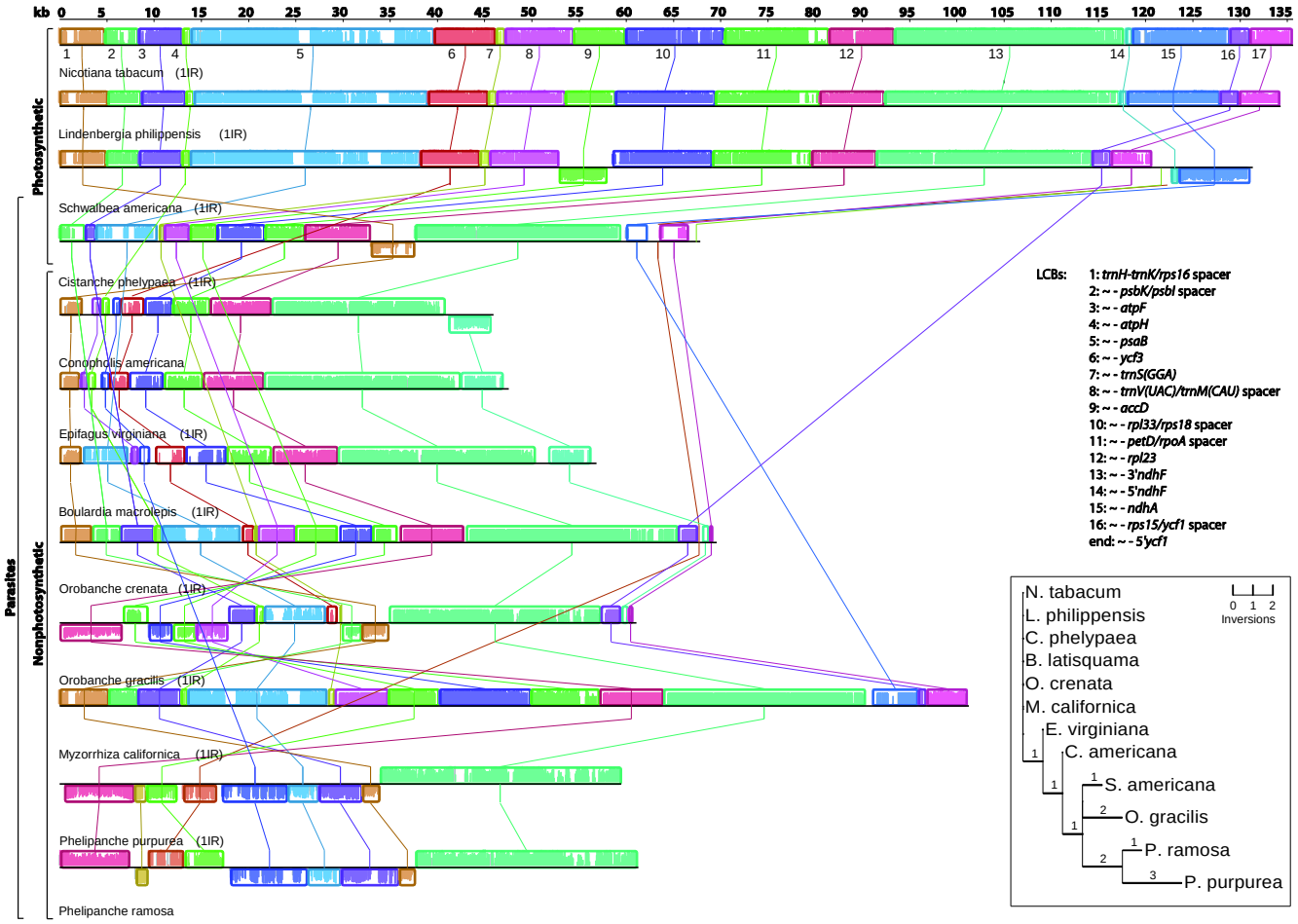
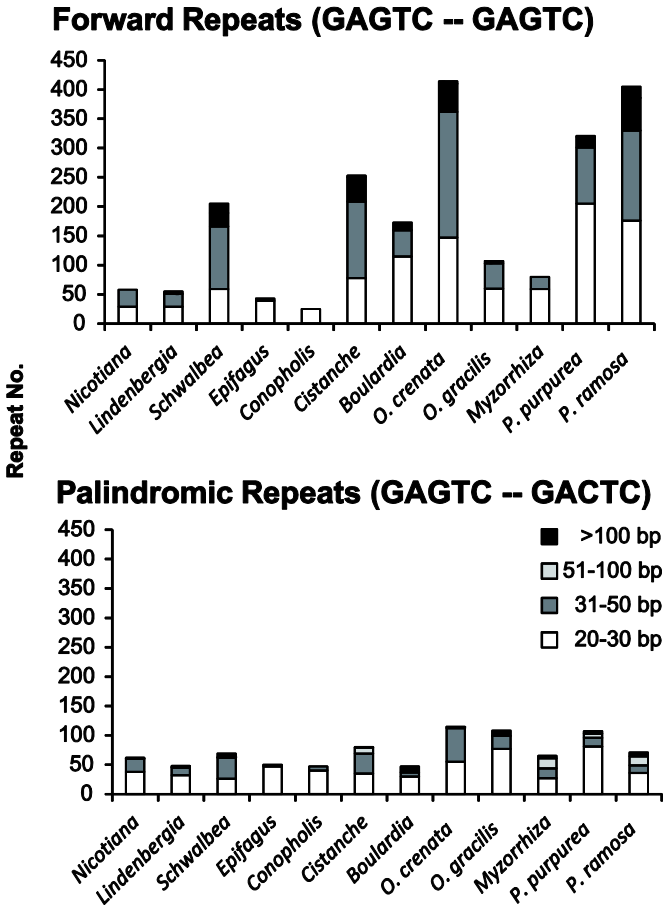
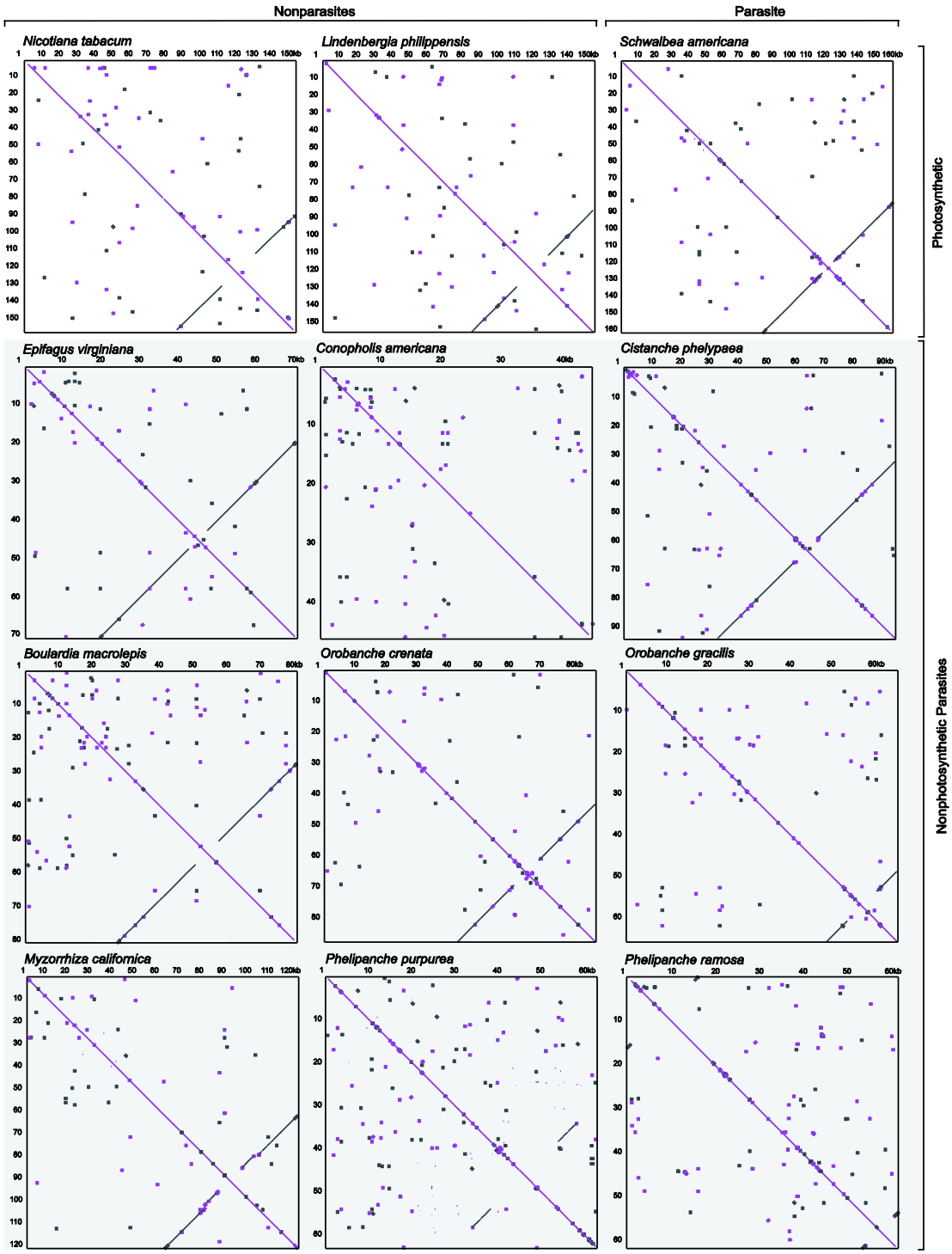


Supplemental Figure 1. Locally co-linear blocks (LCBs) inferred with progressiveMauve.

