

Strong purifying selection in the silicon transporters of marine and freshwater diatoms

Andrew J. Alverson¹

The University of Texas at Austin, Plant Biology Graduate Program, 1 University Station (A6700), Austin, Texas 78712

Abstract

Marine and freshwater diatoms show several important differences in silicon physiology. In addition to containing an order of magnitude more silica in their cell walls, freshwater diatoms also appear to have a less efficient silicic acid uptake mechanism. A novel set of silicon transporters (SITs), encoded by a small gene family, import silicic acid from the environment into the diatom cell. Some evidence suggests that the disparity in uptake efficiency between marine and freshwater diatoms might be attributable to different demands on SITs in the two environments. To test this hypothesis, partial *SIT* genes were cloned and sequenced from 45 species of Thalassiosirales, a diatom lineage with high diversity in marine and freshwaters. Phylogenetically based codon substitution models were used to test whether *SITs* from marine and freshwater taxa were under similar selective constraints and whether codons in different structural locations of the protein were under similar functional constraints. Purifying selection is the predominant evolutionary force acting on *SITs*, irrespective of location in the protein, and differences in efficiency of silicic acid uptake between marine and freshwater diatoms are not due to sequence differences in *SITs*.

Diatoms are best known for their resilient and intricately ornamented cell walls of opaline silica, called “frustules.” Silicon comprises ~25% (by weight) of the earth’s crust (Iler 1979), and it has been estimated that each silica atom, of the $\sim 1.7 \times 10^{12}$ kg annually input into the world’s oceans, is cycled through nearly 40 diatoms before burial in marine sediments (Treguer et al. 1995). The predominant form of dissolved silicon in marine and freshwaters is undissociated orthosilicic acid, $\text{Si}(\text{OH})_4$, which is also the form used by most diatoms (Paasche 1980; Del Amo and Brzezinski 1999). Diatoms are the primary biological mediators of the silica cycle in modern oceans (Conley 2002), where surface concentrations are extremely low, typically 1–10 $\mu\text{mol L}^{-1}$, and are thought to have been so since the marked diversification and ecological expansion of diatoms beginning in the Cenozoic (Martin-Jezequel et al. 2000; Katz et al. 2004). Silicon concentrations are generally much greater in freshwaters, with values reaching 100 $\mu\text{mol L}^{-1}$ in lakes (Paasche 1980) and 150 $\mu\text{mol L}^{-1}$ in rivers (Treguer et al. 1995). Most evidence suggests that diatoms are ancestrally marine, so the many diatom lineages with representatives in freshwaters are the result of independent colonization events. A large body of evidence drawn from a broad sampling of these lineages shows several fundamental

differences in the ways marine and freshwater diatoms use silicon.

Marine and freshwater diatoms differ in both total silica content and efficiency of silicic acid utilization. Freshwater diatoms contain approximately one order of magnitude more silica in their cell walls than marine diatoms, a difference that supercedes the confounding effects of cell biovolume, habitat, and comparisons of natural populations versus cell cultures (Conley et al. 1989). Culture experiments have demonstrated that diatoms grown under limiting silicon concentrations have thinner frustules (Harrison et al. 1977; Brzezinski et al. 1990), so the difference in silica content between marine and freshwater diatoms might reflect the disparity in silicon availability between the two habitats. Sponges and radiolarians also become less silicified in response to low silicon, a trend that has occurred over geological time and is associated with depleted oceanic silicon coincident with the proliferation of diatoms (Moore 1969; Harper and Knoll 1974). The difference in silica content between marine and freshwater diatoms might also reflect gross osmotic differences between the two habitats. Data from another set of culture experiments showed that a freshwater culture strain of *Cyclotella meneghiniana* grew faster (cell divisions per day) and contained less silica in high-salinity treatments (Tuchman et al. 1984). These two responses—increased growth rate and reduced silicification—are probably not independent. Silicon transport is coordinated with cell division in most diatoms, so transport peaks during the G2 and M phases of the cell cycle in concert with valve synthesis (Sullivan 1977; Brzezinski et al. 1990; Brzezinski 1992). Because salinity can promote diatom growth (i.e., cell division rate) (Guillard and Ryther 1962; Olsen and Paasche 1986), the difference in silica content between marine and freshwater diatoms might be a consequence of the hastened cell cycle in marine diatoms, which simply allows less time for silicon transport and deposition (Flynn and Martin-Jezequel 2000; Martin-Jezequel et al. 2000). Intuitively, a lower overall silica budget should be

¹ Present address: Indiana University, Department of Biology, 142 Jordan Hall, 1001 East Third Street, Bloomington, Indiana 47405.

Acknowledgments

Mark Hildebrand and Kim Thamatrakoln provided technical advice and valuable insight on SIT structure and function. Ed Theriot provided valuable input about all aspects of the project, and Gwen Gage provided technical assistance with figures. The comments of two anonymous reviewers helped improve earlier versions of this manuscript. This research was supported by an NSF PEET grant (DEB-0118883) and an NSF Doctoral Dissertation Improvement Grant (DEB-0407815).

advantageous for diatoms in the marine environment, where silicon concentrations are extremely low and potentially growth limiting. A potentially greater selective force on marine diatoms to maintain lightly silicified cells with reduced sinking rates also is present because, unlike diatoms in freshwaters, marine diatoms have less chance of being resuspended after sinking below the well-mixed photic zone (Conley et al. 1989). Although provocative, there is no direct support for this hypothesis (Martin-Jezequel et al. 2000).

A second fundamental difference between marine and freshwater diatoms is in their efficiency of silicic acid use. Uptake of silicic acid by diatoms follows the Michaelis–Menten saturation function, and silicon-limited growth rate (a surrogate measure for uptake) follows the Monod function (Martin-Jezequel et al. 2000). The inferred half-saturation constants (K_s for uptake, K_μ for growth) from these experiments measure a diatom's enzymatic affinity for silicic acid: diatoms with lower K_s and K_μ values are predicted to have a greater enzymatic affinity for silicic acid. These values are commonly used to predict the outcome of interspecific competition for low silicon, whereby the species with the lowest K_s or K_μ is predicted to competitively displace other species (Tilman 1977, 1981). Half-saturation constants for species of Thalassiosirales, which are the focus of this study, reflect the variation across diatoms as a whole, with K_s values of 0.2–7.6 $\mu\text{M L}^{-1}$ and K_μ values of 0.02–8.6 $\mu\text{M L}^{-1}$ (reviewed in Martin-Jezequel et al. 2000). These wide ranges reflect the disparity between marine and freshwater diatoms, with the clear trend that marine species have drastically lower K_s and K_μ values than freshwater species (Paasche 1980; Martin-Jezequel et al. 2000). Taken together, these data suggest that marine diatoms might be constrained to have a lower overall silica budget and substantially greater affinity for silicic acid, possibly because of its scarcity in modern oceans. In contrast, the higher overall silica budget and increased kinetic parameters of freshwater diatoms suggest that this constraint might be relaxed in freshwaters, which tend to have a surplus of silicic acid.

