F) Supplemental Text and Figures

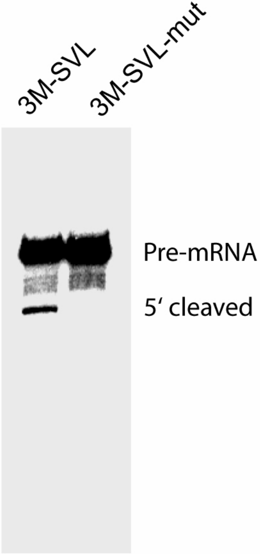


Figure S1. Mutant RNA substrates are defective in 3’ cleavage. 3M-SVL and 3M-SVL-mut RNA substrates were used in cleavage assays. Purified RNAs were resolved by 6% denaturing gel and visualized by phosphorimagery. Pre-mRNA and the 5’ cleaved product were marked.

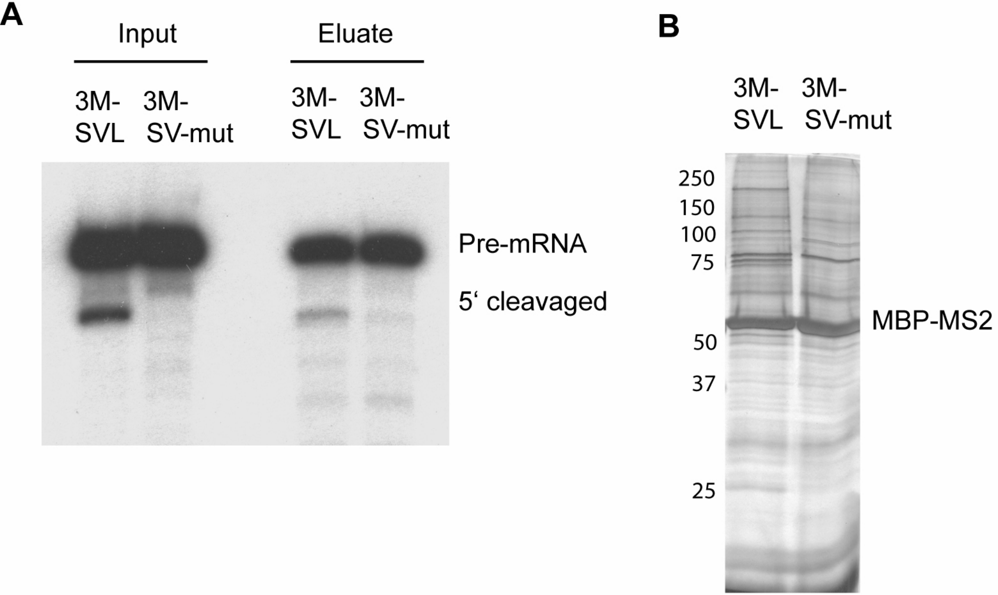


Figure S2. Purification of 3’ processing complexes assembled under cleavage conditions.

A. 3M-SVL or mutant (3M-SVL mut) RNAs were isolated from input cleavage reaction mixtures (input) or from eluted complexes after affinity purification (eluate), resolved on a 6% denaturing gel, and visualized by using a phosphorimager. Pre-mRNA and the 5’ cleaved products were marked. B. Proteins in complexes assembled on 3M-SVL or mutant (3M-SVL mut) RNAs after affinity purification were resolved on SDS-PAGE and visualized by silver staining. The position of the MBP-MS2 protein is marked.

Figure S3. Unique peptides from WDR33 detected in mass spectrometry analyses of the CPSF73 complex.

