Drosophila melanogaster genes for U1 snRNA variants and their expression during development

Patrick C.H.Lo and Stephen M.Mount*
Department of Biological Sciences, Columbia University, New York, NY 10027, USA

Received August 9, 1990; Revised and Accepted October 25, 1990 EMBL accession nos X53541–X53544

ABSTRACT

We have cloned and characterized a complete set of seven U1-related sequences from *Drosophila melanogaster*. These sequences are located at the three cytogenetic loci 21D, 82E, and 95C. Three of these sequences have been previously studied: one U1 gene at 21D which encodes the prototype U1 sequence (U1a), one U1 gene at 82E which encodes a U1 variant with a single nucleotide substitution (U1b), and a pseudogene at 82E. The four previously uncharacterized genes are another U1b gene at 82E, two additional U1a genes at 95C, and a U1 gene at 95C which encodes a new variant (U1c) with a distinct single nucleotide change relative to U1a. Three blocks of 5′ flanking sequence similarity are common to all six full length genes. Using specific primer extension assays, we have observed that the U1b RNA is expressed in *Drosophila* K_c cells and is associated with snRNP proteins, suggesting that the U1b-containing snRNP particles are able to participate in the process of pre-mRNA splicing. We have also examined the expression throughout *Drosophila* development of the two U1 variants relative to the prototype sequence. The U1c variant is undetectable by our methods, while the U1b variant exhibits a primarily embryonic pattern reminiscent of the expression of certain U1 variants in sea urchin, *Xenopus*, and mouse.

INTRODUCTION

The splicing of intervening sequences from pre-mRNAs is a two step process that occurs within the large, multi-component nuclear complex known as the spliceosome (for reviews see 1,2). Among the components of the spliceosome are small nuclear ribonucleoprotein particles, or snRNPs (for reviews see 3,4). Each snRNP particle consists of one or two different members of the U family of small nuclear RNAs (snRNAs) complexed with approximately ten proteins (5). One of these snRNPs, the U1 snRNP, is essential for the interaction of the spliceosome with the 5′ splice site of introns (6–8). This interaction is mediated by the basepairing of the highly conserved 5′ end of U1 snRNA with the 5′ splice site, as was demonstrated genetically by Zhuang and Weiner (9).

It is possible that in the process of alternative splicing the choice of splice sites is affected by heterogeneity in the U1 snRNP population. One such source of heterogeneity could be the U1 snRNA molecule. In higher eukaryotes, there is variation in the coding sequences of the true U1 genes within particular species (10–14). These sequence differences are relatively minor, and none occur in the highly conserved 5′ end of the molecule that interacts with the 5′ splice sites of primary transcripts.

Evidence that U1 sequence variants may be functionally significant comes from observations that differential regulation of their expression occurs during development. This has been observed in sea urchins (15), *Xenopus* (11,16), and mice (17). In general, there seems to be two major groups of U1 snRNAs: an adult class of U1 sequences (referred to as U1a) and an embryonic class (referred to as U1b). The U1a species predominates throughout development, especially during later stages, and it is almost exclusively expressed in various terminally differentiated cell lines. In contrast, the U1b species is abundant only during the embryonic stages of development and is the major variant present in cell lines that possess the ability to differentiate.

The number of copies of true U1 genes per haploid genome varies anywhere from 10 in chickens, 30 in humans, 40 in mice, up to a high of 1,000 in *Xenopus* (for a review see 19). In addition, a significant proportion of U1-related sequences in most higher eukaryotic species are pseudogenes. By contrast, *Drosophila melanogaster* was known to have relatively few U1-related sequences, on the order of six to seven (18), making it a suitable organism in which to examine the differential regulation of U1 sequence variants.

Mount and Steitz (20) determined the sequence of a U1 RNA, isolated by immunoprecipitation with anti-(U1)RNP antibodies from the *Drosophila* K_c cell line, by direct RNA sequencing methods, and obtained a single genomic clone of what appeared to be the gene that coded for that U1 sequence. *In situ* hybridization showed that this clone was derived from the salivary gland chromosomal band 21D and that *Drosophila* has U1-related sequences at only three cytological locations: 21D, 82E, and 95C (M.L. Pardue, unpublished). This was consistent with results from hybridization with iodinated snRNAs (18). Alonso *et al.* (21) subsequently isolated and sequenced a U1 gene from 21D that was identical to the gene first isolated by Mount and Steitz (20). Southern analyses by all of these groups proved to be

* To whom correspondence should be addressed

inconclusive, but generally indicated that the Drosophila genome has about seven U1-related sequences. Two additional U1 sequences from the 82E locus were isolated by Kejzlarova-Lepesant *et al.* (22). One of these appeared to be a pseudogene with the first 50 nucleotides of the U1 coding sequence deleted. The other was a sequence variant that possessed a full-length coding sequence with a single nucleotide change: a G to T transversion at position 134. In order to distinguish this sequence variant from the original prototype sequence identified in K_c cells, we have designated the prototype U1 sequence as U1a and the variant as U1b.

Here we report on the isolation and characterization of what we believe to be a complete set of U1-related sequences from *Drosophila melanogaster*. In particular, we have cloned the U1-related sequences present at the 95C locus and discovered a new U1 sequence variant (U1c). We have examined the abundance through development of these variants and found that one (U1b) is an embryonic form while the other (U1c) shows no detectable expression.

MATERIALS AND METHODS

Preparation of U1 probe

The plasmid pPL1 was generated by the insertion of the 119 bp *Eco*RI-*Hin*dIII fragment of pDmU1.4d (23) into *Eco*RI-*Hin*dIII-digested pIBI24 (International Biotechnologies Inc.). The resulting plasmid contains nucleotides 50−140 of U1a flanked upstream by 24 basepairs (nucleotides 0−24) and downstream by 5 basepairs (nucleotides 25−29) of pBR322 sequence, oriented such that transcription from the T7 promoter of pIBI24 will generate an RNA with the U1 sequence in the sense orientation.

The probe for screening the genomic library was prepared from the T7 transcription of *Hin*dIII-linearized pPL1 in a reaction volume of 20 μl under the following conditions: 1.5 μg template DNA, 40 mM Tris-HCl (pH 8.0), 20 mM $MgCl_2$, 10 mM NaCl, 0.5 mM each of ATP, GTP, and UTP, 12 μM CTP, 7 μCi/μl ^{32}P-α-CTP (6 μM), and 40 units of T7 RNA polymerase (U.S. Biochemical) at 37° C for 30 minutes. The reaction was terminated by the addition of 1.6 μl of 250 mM EDTA (pH 8.0) and 1.0 μl of 10% SDS.

Screening of a genomic library

A genomic library of *Sau*3A partially-digested Oregon R *Drosophila melanogaster* DNA cloned into *Bam*H1-digested λEMBL4 was a gift of Mike Goldberg. Duplicate sets of nitrocellulose filters representing 120,000 plaques each were screened according to Maniatis *et al.* (24) using as a probe the ^{32}P-labelled RNA described above, which was added directly to the hybridization solution.

Clones were assigned to specific loci based on Southern blot analysis using either the 21D-specific probe pDmU1.4b (23) or the 82E-specific probe pDm6B (22; a kind gift of Dr. Jean-Antoine Lepesant).