Silicic acid uptake by diatoms is controlled by a small gene family of silicon transporters (SITs) originally characterized from the biraphid pennate diatom *Cylindrotheca fusiformis* (Hildebrand et al. 1997, 1998). SITs have been found in all diatom species investigated for them (Hildebrand et al. 1998; Sherbakova et al. 2005; Thamtrakoln et al. 2006), and a SIT homolog was also reported from chrysophytes (Likoshway et al. 2006). *Cylindrotheca fusiformis* had five distinct SITs, each comprising 10 membrane-spanning segments thought to be the site of silicic acid transport and a coiled-coil motif at the carboxy terminus thought to help regulate SIT activity and localization (Hildebrand et al. 1998). SITs from all Thalassiosirales species investigated lack a C-terminal coiled-coil motif (Thamtrakoln et al. 2006). The regular, coordinated expression of different SITs through the cell cycle suggests some degree of subfunctionalization among paralogs, perhaps allowing for fine control of the timing, affinity, and capacity of transport (Hildebrand et al. 1998; Martin-Jezequel et al. 2000; Thamtrakoln et al. 2006). A

comparative analysis of SIT sequences from eight diverse diatom species identified a conserved motif thought to be important for silicon transport and led to the development of a testable transport model (Thamtrakoln et al. 2006). Surprisingly, the same study showed no clear differences in SITs between marine and freshwater species, though low taxonomic sampling might have limited the statistical power necessary to detect any differences (Thamtrakoln et al. 2006).

High species diversity in both marine and freshwaters and the availability of a densely sampled phylogenetic hypothesis (Alverson 2006) make Thalassiosirales an excellent system for studying the gene phylogeny and molecular evolution of SITs, especially with regard to important differences in silicon physiology between marine and freshwater species. The scarcity of silicic acid in oceans coupled with high demand by diatoms led to the hypothesis that marine diatoms are constrained to maintain a highly efficient uptake system, consistent with observing purifying (negative) selection on SITs. This constraint might have been relaxed upon colonization of freshwaters in which silicic acid concentrations are much greater, resulting in neutral (or near-neutral) evolution of SITs. Alternatively, SITs in freshwater taxa might be evolving under diversifying (positive) selection in response to the vastly different nutrient and osmotic environment in freshwaters. To test these hypotheses, partial SIT genes were amplified by polymerase chain reaction (PCR), cloned, and sequenced from 45 species (26 marine, 19 freshwater) sampled across the Thalassiosirales phylogeny. The inferred gene phylogeny was used to test a number of hypotheses with codon substitution models. First, random-sites models (Yang et al. 2000; Yang and Swanson 2002) were used to test for sites under positive selection without regard to the secondary structure of the protein. Second, exposed (noncytosolic), transmembrane, and internal (cytosolic) SIT segments probably serve different functions in transporting silicic acid across the plasma membrane, so fixed-sites models (Yang and Swanson 2002) were used to test whether sites from these different segments were under similar selective constraints. Finally, clade models (Bielski and Yang 2004) were used to test whether SITs from marine and freshwater taxa were under similar selective constraints. The overall approach of this study was to integrate perspectives from ecology, physiology, and molecular biology to better understand the underlying evolutionary basis of an important ecological trait in diatoms.

Materials and methods

Cell culture and laboratory methods—Diatom cell culture techniques and DNA extraction methods followed Alverson and Kolnick (2005). Degenerate primers were designed from an alignment of full-length SIT sequences identified in the complete genome sequence of *Thalassiosira pseudonana* (Armbrust et al. 2004). Degenerate primers SIT22+ (GGI MGK CAG TTC ATG GTI CT) and SIT1037– (CCR ACG TAI WCT TCG TCG TA) were used to amplify 900–1,000 nucleotide fragments under the following conditions:

96°C for 5 min, 37 cycles of (94°C for 50 s, 50°C for 50 s, 72°C for 60 s), and final extension at 72°C for 7 min. Amplicons were cloned with a TOPO TA Cloning Kit (Invitrogen) following the instructions of the manufacturer but with one-third the reaction size. Transformant *Escherichia coli* was spread onto Luria–Bertani plates containing X-gal (40 mg L⁻¹) and kanamycin (50 µg mL⁻¹) and incubated for 12–16 h. Several transformant colonies per plate were transferred directly to a PCR cocktail and screened for the insert with provided M13 primers and a “hotstart” protocol: 94°C for 10 min, 72°C for 5 min, 37 cycles of (94°C for 50 s, 48°C for 30 s, 72°C for 70 s), and final extension at 72°C for 7 min. Amplicons were purified with QIAquick PCR Purification columns (Qiagen). Forward and reverse strands were cycle sequenced with BigDye (Applied Biosystems) with the PCR primers, and sequences were resolved with an ABI 3700 DNA Analyzer (Applied Biosystems).

Multiple sequence alignment and phylogenetic analysis—SIT nucleotide sequences had very few nucleotide insertions and deletions and were aligned manually with MacClade (ver. 4.08; Maddison and Maddison 2003). Nucleotide sequences were converted to their predicted AA sequences with MacClade for secondary structure prediction. All phylogenetic analyses were based on the nucleotide alignment. MrModeltest (Nylander 2004) was used to determine the most appropriate model of DNA sequence evolution. Hierarchical likelihood ratio tests and the Akaike Information Criterion both favored the general time-reversible model with parameters for gamma-distributed rate heterogeneity (Γ) and a proportion of invariant sites (I). The maximum likelihood tree was found by GARLI (Zwickl 2006), with all model parameters estimated from the analysis. Default settings were used, and the algorithm was run multiple times with different combinations of random number seeds and starting tree topologies to verify convergence. Final branch lengths and log likelihood scores were calculated with PAUP* (ver. 4.0b10; Swofford 2001). The computational speed of GARLI made it possible to assess branch support with a full-optimization maximum likelihood bootstrap analysis. Five hundred pseudoreplicates were performed with default settings, and bootstrap proportions were calculated by computing a 50% majority rule consensus tree with PAUP*.

SIT secondary structure—Four tools for prediction of transmembrane domains were used to determine the SIT topology: HMMTOP 2.0 (Tusnady and Simon 2001), Phdhtm (Rost et al. 1996), SOSUI (Hirokawa et al. 1998), and TMAP (Persson and Argos 1997). There was overall agreement among the four methods about the core SIT structure, and final transmembrane boundaries were based on a consensus among the methods. Greater weight was given to HMMTOP 2.0 and Phdhtm, which returned similar results and have been shown to be accurate predictors of topology (Cuthbertson et al. 2005).

