Supplemental data

Photosynthetic trichomes contain a specific Rubisco with a modified pH-dependent activity

Raphaëlle Laterre, Mathieu Pottier, Claire Remacle and Marc Boutry

Supplemental Table S1 – MS-MS identification of the 49 most abundant trichome-specific protein spots.

Spot[1]	EST accession[2]	Corresponding protein name
Metabolism		
20	FG175097	Glutamine synthetase
4	EB432526	Nucleoside diphosphate kinase
24	FG194360	3-isopropylmalate dehydrogenase large subunit
22	FG194360	3-isopropylmalate dehydrogenase large subunit
26	FG139607	D-cysteine desulfhydrase
42	DV157962	Rubisco small subunit
15	EB679897	NAD-dependent dehydrogenase
43	EB679897	NAD-dependent dehydrogenase
Energy		
27	DV162381	Phosphoglycerate kinase
33g	FG133688	ATP synthase β subunit
33d	FG135901	ATP synthase β subunit
3d	FS414373	ATP synthase d subunit
30	EB680336	Aspartate aminotransferase
13	FG198580	Light harvesting chlorophyll a/b-binding protein
Protein destination and storage		
34d	DW000222	Chaperonin-60 beta subunit
Disease/defense		
1	FS387666	Putative PR-10 type pathogenesis-related protein
2	FS387666	Putative PR-10 type pathogenesis-related protein
3g	FS387666	Putative PR-10 type pathogenesis-related protein
6	FS387666	Putative PR-10 type pathogenesis-related protein
32	AM817165	Pathogenesis-related protein
9	FG194579	Thioredoxin peroxidase
10g	FG194579	Thioredoxin peroxidase
5	FS429935	Thioredoxin peroxidase
19	EB677336	Fe-superoxide dismutase
18	EB677336	Fe-superoxide dismutase
8	EH622457	Superoxide dismutase
7	EH624278	Glutathione peroxidase
17	FS374016	Glutathione S-transferase
16	FS374016	Glutathione S-transferase
14b	FS411447	Glutathione S-transferase
14h	FS411447	Glutathione S-transferase
44	EB447814	Glutathione S-transferase
10d	EB440026	Peroxiredoxin
11	EB440026	Peroxiredoxin
Secondary metabolism		
34g	FG187107	Cembratrienol synthase 2b
35	FG187107	Cembratrienol synthase 2b
36	FG187107	Cembratrienol synthase 2b
37	FG187107	Cembratrienol synthase 2b
28	EH615865	Chalcone synthase
12	EB426860	Chalcone isomerase
25	FG175910	1-deoxy-D-xylulose-5-phosphate reductoisomerase
23	DV998901	1-hydroxy-2-methyl-butenyl 4-diphosphate reductase
21	FG187212	1-hydroxy-2-methyl-butenyl 4-diphosphate reductase
38	FG175170	(E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase
39	FG149054	(E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase
40	FG149054	(E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase
41	FG149054	(E)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase
Unknown		
29	AM816756	No identification
31	AM816756	No identification

[1] Numbering corresponds to that of Supplemental Fig 2. [2] Closest EST from the translated *N. tabacum* NCBI EST database

Supplemental Table S2 – List of the protein sequences used for the phylogenetic analysis of the RbcS. A, a protein BLAST search was performed (5 May 2015) using the Nt-RbcS-T mature protein sequence as a query. All of the proteins displaying a score higher than 200 are listed. Although the rice Rubisco small subunit, OsRbcS1 displays a BLAST score lower than the 200 threshold, its sequence was added to this analysis since its corresponding gene also exhibits a specific expression pattern (Morita *et al*., 2014). Because, no sequence from C4 plants was obtained through the BLAST analysis, the two RbcS sequences of *Zea mays*, a C4 plant, were also included. The "Secretory organs" column mentions whether glandular trichomes (GT), trichomes without more information (T) secretory cavities (SC), nectaries (N) or colleters (C) have been reported for each species. B, for each species listed in A, the RbcS protein showing the highest identity with the protein listed in A was retrieved. In addition, all of the RbcS sequences from *Arabidopsis* were listed.

Abbreviation	Accession number (NCBI)	Species	Secretory organs	References[1]
A. Cluster T				
NtRbcS-t	DV157962.1	*Nicotiana tabacum*	GT	1
Nsy-t	XP_009794031.1	*Nicotiana sylvestris*	GT	1
Nto-t	XP_009608851.1	*Nicotiana tomentosiformis*	GT	1
Stu-t	XP_006344531.1	*Solanum tuberosum*	GT	2
Sly-t	XP_004243079.1	*Solanum lycopersicum*	GT	3
Egr-t	XP_010046977.1	*Eucalyptus grandis*	SC	4
Lja-t1	AFK34410.1	*Lotus japonicus*	T	5
Cca-t	CDP02988.1	*Coffea canephora*	GC	6
Egu-t	EYU37322.1	*Erythranthe guttata*	T	7
Csi-t1	KDO43856.1	*Citrus sinensis*	SC	8
Jcu-t1	XP_012074043.1	*Jatropha curcas*	T	9
Ccl-t1	XP_006453353.1	*Citrus clementina*	SC	8
Jcu-t2	KDP36388.1	*Jatropha curcas*	T	9
Ccl-t2	XP_006453352.1	*Citrus clementina*	SC	8
Csi-t2	KDO43855.1	*Citrus sinensis*	SC	8
Lja-t2	AFK37376.1	*Lotus japonicus*	T	5
Gra-t	KJB20540.1	*Gossypium raimondii*	GT	10
Gar-t	KHG02980.1	*Gossypium arboreum*	GT	10
Atr-t	XP_006856657.1	*Amborella trichopoda*	T	11
Peu-t	XP_011043951.1	*Populus euphratica*	GT	12
Ptr-t	ABK96070.1	*Populus trichocarpa*	GT	12
Vvi-t	XP_002270117.1	*Vitis vinifera*	GT	13
Sin-t	XP_011069924.1	*Sesamum indicum*	GT	14
Ppe-t	XP_007223576.1	*Prunus persica*	C	15
Mno-t	XP_010094223.1	*Morus notabilis*	GT	16
Fve-t	XP_004296442.1	*Fragaria vesca*	GT	17
Tca-t	XP_007014245.1	*Theobroma cacao*	GT	18
Pmu-t	XP_008223772.1	*Prunus mume*	C	15
Pbr-t	XP_009336182.1	*Pyrus bretschneideri*	N	19
Gso-t	KHN22693.1	*Glycine soja*	GT	20
Nnu-t	XP_010275523.1	*Nelumbo nucifera*	GT	21
Mtr-t	XP_003610095.1	*Medicago truncatula*	GT	22
Gma-t	NP_001235868.1	*Glycine max*	GT	23
Pda-t	XP_008803597.1	*Phoenix dactylifera*	T	24
Egi-t	XP_010914155.1	*Elaeis guineensis*	T	25
Tha-t	XP_010526347.1	*Tarenaya hassleriana*	GT	26
OsRbcS1	BAD38596.1	*Oryza sativa*	-	27
B. Cluster M				
Nta	CAA26208.1	*Nicotiana tabacum*		
Nsy	P22433.1	*Nicotiana sylvestris*		
Nto	XP_009625434.1	*Nicotiana tomentosiformis*		
Stu	ABY21254.1	*Solanum tuberosum*		
Sly	BAA01888.1	*Solanum lycopersicum*		
Egr	XP_010044264.1	*Eucalyptus grandis*		