Subcloning of genomic fragments

The 3.4 kb *Eco*RI fragment from λD5-8 with the U1-82.3 gene was subcloned into *Eco*RI-cut pIBI24 to generate pPL2. The 8.9 kb *Eco*RI fragment from λD5-8 containing the U1-82.2 pseudogene and the U1-82.1 gene was subcloned into *Eco*RI-cut pIBI24 to generate pPL3. The 2.3 kb *Bam*HI fragment containing the U1-82.2 pseudogene from the insert of pPL3 was then further subcloned into *Bam*HI-digested pIBI24 to create pPL5, which is equivalent to the subclone pDm6Y of Kejzlarova-Lepesant *et al.* (22). The 2.5 kb *Bam*HI fragment containing the U1-82.1 gene from the insert of pPL3 (which is actually the *Bam*H1-*Eco*RI fragment of λD5-8 with a *Bam*H1 site from the pIBI24 polylinker) was subcloned into *Bam*HI-digested pIBI24 to create pPL6, which is equivalent to the subclone pDm6A of Kejzlarova-Lepesant *et al.* (22).

The 2.6 kb *Eco*RI fragment of λD5-50 containing the U1-95.1 gene was subcloned into *Eco*RI-cut pIBI24 to generate pPL4. The 1.9 kb *Eco*RI fragment from λD5-50 containing U1-95.2 and U1-95.3 was used for sequencing without prior subcloning.

In situ hybridizations

Salivary gland squashes and *in situ* hybridizations were carried out according to Ashburner (25), using specific genomic clones as probes.

Sequencing of U1-related sequences

Sequencing was carried out by the dideoxynucleotide chain termination method (27) with modified T7 polymerase (Sequenase® , U.S. Biochemicals; 28), according to the strategy of Bankier and Barrell (26) and using Beckman Microgenie® software.

Genomic Southerns

Genomic DNA from the various strains was isolated as described by McGinnis *et al.* (29), with the omission of the RNase treatment. 3 μg of genomic DNA from each strain was digested with the restriction enzymes indicated in Figure 4, electrophoresed on a 0.8% agarose gel, and Southern blotted according to Maniatis *et al.* (24). Separate blots were individually probed with various plasmid subclones labelled with ^{32}P according to Feinberg and Vogelstein (30). The subclone pPL1 was used to identify all restriction fragments with U1-related sequences. pDmU1.4b (20) was used to identify the fragments with the U1-21.1 gene, and pPL4 to identify those with the U1-95.1 gene. For the U1b genes, uniformly-labelled, single-stranded probes were synthesized as described in Maniatis *et al.* (24) from specific M13 single-stranded clones obtained from shotgun sequencing. The probe for the 82.1 gene was synthesized from the M13 clone D41-053 and had sequences starting from +311 basepairs downstream of the gene to the *Cla*I site at +601. The probe for the 82.3 gene was derived from the M13 clone D35-130 and had sequences starting from +82 basepairs downstream to the *Hha*I site at +282.

The strains used included two wild type strains (Canton S and Oregon R, both obtained from the Rubin laboratory) and three strains carrying isogenic third chromosomes. Chromosomes were isogenized using the w^{1118}; *TM3/CxD* strain (31). The complete genotypes of these three strains are 1) se^{EC4}, which carries a third chromosome extracted from wild flies collected in El Cerrito, California; 2) w^{1118}; P[w^{aRsLTR}]21 (31) and 3) st^{85e}, a strain carrying a spontaneous allele of *scarlet*.

Preparation of T7-U1 RNA variants

Templates for the T7 transcription of RNAs encoding each of the three U1 sequence variants were synthesized by PCR. The oligonucleotides used were synthesized on an Applied Biosystems 280A DNA synthesizer. The sequence of the 5′ oligonucleotide (OLIGO 6; 5′-ATTTAGGTGACACTATAATACTTACC-TGGCGT-3′) consisted of the 17 nucleotide SP6 promoter joined to the first 15 nucleotides of the *Drosophila* U1 sequence. The

sequence of the 3' oligonucleotide was complementary to the final 15 nucleotides of the *Drosophila* U1 sequence. The PCR reactions were carried out essentially as described by Saiki *et al.* (32). The PCR products were gel-purified, phenol-chloroform extracted twice, and ethanol precipitated prior to T7 transcription.

The T7 transcriptions were carried out as described before in a 100 μl volume with 120 ng of a T7-U1 DNA template, 0.5 mM each of ATP, CTP, GTP, and UTP, 100 units of RNasin (Promega Biotec), and 85 units of T7 RNA polymerase. These were incubated at 37° C for 2 hours, after which 1 unit of RNase-free DNase (Promega Biotec) was added and the incubation continued for another 15 minutes. The reactions were phenol-chloroform and chloroform extracted and ethanol precipitated.

Isolation of RNAs from Drosophila melanogaster

Total or poly A^- RNA from the indicated developmental stages of *Drosophila melanogaster* Oregon R flies were prepared basically as described by O'Hare *et al.* (33), using a Brinkmann 'polytron' homogenizer.

Immunoprecipitation of U1 snRNA from K_c nuclear extracts

Nuclear extracts from *Drosophila* K_c cells, prepared essentially as described in Dignam *et al.*, (34), were the kind gift of Ming Guo. The immunoprecipitation of snRNPs from a 40 μl aliquot of K_c nuclear extract were carried out with 10 μl of human anti-Sm antisera (35; a kind gift of Zhen-Qiang Pan). The procedure is as described by Steitz (36) using protein A-Sepharose beads (Pharmacia) and IPP buffer [0.5 M NaCl, 0.1% NP-40, 10 mM Tris (pH 7.9)], with the addition of a proteinase K treatment step. Prior to the phenol-chloroform extractions of the samples, 400 μl of proteinase K solution [1.7% SDS, 300 mM NaCl, 40 mM Tris (pH 7.5), 0.5 mg/ml proteinase K (Boehringer Mannheim)] was added to each sample and these were incubated at 30° C for 30 minutes. The remainder of the procedure is as described, and the resulting immunoprecipitated RNAs were resuspended in 20 μl of double distilled water.

Primer extension assays for the U1 sequence variants

For the developmental assay of the U1b variant relative to the non-U1b (U1a and U1c) sequences, 2 μg of poly A^- or total RNA from the indicated stages of *Drosophila* development was annealled to 120 fmol of a ^{32}P-end labelled oligomer (OLIGO 4) complementary to the last 28 nucleotides of the *Drosophila* U1 molecule in the presence of 50 mM Tris-Cl (pH 8.4), 60 mM NaCl, and 10 mM dithiothreitol. For the same assay of the anti-Sm-precipitated U1 snRNAs, the annealling consisted of approximately 30 ng of the immunoprecipitated RNA. The samples were heated at 95° C and allowed to slowly cool to about room temperature. These were then extended with MMLV reverse transcriptase (Bethesda Research Laboratories) in the presence of 6 mM $MgCl_2$ and 375 μM each of dATP, dGTP, dTTP, and ddCTP at 37° C for 30 min. The reactions were stopped by the addition of a formamide loading buffer and electrophoresed on a 15% acrylamide/8.3 M urea sequencing gel. Dried gels were exposed to XAR-5 film with a Quanta III intensifying screen.

The assay for the expression of the U1c variant relative to the non-U1c (U1a & U1b) sequences during development is identical to the above assay, except for the extension step, which was done with AMV reverse transcriptase (Boehringer Mannheim) in the presence of 6 mM $MgCl_2$, 375 μM each of dATP, dCTP, dTTP, and 94 μM of ddGTP at 42° C for 30 min.

