat C are represented in such transcripts (Fig. 3c). An unusual feature, however, is that the strand of repeat C represented in the ISp1 transcript, as shown below, is represented as well in a prominent set of discrete low-molecular-weight RNAs from 400 to 600 nucleotides in length (Fig. 3a). No such transcripts are displayed by the complementary probe (Fig. 3c). Figure 3a indicates that the prevalence of these small RNAs in egg poly(A) RNA is far greater than that of any of the large interspersed poly(A) RNAs. However, dramatic developmental changes in the population of these small, asymmetrically represented maternal transcripts occur early in embryogenesis. Thus, by the blastula stage they have clearly decreased in amount and seemingly have been altered from about 600 to 400 nucleotides in length as well. At later stages, these transcripts are barely evident. All repeat C probes representing the appropriate strand display these small early embryo transcripts (Fig. 3a), which we discuss in further detail elsewhere (F. Calzone et al., in prep.).

Figure 3b demonstrates that the single-copy probes 2 and 32 display exclusively the 3.7-kb ISp1 transcript (no reaction is observed with the complementary single-copy probes). The ISp1 transcript was seen earlier in a similar experiment of Posakony et al. (1983) using the single-copy 10HB probe, which is included in our probe 2. At the relatively exacting hybridization criterion, utilized probes 78 and 53, which contain sequences of repeat E, also reveal only the single ISp1 transcript (Fig. 3b). In contrast, this observation emphasizes the relatively low sequence divergence of the C family of transcribed repetitive sequences (and possibly D as well) compared with the repeat E transcripts. By comparison, in a lower criterion blot hybridization (20% formamide, 0.75 M NaCl, 40°C), Posakony et al. (1983) detected a
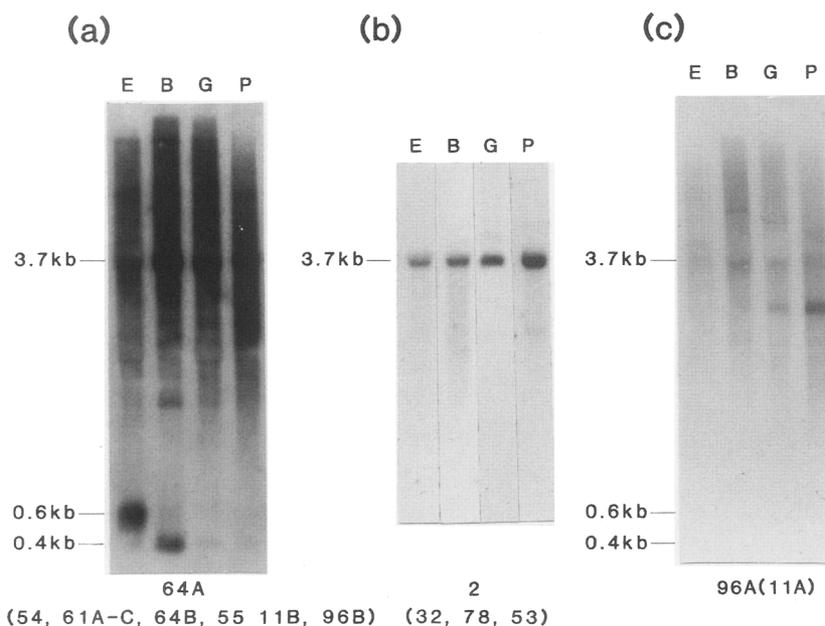
Figure 3. RNA gel blot hybridizations carried out with single-stranded probes. The map of the transcription unit and flanking regions showing the location and orientation of single-stranded probes is shown in Fig. 1a (RNA gel blot probes). Half-arrows pointing left indicate antisense probes with respect to the ISp1 transcript. (a–c) RNA gel blot hybridizations. Each lane contained 2 μg of poly(A) RNA from unfertilized egg (E)-, mesenchyme blastula (B)-, gastrula (G)-, or pluteus (P)-stage embryos. Criterion conditions were 50% formamide, 45°C, 0.75 M Na⁺ for RNA probes; 0.75 M Na⁺, 71°C for DNA probes. The probe used for each experiment is shown below the respective blot, and probes listed in parentheses detected a similar set of transcripts. None of the probes indicated in Fig. 1a other than those shown revealed any transcripts under the same conditions.

Calzone et al.

typical set of large interspersed maternal poly(A) RNAs (i.e., including ISp1), using the repeat E probe indicated in black in the diagram of Figure 1a. The most important conclusion that can be drawn from Figure 3b is that no smaller transcripts including the same single-copy sequences appear during embryonic development. Because probe 32 extends beyond the 3' terminus of the transcript (see below), this in itself excludes the possibility that ISp1 serves as a precursor for a spliced embryonic RNA that includes any portion of either the single copy region or repeat E.

S1 and exonuclease VII (ExoVII) nuclease protection experiments were carried out with many of the same single-stranded probes, as shown in Figure 1a. The full length of probe 2 (Fig. 4a) and nearly the full length of probe 32 (Fig. 4b) were protected from nuclease attack by maternal poly(A) RNA. The location of the 3' end of the S1 transcript within probe 32 was obtained from the DNA sequence of a 3' terminal cDNA clone (ISp1 3'; see Fig. 1). Because a unique fragment of probe 32 is protected, the 3' end of the transcript occurs at a discrete

termination site in the sequence, at least within the resolution afforded by this measurement. Probe 11A, which contains repeats C, D, and E, yielded a series of three large protected fragments when reacted with egg poly(A) RNA and then digested with ExoVII, an enzyme of choice because it does not attack the internal mismatches expected of repetitive sequence RNA–DNA duplexes. This experiment is shown in Figure 4c. The summary diagram presented in Figure 1a displays these three fragments as an overlapping series extending from the 3' terminus (with respect to the transcript) of probe 11A, through repeat E for various distances into repeats D and C in the 5' direction. This was demonstrated in a mapping experiment (not shown) in which the protected fragments were renatured with the complementary sequence, and the location of several internal restriction sites was determined (for method and details, see Calzone et al. 1987). Because probes for repeat E at high criterion detect only the 3.7-kb ISp1 transcript, the probable explanation for the formation of these different 5' termini is competition for the probe sequence in the re-
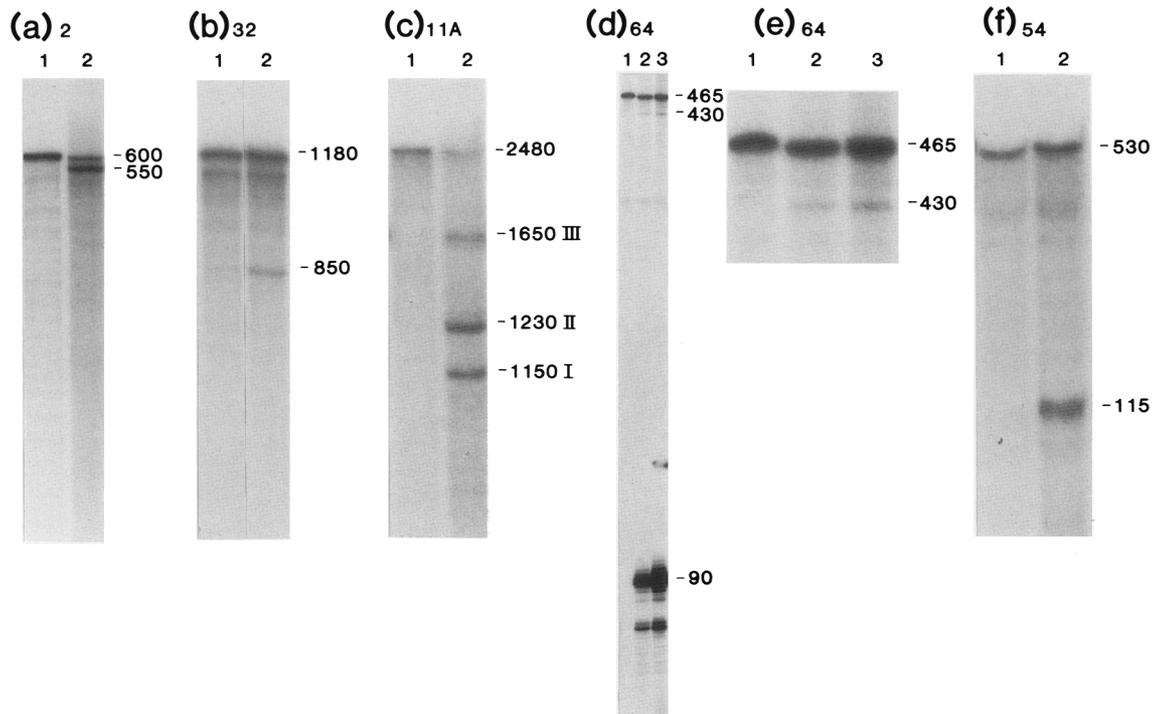
Figure 4. Nuclease S1 and Exo VII mapping studies carried out with unfertilized egg poly(A) RNA. The genomic sequences included in each single-stranded DNA probe used for these probe protection experiments are shown in Fig. 1a (nuclease mapping probes). The probes were labeled internally as described in Materials and methods. Not shown in Fig. 1a are the 5'- and 3'-vector sequence tails included in each probe that allowed fragments protected by RNA to be distinguished from full-length probe transcripts protected by the residual M13 template included in most probe preparations. Probe orientations with respect to the ISp1 transcript are indicated as for RNA gel blot reactions. Of those shown in Fig. 1a, only probes 2, 11A, 32, 54, and 64 contained sequences protected from nuclease attack by poly(A) RNA from the unfertilized egg. The sequences included in the protected probe fragments shown in (a–f) are indicated by the blackened boxes in the probe maps shown in Fig. 1a. Probe 2 (a), and probe 32 (b), containing single-copy sequences in the 3' portion of the ISp1 transcript were reacted with yeast RNA (lane 1) and egg poly(A) RNA (lane 2) and treated with nuclease S1. (c) Probe 11A was reacted with yeast RNA (lane 1) or egg poly(A) RNA (lane 2) and treated with Exo VII. Only fragments greater than 500 nucleotides were retained on this gel. (d) Lanes 2 and 3 contain the fragments of probe 64 protected from Exo VII digestion by 5 and 10 μg of egg poly(A) RNA respectively. The yeast RNA control is shown in lane 3. (e) An enlargement of the top section of the autoradiogram in d is shown. (f) The fragment of probe 54 protected from nuclease S1 by egg poly(A) RNA is shown in lane 2. The yeast RNA control is lane 1.