Tests for positive and divergent selection—Codon models that consider the ratio of nonsynonymous (dn, changes the

amino acid) to synonymous (ds, does not change the amino acid) rates of nucleotide substitution are commonly used to study the molecular evolution of protein coding DNA sequences (Yang 2002). These models can be used to identify codons evolving under different dn : ds substitution rate ratios ($\omega = \text{dn} : \text{ds}$), whereby $\omega < 1$, $\omega = 1$, and $\omega > 1$, indicate purifying (negative) selection, neutral evolution, and diversifying (positive) selection, respectively (Yang 2002). For example, $\omega > 1$ indicates that amino acid-altering substitutions are being fixed at a higher rate than those that do not change the amino acid, which provides compelling evidence for positive selection. Standard codon models were fit to the dataset with the PAML software package (Yang 1997), and likelihood ratio tests were used to compare the relative fit of the different hierarchically nested models. Likelihood ratio test statistics were compared to a chi-square distribution to determine whether the more complex model provided a significantly better fit than the more restricted model. Model descriptions and model comparison strategies for these analyses are outlined in original articles on the subject (e.g., Muse and Gaut 1994; Yang and Swanson 2002; Bielawski and Yang 2004). Model nomenclature follows Yang and Swanson (2002) and Bielawski and Yang (2004). Random-sites models M0, M1a, M2a, M3, M7, and M8 were fit to the dataset; these models identify variation in ω among sites but not among lineages. Likelihood ratio tests to identify positively selected sites compared the following pairs of models: null model M1a (neutral) to alternative model M2a (selection) and null model M7 (beta) to alternative model M8 (beta& ω). In addition to PAML analyses, the single-likelihood ancestor counting (SLAC) method implemented in HyPhy was used to characterize ω among sites and to identify sites evolving under positive selection (Kosakovsky Pond and Frost 2005; Kosakovsky Pond et al. 2005). This analysis was used to independently verify the PAML results and determine whether inferences were sensitive to variation in the rate of synonymous substitution among sites (Kosakovsky Pond and Frost 2005).

Internal and external segments are directly exposed to ambient internal and external silicon concentrations, respectively, which in part govern the influx or efflux of silicon via membrane-spanning segments (Martin-Jezequel et al. 2000). It is possible that sites in these different structural classes experience different selection pressures, a hypothesis that was tested with fixed-sites models (Yang and Swanson 2002). For fixed-sites models, nucleotide sequences were divided into one class of sites from internal and external segments and a second class of sites from transmembrane segments of the protein. Codon models A–E (Yang and Swanson 2002) were fit to the dataset using PAML; likelihood ratio tests to identify positively selected sites compared the following two pairs of models: null model B (or C) to alternative model D and null model C to alternative model E.

With no a priori expectation of positive selection or relaxed evolutionary constraint in freshwater lineages, the second set of analyses used “clade” models that provide a more general test for divergent selection in the lineages of interest (Bielawski and Yang 2004), rather than branch-site

models designed more specifically for tests of positive selection (Yang and Nielsen 2002). Clade model D (Bielawski and Yang 2004) was fit to the dataset and tested against null model M3 to determine whether a proportion of the sites in freshwater *SITs* were under different evolutionary constraints than marine *SITs*.

All PAML analyses were run from multiple starting values of ω to verify convergence. In addition to the extraordinarily divergent “*SIT3*” types recovered only from *T. pseudonana* (Thamatrakoln et al. 2006), two partial sequences (DQ482555 and DQ482533) were excluded from selection analyses.

Results

Alignment, sequence characterization, and SIT secondary structure—The final alignment contained 97 unique sequences (Web Appendix 1, http://www.aslo.org/lo/toc/vol_52/issue_4/1420a1.pdf). Two large indels were present in the alignment. One highly divergent *SIT* paralog from *T. pseudonana* (TpSIT3 in Thamatrakoln et al. 2006) shared a 4-AA indel (AA residues 136–139) with *SITs* from *P. glacialis*, *P. pseudodenticulata*, and *L. annulata*. The latter three species comprise the sister clade to all other Thalassiosirales (Alverson 2006). Together, the highly divergent *SIT3* sequence in *T. pseudonana*, its shared indel with basal species of Thalassiosirales, and its sister relationship to these species in unrooted trees (results not shown) support the hypothesis that the similarity of TpSIT3 to other *SIT* types traces to a gene duplication that predated the origin of Thalassiosirales. Thus, the *SIT* gene tree was rooted at this split (Fig. 1). The *SIT3* type of *T. pseudonana* was subsequently removed from selection analyses because its high divergence—even among diatoms as a whole (Thamatrakoln et al. 2006)—suggests that it might not function directly in silicon transport. Finally, *Thalassiosira oceanica* had a 72-nt intron, in frame, between transmembrane segments 3 and 4.

Despite minor discrepancies about exact boundaries of the transmembrane segments, there was overall agreement among the four methods about the core *SIT* structure. HMMTOP 2.0, Phdhtm, and SOSUI predicted seven transmembrane segments. SOSUI, which considers single sequences rather than multiple sequence alignments, predicted five for some sequences, and TMAP consistently predicted five. The five transmembrane domains predicted by SOSUI and TMAP were always a subset of the seven segments predicted by HMMTOP 2.0, Phdhtm, and SOSUI. Thus, the final consensus model had seven transmembrane segments (AA residues 17–35, 50–70, 106–125, 145–167, 204–224, 250–267, 274–294) separated by four internal (AA residues 1–16, 71–105, 168–203, 268–273) and four external segments (AA residues 36–49, 126–144, 225–249, 295–300) (Fig. 2). These seven transmembrane segments correspond to segments 4–10 predicted by Thamatrakoln et al. (2006).

SIT phylogeny—Three *SITs* were found in the complete genome sequence of *T. pseudonana* (Armbrust et al. 2004), and likewise multiple *SIT* types (typically ≤ 3) were found

in nearly every species examined, all of which appeared orthologous to the *SIT1* and *SIT2* types from *T. pseudonana* (Armbrust et al. 2004; Thamatrakoln et al. 2006). No orthologs of the *SIT3* type from *T. pseudonana* were recovered from any other species of Thalassiosirales, which included sampling of *T. pseudonana*'s closest known relatives (Alverson 2006). In many cases, multiple types recovered from the same species or culture strain grouped together (Fig. 1), making it difficult to discern allelic variation from lineage-specific gene duplications. Obvious gene duplication events were identifiable, especially in the densely sampled freshwater lineage (clade D, Fig. 1). For example, distantly related *SIT* types were found in *Discostella pseudostelligera*, *Stephanodiscus reimerii*, and *Stephanodiscus yellowstonensis* (Fig. 1); the relationship between distantly related types within these species most likely traces to an earlier gene duplication (Fig. 1). In another example, the expected pattern of orthology was recovered in the two *SIT* types from the closely related species *Stephanodiscus agassizensis* and *Stephanodiscus neoastreae* (Fig. 1).

The *SIT* gene phylogeny was grossly similar to the underlying species phylogeny inferred from nuclear and chloroplast DNA (Alverson 2006). *SITs* from the 19 freshwater species sorted into four separate clades spaced throughout the gene tree (Fig. 1). Two of these clades (Fig. 1, A and B) represented different paralogous gene lineages nested within the *Cyclotella meneghiniana* species complex, and one clade included the three *SIT* types recovered from the secondarily freshwater species, *Thalassiosira gessneri* (Fig. 1, C). The other freshwater clade contained *SIT* types from *Discostella*, *Cyclostephanos*, and *Stephanodiscus*, which together with several unrepresented genera compose the most species-rich lineage of freshwater thalassiosiroids (Fig. 1, D).