RESULTS

Isolation and characterization of Drosophila U1 sequence variants

We screened at a high stringency approximately ten genome equivalents of a genomic library of Oregon R DNA with a probe from the *Drosophila* U1a coding region (see Materials and Methods). A total of 61 positives were isolated from the primary screen, and of these 25 were purified through a secondary screen. Ten of these latter clones were then chosen at random for detailed characterization by restriction digests and Southern analyses. Southern blots using probes that were specific to the U1 genes at either 21D (20) or 82E (22), or that contained only the U1 coding region allowed these clones to be divided into three groups of identical or overlapping clones. One group consisted of four phage from 21D, a second of three phage from 82E, and the third of three phage which did not hybridize to either of the locus-specific probes, which suggested that they contained the U1-related sequences from the 95C locus. For both the 82E and putative 95C groups, we were able to identify for further study one phage that contained all the U1-hybridizing sequences that could be found in that particular group of phage. Restriction maps of the representative clone from 82E (λD5-8) and that from 95C (λD5-50) are shown in Figure 1. A restriction map of a previously isolated genomic clone from the 21D locus (37) that contains the U1-21.1 gene (S.M. Mount, unpublished) is also presented in Figure 1.

In situ hybridizations to salivary gland chromosomes using λD5-50 as a probe verified that it originated from the 95C locus

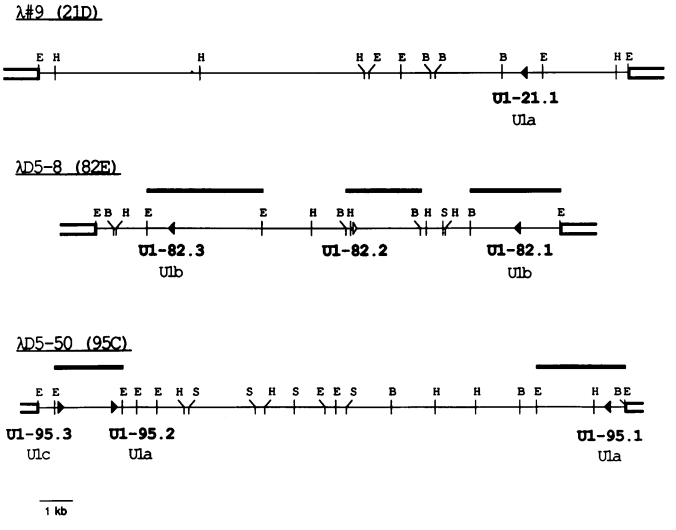

Figure 1. Restriction maps of genomic clones. The restriction maps of the inserts from the genomic clones λD5-8 and λD5-50 are shown. The bars over the maps of these two clones indicate the restriction fragments that were sequenced. For completeness, a restriction map of a genomic clone from 21D (37) containing the U1-21.1 gene (S.M. Mount, unpublished) is also shown. Solid arrowheads indicate the position and orientation of U1 genes within a restriction fragment. The open arrowhead indicates the position and orientation of the pseudogene. The restriction sites are: B= *Bam*HI; E= *Eco*RI; H= *Hin*dIII; and S= *Sal*I. Two other previously isolated genomic clones from the 21D locus containing a copy of the U1-21.1 gene, U14 and U17 (21), extend approximately 9 kb further to the right of λ # 9 and neither contains any other U1-related sequence. It should be noted that the restriction map of the phage Dm525 of Kejzlarova-Lepesant *et al.* (22) containing the U1-82.1 and U1-82.2 sequences differs from the restriction map of phage λD5-8 which contains these two sequences and the U1-82.3 sequence.

(data not shown). The clone λD5-8 hybridized strongly to 82E, and less strongly to the chromocenter and to another band proximal to 82E (data not shown). This localization of the weaker hybridization suggests that the λD5-8 clone from 82E contains a moderately repetitive element that is found in β-heterochromatin.

Five restriction fragments containing U1-related sequences present in λD5-8 and λD5-50 were sequenced, and a total of six U1-related sequences were found. Each of the two representative phage contained three of these sequences (Figure 1). In order to systematize the designation of the U1-related sequences, a two part number was assigned to each sequence. The first number indicates the sequence's cytological location and the second distinguishes the various sequences at a particular cytological locus.

U1-82.1 and 82.2 correspond to U1 sequences previously isolated from the 82E locus and studied by Kejzlarova-Lepesant *et al.* (22). U1-82.3 is also derived from the 82E locus and possesses a U1b coding region. Two of the genes from the 95C locus, U1-95.1 and U1-95.2, were found to contain U1a coding regions. The third gene at 95C, U1-95.3, represents a new U1 sequence variant, U1c, consisting of a single T to C transition (relative to U1a) at position 123 of the U1 coding sequence (see Figure 2).

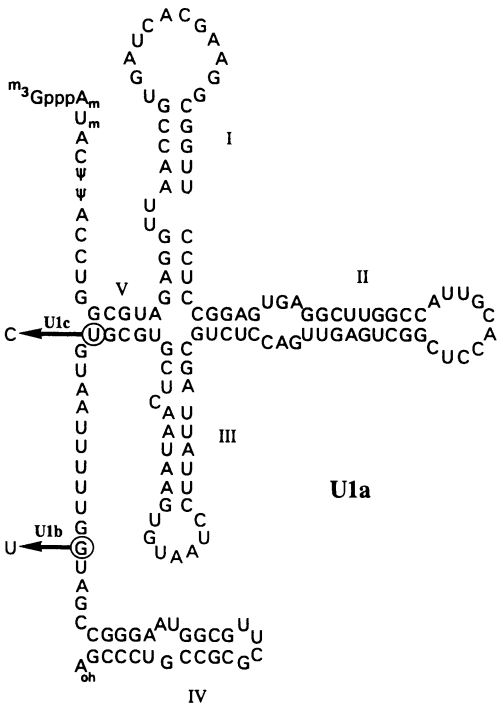

Figure 2. Predicted RNA sequence variants. RNA sequence changes inferred from the DNA sequences of U1-82.1 (U1b), U1-82.3 (U1b), and U1-95.3 (U1c) are shown on the standard secondary structure of Drosophila U1a (20).

Flanking sequences of the U1 genes

An examination of upstream sequences of the six *Drosophila* U1 genes shows three blocks of considerable homology (see Figure 3). Beck *et al.* (38) had first noted that the 5′ flanking sequence of the single U1 gene that they had isolated from 21D (U1-21.1) shared two blocks of homology (at −61 to −41 and −32 to −25) with the 5′ flanking sequences of the three *Drosophila* U2 genes. These two conserved sequence elements were later observed in the 5′ flanking sequence of the full-length U1 sequence isolated from 82E (U1-82.1) by Kejzlarova-Lepesant *et al.* (22). Saba *et al.* (39) have also observed that these two blocks of homology are present in the upstream sequences of the two U4 genes of *Drosophila* that they isolated and characterized. The four additional U1-related sequences that we have characterized in this study also possess these two conserved 5′ sequence elements. These twelve upstream sequences all have between 17 to 21 matches to the 21 nucleotide consensus TAATTCCCAA-CTGGTTCTRGC at −61 to −41 and between 6 to 8 matches to the 8 nucleotide consensus CATGGARA at −32 to −25. Thus, all known RNA polymerase II-transcribed U snRNA genes that have been sequenced in *Drosophila* possess these two highly conserved sequence elements in their 5′ flanking sequence.

In addition, we have noticed that there is an absolutely conserved pentanucleotide, AAAGC, located immediately upstream of the presumed transcriptional start site of all the *Drosophila* U1 genes (Figure 3). This sequence element is unique to the U1 genes and is not present immediately upstream of the transcriptional start sites of any of the U2 or U4 genes of *Drosophila*.

