Progress in Molecular and Subcellular Biology 12

Edited by
P. Jeanteur Y. Kuchino
W. E. G. Müller (Managing Editor)
P. L. Paine

Springer-Verlag

Progress
12 in Molecular and
Subcellular Biology

Edited by
P. Jeanteur, Y. Kuchino
W.E.G. Müller (Managing Editor)
P.L. Paine

With 20 Figures

Springer-Verlag
Berlin Heidelberg New York
London Paris Tokyo
Hong Kong Barcelona
Budapest

Contents

Synthesis of Small Nuclear RNAs
R. Reddy and R. Singh

Developmental Regulations of Heat-Shock Protein Synthesis in Unstressed and Stressed Cells

O. BENSAUDE, V. MEZGER, and M. MORANGE

The Interactions of Water and Proteins in Cellular Function

J.G. WATTERSON

Synthesis of Small Nuclear RNAs

R. REDDY and R. SINGH[1]

1 Introduction

There are seven abundant and several less abundant capped small nuclear RNAs characterized in mammalian cells. These RNAs are all capped on their 5′ ends and were designated U snRNAs because the U1-U3 snRNAs initially studied were rich in uridylic acid (Hodnett and Busch 1968). These capped snRNAs play important roles in the processing of nuclear precursor mRNAs and precursor rRNAs (reviewed in Busch et al. 1982; Brunel et al. 1985; Green 1986; Padgett et al. 1986; Maniatis and Reed 1987; Guthrie and Patterson 1988; Steitz 1988; Steitz et al. 1988; Zieve and Sauterer 1990). The functions of the U snRNAs are summarized in Table 1. While the roles of U snRNAs in the processing of eukaryotic precursor RNAs are well established, U5 snRNA was recently shown to have the potential to transform cells in vitro (Hamada et al. 1989), suggesting multiple roles for the U snRNAs. Each HeLa cell contains a total of approximately 2–3 million copies of U snRNAs (Weinberg and Penman 1968), and it is estimated that each of the U1 and U2 snRNA genes is transcribed every 2–4 s, generating the large amounts of U snRNAs found in mammalian cells (Skuzeski et al. 1984; Mangin et al. 1986; reviewed in Dahlberg and Lund 1988); hence, the snRNA genes have very strong promoters compared to many other cellular genes.

2 Two Classes of U snRNA Genes

Based on the type of cap structure present on their 5′ ends, U snRNAs are divided into two classes. The trimethylguanosine (TMG) cap-containing U snRNAs include U1-U5 and U7-U14 snRNAs; and the methyl (mepppG) cap-containing U snRNAs include U6 and 7SK RNAs. These two cap structures are shown in Fig. 1. Although many small nuclear RNAs are capped, not all small RNAs in the nucleus contain cap structures. For instance, human RNaseP (H1) RNA (Baer et al. 1990) and 7SM (7–2/MRP) RNA (Hashimoto and Steitz 1983; Yuan et al. 1989) do not contain cap structures.

The first U snRNA gene to be isolated and characterized was the U3 snRNA gene from slime mold *Dictyostelium* (Wise and Weiner 1980). To date, approxi-

[1]Department of Pharmacology, Baylor College of Medicine, Houston, TX 77030, USA

