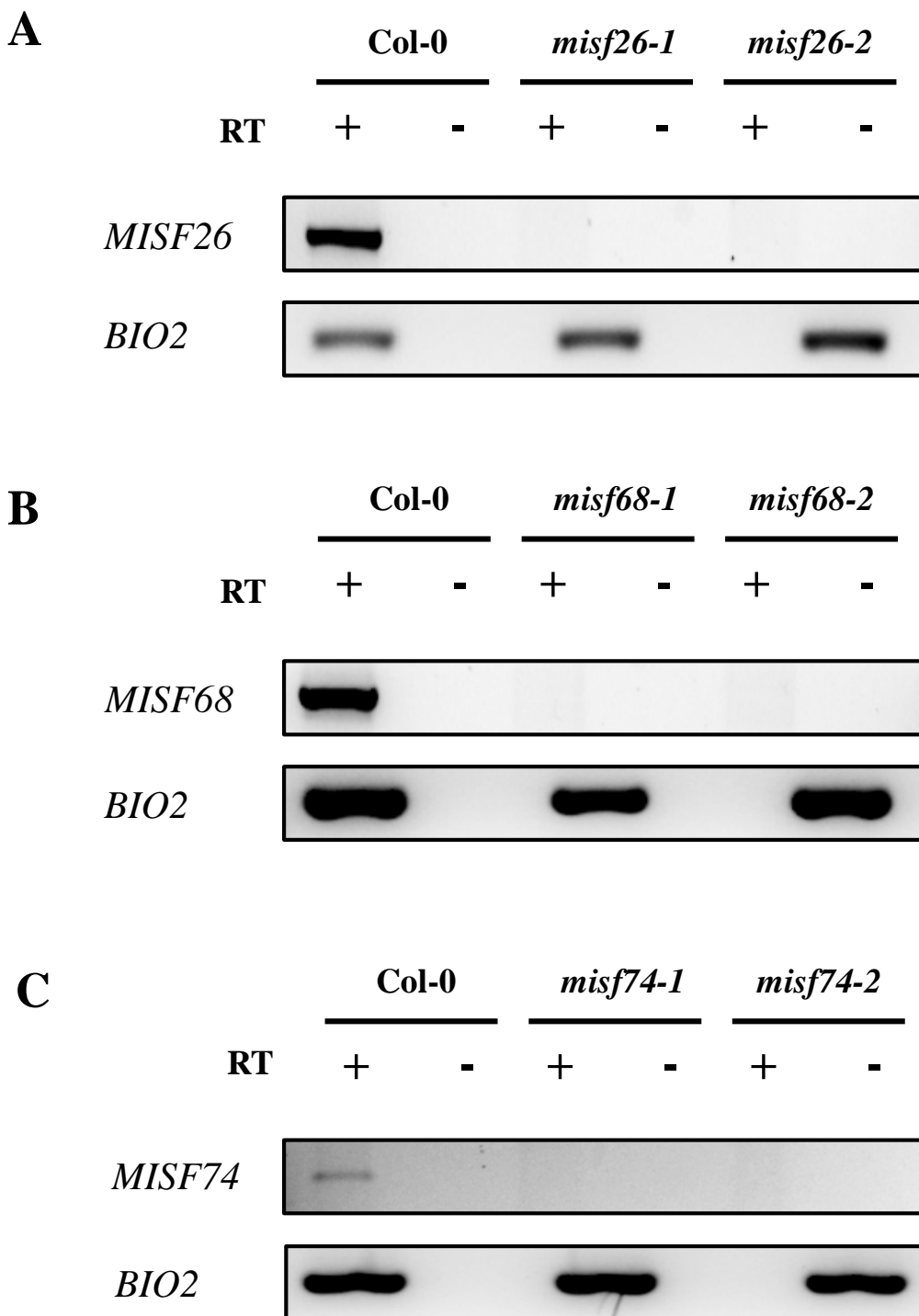
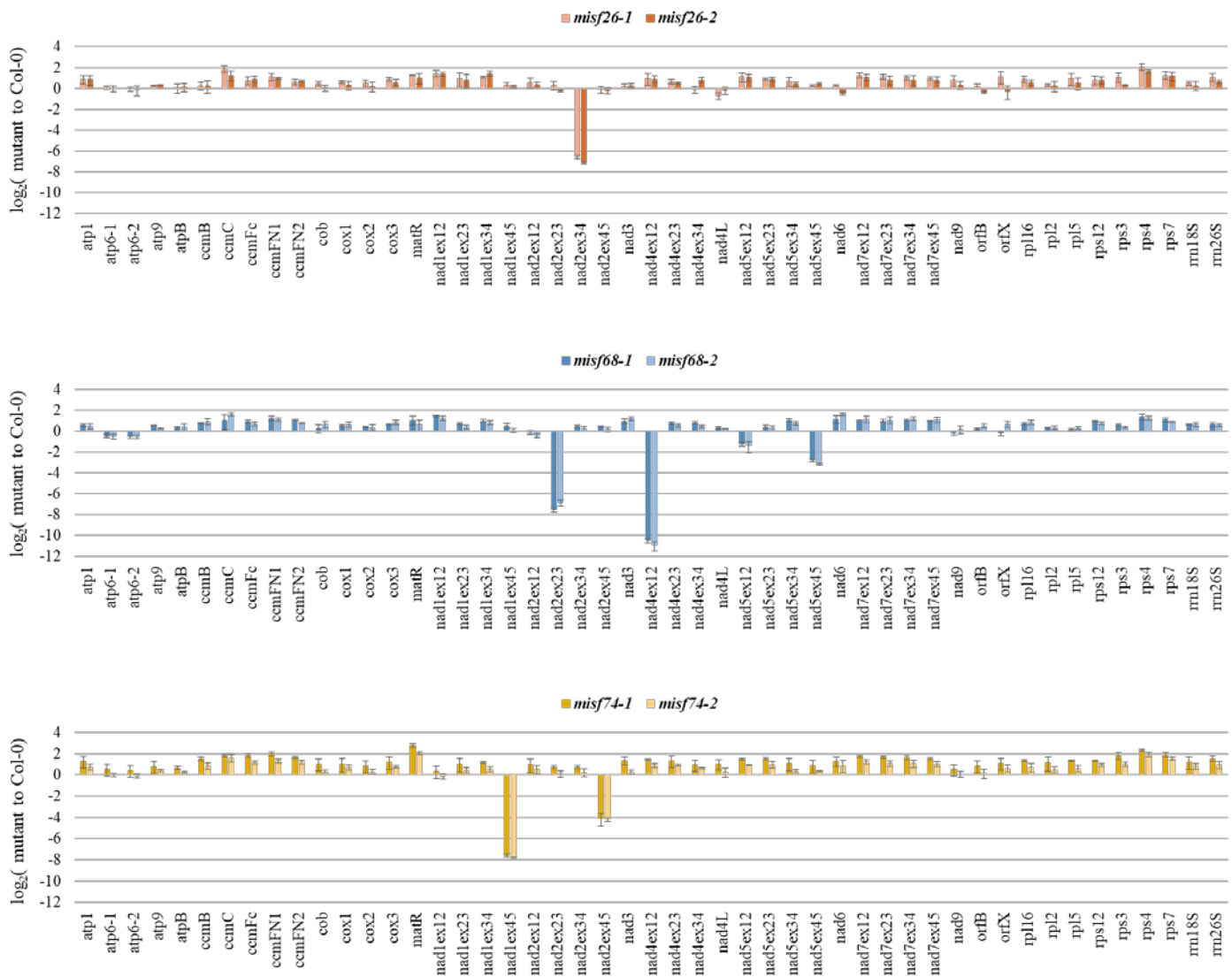


