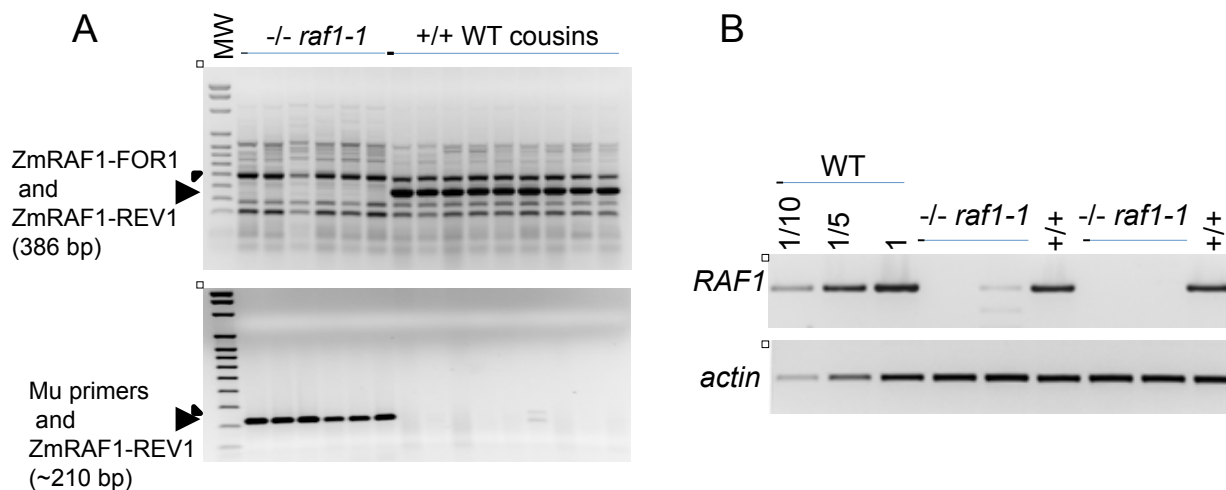


**Supplemental Figure 1.** Accumulation of the *BSD2* transcript in *raf1*.

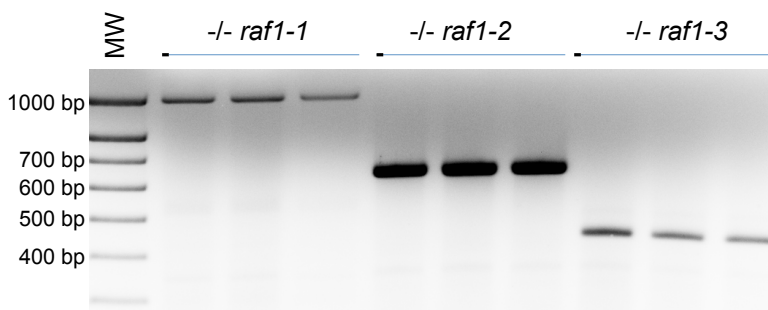
RT-PCR was performed using primers ZmBSD2-SQF and ZmBSD2-SQR (Supplemental Table 3). Actin was used as the control gene amplified using primers ZmActin1-SQF and ZmActin1-SQR. Samples with genotype +/- are siblings of the parents of *raf1-1* homozygous (-/-) mutants. *bsd2* mutant was used as a negative control.



**Supplemental Figure 2.** Validation of *RAF1* gene identification.

**(A)** Gene-specific PCR validating *RAF1* identification. The higher and lower molecular weight products appear to result from mis-priming by the ZMRAF1-FOR1 primer.

**(B)** Accumulation of *RAF1* transcripts in *raf1-1*. RT-PCR was performed using ZmRAF1-FOR1 and ZmRAF1-REV1 primers. Actin was used as the control gene.



**Supplemental Figure 3.** Gene-specific PCR validating the locations of the three *RAF1* insertions.

The primers *Mu* and ZmRAF1-REV2 amplified the region between each insertion and 23 bp upstream of the stop codon.