Table 1. Functions of U snRNAs

RNA	Function	Reference
U1	Splicing of nuclear pre-mRNAs; binds specifically to the 5′ splice site	Mount et al. (1981); Steitz et al. (1988); Zhuang and Weiner (1986); Guthrie and Patterson (1988)
U2	Splicing of nuclear pre-mRNAs; binds specifically to the branch point	Parker et al. (1987); Steitz et al. (1988)
U3	Processing of pre-ribosomal RNA; binds specifically near the 5′ end of 45S RNA	Kass et al. (1990); Maser and Calvet (1989); Stroke and Weiner (1989)
U4	Splicing of nuclear pre-mRNAs; binds specifically to U6 snRNA	Berget and Robberson (1986); Black and Steitz (1986); Steitz et al. (1988)
U5	Splicing of nuclear pre-mRNAs; binds specifically to the 3′ splice site	Chabot et al. (1985); Tazi et al. (1986; Steitz et al. (1988)
U6	Splicing of nuclear pre-mRNAs; binds specifically to U4 snRNA	Berget and Robberson (1986); Black and Steitz (1986); Brown and Guthrie (1988); Steitz et al. (1988)
U7	3′-end formation of histone pre-mRNAs	Schaufele et al. (1986); Mowry and Steitz (1987)
U8	Not known; probably involved in pre-rRNA processing	Reddy et al. (1985); Tyc and Steitz (1989)
U11	Not known; probably involved in polyadenylation	Christofori and Keller (1988); Montzka and Steitz (1988)
U12	Not known; probably involved in nuclear pre-mRNA processing	Montzka and Steitz (1988)
U13	Not known; probably involved in pre-rRNA processing	Tyc and Steitz (1989)
U14	Disrupts production of 18S rRNA; probably involved in pre-rRNA processing	Li et al. (1990)
7SK	Not known; probably involved in nuclear pre-mRNA processing	Gupta et al. (1990b)

mately 100 snRNA genes from diverse species have been isolated and characterized. These include genes from human, rat, mouse, chicken, frog, *Drosophila*, sea urchin, *Trypanosome, C. elegans*, slime mold, yeast, and several plants. The U snRNA genes encoding the TMG-capped U1-U5 and U7-U14 snRNAs have many common features, as summarized in Table 2. U6 and 7SK RNA genes belong to another class, and they differ from the other U snRNA genes in several important respects; the main difference is that these two genes are transcribed by RNA polymerase III (pol III) in contrast to the transcription of TMG-capped snRNAs by RNA polymerase II (pol II).

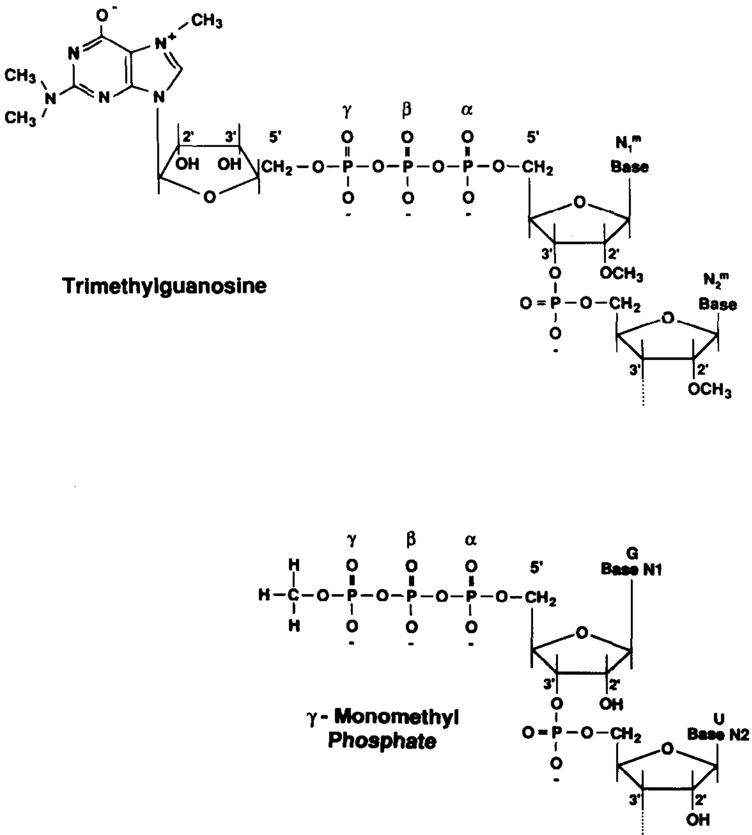

Fig. 1. Cap structures in U snRNAs. *Top*: TMG-cap structure found in U1–U5 and U7–U14 snRNAs. The 2′-O-methylations occur only in higher eukaryotes, such as rat and HeLa cells, but not in amoeba or dinoflagellates. *Bottom*: MepppG-cap structure found in U6 and 7SK snRNAs. Diagrammatic representation of the methylated γ-phosphate of the 5′ nucleotide G (nucleotide N1) of human U6 and 7SK snRNA. The 2′,3′, and 5′ represent the carbon moieties of the ribose sugar