Name	Sequence 5'-3'	Experiment
GWPPR-P26-7	CAAAAAAGCAGGCTAAAATGGCGTCAGCTTTGCGCCG	GFP expression
GWPPR-P26-8	CAAGAAAGCTGGGTCAGAAGATGAGATCTCATAAG	
GWPPR-P68-5	CAAAAAAGCAGGCTAAAATGATGTATGTATCTGCTCG	
GWPPR-P68-6	CAAGAAAGCTGGGTCACGCTAACATATGTGTTAC	
GWPPR-P74-3	CAAAAAAGCAGGCTAAAATGATTCGCCGGCCGATCTA	
GWPPR-P74-6	CAAGAAAGCTGGGTCAGAATGGCGGAAATTGGGCTG	
PPR-P26-3	CGAACGAGATTACGATTCGG	Genotyping
PPR-P26-1	TTTGGTGGGAATGGAAGTTGAC	
PPR-P26-2	TGTTTAGGGTTTTGGAAGAGATG	
PPR-P68-1	TGCAATTTTGAATAGTTCCGG	
PPR-P68-2	CCGAGTTGTAAGTGCTCGAAG	
PPR-P68-3	TTCAGTTTTTCAAATGGGCTG	
PPR-P68-4	GCTCATTGCTGCTTCATATC	
PPR-P74-1	GCATGCCAGAATACGGTTATG	
PPR-P74-2	ATCCACATCAGGGACAACATC	
PPR-P74-4	TAACTGAGATACCGACCACC	
PPR-P74-7	AGCTTTTCAGGATCAGGAGGA	
LB-Gabi1	CCCATTGGACGTGAATGTAGACAC	
LB-Salk2	GCTTTCTCCCTTCCTTTCTC	
LB3-Sail	TAGCATCTGAATTTCATAACCAATCTCGATACAC	
RT-P26-1	TGGAGGATTTACGAATGTGCGA	RT-PCR
RT-P26-2	TTCATCGGGCACAACCCTT	
RT-PR68-1	TCAGTGC GGATGTTGGATGA	
RT-PR68-2	CACTCGACCCACTTTGCCTA	
RT-P74-1	CGCAATCGGATCCTCTCCTC	
RT-P74-2	ACCTATCAGGCACGAATCCTT	
nad1e7	CATCACCTAGCAAGCCTAAC	northern blot, nad1 probe
nad1e8	TAAGGAAGCCATTGAAAGGTG	
nad1 int4-1	ACGGGAAAGTTGACTCCTGG	northern blot, nad1 intron 4 probe
nad1 int4-2	AGGGCTTCCCTCAGTTCAGT	
nad2e1	GCAGAATTCGTTCCGGATC	northern blot, nad2 probe
nad2e8	TATGAACTGAGTGCCATTTGA	
nad2 int2a-1	ATGACGAAAGGAGAGTCGGC	northern blot, nad2 intron 2a probe
nad2 int2a-2	TTCTGGCCCTGTGAACATC	
nad2 int2b-1	CCAACGGGGAATAGAAGCGA	northern blot, nad2 intron 2b probe
nad2 int2b-2	ATACGGCTCGCACAAGACT	
nad2 int3-1	TTAGTCAATGGGGCAGCAGG	northern blot, nad2 intron 3 probe
nad2 int3-2	TCACACGAAAGGGAACGAGG	
nad2 int4-1	GGGAATGCATGCTACGAAAGA	northern blot, nad2 intron 4 probe
nad2 int4-2	TTCCGCCCCCAATGAGATG	
nad4F	ATTCTATGTTTTCCCGAAAGC	northern blot, nad4 probe
nad4R	TGAAATTTGCCATGTTGCAC	
nad4 int1-1	CCCGCCCAATCCAATAAGA	northern blot, nad4 intron 1 probe
nad4 int1-2	AGACCTGGACGATTATCCCT	
nad5F	AACATTGCAAAGGCATAATGA	northern blot, nad5 probe
nad5R	CCATGGATCTCATCGGAAAT	
nad5 intron4-1	TCCCTATGGACAAGGGGACA	northern blot, nad5 intron 4 probe
nad5 intron4-2	ACTCTTCATTCTCCCGTGCT	

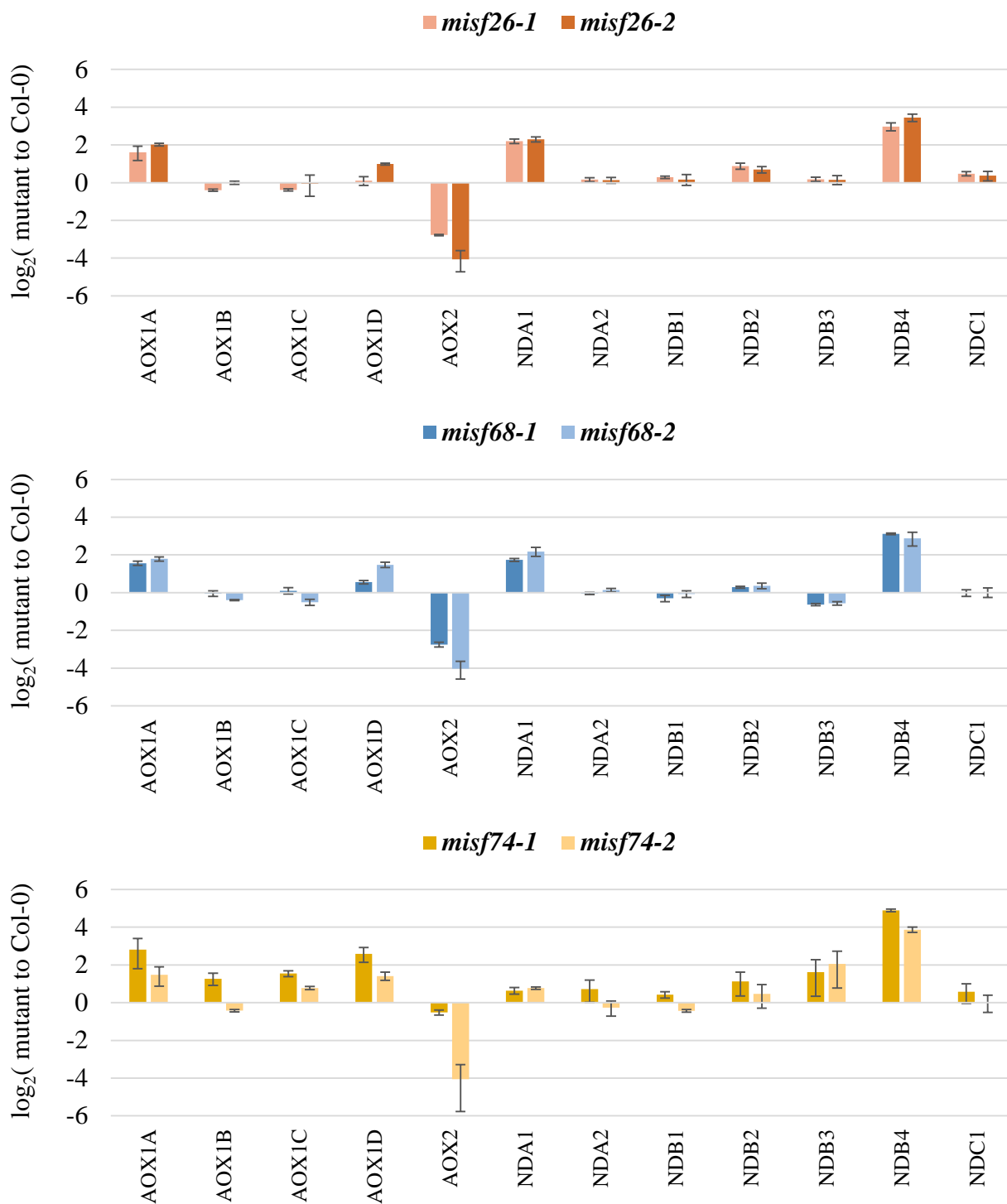
Supplementary Table S1: Oligonucleotides used in this study.