gion of repeat C by the various other transcripts in which this sequence is represented (cf. Fig. 3a). This competition is reduced in reactions with shorter probes at higher ratios of probe to RNA (Fig. 4d,e).

Although the whole length of probe 64 is protected from S1 nuclease digestion (Fig. 4d,e), the adjacent probe, 54, is only partially protected (see Fig. 1a and Fig. 4f). Additional data not shown indicate that the discrete band observed in Figure 4f defines approximately the 5′ terminus of the 3.7-kb ISp1 transcript, which can be seen to lie at or very near the 5′ border of repeat element C. On the presumption that the protection of probe 54 shown is due to ISp1, which seems most likely but is not proven, the accuracy of the determination of the start site location of this transcript is probably about ±10 nucleotides. A more exact definition of this site by primer extension could not be achieved satisfactorily due to competitive reactions from other transcripts sharing the 5′-repetitive sequences of ISp1. Note that some smaller protected fragments deriving from other transcripts containing repeat C (F. Calzone et al., in prep.) are also formed; in some cases (e.g., Fig. 4d), these are included in the gels shown, whereas in others (Fig. 4b), they had run off by the time the desired separation of the fragments of interest had occurred.

In summary, the ISp1 transcript is a discrete 3.7-kb poly(A) RNA that is completely colinear with the genomic DNA from which it is transcribed. The experiments of Figures 3 and 4 exclude several possibilities. Among these are that the transcript has either a nonunique 5′ or 3′ terminus (unless alternative initiation and termination sites are so close together as not to be resolvable in these gels), that during embryonic development the transcript is spliced so as to form a smaller derivative including any of the ISp1 single-copy sequence, or that the transcript is both spliced within the repeat CD region and processed by removal of repeat E and/or the 3′ single-copy region. In the latter case, S1 nuclease and ExoVII nuclease digestion would have yielded a different set of protected fragments from the repeat CD region which, in fact, they did not. Only two nonexclusive alternatives remain. The 3.7-kb nonspliced ISp1 RNA could be the sole maternal poly(A) RNA product of this transcription unit or, in addition, short, nonspliced maternal transcripts about 600 nucleotides long could be formed from it, such as that observed in Figure 3a. This would imply an alternative chain termination process that is sometimes inactive, thus yielding the 3.7-kb RNA. However, though not excluded, we have no direct evidence that any of the 400 to 600 nucleotide transcripts derive from this particular transcription unit, but the 3.7-kb ISp1 RNA must derive from it.

Sequence of the ISp1 transcript and flanking genomic regions

The sequence of the ISp1 region of the genome is given in Figure 5. The approximate start of transcription (from experiments such as that shown in Fig. 4f) maps to position 2990 in the sequence shown. The transcript termi-

nates at position 6460 and is thus about 3470 nucleotides in length, not counting the poly(A) tail. Because maternal RNA poly(A) tails in *S. purpuratus* are usually 100–200 nucleotides long, the 3.7-kb length measured in denaturing gels (e.g., Fig. 3b) is close to expectation. The sequence AATAAA usually found near the terminus of poly(A) transcripts (Proudfoot and Brownlee 1976) is located 34 nucleotides before the site of polyadenylation, which was identified in an oligo(dT)-primed cDNA clone, the terminal sequence of which is identical to that shown in Figure 5.

Within the ISp1 transcript, there are no open reading frames of significant length in any of the three possible translational phases. This is illustrated graphically in Figure 6, in which the location of all the stop codons in each potential reading frame is plotted. The distances between stop codons are not significantly different than would be predicted to occur in a random sequence, nor were significant open reading frames detected in the nontranscribed strand of ISp1 or in 5′- and 3′-flanking DNA. Using the first available AUG initiation codon at position 3010 in reading frame 2, it appears that the longest polypeptide that could be encoded in the ISp1 sequence is 30 amino acid residues. Because the transcript does not give rise to spliced derivatives during embryogenesis, it must be concluded that it cannot function during early development as either a message or message precursor.

Near the likely location of the transcriptional start site (Fig. 5) is found a 5 nucleotide sequence, TATCA, which is identical to a sequence beginning 1 nucleotide from the cap site of the *S. purpuratus* CyIIIa actin transcript (Akhurst et al. 1987). This sequence is located at position 2979–2984, i.e., 6–11 nucleotides upstream of the nominal ISp1 start site but within the range that the limited accuracy of the S1 nuclease protection experiment requires. Upstream of this are located several sequence elements typical of proximal promoter regions, including TATA and CCAAT boxes, and possibly an octamer sequence, as indicated in Figure 5.

The locations of the interspersed repetitive elements (A–F) are indicated on the right of the sequence shown in Figure 5. Three pairs of 22- to 26-bp direct repeat sequences were found in the 5′-flanking sequences of ISp1. These could potentially have been involved in the frequent genomic rearrangements in this region noted above. The sequence homology within each pair of direct repeats averaged >80% (other less well matched direct repeats in the 5′-flanking sequence have not been marked in the genomic sequence). At the 3′ end, repeat F can be seen to begin about 200 nucleotides beyond the terminus of the ISp1 transcript. Though nucleotide substitutions, deletions, and additions are common at certain positions, repeat F can be seen to be constructed of a basic 25-nucleotide monomer that is tandemly reiterated.

Prevalence of the ISp1 transcript during development

The RNA gel blot shown in Figure 3b provides a suggestion that the ISp1 transcript begins to accumulate anew,

