Zhao et al, Supplemental material

Table S1: Conserved eukaryotic cleavage and polyadenylation related factors*

Factors (mammalian/yeast)	Subunits (mammalian/yeast)	Factors (plants)	Subunits (Arabidopsis)	Arabidopsis gene IDs
CPSF/CPF	CPSF-160/Yhh1p	CPSF	AtCPSF160	At5g51660
CPSF/CPF	CPSF-100/Ydh1p	CPSF	AtCPSF100	At5g23880
CPSF/CPF	CPSF-73/Ysh1p	CPSF	AtCPSF73-I	At1g61010
unknown/none	RC-68/none	CPSF	AtCPSF73-II	At2g01730
CPSF/CPF	CPSF-30/Yth1p	CPSF	AtCPSF30	At1g30460
CPSF/CPF	hFip1/Fip1p	CPSF	AtFIP5	At5g58040
unknown/CPF	hPfs2/Pfs2	CPSF	AtFY	At5g13480
CstF/CF IA	CstF-77/Rna14p	CstF	AtCstF77	At1g17760
CstF/CF IA	CstF-64/Rna15p	CstF	AtCstF64	At1g71800
CstF/none	CstF-50/None	CstF	AtCstF50	At5g60940
CstF/none	CFIm-25/None		None	
CstF/none	CFIm-68/None		None	
CF IIm/CF IA	hPcf11/Pcf11p	Unknown	AtPCFS1	At1g66500
			AtPCFS4	At4g04885
			AtPCFS5	At5g43620
CF IIm/CF IA	hClp1/Clp1p	Unknown	AtCLPS3	At3g04680
			AtCLPS5	At5g39930
unknown/CPF	Symplekin/Pta1p	Unknown	AtSYM5	At5g01400

*Note: Factors that are not discussed in this paper, such as poly(A) polymerase (PAP), RNA polymerase II C-terminal domain (Pol II CTD), and poly(A) binding protein II (PABP II), are not listed in this table. For a detailed list of eukaryotic cleavage and polyadenylation factors, please refer to Mandel et. al., 2008.

Zhao et al, Supplemental material.

Table S3. Polyadenylation-related proteins purified via TAP-fused AtCPSF and AtCLPS3

Identified proteins	Locus ID	Functions
From TAP-vector purified plant proteins		
None		
From TAP-AtCPSF30 purified plant proteins		
AtCPSF30 [1], [2]	At1g30460	The 30 kDa subunit of cleavage and polyadenylation specificity factor (CPSF); interacts with itself and with calmodulin; binding to RNA is inhibited by calmodulin in a calcium-dependent fashion (Delaney et al., 2006; Addepalli and Hunt, 2007; Hunt et al., 2008)
AtFIPS5 [1], [3]	At5g58040	A subunit of the polyadenylation apparatus that interacts with and stimulates the activity of poly(A) polymerase; RNA-binding protein; interacts with several polyadenylation factor subunits and coordinates a number of polyadenylation factor subunits with PAP and RNA (Forbes et al., 2006; Hunt et al., 2008)
From TAP-AtCPSF73-I purified plant proteins		
AtCPSF73-I [1], [2]	At1g61010	The 73 kDa subunit of CPSF; contains RNA-metabolizing domain (metallo-beta-lactamase) (Ryan et al., 2004; Dominski et al., 2005; Mandel et al., 2006; Xu et al., 2006; Hunt et al., 2008)
AtCPSF100 [1], [2]	At5g23880	The 100 kDa of CPSF; localized in nucleus; potentially functions in protein and DNA binding (Herr et al., 2006; Hunt et al., 2008; Mandel et al., 2008)
AtCPSF160 [2]	At5g51660	The 160 kDa subunit of CPSF; localized in nucleus; has nucleic acid binding activity (Hunt et al., 2008; Mandel et al., 2008)
AtFY [2]	At5g13480	A protein with similarity to yeast Pfs2p; an mRNA processing factor; involved in regulation of flowering time; affects FCA mRNA processing; has protein binding domains (Herr et al., 2006; Hunt et al.,

		2008)
AtCLPS3 [1]	At2g06480	pre-mRNA cleavage family complex protein (Hunt et al., 2008; Xing et al., 2008a)
From TAP-AtCPSF73-II purified plant proteins		
AtCPSF30 [1]	At1g30460	The same as described above
AtCPSF73-II [1] , [3]	At2g01730	A homolog of AtCPSF73-I that plays an essential role in the development of female gametophyte and embryo (Hunt et al., 2008)
AtCPSF100 [2]	At5g23880	The same as described above
AtCPSF160 [2]	At5g51660	The same as described above
AtFY [2]	At5g13480	The same as described above
From TAP-AtCPSF100 purified plant proteins		
CPSF100 [1]	At5g23880	The same as described above
CPSF160 [1]	At5g51660	The same as described above
From TAP-AtCLPS3 purified plant proteins		
AtCLPS3 [1], [2]	At3g04680	The same as described above
AtPCFS4 [1] , [3]	At4g04885	A homolog of yeast polyadenylation factor Protein 1 of Cleavage Factor (Pcf11p); involves in mRNA polyadenylation; regulates FCA (AT4G16280) mRNA polyadenylation; promotes flowering time (Herr et al., 2006; Hunt et al., 2008; Xing et al., 2008b)
AtSYM5 [1] , [3]	At5g01400	A subunit of CPSF; Symplekin/Pta1 homologue that has potential to interact with either ESP1 or AtCstF64; involves in posttranscriptional gene silencing by RNA and RNA processing (Herr et al., 2006)
From TAP-AtFY purified plant proteins		
AtFY [1], [2]	At5g13480	The same as described above
AtCPSF73-II [1] , [3]	At2g01730	The same as described above
AtCPSF100 [1], [2]	At5g23880	The same as described above
AtCPSF160 [2]	At5g51660	The same as described above

Note:
[1]. Polyadenylation-related proteins identified by mass-spectrometry
[2]. Polyadenylation-related proteins identified by antibodies
[3]. Antibody not available

Table S4. Sequence information of MS identified polyadenylation related protein factors.