Numbers below the LCBs in *Nicotiana* denominate the border of each LCB. Brackets to the left provide lifestyle information for the different taxa. Where present, the large inverted repeat region B was removed, which is indicated in the species names by an '(1IR)'-suffix (see Material and Methods in the main text). The inset shows the inferred number of rearrangements/inversions.

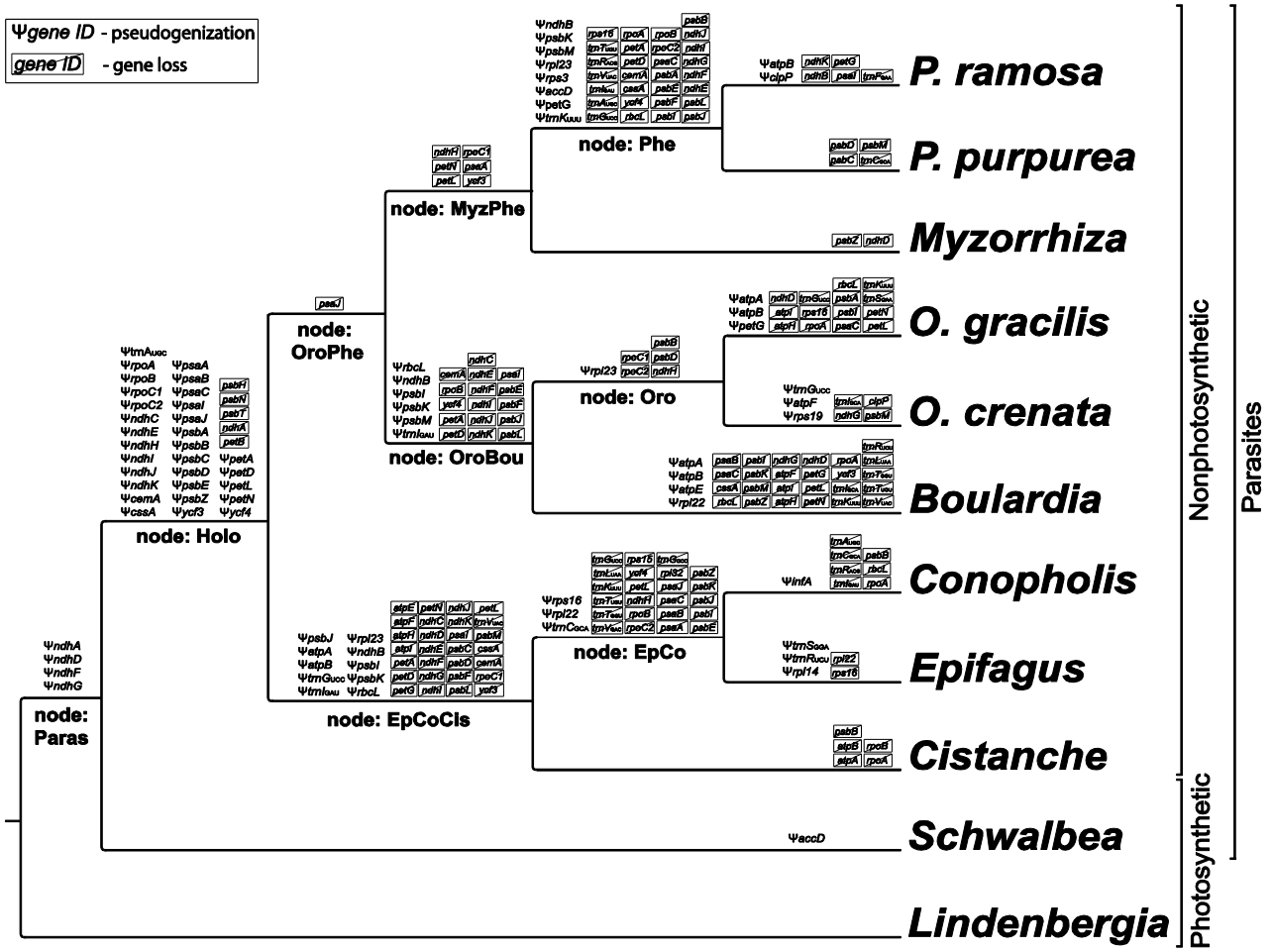


Supplemental Figure 2. Proportion and length of repetitive DNA in plastid genomes of Orobanchaceae.
The proportion of different size classes of forward and palindromic (reverse complement) repeats is illustrated for photosynthetic and nonphotosynthetic Orobanchaceae as well as tobacco as a closely related outgroup taxon. A definition, i.e., the direction, is given for each repeat type.



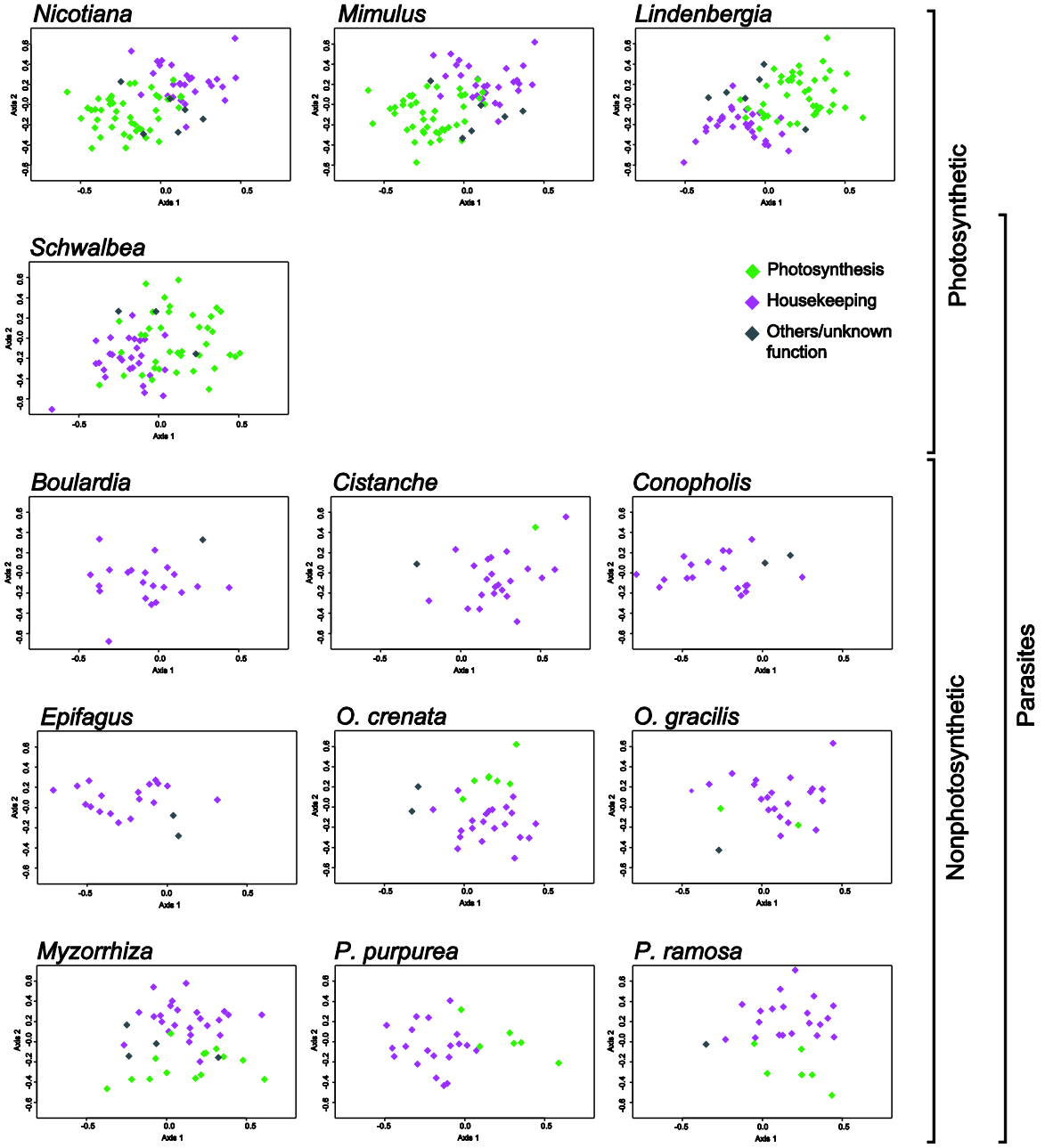
Supplemental Figure 3. Self-self dotplots of plastid genomes of Orobanchaceae and of *Nicotiana* illustrating the number and distribution of small and large plastid repeats.

Direct repeats are illustrated as purple dots or lines whereas inverted repeats are shown as dark gray dots or lines. Brackets on top indicate lifestyle for the three photosynthetic species (top row); nonphotosynthetic parasites are shown on light gray background.

