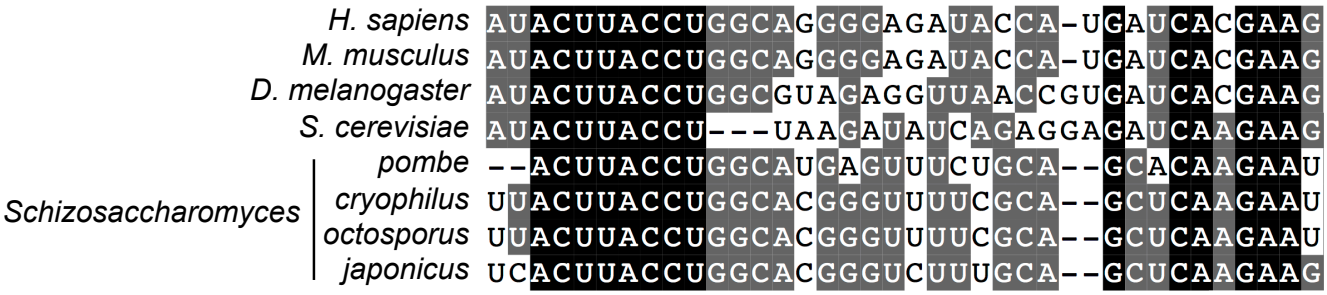
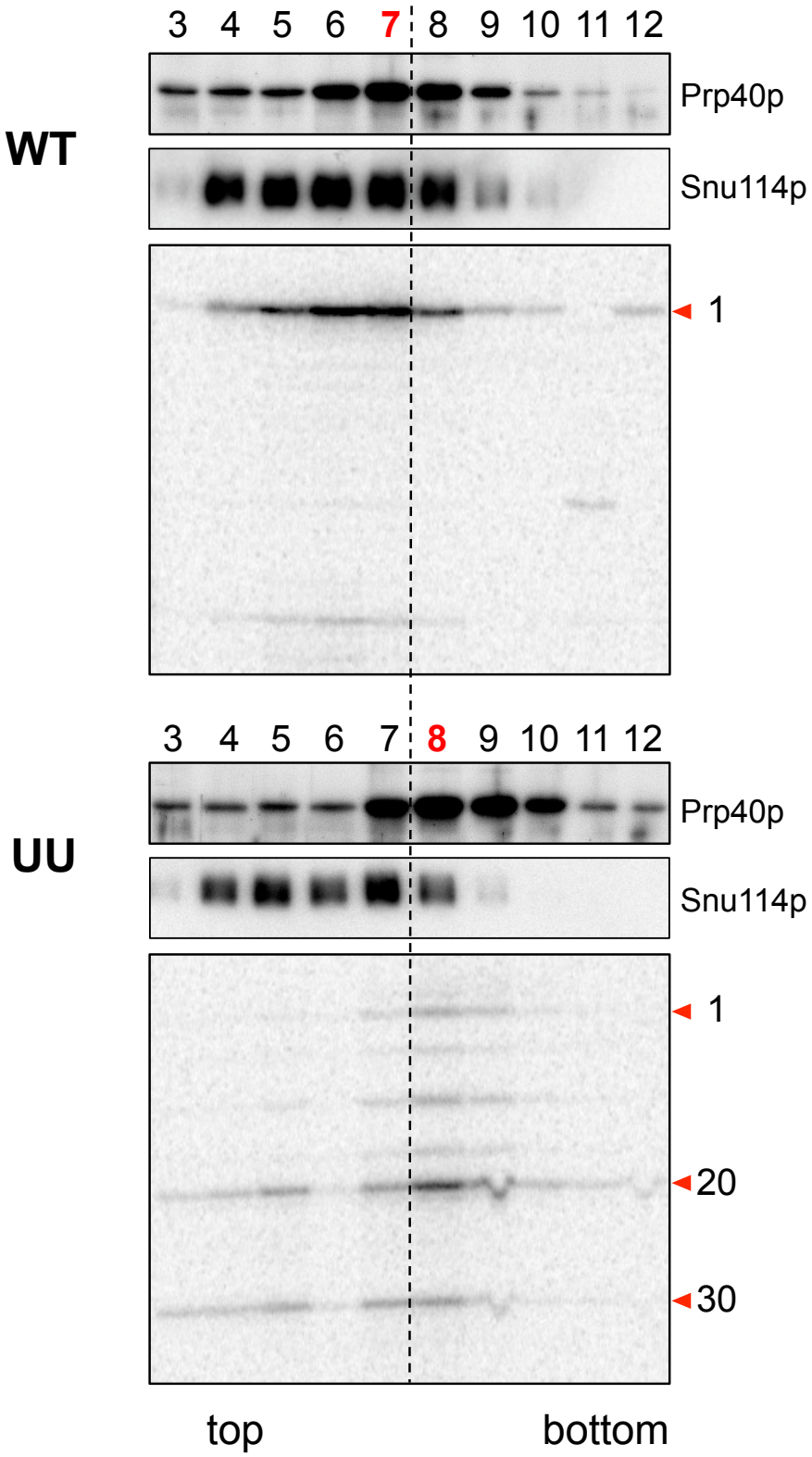