The peptides high-lighted in red were confidently identified by MS**.**

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCPSF30	AtCPSF30	26

```
  1 MGTSVQVTPL CGVYNENPLS YLVSIDGFNF LIDCGWNDLF DTSLLEPLPR
 51 VASTIDAVLL SHPDTLHIGA LPYAMKQLGL SAPVYATEPV HRLGLLTMYD
101 QFLSRKQVSD FDLFTLDDID SAFQNVIRLT YSQNYHLSGK GEGIVIAPHV
151 AGHMLGGSIW RITKDGEDVI YAVDYNHRKE RHLNGTVLQS FVRPAVLITD
201 AYHALYTNQT ARQQRDKEFL DTISKHLEVG GNVLLPVDTA GRVLELLLIL
251 EQHWSQRGFS FPIYFLTYVS SSTIDYVKSF LEWMSDSISK SFETSRDNAF
301 LLRHVTLLIN KTDLDNAPPG PKVVLASMAS LEAGFAREIF VEWANDPRNL
351 VLFTETGQFG TLARMLQSAP PPKFVKVTMS KRVPLAGEEL IAYEEEQNRL
401 KREEALRASL VKEEETKASH GSDDNSSEPM IIDTKTTHDV VGSHGPAYKD
451 ILIDGFVPPS SSVAPMFPYY DNTSEWDDFG EIINPDDYVI KDEDMDRGAM
501 HNGGDVDGRL DEATASLMLD TRPSKVMSNE LIVTVSCSLV KMDYEGRSDG
551 RSIKSMIAHV SPLKLVLVHA IAEATEHLKQ HCLNNICPHV YAPQIEETVD
601 VTSDLCAYKV QLSEKLMSNV IFKKLGDSEV AWVDSEVGKT ERDMRSLLPM
651 PGAASPHKPV LVGDLKIADF KQFLSSKGVQ VEFAGGGALR CGEYVTLRKV
701 GPTGQKGGAS GPQQILIEGP LCEDYYKIRD YLYSQFYLL
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCPSF30	AtFIPS5	1

```
   1 MEEDDEFGDL YSDVLQPFQP PVVLPPPPPL PHRSIDLNLR SQDQDVSEPN
  51 SAPISRVSDN DAVKLSTQDA TRQAIVDGGG DDKDMSFDIE EPDADSTPTI
 101 PGLFVTGALP GLATDRGVSQ VTTRIEQQVG GGGDGGYGGQ GEGDDWDSDS
 151 EDDLQIVLND SSRNVMIGGA DRRSRMGDNE DDDDEDDEDP LVIVADTDPN
 201 QPMEEQMWGE DGLQGIEGDG KDGGEAGKGS GPGGATGPPK AGYSSHGYHP
 251 FHSQFKYVRP GAAPIPGGAA SVGGPSSGQV RPPANLGPMA GRGRGDWRPL
 301 GMRNASAAQK GFHQPWGSNT AGRGLDFTLP SHKTIFEVDI DSFEEKPWRY
 351 PGVEMTDYFN FGLNEESWKD YCKQLDQHRI QTTMQSRIRV YESGRTDQGY
 401 DPDLPPELAA ATGAQGVPVD SSNLVKPDSV QGDSAKVPAN VRPTLPPGRP
 451 IPVETGSGER LPSIDTRAPR MRDLDAIIED SHEDEPSGEN GTDQADSSLP
 501 GENVPVETSY VNNKRPDTES AEHSPAQDEP HKNLLKKQDD EISRSTDSGQ
 551 SFRSSSPVGD RGTRSSSVDR EDVGGEAGKD AEMGEELKMS FTSPQSAVQE
 601 DDGGESKTER SSESSKARSG SHRDFQQEED VIQDKHSSRP ANNRKQYDNN
 651 APHQSRKNQD RGKEMERTRA ASKGGRENSN PHMELDSTYI YSIASREDFD
 701 KRKERDVDGA VWRRKEDDPY SRRGGDEGSR KRDREDDPGF RQRGKMRENE
 751 IRSKDDQVPS RKHMDDAGMR NIYEPDDHIN KRRKDEEYLR RSRPEKNEIS
 801 YGQRESMSRV KRERDDRLEH QKRDVQHKIR DDFDDHGSLR QRDDIYMQRD
 851 GNERLRERDV LDKLKLPHED GISARGRERQ VAVRGHRGSE DRSSRMKDEY
 901 KASDKEHVTK DTLRHAKQTK RRDYPGEESS SHHRGHEDFS ARTDNIVNNE
 951 KKPRQERTGA KIDKFIDTLD GQRLQDRKHK DSRRKIKEQR EGTESLSKQG
1001 EQNGSSVVTG SKGTNDARNC RSEIPHQPNT AKRHKENASS GDEIHDSKRG
1051 RTKLERWASH KEREDAVSAK SSSISSKLEE KENNTNGRLS EPVHGSIGKS
1101 RDVTEEKIGH DLADTKDGSE KGPGDRHLDT VEKLKKRSER FKLPMPTEKD
1151 TTGVKKMESE TLPSAKIEGP VDSEGEYVWD ERSCVRIGRE YA
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCPSF73-I	AtCPSF73-I	50

```
  1 MASSSTSLKR REQPISRDGD QLIVTPLGAG SEVGRSCVYM SFRGKNILFD
 51 CGIHPAYSGM AALPYFDEID PSSIDVLLIT HFHIDHAASL PYFLEKTTFN
101 GRVFMTHATK AIYKLLLTDY VKVSKVSVED MLFDEQDINK SMDKIEVIDF
151 HQTVEVNGIK FWCYTAGHVL GAAMFMVDIA GVRILYTGDY SREEDRHLRA
201 AELPQFSPDI CIIESTSGVQ LHQSRHIREK RFTDVIHSTV AQGGRVLIPA
251 FALGRAQELL LILDEYWANH PDLHNIPIYY ASPLAKKCMA VYQTYILSMN
301 DRIRNQFANS NPFVFKHISP LNSIDDFNDV GPSVVMATPG GLQSGLSRQL
351 FDSWCSDKKN ACIIPGYMVE GTLAKTIINE PKEVTLMNGL TAPLNMQVHY
401 ISFSAHADYA QTSTFLKELM PPNIILVHGE ANEMMRLKQK LLTEFPDGNT
451 KIMTPKNCES VEMYFNSEKL AKTIGRLAEK TPDVGDTVSG ILVKKGFTYQ
501 IMAPDELHVF SQLSTATVTQ RITIPFVGAF GVIKHRLEKI FESVEFSTDE
551 ESGLPALKVH ERVTVKQESE KHISLQWSSD PISDMVSDSI VALILNISRE
601 VPKIVMEEED AVKSEEENGK KVEKVIYALL VSLFGDVKLG ENGKLVIRVD
651 GNVAQLDKES GEVESEHSGL KERVRVAFER IQSAVKPIPL SAS
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCPSF73-I	AtCLPS3	19