3 RNA Polymerases Transcribing U snRNA Genes

3.1 TMG-Capped snRNAs

Several lines of evidence suggest that pol II is responsible for the synthesis of TMG-capped U snRNAs. (1) The synthesis of these RNAs is inhibited by low concentrations of α-amanitin in whole animals (Ro-Choi et al. 1976), cultured cells (Frederiksen et al. 1978; Chandrasekharappa et al. 1983), isolated nuclei (Roop et al. 1981; Lobo and Marzluff 1987), cell-free extracts (Morris et al. 1986; Lund and Dahlberg 1989; Southgate and Busslinger 1989), and frog oocytes (Murphy et al. 1982; Mattaj and Zeller 1983; Skuzeski et al. 1984; Reddy et al. 1987). (2) Gram-Jensen et al. (1979) used a cell line containing an altered pol II which is 800 times

Table 2. Two types of U snRNA genes

Characteristic	TMG-capped	mepppG-capped
Examples	U1-U5, U7-U14	U6, 7SK
5′-end cap	TMGpppA/G	CH_3-O-pppG/A
Synthesized by	Polymerase II (B)	Polymerase III (C)
Transcription factors	Share with mRNA genes	Share with mRNA genes
3′-end formation requires	3′ Box and a compatible snRNA promoter	T-stretch
Introns	Not present (with one exception)	Not present (with one exception)
Initiation nucleotide	Purine	Purine
PSE at –45 to –70[a]	Present and required	Present and required
DSE around –250[a]	Functions as enhancer	Functions as enhancer
Intragenic promoter	None	None

[a]PSE and DSE stand for Proximal and Distal Sequence Element, respectively. Mammalian U snRNA genes have sequences downstream of –50 that are important for PSE function (Murphy et al. 1987a; reviewed in Dahlberg and Lund 1988); U6 and 7SK snRNA genes, as well as plant U snRNA genes, contain an essential TATA-motif at –30 region (reviewed in Geiduschek and Tocchini-Valentini 1988; also see Sect. 5.3).

more resistant towards inhibition by α-amanitin than the wild-type enzyme. In these cells, the synthesis of U1, U2, and U3 snRNAs was not inhibited by high concentrations of α-amanitin. Furthermore, the synthesis of these U snRNAs is inhibited at nonpermissive temperature in the cells that contain a temperature-sensitive pol II (Hellung-Larsen et al. 1980). (3) The synthesis of U1 and U2 snRNAs is sensitive to 5,6-dichloro-1-β-D-ribofuranosyl benzimidazole, which is a specific inhibitor of transcription by pol II at low concentrations (Hellung-Larsen et al. 1981) and the primary transcripts of U1 snRNA, like mRNAs, are co-transcriptionally capped with m^7G (Eliceiri 1980; Skuzeski et al. 1984; Mattaj 1986). (4) Antibodies against the large subunit of the pol II inhibit the synthesis of U1 snRNA in the frog oocytes (Thompson et al. 1989). (5) Finally, Pol III is unlikely to be involved in the synthesis of TMG-capped U snRNAs because a U cluster (AUUUUUG as Sm antigen-binding site) is present within the transcribed portion of a large number of these genes and this signal results in termination of pol III-mediated transcription. All these data show that TMG-capped snRNAs are synthesized by pol II or by an RNA polymerase closely related to pol II. Although studies have been carried out on the synthesis of only some TMG-capped snRNAs, it is likely that other TMG-capped snRNAs are also synthesized by pol II.