Lja-1	AFK45650.1	*Lotus japonicus*
Cca	CBY05833.1	*Coffea canephora*
Egu	EYU38174.1	*Erythranthe guttata*
Csi	XP_006482254.1	*Citrus sinensis*
Jcu	ADB85091.1	*Jatropha curcas*
Ccl	XP_006430779.1	*Citrus clementina*
Gra	AFS41724.1	*Gossypium raimondii*
Gar	AFS41722.1	*Gossypium arboreum*
Atr	XP_006842277.2	*Amborella trichopoda*
Peu	XP_011035062.1	*Populus euphratica*
Ptr	XP_002305162.2	*Populus trichocarpa*
Vvi	XP_002276967.1	*Vitis vinifera*
Sin	XP_011099669.1	*Sesamum indicum*
Ppe	XP_007215051.1	*Prunus persica*
Mno	XP_010103758.1	*Morus notabilis*
Fve	XP_004303137.1	*Fragaria vesca*
Tca	XP_007033270.1	*Theobroma cacao*
Pmu	XP_008244506.1	*Prunus mume*
Pbr	XP_009357812.1	*Pyrus bretschneideri*
Gso	KHN47663.1	*Glycine soja*
Nnu	XP_010263805.1	*Nelumbo nucifera*
Mtr	XP_003618735.1	*Medicago truncatula*
Gma	AAG24884.1	*Glycine max*
Pda	XP_008797838.1	*Phoenix dactylifera*
Egi	XP_010916368.1	*Elaeis guineensis*
Tha	XP_010521010.1	*Tarenaya hassleriana*
OsRbcS2	BAT16666.1	*Oryza sativa*
OsRbcS3	BAT16784.1	*Oryza sativa*
OsRbcS4	BAT16793.1	*Oryza sativa*
OsRbcS5	BAT16789.1	*Oryza sativa*
Zma1	NP_001105294	*Zea mays*
Zma2	NP_001146043	*Zea mays*
Ath-1	AAD10655.1	*Arabidopsis thaliana*
Ath-2	CAA32702.1	*Arabidopsis thaliana*
Ath-3	CAA32700.1	*Arabidopsis thaliana*
Ath-4	CAA31948.1	*Arabidopsis thaliana*

[1]References

1. Sierro N, Battey JND, Ouadi S, Bovet L, Goepfert S, Bakaher N, Peitsch MC, Ivanov NV (2013) Reference genomes and transcriptomes of *Nicotiana sylvestris* and *Nicotiana tomentosiformis.* Genome Biol 14: R60

2. Lyshede OB (1980) The ultrastructure of the glandular trichomes of *Solanum tuberosum*. Annals Bot 46: 519-526

3. Bergau N, Bennewitz S, Syrowatka F, Hause G and Tissier A (2015) The development of type VI glandular trichomes in the cultivated tomato *Solanum lycopersicum* and a related wild species *S. habrochaites*. BMC Plant Biology 15:289

4. Goodger JQD, Heskes AM, Mitchell MC, King DJ, Neilson EH, Woodrow IE (2010) Isolation of intact sub-dermal secretory cavities from *Eucalyptus* (2010) Plant Methods 6: 20

5. Stenglein SA, Colares MN, Arambarri AM, Novoa MC, Vizcaino CE, Katinas L (2003) Leaf epidermal microcharacters of the Old World species of *Lotus* (Leguminosae: Loteae) and their systematic significance. Aust J Bot 51: 459-469

6. Mayer JLS, Carmello-Guerreiro SM, Mazzafera P (2013) A functional role for the colleters of coffee flowers. AoB Plants 5: plt029

7. Nesom GL (2013) Two new species of *Erythranthe* Sect. *Mimulosma* (Phrymaceae) from California. Phytoneuron 43: 1-10

8. Bennici A, Tani C (2004) Anatomical and ultrastructural study of the secretory cavity development of *Citrus sinensis* and *Citrus limon*: evaluation of schizolysigenous ontogeny. Flora 199: 464-475

9. Rai K, Tiwari E (2012) Epidermal studies in identification of *Jatropha* species. IOSR J Pharmacy Biol Sci 2: 41-51

10. Turley RB, Vaughn KC (2012) Differential expression of trichomes on the leaves of upland cotton (*Gossipum hirsutum* L.). J Cotton Sci 16:53-71

11. Endress PK, Igersheim A (2000) The reproductive structures of the basal angiosperm *Amborella trichopoda* (Amborellaceae). Int J Plant Sci 161: S237-S248

12. Trealase W (1881) The foliar nectar glands of *Populus*. Bot Gaz 6: 284-290

13. Ma ZY, Wen J, Ickert-bond SM, Chen LQ, Liu XQ (2016) Morphology, structure, and ontogeny of trichomes of the grape genus (*Vitis*, Vitaceae) Frontiers Plant Sci 7: 704

14. Zhigila DA, Sawa FBJ, Abdul SD (2015) Systematic significance and pharmaceutical potentials of trichomes in accessions of *Sesamum indicum* L. J Med Plants Res 9: 385-394

15. Chin SW, Lutz S, Wen J, Potter D (2013) The bitter and the sweet: inference of homology and evolution of leaf glands in *Prunus* (Rosaceae) through anatomy, micromorphology, and ancestral-character state reconstruction. Int J Plant Sci 174: 27-46

16. Singhal BK, Baqual MF, Khan MA, Bindroo BB, Dhar A (2010) Leaf surface scanning electron microscopy of 16 mulberry genotypes (*Morus* spp.) with respect to their feeding value in silkworm *Bombyx mori* L.) rearing. Chil J Agric Res 70: 191-198

17. Figueiredo AST, Resende JTV, Morales RGF, Gonçalves APS, Da Silva PR (2013) The role of glandular and non-glandular trichomes in the negative interactions between strawberry cultivar and spider mite. Arthropod-Plant Interactions 7: 53-58