Supplementary Figure S1: RT-PCR analysis of each *MISF* transcript in wild-type and *misf* mutants. cDNA prepared from both wild-type and *misf* plants were PCR amplified using primers located on both sides of the T-DNA insertion sites corresponding to each *misf* mutation, and the resulting amplification products were size fractionated on agarose gel. The amplification of *BIO2* cDNA was done separately to control the efficiency of the RT reaction. Reverse transcriptase was either added (RT+) or omitted (RT-) from the initial RT reaction.



Supplementary Figure S2: Various mature mitochondrial transcripts under-accumulate in *misf* mutants. The steady state levels of mature mitochondrial mRNAs measured by quantitative RT-PCR in Col-0 and *misf* mutant plants. The histograms show the \log_2 ratio of mature transcripts in mutants to the wild-type (Col-0). A single PCR was considered for mRNAs carrying no introns, whereas the accumulation of individual exons was analyzed for intron-containing transcripts. Three technical replicates of two independent biological repeats were used for each genotype; standard errors are indicated. The data were normalized to the nuclear 18S rRNA gene.



Supplementary Figure S3: Quantitative RT-PCR measuring the steady-state levels of mRNAs encoding alternative oxidases (AOX) and NADH-dehydrogenases (NDA, NDB and NDC). Three technical replicates of two independent biological repeats were used for each genotype; standard errors are indicated.

Family	PPR										Maturase			Helicase		CRM	PORR	RCC1-like	mTERF	RAD52-like	
	ABO5	ABO8	BIR6	MISF26	MISF68	MISF74	OTP43	OTP439	SLO3	TANG2	nMAT1	nMAT2	nMAT4	PMH2	ABO6	mCSF1	WF9	RUG3	mTERF15	ODB1	
splicing factors																					
nad1 intron 1																					
nad1 intron 2																					
nad1 intron 3																					
nad1 intron 4																					
nad2 intron 1																					
nad2 intron 2																					
nad2 intron 3																					
nad2 intron 4																					
nad4 intron 1																					
nad4 intron 2																					
nad4 intron 3																					
group II introns																					
nad5 intron 1																					
nad5 intron 2																					
nad5 intron 3																					
nad5 intron 4																					
nad7 intron 1																					
nad7 intron 2																					
nad7 intron 3																					
nad7 intron 4																					
ccmFC intron																					
cox2 intron																					
rpl2 intron																					
rps3 intron																					

Supplementary Figure S4: Listing of known nucleus-encoded proteinaceous factors involved in the splicing of mitochondrial group II introns in Arabidopsis. Different nucleus-encoded protein factors assisting mitochondrial intron splicing has been identified *via* genetic screens in Arabidopsis so far. Bold black indicate trans-spliced introns. Introns whose splicing is affected by the indicated trans-factors are indicated by orange boxes. Corresponding references are cited in the introduction.

A		aa 5	D	N	R	T	S	T	N	N	S	R	D				
		aa 35	S	N	S	D	N	Y	S	D	G	N	R				
		Prediction	N	C/U	N	G	A	N	C	U/C	N	N	N	<i>p</i> -value			
<i>nad2</i> intron 3	BS1	G	C	A	G	A	C	C	U	U	U	C		4.84E-04			
	BS2	G	C	G	G	A	C	C	U	U	U	U		4.84E-04			
	BS3	G	C	G	G	A	G	C	C	C	G	U		8.55E-04			
B		aa 5	M	S	N	S	T	T	N	N	S	C	A	N	N	N	
		aa 35	S	T	D	T	T	D	T	S	C	S	D	D	D	D	
		Prediction	N	G/U	U/C	G/U	C	G	C/U	C	N	C/U	U/G	U/C	U/C	U/C	<i>p</i> -value
<i>nad2</i> intron 2	BS1	C	U	U	G	G	G	C	U	G	U	G	U	C	U		1.35E-04
	BS2	C	U	U	U	C	U	U	C	U	U	U	C	U	U		4.18E-04
	BS3	C	G	U	C	C	G	C	C	U	A	U	C	U	U		5.53E-04
	BS4	C	U	U	C	C	G	G	C	U	G	G	C	U	U		7.44E-04
	BS5	C	U	U	U	C	A	U	U	U	G	G	C	U	U		7.55E-04
	BS6	C	C	U	G	C	A	C	C	U	U	A	U	U	U		8.48E-04
<i>nad4</i> intron 1	BS1	G	U	U	G	C	G	U	U	C	G	G	U	U	C		2.19E-04
	BS2	U	U	U	U	C	A	G	U	A	U	G	U	C	U		7.97E-04
C		aa 5	T	L	T	R	N	K	T	N	R	N					
		aa 35	H	T	P	N	D	D	D	N	H	H					
		Prediction	A/G	U	A	N	U/C	G/C	G	C/U	N	C/U	<i>p</i> -value				
<i>nad1</i> intron 4	BS1	G	U	A	G	U	C	G	G	C	C		1.41E-04				
<i>nad2</i> intron 4	BS1	A	U	C	U	U	C	G	C	C	C		6.82E-04				

Supplementary Figure S5: Predictions of potential RNA target sites for each MISF protein according to the PPR code. The amino acids at positions 5 and 35 were extracted from each MISF PPR repeat and are listed from N- to C-terminus. These combinations were used to identify potential RNA recognition sites of each MISF protein within their intron targets according to the PPR code. The RNA targets within introns were shown in different background colors, and nucleotides that match the predicted RNA recognition code are marked in red. The *P*-values were determined using the FIMO program as previously described (Wang et al., 2017a). (A), (B) and (C) show the results for MISF26, MISF68 and MISF74, respectively.

A

A.thalianaMASALRRLEQQQWRYLVQSQSTRSPKLIHGFFSFSSKTNPNPNKQQQLIDYISKSLQSNDFWETSTHFFSS.....IDLSDSLIETHLIRFK 88
 B.rapaMRHLGKQ.CREFTLISKRSPRRTLGFFSSHSKD.....QQQLLVAISELQNRDNFDTSTHFFSS.....ANLNSLWQDLIRFK 77
 C.papayaMNSIRLPPPISSLITRLLKHTHTQIPNTAT.....TSLSAICDSLRAASNWETLSAIFSS.....IHLTDSLINTLILQK 73
 T.cacaoMALLNRRKLGSLV...TSIHNPFCALPKSIIHSHS.LSQLNNDNSIVNSISDSFKKTHNNGRTTKHFFSS.....VQLTHSLQVQLQK 84
 P.trichocarpaMALLRRTPFSLISTPLKYSIHPVCSWFVARELHD...GKTESDTPVSSICDSLRRGYNNDTNRKFFSS.....LQNLNLLKNNVLEK 85
 E.guttataMATFVGIISRKKEPIPFPLIHLTKPELNGNT.....SSSSITSLIDSLHERRSKTKSEKFFSS.....LKFTHQIQVQLQK 78
 S.bicolorMAVNA.....VARRGGG.....LHGRTPASALFTAATHAT.PQHISHYLAHQPRATWEASAFAPAG.....AVSHGHVDAVLLSIA 71
 Z.maysMAVNAG.....GVVARRGGG.....LLGRRPASALSTATTAT.PQHISHYLAHQPRATWEASAFAPAG.....EMSHGHVDAVLLSIA 75
 O.sativaMALIAKSG...GGVLAQRGGGGNLFGLAAASALSASTSTAATTPQRISHYLAHQPRATWEASAFAPAG.....AADH...VDAVLLSIA 80
 S.italica MTVRSWGLSMALHAANCGRVARRG.....LGRALASALL.CTTATT.PQRISHYLAHQPRATWEASAFAPAG...GGAVPHGQVDAVLLSIA 87
 B.distachyonMAAKSG...GRLVARRGG.....AAVCAAAATALPTSTTAS..PQHIAHYLAHNPRVTEASAFAPAG...EPTAAVAAAAPDRQVDAVLLSIA 80