```
  1 MAYGGPSMNP PALSGAVPGS ANLKQVKLER ESELRIEVSE EPLRLRVVNG
 51 TAEIFGSELP PEIWRTFPPR MKFAVFTWYG ATIEMDGVTE TDYTADETPM
101 VSYINVHAIL DARRRFAKAS TSNDPESSQG PRVIVVGPTD SGKSTLTKML
151 LSWAAKQGWR PTFVDLDVGQ GSITIPGSIA AAPIEMPLDP VEGFPLDMAL
201 VYYYGHASPN MNVELYKALV KELAQVLEKQ FVGNPESRAA GMVINTMGWI
251 EGIGYELLLH AIDTFNASVV LVLGQEKLFS RLKDVLRSKS NVDVVKLHKS
301 GGVVARVKEV RKRSRNFKIQ EYFYGLSKEL SPYANTSSFS DLQVFRIGGG
351 PQAPKSALPA GSTSVSNPLR VTPVNIDDRD LLHSVLAVSY AEEPDQIISS
401 NVSGFVYVTE VNVQKKKITY LAPSPGTLPS KLLVAGSLAW LESV
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCPSF73-I	AtCPSF100	6

```
  1 MGTSVQVTPL CGVYNENPLS YLVSIDGFNF LIDCGWNDLF DTSLLEPLPR
 51 VASTIDAVLL SHPDTLHIGA LPYAMKQLGL SAPVYATEPV HRLGLLTMYD
101 QFLSRKQVSD FDLFTLDDID SAFQNVIRLT YSQNYHLSGK GEGIVIAPHV
151 AGHMLGGSIW RITKDGEDVI YAVDYNHRKE RHLNGTVLQS FVRPAVLITD
201 AYHALYTNQT ARQQRDKEFL DTISKHLEVG GNVLLPVDTA GRVLELLLIL
251 EQHWSQRGFS FPIYFLTYVS SSTIDYVKSF LEWMSDSISK SFETSRDNAF
301 LLRHVTLLIN KTDLDNAPPG PKVVLASMAS LEAGFAREIF VEWANDPRNL
351 VLFTETGQFG TLARMLQSAP PPKFVKVTMS KRVPLAGEEL IAYEEEQNRL
401 KREEALRASL VKEEETKASH GSDDNSSEPM IIDTKTTHDV VGSHGPAYKD
451 ILIDGFVPPS SSVAPMFPYY DNTSEWDDFG EIINPDDYVI KDEDMDRGAM
501 HNGGDVDGRL DEATASLMLD TRPSKVMSNE LIVTVSCSLV KMDYEGRSDG
551 RSIKSMIAHV SPLKLVLVHA IAEATEHLKQ HCLNNICPHV YAPQIEETVD
601 VTSDLCAYKV QLSEKLMSNV IFKKLGDSEV AWVDSEVGKT ERDMRSLLPM
651 PGAASPHKPV LVGDLKIADF KQFLSSKGVQ VEFAGGGALR CGEYVTLRKV
701 GPTGQKGGAS GPQQILIEGP LCEDYYKIRD YLYSQFYLL
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCPSF73-II	AtCPSF30	10

```
  1 MEDADGLSFD FEGGLDSGPV QNTASVPVAP PENSSSAAVN VAPTYDHSSA
 51 TVAGAGRGRS FRQTVCRHWL RGLCMKGDAC GFLHQFDKAR MPICRFFRLY
101 GECREQDCVY KHTNEDIKEC NMYKLGFCPN GPDCRYRHAK LPGPPPPVEE
151 VLQKIQQLTT YNYGTNRLYQ ARNVAPQLQD RPQGQVPMQG QPQESGNLQQ
201 QQQQQPQQSQ HQVSQTLIPN PADQTNRTSH PLPQGVNRCV QSPKVFNWVL
251
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCPSF73-II	AtCPSF73-II	35

```
  1 MAIDCLVLGA GQEIGKSCVV VTINGKKIMF DCGMHMGCDD HNRYPNFSLI
 51 SKSGDFDNAI SCIIITHFHM DHVGALPYFT EVCGYNGPIY MSYPTKALSP
101 LMLEDYRRVM VDRRGEEELF TTTHIANCMK KVIAIDLKQT IQVDEDLQIR
151 AYYAGHVLGA VMVYAKMGDA AIVYTGDYNM TTDRHLGAAK IDRLQLDLLI
201 SESTYATTIR GSKYPREREF LQAVHKCVAG GGKALIPSFA LGRAQELCML
251 LDDYWERMNI KVPIYFSSGL TIQANMYYKM LISWTSQNVK EKHNTHNPFD
301 FKNVKDFDRS LIHAPGPCVL FAIPGMLCAG LSLEVFKHWA PSPLNLVALL
351 GYSVAGTVGH KLMAGKPTTV DLHNGTKVDV RCKVHQVAFS PHTDAKGIMD
401 LTKFLSPKNV VLVHGEKPSM MILKEKITSE LDIPCFVPAN GETVSFASTT
451 YIKANASDMF LKSCSNPNFK FSNSTQLRVT DHRTADGVLV IEKSKKAKIV
501 HQDEISEVLH EKNHVVSLAH CCPVKVKGES EDDDVDLIKQ LSAKILKTVS
551 GAQIHESENC LQVASFKGSL CLKDKCMHRS SSSSSEAVFL CCNWSIADLE
601 LGWEIINAIK LNH
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCPSF100	AtCPSF100	40