Molecular evolution of marine and freshwater SITs—Random-sites models were used to determine whether ω varied among sites (Nielsen and Yang 1998; Yang et al. 2000). Model M0 assumes a single, average ω across all sites and all branches in the tree, and the estimate of $\omega = 0.069$ indicated that purifying selection is the predominant evolutionary force acting on *SITs* (Table 1). Model M1a, which assumes one class of conserved sites ($0 < \omega < 1$) and one class of neutrally evolving sites ($\omega = 1$), provided a significantly better fit to the data than M0 (Tables 1, 2). Model M2a, which adds an additional site class of $\omega > 1$ to accommodate sites evolving under positive selection, did not improve over M1a ($l_{M1a} = l_{M2a} = -26,343.967$) (Tables 1, 2), indicating that no sites are evolving under positive selection. Model M3 fits $k = 2$ or $k = 3$ ω classes estimated from the data. Both variants of model M3 provided a significant improvement over M0, indicating significant variation in ω among sites (Tables 1, 2). Model M3 ($k = 3$) indicated that 10% of sites are under substantially reduced functional constraint ($\omega = 0.404$) compared with the remaining 90% of sites ($\omega < 0.112$) (Tables 1, 2). Model M7 fits a beta-distribution of ω to the data that is limited to the interval (0, 1); M8 adds an additional site class of $\omega > 1$ to M7, so the M7–M8

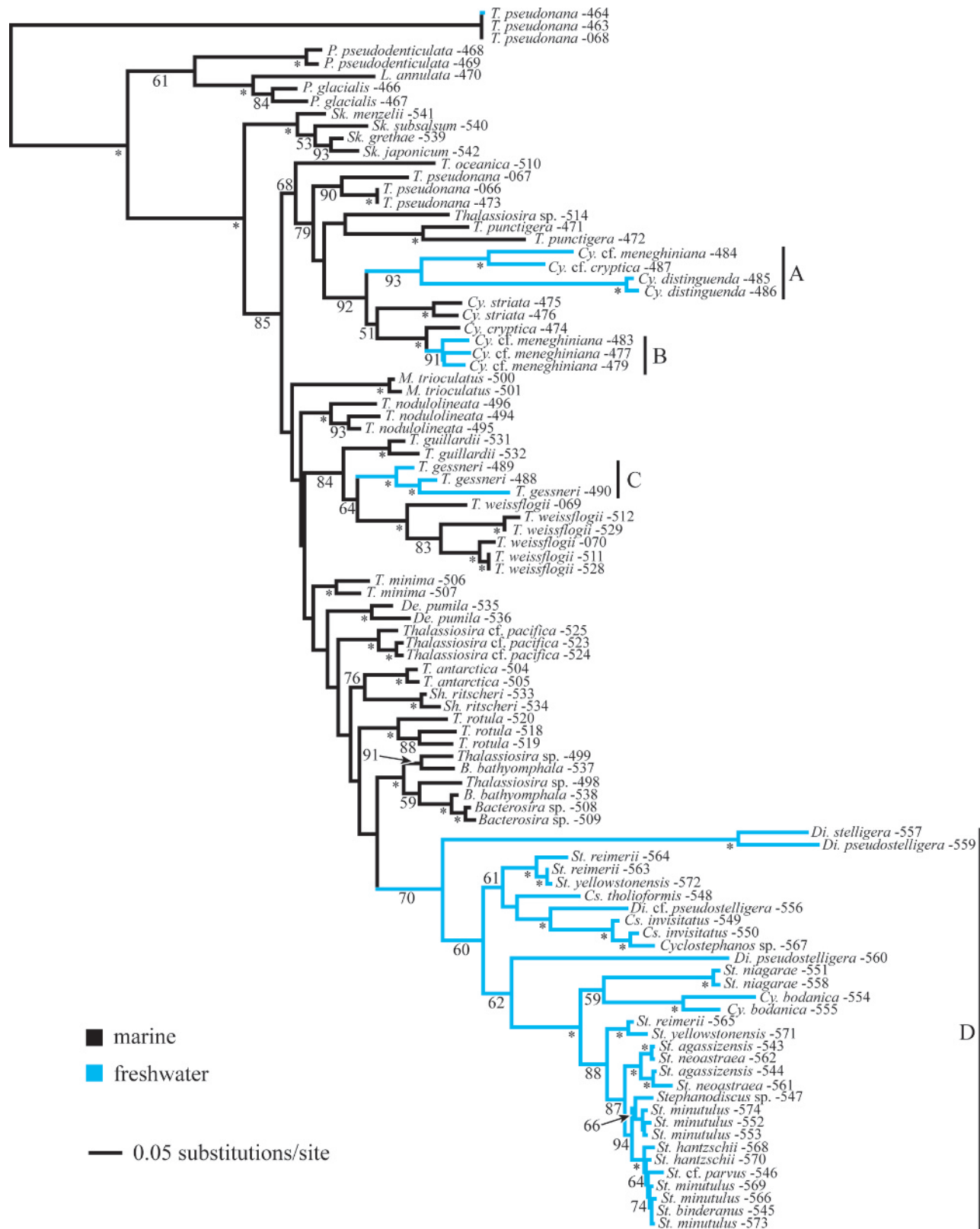


Fig. 1. Gene phylogeny of silicon transporter genes (*SITs*) from thalassiosiroid diatoms. Each *SIT* is identified by species name and the last three digits of its GenBank accession ID (Web Appendix 1). *SITs* from freshwater species (blue) were treated as “foreground” branches in tests for divergent selection (see text). Bootstrap support values greater than 95% are shown with an asterisk (*). Four freshwater *SIT* lineages (A–D) are labeled to facilitate discussion in the text.

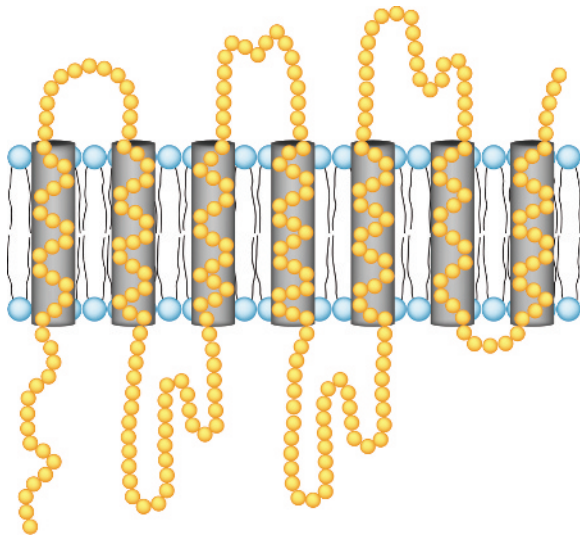


Fig. 2. Secondary structure of the last seven transmembrane segments of silicon transporters in Thalassiosirales, inferred from an alignment of partial *SIT* gene sequences. Amino acid residues are shown by circles, and transmembrane boundaries are demarcated by cylinders embedded within the plasma membrane. The upper segments connecting transmembrane domains are noncytosolic, and the lower segments represent internal, cytosolic segments.

comparison provides another test of positive selection. Models M7 and M8 fit the data equally well ($l = -25,841.426$), providing further evidence that no sites are evolving under positive selection (Tables 1, 2). The estimated beta distribution $\beta(0.49, 3.98)$ from M7 and M8 is an extreme “L” shape, indicating that most sites have

$\omega \approx 0$. Finally, the SLAC method implemented in HyPhy validated results of the PAML analyses, identifying 261 sites under purifying selection and no sites under positive selection.