Calzone et al.

```
   1 TGTAGAATAT GAATACAAAA AAGCTCATTT TAGAGCTAGA ACAGGTTCAA CTCAGAATAC    Repeat A
                 (1)                                              (1)
  61 AAAAAAAGTC TTTTAAGAGC ATTCTGCCAT GAACGTGTGC TACAGTTAAA TAAGTTCCAA
 121 TTTTGATATG CAAACATCAT TGATAATGAT TTATCAATTT TCAGGCTTAA ATTAATCCAA
 181 GGCCACCAAG GCTATGCCTT TGATGCCCCT TTGATTGGCC TCAAATGCCC CTCTCATTGG
                          (2)              (2)
 241 CCTTGGAGAT TTTGTGTGCC CTTTCCAATT CTTCCTATTT GTGCTGTAAT ATAGAAATGT
 301 GCCCTACAGA AAAGTGGCCT TGCATGCCCT AAAAATATTG AAATTTAAGG CCTGAATTTT
 361 GAATGATTTT ACATGTAATT GTAAAATGAA CTGTACAAAA AATGATTCCA GTTATATTGC
 421 AAAAAAATAA ATTCAGTCAT AACCATGACG AAAGTGTCCA TCTTAAAATT TGATTGCTTT
 481 GTACTATTCA ACGTTTATTC TAATTTAGTT AGATTTGTTT ATGGGAGAAA TGTGTCTGAG
 541 AGAGTAATGT TCTAGATGAT AAAGACTAAT TGACAAGAAG ATAGACACGT GACTGAAAAG
 601 AGAGGTTAAC AAACGAATTA TCAGATGTGA TTAATGGAGA AAGAAAAAGA GAAATAGAGA
 661 GAGAGAGAGA GAGAGAAAGA GAGATAGTGA TCGCTAGAGA GAGAGAGAGA GATACAAAGA
 721 ATTGAGAAGA AATTGTACAT AGTGCTTGGT GNGCCCCCCC CCCCCCAACA AATGGTATAC
 781 CAATATACAT TTGATCTTAG ACAAATAAAT AAACAAAATG GATTTCACTC ACATTCTCCC
 841 CCACAGGAAG TTGAAAGAAC ACCTTTTGTC TATTTTTATA TTTTTGGACT CAAAGAGGCGC
 901 AAAACATAAT GATATAACCC ATTAATGTGC AAGAAATATAA CCCATTAATG TGCAAGGAAT
                       (3)
 961 TGAGCTTGTG AAGAAATAGA TGGTTGTTAA GAGGGTCATA CTAATTCATC AAAATTTTCC
1021 TCCATTGCAT GACTATGCGC CGAGCGCGTT CACGCCCTTC TGCTGGGAAC AAATTTTGCT
1081 AAAGTGTAGT TTTCCTTCCC GGGAACAAAA ACACTGCTTT ACCGTAAATT TGTTCCCAGC
1141 AGAAGGGGTG AACGGAAAAG TTGTTTATAT CTTTAAAAGT TTTTGCACCA TTTCATATCA
1201 TATTTAGGTT GATGATTCTT GAATAAGCTA TTTATCATC GGGTGAAATT TAGAAACATT
1261 TAAACAGTTT ACTTCAAGAA TGTCACCTGC AATGAGGAGC ATATTCTTCA AACGCGATTT
1321 TCTCGAACNN NNNTTTTCGT GCGACATGCG CTTACGCCCT CG

                         CCGGTCCC GGCACACGAC    Repeat B
1381 GGATATCTAT TTACCAAATA TACAATAGTA GGCGGTTTCC CCGTCAAAAT GTTACGTCAT
1441 GTAATATTCA AACATATTCT CTTCACATGA TTGGTAATAA GACGTTCCAT TCATTTTTTG
1501 AAATTGGGCG TTTCCAATAC TGGTCTTTAC CTTTAACCAA ATTGATCAAA GTGTCAAATC
1561 AAAAAAACC TTGGTAGACT AAGCTATCCT GGATATCGGG GGAAACGAGT TCCTCTTTTA
1621 CAAGTCAGGG AATTTCAGAA AAAGAGCTAA AATTAGGGAA AATCAGAGAA TTTCATTCAC
1681 CCCTAAAAGA AAGTTAATAA GTTGCTTATA TGTTTGTCTC TTGTTCAAAC AGCATCATTT
1741 GGGTGATAA

             G CTATGGTGGG TTCCCAAATG TTTACATTGT TTCCATTAAT GATCCTGTAT    Single Copy
                                                                          DNA
1801 TTTGCAACTC CATCAGCGGC AAATTCCAGG AAAACCCTGG AAAAATCAGA GTTTTTAATA
1861 AATGAAATGC GTGGAAACCA GGGCCCCATC TTACAAAAG TTGATATCAA TCGTAACTAT
1921 GTACATAAGA AATGGGAGAC TGCAAACTCG CGATTCATTG CAGGGCTTGC GATTGCTTTG
1981 GATCAATCGT AACTCTTTGT AAGACGGGGC TCGTACTGAT TACAATGGTA CTTAGATTTG
2041 AATCTGTTTT CCTTTAAATT TTTAGAATCT GTATAGTGTA CCTGAGGACG TCTCCAGAGC
2101 ACTGCTACGA TCGGATTCAG AAGAGACACA GACGAGAAGA GACGGGAAAT CTCGATAGAT
2161 TACCACACTA AGCTGACACG AGCTATGCGA AGATTGGCTC ATCCATGGAA ATAGGTTGGT
2221 CTCTTCCTGC GCCTGTTGTG GTAAGACCCT TTTAGAGTAC TCGGTAGTCA AATTCCATCT
2281 TTTTAGCCGT AAAAGACTTT GTTGTTTTTG TGGTTCCATA GTGTAATACT CTTGGCTGCA
2341 GAGTGACATT GATTTATGAG TTGTGGATAG AAGTAAAGTA TAAGGAGTAAG CCCATGATGT
2401 TTTATATCAA TATACCTCAA AATGATTCAA AATTCTTGCT ATTAAATGAA TTATGTATGT
2461 AGTATACCAT TAACAACTAA ACCATTCATA TCCTTACCAG GCTTTTTAAA GTTCTACGAA
2521 ATGTGAAAAA TTGATTTGTC CTAATTTATG TTTCTTGGAG AATGTGAACA TGAAAAGCTT
2581 ATGCTAAAAG TATTAAAACAG AAGGCCTTAA TTCAAACAGT CCACCTTTCA TTGCTACCTT
2641 ACGAAAAACA GAAAAAACAT CAAATTTACA AGGGGTATAA ATAATTTTAG TTTATTATCT
2701 GTACATTTGA ATATTTTTGT GCATATGTGA AAAAATATGT ACCAATGAGT TTTCTACGAA
          (4)                                   (5)
2761 TTTCATCTTG GATAAGGAGT GAGCAACCAG AATGACGAAG TTGTCCATCA GATTCACTGT
2821 AAATTTCTCT TCAAACTGGT ATCAACGCATT CAAAAACTAA GTCTAGAGTT CTCTAGTTAT
2881 TCCTCATAAT TTGGAGAGCT GAAGCAAGGT GCTACCTACA TGTACATGTA TATGTGCACT
          (6)
2941 GCATTGTCAT GATTTAATTG AAGGAAAATC CACTCTTTGAT ATCAGATTG
          (6)                              (7)

                                 C TTTGAAAGAA    Repeat C
3001 ATCAGAACAA ATACAGAAAG CAGATCGGTG AAAGTTTGCA AGAAAACGGC GCGACAATTA
3061 GAAGTTATGA ATTTTGAAAA GTTATCCGGT AATCATATGT ATAGGAATTC CAAACTGGCT
3121 AAATTGATGT TTTTTTGTTT ACTGTCCCAT TTGTTCCATA ATGCGAAATA AAATGCGACT
3181 TTTCAAAAGG TCAAGAGGTC TTTTCTTTTA AGAGGACCTA TTATTTAAAA AATGAGATCC
3241 GCTATCCATA ATATATTTCA GTAGTACCAC AAGGGAAAAT GCCATTTGGA AAAAACTGAA
3301 ATAAATGATA TTGTTTTGAC ATGACAAACA AGAGAGATGT CTGTGAGACT ATGCCACAAT
3361 CCACTTACCT TATTGCCAGT TTGATATTTG CTTATACACC CTGTATAGTG TAGGTAGTGA
3421 ATAGTGATTT CATATTTGAA ACATTAAAAG ACGTTAAAG ATGGGAGTTG TTGCTACCTT
3481 GTTTGGGGTT CAGAAATAAG GGATAGAGCA ACGTTGATCT GGCGCTGCTC AGTGGCTGCC
3541 GGGCTCACGA TCAATTGGGC AAAAAGAATT TCCGTTTCGG TTTAATAAAA TTT

                                        GGAGTGG    Repeat D
3601 ACTATTAAGC TATTATTAATT ATTAGTATTA TTATTCAATAA CTTTCTTATT TTTGGTCCAC
3661 TTTATTTCAA ACTTTTGCCG TTCTGATTTT CTTATTTTTC TGCTTTTTGT ACAAAAAACC
3721 TCAAAAATGA GGGTTGGACT GCTCTTTATG TTGTGTAGTG TAGACATTTA TCCTATGTTT
3781 TAGTTGTAAC TGTAATTTTG TTTATAAATA TGCCATAAAT CTGATTTTAT GTTTTTATTCT
3841 CTACAGGTTA TTGATTCCAA CAATTCAGTG GAACAAATGC TGCAATATTA CAAGGATAAC
3901 AAGGAACATA TTTCCTTAC AGGACAAGAT AGGTCATTGT ACAGAGATAA TAAAACTCAC
3961 ATAATTGGAC AAAAGCTACA TGTATCATGG AGCATCTTCC TAGCAGGAAA CACAAACACA
4021 TCAGCGACCA AAGTTTAGAC T

                         TCGACAAGG ATTCATTCTG TAGGTATGCA TTCATCAGGA    Repeat E
4081 GCATTTATTG ACTGTTTAAT AACTGTATAT GGTGGGATTT TTGTACAGT GTGGGAGTTT
4141 CAGTTGTCCA GATTGGCAAA TTGTCACAGT GTGGAATGAT TGAATCGCCA AACTAGAATC
4201 AAAACCTTTA AATTTTATGA TTGCCACATA TCACAATAGC AACTACTGTA GAAAAATCAT
4261 CCTCGGGCAA TTTCAGTAGC CTAGTGCGTC ACCTACCTAC TGACAATGAG GTATTACGTA
4321 ACCTGAACAT GATTTCTATG GGACATACAA TTGGGTATTG AAATTGCCCT CAAGTTGCAT
4381 TTTCAAAAAA ATAGTTCGGA TCGTGTAACA TGCTTCTTGC CTTGAAAAAT GCTGTCAGTGT
4441 CATGAAAAGG CGAAGAGTTC CTGGGTTTAT AGTTCGTTCG TTCGTTATCAG TTTGCTTCAG
4501 ATTAATGTAA TCTTGATCTT CGGATTGAGTT GCTACAACCG GGTATGTCAG ACACGGTTCC
4561 CCGGGGTGCA AACCAAGAGT CCTGGGTTTA TAAGGAGAGA GGATATGCAG TGGACCGCAG
4621 TGGACAGTAG ACAGGATCAG AATTTGTGCC AACCTCTCGA TGAAGCTATT GCTGCAAACT
4681 CTTTGAGTGT AGCATTCCAA CTTTGAAACA GGGGGGCATT TCACAAAACT TATCATCAGT
4741 GACAAATGAC AGTTTTTTTT ATAAGCTACT GGAGTCCTTG CTTATGATTG GTTATCAGCA
```

```
4801 GATTTGTTTC TGATTTTTGT CATTTTTCAT TGAAAAAAAG GCTTTGTGAA ACAGGTCCCT
4861 GGTTTACTAA AGGAGAGTTT AAAAATCCAT CCGGATCAAG TGGATTGAAA TCCTTCTCTT
4921 AATAGGGGTC TCTTGTTATT CACTATGCAC CTACATGTCAT CTTTGATCCA GTGATCTCGA
4981 GCAGCTGACA CAAGGAAGTT GTTTGAAGAA TCCATTGCCT TTAGAAAATC CACAGGAGGT
5041 GACGAAAGGG GATGG

                   AAGCT TGCTACCAAG TTTCAAGGTT TCAAGAATCC AAAACTACAT    Single Copy
                                                                          DNA
5101 GTTAGGATTA CTCCCTACCT ATGAAAGTGT TTTGTAGCAA GTTCACATTT GCTGACATAC
5161 AGATAAAATA TCTTGAAATT CTCTTACCTA TTCCAAGAGG ACAATTTCTT TTTCTTTTGT
5221 AAACCCTTTT AGACTATTTC TAATCATTAT GTCACACACA TTTTCATTGT GTTTCATTCC
5281 GTGGCTATAC ACTGGCGTTG GATATTCCAT AAATCTAGAG TAGATTTACC TGTTCCCCCA
5341 AATTAGACCT TATCATTATG ATAAACAACA AGAGAAATGC AGAAGGGGAA AATTTATATT
5401 TCTTCTTCTT CAGATTCTTT TTTAAACAGT TGATCAGAAG GGTTGGCAAA AAAAAAATTA
5461 TTGGAGCACC CCTCTTTAAA ACCACACCCT GTCATGTCTT TTGTTTACCA ATGATTTGAC
5521 AATCTTGGAT TAGTTGGATA TTGAACCATT TATTATAATC AACATGTACA TGTATGTAAA
5581 TATATTTAAA AGTATTGAAC AGAAATGCAG ATCTTCCAGA GTTTTTCATC TTAAATTCTC
5641 ATTTTGCCAT ATAATAGTAA CTCAGAAGAT GTATCATCGT GTGAATATGCG TTTGATATGG
5701 GTTTAGTTTT TTTAATCTGT TACAAATAAT TTTGTGTGAT GTTCTTAAAA ATGAATTGAT
5761 GTACGGGTGT ATTGTACAAT ATATAATTTA TTTTGGAAAA CAGAGTGAGA GTATTCAAGA
5821 GAGAGAAAAA AAAGATCAAA ATCAACAGTA AAGACAATCA TCATAAATAT GAAATTAAAT
5881 ATTGTGGATG CATGTACGTA TACTTATCTT ATTATATTGA TTTATATCAT TGTTATCTTA
5941 CGGGCTATAC AGGTGAAGCA CAACGATAGT CATGCACCAT TGGTTATCTC ATGTCTTACC
6001 ATGGATAGAT TATCCCACTT AGAATTGAAA GTAGAAGAAG GTATTTGGAC TAAGAATGGT
6061 TTCTTCATC AAGACAGAAT GGAAACCCAT TTGTGAAGTT TTTATTATTT TTATATGGGA
6121 TATTTACTTG AGTCCAATCA AAAGGTGAAA CAAGCATTAT CTCTTTACTA TAATAATGAC
6181 ATCTACTTGA AGTTCATTTA TGTCAAATGA TTACAAAAGT CGGTTGCACA AGGGCTGAAAG
6241 TTAAATTGGT TTAAAGATGC TTTTAAACCA CATGTATTTG TTTCCTCAAG ATTTCCCTGT
6301 TAATTTAAAG GTTGTCTTGT GGTAGTATAT GTGTAATGTC ACAGATGTAT TCTGTTCTTA
6361 GGTATGCGTA TATTGTGTCA TCACAATAGT GTTGTCAATA CTATCTTTTT TTAGTTGTGA
                                                                          (9)
6421 ATAAATTATT GATAATTACG ATCCAGAAAT GTGGTTGATT GGT
                                                   (10)

                                              TTTTTGG TTTAAATGTG    Single Copy
                                                                    DNA
6481 TGGCAAGTTA ATATTTATGGA GAGATAATGG ATAAAGGAAT AGTCTGCAAG CACTGGTACA
6541 GTACAGTTGA CATGGTCAAG ATGTTTTGTG TCTACTCTGC TGGGGAAATC AAGGGTGTGG
6601 GTCCAAGTCC TTGGTTGAAG GCACTGATGT TTTGATGCAA GGCATTTTAA TGTTATCTGC
6661 ATGGCTTTTT CTCACCCAGG TAACGTTGA ATAACTGGTA GGATGTAAAT GTTATGATGG
6721 GTGGTTTATC ACTATGTGTA CATGTACTTG TAAAATGGTG GCAAACAAGG GCATGTAGGG
6781 C

             AAGTTGCGG TTGTGAAATG ATGACTGGTT GAGGAGTGTA AAATGATGAC TAGGTGAGCA    Repeat F
                      *            *                   *
6841 GTGTAAAATG ATGGCTGGGT GAGGGTGTAA AATGACTGGT TGAGGAGTGT AAAATGATGA
            *              *
6901 CTGGCTGGGG GTGTAAAATG ATGACTGGCT GAGGGGTGTA AAATGACTGG TTGAGGAGGT
            *              *
6961 GTAAAATGAT GACTGGCTGA GGGTGTAAA ATGATGACTG GTTGAGGAGG TGTAAAATGA
            *              *
7021 TGGCTGGGCG AGGGTGTAAA ATGACTGGTT GAGGAGTGTA AATGATGACT GGCTGAGGGG
            *              *
7081 TGTAAAATGA TGACTGGTTG AGGAGTGGTA AAATGATGAC TGGGTGAGGG GTGTAAAATG
            *              *
7141 ACTGGTTGAG GAGTGTAAAA TGATGACTGG CTGAGGGGTG TAAAATGATGA CTGGTTGAGG
                          *                                 *
7201 AGTGTAAAAT GATGACTGGC TGAGGGGTGT AAAATGATGA CTGGCTGAGG GGTGTAAAATG
                          *                                 *
7261 ATGACTGGTT GAGGAGTGTA AAATGATGAC TGGCTGAGGG GTGTCCGGGT GAGGAGTGTA
                          *                                 *
7321 AAATGATGAC TGGCTGGGAA GTGTAAAATG ATGATCGGGT GAGGGGTGTA AAATGATGAC
                          *                                 *
7381 TGGGTGAGTA GTGTAAAATG ATGATCGGGT GAGGGGTGTA AAATGATGAC CGGCTGCTGG
                          *                                 *
7441 AACTGTCAAA ATGACCGACT GAAGAGTGTA AAATAATACT GGTGAGGGTG TAAAATGATG
                          *                                 *
7501 ACTGCTAGGG TTTAAAATAT GTGACTGTTA GATGTAAAAT GATGCTGCT GGGAGTGTAA
                          *                                 *
7561 AATGATGACT GGGTGAGTAG TGTAAAATGA TGACTGGTTG AGGAGTGTAA AATGATGACT
                          *                                 *
7621 GGCTGGGGAG TGTAAAATGA TGACTGGGTG AGGGGTGTAA AATGTTGGCC TGGTTGAGGA
                          *                                 *
7681 GTGTAAAATG ATGACCGACC GAGGTGTGTA AAGTGATGAC TGGCTTGT
                          *

                                              AA GCGTTGCATA    Single Copy
                                                               DNA
7741 GAAGCTGGCT TTCATTGACT TGAAATAAGC AAAATATCAA GTGCTTTATT TCTTCTGTCT
7801 TTTTTCACTT GTGATTCCTT CAGCCAAGCC TAGGTTTGAA ACTGAAAAAA TGTATGCCAT
7861 AAAGGAGCA GAGCTTTTGT TTGTGAAGCA GGGGAGGCG ACCCGGAAGG GGGGGGGGGC
7921 ACAATTCCAC CACTTTTTTC AAGTACTCCT TTAGGCTCTA GTGCCTCCAC TTTACATAGA
7981 AAAAGTGCCC CTTATTATAC CAGTCAATCA TTCTTAATGA TGAAGAGCAC AATGTACTAC
8041 ATTGGAAATG AATTGATATA TATATATATA TATTTTTTTT TTTTGGGGG GGGGGGATTT
8101 TGCTCTCTCT GCTCATGTAT ATTTGATCCG TTTTTTGGGG GGGGGGGGG GAATCGGGGT
8161 TGTCGGGGAA GGNNGAGNCG GAGAGAGAAG AAGATAAAGT AAAATCAGGT AACCTTATTT
8221 GTCCGTTATA AAAAACTCAT GATAACAGCA TATTTTGTGC TGCTTAAGTC CTATTTTCAT
8281 ATTCTATTTT TAACACAGAT GTACAAAGTT TCAATATTAG AGGTAAAATTC CTTGGTCTGA
8341 AATATGTTTG TTTTTACGGG GTTCCACATG TATTTGTAA CTTATACACT GCATTGAACT
8401 AGGAAATAAA TAAAGCNNNT TTTATCCCAT CTAAAATCCA ATATTTATTA TTAAGCAATA
8461 ATTGTTAAAG CATATTGACA TTACATTTAT ATTACTCTAC AACTCATTAC CAATTACGTC
8521 ACTTCAGAAA GCTCTATTTA AATTGAATTG AATATAAATA CATATTCAGG GTATAGTGAG
8581 CAGCACACT GCAGTCAGGC CTCACGATAC AGTGGATGGA AGAGGCAAAG GACCGACTTA
8641 GATTCAAATT ACCGAAAAAA TATCTCGTTA CCGCGATAAC TTGGAAAGTT TTGAAGCAAT
8701 TCAAACCAA CTTGGATGAA GTATAGGATA GTTATACACG GATGAACTGA TTAAAATGTA
8761 TGATTCAGAT ATAAACAAGG GGAAAAGACA GGATATACCT GACAATATTG CCAAATAACT
8821 ATCACAATCT CAATGCTCAT GCAAGCGTTA ATCAAGATTT TTTTAACAGC TTTAAAAGCT
8881 T
```