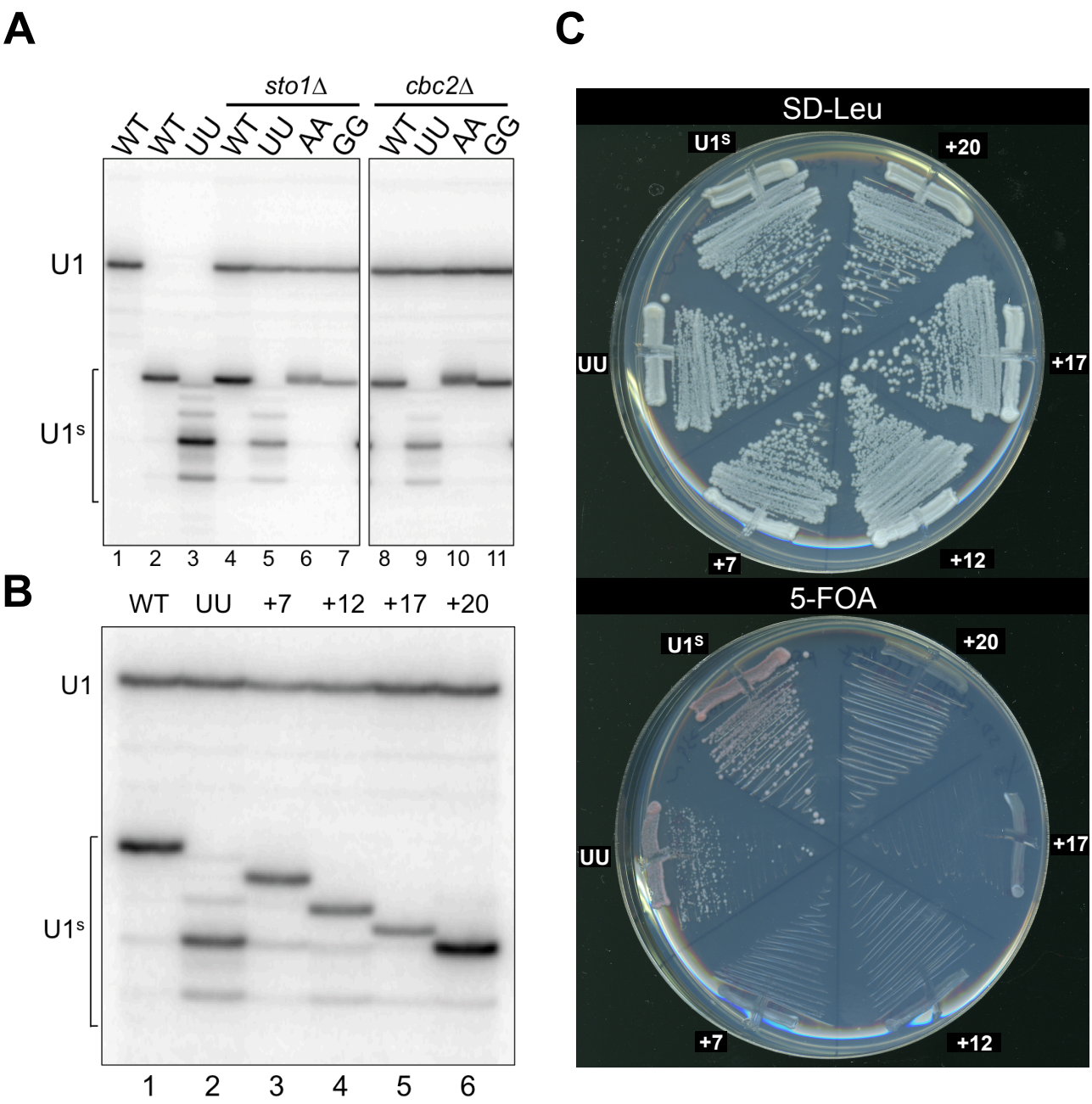


