Supporting Information

Liu et al. 10.1073/pnas.0905841106

SI References

1. McDonald SM, Sarno D, Scanlan DJ, Zingone A (2007) Genetic diversity of eukaryotic ultraphytoplankton in the Gulf of Naples during an annual cycle. *Aquat Microb Ecol* 50:75–89.
2. Frada M, Not F, Probert I, de Vargas C (2006) CaCO3 optical detection with fluorescent in situ hybridization: A new method to identify and quantify calcifying microorganisms from the oceans. *J Phycol* 42(6):1162–1169.
3. Chao A, Shen T (2003–2005) Program SPADE (species prediction and diversity estimation).
4. Aubry M-P, Bord D (2009) Reshuffling the cards in the photic zone at the Eocene/Oligocene boundary. *The Late Eocene Earth: Hot House, Ice House, and Impacts*, eds Koeberl C, Montanari A (*Geo Soc Am Bull*, Boulder, CO).
5. Tamura K, Dudley J, Nei M, Kumar S (2007) MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* 24:1596–1599.
6. Zubkov MV, Tarran GA (2008) High bacterivory by the smallest phytoplankton in the North Atlantic Ocean. *Nature* 455(7210):224–226.
7. Nygaard K, Tobiesen A (1993) Bacterivory in algae: A survival strategy during nutrient limitation. *Limnol Oceanogr* 38(2):273–279.
8. Tillmann U (1998) Phagotrophy of a plastidic haptophyte, Prymnesium patelliferum. *Aquat Microb Ecol* 14:155–160.
9. Aubry M-P (2009) A sea of lilliputians. *Extinction, Dwarfing and the Lilliput Effect*, eds Twitchett R, Wade BS (Palaeogeography, Palaeoclimatology, Palaeoecology, Special Publication), in press.