Figure 5. The genomic sequence of ISp1 allele B1 and its flanking DNA. Locations of the repetitive sequence elements referred to in the text are indicated on the right. Specific sequence elements noted in the text are underlined and marked by a number under the first nucleotide of each element. (1–3) Direct repeats (22–26 nucleotides). (4) An 8-nucleotide sequence homologous at 7/8 positions, with the octamer sequence detected in the testis H2B gene of *Psammechinus miliaris* (Barberis et al. 1987). (5) An 8-nucleotide sequence homologous at 7/8 positions with the 'CCAAT' element present in the testis H2B gene of *P. miliaris*. (6) 'TATA' homologies. (7) A 5-nucleotide sequence homologous with the 5-nucleotide 5'-cap sequence of the *S. purpuratus* CyIIIa actin gene primary transcript (Akhurst et al. 1987). (8) The approximate start of the ISp1 transcript. (9) A consensus poly(A) termination signal (Proudfoot and Brownlee 1976). (10) The 3' terminus of the ISp1 transcript. An asterisk (*) marks the first nucleotide of the imperfect 25-bp sequence tandemly repeated 40 times in repeat F.
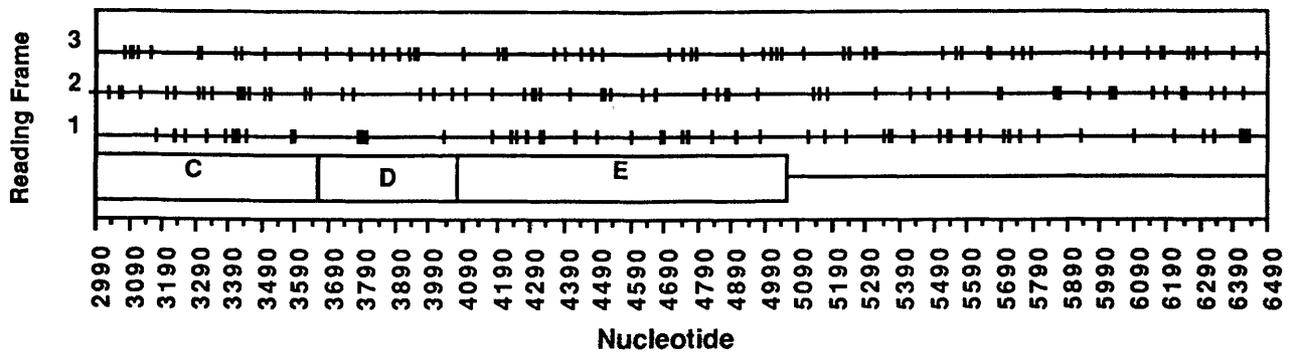
Figure 6. Location of translational stop codons in the ISp1 sequence. All three possible reading frames are shown, and the sequence organization of the transcription unit (see Figs. 1 and 2) is plotted for reference.