A.thaliana NPET.....AKCAHSFFWSSHTRNLRHG.....IKSYALTHIVVARRLILARALIESSLN.....SPDSDLVSLDLYEIS. 160
 B.rapa HPET.....AKRATFFWSSHARNLRHG.....VSSYALTHIVVARRLILARALIESSLN.....SSPDSLDDSLDLYEVS. 149
 C.papaya EPTD.....AKRAGFFWSSAKTKNHPHG.....IRSYIAIHLVARRLILARALIESSLKNTVIVH.GSREKFSVSLDLYEIVT. 151
 T.cacao QPEH.....ARSALNFFWSSAKSQNFKHQ.....IYSYIAIHLVHAKQLPBAKILHSAKLTSA...DSTRSCILESLDLYEIVV. 159
 P.trichocarpa EPTD.....AKRAGFFWSSAR.RNFVHG.....VQSYCLMHIIDQARIMDAQALLESLLKRSVG...DPTKFLVLSLSSYKII. 159
 E.guttata EFIN.....SRKANFFWSSRKMNFVHG.....LSTYCLMHIIDQARIKAKALIESSLIKDFD.D.GDSRMLDVLSDLYEIV. 155
 S.bicolor RHPHASP.....PVRKNAFFWSSAAAASSSS.....HSLRSYCLLVHLSRAAFRNASVLESAISRHSSSS...AFASCFILFFAAYEDSG 155
 Z.mays RHPHASP.....PVRKNAFFWSSAAAASSSSPPPPSSSHLSRSYCLLVHLSRAAFRNASVLESAISRHSSSS...AFASCFILFFAAYEDSG 167
 O.sativa KHLHSSSSSYSPELVARNALFFWSSAAAASSSSST.....PHTLRAYCLLVHLSRAAFIRNASVLESAIAKHSSSS...PASAFILFFAAYEDSG 169
 S.italica RHPHASP.....PVRKNAFFWSSAAAAVAVLPS.....SSHLSRSYCLLVHLSRAAFRNASVLESAIAKHSSSS...VPASSFLFFAAYEDSG 176
 B.distachyon KNSSPSSS...ETIAKNAHSFFWSSAAAASSPS.....PHLSRSYCLLVHLSRAAFIRNASVLESAIAKHSSSS...PASSFLFFAAYEDSG 164

A.thaliana .SSTPLVFDLIVQCAKIRYLELGFDFVKRCDGQFTLVSITNTLIIYSSKSKIDDLVIRIMECAIDARIVENEITRIMIQVLCRGRIRKVEVVDLLDR 259
 B.rapa .CSTPLAFDILVQCAKIRLLELGSDFVKNITDRGFSLSVITNTLIIHVSQSNRIDLVIRIMEAIVDARVINEITRIMISALCKEGRIRKVEVLDK 248
 C.papaya .NSIPLVLDLIVQCAKIRMIIEIGFDVRRIDEHGFTLVSISNLLIVVQKSDRNDLVIRIMEHIERIVINEITRIMVSALCKEGRIRKVEVLDK 250
 T.cacao .GSSPLVFDLIVQCAKIRLLELGFDFVKRCDGQFTLVSITNTLIIHVSQSNRIDLVIRIMEAIVDARVINEITRIMISALCKEGRIRKVEVLDK 258
 P.trichocarpa .ISSPLVFDLIVQCAKIRMIIEIGFDVRRIDEHGFTLVSISNLLIVVQKSDKSPKAKIYEHMLHRTVINEATIESMISALCKEGRIRKVEVLDK 258
 E.guttata .ESVPLVFDLIVQCAKIRRMVDDIDAKKIRSHDFPLSVISNTLIIHVMIKSEKRLVSVYEHMISBRMCPNEMTRIMVSALCKEGRIRKVEVLDK 254
 S.bicolor TAAITRGLHLIVHAYARARLPEALEAQRVYIARRGVPLSLSAANAALAAQAQSGFVAVVEVHELMTLQRVYMANQSTVELVIGVLSREGALARTALVER 265
 Z.mays TAAITRGLHLIVHAYARARLPEALEAQRVYIARRGVPLSLSAANAALAAQAQSGFVAVVEVHELMTLQRVYMANQSTVELVIGVLSREGALARTALVER 257
 O.sativa TAAITRGLHLIVHAYARARLPEALEAQRVYIARRGVVPLSLSAANAALAAQAQSGFVAVVEVHELMTLQRVYMANQSTVELVIGVLSREGALARTALVER 269
 S.italica TAAITRGLHLIVHAYARARLPEALEAQRVYIARRGVPLSLSAANAALAAQAQSGFVAVVEVHELMTLQRVYMANQSTVELVIGVLSREGALARTALVER 276
 B.distachyon TAAITRGLHLIVHAYARARLPEALEAQRVYIARRGVPLSLSAANAALAAQAQSGFVAVVEVHELMTLQRVYMANQSTVELVIGVLSREGALARTALVER 264

A.thaliana IICGRQL.FSIVVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 358
 B.rapa IICGRCS.FPVIIVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 347
 C.papaya IICGRCS.FPVIIVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 349
 T.cacao IICGRCS.FPVIIVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 357
 P.trichocarpa IICGRCS.FPVIIVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 357
 E.guttata IICGRCS.FPVIIVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 354
 S.bicolor IICGRCS.FPVIIVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 354
 Z.mays IICGRCS.FPVIIVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 356
 O.sativa IICGRCS.FPVIIVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 364
 S.italica IICGRCS.FPVIIVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 368
 B.distachyon IICGRCS.FPVIIVNTSLVFRVLEEMRIEESMSLKRLLMKNMVVDITIGSIVVYAKAREGDIVSARKVFDMLQRGFSANFSVYIVFVRVCEKEDVKEA 375