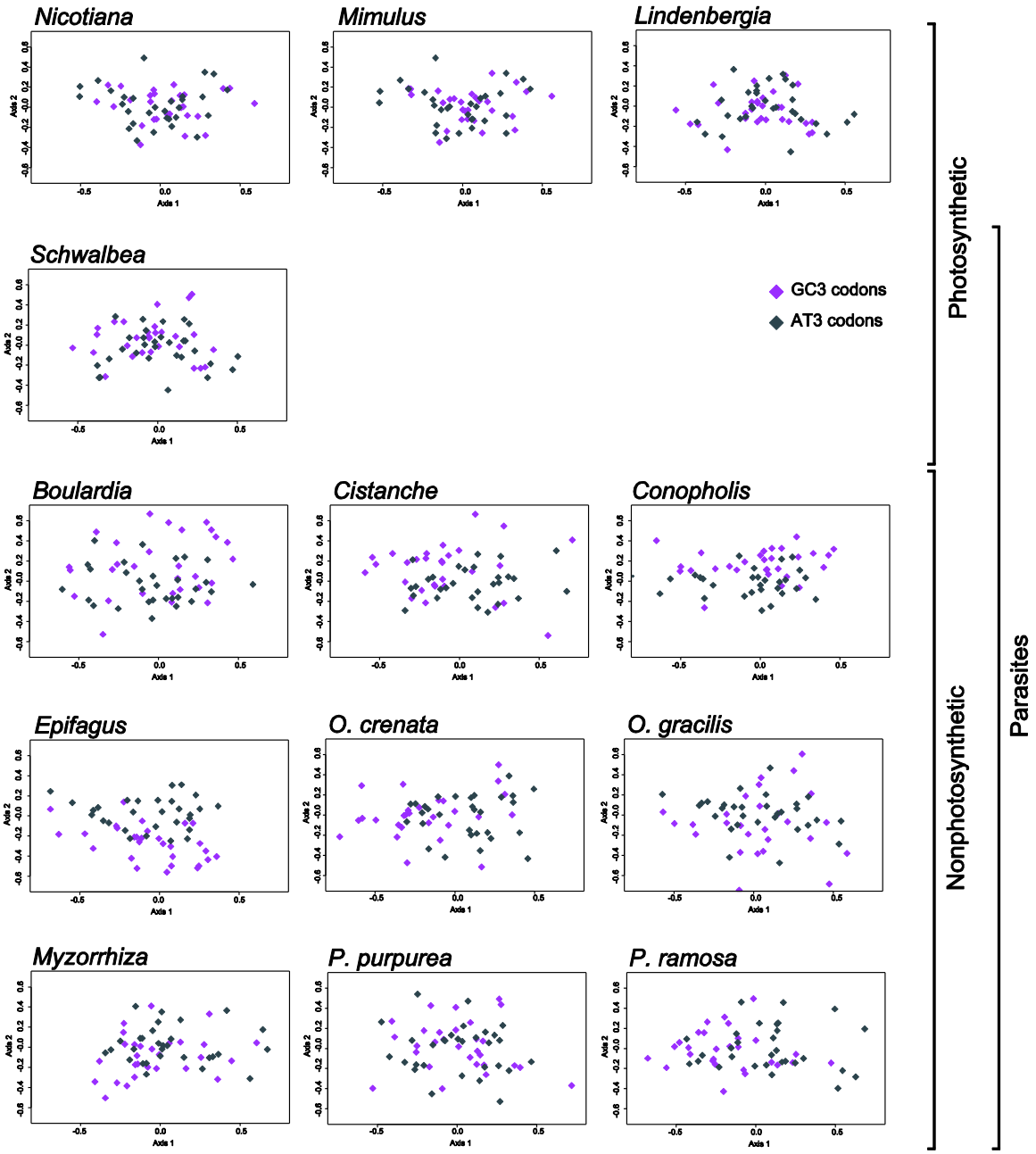


Supplemental Figure 4. Graphical summary of plastid gene losses as inferred by ancestral state reconstruction.

Ψ denotes the pseudogenization of a specified gene and a crossed box indicates the deletion of a specified genic region, without discriminating between loss-of function and loss-of-pseudogene deletion. Refer to Supplemental Tables S3 and S4 for additional information such as probability scores from the ancestral state reconstruction. Brackets to the right indicate different lifestyle.



Supplemental Figure 5. Results of coalescence analyses of codon usage (CoA-CU) for intact plastid genes.
Genes were grouped according to their main function, and CoA-CU was computed for all intact genes from four photosynthetic plants (upper two rows) and nine nonphotosynthetic plants. Brackets to the right the indicate different lifestyles. Genes involved in photosynthetic pathways are shown in green, genes of housekeeping function are in magenta, and gray dots indicate genes of other pathways or of a currently unknown function.



Supplemental Figure 6. Results of coalescence analyses of codon usage (CoA-CU) for 59 degenerated plastid codons.
CoA-CU was performed for codons from all intact genes from four photosynthetic plants (upper two row) and nine nonphotosynthetic plants (middle and bottom row). Brackets to the right indicate the different lifestyles. Codons ending in AT are illustrated as dark gray dots; and magenta indicates codons ending in GC.

Supplemental Table 1. Plant material used for plastome sequencing.

Taxon name	Source and voucher[a]	Sequencing method[b]	Accession number
Boulardia latisquama F.W. Schultz	Spain, Cap de la Nao (WU: S. Wicke s.n., 22. Apr. 2008)	FSS and FSP	pending review by ENA/EMBL
Cistanche phelypaea (L.) Cout.	Spain, between Murcia and Calasparra (WU: S. Wicke s.n., 25 Apr. 2008)	FSS	pending review by ENA/EMBL
Conopholis americana (L.) Wallr.	U.S.A., Pennsylvania (PAC: dePamphilis s.n.)	FSS	pending review by ENA/EMBL
Lindenbergia philippensis (Cham. & Schltdl.) Benth.	U.S.A., cultivated at PennState University (PAC: S. Wicke LP60/LP61, 29. Sept. 2009)	WGSP	pending review by ENA/EMBL
Myzorrhiza californica (Cham. & Schltdl.) Rydb.	U.S.A., cultivated at PennState University (PAC: S. Wicke Oc54, 18. Aug. 2009)	FSS	pending review by ENA/EMBL
Orobanche crenata Forssk.	Germany, cultivated on *Vicia faba* at the Botanical Garden Bonn (BONN: S. Wicke OC41)	FSS; WGSP	pending review by ENA/EMBL
Orobanche gracilis Sm.	Austria, Lower Austria, parasitizing *Chamaecytisus* sp. (WU: G.M. Schneeweiss 7)	WGSP	pending review by ENA/EMBL
Phelipanche purpurea (Jacq.) Soják	Germany, cultivated on *Achillea millefolium* at the Botanical Garden Bonn (BONN: S. Wicke Op38/39)	FSS and FSP; WGSP	pending review by ENA/EMBL
Phelipanche ramosa (L.) Pomel	Germany, cultivated on tomato at the Botanical Garden Bonn (BONN: S. Wicke Pr52/53)	FSS and FSP; WGSP	pending review by ENA/EMBL
Schwalbea americana L.	U.S.A., GA, Newton (PAC: S. Wicke & C.W. dePamphilis Sa57)	WGSP	pending review by ENA/EMBL

[a] herbarium acronyms according to Thiers (2013)

[b] WGSP – whole genome shotgun pyrosequencing, FSS – fosmid clone shotgun Sanger-sequencing, FPS – fosmid-clone pyrosequencing

Supplemental Table 2. Organization of transcription units in angiosperm plastid genomes.

Plastid gene or transcription unit	Trans-cription type	Model organism	Remarks and references
psbA	MC	*Nicta*	Sugita and Sugiura (1996) - dicistronic transcript in mustard described by Nickelsen and Link (1991)
trnK-matK	MC/DC	*Nicta*	Sugita et al. (1985)
rps16	MC	*Nicta*	Shinozaki et al. (1986)
psbK-I	PC	*Nicta*	Sugita and Sugiura (1996)
trnSgcu	MC	*Nicta*	Sugita and Sugiura (1996)
trnGucc	MC	*Nicta*	Sugita and Sugiura (1996)
rps2-atpA-F-H-I	PC	*Nicta* *Spinacia*	Ohto et al. (1988), Miyagi et al. (1998) - *atpH/I* also transcribed independently according to Sugita and Sugiura (1996)
trnCgca	MC	*Nicta* *Oryza*	Sugita and Sugiura (1996) Kanno and Hirai (1993)
rpoB-C1-C2	PC	*Nicta*	Shinozaki et al. (1986)
petN	MC	*Nicta* *Oryza*	Sugita and Sugiura (1996), Legen (2002) Kanno and Hirai (1993)
psbM	MC	*Nicta*	Wakasugi et al. (1992)
trnEuuc-Ygua-Dguc	PC	*Nicta*	Ohme et al. (1985)
trnTggu	MC	*Nicta*	Sugita and Sugiura (1996)
trnSuag	MC	*Spinacia*	Gruissem et al. (1986)
psbD-psbC-trnSuag-psbZ	PC	*Nicta*	Sugita and Sugiura (1996) - *psbC* also transcribed independently - $trnS_{UAG}$ transcribed independently
trnGgcc	MC	*Nicta*	Sugita and Sugiura (1996)
trnfMcau	MC	*Oryza*	Kanno and Hirai (1993)
rps14-psaA-psaB	PC	*Nicta*	Ohto et al. (1998)
trnSgga	MC	*Nicta* *Oryza*	Ohto et al. (1998) Kanno and Hirai (1993) - maybe co-transcribed with *ycf3*-operon
ycf3-trnSgaa-rps4-trnTugu	PC	*Nicta* *Oryza*	Ohto et al. (1998) Kanno and Hirai (1993) - unclear how many genes transcribed may include $trnS_{GGA}$
trnLuaa	MC	*Oryza*	Kanno and Hirai (1993) - maybe co-transcribed with $trnF_{GAA}$
trnFgaa	MC	*Oryza*	Kanno and Hirai (1993) - maybe co-transcribed with $trnL_{UUA}$
ndhC-K-J	PC	*Nicta*	Matsubayashi et al. (1987)
trnVuac	MC	*Nicta* *Oryza*	Sugita and Sugiura (1996) Kanno and Hirai (1993)
trnMcau	MC	*Oryza*	Kanno and Hirai (1993)
atpB-E	DC	*Nicta*	Sugita and Sugiura (1996)
rbcL	MC	*Nicta*	Sugita and Sugiura (1996)

