

Supplemental Figure 1. Unrooted tree demonstrating orthology among maize, rice, and *Arabidopsis* CFM2.

A phylogenetic tree that includes the complete CRM protein families from rice and *Arabidopsis* was reported previously (Barkan et al, 2007) and defines the CRS1 subfamily. To generate the tree shown here, all CRS1 subfamily members from rice (Os) and *Arabidopsis* (At), and the characterized members of the subfamily from maize (Zm) were aligned in Clustal X. An unrooted tree was generated in PAUP 4.0, using the neighbor-joining method with 1000 bootstrap replicates.

A

↓

```

ZmCFM2 1 MLLSFSPHFSS--LILSPSTSACK---PHARLRPVHASASAS-----TSPPELLGKLSALRRISDKLRLSLGYLETVS--E
OsCFM2 1 MLLFLFLPHSPPLLPAAASRTRPP---PRLLLPPIHASP-----SPELLAKLSALRRISDKLRLSLGYLEBADHPEA
AtCFM2 1 MLLFLFHQQPLIAKTFPDRIFPPFLV---PNTLVSRNNVSRANSNGIFCSSASGRKTLFQSAIQRIARLRLSLGYLEEKHSDP

ZmCFM2 68 PPTPAPNKSGDAPSPEGEIFVPTPAQLPRHRVGSSTLDPSSWAT-----GDGEASSTSRORRRGRGRDASGS---PSAFP
OsCFM2 67 APGPAAPEAGAGASPEGEIFVPTPAQLPRHRVGSSTLDPSSWAT-----GDGEGAAASRRRRRG-GRDSSAA---ASAFP
AtCFM2 81 TRRTIGEESGKN-SPGEIFVPLPKQLPIHRVGTITDPSWATPSYVPVKKP---GSGTAISRHELKRVVKKKETEEMERKKEKVP

ZmCFM2 137 SAAETALPRDELRRLOGICIRVRKRLKVKAGCTEGIVNGIHERWRNAEVVKLRCEVDVWAMMMRRHEILERKKGGLVIV
OsCFM2 135 SAAELALPRDELRRLOGACIRLRNRLKVKAGVTEGIVNGIHERWRNAELVKIRCDVDSAMMMKRTHEILERKKTGGLVIV
AtCFM2 160 SLAELTLPFAELRRLRTVGIRLTKKLLKIKGAGCTEGIVNGIHERWRTEVVKIFCEDISRMMMKRTHDVLERTKGGGLVIV

ZmCFM2 217 RSGSIIILYRGNTVYFYFHHSEKRVDSFLDKESDQNSGDEDETSSQHGSSHEKSSENP-VVACAEQIHVGEKNSQTI
OsCFM2 215 RSGSIIILYRGTDYKYPYFHDREMKNDMD--ESSEHTSSDDEADLAIASEQSGSEEDSDNPAEHSNHTEEGGDLTRR
AtCFM2 240 RSGSKIIILYRGVNYQYFYFVS-----DRLLAHEAASGASSMDQG-----

ZmCFM2 296 YLNQSLSEKEDTNHPVSSIKRVLDADEGNLDIRAGNPEQHVRLQENTHPDSFNKFGPRDRSSLVAGVGSQNKFRLOLP
OsCFM2 293 FGVDALEGNLDIGSAEQSIN-----SATKQQAILHSTNVSRSEISGRARSTLVAGVGSQNKFRLOLP
AtCFM2 279 -VVDSREK-----QSIASSASITNKMVKPMLTQGVGSPDKVRFQLP

ZmCFM2 376 GEVLAEEADKLLDGLGPRFSGWVCYDVPVDADLLPAIVPGYRFRLLLPSPVPPKLTDRREMTILRRLAHALPFHYALG
OsCFM2 358 GEVLAEEADKLLDGLGPRFSDWVCYDVPVDADLLPAIVPGYR-----
AtCFM2 321 GEVQLVEEADRLLEGLGPRFDWVCYDVPVDGDLPAIVPDRRFRLLPYGVSPKLTDEEMTTTTRRLGRPLCFHFLG

ZmCFM2 456 RSSNLQGLAASMIKWERCEVAKIALKRDAHNTDSELITEVVKEL-----TGGTLLSRDKESIVFVRGKDFLPPAV