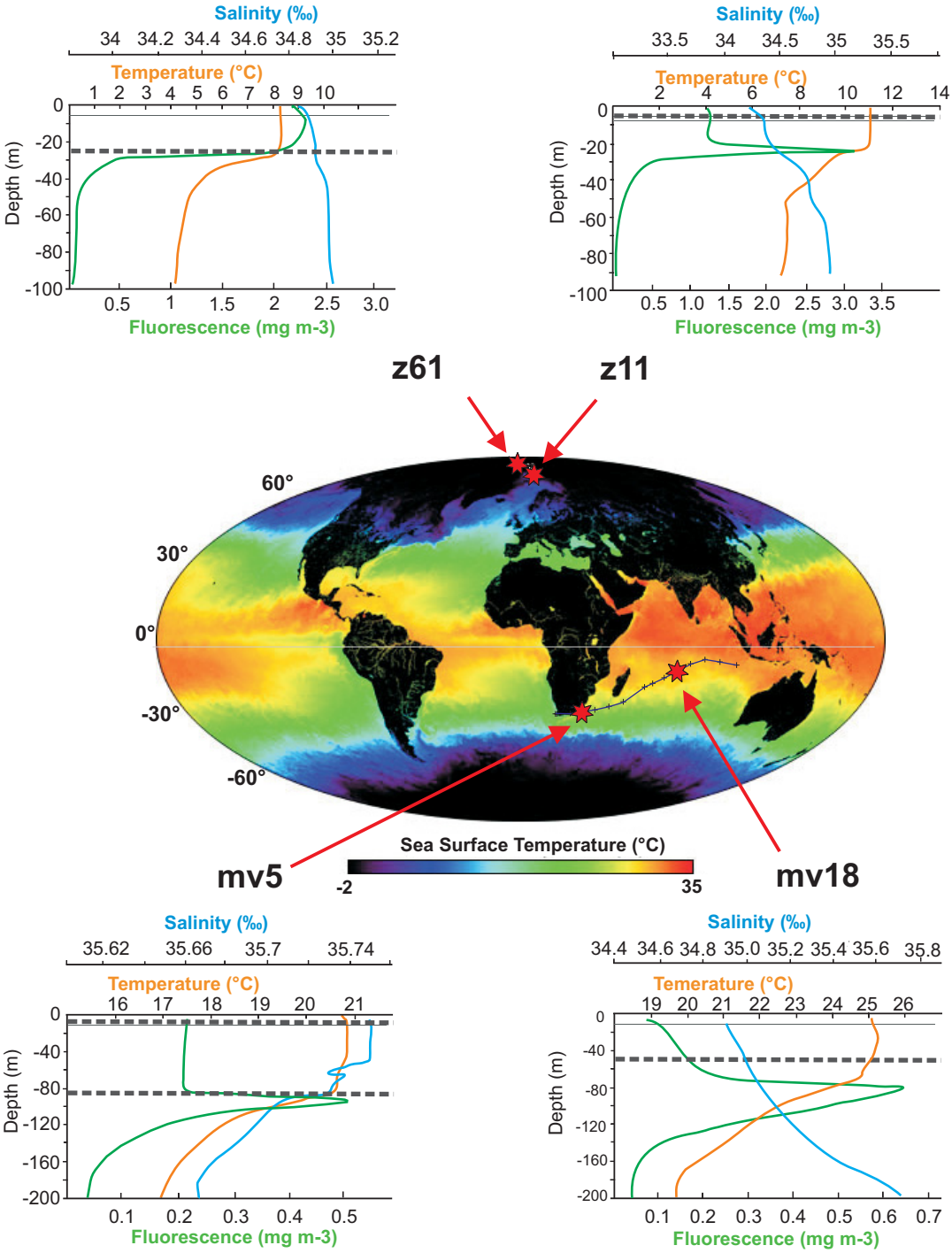

Fig. S1. Cruise tracks and sampling locations. The 4 stations are marked with red stars. Temperature, salinity, and fluorescence profiles down to 100 (z) or 200 m (mv) depth are given for each station. Dotted lines indicate the depths at which water used for DNA extraction and rDNA sequencing was sampled with Niskin bottles. Global sea surface temperature corresponds to a monthly (May 2001) composite of data captured by the satellite *MODIS* (http://modis.gsfc.nasa.gov/). Further details shown in Table S1 below.

