When Naked Became Armored: An Eight-Gene Phylogeny Reveals Monophyletic Origin of Theca in Dinoflagellates

Russell J. S. Orr1, Shauna A. Murray2,3, Anke Stüken1, Lesley Rhodes4, Kjetill S. Jakobsen1,5

1 Microbial Evolution Research Group (MERG), Department of Biology, University of Oslo, Oslo, Norway, 2 Ecology and Evolution Research Centre and School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, Australia, 3 Sydney Institute of Marine Science, Mosman, New South Wales, Australia, 4 Cawthron Institute, Nelson, New Zealand, 5 Centre for Ecological and Evolutionary Synthesis (CEES), Department of Biology, University of Oslo, Oslo, Norway

Abstract

The dinoflagellates are a diverse lineage of microbial eukaryotes. Dinoflagellate monophyly and their position within the group Alveolata are well established. However, phylogenetic relationships between dinoflagellate orders remain unresolved. To date, only a limited number of dinoflagellate studies have used a broad taxon sample with more than two concatenated markers. This lack of resolution makes it difficult to determine the evolution of major phenotypic characters such as morphological features or toxin production e.g. saxitoxin. Here we present an improved dinoflagellate phylogeny, based on eight genes, with the broadest taxon sampling to date. Fifty-five sequences for eight phylogenetic markers from nuclear and mitochondrial regions were amplified from 13 species, four orders, and concatenated phylogenetic inferences were conducted with orthologous sequences. Phylogenetic resolution is increased with addition of support for the deepest branches, though can be improved yet further. We show for the first time that the characteristic dinoflagellate thecal plates, cellulosic material that is present in the sub-cuticular alveoli, appears to have had a single origin. In addition, the monophyly of most dinoflagellate orders is confirmed: the Dinophysiales, the Gonyaulacales, the Prorocentrales, the Suessiales, and the Syndiniales. Our improved phylogeny, along with results of PCR to detect the sxtA gene in various lineages, allows us to suggest that this gene was probably acquired separately in Gymnodinium and the common ancestor of Alexandrium and Pyrodinium and subsequently lost in some descendant species of Alexandrium.

Citation: Orr RJ, Murray SA, Stüken A, Rhodes L, Jakobsen KS (2012) When Naked Became Armored: An Eight-Gene Phylogeny Reveals Monophyletic Origin of Theca in Dinoflagellates. PLoS ONE 7(11): e50004. doi:10.1371/journal.pone.0050004

Editor: Senjie Lin, University of Connecticut, United States of America

Received May 22, 2012; Accepted October 15, 2012; Published November 19, 2012

Copyright: © 2012 Orr et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: This study was financially supported by the Research Council of Norway (RCN) grant 186292/K40 to KSJ, and by a PhD grant to RJSO from the Molecular Life Science program (MLS), University of Oslo. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: k.s.jakobsen@bio.uio.no

Introduction

Approximately 2000 species of living dinoflagellates are known, most of which are found in marine habitats [1]. Species vary widely, in characteristics such as cell morphology and modes of nutrition (e.g., autotrophy, heterotrophy, mixotrophy, symbiosis, and parasitism) [1,2,3]. Dinoflagellate taxonomy is based on morphological characters such as the presence of a dinokaryon, and the arrangement and shape of thecal plate-containing amphimesmal vesicles. A dinokaryon, a modified nucleus containing permanently condensed fibrillar chromosomes [1,4,5,6,7,8], is present in the “core” dinoflagellates, but lacking from the “pre-dinoflagellate” lineages Oxyrrhinecae and the Syndiniales [3,9]. The Blastodiniales and Noctilucales lack a dinokaryon during particular life cycle stages. For this reason, it has been hypothesized that these lineages are basal [6,10,11,12], although recent evidence suggests that the Blastodiniales may have diverged more recently [13,14,15].

The arrangement of the thecal plate bearing amphimesial vesicles is an important character in distinguishing clades of dinoflagellates [16]. The thecate (armored) orders (Dinophysiales, Gonyaulacales, Prorocentrales and Suessiales) have comparatively fewer, large amphimesial vesicles in distinctive patterns, with cellulosic material in the vesicles. Athecate (unarmored or naked) taxa, however (Gymnodiniales, Noctilucales and Syndiniales) often contain hundreds of alveoli lacking cellulosic material, and therefore relationships are determined based on other features, such as the presence and shape of grooves on the cell surface or on the cell apex, and the shape of the epicone [17,18,19,20].

The monophyly of dinoflagellates and their sister relationships to the Apicomplexa have been established from previous dinoflagellate phylogenies, as well as global eukaryotic phylogenies [16,21,22,23,24,25,26]. Fossil evidence suggests that these groups diverged earlier than 400 Ma [27], and that the species Alexandrium tamarense is a fairly recent dinoflagellate lineage, emerging between 23–45 Ma [28]. The phylogenetic relationship between the dinoflagellate orders, however, is unresolved, with a lack of statistical support for the phylogenetic backbone [5,16,23,24,29,30].

Early molecular phylogenetic studies of dinoflagellate relationships were based on ribosomal rDNA, either partial large-subunit (LSU) [5,17,31], or most frequently, small-subunit (SSU) [5,24,32,33]. However, the low proportion of informative char-
acters resulted in poor resolution despite broad taxon sampling [5,24,33]. Concatenated rDNA inferences have added more resolution, proving useful in the interpretation of genus level relationships [5,29,30,34,35]. However, inter-order relationships remain unclear, with deep branches receiving little or no statistical support, making trends difficult to infer [5,16,36]. Recently, the use of protein genes for phylogenetic inference of diinnoflagellates has increased, in particular actin, alpha- and beta-tubulin [22], hsp90 [16,24], and the mitochondrial cytochrome genes [23]. However, as yet few have inferred a broad diinnoflagellate phylogeny based on more than two concatenated genetic markers [16,25]. Presently, sequence data is only available for approximately 10% or less of the known diinnoflagellate species diversity [29]. The identification of the marine alveolate lineages (MALV), gives an insight into the large parasitic Syndiniales diversity [37,38,39,40,41,42]. A bias toward the photosynthetic taxa also exists, as a large proportion of heterotrophic species, which make up approximately 50% of the true diinnoflagellate lineage [43], are difficult or impossible to culture.