Both full-length U1 genes at 82E have nearly identical flanking sequences

The sequences flanking the two full-length genes from 82E, U1-82.1 and U1-82.3, are identical for 375 nucleotides upstream of the gene and identical for 93 of 96 nucleotides downstream. Furthermore, there is a 237 bp region of near identity (92%) further downstream of both genes, beginning at +303 relative to U1-82.1 and +174 relative to U1-82.3. These regions of near identity in the flanking sequences of U1b genes seem to suggest that one of these genes was derived from the other through a recent duplication event.

Copy number of U1 sequences in Drosophila melanogaster

We attempted to determine the copy number of U1-related sequences in *Drosophila melanogaster* by carrying out Southern analyses on two wild-type strains (Canton S and Oregon R) and three strains that had been isogenized for chromosome III. Genomic DNAs from these strains were restricted with *Hpa*I, *Eco*RI, or *Bam*HI and *Eco*RI, blotted, and probed with a sequence that corresponded to nucleotides 50−140 of the U1a coding region. Since there is only one *Hpa*I site (at position 23) in each of the full-length U1 sequences the number of *Hpa*I bands which are detected should be equal to the number of full-length genes plus the one pseudogene, provided that the latter is either separated by a *Hpa*I site or is upstream from the 5′ end of an adjacent full-length sequence. In the case of the U1-82.2 pseudogene located on λD5-8, it is separated by a *Hpa*I site from either flanking full-length U1 gene (data not shown). An examination of the *Hpa*I digests (Figure 4) reveals variation in the number of bands from three to seven between strains. However, the uppermost band (marked by the open arrowhead) probably results from partial digestion since its intensity relative to other bands varies between lanes and between repetitions of the same experiment (data not shown).

We also hybridized these replicate blots with probes that were specific for various U1 genes (data not shown). A probe specific for the U1-95.1 gene detected a *Hpa*I band from all strains except st^{85e} that comigrated with a band also seen with a probe specific for the U1-21.1 gene. Therefore, it is neccesary to add one to

Figure 3. U1 coding and flanking sequences. The coding and flanking sequences of the seven U1 genes of *Drosophila melanogaster* are shown aligned by the U1 coding region. In addition, the 5′ flanking sequences of the *Drosophila melanogaster* U2 (38) and U4 genes (39) are shown. Note the conserved regions of −61 to −41, −32 to −25, and −5 to −1 (denoted by outlining) shared among the 5′ flanking sequences of all the U1 genes. The regions at −61 to −41 and -32 to −25 are also conserved in the U2 and U4 genes, while the region at −5 to −1 is absolutely conserved only among the U1 genes.

the number of *Hpa*I bands in these lanes in order to obtain the correct number of U1-related sequences for those strains. In the st^{85e} lane, the comigrating bands present in the other strains are resolved into a closely spaced doublet, with the upper band corresponding to the 21.1 gene and the lower band to the 95.1 gene. Probes specific for the U1b genes from the 82E locus (82.1 and 82.3) were also used. These probes were derived from the 3′ flanking sequences of either U1b gene; both possessed sequences unique to each gene and different amounts of sequence from the 237 bp region of near identity in the 3′ flank of these genes (see Materials and Methods). Blots hybridized with these probes showed that each strain contains at least one copy of a U1b gene from the 82E locus. Despite some cross-hybridization, it was possible to specifically assign U1b genes with *Hpa*I fragments based on the differential levels of hybridization to the two probes. The other HpaI bands were assigned to the remaining U1-related sequences based on their sizes, since these were the same as the sizes of the respective HpaI bands derived from the genomic clones (data not shown). It should be noted that a copy of the U1-82.1 gene comigrates with the U1-95.2 band in the Canton S lane.

Overall, it appears from these Southern analyses that the number of U1-related sequences varies between five in the EC4, Canton S, and st^{85e} strains and seven in the Oregon R and V21 strains, with the variation being in the number of sequences at 82E. Since we have isolated a total of seven U1-related sequences from our screen of an Oregon R genomic library, we are confident that we have isolated all the possible U1 sequences from *Drosophila melanogaster*.

Developmental patterns of expression of the U1 variants

Having characterized what we believe to be a complete set of *Drosophila* U1 genes, we next examined the expression of the three different U1 sequences during development. A set of specialized primer extension assays were devised that could discriminate either variant (U1b or U1c) from the other two sequences (see Figure 5). Extension of a 28-mer complementary to the very 3′ end of *Drosophila* U1 in the presence of dATP, dGTP, dTTP, and ddCTP results in a 32 nucleotide extension product from U1b RNA and a shorter 31 nucleotide extension product from non-U1b (U1a/U1c) RNA. The assay for the U1c variant is analogous, with the extension carried out in the presence of dATP, dCTP, dTTP, and ddGTP. The U1c RNA will generate a 42 nucleotide product while non-U1c RNA will generate a 44

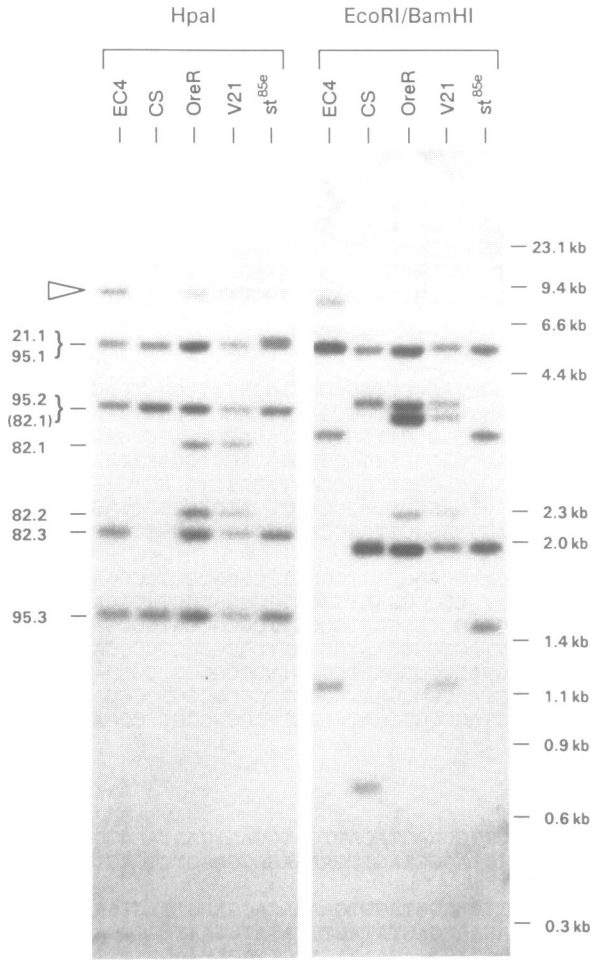

Figure 4. Genomic Southerns of various strains Genomic DNA from the indicated strains were digested with either *Hpa*I, or *Bam*HI and *Eco*RI. The blot was hybridized to the probe containing the U1a coding sequence from 50 to 140. The assignment of specific U1-related sequences to *Hpa*I bands is shown on the left of the figure (see text). The positions of size markers are indicated on the right side. The open arrowhead indicates the *Hpa*I band that is the result of partial digestion. The strains are EC4= e^{EC4}, CS= Canton S, OreR= Oregon R, V21= w^{1118}; P[w^{aRsLTR}]21, and st^{85e}= st^{85e}.