```
  1 MGTSVQVTPL CGVYNENPLS YLVSIDGFNF LIDCGWNDLF DTSLLEPLPR
 51 VASTIDAVLL SHPDTLHIGA LPYAMKQLGL SAPVYATEPV HRLGLLTMYD
101 QFLSRKQVSD FDLFTLDDID SAFQNVIRLT YSQNYHLSGK GEGIVIAPHV
151 AGHMLGGSIW RITKDGEDVI YAVDYNHRKE RHLNGTVLQS FVRPAVLITD
201 AYHALYTNQT ARQQRDKEFL DTISKHLEVG GNVLLPVDTA GRVLELLLIL
251 EQHWSQRGFS FPIYFLTYVS SSTIDYVKSF LEWMSDSISK SFETSRDNAF
301 LLRHVTLLIN KTDLDNAPPG PKVVLASMAS LEAGFAREIF VEWANDPRNL
351 VLFTETGQFG TLARMLQSAP PPKFVKVTMS KRVPLAGEEL IAYEEEQNRL
401 KREEALRASL VKEEETKASH GSDDNSSEPM IIDTKTTHDV VGSHGPAYKD
451 ILIDGFVPPS SSVAPMFPYY DNTSEWDDFG EIINPDDYVI KDEDMDRGAM
501 HNGGDVDGRL DEATASLMLD TRPSKVMSNE LIVTVSCSLV KMDYEGRSDG
551 RSIKSMIAHV SPLKLVLVHA IAEATEHLKQ HCLNNICPHV YAPQIEETVD
601 VTSDLCAYKV QLSEKLMSNV IFKKLGDSEV AWVDSEVGKT ERDMRSLLPM
651 PGAASPHKPV LVGDLKIADF KQFLSSKGVQ VEFAGGGALR CGEYVTLRKV
701 GPTGQKGGAS GPQQILIEGP LCEDYYKIRD YLYSQFYLL
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCPSF100	AtCPSF160	30

```
   1 MSFAAYKMMH WPTGVENCAS GYITHSLSDS TLQIPIVSVH DDIEAEWPNP
  51 KRGIGPLPNV VITAANILEV YIVRAQEEGN TQELRNPKLA KRGGVMDGVY
 101 GVSLELVCHY RLHGNVESIA VLPMGGGNSS KGRDSIILTF RDAKISVLEF
 151 DDSIHSLRMT SMHCFEGPDW LHLKRGRESF PRGPLVKVDP QGRCGGVLVY
 201 GLQMIILKTS QVGSGLVGDD DAFSSGGTVS ARVESSYIIN LRDLEMKHVK
 251 DFVFLHGYIE PVIVILQEEE HTWAGRVSWK HHTCVLSALS INSTLKQHPV
 301 IWSAINLPHD AYKLLAVPSP IGGVLVLCAN TIHYHSQSAS CALALNNYAS
 351 SADSSQELPA SNFSVELDAA HGTWISNDVA LLSTKSGELL LLTLIYDGRA
 401 VQRLDLSKSK ASVLASDITS VGNSLFFLGS RLGDSLLVQF SCRSGPAASL
 451 PGLRDEDEDI EGEGHQAKRL RMTSDTFQDT IGNEELSLFG STPNNSDSAQ
 501 KSFSFAVRDS LVNVGPVKDF AYGLRINADA NATGVSKQSN YELVCCSGHG
 551 KNGALCVLRQ SIRPEMITEV ELPGCKGIWT VYHKSSRGHN ADSSKMAADE
 601 DEYHAYLIIS LEARTMVLET ADLLTEVTES VDYYVQGRTI AAGNLFGRRR
 651 VIQVFEHGAR ILDGSFMNQE LSFGASNSES NSGSESSTVS SVSIADPYVL
 701 LRMTDDSIRL LVGDPSTCTV SISSPSVLEG SKRKISACTL YHDKGPEPWL
 751 RKASTDAWLS SGVGEAVDSV DGGPQDQGDI YCVVCYESGA LEIFDVPSFN
 801 CVFSVDKFAS GRRHLSDMPI HELEYELNKN SEDNTSSKEI KNTRVVELAM
 851 QRWSGHHTRP FLFAVLADGT ILCYHAYLFD GVDSTKAENS LSSENPAALN
 901 SSGSSKLRNL KFLRIPLDTS TREGTSDGVA SQRITMFKNI SGHQGFFLSG
 951 SRPGWCMLFR ERLRFHSQLC DGSIAAFTVL HNVNCNHGFI YVTAQGVLKI
1001 CQLPSASIYD NYWPVQKIPL KATPHQVTYY AEKNLYPLIV SYPVSKPLNQ
1051 VLSSLVDQEA GQQLDNHNMS SDDLQRTYTV EEFEIQILEP ERSGGPWETK
1101 AKIPMQTSEH ALTVRVVTLL NASTGENETL LAVGTAYVQG EDVAARGRVL
1151 LFSFGKNGDN SQNVVTEVYS RELKGAISAV ASIQGHLLIS SGPKIILHKW
1201 NGTELNGVAF FDAPPLYVVS MNVVKSFILL GDVHKSIYFL SWKEQGSQLS
1251 LLAKDFESLD CFATEFLIDG STLSLAVSDE QKNIQVFYYA PKMIESWKGL
1301 KLLSRAEFHV GAHVSKFLRL QMVSSGADKI NRFALLFGTL DGSFGCIAPL
1351 DEVTFRRLQS LQKKLVDAVP HVAGLNPLAF RQFRSSGKAR RSGPDSIVDC
1401 ELLCHYEMLP LEEQLELAHQ IGTTRYSILK DLVDLSVGTS FL
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCPSF100	AtFY	20

```
  1 MYAGGDMHRG SQMPQPPMMR QSSASSTNIN PDYHHPSGPF DPNVDSFGAK
 51 RMRKHTQRRA VDYTSTVVRY IQARTWQRDS RDRTTLQPTP AAAVDMLPTV
101 AYSDNPSTSF AAKFVHASLN KNRCSINRVL WTPSGRRLIT GSQSGEFTLW
151 NGQSFNFEMI LQAHDQPIRS MVWSHNENYM VSGDDGGTLK YWQNNMNNVK
201 ANKTAHKESI RDLSFCKTDL KFCSCSDDTT VKVWDFTKCV DESSLTGHGW
251 DVKSVDWHPT KSLLVSGGKD QLVKLWDTRS GRELCSLHGH KNIVLSVKWN
301 QNGNWLLTAS KDQIIKLYDI RTMKELQSFR GHTKDVTSLA WHPCHEEYFV
351 SGSSDGSICH WIVGHENPQI EIPNAHDNSV WDLAWHPIGY LLCSGSNDHT
401 TKFWCRNRPA DNPRDVLMQN QGYNEQGFGR QPDNFQPSEA SPIPGAFVPG
451 LTRNEGTIPG IGIAMPFDAS SQGDHKQPLP GSMALGAPPL PPGPHPSLLG
501 SGQQQGYQQQ QQHQGHPQQM LPMPNMPHHQ LPPSSHMPLH PHHLPRPMQM
551 PPHGHMPPPS MPMSHQMPGS MGMQGGMNPQ MSQSHFMGAP SGVFQGQPNS
601 GGPQMYPQGR GGFNRPQMIP GYNNPFQQQQ QPPLPPGPPP NNNQQHQ
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCLPS3	AtCLPS3	37

```
  1 MAYGGPSMNP PALSGAVPGS ANLKQVKLER ESELRIEVSE EPLRLRVVNG
 51 TAEIFGSELP PEIWRTFPPR MKFAVFTWYG ATIEMDGVTE TDYTADETPM
101 VSYINVHAIL DARRRFAKAS TSNDPESSQG PRVIVVGPTD SGKSTLTKML
151 LSWAAKQGWR PTFVDLDVGQ GSITIPGSIA AAPIEMPLDP VEGFPLDMAL
201 VYYYGHASPN MNVELYKALV KELAQVLEKQ FVGNPESRAA GMVINTMGWI
251 EGIGYELLLH AIDTFNASVV LVLGQEKLFS RLKDVLRSKS NVDVVKLHKS
301 GGVVARVKEV RKRSRNFKIQ EYFYGLSKEL SPYANTSSFS DLQVFRIGGG
351 PQAPKSALPA GSTSVSNPLR VTPVNIDDRD LLHSVLAVSY AEEPDQIISS
401 NVSGFVYVTE VNVQKKKITY LAPSPGTLPS KLLVAGSLAW LESV
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCLPS3	AtPCFS4	16