18. Bailey BA, Strem MD, Wood D (2009) *Trichoderma* species form endophytic associations within *Theobroma cacao* trichomes. Mycological Res 113: 1365-1376

19. Qinggang XHC, Zhenghai H (1998) Developmental and anatomical studies of the floral nectaries in *Pyrus bretschneideri*.J. Laiyang Agricultural College 1998-01

20. Healy RA, Palmer RG, Horner HT (2009) Multicellular secretory trichome development on soybean and related *Glycine* gynoecia? Int J Plant Sci 170: 444-456

21. Hayes V, Schneider EL, Carlquist S (2000) Floral development of *Nelumbo nucifera* (Nelumbonaceae). Int J Plant Sci 161: S183-S191

22. Aziz N, Paiva NL, May GD, Dixon RA (2005) Transcriptome analysis of alfalfa glandular trichomes. Planta 221: 28-38

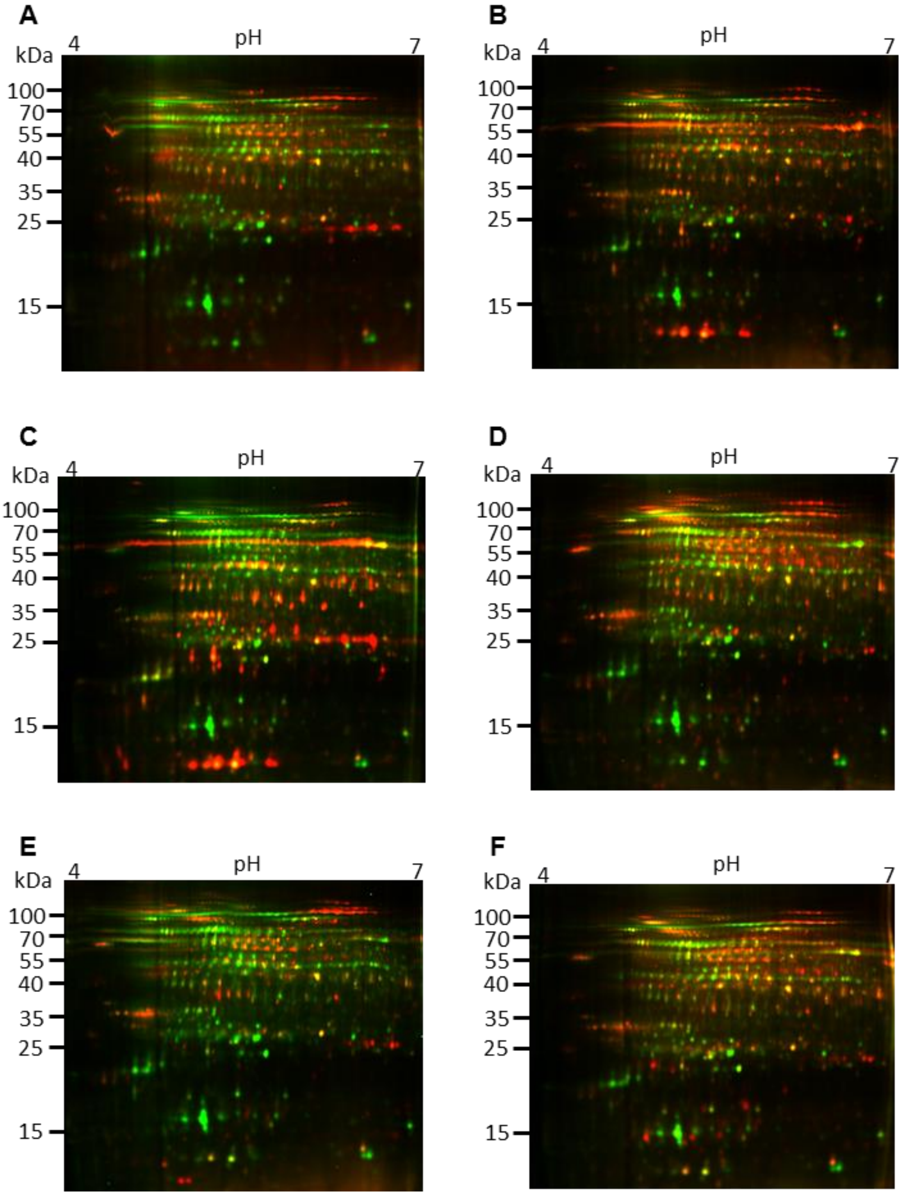
23. Franceschi VR, Giaquinta RT (1983) Glandular trichomes of soybean leaves: cytological differentiation from initiation through senescence. Bot Gaz 144: 175-184

24. Karakish EAK (2008) Studies in the leaf anatomy of some species of Phoenix (Family Araceae). Thesis. Umm Al-Qura University

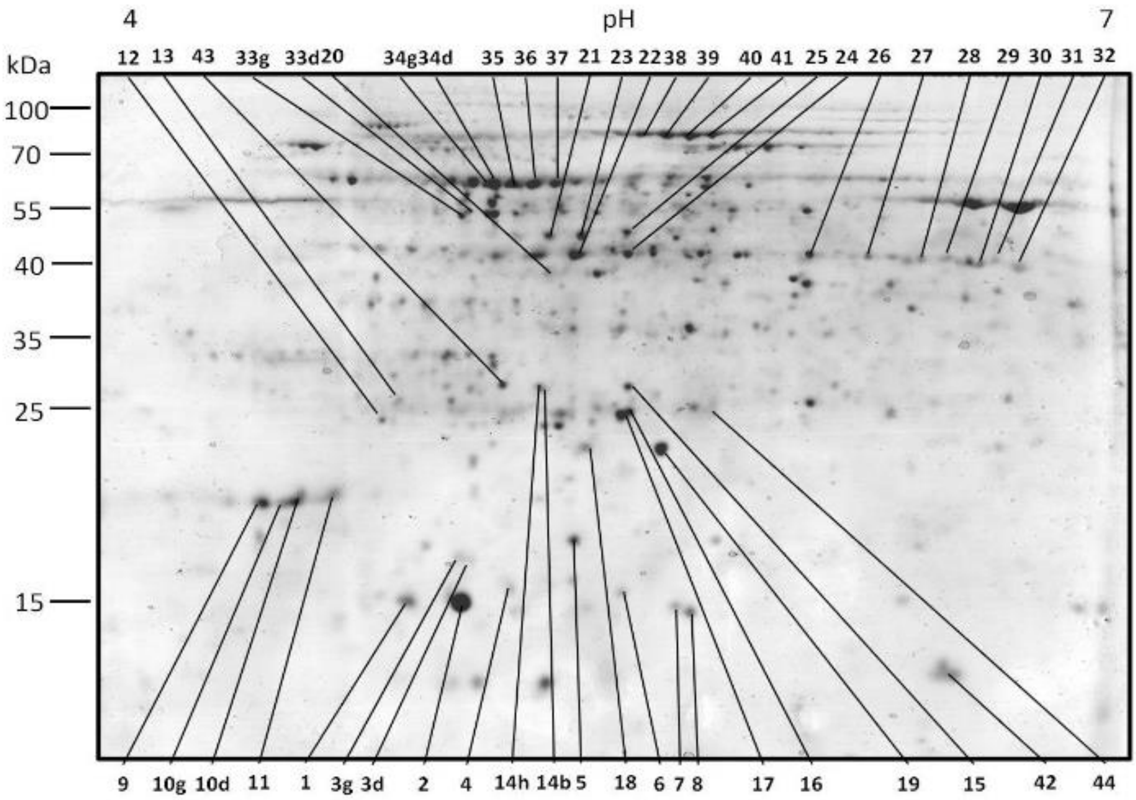
25. Luis ZG, Bezerra KM, Scherwinski-Pereira JE (2010) Adaptability and leaf anatomical features in oil palm seedlings produced by embryo rescue and pre-germinated seeds. Braz J Plant Physiol 22: 209-215