3.2 MepppG-Capped snRNAs

There is much evidence to support the involvement of pol III in the synthesis of U6 and 7SK RNAs. (1) Low concentrations of α-amanitin, sufficient to inhibit the synthesis of mRNAs and TMG-capped U snRNAs, had no inhibitory effect on the synthesis of U6 RNA in frog oocytes, or in vitro (Kunkel et al. 1986; Reddy et al. 1987; Krol et al. 1987), or in isolated nuclei (Kunkel et al. 1986). (2) In U6 snRNA genes, the signal for transcription termination is a T-cluster (Das et al. 1988) similar to the functional termination signal in 5S RNA gene (Bogenhagen and Brown 1981). (3) The transcription of U6 snRNA is competed by other pol III genes like 5S and tRNA genes both in vitro (Reddy et al. 1987) and in vivo (Carbon et al. 1987). (4) The U6 snRNA associates with La antigen (Rinke and Steitz 1985; Reddy et al. 1987) which may be a pol III transcription termination factor (Gottlieb and Steitz 1989). (5) A mutant yeast strain with temperature-sensitive defect in the large subunit of pol III, which results in defective transcription of tRNA and 5S RNA genes, was also defective in U6 snRNA transcription (Moenne et al. 1990). (6) Tagetitoxin, a specific inhibitor of transcription by pol III at low concentrations, inhibits the synthesis of U6 snRNA (Steinberg et al. 1990). All these data show that U6 snRNA genes are transcribed by pol III. Although the involvement of pol III in the synthesis of 7SK RNA is well documented (Zieve et al. 1977; Murph et al. 1986, 1987b; Kruger and Benecke 1987), the mepppG cap structure in 7SK RNA was only recently identified (Gupta et al. 1990b). It is significant to note that U3 snRNAs from tomato, pea and *Arabidopsis* do not contain the TMG cap structure and may contain the mepppA cap structure (Kiss and Solymosy 1990), and that U3 RNA in *Arabidopsis* is synthesized by pol III and not by pol II (Waibel et al. 1990). These observations are consistent with the notion that TMG-capped U snRNAs are pol II products and mepppG/A capped snRNAs are pol III products. All other known small nuclear RNAs, including RNaseP (Baer et al. 1990), add, MRP/7–2 (Hashimoto and Steitz 1983), and Alu-related B1, B2, and B3 RNAs (reviewed in Jelinek and Schmidt 1982), are synthesized by pol III.

4 Organization of U snRNA Genes

The organization and copy numbers of U snRNA genes vary from organism to organism. In general, the gene copy number correlates well with the abundance of each U snRNA; however, the copy number varies widely. In higher eukaryotes, most U snRNAs are represented by a multigene family. In lower eukaryotes, such as yeasts, most U snRNAs are coded for by single copy genes. The gene organization and copy number of U snRNA genes in different organisms is summarized in Table 3.

4.1 Human

Real genes have been characterized from the human genome for U1 (Manser and Gesteland 1981, 1982; Lund and Dahlberg 1984), U2 (VanArsdell and Weiner 1984; Westin et al. 1984), U3 (Suh et al. 1986; Yuan and Reddy 1988), U4 (Bark et al. 1986), and U6 snRNAs (Kunkel et al. 1986). Most, and perhaps all, of the human U1 snRNA genes are present as a tandem repeat on the short arm of the

Table 3. Copy number, organization and localization of U snRNA genes. (Slightly modified and updated from Dahlberg and Lund 1988)