```
  1 MDSEKILNPR LVSINSTSRK GMSVELPQKP PPPPSLLDRF KALLNQREDE
 51 FGGGEEVLPP SMDEIVQLYE VVLGELTFNS KPIITDLTII AGEQREHGEG
101 IANAICTRIL EAPVEQKLPS LYLLDSIVKN IGRDYGRYFS SRLPEVFCLA
151 YRQAHPSLHP SMRHLFGTWS SVFPPPVLRK IDMQLQLSSA ANQSSVGASE
201 PSQPTRGIHV NPKYLRRLEP SAAENNLRGI NSSARVYGQN SLGGYNDFED
251 QLESPSSLSS TPDGFTRRSN DGANPSNQAF NYGMGRATSR DDEHMEWRRK
301 ENLGQGNDHE RPRALIDAYG VDTSKHVTIN KPIRDMNGMH SKMVTPWQNT
351 EEEEFDWEDM SPTLDRSRAG EFLRSSVPAL GSVRARPRVG NTSDFHLDSD
401 IKNGVSHQLR ENWSLSQNYP HTSNRVDTRA GKDLKVLASS VGLVSSNSEF
451 GAPPFDSIQD VNSRFGRALP DGTWPHLSAR GPNSLPVPSA HLHHLANPGN
501 AMSNRLQGKP LYRPENQVSQ SHLNDMTQQN QMLVNYLPSS SAMAPRPMQS
551 LLTHVSHGYP PHGSTIRPSL SIQGGEAMHP LSSGVLSQIG ASNQPPGGAF
601 SGLIGSLMAQ GLISLNNQPA GQGPLGLEFD ADMLKIRNES AISALYGDLP
651 RQCTTCGLRF KCQEEHSKHM DWHVTKNRMS KNHKQNPSRK WFVSASMWLS
701 GAEALGAEAV PGFLPTEPTT EKKDDEDMAV PADEDQTSCA LCGEPFEDFY
751 SDETEEWMYK GAVYMNAPEE STTDMDKSQL GPIVHAKCRP ESNGGDMEEG
801 SQRKKMRS
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtCLPS3	AtSYM5	3

```
   1 MASYSRARLK DLANSAKSAT ELPPKLQRLR YMRRDLQKDD SVFPTELLPH
  51 LFDLLSDQFG AVRKFVAEIL GEIGLKYVEL IPEIVPLLIK SLEDETPAVA
 101 RQVIACGADL FRSTLERVAV QGLHSSELND LLESSWTWLI KFKDEICSVA
 151 FKQGNSGVKL CAMKFVEALI LLYTPHEGIE ADFNISILRG GHPVLKIGDL
 201 SIEASQKLGL LLDQLRHPAA KSLNSSTIIV LINSLSSVAK KRPAYCGRIL
 251 PVLLSLDPLS FLKGVYAAAT NLALKTVFLS CLKCTHPAAA PDRLTSALKE
 301 IEGGGQAAKA KDLFYKTNGS IQDKDSVEDT KVSVEENPLC ASSDVAESNL
 351 SRKRSGSEYN IDLNGDASDG KRARITPSVS EESTDGLNGN DGVSLPRVAS
 401 TSTGPSDSRG VSDSGPAQQL VGLFGTLVSQ GEKAIGSLEI LISSISADLL
 451 TDVVMANMHN IPPNCSSYAD GTDELVMNMC IVGSDAQIKY PPSFVAGVLS
 501 LSTAFPPIAA LINPHNEDEE VYSVHVDQQM FPAEDARTPP GLLATCDTSF
 551 PENEESNTVS PQNVHYIGNR ESGIPGLESS AQHDGSGALV TNVLSSTNVE
 601 AASKNQNASF SGKLLVDVIP SMSVDKLEEF SPKAVGTVAS ASQFVLPKIS
 651 APVVDLSDEE KDSLQKLVFL RIVEAYKQIS MSGGSQLRFS LLAHLGVEFP
 701 SELDPWKILQ EHVLSDYLNH EGHELTVRVL YRLYGEAEAE QDFFSSTTAA
 751 SAYESFLLTV AEALRDSFPP SDKSLSKLLG DSPHLPKSVL MLLESFCCPG
 801 SGEVEKDLQH GDRVTQGLSA VWSLILMRPG IRNDCLNIAL QSAVHHLEEI
 851 RMKAIRLVAN KLYSLSFITE QIEEFAKDRL FSVVSDDCDK MDLDLKSPPN
 901 KPQHSISGMS METPSEATSS STSVTEAQRC LSLYFALCTK VLRIFTILRL
 951 MTNLVFNIYK NASDPVKQAI HLQIPILVRT MGSSSELLKI IADPPSGSDN
1001 LLIQVLQTLT EGPTPSSELI LTIRKLFDTR IKDVEILFPI LPFLPRDDVL
1051 RIFPHMVNLP MEKFQVALSR VLQGSSQSGP VLSPSEALIA IHSIDPARDG
1101 IPLKQVTDAC NTCFAQRQTF TQQVLAGVLN QLVQQIPLPM LFMRTVLQAI
1151 GAFPALSDFI LEILSRLVSK QIWKYPKLWV GFLKCTQTTQ PQSYKVLLQL
1201 PPLQLGNALT KIPALRAPLT AHASQPEIQS SLPRSTLAVL GLVPDSQGTQ
1251 TSQVQANETQ TSQEQQQQQA SEPQQTSQSQ QVSVPLSHSQ VDHQEPSQVV
1301 ASQSQSSPIG TVQSAMSQSQ NSPIDTGRSE MSQSQNSPID TGRSEMSQSQ
1351 NSPIDTGRSE MSQSQNSPID TGRSEMSESQ SSPIGQSQSS PIGTGQSDMS
1401 QTPQVSDSSA PEPTSHTRTS DPQASSQTLR DDDEKIDDTA TSENEVTEIE
1451 KSKESSEEEE EEEEEEE
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtFY	AtFY	2