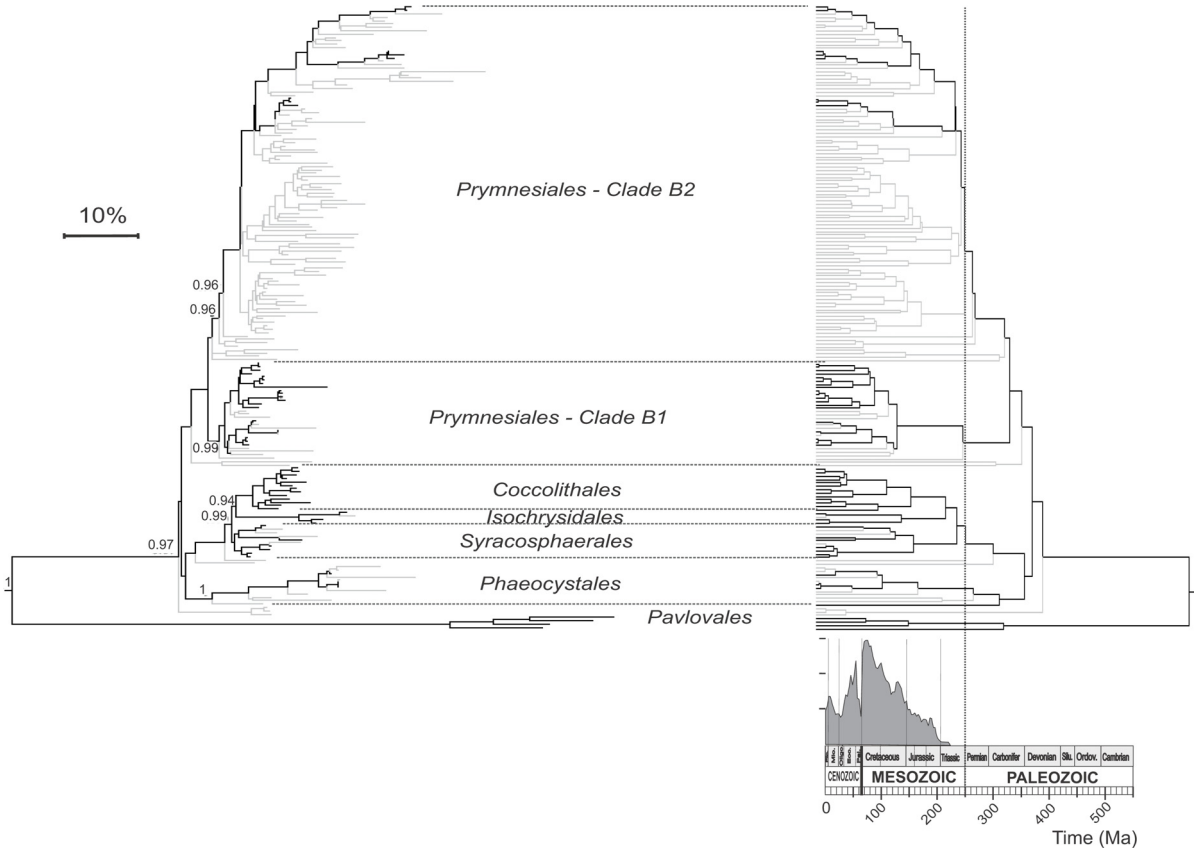

Fig. S2. Diversification of the haptophytes along geological time. Relaxed clock analysis calibrated on the coccolithophore fossil record. The tree on the *Left* shows the pattern of diversification; numbers represent Bayesian posterior probabilities of key divergences only. The tree on the *Right* shows the corresponding divergence times. Branches are color coded: black for sequences obtained from cultured and previously described haptophytes; gray for environmental and previously unknown sequences.