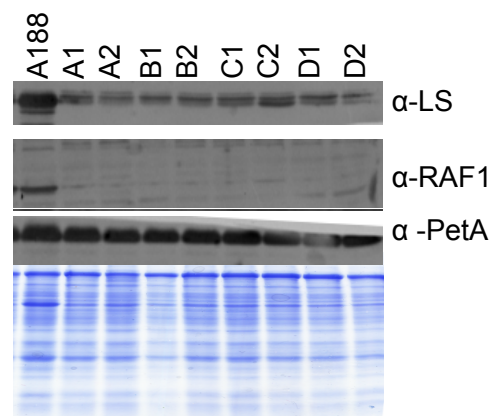
**A**

Ear	female parent(+)	male parent(+)	segregating green F <sub>1</sub>	segregating pale green F <sub>1</sub>	<i>P</i> value
1	<i>raf1-1-1</i>	<i>raf1-2-1</i>	34	13	0.67
2	<i>raf1-1-2</i>	<i>raf1-2-1</i>	38	10	0.5
3	<i>raf1-1-2</i>	<i>raf1-2-3</i>	33	10	0.79
4	<i>raf1-1-3</i>	<i>raf1-3-6</i>	32	14	0.39
5	<i>raf1-2-1</i>	<i>raf1-1-1</i>	22	5	0.44
6	<i>raf1-2-2</i>	<i>raf1-1-3</i>	32	12	0.73
7	<i>raf1-2-4</i>	<i>raf1-1-3</i>	22	2	0.06
8	<i>raf1-2-5</i>	<i>raf1-1-1</i>	30	9	0.78
9	<i>raf1-2-6</i>	<i>raf1-1-2</i>	32	11	0.93
10	<i>raf1-3-1</i>	<i>raf1-1-2</i>	26	2	0.03
11	<i>raf1-3-2</i>	<i>raf1-1-1</i>	33	10	0.79
12	<i>raf1-3-3</i>	<i>raf1-1-3</i>	36	7	0.19
13	<i>raf1-3-5</i>	<i>raf1-1-1</i>	19	8	0.58
14	<i>raf1-3-2</i>	<i>raf1-2-4</i>	10	2	0.5

**B**

Ear ID	cross		<i>raf1</i> alleles
	female	male	
A	LF4.7	LF3.4	1×2
B	4DT	31-6-10	2×1
C	108-2-3	31-6-10	3×1
D	31-6-10	7LS	1×3

**C**



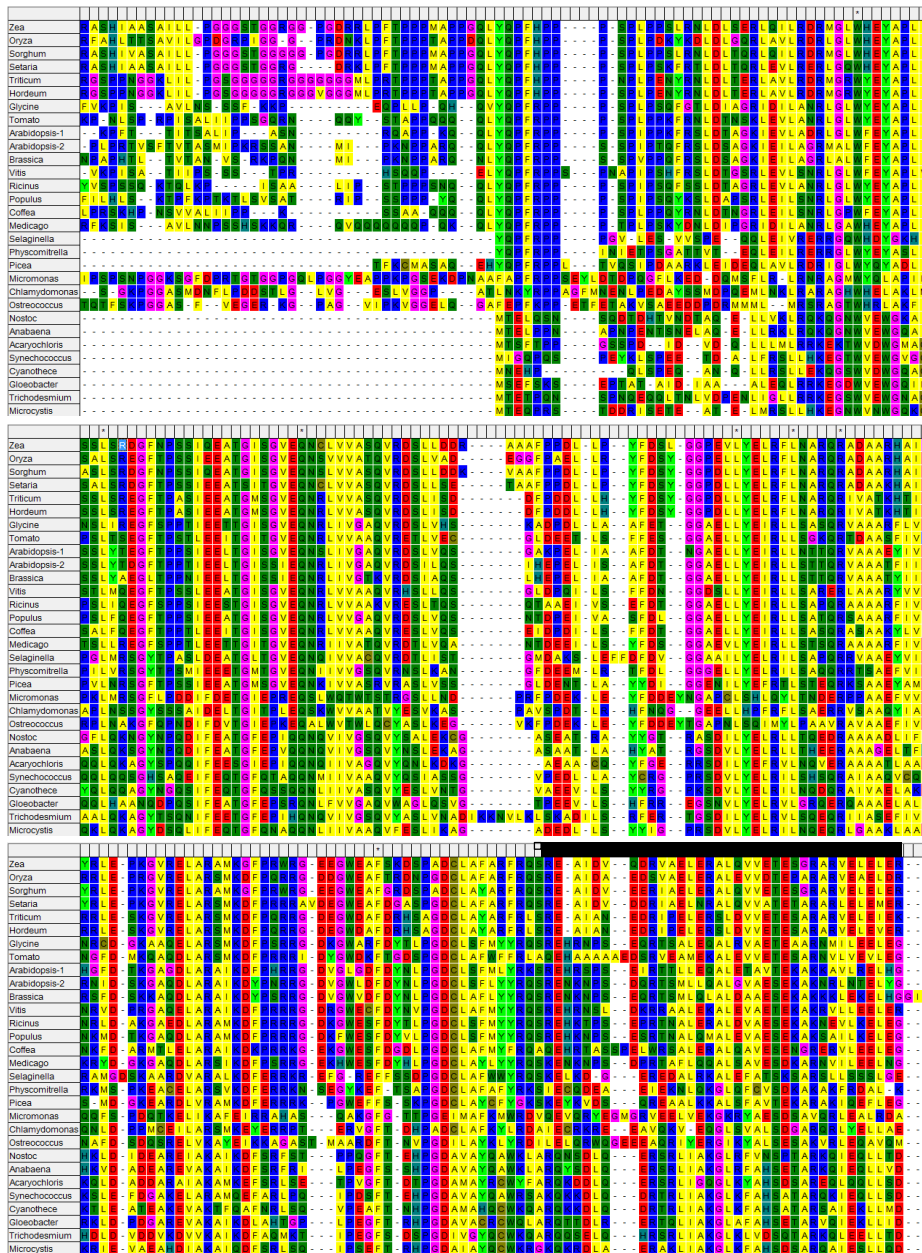
**Supplemental Figure 4.** Complementation crosses confirm that *raf1* encodes the regulator of Rubisco accumulation.