OsCFM2 402 RSSNLQGLAASMIKWERCEVAKVAIKRGAENIDSDILEKLLKGL-----TGGTLLSRDNESIVFVRGKDFLPPAV
AtCFM2 401 RNRNLQGLAVAVIKLWEEKCELAKTAVKRVQNTNSLEMAELKVVGLLLVIKWLTTGGTLLSRDKDFIVFVRGKDFLPPAV

ZmCFM2 527 SLATEKRRKLGSSSTIYKAKPGIEESMPTQNDVSVLKVSDDVSVHVRE-----EGTSVTENRAESLNTVAKDVETRLSOA
OsCFM2 473 SLATEKRRKIGNSTISNPKLFDKSTP-QNSSKLMKMATDVSLDGHCYEEKKKHDETAVSDNRAESLNVFQNVEARLSOA
AtCFM2 481 SSAIEERRRQTMIMENS SVGHKLTENESEEIKPRAVKEDIIELEAKD--QKDHIQTHQMKSRQRNSPEALLEKTSMKLSMA
Atcfm2-1

ZmCFM2 600 IAEKAKAEKLEIELEKASPLSKAEV-REITSEDERMYLRKVLGKMKQFLLLRGRGVFDGTIENMHLHWKYRELVKIICKE
OsCFM2 552 IAEKKTKEKLEIELEMSSEPSRAET-REVISEDERMYLRKVLGKMKSFLLLRGRGVFDGTIENMHLHWKYRELVKIICKE
AtCFM2 559 LEKAKANAELVADLENRESQLSIDIKKCHTNDKMYLRKVLGKMKPFLLRGRGVFDGTIENMHLHWKYRELVKIICNE

ZmCFM2 679 HRLEDVEYAARTLEAESGGILVAVEKVSKGHAITVYRGKNYKRPKLRPKTLLSKRDALKRSLENORCKSLKVVHLKLSK
OsCFM2 631 HNIKDVVEYAARTLEAESGGILVAVERVSKAHAIIVYRGKNYQRPSTLRPKSLNKKDALKRSVEVQRYKSLKHLVNLKSK
AtCFM2 639 YSIEAAHVAEILEAESGGILVAVEMVSKGYAIVYRGKNYERPQCLRPQTLLSKREALKRSVEVQRKSLKHLVHLKLSN

ZmCFM2 759 NIDVLRDQ-----MNSSYHKMDHPVSNVTLTQQQDEEMPEVAPMSSEPEVEKWTSS--VEIDRALDLT
OsCFM2 711 NIDVLRDQMFVKQMEVQPVPTTNGMNSGHNQGILDNLVNSGTLVDKKEEVEVSLPECAKSVVVECCSSGESETRGTSVLT
AtCFM2 719 NIEEENRQ-----LVEDSATNWTWSDGESNMMVVEETENQHTPEKAREKIELGYSDDLVSVPSSGEEEN

ZmCFM2 821 KSGVPEVDMQSK-VCFNKLEDDSSATAGPCLTGSTSIAASSYNI LRHQNRSSVTSSPD-GRYEG---APSKVVDPAKL
OsCFM2 791 KSGVPLDVMQNKLLCFSKHTDDLSETTSSCLTSTSSSES--THQPLSSVMHNSDSHRVSGSKFVGTLPVHELKL
AtCFM2 783 WEDDSEGEVDPLTTSSQEYQDESESASSORHEGNSLDSSTANLSVFAETGSANASSFHDRSLPHNSFLNANRKLPGSSGT
Atcfm2-2

ZmCFM2 895 DAESLSVSPRAAP-----LSNQRLLVRKQALQMKKRPVLSIGRNNAITGVAKTKTHFKKHPLAIVNIKNRADGTP
OsCFM2 868 DEKS-SQLPSAAP-----LSNRERLLVRKQALQMKKRPVLAAGRNNVITGVAKAKTKTHFKKHPLAIVNIKNRADGTP
AtCFM2 863 SSGSISALRERKSENDGLVTDLSNRERLLVRKQALQMKKRPFAVGRSNVVTGLARTLKMHFQKNPLAIVNVKGRANGTS

ZmCFM2 968 IQQLISELEEAATGSLVLSRETNKVILYRGWGAEVAQNS-SRESSTDE-----GEKEVISPOLLEAIRLECGLLP
OsCFM2 940 IQQLISELEEAATGSLVLSREPNKVILYRGWGAEVAQNSLSGNNSTEQ-----VEKEVISPOLLEAVRLECGLLP
AtCFM2 943 VOEVHAKKEETGALLVSOEPSKVIYLRGWGAEEEMKSFYFNNNVKSINLPSTRSFVDDPPHVSVALLEAIRLECGLL--

ZmCFM2 1036 ADSG
OsCFM2 1009 GESE
AtCFM2 ----

