Additional File 1 for:

The *C. elegans* 3'UTRome V2: an updated genomic resource to study 3'UTR biology

Steber HS ^{1, 2}, Gallante C³, O'Brien S², Chiu P.-L⁴, Mangone M*^{1, 2}.

¹Molecular and Cellular Biology Graduate Program, School of Life Sciences 427 East Tyler Mall Tempe, AZ 85287 4501.

²Virginia G. Piper Center for Personalized Diagnostics, The Biodesign Institute at Arizona State University, 1001 S McAllister Ave, Tempe, AZ, USA

³Barrett, The Honors College, Arizona State University, 751 E Lemon Mall, Tempe, AZ 85281

⁴Center for Applied Structural Discovery, The Biodesign Institute at Arizona State University, 1001 S McAllister Ave, Tempe, AZ, USA

 To whom correspondence should be addressed. Tel: +1(480) 965-7957; Fax: +1(480) 965-3051; Email: mangone@asu.edu

Present Address: Marco Mangone, Arizona State University, Biodesign Institute Building A, 1001 S McAllister Ave, Tempe, AZ 85281 USA

This PDF includes:

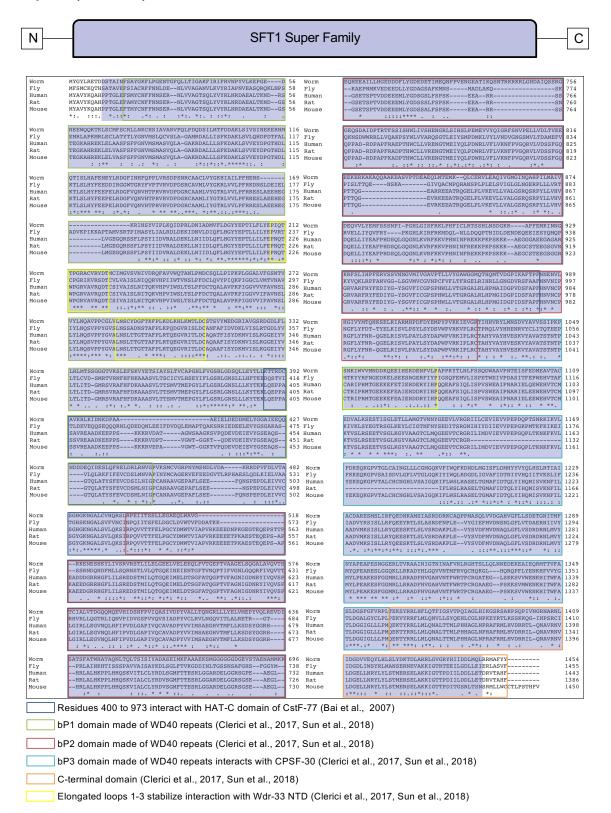
Figures S1-S9

Table of Contents

Additional File 1

-	Protein alignment of members of the CPC complex in five species
Figure S2:	Results of the RNAi experiments of the <i>C. elegans</i> CPC14
Figure S3:	Bioinformatic pipeline used in this study15
Figure S4:	PAS site usage in genes with multiple 3'UTR isoforms16
Figure S5:	Detection of the 'UGUA' element in <i>C. elegans</i> 3'UTRs17
Figure S6:	Nucleotide binding site of the human CPSF160-WDR33-CPSF30 complex18
Figure S7:	In vivo cleavage assay for M03A1.319
Figure S8:	In vivo cleavage assay for Y106G6H.920
Figure S9:	<i>In vivo</i> cleavage assay for <i>ges-1</i> 21

cpsf-1 (CPSF160)



Supplemental Fig. S1: Protein alignment of members of the CPC complex in five organisms. Amino acid sequence alignment for the members of the CPC from five organisms produced with Clustal Omega Multiple Sequence Alignment. Conserved domains produced with the Batch Conserved Domain Search on NCBI are represented by the highlighted regions of each figure. Known domains determined from previously published literature are outlined.



]_	CPSF2-like MBL Fold beta-CAS	P)—(CPSF100 C
	MTSIIKLHTISGAMDESPPCYILQIDDVRILLDCGWDEKFDANFIKELKRQVHTLDAVLL MTSIIKLTTLSQOQEESALCYLLQVDEFRFLDCGWDEHFSWDIIDSLRKHVHQIDAVLL MTSIIKLTTLSGVQEESALCYLLQVDEFRFLDCGWDEHFSVDIIDSLRKHVHQIDAVLL	60 60 60 60 60	Worm Fly Human Mouse Rat	KASFFKTTKKSFPHPPYIEEKVKHDDYGEVIKPEDYTVISKIDLRKGONKD 5: HSGFFKSNKRHHVMFPYHEEKVKCDEYGEIINLDDYRIADATGYEFVPHEEQNKENVKKE 5 KGSFFKQAKKSYPMPAPEERIKWDEYGEIKPEDFUPPLQATEEEKSKL 5 KGSFFKQAKKSYPMPAPEERIKWDEYGEIKPEDFUPPLQATEEEKSKL 5 EEKSKL 5
	SHPDAYHLGALPYLVGKLGLNCPIYATIPVFKMGQMFMYDLYMSHFNMGDFDLFSLDDVD SHPDPLHLGALPYAVCKLGLNCAIYATIPVYKMGQMFMYDLYQSRHNTEDFTLFTLDDVD SHPDPLHLGALPFAVCKLGLNCAIYATIPVYKMGQMFMYDLYQSRHNTEDFTLFTLDDVD	120	Worm Fly Human Mouse Rat	EPVVVKKREEEEEVYNPNDHVEEMPTKCVEFKNRVEVSCRIEFIEYEGISDGESTKKLLA 5 EPGIGAEQQANGGIVDNDVQLLEKPTKLISORKTIEVNAQVQRIDFEGRSDGESNLKILS 5 ESGLTNOEEPHQDLSOVPTKCISTEFIEIRANVTYIDVEGRSDGESKKIIN 5 ESGLTNGEEPHQQLSOVPTKCVSATESIEIRANVTYIDVEGRSDGSIKKIIN 5 # : :::::::::::::::::::::::::::::::::
	TAFEKITQLKYNQTVSLKDKGYGISITPLNAGHMIGGTIWKIVKVGEEDIVYATDFNHKK AAFDKIOOLKESOIVNLKGKGHGLSITPLPAGHMIGGTIWKIVKOGEEEIVYAVDFNHKE	180 180	Fly Human Mouse	GLLPRQIIVVHGSRDDTRDLVAYFADSGFDTTMLKAPEAGALVDASVESFIYQVALSDAL GLPRQIIVHGTABGTQVVARHCSQNVGARVFTPQKGEIIDVTSEIHIYQVRLHEDL GMKPRQLIIVGFPEASOLAECCRAFGGKDIKVYMPKLHETVDATSETHIYQVRLKDSL GMKPRQLIIVHGPEASOLAECCRAFGGKDIKVYMPKLHETVDATSETHIYQVRLKDSL GMKPRQLIIVHGPEASOLAECCRAFGGKDIKVYMPKLHETVDATSETHIYQVRLKDSL i**:*:*:*****:*:*
	ERHLNGCSFDNFNRPHLLITGAHHISLPOMRRKDRDECLVTKILRTVRQKGDCNIVIDTA ERHLSGCELDRLQRFSLLTDAYNAQYQQARRRADEBKLMTNILGTVRUNGNVLLAVDTA EIHLNGCSLEMLSRFSLLTDSFNATYQPRKRQREDCLLTNUETLRGONVLLAVDTA EIHLNGCSLEMLSRFSLLTDSFNATYQPRKRQREDCLLTNUETLRGONVLLAVDTA HILLNGCSLEMLSRFSLLTDSFNATYQPRKRQREDCLLTNUETLRGONVLLAVDTA *******	240 240 240	Fly Human Mouse	LADIQFKEVSEGNSLAWIDARVMEK-EAIDNMLAVGTSNLMIDDKNREEDVNDQEENGAT 7C VSGLQFCKK-DAEVANUDGKLCMRVKAIEAPHDVTV
	GRVLELAHMLDQLWKNKESGLMAYSLALLNNVSYNVIEFAKSQIEWMSDKLTKAF-EGAR	299 299 299	Worm Fly Human Mouse Rat	RGEGNAEP HEIGENGSQESLAISESGKEVENGHTNDSRTKKGTKGKIRGNLILDPLPKRL 76 -DDGEDSE MQVEAFSDSSVIAQQKAMKSL-FGDDEKFCHEESSIIFTLEPLPPHE 70 -DDGEDSE MQVDAFSDSSMAQQKAMKSL-FGDDEKELGEEFSIIFTLEPLPPHE -DDGEDSE MQVDAFSDSSMAQQKAMKSL-FGDDEKELGEEFSVIFTLEPLPPHE 70 -DDGEDSE MQVDAFSDSSVIAQQKAMKSL-FGDDEKELGEESVIFTLEPLPPHE 70 - * *
	NNPFQFKHIQLCHSLADVYKLPAGPKVVLASTPDLESGFTRDLFVQWASNANNSIILTTR NNPFQFRHLSLCHGLSDLARVP-SFKVVLASQPDLECGFSRDLFIQWCQDFKNSIILTYR NNFFQFRHLSLCHGLSDLARVP-SFKVVLASQPDLECGFSRDLFIOWCOPFKNSIILTYR	359 358 358	Worm Fly Human Mouse Rat	IPIHQAVFVNDPKLSDFKNLLTDKGYKAEFLSGTLLINGGNCSIRRNDTGVFQMEGAFTK 82 IPIHNSVLINELKLSDFKQTLMENNINSFFGGVLMCSNOTLALRRVDAGKVAMEGCLSE 74 VPGGQSVFMEPRLSDFKQVLLBEGIQAFFVGGVLVCNNQ-VAVRRTETGRIGLEGCLCQ 76 VPGGQSVFMEPRLSDFKQVLLBEGIQAFFVGGVLVCNNQ-VAVRRTETGRIGLEGCLCQ 76 I**:::::::::::::::::::::::::::::::::::
	PASTLARKLVNMAERANDGVLKHEDRLISLVVKKRVALEGELLEYKRRKAERDAETR TSPGTLANELVENCA	410 408 408	Fly Human	DYYKLRRLFYDQFAVL 843 EYYKIRELLYEQYATV 756 DFYRKRDLLYEQYATV 784 DFYRKRDLYEQYATV 782 DFYRKRDLYEQYATV 782 If*if********* 782
	LEQSKE-ADIDSSDESDIEDIDQPSAHKTKHDLMMKGEGSR LEQSKE-ADIDSSDESDVEEDVDQPSAHKTKHDLMMKGEGSR	477 445 449 449 449		

Beta-CASP domain (Mandel et al., 2006)

cpsf-3 (CPSF-73)

Г

•			
N	Lactamase βbeta-CASP	_(F	Pfam CPSF-73/CPSF-100 C
Worm		Worm	IIAGYCVEGTLAKHILSEPEEIVSLSGEKLPMRMQVGYVSFSAHTDYHQTSNFVKALKPP 413
Human Mouse	MSAIPAEESDQLLIRPLGAGQEVGRSCIILEFKGRKIMLDCGIHPGLEGMDALP 54 MSAIPAEESDOLLIRPLGAGOEVGRSCIILEFKGRKIMLDCGIHPGLEGMDALP 54	Human Mouse	IIAGYCVEGTLAKHIMSEPEEITTMSGQKLPLKMSVDYISFSAHTDYQQTSEFIRALKPP 412 IIAGYCVEGTLAKHIMSEPEEITTMSGQKLPLKMSVDYISFSAHTDYQQTSEFIRALKPP 412
Rat	MSAIPAEESDQLLIRPLGAGQEVGRSCIILEFKGRKIMLDCGIHPGLEGMDALP 54	Rat	IIAGYCVEGTLAKHIMSEPEEITTMSGQKLPLKMSVDYISFSAHTDYQQTSEFIRALKPP 412
Fly	MTQATCDARMPDEESDLLQIKPLGAGQEVGRSCIMLEFKGKKIMLDCGIHPGLSGMDALP ::** * : ***:*************************	Fly	IIAGYCVEGTLAKAVLSEPEEITTLSGQKLPLNMSVDYISFSAHTDYQQTSEFIRLLKPT ************************************
Worm	FVDFVEIENIDLLLITHFHLDHCGALPWLLQKTAFQGKCFMTHATKAIYRMLLGDYVRIS 113	Worm	HLVLVHGELHEMSRLKSGIERQFQD-DNIPIEVHNPRNTERLQLQFRGEKTAKVIGKLAQ 472
Human	YIDLIDPAEIDLLLISHFHLDHCGALPWFLQKTSFKGRTFMTHATKAIYRWLLSDYVKVS 114	Human	HVILVHGEQNEMARLKAALIREYEDNDEVHIEVHNPRNTEAVTLNFRGEKLAKVMGFLAD 472
Mouse Rat	YIDLIDPAEIDLLLISHFHLDHCGALPWFLQKTSFKGRTFMTHATKAIYRWLLSDYVKVS 114 YIDLIDPAEIDLLLISHFHLDHCGALPWFLQKTSFKGRTFMTHATKAIYRWLLSDYVKVS 114	Mouse Rat	HVILVHGEQNEMARLKAALIREYEDNDEVHIEVHNPRNTEAVTLNFRGEKLAKVMGFLAD 472 HVILVHGEQNEMARLKAALIREYEDNDEVHIEVHNPRNTEAVTLNFRGEKLAKVMGFLAD 472
Fly	VVDLIEADEIDLISSH HLDHCGALEWFIMKYSFKGRCFMTHATKAIYRWMLSDVIKIS ::*::::::::::::::::::::::::::::::::::		HVVLV GEONEMSRLKLALQREYEADASTDIKFYNPRNTHAVIDLYFRGERTARVNOSIAA 478 *::***** :**:*** : *::: *::***********
Worm	KYGGPDRNQLYTEDDLEKSMAKIETIDFREQKEVNGIRFWPYVAG <mark>H</mark> VLGACQFMIEIAGV 173		RVPENNETISGVLVKNNFSYSIMVPEELGSYTSLRISSLEQRMSVHYSGSLKLLIFNLQQ 532
Human		Human	KKPEQGQRVSGILVKRNFNYHILSPCDLSNYTDLAMSTVKQTQAIPYTGPFNLLCYQLQK 532
Mouse Rat		Mouse Rat	KKPEQGQRVSGILVKRNFNYHILSPCDLSNYTDLAMSTVKQTQAIPYTGPFYLLYYQLQK 532 KKPEQGQRVSGILVKRNFNYHILSPCDLSNYTDLAMSTVKQTQAIPYTGPFYLLYYQLQK 532
Fly	NISTEQMLYTEADLEASMEKIETINFHEERDVMGVRFCAVIACHVLGAAMFMIEIAAI 178 : . : **** *** ** ****::: ::* *::* ********		KNSEVGSKLSGVLVKRDFKYHLLAPSDLGKYTDMSNSVVTQRQSIPWGSSLSTLELLLDR 538 : * :**:***.:*.* :: * :**: : * :: * :: * :: * :: * *::
Worm	RVLYTCDFSCLEDRHLCAAEIPPITPQVLITESTYCTQTHEDRAVREKRFTQMVHDIVTR 233		LNDDACLIQNIKLKEISKKGSVTQAITVFQGKVNVTVYGNDHVVVVRWDSNPVYDMYADS 592
Human Mouse		Human Mouse	LTGDVEELEIQEKPALKVFKNITVIQEPGMVVLEWLANPSNDMYADT 579 LTGDVEELEIOEKPALKVFKSITVVOEPGMVVLEWLANPSNDMYADT 579
Rat		Rat	LTGDVEELEIQEKPALKVFKSITVVQEPGMVVLEWLANPSNDMYADT 579
Fly	KILYTCD SRQEDRHLMAAEVPPMKPDVLITESTYCTHIHEKREDRENRFTSLVQKIVQQ 238	Fly	IGAGCVEV-LEAERKLRVFGCIELTVEQKIIVMEWQATHVNDVYADA 584
Worm	GGRCLIPAFAIGPAQELMLILDEYWESHQELHDIPVYYASSLAKKCMSVYQTFVNGMNSR 293		VVAAILHAQANPVPDKYLPSNSSFPQFNTAIEGMVKHICGDDVSIVMSERGLLAQF 648
Human Mouse	GGRGLIPVFALGRAQELLLILDEYWQNHPELHDIPIYYASSLAKKCMAVYQTYVNAMNDK 292		VTTVILEVQSNPKIRKGAVQKVSKKLEMHVYSKRLEIMLQDIFGEDCVSV-KDDSILSVT 638
Rat	GGRGLIPVFALGRAQELLLILDEYWQNHPELHDIPIYYASSLAKKCMAVYQTYVNAMNDK 292 GGRGLIPVFALGRAQELLLILDEYWQNHPELHDIPIYYASSLAKKCMAVYQTYVNAMNDK 292		VTTVILEVQSNPKIRKGAVQKVSKKLEMHVYSKRLEVMLQDIFGEDCVSV-KDDSVLSVT 638 VTTVILEVQSNPKIRKGAVQKVSKKLEMHVYSKRLEVMLQDIFGEDCVSV-KDDSVLSVT 638
Fly	3GRCLIPVFALGRAQELLLILDEFWSQNPDLHEIPIYYASSLAKKCMAVYQTYINAMNDR 298		VLACIMOSELGGTNLKGATKQTKSEDSRFRECLIETLODTFGDNCVPKMFEGDLLPVT 642 * : *:. : . * : . : : : : : : : : : : :
Worm	IQKQIAVKNPFIFKHVSTLRGMDQFEDAGPCVVLATPGMLQSGFSRELFESWCPDTKNGC 353	Worm	EEDGRRLLVEGSSDGPVMMGGDDPMDDPTTSHLLQNLTEKMRQIVTTNTEVNEIDDMEC 707
Human	IRKQININNPFVFKHISNLKSMDHFDDIGPSVVMASPGMMQSGLSRELFESWCTDKRNGV 352		V-DGKTANLNLETRTVECEEGSEDDESLREMVELAAQRLYEALTPVH684
Mouse Rat	IRKQININNPFVFKHISNLKSMDHFDDIGPSVVMASPGMIQNGLSRELFESWCTDKRNGV 352 IRKQININNPFVFKHISNLKSMDHFDDIGPSVVMASPGMIQNGLSRELFESWCTDKRNGV 352		V-DGKTANINLETRAVECEEGSEDDESLREMVELAAQRLYEALTPVH684 V-DGKTANINLETRAVECEEGSEDDESLREMVELAAQRLYEALTPPVH685
Fly	IRRQIAVNNFFVFRHISKLKGIDHFEDIGPCUIMASPGMMQSGLSRELFSWCTDFKNGV		V-SGKRAEINLETL-AISCAEDDVLRQULNTTVQKLHQTLVSAL684 .*: ::.: :: :: ::::::::::::::::::::::::
N	Metallo-beta-lactamase domain (Mandel et al., 2006)		Ligands to Zn1 (Mandel et al., 2006)
E	Beta-CASP domain (Mandel et al., 2006)		Ligands to Zn2 (Mandel et al., 2006)
	Bridging ligand to Zinc atoms (Mandel et al., 2006)		General acid for catalysis (Mandel et al., 2006)

cpsf-4 (CPSF-30)