**(A)** Genetic analysis of crosses between heterozygotes carrying the indicated alleles. The *P*-value was derived from a Chi-squared test, based on the phenotypic segregation patterns of F<sub>1</sub> progeny.

**(B)** Crosses used for the immunoblot analysis in panel **(C)**.

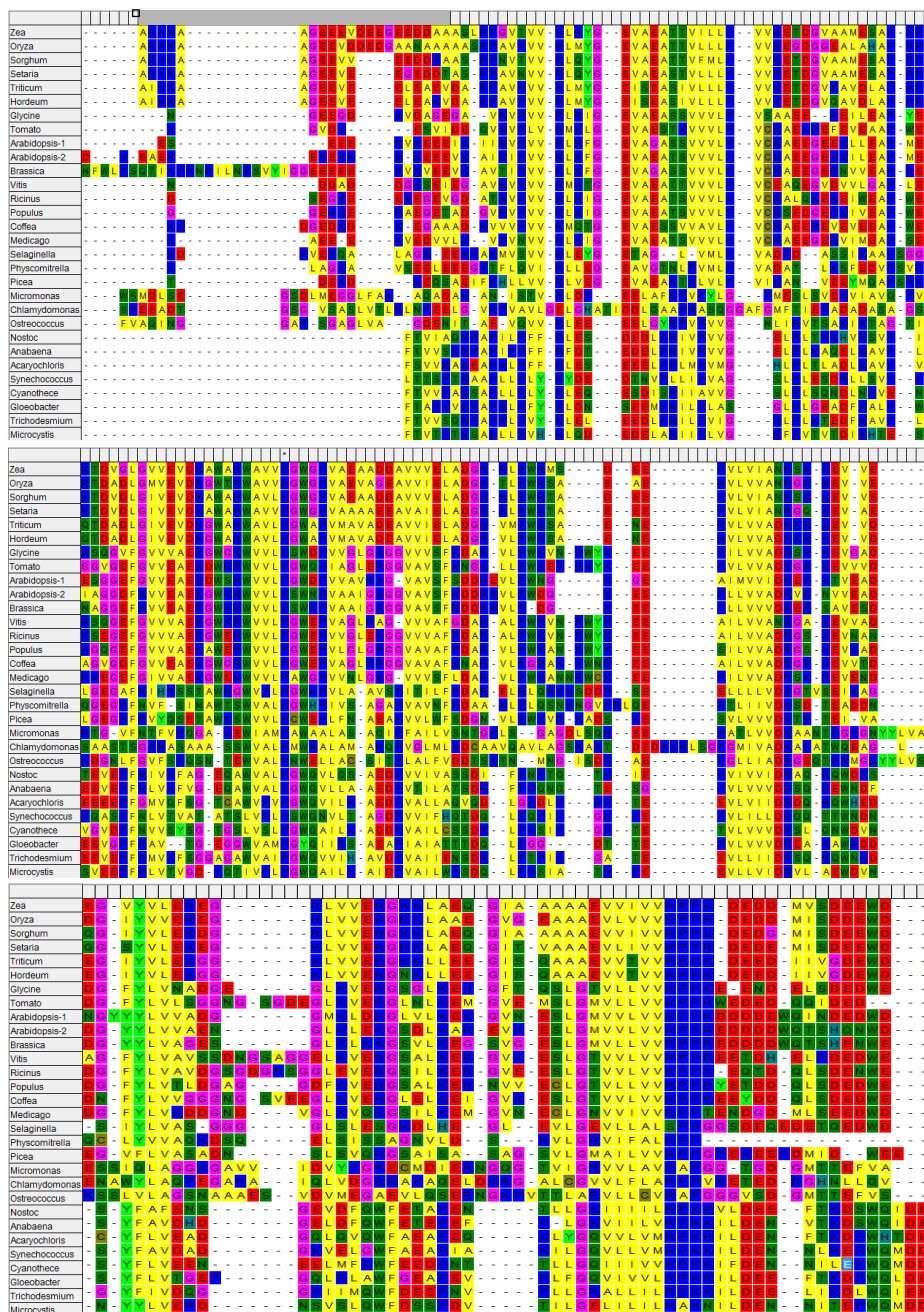
**(C)** Immunoblot analysis of representative complementation cross progeny as given in panel **(B)**.