26. Williams LAD, Vasques E, Reid W, Porter R, Kraus W (2003) Biological activities of an extract from *Cleome viscosa* L. (Capparaceae). Naturwissenschaften 90: 468-472

27. Li W, Wu J, Weng S, Zhang D, Zhang Y, Shi C (2010). Characterization and fine mapping of the glabrous leaf and hull mutants (gl1) in rice (Oryza sativa L.). Plant Cell Rep 29: 617-627



Supplemental Figure S1 – 2D-DIGE of soluble proteins of tall glandular trichomes compared to other tissues. Proteins of tall glandular trichomes were compared to proteins from roots (A), germinating seeds (B), leaves (C), seed pods (D), stems (E, and flowers (F). Soluble proteins (20 µg) isolated from trichomes and other tissues were labeled with the Cy3 (green) and Cy5 (red) fluorescent dyes, respectively. As an internal control, 10 µg of both compared protein samples were mixed and tagged with the Cy2 (blue) fluorescent dye (not shown).



Supplemental Figure S2 – Preparative 2-DE of soluble tall glandular trichome proteins. Proteins (500 µg) from tall glandular trichomes were analyzed by 2D gel electrophoresis and stained with colloidal blue. Spots that are numbered correspond to trichome-specific proteins and were analyzed by mass spectrometry. Their identification is given in Supplemental Table S1.

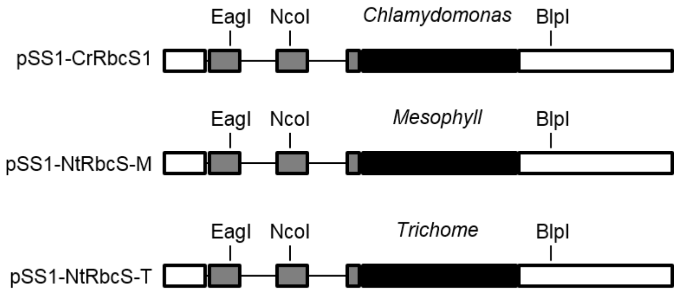
A

```
ATGATGACCTGGAACCCGATCGACAACAAGAAGTTCGAGACCCTGTCCTACCTGCCGCCGCTGTCCG
AGGACTCCATCGCCAAGGAGGTCGACTACATGATCCAGAAGGGCTGGATCCCGTGCCTGGAGTTCGA
CCAGGTCGGCTACGTCCGCCGCGAGAACTCCTCCATGCCGGGCTACTACGACGGCCGCTACTGGACC
CTGTGGAAGCTGCCGATGTTCGGCTGCAACGACTCCTCCCAGGTCCTGAACGAGATCCAGGAGTGCA
AGAAGGCCTACCCGAACGCCTTCATCCGCTGCCTGGCCTTCGACAACGTCAAGCAGGTCCAGTGCAT
GGCCTTCCTGATCCAGAAGCCGGCCGCCGCCTAA
```

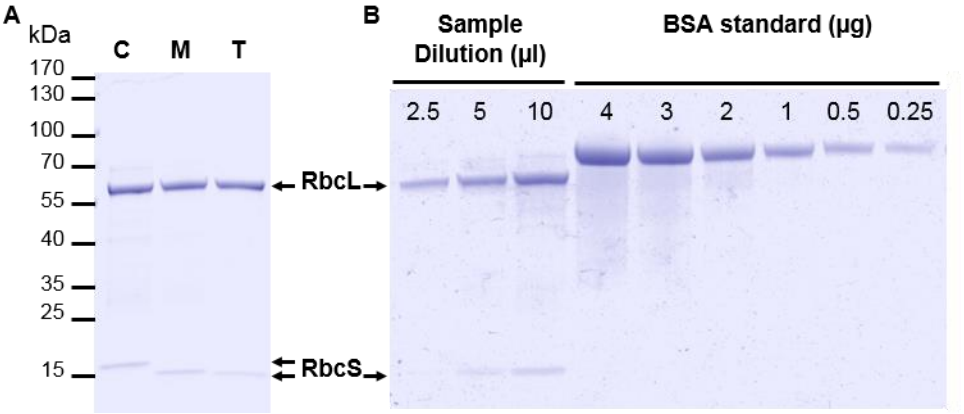
B

```
ATGATGGTCTGGCCGCCGATCAACAAGAAGAAGTACGAGACCCTGTCCTACCTGCCGGACCTGTCCC
AGGAGCAGCTGCTGTCCGAGGTCGAGTACCTGCTGAAGAACGGCTGGGTCCCGTGCCTGGAGTTCGA
GACCGAGCACGGCTTCGTCTACCGCGAGAACAACAAGTCCCCGGGCTACTACGACGGCCGCTACTGG
ACCATGTGGAAGCTGCCGATGTTCGGCTGCACCGACGCCACCCAGGTCCTGGCCGAGGTCGAGGAGG
CCAAGAAGGCCTACCCGCAGGCCTGGATCCGCATCATCGGCTTCGACAACGTCCGCCAGGTCCAGTG
CATCTCCTTCATCGCCTACAAGCCGGAGGGCTACTAA
```

Supplemental Figure S3. Synthetic coding sequences of NtRbcS-T and Nt-RbcS-M for expression in *Chlamydomonas.* The coding sequences for the mature small subunit NtRbcS-T (A) and NtRbcS-M (B) were optimized for expression in *Chlamydomonas*, and engineered to replace the plant Lys2 and Gln2, respectively, with the *Chlamydomonas* Met to facilitate transit-peptide processing. These sequences were used to exactly replace the *Chlamydomonas rbcS1* mature-protein coding sequence of pSS1-ITP (Genkov *et al*., 2010) to generate the plasmids pSS1-NtRbcS-T and pSS1-NtRbcS-M.