Fixed-sites models were fit to the dataset to determine whether transmembrane sites were under different selective constraints than exposed sites on the cell interior and exterior. The first null model fit to the dataset corresponds to random-sites model M0, which assumes a single, homogeneous value of ω across all sites, regardless of position in the protein secondary structure. Model B, which assumes a different rate of evolution between the two site classes (transmembrane vs. exposed), actually provided a worse fit to the data than M0, indicating that sites in the two classes have similar rates of evolution (Tables 3, 4). Increasingly complex fixed-sites models add additional parameters to account for potential variation in codon frequency (model C), ω (model D), and transition:transversion rate (model E) (Table 3). Although each sequentially more complex model provided a significantly better fit to the data (e.g., C better than B, D better than C, etc.), their surprising lack of improvement over the homogeneous M0 model is the most telling result, indicating that exposed and transmembrane sites are evolving under similarly strong purifying selection (Tables 3, 4).

Clade models (Bielawski and Yang 2004) were used to test for divergent selection between marine and freshwater lineages. For these analyses, all freshwater *SITs* and subtending ancestral branches were treated as “foreground” lineages, and marine branches were treated as “background” lineages (Fig. 1). Again, the null model M0 (Tables 1–3) shows the dominant role of purifying selection in *SIT* evolution ($\omega = 0.069$), and models M3 ($k = 2$) and

Table 1. Number of free parameters (p), log likelihood (l), and parameter estimates for random-sites and clade models. See text for model details.

Model	p	l	Parameter estimates*	Positively selected sites
M0 (homogeneous)	1	-26,724.349	$\omega = 0.069$	Not allowed
M1a (neutral)	2	-26,343.967	$\omega_0 = 0.060, p_0 = 0.92$ $\omega_1 = 1.000, p_1 = 0.08$	None
M2a (selection)	4	-26,343.967	$\omega_0 = 0.060, p_0 = 0.92$ $\omega_1 = 1.000, p_1 = 0.06$ $\omega_2 = 1.000, p_2 = 0.02$	None
M3 (discrete, $k = 2$)	3	-25,978.351	$\omega_0 = 0.028, p_0 = 0.73$ $\omega_1 = 0.221, p_1 = 0.27$	None
M3 (discrete, $k = 3$)	5	-25,854.429	$\omega_0 = 0.018, p_0 = 0.58$ $\omega_1 = 0.112, p_1 = 0.32$ $\omega_2 = 0.404, p_2 = 0.10$	None
M7 (beta)	2	-25,841.426	$\beta(0.49, 3.98)$ $\omega < 0.40, p = 1.00$	Not allowed
M8 (beta& ω)	4	-25,841.426	$\beta(0.49, 3.98)$ $\omega_0 < 0.40, p_0 = 1.00$	None
D (clade, $k = 2$)	4	-25,969.303	$\omega_0 = 0.028, p_0 = 0.73$ $\omega_{\text{marine}} = 0.251,$ $\omega_{\text{freshwater}} = 0.179, p_1 = 0.27$	None
D (clade, $k = 3$)	6	-25,835.950	$\omega_0 = 0.019, p_0 = 0.60$ $\omega_1 = 0.120, p_1 = 0.31$ $\omega_{\text{marine}} = 0.542,$ $\omega_{\text{freshwater}} = 0.266, p_2 = 0.09$	None

* ω , dn:ds ratio; p_n , proportion of sites in ω class n ; $\beta(n_1, n_2)$, shape parameters (n_1, n_2) for beta distribution of ω .

Table 2. Likelihood ratio test statistics ($2\Delta l$), degrees of freedom (df), and p -values for comparisons of random-sites and clade models. See text for model details.

	$2\Delta l$	df	p -value
M0 vs. M1	760.8	1	<0.0001**
M1 vs. M2	0.0	2	1.0
M0 vs. M3 ($k = 2$)	1,492.0	2	<0.0001**
M0 vs. M3 ($k = 3$)	1,739.8	4	<0.0001**
M3 ($k = 2$) vs. M3 ($k = 3$)	247.8	2	<0.0001**
M3 ($k = 2$) vs. D ($k = 2$)	18.1	1	<0.0001**
M3 ($k = 3$) vs. D ($k = 3$)	37.0	1	<0.0001**
D ($k = 2$) vs. D ($k = 3$)	266.7	2	<0.0001**

** Significant at $p < 0.05$.

M3 ($k = 3$) indicate significant variation in ω among sites (Tables 1, 2). These models do not account for potential variation in ω among lineages, as might be expected between marine and freshwater *SITs*; therefore, clade model D (Bielawski and Yang 2004) was used to test for divergent selection between marine and freshwater *SITs*. Model D ($k = 2$) fits one ω site class that is invariable among marine and freshwater lineages and one site class with variable ω among marine and freshwater lineages (Table 1). Model D ($k = 3$) fits two ω site classes that are invariable among lineages and one class with variable ω between marine and freshwater lineages. Both variations of Model D fit the data significantly better than models that do not account for variation in ω among lineages (Tables 1, 2). Model D ($k = 3$), which provided the best fit, showed that 9% of sites were under different selective constraints in marine and freshwater *SITs* (Table 1). Surprisingly, these divergent sites were under stronger purifying selection in freshwater *SITs* ($\omega = 0.266$) than in marine *SITs* ($\omega = 0.542$).

Discussion

A large and longstanding body of evidence shows several important differences in the ways marine and freshwater diatoms use silicon (e.g., Olsen and Paasche 1986; Conley et al. 1989; Martin-Jezequel et al. 2000). The more efficient silicic acid uptake kinetics of marine diatoms was recognized as early as 1980 (Paasche 1980), and it was later shown that freshwater diatoms contain fully one order of magnitude more silica in their cell walls than marine

diatoms (Conley et al. 1989). The underlying causes of these differences are largely unknown. The characterization of silicon transporter proteins and the genes that encode them provided the first insight into the molecular basis of silicon uptake by diatoms (Hildebrand et al. 1997, 1998). The complete genome sequence of *T. pseudonana* then made it practical to explore the molecular evolution of *SITs* in Thalassiosirales, a lineage with high species diversity in marine and freshwater habitats. In addition, many of the foundational experiments on diatom silicon physiology used Thalassiosirales (e.g., Paasche 1973; Tilman and Kilham 1976; Brzezinski 1992), further compelling its use in a comparative analysis of marine and freshwater *SITs*.