A.thaliana ERLLSEMEESVSRFYDETENCIGCFAR.FEWEKGLKEYCEVVTFRGLMESCASFNEVNVKSVSKIENVRANEITTKSIDHGFVEDEHTYSHLIRGFIEG 457
 B.rapa EKLIAMESSVSRFYDETENCIGCFAR.FEWEKGLKEYCEVVTFRGLMESCASFNEVNVKSVSKIENVRANEITTKSIDHGFVEDEHTYSHLIRGFIEG 446
 C.papaya NRIIDEMEIMGLPYDETENHILHGOAR.AGRVEESNLCEMVKQTLBSCSAFNEVLEKVNAIEDLRANGLITILLEAGFIPDEITYSHLIRGHVKG 448
 T.cacao ENVLREMNGLPYDETENHILHGOAR.AGRVEESNLCEMVKQTLBSCSAFNEVLEKVNAIEDLRANGLITILLEAGFIPDEITYSHLIRGHVKG 456
 P.trichocarpa NQIQEMLKPLPYDETENHILHGOAR.AGRVEESNLCEMVKQTLBSCSAFNEVLEKVNAIEDLRANGLITILLEAGFIPDEITYSHLIRGHVKG 456
 E.guttata VGLFBEESLGLPYDETENHILHGOAR.AGRVEESNLCEMVKQTLBSCSAFNEVLEKVNAIEDLRANGLITILLEAGFIPDEITYSHLIRGHVKG 454
 S.bicolor AQLHCEMLSMGAPYDATYNHILHGOAR.CQMKEGLAYFDNHEEGFVLDIGSCNEMLEGLCNAEVEHRRANVITAMMDGLLPPQDITYSHLIRGHVKG 453
 Z.mays AQLHCEMLSMGAPYDATYNHILHGOAR.CQMKEGLAYFDNHEEGFVLDIGSCNEMLEGLCNAEVEHRRANVITAMMDGLLPPQDITYSHLIRGHVKG 465
 O.sativa MQLFBEESLGLPYDATYNHILHGOAR.CQMKEGLAYFDNHEEGFVLDIGSCNEMLEGLCNAEVEHRRANVITAMMDGLLPPQDITYSHLIRGHVKG 467
 S.italica MQLFBEESLGLPYDATYNHILHGOAR.CQMKEGLAYFDNHEEGFVLDIGSCNEMLEGLCNAEVEHRRANVITAMMDGLLPPQDITYSHLIRGHVKG 474
 B.distachyon MQLFBEESLGLPYDATYNHILHGOAR.CQMKEGLAYFDNHEEGFVLDIGSCNEMLEGLCNAEVEHRRANVITAMMDGLLPPQDITYSHLIRGHVKG 462

A.thaliana NDIDQALFLYEMERYKMSPGFEVFRSIVGICCKRVEAGERYLKIMKKRLIEENADIDALAKAFQIGDKTNADRVYNEISVSR..... 544
 B.rapa NDIEQALFLYEMERYKMSPGFEVFRSIVGICCKRVEAGERYFRIMKRLIQENAEINAEALAKAFRTDKNTNADRVYNEISVSR..... 533
 C.papaya GEVEEAFKLYEMERYKMSPGFEVFRSIVGICCKRVEAGERYFRIMKDRFMVSVVDVHAAVGSYCRKGDKKRAGVYVREMMSEGLNPSGLYDLEDSE 548
 T.cacao GNIQOQVFLYEMERYKMSPGFEVFRSIVGICCKRVEAGERYFRIMKDRSIVLSEIDNEALITGHFEKGDGTAGIYNEVARGMRFHWGNFTKEV 556
 P.trichocarpa NQIQEMLKPLPYDETENHILHGOAR.AGRVEESNLCEMVKQTLBSCSAFNEVLEKVNAIEDLRANGLITILLEAGFIPDEITYSHLIRGHVKG 556
 E.guttata DDEVLGLFLYEMERYKMSPGFEVFRSIVGICCKRVEAGERYFRIMKDRSIVLSEIDNEALITGHFEKGDGTAGIYNEVARGMRFHWGNFTKEV 541
 S.bicolor GNAQGVNLYHEMHRGLDPVEVFTLTKGICCGNLKEAEFFIMMKKKTVAITSDLDMLISGCEKRNKRALWLYDMVAMNERLVPASAEFTMML 553
 Z.mays GNAQGVNLYHEMHRGLDPVEVFTLTKGICCGNLKEAEFFIMMKKKTVAITSDLDMLISGCEKRNKRALWLYDMVAMNERLVPASAEFTMML 565
 O.sativa GDAQGIILYHEMHRGLNIGVDVSSILRACCGDGLKEAEFFLAILERKLLATSEINDLISGNCCKGNTKRALWYDMVAMNERLVPASAEFTMML 567
 S.italica GDAEIVRYHEMHRGLNIGVDVSSILRACCGDGLKEAEFFLAILERKLLATSEINDLISGNCCKGNTKRALWYDMVAMNERLVPASAEFTMML 574
 B.distachyon GDSQGIILYHEMHRGLNIGVDVSSILRACCGDGLKEAEFFLAILERKLLATSEINDLISGNCCKGNTKRALWYDMVAMNERLVPASAEFTMML 562

A.thaliana 544
 B.rapa 533
 C.papaya TVNMKALSQAQKIEYR... 564
 T.cacao TSMKDIAKTIPER... 569
 P.trichocarpa 556
 E.guttata 541
 S.bicolor VRRVIKVNVCSPDS. 568
 Z.mays VRRVIKVNNTYSPDS. 580
 O.sativa VRRVIKPKSTCSPNC. 582
 S.italica VRRVIKVNYPNN. 589
 B.distachyon VRRVIRVTTCSPT. 577

B

A.thalianaMMYVSARS....GSARRSSTSLRHSORFHOTENEIVGMPSTVNH...ESEKPEQEKWKLSDKFSVRMIDERFIRIVKIFKVGDAE 81
 B.rapaMAYITRSR....ILVRRSVSTLPHDSORFHOTENEIVGMPSTVNH...EYENPEQEKWKLSDKFSVRMIDERFIRIVKIFKVGDAE 81
 C.papayaMVVRS....FVLRHSTSLRHSORFHOTENEIVGMPSTVNH...DNRVENLPMNR.KLLRDLRVRIDERFIRIVKIFKVGDAE 79
 T.cacaoMRTMPS....SALRRSSTSLRHSORFHOTENEIVGMPSTVNH...KKAQDLAVKR.KKP...SVRTIDERFIRIVKIFKVGDAE 77
 P.trichocarpaMVVLRH....IASRRSSTSLRHSORFHOTENEIVGMPSTVNH...KKAQDLAVKR.KKP...SVRTIDERFIRIVKIFKVGDAE 82
 E.guttataMRVEGKMLAKS....IASSRSTSLRHSORFHOTENEIVGMPSTVNH...KKAQDLAVKR.KKP...SVRTIDERFIRIVKIFKVGDAE 84
 S.bicolorMARRG...VQLLVARGSSSPGSRRIHOTENEIVGMPSTVNH...QSEEAIAITVP...HYTSVRVIDERFIRIVKIFKVGDAE 78
 Z.maysMMTRCG...VQLLVARGSSSPGSRRIHOTENEIVGMPSTVNH...QSEEAIAITVP...HYTSVRVIDERFIRIVKIFKVGDAE 79
 O.sativaMASRAP...CLLAVRGSASSPGLARRHOTENEIVGMPSTVNH...RNEDAVAALSP...HYTSVRVIDERFIRIVKIFKVGDAE 78
 S.italicaMARRG...AQLLARRSSTSLRHSORFHOTENEIVGMPSTVNH...QSEEAIAITVP...HYTSVRVIDERFIRIVKIFKVGDAE 78
 B.distachyon MVTACAVTAPATVPVPLPSPPHRSRRRPEMARLLAVRGSASSPGLARRHOTENEIVGMPSTVNH...RSEEAIVAPIER...HYTSVRVIDERFIRIVKIFKVGDAE 105

A.thalianaALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 191
 B.rapaALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 191
 C.papayaALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 191
 T.cacaoALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 191
 P.trichocarpaALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 191
 E.guttataALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 191
 S.bicolorALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 191
 Z.maysALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 191
 O.sativaALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 191
 S.italicaALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 191
 B.distachyonALEVLRVVDLRLRVEVLDLDELINVRILFFWAKGRFNEFHCHCSTYTLRCLCEARLYGCVYRTOEIVNTYVSVSFAVLSLVRALGPAKMYKSAISVFGARGR 215