>gi|56243590|ref|NP_060853.3| WD repeat domain 33 isoform 1 [Homo sapiens]

MATEIGSPPRFFHMPRFQHQAPRQLFY**KRPDFAQQQAMQQLTFD**GKRMRKAVN**RKTIDYN**
PSVIKYLENRIWQRDQRDM**RAIQPDAGYYNDLVPPIGMLNNPMNAVTTKF**VRTSTNKV**KC**
PVFVVRWTPEGR**RLVTGASSGEFTLWNGLTFNFETILQAHDSPVRA**MTWSHNDMWMLTAD
HGGYV**KYWQSNMNNVKM**FQAHKEAIREASFSPTDNKFATCSDDGTVRIWDFLRCHEERIL
RGHGADVKCVDWHPTKGLVVSGS**KDSQQPIKFWDPKT**GQSLATLHAHKNTVMEV**KLNLNG**
NWLLTASRDHLCKLFDI**RNLKEELQVFRG**HK**KEATAVAWHPVHEGLFASGGSDG**SLLFWH
VGVEKEVGGMEMAHEGMIWSLAWHPLG**HILCSGSNDHTSKF**WTRNRPGDKM**RDRYNLNLL**
PGMSEDGVEYDDLEPNSLAVIPGMGIPEQLKLAMEQEQMGKDESNEIEMTIPGLDWGMEE
VMQKDQKKVPQKKVP**YAKPIPAQFQQAWMQNKVPIPAPNEVLNDRK**EDIKLEEK**KKTQAE**
IEQEMATLQYTNPQLLEQLKIERLAQKQVEQIQPPPSSGTPLLGPQPFPGQGPMSQIPQG
FQQPHPSQQMPMNMAQMGPPGPQGQFRPPGPQGQMGPQGPPLHQGGGPQGFMGPQGPQG
PPQGLPRPQDMHGPQGMQRHPGPHGPLGPQGPPGPQGSSGPQGHMGPQGPPGPQGHIGPQ
GPPGPQGHLGPQGPPGTQGMQGPPGPRGMQGPPHPHGIQGGPGSQGIQGPVSQGPLMGLN
PRGMQGPPGP**RENQGPAPQGMIMGHPPQEMRG**PHPPGGLLGHGPQEMRGPQEI**RGMQGPP**
PQGSMLGPPQELRGPPGSQSQQGPPQGSLGPPPQGGMQGPPGPQGQQNPARGPHPSQGPI
PFQQQ**KTPLLGDGPRAPFNQEGQSTGPPPLIPGLGQQGAQGRIPPLNPGQGPGPNKG**DSR
GPPNHHMGPMSERRHEQSGGPEHGPERGPFRGGQDCRGPPDRRGPHPDFPDDFSRPDDFH
PDKRFGHRLREFEGRGGPLPQEEKWRRGGPGPPFPPDHREFSEGDGRGAA**RGPPGAWEGR**
RPGDERFPRDPEDPRFRGRREESFRRGAPPRHEGRAPP**RGRDGFPGPEDFGPEENFDASE**
EAARGRDLRGRGRGTPRGGR**KGLLPTPDEFPRF**EGGRKPDSWDGNREPGPGHEHFRDTPR
PDHPPHDGHSPASRE**RSSSLQGMDMASLPPRK**RPWHDGPGTSEH**REMEAPGGPSEDRG**GK
GRGGPGPAQRVPKSGRSSSLDGEHHDGYHRDEPFGGPPGSGTPSRGGRSGSNWGRGSNMN
SGPPRRGASRGGGRGR

*Sequences covered by unique peptides detected by mass spectrometry were underlined.

*Total number of unique peptides detected: 55

*Total number of spectrum: 521

*Coverage: 38.4%

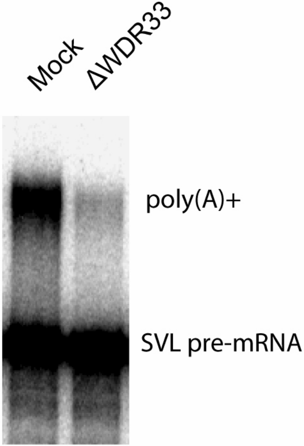
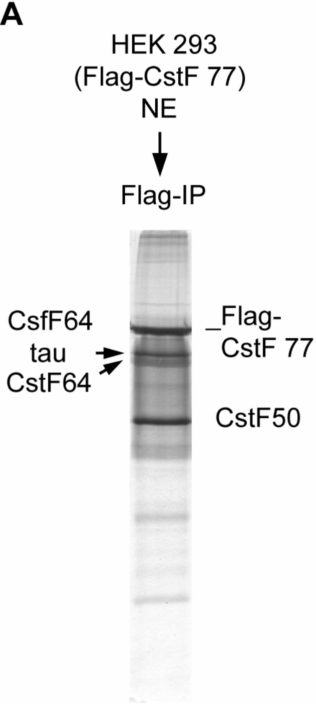


Figure S4. Depletion of WDR33 abolishes polyadenylation.
Mock-depleted (mock) and WDR33-depleted (ΔWDR33) NE were used in polyadenylation assays with SVL substrate. Purified RNAs were resolved on a 6% denaturing gel and visualizedusing a Phosphorimager. Pre-mRNA and poly(A)+ RNAs are marked.