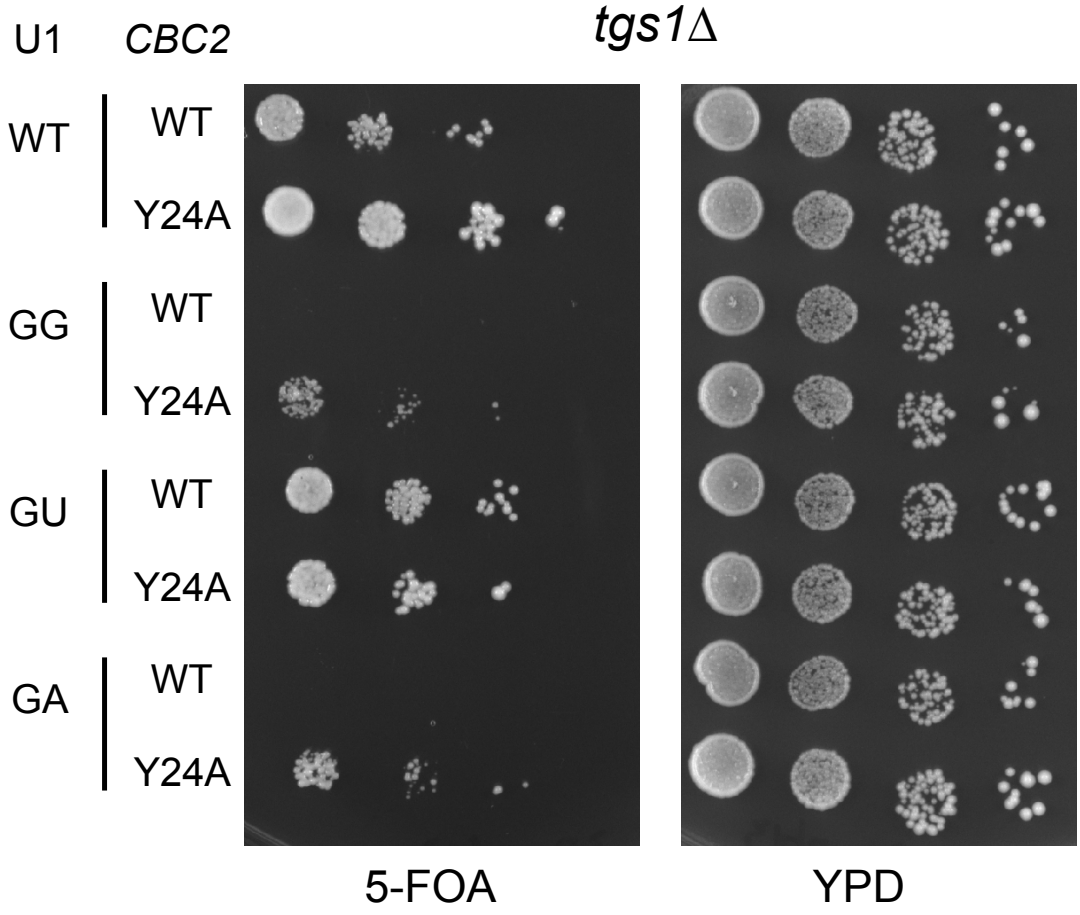
Supplementary Figure S1. Sequences of the 5'-end dinucleotide of the U1 snRNAs in four *Schizosaccharomyces* species are distinct from the highly conserved AU dinucleotide. U1 snRNA sequence of *Schizosaccharomyces* species were downloaded from EnsemblFungi. *Sp*: SPSNRNA.01; *Sc*: EFSCRG00000000143; *So*: EFSOCG00000000255 and *Sj*: EFSJAG00000000237.