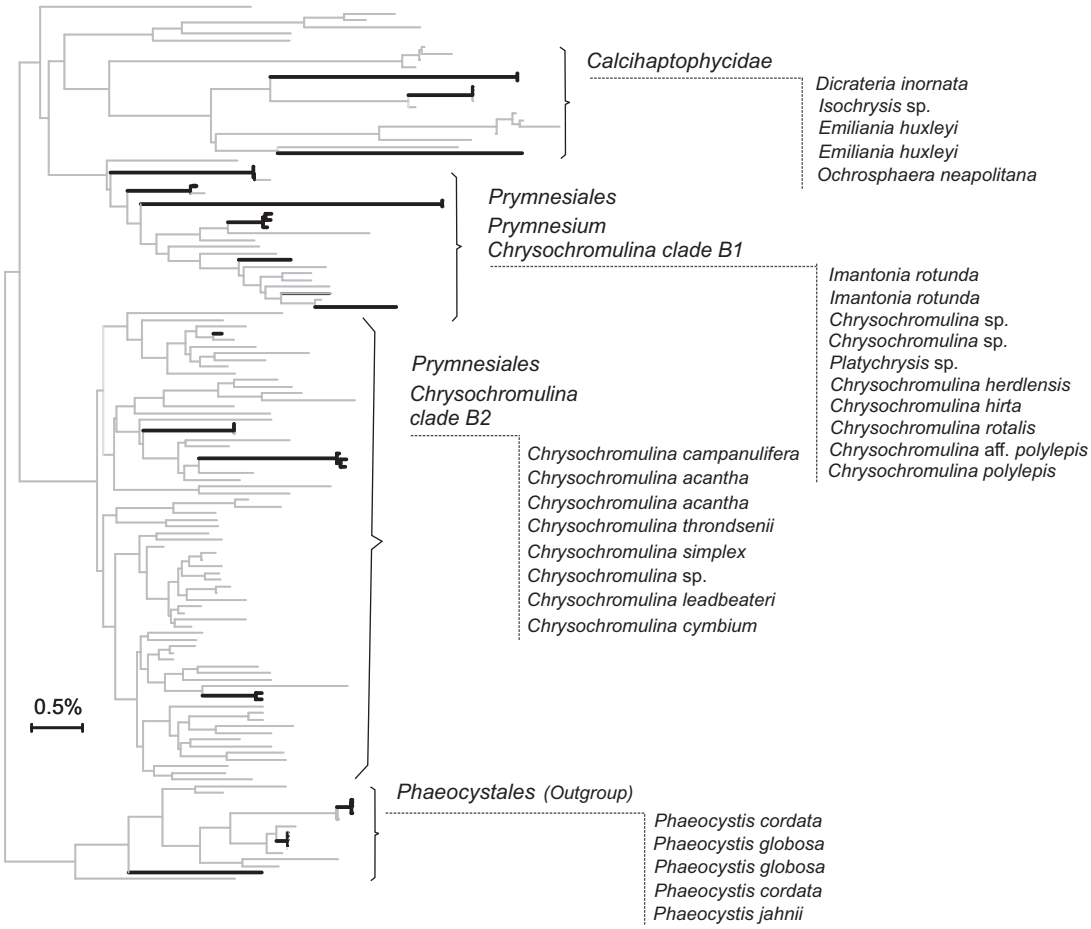

Fig. S3. Chloroplastic ''view'' of eukaryotic and haptophyte diversity from the Gulf of Naples, Mediterranean Sea. McDonald and collaborators (1) used chloroplastic-biased 16S rDNA primers to explore 6 environmental clone libraries over an annual cycle. Forty-six percent of the retrieved unique eukaryotic sequences and 73% of the total eukaryotic OTUs belonged to the Haptophyta. This overwhelming haptophyte biodiversity is reanalyzed here using the neighbor-joining phylogenetic method based on a Tajima-Nei distance matrix. Gray branches represent the unveiled environmental diversity, integrated into taxonomically known 16S rDNA data (black branches). Note that several taxonomic inconsistencies were removed as compared to the original data set presented in ref. 1.

