

Supplemental Table S1. The distribution of internal stop codons

Species	Codon	Non-intron gene	One-intron gene			Two-intron gene			Total
			Exon1	Exon2	Total	Exon1	Exon2	Exon3	
<i>S. uncinata</i>	TAA	0	0	0	0	0	0	0	0
	TGA	5	0	0	0	0	0	0	5
	TAG	1	0	0	0	0	0	0	1
	NNG/NNAs	0.20	-	-	-	-	-	-	0.20
<i>C. lindheimeri</i>	TAA	25	1	1	2	0	0	0	27
	TGA	8	1	1	2	0	0	0	10
	TAG	1	0	0	0	0	0	0	1
	NNG/NNAs	0.03	0.00	0.00	0.00	-	-	-	0.03
<i>A. spinulosa</i>	TAA	42	1	0	1	0	1	0	44
	TGA	19	2	2	4	0	0	0	23
	TAG	6	0	0	0	0	0	0	6
	NNG/NNAs	0.10	0.00	0.00	0.00	-	0.00	-	0.00

Supplemental Table S2. The statistic analysis of SCUB frequencies

Taxonomy	Species	SCUB based on amino acids			SCUB based on NNA/T and NNC/G		
		Mean	CV	P value	NNA/T	NNC/G	P value
Archaeplastida	<i>P. haitanensis</i>	0.250	0.329	5.23E-18	39837	8885	0
	<i>G. taiwanensis</i>	0.220	0.236	1.15E-21	39775	8405	0
	<i>C. caldarium</i>	0.295	0.269	8.03E-18	35306	9914	0
Chlorophyta	<i>O. viridis</i>	0.420	1.021	2.39E-05	18855	5322	0
	<i>O. cardiacum</i>	0.183	0.350	1.72E-20	32024	5482	0
	<i>P. minor</i>	0.329	0.977	8.71E-08	15661	3517	0
Charophyta	<i>M. viride</i>	0.117	0.435	9.30E-23	21811	2519	0
	<i>Z. circumcarinatum</i>	0.244	0.341	5.14E-18	20989	5002	0
	<i>C. globosum</i>	0.157	0.319	1.65E-22	21706	3140	0
Bryophyte	<i>M. polymorpha</i>	0.106	0.417	7.08E-24	20341	1911	0
	<i>S. ruralis</i>	0.116	0.411	3.04E-23	17254	1776	0
	<i>P. patens</i>	0.119	0.397	2.83E-23	19973	2124	0
Pteridophyte	<i>S. moellendorffii</i>	0.748	0.330	0.0004	12290	9134	4.1E-103
	<i>S. uncinata</i>	0.918	0.327	0.2628	11894	10675	4.889E-16
	<i>C. lindheimeri</i>	0.577	0.264	1.31E-09	16228	9008	0
	<i>A. spinulosa</i>	0.472	0.268	2.05E-12	17373	8111	0
Gymnosperms	<i>C. oliveri</i>	0.318	0.286	1.39E-16	18176	5866	0
	<i>G. parvifolium</i>	0.391	0.296	5.00E-14	14018	5394	0
	<i>G. biloba</i>	0.374	0.387	1.21E-12	15267	5684	0
Monocotyledon	<i>S. bicolor</i>	0.381	0.290	1.69E-14	13712	5224	0
	<i>Z. mays</i>	0.405	0.297	1.37E-13	15193	6187	0
	<i>O. sativa</i>	0.406	0.242	5.18E-15	14889	6113	0
Dicotyledon	<i>A. thaliana</i>	0.349	0.267	4.51E-16	18543	6530	0
	<i>G. max</i>	0.325	0.287	2.62E-16	18477	6049	0
	<i>P. trichocarpa</i>	0.370	0.265	1.78E-15	19156	7156	0

Supplemental Table S3. The comparison on the ratios of NNC/NNG to NNA/NNT with the ratios of C and G to A and T in the gene body, intron, and whole genome sequences