```

B

GXXG

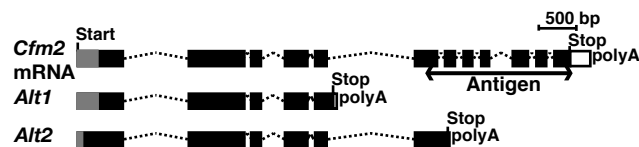
```

ZmCRM1 LPRDELRRLOGICIRVR--KRLKVKAGITGIVNGIHERWRNAEVVKLRCE-DVWAMNM
AtCRM1 LPPAELRRLRTVGIRLT--KRLKIKGAGITGIVNGIHERWRTEVVKIFCE-DISRMM
ZmCRM3 ISEDERMYLRKVLGKMK--QFLLLRGRGVFDGTIENMHLHWKYRELVKIICKE-ERHLEDV
AtCRM3 ITNDEKMYLRKVLGKMK--PFLLRGRGVFDGTIENMHLHWKYRELVKIICN-EYSIEAA
ZmCRM2 LTDREMTILRRLAHALP--PHYALGRSSNLQGLAASMIKWERCEVAKIALKRDAHNTDS
AtCRM2 LTDDEMTTIRRLGRPLP--CHFALGRNRLQGLAVAVKLVKWEKCELAKTAVKRVQNTNS
ZmCRM4 LSNQRLLVRKQALQMKKRPVLSIGRNNAITGVAKTKTHFKKHPLAIVNIKNRADGTP
AtCRM4 LSNRRLILVRKQALQMKKRPFAVGRSNVVTGLARTLKMHFQKNPLAIVNVKGRANGTSV

ZmCRM1 RRTHETLERK-----TGGLVIWRS---GSTIILYR 226
AtCRM1 KRTHDVLTK-----TGGLVIWRS---GSKTILYR 249
ZmCRM3 EYAARTLEAE-----SGGLVAVKVSQKHAITVYR 715
AtCRM3 HKVAETLEAE-----SGGLVAVEMVSKGYAIVYR 675
ZmCRM2 ELITEVVKEL-----TGGLLSRD---KESIVFYR 517
AtCRM2 ELMAEETKVVGLLLVIKWLTTGGTLLSRD---KDFIVLYR 471
ZmCRM4 QQLISELEEA-----TGSVLSRE---TNKVILYR 995
AtCRM4 QEVIAKLKEE-----TGALLVSOE---PSKVIYLR 970

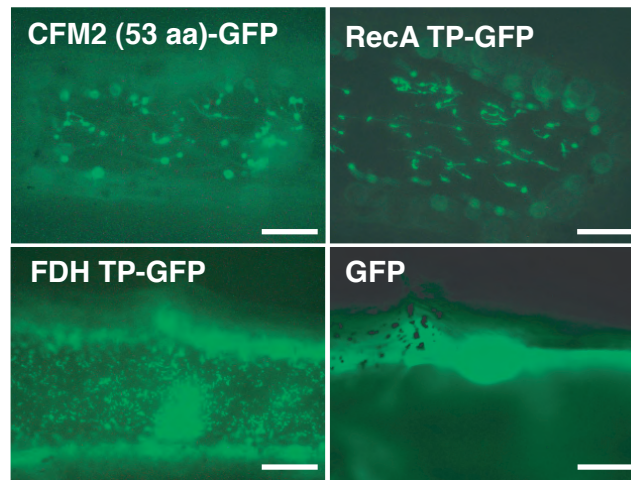
```