B

>gi|14149675|ref|NP_056050.1| CstF 64, tau [Homo sapiens]

MSSLAVRDPAMDRSL**RSVFVGNIPYEATEEQLKDIFSEVGSVVSFRL**
VYDRETGKP**KGYGFCEYQDQETALSAMRN**LNGREFSGRALRVDNAAS
EKNKEEL**KSLGPAAPIIDSPYGDPIDPEDAPESITRAVASLPPEQMF**
ELMKQMKLCVQNSHQEA**RNMLLQNPQLAYALLQAQVVMRIMDPEIAL**
KILHRKIHVTPLIPG**KSQSVSVSGPGPGPGPGLCPGPNVLLNQQNPP**
APQPQHLARRPV**KDIPPLMQTPIQGGIPAPGPIPAAVPGAGPGSLTP**
GGAMQPQLGMPGVGPVPLERGQVQMSDPRAPIP**RGPVTPGGLPPRGL**
LGDAPNDPRGGTLLSVTGEVEPRGYLGPPHQGPPMHHASGHDTRGPS
SHEMRGGPLGDPRLLIGEPRGPMIDQRGLPMDGRGGRDSRAMET**RAM**
ETEVLETRVMER**RGMETCAMETRG**MEARGMDARGLEMRGPVPSS**RGP**
MTGGIQGPGPINIGAGGPPQGPRQVPGISGVGNPGAGMQGTGIQGTG
MQGAGIQGGGMQGAGIQGVSIQGGGIQGGGIQGAS**KQGGSQPSSFSP**
GQSQVTPQDQEKAALIMQVLQLTADQIAMLPPEQRQSILILKEQIQK
STGAS

*Sequences covered by unique peptides detected by mass spectrometry were underlined.
*Total number of unique peptides detected: 50
*Total number of spectrum: 1013
*Coverage: 56.3%

Figure S5. CstF64 tau is a component of the CstF complex.
(A) Immuno-purification of the CstF complex. NE was made from a stable HEK293 cell line expressing Flag-CstF77, and IP was performed using anti-Flag antibodies. Purified proteins were resolved and stained with silver. CstF components are marked.
(B) Mass spectrometry analyses of the CstF complexdetect CstF64 tau.

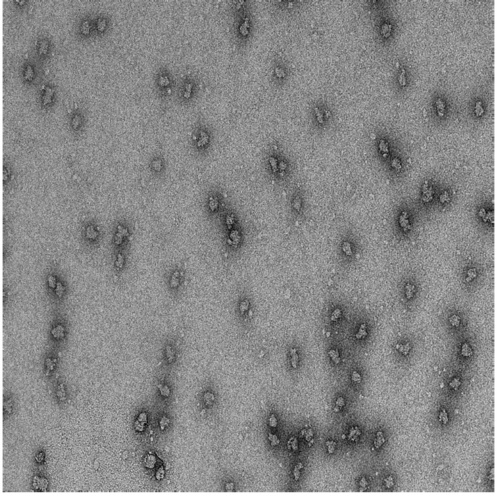


Figure S6. Tilted image of the negatively stained 3' processing complexes. A 50° tilted image of the negatively stainedparticles. The tilted images indicate that the sample is fully sandwiched between carbon membranes, and the staining seems homogeneous.