A.thalianaCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 300
 B.rapaCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 300
 C.papayaCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 300
 T.cacaoCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 300
 P.trichocarpaCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 300
 E.guttataCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 300
 S.bicolorCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 300
 Z.maysCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 300
 O.sativaCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 300
 S.italicaCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 300
 B.distachyonCKPSTSTYNSVIMLHQGCHHEKVHELYNEMCNEG.HCPDITVYSALISYKLRNDSAIRFEEMDNCMCPKRIYTHLGLVYFKVGGVERALDFFEMKRAKGS 324

A.thalianaETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 410
 B.rapaETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 410
 C.papayaETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 410
 T.cacaoETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 410
 P.trichocarpaETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 410
 E.guttataETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 410
 S.bicolorETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 410
 Z.maysETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 410
 O.sativaETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 410
 S.italicaETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 410
 B.distachyonETVVTYTELRIGCGGRVDEAYGEYKMLRDLPLDVLVFNLMNIIIGVGRVEELTNVSEEMGMWRCHTIVSVSYVIRKAFESKAVHSEVSWEDRKAADSVSPSEF 434

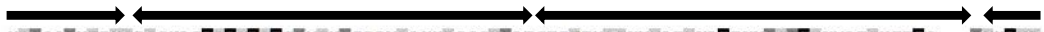
A.thalianaTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 520
 B.rapaTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 520
 C.papayaTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 520
 T.cacaoTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 520
 P.trichocarpaTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 520
 E.guttataTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 520
 S.bicolorTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 520
 Z.maysTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 520
 O.sativaTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 520
 S.italicaTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 520
 B.distachyonTYSIIDGCKNRVEKALLLEEMDEHGFPFPAAAYCSLINDLGRAKRYEAARELEKELKENFNVSSRYAVMIKIHGRCGLSEAVDLEEMKKGSGGPPVYVYNAL 544

A.thalianaMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 629
 B.rapaMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 629
 C.papayaMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 629
 T.cacaoMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 629
 P.trichocarpaMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 629
 E.guttataMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 629
 S.bicolorMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 629
 Z.maysMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 629
 O.sativaMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 629
 S.italicaMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 629
 B.distachyonMSGVVRAGINAEASLIRKVEENCGIARVNSHNINLNGARAGCAARFAEMFETIHSGIKRFDVYNTVLCFAHGMEEBAAKMMEMKDKGCFYDIIPYSSILRAI 653

A.thalianaKNVHEKDDVSS 641
 B.rapaKNVHVKDVL.. 640
 C.papayaKNVDEHNPVI.. 637
 T.cacaoKNVIEDRSLSLF 636
 P.trichocarpaKNVIEDDEPNFP 641
 E.guttataKNVIEDCKLVSS 645
 S.bicolorKNVIEHYTGQGC 645
 Z.maysKNVIEHCDP... 638
 O.sativaKNVIEOE..... 632
 S.italicaKNVIEHE..... 632
 B.distachyonKNVIEHE..... 659

C

A.thaliana	MIRRPIDFAAV.....FRHLTSPSTSSRFLLFYSSEHEARKPIVSN.....PKSPICSTFTVCKLIASQSDPLLAREIFD	72
B.rapa	MIRRIQNSSAV.....SRHLTSPISLFSRFLFSSSENNTHPEIVSNRN.....PKSPFCSFTVCKLIASQSDPLLAREIFD	74
C.papaya	MQRLLNSSKAMITLT...SSHFTFLPKSLIFFYSYSPESPQKNQQQLSSASASAN...SAIGSFTVCKLIASQSDPLLAREIFD	83
P.trichocarpa	MHKPFLVCKILLTTPPRTTRVPLPKQSLFFYSSSHIIHQHRELEPDSHPNANT...KSPICSTFTVCKLIASQSDPLLAREIFD	88
T.cacao	MQQALFSSSKTV.....SRSLTFPPKPPFLFRSSSSSPGPPHKKQ.....PPRTC.TSAIGSPARVFKLIASQSDPLLAREIFD	74
E.guttata	MYHASLS.....YSLLSKVRKPSFYCLTTLSSHQKQEPNDRFRNEKQEPKVDKQPSIGSPARICKLIASQSDPLLAREIFD	78
S.bicolorMLSPPK.....ISPARLHKIVTSCDPLLARELVT	31
Z.maysMLSPPPA.....TSPARLHKIVTSCDPLLARELVN	31
S.italicaMLSPPPA.....ISPARLHKIVTSCDPLLARELVT	31
B.distachyonMLSPPPA.....ISPARLHKIVTSCDPLLARELVT	31
O.sativaMLSPRAPP.....ISPARLHKIVTSCDPLLARELVT	34



A.thaliana	YASQCENFRHRSSSHLLILKIKGRCYFNIDVDLAKHRSSGYLTGEIFTYLIKVYAEAKLPEKVLSTFYKMLEFNFTFQ...PKHLNR	159
B.rapa	YASQHSERHRSQSSHLVLLKIGRSRHFNIIDVDLAKHRSSGYVVTGELFTYLIKVYAEAELEPKAKTFYKMLEFNFTFQ...PKHLNR	161
C.papaya	YAFRQOQGERHSYSSYLILKIKGRCYFNIIDVDLILSLKTKQRYITPILFSYVVKIYAEADLPKALKTFFYKIRFEGCFL...PKHLNR	170
P.trichocarpa	YASRQNEQHSYSSYLILKIKGRCYFNFIDVDLLDLKSKNYVVTQTFESYIINIYKGNLPEALKIFTYTILKFDKCNFS...PKHLNR	175
T.cacao	YASNQLGERHSYSSFLVLLKIGRSRKHFSIVDDLLIRLKTDRYVPTLFESYLIKVYAEANLPERALKTFFYKMLEFNIRFL...PKHLNR	161
E.guttata	LASRQNERHSYATCTHLLKIGRSRHFSSIMQALLSSLKSKDYSSISPLETHIIQIYGDAKMPKALKTFFYTILEFNIRFL...TKQDNC	165
S.bicolor	VTS...FTTTPHPATLHSLRLRPAR.RRDHHPHALALLRRLPSPSPRLLPLLLAVLRLRRFPQLFLSTFNSLFFVSGSPFLPLHPELRLR	118
Z.mays	VTS...FTTTPHPATLHSLRLRPAR.RRDHHPHALALLRRLPSPSPRLLPLLLAAVLRRLRRFPQLFLSTFNSLFFVSGSPFLPLHPELRLR	118
S.italica	VTS...FTTTPHPATLHSLRLRPAR.RRDHHPHALALLRRLPSPSPRLLPLLLAVLRLRRFPQLFLSTFNSLFFVSGSPFLPLHPELRLR	118
B.distachyon	VTS...FTTTPHPATLHSLRLRPAR.RRDHHPHALALLRRLPSPSPRLLPLLLSVLRLRRFPQLFLSTFNSLFFVSGSPFLPLHPELRLR	118
O.sativa	VTA...FTTAPHSTLHSLRLRPAR.RRDHHPHALALLRRLPSPSPRLLPLLLSALRLRRFPQLFLSTFNSLFFVSGSPFLPLHPELRLR	121