Supplemental Figure 2. Multiple sequence alignments of CFM2 proteins and CRM domains. The alignments were generated in ClustalW and shaded with BoxShade. Identical residues are shaded in black and similar residues in gray. **(A)** Alignment of AtCFM2, OsCFM2, and ZmCFM2. The predicted transit-peptide cleavage site in ZmCFM2 is indicated by the arrow. The positions of the T-DNA insertions in the Arabidopsis mutants analyzed here are indicated with triangles. The CRM domains are underlined. We believe this rice gene model (from TIGR) is erroneous as it is not supported by cDNAs and use of alternative splice junctions improves colinearity with the maize and Arabidopsis proteins. **(B)** Alignment of the CRM domains in AtCFM2 and ZmCFM2. The residues corresponding to the “GxxG” loop proposed to contribute to RNA binding (Barkan et al, 2007; Ostheimer et al, 2002) are indicated.

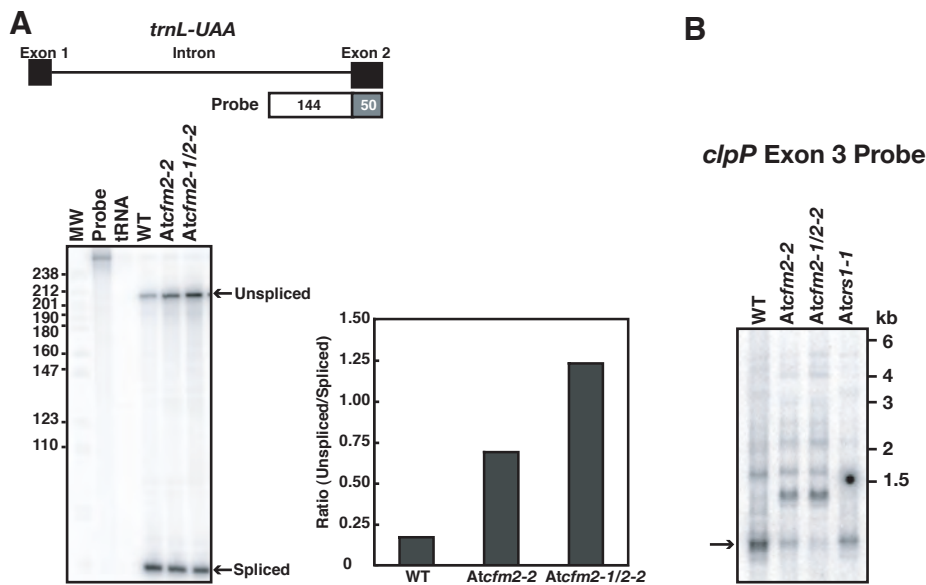


Supplemental Figure 3. Maize *cfm2* cDNA variants.

Exons are indicated with boxes and introns with dotted lines. The gray exon segments at the 5' ends were inferred from genomic DNA sequence and encode the start codon. The top diagram shows the cDNA encoding the full-length protein, which is best conserved with the predicted rice and Arabidopsis orthologs and is shown in Supplemental Figure 2. Two mRNA variants resulting from the use of alternative polyadenylation sites were detected as cDNAs and are diagrammed below. Alt1 and Alt2 encode proteins ending within and after the third CRM domain, respectively. These RNAs were not detected by RNA gel blot hybridization of seedling leaf RNA; the cDNAs may be derived from non-leaf tissues that contributed to the mixed tissue cDNA library. The smaller protein detected in leaf by CFM2 antisera is not derived from these alternative mRNAs because the Alt1 form lacks sequences encoding the antigen, and affinity purification of the sera against a peptide that is not encoded by Alt2, which in any case overlapped the antigen only slightly, did not change the ratio of the two bands detected.



Supplemental Figure 4. Chloroplast localization of CFM2-GFP in onion epidermal cells. A fusion protein consisting of the N-terminal 53 amino acids of maize CFM2 fused at their C-terminus to GFP was transiently expressed in onion root epidermal cells. GFP fused to the transit peptides of chloroplast RecA and mitochondrial FDH were used to visualize chloroplasts and mitochondria, respectively. The panel labeled GFP shows results of transformation with pOL-LT, the vector encoding GFP alone. Bars =5 μ m.



Supplemental Figure 6. Confirmation of splicing defects in *AtCFM2* mutants.