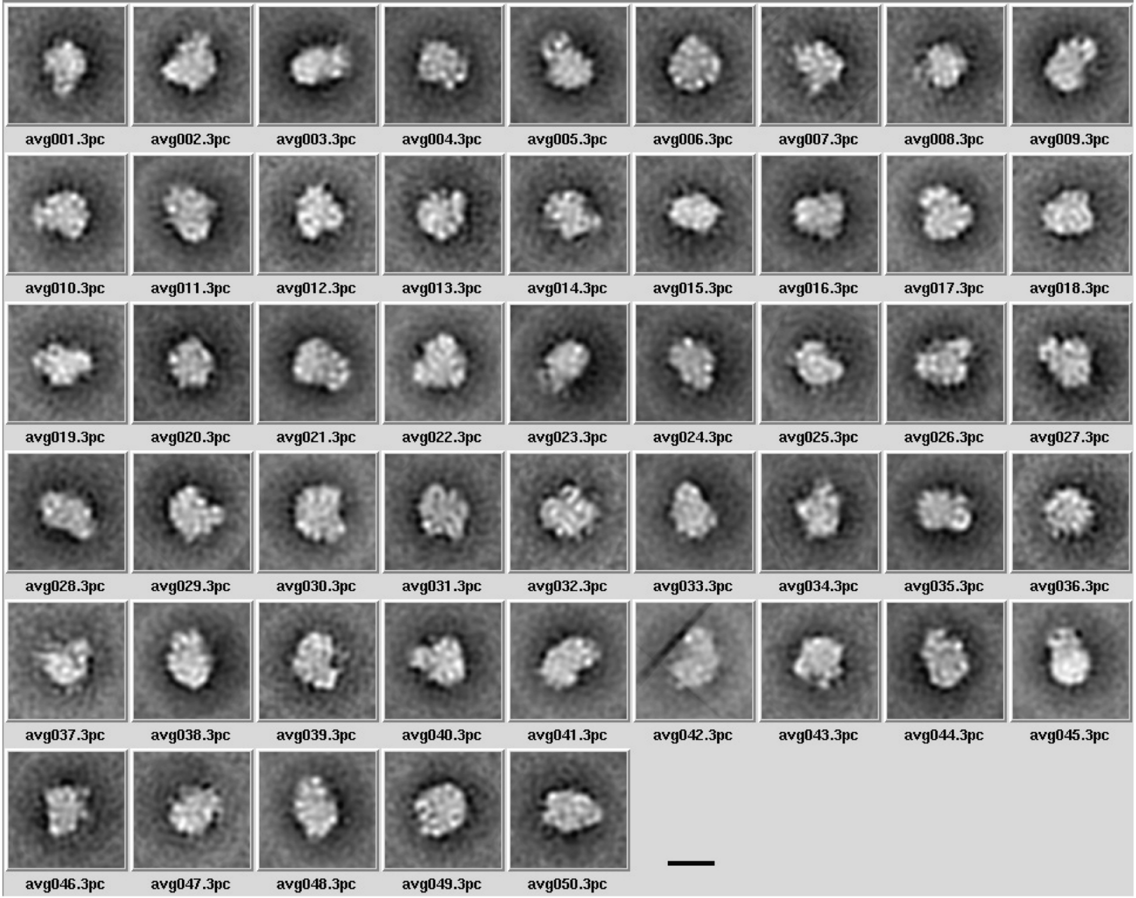


Figure S7. Class average images of the 3’ processing complex using SPIDER. 3,671 images of negatively stained single particles were classified into 50 groups and class averages were obtained after reference-free alignment using SPIDER. Scale bar, 20nm.