N		AIR1	- C
Worm Mouse Rat Human Fly	MQLMQRIFAPCTMTDLVSNITAPMTDVEEALFNQRMKRAFFRDIDRS 50 MQEIIASVDHIKFDLEIAVEQQLGAQPLPFPGNDKS 36 	Worm FCLQGPDCQYAHPSFCLPSFENIAVSHARPTYSQAITCHHKCHERGHKATTCPHLPGOTP Mouse FCPEGPSCKFMHPRFELPMGTTE[pPLPQQTQPPTKQSNNPPLQRSSSL] Rat FCPEGPSCKFMHPRFELPMGTTE[pPLPQQTQPPTKQSNNPPLQRSSSL] Human FCPEGPSCKFMHPRFELPMGTTE[pPLPQQTQPPTKQSNNPPLQRSSSL] Fly FCPEGPSCKHMHPHFELPPLAELSKDQLHKLPTCHYCGELGHKANSKQYVGSLH *::** : :	- 204 - 204 - 204
Worm Mouse Rat Human Fly	GKAVCRKNKLKMCPFGPTCPLRHIDGEKAVVCKHWLRGLCKKGDQCEFLHEYDLTMP 106 GAAVCEFFLKAACGKGMCPFRHISGEKTVVCKHWLRGLCKKGDQCEFLHEYDMTMP 94 GAAVCEFFLKAACGKGMCPFRHISGEKTVVCKHWLRGLCKKGDQCEFLHEYDMTMP 94 IAAVCEFFLKAACGKGMCPFRHISGEKTVVCKHWLRGLCKKGDQCEFLHEYDMTMP 94 IAAVCEFFLKAACGKGMCPFRHISGEKTVVCKHWLRGLCKKGDQCEFLHEYDMTMP 94	Mouse QLTSQNSSPNQQRAPQVIGVMQSQNSSAGNRCPRPLEQVTCYKCC Rat QLTSQNSSPNQQRAPQVIGVMQSQNSSAGNRCPRPLEQVTCYKCC Human QLTSQNSSPNQQRPQVIGVMQSQNSSAGNRCPRPLEQVTCYKCC	E 250 E 250 E 250
Worm Mouse Rat Human Fly	ECFFFSKYSACSNRECPFHIDPETKMKDCPWYDRGFCRHGPYCKHRHRRACCPNYLAG ECYFYSKFGECSNKECPFLHIDPESKIKDCPWYDRGFCKHGPLCRHRHTRAVICVNYLVG 5000000000000000000000000000000000000	Mouse KGHYANRCTKGHLAFLSGQ	
	Zinc Finger 1 (Clerici et al., 2017) Zinc Finger 2, contacts A1 and A2 of PAS (Clerici et al., 2 Zinc Finger 3, contacts A4 and A5 of PAS (Clerici et al., 2 Zinc Finger 4 (Clerici et al., 2017) Zinc Finger 5 (Clerici et al., 2017) Zinc Knuckle (Clerici et al., 2017)	Contact Wdr-33 and CPSE-160	
fipp N	-1 (FIP-1) Fip1		- C

Fly	MADDTNEDSWLYGTSNPDSTTGELGNGGDSLTAEHESAAAEAOAL 45 Fly	
Worm	ADDTNEDSWLYGTSNPDSTTGELGNGGDSLTAEHESAAAEAQAL 45 Fly AV 21 Worm	SRGGLGPMAPNFPLPGASEEPFFHEPEPFDYGYEPTQESQWNNDNAGWVPSGIK 426
Human	MSAGEVERLVSELSGGTGGDEEEEWLYGGPWDVHVHSDLAKD- 42 Humar	GTAPPTSVADAPPGVDMSSDLPPGVESAPGASVAPLLPGGLDLN 315
Mouse	MSAGEVERLVSELSGGTGGDEEEEWIGGFWDHVHSDLAKD- 41 Mouse	
Rat	MSAGEVERLV-ELSGGTGGDEEEEWIGGFWDHVHSDLAK	
Nac	. :: :	GRAESPDLRRLPGAIDVIGQTITISRVEGRRRANENSNIQVLSDRSATEVDNN 351
		*. *.: : :.
Fly	AELVAGEKPGASAGTNSTEDSPRCPKEEVPEYSEFDDPAQEMEEDEDALPNTDSRSR 102 Fly	-ELTPGHAHMQQPPPGMPPPGMSVPPPQMGGPPPNLRGIMPPNMRMPPNMNMGPPPGMMM 48
Worm	APVEDSENPDVIQLDDDVIPEEDVPDDEETAEN-LENLE 59 Worm	
Human	ERPEEENASANPPSGIE 65 Huma	
Mouse	ERPEEENASANPPSGIE 64 Mous	
Rat	ERPEEENASANPPSGIE 64 Rat	FSKPPPFFPPGAPPTHLPPPPFLPPPPTVSTAPPLIPPPGFPP 39
	:*: :::*:: *	* ** :** . : ****:
Fly	RERDRERDRGRCADRADARSSPDPEDDEMSDGPAARRERNGSGSDDDEDDDSDDDINVVI 162 Fly	GGNAP-POMRMGMAPPORLAMGDRGAYDDDRERRRREKEKLLKKDOLRKDFLDMLRERHD 54
Worm	NVLDDSEIIAENSGEADQEVE-EEDNPFADDDDSDEDGGGVQVTI 103 Worm	PPPOFOOHSRAGFGPGPVGAASAPRS 37
Human	DETAENGVPKPKVT-ETEDDSDSDSDDDEDDVHVTI 100 Huma	n PPGAPPPSLIPTIESGHSSGYDSRSARAFPYGNVAFPHL 43
Mouse	EEAAENGVAKPKVT-ETEDDSDSDSDDDEDDVHVTI 99 Mous	e PPGAPPPSLIPTIESGHSSGYDSRSARAFPYGNVAFPHL 43
Rat	EEAAENGVAKPKVT-ETEDDSDSDSDDDEDDVHVTI 99 Rat	PPGAPPPSLIPTIESGHSSGYDSRSARAFPYGNVAFPHL 43
	: : * :*.:.*:::*.*	*.: * :
Fly	GDIKQAPSTYNIKQRPNLLAGGTGAAGDKAKPAGQAGKFSIEDFEGAGTINGVAVH 218 Flv	IERHTRWYDIKKKFEADPRYRALDSSYREEYFEDYLHLLKEEKRKERDLKERERHRDKER 60
Worm	RKMEPTEKPAAROGKLDLDTTATINDKPIY 133 Worm	
Human	GDIKTGAPQYGSYGTAPVNLNIKTGGRVYQTTGTKVKGVDLDAPGSINGVPLL 153 Huma	
Mouse	GDIKTGAPOYGSYGTAPVNLNIKAGGRVYGNTGTKVKGVDLDAPGSINGVPLL 152 Mous	
Rat	GDIKTGAPQYGSYGTAPVNLNIKAGGRVYGNTGTKVKGVDLDAPGSINGVPLL 152 Rat	TSSAPSWPSLVDTTKOWDYYARREKDRDRDRERDRDRER 47
	*:: .:**. :	
Fly	EFSIDSLEEKPWRKPGADITDYFNYGFNEETWRAYCERQKRFRVAESGVGLASLTQNVNQ 278 Fly	SRDKDKDKDKEKDKDKDKDKEKDKEKDKDKDKEKEKDKESSRRERS 65
Worm	DLDLAQMEDRPWRKPGADITDYFNYGFTEETWNLYCERQKKLRIEFAGNQKAANEALFSS 193 Worm	
Human	EVDLDSFEDKPWRKPGADLSDYFNYGFNEDTWKAYCEKQKRIRMGLEVIPVTSTTNKITA 213 Huma	
Mouse	EVDLDSFEDKPWRKPGADLSDYFNYGFNEDTWKAYCEKQKRIRMGLEVIPVTSTTNKITA 212 Mous	
Rat	EVDLDSFEDKPWRKPGADLSDYFNYGFNEDTWKAYCEKQKRIRMGLEVIPVTSTTNKITA 212 Rat	ERDRDRERERTRERERERDHSPTPSVFNSDEERYRYREYAERGYE-RHRASREKEE 52
	···· ··*··*******···******************	.**:*: . : ::* * <mark>:</mark> : . : . *
Fly	NAPIGILTDGGMGMGPPGMHSIQSMVGMGGESGM 312 Fly	RSREKSSRRKSKSREKDRSERSSKSSSSNTGGSSSRSEKKKSHRKDKEEDD 70
Worm	IKIANPLANPVMNTTSSVVKVLTDNGGRFKQHVHQSAAPTPLMNDQVIRTVISGNNQSAP 253 Worm	n R-RERRGDDEDRKKRSRRGDEEEESSSGRKERKEKSSR-SRHEDEESSTGVVVK 50
Human	EDCTMEVTPGAEIQDGRFNLFKVQQGRTGNSEKETAL-PSTKAEFTSPP 261 Huma	an RHRERRHREKEETRHKSSRSNSRRRHESEEGDSHRRHKHKKSKR-SKEGKEAGS 57
Mouse	EDCTMEVTPGAEIQDGRFNLFKVQQGRTGNSEKEAAL-PSTKAEFTSPP 260 Mous	se RHRERRHREKEETRHKSSRSNSRRRHESEEGDSHRRHKHKKSKR-SKEGKEAGS 58
Rat	EDCTMEVTPGAEIQDGRFNLFKVQQGRTGNSEKEAAL-PSTKAEFTSPP 260 Rat	RHRERRHREKEETRHKSSRSNSRRRHESEEGDSHRRHKHKKSKR-SKEGKEAGS 58
	* * . *. *	* ** *: *.* .* ::* : .::** * .::
Fly	OMPPPGMPPPMMOHOSRSGMGLMORPPRPISTTGGDRGGDRERAVKENAIOVMTAECREY 372 Flv	701
Worm	SLMDFTRPPPGMSMPPPM 271 Worn	n EEIPDDE 513
Human	SLFKTGLPPSRNSTSSOSOTSTASRKANSSVGKWODRY 299 Huma	
Mouse	SLFKTGLPPSRNSTSSOSOTSTASRKASSSVGKWODRY 298 Mous	
Rat	SLFKTGLPPSRNSTSSQSQTSTASRKASSSVGKWQDRY 298 Rat	EPVPEOESTEAAPAE 595
	.: ** .	

Conserved Domain contacts CPSF-30 (Clerici et al., 2017) RE/D region interacts with CFIm68/59 (Zhu et al., 2018)

CPSF Complex

pfs-2 (WDR33)