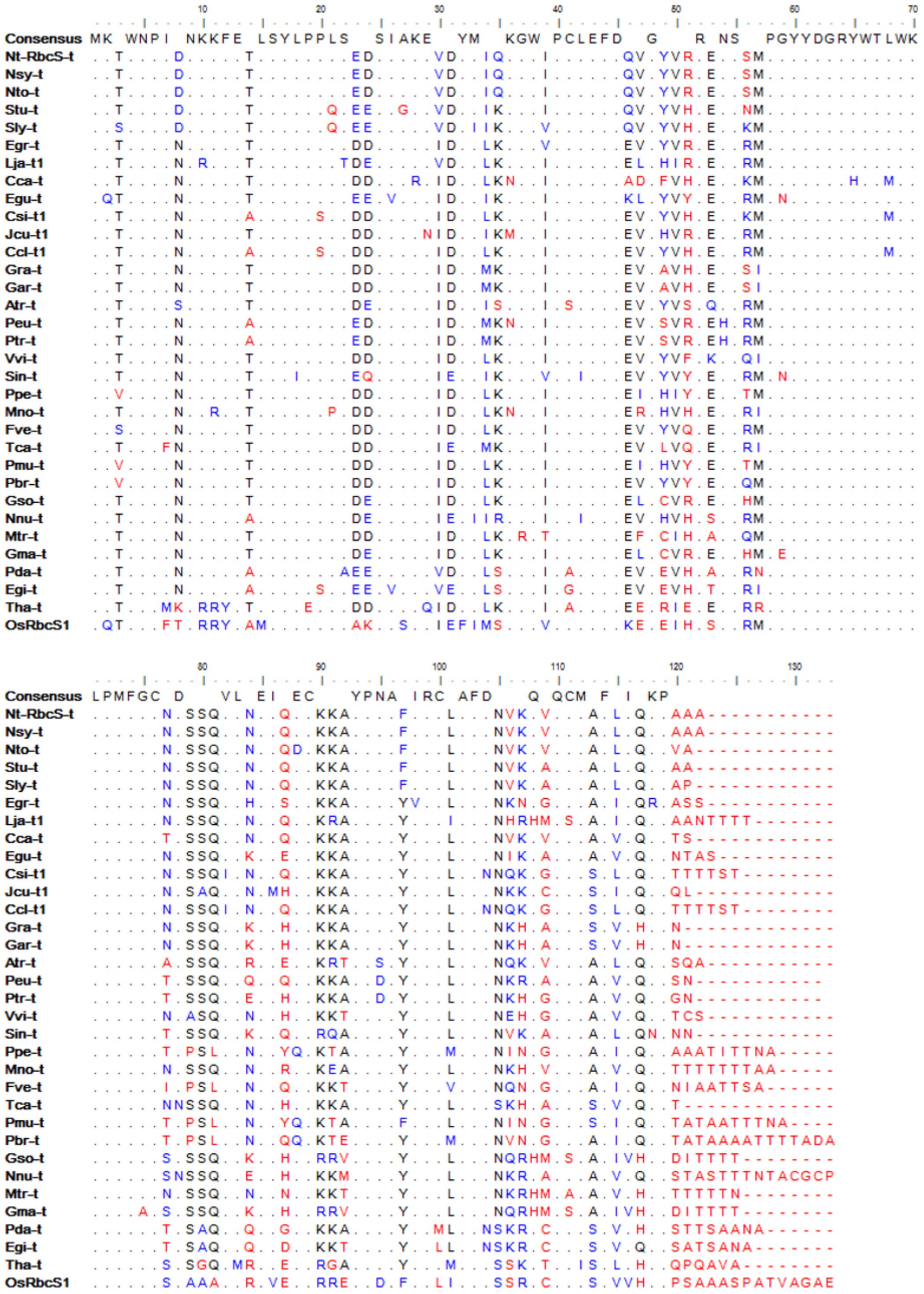
Supplemental Figure S4. Engineered plasmids containing the wild-type *Chlamydomonas rbcS1* gene (pSS1-CrRbcS1) and the *N. tabacum* mesophyll- and trichome-expressed *rbcS* cDNA sequences (pSS1-NtRbcS-M and pSS1-NtRbcS-T). The regions encoding the transit peptide and mature small subunit are shown as *gray* and *black boxes*, respectively, and introns are denoted by black lines. Introns are inserted into the transit peptide-encoding region. White boxes represent part of the plasmid. Figure adapted from Genkov *et al.*, 2010.



Supplemental Figure S5. Electrophoretic analysis of the purified Rubisco from *Chlamydomonas* transformants. A, Purified Rubisco (1 µg) from the strain expressing pSS1-CrRbcS1 (C), pSS1-NtRbcS-M (M), or pSS1-NtRbcS-T (T) were analyzed by SDS-PAGE and gel stained with Coomassie blue. B, Example of Rubisco quantification on gel. Increasing quantities of Rubisco and bovine serum albumin (BSA) used a standard were analyzed by SDS-PAGE. The gel pictures were then analyzed using ImageJ software.

Supplemental Figure S6. Comparison of the Rubisco amount in trichome and leaf samples of *N. tabacum*. Soluble proteins (pool of the four biological repetitions of Fig. 6B) from leaves cleared of trichomes and from trichomes were analyzed by SDS-PAGE and western blotting using an anti-RbcL antibody. Relative quantification of the Rubisco determined in both samples using ImageJ software showed that trichome samples contains 4.5 times less Rubisco that leaf samples.