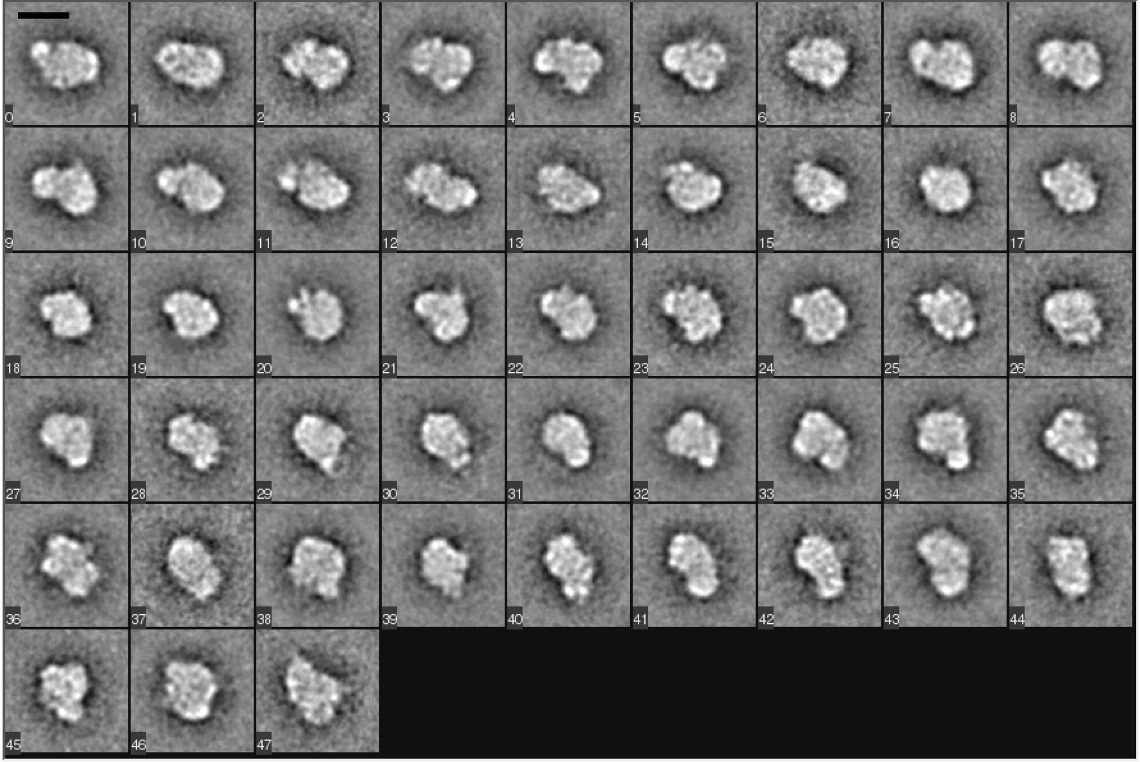


Figure S8. Class average images of the 3' processing complex obtained with EMAN. 3,671 images of negatively stained single particles were classified into 47 groups (~40-110 particles in each group) and class averages were obtained after reference-free alignment using EMAN. Scale bar, 10nm.

Table S1. Protein composition of the human pre-mRNA Cleavage complexes

Protein Name	Accession #	Mol. Weight	# peptides	
			L3	SVL
CPSF160	NP_037423	160822	18	17
CPSF100	gi 51338827	88487	15	9
CPSF73	NP_057291	77486	5	3
CPSF30	NP_006684	30124	1	0
hFip1	NP_112179	66526	12	12
CstF77	NP_001317	82922	16	16
CstF64	NP_001316	60959	12	10
CstF50	NP_001315	48358	5	6
CF Im 25	NP_008937	26227	14	7
CF Im 59	NP_079087	52050	11	5
CF Im 68	NP_008938	59209	12	7
Symplekin	NP_004810	126500	11	3
PABP 1	NP_002559.1	70324	5	1
WDR33	NP_060853	145921	10	7
RBBP6	NP_008841	201563	6	0
PP1 beta	NP_002700	37187	1	3
DNA topoIIalpha	NP_001058	174384	4	5
PARP1	NP_001609	113135	5	2
DNA-PK	NP_008835	469093	14	8
Ku 70	NP_001460	69843	4	2
Ku 86	NP_066964	82705	5	3
MDC1	NP_055456	226643	1	1
THO complex subunit 4	NP_005773	26757	3	4
THO complex subunit 6	AAH03118	32891	1	1
FACT complex large subunit	NP_003137	119914	2	2
NELF B subunit (BRCA1 coactivator)	NP_060853	61640	2	9
NELF E subunit	NP_002895.3	43240	1	1
BAF53a	NP_004292	47461	1	2
BRG1/SMARCA4	NP_003063	184644	1	1
BAF170	NP_003066	132879	1	0
BAF155	NP_003065	122753	1	0
BAF60b/SMARCAD	NP_003068	54945	1	0
Enhancer of rudimentary	NP_004441	12259	1	2