```
   1 MYAGGDMHRG SQMPQPPMMR QSSASSTNIN PDYHHPSGPF DPNVDSFGAK
  51 RMRKHTQRRA VDYTSTVVRY IQARTWQRDS RDRTTLQPTP AAAVDMLPTV
 101 AYSDNPSTSF AAKFVHASLN KNRCSINRVL WTPSGRRLIT GSQSGEFTLW
 151 NGQSFNFEMI LQAHDQPIRS MVWSHNENYM VSGDDGGTLK YWQNNMNNVK
 201 ANKTAHKESI RDLSFCKTDL KFCSCSDDTT VKVWDFTKCV DESSLTGHGW
 251 DVKSVDWHPT KSLLVSGGKD QLVKLWDTRS GRELCSLHGH KNIVLSVKWN
 301 QNGNWLLTAS KDQIIKLYDI RTMKELQSFR GHTKDVTSLA WHPCHEEYFV
 351 SGSSDGSICH WIVGHENPQI EIPNAHDNSV WDLAWHPIGY LLCSGSNDHT
 401 TKFWCRNRPA DNPRDVLMQN QGYNEQGFGR QPDNFQPSEA SPIPGAFVPG
 451 LTRNEGTIPG IGIAMPFDAS SQGDHKQPLP GSMALGAPPL PPGPHPSLLG
 501 SGQQQGYQQQ QQHQGHPQQM LPMPNMPHHQ LPPSSHMPLH PHHLPRPMQM
 551 PPHGHMPPPS MPMSHQMPGS MGMQGGMNPQ MSQSHFMGAP SGVFQGQPNS
 601 GGPQMYPQGR GGFNRPQMIP GYNNPFQQQQ QPPLPPGPPP NNNQQHQ
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtFY	AtCPSF100	8

```
  1 MGTSVQVTPL CGVYNENPLS YLVSIDGFNF LIDCGWNDLF DTSLLEPLPR
 51 VASTIDAVLL SHPDTLHIGA LPYAMKQLGL SAPVYATEPV HRLGLLTMYD
101 QFLSRKQVSD FDLFTLDDID SAFQNVIRLT YSQNYHLSGK GEGIVIAPHV
151 AGHMLGGSIW RITKDGEDVI YAVDYNHRKE RHLNGTVLQS FVRPAVLITD
201 AYHALYTNQT ARQQRDKEFL DTISKHLEVG GNVLLPVDTA GRVLELLLIL
251 EQHWSQRGFS FPIYFLTYVS SSTIDYVKSF LEWMSDSISK SFETSRDNAF
301 LLRHVTLLIN KTDLDNAPPG PKVVLASMAS LEAGFAREIF VEWANDPRNL
351 VLFTETGQFG TLARMLQSAP PPKFVKVTMS KRVPLAGEEL IAYEEEQNRL
401 KREEALRASL VKEEETKASH GSDDNSSEPM IIDTKTTHDV VGSHGPAYKD
451 ILIDGFVPPS SSVAPMFPYY DNTSEWDDFG EIINPDDYVI KDEDMDRGAM
501 HNGGDVDGRL DEATASLMLD TRPSKVMSNE LIVTVSCSLV KMDYEGRSDG
551 RSIKSMIAHV SPLKLVLVHA IAEATEHLKQ HCLNNICPHV YAPQIEETVD
601 VTSDLCAYKV QLSEKLMSNV IFKKLGDSEV AWVDSEVGKT ERDMRSLLPM
651 PGAASPHKPV LVGDLKIADF KQFLSSKGVQ VEFAGGGALR CGEYVTLRKV
701 GPTGQKGGAS GPQQILIEGP LCEDYYKIRD YLYSQFYLL
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtFY	AtCLPS3	5

```
  1 MAYGGPSMNP PALSGAVPGS ANLKQVKLER ESELRIEVSE EPLRLRVVNG
 51 TAEIFGSELP PEIWRTFPPR MKFAVFTWYG ATIEMDGVTE TDYTADETPM
101 VSYINVHAIL DARRRFAKAS TSNDPESSQG PRVIVVGPTD SGKSTLTKML
151 LSWAAKQGWR PTFVDLDVGQ GSITIPGSIA AAPIEMPLDP VEGFPLDMAL
201 VYYYGHASPN MNVELYKALV KELAQVLEKQ FVGNPESRAA GMVINTMGWI
251 EGIGYELLLH AIDTFNASVV LVLGQEKLFS RLKDVLRSKS NVDVVKLHKS
301 GGVVARVKEV RKRSRNFKIQ EYFYGLSKEL SPYANTSSFS DLQVFRIGGG
351 PQAPKSALPA GSTSVSNPLR VTPVNIDDRD LLHSVLAVSY AEEPDQIISS
401 NVSGFVYVTE VNVQKKKITY LAPSPGTLPS KLLVAGSLAW LESV
```

Bait protein	Interacting protein identified by MS	Sequence coverage (%)
AtFY	AtCPSF73-II	2