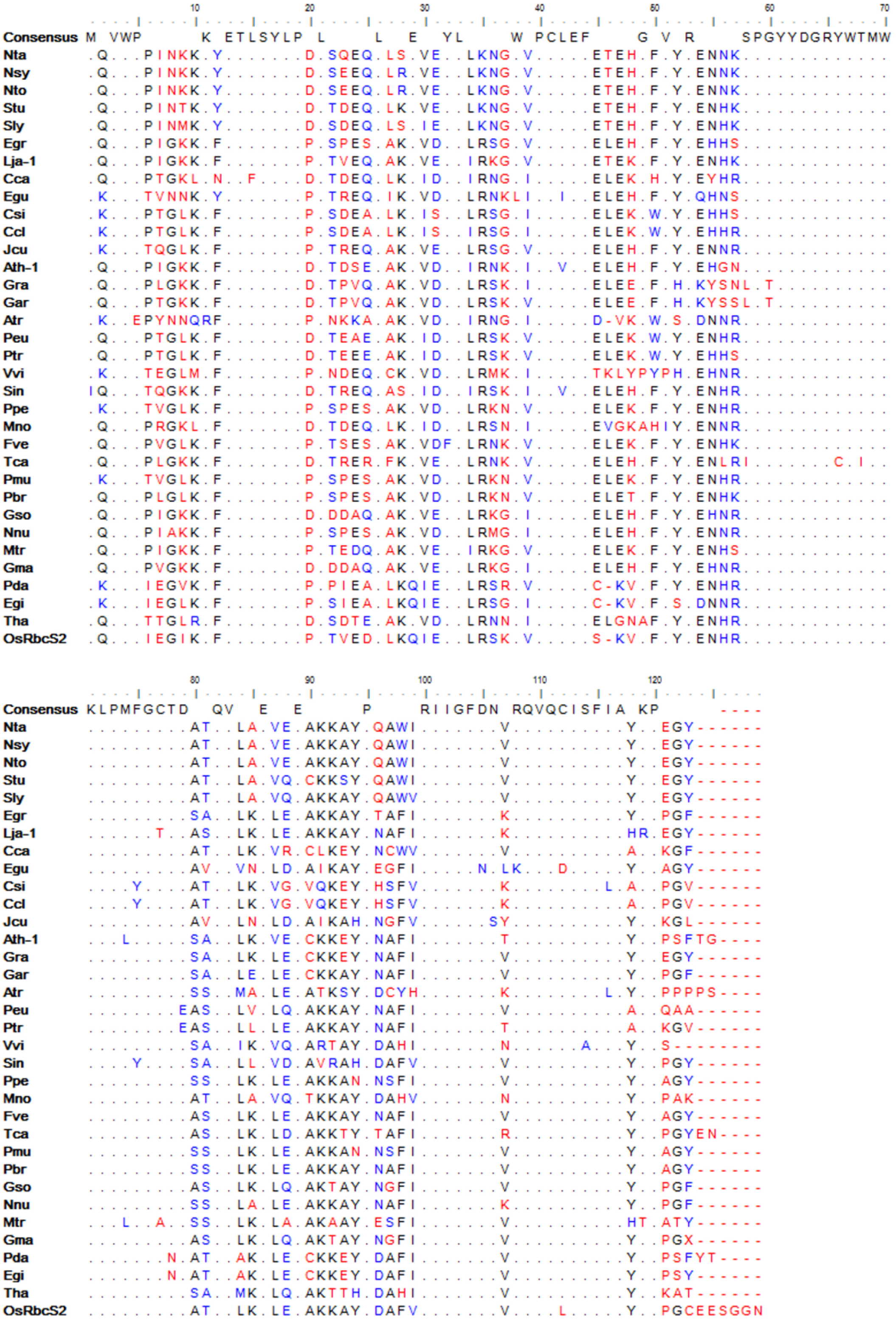
Cluster T



```
                    10        20        30        40        50        60        70
            ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Consensus   MK WNPI NKKFE LSYLPPLS  SIAKE  YM  KGW PCLEFD  G   R NS  PGYYDGRYWTLWK
Nt-RbcS-t   ..T....D.....T........ED.....VD..IQ...I......QV.YVR.E..SM.............
Nsy-t       ..T....D.....T........ED.....VD..IQ...I......QV.YVR.E..SM.............
Nto-t       ..T....D.....T........ED.....VD..IQ...I......QV.YVR.E..SM.............
Stu-t       ..T....D.....T......Q.EE..G..VD..IK...I......QV.YVH.E..NM.............
Sly-t       ..S....D.....T......Q.EE.....VD.IIK...V......QV.YVH.E..KM.............
Egr-t       ..T....N.....T........DD.....ID..LK...V......EV.YVR.E..RM.............
Lja-t1      ..T....N.R...T.......TDE.....VD..LK...I......EL.HIR.E..RM.............
Cca-t       ..T....N.....T........DD...R.ID..LKN..I......AD.FVH.E..KM.......H..M..
Egu-t       .QT....N.....T........EE.V...ID..IK...I......KL.YVY.E..RM.N...........
Csi-t1      ..T....N.....A.....S..DD.....ID..LK...I......EV.YVH.E..KM..........M..
Jcu-t1      ..T....N.....T........DD....NID..IKM..I......EV.HVR.E..RM.............
Ccl-t1      ..T....N.....A.....S..DD.....ID..LK...I......EV.YVH.E..RM..........M..
Gra-t       ..T....N.....T........DD.....ID..MK...I......EV.AVH.E..SI.............
Gar-t       ..T....N.....T........DD.....ID..MK...I......EV.AVH.E..SI.............
Atr-t       ..T....S.....T........DE.....ID..IS...I.S....EV.YVS.Q..RM.............
Peu-t       ..T....N.....A........ED.....ID..MKN..I......EV.SVR.EH.RM.............
Ptr-t       ..T....N.....A........ED.....ID..MK...I......EV.SVR.EH.RM.............
Vvi-t       ..T....N.....T........DD.....ID..LK...I......EV.YVF.K..QI.............
Sin-t       ..T....N.....T...I....EQ.....IE..IK...V..I...EV.YVY.E..RM.N...........
Ppe-t       ..V....N.....T........DD.....ID..LK...I......EI.HIY.E..TM.............
Mno-t       ..T....N..R..T......P.DD.....ID..LKN..I......ER.HVH.E..RI.............
Fve-t       ..S....N.....T........DD.....ID..LK...I......EV.YVQ.E..RM.............
Tca-t       ..T...FN.....T........DD.....IE..MK...I......EV.LVQ.E..RI.............
Pmu-t       ..V....N.....T........DD.....ID..LK...I......EI.HVY.E..TM.............
Pbr-t       ..V....N.....T........DD.....ID..LK...I......EV.YVY.E..QM.............
Gso-t       ..T....N.....T........DE.....ID..LK...I......EL.CVR.E..HM.............
Nnu-t       ..T....N.....A........DE.....IE.IIR...I..I...EV.HVH.S..RM.............
Mtr-t       ..T....N.....T........DD.....ID..LK.R.T......EF.CIH.A..QM.............
Gma-t       ..T....N.....T........DE.....ID..LK...I......EL.CVR.E..HM.E...........
Pda-t       ..T....N.....A.......AEE.....VD..LS...I.A....EV.EVH.A..RN.............
Egi-t       ..T....N.....A.....S..EE.V...VE..LS...I.G....EV.EVH.T..RI.............