Organism	RNA	Copy #[a]	Cloned genes	Organization	Location	Reference
Human	U1	~30	HSD1–7 HU1–1 cosD1, cosD21	Loosely clustered >44 kb apart, large tandem repeat unit	1p36	Manser and Gesteland (1981, 1982) Buckland et al. (1983) Lund and Dahlberg (1984) Bernstein et al. (1985)
	U2	10–20	U2.24A,B U2/6	Tightly clustered in one tandem array; 6.2 kb repeat unit	17q21 q22	Van Arsdell and Weiner (1984) Westin et al. (1984)
	U3	7–10	U3–1–4	If clustered, >10 kb apart	ND	Suh et al. (1986) Yuan and Reddy (1989)
	U4	100	U4C and U4B-like	Cloned genes tightly clustered	ND	Bark et al. (1986)
	U6	ND	HU6	If clustered, >10 kb apart	ND	Kunkel et al. (1986)
Rat	U1	~50	6–6A,B	Cloned genes 3.6 kb apart, in opposite orientation	ND	Watanabe-Nagasu et al. (1983)
	U2	40	RU2–3	If clustered, >10 kb apart	ND	Tani et al. (1983)
	U3	5–10	U3D, U3B.4,7 each type: 1-few	If clustered, >10 kb apart	ND	Stroke and Weiner (1985)
Mouse	U1	20–40 each type: 5–10	U1.1,2 (U1b2) U1a-236 (U1a1) U1b-136 (U1b2) U1b-453,550 (U1b6)	Inverted repeat, U1.1 and U1.2 5.0 kb apart Genes for each type loosely clustered,, more than 5–10 kb apart	3(U1b2,b3) 11 (U1a1) 12 (U1a2)	Marzluff et al. (1983) Blatt et al. (1988) Michael et al. (1986) Howard et al. (1986) Lund and Nesbitt (1988)
	U2	~10	U2 U2.47	One locus contains inverted repeat with 2 genes, 3.8 kb apart; Another locus one gene	ND	Nojima and Kornberg (1983) Moshier et al. (1987, 1988)

	U3	6–7		Two genes are 5kb apart	ND	Mazan and Bachelleri (1988)
	U6	2	U6–52	If clustered, >9 kb apart	ND	Ohshima et al. (1981) Yuan and Reddy (1988)
Chicken	U1	6–10	U1.2.5 U1–52a,b,c	3 genes within 5kb; 1.8 kb apart	ND	Roop et al. (1981) Early et al. (1984)
	U2	35–40	U2–6	Tightly clustered in tandem array, 5.35 kb apart	ND	Korf and Stumph (1988)
	U4	2	U4B,U4X	Cloned genes 465 bp apart	ND	Hoffman et al. (1986)
Xenopus	U1	Minor family (adult) ~50 Major family (embryonic) ~1000	X1U1.3 X1U1.8 X1U1b1, X1U1b2	Cloned locus has 3 genes within 5 kb Tightly clustered in large tandem array(s), 1.85 kb repeat unit with 1 copy each of the b1 and b2 genes	ND	Zeller et al. (1984) Mattaj and Zeller (1983) Lund et al. (1984) Krol et al. (1985) Ciliberto et al. (1985)
	U2	~500–1000	XLU2–5	Tightly clustered in large tandem array, 830 bp repeat unit	ND	Mattaj and Zeller (1983)
	U5	~100	XlU511H	Major family is a tandem repeat; 583 bp apart	ND	Kazmaier et al. (1987)
	U6	~600	XtU6–2	Tightly clustered in large tandem arrays, 1.6 and 1 kb repeat units		Krol et al. (1987)
Drosophila	U1	7	U14 DmU1.4 Dm6A	Dispersed	11B,21E 61A,82E 95C	Alonso et al. (1984b) Mount and Steitz (1981) Kejzlarova-Lepesant et al. (1984); Saluz et al. (1983, 1988)