i.e., as a zygotic transcript, late in development. To obtain quantitative measurements of its prevalence at various stages, we carried out single-strand probe excess titrations, as described by Lee et al. (1986). These measurements were performed with a labeled antisense RNA that contained the single-copy sequence of probe 2 (cf. Fig. 1a). Titrations of total RNA isolated from the unfertilized egg and at various stages of development are shown in Figure 7a, and the levels of the ISp1 transcript calculated from these data are listed in Table 1. The unfertilized egg contains about 400 transcripts of ISp1. The level of this transcript increases to about 900 copies in the pluteus, i.e., somewhat less than one transcript per cell, on the average. The increase in ISp1 content, which is greater than twofold, clearly indicates synthesis, though at a low rate that could not be directly detected by metabolic labeling in vivo (F. Calzone, unpubl.). The prevalence of this transcript is typical of that of the rare or complex class of transcripts in the sea urchin egg and embryo (reviewed by Davidson 1986).

The concentration of the ISp1 transcript in the nuclear and cytoplasmic fractions of total RNA was also determined, as illustrated in Figure 7b. Table 1 shows that the large majority of the ISp1 transcript in the early gastrula is located in the cytoplasm. A similar experiment carried out on egg homogenates from which the pronucleus had been quantitatively removed by centrifugation confirms that >80% of ISp1 RNA is cytoplasmic before fertilization as well, as is the bulk of the interspersed maternal poly(A) RNA.

Figure 7. Measurements of ISp1 prevalence by single-strand probe excess titration. A ^{32}P-labeled RNA antisense probe (10HB) containing the single-copy sequences of probe 2 (Figs. 2–4) was utilized to measure the concentration of the ISp1 3.7-kb transcript in RNAs from the unfertilized egg and embryo, as described in Materials and methods. These data were used for calculation of the transcript prevalences listed in Table 1. (a) Measurements of ISp1 transcript concentrations in total RNAs. The total RNAs represented by each symbol are unfertilized egg (●), 7-hr cleavage stage embryo (○), mesenchyme blastula (×), 35-hr gastrula (△), and pluteus (■). (b) Measurements of ISp1 transcript concentration in 44-hr gastrula nuclear and cytoplasmic RNAs. (●) Gastrula nuclear RNA hybridization values; (○) cytoplasmic RNA.

Discussion

The ISp1 transcript is not an unprocessed pre-mRNA

Long interspersed transcripts constitute a major fraction of the mass of poly(A) RNA stored in both echinoderm and amphibian eggs, and since the discovery of this class of maternal RNA several years ago, its possible functional significance has been a matter of speculation (reviewed by Davidson 1986). The lack of direct transla-
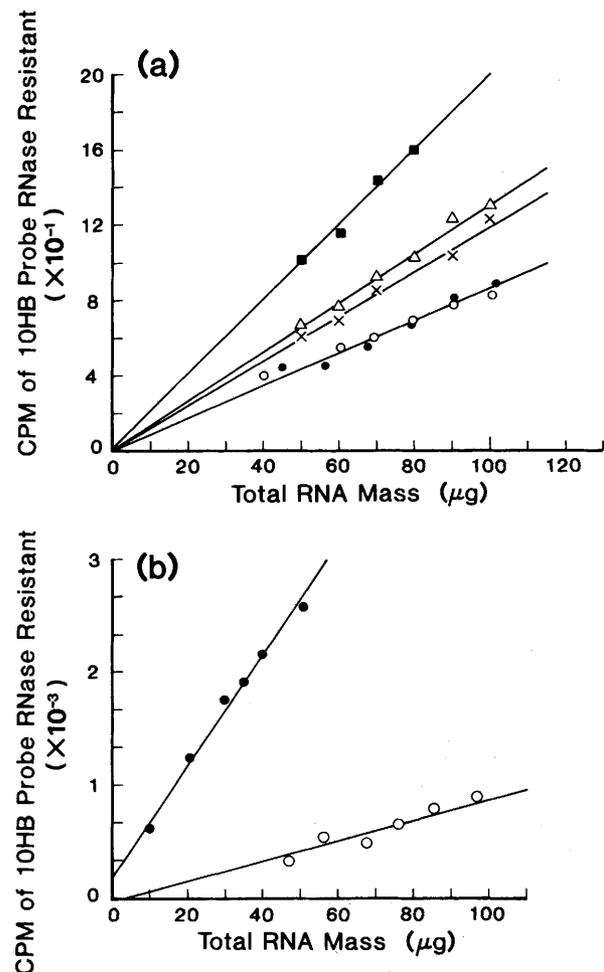
Table 1. *Titration measurements of ISp1 transcript prevalence*

RNA source[a]	Probe hybridized per µg RNA[b] (cpm/µg)	ISp1 transcripts per embryo[c]	
		(pg/embryo)	(molecules/ embryo)
Unfertilized egg	8.67	8.0×10^{-4}	400
Embryo (7 hr)	8.57	7.9×10^{-4}	390
Blastula	12.0	1.1×10^{-3}	540
Gastrula (35 hr)	13.2	1.2×10^{-3}	610
Gastrula (44 hr) (cytoplasm)	11.1	1.60×10^{-3}	833 (91%)[d]
Gastrula (44 hr) (nucleus)	48.4	1.70×10^{-4}	81 (9%)[d]
Pluteus	20.1	1.9×10^{-3}	910

[a] Except where indicated, RNA was isolated from whole unfertilized egg or embryo.