```
  1 MAIDCLVLGA GQEIGKSCVV VTINGKKIMF DCGMHMGCDD HNRYPNFSLI
 51 SKSGDFDNAI SCIIITHFHM DHVGALPYFT EVCGYNGPIY MSYPTKALSP
101 LMLEDYRRVM VDRRGEEELF TTTHIANCMK KVIAIDLKQT IQVDEDLQIR
151 AYYAGHVLGA VMVYAKMGDA AIVYTGDYNM TTDRHLGAAK IDRLQLDLLI
201 SESTYATTIR GSKYPREREF LQAVHKCVAG GGKALIPSFA LGRAQELCML
251 LDDYWERMNI KVPIYFSSGL TIQANMYYKM LISWTSQNVK EKHNTHNPFD
301 FKNVKDFDRS LIHAPGPCVL FATPGMLCAG FSLEVFKHWA PSPLNLVALP
351 GYSVAGTVGH KLMAGKPTTV DLYNGTKVDV RCKVHQVAFS PHTDAKGIMD
401 LTKFLSPKNV VLVHGEKPSM MILKEKITSE LDIPCFVPAN GETVSFASTT
451 YIKANASDMF LKSCSNPNFK FSNSTQLRVT DHRTADGVLV IEKSKKAKIV
501 HQDEISEVLH EKNHVVSLAH CCPVKVKGES EDDDVDLIKQ LSAKILKTVS
551 GAQIHESENC LQSREKEESQ FSSLLMSSYS SDSTAARDQH APLLRPRHDG
601 SFSSSSSSAR PTALAVLLGR ITGHRAPSML VRETAARALE ERRIDWGYSK
651 PVVAADILWN AALVLASAVM LVGTVEERPN EPIRVWICVY GLQCLFHVVL
701 VWSEYWRRNS TRRARDLESY DHEDYNIEYD YEQDSDDNST TYRLSVIFLA
751 IDVFFAVFCV VLACLVGIAL CCCLPCIIAL LYAVAGTNLE TPFLAGFIQE
801 GVSEAELGVL PLYKFKAFHS NEKNITGPGL LHMSEFI
```

Zhao et al., Supplemental Materials and Methods

Constructions of fusion proteins

The Gateway system (Invitrogen Inc.) was used for the cloning of the Arabidopsis CPSF cDNA sequences (Xu et al., 2006) into the TAP tag containing binary vectors (TAPi; Rohila et al., 2004; a gift from Dr. Michael Fromm, University of Nebraska). Briefly, the PCR-amplified CPSF cDNA sequences were first cloned into the pENTR clones, as described by the manufacturer. After being confirmed by sequencing, the CPSF cDNAs were fused to the binary vectors through LR recombination reactions. After fusion, the TAPi tag sequences were either located at the C-terminus of the protein coding sequences, such as AtCPSF30, AtCPSF73-I, AtCLPS3, and AtFY, or at the N-terminus, as in the case of AtCPSF73-II and AtCPSF100 (Fig. 1). The cloning of AtCLPS3 and AtFY was described in Xing et al. (2008a and b). Plasmids containing the insertions encoding the TAP-fused proteins were transformed into the *Agrobacterium* strain GV3505 by electroporation (Xu and Li, 2008).

Cell culture conditions

Arabidopsis thaliana cell culture (MM1 from *Landsberg erecta,* a gift from Dr. Chris Makaroff, Miami University) was maintained and transformed as described by Menges and Murray (2004), with modifications. Briefly, wild-type or transgenic cell cultures were maintained in 250 ml flasks containing 50 ml MS medium (pH 5.7, 4.5 g Murashige Skoog basal salt mixture, 0.112 g vitamin B5, 30 g sucrose, 0.5 mg alpha-Naphthylacetic acid (NAA), and 0.05 mg kinetin, per liter). The cell cultures were rotated on a shaker at 130 rpm (23 ℃; 16/8 hr photoperiod). The cells were sub-cultured every 7 days by transferring 10% (5 ml) of old cell culture into fresh media. The system can be scaled up to 1000 ml cell culture in 5-liter flasks when large amounts of cells are needed for proteomic study.

Transformation of *Arabidopsis* cell cultures

To stably introduce the TAP-fused CPSF constructs into *Arabidopsis* suspension cell cultures, a modified *Agrobacterium*-mediated transformation method was used (Menges and Murray, 2004). For each transformation, 10 ml of an early stationary phase cell

culture (seven days after fresh sub-culturing) was sub-cultured in 50 ml MS medium for two days. Ten ml of this exponentially growing cell culture was again sub-cultured into 50 ml fresh MS medium and was used for transformation. The day before transformation, *Agrobacterium* was prepared by inoculating into Luria-Bertani (LB) media containing 40 μg/ml of tetracycline and shaken overnight at 28℃. The next day, 100 μl of *Agrobacterium* was washed three times with 1 ml MS medium (by centrifugation and re-suspension) and re-suspended in 900 μl MS medium.

For each transformation, 100 μl washed *Agrobacterium* and 100 μl (50 mM) acetosyringone (3', 5'-dimethoxy-4'-hydroxyacetophenone, final concentration 100 μM; Acros Organic Inc.) were added to the two-day-old fresh sub-culture mentioned above, and the mixture was co-incubated at 23°C for two days. Two days later, the infected cell culture was transferred into a 50-ml Falcon tube and centrifuged for 5 min at 387 rpm, without applying brake force. The cells were then gently washed 3 times with MS medium and cultured in 20 ml of MS medium, supplemented with Cefotaxime (Bioplus Inc.; 100 μg/ml final concentration), which eliminates remaining *Agrobacterium*. After 3 days, the culture was washed with MS medium three times and re-suspended in 10 ml fresh MS medium. Then, 2.5 ml cell culture was spread on a MS plate containing 0.8% (W/V) agar, 100 μg/ml cefotaxime and 800 μg/ml glufosinate ammonium (Sigma Inc.). Plates were kept under light for 2 to 3 weeks. When calli were about the diameter of 0.5 centimeters, a single callus from each transformation was selected, crunched by sterilized pipette tips, and cultured with 10 ml of MS medium containing 100 μg/ml cefotaxime and 800 μg/ml glufosinate ammonium. The expressions of the TAP-fused proteins were detected by Western blotting using an antibody specifically against the TAP tag (peroxidise anti-peroxidase soluble complex antibody, Sigma 1291). For each assay, cells from a 1 ml cell culture were collected by spinning. About 20 to 50 μl of 1X SDS-PAGE loading buffer was added to the pellet and boiled at 100 ^{o}C for 5 min. The samples were separated on a 12% SDS-PAGE gel. The proteins were transferred to a polyvinylidene fluoride (PVDF) membrane for detection, as described below.

TAP purification

TAP purification was done as described by Rohila et al. (2004), with slight modification. About 100 grams of cells (fresh weight) were used for a typical mass