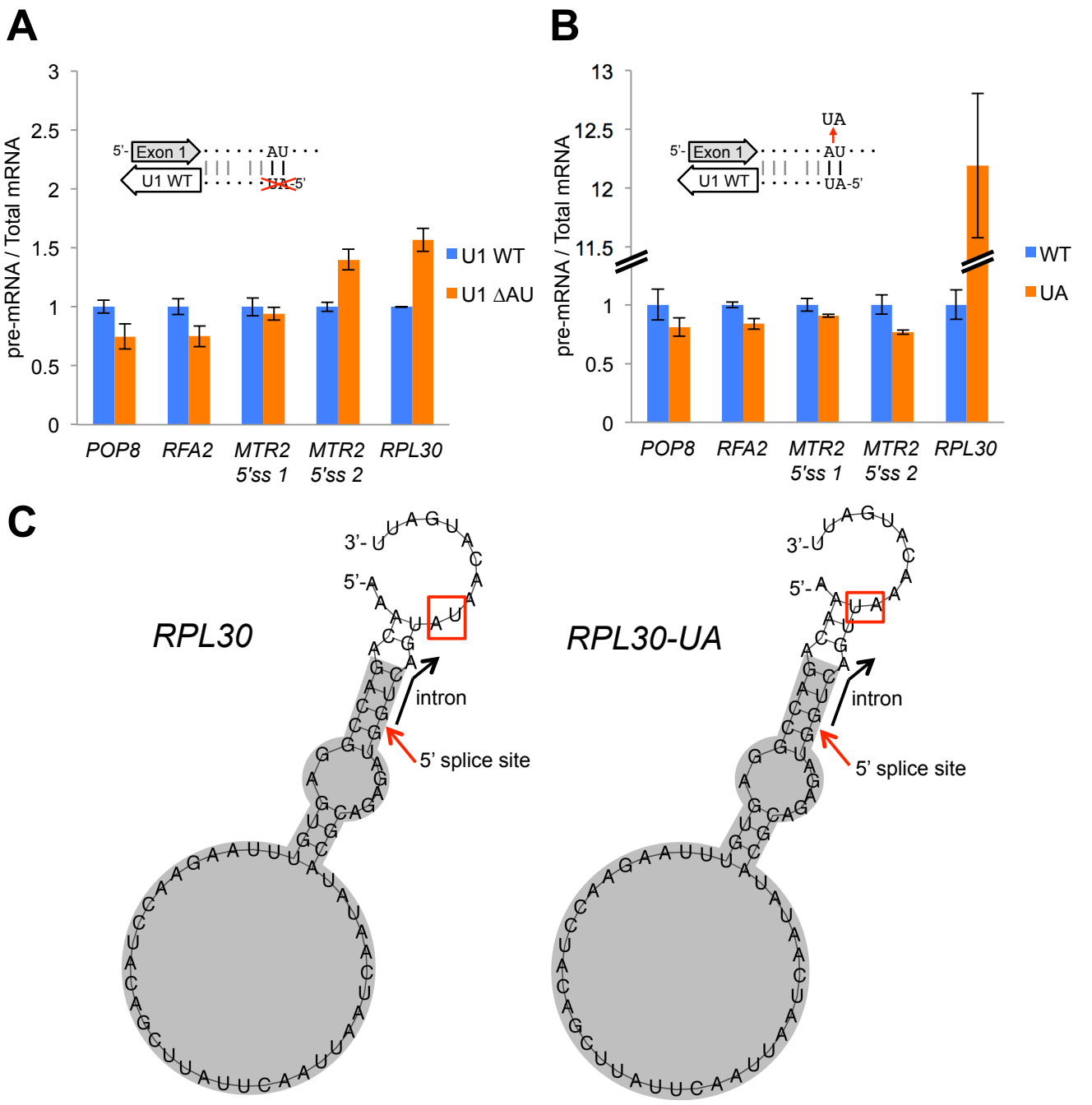
Supplementary Figure S2. Sucrose gradient analysis of the U1 snRNP in the wild-type (top) and the UU mutant strains (bottom). U1 snRNP, as detected by anti-Prp40p antibody, and U1 snRNA, as detected by primer extension co-peak at the fraction 7 in the wild-type strain (top). In contrast, U1 snRNP in the UU mutant is shifted to fraction 8 (bottom). Snu114p is a U5-snRNP specific component, thus serving as a proxy for U5 snRNP. Position of the 5'-end of U1 snRNAs are numbered.