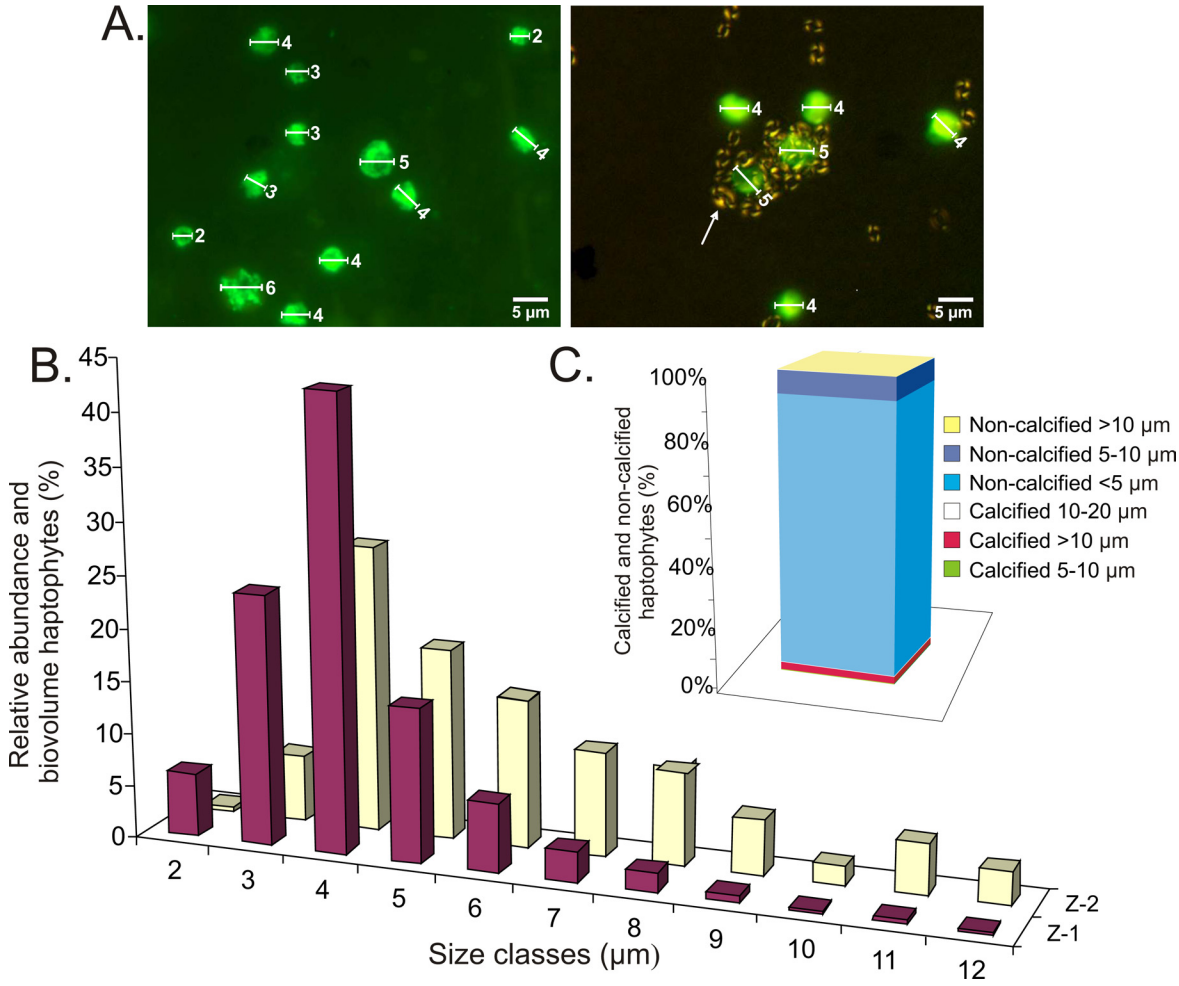

Fig. S4. Abundance, size, and biovolume in noncalcifying and calcifying haptophytes. (*A*) Haptophyte cells from 28 plankton samples from various depths in North Pacific, Mediterranean Sea, and North Atlantic waters (see Table S4) were CODFISHed ($CaCO_3$ optical detection with haptophyte-specific fluorescent in situ hybridization) (2) and cell diameters were measured from 548 individuals (calibration bars, μm units). The white arrow in the *Right* panel points to a single $CaCO_3$ coccolith displaying typical light polarization pattern and allowing the detection of calcifying versus noncalcifying cells. The microscopy field shown in the *Left* panel displays 12 noncalcifying cells. (*B*) Relative abundance (Z-1) and relative biovolume (Z-2, estimated as a sphere [$\frac{4}{3}\pi r^3$]) of noncalcifying haptophytes in various size classes. (*C*) Relative abundance of different size fractions of noncalcifying and calcifying haptophytes. Note that the extensive diversity of LSU rDNA types reported herein and recovered from <3-μm filtrates, may partly come from larger, delicate nanoplanktonic cells disrupted by the vacuum pumping filtration process.