Tha-t       ..T...MK.RRY.T....E...DD....QID..LK...I.A....EE.RIE.E..RR.............
OsRbcS1     .QT...FT.RRY.AM.......AK..S..IEFIMS...V......KE.EIH.S..RM.............

                    80        90       100       110       120       130
            ....|....|....|....|....|....|....|....|....|....|....|....|...
Consensus   LPMFGC D   VL EI EC   YPNA IRC AFD   Q QCM F I KP
Nt-RbcS-t   ......N.SSQ..N..Q..KKA....F...L...NVK.V...A.L.Q..AAA-----------
Nsy-t       ......N.SSQ..N..Q..KKA....F...L...NVK.V...A.L.Q..AAA-----------
Nto-t       ......N.SSQ..N..QD.KKA....F...L...NVK.V...A.L.Q..VA------------
Stu-t       ......N.SSQ..N..Q..KKA....F...L...NVK.A...A.L.Q..AA------------
Sly-t       ......N.SSQ..N..Q..KKA....F...L...NVK.A...A.L.Q..AP------------
Egr-t       ......N.SSQ..H..S..KKA....YV..L...NKN.G...A.I.QR.ASS-----------
Lja-t1      ......N.SSQ..N..Q..KRA....Y...I...NHRHM.S.A.I.Q..AANTTTT-------
Cca-t       ......T.SSQ..N..Q..KKA....Y...L...NVK.V...A.V.Q..TS------------
Egu-t       ......N.SSQ..K..E..KKA....Y...L...NIK.A...A.V.Q..NTAS----------
Csi-t1      ......N.SSQI.N..Q..KKA....Y...L..NNQK.G...S.L.Q..TTTTST--------
Jcu-t1      ......N.SAQ..N.MH..KKA....Y...L...NKK.C...S.I.Q..QL------------
Ccl-t1      ......N.SSQI.N..Q..KKA....Y...L..NNQK.G...S.L.Q..TTTTST--------
Gra-t       ......N.SSQ..K..H..KKA....Y...L...NKH.A...S.V.H..N-------------
Gar-t       ......N.SSQ..K..H..KKA....Y...L...NKH.A...S.V.H..N-------------
Atr-t       ......A.SSQ..R..E..KRT..S.Y...L...NQK.V...A.L.Q..SQA-----------
Peu-t       ......T.SSQ..Q..Q..KKA..D.Y...L...NKR.A...A.V.Q..SN-----------
Ptr-t       ......T.SSQ..E..H..KKA..D.Y...L...NKH.G...A.V.Q..GN-----------
Vvi-t       ......N.ASQ..N..H..KKT....Y...L...NEH.G...A.V.Q..TCS-----------
Sin-t       ......T.SSQ..K..Q..RQA....Y...L...NVK.A...A.L.QN.NN------------
Ppe-t       ......T.PSL..N..YQ.KTA....Y...M...NIN.G...A.I.Q..AAATITTNA-----
Mno-t       ......N.SSQ..N..R..KEA....Y...L...NKH.V...A.V.Q..TTTTTTTAA-----
Fve-t       ......I.PSL..N..Q..KKT....Y...V...NQN.G...A.I.Q..NIAATTSA------
Tca-t       ......NNSSQ..N..H..KKA....Y...L...SKH.A...S.V.Q..T-------------
Pmu-t       ......T.PSL..N..YQ.KTA....F...L...NIN.G...S.I.Q..TATAATTTNA----
Pbr-t       ......T.PSL..N..QQ.KTE....Y...M...NVN.G...A.I.Q..TATAAAATTTTADA
Gso-t       ......S.SSQ..K..H..RRV....Y...L...NQRHM.S.A.IVH..DITTTT--------
Nnu-t       ......SNSSQ..E..H..KKM....Y...L...NKR.A...A.V.Q..STASTTTNTACGCP
Mtr-t       ......N.SSQ..N..N..KKT....Y...L...NKRHM.A.A.V.H..TTTTTN--------
Gma-t       ....A.S.SSQ..K..H..RRV....Y...L...NQRHM.S.A.IVH..DITTTT--------
Pda-t       ......T.SAQ..Q..G..KKA....Y..ML..NSKR.C...S.V.H..STTSAANA------
Egi-t       ......T.SAQ..Q..D..KKT....Y..LL..NSKR.C...S.V.Q..SATSANA-------
Tha-t       ......S.SGQ.MR..E..RGA....Y...M...SSK.T..IS.L.H..QPQAVA--------
OsRbcS1     ......S.AAA..R.VE..RRE..D.F..LI...SSR.C...S.VVH..PSAAASPATVAGAE
```