I]—	WDR40		_(Pr	o-Rich	-[
m	MTAVSNIASFNMNGGMMRGNQMPNVTLTIQPS-TSSMQNSQPRIMNNHHHPH	51	Worm	SYGGQGTLQGSYAS-	67
an	MDVSQPPPQLLTAPSALATNFTSLPPPNMGGQHYRHYHPHHGSN	44	Fly Human	DRDRERERARERERDQSGGRPNPWASNAPWSNNGNPGGNGN-GSNGFSN-NGAESF- PQGMQRHPGPH-GPLGPQGPPGPQGSSGPQGHMGPQGPPGPQGHIGPQGPPGPQGHLGPQ	76
se	MATEIGS-PPRFFHMPRFQHQAP MATEIGS-PPRFFHMPRFQHQAP	22	Mouse	PQGMQRHPGPH-GPLGPQGPPGPQGSSGPQGHMGPQGPPGPQGHIGPQGPPASQGHMGPQ	73
	MATEIGS-PPRFFHMPRFQHQAP ::.: . * : * *	22	Rat	PQGMQRHPGPH-GPLGPQGPPGPQGSSGPQGHMGPQGPPGPQGHIGPQGPPAPQGHMGPQ * . * . * * * * * :* :* .	73
n	NRFQREHVMPDVMGDGPGRRLRKNVANVRRHVDYVSTVLNHCENRLWQ	99	Worm	QQG-PPPRQIQHIPSDIDYRTAPSSSNGSTGDVDMRTMVP	71
n	<pre>KHGYNQFKPFMPGGFQRPFGMSQDDFDGKRLRKSVMRKTVDYNASIIKALENRLYQ RQLFYKRPDFAQQQAMQQLTFDGKRMRKAVNRKTIDYNPSVIKYLENRIWQ</pre>	100 73	Fly Human	-NDRDRGGGGQRRRRGINNSGPGNGGGSGGGG-GGSGRNRGRNNNRRY GPPGTQGMQGPPGPRGMQGPPHPHGIQGG-PGSQGIQGPVSQGPLMGLNPRGMQGP	80
se	RQLFYKRPDFAQQQAMQQLTFDGKRMRKAVNRKTIDYNPSVIKYLENRIWQ	73	Mouse	GPPGTQGMQGPPGPRGMQGPPHPHGIQGG-PASQGIQGPLMGLNPRGMQGP	
	ROLFYKRPDFAQQQAMQQLTFDGKRMRKAVNKKTIDYNPSVIKYLENRIWO RQLFYKRPDFAQQAMQQLTFDGKRMRKAVNKKTIDYNPSVIKYLENRIWO *: *: *!!** *: *: *!!	73	Rat	GPPGTQGMQGPPGPRGMQGPPHPHGIQG-G-PTSQGIQGPLMGLNPRGMQGP * * :: * *	71
n	YGKO-RILOOPDILYOOYAVPADSTPDVPVDCILTKFIRTAMNKVKCPVYSVCWSPEGKR	158	Worm	SG	
in	RDYRDRLALQPDSIYVPHMLPPSAYLDNPSNAVTTRFVKTATNKMRC <mark>PIFTLAWTPEGRR</mark> RDQRDMRAIQPDAGYYNDLVPPIGMLNNPMNAVTTKFVRTSTNKVKC <mark>PVFVVRWTPEGRR</mark>	160	Fly Human		
in se	RDQRDMRAIQPDAGIINDLVFFIGMLNNPMNAVITIKFVRISTNKVKCFVFVVRWTPEGRR	133	Mouse	PGPRENQGPAPQGMIMGHPPQEMRGPHPPGGLLGHGPQEMRGPQEIRGMQGPPPQGSMLG PGPRENQGPAPQGLMIGHPPQEMRGPHPPSGLLGHGPQEMRGPQEMRGMQGPPPQGSMLG	84
	RDQRDMRAIQPDAGYYNDLVPPIGMLNNPMNAVTTKFVRTSTNKVKC <mark>PVFVVRWTPEGRR</mark> .: *** * :* .: * :.: *::*: **::* **::*	133	Rat	PGPRENQGPAPQGIMIGHPPQEMRGPHPPSGLLGHGPQEMRGPQEMRGMQGPPPQGSMLG	84
n	LITGCQTGEFTLWNGTAFNFETILQAHDSAIRALKWASNEQWLLSAD	218	Worm	EPEHWRGPPPVSHQQQSQQQHPPPI-NMQRMDPRRDPRMLSGRSD	76
an	LVTGASSGEFTLWNGLTFNFETILQAHDISVRTMVWSHNDSWMVTGDHGGYVKYWQSNMN LVTGASSGEFTLWNGLTFNFETILQAHDSPVRAMTWSHNDMWMLTADHGGYVKYWQSNMN	220	Fly Human	PPQELRGPPGSQSQQGPPQGSLGPPPQGGMQGPPGPQGQQNPARGPHPSQGPIPFQQQKT	
in se	LVTGASSGEFTLWNGLTFNFETILOAHDSPVRAMTWSHNDMWMLTADHGGYVKYWOSNMN	193	Mouse	PPQELRGPSG SQGQQGPPQGSLGPPPQGGMQGPPGPQGQQNPARGPHPSQGPIPFQQQKA	90
	LVTGASSGEFTLWNGLTFNFETILQAHDSPVRAMTWSHNDMWMLTADHGGYVKYWQSNMN *:**:******** :**********************	193	Rat	PPQELRGPSGSQGQQGPPQGSLGPPPQGGMQGPPGPQGQQNPARGPHPSQGPIPFQQQKA	90
n	NAHMFSAHKDEAIRGLAFAPTDVKFATASDDGTARVWDFARYTEERVLRGHGAEVRCIDW	278	Worm	QLSPSGPPPQQQQSSAQQGGKNQWMPQFEAGQNQNQGNYAGNRGGGV	80
an	NVKMYQAHK-EAIRGISFSPTDSKFVSGSDDGTLRIWDFMRCQEERVLRGHGADVKCVHW NVKMFQAHK-EAIREASFSPTDNKFATCSDDGTVRIWDFLRCHEERILRGHGADVKCVDW	279 252	Fly Human	PLLGDGPRAPFNOEGOSTGPPPLIP-GLGQO-GAOGRIPPLNPGOGPG	
se	NVKMFQAHK-EAIREASFSPTDNKFATCSDDGTVRIWDFLRCHEERILRGHGADVKCVDW	252	Mouse	PLLGDGPRAPFNQEGQSTGPPPLIP-GLGQQ-GAQGRIPPLNPGQGPG	94
	NVKMFQAHK-EAIREASFSPTDNKFATCSDDGTVRIWDFLRCHEERILRGHGADVKCVDW *.:*:.*** **** :*:*** **.: ***** *:**** * ***:******:*:*:*	252	Rat	PLLGDGPRAPFNQEGQSTGPPPLIP-GLGQQ-GAQGRIPPLNPGQGPG	94
n	HPTKGLIATGSRDTQQPVKIWDPKSGSCLATLQEHKSSVMAVEFNKNGNWLLTGGRDHLV			GGRGRGQPY	
an	HPQKGMIVSGSKDNQQPIKIWDPKSGIALATLHAHKSTVMDLKWNDNGNWLVTASRDHLL HPTKGLVVSGSKDSQQPIKFWDPKTGOSLATLHAHKNTVMEVKLNLNGNWLLTASRDHLC	339 312	Fly Human	- PNKGDSRGPPNHHMGPMSERRHEQSGGPEHGPERGPFRGGQDCRGPPDRRGPHPDFPDDF	80
se	HPTKGLVVSGSKDSQQPIKFWDPKTGQSLATLHAHKNTVMEVKLNLNGNWLLTASRDHLC	312	Mouse	PNKGDTRGPPNHHLGPMSERRHEQSGGPEHGPDRGPFRGGQDCRGPPDRRGSHPDFPDDF	10
	HPTKGLVVSGSKDSQQPIKFWDPKTGQSLATLHAHKNTVMEVKLNLNGNWLLTASRDHLC ** **::.:**:*:*:*:*:*:*:*:*:*:*:*:*:*:*:	312	Rat	PNKGDSRGPPNHHLGPMSERRHEQSGGPEHGPDRGPFRGGQDCRGPPDRRGSHPDFPDDF	10
m	KMYDIRMMK-EMRTYRAHKKEVISLAWHPIHEGLFVSGGGDGSIVYWMVDGEKEIGLLEH	397	Worm		
an	KLFDIRNLREEVQVFRGHKKEASSVSWHPIHEGLFCSGGSDGSILFWNVGTDKEIGCVET KLFDIRNLKEELQVFRGHKKEATAVAWHPVHEGLFASGGSDGSLLFWHVGVEKEVGGMEM	399	Fly	SRPDDFHPDKRFGHRLREFEGRGGPLPQEEKWRRGGPGPPFPPDHREFSEGDGRGAARGP	
se	KLFDIRNLKEELQVFRGHKKEATAVAWHPVHEGLFASGGSDGSLLFWHVGVEKEVGGMEM	372	Mouse	-RPDDFHPDKRFGHRLREFEGRGGPLPQEEKWRRGGPGPPFPPDHREFNEGDGRGAARGP	10
	KLFDIRNLKEELQVFRGHKKEATAVAWHPVHEGLFASGGSDGSLLFWHVGVEKEVGGMEM *::*** :: *::.:*.***: :::****** ***.**** ***.***::* *. :**:* **	372	Rat	SRPDDFHPDKRFGHRLREFEGRGGPLPQEEKWRRGGPGPPFPPDHREFNEGDGRGAARGP	10
m	AHDQAIWSMKWHPLGHILATGSNDNNTKFWARNRPGDTVEDIFGLSNTNMIG-HLDKERE	456	Worm		83
an	AHDSIVWTLAWHPLGHILCSGSNDHTIKFWTRNRPGDLMRDKYNLNTLPASLAALDECEY AHEGMIWSLAWHPLGHILCSGSNDHTSKFWTRNRPGDKMRDRYNLNLLPGMSEDGVEY	459	Human	SRPDDFHPDKRFGHRLREFEGRGGPLPQEEKWRRGGPGPPFPPDHREFSEGDGRGAARGP	1
se	AHEGMIWSLAWHPLGHILCSGSNDHTSKFWTRNRPGDKMRDRYNLNLLPGMSEDGVEY	430	Mouse	-RPDDFHPDKRFGHRLREFEGRGGPLPQEEKWRRGGPGPPFPDHREFNEGDGRGAARGP SRPDDFHPDKRFGHRLREFEGRGGPLPQEEKWRRGGPGPPFPDHREFNEGDGRGAARGP	
	AHEGMIWSLAWHPLGHILCSGSNDHTSKFWTRNRPGDKMRDRYNLNLLPGMSEDGVEY **: :*:: ********.: ***: ***: .**: *:****** :.* :.	430	Rac	SKEDUT NEDIKT GIKLIKET EGROGELEVEENKKOOFGETT FEDIKET NEODOKOAAKGE	1
m	PRMAPPKPSIETQETYRPDTFIPGMGLDEHLYEQLNRDHNMMTTDSTLLVPD-DLTRQ	513	Worm		8
an	DTADKGFIPGLDLDPS DDLEPNSLAVIPGMGIPEQLKLAMEQEQMGKDESNEIEMTIPGLDWGME	494	Human	PGAWEGRRPGDERFPRDPEDPRFRGRREESFRRGAPPRHEGRAPPRGRDGFPGPEDFGPE	
se	DDLEPNSLAVIPGMGIPEQLKLAMEQEQMGKDESSEIEMTIPGLDWGME	479	Mouse	eq:pgawegrrpgddrfprdpddrfrgrreesfrrgapprhegrapprgrdnfgpe	
	DDLEPNSLAVIPGMGIPEQLKLAMEQEQMGKDESNDIEMTIPGLDWGME :.***** ::: :: :: :: :: :: :: ::	479	Rat	PGAWEGRRPGDERFPRDPDDPRFRGRREESFRRGAPPRHEGRAPPRGRDNFPGPDDFGPE	1
m	NFAPMIGAKRTLIKQPPAKKAQRQFERMWNNSKGIGAGSDD-FTTMKGGLGREDAEG	569	Worm		8
an	KAINDRDREKKVPYSKPIPRNFQVNWAGPRRADDPSV-L-SIHEELAAREAKDTTN EVMQKDQKKVPQKKVPYAKPIPAQFQQAWMQNKVPIPAPNEVLNDRKEDIKLEEKKKTQA	548 539	r⊥y Human	ENFDASEEAARGRDLRGRGRGTPRGG-RKGLLPTPDEFPRFEGGRKPDSWDGNREPGPGH	8 1
se	EVMQKDQKKVPQKKVPYAKPIPAQFQQAWMQNKVPIPAPNEVLNDRKEDIKLEEKKKTQA EVMQKDQKKVPQKKVPYAKPIPAQFQQAWMQNKVPIPAPNEVLNDRKEDIKLEEKKKTQA	539	Mouse	EGFDASDEAARGRDLRGRGRGTPRGGSRKCLLPTPDEFPRFEGGRKPDSWDGNREPGPGH EVFDASDEAARGRDLRGRGRGTPRGGSRKCLLPTPDEFPRFEGGRKPDSWDGNREPGPGH	1
	* * :* : * : * :				
n	AQFGPSKSFLGPPTTGGSLLGP GLNHQQISENTMEGILASGMLNGLDPQTIVGVLVYNRFIRVHPDSRLMAAIRQGAEFLNK	608	Fly		8
an	EIEOEMATLOYTNPOLL-EOLKIERLAOKOVE-OIOPPPSSGTPLLGP	585	Human	EHFRDTPRPDHPHDGHSPASRERSSSLQGMDMASLPPRKRPWHDGPGTSEHREMEAPGG	
se	EIEQEMATLQYTNPQLL-EQLKIERLAQKQAD-QIQPPPSSGTPLLGP EIEQEMATLQYTNPQLL-EQLKIERLAQKQAD-QIQPPPSSGTPLLGP	585	Rat	EHFRDAPRPDHPPHDGHSPASRERSSSLQGMDMASLPPRKRPWHDGSGTSEHREMEAQGG EHFRDAPRPDHPPHDGHSPASRERSSSLQGMDMASLPPRKRPWHDGSGTSEHREMEAQGG	1:
	: . : * :*.				
m	SQPRPQEFRSVPPPQQQQGPPPNW		Worm Fly		
an	OPFPGOGPMSOIPOGFOOPHPSOOMPMNMAOMGPPGPOGOFRPPGPOGOMG	636	Human	PSEDRGGKGRGGPGPAQRVPKSGRSSSLDGEHHDGYHRDEPFGGPPGSGTPSRGGRSGSN	
se	OPFSGQGPISQIPQGFQ0PHPSQOMPL-VPOMGPPGPQGOFRAPGPQGQMG QPFSGQGPMSQIPQGFQ0PHPSQOMPL-VAQMGPPGPQGQFRAPGPQGQMG	635 635	Mouse Rat	PSEDRGSKGRGGPGPSQRVPKSGRSSSLDGDHHDGYHRDEPFGGPPGSSSSSRGARSGSN PSEDRGSKGRGGPGPSQRVPKSGRSSSLDGDHHDGYHRDEPFGGPPGSSSSSRGGRSGSN	
m	PGQGYGPPGRMPGQGYGPPGRM	635	Norm	820	
-m				807	
	POCEDE HOCCCCEDOCENCEDOCEDOCE DE DODNIC	672	Human	WGRGSNMNSGPPRRGASRGGGRGR 1336	
an se		675	Mana	WGRGSNMNSGPPRRGTSRGSGRGR 1330	

N-terminal domain, contacts CPSF-160 and CPSF-30 (Clerici et al., 2017)

WD40 repeats, contact U3 and A6 of PAS (Clerici et al., 2017)

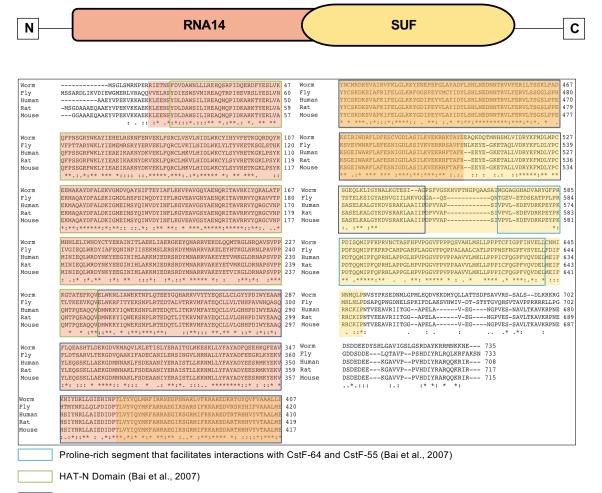
Г

CstF Complex

cpf-1 (CstF50)