[b] These values are the slopes extracted directly from the titration data shown in Fig. 7.

[c] The mass (M) of ISp1 transcript was calculated as M = (C)(2800)/(S)(F), where (C) is the slope of the titration curve, (2800) is the mass of RNA in the egg or embryo in pg (Goustin and Wilt 1982), (S) is the specific activity of the probe, and (F) is the fraction of the length of the ISp1 transcript represented by the genomic sequence in the probe. Probe-specific activities ranged from 1.15×10^8 to 2.06×10^8 cpm/µg. The efficiency of isolation of total, nuclear, and cytoplasmic RNA was reproducibly found to be 50–60% (Lee 1986; Lee et al. 1986). In the experiment for gastrula 44-hr RNA, 7.4% of total embryo RNA was present in the nuclear fraction. To calculate the number of molecules of ISp1 nuclear transcripts, a value of 207 pg was used for the total amount of nuclear RNA per embryo (2800 × 0.074). The value taken for total cytoplasmic RNA was thus 2600 pg.

[d] The values in parentheses are the percentages of total ISp1 transcripts in the cytoplasmic and nuclear functions, respectively.

tional competence of these RNAs (Richter et al. 1984), their mean molecular length, which much exceeds that of bona fide maternal message (Posakony et al. 1983), and the inclusion of repetitive sequences in internal locations are all characteristics shared with the primary nuclear transcripts of many genes that code for proteins. Pre-mRNAs are also typically polyadenylated, and in organisms whose genomes display short period repeat sequence interspersion, repetitive sequence elements occur frequently within introns (see, e.g., Cochet et al. 1979; Wahli et al. 1981; Monson et al. 1982; Schibler et al. 1982; Tsukada et al. 1982; Maroteaux et al. 1983; Ryffel et al. 1983; Suske et al. 1983). Thus, one possibility is that the interspersed maternal transcripts are unprocessed oocyte pre-mRNAs that have found their way into the cytoplasm and have been stabilized there, as are stored maternal mRNAs. Two different ancillary interpretations that have been suggested are that such nuclear pre-mRNAs may simply have 'leaked' into the oocyte cytoplasm, in which case they may well be expected to have no function after fertilization or, on the other hand, that they are specifically transported to the cytoplasm during oogenesis and in the course of

cleavage distributed to the embryonic cells, in which they are destined to be productively processed. Maternal interspersed RNAs are thus envisioned as a special, post-transcriptionally regulated potential source of embryonic messages. Possibly suggestive of this is an observation of Ruzdijic and Pederson (1987), who reported that a small fraction (≤1%) of interspersed maternal RNAs bearing a region of repeat E can be immunoprecipitated with anti-U1 snRNP antibody. None of these propositions can convincingly apply to the ISp1 RNA, however, as it is extremely unlikely that this is, or would ever serve as, a message precursor. Thus, the ISp1 transcript lacks any significant open reading frames beyond what would be expected on the basis of chance. Nor are there convincing donor and acceptor canonical splice sites within the ISp1 sequence that would yield a significant open reading frame. Furthermore, in Figures 3 and 4, we show that during embryogenesis no spliced derivatives are in fact formed from the ISp1 RNA that include either the single-copy region or the 5' repetitive portion of the transcript. The demonstration that this particular interspersed maternal poly(A) RNA cannot be interpreted as a pre-mRNA certainly does not preclude the possibility that other unprocessed message precursors are stored in the egg cytoplasm. However, although deliberate searches for unspliced precursors of a number of specific cloned mRNAs have been carried out in amphibian and echinoderm egg cytoplasm (reviewed by Davidson 1986; Rosenthal and Ruderman 1987), no positive indications have been uncovered. The minimum inescapable conclusion from the present work is that long maternal poly(A) RNAs bearing interspersed repeat sequences do not necessarily indicate stored unprocessed pre-mRNAs, and the most likely conclusion is that maternal RNA does not, in fact, include a significant complement of primary transcripts of this nature.

The ISp1 transcript is not a readthrough product initiated at an oogenetic mRNA promoter

The germinal vesicles of urodele oocytes contain long transcripts initiated at the histone gene promoters and continuing on into the adjacent 3'-flanking sequences (Diaz et al. 1981; Gall et al. 1981; Stephenson et al. 1981; Diaz and Gall 1985). These readthrough transcripts are confined to the oocyte nucleus, and they are evidently the consequence of failure or inefficiency of a sequence-specific 3'-processing mechanism known to be required for production of mature histone mRNAs in both amphibian and echinoderm systems (Krieg and Melton 1984; Birnstiel et al. 1985). A reasonable generalization might be that interspersed maternal poly(A) RNAs are also readthrough products, deriving from a variety of other promoters utilized for maternal mRNA synthesis during oogenesis. They would then naturally include downstream repetitive sequences, and, because they eventually encounter a randomly occurring polyadenylation site, they might be exported from the germinal vesicle and stored in the cytoplasm (Wickens and Gurdon 1983). However, the absence of significant open

reading frames at the 5' end of the ISp1 transcript also obviates this explanation. Furthermore, as shown in Figures 3b and 7 and Table 1, exactly the same transcript is generated in somatic cells of the late embryo, in which it is also largely transported to the cytoplasm. The present analysis thus offers no support for the hypothesis that interspersed maternal transcripts are special oocyte readthrough products initiated at the same sequence locations as are bona fide maternal mRNAs.

Other possibilities

The promoter of the ISp1 transcript is active both during oogenesis and in the embryo and could be functional in other cell types as well. The transcript begins at a precise site within the sequence, or at least within a range of a few nucleotides. The transcript also terminates precisely, the usual distance beyond a canonical polyadenylation signal. In all respects for which the data permit comparison, the ISp1 transcript is typical: Other maternal interspersed transcripts identified with single-copy probes by Posakony et al. (1983) are also of discrete length; the low prevalence (cf. Costantini et al. 1978, 1980) and the nontranslatability here confirmed by the primary sequence of Figure 5 are both characteristic of this class of maternal RNA. ISp1 is representative in another sense as well, as it is a member of the large family of interspersed maternal RNAs that share repeat C (Fig. 3a,c). Other data reported separately (F. Calzone et al., in prep.) show that repeat C is usually or always located at the 5' end of the set of large transcripts that constitute this family which, as a whole, accounts for as much as 2% of the total interspersed maternal poly(A) RNA. If such transcripts are not pre-mRNAs or readthrough products of otherwise unexceptional structural genes active in the oocyte, what are they?

An interesting though rather agnostic argument is that interspersed maternal RNAs are likely to have no functional significance, in that they derive from promoters, the transcripts of which are not functionally utilized at any stage of life. Such promoters might have been carried about during evolution in association with transposable repeat sequences, perhaps such as repeat C, inserting randomly in the genome and generating transcripts which, in the peculiar milieu of the egg, accumulate in stable form as do 'useful' RNAs. Despite their aggregate mass, which is anything but inconsiderable, such transcript species might be tolerated by virtue of their low individual sequence concentrations. In this connection, however, it is useful to recall that the typical prevalence of interspersed maternal poly(A) RNA species is of the same magnitude as that of most bona fide mRNA species in the sea urchin embryo or, for that matter, in most animal cells, i.e., one or a few molecules per cell or (for the egg) per cell equivalent (reviewed in Davidson 1986). It is not easy to demonstrate that something has no function. However, it is scarcely inconceivable that there might remain functional roles for long heterogeneous RNAs, of which molecular biology is yet entirely unaware. With these alternatives having been

noted, there are two more precise possibilities to consider.

The sharp developmental regulation of the 400–600 nucleotide small poly(A) RNAs observed in Figure 3a may provide an interesting access to the mystery with which we are concerned here. The regulated change in length from ~600 to ~400 nucleotides and the early and quantitative disappearance of these prevalent small transcripts is most unusual and clearly implies some function (though not mRNA function) either in oogenesis or pregastrular development. Perhaps the real functional role of transcription units such as that also producing ISp1 RNA is to generate these small transcripts, the accumulation of which would depend on a chain termination processing step that operates only during oogenesis. The control of the small RNA synthesis would thus depend on an oocyte-specific post-transcriptional mechanism, whereas the long interspersed RNAs would represent the default product that continues to be synthesized, e.g., in late embryos, in the absence of this mechanism. A related possibility is that in the growing oocyte, relevant examination of which has not yet been carried out, splicing yielding a translatable product does indeed occur. Thus, species of RNA containing regions of the long transcripts spliced to coding sequences located far downstream could be formed but have disappeared by the end of oogenesis when our observations begin. The necessary post-translational mechanism would again exist only in the oocyte, and the interspersed RNAs detected in the mature egg would represent the unspliced default moieties, as above. In summary, we raise the possibility that understanding interspersed maternal poly(A) RNAs might require refocusing on other transcription products of the same promoters, and on post-transcriptional mechanisms specific to the growing oocyte.