Plastid gene or transcription unit	Trans-cription type	Model organism	Remarks and references
accD	MC	*Nicta*	Hajdukiewicz et al. (1997)
psaI-ycf4-cemA-petA	PC	*Nicta*	Shinozaki et al. (1986), Świątek (2002)
psbE-F-L-J	PC	*Nicta*	Sugita and Sugiura (1996)
trnWcca	MC	*Oryza*	Kanno and Hirai (1993)
trnPuug	MC	*Oryza*	Kanno and Hirai (1993)
petL-G-trnWcca-Pugg-psaJ-rpl33-rps18	PC	*Nicta*	Sugita and Sugiura (1996), Ohto et al. (1998)
clpP-rps12-rpl20	PC	*Nicta*	Ohto et al. (1998)
psbN	MC	*Nicta* *Oryza*	Wakasugi et al. (1992) Kanno and Hirai (1993)
psbB-T-N-H-petB-D	PC	*Nicta* *Oryza*	Sugita and Sugiura (1996) Kanno and Hirai (1993) - *psbN* also transcribed from own promoter
rpoA-rps11-rpl36-infA-rps8-rpl14-16-rps3-rpl22-rps19-rpl2-rpl23	PC	*Nicta*	Tanaka et al. (1986), Ohto et al. (1998)
trnIcau	MC	*Nicta*	Sugita and Sugiura (1996)
ycf2	MC	*Nicta*	Hajdukiewicz et al. (1997), Drescher et al. (2000)
trnLcaa	MC	*Oryza*	Kanno and Hirai (1993)
ndhB	MC	*Nicta*	Matsubayashi et al. (1987)
rps7-12	DC	*Nicta*	Ohto et al. (1998)
trnVgac-rrn16-trnIgau-Augc-rrn23-rrn4.5-rrn5	PC	*Nicta*	Shinozaki et al. (1986)
trnRacg	MC	*Oryza*	Kanno and Hirai (1993) - co-transcribed with rDNA operon in *Brassica napus* (Leal-Klevezas et al. 2000)
trnNguu	MC	*Oryza*	Kanno and Hirai (1993)
ndhF	MC	*Nicta*	Sugita and Sugiura (1996)
trnLuag	MC	*Nicta*	Sugita and Sugiura (1996)
rpl32-trnLuag-cssA	PC	*Nicta*	Sugita and Sugiura (1996)
rps15-ndhH-A-I-G-E-psaC-ndhD	PC	*Nicta*	Ohto et al. (1998)
ycf1	MC	*Nicta*	Drescher et al. (2000), Drescher (2003)

Abbreviations: MC – monocistronic, DC – dicistronic; PC – polycistronic; Nicta – *Nicotiana tabacum*.

Supplemental Table 3. Detailed results from multiple regression analyses.

Residuals									Coefficients			t-statistics		F-statistics			Breusch-Pagan test (heteroscedasticity)			AIC
Min	1Q	Median	3Q	Max	Std. error	df	$R^2_{mult.}$	$R^2_{adj.}$	Variable	Estimate	Std. error	t	Pr(>|t|)	F	df	p-value	BP	df	p-value	AIC
Regression model 1: survival time ~ distance (V1) + operon localization (V2) + strandedness (V3) + gene length (V4) + Intercept (I)																				
-5.329	-1.402	0.081	1.548	5.069	2.374	73	0.181	0.136	I	7.316	0.709	10.350	<0.001	4.027	4 & 73	0.005	8.281	4	0.082	363.1
									V1	-0.001	<0.001	-3.029	0.003							
									V2	-1.508	0.616	-2.448	0.017							
									V3	-0.109	0.583	-0.185	0.853							
									V4	<0.001	<0.001	0.387	0.700							
Regression model 2: survival time ~ distance (V1) +operon (V2) + strandedness (V3) + Intercept (I)																				
-5.472	-1.424	0.156	1.465	5.156	2.360	74	0.179	0.146	I	7.436	0.631	11.782	<0.001	5.381	3 & 74	0.002	8.936	3	0.030	361.2
									V1	<-0.001	<0.001	-3.022	0.004							
									V2	-1.501	0.612	-2.452	0.017							
									V3	-0.179	0.550	-0.325	0.746							
Regression model 3: survival time ~ distance (V1) + operon (V2) + gene length (V3) + Intercept (I)																				
-5.472	-1.425	0.156	1.465	5.156	2.358	74	0.180	0.147	I	7.247	0.597	12.147	<0.001	5.428	3 & 74	0.002	7.674	3	0.053	361.1
									V1	<-0.001	<0.001	-3.114	0.003							
									V2	-1.490	0.605	-2.465	0.016							
									V3	<0.001	<0.001	0.471	0.639							
Regression model 4: survival time ~ distance (V1) + operon (V2) + Intercept (I)																				
-5.354	-1.397	0.055	1.526	5.113	2.346	75	0.178	0.156	I	7.342	0.558	13.158	<0.001	8.115	2&75	0.001	8.248	2	0.016	359.3
									V1	<-0.001	<0.001	-3.098	0.003							
									V2	-1.467	0.599	-2.447	0.017							
Regression model 5: survival time ~ distance (V1) + gene length (V2) + Intercept (I)																				
-5.523	-1.580	0.172	1.812	6.236	2.437	75	0.113	0.089	I	6.229	0.445	13.998	<0.001	4.781	2&75	0.011	12.121	2	0.002	365.3
									V1	<-0.001	<0.001	-3.092	0.003							
									V2	<0.001	<0.001	0.265	0.792							
Regression model 6: survival time ~ distance (V1) + strandedness (V2) + Intercept (I)																				
-5.382	-1.594	0.170	1.807	6.252	2.438	75	0.112	0.089	I	6.274	0.430	14.574	<0.001	4.747	2&75	0.011	12.999	2	0.002	365.3
									V1	<-0.001	<0.001	-3.072	0.003							
									V2	0.053	0.560	0.094	0.925							
Regression model 7: survival time ~ operon (V1) + gene length (V2) + Intercept (I)																				
-6.623	-1.518	0.337	2.003	3.919	2.491	75	0.073	0.048	I	6.617	0.593	11.160	<0.001	2.952	2&75	0.058	0.648	2	0.723	368.7
									V1	-1.551	0.638	-2.430	0.018							
									V2	<0.001	0.000	0.158	0.875							
Regression model 8: survival time ~ operon (V1) + strandedness (V2) + Intercept (I)																				
-6.867	-1.693	0.231	2.106	3.744	2.485	75	0.078	0.053	I	6.867	0.634	10.828	<0.001	3.163	2&75	0.048	0.292	2	0.864	368.3
									V1	-1.611	0.643	-2.504	0.014							
									V2	-0.371	0.575	-0.645	0.521							
Regression model 9: survival time ~ distance (V1) + Intercept (I)																				
-5.406	-1.575	0.171	1.782	6.272	2.422	76	0.112	0.101	I	6.294	0.369	17.050	<0.001	9.610	1&76	0.003	12.416	1	0.000	363.3
									V1	<-0.001	<0.001	-3.100	0.003							
Regression model 10: survival time ~ operon (V1) + Intercept (I)																				
-6.655	-1.529	0.346	2.002	3.887	2.475	76	0.073	0.060	I	6.655	0.540	12.320	<0.001	5.956	1&76	0.017	0.556	1	0.456	366.7
									V1	-1.542	0.632	-2.441	0.017							
Regression model 11: survival time ~ gene length (V1) + Intercept (I)																				
-5.539	-1.884	0.168	2.127	3.490	2.571	76	<0.001	-0.013	I	5.541	0.406	13.632	<0.001	0.002	1&76	0.964	0.556	1	0.456	372.6
									V1	<-0.001	<0.001	-0.046	0.964							
Regression model 12: survival time ~ strandedness (V1) + Intercept (I)																				
-5.585	-1.841	0.081	2.081	3.547	2.570	76	0.001	-0.012	I	5.585	0.387	14.417	<0.001	0.050	1&76	0.823	0.243	1	0.622	372.6
									V1	-0.132	0.587	-0.225	0.823							

Supplemental Table 4. Results of pairwise Wilcoxon tests (*p*-values) evaluating differences in GC content and nucleotide composition of coding regions between non-parasites and parasites.