Taxomony	Species	CDS NNC/G to NNA/T	Gene			Intron			Genome		
			CG to AT	P value (a)	P value (b)	CG to AT	P value (a)	P value (b)	CG to AT	P value (a)	P value (b)
Archaeplastida	<i>P. haitanensis</i>	0.223	0.511	0.000	0.000	-	-	-	0.492	0.000	0.000
	<i>G. taiwanensis</i>	0.211	0.467	0.000	0.000	-	-	-	0.440	0.000	0.000
	<i>C. caldarium</i>	0.281	0.497	0.000	0.000	-	-	-	0.487	0.000	0.000
Chlorophyta	<i>O. viridis</i>	0.282	0.606	0.000	0.000	0.661	2.02E-46	1.94E-151	0.680	0.000	0.000
	<i>O. cardiacum</i>	0.171	0.422	0.000	0.000	0.542	0.000	0.000	0.419	0.000	0.000
	<i>P. minor</i>	0.225	0.550	0.000	0.000	-	-	-	0.535	0.000	0.000
Charophyta	<i>M. viride</i>	0.115	0.453	0.000	0.000	-	-	-	0.432	0.000	0.000
	<i>Z. circumcarinatum</i>	0.238	0.527	0.000	0.000	0.389	0.000	1.62E-65	0.451	0.000	0.000
	<i>C. globosum</i>	0.145	0.430	0.000	0.000	0.248	0.000	1.25E-47	0.421	0.000	0.000
Bryophyte	<i>M. polymorpha</i>	0.094	0.386	0.000	0.000	0.253	0.000	3.26E-149	0.405	0.000	0.000
	<i>S. ruralis</i>	0.103	0.390	0.000	0.000	0.220	0.000	1.09E-83	0.396	0.000	0.000
	<i>P. patens</i>	0.106	0.394	0.000	0.000	0.231	0.000	3.93E-104	0.399	0.000	0.000
Pteridophyte	<i>S. moellendorffii</i>	0.743	1.019	1.02E-02	6.27E-91	1.010	6.47E-01	1.07E-31	1.041	2.56E-14	1.19E-115
	<i>S. uncinata</i>	0.898	1.187	1.06E-112	2.38E-74	1.259	6.55E-85	1.26E-80	1.215	8.34E-297	2.53E-99
	<i>C. lindheimeri</i>	0.555	0.729	0.000	2.42E-76	0.709	2.29E-57	4.23E-22	0.746	0.000	5.67E-98
	<i>A. spinulosa</i>	0.467	0.681	0.000	4.19E-140	0.664	5.53E-81	2.05E-43	0.679	0.000	3.98E-150

Gymnosperms	<i>C. oliveri</i>	0.323	0.562	0.000	9.29E-248	0.515	1.12E-179	4.46E-63	0.544	0.000	4.78E-236
	<i>G. parvifolium</i>	0.385	0.587	0.000	4.15E-124	0.516	5.06E-118	7.28E-19	0.618	0.000	4.46E-170
	<i>G. biloba</i>	0.372	0.643	0.000	4.29E-222	0.610	2.95E-85	2.73E-62	0.655	0.000	2.14E-265
Monocotyledon	<i>S. bicolor</i>	0.381	0.630	0.000	1.22E-173	0.548	1.04E-177	1.09E-41	0.626	0.000	9.58E-187
	<i>Z. mays</i>	0.407	0.639	0.000	1.13E-160	0.552	3.29E-172	2.95E-31	0.625	0.000	1.08E-158
	<i>O. sativa</i>	0.411	0.643	0.000	3.42E-157	0.563	1.43E-164	1.29E-33	0.639	0.000	2.81E-166
Dicotyledon	<i>A. thaliana</i>	0.352	0.577	0.000	1.02E-212	0.500	3.63E-284	4.23E-47	0.570	0.000	4.87E-219
	<i>G. max</i>	0.327	0.551	0.000	6.83E-226	0.457	0.000	2.67E-42	0.547	0.000	3.72E-237
	<i>P. trichocarpa</i>	0.374	0.590	0.000	1.06E-195	0.506	6.86E-261	1.12E-35	0.579	0.000	9.67E-195

Supplemental Table S4. The statistical analysis of SCUB frequency based on intron number and exon position