Materials and methods

Eggs and embryos

Gametes of S. purpuratus were collected by intracoelomic injection of 0.5 M KCl. Before fertilization, eggs were washed several times with fresh Millipore-filtered sea water. Embryos were cultured at a concentration of 1×10^4/ml in Millipore-filtered sea water, containing 40 µg/ml gentamicin at 15°C, with constant aeration and stirring, as described by Smith et al. (1974), Hough-Evans et al. (1977), and Leahy (1986). Embryos at 7 hr contained 32–64 cells. The embryos hatched approximately 18.5 hr after fertilization. Blastula-stage embryos were collected at 20 hr. Gastrula RNA was isolated from embryos harvested at 36 hr, and plutei were collected at 65 hr.

Isolation of RNA and DNA

Total RNA was isolated from unfertilized eggs and embryos, as described by Posakony et al. (1983) and Lee et al. (1986). Methods for preparation of nuclear and cytoplasmic RNA fractions from gastrula-stage embryos are described by Lee (1986). Poly(A) RNA was separated from total RNA, as described by Posakony et al. (1983). High-molecular-weight genomic DNA was prepared from the sperm of individual S. purpuratus, as described by Lee et al. (1984).

Calzone et al.

Isolation of genomic and cDNA clones

Genomic clones of *S. purpuratus* individuals B and F were obtained from a *Mbo*I partial digest of each individual's sperm DNA cloned separately into λEMBL3 (Firschauf et al. 1983). The genomic inserts used to prepare each library ranged from 15 to 20 kb in size and were isolated by sucrose gradient centrifugation of *Mbo*I-digested DNA, as described by Maniatis et al. (1978). The *Mbo*I genomic fragments were ligated into the *Bam*HI sites of λEMBL3 after treatment of the vector with *Eco*RI and *Bam*HI and removal of the short *Eco*RI–*Bam*HI linkers by ammonium acetate–isopropanol precipitation, as suggested by Firschauf et al. (1983). All other steps in the construction of the libraries employed standard techniques. Some genomic clones were isolated from an *Eco*RI partial library of individual F sperm DNA, constructed by Howard T. Jacobs in this laboratory. The single-copy DNA probe used to identify ISp1 clones in each genomic DNA library was a 550-bp *Hin*dIII–*Bam*HI fragment of the maternal cDNA clone 2109A-10 (Posakony et al. 1983). The probe was labeled by the end-fill reaction with [^{32}P]dNTPs. The hybridization conditions employed are described below. A 3' cDNA clone of ISp1 was isolated from a pluteus-stage λgt11 cDNA library constructed in this laboratory (Sucov et al. 1987) using probe 32, which is described in Results (Figs. 2–4).

DNA and RNA gel blot hybridization reactions

Detailed protocols for DNA and RNA gel blot hybridization reactions have been published previously (Posakony et al. 1983; Lee et al. 1984; Shott et al. 1984). GeneScreen Plus (Dupont/NEN) was substituted for nitrocellulose for RNA gel blots using the transfer protocol recommended by the manufacturer. The basic buffer for all blot hybridization reactions was: 5× SET (0.15 M NaCl, 0.01 M Tris–HCl at pH 7.8, 0.002 M EDTA); 5× Denhardt's solution [0.02% (wt/vol) each of Ficoll (Sigma) polyvinylpyrrolidone and bovine serum albumin (Pentax), as described by Denhardt (1966)]; 0.2% (wt/vol) SDS, 0.05 M phosphate buffer (pH 6.8), 25 µg/ml poly(A), 25 µg/ml poly(C), and 12.5 µg/ml denatured calf thymus DNA. The concentration of SDS for blots with GeneScreen Plus membranes was increased to 1% (wt/vol). Blots were reacted with 1×10^5 to 5×10^5 dpm/ml of denatured probe overnight at 68°C for DNA hybrids and 71°C for RNA/DNA duplexes. When formamide at 50% (wt/vol) was included in the hybridization buffer for RNA gel blots, the temperature was lowered to 45°C. After incubation, the blots were washed 1× for 30 min with 5× SSC (0.15 M NaCl, 0.015 M sodium citrate); 2× for 30 min with 2× SSC, 0.1% SDS; and 2× with 0.2× SSC, 0.1% SDS at 68°C or 71°C. To prepare ^{32}P-labeled RNA probes, genomic DNA fragments were inserted into the Riboprobe vectors (Promega Biotec), pSP64, pSP65, pGem1, or pGem2. After linearization of each clone with an appropriate restriction enzyme, high specific activity probes were generated using the protocol recommended by the manufacturer. Most DNA probes were prepared with M13 subclones generated for DNA sequencing (see below) and the 15-mer sequencing primer (New England Biolabs) to prime the extension reaction. The labeled inserts were released from M13 vector with appropriate restriction enzymes. In a few cases, DNA fragments were labeled by the end-fill reaction after release from pUC9 (Vieira and Messing 1982) subclones with appropriate enzymes. All labeled DNA inserts were purified by gel electrophoresis. DNA and RNA probe specific activities were usually in the range of 5×10^8 to 10×10^8 dpm/µg.

Nuclease S1 and Exo VII mapping

Single-stranded DNA probes labeled internally with ^{32}P were employed for all nuclease mapping experiments. The probes were synthesized using the M13 subclones, pentadecamer sequencing primer, and Klenow fragment *Escherichia coli* DNA polymerase, as described by Calzone et al. (1987). One of the four nucleotides in the synthesis reaction was [^{32}P]dATP (400 Ci/mmole/Amersham). For restriction mapping experiments with probe 11A the amount of [^{32}P]dATP in the synthesis was reduced so as to preferentially label the 5' region of the probe. The methods for hybridization reactions, nuclease S1 and Exo VII treatment, restriction mapping of protected fragments, and preparation of samples for electrophoresis in denaturing urea–acrylamide gels have been described in detail by Calzone et al. (1987). Hybridization reactions contained 10–20 µg of sea urchin poly(A) RNA or yeast total RNA, 0.4 M NaCl, 25 mM PIPES (pH 6.8), 0.2 mM EDTA, and 50% or 80% formamide, depending on the length and base composition of the probe sequence. The amount of probe in each reaction was sufficient for >95% reaction after incubation for 15–18 hr at an appropriate temperature (30–50°C).

RNA titration reactions

Detailed protocols for titrating RNA transcripts with RNA probes and RNases A and T1 have been published previously (Lee et al. 1986). A ^{32}P-labeled RNA probe containing the single-copy sequences present in probe 2 (Figs. 2, 3, and 4) was synthesized using a pSP64 (Promega Biotec) subclone. Hybridization reactions contained 50% formamide, 25 mM PIPES (pH 6.8), 0.4 M NaCl, 1 mM EDTA, and an excess of probe over ISp1 transcript. After incubation for 15–18 hr at 50°C, the quantity of RNA/RNA duplex was assayed by treatment with RNases A and T1 in 0.375 M NaCl, 10 mM Tris (pH 7.4), and 1 mM EDTA, followed by precipitation in 4% trichloroacetic acid, as described by Lee et al. (1986).

DNA sequencing

The nucleotide sequence of the genomic region containing ISp1 was obtained using the dideoxy chain termination method of Sanger et al. (1977). Genomic DNA fragments were subcloned into the M13 sequencing vectors MP10, MP11, MP18, or MP19 (Messing et al. 1981). The 15-mer sequencing primer (New England Biolabs) was used to prime DNA synthesis.

Acknowledgments

References

Akhurst, R.J., F.J. Calzone, J.J. Lee, R.J. Britten, and E.H. Davidson. 1987. Structure and organization of the CyIII actin gene subfamily of the sea urchin, *Strongylocentrotus purpuratus*. *J. Mol. Biol.* **194:** 193–203.

Anderson, D.M., J.D. Richter, M.E. Chamberlin, D.H. Price, R.J. Britten, L.D. Smith, and E.H. Davidson. 1982. Sequence organizaton of the poly(A) RNA synthesized and accumulated in lampbrush chromosome stage *Xenopus laevis* oocytes. *J. Mol. Biol.* **155:** 281–309.

Barberis, A., G. Superti-Furga, and M. Busslinger. 1987. Mutually exclusive interaction of the CCAAT binding factor and displacement protein with overlapping sequences of a histone gene promotor. *Cell* **50:** 347–359.

Birnstiel, M.L., M. Busslinger, and K. Strub. 1985. Transcription termination and 3' processing: The end is in site. *Cell* **41:** 349–359.

Calzone, F.J., R.J. Britten, and E.H. Davidson. 1987. Mapping of gene transcripts by nuclease protection assays and cDNA primer extension. *Methods Enzymol.* **152:** 611–632.