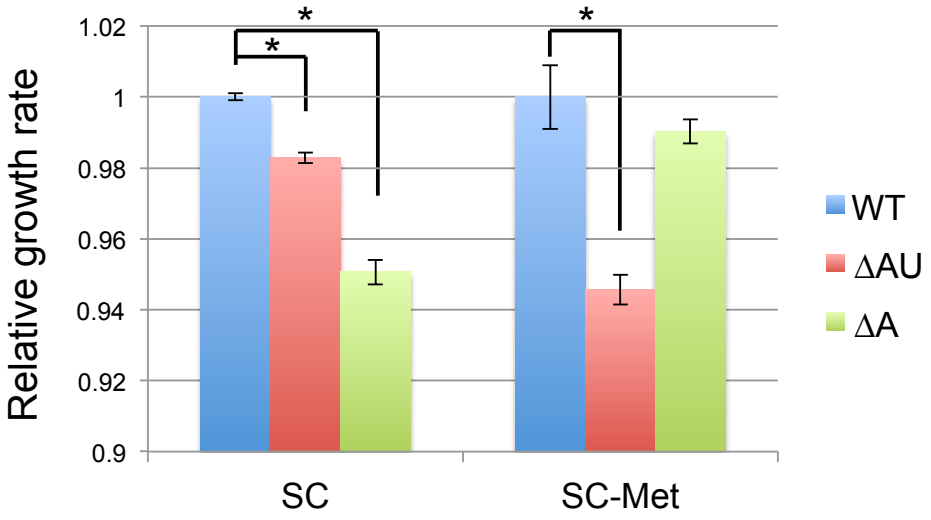
Supplementary Figure S3. 5'-end truncated U1 snRNA species produced in UU mutant are functionally defective. (**A**) The 5' ends of U1 snRNA species were mapped by primer extension in cells expressing either a full-length U1 snRNA (lane 1), a shortened form of U1 snRNA (U1^s) (lanes 2-3), or both (lanes 4-11). The 5'-end mutations (UU, AA, and GG) were engineered in the context of U1^s. The relevant chromosomal genetic backgrounds (*sto1*Δ and *cbc2*Δ) are indicated. Note that the UU mutant produced multiple species of shortened U1 snRNAs, with prominent starts from +20 and +30 (see also Figure 2). (**B**) Production of U1^s snRNA species form +7 (lane 3), +12 (lane 4), +17 (lane 5), and +20 (lane 6) in the presence of full-length U1 snRNA. (**C**) The 5'-end shortened U1 snRNA species fail to complement *snr19*Δ. Cells co-expressing a full-length U1 snRNA and a specific form of U1^s grew well on SD-Leu plate (top panel), but were inviable upon 5-FOA challenge, which counterselected the plasmid expressing the full-length U1 snRNA.