Taxomony	Species	Intron number	Exon position		
			2-exons	3-exons	11-exons
Archaeplastida	<i>P. haitanensis</i>	-	-	-	-
	<i>G. taiwanensis</i>	-	-	-	-
	<i>C. caldarium</i>	-	-	-	-
Chlorophyta	<i>O. viridis</i>	3.8189E-10	0.0177	-	-
	<i>O. cardiacum</i>	1.9696E-36	0.0474	0.7807	0.3611
	<i>P. minor</i>	-	-	-	-
Charophyta	<i>M. viride</i>	-	-	-	-
	<i>Z. circumcarinatum</i>	0.8791	0.1936	0.9864	
	<i>C. globosum</i>	0.0001	0.0625		
Bryophyte	<i>M. polymorpha</i>	0.9112	0.0045	0.3619	
	<i>S. ruralis</i>	0.0580	0.9101	0.7858	
	<i>P. patens</i>	0.1051	0.3949	0.1131	
Pteridophyte	<i>S. moellendorffii</i>	0.9888	0.3252	0.3700	
	<i>S. uncinata</i>	0.3542	0.9604	0.7640	
	<i>C. lindheimeri</i>	0.0736	0.6375	0.7797	
	<i>A. spinulosa</i>	0.2487	0.2444	0.7502	
Gymnosperms	<i>C. oliveri</i>	0.4815	0.6688	0.9153	
	<i>G. parvifolium</i>	0.4544	0.0234	0.8574	
	<i>G. biloba</i>	0.4365	0.9917	0.1137	
Monocotyledon	<i>S. bicolor</i>	0.3969	0.1570	0.3659	
	<i>Z. mays</i>	0.0861	0.2145	0.9869	
	<i>O. sativa</i>	0.4340	0.2158	0.9557	
Dicotyledon	<i>A. thaliana</i>	0.0630	0.2570	0.2791	
	<i>G. max</i>	0.3050	0.4592	0.1130	
	<i>P. trichocarpa</i>	0.1173	0.2421	0.2063	

Supplemental Table S5. The statistical analysis of the association between the DNA methylation induced conversion of C to T and SCUB frequency

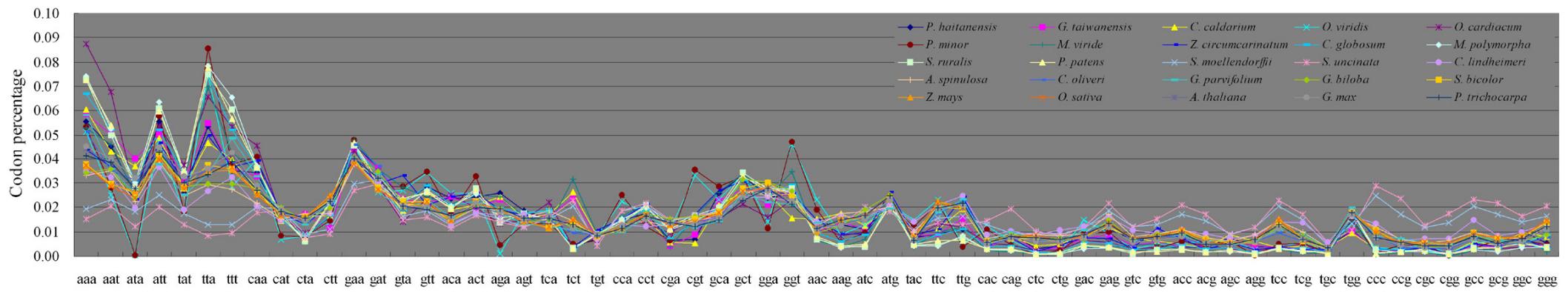
Taxonomy	Species	The second-third nucleotides		The third and next first nucleotides	
		Combination	P value	Combination	P value
Pteridophyte	<i>S. moellendorffii</i>	NCG/NCC vs NAG/NAC	0.0193	NC G/NG G vs NC A/NG A	1.3418E-04
	<i>S. uncinata</i>	NCG/NCC vs NAG/NAC	0.0299	NC G/NG G vs NC A/NG A	4.6580E-03
	<i>C. lindheimeri</i>	NCG/NCC vs NAG/NAC	9.8025E-04	NC G/NG G vs NC A/NG A	1.0950E-03
	<i>A. spinulosa</i>	NCG/NCC vs NAG/NAC	3.6793E-17	NC G/NG G vs NC A/NG A	6.9371E-03
Gymnosperms	<i>C. oliveri</i>	NCG/NCC vs NAG/NAC	1.7837E-18	NC G/NG G vs NC A/NG A	6.3845E-03
	<i>G. parvifolium</i>	NCG/NCC vs NAG/NAC	0.0430	NC G/NG G vs NC A/NG A	0.3198
	<i>G. biloba</i>	NCG/NCC vs NTG/NTC	1.9690E-23	NC G/NG G vs NC A/NG A	1.2030E-03
Monocotyledon	<i>S. bicolor</i>	NCG/NCC vs NAG/NAC	9.4962E-17	NC G/NG G vs NC A/NG A	7.3635E-09
	<i>Z. mays</i>	NCG/NCC vs NTG/NTC	1.6051E-20	NC G/NG G vs NC A/NG A	2.7782E-10
	<i>O. sativa</i>	NCG/NCC vs NTG/NTC	3.5284E-19	NC G/NG G vs NC A/NG A	2.2764E-08
Dicotyledon	<i>A. thaliana</i>	NCG/NCC vs NAG/NAC	9.2478E-13	NC G/NG G vs NC A/NG A	1.9299E-10
	<i>G. max</i>	NCG/NCC vs NAG/NAC	4.1584E-13	NC G/NG G vs NC A/NG A	3.4352E-11
	<i>P. trichocarpa</i>	NCG/NCC vs NAG/NAC	1.4345E-19	NC G/NG G vs NC A/NG A	5.6690E-09