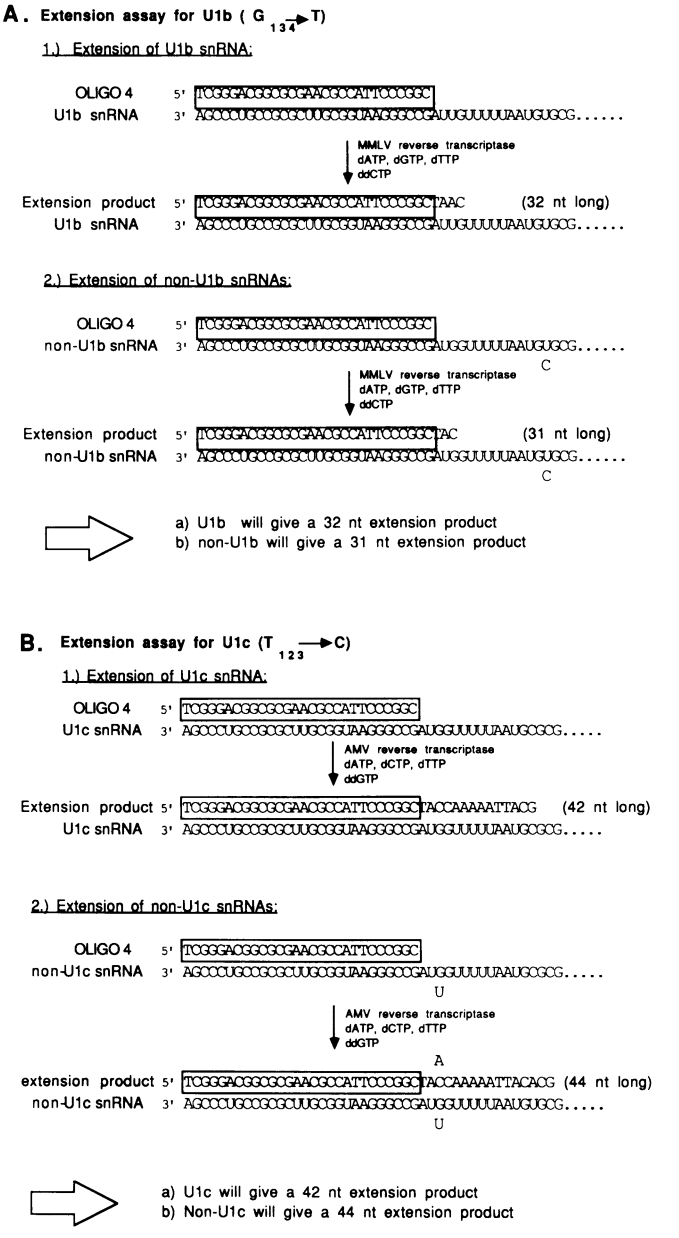

Figure 5. Primer extension assays for the *Drosophila* U1 variants. Using a DNA oligomer that is complementary to the 28 nucleotides at the 3′ end of all *Drosophila* U1 variants (OLIGO 4), it is possible to obtain differently sized extension products for each sequence variant if the appropriate mix of three dNTPs and one ddNTP is used with reverse transcriptase. Panel A summarizes the reaction conditions and expected products for the U1b assay used in Figure 6B, C, and D. Panel B summarizes the U1c assay used in Figure 6A.

nucleotide product. It should be noted that we have used these primer extensions to quantitate the ratios of either U1b to non-U1b RNAs or U1c to non-U1c RNAs, and not to measure the absolute amounts of the individual U1 sequences. We also cannot determine by these assays which of the multi-copy U1 variant genes are actually being transcribed *in vivo*.

These assays were first tested on T7 RNAs transcribed from templates containing one of the three *Drosophila* U1 sequences (see Materials and Methods). In all cases, the expected extension products were observed and for the samples that contained a mixture of U1a with either U1b or U1c, the ratio of extension products was the same as the ratio of the different RNAs present in the sample being assayed (Figure 6A,D; data not shown).

We then used these assays to measure the relative amounts of the U1 variants relative to U1a in RNA samples from various stages of *Drosophila* development (Figure 6). Within the limits of detection of our assay, there is either an extremely low amount of U1c (less than 5% of the level of U1a/U1b) or none of this variant being expressed *in vivo* (Figure 6A). Since this is the case, we can consider the non-U1b class of RNAs to be composed almost totally of the U1a species.

Using our primer extension assay it was observed that the abundance of U1b does vary relative to U1a during development. During early embryogenesis (see Figure 6C), the ratio of U1b to U1a is close to 1:1, but then this ratio decreases to about 1:2 by 12 hours. This ratio of U1b to U1a is maintained until the 3rd instar larval stage, at which point at least a two-fold decrease in the ratio occurs (Figure 6B). This ratio remains constant throughout the remainder of development, except in late female adults, where a much higher ratio of U1b to non-U1b is again