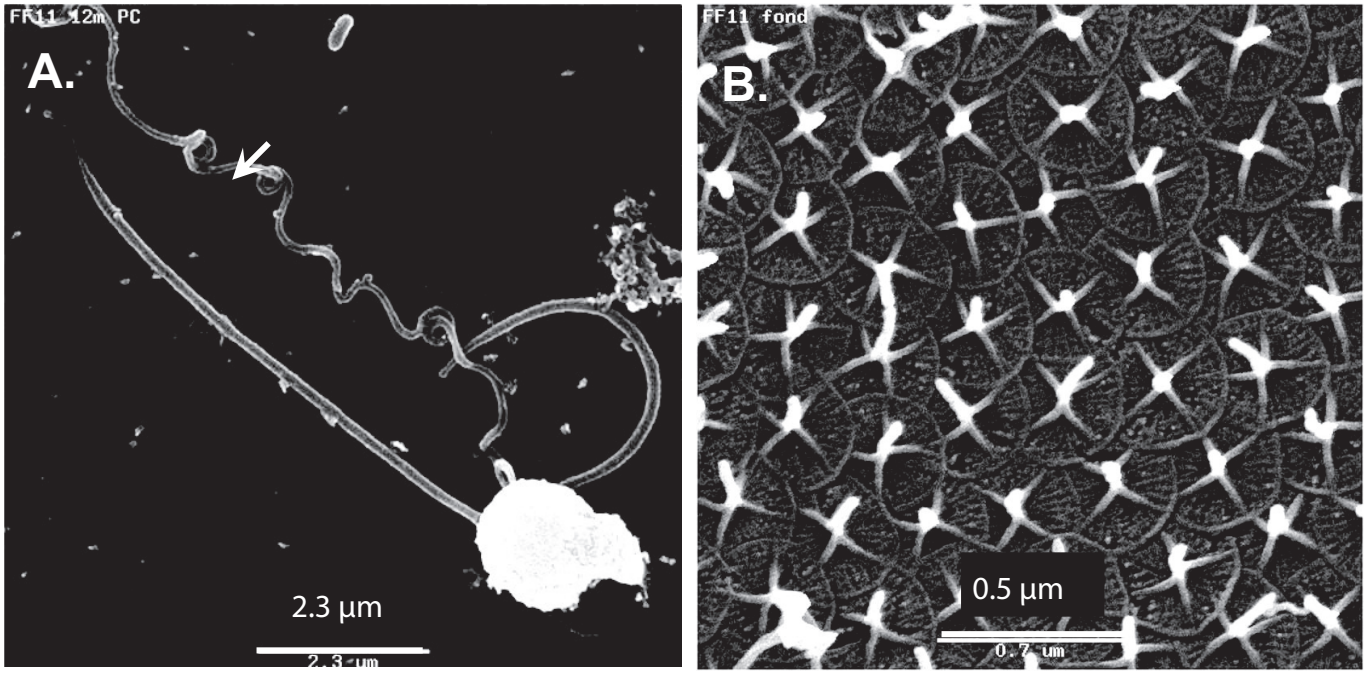

Fig. S5. Tiny Chrysochromulina: haptonema and scales. (*A*) Whole-mounts transmission electron microscopy picture of a tiny *Chrysochromulina* sp. collected from 24 m depth in the Bay of Banyuls/Mer, France, on May 15, 2001 (M.-J. Chrétiennot-Dinet, personal communication). Such unidentified species from the genus *Chrysochromulina* are very common and diverse in oligotrophic water but do not grow in current culture media. White arrow indicates the *haptonema*, which can be used to capture bacteria; bacterial ingestion is commonly observed in tiny haptophytes from both culture and field samples (6–9). (*B*) Typical organic scales covering the surface of cells within the genus *Chrysochromulina* (here the species *C. ephippium*). Both images were graciously contributed by M.-J. Chrétiennot-Dinet.

Table S1. Basic features of the samples analyzed in this paper

Library name	Cruise	Station	Lat, °	Long, °	Depth, m	Date	Vol seawater, L	Temperature, °C	Salinity, ‰	Fluorescence, mg/m^3	No. sequences retrieved
z11_11	ARCTIC	z11	72.5	19.6	5	8/25/2002	5	11.05	34.33	1.21	153
z61_43	ARCTIC	z61	76.32	7.98	25	8/29/2002	5	8.17	34.91	1.98	95
mv5_19	VANC 10MV	mv5	34.35	37.68	7	5/18/2003	15	20.81	35.75	0.22	144
mv5_21	VANC 10MV	mv5	34.35	37.68	85	5/18/2003	15	20.55	35.73	0.50	160
mv18_59	VANC 10MV	mv18	17.17	83.67	50	6/1/2003	15	25.03	35.04	0.24	122

Note that sea water was prefiltered through 3-μm membrane filters and collected onto 0.2-μm membrane filters.

Table S2. LSU rDNA total diversity estimates for each library using the Chao1 and ACE statistics and 2 divergence cutoffs

	rDNA sequence diversity estimate			
	Unique rDNA		3% divergence cutoff	
Sample	Chao1	ACE	Chao1	ACE
mv5_19	1099 (537–2,401)	1412 (682–3,076)	169 (114–288)	287 (158–595)
mv5_21	1098 (587–2,183)	1408 (790–2,606)	249 (158–446)	347 (210–622)
mv18_59	1147 (527–2,664)	1084 (554–2,230)	250 (154–463)	362 (220–642)
z11_11	509 (268–1,062)	756 (388–1,585)	88 (52–200)	90 (56–179)
z61_43	325 (157–769)	414 (198–956)	26 (19–62)	30 (20–66)

95% confidence intervals are given in parentheses. The Chao1 statistics give a lower bound of species richness estimation, while ACE scores indicate point estimation of species richness (3).