Supplementary Figure S4. Rescue of the *GG/tgs1*Δ and *GA/tgs1*Δ synthetic lethality by CBC2-Y24A mutation at 30˚C. See also Figure 6A.



Supplementary Figure S5. Disrupting the hypothetical basepairing between the 5'-end AU dinucleotide of the U1 snRNA and intron position 7-8 does not affect splicing efficiency. (**A**) Deleting the 5'-end AU dinucleotide of the U1 snRNA did not significantly affect splicing of selected introns. (**B**) Altering intron positions 7-8 did not affect splicing except in *RPL30*. For each essential gene, we constructed two strains harboring either wild-type or mutated intron (UA) in the context of its chromosomal deletion. Data in (**A**) and (**B**) represent the mean ± s.e.m. of three biological replicates. The significant accumulation of intron-containing *RPL30-UA* transcript may result from a hyper-stabilized secondary structure known to inhibit *RPL30* splicing (see below). (**C**) Predicted secondary structure of *RPL30* exon 1 and part of intron sequence. The *RPL30* RNA folds into a well-studied structure (gray-boxed) that blocks the 5'ss and inhibits its own splicing. Substitution of *RPL30* intron position 7-8 from AU to UA (red-boxed) is expected to further stabilize this structure by introducing one additional base pair (or three more, in comparison to the experimentally validated stem structure covered by the gray region). Structures were predicted by *RNAfold* WebServer (http://rna.tbi.univie.ac.at//cgi-bin/RNAWebSuite/RNAfold.cgi).



Supplementary Figure S6. Loss of the 5'-end AU dinucleotide of the U1 snRNA results in a significant reduction of cell fitness. Fitness was measured for the ΔAU and the ΔA mutants grown in synthetic complete medium (SC) and in SC medium dropping out of methionine (SC-Met), respectively. Note that both these two mutant strains produce a U1 snRNA transcript missing the 5'-end AU dinucleotide. Data represent the mean ± s.e.m. of three biological replicates. *: $p < 0.01$, two-tailed t-test.

Supplementary Table S1: Oligonucleotides used in this study.