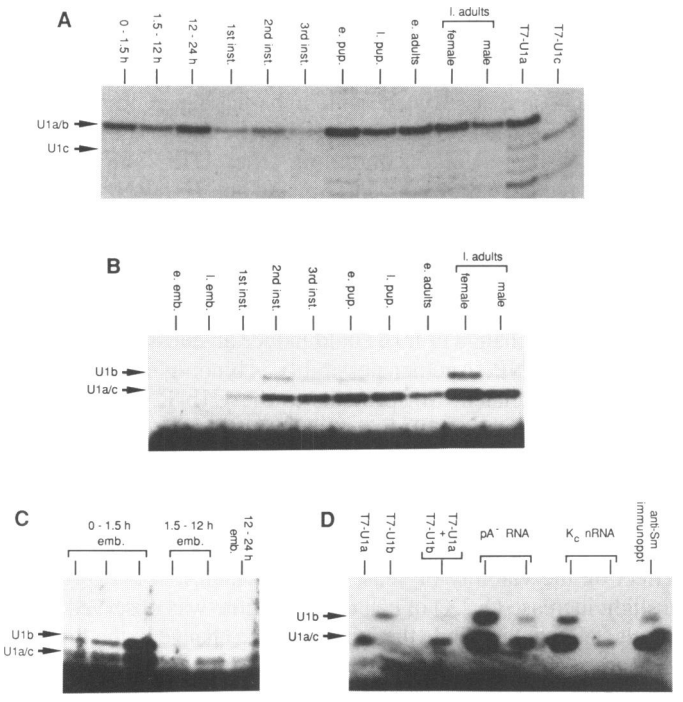

Figure 6. Developmental expression patterns of the U1 variants. The primer extension assays described in the Figure 5 were used to assay the relative expression levels of the different U1 sequences during development in *Drosophila melanogaster*. **A.** Assay of the expression of the U1c variant relative to the non-U1c (U1a & U1b) sequences throughout development. The developmental stages are: 0−1.5 h = 0−1.5 h embryos; 1.5−12 h = 1.5−12 h embryos; 12−24 h = 12−24 h embryos; 1st inst.= 1st instar larvae; 2nd inst.= 2nd instar larvae; 3rd inst.=3rd instar larvae; e. pup.= early pupae; l. pup.= late pupae; e. adults= early adults (less than 1 day post-eclosion); l. adults = late adults; T7-U1a= T7-transcribed U1a sequence; T7-U1c = T7-transcribed U1c sequence. **B.** Assay of the expression of the U1b variant relative to the non-U1b (U1a & U1c) sequences in poly A⁻ RNA throughout development. The samples are as described above, except for the embryonic samples, which are: e. emb.= early embryos (0−12 hour); l. emb.= late embryos (12−24 hour). The extremely dark band at the bottom of this panel and panels C and D is due to unextended primer. **C.** Assay of the expression of the U1b variant relative to the non-U1b (U1a & U1c) sequences during embryogenesis. Lanes 1−3 represent multiple, independently isolated samples of total RNA from the 0−1.5 hour time period; lanes 4−5 are independent samples from the 1.5−12 hour time period; and lane 6 is a single sample from the 12−24 hour time period. **D.** Assay of the expression of the U1b variant relative to the non-U1b (U1a & U1c) sequences in the total nuclear RNA and anti-Sm-immunoprecipitate of nuclear extracts from *Drosophila* K_c cells. The samples are: T7-U1a= T7-transcribed U1a sequence; T7-U1b= T7-transcribed U1b sequence; T7-U1a+T7-U1b= a 5:1 mixture, respectively, of the two T7-transcribed U1 sequences; pA⁻ RNA= 18.5 and 1.9 μg, respectively, of pA⁻ RNA from adult flies; K_c nRNA= 2.5 and 0.5 μg, respectively, of total RNA extracted from K_c cell nuclei; anti-Sm immunoppt= 30 ngs of anti-Sm-precipitated RNAs from K_c nuclear extracts.

seen. This higher ratio is presumably due to a maternal deposition of relatively more U1b in the developing oocytes, consistent with the observation of relatively high U1b levels in embryos less than 1.5 hours old.

U1b is expressed in K_c cells and is anti-Sm-precipitable

Expression of the U1b variant in flies left open the question of whether this class of U1 molecules is fully functional with respect to the role of the U1 snRNP in recognizing 5′ splice sites. The U1 snRNP normally contains U1 snRNA complexed with the common snRNP proteins and three U1 snRNP-specific proteins: A, C, and 70K. The Sm proteins common to all the spliceosomal snRNPs (U1, U2, U4/U6, and U5) are recognized by anti-Sm antisera that are derived from patients with certain autoimmune disorders. U1 snRNP-specific proteins are recognized by anti-(U1)RNP antisera, which allows for the selective immunoprecipitation of U1 snRNP particles. We asked if the U1b variant was present in *Drosophila* K_c cells and if it was in an anti-Sm-precipitable form. An association of the U1b molecules with Sm proteins would indicate that this sequence variant was at least competent to form an RNP particle of some sort. Nuclei were prepared from *Drosophila* K_c cells and total RNA was prepared from these nuclei. The primer extension assay for U1b detected this variant in the K_c nuclear RNA at a level that was approximately one-tenth that of the U1a levels (Figure 6D), indicating that U1b was being expressed as a minor species in this cell line. Nuclear extracts from K_c cells were then immunoprecipitated with a human anti-Sm antiserum and the RNA extracted from the immunoprecipitate was analyzed by the primer extension assay (Figure 6D). From this experiment, it was clear that the U1b variant was present in the *Drosophila* K_c cells in an anti-Sm-precipitable form and that it also represented about one-tenth of the amount of U1 present in the immunoprecipitate.

Direct RNA sequence analysis of K_c cell U1 RNA (20) immunoprecipitated with anti-(U1)RNP serum had previously indicated only the U1a species. However, re-examination of the RNA fingerprints and secondary digests of the RNase T1 oligonucleotides from that study confirmed that the U1b variant was also present (data not shown). The original secondary (RNase A) digest of RNase T1 spot #11 (Figure 3 of reference 20) shows, in addition to the nucleotides expected from the oligonucleotides CUUG and UUCG, the dinucleotide AG as a minor product (data not shown). This dinucleotide would be a predicted product of the pancreatic RNase A digestion of the UUAG oligonucleotide obtained from the RNase T1 digestion of U1b, and its amount relative to the other expected nucleotides is roughly consistent with the level of U1b that we have observed here (one-tenth that of U1a).

DISCUSSION

We have cloned and characterized what we believe to be the complete set of U1 snRNA genes from *Drosophila melanogaster*. Four additional U1 genes not previously identified were isolated, including three from the 95C locus, and one of these was found to code for a previously undescribed U1 sequence variant (U1c). In all, a total of seven U1-related sequences have been isolated to date from the Oregon R strain of *Drosophila melanogaster*, which is in good agreement with the copy number determined by Saluz *et al.* (18) and by ourselves. Based on *in situ* hybridizations to salivary gland polytene chromosomes carried out by M.L. Pardue (unpublished), Saluz *et al.* (18), and Kejzlarova-Lepesant *et al.* (22), these seven U1-related sequences are located at only three cytological loci in the *Drosophila* genome: 21D, 82E, and 95C. We have found no evidence for the presence of other U1-related sequences at 11B or 61A, as initially reported by Alonso *et al.* (21). As was pointed out by McClelland (40), the probe used by Alonso *et al.* (21) contained the LSP1-β gene that is adjacent to the U1 gene at 21D. This gene is similar to the LSP-1α gene at 11A and the LSP1-γ gene at 61A, and it is these two loci that are detected by the probe utilized by Alonso *et al.* (21) for their *in situ* hybridizations. This erroneous assignment of U1 genes to 11A and 61A has been

reported in some subsequent surveys (41,42) and it is our hope that further references to these incorrect sites will be avoided.

All six full-length *Drosophila* U1 genes possess two highly conserved blocks of sequence in their 5′ flank that are also found in the upstream sequences of the *Drosophila* U2 (38) and U4 (39) genes. In addition, all of the U1 genes possess an absolutely conserved pentanucleotide, AAAGC, immediately preceding the presumed start of transcription, which is not seen in the U2 or U4 genes. This pentanucleotide is found at the corresponding position in the 5′ flank of some U1 genes from *P. polycephalum* (43), mouse (44,45), rat (46), and chicken (47,48). However, it has not been found in the U1 genes analyzed to date from tomato (49), soybean (50), *C. elegans* (51), sea urchin (13,52), *X. laevis* (53−55), and human (56,57). The significance of these blocks of 5′ sequence similarity is unclear since there have been no functional studies of the proximal sequence elements involved in U snRNA transcription in *Drosophila*. It is also not certain what relationship these conserved 5′ sequence elements in these *Drosophila* U snRNA genes have with the functionally defined upstream sequence elements of vertebrate spliceosomal U snRNA genes (for reviews see 19,58), such as the snRNA TATA-like box (or proximal sequence element) and the snRNA enhancer sequence (or distal sequence element). However, two strongly conserved sequence elements at −70/−80 and −35 relative to the *Arabidopsis thaliana* U2 and U5 genes have been shown to be required for transcription (59,60), and the two *Drosophila* 5′ sequence elements furthest upstream are located at roughly the same positions.

The seven U1-related sequences code for three possible U1 RNA sequences, of which two, U1a and U1b, are expressed *in vivo* and in *Drosophila* K_c cells at levels detectable by our primer extension assays. The U1c variant was not detected by our assay at any stage of development in *Drosophila*, but it remains a possibility that extremely low levels are expressed which are undetectable by our methods.

The sequence changes in U1b and U1c relative to U1a are not expected to greatly affect the potential secondary structure of *Drosophila* U1 snRNA (see Figure 2). The U to C transition at position 123 of U1c is located at the base of the closing stem of the U1 cloverleaf, replacing a G·U basepair with a stronger G-C basepair. This sequence change does not alter any of the highly conserved U1 sequence elements such as the 5′ end.