Table S3. Identification and origin of the haptophyte strains isolated, cultured, and characterized by electron microscopy and LSU rDNA D1–D2 sequencing, used in our study to anchor environmental genetic diversity

Accession no.	Species	Culture strain	Culture collection	Isolation, source
EU729452	*Platychrysis* sp.	RCC1385	Roscoff culture collection (RCC), France	Mediterranean, Spain
EU729451	*Platychrysis pienaarii*	RCC1392	RCC, France	Unknown
EU729449	*Prymnesium* sp.	RCC1440	RCC, France	Mediterranean, Tunisia
EU729447	*Prymnesium zebrinum*	RCC1432	RCC, France	N Atlantic, France
EU729448	*P. zebrinum*	RCC1438	RCC, France	N Atlantic, France
EU729446	*Prymnesium* sp.	RCC1446	RCC, France	Unknown
EU729445	*Prymnesium* sp.	RCC1443	RCC, France	N Atlantic, Spain
EU729444	*P. calathiferum*	CCMP707	Center for culture of marine phytoplankton (CCMP), USA	North Island, New Zealand
EU729450	*Prymnesium* sp.	RCC1450	RCC, France	Mediterranean, Tunisia
AF289038	*P. patelliferum*	unknown	Unknown	Unknown
EU729443	*P. parvum*	RCC1434	RCC, France	N Atlantic, English Channel
EU729442	*Prymnesium* sp.	CCMP711	CCMP, USA	N Atlantic, Maine, USA
EU729441	*Chrysochromulina* sp.	RCC1184	RCC, France	N Atlantic, France
EU729458	*Platychrysis pigra*	RCC1390	RCC, France	Mediterranean, France
EU729457	*Imantonia rotunda*	RCC1343	RCC, France	N Atlantic, France
EU729456	*Chrysochromulina brevifilum*	S-3	Algobank culture collection, France	N Atlantic, Spain
EU729455	*C. ericina*	CCMP283	CCMP, USA	N Altantic, Gulf of Maine, USA
EU729454	*C. hirta*	S-17	Algobank culture collection, France	N Atlantic,Spain
EU729453	*C. cf herdlensis*	CCMP284	CCMP, USA	49.87° N 142.67° W
EU729440	*C. camella*	CCMP289	CCMP, USA	29.97° N 63.86° W
EU729439	*C. camella*	RCC1185	RCC, France	N Atlantic, France
EU729438	*Chrysochromulina* sp.	S-14	Algobank culture collection, France	N Atlantic, Spain
EU729437	*C. acantha*	S-6	Algobank culture collection, France	N Atlantic, Spain
EU729436	*C. throndsenii*	S-5	Algobank culture collection, France	N Atlantic, Spain
EU729435	*Chrysochromulina* sp.	No code available	RCC, France	Unknown
EU729434	*C. simplex*	RCC1193	RCC, France	N Atlantic, Spain
DQ980469	*Chrysochromulina* sp.	NIES 1333	NIES collection, Japan	Pacific, Japan
EU729460	*Calcidiscus* sp.	RCC1157	RCC, France	Mediterranean, Spain
EU502878	*Calcidiscus* sp.	RCC1147	RCC, France	S Atlantic, Namibia
EU729463	*Umbilicosphaera hulburtiana*	RCC1474	RCC, France	S Atlantic, South Africa
EU729461	*U. sibogae*	RCC1468	RCC, France	Mediterranean, Spain
EU729462	*U. foliosa*	RCC1470	RCC, France	N Atlantic, Puerto Rico
EU729464	*Coccolithus braarudii*	AC613	Algobank culture collection, France	N Atlantic, English Channel
EU502875	*Jomonlithus litoralis*	RCC1354	RCC, France	Mediterranean, Spain
EU502872	*Hymenomonas globosa*	RCC1338	RCC, France	N Atlantic, English Channel
EU729469	*Ochrosphaera neapolitana*	RCC1359	RCC, France	N Atlantic, English Channel
EU729468	*Pleurochrysis dentata*	RCC1400	RCC, France	New Mexico, USA
EU729467	*Stauronertha neohelis*	RCC1206	RCC, France	N Atlantic, Guadeloupe
EU729466	*Calyptrosphaera sphaeroidea*	RCC1178	RCC, France	North Sea, Norway
EU729465	*Helladosphaera* sp.	RCC1182	RCC, France	Pacific, Japan
EU502879	*Syracosphaera pulchra*	RCC1460	RCC, France	Mediterranean, Spain
EU729471	*Coronosphaera mediterranea*	RCC1204	RCC, France	S Atlantic, South Africa
EU729473	*Helicosphaera carteri*	RCC1333	RCC, France	S Atlantic, South Africa
EU729472	*Scyphosphaera apsteinii*	RCC1455	RCC, France	Mediterranean, Spain
EU729470	*Algirosphaera robusta*	RCC1128	RCC, France	Mediterranean, Spain
EU729476	*Gephyrocapsa oceanica*	RCC1289	RCC, France	Mediterranean, Spain
EU729475	*Dicrateria* sp.	RCC1207	RCC, France	Mediterranean, Morocco
EU729474	*Isochrysis galbana*	RCC1348	RCC, France	N Atlantic, Irish Sea
EU729459	*Phaeocystis cordata*	CCMP 2495	CCMP, USA	Mediterranean, Italy
AF289040	*P. antarctica*	Unknown	Unknown	Unknown
EU502882	*Phaeocystis* sp.	AC618	Algobank culture collection, France	N Atlantic, English Channel
EU729479	*Exanthemachrysis gayraliae*	RCC1523	RCC, France	N Atlantic, English Channel
EU729478	*Rebecca salina*	RCC1545	RCC, France	N Atlantic, English Channel
EU729477	*Pavlova virescens*	RCC1535	RCC, France	N Atlantic, France
EU502883	*P. pinguis*	RCC1538	RCC, France	Mediterranean, France