		Nic	*Mim*	*Lin*	*Sch*	*Epi*	*Con*	*Cis*	*Bou*	*Ocr*	*Ogr*	*Myz*	*Ppu*	*Pra*
GC	***Auc***	1.000	0.331	1.000	1.000	0.007	0.001	0.007	0.001	<0.001	0.001	0.007	<0.001	<0.001
	Nic	NA	0.034	1.000	1.000	0.006	0.001	0.002	0.001	<0.001	<0.001	0.004	<0.001	<0.001
	Mim	NA	NA	0.042	0.001	0.004	0.001	0.004	0.001	<0.001	<0.001	0.029	<0.001	<0.001
	Lin	NA	NA	NA	0.360	0.008	0.002	0.003	0.003	0.003	<0.001	0.085	0.003	0.003
	Sch	NA	NA	NA	NA	0.002	0.001	0.001	0.001	<0.001	<0.001	0.001	<0.001	<0.001
GC1	***Auc***	0.035	1.000	1.000	1.000	0.003	0.001	0.073	0.001	0.001	0.001	0.093	<0.001	<0.001
	Nic	NA	0.063	0.063	0.063	0.003	0.001	0.040	0.002	0.001	<0.001	0.008	<0.001	0.001
	Mim	NA	NA	0.811	0.811	0.001	0.001	0.007	0.001	0.001	0.001	0.028	<0.001	<0.001
	Lin	NA	NA	NA	0.375	0.001	<0.001	0.005	<0.001	<0.001	0.001	0.005	<0.001	<0.001
	Sch	NA	NA	NA	NA	0.004	<0.001	0.048	0.004	0.001	<0.001	0.048	<0.001	<0.001
GC2	***Auc***	0.309	0.309	0.309	0.097	0.001	0.002	0.003	0.003	<0.001	<0.001	0.010	<0.001	<0.001
	Nic	NA	1.000	1.000	1.000	0.003	0.003	0.004	0.005	<0.001	<0.001	0.059	<0.001	<0.001
	Mim	NA	NA	0.891	0.891	0.001	0.001	0.003	0.001	<0.001	<0.001	0.151	<0.001	<0.001
	Lin	NA	NA	NA	0.280	0.001	0.001	0.001	0.005	<0.001	<0.001	0.022	<0.001	<0.001
	Sch	NA	NA	NA	NA	0.004	0.003	0.030	0.030	<0.001	<0.001	0.174	<0.001	<0.001
GC3	***Auc***	0.309	0.309	0.309	0.097	0.001	0.002	0.003	0.003	<0.001	<0.001	0.010	<0.001	<0.001
	Nic	NA	<0.001	0.662	0.170	0.011	0.076	0.170	0.021	0.002	0.003	0.294	0.003	0.015
	Mim	NA	NA	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
	Lin	NA	NA	NA	0.041	0.002	0.005	0.046	0.003	0.001	0.007	0.087	0.001	0.003
	Sch	NA	NA	NA	NA	0.002	0.002	0.004	0.002	<0.001	0.001	0.002	<0.001	<0.001
A1	***Nic***	NA	0.048	0.072	0.040	0.002	0.002	0.088	0.008	0.006	0.006	0.072	<0.001	0.002
	Mim	NA	NA	0.260	0.547	0.002	0.002	0.157	0.006	0.010	0.002	0.207	<0.001	0.001
	Lin	NA	NA	NA	0.300	0.002	0.001	0.009	0.001	0.002	0.002	0.040	<0.001	<0.001
	Sch	NA	NA	NA	NA	0.011	0.004	0.016	0.011	0.003	0.002	0.165	<0.001	<0.001
A2	***Nic***	NA	1.000	1.000	1.000	0.153	0.067	0.512	0.722	0.004	0.012	1.000	<0.001	<0.001
	Mim	NA	NA	1.000	0.615	0.010	0.014	0.268	0.615	0.005	0.004	1.000	<0.001	<0.001
	Lin	NA	NA	NA	0.714	0.027	0.011	0.199	0.956	0.005	0.005	0.956	<0.001	<0.001
	Sch	NA	NA	NA	NA	0.562	0.764	1.000	1.000	0.002	0.014	1.000	0.004	0.004
A3	***Nic***	NA	0.742	1.000	0.126	0.024	0.434	0.411	0.104	0.031	0.008	1.000	0.009	0.015
	Mim	NA	NA	0.108	0.007	0.108	0.477	0.477	0.063	0.051	0.003	0.798	0.002	0.005
	Lin	NA	NA	NA	0.058	0.016	0.050	0.050	0.008	0.004	0.001	0.395	<0.001	0.001
	Sch	NA	NA	NA	NA	0.022	0.116	0.116	0.012	0.004	0.002	0.116	0.001	0.002
T1	***Nic***	NA	1.000	0.813	0.465	0.038	0.021	0.550	0.071	0.022	0.092	0.550	0.020	0.047
	Mim	NA	NA	0.337	0.767	0.007	0.007	0.119	0.012	0.003	0.024	0.119	0.003	0.007
	Lin	NA	NA	NA	0.631	0.015	0.016	0.577	0.070	0.015	0.034	1.000	0.020	0.066
	Sch	NA	NA	NA	NA	0.061	0.054	0.305	0.305	0.143	0.085	0.305	0.143	0.231
T2	***Nic***	NA	1.000	1.000	1.000	0.005	0.015	0.053	0.009	0.001	0.144	0.155	0.224	0.167
	Mim	NA	NA	1.000	1.000	0.003	0.024	0.010	0.003	0.003	0.029	0.153	0.124	0.024
	Lin	NA	NA	NA	0.961	0.004	0.018	0.020	0.003	0.002	0.020	0.069	0.096	0.038
	Sch	NA	NA	NA	NA	0.008	0.008	0.018	0.002	0.001	0.008	0.048	0.018	0.018

		Nic	*Mim*	*Lin*	*Sch*	*Epi*	*Con*	*Cis*	*Bou*	*Ocr*	*Ogr*	*Myz*	*Ppu*	*Pra*
T3	***Nic***	NA	0.247	0.757	0.757	0.027	0.357	0.757	0.357	0.005	0.238	0.757	0.247	0.357
	Mim	NA	NA	0.426	0.004	0.027	1.000	1.000	1.000	0.027	0.937	1.000	0.937	0.937
	Lin	NA	NA	NA	0.124	0.008	0.125	0.125	0.096	<0.001	0.037	0.125	0.125	0.125
	Sch	NA	NA	NA	NA	0.003	0.004	0.013	0.004	<0.001	0.004	0.013	0.005	0.013
C1	***Nic***	NA	0.256	0.583	1.000	0.020	0.013	0.256	0.025	0.026	0.044	0.004	<0.001	0.001
	Mim	NA	NA	0.084	0.337	0.007	0.005	0.337	0.009	0.337	0.078	0.034	<0.001	0.001
	Lin	NA	NA	NA	0.557	0.004	0.001	0.017	0.001	0.013	0.017	0.003	<0.001	<0.001
	Sch	NA	NA	NA	NA	0.015	0.015	0.056	0.049	0.049	0.015	0.015	<0.001	<0.001
C2	***Nic***	NA	1.000	1.000	0.346	0.005	0.005	0.069	0.209	0.037	0.041	0.219	0.006	0.004
	Mim	NA	NA	0.373	0.089	0.001	0.001	0.022	0.022	0.002	0.001	0.020	0.002	0.002
	Lin	NA	NA	NA	0.047	0.001	0.001	0.010	0.020	0.003	0.003	0.010	0.002	0.003
	Sch	NA	NA	NA	NA	0.002	0.004	0.399	0.322	0.006	0.009	0.399	0.001	0.003
C3	***Nic***	NA	0.136	0.909	0.909	0.148	0.178	0.909	0.909	0.012	0.025	0.498	0.001	0.178
	Mim	NA	NA	0.711	<0.001	0.131	0.421	1.000	1.000	0.015	0.112	1.000	0.001	0.421
	Lin	NA	NA	NA	0.034	0.010	0.034	0.039	0.193	0.001	0.002	0.193	0.001	0.016
	Sch	NA	NA	NA	NA	0.001	0.006	0.006	0.008	<0.001	0.002	0.006	<0.001	0.002
G1	***Nic***	NA	0.236	0.236	0.011	0.014	0.006	0.100	0.006	0.006	0.006	0.236	0.011	0.006
	Mim	NA	NA	0.463	0.016	0.004	0.005	0.005	0.001	0.001	0.001	0.088	0.001	0.001
	Lin	NA	NA	NA	0.137	0.006	0.004	0.009	0.003	0.002	0.003	0.349	0.002	0.002
	Sch	NA	NA	NA	NA	0.167	0.134	0.209	0.167	0.007	0.010	0.747	0.002	0.001
G2	***Nic***	NA	1.000	1.000	1.000	0.035	0.024	0.021	0.134	0.001	0.001	0.473	0.016	0.020
	Mim	NA	NA	1.000	1.000	0.038	0.026	0.045	0.005	0.003	<0.001	0.924	0.001	0.002
	Lin	NA	NA	NA	0.875	0.030	0.027	0.030	0.041	0.002	0.001	0.875	0.002	0.005
	Sch	NA	NA	NA	NA	0.099	0.050	0.062	0.099	0.001	0.001	0.555	0.001	0.001
G3	***Nic***	NA	1.000	1.000	0.173	0.030	0.087	0.354	0.011	0.019	0.031	1.000	0.034	0.022
	Mim	NA	NA	0.146	0.002	0.049	0.146	0.254	0.016	0.032	0.045	0.964	0.124	0.003
	Lin	NA	NA	NA	0.070	0.007	0.010	0.070	0.001	0.003	0.010	0.185	0.007	0.001
	Sch	NA	NA	NA	NA	0.010	0.008	0.067	0.004	0.001	0.003	0.067	0.004	0.001

Differences in GC content at the first, second, and third codon position and in nucleotide composition per codon were evaluated by pairwise Wilcoxon tests (with sequential Bonferroni correction) between photosynthetic plants and parasites. Abbr: Auc – *Aucuba japonica*, Nic – *Nicotiana tabacum*, Lin – *Lindenbergia philippensis*, Sch – *Schwalbea americana*, Epi – *Epifagus virginiana*, Con – *Conopholis americana*, Cis – *Cistanche phelypaea*, Bou – *Boulardia latisquama*, Ogr – *Orobanche gracilis*, Ocr – *O. crenata*, Myz – *Myzorrhiza californica*, Ppu – *Phelipanche purpurea*, Pra – *P. ramosa*; NA – not tested.

Supplemental Table 5. Results of unpaired Wilcoxon tests (p-values) evaluating differences in GC content of coding regions between nonparasites and parasites.