A comparison of sequence changes associated with embryonic U1's of the four species examined reveals no common base changes. The majority of changes detected in the mouse embryonic U1 molecules reside in stem-loop II, with some changes in loop I and stem-loop IV (17). In *Xenopus*, the sequence changes associated with the two known embryonic U1's occur in loop I or stem II (53,55). The various sequence changes which occur in the embryonic class of U1 snRNAs from these two species may be functionally significant since it is known that the highly conserved loop I and II sequences are required for the binding of the 70K (61,62) and A (63,64) protein, respectively.

The change found in sea urchin embryonic U1 is the only example from another species that may be related to the change in *Drosophila* U1b. It is a U to C transition located immediately upstream of the metazoan consensus Sm binding site ($RAU_{3-4}NUGR$; see 65) of the sea urchin U1 sequence. The G to U transversion in *Drosophila* U1b is located at the extreme 3′ end of the *Drosophila* Sm consensus sequence ($RAU_{5-6}GG$), and changes the second G, which is invariant in all known *Drosophila* U RNAs except U1b. We have derived this *Drosophila* consensus sequence from the Sm binding sites of *Drosophila* U1 (except for U1b), U2 (66,67), U4 (39), and U5 (68). Therefore, both the *Drosophila* and sea urchin embryonic U1 variants are associated with a base change immediately adjacent to a border of the Sm binding site of those molecules.

Since U1 snRNPs containing the U1b snRNA are immunoprecipitated by both anti-Sm and anti-(U1)RNP antisera, we conclude that there are no gross alterations in the structure of such particles. It is still possible that there are subtle structural changes in U1b-containing snRNPs which are not detected by immunoprecipitation and yet can cause functional differences. The sequence change in U1b could induce an allosteric alteration of the U1 snRNP. Another possibility is binding of variant, but antigenically related, Sm proteins to the U1b RNA.

Jones and Guthrie (69) have analyzed the effects of various mutations in the Sm binding site of the yeast U5 molecule on cell viability. One mutation that they generated in the yeast U5 Sm binding site, from AUUUUUUGG to AUUUUUUGU, mimics the alteration found in U1b. This mutation had a small effect on the viability of yeast cells in which it was expressed; a slight increase (1.2X) of cell doubling time was observed. This result suggests that while there is no drastic change in the function of the yeast U5 molecules which possess this variation in their Sm binding site, the *Drosophila* U1b sequence variant may create functionally distinct snRNP particles.

We have observed that there is one U1 sequence (U1a) which, while it is present throughout *Drosophila* development, is the predominant species during the later stages of development, i.e. pupation and adulthood. This pattern of abundance of U1 is therefore like that of the adult class of U1's seen in the other eukaryotic species that have exhibited differential patterns of abundance of U1 sequence variants. Similarly, the other U1 variant which is expressed, U1b, shows a pattern of abundance similar to that of the embryonic class of U1 sequences in previously described species since it shows its highest levels relative to U1a during early oogenesis and embryogenesis. It should be noted that while its relative amounts are significant, at no point does it appear that the level of U1b ever exceeds that of U1a in the whole organism.

One possible reason for the relatively higher abundance of a separate class of U1 variants during early development is that the U1 snRNPs which contain these variants are somehow altered with regard to their function in the splicing of pre-mRNAs. This alteration of function could conceivably result in alternative patterns of splicing of various primary transcripts, which would be one possible mechanism for the control of gene expression during early stages of development in *Drosophila*. On the other hand, the sequence changes that have been observed in all the embryonic U1 snRNAs of various species may be neutral with respect to the function of the U1 molecule, and their differential pattern of expression during development may simply be a consequence of differential regulation of functionally identical genes at the transcriptional or post-transcriptional level.

It appears that of the species of higher eukaryotes that have exhibited differential expression of an embryonic variant of U1, *Drosophila* is the best experimental organism in which to examine the functional significance of this phenomenon. It is the only one of the four species with differential regulation where all the U1 variants involved have been isolated and sequenced. In addition, since we have a number of strains (EC4, Canton S, and st^{85e}) that possess only one copy of a U1b gene, it should be possible through either classical or molecular genetic means to determine

the effect on *Drosophila* of deleting the sole U1b gene. In this manner it should be possible to ascertain if these embryonic variants have any functional significance.

ACKNOWLEDGEMENTS

We thank Dr. Jean-Antoine Lepesant for the subclones pDm6A, pDm6Y, and pDm6B, Mike Goldberg for the Oregon R genomic library, Zhen-Qiang Pan for the human anti-Sm serum, Ming Guo for the K_c nuclear extract, Ricardo Mancebo for poly A^- RNAs, and Elana Goltsman for performing the *in situ* hybridizations. We are also grateful to Carol Prives and Daniel Kalderon for critically reading this manuscript. We acknowledge Mary Lou Pardue and Joan Steitz for their contributions to unpublished results cited in this work.

This work was supported by Public Health Service grant GM 37991 from the National Institute of General Medical Sciences, by a National Science Foundation Presidential Young Investigator award to S.M.M., and by Basil O'Connor Starter Scholar research award 5-360 from the March of Dimes Birth Defects Foundation. P.C.H.L. was supported by NIH predoctoral training grant T32 GM07216.