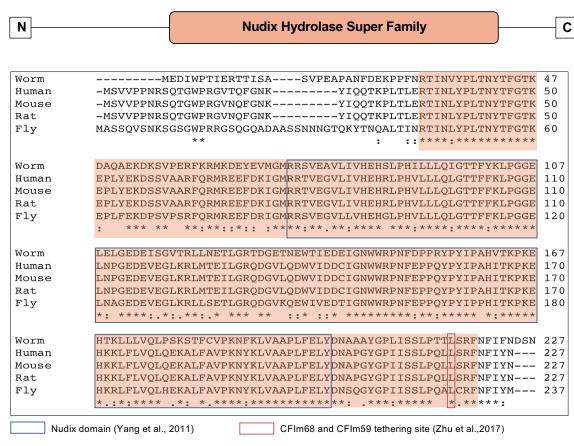
N	WD40 Domain	<u> </u>	(CSTF1 D	imer	
Human Mouse Rat		ANGLINEIKPQSV-CAPSEQLLHLIK ANGLINEIKPQSV-CAPSEQLLHLIK ANGLINEIKPQSV-CAPSEQLLHLIK YAMELSM-LVKADQ-CAPSERLLHVMI	55 1 55 1 55 1	Worm Human Mouse Rat Fly	LVATDHPTVRLYNIETAQAYASAN LVGTQHPTLRLYDINTFCCFVSCN LVGTQHPTLRLYDINTFCCFVSCN LVGTQHPTLRLYDINTFQCFVSCN AIGTEHNVLRYYDVATTCCFVSCN :*:**:	PQDQHTDAICSVNYNSSANMY PQDQHTDAICSVNYNPSANMY PQDQHTDAICSVNYNPSANMY PSQQHKAGVTCVKYSPTGKLY	YVTGSKDGCIKLWDGV YVTGSKDGCIKLWDGV YVTGSKDGCIKLWDGV YATGSYDGDIKIWDGI
Human Mouse Rat	IAKOVEDPDYKAEEKSSPMQFEPVSSGLDLEYI LGMENDDTAVYYAIGRSDTVAPGTGIDLEFI LGMENDDTAVYYAIGRSDTVAPGTGIDLEFI LGMENDDTAVYAIGRSDTVAPGTGIDLEFI LGMENDDTAVYAIGRSDDVLFGIDLEFI	DADVQPVSSEPSEYETIYLTVHKAPCR DADVQTMSPEASEYETCYVTSHKGPCR DADVQTMSPEASEYETCYVTSHKGPCR DADVQTMSPEASEYETCYVTSHKGPCR	112 I 112 I 112 I	Worm Human Mouse Rat Fly	SNRCVETFKRAHDGSSICSAKFTK SNRCITTFEKAHDGAEVCSAIFSK SNRCITTFEKAHDGAEVCSAIFSK SNRCITTFEKAHDGAEVCSAIFSK SGRCITTIAEAHGGAAICSLEFTR *.**: *: .**.*: :** *:	NGKYILTSGMDSIVKLWELST NSKYILSSGKDSVAKLWEIST NSKYILSSGKDSVAKLWEIST NSKYILSSGKDSVAKLWEIST NGKYLLSSGMDSLVYLWELCT	TNRCLIVYTGAGATGA TGRTLVRYTGAGLSGR TGRTLVRYTGAGLSGR TGRTLVRYTGAGLSGR TSRPIQTYTGAGTTGK
Human Mouse Rat	AAAPNSDGSLVATGSADCSIKIMDVERILAREKH VATYSRDQQLIATGSADASIKILDTERMLAKSAN VATYSRDQQLIATGSADASIKILDTERMLAKSAN VATYSRDQQLIATGSADASIKILDTERMLAKSAN AGAPSCDGSLVATGSVDASIKILDVERMLAKSAN	IPIEVMMNETAQQNMENHPVIRTLYDH IPIEVMMNETAQQNMENHPVIRTLYDH IPIEVMMNETAQQNMENHPVIRTLYDH	172 172 172	Worm Human Mouse Rat Fly	QDFATNASPNHNEDYYLFPDEKSG QVHRTQAVFNHTEDYYLLPDERTI QVHRTQAVFNHTEDYILLPDERTI QVHRTQAVFNHTEDYILPDERTI OEHQTEAVFNHTEDYYLFPDERTI * *:* *******	SLCCWDSRTAERRNLLSLGHN SLCCWDSRTAERRNLLSLGHN SLCCWDSRTAERRNLLSLGHN SLCSWNSRNGCRLTLNSLGHN	NNIVRCIVHSPTNPGF NNIVRCIVHSPTNPGF NNIVRCIVHSPTNPGF NGPVRYITHSPNGPAF
Human Mouse Rat Fly	VDDVNTVIFHPRDSILISGSNDKTVKLFDFSKT VDEVTCLAFHPTEQILASGSRDYTLKLFDYSKPS VDEVTCLAFHPTEQILASGSRDYTLKLFDYSKPS VDEVTCLAFHPTEQILASGSRDYTLKLFDYSKPS TDEVSVLFHPKEHILASASRDGTVKLFDIAKP .*:*. : *** : ** *.**	SAKRAFKYIQEAEMLRSISFHPSGDFI SAKRAFKYIQEAEMLRSISFHPSGDFI SAKRAFKYIQEAEMLRSISFHPSGDFI SVKKAHKVFTDCEPVLCLSFHPTGDYV	232 232 232	Worm Human Mouse Rat Fly	MTGSDDHRARFWYRKPATNEH MTCSDDFRARFWYRRSTTD MTCSDDFRARFWYRRSTTD LTCSDDFRARFWYRRSTTD LTCSDDFRARFWYRRSTNO ;* ***, *******;	430 431 431 431 424	
<i>cpf-2</i> N	RRM SF	CSTF2	Hing	ge)	CST	IF C
N Worm Fly	RRM SF	TIRSIFSKAGNVLSIKMVHDRETGKPM KLKEIFSEVGPVLSLKLVFDRESGKPM	(G 60 (G 58	Worm Fly	-MGHPQAPQYGQNY GRGMDQDLRASLPNPVP	GQPV	/APQQYKPPPQQQI
Worm Fly Human Mouse Rat	RRM SF	TIRSIFSKAGNVLSIKMVHDRETGRP KLKEIFSEVGPVLSLKLVFDRESGRP JLKDIFSEVGFVVSFRLVVDRETGRP JLKDIFSEVGFVVSFRLVVDRETGRP LLKDIFSEVGFVVSFRLVVDRETGRP	G 60 G 58 G 58 G 58 G 58 G 58 S 58 S 58 S 58	Worm Fly Human Mouse Rat	GRGMDQDLRASLPNPVP PMQDPRAAMQRGSLPANVPTPRGI PMQDPRAAMQRGALPTNVPTPRGI PMQDPRAAMQRGALPANVPTPRGI		JAPQQYKPPPQQQ PPCMD /EPRGYLGPPHQG /EPRAYLGPPPPHQG /EPRAYLGPPPPPHQG *
Worm Fly Human Mouse	RRM SF 	TIRSIPSKAGNULSIKMVHDRETGKPP KILKEIPSEVGPULSILKLVFDRETGKP JLKDIFSEVGPUSFLLVDRETGKP JLKDIFSEVGPUSFLLVDRETGKP ILKDIFSEVGPUSFLLVDRETGKP ILKDIFSEVGPUSFLUVDRETGKP ILKDIFSEVGPUSFLUVDRETGKP MACTEKSRMEMGQL-LQG-PQVENYU MASEKNKELKSL-GTGAPULSPVC MASEKNKELKSL-GTGAPULSPVC	G 60 G 58 G 58 G 58 G 58 G 58 F* 3P 120 3E 116 3E 117 3E 117	Worm Fly Human Mouse Rat Worm Fly Human Mouse	GRGMDQDLRASLPNPVP PMQDPRAAMQRGSLPANVPTPRGI PMQDPRAAMQRGALPTNVPTPRGI		/APQQYKPPPQQQ PPGQI /EPRAYLGPPHQGI /EPRAYLGPPPPHQG /EPRAYLGPPPPHQG -LC
Worm Fly Human Mouse Rat Worm Fly Human Mouse	RRM SF	TIRSIPSKAGNULSIKMUHDRETGKPP KILKEIPSEVGPULSLKLVFDRESGKPP JLKDIFSEVGPUSFLLVYDRETGKP JLKDIFSEVGPUSFLLVYDRETGKP :	G 60 G 58 G 58 G 58 G 58 G 58 110 117 117 117 117 117 117 117 117 117 117 117 117 117 117 117 117 117 117 117	Worm Fly Human Mouse Rat Worm Fly Human Mouse Rat Worm Fly Human Mouse	GRGNDODLRASLPNPVP PMODPRAAMQGGSLAVVPTPRGI PMODPRAAMQRGALPANVPTPRGI PMODPRAAMQRGALPANVPTPRGI PVOMRPPVQQ R		/APQQYKPPPQQQ PPGQD /EPRAYLGPPPPHQG /EPRAYLGPPPPHQG /EPRAYLGPPPPPHQG /EDRAGGRDPRGIDARG /LDARGGRDPRGIDARG /LDARGGRDPRGIDARG /LDARGGRDPRGIDARG /LDARGGRDPRGIDARG /LARGGRDPRGIDARG /LARGGRDPRGIDARG /LARGGRDPRGIDARG /LARGARGMDTRGPPG /LARGARGMDTRGPPG
Worm Fly Human Mouse Rat Worm Fly Human Mouse Rat Worm Fly Human Mouse	RRM SF	TIRSIFSKAGNVLSIKHVHDRETGKPF KLKEIFSSVGPVLSLKLVPDRESGKP JLKDIFSSVGPVVSFRLVVDRETGKP JLKDIFSSVGPVVSFRLVVDRETGKP JLKDIFSSVGPVVSFRLVVDRETGKP SAAGGMINHEEFGSSSNAPAPVE]NPYC NACTEKSMEHQD-L0G-P0VENYC NAASEKNEKELSL-GTGAPVIESPYC NAASEKNEKELSL-GTGAPVIESPYC NAASEKNEKELSL-GTGAPVIESPYC NAASEKNKELGKSL-GTGAPVIESPYC SLKNNPSELKKELVEHPOIAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLLQNPOLAYALLOAC CVGNSPQEANNLQNPOLAYALLOAC CVGNSPQEANNLQNPOLAYALLOAC CVGNSPQEANNLQNPOLAYALLOAC CVGNSPQEANNLQNPOLAYALLOAC CVGNSPQEANNLQNPOLAYALLOAC CVGNSPQEANNLQNPOLAYALLOAC CVGNSPQEANNLQNPOLAYALLOAC CVGNSPQEANNLQNPOLAYALLOAC CVGNSPQEANNLQNPOLAYALLOAC	G 60 G 58 G 58 G 58 G 58 G 58 G 58 C 58 C 58 C 58 C 58 C 58 C 58 C 58 C	Worm Fly Human Mouse Rat Worm Fly Human Mouse Rat Worm Fly Human Mouse Rat	GRGNDODLRASLPNPVP PMODPRAAMQGGLAVNVPTPRGI PMODPRAAMQGGALPANVPTPRGI PMODPRAAMQRGALPANVPTPRGI PVONRPPV00RAQMPP000GVP00 PMHHVPGHESRGPPHDRRGGPLJ PMHHVPGHEGRGPPPHDMRGGPLJ PMHHVPGHEGRGPPPHDMRGGPLJ 		ZAPQQYKPPPQQQ PPGQI /EPRAYLGPPPPHQGI /EPRAYLGPPPPHQGI /EPRAYLGPPPPPHQGI /EPRAYLGPPPPPHQGI /EPRAYLGPPPPHQGI /EDRGGRDPRGLDARG /LDARG

suf-1 (CstF77)



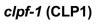
HAT-C Domain that interacts with CPSF-160 (Bai et al., 2007)

cfim-1 (CFim25)



cfim-2 (CFim68)

N —	RRM CFIm68	PABP-1234
Worm Fly Mouse Human Rat	MAELD-EAALLGGGNEQHDGPIDENALLDGKELKEEDIDDLYDEAIAPTNSTESAKPVS-58 Worm MADUVLDLYAEDLDKDFAGQAQDEFGGGGVDLYDDIGGPTESAASGGGGGG Fly BAEYGGHDJDLYDDVISFSANNGAPEDR-30 Mouse MADGVDHIDIYAD	LGSMGVQPLMQNNTAM-RPPINGLPPV
Worm Fly Mouse Human Rat	-PTITTVTAPTAGIGAKPATSSEGR#YCCYVGNLLWYTTDADLLK TPSADGAAGPGSGEPGERNSGGPNGVIHQSSGSLTPTMNREYQLYVGNLTWWTTDQDIAN -PYMD-TLPPTVGDD-VGKAAPNVVYTYTGKIALVIGNLTWWTTDEDLTE -DYMD-TLPPTVGDD-VGKGAAPNVVYTYTGKIALVIGNLTWWTTDEDLTE 99 Human -PYMD-TLPPTVGDD-VGKGAAPNVVYTYTGKIALVIGNLTWWTTDEDLTE *	HVNPQNFPGLQGTVLSDAEFEDVMTRNQTVSSSAIARAITDASVGDIKG 405 PRGPWPPPQGKPPGPFPDPQQMGPQLTEVEFEEVMSRNRTVSSSAIARAVSDAAAGEYSS 477 RPPPVDRGDVGPPGREMDTARTPLSEAEFEEIMNRNRAISSSAISRAVSDASAGDYGS 423 RPPPVDRGDVGPPGREMDTARTPLSEAEFEEIMNRNRAISSSAISRAVSDASAGDYGS 443 RPPPVDRGDVGPPGREMDTARTPLSEAEFEEIMNRNAISSSAISRAVSDASAGDYGS 443 RPPPVDRGDVGPPGREMDTARTPLSEAEFEEIMNRNAISSSAISRAVSDASAGDYGS 443 RPPPVDRGDVGPFGREMDTARTPLSEAEFEEIMNRNAISSSAISRAVSDASAGDYGS 443
Worm Fly Mouse Human Rat	ALQSTGLARSQFADMKFFENRTNQQSKGYALLVLNSDAAVKQTMEILPTKTIHGQSPTVL SLADIGUSDLGEVKFFENRANQQSKGFSVISLGSESSLAAVLOQLPKKEMGQAPVVT 169 Fly AVMSLGVNDILEIKFFENRANQQSKGFALVGVGSEASSKKIMDLLPKRELHGQNPVVT 157 Human AVMSLGVNDILEIKFFENRANQGSKGFALVGVGSEASSKKIMDLLPKRELHGQNPVVT 157 Rat	ASETIIPAIQLIKNSSIGHDERCRQLVYGLEHPIAGLESKGYSSRGKKSHRDRSRSBERDR 465 AIETIVTAISLIKQSRVAHDERCRILISSLQDTLGIELAKSYNRRÅRS-RSRERSHRS-R55 AIETIVTAISLIKQSRVSADDRCKVLISSLQDCLHGIESKSVGSGÅRERSRERNHSRR 50 AIETIVTAISLIKQSRVSADDRCKVLISSLQDCLHGIESKSVGSGÅRERSRERRHSRSR 50 AIETIVTAISLIKQSRVSADDRCKVLISSLQDCLHGIESKSVGSGÅRERSRERRHSRSR 50 AIETIVTAISLIKQSRVSADDRCKVLISSLQDCLHGIESKSVGSGÅRERSRERRHSRSR 50 AIETIVTAISLIKQSRVSADDRCKVLISSLQDCLHGIESKSVGSGÅRERSRERRHSRSR 50 AIETIVTAISLIKQSRVSADDRCKVLISSLQDCLHGIESKSVGSGÅRERSRERRHSRSR 50 AIETIVTAISLIKQSRVSADDRCKVLISSLQDCHGIESKSVGSGÅRERSRERRHSRSR 50 AIETIVTAISLIKQSRVSADDRCKVLISSLQDCHGIESKSVGSGÅRERSRERRHSRSR 50 AIETIVTAISLIKQSRVSADDRCKVLISSLQDCHGIESKSVGSGÅRERSRERRHSRSR 50 AIETIVTAISLIKQSRVSADDRCKVLISSLQDCHGIESKSVGSGÅRERSRERRHSRSR
Vorm 71y 4ouse 4uman Rat	SYNKTNQAKLEDAQAKNOTRPDVKKKGFEGCVNMGTIRIGAGGQTGR 210 WOrm YPSKQALTQFESLQKTRHVPPQQNGPFRGPAP5MGGGPMPTGHPGGP 218 Fly PCNKQFLSQFEMQSKRTTQSGGMSGEGKAGPPCGGSRAAFPQGGGRGKGRFPGAVPGGD 215 Human PCNKQFLSQFEMQSRKTTQSGQMSGEGKAGPPCGGSRAAFPQGGRGKGRFPGAVPGGD 215 Rat	XRRRRSPTBERSYSRSYSREPSPRRRY
Norm Fly Mouse Human Rat	TGTSVSGRSGPPPLMMQQVRPTPLMSQPTSLPSNLN 246 Worm QGGRPGHPPRGNNSIMQCQVRQUGVPQVGGPNSGPPRQPPRHPQGGLMGNQQ 276 Fly RPFGRAGGPPPPPRAGOTPPPPLGPPC-PCPC 4251 Mouse RPFGPAGFGGPPPPFPAGOTPP	489 SRSVNPRKSPEPVVAEAAEAPSSKRYYEDRERYRSSDRERRDRDRDRDRENEDRDRREE 647
Yorm Yly Mouse Muman Rat	QAPQMRLQINGQ0VPLMNR-APVPQQTMLGNGPLGGNNQ0IQQ0PQMMMGQQ 297 Worm PPPRYESAGGWFCQRPCGPRGPPNGPPQRPHPQGFMCMVPKRPAGED 326 Fly PPPCQVLPPFLAGPPNUFPCQPFCQGPLGPLPCQ 75 Mouse PPPPGQVLPPFLAGPPNKGDRPPPVLFPQ0FCQPFLGPLPCQ 295 Human PPPPGQVLPPFLAGPPNKGDRPPPVLFPQ0FCQPFLGPLPCQ 295 Rat	489 HRBRH 652 HR 531 HR 551 HR 551
lorm 'ly louse luman tat	VRPMMQTSMGIQPMMGNNAPPPMNNQFQNRPPQGPRGLPPAPGPGGPHGA 357 NPPPVPGVGPPPGPPPPMGGFPPPGPFQPCPLCPPLTLAPPHLPGPPGAPP 332 PPPPVPGVGPPPCPPPQGGPPPPFGPPPRFQPCLGPPLTLAPPHLPGPPGAPP 352 PPPPVFGVGPPPGPPPPGGPPPPFGPPPPGGPPGPPGPPGPPGPP352 **	



N	CLP1 Super	Family	CLP1 P	C
Worm Fly Human Rat Mouse	MSEDCKDYTLESDSELFFEIEQKDAKUVSLVSGFAELFGYELVKKRYFFGV 54 MGEEANDDKKFTTKFELERFELRFEVEASQ-SVQLELLTGMAEIFGTELTRNKFFFDA 59 MSEESNDDKKTTKFFELERFETELRFEVEASQ-SVQLELLAGMAEIFGTELTRNKFFFDA 59	Fly VKGSGYA Human VKGSGYQ Rat VKGYGYQ Mouse VKGYGYQ	HLLHAAKAYGACAIFVLDQERLYI ALVHAASAFEVDVVVVLDQERLYI ALVHAASAFEVDVVVVLDQERLYI ALVHAASAFEVDVVVVLDQERLYI	SDISKELPEVURLTHUVEKSGUVERNICTE 293 HELLROVEKUNULLEKSGUVERSKELR 288 HELROLHFVRTVLLPKSGUVERSKELR 293 HELKROLHFVRTVLLPKSGUVERSKELR 293 IIIIK IIIK IIIIK 293
Worm Fly Human Rat Mouse	KSRVAAFTWKGATIELUGTTESAVVABSTPMVIYLNIHAAMEEVPKKKAEDAAGNSNKKK GAKVAIFTYGGVLHVSGKMDVYISKETPMVQVVNCHAALEQFRMEATEKDFY GAKVAVFTWHGCSVQLSGKTEVAVVSKDTPMLIVINTHTALEQNRRQAIKEER 113 GAKVAVFTWHGCSLQLSGKTEVAVVSKDTPMLIVINTHTALEQNRRQAIKEER 113 GAKVAVFTWHGCSLQLSGKTEVAVVSKDTPMLIVINTHTALEQNRRQAIKEER 113 ::***::::::::::::::::::::::::::::::	Fly HEARDQR Human RECRDER Rat RECRDER Mouse RECRDER	IKEYFYGNTRAPFYPFSFEVKFQI IREYFYGF-RGCFYPHAFNVKFSI IREYFYGF-RGCFYPHAFNVKFSI IREYFYGF-RGCFYPHAFNVKFSI	JVTLCKIGAEQLEDSCLEPEGNEVENHETKI 353 JLRLYKIGAPELPDSCHPIGHKAEDNKTKV 348 JVKIYKVGAPTIEDSCLEPLGNSQEDNQLKI 352 JVKIYKVGAPTIEDSCLEPLGNSQEDNQLKI 352 JVKIYKVGAPTIEDSCLEPLGNSQEDNQLKI 352 JVKIYKVGAPTIEDSCLEPLGNSQEDNQLKI 352 JVKIYKVGAPTIEDSCLEPLGNSQEDNQLKI 352 JVKIYKVGAPTIEDSCLEPLGNSQEDNQLKI 352
Worm Fly Human Rat Mouse	GPRLLLVGPTDVGKTTVSRILCNYSVRQGRTPIFVELDVGQNSVSVPGTVAAVLVQKTAD 173 GPVAMVVGPNDVGKSTLCKILLNYAVKVGRRPLYADLDVGQGSVATLIERPAD 168 GPRVMVVGPTDVGKSTVCLLLNYAVKVGRRPTVVELDVGQGSVSIPGTVGALYIERPAD 173 GPRVMVVGPTDVGKSTVCLLLNYAVKLGRRPTVVELDVGQGSVSIPGTVGALYIERPAD 173 GPRVMVVGPTDVGKSTVCLLLNYAVKLGRRPTVVELDVGQGSVSIPGTVGALYIERPAD 173 GPRVMVVGPTDVGKSTVCLLLNYAVKLGRRPTVVELDVGQGSVSIPGTVGALYIERPAD 173 ** :::**********************************	Fly VAVTPTP Human VPVTPGRI Rat VPVTPGRI Mouse VPVTPGRI	ALIHHVLALSFAESVEDDVIGTN\ DMVHHLLSVSTAEGTEENLSETS\ DMVHHLLSVSTAEGTEENLSETS\	JFGFCLVTEVDLEKRTMSILCPQRT-IP-S 411 AGGFCCVTEVDHERQAVMLLSPQPRLIPPN 408 AGGFUVTSVDLEHQVTTVLSPAPRPLIPKN 412 AGGFUVTSVDUEHQVTTVLSPAPRPLIPKN 412 *** **.**:*:::*. * * .
Worm Fly Human Rat Mouse	VEEGFSIQAPLVYHFGSTTPGTNIKLYNKITSRLADVFNQRCEVNRRASVSGCVINTCGW 233 1	Fly ALLLWSE Human FLLI-MD Rat FLLI-MD	TTHLDDQIKR 428 LQFMDNHF 423 TRFMDIK 425 IRFMDIK 425 IRFMDIK 425 IRFMDIK 425	
	N-terminal domain (Dikfidan et al., 2014)			
	Polynucleotide kinase domain (Dikfidan et al., 2014)			
	C-terminal domain (Dikfidan et al., 2014)			
	Residues that crosslink to Pcf-11 (Schäfer et al., 2018)			

