

**TABLE S2.** List of the RBCL sequences used in phylogenetic analysis. Accession numbers for GenBank entries and contig ids for MMETSP\* are listed.

	ORGANISM	SOURCE (Database)	ACCESSION NUMBER/CONTIG	
Euglenozoa	<i>Euglena longa</i>	GenBank	NP_074963	
	<i>Euglena gracilis</i> Z	GenBank	NP_041936	
	<i>Euglena viridis</i> NJ001	GenBank	YP_007517056	
	<i>Eugleniformis proxima</i>	GenBank	YP_009032761	
	<i>Eutreptiella gymnastica</i> K-0333	GenBank	YP_006234185	
	<i>Eutreptiella gymnastica</i> CCMP1594	MMETSP	CAMNT_0046485463	
	<i>Eutreptia viridis</i>	GenBank	AEQ94204	
	<i>Colacium vesiculosum</i>	GenBank	AEW12959	
	<i>Monomorphina aenigmatica</i>	GenBank	YP_007317188	
	<i>Strombomonas acuminata</i>	GenBank	AEW12993	
	<i>Cryptoglana skujai</i>	GenBank	AKL39014	
	<i>Euglena gracilis bacillaris</i>	GenBank	AKL82391	
	<i>Euglena viridis</i> SAG 224-17d	GenBank	AKL79017	
	<i>Euglenaria anabaena</i>	GenBank	AKJ83339	
	<i>Monomorphina parapryum</i>	GenBank	AKL78929	
	<i>Trachelomonas volvocina</i>	GenBank	AKL82413	
	Chloroplastida	<i>Chlamydomonas reinhardtii</i>	GenBank	NP_958405
<i>Chlamydomonas moewusii</i>		GenBank	ABU88322	
<i>Arabidopsis thaliana</i>		GenBank	NP_051067	
<i>Oryza sativa</i>		GenBank	YP_654221	
<i>Ostreococcus tauri</i>		GenBank	YP_717262	
<i>Glycine max</i>		GenBank	YP_538747	
<i>Medicago truncatula</i>		GenBank	YP_001381744	
<i>Micromonas</i> sp. RCC299		GenBank	ACO55540	
<i>Zea mays</i>		GenBank	NP_043033	
<i>Vitis vinifera</i>		GenBank	YP_567084	
<i>Selaginella moellendorffii</i>		GenBank	ADH10411	
<i>Volvox carteri</i>		GenBank	ACY06055	
<i>Physcomitrella patens</i>		GenBank	NP_904194	
<i>Pyramimonas amyliifera</i>		MMETSP	CAMNT_0041901623	
<i>Pyramimonas obovata</i>		MMETSP	CAMNT_0006873903	
<i>Pyramimonas parkeae</i>		GenBank	YP_002600925	
<i>Dunaliella salina</i>		GenBank	ACS95083	
<i>Haematococcus pluvialis</i>		GenBank	ACJ64122	
<i>Chloromonas</i> sp. ANT3		GenBank	AAD00447	
<i>Chlorella vulgaris</i>		GenBank	NP_045897	
<i>Auxenochlorella protothecoides</i>		GenBank	YP_009019341	
<i>Aneura mirabilis</i>		GenBank	YP_001687224	
<i>Orobanche corymbosa</i>		GenBank	AAC49722	
<i>Orobanche fasciculata</i>		GenBank	AAC49723	
<i>Orobanche californica</i>		GenBank	YP_009108392	
<i>Harveya capensis</i>		GenBank	AAAY98036	
<i>Harveya stenosiphon</i>		GenBank	AAAY98035	
<i>Harveya coccinea</i>		GenBank	AAAY98034	
Glaucophyta		<i>Cyanophora paradoxa</i>	GenBank	NP_043240
		<i>Cyanoptyche gloeocystis</i>	MMETSP	CAMNT_0041991067
Cyanobacteria		<i>Synechococcus</i> sp. JA-3-3Ab	GenBank	Q2JV67
		<i>Gloeobacter violaceus</i>	GenBank	NP_925102
		<i>Anabaena variabilis</i>	GenBank	ABA23512
	<i>Acaryochloris marina</i>	GenBank	B0CCC1	
	<i>Cyanothece</i> sp. PCC 7425	GenBank	B8HQS5	
Chlorarachniophyta	<i>Bigelowiella natans</i>	GenBank	YP_778601	
	<i>Lotharella</i> sp. CCMP622	GenBank	AGY61412	

\*MMETSP - Marine Microbial Eukaryote Transcriptome Sequencing Project (<http://marinemicroeukaryotes.org/>)