A



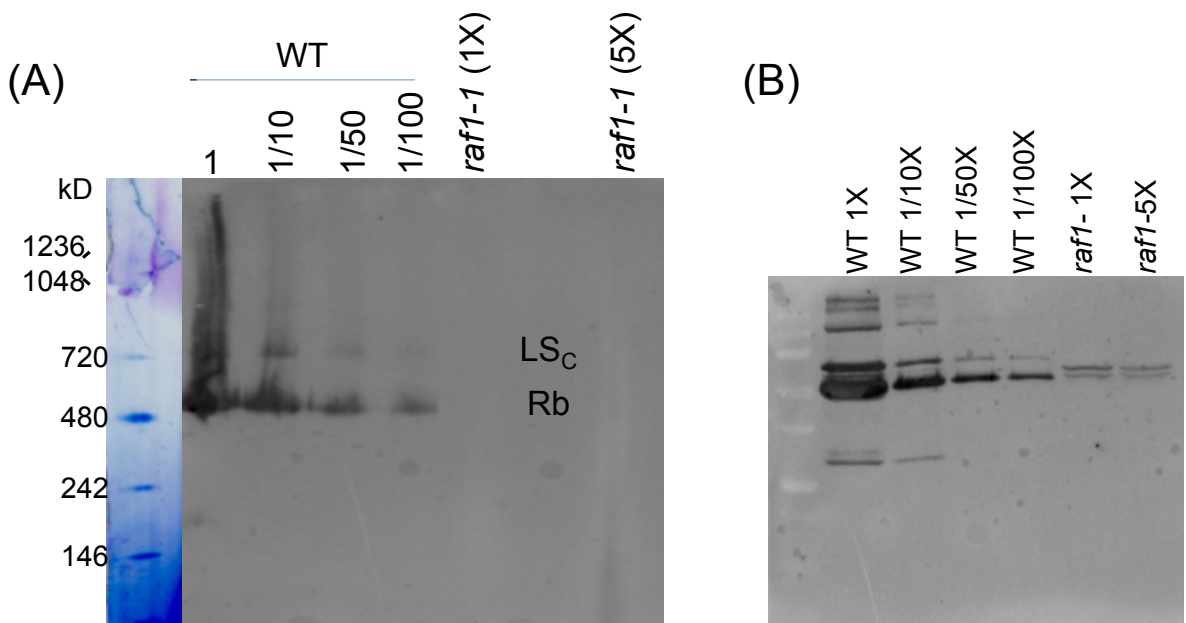
Supplemental Figure 5. Alignment of RAF1 homologues.

Page 1 of 2; see legend on page 2.



**Supplemental Figure 5.** Alignment of RAF1 homologues.

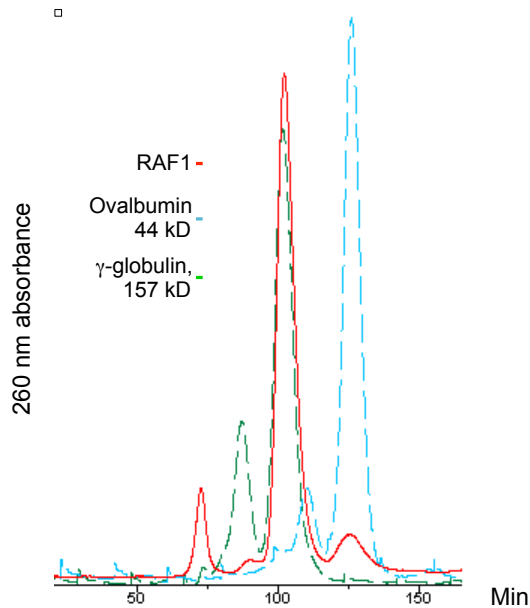
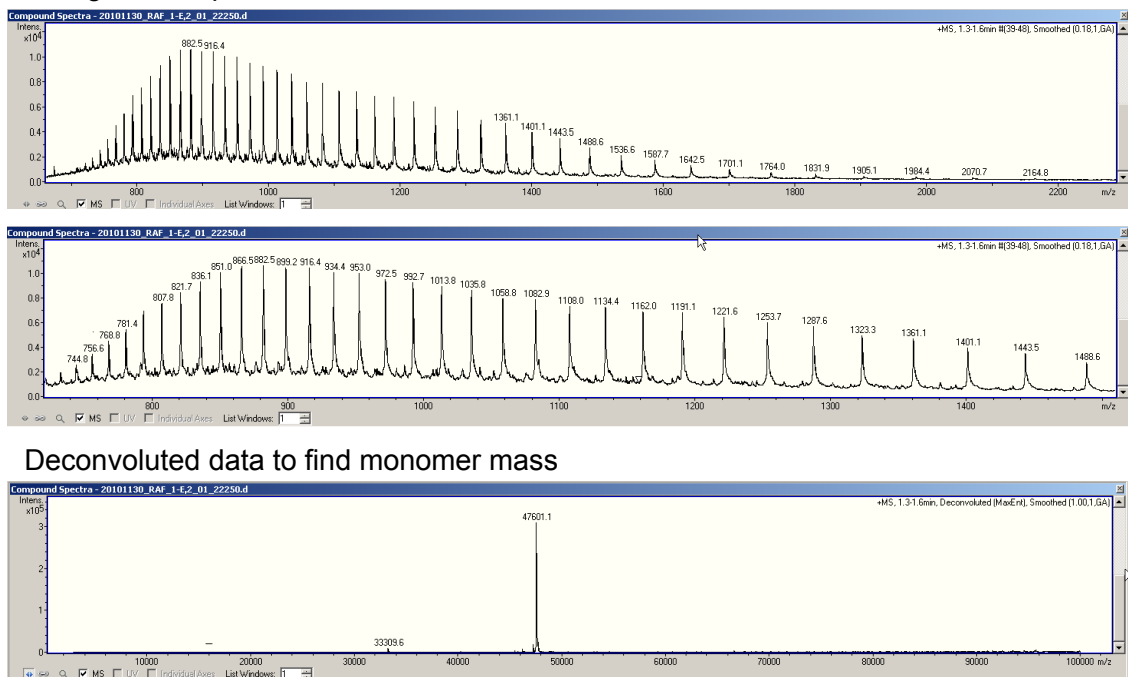
Alignments were performed using the ClustalW option of the MEGA4 program, after deletion of the predicted chloroplast transit peptide from the plant proteins. Amino acids are color coded as hydrophobic, yellow; polar, green; acidic, red; basic, blue; tyrosine, light green; glycine, pink; proline, blue; cysteine, brown; and histidine, light blue. The first 35 amino acids of a 60 amino acid domain identified by MGD as having homology to Docking domain A of the erythromycin polyketide synthase are highlighted with a black bar, whereas the last 21 amino acids, highlighted in gray, are conserved only in monocots and are absent from cyanobacterial proteins.