1-	COG5222 SOBP)		PTZ00121 SF-C11
orm ly	MSSIHYKFRAELDYKTLQFDGLHIRGEQLVREICAKENL-KLELFELQLQNA -MSYHYKFKSTLRPDTIFDGLHISVGDLKREIVQQKRLGKIIDFDLQITAN	QSKEEYKD 59	Worm Fly	SSTINTADEDDESSKKMKKHKKSKKNKKHHRKEEDGDEDEERKRKHKKHKKEKKSKKEK SSKGDSSKKRGENRHEAPRKRHRSRSISKEPK
uman at ouse	MSCUHYKFSSKLNYDTVFFOGLHISLCDLKKQIMGREKL-KAADCDLQTTMA MSCUHYKFSSKLNYDTVFFOGLHISLCDLKKQIMGREKL-KAADSDLQITMA MSCUHYKFSSKLNYDTVFFOGLHISLCDLKKQIMGREKL-KAADSDLQITMA .:**** : *::.*: ****** :* ::* ::* ::* ::	QTKEEYTD 59 QTKEEYTD 59	Human Rat Mouse	NQKDNYKSKEKESEN
orm Y iman it ouse	D-ELIPENSSIIVQRFPRKDAAKVQKVQAGVNSGMVNQLDATSSFLDPSSHI DGELIPENTTLIISRIPIAHPFKGWEPPAAENAFSA	PAKQDNFN 105 SLAQLTKT 117 T 102	Worm Fly Human Rat Mouse	RDEDEDOLDTEKKEKK
orm -y	FENMDEAERLNHIRDQSTRAYDQSNFRRRQPGIMTGPPPPTYTCNRCS MDLSKNQGTEEDKIQAMMMQSTVDYDPKTYHRIKGQSQVGEVPASYRCNKCK	KSGHWIKN 165	Worm Fly Human	SLDEDKVELDKNFADKKEKKKKWNEDEEDIFEDRKEELPKESDRRDRKDRKDR ESSFKTPEKSHDDYLTAKARIMSAQPVI SVSEKOKRERDKFKAKGDKRKKNCSAVSKKENIVKPAKGPQEKVDGEBERSPRSEPPI
uman it ouse	ANLAEANASEEDKIKAMMSOSGHEVDFINYNKKP-LGPPPPSYTCFRCG ANLAEANASEEDKIKAMMSOSGHEVDFINYNKKTLVGPPPPSYTCFRCG ANLAEANASEEDKIKAMMSOSGHEVDFINYNKKTLVGPPPPSTCFRCG . * :::: ** ** .:: * *:* ** .::	KPGHYIKN 159 KPGHYIKN 174	Rat Mouse	SYSEADAREDARFANGUATANAUDIATANAUDSAYGANGYU YAFANGYQGAYUGGAGAFADIFI SYSEADAREDARFAYGANITAKKADGASATAKKADNULKPSKGPQOKUGGAKSPRSEPPI SVSDKDKREKDKPKVKSDKTKRKSDGSATAKKDNULKPSKGPQEKUDGDREKSPRSEPPI
orm y iman	CPMLNTERTTGIPSQELMETTV-DDPDAMMHPSGKYVIPIMHW CPFVGGKDQQEVKRNTGIPRSFRDKPDAENESADFVLP CPTNGDKNFESGPRIKKSTGIPRSFMMEVKDPNMKGAMLTNTGKYAIPTIDA	204 EAYAIGKK 233	Worm Fly Human	DDRKERRHERDSQKIDEQDRKKERKRDRETEAYD-SDKLQAPKTKVKREDKKRDDRKDYE NDTEME KTAKEETFKTDNTKSSSSSOKDEKITGTPKRAHSSASKEHQETFVVKEEVVKKKDYS
at ouse	CPTNGDKNFESGPRIKKSTGIPRSFMMEVKDPNMKGAMLTNTGKYAIPTIDA CPTNGDKNFESGPRIKKSTGIPRSFMMEVKDPNMKGAMLTNTGKYAIPTIDA ** . *:.**** : :	EAYAIGKK 234	Rat Mouse	KKAKEEATKVDSVKPSS-SQKDEKVTGTPFKAHSKSAKEHQEAKPVKEEKAKKDCS KKAKEEATKIDSVKPSSSSQKDEKVTGTPFKAHSKSAKEHQEAKPAKDEKVKKDCS . : : : : : : : : : : : : : : : : : : :
orm ly uman at ouse	RKNEDGSSFPAQTSRKVPPELLCPICQSLFKEATUTSCCGNSYCOLOCE NOVQNEIPEDLCICICNEIPEDAVITCGCGSSFCDCUV EKPPFLPEEPSSSSEDDPIPELLCLICKDIWTDAVVIPCCGNSYCOLOCE EKPPFLPEEPSSSSEDDPIPELLCLICKDIWTDAVVIPCCGNSYCOLOCE EKPPFLPEEPSSSSEDDPIPELLCLICKDIWTDAVVIPCCGNSYCOLOCE	TSLLESED 251 TALLESDE 293 TALLESDE 279 TALLESDE 294	Worm Fly Human Rat Mouse	RORER RED YEKEKSERKES DEDNEKEKOPERETVEKEHEKDKERKE KOKAE KRINKEDK
orm ly uman	QKCPGADCGKDISITSIIPNKTLRDAAAAWLSATGPGAPTTPQIVPEPEQ- SECPDCK-EKNCSPGSLIPNRFLRNSVNAFKNETGYNKSAAKPAAVKNEEKP HTCTTCH-QNDVSPDALIANKFLRQAVNNFKNETGYTKRLRKQIPPPPPPIP	PVEKEVEK 310	Worm Fly Human	IVEKESEKPRKSVHERMQKADSSTSSSSRTTTAPSLERKPVSFTVA-SSKPTTNIRV VLS PRA QSPSIEINAA IVK -PSPKRKMEPDTEKMDRTPEKDKISL-SAPAKKIKIMRE
at ouse	HTCPTCH-QNDVSPDALIANKFLRQAVNNFKNETGYTKRLRKQL HTCPTCH-QNDVSPDALIANKFLRQAVNNFKNETGYTKRLRKQL ** :: * :: * :: * :: * *: * *: **	APRPIIQR 338 PPRPIMQR 353	Rat Mouse	IVKPSPKRKMRGDVEKLERTPEKDKSASSTTPAKKIKLNRE IVKPSPKRKMEGDVE
orm ly uman	IGLKAPSSSQSQITPSGISPGST KPVAEV-EPEETEVKPEKQ-KESETNGSNPPKS NLQPLMRSPISRQQDPLMIPVTSSSTHPAPSISSLTSNQSSLAPPVSGNPSS	-ESPEPPA 348 APAPVPDI 412	Worm Fly Human	QL
at ouse	NLQPLMRSPISRQQDPLMIPVTSSSAHSTPSISSLTSNPSSLAPSVPGNPSS NLQPLMRSPISRQQDPLMIPVTSSSAHSAPSISSLTSNPSALAPSVSGNPSS **		Rat Mouse	TGKKIGNAENPSTTKEPSEKLESTSNKIKQEKVKGKARRKVAGTEGPSSTLVDYTSTSST TGKKIGNAENASTTKEPSEKLESTSSKIKQEKVKGKAKRKVAGSEGSSSTLVDYTSTSST
orm Ly uman	TTLTSV-SSGTSL TTEPS-QKEKDKYDSDYEDNITIKMPQPAADS TATVSISVHSEKSDGPFRDSDNKILPAAALASEHSKGTSSIAITALMEEKGY	QVPVLGTP 472	Worm Fly Human Rat	KEQEDEERSKEDREKKETDVESIGEKEKSSSREVPKESUDVHKESTLYEN SPTHNATEWYRSHSILTVGAASDDNIGPKSKLSERNSVILSWEIDENILG GSSVRKSEEKTDTKRTVIKTMEEYNNDNTAPAEDVIIMIQVPQSKWDKDDFESE
at ouse	TATVSISVHSEKSDGPFRDSDNKLLPAAALASEHSKGASSIAITALMEEKGY TATVSISVHSEKSDGPFRDSDNKLLPAAALTSEHSKGASSIAITALMEEKGY * : *	QVPVLGTP 473	Mouse	GGSFVRKSEEKTDTKRTVIKTMEEYNNDNTAPAEDVIIMIQVPQSKWDKDDFESE GGSFVRKSEEKTDTKRTVIKTMEEYNNDNTAPAEDVIIMIQVPQSKWDKDDFESE * *
orm ly uman at		DRSDYVSD 408 ERSCYRSI 532	Worm Fly Human Rat	LEDSSKKAAGASDDPSEITSDVLRKAENA IFAKAINA IRPMEFQVII- E-EDVKSTQPISSVGKPASVIKNVSTKPSNIVKYPEKESEPSEKIQKFTKDVSHEIIQHE EEDDVRTTQPIQSVGKPSSIIKNVTTKPSATVKCTEKKNEQSEKIQKITKEVSHEVIQHE
ouse	SLLGQSLLHGQLIPTTGPVRINAARGGGRPGWEHSNKLGYLVSPPQQIRRGI :: *.	ERSCYRSI 533	Mouse	E-EDVKTTOPIOSVGKPSSIIKNVTTKPSATAKYTEKESEQPEKLOKLPKEASHELMOHE
orm Ly iman at ouse	NRCRIHISERSQRTQCFSLPATFVTFVPPP NRCRIHISERSQRTQCFSLPATFVTFVPPP NRCRIHISERSQRTQCFSLPATFVTFVPPP NRCRIHISERSQRTQCFSLPATFVTFVPPPP PLYI	KPPYMQ 462 PPPPHTLP 573 PPPPHTLP 559	Fly Human Rat Mouse	-NSKDNSKDRSVVRSDKDRSSSPRRNNSSRSVKDRLGTKISN- VKSSKNSASSEKGTNDDYSVLEKENEKKNNSTQPERESNLDRLNBQGNFKBLGQS LKSSKGSVSEKGRAKDORHASEKDNPDKKRSIQDPERSTVDRLSEQGHFKTLGQS LRSSKGSASSEKGRAKDREHSGSEKDNPDKKRSGAQDDKESTVDRLSEQGHFKTLGS
orm ly uman	LPPGIPGLPQFGLPPPGVPGLSATVLPQHQSMPLNYGMPLFSAGFPPA' MQRGPPPMHMMSHHMPAYNNGFNNMGQRPPLSYVPYQNQSVHPMRAPYG LPPGVPPPPFSPQF-PFGQPPPAGYSVPPFGFP-	SAGGGMNM 519	Worm Fly Human	- -DRSRSRDKSKGRRRAARSSDDDANRGRSDRHGSRKRDNRSRDRAAPSE SKEARTSKHDSTRASSNKDFTPNRDKKTDYDTREYSSSKRBDENELTRRKDSPSR
at ouse	LPPGVPPPQFSPQF-PPGQPPPAGYSVPPPGFP LPPGVPPPQFSPQF-PPGQPPPAGYSVPPPGFP * * * * * * *		Rat Mouse	SKETRTSEKHESVRGSSNKDFTPGRDKKTDYDSRDYSSSKRADERSELTRRKDSPPR SKETRTSEKHESVRGSSNKDFTPGRDKKVDYDSRDYSSSKRADERGELARRKDSPPR
orm ly uman at ouse	AISDEWN	RFR 569 EFTNDFAK 667 EFTNDFAK 653	Worm Fly Human Rat Mouse	
orm ly uman at ouse		SRSHSRSY 727 SRSHSRSY 713	Worm Fly Human Rat Mouse	NNSDDSDRRAAKHTKSSDSRVVSSVTAVVAPFKPCRPDNPFRKFVDT
orm	EKERPRRADEHRRDRDRDRERDRDRSHRDVRSSGRSKDIKASSSHRRDR	DARRKDRR 596	Worm	
ly uman at ouse	SRSPEKREHSVERHIEHPRSSROPNDGS SRSPPYPRGRGKSRNYRSRSRSHGYHRSRSRSPYRRYHSRSRSF SRSPPYPRGRGKSRNYRSRSRSHGYHRSRSRSPYRRYHSRSRSF SRSPPYPRGRGKSRNYRSRSRSHGYHRSRSRSPYRRYHSRSRSF	AFRGQSPT 768	Fly Human Rat Mouse	
orm ly	. *.*.* * * : : . RDDVRKKERREKREEEDDDQKTKDAESKDEDEIDIDGIIAEYGNVQA KSPGGRIKRSGHRSASPKPGYKSDYRD		Worm Fly	SDNGMEHRKQRDKKLKKHSK-YSSTDSLKSEKRKDPKSKKKSKILK
uman at ouse	KRNVPQGETEREYFNRYREVPPPYDMKAYYGRSVDFRDPF KRNVPQGETEREYFNRYREVPPPYDIKAYYGRSVDFRDPF KRNVPQGETEREYFNRYREVPPPYDIKAYYGRSVDFRDPF . :*	EKERYREW 830 EKERYREW 816 EKERYREW 831	Human Rat Mouse	VOUGISENOSHSSPSVSPSRSHSPSGSOTESHSSASSAESODSKKKKKKKKKKKKKKKK SOPGADRSOSOSSPSVSPSRSHSPSGSOTESHSSSASSAESODSKKKKKKKKKKKKKKKK SOPGADRSOSOSSPSVSPSRSHSPSGSOTERSHSSASSAGSODSKKKKKKKKKKKKKKKKKKKKKK
orm ly uman at ouse	OPTDGDAQVADENAQNBEDSTSFKEESVSFKPETHEEDADEE 	D 691 PFTRGRRE 881 PFTRGRRE 867	Worm Fly Human Rat Mouse	1128 KKKKSKK- 1231 KNKKHKRHACTEVELEKSOKHKHKKKSKNNDEKEKEKEDOQVKSSTV 1792 KHKNHKHAOVGODVESOKNHKHKKKAKAKNDE-KEKDOQVKSSTV 1774 KHKKHKHAGADGDVEKSQKHKHKKKKAKKNKE-KE-KEKDOQVKSSTV 1790
orm ly uman at ouse	DTEVUVRDKSIDPVYQAMETSEAEVAETKEESVPVEEDEEPENHDEDVED GYRNKHPTS DYVGGSHRSRNIGSNYPEK	-SEASQ 705 -LSARDGH 908 -ISTRDSH 894		
	Domain With No Name (DWNN), ubiquitin-like	e (Pugh et a	al., 2006	b) p53-binding domain (Pugh et al., 2006)
	RING finger domain (Chibi et al., 2008, Lee e Proline-rich domain (Pugh et al., 2006)	t al., 2014)		

Zinc knuckle (Lee et al., 2014)

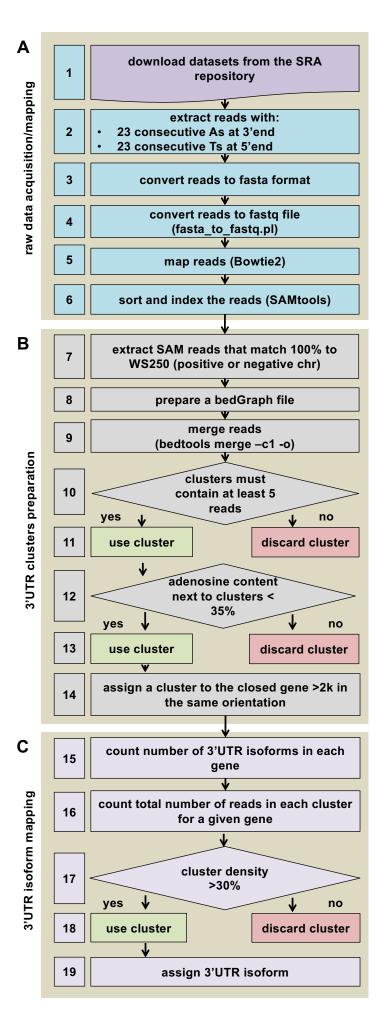
pap-1 (PAP1N)