A well-resolved diinnoflagellate phylogeny is essential to understanding the evolution of toxin synthesis in this phylum. Approximately 100 known species of diinnoflagellates produce a variety of toxins, that can accumulate in the water column as harmful algal blooms (HABs) [44]. Saxitoxin (STX), and its analogues, is one such toxin that can cause severe symptoms upon consumption of vector species [45,46]. STX is synthesized by eukaryotic marine diinnoflagellates and freshwater cyanobacteria [47,48]. The toxins appear to be synthesized by similar processes in cyanobacteria and diinnoflagellates [49]. The biosynthetic pathway and genes responsible for STX-synthesis are known from cyanobacterial species [50,51,52,53,54]. The genes common between these clusters have been defined as “core” genes [46,55]. One such core gene, sxtA, the unique starting gene of STX synthesis, has recently been identified in the diinnoflagellates Gymnodinium catenatum and multiple species within the genus Alexandrium [56]. The origin of this gene cluster within the diinnoflagellates may have occurred by way of a horizontal gene transfer (HGT) event between an ancestral STX-producing bacterium and the diinnoflagellates, before Alexanderion and Pyrodinium diverged [56]. Thus the ability to produce STX may have been secondarily lost for some descendent species. As the sxtA sequence of Gymnodinium catenatum, in the order Gymnodiniales, branches with Alexandrium, an independent acquisition of STX from a diinnoflagellate-dinoflagellate transfer has been postulated [56]. As the phylogenetic relationship between diinnoflagellate species remains unresolved, trends in the evolution of the genetic basis for the synthesis of STX or other toxins cannot be established [23].

The aim of this study is to improve the resolution of the diinnoflagellate phylogeny by sampling a broad range of both taxa and genes using concatenated alignments. This will allow us to address relationships between orders and identify possible phylogenetic trends in the evolution of STX production and other major phenotypic characters, such as morphological traits. To achieve this, 55 sequences for eight molecular markers were amplified, from 13 species, spanning four orders. A concatenated phylogenetic approach was used with all orthologous database sequences. Furthermore, we tested 20 species from five orders for presence of sxtA1 and sxtA4.

**Methods**

Our use of order, family and genus-level names follows the taxonomic revision of Fensome et al. (1993), and its recent update Fensome et al. (2008), made publically available online (http://dinoflag.spmu.ca/) [6,57]. However, the following amendments are used: “core” diinnoflagellates for dinokaryota [58], family Kareniaee [17,59], the inclusion of Brachidinium in Kareniaee [60], genus Thoxosphaera [61], genus Adenoides [62], Buchelia bullica (syn. Woloszynskia halophila) [63], Pelagodinium beii (syn. Gymnodinium beii) [64], and Protodinium simplex (syn. Gymnodinium simplex) [58].

**Culturing**

The diinnoflagellate strains used in this study (Table 1) were grown in L1 media [65] or GSe media [66] at 16–25°C. In addition, Polarella glacialis CCMP2088 was grown at 5°C, all with a 12:12 h light-dark photoperiod and a photon irradiance of ~100 mmol photons m−2 s−1. Strains were not maintained axenically. The identity of each strain was confirmed by amplifying the 18S rDNA gene using the primer pairs NSF1 - HSS6 [152R] and 18sF8 - ITS801 [67,68,69]. Inclusion of Pyrodinium bahamense in the study would be desirable as it produces saxitoxin, however no culture was available to us.

**DNA and RNA Isolation, cDNA Synthesis, PCR Amplification, Sequencing and Assembly**

Genomic DNA and total RNA were isolated from 20 ml cultures in the exponential growth phase, centrifuged for 2 min at 12,000×g, washed with PBS and bead-beaten on dry ice with the FastPrep-24 from Medinor (20 s, speed 4) using 1.4 mm beads (Medinor). For DNA the CTAB method [70] or Invitrogen ChargeSwitch gDNA plant kit (Invitrogen) were utilized. For total

**Table 1. List of dinoflagellate strains used in this study and whether sxtA1 and sxtA4 were PCR amplified.**

| Species/Taxon | Strain   | sxtA1 | sxtA4 |
|---------------|----------|-------|-------|
| Adenoides eludens | CCMP1891 | n.d.  | n.d.  |
| Alexandrium fundyense | CCMP1719 | + [56] | + [56] |
| Alexandrium minutum | CCMP113 | + [56] | + [56] |
| Amphidinium carterae | UI0081 | n.d.  | n.d.  |
| Amphidinium massartii | CS-259 | n.d. [56] | n.d. [56] |
| Amphidinium moomonorum | CAWD161 | n.d.  | n.d.  |
| Azadinium spinosum | RCC2538 | n.d.  | n.d.  |
| Ceratium longipes | CCMP1770 | n.d.  | n.d.  |
| Coolia mononis | CAWD98 | n.d.  | n.d.  |
| Gambierdiscus australis | CAWD148 | n.d.  | n.d.  |
| Gymnodinium aureolum | SCCAP K-1561 | n.d.  | n.d.  |
| Heterocapsa triqueta | RCC2540 | n.d.  | n.d.  |
| Karlodinium venicum | RCC2539 | n.d.  | n.d.  |
| Lepidodinium chlorophorum | RCC2537 | n.d.  | n.d.  |
| Lingulodinium polyedrum | CCMP1931 | n.d.  | n.d.  |
| Pentapharsodinium dalei | SCCAP K-1100 | n.d.  | n.d.  |
| Polarella glacialis | CCMP2088 | n.d.  | n.d.  |
| Prorocentrum micans | UI0292 | n.d.  | n.d.  |
| Prorocentrum minimum | UI0085 | n.d.  | n.d.  |
| Protoceratium reticulatum | CAWD99 | n.d.  | n.d.  |
| Pyrocystis noctiluca | CCMP732 | n.d.  | n.d.  |
| Scrippsiella trochoideae | BS-46 | n.d.  | n.d.  |
| Thecadinium kofoidii | SCCAP K-1504 | n.d.  | n.d.  |

n.d. not detected. + amplified sequence.
doi:10.1371/journal.pone.0050004.t001
RNA the Invitrogen ChargeSwitch TotalRNA cell kit (Invitrogen) or RNaseasy Plant Mini kit (Qiagen) were used in accordance with supplied protocol. First strand cDNA was synthesized with Invitrogen 3’ RACE system (Invitrogen) following the high GC protocol and utilizing the (AP) adapter primer, or with Invitrogen Superscript First-Strand Synthesis System (Invitrogen). DNA, RNA and cDNA quality was checked with a NanoDrop spectrophotometer (ThermoScientific).