Cluster M



```
                  10        20        30        40        50        60        70
          ....|....|....|....|....|....|....|....|....|....|....|....|....|....|
Consensus M VWP     K ETLSYLP L    L  E  YL    W PCLEF    G V R    SPGYYDGRYWTMW
Nta       .Q...PINKK.Y.......D.SQEQ.LS.VE..LKNG.V.....ETEH.F.Y.ENNK.............
Nsy       .Q...PINKK.Y.......D.SEEQ.LR.VE..LKNG.V.....ETEH.F.Y.ENNK.............
Nto       .Q...PINKK.Y.......D.SEEQ.LR.VE..LKNG.V.....ETEH.F.Y.ENNK.............
Stu       .Q...PINTK.Y.......D.TDEQ.LK.VE..LKNG.V.....ETEH.F.Y.ENHK.............
Sly       .Q...PINMK.Y.......D.SDEQ.LS.IE..LKNG.V.....ETEH.F.Y.ENHK.............
Egr       .Q...PIGKK.F.......P.SPES.AK.VD..LRSG.V.....ELEH.F.Y.EHHS.............
Lja-1     .Q...PIGKK.F.......P.TVEQ.AK.VE..IRKG.V.....ETEK.F.Y.ENHK.............
Cca       .Q...PTGKL.N..F....D.TDEQ.LK.ID..IRNG.I.....ELEK.H.Y.EYHR.............
Egu       .K...TVNNK.Y.......P.TREQ.IK.VD..LRNKLI..I..ELEH.F.Y.QHNS.............
Csi       .K...PTGLK.F.......P.SDEA.LK.IS..LRSG.I.....ELEK.W.Y.EHHS.............
Ccl       .K...PTGLK.F.......P.SDEA.LK.IS..IRSG.I.....ELEK.W.Y.EHHR.............
Jcu       .K...TQGLK.F.......P.TREQ.AK.VE..LRSG.V.....ELEH.F.Y.ENNR.............
Ath-1     .Q...PIGKK.F.......D.TDSE.AK.VD..IRNK.I..V..ELEH.F.Y.EHGN.............
Gra       .Q...PLGKK.F.......D.TPVQ.AK.VD..LRSK.I.....ELEE.F.H.KYSNL.T..........
Gar       .Q...PTGKK.F.......D.TPVQ.AK.VD..LRSK.V.....ELEE.F.H.KYSSL.T..........
Atr       .K..EPYNNQRF.......P.NKKA.AK.VD..IRNG.I.....D-VK.W.S.DNNR.............
Peu       .Q...PTGLK.F.......D.TEAE.AK.ID..LRSK.V.....ELEK.W.Y.ENHR.............
Ptr       .Q...PTGLK.F.......D.TEEE.AK.ID..LRSK.V.....ELEK.W.Y.EHHS.............
Vvi       .K...TEGLM.F.......P.NDEQ.CK.VD..LRMK.I.....TKLYPYPH.EHNR.............
Sin       IQ...TQGKK.F.......D.TREQ.AS.ID..IRSK.I..V..ELEH.F.Y.ENHR.............
Ppe       .K...TVGLK.F.......P.SPES.AK.VD..LRKN.V.....ELEK.F.Y.ENHR.............
Mno       .Q...PRGKL.F.......D.TDEQ.LK.ID..LRSN.I.....EVGKAHIY.ENNR.............
Fve       .Q...PVGLK.F.......P.TSES.AK.VDF.LRNK.V.....ELEK.F.Y.ENHK.............
Tca       .Q...PLGKK.F.......D.TRER.FK.VE..LRNK.V.....ELEH.F.Y.ENLRI.......C.I..
Pmu       .K...TVGLK.F.......P.SPES.AK.VD..LRKN.V.....ELEK.F.Y.ENHR.............
Pbr       .Q...PLGLK.F.......P.SPES.AK.VD..LRKN.V.....ELET.F.Y.ENHK.............
Gso       .Q...PIGKK.F.......D.DDAQ.AK.VE..LRKG.I.....ELEH.F.Y.EHNR.............
Nnu       .Q...PIAKK.F.......P.SPES.AK.VD..LRMG.I.....ELEH.F.Y.ENNR.............
Mtr       .Q...PIGKK.F.......P.TEDQ.AK.VE..IRKG.V.....ELEK.F.Y.ENHS.............
Gma       .Q...PVGKK.F.......D.DDAQ.AK.VE..LRKG.I.....ELEH.F.Y.EHNR.............
Pda       .K...IEGVK.F.......P.PIEA.LKQIE..LRSR.V.....C-KV.F.Y.ENHR.............
Egi       .K...IEGLK.F.......P.SIEA.LKQIE..LRSG.I.....C-KV.F.S.DNNR.............
Tha       .Q...TTGLR.F.......D.SDTE.AK.VD..LRNN.I.....ELGNAF.Y.ENHR.............
OsRbcS2   .Q...IEGIK.F.......P.TVED.LKQIE..LRSK.V.....S-KV.F.Y.ENHR.............

                  80        90       100       110       120
          ....|....|....|....|....|....|....|....|....|....|....|....
Consensus KLPMFGCTD  QV  E  E     P    RIIGFDN RQVQCISFIA KP     ----
Nta       .........AT..LA.VE.AKKAY.QAWI.......V..........Y..EGY------
Nsy       .........AT..LA.VE.AKKAY.QAWI.......V..........Y..EGY------
Nto       .........AT..LA.VE.AKKAY.QAWI.......V..........Y..EGY------
Stu       .........AT..LA.VQ.CKKSY.QAWI.......V..........Y..EGY------
Sly       .........AT..LA.VQ.AKKAY.QAWV.......V..........Y..EGY------
Egr       .........SA..LK.LE.AKKAY.TAFI.......K..........Y..PGF------
Lja-1     ......T..AS..LK.LE.AKKAY.NAFI.......K..........HR.EGY------
Cca       .........AT..LK.VR.CLKEY.NCWV.......V..........A..KGF------
Egu       .........AV..VN.LD.AIKAY.EGFI.....N.LK...D.....Y..AGY------
Csi       ....Y....AT..LK.VG.VQKEY.HSFV.......K........L.A..PGV------
Ccl       ....Y....AT..LK.VG.VQKEY.HSFV.......K..........A..PGV------
Jcu       .........AV..LN.LD.AIKAH.NGFV......SY..........Y..KGL------
Ath-1     ...L.....SA..LK.VE.CKKEY.NAFI.......T..........Y..PSFTG----
Gra       .........SA..LK.LE.CKKEY.NAFI.......V..........Y..EGY------
Gar       .........SA..LE.LE.CKKAY.NAFI.......V..........Y..PGF------
Atr       .........SS..MA.LE.ATKSY.DCYH.......K........L.Y..PPPPS----
Peu       ........EAS..LV.LQ.AKKAY.NAFI.......V..........A..QAA------
Ptr       ........EAS..LL.LE.AKKAY.NAFI.......T..........A..KGV------
Vvi       .........SA..IK.VQ.ARTAY.DAHI.......N......A...Y..S--------
Sin       ....Y....SA..LL.VD.AVRAH.DAFV.......V..........Y..PGY------
Ppe       .........SS..LK.LE.AKKAN.NSFI.......V..........Y..AGY------
Mno       .........AT..LA.VQ.TKKAY.DAHV.......N..........Y..PAK------
Fve       .........AS..LK.LE.AKKAY.NAFI.......V..........Y..AGY------
Tca       .........AS..LK.LD.AKKTY.TAFI.......R..........Y..PGYEN----
Pmu       .........SS..LK.LE.AKKAN.NSFI.......V..........Y..AGY------
Pbr       .........SS..LK.LE.AKKAY.NAFI.......V..........Y..AGY------
Gso       .........AS..LK.LQ.AKTAY.NGFI.......V..........Y..PGF------
Nnu       .........SS..LA.LE.AKKAY.NAFI.......K..........Y..PGF------
Mtr       ...L..A..SS..LK.LA.AKAAY.ESFI.......V..........HT.ATY------
Gma       .........AS..LK.LQ.AKTAY.NGFI.......V..........Y..PGX------
Pda       .......N.AT..AK.LE.CKKEY.DAFI.......V..........Y..PSFYT----
Egi       .......N.AT..AK.LE.CKKEY.DAFI.......V..........Y..PSY------
Tha       .........SA..MK.LQ.AKTTH.DAHI.......V..........Y..KAT------
OsRbcS2   .........AT..LK.LE.AKKAY.DAFV.......V....L.....Y..PGCEESGGN
```

Supplemental Figure S7. Alignment of the RbcS amino acid sequences. Amino acid sequences of the mature proteins from clusters T and M as listed in Table S2 were separately aligned using the ClustalW Multiple alignment, and a consensus sequence was generated for positions where at least 85% of the residues are conserved. Similar residues are in blue, identical residues are in black, and different residues are in red. Residues identical to the consensus sequence are indicated with a dot.