A.thaliana	IIDVIVSHRGYLQKAFELFKSSRLHGVN...TRYNYIMRAFCLNDDLSIAYLQKGLMERVVVVDVSKYIICGHCRSGOVNGAME	246
B.rapa	IIEVIVSHRGYLQKALELFKSARLHGVN...TRYNYIMRAFCLNDDLSIAYLQKGLMERVVVVDVSKYIICGHCRSGOVNGAME	248
C.papaya	IIEVIVSHRSYLRPFDFERSAHRHGVLEN...IQSYNYIMRAFCLNGDISIAYSLNKMFKRVLVPOVESYRILMGCICRSGOVNGAVD	257
P.trichocarpa	IIEVIVSHHNYIKPFAFDELKDAHTYDVFN...TKSYNYIMRAFCLNGQISMAYSLNQMFKRVVMPVESYRILMGCICRSGOVNGAVD	262
T.cacao	IIEVIVSHRNLMFAFDELKNAHKHGVLEN...TKSYNYIMRAFCLNGDLSVAYKLNKMFERVVVVDVSKYIIMGCICRSGOVNRAVD	248
E.guttata	IIEVIVANNFLRRAFDLFRAAHKHGVSAN...TKSYNYIMRAFCLSGDLSIAYTLNQMFKRVVLPOVESYRILMGCICRSGOVNRAVD	252
S.bicolor	IIGVIVSSPSPHFPSALHLLRRLVSSRLPPLAPLVLASHNIIIDAAARSGHVAVLSLHRLRSLVSPADYVYRITQCSICRGAOVRTAAT	208
Z.mays	IIGVIVSSPSPHFPSALHLLRRLVSSRLPPLAPLVLASHNIIIDAAARSGHVAVLSLHRLRSLVSPADYVYRITQCSICRGAOVRTAAT	208
S.italica	IIGVIVSSPSPHFPSALHLLRRLVSSRLPPLAPLVLASHNIIIDAAARSGHVAVLSLHRLRSLVSPADYVYRITQCSICRGAOVRTAAT	208
B.distachyon	IIGVIVSSPSPHFPSALHLLRRLVSSRLPPLAPLVLASHNIIIDAAARSGHVAVLSLHRLRSLVSPADYVYRITQCSICRGAOVRTAAT	208
O.sativa	IISVIVSSPSPHFPSALHLLRDLVSTRPLPEPLVLASHNIIIDAAARSGHVAVLSLHRLRSLVSPADYVYRITQCSICRGOVHTAAT	211



A.thaliana	LLDDMLNKGFEVDRTSYTLLNSLCKRKTLEAYKLLCRKVKVGCNPDVHYNTVITGECREGRAMDARKVLLDDMLNSGCSNLSVSYRTL	336
B.rapa	LLEDMLNKGFEVDRTSYTLLNSLCKRKTLEAYKLLCRKVKVGCNPDVHYNTVITGECREGRAMDARKVLLDDMLNSGCSNLSVSYRTL	338
C.papaya	LLEDMLNKGFEVDRTSYTLLNSLCKRKTLEAYKLLCRKVKVGCNPDVHYNTVITGECREGRAMDACKVLEDMPNSGCLNLSVSYRTL	347
P.trichocarpa	LLEDMLNKGFEVDRTSYTLLNSLCKRKTLEAYKLLCRKVKVGCNPDVHYNTVITGECREGRAMDACKVLEDMPNSGCMNLSVSYRTL	352
T.cacao	LLEDMLNKGFEVDRTSYTLLNSLCKRKTLEAYKLLCRKVKVGCNPDVHYNTVITGECREGRAMDACKVLEDMPNSGCLNLSVSYRTL	338
E.guttata	LLEDMLNKGFEVDRTSYTLLNSLCKRKTLEAYKLLCRKVKVGCNPDVHYNTVITGECREGRAMDACKVLEDMPNSGCLNLSVSYRTL	342
S.bicolor	LLEDMLHRGIPADPRLAYTVLNLALCRKQCLEAYRLLCINRGRGVSVDVHYNTVITGECREGRAMDACKVLEDMPNSGCLNLSVSYRTL	298
Z.mays	LLEDMLHRGIPADPRLAYTVLNLALCRKQCLEAYRLLCINRGRGVSVDVHYNTVITGECREGRAMDACKVLEDMPNSGCLNLSVSYRTL	298
S.italica	LLEDMLHRGIPADPRLAYTVLNLALCRKQCLEAYRLLCINRGRGVSVDVHYNTVITGECREGRAMDACKVLEDMPNSGCLNLSVSYRTL	298
B.distachyon	LLEDMLHRGIPADPRLAYTVLNLALCRKQCLEAYRLLCINRGRGVSVDVHYNTVITGECREGRAMDACKVLEDMPNSGCLNLSVSYRTL	298
O.sativa	LLEDMLHRGIPADPRLAYTVLNLALCRKQCLEAYRLLCINRGRGVSVDVHYNTVITGECREGRAMDACKVLEDMPRESGCAENAVTYTAV	301



A.thaliana	IGGLDQGMFDEGKRYLEEMISKGFSPHESVSNCLVKGCGSFGKVEBACDVVEVMKNGETIHSDDTMEWVPLIINEDESEKIKLLEDA	426
B.rapa	IGGLDQGMFDEGKRYLQEMISKGFSPHESVSNCLVKGCGSFGKVEBACDVVELVMKNGDALHSDDTMEWVPLIIPVCKEDVSEKIQRELEDA	428
C.papaya	VGGLDQDRIEGEAKRYLEEMLSKGFSPHESIALVGEVGNVGIKIEACVLVEKMLNHGEAHHMDTWTIIPMIGVEDDRVCMEEILDKV	437
P.trichocarpa	VGGLDQGMFDEAKSHLEEMMKGFSPHESAVSNALIKGFENVGKIEACGVVEELLKHGEAHHMDTWTIIPMIGVEDDLQRIEGELEDA	442
T.cacao	IGGLDQGMFDEAKKMEBMLKGFSPHESVSHTLVKGCGSFGKVEBACDVVEVMKNGEIVGFGEMLYKGEVHHMDTWLIIIPRIEEDYETERMGEILEEV	428
E.guttata	VGGLDQGMFDEAKTYVVKVMISKGFSPHESIVHILVTGCRIGKIEBACEVLCCELLGHGNSPHDTWAEIHLTVGEEYNTAMGDVVKQV	432
S.bicolor	VNGLGVSLYDKAEAYLVDMVGNLVPHEVSVFHSVIRKCCVTVGRVBAQAQMSWMLDLGVVPHVESWSSVIRCVCKDED...CIEABLLQI	386
Z.mays	VNGLGVSLYDKAEAYLVDMVGNLVPHEVSVFHSVIRKCCVTVGRVBAQAQMSWMLDLGVVPHVESWSSVIRCVCKDED...CIEABLLQI	386
S.italica	VNGLGVSLYDKAEAYLVDMVGNLVPHEVSVFHSVIRKCCVTVGRVBAQAQMSWMLDLGVVPHVESWSSVIRCVCKDED...CIEABLLQI	386
B.distachyon	VNGLGISLFDKAEAYLVDMVGGGIIIPHEVSVFHSVIRKCCVTVGRVBAQAQMSWMLDLGIPPEAETWSSVIRCVCKDED...YIEVILLHM	386
O.sativa	VNGLGVNLYDKAEAYLDMGLKGLIIPHEVSVLHSVIRKCCVAVGRVBAQAQMMTRMLDLGMVPEAETWSSVIRSVSDEE...NVEVLLQV	389



A.thaliana	VKEEITGDTRIVDVG.IGLGSYLSSKLMKRNARERRRL.....	466
B.rapa	MKVEISGDTRIVDAGIIGLGSYLSSKRMKRNARERRRL.....	468
C.papaya	IKIEITGNTRIVDAG.IGLEGLYLVKKIQSKLWRA.....	470
P.trichocarpa	KKVELKGDTRIVEAG.IGLEEYLKIKRTQOKAWRY.....	475
T.cacao	MKVEIKRTRIVDAG.TGLEDYLIKIRISRSKRP.....	461
E.guttata	LKVEIKPNTRIVDVG.AALEEYLTKKIKTRDQGVVYSCRGDGVSLSHASCHSAC	485
S.bicolor	VG.R.....	389
Z.mays	VG.RQRGSNMISDRVP.....	401
S.italica	VT.GRRHGSSTSSTLK.....	401
B.distachyon	MQEQKRCNSNMISKSTW.....	402
O.sativa	MKGIKHRSNINRSRTWT.....	406