Supplemental Table S6. The statistical analysis of the association between the DNA methylation induced conversion of C to T and SCUB frequency based on special amino acids

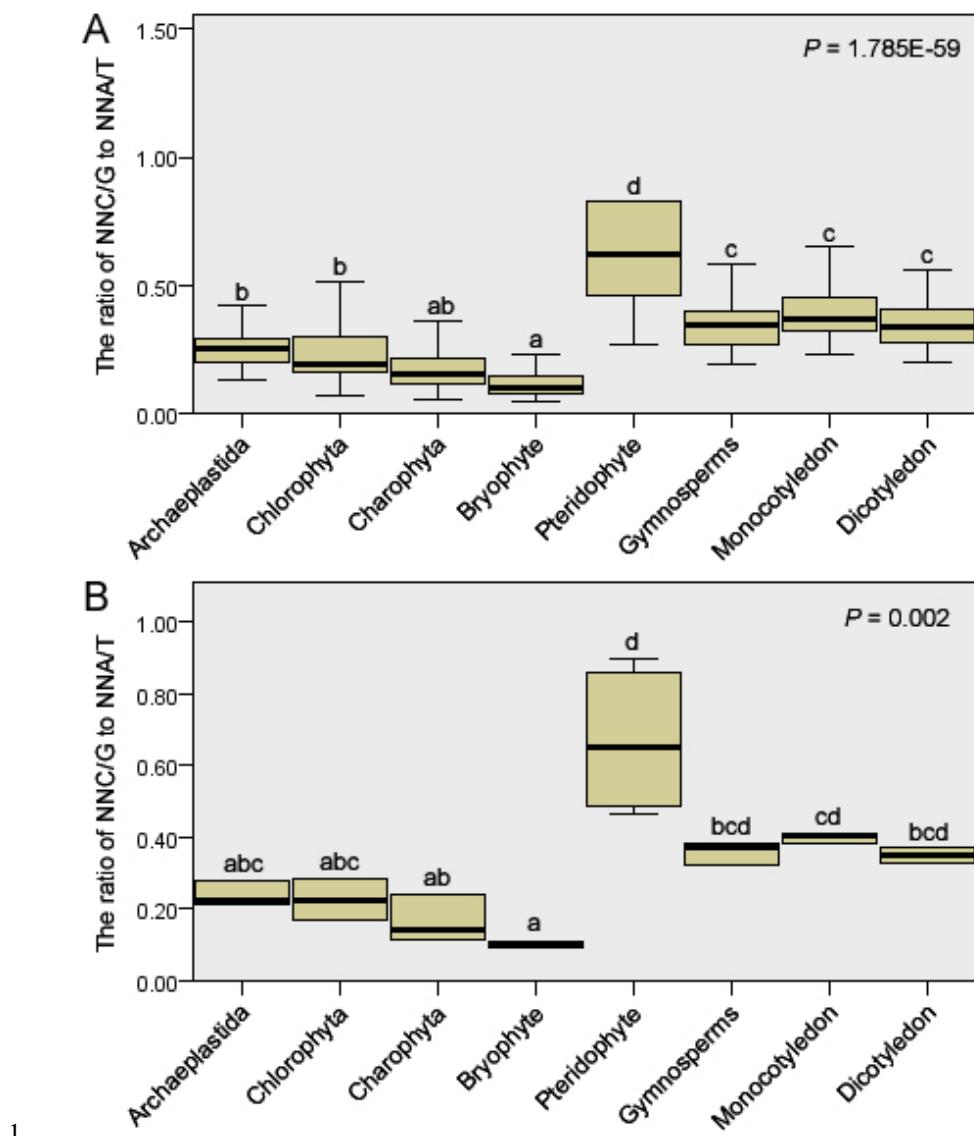
Taxonomy	Species	Means of NCG/NCC ratios	Means of NXG/NXC ratios	P value
Archaeplastida	<i>P. haitanensis</i>	0.895	0.793	0.657
	<i>G. taiwanensis</i>	0.839	0.923	0.799
	<i>C. caldarium</i>	0.816	0.836	0.910
Chlorophyta	<i>O. viridis</i>	1.170	1.162	0.984
	<i>O. cardiacum</i>	1.259	1.060	0.455
	<i>P. minor</i>	1.715	1.265	0.576
Charophyta	<i>M. viride</i>	1.017	0.725	0.570
	<i>Z. circumcarinatum</i>	0.678	1.553	0.286
	<i>C. globosum</i>	1.010	0.784	0.378
Bryophyte	<i>M. polymorpha</i>	0.844	0.893	0.809
	<i>S. ruralis</i>	0.710	1.179	0.299
	<i>P. patens</i>	1.027	1.033	0.991
Pteridophyte	<i>S. moellendorffii</i>	0.771	1.077	0.003
	<i>S. uncinata</i>	0.827	1.190	0.033
	<i>C. lindheimeri</i>	0.724	1.122	0.041
	<i>A. spinulosa</i>	0.689	1.225	0.013
Gymnosperms	<i>C. oliveri</i>	0.590	1.159	0.006
	<i>G. parvifolium</i>	0.834	1.104	0.136
	<i>G. biloba</i>	0.585	1.217	0.002
Monocotyledon	<i>S. bicolor</i>	0.584	1.141	0.043
	<i>Z. mays</i>	0.587	1.149	0.056
	<i>O. sativa</i>	0.591	1.138	0.091
Dicotyledon	<i>A. thaliana</i>	0.648	1.184	0.030
	<i>G. max</i>	0.625	1.249	0.019
	<i>P. trichocarpa</i>	0.620	1.190	0.018

Supplemental Table S7. The statistic analysis of usage bias of stop codons and internal stop codons