The genes amplified in this study 18S rDNA (Small subunit), 5.8S rDNA, 28S rDNA (Large subunit), actin, beta-tubulin, cytochrome b (cob), cytochrome c oxidase subunit I (coxI), and heat-shock protein 90 (hsp90), were determined from dinoflagellate sequence availability within NCBI. Mixing of gDNA and cDNA sequences for phylogenetic inference may produce invalid results due to widespread mRNA editing in dinoflagellates [23]. Thus, mRNA was utilized for cob and coxI, as only a single functional sequence is reported, in comparison to multiple genomic copies [23]. Likewise, actin in dinoflagellates is present in a variable number of copies in the genome, including pseudogenes; therefore mRNA was again favored [71]. Beta-tubulin (mRNA) and hsp90 (gDNA) were determined from sequence availability. Template was only PCR amplified for the genes and strains lacking Genbank sequence data. cDNA/ gDNA template was PCR amplified using Qiagen HotStar Taq Plus polymerase (Qiagen), Bioline Mytag polymerase (Bioline) or BD Advantage 2 polymerase (Clontech) in the presence of 10% BSA in a MJ Research PTC-200 Thermo Cycler (MJ Research) with the following PCR conditions: an initial denaturing before 35 cycles of (1) 30 sec denaturing, (2) 30 sec annealing (variable temperature, see Table S1 for T_M), and (3) 1–2 minute extension, with a final 10 minute extension at the same temperature. PCR products were gel excised using Promega Wizard SV Gel and PCR Clean-Up System (Promega), before direct sequencing with an ABI3730 DNA analyzer (Applied Biosystems) using combinations of primers (Table 2; Table S1). The universal primers used in this study have been designed using Primaclade based on alignments constructed from multiple orthologous dinoflagellate sequences. [72]. Melting temperature (T_M) was calculated using OligoCalc [73]. Sequences were quality checked and assembled using the Phred/Phrap/Consed [74] package under default settings. Additional manual editing was performed in MacCladev4.07 [75]. The presence of sxtI1 and sxt4H genes were tested for all dinoflagellate strains following the protocol described in [56]. The sxtI1 fragment was amplified with primers sxt001 & sxt002 (~500 bp) and the sxt4H fragment with the primers sxt007 & sxt008 (~750 bp) (Table S1). A positive gDNA control from A. fundyense CCMP1719 or A. minutum CCMP113 was utilized in all sxtI PCRs.

Phylogenetic Inferences

All sequences generated in this study as well as dinoflagellate orthologous sequences in the NCBInr nucleotide and EST databases http://www.ncbi.nlm.nih.gov/ (as of 12.2011) for each gene were separated into their respective datasets. The three-rDNA genes (18S, 5.8S, and 28S) were separately aligned using the MAFFTv6 Q-INS-I model [76,77,78], considering secondary RNA structure (default parameters used). The five protein coding datasets (actin mRNA, beta-tubulin mRNA, cob mRNA, coxI mRNA, and hsp90 gDNA) were separately aligned at the nucleotide level based on the corresponding amino acid alignment, as to maintain codon integrity, inferred with MAFFTv6 G-INS-I model (default parameters used). To increase phylogenetic signal, allowing for synonymous substitutions, the nucleotide sequence (3rd codon removed) was used for subsequent inferences. Outgroup taxa (Apicomplexa) was established from previous dinoflagellate phylogenies [16,23,24], as well as global eukaryotic phylogenies that concur in placing this as the closest extant relative to the dinoflagellates [25,26]. In-group taxa (dinoflagellata) required both 18S and 28S rDNA sequence data, thus Blastodinium, having only 18S rDNA was excluded. The only exception was Ceratocorys horrida; this species’ 28S rDNA sequence is not available, though as it had available cytochrome c oxidase sequences, and as the only representative from this family, its inclusion was considered important. The resulting single gene alignments were subsequently checked manually using MacCladev4.07 [75]. The eight separate alignments were then checked with Gblocks v0.91b [79], under the least stringent parameters (small final block, gap positions in final block and less strict flanking), to exclude poorly aligned positions and divergent regions from subsequent phylogenetic inferences. The alignments were then concatenated into the following supermatrices; (1) rDNA; (18S+5.8S+28S), (2) rDNA+nuclear protein; (18S+5.8S+28S+actin+beta-tubulin+hsp90) and (3) rDNA+mitochondrial+nuclear protein; (18S+5.8S+28S+cob+coxI+actin+beta-tubulin+hsp90), a reduced dataset was additionally constructed from the previous two, excluding taxa with only rDNA signal (lacking protein coding gene data); done to evaluate effects of missing characters and taxon sampling on the inferences.

The supermatrix (concatenated) approach provides support not always apparent with fewer genes [80]. To reduce missing data and improve phylogenetic placement, concatenated Hematodinium sequences were a composite (in silico chimeric) of closely related intra-genus species (H. perez, Hematodinium sp., Hematodinium sp. ex Callinectes sapidus, and Hematodinium sp. ex Nephrops norvegicus; Table S2).

Taxa have not been excluded from the inferred supermatrices with “missing characters” as a criteria, as phylogenetic estimates including incomplete taxa show little evidence to support taxa exclusion based on missing data [81]. Addition of incomplete taxa, even <10% complete, can be equally beneficial to a phylogeny as 100% complete taxa, improving resolution at the genus level, placing with strong statistical support, and even subdividing misleading long branches [80,81,82,83]. The critical factor for taxa placement is not character absence, but the quality and number of those present [81]. All concatenated datasets were then analyzed with MODELC/TEST [84] to establish the optimal model of nucleotide evolution; for all alignments the (General Time Reversible) GTR model was preferred for both the Akaike and Bayesian information Criterion (AIC and BIC). Maximum Likelihood (ML) analyses were performed with RAXML-VI-HPCv7.2.6, GTRCAT model with 25 rate categories [85,86]. The most likely topology was established from 100 separate searches and bootstrap analyses were performed with 500 pseudoreplicates. Bayesian analyses were carried out with MrBayes MPI version 3.1.2 [87,88]. Trees were generated from two independent runs with one heated and one cold chain in the Markov Chain Monte Carlo (MCMC) with 40,000,000 generations, sampling every 1000. Analyses ran until the average standard deviation of split frequencies were <0.01. Burn-in trees were set based on the assessment of likelihood plots and convergence diagnostics implemented in MrBayes. The Potential Scale Reduction Factor (PSRF) values for all inferences were ~1.0, indicating a good posterior probability distribution sample. The majority rule tree and posterior probabilities for each inference was constructed from a consensus of the sampled post burn-in trees. Topological congruence between the inferred phylogenies were calculated...
The Phylogeny of Dinoflagellates

All inferred dinoflagellate phylogenies show good topological congruence with an $\ell_{\text{long}}$ value $<0.05$ (Fig. 1, 2, 3, 4). Also the comparison of topologies inferred from separate rDNA and protein coding gene datasets demonstrated good congruence (Fig. S2). Removal of long-branching taxa had minimal rDNA impact (data not shown). In addition, the inference of the corresponding translated supermatrices had minimal topological

using the $\ell_{\text{long}}$ index: http://max2.esc.u-psud.fr/bases/upresa/pages/devienne/index.html [89].

N. scintillans was excluded from the presented concatenated analyses as its cryptic and inconsistent placement reduced phylogenetic support. However, its “most probable” placement was determined from parallel Bayesian inferences.