N	PAP Central PAP RI	٩N	Bindi	ng
Worm Fly Human Mouse Rat	MSATEKDKTPLLGVSQPISLAHPDSKDIAQTTLLIETLKKFGSY MWNSEPTHRQHHQHNCHSTGGCPARQLGMTSAISLAEPRPEDLQRTDELRGSLEPYNPV PPPVTTQGSQQTQPPQRHYGITSPISLAAPKETDCLITQKLIETLKPFGV MPFPVTTQGSQQTQPPQRHYGITSPISLAAPKETDCLLTQKLIETLKPFGV MPFPVTTQGSQQTQPPQRHYGITSPISLAAPKETDCLLTQKLIETLKPFGV MFFPVTTQGSQTQPPQRHYGITSPISLAAPKETDCLLTQKLIETLKPFGV MFFPVTTQGSQTQPPQRHYGITSPISLAAPKETDCLLTQKLIETLKPFGV	60 52 52	Worm Fly Human Mouse Rat	AEHAKTLDLTNEIQRFKTNVELQASNVKGIGPNCQVQIDMFYVKRNSLIQVISAAD 499 ERSENLNVDLTBSIQNTEHVMHGVNIKMLKEG-MTIDAHVKKRQLGLLVLDSDF 530 KKTENSENLSVDLTVDIGSFTDYTVRQAINSKMFFEDMKIAAMHVKRKQLGLLFNHV 501 KKTENSENLSVDLTVDIGSFTDYTVRQAINSKMFFEDMKIAAMHVKRKQLHQLLFSHV 501
Worm Fly Human Mouse Rat	EPKEETEQRMEVLRNLNRLVKEWVKNVTAM-KIPNGEGVNAGGKLPTPGSYRLGVHSGA ESODELNHRMEILAKLNTLVKGWVKEISVSKNMPESAAEKLGGKITPTGSYRLGVHRKGA EEEEELQRAILILGKLNNLVKEWIREISSKNLPGSVIENVGGKITPTGSYRLGVHRKGA EEEEELQRAILILGKLNNLVKEWIREISESKNLPGSVIENVGGKITPTGSYRLGVHRKGA EEEEELQRAILILGKLNNLVKEWIREISESKNLPGSVIENVGGKITPTGSYRLGVHRKGA * :: * :: * : * * * * * * * * * : * : : : * * : *	120 112 112	Worm Fly Human Mouse Rat	LRRGRWKKVVPIATNTSVSSSTPRSVVRTTSTSSVPTTPTGLAAPKTPLSASVSATNEP 559 LKRERKSMESHNNTNTLLAN-RKRLSTELAGSO-DFLPF-GQOPSSGN 576 LQKKKKHSTEGVKLTLALNDSSLDLSMDSNDNSVPSPT-SAM-KTSPLNSSGSS 553 LQKRKHSTEGVKLAALNDSSLDLSMDSDNSMSVPSPT-SAM-KTSPLNSSGSS 553
Worm Fly Human Mouse Rat	DIDTLAVUPRHIDRSDFFTSFKEMLNNDPNUTELHGVEEAFUPUMKLKYSGVELDILFAR DIDALCVAPRNIERDYPOSFFEVLKKQPEUTECRSVEEAFUPUIKMPOGIEIDLLFAR DIDALCVAPRHUDRSDFFTSFVDKLKQEEVKDLARVEEAFUPUIKLCPGIEIDILFAR DIDALCVAPRHUDRSDFFTSFVDKLKQEEVKDLARVEEAFUPUIKLCPGIEIDILFAR DIDALCVAPRHUDRSDFFTSFVDKLKQEEVKDLARVEEAFUPUIKLCPGIEIDILFAR	180 172 172	Worm Fly Human Mouse Rat	DSTINGTPLSRKR-SMDEESSTIVIS0ISDESVPKKKTRDDILEENRVSMV 609 RGRDSGAKIQRLSDSLTEENSNASSDMGACTPTPTTTAQLSAPSFKSSGKNGSE
Worm Fly Human Mouse Rat	LALKEVPDTQELSDDNLLRNLDQESVRSLNGCRVAEQLLKLVPRQKEFCVTLRAIKLWAK LSLKEIPDDFDLRDDNLLRNLDHRSVRSLNGCRVTDEILALVPNIENFRLALETIKLWAK LALQTIPEDLDLRDDSLLKNLDICCISLNGCRVTDEILHUPPNIENFRLTLRAIKLWAK LALQTIPEDLDLRDDSLLKNLDICCISLNGCRVTDEILHUPPNIENFRLTLRAIKLWAK LALQTIPEDLDLRDDSLLKNLDICCISLNGCRVTDEILHUPPNIENFRLTLRAIKLWAK	240 232 232	Worm Fly Human Mouse Rat	VEVSNVVVEQRTKVVQEIVDLQADNGLNTSNGLEASEQKMEVPQSV
Worm Fly Human Mouse Rat	NHGIYSNSMGFFGGITWAILVARACQLYPNASPSRLVHRMFFIFSTWTWPHPVVLNEMNN KHGIYSNSLGYFGGUTWAMLVARTCQLYPNAAAATLVHKFFLVFSKWKKPNPVLLKOPED RHNIYSNILGFLGGUSWAMLVARTCQLYPNAIASTLVHKFFLVFSKWKKPNPVLLKOPEE RHNIYSNILGFLGGUSWAMLVARTCQLYPNAIASTLVHKFFLVFSKWKWPNPVLLKOPEE RHNIYSNILGFLGGUSWAMLVARTCQLYPNAIASTLVHKFFLVFSKWKWPNPVLLKOPEE 	300 292 292	Fly	
Worm Fly Human Mouse Rat	DRNDIPTLCELVWDPRRKNTDRFHVMPIITPAFPEQNSTHNVTRSTATVIKNEICEALEI VNLRFQVWDPRVNASDRYHLMPIITPAFPQQNSTFNVSSSTKKVLITEFRNGUNI CNLNLFVWDPRVNFSDRYHLMPIITPAFPQQNSTYNVSVSTRWVMVEFKQGLAI CNLNLPVWDPRVNFSDRYHLMPIITPAFPQQNSTYNVSVSTRWVMVEFKQGLAI	355 347 347	Worm Fly Human Mouse Rat	655 QQTATSTTMISSSNIAHSNYTQHSTHNNNGGGGGSHYNKTYIDDDDEDNNQPLLTPNNQ SSPHKSTRO 679 SSPHK
Worm Fly Human Mouse Rat	CRD ISEGKSKWTALFEEVNFFSRYKHFIALIMAAPNEEEELNYGGFLESR IRLLVQSLEF TDEILLSRAPKSKLFEANSFFFRYKHFIYLLVNSQFADDHLEWCGLVESKVRLICGLER TDEILLSRAPKSKLFEANNFFOKYKHYIYLLASAPTERQRLEWVGLVESKIRILVGSLEF TDEILLSRAEWSKLFEANNFFOKYKHYIYLLASAPTEKQRLEWVGLVESKIRILVGSLEK TDEILLSRAEWSKLFEANNFFOKYKHYIYLLASAPTEKQRLEWVGLVESKIRILVGSLEK	415 407 407 407	Fly Human Mouse	655 TQQQQKAVRITATSASFATSVPVAASAAPPAPTISL
Worm Fly Human Mouse Rat	NQDIIIAHNDPNKHKPSPNAKF	475 443 443		

symk-1 (Symplekin)

N	DUF3453			Symplekin C
Worm	MDYIQGLNEENETASERIGEALKEARDAETIEKKLLSLSTAM	42	Worm	-FYKVLLETPLLTPNAIERLKQVCLAKE-NEHGMAMLRELIMTRNRQRPQLLQFLFGLFF 7
Fly	MDSIIGRSQFVS-ETANLFTDEKTATARAKVVDWCNELVIASP-STKCELLAKVQ		Fly	LIRRVYLEAPILPEVSIGHLVQLSLDDEFSQHGLELIKDLAVLRPPRKNRFVRVLLNFSV 7
Human	MASGSGDSVTRRSVASQFFTQEEGPGIDGMTTSERVVDLLNQAALITN-DSKITVLKQVQ		Human	-FTKVVLEAPLITESALEVVRKYCEDESRTYLGMSTLRDLIFKRPSRQFQYLHVLLDLSS 7
Mouse Rat	MASSSGDSVTRRSVASQFFTQEEGPSIDGMTTSERVVDLLNQAALITN-DSKITVLKQVQ MASGSGDSVTRRSVASQFFTQEEGPGIDGMTTSERVVDLLNQAALITN-DSKITVLKQVQ	59	Mouse Rat	-FTKVVLEAPLITESALEVIRKYCEDESRAYLGMSTLGDLIFKRPSRQFQYLHVLLDLSS 7 -FTKVVLEAPLITESALEVIRKYCEDESRAYLGMSTLGDLIFKRPSRQFQYLHVLLDLSS 7
Kdt	* * ** *. *: :: : ::* *.	59	Nac	: :* **:*:: :: :: : *: ::* . * :: :::.*:.
Worm	HLLIDPSLSISILDNFLTEMLEFAELNDSRILCLLVDFLLKASAKDFTLCNKTVERYSFY			MERPELRSSCLEVVKELCYLPF-IRSSLSDQARMQIHDCLQESPPMYMRSSEDSD 7
Fly	ETVLGSCAELAEEFLESVLSLAHDSNMEVRKQVVAFVEQVCKVKVELLPHVINVVSML	111	Fly	HERLDLRDLAQAHLVSLYHVHKILPARIDEFALEWLKFIEQESPPAAVFSQDFGRPTEEP 7
Human Mouse	ELIINKDPTLLDNFLDEIIAFQADKSIEVRKFVIGFIEEACKRDIELLKKLIANLNML ELIINKDPTLLDNFLDEIIAFQADKSIEVRKFVIGFIEEACKRDIELLKKLIANLNML			HEKDKVRSQALLFIKRMYEKE-QLREYVEKFALNYLQLLVHPNPPSVLFGADKDTE-VAA 7 HEKDRVRSQALLFIKRMYEKE-QLREYVEKFALNYLQLLVHPNPPSVLFGADKDTE-VAA 7
Rat	ELIINKDFTLLDNFLDEIIAFQADKSIEVRKFVIGFIEEACKRDIELLLKLIANLNML ELIINKDPTLLDNFLDEIIAFQADKSTEVRKFVIGFIEEACKRDIELLLKLIANLNML	117	Pat	HEKDRVRSQALLFIRRMIEKE-QLREIVERFALNILQLLVHPNPPSVLFGADKDIE-VAA 7 HEKDRVRSQALLFIRRMYEKE-QLREYVEKFALNYLQLLVHPNPPSVLFGADKDIE-VAA 7
.u.c		11/	1000	*: :* : : : :* :: :.** :.
√orm	LIPNKSIKRYESVIKRVVVASTNLYPIVLEFAIMDKNDNAESCWDAFNLLKNRICMLV	160	Worm	QWTDEMYKNSLAVYSTLMPSDPL-LLIPLASVYAQSTNVFKRVVLRSLEPVFRQLSQE 8
Fly	LRDNSAQVIKRVIQACGSIYKNGLQYLCSLMEPGDSAEQAWNILSLIKAQILDMI	166	Fly	DWREDTTKVCFGLAFTLLPYKPEVYLQQICQVFVSTSAELKRTILRSLDIPIKKMGVESP 8
Human Mouse	LRDENVNVVKKAILTMTQLYKVALQWMVKSRVISELQEACWDMVSAMAGDIILLL LRDENVNVVKKAILTMTOLYKVALOWMVKSRVISDLOEACWDMVSSMAGEIILLL			PWTEETVKQCLYLYLALLPQNHK-LIHELAAVYTEAIADIKRTVLRVIEQPIRGMGMNSP 8 PWTEETVKOCLYLYLALLPONHK-LIHELAAVYTEAIADIKRTVLRVIEOPIRGMGMNSP 8
Rat	LRDENVNVVKKAILIMIGLIKVALQWMVKSKVISDLQEACWDMVSSMAGEIILLL LRDENVNVVKKAILIMIGLYKVALQWMVKSRVISDLQEACWDMVSSMAAEIILLL			PWTEETVKQCLYLYLALLPQNHK-LIHELAAVYTEAIADIKKTVLKVIEQPIRGMGMNSP 8
	:*::: .: * :: .: * ::			* :: * .: : :*:* . : :. *:: :.:** :: :: :.
Vorm	SDDHEGVRTVTVKFLEALILCQSPKPRELATGSNISWAREANTRFNRISLSDVPRSHRFL			MVISLIEDCPYGAETLVARLVVLLTERIT-PSTDLIQKLKILHDERKMDIRALLPIIGGL 9
7ly Human	DNENDGIRTNAIKFLEGVVVLQSFADEDSLKRDGDFSLADVPDHCTLF DSDNDGIRTHAIKFVEGLIVTLSPRMADSEIPRRQEHDISLDRIPRDHPYI	214	Fly	TLLQLIEDCPKGMETLVIRIIYILTERVPSPHEELVRRVRDLYQNKVKDVRVMIPVLSGL 8 ELLLLVENCPKGAETLVTRCLHSLTDKVP-PSPELVKRVRDLYHKRLPDVRFLIPVLNGL 8
iuman louse	DSDNDGIRTHAIRFVEGLIVTLSPRMADSEVPRRQEHDISLDRIPRDHPYI DSDNDGIRTHAIKFVEGLIVTLSPRMADSEVPRRQEHDISLDRIPRDHPYI			ELLLLVENCPKGAETLVTRCLHSLTDKVP-PSPELVKRVRDLYHKRLPDVRFLIPVLNGL 8 ELLLLVENCPKGAETLVTRCLHSLTDKVP-PSPELVKRVRDLYHKRLPDVRFLIPVLNGL 8
Rat	DSDNDGIRTHAIKFVEGLIVTLSPRMADSEVPRRQEHDISLDRIPRDHPYI	223	Rat	ELLLLVENCPKGAETLVTRCLHSLTDKVP-PSPELVKRVRDLYHKRLPDVRFLIPVLNGL 8
	:::*:*:*:::: * : : :**::*:::			:: *:*:** * **** * : **::: * :*:::: *:.:: *:::**
lorm	SYHKTQLEAEENFSALLKQTTVAEATSQNLITVIESLCMITRCRPQWENALPRVFDVIKA			EREEVVRLIPTFIFRAEYQKSVNVLFRKLYTVRDPQTGNLVFDPIEVIKEYHKIEPKN 9
'ly	RREKLQEEGNNILDILLQFHGTTHISSVNLIACTSSLCTIAKMRPIFMGAVVEAFKQ QYNVLWEEGKAALEQLLKFMVHPAISSINLTTALGSLANIARQRPMFMSEVIQAYET	271	Fly	TRSELISVLPKLIKLNPAVVKEVFNRLLGIGAEFAHQ-TMAMTPTDILVALHTIDTSV 9 EKKEVIQALPKLIKLNPIVVKEVFNRLLGTQHGEGNSALSPLNPGELLIALHNIDSVK 9
luman louse	QYNVLWEEGKAALEQLLKFMVHPAISSINLTTALGSLANIARQRPMFMSEVIQAYET OYNVLWEEGKAAVEOLLKFMVHPAISSINLTTALGSLANIARORPMFMSEVIOAYET			EKKEVIQALPKLIKLNPIVVKEVFNRLLGTQHGEGNSALSPLNPGELLIALHNIDSVK 9 EKKEVIQALPKLIKLNPIVVKEVFNRLLGTOHGEGNSALSPLNPGELLIALHNIDSVK 9
Rat	QINVLWEEGKAALEQLLKFMVHPAISSINLITALGSLANIARQRFMFMSEVIQAIEI QYNVLWEEGKAALEQLLKFMVHPAISSINLTTALGSLANIARQRFMFMSEVIQAYET			EKKEVIQALPKLIKLNPIVVKEVFNKLLGTQHGEGNSALSFLNPGELLIALHNIDSVK 9
	· *.: **: :* ** : **. *:: ** : . : .:. :			·.*:: :*::* *: :*::* : * ::: *.*:
∛orm	LHSNVPPMLSKGQVKFLRKSFKYNLLRFLKLPASVPLQQKITTMLTNYLGASPREVQQSI			DNEAELLVNNLEFLFDPALLKPDTASQAIEAVFKWENVPFLFLHSLYTLFHKFKTFESFV 1
Fly	LNANLPPTLTDSQVSSVRKSLKMQLQTLLKNRGAFEFASTIRGMLVD-LGSSTNEIQKLI			CDIKAIVKATSLCLAERDLYTQEVLMAVLQQLVEVTPLPTLMMRTTIQSLTLYPRLANFV 1
Human Mouse	LHANLPPTLAKSQVSSVRKNLKLHLLSVLKHPASLEFQAQITTLLVD-LGTPQAEIARNM LHANLPPTLAKSQVSSVRKNLKLHLLSVLKHPASLEFQAQITTLLVD-LGTPQAEIARNM	339	Human	CDMKSIIKATNLCFAERNVYTSEVLAVVMQQLMEQSPLPMLLMRTVIQSLTMYPRLGGFV 1 CDMKSIIKATNLCFAERNVYTSEVLAVVMQQLMEQSPLPMLLMRTVIQSLTMYPRLGGFV 1
Rat	LHANLPPTLARSQVSSVRKNLKLHLLSVLKHPASLEPQAQITTILVD-LGTPAELARMM *::*:** *:**. :*::*:* .** .:: *: :*: :			CDMKSIIKAINLEFABRWYTSEVIAVWQQLHEQSFDHILMFNYQSIMPFRUGEV 1 : : : : : : : : : : : : : : : : : : :
Worm	PPELIQKIAPPRPPQHPAEPVAKRPKIQNQIFEDDDDDDDEAGPSTSTV-NAKDARTE			ANLFYKVTEKKMYQQSDRWKQAFFKCIKELKTKAYPAVITFLSFEEYEELKEVL 1
Fly	PKMDKQEMARRQKRILENAAQSLAKRARLACE-QQDQQQREMELDTEELERQKQK			MNLLQRLIIKQVWRQKVIWE-GFLKTVQRLKPQSMPILLHLPPAQLVDALQQCPDLRPAL 1
Human Mouse	PSSKDTRKRPRDDSDSTL-KKMKLEPNLGEDDEDKDLEPGPSGTSKASAQISGQS PSSKDSRKRPRDDTDSTL-KKMKLEPNLGEDDEDKDLEPGPSGTSKASAQISGQS			MNILSRLIMKQVWKYPKVWE-GFIKCCQRTKPQSFQVILQLPPQQLGAVFDKCPELREPL 1 MNILARLIMKQVWKYPKVWE-GFIKCCQRTKPQSFQVILQLPPQQLGAVFDKCPELREPL 1
Rat	PSSKDSKKRFRDDTESTL-KKMKLEPNLGEDDEDKDLEPGPSGTSKASAQISGQS PSSKDSRKRPRDDTESTL-KKMKLEPNLGEDDEDKDLEPGPSGTSKASAQISGQS	303	Pat	MNILARLIMKQVWKIPKVWE-GFIKCCQRTKPQSFQVILQLPPQQLGAVFDKCPELREPL 1 MNILARLIMKQVWKYPKVWE-GFIKCCQRTKPQSFQVILQLPPQQLGAVFDKCPELREPL 1
tu c	* : *: :: : : :*::: : * : .	575	Muc	*::::: *:::: *: .*:: :::: :::: :::: *:
Vorm	AIDMTAKFI-MECLNHETVMNLVKISLYTLPSEMPAAFASSYTPIANAGTEPNRQELSEL			GNMDEKIKEELHDKERENR- 1
7ly	${\tt STRVNEKFLAEHFRNPETVVTLVLEFLPSLPTEVPQKFLQEYTPIREMSIQQQVTNISRF}$	444	Fly	SEYAESMQDEPMNGSGITQQVLDIISGKSVDVFVTDESGGYISAEHIKKEAPDPSEISVI 1
luman louse	DTDITAEFL-QPLLTPDNVANLVLISMVYLPEAMPASFQAIYTPVESAGTEAQIKHLARL DTDITAEFL-OPLLTPDNVANLVLISMVYLPETMPASFOAIYTPVESAGTEAOIKHLARL			LAHVRSFTPHQQAHIPNSIMTILEASGKQP-EAKEAPAGPLEEDD- 1 LAHVRSFTPHQQAHIPNSIMTILEATGKQP-EVKEAPSGPLEEDD- 1
louse lat	DTDITAEFL-QPLLTPDNVANLVLISMVYLPETMPASFQAIYTPVESAGTEAQIKHLARL			LAHVRSFTPHQQAHIPNSIMTILEASGKQEP-EVKEAPSGSLEEDD- 1
	:. :*: . :.* .** : ** :* * ***: : : .:::			* * : *:
orm	MAVQMTNKEIGPGYEWLQQQRKKEYEARNKARSEGMAIAQTP			ERDKRLRREEKKEKEREKE-RTRESGKERS 1
ly	FGEQLSEKRLGPGAATFSREPPMRVKKVQAIESTLTAMEVDEDAVQK	491	r LY Human	STVPVLTSLVPLPVPPPIGSDLNQPLPPGED 1 -LEPLTLAPAPAPRPPODLIGLRLAQEKALKROLEEEOKLKPGGVGAP 1
luman louse	MATQMTAAGLGPGVEQTKQCKEEPKEEKVV-KTESVLIKRRLSAQGQAISVVGSLSSMSP MATQMTAAGLGPGVEQTKQCKEEPKEEKVV-KPESVLIKRRLSVQGQAISVVGSQSTMSP	511	Mouse	-LEPLTLAPAPAPRPPQDLIGLRLAQEKALKRQLEEEQKLKPGGVGAP 1 -LEPLALALAPAPAPAPAPAPAPRPPQDLIGLRLAQEKALKRQLEEEQKQKPTGIGAP 1
lat	MATQMTAAGLGPUVEQIKQCKEEPKEEKVV-KPESVLIKKRLSVQGQAISVVGQSINSF MATQMTAAGLGPUVEQIKQCKEEPKEEKVV-KPESVLIKKRLSVQGQAISVVGSQSINSF :. *:: :** .: *.			-LEPLALALAPALAPAPAPRPPQDLIGLRLAQEKALKRQLEEEQKQKPTGVGAP 1
orm	IHEPN-MSNRVPAQIVKQSLQE-INTLPVIQKAKKAFNLVEEAVVFDDKEAAEMFEL			SRR 1
ly	LSEEEFQRKEEATKKLRETMERAKGEQTVIEKMKERAKTLKLQEITKPLPRNLKEKFLTD	551	Fly	1
luman louse	LEEEAPQAKRRPEPIIPVTQPRLAGAGGRKKI FRLSDVLKPLTDAQVEAMKLG LEEEVPOAKRRPEPIIPVTOPRLAGAGGRKKI FRLSDVLKPLTDAOVEAMKLG			SSSSPSPSPSARPGPPPSEEAMDFREEGPECETPGIFISMDDDSGLTEAALLDSSLE 1 AACVSSTPSVPAAARAGPTPAEEVMEYREEGPECETPAIFISMDDDSGLAETTLLDSSLE 1
at	LEEEVPQARKREF11PVTQFRLAGAGGRKK1FRLSDVLKFLTDAQVEAMKLG LEEEVPQAKRRPEPIIPVTQFRLAGAGGRKKIFRLSDVLKFLTDAQVEAMKLG : * :. : * :.* :. : : : : : : : : : : :			AACVSSTPSVFAARAGFTPAEEVMEYREGGPECETFAIFISMDDDSGLAETTLLDSSLE 1 TSSVSSTPLVGPAARAGPTPAEEVMEYREGGPECETPAIFISMDDDSGLAETTLLDSSLE 1
	AYESVLQAERRVVAGGARLMYQKLVVRLTTRFWEDCTPFEEKLIEFVLADHKKRNDLALL			
lorm 'ly	AYESVLQAERRVVAGGARLMYQKLVVRLTTRFWEDCTPFEEKLIEFVLADHKKRNDLALL AVRRILNSERQCIKGGVSSKRRKLVTVIAATFPDNVRYGIMEFILEDIKQRIDLAFS			1143 1165
uman	AVKRILRAEKAVACSGAAQVRIKILASLVTQFNSGLKAEVLSFILEDVRARLDLAFA	621	Human	GPLPKETAAGGLTLKEERSPQTLAPVGEDAMKTPSPAAEDAREPEAKGNS 1274
louse	AVKRILRAEKAVACSGAAQVRIKILASLVTQFDSGFKAEVLSFILEDVRARLDLAFA	621	Mouse	GPLPKEAAAVGSSSKDERSPQNLSHAVEEALKTSSPETREPESKGNS 1284
lat	AVKRILRAEKAVACSGAAQVRIKILASLVTQFDSGFKAEVLSFILEDVRARLDLAFA * . :*.:*: .*. *:. ::: * ::.*:* * : * ***:	621	Rat	GPLPKEAAAVGPSSKDERSPONLSHAVEEALKTSSPEAREPESKGNS 1280
lorm	WLCELYAQYQGYSNCALFMKEMIAGQEGLTQAQRLDRYDQAMCKMLDAMLERNMEKEAL-	672		
ly	WLFEEYSLLOGFTRHTYVKTENRPDHAYNELLNKLIFGIGERCDHKDKII	658		
luman	WLYQEYNAYLAAGASGSLDKYEDCLIRLLSGLQEKPDQKDGI-	663		
louse	WLYQEYNAYLAAGTSGTLDKYEDCLICLLSGLQEKPDQKDGI-			
Rat	WLYQEYNAYLAAGPSGTLDKYEDCLICLLSGLQEKPDQKDGI-	662		