homolog (repressor)				
Bre1A	NP_149974	113977	1	3
U1-70K	NP_003080	70082	2	1
U4/U6.U5 tri-snRNP associated protein 1	NP_055317	55181	4	2
U4/U6.U5 tri-snRNP associated protein 2	NP_005137	90255	1	2
SF3b155 (SAP155)	NP_036565	145815	6	3
P54/NRB	NP_004759	53542	1	2
ASF/SF2	NP_008855	27745	1	2
SR-A1	NP_067051	139296	1	?
SRp38	NP_473357	31345	1	1
SRrp86 (SR12)	NP_631907	71650	1	2
SRm300	NP_057417	299676	6	3
(KSRP) KH-type splicing regulatory protein (also mRNA turnover)	NP_003676	73161	2	2
hnRNPA0	NP_006796.1	30841	4	1
hnRNP A3	NP_005749.1	29357	4	2
hnRNP U-like (E1B5-associate protein 5)	NP_008971.2	95739	3	2
hnRNP D	NP_005454.1	46437	2	1
RNA helicase A (DEAD/H box-9)	NP_001348	142069	7	1
DEAD/H box-39	NP_005795.2	49130	2	1
DEAD/H box-49	NP_-61943	54226	1	1
DEAH box-8	NP_004932	139314	1	1
DEAD box-21	NP_004719	87344	2	2
DEAD/H box-15	NP_001349	92829	2	2
DEAD box 42	NP_987095	102975	1	1
Exosome 10		100831	3	1
RRP4		32789	0	1
RRP40		29441	0	1
RRP41		26252	2	0
RRP42		31835	1	0
RRP43		30040	1	0
EEF1-alpha	NP_001393	50141	3	2
EEF1-epsilon	NP_004271.1	37974	1	1
EIF 4B	NP_001408	69224	4	2
40S ribosomal protein S5	NP_001000.2	22745	1	1
40S ribosomal	NP_001008.1	17091	1	2

protein S13				
40S ribosomal protein S24	NP_148982.1	15069	1	1
40S ribosomal protein S26	NP_001020.2	13015	1	1
Similar to 60S ribosomal protein L7a	NP_000963.1	29996	1	1
60S ribosomal protein L8	NP_000964.1	28025	2	2
60S ribosomal protein L9	NP_000652.2	21863	2	1
60S ribosomal protein L24	NP_000977.1	17779	2	1
60S ribosomal protein L27	NP_000979.1	15798	1	1
zinc finger CCHC domain-containing protein 8 (ZCCHC8)	NP_060082	79375	8	3
Scaffold attachment factor B	NP_002958.2	102768	3	2
TAR DNA-binding protein	NP_031401.1	33730	3	2
Lamin-A/C	NP_005563.1	74140	3	2
Clathrin heavy chain	NP_004850	191613	3	5
Lamina-associated polypeptide 2 alpha (Thymopoietin)	NP_003267.1	75361	3	2
Phenyalanyl-tRNA synthetase beta	NP_00678.2	66130	1	1

*Components of multi-subunit complexes that are present in only one purified complexe are listed and lightly shaded.

Table S2: Protein compositions of the CPSF and CstF subcomplexes, and comparison with that of the human pre-mRNA 3’ processing complex.

Protein Name	Accession #	Yeast homolog	Motifs	Cal. Mass	# of peptides			
					L3	SVL	CPSF73-3Flag	Flag-CstF77
KNOWN POLYADENYLATION FACTORS								
CPSF Complex								
CPSF160 (CPSF1)	NP_037423	CFT1	SFT1	160822	80	72	87	22
CPSF100 (CPSF2)	NP_059133.1	CFT2	β-CASP	88487	59	52	83	24
CPSF73 (CPSF3)	NP_057291	YSH1	β-CASP	77486	42	22	81	9
CPSF30 (CPSF4)	NP_006684	YTH1	Zinc finger	30124	18	13	19	2
hFip1	NP_112179	FIP1		66526	23	19	35	13
CstF Complex								
CstF77 (CSTF3)	NP_001317	RNA14	HAT	82922	60	54	3	118
CstF64 (CSTF2)	NP_001316	RNA15	RRM	60959	31	22	11	45
CstF50 (CSTF1)	NP_001315		WD repeats	48358	21	18	0	48
CF Im Complex								
CF Im 68 (CPSF6)	NP_008938		RRM	59209	15	10	0	0