Diatom *SITs*, including those of *T. pseudonana*, are comprised of a sequence of 10 transmembrane segments connected by alternating internal and external segments (Thamatrakoln et al. 2006). In general, diatoms undergo three different modes of silicic acid uptake: surge, externally controlled, and internally controlled (Conway et al. 1976; Conway and Harrison 1977; Martin-Jezequel et al. 2000). Externally and internally controlled uptakes are governed by external and internal silicic acid concentrations, respectively, so these exposed segments of the protein might be involved in silicon detection and the regulation of externally and internally controlled uptake. Exposed segments encounter a broad range of osmotic, silicic acid, and nutrient conditions, so one hypothesis was that some sites in exposed segments might have undergone positive selection to optimize their function in the vastly different conditions encountered in marine and freshwaters. In contrast, transmembrane segments are thought to mediate influx and efflux directly and so might be under tighter selective constraints. By identifying sites in these different structural regions, random- and fixed-sites models could be used to test these hypotheses directly. No sites showed evidence of positive selection, regardless of location in the secondary structure of the protein. Random-sites models showed evidence for some variability in constraint among sites, with 90% of sites under strong purifying selection ($0.018 < \omega < 0.112$) and the remaining 10% under somewhat weaker constraint ($\omega = 0.404$). Although *SIT* secondary structure was predicted as accurately as possible, exact boundaries of transmembrane segments cannot be determined algorithmically without some error (Cuthbertson et al. 2005). Naturally, fixed-sites models will perform poorly when sites are not partitioned into the correct

Table 3. Number of free parameters (p , including parameters for codon frequencies), log likelihood (l), and parameter estimates for fixed-site models that distinguished internal and external segments (partition 1) from transmembrane segments (partition 2). See text for model details.

Model*	p	l	rs	κ	ω
A (homogeneous)	11	-26,724.349	—	1.50	0.0694
B (different rs)	12	-27,297.443	0.996	1.52	0.0689
C (different $rs, \pi s$)	21	-27,327.963	0.895	1.51	0.0683
D (different rs, κ, ω)	14	-27,294.354	0.972	$\kappa_1 = 1.46$ $\kappa_2 = 1.62$	$\omega_1 = 0.0653$ $\omega_2 = 0.0749$
E (different $rs, \kappa, \omega, \pi s$)	23	-27,324.504	0.880	$\kappa_1 = 1.43$ $\kappa_2 = 1.66$	$\omega_1 = 0.0654$ $\omega_2 = 0.0734$

* rs , nucleotide substitution rate; πs , codon frequencies; κ , nucleotide transition : transversion ratio; ω , dn : ds ratio.

Table 4. Likelihood ratio test statistics ($2\Delta l$), degrees of freedom (df), and p -values for comparisons of fixed-sites models. See text for model nomenclature.

	$2\Delta l$	df	p -value
A vs. B	-1,146.2	1	1
B vs. C	61.0	9	1
B vs. D	6.2	2	0.046**
C vs. E	6.9	2	0.032**

** Significant at $p < 0.05$.

structural classes (Yang and Swanson 2002). Importantly, however, fixed- and random-sites models perform equally well in the identification of positively selected sites (Yang and Swanson 2002). In light of this and the similar finding by both model types of strong purifying selection across sites in this dataset, any error in the *SIT* structure prediction probably had little effect on the overall conclusions of this study.

Among the many differences between marine and freshwaters, freshwaters generally are replete with silicic acid ($\sim 100 \mu\text{mol L}^{-1}$) compared with the uniformly low surface concentrations across most of the world's oceans ($\sim 10 \mu\text{mol L}^{-1}$) (Paasche 1980; Treguer et al. 1995; Martin-Jezequel et al. 2000). Low oceanic silicic acid concentrations are attributed to high demand by diatoms, which presumably need to maintain a high affinity for silicic acid to compete successfully for this potentially growth-limiting resource. The extremely efficient silicic acid uptake kinetics of marine diatoms are consistent with this hypothesis (Paasche 1980; Martin-Jezequel et al. 2000) and led to the prediction that *SITs* of marine species are under strong purifying selection to maintain optimal function. The situation in freshwaters is quite different, where silicic acid concentrations are much greater, and the silicic acid uptake kinetics of freshwater diatoms are much greater—representing sluggish, inefficient uptake compared with their marine counterparts (Paasche 1980; Treguer et al. 1995; Martin-Jezequel et al. 2000). This evidence led to two alternative hypotheses. First, competition for low, growth-limiting silicic acid in oceans might act as constraint on *SIT* evolution. This constraint might be relaxed in freshwaters, where silicic acid is considerably more abundant. Relaxation of this constraint might manifest as near-neutral evolution on freshwater *SITs* (i.e., $\omega_{\text{marine}} < \omega_{\text{freshwater}} \leq 1$). Alternatively, some residues in freshwater *SITs* might be under positive selection to optimize their function in a drastically different osmotic and nutrient environment. In fact, the magnitude of purifying selection was fairly uniform across marine and freshwater *SITs*, so each of these hypotheses was rejected. Surprisingly, results actually showed some evidence for selection in the opposite direction of what was expected, with 9% of sites under substantially stronger purifying selection in freshwater *SITs*.

Taken together, these results indicate that differences in the efficiency of silicic acid uptake between marine and freshwater diatoms are not attributable to differences in *SIT* coding sequences. One possible explanation is that *SIT*

structure and function had already been optimized in the common ancestor of Thalassiosirales, and purifying selection has suppressed substantive deviations from the ancestral *SIT* type. Surprising evidence was found that freshwater *SITs* might be under stronger selective constraints than marine *SITs*, discounting initial hypotheses that either positive selection had optimized them for function in freshwater or high silicic acid concentrations had relaxed the constraint imposed by its growth-limiting concentrations in the ancestral marine environment. Analyses that tested specifically for positive selection in the ancestral branches subtending freshwater *SIT* lineages showed no evidence for episodic changes in *SITs* upon colonization of freshwater (results not shown), which further suggests that freshwaters do not alter the selective constraints on *SITs*.

A few caveats on the interpretation of these results should be considered. First, this study examined only seven of the 10 total transmembrane segments. More importantly, the long, external N-terminus was not included in these analyses, so any inferences made here might not apply to this different functional domain. Second, inferences about the gene phylogeny are limited because of potential biases introduced by the experimental protocol. It is clear, for example, that PCR primers preferentially amplified *SITs* orthologous to the *SIT1* and *SIT2* types of *T. pseudonana* (Thamatrakoln et al. 2006). In fact, *SIT3*-specific primers did not amplify *SIT3* orthologs in species other than *T. pseudonana*. Thus, the potentially important role of the *T. pseudonana* *SIT3* type, if present throughout Thalassiosirales, was not addressed in this study. Inferences about the gene phylogeny were limited by the experimental approach because phylogenetic patterns inferred as gene duplications or losses could simply be caused by incomplete or biased sampling of *SITs* from different species (e.g., by PCR bias or undersampling of clones). A broad-scale Southern hybridization study across Thalassiosirales would provide valuable insights into the *SIT* gene phylogeny by identifying putative gene duplications and losses that could be targeted for follow-up study. A fully sampled gene phylogeny might reveal whether paralogous gene lineages diversified after gene duplication (Bielawski and Yang 2004). The regular, coordinated expression of different *SITs* over the cell cycle suggests some degree of sub-functionalization among *SIT* types (Hildebrand et al. 1998), a hypothesis that could be tested comparatively with a fully sampled gene phylogeny (Bielawski and Yang 2004).