Supplementary Figure S6: Multiple sequence alignments of MISF protein homologs from a selection of dicot and monocot plant species. Amino acid sequence alignments of three MISF proteins from a representative selection of dicot (*Arabidopsis thaliana*, *Brassica rapa*, *Carica papaya*, *Theobroma cacao*, *Populus trichocarpa*, *Erythranthe guttata*) and monocot (*Oryza sativa*, *Zea mays*, *Sorghum bicolor*, *Setaria italic*, *Brachypodium distachyon*) plant species. Identical and non-identical amino acids were shaded in black and grey respectively. PPR motifs are indicated by double arrow lines. Grey bars represent protein regions corresponding to mitochondrial targeting sequences, according to TargetP predictions. (A), (B) and (C) show the corresponding results from MISF26, MISF68 and MISF74 alignments.

A.th_nad2_3'-int2	0
B.rapa_nad2_3'-int2	0
C.papaya_nad2_3'-int2	0
E.quattara_nad2_3'-int2	160
Z.mays_nad2_3'-int2	0
O.sativa_nad2_3'-int2	0
A.th_nad2_3'-int2	63
B.rapa_nad2_3'-int2	65
C.papaya_nad2_3'-int2	63
E.quattara_nad2_3'-int2	320
Z.mays_nad2_3'-int2	63
O.sativa_nad2_3'-int2	0
A.th_nad2_3'-int2	234
B.rapa_nad2_3'-int2	215
C.papaya_nad2_3'-int2	230
E.quattara_nad2_3'-int2	473
Z.mays_nad2_3'-int2	237
O.sativa_nad2_3'-int2	123
A.th_nad2_3'-int2	341
B.rapa_nad2_3'-int2	344
C.papaya_nad2_3'-int2	341
E.quattara_nad2_3'-int2	598
Z.mays_nad2_3'-int2	346
O.sativa_nad2_3'-int2	204
A.th_nad2_3'-int2	406
B.rapa_nad2_3'-int2	514
C.papaya_nad2_3'-int2	515
E.quattara_nad2_3'-int2	753
Z.mays_nad2_3'-int2	430
O.sativa_nad2_3'-int2	494
A.th_nad2_3'-int2	643
B.rapa_nad2_3'-int2	674
C.papaya_nad2_3'-int2	678
E.quattara_nad2_3'-int2	709
Z.mays_nad2_3'-int2	645
O.sativa_nad2_3'-int2	595
A.th_nad2_3'-int2	804
B.rapa_nad2_3'-int2	827
C.papaya_nad2_3'-int2	816
E.quattara_nad2_3'-int2	1062
Z.mays_nad2_3'-int2	816
O.sativa_nad2_3'-int2	736
A.th_nad2_3'-int2	942
B.rapa_nad2_3'-int2	964
C.papaya_nad2_3'-int2	972
E.quattara_nad2_3'-int2	1213
Z.mays_nad2_3'-int2	945
O.sativa_nad2_3'-int2	896
A.th_nad2_3'-int2	1111
B.rapa_nad2_3'-int2	1124
C.papaya_nad2_3'-int2	1075
E.quattara_nad2_3'-int2	1392
Z.mays_nad2_3'-int2	1195
O.sativa_nad2_3'-int2	1043
A.th_nad2_3'-int2	1256
B.rapa_nad2_3'-int2	1252
C.papaya_nad2_3'-int2	1251
E.quattara_nad2_3'-int2	1516
Z.mays_nad2_3'-int2	1251
O.sativa_nad2_3'-int2	1156
A.th_nad2_3'-int2	1409
B.rapa_nad2_3'-int2	1401
C.papaya_nad2_3'-int2	1437
E.quattara_nad2_3'-int2	1647
Z.mays_nad2_3'-int2	1401
O.sativa_nad2_3'-int2	1352
A.th_nad2_3'-int2	1559
B.rapa_nad2_3'-int2	1569
C.papaya_nad2_3'-int2	1572
E.quattara_nad2_3'-int2	1546
Z.mays_nad2_3'-int2	1498
O.sativa_nad2_3'-int2	1498
A.th_nad2_3'-int2	1720
B.rapa_nad2_3'-int2	1712
C.papaya_nad2_3'-int2	1747
E.quattara_nad2_3'-int2	1758
Z.mays_nad2_3'-int2	1689
O.sativa_nad2_3'-int2	1647
A.th_nad2_3'-int2	1866
B.rapa_nad2_3'-int2	1863
C.papaya_nad2_3'-int2	1838
E.quattara_nad2_3'-int2	1898
Z.mays_nad2_3'-int2	1839
O.sativa_nad2_3'-int2	1772
A.th_nad2_3'-int2	2022
B.rapa_nad2_3'-int2	2019
C.papaya_nad2_3'-int2	2045
E.quattara_nad2_3'-int2	2053
Z.mays_nad2_3'-int2	1994
O.sativa_nad2_3'-int2	1524
A.th_nad2_3'-int2	2171
B.rapa_nad2_3'-int2	2167
C.papaya_nad2_3'-int2	2200
E.quattara_nad2_3'-int2	2104
Z.mays_nad2_3'-int2	2153
O.sativa_nad2_3'-int2	2075
A.th_nad2_3'-int2	2250
B.rapa_nad2_3'-int2	2302
C.papaya_nad2_3'-int2	2315
E.quattara_nad2_3'-int2	2317
Z.mays_nad2_3'-int2	2310
O.sativa_nad2_3'-int2	2253
A.th_nad2_3'-int2	2382
B.rapa_nad2_3'-int2	2392
C.papaya_nad2_3'-int2	2392
E.quattara_nad2_3'-int2	2392
Z.mays_nad2_3'-int2	2392
O.sativa_nad2_3'-int2	2392
A.th_nad2_3'-int2	2499
B.rapa_nad2_3'-int2	2499
C.papaya_nad2_3'-int2	2499
E.quattara_nad2_3'-int2	2499
Z.mays_nad2_3'-int2	2499
O.sativa_nad2_3'-int2	2499

B55

B56

B54

Supplementary Figure S7: Multiple sequence alignments analyzing the conservation of the different predicted RNA targets of each MISF protein. Multiple sequence alignments of the different MISF protein intron targets from a representative selection of dicot (*Arabidopsis thaliana*, *Brassica rapa* and *Carina papaya*) and monocot (*Oryza sativa*, *Zea mays*) species. The sequences corresponding to the MISF predicted binding sites are shown in black boxes. (A): Alignment of *nad2* intron 3 sequences. BS1 to BS3 represent the three putative RNA binding sites of MISF26 in *nad2* intron 3. (B) and (C): Sequence alignment of 5'- and 3'-portions of *nad2* intron 2 (trans-intron). The six potential RNA binding sites of MISF68 are shown as BS1 to BS6. (D): Sequence alignment of *nad4* intron 1. The two potential RNA binding sites of MISF68 are shown as BS1 to BS2. (E): Alignment of *nad1* intron 4 sequences. Since *nad1* intron 4 is fragmented in certain plant species, the sequences used for the alignment correspond only to the 300 nucleotides downstream of the *matR* open reading frame. The only putative RNA binding site of MISF74 is shown as BS1. (F): Alignment of *nad2* intron 4 sequences. BS1 represents the single putative RNA binding site found for MISF74 in *nad2* intron 4.