Gene-ID	Subset of nonparasitic taxa	Subset of parasitic taxa	GC1	GC2	GC3
accD	Mg Lp Am Oe Ab Nt Sl Ca No	Ev Co Bm Oc Og Mc	0.316	0.003	0.018
atpA	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Oc Mc Pp Pr	0.003	0.016	0.032
atpB	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Oc Mc Pp Pr	0.003	0.005	0.075
atpE	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Oc Og Mc Pp Pr	0.082	0.021	0.003
atpF	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Oc Og Mc Pp Pr	0.022	0.016	0.173
atpH	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Oc Mc Pp Pr	0.007	0.792	0.023
atpI	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Oc Mc Pp Pr	0.113	0.027	0.500
clpP	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Bm Oc Og Mc Pp	0.011	0.393	0.285
infA	Mg Lp Am Jn Oe Ca No	Sa Cp Bm Oc Mc Pp Pr	0.006	0.006	0.032
matK	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.001	0.001	0.040
rpl33	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Co Cp Bm Oc Og Mc Pp Pr	0.007	0.002	0.060
rpl16	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.003	0.000	0.520
rpl2	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.001	0.001	0.289
rpl20	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.006	0.001	0.075
rpl22	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Cp Oc Og Mc Pp Pr	0.525	0.807	0.007
rpl23	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Cp Oc Og Mc Pp Pr	0.090	0.419	0.297
rpl32	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Cp Bm Oc Og Mc Pp Pr	0.894	0.068	0.020
rpl33	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.016	0.233	0.049
rpl36	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.065	0.025	0.876
rps2	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.238	0.003	0.427
rps3	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc	0.001	0.005	0.021
rps4	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.023	0.001	0.059
rps7	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.004	0.001	0.025
rps8	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.002	0.006	0.005
rps11	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.001	0.001	0.210
rps12	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.027	0.267	0.405
rps14	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.124	0.646	0.003
rps15	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Cp Bm Oc Og Mc Pr	0.001	0.107	0.883
rps16	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Cp Oc Mc	0.027	0.777	0.009
rps18	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Oc Og Mc Pp Pr	0.001	0.305	0.239
rps19	Mg Lp Am Jn Oe Ab Nt Sl Ca No	Sa Ev Co Cp Bm Og Mc Pp Pr	0.189	0.003	0.024

P-values are sequential Bonferroni corrected. Taxon abbreviations: Mg – *Mimulus*, Lp – *Lindenbergia*, Am – *Antirrhinum*, Jn – *Jasminum*, Oe – *Olea*, Ab – *Atropa*, Nt – *Nicotiana*, Sl – *Solanum*, Ca – *Coffea*, No – *Nerium*, Sa – *Schwalbea*, Ev – *Epifagus*, Co – *Conopholis*, Cp – *Cistanche*, Bm – *Boulardia*, Oc – *Orobanche crenata*, Og – *O. gracilis*, Mc – *Myzorrhiza*, Pp – *Phelipanche purpurea*, Pr – *P. ramosa*

Supplemental Table 6. Codon usage in photosynthetic and nonphotosynthetic Orobanchaceae.

		Proportion of codon usage (in %)													
AA	**Codon**	***Nic***	***Mim***	***Lin***	***Sch***	***Epi***	***Con***	***Cis***	***Bou***	***Ocr***	***Ogr***	***Myz***	***Ppu***	***Pra***	**Remarks**
Phe	**UUU**	66.7	68.8	67.9	64.6	78.6	76.8	67.7	79.6	67.2	74.5	66.7	82.1	70.2	
	UUC	33.3	31.2	32.1	35.4	21.4	23.2	32.3	20.4	32.8	25.5	33.3	17.9	29.8	lost in *Pra*
Leu	**UUA**	32.5	32.5	31.7	29.8	35.6	34.7	27.5	38.0	32.7	38.6	30.1	41.7	36.4	lost/pseudogenized in *Epi, Con, Bou*
	UUG	20.1	19.9	19.7	22.2	17.5	17.8	23.0	20.2	21.8	23.0	19.9	20.2	18.8	
	CUU	21.6	21.3	21.3	20.3	21.1	19.5	22.4	18.4	21.8	17.8	22.8	18.4	20.4	
	CUC	6.9	5.9	5.7	6.8	6.3	5.7	5.5	4.0	4.3	2.2	6.1	2.7	5.7	
	CUA	12.1	13.1	13.8	12.8	10.7	12.6	12.8	9.4	12.0	10.6	12.5	9.4	11.6	
	CUG	6.2	6.5	6.8	7.4	5.6	6.2	6.7	6.5	5.6	4.4	6.9	4.9	5.0	
Ile	**AUU**	49.5	49.8	49.9	50.8	43.0	42.7	48.7	52.8	51.8	49.0	49.6	53.3	50.2	
	AUC	20.1	19.3	20.3	20.6	12.3	14.8	18.5	12.2	15.5	11.8	19.0	7.8	12.3	lost/pseudogenized in all holoparasites but *Myz*
	AUA	30.4	30.9	29.8	28.6	44.7	42.5	32.8	35.1	32.7	39.2	31.4	38.9	37.5	
Met	**AUG**	100	100	100	100	100	100	100	100	100	100	100	100	100	
Val	**GUU**	37.1	38.0	38.6	36.0	37.9	37.8	35.6	36.0	39.3	36.4	35.7	36.3	41.7	
	GUC	11.9	11.6	11.4	12.2	12.1	9.1	12.0	12.2	9.1	9.7	11.0	9.7	10.2	lost in *Epi, Con*
	GUA	38.1	37.7	37.1	37.5	38.2	35.7	34.1	37.6	40.1	41.0	40.7	42.3	35.5	lost/pseudogenized in all holoparasites but *Myz, Ocr*
	GUG	13.0	12.7	13.0	14.3	11.8	17.4	18.4	14.3	11.5	12.8	12.7	11.6	12.7	
Tyr	**UAU**	80.3	81.4	82.0	81.6	82.5	79.8	83.6	86.9	85.3	86.6	81.4	82.2	82.7	
	UAC	19.7	18.6	18.0	18.4	17.5	20.2	16.4	13.1	14.7	13.4	18.6	17.8	17.3	
His	**CAU**	77.2	77.5	77.9	77.4	74.7	80.2	79.3	80.2	79.2	80.6	77.1	78.8	77.4	
	CAC	22.8	22.5	22.1	22.6	25.3	19.8	20.7	19.8	20.8	19.4	22.9	21.3	22.6	
Gln	**CAA**	75.6	76.9	76.3	73.7	81.2	79.9	77.9	85.3	79.5	82.2	79.4	82.7	76.4	
	CAG	24.4	23.1	23.7	26.3	18.8	20.1	22.1	14.7	20.5	17.8	20.6	17.3	23.6	
Asn	**AAU**	77.0	78.7	77.4	76.4	80.3	79.6	74.7	75.9	81.3	83.0	78.5	79.4	78.0	
	AAC	23.0	21.3	22.6	23.6	19.7	20.4	25.3	24.1	18.7	17.0	21.5	20.6	22.0	
Lys	**AAA**	75.5	76.8	76.3	74.9	81.1	80.3	75.5	80.1	78.3	81.4	77.4	84.1	79.6	lost/pseudogenized in *Epi, Con, Bou, Ogr, Ppu, Pra*
	AAG	24.5	23.2	23.7	25.1	18.9	19.7	24.5	19.9	21.7	18.6	22.6	15.9	20.4	
Asp	**GAU**	79.6	81.2	81.1	78.7	80.2	82.4	82.0	79.8	83.2	76.7	80.4	83.4	82.3	
	GAC	20.4	18.8	18.9	21.3	19.8	17.6	18.0	20.2	16.8	23.3	19.6	16.6	17.7	
Glu	**GAA**	75.6	77.7	77.2	73.8	75.9	76.2	73.6	77.9	73.1	75.6	73.6	78.4	73.9	
	GAG	24.4	22.3	22.8	26.2	24.1	23.8	26.4	22.1	26.9	24.4	26.4	21.6	26.1	
Ser	**UCU**	29.7	29.2	29.4	29.1	26.1	22.1	27.5	26.7	30.1	31.5	27.6	34.2	31.3	
	UCC	15.1	14.7	15.7	17.1	13.4	14.8	17.1	13.9	15.4	14.1	16.8	12.1	13.0	pseudogenized in *Epi*
	UCA	19.5	19.1	19.1	17.5	24.3	25.7	21.3	19.4	21.5	20.4	19.8	17.9	21.0	
	UCG	9.4	9.8	9.5	10.5	8.9	9.1	9.8	5.9	9.1	5.6	9.6	8.6	8.5	
	AGU	20.7	21.7	21.2	20.2	23.0	23.4	19.5	29.5	20.5	25.6	21.6	24.6	23.9	
	AGC	5.8	5.5	5.2	5.5	4.3	4.8	4.8	4.5	3.3	3.0	4.5	2.6	2.3	
Pro	**CCU**	38.9	38.7	38.4	37.9	35.5	39.2	31.0	39.7	40.8	41.4	38.7	37.6	35.8	
	CCC	18.7	19.6	18.7	20.2	23.8	27.4	23.8	27.3	22.4	19.3	21.2	21.3	23.7	
	CCA	28.9	29.0	28.2	25.5	31.9	23.2	31.4	28.1	26.0	32.1	27.6	30.9	28.9	
	CCG	13.4	12.8	14.6	16.4	8.9	10.3	13.8	5.0	10.8	7.1	12.5	10.1	11.6	