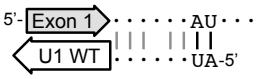
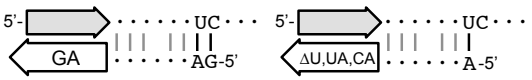
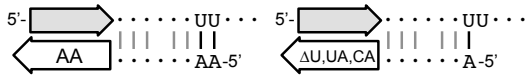
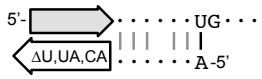
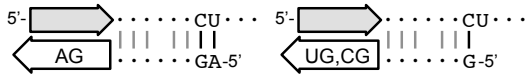
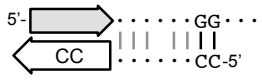
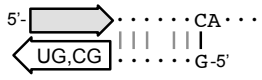
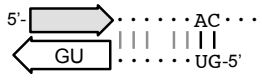
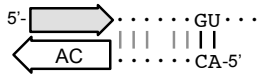
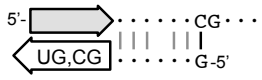
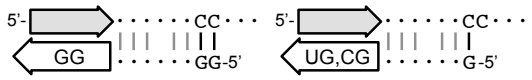
Gene	Experiment	Sense oligonucleotide	Antisense oligonucleotide	Note
SNR19	Primer extension		GGTGTCAAACTTCTCCAGGC	Position 266-285
	Northern blotting	GATCAAGAAGTCCTACTGATCAAACATGCG	CACCCGTTCCTACCAAGACCTTCC	For PCR DIG probe, position 26-543
	Quantitation of TMG-capped snRNA	CGGCAGATTCGAATGAACTTAA	ACAATCCCGACCAAATAATCTCA	Position 388-455
SNR20	Primer extension		AAAGTAGTGAGACCTGACATTAGCGG	Position 172-197
	CC formation		CAGATACTACACTTG	RB60, position 29-43
	Quantitation of TMG-capped snRNA	GCCGGCGGCATCAA	GGGTCGCGACGTCTCTAACT	Position 1037-1103
ACT1	ChIP primer set 1	CCAAGAAGAAAAAGAAAAGGTCAATC	AAAGCAGTAAGCGTGAAAAATCTAAA	Position (-122)-(-36)
	ChIP primer set 2	TCTTCCCAAGATCGAAAATTTACTG	GGACCGTGCAATTCTTCTTACAG	Position (-33)-64
	ChIP primer set 3, Splicing efficiency	GATTGCTTCATTCTTTTTGTTGCTAT	CTTTACACATACCAGAACCGTTATCAA	Position 280-363, also used for splicing efficiency assay (intron)
	ChIP primer set 4, Splicing efficiency	CTCGTGCTGTCTTCCCATCTATC	TTGACCCATACCGACCATGA	Position 388-455, also used for splicing efficiency assay (exon 2)
	ChIP primer set 5	CCGTGACATCAAGGAAAAACTATG	CAGCGGTTTGCATTTCTTGTT	Position 935-999
	ChIP primer set 6	TCCCAGGTATTGCCGAAAGA	TCATGGAAGATGGAGCCAAAG	Position 1225-1284
RPS9B	Splicing efficiency	TTTTCCAAGGCGGATACCTAATT	GGCCCAAGTAAGTGCTCATAAATATT	Intron region, 281-353
	Splicing efficiency	AAACGCTGCTAGAAAGGCTGAA	TTTATTCTTCGTCTTCGGCTTCTT	Exon 2, 938-1002
RPS10A	Splicing efficiency	GACAGCACAAGGCCATCGA	TGCCCAAATATAGGACGTATACCA	Intron region, 272-334
	Splicing efficiency	TTCAAGAAAGAAACCCAACTCAAA	AGAGAAAATTCAATTGGAGCTACGAAT	Exon 2, 712-785
RPS10B	Splicing efficiency	AGTACGGCAGTGCATGAATCC	GACGCGTGAAATAATGGAATTCT	Intron region, 199-264
	Splicing efficiency	TTCAAGAAAGAAACCCATCCCAAA	ATAGAAAATAGTTTACTTAAACGTAACAACGC	Exon 2, 685-762
RPS11A	Splicing efficiency	CGACTAATAGAATTTGCCAGTCAAA	ACCCAATCAACTATCGTAAACGTAAT	Intron region, 190-321
	Splicing efficiency	GATTCAACGTTGTCAAGGTCTCTGCT	CGTAAGTCTTATGCTGATTCATAACCAGCT	Exon 2, 746-839
RPS11B	Splicing efficiency	CCAAGTAAGGAGGATCTCCATCA	TTAGGCGAAAAGGAAATCGTTAA	Intron region, 408-485
	Splicing efficiency	GATTCAACGTCGTCAAGGTCTCT	CCTACGTTTTATCCTCCACTTATTATTATATCA	Exon 2, 918-1034
RPL16A	Splicing efficiency	GCCGCAGGATTTGCTCATAG	CTATCTCCATGTTTTAGCGTATACACAA	Intron region, 78-208
	Splicing efficiency	CAAATTGGAAGCAAAGAGAAAGGTTTC	GCAGTAGCATTAGCAGAGGCAACTTTC	Exon 2, 752-844
MUD1	Splicing efficiency	CGTTTCCTTAAATCTTCTTTCTACTA	CTGGCCTACTGGGTAAATTTT	Intron region, 32-129
	Splicing efficiency	AAATTTGCCAAGCGGCACTA	CAACTAAAGCCTCATTGCCAAGT	Exon 2, 785-849
SAE3	Splicing efficiency	TGAATGGCAACCTGATAAAGATGT	GCCGACTGGAAAAAATAGCTACA	Intron region, 62-139
	Splicing efficiency	ATTGAGAGATGCCGGTTTAAGG	CTTAATTTTGCATTGCTTTTCATCAG	Exon 2, 260-323

(Continued on next page)

Supplementary Table S1 (continued): Oligonucleotides used in this study.