Previous phylogenies based on the mitochondrial cytochrome genes, cob and cox1, place Heterocapsa basal within the dinoflagellate lineage [23,90], a possible phylogenetic artifact as a result of a faster mutation rate [23] As this position is inconsistent with morphological data and phylogenies without mitochondrial genes [16,30], we investigated this further. A simple distance-based comparative rate test was used to measure the divergence of the different genes for Heterocapsa triquetra to that of the ingroup (“core” dinoflagellates) [91,92]. In this context, the comparative rate was defined as the ratio of the pairwise distances of Heterocapsa to the ingroup taxa, compared with the mean distance between the same ingroup taxa. Here we considered A. carterae, A. eludens, A. minutum, A. spinosum, G. aureolum, K. veneficum, P. glacialis, P. minimum, and S. trochoideae to form the ingroup. The distance from Heterocapsa to the ingroup taxa was divided by the mean distance of the ingroup taxa to each other. The pairwise distances between all taxa were calculated using RAxML [83,86] with the -x option and GTRGAMMA model for each individual gene alignment as well as the rDNA, nuclear protein and cytochrome concatenated alignments. Subsequently, cob and cox1 solely for Heterocapsa were excluded from inferences.

All model estimation and phylogenetic analyses were done on the freely available Biportal [93] at the University of Oslo (http://www.bioportal.uio.no/).
Figure 1. Phylogenetic tree of dinoflagellates inferred from rDNA. Concatenated phylogeny, inferred from 18S + 5.8S + 28S rDNA (2900 characters). The tree is reconstructed with Bayesian inference (MrBayes). Numbers on the internal nodes represent posterior probability and bootstrap values (> 50%) for MrBayes and RAxML (ordered; MrBayes/RAxML). Black circles indicate a posterior probability value of 1.00 and bootstrap >90%. N. scintilans is represented with a dashed branch as this taxon was excluded from the inference; alternatively its most “probable” placement...
impact (data not shown). Resolution was limited for the eight single gene phylogenies (Fig. S4, S5, S6, S7, S8, S9, S10, S11).

For interpretation of the phylogenetic inferences (Fig. 1, 2, 3, 4; Table 3), statistical support is defined as: full 1.00PP/100BP, high >90BP, moderate >65BP, and low >50BP. Dinoflagellate in-group monophyly was inferred for all datasets with support varying from moderate, to high (Fig. 2, 3, 4; Table 3). The first, and most basal, dinoflagellate order to diverge from the main branch was the Oxyrrhiniaceae (Fig. 3–4). This position was unsupported, with Syndiniales being alternatively recovered as the most basal in the rDNA+nuclear protein dataset (Fig. 2). Both orders formed a sister relationship in the rDNA inference (Fig. 1). The Syndiniales clade was constantly recovered with moderate support (Fig. 1, 2, 3; Table 3). The “pre-dinoflagellate” Oxyrrhiniaceae and Syndiniales lineages were excluded from the “core” dinoflagellates with support varying from low to full (Fig. 1, 2, 3, 4; Table 3).

The first “core” dinoflagellate order to diverge was the athecate Gymnodiniolaxes (Fig. 1, 2, 3, 4), branching paraphyletic in the largest dataset (Fig. 3), being divided into five sub-clades. The first sub-clade to diverge from the Gymnodiniolaxes, and the most basal “core” dinoflagellate, was the highly supported genus *Amphidinium* (Fig. 1, 2, 3, 4; Table 3). *Amphidinium* placed as the sister group to the high to fully supported genus *Gyrodinium* in all but the rDNA inference, where they formed a monophyletic relationship (Fig. 1, 2, 3; Table 3). The low to fully supported family Karentiaceae [17] was next to diverge in all but the rDNA inference, where it alternatively placed terminal to the genus *Gymnodinium sensu stricto* [17]. The three previous Gymnodiniolaxes clades were basal to *Gymnodinium sensu stricto* (Fig. 2–3), excluded with moderate support (1.00/60) in the rDNA+nuclear protein dataset (Fig. 3). The *Gymnodinium sensu stricto* and several other very closely related genera formed a high to fully supported clade for all datasets (Fig. 1, 2, 3, 4; Table 3). The position of the genus *Akalasico* was unstable, placing within the Peridiniales with few inferred characters (Fig. 1–2). Increasing character number resulted in an unstable position of *Akalasico*, and thus Gymnodiniolaxes in the largest dataset (Fig. 3; Table 3), placing as the sister lineage in all but the rDNA inference, where it alternatively was positioned as the most basal “core” dinoflagellate (Fig. 1).

The basal athecate lineages (Gymnodiniolaxes, Noctilucales, Oxyrrhiniaceae and Syndiniales) were excluded from the monophyletic athecate (Dinophysiales, Gonyaulacales, Peridiniales, Prorocentrales and Suessiales) with low support (Fig. 3–4; Table 3). This division was found only in the dataset with the most characters (eight genes). With fewer inferred characters, the position of *Akalasico* within the Peridiniales resulted in Gymnodiniolaxes being recovered as a polyphyletic order (Fig. 1–2). Additionally, the support for the thecate/athecate split was reduced when a narrower taxon sample was inferred (Fig. 3–4; Table 3). Within the thecate dinoflagellates, the Peridiniales and Prorocentrales were recovered as an unsupported clade (Fig. 3–4), though this was excluded from the Dinophysiales, Gonyaulacales and Suessiales clade with low support in the largest dataset (Fig. 3; Table 3). The Prorocentrales was recovered as a monophyly with low to moderate support when inferred including the mitochondrial genes (Fig. 3–4; Table 3). The incerta sedis genus, *Adriales* [62], which placed as the sister lineage to Prorocentrales, was monophyletic to the prorocentroid clade with the exclusion of mitochondrial cytochrome genes (Fig. 1–2). The high to fully supported genus *Heterocapsa* (Fig. 1, 2, 3; Table 3) was consistently recovered with a placement directly basal to the main Peridiniolaxes clade in all phylogenies excluding the cytochrome genes for this lineage (Fig. 1, 2, 3, 4). The inclusion of the mitochondrial clades resulted in *Heterocapsa* being recovered as the most basal “core” dinoflagellate (Fig. S1). *Heterocapsa* and *Adriales* were constantly recovered paraphyletic to main Peridiniolaxes clade (Fig. 1, 2, 3, 4). Though, *Azadinium* (Fig. 1) and *Kryptoperidinium* were also recovered paraphyletic to the main Peridiniolaxes monophyly.

The monophyly of Dinophysiales, Gonyaulacales and Suessiales were recovered with broad taxon sampling, with the largest dataset adding support to this relationship (Fig. 1, 2, 3; Table 3). When the taxon sample was reduced, the Suessiales was excluded from this monophyly, alternatively placing as the most basal thecate order, though the placement was unsupported (Fig. 4). The branching pattern for the Dinophysiales, Gonyaulacales and Suessiales clade is uncertain, with no support for a relationship. However, the Dinophysiales was recovered as the unsupported sister clade to the Gonyaulacales when inferring using all eight genes (Fig. 3–4). The Suessiales were monophyletic with high to full support for all datasets, with a more resolved internal branching pattern when inferring more genes (Fig. 1, 2, 3, 4; Table 3). The Dinophysiales, likewise, receive high to full support for their monophyly, with a resolved internal branching pattern for all datasets (Fig. 1, 2, 3, 4; Table 3). The monophyletic Gonyaulacales received low to high support with the addition of inferred characters, though the internal branching pattern was not fully resolved (Fig. 1, 2, 3, 4; Table 3).