Table 3. *(Continued)*

Organism	RNA	Copy #[a]	Cloned genes	Organization	Location	Reference
	U2	4–5	U2 131A, 131B U2 141A, 141B	Two unlinked clusters, each containing 2 genes within 3–5 kb	34BC 84C	Alonson et al. (1983, 1984a)
	U3	ND	DU3–1		ND	Akao et al. (1986)
	U4	3	U4–1, U4–2	Dispersed	39B, 40AB	Saba et al. (1986); Saluz et al. (1988)
	U5	7		Dispersed	14B,23D, 34A,35EF, 39B,63A	Saluz et al. (1988)
	U6	3	DU6–1, 6–2, 6–3	Closely linked, 500 bp apart, same orientation	96A	Das et al. (1987); Saluz et al. (1988)
Sea urchin	U1	20	LvU1.1 LvU1.2	Tightly clustered in large tandem arrays, 1.4 kb repeat unit; 2 types of repeat	ND	Brown et al. (1985) Nash and Marzluff (1988) Card et al. (1982) Yu et al. (1986)
	U2	20		1.1 kb repeat unit	ND	Card et al. (1982)
	U7	5	U7	5 genes within 9.3 kb	ND	Lorenzi et al. (1986)
Dictyostelium	U3	5	D2.1	Dispersed	ND	Wise and Weiner (1980)
Trypanosome	U2	1	U2		ND	Tsuchidi et al. (1986);
(T. brucei)	U4	1	U4		ND	Mottram et al. (1989)
	U6	1	U6		ND	Mottram et al. (1989)
C. elegans	U1	~11		Clusters of 2–3 genes or single copies dispersed in the genome	ND	Thomas et al. (1990)
	U2	~12			ND	Thomas et al. (1990)
	U4	~6			ND	Thomas et al. (1990)
	U5	~9			ND	Thomas et al. (1990)
	U6	~10			ND	Thomas et al. (1990)

Phaseolus	U1	1-few	U1	If more than one >14 kb apart	ND	van Santen and Spritz (1987)
Soybean	U1	2–4	U1a,b	Dispersed and tandem repeats		van Santen et al. (1988)
Tomato	U1	8 or more	U1.1 to U1.8	In 5 loci		Abel et al. (1989)
Arabidopsis	U2	10–15	U2.1 to U2.9	Do not appear to be clustered	ND	Vankan and Filipowicz (1988)
Maize	U2	25–40	U2–27			Brown and Waugh (1989)
Pea	U2	Many				Hanley and Schuler (1989)
Tomato	U3	1-few	U3	Genes and pseudogenes linked	ND	Kiss and Solymosy (1990)
Arabidopsis	U5	8–9	U5	Some gene linked	ND	Vankan et al. (1988)
Arabidopsis	U6		U6		ND	Waibel et al. (1990)
Tomato	U6		U6		ND	Szkukalek et al. (1990)
Yeast	U2	1	pMa2		ND	Brennwald et al. (1988)
(*S. pombe*)	U3	2		If linked, >8 kb apart	ND	Porter et al. (1988)
	U6	1			ND	Tani and Ohshima (1989)
Yeast (*S. cerevisiae*)	U1	1	SNR 19		ND	Siliciano et al. (1987a); Kretzner et al. (1987)
	U2	1	SNR20		ND	Ares (1986)
	U3	2	SNR17a,b	Genetically unlinked	ND	Hughes et al. (1987); Myslinski et al. (1990)
	U4	1	SNR3, 14		ND	Siliciano et al. (1987b)
	U5	1	SNR7		ND	Patterson and Guthrie (1987)
	U6	1	SNR6		ND	Brow and Guthrie (1988)
	U14	1	SNR128	67 bp apart from SNR190	ND	Zagorski et al. (1988)
	SNR3,4,5 8,9,10	1 each	SNR3,4,5,7, 8,9, 1sr1 (snR20)	Dispersed	ND	Tollervey et al. (1983); Wise et al. (1983) Parker et al. (1988)
	SNR 190	1		67 bp apart from U14	ND	Zagorski et al. (1988)

[a]Copy # indicates the number of genes per haploid genome.

chromosome 1 (Lund et al. 1983). This result, obtained from human/rodent hybrid cells, was confirmed by in situ hybridization; these studies showed that U1 genes are clustered at 1p36 (Naylor et al. 1984; Lindgren et al. 1985b). The human U2 genes have been localized in 17q21–22 region (Lindgren et al. 1985a). Interestingly, the locations of U1 and U2 genes correspond to viral chromosome modification sites (Lindgren et al. 1985a). The flanking regions of up to 20 kb on either side of human U snRNA genes are highly conserved (e.g., Manser and Gesteland 1982; Bernstein et al. 1985). Since the pseudogenes for U snRNAs are abundant in the human genome (Denison et al. 1981), these conserved flanking regions were successfully used to estimate the true gene copy number in the midst of abundant pseudogenes (Lund and Dahlberg 1984).