		Proportion of codon usage (in %)													
AA	**Codon**	***Nic***	***Mim***	***Lin***	***Sch***	***Epi***	***Con***	***Cis***	***Bou***	***Ocr***	***Ogr***	***Myz***	***Ppu***	***Pra***	**Remarks**
Thr	**ACU**	39.3	42.1	40.6	39.7	38.4	36.8	38.1	38.0	40.8	38.1	39.6	43.4	37.7	
	ACC	19.6	18.3	18.6	21.2	18.7	18.4	18.8	22.9	17.6	20.3	20.2	20.6	23.0	lost in *Epi, Con, Bou*
	ACA	30.5	29.2	28.6	27.8	32.9	31.9	31.2	29.1	29.8	32.0	29.7	28.9	28.4	lost in *Epi, Con, Bou, Ppu, Pra*
	ACG	10.6	10.4	12.2	11.3	10.0	12.9	11.9	10.1	11.8	9.6	10.5	7.0	10.9	
Ala	**GCU**	44.7	44.6	44.2	41.9	36.7	37.7	37.7	33.5	43.9	38.0	40.5	40.8	40.8	
	GCC	17.2	15.4	17.0	17.7	16.7	19.4	17.2	18.8	15.8	17.6	19.3	18.7	21.7	
	GCA	28.0	28.2	27.4	27.4	36.3	35.6	35.8	38.2	29.8	36.6	28.0	32.0	29.6	lost/pseudogenized in all holoparasites but *Myz*
	GCG	10.0	11.9	11.4	12.9	10.4	7.3	9.3	9.4	10.5	7.9	12.2	8.5	8.0	
Cys	**UGU**	75.2	76.4	76.5	77.3	74.5	73.5	77.0	72.0	83.0	79.6	75.6	79.2	70.8	
	UGC	24.8	23.6	23.5	22.7	25.5	26.5	23.0	28.0	17.0	20.4	24.4	20.8	29.2	lost in *Epi, Con, Ppu*
Trp	**UGG**	100	100	100	100	100	100	100	100	100	100	100	100	100	
Arg	**CGU**	22.0	22.0	21.1	21.4	19.5	17.0	19.0	23.0	22.4	26.6	20.1	23.8	20.8	lost in *Con, Ppu, Pra*
	CGC	6.3	6.8	7.3	8.7	6.9	7.0	5.8	7.1	5.0	4.3	6.1	6.7	7.6	
	CGA	25.1	21.6	22.4	23.2	16.8	18.4	21.5	20.8	21.8	22.0	19.3	19.0	19.6	
	CGG	7.3	8.1	8.4	8.7	7.1	6.4	7.7	5.0	7.0	6.6	7.9	7.0	6.9	
	AGA	28.9	31.4	31.2	28.5	36.5	36.2	32.3	36.0	33.5	31.5	35.2	35.0	33.7	pseudogenized in *Epi*
	AGG	10.4	10.1	9.6	9.4	13.2	14.9	13.7	8.1	10.4	9.0	11.5	8.4	11.4	
Gly	**GGU**	32.3	32.7	32.5	32.2	33.4	30.8	28.8	30.6	32.8	32.6	31.4	36.3	33.1	
	GGC	11.7	10.8	10.5	11.5	7.3	5.9	7.5	9.9	8.5	8.4	8.5	8.9	10.4	
	GGA	38.8	39.3	39.3	37.6	40.7	40.3	41.8	41.9	39.8	39.6	40.3	40.9	39.5	lost/pseudogenized in *Epi, Con, Cis, Bou, Ocr*
	GGG	17.1	17.1	17.7	18.7	18.5	23.0	21.9	17.6	18.9	19.4	19.8	13.9	17.1	
TER	**UAA**	53.8	61.5	53.1	47.3	57.1	63.3	64.5	82.6	70.6	73.1	54.2	75.0	76.7	
	UAG	23.1	17.9	24.7	31.1	28.6	30.0	22.6	17.4	14.7	23.1	20.8	15.6	10.0	
	UGA	23.1	20.5	22.2	21.6	14.3	6.7	12.9	0.0	14.7	3.8	25.0	9.4	13.3	

The proportion (in %) of used codons for all amino acids (AA) and stop codons is summarized for photosynthetic and nonphotosynthetic Orobanchaceae, and for *Nicotiana*. The preferred codon is highlighted in blue. Abbreviations: Nic – *Nicotiana tabacum*, Lin – *Lindenbergia philippensis*, Sch – *Schwalbea americana*, Epi – *Epifagus virginiana*, Con – *Conopholis americana*, Cis – *Cistanche phelypaea*, Bou – *Boulardia latisquama*, Ogr – *Orobanche gracilis*, Ocr – *O. crenata*, Myz – *Myzorrhiza californica*, Ppu – *Phelipanche purpurea*, Pra – *P. ramosa*; NA – not tested; TER – stop.

Supplemental Method 1. Details regarding gene- and species-specific annotation

The classification of genes as putatively functional or as pseudogenes is based solely on evidence from the DNA level and must therefore be treated as preliminary. As of writing this manuscript, expression of genes is still validated in our labs.

Schwalbea: ycf1 shows extreme sequence divergence and was therefore excluded from codon usage analyses (CU); sequencing/assembly error is likely, because a frameshift occurs after a series of homopolymer stretches. *ccsA* contains putative sequencing/assembly-errors and was therefore excluded from CU. *accD* may use an alternative start codon (expression analysis is underway).

Epifagus: If *matK* is annotated as in Young and dePamphilis (2000) it would contain 37 stop codons; if it is annotated as in Wolfe et al. (1992) no stop codons are present, but the ORF is 189 bp shorter than in other clades of holoparasitic Orobanchaceae; we used the latter annotation for CU analysis.

Conopholis: If *matK* is annotated as in Young and dePamphilis (2000) it would contain 37 stop codons; if it is annotated as in Wolfe et al. (1992) it contains 6 premature stop codons accumulating towards the 3'-end after a homopolymer stretch, but the ORF is 189 bp shorter than in other clades of holoparasitic Orobanchaceae; we used the latter annotation for CU analysis. The inferred premature stop codons may be due to sequencing/assembly errors; amino acid sequence is highly similar to that of *Epifagus*.

Cistanche: If *matK* is annotated as in Young and dePamphilis (2000) it is 189 bp longer than if annotated as in Wolfe et al. (1992); we used the former annotation for CU. *matK* contains 5 premature stops accumulating towards the 3'-end, which possibly is due to sequencing/assembly errors after several homopolymer stretches (similar to the situation in *Schwalbea*).

Boulardia: *ycf1* may be a pseudogene, because it has several premature stop codons and shows numerous deletions; however, potential sequencing/assembly errors in homopolymer regions cannot be entirely excluded, even after Sanger re-sequencing.

Myzorrhiza: *rpl22* may be a pseudogene due to large indels and a high substitution rate, it was assigned functional in ancestral state reconstruction and was included in CU.

Phelipanche purpurea: *atpA* contains 3 untranslatable codons (because of ambiguous nucleotide calls); *atpH* contains 1 untranslatable codon; *matK* contains 3 untranslatable codons. *ycf1* is fragmented with unclear gene start and gene stop and it contains several frameshifts or stop codons in those regions that can be unambiguously identified as *ycf1* fragments (maybe a pseudogene).

Phelipanche ramosa: *matK* contains 1 premature stop and 3 untranslatable codons; *atpF* has an unclear stop codon, because it contains 3 potential stops in a row: current annotation assumes the first stop codon as the true one; *rpl2* contains 4 premature stops; *ycf1*: see *P. purpurea* for details; expression analyses for *clpP* and *accD* is underway.

Supplemental Method 2. Details of experimental procedures used for plastid genome sequencing.

Fosmid library construction and library sorting

DNA extraction and purification (with subsequent RNA digest) was performed following the CTAB-based protocol of McNeal et al. (2006) using fresh and young flower tissue. Representative fosmid libraries were constructed for *Boulardia latisquama*, *Cistanche phelypaea, Conopholis americana, Myzorrhiza californica, Orobanche crenata, Phelipanche purpurea* using the CopyControl™ (HTP) Fosmid Library Production Kit (EPICENTRE® Biotechnologies) following the manufacturer's instructions with slight modifications. Five to 10 µg of freshly extracted total genomic and unsheared DNA was size selected by electrophoretic separation on a 30 cm long, 1% low melting point agarose gel. The gel was run overnight for a minimum of 16 hours at 8–12°C in freshly prepared 1× TAE buffer at 65 V. The gel was post-stained with GelStar® Nucleic Acid Gel Stain (Lonza), and the DNA was detected on a blue-light transilluminator. DNA fragments larger than undigested lambda DNA were excised using sterilized cover slips. Gel extraction and DNA purification was carried out as instructed in the Fosmid Library Production Kit using agarase and ethanol/sodium acetate precipitation. At least 0.5–1 µg of pure genomic DNA were ligated into the fosmid vector pcc1FOS or pcc2FOS, respectively, for a minimum of 2 hours at room temperature followed by a subsequent overnight incubation at 4–8°C. After inactivation of ligase (10–15 min at 71°C), the vector-DNA concatemers were directly used for phage packaging and transduction. Size selection of genomic DNA was omitted for *Myzorrhiza*, where purified DNA was directly ligated into pcc1FOS. Selection and titering of fosmid clones was carried out using LB media and agar supplemented with 12.5 µg/ml chloramphenicol and 10 µg/ml cycloheximid. Between 2,000 and 3,000 fosmid clones were plated and grown on 24×24 cm LB-agar trays for 8–10 hours at 37°C. Libraries were sorted into 384-well plates filled with LB-freezing medium (Sambrook and Russell 2001) using a QPIX colony picker. Fosmid clones were redundantly arrayed on 22×22 cm positively-charged nylon membranes (Performa Nylon Filters, Genetix Ltd.) using a gridding robot (QPIX II, MicroGrid II) in 3×3, 4×4, or 5×5 offset double spotting pattern. Colony lysis/denaturation, neutralization, and fixation of DNA onto the filters were performed as suggested by the membrane manufacturer (Genetix Ltd.).