Taxonomy	Species	Stop codons			Internal stop codons		
		TAA+TGA	TAG	P value	TAA+TGA	TAG	P value
Archaeplastida	<i>P. haitanensis</i>	173	38	1.489E-20			
	<i>G. taiwanensis</i>	181	52	2.886E-17			
	<i>C. caldarium</i>	153	44	8.105E-15			
Chlorophyta	<i>O. viridis</i>	77	16	2.526E-10			
	<i>O. cardiacum</i>	89	10	2.025E-15			
	<i>P. minor</i>	76	6	1.074E-14			
Charophyta	<i>M. viride</i>	89	16	1.048E-12			
	<i>Z. circumcarinatum</i>	87	16	2.637E-12			
	<i>C. globosum</i>	90	8	1.199E-16			
Bryophyte	<i>M. polymorpha</i>	84	5	5.570E-17			
	<i>S. ruralis</i>	74	5	8.287E-15			
	<i>P. patens</i>	80	5	4.123E-16			
Pteridophyte	<i>S. moellendorffii</i>	41	11	3.179E-05			
	<i>S. uncinata</i>	42	10	9.097E-06	5	1	0.102
	<i>C. lindheimeri</i>	58	24	1.736E-04	33	1	4.066E-08
	<i>A. spinulosa</i>	67	20	4.681E-07	61	6	1.826E-11
Gymnosperms	<i>C. oliveri</i>	66	16	3.360E-08			
	<i>G. parvifolium</i>	56	10	1.494E-08			
	<i>G. biloba</i>	64	20	1.580E-06			
Monocotyledon	<i>S. bicolor</i>	64	20	1.580E-06			
	<i>Z. mays</i>	83	28	1.786E-07			
	<i>O. sativa</i>	78	30	3.860E-06			
Dicotyledon	<i>A. thaliana</i>	64	21	3.101E-06			
	<i>G. max</i>	63	20	2.360E-06			
	<i>P. trichocarpa</i>	71	27	8.803E-06			