The distance-based comparative rate test measured sequence divergence between the different *Heterocapsa* genes (Fig. S3). *Heterocapsa* was compared to the “core” dinoflagellate mean, so if genes were homogenous the distance ratio would be approximately one. Values greater than one indicate a more divergent gene, with values below one indicating less divergent gene. The distance ratios indicated that *Heterocapsa* rDNA (18S+5.8S+28S) and nuclear protein gene (*actin*,*beta-tubulin*,*hsp90*) divergence was less than that of the “core” dinoflagellate mean (1.0), with a 50th percentile range of 0.14360–0.73360, and a 25th–75th percentile range of 0.04607–0.77490. The nuclear protein genes diverged approximately 2.39 faster than that of the rDNA. In comparison, the mitochondrial genes were approximately 3 times more divergent than the “core” dinoflagellate mean, as well the rDNA and nuclear protein genes for *Heterocapsa*, with a 50th percentile of 3.58400–3.64000 and a 25th–75th percentile of 3.41000–3.68100. This equates to a divergence rate for *Heterocapsa* mitochondrial genes approximately 10 times that of their rDNA genes, and 4 times that of their nuclear protein genes.

**SxtA Detection**

*SxtA1* and *sxtA4* were not detected in 20 species and five orders. This included two additional orders (Peridiniales and Suessiales) to those already reported (Table 1) [56]. The 18S rDNA control was amplified for all tested species as were the *sxtA* (1/4) positive controls.
Discussion

An Improved Dinoflagellate Phylogeny

The phylogenies presented here, with the broadest taxon sampling and largest number of inferred phylogenetic informative nucleotide positions to date, improves the resolution of dinoflagellate in-group relationships. As in previous molecular phylogenetic studies, the dinoflagellates are recovered as a monophyletic lineage [16,21,22,23,24], with the inferred outgroup taxa. We add statistical support to the phylogenetic backbone allowing us to infer previously unseen relationships and trends between dinoflagellate orders. For the first time, we find that the thecate dinoflagellates have a supported monophyletic origin (Fig. 3–4), diverging from an athecate ancestor. The increased resolution is mostly congruent with the subdivision of orders based on morphological characters, thus broadly supporting the classification based on plate tabulation patterns [6,94]. Despite this, resolution for some nodes, for example Akashiwo sanguinea, can be improved yet further.

This result highlights the importance of using broad taxon sampling, whilst in parallel increasing character number, to further resolve dinoflagellate evolutionary relationships [95,96,97]. Previously, “missing characters” has been used as a criterion for taxa exclusion from alignments [16]. In agreement with Wiens (2006), we find little evidence to support the exclusion of taxa based on missing data [81]. Further, we find that the exclusion of such taxa negatively impacts topological resolution, as well as statistical support; with a reduced taxon sample resulting in a more divergent placement of Suessiales, reduced support for the thecate/athecate split, and the recovery of Kryptoperidinium external to main Peridiniales clade (Fig. 1, 2, 3, 4; Table 3).

Previous phylogenies inferred with the mitochondrial cytochrome genes cox1 and cox1, found Heterocapsa to be the most basal “core” dinoflagellate [23,90]. We also found Heterocapsa to be in a basal position when inferring the phylogeny based on mitochondrial genes (Fig. S1). However, the exclusion of cox1 and cox1 for Heterocapsa resulted in a position congruent with both morphological data and phylogenies without mitochondrial genes [16,30]. The basal position of Heterocapsa recovered with mitochondrial genes has been hypothesized as a possible artifact, a result of a faster mutation rate [23]. Indeed, the cox1 and cox1 sequence for species of Heterocapsa were found to be highly diverged compared to that of all other dinoflagellates with a divergence rate approximately 3 times higher, and 4 times that of its own nuclear protein genes. This may result in this lineage being repelled from the in-group and, in contrast, artificially attracted to the out-group. The mitochondrial genes are promising markers for interpreting dinoflagellate evolutionary history, inferring the monophyly of Prorocentrales, improving resolution and increasing support [23]. Thus the exclusion of these markers for all taxa seems to directly oppose the goal of inferring a more resolved phylogeny of the dinoflagellates. Similarly, the exclusion of Heterocapsa from inferences would also reduce resolution. Accordingly, and until either the evolution of mitochondria in Heterocapsa is fully understood, or enough genetic markers are available to dilute this incongruent signal supported by the trend of mRNA editing of these genes [90], the exclusion of these markers solely for this lineage is warranted. Subsequent exclusion results in a general increase in resolution and support (Fig. 3–4).

The most basal dinoflagellate lineage could not be inferred with certainty from this analysis, with support for the branching pattern between Oxyrrhinaceae and Syndiniales being somewhat inconclusive. A recent study of dinoflagellates based on multiple morphological characters found that the Oxyrrhinaceae was basal to the Syndiniales, which is in line with the results of our largest dataset (Fig. 3) [9]. This is the first phylogeny to conclusively show the sister relationship of these orders, when inferred with broad dinoflagellate taxon sampling [91]. Oxyrrhinaceae and Syndiniales have been excluded from the “core” dinoflagellates, as both lack a dinokaryon [36,90]. However, previous dinoflagellate phylogenies have lacked either Oxyrrhinaceae [5,30] or Syndiniales [16,23], with a highly derived position for Oxyrrhis when inferred together [36]. The ancestral position of Oxyrrhinaceae and Syndiniales, both lacking theca, support an athecate dinoflagellate ancestor [5].

In this study, the Gymnodiniales was found to be the most basal “core” dinoflagellate order. Gymnodiniales could be hypothesized to be the sister group to Oxyrrhinaceae and Syndiniales, as the small alveoli are homologous between these orders [5]. However, previous phylogenies have shown the order to be polyphyletic [5,16,23,30]. In contrast, we found the order to show a paraphyletic branching pattern when inferred based on a larger number of genes (Fig. 3–4). This result indicates that it is unlikely that athecate lineages have had thecate ancestors [5,36,94,99]. Similar to previous phylogenies, a polyphyly was observed when the phylogeny was inferred based on fewer characters, as Akashiwo showed an affinity to the Peridiniales (Fig. 1–2).

The genus Amphidinium was consistently recovered as the most basal “core” dinoflagellate, in all but the phylogeny based on rDNA (Fig. 1), where it shared this position with Gymnodinium. This position concurs with previous studies [23,100]. Unlike the finding for Heterocapsa, our results showed no inconsistencies between the mitochondrial and nuclear protein gene phylogenetic signals in Amphidinium [23].