CF Im 59	NP_079087		RRM	52050	15	12	0	0
Cf Im 25 (CPSF5)	NP_008937			26227	26	21	0	0
CF IIm Complex								
Pcf11	NP_056969.2	PCF11	CID	173050	5	3	0	0
Other known polyadenylation factors								
Symplekin	NP_004810	PTA1		126500	51	39	31	0
Symplekin variant	BAE06092.1			118793	48	38	33	0
PAPOLG	NP_075045	PAP	RRM	82803	2	3	0	0
PABPC1	NP_002559.2	PAB1	RRM	70671	27	33	3	0
PABPC4	NP_003810.1		RRM	70783	20	23	3	0
PABPN1	NP_004634		RRM	31749	14	12	0	0
(Putative) homologues of yeast polyadenylation factors								
CstF64 tau (CSTF2T)	NP_056050.1	SPAC644.16 (S. pombe)	RRM	64436	21	16	8	50
WDR33	NP_060853	PFS2	WD repeats	145921	49	45	55	10
RBBP6	NP_008841	MPE1	RS, DWNN	201563	4	3	0	0
PP1 alpha	NP_002699.1	GLC7	Phosphatase	37512	2	7	2	0
PP1 beta	NP_002700	GLC7	Phosphatase	37187	2	5	0	0
DNA Damage Response Factors								
DNA-PK	NP_008835		Kinase	469093	21	15	0	0
Ku 70	NP_001460	YKU70	DNA helicase	69843	4	7	0	0
Ku 86/XRCC5	NP_066964	YKU86	DNA helicase	82705	6	13	0	0

PARP-1	NP_001609			113135	5		0	0
RNAP II and associated factors								
Pol II large subunit Rpb1	NP_000928.1	RPB1	Polymerase	217204	2	2	2	0
Pol II B	NP_000929.1	RPB2		133896		2	4	0
Rpb5 (Pol II E)	NP_002686.2	RPB5		24611	2		0	0
Rpb11	AF468111_1	RPB11		17208		2	0	0
Transcription Factors								
TF II, I	NP_127492.1			112416	8	2	0	0
TAF15	AAH46099.1		RRM, zf-RanBP	52061	6	5	0	0
Integrator Complex								
INTS2	NP_065799.1			134346	2	2	0	0
INTS3	NP_075391.3			118013	4	9	0	0
INTS4	NP_291025.3			108171	2	3	0	0
INTS6	NP_036273.1			100390	2	3	0	0
INTS7	NP_056249.1			106834	2	2	0	0
INTS8	NP_060334.2			113088		3	0	0
INTS9	NP_060720.1		YSH1 homology	73815		4	0	0
INTS10	NP_060612.2			82236		2	0	0
PAF Complex								
Cdc73	AAH14351.2	CDC73		46675	2	2	0	0
FACT Complex								

FACT complex large subunit	NP_009123	SPT16		119914	4	4	0	0
Splicing Factors								
p54nrb	NP_031389.3			54232	11	3	2	0
PSF	NP_005057.1			76150	10	4	3	0
PUF60	NP_510965.1		RRMs	59876	3	9	0	0
UAP56	NP_004631.1	SUB2	RNA Helicase	48991	12	9	3	0
SF1	NP_004621.2	BBP		68330	5	5	0	0
SF3A, subunit 1	NP_005868.1	PRP21	Surp	88886	9	15	0	0
SF3A, subunit 2	NP_009096.2	PRP11		49256	2	4	0	0
SF3A60	CAA57388	PRP9		58777	3	18	0	0
SF3B, subunit 1	NP_036565.2	HSH155		145830	16	34	5	0
SF3B, subunit 2	NP_006833.2	CUS1		100228	2	21	3	0
SF3B, subunit 4	NP_005841.1	HSH49		44386	3	5	0	0
SF3B, 14 kDa	NP_057131.1	SNU17		14585		5	0	0
Prpf38b	NP_060531.1			64441	2	8	0	0
SRm300	NP_057417.2		RS	299676	13	15	0	0
U1 70K	NP_003080.2	SNP1	RRM	51557	7	7	3	0
U2AF65	NP_009210.1	MUD2	RRM	53501	6	11	0	0
U2AF35	NP_006749.1		RRM	27872	2	3	0	0
PRP19	NP_055317.1	PRP19	RING, WD	55181	5	3	6	0
Exosome								
SKIV2L2 (hMTR4)	NP_056175.2	MTR4	DEXD	117933	29	28	0	0
Translation Factors								