Repeated measurements from a diverse set of diatom species show that freshwater species have strikingly higher K_s and K_u values for silicic acid than marine species (Paasche 1980; Martin-Jezequel et al. 2000), suggesting a less efficient uptake mechanism. Results from this study suggest that this difference is not attributable to differences in the gene sequences of marine and freshwater *SITs*, which makes this longstanding question even more vexing. Future investigations might examine the relative expression levels of *SITs* in marine and freshwater species, which would require comparisons of orthologous *SITs*. To this same end, a comparative analysis of promoter sequences might

also be informative. Finally, silicic acid transport is sodium-coupled in the marine pennate species *Nitzschia alba* (Bhattacharyya and Volcani 1980), whereas transport appeared sodium and possibly potassium coupled in the freshwater pennate species *Navicula pelliculosa* (Sullivan 1976). Some experimental evidence and comparative sequence data suggest a possible role of zinc in silicic acid transport (Rueter and Morel 1981; Grachev et al. 2005). If silicic acid transport is strictly sodium coupled, the marine environment might simply be more favorable to uptake. Notably, the euryhaline species *Thalassiosira pseudonana* had significantly reduced K_s values for silicic acid in high salinity compared with freshwater treatments (Olsen and Paasche 1986). Follow-up studies on freshwater species sampled across the diatom phylogeny should verify the exact ionic coupling of silicic acid transport in freshwater diatoms. Differences in ionic coupling might underlie some of the disparity in K_s values between marine and freshwater species. If silicic acid transport is, for example, found to be potassium coupled in a number of distantly related freshwater lineages, this would provide strong evidence for an adaptive change correlated with the colonization of freshwaters.

A central goal in diatom research is to discover the underlying molecular bases of those physiological traits that make diatoms so important in the structure and function of aquatic ecosystems, as well as in the global cycling of biologically important elements, particularly carbon and silica. The complete genome sequence of *T. pseudonana* was a milestone in this effort (Armbrust et al. 2004), and it along with a growing number of other studies in the Thalassiosirales are establishing this group as the premier system for investigating important questions in diatom biology. Availability of a densely sampled phylogenetic hypothesis that includes a diversity of marine and freshwater species (Alverson 2006) make this group particularly suited to powerful comparative studies, including those focused on understanding other physiological differences between marine and freshwater diatoms.

References

- ALVERSON, A. J. 2006. Phylogeny and evolutionary ecology of thalassiosiroid diatoms. Ph.D. thesis, The Univ. of Texas at Austin.
- , AND L. KOLNICK. 2005. Intragenomic nucleotide polymorphism among small subunit (18S) rDNA paralogs in the diatom genus *Skeletonema* (Bacillariophyta). *J. Phycol.* **41**: 1248–1257.
- ARMBRUST, E. V., AND OTHERS. 2004. The genome of the diatom *Thalassiosira pseudonana*: Ecology, evolution, and metabolism. *Science* **306**: 79–86.
- BHATTACHARYYA, P., AND B. E. VOLCANI. 1980. Sodium-dependent silicate transport in the apochlorotic marine diatom *Nitzschia alba*. *Proc. Natl. Acad. Sci. USA* **77**: 6386–6390.
- BIELAWSKI, J. P., AND Z. H. YANG. 2004. A maximum likelihood method for detecting functional divergence at individual codon sites, with application to gene family evolution. *J. Mol. Evol.* **59**: 121–132.
- BRZEZINSKI, M. A. 1992. Cell-cycle effects on the kinetics of silicic acid uptake and resource competition among diatoms. *J. Plankton Res.* **14**: 1511–1539.
- , R. J. OLSON, AND S. W. CHISHOLM. 1990. Silicon availability and cell-cycle progression in marine diatoms. *Mar. Ecol. Prog. Ser.* **67**: 83–96.
- CONLEY, D. J. 2002. Terrestrial ecosystems and the global biogeochemical silica cycle. *Global Biogeochemical Cycles* **16**: 1121, DOI:10.1029/2002GB001894.
- , S. S. KILHAM, AND E. THERIOT. 1989. Differences in silica content between marine and freshwater diatoms. *Limnol. Oceanogr.* **34**: 205–213.
- CONWAY, H., AND P. HARRISON. 1977. Marine diatoms grown in chemostats under silicate or ammonium limitation IV. Transient response of *Chaetoceros debilis*, *Skeletonema costatum*, and *Thalassiosira gravida* to a single addition of the limiting nutrient. *Mar. Biol.* **43**: 33–43.
- , ———, AND C. DAVIS. 1976. Marine diatoms grown in chemostats under silicate or ammonium limitation II. Transient response of *Skeletonema costatum* to a single addition of the limiting nutrient. *Mar. Biol.* **35**: 187–199.
- CUTHERBERTSON, J. M., D. A. DOYLE, AND M. S. P. SANSOM. 2005. Transmembrane helix prediction: A comparative evaluation and analysis. *Protein Eng. Des. Sel.* **18**: 295–308.
- DEL AMO, Y., AND M. A. BRZEZINSKI. 1999. The chemical form of dissolved Si taken up by marine diatoms. *J. Phycol.* **35**: 1162–1170.
- FLYNN, K. J., AND V. MARTIN-JEZEQUEL. 2000. Modelling Si-N-limited growth of diatoms. *J. Plankton Res.* **22**: 447–472.
- GRACHEV, M., T. SHERBAKOVA, Y. MASYUKOVA, AND Y. LIKHOSHWAY. 2005. A potential zinc-binding motif in silicic acid transport proteins of diatoms. *Diatom Res.* **20**: 409–411.
- GUILLARD, R. R. L., AND J. H. RYTHER. 1962. Studies of marine planktonic diatoms. I. *Cyclotella nana* Hustedt and *Detonula confervacea* (Cleve) Gran. *Can. J. Microbiol.* **8**: 229–239.
- HARPER, H. E., AND A. H. KNOLL. 1974. Silica, diatoms, and Cenozoic radiolarian evolution. *Geology* **3**: 175–177.
- HARRISON, P. J., H. L. CONWAY, R. W. HOLMES, AND C. O. DAVIS. 1977. Marine diatoms grown in chemostats under silicate or ammonium limitation. 3. Cellular chemical composition and morphology of *Chaetoceros debilis*, *Skeletonema costatum*, and *Thalassiosira gravida*. *Mar. Biol.* **43**: 19–31.
- HILDEBRAND, M., K. DAHLIN, AND B. E. VOLCANI. 1998. Characterization of a silicon transporter gene family in *Cylindrotheca fusiformis*: Sequences, expression analysis, and identification of homologs in other diatoms. *Mol. Gen. Genet.* **260**: 480–486.
- , B. E. VOLCANI, W. GASSMANN, AND J. I. SCHROEDER. 1997. A gene family of silicon transporters. *Nature* **385**: 688–689.
- HIROKAWA, T., S. BOON-CHIENG, AND S. MITAKU. 1998. SOSUI: Classification and secondary structure prediction system for membrane proteins. *Bioinformatics* **14**: 378–379.
- ILER, R. K. 1979. The chemistry of silica: Solubility, polymerization, colloid and surface properties, and biochemistry. Wiley.
- KATZ, M. E., Z. V. FINKEL, D. GRZEBYK, A. H. KNOLL, AND P. G. FALKOWSKI. 2004. Evolutionary trajectories and biogeochemical impacts of marine eukaryotic phytoplankton. *Ann. Rev. Ecol. Syst.* **35**: 523–556.
- KOSAKOVSKY POND, S. L., AND S. D. W. FROST. 2005. Not so different after all: A comparison of methods for detecting amino-acid sites under selection. *Mol. Biol. Evol.* **22**: 1208–1222.
- , AND S. V. MUSE. 2005. HyPhy: Hypothesis testing using phylogenies. *Bioinformatics* **21**: 676–679.
- LIKOSHWAY, Y. V., Y. A. MASYUKOVA, T. A. SHERBAKOVA, D. P. PETROVA, AND M. A. GRACHEV. 2006. Detection of the gene responsible for silicic acid transport in chrysophycean algae. *Dok. Akad. Nauk SSSR* **408**: 845–849.