Gene	Experiment	Sense oligonucleotide	Antisense oligonucleotide	Note
POP8	Splicing efficiency	GGGGAAAAAGACTTTTAGAGAATGG	GTTTATTTCTCTTCGGTAGTTGAGACTCA	Intron region, 3-97
	Splicing efficiency	CATTCGATCAAGATGTGGACGAT	TTATTTAACCATTGCCTCCATGTC	Exon 2, 128-194
RFA2	Splicing efficiency	GTATGTATTAGTGCTAGGAATTGGAGGAT	GAATCCCATCCATTGTTAGTAATTTTAAC	Intron region, 8-107
	Splicing efficiency	TGTGAGGGCAAAGACGCTAA	AGGATTGCGAGATCAATGGAA	Exon 2, 772-829
MTR2	Splicing efficiency	TGATAGACCATGACTCCCTGTATGTAT	GACACGTTCAACTTTTCTTGGTTAGT	Wild-type 5'ss 1, Intron region, (-236)-(-167)
	Splicing efficiency	TGATAGACCATGACTCCCTGTATGTTA	GACACGTTCAACTTTTCTTGGTTAGT	5'ss 1 UA mutant, Intron region, (-236)-(-167)
	Splicing efficiency	AGAAAAGTTGAACGTGTCGTACGTAT	CGTTTGCGCATCCTCTTTTTA	Wild-type 5'ss 2, Intron region, (-185)-(-99)
	Splicing efficiency	AGAAAAGTTGAACGTGTCGTACGTTA	CGTTTGCGCATCCTCTTTTTA	5'ss 2 UA mutant, Intron region, (-185)-(-99)
	Splicing efficiency	GCAACGTCAATTGCAAAGTCA	CGTCTTGCCCCATCTTGTCT	Exon 2, 287-346
RPL30	Splicing efficiency	TCCTGCTCTTGTGTTTGAGAGAGT	GCATATCTCACTTTTATTTCACAGTCATG	Intron region, 98-170
	Splicing efficiency	ACTTCCAAGGTGGTAACAACGAA	AAATAGAGACAACACCGACTCTGAATAA	Exon 2, 436-507

Supplementary Table S2: Hypothetical basepairing of U1 5'-end variants and the intron positions 7-8.

Position 7-8 of introns	No. of pertinent 5'ss	Introns in essential genes	Base pairing with U1 5'-end variants[a]	Essential genes
AU	45 (14.3%)	15 (23.8%)		*ARP9*, *GPI15*, *MTR2*[b], *POP8*, *PRE3*, *RFA2*, *RPL30*, *TAD3*[c], *YML6*
UA	56	14		*COF1*, *DBP2*, *ERD2*, *GLC7*, *NCB2*, *NOG2*, *RPL33A*, *RPL42A*, *RPS26A*, *SAR1*, *SEC27*, *YOS1*[c], *YRA1*, *YSF3*
UC	22	8		*ACT1*, *EFB1*, *GCR1*, *LSM2*, *MOB2*, *RPL18A*, *TFC3*, *TUB1*
UU	44	5		*BOS1*, *KEI1*, *NSP1*, *RPS13*, *SMD2*
UG	15	4		*ARP2*, *ERV1*, *MOB1*, *SEC14*
GA	15	4		*BET1*, *PFY1*, *RPL25*, *RPL32*
AG	12	4		*BIG1*, *RPL28*, *SEC17*, *UBC9*
AA	33	3		*PRP5*, *SFT1*, *YOS1*[c]
GC	10	2		*BET4*, *PMI40*
CU	8	2		*ECM9*, *SPT14*
GG	8	1		*KIN28*
CA	8	1		*RPL17A*
AC	20	0		
GU	12	0		
CG	4	0		
CC	2	0		
Total	314	63		

a. Baseparing between intron positions 7-8 and the "observed" 5'-end sequence of U1 snRNA (**Fig. 2**).
b. MTR2 contains six 5' UTR introns from permutations of two 5' ss and three 3' ss.
c. Contains two introns.

Supplementary Table S3: Biding free energy of CBC and the cap analogue m^7GpppX

		ΔG° (kcal/mol)		
CBC20	m^7GpppX	Exp. Data	Modeling	$\Delta\Delta G^\circ$
WT	G	-11.3 ± 0.1	-14.4	
Y138A	G	-10.5 ± 0.1	-13.4	0.8 / 1.0
WT	A	N.A.	-14.3	0.1
	C	N.A.	-13.7	0.7
	U	N.A.	-13.7	0.7

Exp. Data is adapted from Worch *et al.* (2009).
$\Delta\Delta G^\circ$ is the difference of ΔG° in comparing to that of WT vs. m^7pppG.