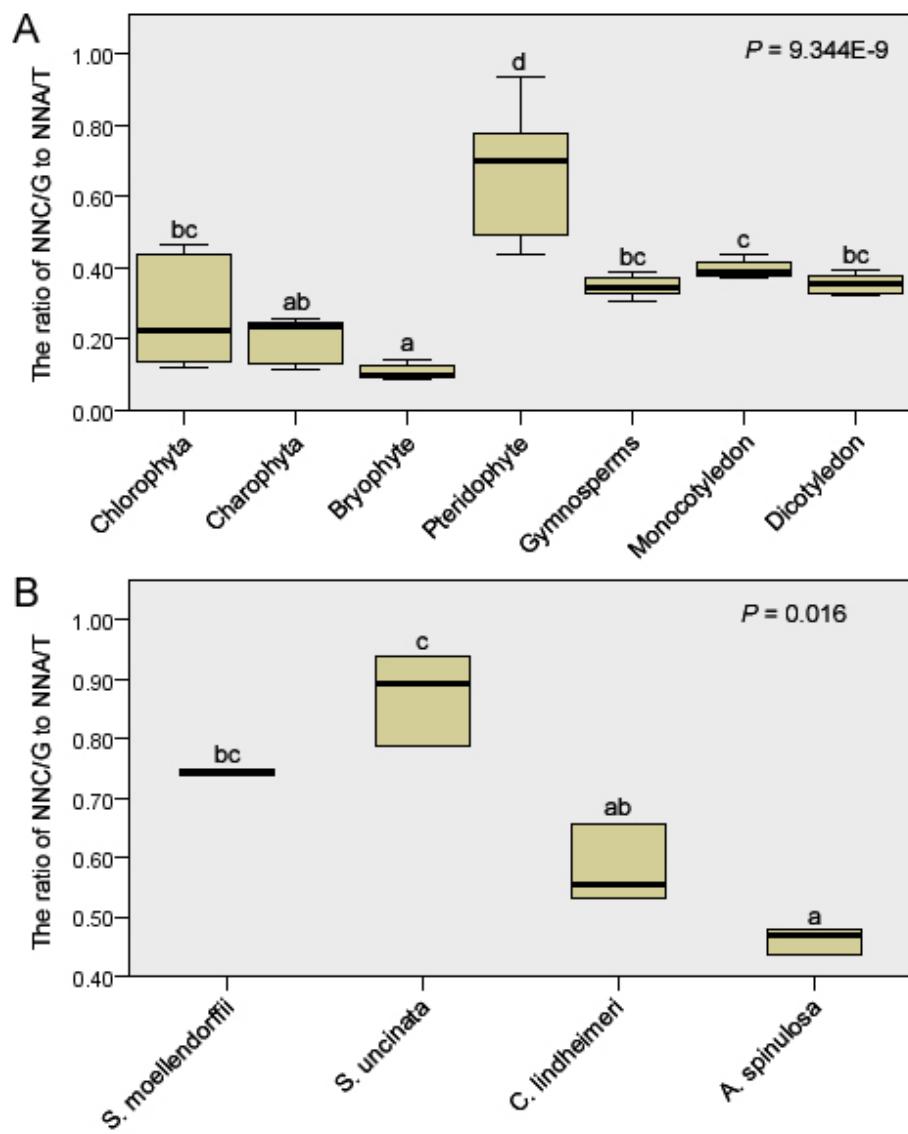


Supplemental Figure S1

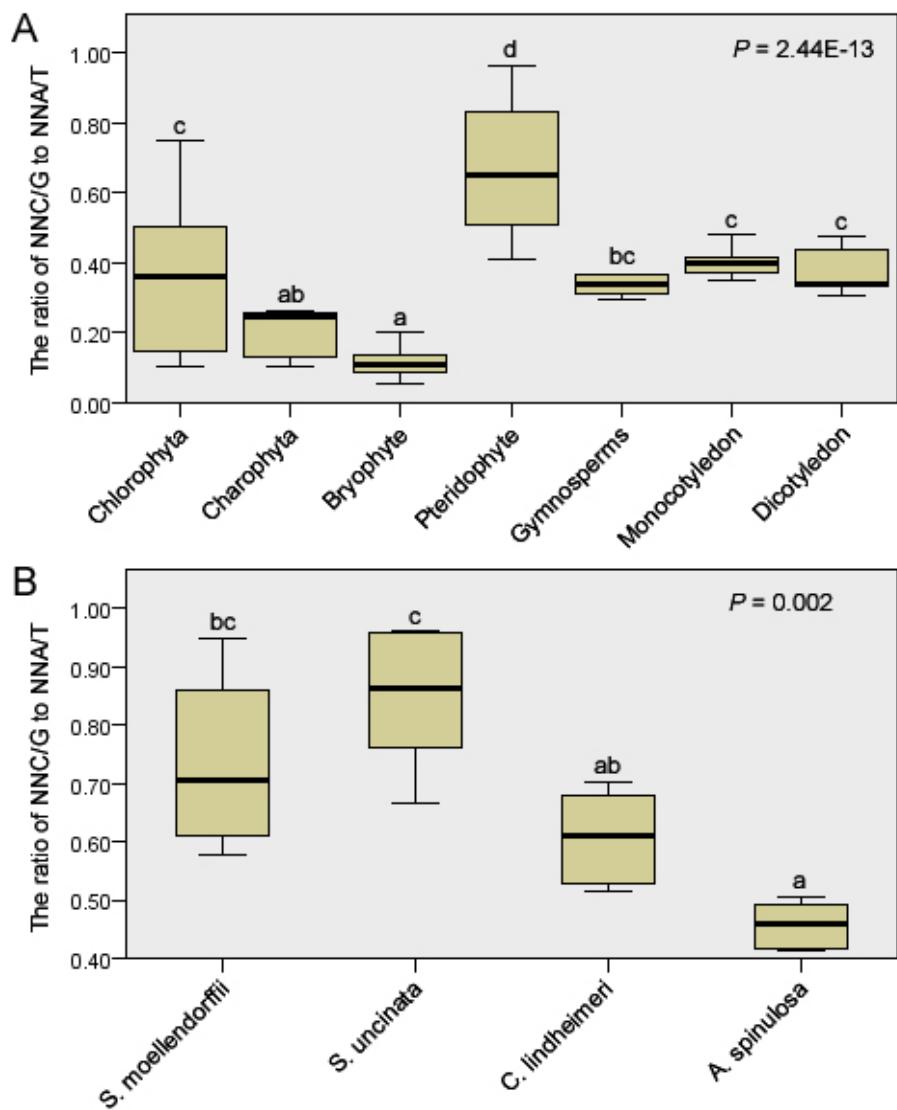


1.

Supplemental Figure S2



Supplemental Figure S3



Supplemental Figure S4