- MADDISON, W. P., AND D. R. MADDISON. 2003. MacClade. Sinauer.
- MARTIN-JEZEQUEL, V., M. HILDEBRAND, AND M. A. BRZEZINSKI. 2000. Silicon metabolism in diatoms: Implications for growth. *J. Phycol.* **36**: 821–840.
- MOORE, T. C. 1969. Radiolaria: Change in skeletal weight and resistance to solution. *Geol. Soc. Am. Bull.* **80**: 2103–2108.
- MUSE, S. V., AND B. S. GAUT. 1994. A likelihood approach for comparing synonymous and nonsynonymous nucleotide substitution rates, with application to the chloroplast genome. *Mol. Biol. Evol.* **11**: 715–724.
- NIELSEN, R., AND Z. H. YANG. 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. *Genetics* **148**: 929–936.
- NYLANDER, J. A. A. 2004. MrModeltest v2. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- OLSEN, S., AND E. PAASCHE. 1986. Variable kinetics of silicon limited growth in *Thalassiosira pseudonana* (Bacillariophyceae) in response to changed chemical composition of the growth medium. *Br. Phycol. J.* **21**: 183–190.
- PAASCHE, E. 1973. Silicon and the ecology of marine plankton diatoms. I. *Thalassiosira pseudonana* (*Cyclotella nana*) grown in a chemostat with silicate as limiting nutrient. *Mar. Biol.* **19**: 117–126.
- . 1980. Silicon, p. 259–284. In I. Morris [ed.], *The physiological ecology of phytoplankton*. Univ. of California Press.
- PERSSON, B., AND P. ARGOS. 1997. Prediction of membrane protein topology utilizing multiple sequence alignments. *J. Protein Chem.* **16**: 453–457.
- ROST, B., P. FARISELLI, AND R. CASADIO. 1996. Topology prediction for helical transmembrane proteins at 86% accuracy. *Protein Sci.* **5**: 1704–1718.
- RUETER, J. G., AND F. M. M. MOREL. 1981. The interaction between zinc deficiency and copper toxicity as it affects the silicic-acid uptake mechanisms in *Thalassiosira pseudonana*. *Limnol. Oceanogr.* **26**: 67–73.
- SHERBAKOVA, T. A., AND OTHERS. 2005. Conserved motif CMLD in silicic acid transport proteins of diatoms. *Mol. Biol.* **39**: 269–280.
- SULLIVAN, C. W. 1976. Diatom mineralization of silicic acid. I. Si(OH)₄ transport characteristics in *Navicula pelliculosa*. *J. Phycol.* **12**: 390–396.
- . 1977. Diatom mineralization of silicic acid. II. Regulation of Si(OH)₄ transport rates during the cell cycle of *Navicula pelliculosa*. *J. Phycol.* **13**: 86–91.
- SWOFFORD, D. L. 2001. PAUP*: Phylogenetic analysis using parsimony (*and other methods). Academic.
- THAMATRAKOLN, K., A. J. ALVERSON, AND M. HILDEBRAND. 2006. Comparative sequence analysis of diatom silicon transporters: Toward a mechanistic model of silicon transport. *J. Phycol.* **42**: 822–834.
- TILMAN, D. 1977. Resource competition between planktonic algae: Experimental and theoretical approach. *Ecology* **58**: 338–348.
- . 1981. Tests of resource competition theory using 4 species of Lake Michigan algae. *Ecology* **62**: 802–815.
- , AND S. S. KILHAM. 1976. Phosphate and silicate growth and uptake kinetics of the diatoms *Asterionella formosa* and *Cyclotella meneghiniana* in batch and semicontinuous culture. *J. Phycol.* **12**: 375–383.
- TREGUER, P., D. M. NELSON, A. J. VANBENNEKOM, D. J. DEMASTER, A. LEYNAERT, AND B. QUEGUINER. 1995. The silica balance in the world ocean: A reestimate. *Science* **268**: 375–379.
- TUCHMAN, M. L., E. C. THERIOT, AND E. F. STOERMER. 1984. Effects of low level salinity concentrations on the growth of *Cyclotella meneghiniana* Kütz. (Bacillariophyta). *Arch. Protistenkd.* **128**: 319–326.
- TUSNADY, G. E., AND I. SIMON. 2001. The HMMTOP transmembrane topology prediction server. *Bioinformatics* **17**: 849–850.
- YANG, Z. H. 1997. PAML: A program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13**: 555–556.
- . 2002. Inference of selection from multiple species alignments. *Curr. Opin. Genet. Dev.* **12**: 688–694.
- , AND R. NIELSEN. 2002. Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Mol. Biol. Evol.* **19**: 908–917.
- , ———, N. GOLDMAN, AND A. M. K. PEDERSEN. 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* **155**: 431–449.
- , AND W. J. SWANSON. 2002. Codon substitution models to detect adaptive evolution that account for heterogeneous selective pressures among site classes. *Mol. Biol. Evol.* **19**: 49–57.
- ZWICKL, D. J. 2006. GARLI. Program distributed by the author. Section of Integrative Biology, The Univ. of Texas at Austin.

Received: 22 August 2006

Accepted: 27 February 2007

Amended: 16 May 2007