**Supplemental Figure 6.** Analysis of LS from total soluble protein.

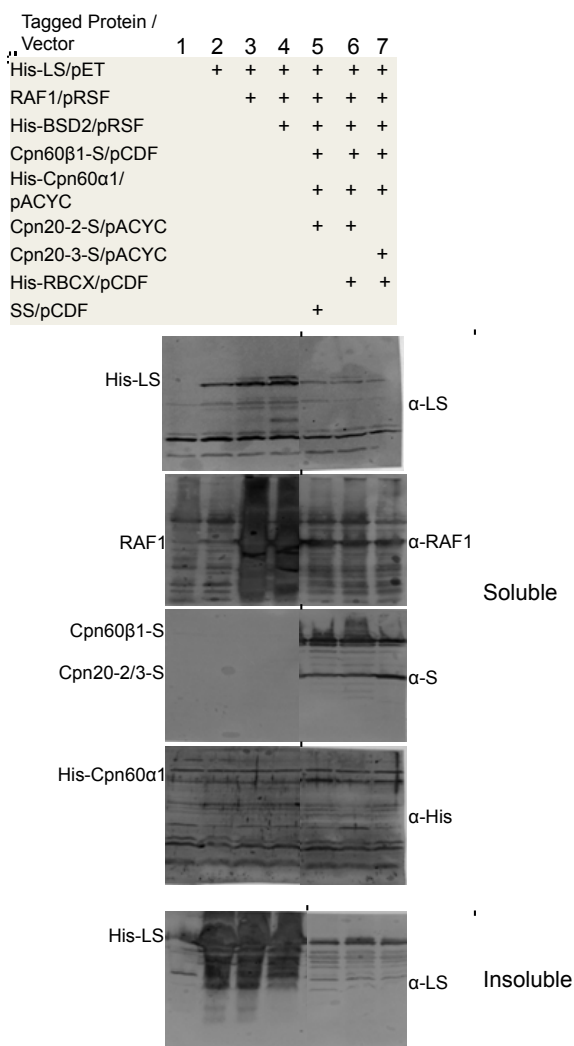
(A) Native gel analysis. Total soluble leaf proteins of 10-day-old WT and *raf1-1* seedlings were extracted from equal surface areas and separated in a 3-12% native gel, which was analyzed by immunoblotting for LS.  $LS_C$  marks the migration of the putative LS-chaperone complex, and Rb the position of Rubisco holoenzyme.

(B) Denaturing gel analysis of the same protein samples.