Fosmid library screening, probe preparation, end-sequencing

Plastid probes of all protein-coding genes and of selected tRNA-gene regions known to be present in the plastome of *Epifagus virginiana* have been PCR-amplified from *Nicotiana tabacum* using custom primers (available upon request). Probes were designed to be 0.2 to 1.5 kb long. GoTaq®Flexi polymerase system (Promega) was employed for PCR with reactions (25 µl) typically containing 1× Flexi reaction buffer, 20 mM MgCl, 0.1 M betaine, 0.20 mM of each dNTP, 10 mM of each amplification primer, 0.1 U *Taq* polymerase, and 10–20 ng of template DNA. Cycling conditions were as follows: 3 min of pre-denaturation; 35 cycles with 30 s of denaturation, 20 s primer annealing at 50–53° C (i.e. 2–3° below calculated primer annealing temperatures), 60–150 s (according to the length of the expected product: 60 s per 1 kb) elongation at 68° C; 10 min of final elongation at 72° C. All probes were agarose gel-purified, and were sequenced at Macrogen Inc. (Seoul, South Korea) prior to fosmid library screening. Southern hybridizations of fosmid filters with plastid gene probes were performed following Sambrook and Russell (2001) with some modifications: Filters were pre-washed in 6× SSPE at 50°C for 30 min and subsequently pre-hybridized at 60 °C overnight in hybridization-buffer (5× SSPE + 5× Denhardt's solution + 0.2 % SDS) containing 10 µg/ml sheared and denatured herring sperm DNA. The probe cocktail (100 ng of 4–6 plastid gene probes mixed at equimolar ratio) was radiolabeled with ^{32}P-dATP using the Prime-a-Gene® Labeling System (Promega); labeling time was extended to two hours. Labeled probes were purified with custom Sephadex G-50 Superfine columns, eluting the samples stepwise with 100 µl 0.5x TE-buffer. The first three consecutive fractions with the highest emission of radiation were pooled, and denatured at 99°C for 10 minutes. Hybridization was carried out overnight at 61° C in fresh hybridization

medium. Filters were washed twice with pre-warmed washing buffer (2x SSC + 0.2 % SDS) at 61°C for 10 min each, followed by a third 5 min washing step with room-temperate washing buffer. Subsequently filters were briefly rinsed with 6x SSPE. Detection of positive signals was performed using a Typhoon 9200 Phosphoimager (GE Healthcare). Positive clones were prepped via alkali lysis (Sambrook and Russell 2001) or using the QIAprep Spin Miniprep Kit (Qiagen) following the manufacturer's instructions. End-sequencing of potential plastid-DNA carrying fosmids was performed at Macrogen (Seoul, South Korea) or at GATC Biotech (Konstanz, Germany).

Shotgun Sanger sequencing and pyrosequencing

Between three and five fosmids were selected for shotgun sequencing. Enriched fosmid DNA was isolated and purified using the NucleoBond® Xtra Midi Kit (Macherey-Nagel). Between 3 and 5 µg of freshly eluted fosmid DNA were precipitated with isopropanol, briefly washed twice with 70% ethanol, and resolved in 1 ml shearing buffer (0.5× TE, pH 8.3 + 10 % glycerol). Fosmid DNA was sheared to fragments of 2–3 kb length. Sheared DNA was precipitated by NaCl/ethanol precipitation, washed twice, and resolved in 50 µl 10 mM Tris-HCl (pH 8.0). DNA was end-repaired using the NEBNext® End Repair Module (New England Biolabs Inc.); DNA was additionally A-tailed if subsequent A/T-cloning was used. Subcloning libraries of each fosmid were produced using either the CloneJET™ PCR Cloning Kit (Fermentas) or the pGEM®-T Easy Vector System I (Promega) in a 3:1 ratio of DNA to cloning vector. Positive clones were sorted into 384-well plates of LB+10% glycerol medium supplemented with the corresponding antibiotics. A minimum of two of these plates were sequenced bidirectionally via Sanger-sequencing at Macrogen (Seoul, South Korea). Alternatively, fosmid DNA was tagged and shotgun-pyrosequenced (12 fosmids in $1/8^{th}$ of a picotiter plate) at the Center for Medical Research, Medical University Graz, Austria. In addition to the fosmid-based approach, plastid un-enriched DNA of *L. philippensis*, and of *S. americana* was 454-pyrosequenced at the Center for Medical Research, Medical University Graz, Austria employing standard GS FLX Titanium series' protocols. For shotgun-pyrosequencing, total DNA was extracted as above. After complete resuspension in 10 mM Tris-buffer, DNA was column-purified using the NucleoSpin® gDNA Clean-up kit (Macherey-Nagel).

Supplemental References

Bennett, J.R., Mathews, S. (2006). Phylogeny of the parasitic plant family Orobanchaceae inferred from phytochrome A. Am. J. Bot. **93:** 1039−1051.

Drescher, A. (2003). *ycf1, ycf14 und RNA-Edierung: Untersuchungen an im Lauf der Plastidenevolution neu hinzu gewonnenen Genen und Eigenschaften.* Dissertation, Ludwig-Maximilian-Universität, München, Germany.

Drescher, A., Ruf, S., Calsa, T., Carrer, H., Bock, R. (2000). The two largest chloroplast genome-encoded open reading frames of higher plants are essential genes. Plant J. **22:** 97−104.

Hajdukiewicz, P.T.J., Allison, L.A., Maliga, P. 1997. The two RNA polymerases encoded by the nuclear and the plastid compartments transcribe distinct groups of genes in tobacco plastids. EMBO J. **16:** 4041−4048.

Kanno, A., Hirai, A. (1993). A transcription map of the chloroplast genome from rice (*Oryza sativa*). Curr. Genet. **23:** 166−174.

Legen, J. (2002). *Gene expression in plastids of higher plants: evolutionary and functional aspects of different RNA polymerases – coordinated assembly of multiprotein complexes.* Dissertation, Ludwig-Maximilian-Universität, München, Germany.

Leal-Klevezas, D.S., Martínez-Soriano, J.P., Nazar, R.N. (2000). Cotranscription of 5S rRNA-tRNA-$Arg^{(ACG)}$ from *Brassica napus* chloroplasts and processing of their intergenic spacer. Gene **253:** 303−311.

Matsubayashi, T., et al. (1987). Six chloroplast genes (*ndhA-F*) homologous to human mitochondrial genes encoding components of the respiratory chain NADH dehydrogenase are actively expressed: determination of the splice sites in *ndhA* and *ndhB* pre-mRNAs. Mol. Gen. Genet. **210:** 385−393.

McNeal, J.R., Leebens-Mack, J.H., Arumuganathan, K., Kuehl, J.V., Boore, J.L., dePamphilis, C.W. (2006). Using partial genomic fosmid libraries for sequencing complete organellar genomes. Biotechniques **41:** 69–73.

Miyagi, T., Kapoor, S., Sugita, M., Sugiura, M. (1998). Transcript analysis of the tobacco plastid operon *rps2/atpI/H/F/A* reveals the existence of a non-consensus type II (NCII) promoter upstream of the *atpI* coding sequence. Mol. Gen. Genet. **257:** 299−307.

Nickelsen, J., Link, G. (1991). RNA-protein interactions at transcript 3′ ends and evidence for *trnK-psbA* cotranscription in mustard chloroplasts. Mol. Gen. Genet. **228:** 89−96.

Ohme, M., Kamogashira, T., Shinozaki, K., Sugiura, M. (1985). Structure and cotranscription of tobacco chloroplast genes for $tRNA^{Glu(UUC)}$, $tRNA^{Tyr(GUA)}$ and $tRNA^{Asp(GUC)}$. Nucl. Acids Res. **13:** 1045−1056.

Ohto, C., Torazawa, K., Tanaka, M., Shinozaki, K., Sugiura, M. (1988). Transcription of ten ribosomal protein genes from tobacco chloroplasts: A compilation of ribosomal protein genes found in the tobacco chloroplast genome. Plant Mol. Biol. **11:** 589−600.

Sambrook, J., Russell, D.W. (2001). *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, N.Y.

Shinozaki, K., et al. (1986). The complete nucleotide sequence of the tobacco chloroplast genome: its gene organization and expression. EMBO J. **5:** 2043−2049.

Sugita, M., Sugiura, M. (1996). Regulation of gene expression in chloroplasts of higher plants. Plant Mol. Biol. **32:** 315−326.

Sugita, M., Shinozaki, K., Sugiura, M. (1985). Tobacco chloroplast $tRNA^{Lys(UUU)}$ gene contains a 2.5-kilobase-pair intron: An open reading frame and a conserved boundary sequence in the intron. Proc. Natl. Acad. Sci. USA **82:** 3557−3561.

Świątek, M. (2002). *Functional analysis of plastid-encoded genes. Application of reverse genetics on Nicotiana tabacum.* Dissertation, Ludwig-Maximilian-Universität, München, Germany.

Tanaka, M., Obokata, J., Chunwongse, J., Shinozaki, K., Sugiura, M. (1987). Rapid splicing and stepwise processing of a transcript from the *psbB* operon in tobacco chloroplasts - Determination of the intron sites in *petB* and *petD*. Mol. Gen. Genet. **209:** 427−431.

Thiers, B. (2013). Index Herbariorum: A global directory of public herbaria and associated staff. New York Botanical Garden's Virtual Herbarium. http://sweetgum.nybg.org/ih/

Wakasugi, T., Meng, B.-Y., Matsubayashi, T., Vera, A., Torazawa, K., Sugiura, M. (1992). Transcription map of the tobacco chloroplast genome. In *Research in Photosynthesis, Proceedings of the IXth International Congress on Photosynthesis, Nagoya, Japan, August 30-September 4, 1992,* N. Murata (ed.). pp. 263−266. Kluwer Academic Publishers, Dordrecht, The Netherlands.

Wolfe, K.H., Morden, C.W., Palmer, J.D. (1992). Function and evolution of a minimal plastid genome from a nonphotosynthetic parasitic plant. Proc. Natl. Acad. Sci. USA **89:** 10648–10652.

Young, N.D., dePamphilis, C.W. (2000). Purifying selection detected in the plastid gene matK and flanking ribozyme regions within a group II intron of nonphotosynthetic plants. Mol. Biol. Evol. **17:** 1933–1941.