

Supplementary Table S2. Organelle DNA sequences (including their lengths, GenBank accession numbers, and noncoding DNA contents) used as queries for BLASTN searches against nuclear genomes. Website addresses are shown in square brackets.

Taxon	Genome Size (nt) ^a	Fraction noncoding ^b	Source / GenBank Accession #
PLASTID			
Land Plants:			
<i>Arabidopsis thaliana</i>	154,478	0.41	NC_000932
<i>Brachypodium distachyon</i>	135,199	0.50	NC_011032
<i>Carica papaya</i> cultivar SunUp	160,100	0.43	NC_010323
<i>Cucumis sativus</i> cultivar Borszczagowski	155,293	0.43	NC_007144
<i>Glycine max</i> cultivar PI 437654	152,218	0.42	NC_007942
<i>Medicago truncatula</i>	124,033	0.42	NC_003119
<i>Oryza sativa</i> subsp. <i>indica</i>	134,496	0.42	NC_008155
<i>Oryza sativa</i> subsp. <i>japonica</i>	134,525	0.42	NC_001320
<i>Physcomitrella patens</i> subsp. <i>patens</i>	122,890	0.33	NC_005087
<i>Populus trichocarpa</i>	157,033	0.39	NC_009143
<i>Selaginella moellendorffii</i>	143,780	0.46	NC_013086
<i>Sorghum bicolor</i> cultivar BTx623	140,754	0.49	NC_008602
<i>Vitis vinifera</i> cultivar Maxxa	160,928	0.44	NC_007957
Green Algae:			
<i>Chlamydomonas reinhardtii</i> strain CC-503	204,159	0.57	FJ423446
<i>Coccomyxa</i> sp. C-169	175,053	NA	Download from DOE JGI [http://genome.jgi-psf.org/Coc_C169_1]
<i>Ostreococcus</i> sp. RCC809	67,238	NA	Download from DOE JGI [http://genome.jgi-psf.org/OstRCC809_2]
<i>Ostreococcus tauri</i> strain OTH95	71,666	0.20	NC_008289
<i>Micromonas pusilla</i> strain CCMP1545	41,811	0.37	NC_012568
<i>Micromonas</i> sp. RCC299	72,585	0.37	NC_012575

<i>Volvox carteri</i> f. <i>nagariensis</i>	461,064	0.80	GU084820
Red Alga:			
<i>Cyanidioschyzon merolae</i> strain 10D	149,987	0.05	NC_004799
Apicomplexans:			
<i>Babesia bovis</i> strain T2Bo	35,107	0.28	AAXT01000007
<i>Eimeria tenella</i> strain Penn State	34,750	0.05	NC_004823
<i>Plasmodium falciparum</i> strain HB3	29,529	0.03	DQ642846
<i>Theileria parva</i> strain Muguga	39,579	0.17	NC_007758
<i>Toxoplasma gondii</i> strain RT	34,996	0.16	NC_001799
Haptophyte:			
<i>Emiliania huxleyi</i> strain CCMP373	105,309	0.13	NC_007288
Stramenopiles:			
<i>Aureococcus anophagefferens</i> strain CCMP1984	89,599	0.86	NC_012898
<i>Phaeodactylum tricornutum</i> strain CCAP 1055/1	117,369	0.13	NC_008588
<i>Thalassiosira pseudonana</i> strain CCMP1335	128,814	0.15	NC_008589
MITOCHONDRIA			
Land Plants:			
<i>Arabidopsis thaliana</i>	366,924	0.795	NC_001284
<i>Carica papaya</i> cultivar SunUp	476,890	0.918	NC_012116
<i>Oryza sativa</i> subsp. <i>indica</i>	491,515	0.876	NC_007886
<i>Oryza sativa</i> subsp. <i>japonica</i>	490,520	0.894	NC_011033
<i>Physcomitrella patens</i> subsp. <i>patens</i>	105,340	0.600	NC_000892
<i>Sorghum bicolor</i> cultivar BTx623	468,628	0.918	NC_008360
<i>Vitis vinifera</i> cultivar "Pinot noir clone ENTAV115"	773,279	0.925	NC_012119

Green Algae:

<i>Chlamydomonas reinhardtii</i> strain CC-503	15,782	0.180	EU306617
<i>Coccomyxa</i> sp. C-169	NA	NA	Download from DOE JGI [http://genome.jgi-psf.org/Coc_C169_1]
<i>Ostreococcus</i> sp. RCC809	43,168	0.079	Download from DOE JGI [http://genome.jgi-psf.org/OstRCC809_2]
<i>Ostreococcus tauri</i> strain OTH95	44,237	0.079	NC_008290
<i>Micromonas pusilla</i> strain CCMP1545	41,691	NA	FJ858268
<i>Micromonas</i> sp. RCC299	47,425	0.175	NC_012643
<i>Volvox carteri</i> forma <i>nagariensis</i>	30,000	0.610	GU048821

Red Alga:

<i>Cyanidioschyzon merolae</i> strain 10D	32,211	0.001	NC_000887
--	--------	-------	-----------

Apicomplexans:

<i>Babesia bovis</i> strain T2Bo	6,005	0.365	NC_009902
<i>Plasmodium falciparum</i> strain NF54	5,967	0.412	NC_002375
<i>Theileria parva</i> strain <i>Muguga</i>	5,895	0.325	NC_011005

Haptophyte:

<i>Emiliana huxleyi</i> strain CCMP373	29,013	0.220	NC_005332
---	--------	-------	-----------

Stramenopiles:

<i>Thalassiosira pseudonana</i> strain CCMP1335	43,827	0.264	NC_007405
--	--------	-------	-----------

^a Genome size statistics came from GenBank's Entrez Organelle Genome database [[http://www.ncbi.nlm.nih.gov/genomes/genlist.cgi?taxid=2759&type=4&name=Eukaryotae Organelles](http://www.ncbi.nlm.nih.gov/genomes/genlist.cgi?taxid=2759&type=4&name=Eukaryotae%20Organelles)].

^b Noncoding DNA content was calculated as follows: genome length minus the collective length of all annotated protein-, rRNA-, and tRNA-coding regions, not including the portions of these regions that are also annotated as introns. This method is contingent on the authors of the GenBank records having properly annotated their entry. If coding regions or introns have been ignored or inaccurately annotated, coding and noncoding DNA content values will be incorrect.