4.2 Rodent

Rat U1 (Watanabe-Nagasu et al. 1983), U2 (Tani et al. 1983), U3 (Stroke and Weiner 1985), mouse U1 (Marzluff et al. 1983; Howard et al. 1986; Michael et al. 1986), U2 (Nojima and Kornberg 1983; Moshier et al. 1987), U3 (Mazan and Bachellerie 1988), and U6 (Ohshima et al. 1981) snRNA genes have been cloned and characterized. As in the case of human U snRNA genes, the flanking regions within each rodent U snRNA gene repeat are highly conserved; however, the 5′ flanking regions of different types of mouse U1 snRNA genes differ widely (Howard et al. 1986). Interestingly, the 5′ flanking sequences in the rat and mouse U1 gene repeats are the same (Moussa et al. 1987); similarly, the 5′ flanking sequences of the rat and mouse U3 snRNA genes are also conserved (Mazan and Bachellerie 1988). However, these flanking sequences differ from the 5′ flanking sequences in the corresponding human U1 or U3 snRNA genes. These data provide evidence for the conservation of snRNA gene repeats in closely related species.

4.3 Chicken

Genes for chicken U1 (Roop et al. 1981; Earley et al. 1984), U2 (Korf and Stumph 1986), and U4 (Hoffman et al. 1986) snRNAs have been isolated and characterized. The chicken U1 and U2 snRNA genes, are present as tandem repeats as in human and rodent genomes. The U1 and U2 snRNA genes are found in very different genomic environments but have similar promoter structures (Korf and Stumph 1986).

4.4 Xenopus

Frog U1 (Zeller et al. 1984; Lund et al. 1984; Krol et al. 1985; Ciliberto et al. 1985), U2 (Mattaj and Zeller 1983); U5 (Kazmaier et al. 1987), and U6 (Krol et al. 1987) snRNA genes have been characterized. The most unusual feature about the *Xenopus*

U snRNA genes is that the gene copy number is very high. More than 1000 copies each of U1 and U2 snRNA genes and about 600 copies of U6 snRNA genes are present in the *Xenopus* haploid genome. The abundance of these genes is similar to the observations made with 5S genes, of which over 20 000 copies are present in the *Xenopus* genome (reviewed in Long and Dawid 1980).

4.5 Drosophila

Drosophila U1 (Mount and Steitz 1981; Alonso et al. 1984b), U2 (Alonso et al. 1983; 1984a), U3 (Akao et al. 1986), U4 (Saba et al. 1986), and U6 (Das et al. 1987) snRNA genes have been characterized. Most of the U snRNA genes in the *Drosophila* genome have been mapped to particular chromosomal loci (Saluz et al. 1988). The genes for *Drosophila* U1, U4, and U5 snRNAs are dispersed, whereas genes for U2 snRNA are in two unlinked clusters. All the three genes for U6 snRNA are closely linked. While flanking sequences in the human U1 gene family are highly conserved, the flanking sequences of *Drosophila* U gene families are not well conserved (Alonso et al. 1984a; Das et al. 1987).