The placement of the Noctilucae has long been questioned [5,16,36]. The lack of a dinokaryon during particular life cycle stages, would suggest that it may be a basal lineage [6,10,11,12]. However, several novel morphological features and incongruent phylogenies contradict this interpretation [5,36]. Noctiluca was excluded from the presented phylogenies, and instead its most “probable” position inferred from parallel Bayesian analyses. This was done to increase resolution for all taxa, as the cryptic and inconsistent placement of Noctiluca reduced phylogenetic support overall. For the bootstrap pseudoreplicates, the position of Noctiluca changed from that of a pre-dinoflagellate lineage to having multiple positions within Gymnodiniales, Peridiniales and Syndiniales. Only the Bayesian inference of the largest dataset found Noctiluca to be the sister to Akashiwo with full support. Thus all inferred positions for this taxon are tentative. Similar to previous phylogenies based on rDNA genes [5,24,29,30], we found Noctilucae to be a basal “core” dinoflagellate (Fig. 1). However, increasing the number of inferred characters resulted in an affinity...
of Noctilucales to Akashiwo, and thus the Gymnodiniales (Fig. 3–4). An association between the Noctilucales and the Gymnodiniales has been shown previously with multiple genetic markers [16]. This is the first phylogeny to show support for the monophyly of thecate dinoflagellate orders (Dinophysiales, Gonyaulacales, Peridiniales, Prorocentrales and Suessiales) and has a direct consequence for the evolutionary origin of thecal plates, implying that this morphological trait arose from a single event. Our results are in contrast to previous phylogenies suggesting that these orders are either poly- or paraphyletic, indicating that the thecal characteristic had evolved, or been lost, repeatedly within the “core” dinoflagellates [5,16,23]. Taylor (2004), using morphological evidence, hypothesized that thecate dinoflagellates may have arisen from thecate ancestors [36]. The plate increase and plate fragmentation hypotheses did not explain the observed trend [33,94,101,102]. However, aspects of the plate reduction model were seemingly supported, for example the basal position of Gymnodiniales [33,103]. Further, the plate reduction hypothesis

![Figure 3. Phylogenetic tree of dinoflagellates inferred from rDNA, mitochondrial and nuclear protein genes.](image)

Concatenated phylogeny, inferred from 18S + 5.8S + 28S + cob + cox1 + actin + beta-tubulin + hsp90 (7138 characters). The tree is reconstructed with Bayesian inference (MrBayes). Numbers on the internal nodes represent posterior probability and bootstrap values (>50%) for MrBayes and RAxML (ordered: MrBayes/RAxML). Black circles indicate a posterior probability value of 1.00 and bootstrap >90%. *N. scintillans* is represented with a dashed branch as this taxon was excluded from the inference; alternatively its most “probable” placement was determined from a parallel Bayesian analysis. The cytochrome genes cob and cox1 for *H. triqueta* were excluded from the inference, a parallel phylogeny including these genes for this taxon can be seen in Figure S1. * Denotes taxa sequences generated from this study. See Table S2 for a full listing of accessions used. Red font indicates sxtA presence and blue font indicates no sxtA detection. Non-ribosomal gene presence for each taxon is represented in brackets behind each species name (a: actin, b: beta-tubulin, c1: cox1, cb: cob, h: hsp90). The phylogenetic support for the thecate/athecate split is highlighted with bold type.

doi:10.1371/journal.pone.0050004.g003

![Figure 4. Phylogenetic tree of dinoflagellates inferred from rDNA, mitochondrial and nuclear protein genes (reduced phylogeny).](image)

Concatenated phylogeny, inferred from 18S + 5.8S + 28S + cob + cox1 + actin + beta-tubulin + hsp90 (7138 characters). This phylogeny was inferred excluding taxa with only rDNA signal; done to evaluate effects of missing characters and taxon sampling on the inference shown in Fig. 3. The tree is reconstructed with Bayesian inference (MrBayes). Numbers on the internal nodes represent posterior probability and bootstrap values (>50%) for MrBayes and RAxML (ordered: MrBayes/RAxML). Black circles indicate a posterior probability value of 1.00 and bootstrap >90%. *N. scintillans* is represented with a dashed branch as this taxon was excluded from the inference; alternatively its most “probable” placement was determined from a parallel Bayesian analysis. The cytochrome genes cob and cox1 for *H. triqueta* were excluded from the inference. * Denotes taxa sequences generated from this study. See Table S2 for a full listing of accessions used. Red font indicates sxtA presence and blue font indicates no sxtA detection. Non-ribosomal gene presence for each taxon is represented in brackets behind each species name (a: actin, b: beta-tubulin, c1: cox1, cb: cob, h: hsp90). The phylogenetic support for the thecate/athecate split is highlighted with bold type.

doi:10.1371/journal.pone.0050004.g004
Table 3. Comparison of the phylogenetic support (Posterior probability and bootstrap) received for the dinoflagellate orders, major lineages, clades and nodes in the inferences of Fig. 1, 2, 3, 4 (ordered; MrBayes/RAxML).

| Figure | Dataset | 1 | 2 | 3 | 4 |
|--------|---------|---|---|---|---|
|        |         | rDNA | rDNA + nuclear protein | rDNA + mitochondrial + nuclear protein | Reduced rDNA + mitochondrial + nuclear protein |
| Taxa/Characters | Dinoflagellates | 104/2900 | 104/5626 | 104/7138 | 56/7138 |
|        | Syndiniales | 0.95/65 | 1.00/95 | 1.00/82 | 1.00/82 |
|        | “Core” dinoflagellates | 0.96/97 | 0.96/97 | 1.00/97 | 1.00/97 |
|        | Gymnodiniales | (polyphyletic) | (polyphyletic) | (polyphyletic) | (polyphyletic) |
|        | Amphidinium | 1.00/91 | 1.00/91 | 1.00/97 | 1.00/92 |
|        | Gyrodinium | 1.00/94 | 1.00/100 | 1.00/98 | – |
|        | Kareniaceae | 1.00/61 | 1.00/85 | 1.00/92 | 1.00/100 |
|        | Gymnodinium sensu stricto | 1.00/97 | 1.00/100 | 1.00/100 | 1.00/100 |
|        | Thecate/athecate split | – | – | 1.00/62 | 1.00/52 |
|        | Prorocentrales | (polyphyletic) | (polyphyletic) | 0.94/76 | 0.94/62 |
|        | Peridiniales | (polyphyletic) | (polyphyletic) | (polyphyletic) | (polyphyletic) |
|        | Heterocapsa | 1.00/99 | 1.00/100 | 1.00/99 | – |
|        | Suessiales | 1.00/90 | 1.00/93 | 1.00/94 | 1.00/100 |
|        | Dinophysiales | 1.00/99 | 1.00/96 | 1.00/96 | 1.00/100 |
|        | Gonyaulacales | 1.00/99 | 1.00/72 | 1.00/99 | 1.00/91 |
|        | Dinophysiales, Gonyaulacales and Suessiales | 1.00/99 | 1.00/99 | 1.00/56 | – |

As has been found previously [23,104], the monophyly of Prorocentrales was only recovered after the inclusion of the mitochondrial cytchrome genes cob and Cox1 (Fig. 2). The support for this monophyly was minimal, with Adenoides, either placing within the prorocentroid clade (Fig. 1–2), or as its sister group (Fig. 3–4). Adenoides has been tentatively placed within Gonyaulacales based on morphological data [105]. However, phylogenetic data (Fig. 3–4) supports the exclusion of this incertae sedis genus from the Gonyaulacales, alternatively suggesting an affinity to either the Peridiniales or the Prorocentrales [5,23,24]. The Peridiniales has been previously recovered as a polyphyletic lineage [5,16,29]. Nevertheless, the result suggests Prorocentrum is a derived taxon linked to the peridinioles [36].