eEF1-alpha	NP_001393.1	TEF1/2	GTPase	50141	14	12	5	0
eEF1-gamma	NP_001395.1	TEF4	GST	50119	5	4	0	0
eIF2A	NP_114414.2	YGR054W		64990	4	4	0	0
eIF 4A	NP_001407.1	TIF1	Helicase	46154	12	6	0	0
eIF3 gamma	NP_003747.1		JAB/MPN	39930	5	6	0	0
eIF3S2	NP_003748.1	TIF34	WD	36502	2	2	0	0
eIF3S5	NP_003745.1		PCI	37564	3	2	0	0
eIF3S6	NP_001559.1			52221	2	4	0	0
eIF3S9 eta	NP_003742.2		RRM	92482	7	7	0	0
eIF3A	NP_003741.1	RPG1		166569	16	9	0	0
eIF4G1	NP_886553.2	TIF4631		175460	2	2	0	0
RACK1/GNB2L1/lung cancer oncogene 7	NP_006089.1	ASC1	WD	35077	5	5	0	0
Factors with known motifs								
RNA binding motif protein 7	NP_057174.1		RRM	30503	4	4	0	0
RNA binding motif protein 9/Ataxin-binding protein/Fox-1	NP_055124		RRM	39515	2	7	0	0
RNA binding motif protein 25	NP_067062.1		RRM	100186	4	16	0	0
RNA binding motif protein 39, isoform a	NP_909122.1		RRM	59380	2	4	2	0
RNA binding motif protein 39, isoform b	NP_004893.1		RRM	58657	2	4	2	0

DEAH box polypeptide 3	NP_001347.3	DBP1	RNA helicase	7244	13	9	0	0
DEAH box polypeptide 5	NP_004387.1	DBP2	RNA helicase	69148	4	5	0	0
DEAH box polypeptide 9	NP_001348.2		RNA helicase	140958	32	3	0	0
DEAH box polypeptide 15	NP_001349.2	PRP43	RNA helicase	90933	3	17	0	0
DEAH box polypeptide 36	NP_065916.1		RNA helicase	114776	12	27	0	0
DEAH box polypeptide 39	NP_005795.2		RNA helicase	49130	8	5	0	0
DEAD box polypeptide 6	NP_004388.1	DHH1	RNA helicase	82432	3	2	0	0
DEAD box polypeptide 17 (p82)	NP_006377.2		RNA helicase	80273	6	4	0	0
Gemin3 (DEAD box polypeptide 20)	NP_009135.3		RNA helicase	92213	4	2	0	0
DEAD box polypeptide 23	NP_004809.2	PRP28 (S. Pombe)	RNA helicase	95583	4	5	0	0
ZCCHC8	NP_060082		Zinc finger	79375	15	16	0	0
EBNA2 coactivator, p100/SND1	NP_055205.2		Staphylococcal nuclease domain, Tudor	101997	12	2	0	0
Interferon-induced protein with tetratricopeptide repeats 1 isform 2	NP_001539.3		TPR	55360	9	7	0	0

Interferon-induced protein with tetratricopeptide repeats 3	NP_001540		TPR	55985	8	5	0	0
Other Factors								
G3BP/Ras-GTPase-activating protein SH3-domain-binding protein	NP_005745.1		Endoribonuclease	52164	5	2	0	0
JUP protein/ gamma-catenin	NP_002221.1			81745	5	3	0	0
PNUTS (PPP1R10)	NP_002705.2		PP1 regulator	99058	6	9	0	0