spectrometry identification. Briefly, cells were ground in liquid nitrogen with acid washed quartz (Sigma Inc.) to fine powder. About 150 ml of extraction buffer (20mM Tris/HCl, pH8.0, 150 mM NaCl, 0.1% NP-40 [Igepal CA-630, Sigma Inc.], 2.5 mM EDTA, 10 mM β -mercaptoethanol, 1 mM phenylmethylsulphonyl fluoride (PMSF), 2 mM Benzamide, 20 mM NaF, and 0.1% [V/V] of protease inhibitors cocktail [Sigma Inc.]) was added to the ground cells and mixed by stirring for about 20 min on ice. The mixture was spun at 10,000xg for 40 min at 4 ^{o}C. The supernatant was transferred to a pre-chilled tube and incubated with 100 µl (bed volume) IgG beads (Immunoglobulin G; GE Health Inc.) at 4°C for 2-4 h. The IgG beads were then collected by passing through a mini column (Bio-Rad Inc.), where the beads were first washed with 10 ml immuno-precipitation-150 (IPP-150) buffer (10mM Tris/HCl, pH8.0, 150mM NaCl, 0.1% NP40, 1 mM PMSF and 1/1000 protease inhibitor cocktail) 3 times, followed by 10 ml tobacco etch viral protease **(**TEV) cleavage buffer (10mM Tris/HCl, pH8.0, 150mM NaCl, 0.1% NP40, 0.5mM EDTA, 1mM DTT, 1 mM PMSF and 1/1000 protease inhibitor cocktail) once. After the washes, 1 ml TEV cleavage buffer, 10 µl 0.1 M DTT, and 10 µl (100 units) of TEV (Invitrogen Inc.) were added to the same mini column. After mixing by gentle flicks, the column was incubated at 4°C overnight with gentle rocking. The next day, the TEV-digested mixture was passed through the mini column, and the flow-through was collected into a new mini column.

One hundred µl (bed volume) calmodulin affinity beads (Invitrogen Inc.) were prepared by equilibrating with 3 ml of calmodulin binding buffer (10 mM Tris/HCl, 150 mM NaCl, 1 mM Mg-acetate, 1 mM imidazole, 2 mM $CaCl_2$, 0.1% NP40 [V/V], 10 mM β-mercaptoethanol, 1 mM PMSF and 1/1000 protease inhibitor cocktail, pH 8.0) 3 times prior to calcium-dependent binding. Three ml of calmodulin binding buffer and 3 µl of 1M $CaCl_2$ were added to the flow-through collected from the previous step (IgG binding). The column was incubated at 4°C for 1 h. After incubation, the column was drained, and the beads were washed with 10 ml of IPP 150 three times. Finally, the proteins were eluted with 1 to 1.5 ml of elution buffer (10 mM Tris/HCl, 150 mM NaCl, 1 mM Mg-acetate, 1 mM imidazole, 2 mM ethylene glycol tetraacetic acid (EGTA), 0.1 % NP40 (V/V), 10 mM β-mercaptoethanol, 1 mM PMSF and 1/1000 protease inhibitor cocktail,

pH 8.0) into the desired number of fractions, which were monitored by absorbance at 280 nm.

Eluted proteins were precipitated by trichloroacetic acid (TCA) and sodium deoxycholate (DOC) before loading onto a SDS-PAGE gel. To each volume of proteins, 1/100 volume of 2% DOC was added and incubated on ice for 30 min. One hundred percent TCA was added to 6% of the final volume, and the mixture was kept on ice for 1 hour. The tubes were then centrifuged at 2,500 g for 45 min at 4ºC. After decanting the supernatant, the pellets were washed with cold 100% acetone and spun at 2,500 g for another 45 min at 4°C. The pellets were dried by a SpecVac for 1 min and dissolved in 10 to 20 μl (depending on the desired concentration) SDS-PAGE loading buffer before being separated by 12 % SDS-PAGE gels. The proteins on the gels were then transferred to a PVDF membrane and detected by antibodies, or analyzed by mass spectrometry as described below.

Trypsin digestion and MS analysis

After being separated by SDS-PAGE and stained by Coomassie blue R-250 (Sigma Inc.), the desired bands were cut from the gel and chopped into small fragments before being transferred into 1.5 ml Eppendorf tubes. Gel fragments were washed with 300 μl 50% methanol (V/V) by vortexing 15 min at room temperature. The supernatant was decanted by spinning. The washing was repeated once. Three hundred μl 50% acetonitrile (ACN; Sigma Inc.)/50mM NH_4HCO_3 (V/V; pH 9.0) was added and vortexed for 30 min. The supernatant was decanted after spinning. Three hundred μl 50% ACN/10mM NH_4HCO_3 (V/V) was added and vortexed for 30 min again. The supernatant was decanted after spinning. Five hundred μl 100% ACN was added and vortexed for 10 min. The supernatant was decanted after spinning. Then the tubes were dried in a SpecVac for 10 min. The dried gel can be stored at -20 °C for the next step.

The second day, 100 μl 10 mM DTT/25 mM NH_4HCO_3 was added to the dried gel and incubated at 56°C for 1 hour. The DTT/ NH_4HCO_3 solution was then removed, and 55 mM iodoacetamide (Sigma Inc.) was added to merge the gel (ca. 25 μl). The samples were incubated in the dark for 45 min at room temperature. Then iodoacetamide was removed, and the samples were washed with 500 μl 25 mM NH_4HCO_3 by vortexing for 10 min. Finally, the liquid was removed, and the gel was dehydrated by adding 100 μl

100% ACN and vortexing 5 min. After drying in a SpecVac for 5 min, the samples were ready for Trypsin digestion.

The digestion was started by adding 250 μl 10mM NH_4HCO_3 (pH 9.0) and 0.5 μg sequence grade Trypsin (Promega Inc.) to the gel. The samples were kept at 37°C overnight. The next day, the solution containing the digested proteins was separated from the gel by transferring to a new tube where final digestion products were pooled. Two hundred and fifty μl 0.1% trifluoric acid (TFA)/water was added to the tubes containing the gels, and the tubes were shaken for 30 min. The liquid was transferred to the tubes where they were pooled. Two hundred and fifty μl 0.1% TFA/30% ACN (V/V) was added to the gel tubes, and the tubes were shaken for 30 min. The supernatant was moved to the tube where samples were pooled. Two hundred and fifty μl 0.1% TFA/60% ACN was added, and the tubes were shaken for 30 min. The supernatant was pooled again as described above. Two hundred and fifty μl 0.1% TFA/90% ACN (V/V) was added, and the tubes were shaken for 30 min. The supernatant was pooled again. Finally, the pooled products were dried to minimal volumes by a SpecVac and reconstituted into 10 μl TFA/water in the case of over-drying. The digested proteins were then sent to a proteomics facility for identification using liquid chromatographic with tandem mass spectrometry (LC/MS/MS, Mass Spectrometry and Proteomics Facility, Ohio State University).

Western blot analysis

After separation by SDS-PAGE, proteins were transferred to a PVDF membrane and detected using chemiluminescence. Briefly the proteins were first detected by primary antibodies raised against the proteins (Xu et al. 2006; Xing et al., 2008a; AtFY antibody was a gift from Caroline Dean, John Innes Centre) at 1:1000 dilution. Then a horse radish conjugated secondary antibody (goat anti-rabbit IgG-HRP; Sigma Inc.) was used at 1:2000 dilution for two hours (Copse and Fowler, 2002). After the incubation with the secondary antibody, the signal was detected using an ECL detection kit as described by the manufacturer's instructions (GE Healthcare Inc).