The monophyly of Dinophysiales, Gonyaulacales and Suessiales was constantly recovered, with support added for the largest dataset (Fig. 1, 2, 3). This relationship has only been recently seen, albeit without support, with the inference of a large taxon sample for a concatenated 18S and 28S rDNA phylogeny [30]. Previously, the Gonyaulacales has either formed an unsupported monophyly with the Dinophysiales [24] or the Suessiales [5,16], though alignments analyzed in previous studies generally did not include broad taxon sampling. The Suessiales was recovered as a supported monophyletic lineage with a resolved internal branching pattern [5,30,32,106]. However, the Suessiales was never recovered as a basal order, in contrast to its position based on mitochondrial genes in Zhang et al. (2007) [23]. This, combined with the position of Heterocapsa within the Peridiniales, found in this study, rather than as a basal dinoflagellate lineage, as it has been found based purely on mitochondrial genes, suggests that the level of cytochrome mRNA editing is a poor character for determining the most basal lineages of dinoflagellates [90]. The Dinophysiales was found to be a fully supported order, with established internal relationships [5,30,107]. The Gonyaulacales was found to be a monophyletic lineage, similar to previous studies [3,24,29,30]. However, the increase in inferred characters added support to this monophyly, increasing resolution for the internal branching pattern.

The Origin of Saxitoxin

The cyanobacterial STX-pathway is thought to have arisen at least 2100 million years ago [55]. The toxins are seemingly synthesized by similar processes in both cyanobacteria and dinoflagellates [49]. The genes responsible for STX-synthesis have been reported in numerous cyanobacterial species [50,51,52,53,54]. SxtA, the unique starting gene of STX-synthesis, has been recently identified in the dinoflagellates [56]. A HGT event between an ancestral STX-producing bacterium and the dinoflagellates has been proposed [56]. This probably occurred before Alexandrium and Pyrodinium diverged within the order.
Gonyaulacales. Thus STX-synthesis may have been secondarily lost for some descendant species. Gymnodinium catenatum, possessing an sxtA sequence that branches within the Alexandrium clade, probably independently acquired STX from a later dinoflagellate-dinoflagellate transfer [56].

The results presented here appear to lend some support to this hypothesis; sxtA was undetected in any Gonyaulacoids directly basal to the moderately supported clade harboring Alexandrium and Pyrodinium (Fig. 3). This seems to suggest, in addition to the previous hypothesis, that any proposed HGT event occurred in a recent ancestor of Alexandrium and Pyrodinium and not deeper within the Gonyaulacales. sxtA was not detected in Ceratium. Though this genus was found to be basal to the Alexandrium and Pyrodinium split in the largest dataset, its exclusion was only moderately supported (Fig. 3). In comparison, it was found to be the basal sister to Alexandrium in both the rDNA and rDNA+ nuclear protein datasets with high support (Fig. 1–2). If a HGT, as proposed, occurred prior to the split of Alexandrium and Pyrodinium, we may expect Ceratium to still possess sxtA. A negative result for this genus may therefore, in congruence with the hypothesis, support a secondary loss of STX-synthesis with sxtA for some descendant species, additionally including Coolia, Gambierdiscus and Pyrocystis (Fig. 2). However, it may suggest that a HGT event occurred in either Alexandrium or Pyrodinium, with one of these genera later acquiring STX via a secondary dinoflagellate-dinoflagellate transfer event. This would in-turn reject any theory of secondary loss for Coolia, Gambierdiscus and Pyrocystis. For species of the genus Alexandrium, multiple instances of secondary loss can explain the phylogenetic pattern of STX evolution within the genus [69]. To further understand this pattern of loss, a more resolved phylogeny of the order Gonyaulacales is vital. The sxtA sequence of *P. bahamense* is currently unavailable, thus sequence comparison remains to be conducted in the future [100].

We were unable to detect sxtA for any species external to Gonyaulacales and Gymnodiniidae. In addition, sxtA was undetected for other species of Gymnodinium sensu stricto tested. The dissimilar morphology of Gymnodiniidae and Gonyaulacales would support a distant relationship, and phylogenetic studies have tended to support this [16,24,36]. The current study adds statistical support to this relationship, suggesting that the acquisition of sxtA in Gymnodinium catenatum was possibly due to a secondary dinoflagellate-dinoflagellate HGT. SxtA was not detected in any additional dinoflagellate species and orders to those already reported (Fig. 3–4; Table 1) [56]. The result further demonstrates the capabilities of the sxtA primers for the detection of environmental STX [56].

**Future Aims and Perspectives**

This study improves dinoflagellate in-group resolution considerably, however some relationships remain unclear. Presently, the Marine Microbial Eukaryotic Transcriptome Project (https://www.marinemicroeukaryotes.org/) are sequencing 142 dinoflagellate strains, spanning eight orders. Once these data become publicly available, it will be possible to further increase the phylogenetic resolution of dinoflagellates. To increase resolution yet further, a focus is needed on *inertia sedis* taxa. These species are either heterotrophic and unculturable, or rare and not available in culture collections. For example the benthic genera Rhinodinium [109], Cabra [110], Pseudoheudinaeum [111] and Hidiostylodinium [112] have unclear family level affinities. *Adenoids* [105], *Pyrodinium* [113], *Pileudinium* [114] and *Tovelliaeae* [115] have unclear order-level affinities. This study demonstrates that improved taxon sampling is as important, if not more important, as increasing the number of inferred genes [95,96,97] in order to obtain a resolved phylogeny.

In relation to STX, the characterization of additional pathway genes is vital to determine the toxin evolution. This is needed to further corroborate the HGT theory and determine where in dinoflagellate evolution this may have occurred. It is important to understand the pattern of STX loss further. For example, which genes have been lost and from what lineages? Have genes been retained and are they being transcribed? Are there remnants of sxtA in the genome of non-toxic species? Such questions highlight the importance of work on non-STX producing species, whilst most focus has been on their toxic sisters.