**A****B** Charge envelope of the Raf monomer**Supplemental Figure 7. Analysis of recombinant RAF1.**

**(A)** Purified  $\gamma$ -globulin, ovalbumin and RAF1 were separated in parallel by size exclusion chromatography. The smaller peaks observed for ovalbumin and  $\gamma$ -globulin are dimers. X-axis, time in min. QToF-MS was used to assess the purity of the three RAF1 peaks. Both the trimeric and monomeric peaks yielded a monomeric mass as 47,601.1 Da, corresponding to rRAF1. The peak migrating near 75 min, however, contained a mixture of proteins and thus cannot be confirmed as a RAF1 complex.

**(B)** QTOF mass spectrometry of the  $\sim$ 150 kDa peak of rRaf to determine the components of the higher molecular complex.



**Supplemental Figure 8.** Co-expression of LS and Rubisco biogenesis factors in *E. coli*.

The RAF1 requirement for the solubility of maize LS was assessed in *E. coli* cells with co-expression of RAF1 and other Rubisco biogenesis factors, including BSD2 (GRMZM2G062788), Cpn60α1, Cpn60β1, Cpn20-2 (GRMZM2G127609), Cpn20-3 (GRMZM2G399284), RBCX (GRMZM2G115476) and SS (GRMZM2G098520). Extracted lysate from untransformed *E. coli* BI21DE3 was used as negative control (lane 1).

**Methods:** Expression from Duet vectors (Novagen) was induced by 1 mM IPTG in BI21(DE3) cells grown to OD of 0.5, for two hr at 28°C. Equivalent amounts of cells were lysed by sonication and fractioned into soluble and insoluble fractions by centrifugation (15,000xg, 30 min at 4°C). Fractions were analyzed by 13% SDS-PAGE, followed by immunoblotting with LS (Agrisera), RAF1, S-tag (GenScript) and His-tag (Sigma) antibodies.



	Species	Locus
1	<i>Zea mays</i>	GRMZM2G457621
2	<i>Oryza sativa</i>	NM_001071247
3	<i>Sorghum bicolor</i>	XM_002448739
4	<i>Triticum aestivum</i>	AK334642
5	<i>Glycine max</i>	AK285150
6	<i>Arabidopsis thaliana-1</i>	AT5G28500
7	<i>Arabidopsis thaliana-2</i>	AT3G04550
8	<i>Brassica rapa</i>	AC189468
9	<i>Vitis vinifera</i>	XM_002268512
10	<i>Ricinus communis</i>	XM_002521916
11	<i>Populus trichocarpa</i>	XM_002319615
12	<i>Coffea canephora</i>	ABZ89185
13	<i>Medicago truncatula</i>	AC146861
14	<i>Selaginella moellendorffii</i>	XM_002964409
15	<i>Physcomitrella patens</i>	XM_001755326
16	<i>Picea glauca</i>	BT117013.1
17	<i>Ostreococcus tauri</i>	XM_003084347
18	<i>Micromonas sp.</i>	XM_002508645
19	<i>Nostoc azollae</i>	CP002059.1
20	<i>Anabaena variabilis</i>	CP000117.1
21	<i>Acaryochloris marina</i>	CP000828.1
22	<i>Synechococcus sp.</i>	CP000951.1
23	<i>Cyanothece sp.</i>	CP001287.1
24	<i>Gloeobacter violaceus</i>	BA000045.2
25	<i>Trichodesmium erythraeum</i>	CP000393.1
26	<i>Microcystis aeruginosa</i>	AM778923.1

**Supplemental Table 1.** Accession numbers of *RAF1* sequences from green species. The maize accession number was obtained from <http://www.maizegdb.org/>. The *raf1* loci from other species were obtained from Genbank, except the *Setaria viridis* sequence, which was kindly provided by Dr. Katrien Devos, Univ. Georgia. The protein sequence for *Selaginella moellendorffii* present in the database is incorrect.