REFERENCES

1. Padgett, R.A., Grabowski, P.J., Konarska, M.M., Seiler, S., and Sharp, P.A. (1986) *Ann. Rev. Biochem.* **55**, 1119-1150
2. Sharp, P.A. (1987) *Science* **235**, 766-771
3. Maniatis, T. and Reed, R. (1987) *Nature* **325**, 673-678
4. Steitz, J.A., Black, D.L., Gerke, V., Parker, K.A., Kramer, A., Frendewey, D., and Keller, W. (1988) In Birnstiel, M.L. (ed.), *Structure and Function of Major and Minor Small Ribonucleoprotein Particles*. Springer Verlag, Berlin, pp. 115-154
5. Luhrmann. R. (1988) In Birnstiel, M.L. (ed.), *Structure and Function of Major and Minor Small Ribonucleoprotein Particles*. Springer Verlag, Berlin, pp. 71-99
6. Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, A., and Steitz, J.A. (1983) *Cell* **33**, 509-518
7. Kramer, A., Keller, W., Appel, B., and Luhrmann, R. (1984) *Cell* **38**, 299-307
8. Black, D.L., Chabot, B., and Steitz, J.A. (1985) *Cell* **42**, 737-750
9. Zhuang, Y. and Weiner, A.M. (1986) *Cell* **46**, 827-835
10. Lerner, M.R. and Steitz, J.A. (1979) *Proc. Natl. Acad. Sci.* **76**, 5495-5499
11. Forbes, D.J., Kirschner, M.W., Caput, D., Dahlberg, J.E., and Lund, E. (1984) *Cell* **38**, 681-689
12. Kato, N. and Harada, F. (1985) *J. Biol. Chem.* **260**, 7775-7782
13. Yu, J.-C., Nash, M.A., Santiago, C., and Marzluff, W.F. (1986) *Nucleic Acids Res.* **24**, 9977-9988
14. Lund, E. (1988) *Nucleic Acids Res.* **16**, 5813-5826
15. Santiago, C. and Marzluff, W.F. (1989) *Proc. Natl. Acad. Sci.* **86**, 2572-2576
16. Lund, E., and Dahlberg, J.E. (1987) *Genes Dev.* **1**, 39-46
17. Lund, E., Kahan, B., and Dahlberg, J.E. (1985) *Science* **229**, 1271-1274
18. Saluz, H.P., Schmidt, T., Dudler, R., Altwegg, M., Stumm-Zollinger, E., Kubli, E., and Chen, P.S. (1983) *Nucleic Acids Res.* **11**, 77-90
19. Dahlberg, J.E. and Lund, E. (1988) In Birnstiel, M.L. (ed.), *Structure and Function of Major and Minor Small Ribonucleoprotein Particles*. Springer Verlag, Berlin, pp. 38-70
20. Mount, S.M., and Steitz, J.A. (1981) *Nucleic Acids Res.* **9**, 6351-6368
21. Alonso, A., Jorcano, J.L., Beck, E., Hovemann, B., and Schmidt, T. (1984) *J. Mol. Biol.* **180**, 825-836
22. Kejzlarova-Lepesant, J., Brock, H.W., Moreau, J., Dubertret, M.-L., Billault, A., and Lepesant, J.-A. (1984) *Nucleic Acids Res.* **12**, 8835-8846
23. Mount, S.M. (1983) Ph.D. thesis. Yale University, New Haven, Connecticut.
24. Maniatis, T., Fritsch, E.F., and Sambrook, J. (1982) *Molecular cloning: A Laboratory Manual.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
25. Ashburner, M. (1989) *Drosophila: A Laboratory Manual.*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp. 30, 44-49
26. Bankier, A.T. and Barrell, B.G. (1986) In *Techniques in Nucleic Acid Biochemistry*, B508. Techniques in the Life Sciences B5. Elsevier Scientific Publishers Ireland, County Clare, pp. 1-34
27. Sanger, F., Nicklen, S., and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci.* **74**, 5463-5467
28. Tabor, S. and Richardson, C.C. (1987) *Proc. Natl. Acad. Sci.* **84**, 4767-4771
29. McGinnis, W., Shermoen, A.W., and Beckendorf, S.K. (1983) *Cell* **34**, 75-84
30. Feinberg, A.P. and Vogelstein, B. (1983) *Anal. Biochem.* **132**, 6-13
31. Mount, S. M., Green, M.M., and Rubin, G.M. (1988) *Genetics* **118**, 221-234
32. Saiki, R.K., Gelfand, D.H., Stoffel, S., Scharf, S.J., Higuchi, R., Horn, G.T., Mullis, K.B., and Erlich, H.A. (1988) *Science* **239**, 487-491
33. O'Hare, K., Levis, R., and Rubin, G.M. (1983) *Proc. Natl. Acad. Sci.* **80**, 6917-6921
34. Dignam, J.D., Lebovitz, R.M., and Roeder, R.G. (1983) *Nucleic Acids Res.* **11**, 1475-1489
35. Fradin, A., Jove, R., Hemenway, C., Keiser, H.D., Manley, J.L., and Prives, C. (1984) *Cell* **37**, 927-936
36. Steitz, J.A. (1989) *Methods in Enzymology* **180**, 468-481
37. Lepesant, J.-A., Levine, M., Garen, A., Lepesant-Kejzlarova, J., Rat, L., and Somme-Martin, G. (1982) *J. Mol. Appl. Genet.* **1**, 371-383
38. Beck, E., Jorcano, J.L., and Alonso, A. (1984) *J. Mol. Biol.* **173**, 539-542
39. Saba, J.A., Busch, H., Wright, D., and Reddy, R. (1986) *J. Biol. Chem.* **261**, 8750-8753
40. MacClelland, A. (1985) *J. Mol. Biol.* **185**, 649
41. Kubli, E. (1987) In O'Brien, S. J. (ed.), *Genetic maps 1987.* Cold Spring Harbor Laboratory Press, Cold Spring Harbor, pp. 366-368
42. Saluz, H.P., Dudler, R., Schmidt, T., and Kubli, E. (1988) *Nucleic Acids Res.* **16**, 3582
43. Myslinski, E., Wilhelm, F.-X., and Branlant, C. (1989) *Nucleic Acids Res.* **17**, 1019-1034
44. Marzluff, W.F., Brown, D.T., Lobo, S., and Wang, S.-S. (1983) *Nucleic Acids Res.* **11**, 6255-6270
45. Howard, E.F., Michael, S.K., Dahlberg, J.E., and Lund, E. (1986) *Nucleic Acids Res.* **14**, 9811-9825
46. Watanabe-Nagasu, N., Itoh, Y., Tani, T., Okano, K., Koga, N., Okada, N., and Ohshima, Y. (1983) *Nucleic Acids Res.* **11**, 1791-1801
47. Roop, D.R., Kristo, P., Stumph, W.E., Tsai, M.J., and O'Malley, B.W. (1981) *Cell* **23**, 671-680
48. Earley, J.M., III, Roebuck, K.A., and Stumph, W.E. (1984) *Nucleic Acids Res.* **12**, 7411-7421
49. Abel, S., Kiss, T., and Solymosy, F. (1989) *Nucleic Acids Res.* **17**, 6319-6337
50. van Santen, V., Swain, W., and Spritz, R.A. (1988) *Nucleic Acids Res.* **16**, 4176
51. Thomas, J., Lea, K., Zucker-Aprison, E., and Blumenthal, T. (1990) *Nucleic Acids Res.* **18**, 2633-2642
52. Brown, D.T., Morris, G.F., Chodchoy, N., Sprecher, C., and Marzluff, W.F. (1985) *Nucleic Acids Res.* **13**, 537-556
53. Zeller, R., Carri, M.-T., Mattaj, I.W., and De Robertis, E.M. (1984) *EMBO J.* **3**, 1075-1081
54. Ciliberto, G., Buckland, R., Cortese, R., and Philipson, L. (1985) *EMBO J.* **4**, 1537-1543
55. Krol, A., Lund, E., and Dahlberg, J.E. (1985) *EMBO J.* **4**, 1529-1535
56. Manser, T. and Gesteland, R.F. (1982) *Cell* **29**, 257-264
57. Bernstein, L.B., Manser, T., and Weiner, A.W. (1985) *Mol. Cell. Biol.* **5**, 2159-2171
58. Parry, H.D., Scherly, D., and Mattaj, I.W. (1989) *Trends in Biochem. Sci.* **14**, 15-19
59. Vankan, P., Edoh, D., and Filipowicz, W. (1988) *Nucleic Acids Res.* **16**, 10425-10440
60. Vankan, P. and Filipowicz, W. (1989) *EMBO J.* **8**, 3875-3882
61. Query, C.C., Bentley, R.C., and Keene, J.D. (1989) *Mol. Cell. Biol.* **9**, 4872-4881
62. Surowy, C.S., van Santen, V.L., Scheib-Wixted, S.M., and Spritz, R.A. (1989) *Mol. Cell. Biol.* **9**, 4179-4186
63. Lutz-Freyermuth, C. and Keene, J. (1989) *Mol. Cell. Biol.* **9**, 2975-2982
64. Scherly, D., Boelens, W., van Venrooij, W.J., Dathan, N.A., Hamm, J., and Mattaj, I.W. (1989) *EMBO J.* **8**, 4163-4170
65. Guthrie, C. and Patterson, B. (1988) *Ann. Rev. Genet.* **22**, 387-419
66. Alonso, A., Jorcano, J.L., Beck, E., and Speiss, E. (1983) *J. Mol. Biol.* **169**, 691-705
67. Alonso, A., Beck, E., Jorcano, J.L., and Hovemann, B. (1984) *Nucleic Acids Res.* **12**, 9543-9550
68. Myslinski, E., Branlant, C., Wieben, E.D., and Pederson, T. (1984) *J. Mol. Biol.* **180**, 927-945
69. Jones, M.H. and Guthrie, C. (1990) *EMBO J.* **9**, 2555-2561