$\begin{array}{c c} cpsf-1\\ (CPSF160) & 2 & 238/5 & 9\\ \hline 3 & 149/1 & 9\\ \hline 3 & 149/1 & 9\\ \hline 1 & 68/37 & 6\\ \hline 2 & 178/28 & 8\\ \hline (CPSF100) & 3 & 221/25 & 8\\ \hline cpsf-4\\ (CPSF30) & 1 & 323/5 & 9\\ \hline cpsf-4\\ (CPSF30) & 3 & 204/3 & 9\\ \hline 1 & 251/76 & 7\\ \hline 2 & 200/64 & 7\\ \hline (CstF-50) & 3 & 185/25 & 8\\ \hline cpf-2\\ (CstF-64) & 3 & 61/3 & 9\\ \hline \end{array}$	3.7 7.9 93.3 64.8 66.4 99.8 88.5 77.6 88.6 66.8 55.8 88.1
$\begin{array}{c c} (CPSF160) & 2 & 238/5 & 9 \\ \hline & 2 & 238/5 & 9 \\ \hline & 3 & 149/1 & 9 \\ \hline & 1 & 68/37 & 6 \\ \hline & 2 & 178/28 & 8 \\ \hline & 2 & 178/28 & 8 \\ \hline & 2 & 21/25 & 8 \\ \hline & 1 & 323/5 & 9 \\ \hline & 2 & 699/17 & 9 \\ \hline & 2 & 699/17 & 9 \\ \hline & 3 & 204/3 & 9 \\ \hline & 1 & 251/76 & 7 \\ \hline & 2 & 200/64 & 8 \\ \hline & 1 & 446/54 & 8 \\ \hline & 2 & 138/14 & 9 \\ \hline & 3 & 61/3 & 9 \\ \hline \end{array}$	9.3 4.8 6.4 9.8 8.5 7.6 8.6 6.8 5.8
$\begin{array}{c c} \mathbf{cpsf-2} \\ (CPSF100) \\ \hline \mathbf{cpsf-2} \\ (CPSF100) \\ \hline \mathbf{cpsf-4} \\ (CPSF30) \\ \hline \mathbf{cpsf-4} \\ (CPSF30) \\ \hline \mathbf{cpf-1} \\ (CstF-50) \\ \hline \mathbf{cpf-2} \\ (CstF-64) \\ \hline \mathbf{cpf-2} \\ (CstF-64) \\ \hline 3 \\ \hline \mathbf{cpf-3} \\ \hline 1 \\ 1$	4.8 6.4 9.8 8.5 7.6 8.6 6.8 6.8 5.8
$\begin{array}{c c} cpsf-2\\ (CPSF100) \hline 2 & 178/28 & 8\\ \hline 3 & 221/25 & 8\\ \hline 1 & 323/5 & 9\\ \hline cpsf-4\\ (CPSF30) \hline 2 & 699/17 & 9\\ \hline 2 & 699/17 & 9\\ \hline 3 & 204/3 & 9\\ \hline \\ cpf-1\\ (CstF-50) \hline 1 & 251/76 & 7\\ \hline 2 & 200/64 & 7\\ \hline 2 & 200/64 & 7\\ \hline 3 & 185/25 & 8\\ \hline 1 & 446/54 & 8\\ \hline \\ cpf-2\\ (CstF-64) & 3 & 61/3 & 9\\ \hline \end{array}$	36.4 39.8 38.5 77.6 38.6 6.8 5.8
$\begin{array}{c c} (CPSF100) & 2 & 178/28 & 8 \\ \hline (CPSF100) & 3 & 221/25 & 8 \\ \hline (cpsf-4) & 1 & 323/5 & 9 \\ \hline (CPSF30) & 2 & 699/17 & 9 \\ \hline (CSFF-50) & 3 & 204/3 & 9 \\ \hline (CstF-50) & 3 & 204/3 & 9 \\ \hline (CstF-50) & 3 & 185/25 & 8 \\ \hline (CstF-64) & 1 & 446/54 & 8 \\ \hline (CstF-64) & 3 & 61/3 & 9 \end{array}$	9.8 98.5 97.6 98.6 6.8 75.8
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	8.5 7.6 8.6 6.8 5.8
$\begin{array}{c c} cpsf-4\\ (CPSF30) \hline 2 & 699/17 & 9\\ \hline 3 & 204/3 & 9\\ \hline \\ cpf-1\\ (CstF-50) \hline 1 & 251/76 & 7\\ \hline 2 & 200/64 & 7\\ \hline 3 & 185/25 & 8\\ \hline \\ cpf-2\\ (CstF-64) \hline 1 & 446/54 & 8\\ \hline 2 & 138/14 & 9\\ \hline 3 & 61/3 & 9\\ \hline \end{array}$	07.6 08.6 06.8 05.8
$\begin{array}{c c} (CPSF30) & 2 & 699/17 & 9 \\ \hline (CPSF30) & 3 & 204/3 & 9 \\ \hline cpf-1 & 1 & 251/76 & 7 \\ \hline (CstF-50) & 3 & 185/25 & 8 \\ \hline cpf-2 & 1 & 446/54 & 8 \\ \hline (CstF-64) & 3 & 61/3 & 9 \end{array}$	08.6 6.8 5.8
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	6.8 5.8
$\begin{array}{c c} cpf-1 \\ (CstF-50) \end{array} & \begin{array}{c} 2 & 200/64 & 7 \\ \hline 3 & 185/25 & 8 \\ \hline \\ cpf-2 \\ (CstF-64) \end{array} & \begin{array}{c} 1 & 446/54 & 8 \\ \hline 2 & 138/14 & 9 \\ \hline 3 & 61/3 & 9 \end{array}$	′5.8
$\begin{array}{c cccc} (CstF-50) & 2 & 200/64 & 7 \\ \hline & & & & & & & & & & & & \\ \hline cpf-2 & & & & & & & & & \\ (CstF-64) & & & & & & & & & & \\ \hline & & & & & & & &$	
$\begin{array}{c c c c c c c c c c c c c c c c c c c $	8.1
cpf-2 (CstF-64) 2 138/14 9 3 61/3 9	
2 130/14 9 (CstF-64) 3 61/3 9	9.2
3 61/3 9	0.8
1 249/23 9	5.3
	1.5
<i>cfim-1</i> (<i>CFIm25</i>) 2 196/20 9	0.7
3 154/2 9	8.7
1 137/1 9	9.3
cfim-2 2 282/27 9 (CFIm68) 2 282/27 9	1.3
3 260/32 8	9.0
1 311/116 7	2.8
Irp-2 2 281/97 7 (CFIm59) 2 281/97 7	4.3
3 222/93 7	'0.5
1 62/4 9	3.9
symk-1 2 19/4 8	2.6
(Symplekin) 2 10/1 9	5.8
1 293/0 1	00
rbpl-1 2 344/0 1 (RBBP6) 2 344/0 1	00
3 116/0 1	00
1 141/5 9	6.6
pcf-11 2 105/0 1 (CPF11) 2 105/0 1	00
3 172/5 9	7.2
1 286/23 9	2.6
clpf-1 2 289/18 9 (CLP1) 0 000/17 0	4.1
(CLPI) 3 208/17 9	2.4
1 5/0 1	00
2 8/0 1	00
<u> </u>	4.7
4 51/1 9	8.1
	00
pkc-3 6 22/0 1	00
(negative 7 53/0 1	00
8 6/0 1	00
9 13/0 1	00
10 19/0 1	00
11 5/0 1	00
12 4/0 1	



Supplemental Fig. S3: Bioinformatic Pipeline used in this study: The pipeline

uses raw transcriptome datasets downloaded from the public repository SRA trace archive to extract and map 3'UTR end clusters to the closest protein-coding genes in the correct orientation.

The pipeline is divided in three large steps: A) Acquisition/Mapping, B) 3'UTR cluster preparation and C) 3'UTR isoforms mapping.

In the acquisition/mapping step, we used custom made Perl scripts to extract reads with 23 consecutive As at the 3' end or 23 consecutive Ts at the 5' end, and then mapped these filtered reads to the WS250 version of the *C. elegans* genome (Bowtie2).

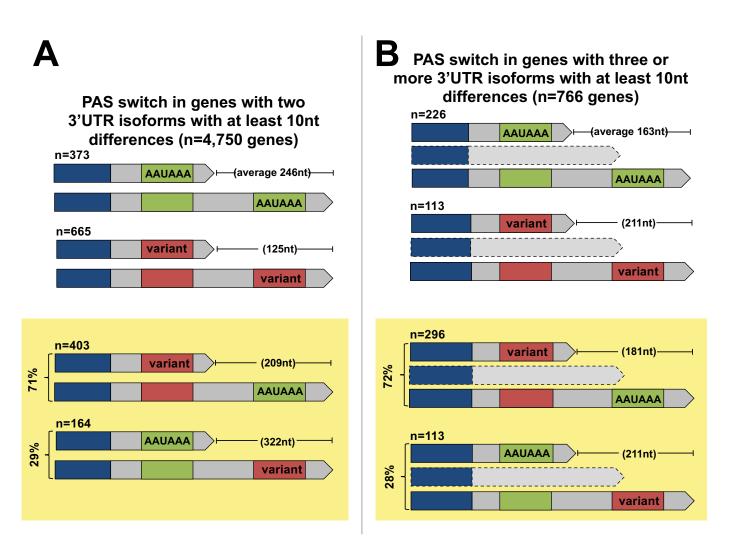
We then sorted and indexed the reads for visualization purposes.