21

## Data from WT sample

Accession number	Annotation	Curated Loc.	Score	Match
<a href="#">GRMZM2G385622_P01</a>	(GRMZM2G062854_P01) CF1 beta subunit (atpB)	thylakoid-peripheral-stromal-side	864	65
<a href="#">GRMZM2G360821_P01</a>	Ribulose biphosphate carboxylase large subunit	plastid stroma	720	54
<a href="#">NP_043022</a>	CF1 $\alpha$ alpha subunit (atpA)	thylakoid-peripheral-stromal-side	700	46
<a href="#">GRMZM2G083841_P01</a>	Phosphoenolpyruvate carboxylase 1 (PEPCase 1)	cytosol	644	75
<a href="#">GRMZM2G113033_P01</a>	Rubisco small subunit-4 (RBCS-4) -1	plastid stroma	443	22
<a href="#">GRMZM2G085019_P01</a>	NADP-malic enzyme 4-1&2 (NADP-ME4)	plastid stroma	414	21
<a href="#">GRMZM2G098520_P01</a>	Rubisco small subunit-4 (RBCS-4) -2	plastid stroma	400	26
<a href="#">GRMZM2G011507_P01</a>	pyruvate phosphate dikinase-1 (PPDK-1) M (C4)	plastid stroma	291	27
<a href="#">GRMZM2G337113_P02</a>	glyceraldehyde-3-phosphate dehydrogenase A-2 (GAPA-2)	plastid stroma	183	15
<a href="#">GRMZM2G048907_P01</a>	CF1 $\gamma$ subunit (atpC)	thylakoid-peripheral-stromal-side	170	17
<a href="#">GRMZM2G083716_P01</a>	Cpn60-beta-1	plastid stroma	153	17
<a href="#">AC215201.3_FGP005</a>	Cpn60-alpha-1	plastid stroma	146	21
<a href="#">GRMZM2G064096_P01</a>	germin-like protein (GER3)		111	10
<a href="#">GRMZM2G046284_P01</a>	fructose-bisphosphate aldolase-2 (SFBA-2)	plastid stroma	102	7
<a href="#">GRMZM2G007263_P01</a>	glyceraldehyde-3-phosphate dehydrogenase B (M-enriched)	plastid stroma	80	8
<a href="#">GRMZM2G106061_P01</a>	elongation factor Tu (EF-Tu-1), plastid	plastid stroma	60	8
<a href="#">GRMZM2G001696_P01</a>	phosphoenolpyruvate carboxykinase	mitochondria	52	2
<a href="#">GRMZM2G041275_P01</a>	ATP synthase beta subunit	mitochondria	48	4
<a href="#">GRMZM2G083016_P01</a>	phosphoglycerate kinase 2 (PGK-2) BS-enriched	plastid stroma	43	6
<a href="#">GRMZM2G089136_P01</a>	phosphoglycerate kinase 1 (PGK-1) M-enriched	plastid stroma	43	8
<a href="#">GRMZM2G018197_P01</a>	Ribosomal protein L4/L1 family	chloroplast/cytosol	42	6
<a href="#">GRMZM2G098290_P01</a>	glutamate-ammonia ligase (GS2), chloroplast	plastid stroma	41	5
<a href="#">GRMZM2G079668_P01</a>	cpHSP70-1 (DnaK homologue)	plastid stroma	41	3
<a href="#">NP_043046</a>	ribosomal protein S18	chloroplast	38	2
<a href="#">GRMZM2G084465_P01</a>	40S ribosomal protein S23		36	3
<a href="#">GRMZM2G007695_P01</a>	Ribosomal protein L4/L1 family	chloroplast/cytosol	35	4
<a href="#">GRMZM2G015419_P02</a>	lipoxygenase 2.3	chloroplast	35	5
<a href="#">GRMZM2G101689_P01</a>	U4/U6 small nuclear ribonucleoprotein PRP4-like protein-like		35	4
<a href="#">GRMZM2G353103_P01</a>	40S ribosomal protein S16		33	1
<a href="#">GRMZM2G016232_P01</a>	Histone H4	nucleus	33	2
<a href="#">GRMZM2G360021_P01</a>	50S ribosomal protein L12-C	plastid ribosome	32	6