The strains are listed following the branching pattern of the tree in Fig. 2*A* (from top to bottom external black branches). All sequences except DQ980469, AF289038, and AF289040 were generated during this study. Note *Stauronertha* is the new genus name for *Cruciplacolithus* (4).

Table 4. Time, space, and depth information for the 28 worldwide samples used to measure total haptophyte cell size (Fig. S4)

Station	Location	Date	Depth, m
Roscoff (France), SOMLIT-Astan	48°46′N, 3°57′W	May to July 2006; Jan., April, June, Oct. 2007	Subsurface ($n = 6$)
Japan, Station A	40°N, 143°E	May 2006	10, 30
Japan, Station B	40°N, 145°E	May 2006	5, 20, 30, 50, 90
Japan, Station E	34°04′N, 140°E	May 2006	10, 25, 40
Japan, Station F	34°26′N, 139°E	May 2006	Subsurface, 30
Japan, Station G	33°21′N, 140°E	May 2006	Subsurface, 20, 70
Villefranche sur mer (France), SOMLIT Point B	43°41′N, 7°19′E	Sept. 2007	Subsurface, 20, 40, 50, 70, 150, 200

For each depth, water was prefiltered through a 60-μm sieve, and planktonic cells were recovered onto 0.2-μm membranes as in ref. 2.

Table S5. SSU versus LSU (D1–D2 fragment) rDNA Tajima–Nei genetic distances for several couples of haptophytes species

Compared cultured strains	SSU rDNA distance, %	LSU rDNA distance, %
Emiliania huxleyi/Gephyrocapsa oceanica	0.0	0.1
Calcidiscus quadriperforatus/C. leptoporus	0.2	3.1
Umbillcosphaera foliosa/Umbilicosphera sibogae	0.4	4.1
Helicosphera carteri/Scyphosphaera apsteinii	1.2	1.7
Syracosphera pulchra/Coronosphaera mediterrancea	1.2	2.2
Pleurochrysis carterae/P. carterae var. *dentata*	1.3	2.4
Jomonlithus litoralis/Hymenomonas coronata	1.8	3.9
Chrysochromulina acantha/C. throndsenii	0.1	0.9
C. ericina/C. simplex	5.7	10.9
C. hirta/C. brevifilum	1.1	4.2
Prymnesium zebrinum/P. parvum	1.9	5.5
Pavlova pinguis/P. virescens	5.4	11.7

Pairs of sequences were automatically aligned by using ClustalX, and the program Mega 4.0 (5) was used to compute genetic distances. As a mean value, LSU rDNA evolves ≈5 times faster than SSU rDNA. Note that closely related species that split in the Pleistocene, such as *E. huxleyi* and *G. oceanica*, cannot be separated by using SSU rDNA sequences. In addition, no intraspecific variability was detected between the several LSU rDNA clones we sequenced from cultured strains.