In the 3'UTR cluster preparation step we extracted SAM reads with 100% match to the WS250, and used them to prepare a new bedGraph file (BEDTools).

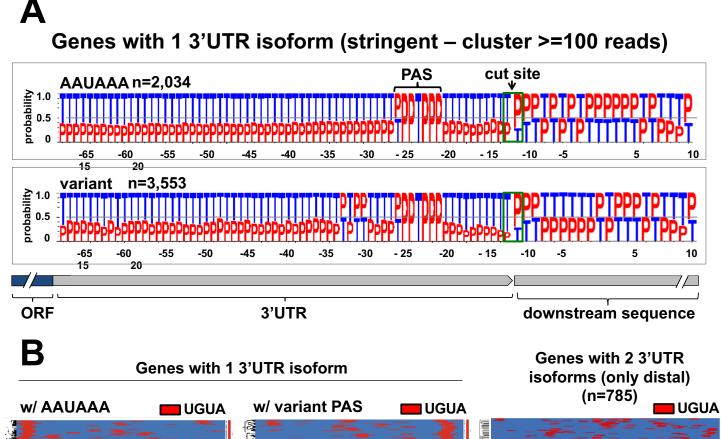
We then merged the reads and discarded the clusters with less than 5 reads. Restrictive parameters for cluster identification and 3'UTR end mapping included the discard of clusters with an Adenosine content of <35% downstream of its end.

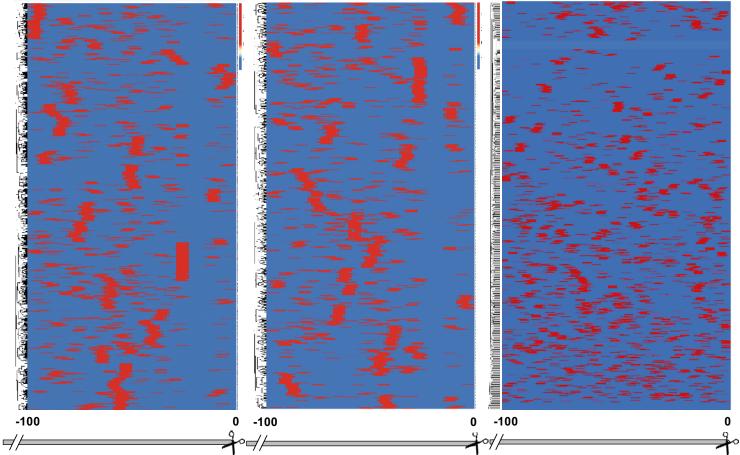
Clusters were assigned to a mapped 3'UTR end and attached to the closest gene with 2,000nt in the same orientation. At the completion of these steps we performed the 3'UTR isoform mapping step, which consists of the counting and assignment the total number of 3'UTR isoforms to a given gene.

We discarded clusters with a density of less than 30% of the total number of reads.

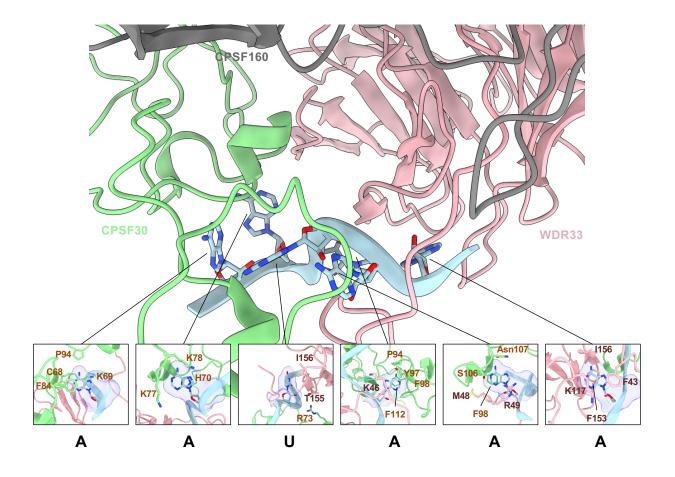


Supplemental Fig. S4: PAS site usage in genes with multiple 3'UTR isoforms. A) In genes with only two 3'UTR isoforms with a difference of at least 10nt between isoforms, 373 pairs of isoforms had canonical PAS elements in both isoforms with an average of 246nt difference between isoforms while 665 pairs had variant PAS elements in both isoforms with an average of 125nt difference between them. In isoform pairs where the type of PAS element switches, 71% have a shorter isoform with a variant PAS element and a longer isoform with a canonical PAS element with an average of 209nt between them while the remaining 29% have a canonical PAS element on the shorter isoform and a variant PAS element on the longer isoform with an average of 322nt between them. B) In genes with three or more 3'UTR isoforms, genes where the longest and the shortest isoform both have canonical PAS elements have an average of 163nt between them while genes where the longest and the shortest isoforms both have variant PAS elements have an average of 211nt between them. 72% of genes switch from a variant PAS elements in the short isoform to a canonical PAS element in the long isoform, with an average of 181nt between them. 28% of genes have canonical PAS elements in the short isoforms and variant PAS elements in the long 3'UTR isoform, with an average of 211nt between the two.





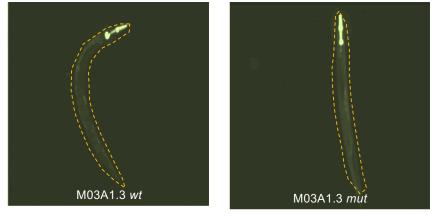
Supplemental Fig. S5: Detection of the 'UGUA' element in *C. elegans* **3'UTRs.** A) Logo plot of genomic region within the cut site in genes with only one 3'UTR isoform and with canonical or variant PAS element. There are not immediate elements visible. (T) = Pyrimidine, (P) = Purine. B) Identification of the 'UGUA' motif (red) within 100nt upstream of the cleavage site in genes with one 17 or two 3'UTR isoforms.



Supplemental Fig. S6: Nucleotide binding site of the human CPSF160-WDR33-CPSF30 complex. Ribbon representation of the cryo-EM structure of human CPSF160-WDR33-CPSF30 complex (PDB code: 6DNF) (Sun et al., 2018). The nucleotides of the bound RNA fragment do not show a specific interaction with either CPSF30 or WDR33. The interactions are mostly established by π - π ring stacking. Color gray shows the CPSF160, pink for WDR33, and light green for CPSF30. Sticks represent the RNA molecules bound with CPSF30 and WDR33. Surfaces in the inlets are for individual nucleotides.

M03A1.3 wt 3'UTR

TGAAAGGACCTGCAGTGTTTTGGGCGATTGGAGTATTCTTCTGCATTGCT GTTGCGTTGTCACTTCTTGTCGTCAATGGATATAAAAATGTA AATGGAATTTTGGAATCTCATCTAATTTATTGATTTTATTGAATACGGGT AGTTTCTGATAATTACTTTGCATTGTAAAAAAAACAAACTTTGTATGAATA AACATATTGAACATCTAA GTGCTTGCGTTTTTTTAAAACTCAACTTTGGTT GCGCATATCTTGGCTCTCTTTAGTTTTTAAAAAATGTCAACTACAGA



B

wt

M03A1.3 wt 3'UTR AATAAACATATTGAACATCT **ϪϹͲϹϹͲͲϹϹϹͲͲͲͲͲͲͲϪϪϪ** AATAAACATATTGAACATCTaaaaaaaaaa #1 AATAAACATATTGAACATTCaaaaaaaaa #2 AATAAACATATTGAACATCTaaa #3 AATAAACATATTGAACATCC<mark>aaaaaaa</mark> #4 AATAAACATATTGAACATCG clone #5 AATAAACATATTGAACAGCaaaaaaaaa #6 AATAAACATATTGAACAGaaaaaaaaaa #7 AATAAACATATTGAACATCCaaaaaaaaaa #8 #9 #10

M03A1.3 mut 3'UTR STOP AATAAACATATTGAACATCTCCGTGCTTGCGTTTTTTTAAA #1 → #2 #3 clone #4 → AATAAACATATTGAACaaaaaaaaa #5 AATAAACATATTGAACATCATaaaaa #6 #7 → #8

Supplemental Fig. S7: *In vivo* cleavage assay for M03A1.3

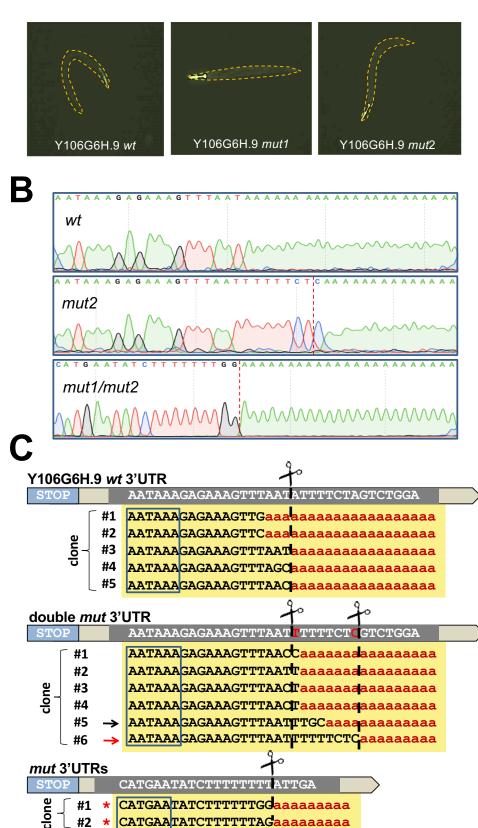
A) M03A1.3 genomic region cloned downstream of the GFP reporter. Blue: terminal portion of the M03A1.3 ORF. Green: STOP codon. Gray: 3'UTR. Red: mutated terminal Adenosine nucleotides. The transgenic worms expressing the pmyo-3::GFP::M03A1.3_3'UTR *wt* and mutant cassette are shown below.

B) At the completion of the experiment, we recovered the total RNA and performed RT-PCR experiments using a forward primer annealing within the GFP ORF and a reverse polydT primer with two anchors containing Invitrogen Gateway adapters. The resultant amplicons were then subcloned in gateway vectors and sequenced to detect the cleavage site. An example of resultant trace files is shown. C) Examples of 10 clones identified in this study for M03A1.3. The removal of the terminal genomic Adenosine nucleotide induces a cleavage site 3nt upstream of the canonical cleavage site in three clones (arrows), which also contain a terminal Adenosine nucleotide. The PAS element is boxed in blue color.

Supplemental Fig. S8

Y106G6H.9 wt 3'UTR

TTTTGAAATGCAAACAACACTCCGCAGTTTTGTTTGGAAAAACGAATTGGT CTACTTCTTCATAAAACATATGCGGTTCAATTGATACTTTTATTTCCATT GGAATTAAATTTAATGAATTGCTTCTTTAAATATTATTTCTATGCATCTG TTCTTCCTTTTGATTCTTCCATGAATATCTTTTTTTÅTTGATCCTACAG GATCGTACAGGATCTTGTCACACTAAAGATATCTACATATTTAATAATGT TCACCTTTGTTTTCTATTCTTCATGCCAATAAAGAGAAAGTTTAAT<mark>A</mark>TTT TCT<mark>A</mark>GTCTGGAATTTTTTATTTTTTAAAAAGCTGTCAACTGACAAATTATTG TCCACGACTTCGTCTGTTATTTTTAGTGAACTAAATGTTAGATCGACAGT



CATGAA TATCTTTTTTAGaaaaaaaaa

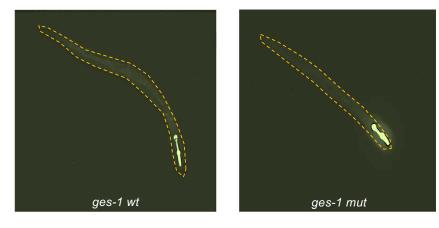
#2

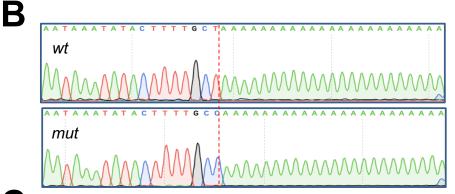
Supplemental Fig. S8: In vivo cleavage assay for Y106G6H.9

A) Y106G6H.9 genomic region cloned downstream of the GFP reporter. Blue: terminal portion of the Y106G6H.9 ORF. Green: STOP codon. Gray: 3'UTR. Red: mutated terminal Adenosine nucleotides. Red Asterisk: position of the cryptic cleavage site (see below). The transgenic worms expressing the pmyo-3::GFP:: Y106G6H.9 3'UTR wt and mutant cassette are shown below. B) At the completion of the experiment we recovered the total RNA and performed RT-PCR experiments using a forward primer annealing within the GFP ORF and a reverse polydT primer with two anchors containing Invitrogen Gateway adapters. The resultant amplicons were then subcloned in gateway vectors and sequenced to detect the cleavage site. An example of resultant trace files is shown.

C) Examples of several clones identified in this study for Y106G6H.9. In the wt we were able to detect two classes of cleavage sites, both ending within 4nt of each other with a terminal Adenosine nucleotide. In the double mutant, the removal of the terminal genomic adenosine induces a cleavage skip in two clones (arrows). In one case (red arrow), the cleavage occurs 20nt downstream of the PAS element. Two of the mutant clones also shown an occurrence of a new cryptic cleavage site 100nt upstream of the natural site (red asterisks), which also contain a terminal Adenosine nucleotide at their 3'end. The PAS element is boxed in blue color.

TCCAACTAATAGTGCCATGCATTCGTCAAACAAGGACGAGCTGTAAAAAT GCAATAAATTTATGTATTTAATTGATTTCG<u>AATAAA</u>TATACTTTTGCT<mark>A</mark>C AAA</mark>TCTTCGGCAAATGCTCATGCTCGATTTCTCCCCGGCCAATTGAGCACC TGTCATTTATCTTGTCATTTTTCCTGTACAAACACTTCTTGCCCCGACCA





9.

C

ges-1 wt 3'UTR					
S	TOP	AATAAATATACTTTTGCTACAAATCTTCGGCAAA			
	∏ #1	AATAAATATACTTTTGCTGGaaaaaaaaaaaaaaaaaaaa			
	#2	AACAAATATACTTTTGCTACaaaaaaaaaaaaaaaaaaaa			
	#3	AATAAATATACTTTTGCTAGaaaaaaaaaaaaaaaaaaaa			
	#4	AATAAATATACCTTTACdaaaaaaaaaaaaaaaaaaaaaa			
clone ^	#5	AATAAATATACTTTTGCCACaaaaaaaaaaaaaaaaaaaa			
- clo	#6	AATAAATATACTTTTGCTaaaaaaaaaaaaaaaaaaaaaa			
	#7	AATAAATATACTTTTGCTACGaaaaaaaaaaaaaaaaaaa			
	#8	AATAAATATACTTTTGCTGaaaaaaaaaaaaaaaaaaaaa			
	#9	AATAAATATACTTTTGGaaaaaaaaaaaaaaaaaaaaaaa			
	#10	AATAAATATACTTTTGCTACaaaaaaaaaaaaaaaaaaaa			

ges-1 mut 3'UTR

0	-	
STOP		AATAAATATACTTTTGCTTCTTCGGCAAA
	∏ #1	AATAAATATACTTTTGCOaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa
	#2	AATAAATATACTTTTGCTACaaaaaaaaaaaaaaaaaaaa
	#3	AATAAATATACTTTTGCTTCaaaaaaaaaaaaaaaaaaaa
	#4	AATAAATATACCTTTGCTTCCCGaaaaaaaaaaaaaaaaa
clone J	#5	AATAAATATACTTTTGCTACaaaaaaaaaaaaaaaaaaaa
	#6	AATAAATATACTTTTGCTTCACaaaaaaaaaaaaaaaaaa
	#7	AATAAATATACTTTTGCTTCGaaaaaaaaaaaaaaaaaaa
	#8	AATAAATATACTTTTGCTGCaaaaaaaaaaaaaaaaaaaa
	#9	AATAAATATACTTTTGCTCCaaaaaaaaaaaaaaaaaaaa
	#10	AATAAATATACTTTTGCTTTGaaaaaaaaaaaaaaaaaaa

Supplemental Fig. S9: *In vivo* cleavage assay for *ges-1*

A) *ges-1* genomic region cloned downstream of the GFP reporter. Blue: terminal portion of the *ges-1* ORF. Green: STOP codon. Gray: 3'UTR. Red: mutated terminal Adenosine nucleotides. The transgenic worms expressing the pmyo-3::GFP::ges-1_3'UTR *wt* and mutant cassette are shown below.

B) At the completion of the experiment we recovered the total RNA and performed RT-PCR experiments using a forward primer annealing within the GFP ORF and a reverse polydT primer with two anchors containing Initrogen Gateway adapters. The resultant amplicons were then subcloned in gateway vectors and sequenced to detect the cleavage site. An example of resultant trace files is shown.

C) Examples of 10 clones identified in this study for *ges-1*. The removal of the terminal genomic Adenosine nucleotide does not alter the cleavage site but makes it more variable. The PAS element is boxed in blue color.