22

Data from *raf1* sample

<a href="#">GRMZM2G083841_P01</a>	Phosphoenolpyruvate carboxylase 1 (PEPCase 1)	cytosol	1409	123
<a href="#">GRMZM2G011507_P01</a>	pyruvate phosphate dikinase-1 (PPDK-1) M (C4)	plastid stroma	565	46
<a href="#">GRMZM2G085019_P01</a>	NADP-malic enzyme 4-1&2 (NADP-ME4)	plastid stroma	462	17
<a href="#">GRMZM2G385622_P01</a>	(GRMZM2G062854_P01) CF1 beta subunit (atpB)	thylakoid-peripheral-stromal-side	449	54
<a href="#">NP_043022</a>	CF1 $\alpha$ alpha subunit (atpA)	thylakoid-peripheral-stromal-side	248	28
<a href="#">GRMZM2G064096_P01</a>	germin-like protein (GER3)		245	13
<a href="#">GRMZM2G097457_P01</a>	C3-type pyruvate phosphate dikinase-2	plastid stroma	192	12
<a href="#">GRMZM2G106061_P01</a>	elongation factor Tu (EF-Tu-1), plastid	plastid stroma	158	7
<a href="#">GRMZM2G158394_P01</a>	extracellular ribonuclease LE precursor		145	12
<a href="#">GRMZM2G007263_P01</a>	glyceraldehyde-3-phosphate dehydrogenase B (M-enriched)	plastid stroma	138	14
<a href="#">GRMZM2G121878_P01</a>	carbonic anhydrase chloroplastic-like	chloroplast	131	23
<a href="#">GRMZM2G337113_P02</a>	glyceraldehyde-3-phosphate dehydrogenase A-2 (GAPA-2)	plastid stroma	124	10
<a href="#">GRMZM2G046284_P01</a>	fructose-bisphosphate aldolase-2 (SFBA-2)	plastid stroma	119	8
<a href="#">GRMZM2G089136_P01</a>	phosphoglycerate kinase		107	20
<a href="#">GRMZM2G083716_P01</a>	Cpn60-beta-1	plastid stroma	101	11
<a href="#">GRMZM2G079668_P01</a>	cpHSP70-1 (DnaK homologue)	plastid stroma	66	3
<a href="#">GRMZM2G047028_P01</a>	phosphoglycerate kinase		56	4
<a href="#">GRMZM2G154218_P01</a>	elongation factor 1 alpha		53	5
<a href="#">GRMZM2G048907_P01</a>	CF1 gamma subunit (atpC)	thylakoid-peripheral-stromal-side	49	2
<a href="#">GRMZM2G134797_P02</a>	nucleoside diphosphate kinase		48	8
<a href="#">NP_043046</a>	ribosomal protein S18	chloroplast	45	4
<a href="#">GRMZM2G057535_P01</a>	elongation factor 1-alpha		45	3
<a href="#">GRMZM2G172369_P03</a>	lysosomal alpha-mannosidase-like		43	1
<a href="#">GRMZM2G008247_P01</a>	beta-D-glucosidase precursor		40	8
<a href="#">AC215201.3_FGP005</a>	Cpn60-alpha-1	plastid stroma	40	7
<a href="#">GRMZM2G119169_P01</a>	60S ribosomal protein L17		40	3
<a href="#">GRMZM2G129246_P01</a>	peroxisomal (S)-2-hydroxy-acid oxidase GLO1-like		39	2
<a href="#">GRMZM2G134264_P01</a>	hypothetical protein		39	4
<a href="#">GRMZM2G024354_P01</a>	60S ribosomal protein L15		32	1
<a href="#">GRMZM2G013478_P01</a>	nucleoside diphosphate kinase 2		31	6
<a href="#">GRMZM2G360021_P01</a>	50S ribosomal protein L12-1	plastid ribosome	30	2
<a href="#">GRMZM2G137868_P01</a>	serine-glyoxylate aminotransferase		29	2
<a href="#">GRMZM2G084465_P01</a>	40S ribosomal protein S23		29	2

**Supplemental Table 2.** Protein data obtained from gel slices corresponding to the position of LS<sub>C</sub>. Accession numbers are from Gramene. "Curated Loc." indicates the annotated subcellular location. "Score," the sum of the highest ion score for each protein; Match, the number of MS/MS spectra that were matched to each protein. The cells highlighted in yellow indicate proteins identified in both genotypes; cells highlighted in green indicate the chaperonins discussed in the main text.

2'

	Primer name	Sequence
1	ZmBSD2-SQF	CGTCTCCCGCCACTGCAAGA
2	ZmBSD2-SQR	TTCATCGAAAGTGCTAAACCGATGAC
3	ZmActin1-SQF	TACGAGCGCTGAGCGGGAGA
4	ZmActin1-SQR	GCTGCTAGGAGCCAGGGACG
5	ZmrbcL-Cod1	TTGCCGCGACAACGGCCTAC
6	ZmrbcL-Rev1	TCGCGCCCTTCGTTACGAGC
7	ZmRBCS2-Cod1	GGCCTACGGCAACAAGAAGTT
8	ZmRBCS2-Rev1	CAAGAATCGACAGTCGTTGC
9	ZmPsbA-5'	CTTCTTCTTGCCCGAATCTG
10	ZmPsbA-3'	TATCGCATTCAATTGCTGCTC
11	EoMu1	GCCTCCATTTGTCGAAT CCC
12	EoMu2	GCCTCTATTTGTCGAATCCG
13	ZmRAF1-FOR1	AGCTCTACC AGCCGTTCCAC
14	ZmRAF1-REV1	CTCGAGCCTGTAGCCGATG
	ZmRAF1-FOR2	GGTCCTGTACGAGCTCCGCT
15	ZmRAF1-REV2	CATGTCGTCCTCGTCCTTGG
16	Zmqactin-1F	GTG ACA ATGGCACTGGAATG
17	Zmqactin-1R	CCA TGCTCAATCGGGTACTT
18	ZmqLS-1F	AAGGGGAACGCGAAATAACT
19	ZmqLS-123R	AGGCTTCTAAAGCCACACGA
20	ZmqSS-9F	GTTTCGTGTACCGCGAGAACT
21	ZmqSS-9R	GTCGATCTAGCCACGGTCTC
22	ZmqME-1F	TGGCAGAGCAGACGTATTTG
23	ZmqME-1R	TGAAGGGAGCCTTTACGAGA
24	ZmqMDH-1F	TCACCTGCTGTTCAA ACTCG
25	ZmqMDH-1R	GGATACAGCGAGTCCTCCAG

**Supplemental Table 3.** Primers used in this work.