(A) Ribonuclease protection assay of *trnL-UAA* splicing in *AtCFM2* mutants. The radiolabeled probe spanned the 3' splice junction, as diagrammed at the top, and included vector-derived sequences that were digested during the assays. The ratios of unspliced to spliced RNAs were quantified with a phosphorimager, and are plotted in the bar graph.

(B) RNA gel blot hybridization showing loss of fully spliced *clpP* mRNA, indicated with an arrow.

Supplemental Table 1. Top-Ranking Fragments in CFM2 RIP-Chip Assays

Fragment name	Fragment Number ^a	α CFM2		Control		P Value ^c
		Median Log ₂ Ratio ^b	<i>n</i> ^b	Median Log ₂ Ratio ^b	<i>n</i> ^b	
trnL-UAA2	88	1.4	20	-4.2	10	4.5E-13
ycf3int1ex1	80	1.3	20	-2.3	10	1.4E-12
ycf3ex2int1	79	1.3	20	-3.0	10	3.8E-10
trnL-UAAint/ex2	89	0.5	15	-4.5	8	5.1E-04
ndhAint	243	-0.2	20	-3.4	10	3.9E-14
ycf3ex1/trnS5'	81	-0.4	20	-2.9	9	6.4E-10
ndhAintex1	244	-0.6	20	-3.3	10	1.0E-12
ycf3int2	78	-0.6	16	-3.4	10	1.8E-07
ndhAex2int	242	-1.3	20	-3.3	10	1.6E-04
ycf3int2ex3	77	-1.6	20	-3.4	10	8.3E-11
ndhBintron	185	-2.1	20	-3.6	9	2.4E-06
rrn16-3prime	200	-2.1	18	-1.8	9	4.4E-01
trnS-GGA1	82	-2.3	12	-3.9	8	1.1E-06
rrn16/trnlint	201	-2.4	12	-2.8	8	5.6E-01
rps12int2ex2	191	-2.4	20	-2.8	9	9.7E-03
rps12int13'	194	-2.5	20	-3.2	10	1.2E-01
trnI-GAUcomplete	203	-2.5	20	-2.3	10	6.3E-01
trnA-UGCcomplete	206	-2.5	12	-2.4	6	7.7E-01
ndhBint2	186	-2.6	12	-3.8	6	9.0E-04
trnI/trnA	205	-2.7	20	-2.0	10	4.7E-02
trnA-int	207	-2.8	20	-2.3	10	1.9E-02
rps12int2A	190	-2.8	20	-3.5	8	4.1E-02
ndhBex2int	184	-2.8	20	-2.8	9	7.2E-02
rps12ex2int13'	193	-2.8	20	-3.8	10	3.8E-03
ndhI/ndhA	241	-2.8	20	-3.6	10	2.8E-03
orf139	176	-3.0	20	-2.0	10	8.7E-03
ndhBex1	187	-3.0	17	-3.7	10	2.1E-02
rps7/rps12ex3	189	-3.0	20	-3.1	10	3.6E-01
trnI-GAUex1/int	202	-3.0	19	-3.2	10	6.0E-01
trnA-UGC/23S5'	209	-3.0	20	-2.6	10	1.7E-03
rps12int2B	192	-3.1	20	-3.4	10	1.2E-01
rpl2ex2int	167	-3.1	20	-3.7	10	6.3E-02
psaC/ndhE/ndhG	236	-3.1	16	-2.2	10	2.9E-02
rps4	84	-3.1	20	-2.4	10	2.3E-01
rpl2int/ex1	169	-3.1	20	-3.3	10	8.7E-02
trnS/rps4	83	-3.2	8	-3.7	8	1.9E-02
trnV-UACex2int	97	-3.2	19	-3.4	8	6.8E-01

Elements ranking in the top 15 % for median normalized enrichment ratio in four replicate assays with CFM2 antibodies are ordered according to the magnitude of their enrichment.

^aFragments on the array are numbered according to chromosomal position. The nucleotide residues on each fragment are described in Array Express (accession number A-MEXP-743) and in Schimitz-Linneweber et al. (2005).

^bMedian (log₂F635/F532) normalized across experiments with α CFM2 and control (α OE16 and α OE23). Replicate experiments constitute a total of *n* replicate spots.

^cP values were calculated with a *t* test (two-tailed, unequal variance) and represent the probability that there is no difference in enrichment between the α CFM2 and control immunoprecipitations.