4.6 Sea Urchin

Several genes for U1 (Brown et al. 1985; Yu et al. 1986; Nash and Marzluff 1988), U2 (Card et al. 1982), and U7 (Lorenzi et al. 1986) have been isolated and characterized. There are multiple copies of genes for each U snRNA and these are tightly clustered. There are two types of U1 gene repeats and both types are transcribed in sea urchin embryos (Yu et al. 1986). Although U snRNA genes from human, rat and mouse are transcribed accurately in frog oocytes (see Sect. 6.2), sea urchin U7 snRNA genes are not expressed faithfully in frog oocytes (Strub and Birnstiel 1986)

4.7 Trypanosome

There has been interest in the structure of U snRNAs from *Trypanosomes* because of the *trans*-splicing and RNA editing that are common in these parasites (Simpson and Shaw 1989). Although several (U2, U4, and U6) capped snRNAs are found in *Trypanosomes* and are required for *trans*-splicing (Tsuchidi and Ullu 1990), they differ significantly from the metazoan U snRNAs (Tsuchidi et al. 1986; Mottram et al. 1989). For example, the Sm-binding site found in metazoan, yeast and plant U2 snRNAs is not present in *Trypanosomal* U2 snRNA. The analog for U1 snRNA has not yet been identified, and if U1 snRNA is even present in *Trypanosomes*, it appears to be a minor RNA or it lacks the TMG cap structure (Mottram et al. 1989). The spliced leader sequences contain TMG cap structure, associate with Sm-antigen, and serve as U1 snRNPs during the *trans*-splicing event (Bruznik et al. 1988; Thomas et al. 1988).

4.8 C. elegans

Nematodes are the only group of organisms in which both *cis*- and *trans*-splicing of nuclear mRNAs are known to occur. The genes for U1, U2, U4, U5, and U6 snRNAs from *C. elegans* have been isolated and characterized. The genes for each U snRNA is represented by a mutigene family and are dispersed randomly in the genome of *C. elegans* (Thomas et al. 1990).

4.9 Yeast

Many U snRNA genes, including U1 (Siliciano et al. 1987a; Kretzner et al. 1987), U2 (Ares 1986), U4 (Siliciano et al. 1987b), U5 (Patterson and Guthrie 1987), and U6 snRNA (Brow and Guthrie 1988), have been isolated from *S. cerevisiae*. With the exception of U3 snRNA genes (Hughes et al. 1987; Myslinski et al. 1990), all U snRNA genes that have been characterized from yeasts are single copy genes and are dispersed in the yeast genome. The U3 snRNA genes from some strains of *S. cerevisiae* contain introns (Myslinski et al. 1990). Genes for U2 (Brennwald et al. 1988). U3 (Porter et al. 1988), and U6 (Tani and Ohshima 1989) snRNAs have also been isolated from *S. pombe*. Interestingly, U6 snRNA genes in *S. pombe* (Tani and Ohshima 1989) and in several related fungi (Frendeway et al. 1990; Reich and Wise 1990) contain introns which resemble the introns found in yeast mRNAs. No pseudogenes have been reported for U snRNAs in the yeast genomes.

4.10 Plants

Genes coding for *Arabidopsis* U2 (Vankan and Filipowicz 1988), U5 (Vankan et al. 1988), and U6 (Waibel et al. 1990), bean U1 (van Santen and Spritz 1987; van Santen et al. 1988), tomato U1 (Abel et al. 1989), tomato U3 (Kiss and Solymosy 1990), U6 (Szkukalek et al. 1990), and maize U2 (Brown and Waugh 1989) snRNAs have been isolated and characterized. The plant U snRNA genes characterized thus far are represented by multigene families and are not closely clustered. Most of the genes that have been isolated were shown to be real genes by the expression of the cloned genes into electroporated plant protoplasts. Several pseudogenes for the plant snRNAs have been reported (e.g., bean U1 and tomato U3 snRNA pseudogenes); however, the pseudogenes do not appear to be as abundant as is the case in mammalian genomes.

4.11 Viral U RNAs

Recently, Lee et al. (1988) discovered that herpesvirus saimiri codes for at least five TMG-capped U snRNAs. These RNAs, in association with Sm antigen, are present as snRNP particles. All the herpesvirus U snRNA genes contain the consensus