Calzone, F.J., H.T. Jacobs, C.N. Flytzanis, J.W. Posakony, and E.H. Davidson. 1985. Interspersed maternal RNA of sea urchin and amphibian eggs. In *Biology of fertilization, The fertilization response of the egg* (ed. C.B. Mertz and A. Monroy), vol. 3, pp. 347–366. Academic Press, Orlando, Florida.

Cochet, M., F. Gannon, R. Hen, L. Maroteaux, F. Perrin, and P. Chambon. 1979. Organisation and sequence studies of the 17-piece chicken conalbumin gene. *Nature* **282:** 567–574.

Costantini, F.D., R.J. Britten, and E.H. Davidson. 1980. Message sequences and short repetitive sequences are interspersed in sea urchin egg poly(A)$^+$ RNAs. *Nature* **287:** 111–117.

Costantini, F.D., R.H. Scheller, R.J. Britten, and E.H. Davidson. 1978. Repetitive sequence transcripts in the mature sea urchin oocyte. *Cell* **15:** 173–187.

Davidson, E.H. 1986. *Gene activity in early development.* Academic Press, Orlando, Florida.

Denhardt, D.T. 1966. A membrane-filter technique for detection of complementary DNA. *Biochem. Biophys. Res. Commun.* **23:** 641–646.

Diaz, M.O. and J.G. Gall. 1985. Giant readthrough transcription units at the histone loci of lampbrush chromosomes of the newt *Notophthalmus. Chromosoma* **92:** 243–253.

Diaz, M.O., G. Barsacchi-Pilone, K.A. Mahon, and J.G. Gall. 1981. Transcripts from both strands of a satellite DNA occur on lampbrush chromosome loops of the newt *Notophthalmus. Cell* **24:** 649–659.

Firschauf, A.M., H. Lehrach, A. Poustka, and N. Murray. 1983. Lambda replacement vectors carrying polylinker sequences. *J. Mol. Biol.* **170:** 827–842.

Galau, G.A., W.H. Klein, M.M. Davis, B.J. Wold, R.J. Britten, and E.H. Davidson. 1976. Structural gene sets active in embryos and adult tissues of the sea urchin. *Cell* **7:** 487–505.

Gall, J.G., E.C. Stephenson, H.P. Erba, M.O. Diaz, and G. Barsacchi-Pilone. 1981. Histone genes are located at the sphere loci of newt lampbrush chromosomes. *Chromosoma* **84:** 159–171.

Goustin, A.S. and F.H. Wilt. 1982. Direct measurement of histone peptide elongation rate in cleaving sea urchin embryos. *Biochim. Biophys. Acta* **699:** 22–27.

Hough-Evans, B.R., B.J. Wold, S.G. Ernst, R.J. Britten, and E.H. Davidson. 1977. Appearance and persistence of maternal RNA sequences in sea urchin development. *Dev. Biol.* **60:** 258–277.

Krieg, P.A. and D.A. Melton. 1984. Formation of the 3' end of histone mRNA by posttranscriptional processing. *Nature* **308:** 203–206.

Leahy, P.S. 1986. Laboratory culture of *Strongylocentrotus purpuratus* adults, embryos, and larvae. *Methods Cell Biol.* **27:** 1–13.

Lee, J.J. 1986. "The genomic organization and expression of the *Strongylocentrotus purpuratus* actin gene family." Ph.D. thesis, California Institute of Technology, Pasadena, California.

Lee, J.J., F.J. Calzone, R.J. Britten, R.C. Angerer, and E.H. Davidson. 1986. Activation of sea urchin actin genes during

embryogenesis. Measurements of transcript accumulation from five different genes in *Strongylocentrotus purpuratus. J. Mol. Biol.* **188:** 173–183.

Lee, J.J., R.J. Shott, S.J. Rose, T.L. Thomas, R.J. Britten, and E.H. Davidson. 1984. Sea urchin actin gene subtypes. Gene number, linkage, and evolution. *J. Mol. Biol.* **172:** 149–176.

Maniatis, T., R.C. Hardison, E. Lacy, J. Lauer, C. O'Connell, D. Quan, G.K. Sim, and A. Efstratiadis. 1978. The isolation of structural genes from libraries of eucaryotic DNA. *Cell* **15:** 687–701.

Maroteaux, L., R. Heilig, D. Dupret, and J.L. Mandel. 1983. Repetitive satellite-like sequences are present within or upstream from three avian protein-coding genes. *Nucleic Acids Res.* **11:** 1227–1242.

Messing, J., R. Crea, and P.H. Seeburg. 1981. A system for shotgun DNA sequencing. *Nucleic Acids Res.* **9:** 309–321.

Monson, J.M., J. Friedman, and B.J. McCarthy. 1982. DNA sequence analysis of a mouse proα1 (I) procollagen gene: Evidence for a mouse B1 element within the gene. *Mol. Cell. Biol.* **2:** 1362–1371.

Posakony, J.W., C.N. Flytzanis, R.J. Britten, and E.H. Davidson. 1983. Interspersed sequence organization and developmental representation of cloned poly(A) RNAs from sea urchin eggs. *J. Mol. Biol.* **167:** 361–389.

Proudfoot, N.J. and G.G. Brownlee. 1976. 3' non-coding region sequences in eukaryotic messenger RNA. *Nature* **263:** 211–214.

Richter, J.D., D.M. Anderson, E.H. Davidson, and L.D. Smith. 1984. Interspersed poly(A) RNAs of amphibian oocytes are not translatable. *J. Mol. Biol.* **173:** 227–241.

Rosenthal, E.T. and J.V. Ruderman. 1987. Widespread changes in the translation and adenylation of maternal messenger RNAs following fertilization in *Spisula* oocytes. *Dev. Biol.* **121:** 237–246.

Ruzdijic, S. and T. Pederson. 1987. Evidence for an association between U1 RNA and interspersed repeat-single copy RNAs in the cytoplasm of sea urchin eggs. *Development* (in press).

Ryffel, G.U., D.B. Muellener, S. Gerber-Huber, T. Wylie, and W. Wahli. 1983. Scattering of repetitive DNA sequences in the albumin and vitellogenin gene loci of *Xenopus laevis. Nucleic Acids Res.* **11:** 7701–7716.

Sanger, F., S. Nicklen, and A.R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci.* **74:** 5463–5467.

Schibler, U., A.-C. Pittet, R.A. Young, O. Hagenbüchle, M. Tosi, S. Gellman, and P.K. Wellauer. 1982. The mouse α-amylase multigene family. Sequence organization of members expressed in the pancreas, salivary gland and liver. *J. Mol. Biol.* **155:** 247–266.

Shott, R.J., J.J. Lee, R.J. Britten, and E.H. Davidson. 1984. Differential expression of the actin gene family of *Strongylocentrotus purpuratus. Dev. Biol.* **101:** 295–306.

Smith, M.J., B.R. Hough, M.E. Chamberlin, and E.H. Davidson. 1974. Repetitive and nonrepetitive sequence in sea urchin hnRNA. *J. Mol. Biol.* **85:** 103–126.

Stephenson, E.C., H.P. Erba, and J.G. Gall. 1981. Histone gene clusters of the newt *Notophthalmus* are separated by long tracts of satellite DNA. *Cell* **24:** 639–647.

Sucov, H.M., S. Benson, J.J. Robinson, R.J. Britten, F. Wilt, and E.H. Davidson. 1987. A lineage-specific gene encoding a major matrix protein of the sea urchin embryo spicule. II. Structure of the gene and derived sequence of the protein. *Dev. Biol.* **120:** 507–519.

Suske, G., M. Wenz, A.C.B. Cato, and M. Beato. 1983. The uteroglobin gene region: Hormonal regulation, repetitive ele-

ments and complete nucleotide sequence of the gene. *Nucleic Acids Res.* **11:** 2257–2271.

Thomas, T.L., J.W. Posakony, D.M. Anderson, R.J. Britten, and E.H. Davidson. 1981. Molecular structure of maternal RNA. *Chromosoma* **84:** 319–335.

Tsukada, T., Y. Watanabe, Y. Nakai, H. Imura, S. Nakanisha, and S. Numa. 1982. Repetitive DNA sequences in the human corticotropin-beta-lipotropin precursor gene region: Alu family members. *Nucleic Acids Res.* **10:** 1471–1479.

Viera, J. and J. Messing. 1982. The pUC plasmids, an M13mp7-derived system for insertion mutagenesis and sequencing with synthetic primers. *Gene* **19:** 259–268.

Wahli, W., I.B. Dawid, G.U. Ryffel, and R. Weber. 1981. Vitellogenesis and the vitellogenin gene family. *Science* **212:** 298–304.

Wickens, M. and J. Gurdon. 1983. Post-transcriptional processing of simian virus 40 late transcripts in injected frog oocytes. *J. Mol. Biol.* **163:** 1–26.

A long, nontranslatable poly(A) RNA stored in the sea urchin Strongylocentrotus purpuratus.

F J Calzone, J J Lee, N Le, et al.

References	This article cites 41 articles, 8 of which can be accessed free at: **http://genesdev.cshlp.org/content/2/3/305.full.html#ref-list-1**
Email Alerting Service	Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or **click here.**