**Supporting Information**

**Figure S1** Phylogenetic tree of dinoflagellates inferred from rDNA, mitochondrial and nuclear protein genes. Concatenated phylogeny, inferred from 18S+5.8S+28S+cob+cox1+actin+beta-tubulin+18S (7138 characters). The tree is inferred as in Figure 3, however, *Heterocapsa* cytochrome has not been excluded. The tree is reconstructed with Bayesian inference (MrBayes). Numbers on the internal nodes represent posterior probability and bootstrap values (>50%) for MrBayes and RAxML (ordered; MrBayes/RAxML). Black circles indicate a posterior probability value of 1.00 and bootstrap >90%. *N. scintillans* is represented with a dashed branch as this taxon was excluded from the inference; alternatively its most “probable” placement was determined from a parallel Bayesian analysis. *Denotes* taxa sequences generated from this study. See Table S2 for a full listing of accessions used. (EPS)

**Figure S2** Phylogenetic tree of dinoflagellates inferred from mitochondrial and nuclear protein genes. Concatenated phylogeny, inferred from cob+cox1+actin+beta-tubulin+18S (4238 characters). The cytochrome genes cob and *cox1* for *H. triqueta* were excluded from the inference. The tree is reconstructed with ML (RAxML). Numbers on the internal nodes represent bootstrap values (>50%). *Denotes* taxa sequences generated from this study. See Table S2 for a full listing of accessions used. The observed topology demonstrates good topological congruence with the equivalent rDNA inference: *l*, *P*-value <0.05. (EPS)

**Figure S3** Comparative rate test of *Heterocapsa* genes. The distance ratio of *Heterocapsa* to nine “core” dinoflagellate (ingroup) taxa compared to mean pairwise distance for the same ingroup taxa, calculated for every gene. When the ratio is greater than one, *Heterocapsa* can be considered more divergent than the “core” dinoflagellate mean. A ratio less than one, the “core” dinoflagellate mean is more divergent than *Heterocapsa*. When the ratio is approximately one divergence between *Heterocapsa* and the “core” dinoflagellate mean is homogenous. The box spans the 25th to 75th percentile with the horizontal bar indicating the 50th percentile. The whiskers include the entire range from 0 to 100 percentile. (EPS)

**Figure S4** 18S rDNA phylogeny (1724 characters): The tree is reconstructed with ML (RAxML). Numbers on the internal nodes represent bootstrap values (>50%). *Denotes* taxa sequences generated from this study. Accessions for each taxon are shown in brackets. (EPS)

**Figure S5** 5.8S rDNA phylogeny (151 characters): The tree is reconstructed with ML (RAxML). Numbers on the internal nodes represent bootstrap values (>50%). *Denotes* taxa sequences
generated from this study. Accessions for each taxon are shown in brackets.

**Figure S6** 28S rDNA phylogeny (1025 characters): The tree is reconstructed with ML (RAxML). Numbers on the internal nodes represent bootstrap values (>50%). * Denotes taxa sequences generated from this study. Accessions for each taxon are shown in brackets.

**Figure S7** Actin mRNA phylogeny (752 characters): The tree is reconstructed with ML (RAxML). Numbers on the internal nodes represent bootstrap values (>50%). * Denotes taxa sequences generated from this study. Accessions for each taxon are shown in brackets.

**Figure S8** Beta tubulin mRNA phylogeny (842 characters): The tree is reconstructed with ML (RAxML). Numbers on the internal nodes represent bootstrap values (>50%). * Denotes taxa sequences generated from this study. Accessions for each taxon are shown in brackets.

**Figure S9** Cytochrome Oxidase 1 mRNA phylogeny (892 characters): The tree is reconstructed with ML (RAxML). Numbers on the internal nodes represent bootstrap values (>50%). * Denotes taxa sequences generated from this study. Accessions for each taxon are shown in brackets.

**Figure S10** Cytochrome B mRNA phylogeny (620 characters): The tree is reconstructed with ML (RAxML). Numbers on the internal nodes represent bootstrap values (>50%). * Denotes taxa sequences generated from this study. Accessions for each taxon are shown in brackets.

**Figure S11** Heat shock protein 90 gDNA phylogeny (1132 characters): The tree is reconstructed with ML (RAxML). Numbers on the internal nodes represent bootstrap values (>50%). * Denotes taxa sequences generated from this study. Accessions for each taxon are shown in brackets.

**Table S1** Primers specifically designed for this study or used in previous studies. Tm calculated using OligoCalc [73]. Annealing site is an approximation and can vary slightly between species. The primer-pairs and PCR annealing temperature used were as follows. 52°C: SL+ DinoActinF2, DinoActinF2+AUAP, SL+ DinoBtubR1, SL+ DinoBtubR2, NSF83+1528R, 18F1574+28R691new, 28F341+28R1318. 54°C: DinoActinF1+AUAP, DinoBtubF1+AUAP, DinoBtubF2+AUAP, Btub305F+AUAP, SL+Btub305R, DinoCYTbF1+AUAP, DinoCYTbF2+AUAP, DinoCOXF2+AUAP, DinoRhsp90F2+DinoRhsp90R2, 56°C: CYTB343F+AUAP, COX211F+ COX1021R, COX631F+AUAP, COX631F+ COX1021R, 18S8F+ ITS R01. 57°C: Actin943F+AUAP, 64°C: Sxt001+Sxt002, Sxt007+Sxt008.

**Table S2** Accession numbers of the species represented in the supermatrices. Accessions amplified from this study are highlighted with an asterisk.

**Acknowledgments**

We would like to thank Marianne Minge and Kamran Shalchian-Tabrizi for support and fruitful discussions. We are grateful to Shuhei Ota and the Marine Biology Program, Department of Biology, UiO for access to cultures. We thank the ABI lab at Dept of Biology, UiO, for sequencing. We thank three anonymous reviewers and the editor for constructive criticism that helped improve the paper.

**Author Contributions**

Conceived and designed the experiments: RJSO KSJ. Performed the experiments: RJSO SAM AS. Analyzed the data: RJSO SAM. Contributed reagents/materials/analysis tools: KSJ LR SAM. Wrote the paper: RJSO SAM. Commented and approved the manuscript: AS KSJ.

**References**

1. Taylor FJR (1987) General group characteristics, special features of interest, short history of dinoflagellate study. In: Taylor FJR, editor. The Biology of dinoflagellates. Oxford: Blackwell Scientific Publications. 790.
2. Hackett JD, Anderson DM, Erdner DL, Bhattacharya D (2004) Dinoflagellates: A remarkable evolutionary experiment. American Journal of Botany 91: 1523–1534.
3. Taylor FJR, Hoppernath M, Saldarriaga JF (2008) Dinoflagellate diversity and distribution. Biodiversity and Conservation 17: 407–418.
4. Dodge JD (1987) Dinoflagellate ultrastructure. In: Taylor FJR, editor. The Biology of Dinoflagellates. Oxford: Blackwell Scientific Publications. 93–119.
5. Saldarriaga JF, Taylor FJR, Cavalier-Smith T, Menden-Deuer S, Keeling PJ (2008) Dinoflagellate diversity and distribution. Biodiversity and Conservation 17: 407–418.
6. Lin S, Zhang H, Zhuang Y, Tran B, Gill J (2010) Spliced leader–based rDNA sequence data, including the erection of three new genera of dinoflagellates. European Journal of Protistology 46: 111–120.
7. Skovgaard A, Massana R, Saiz E (2007) Parasitic species of the genus Blastodinium (Blastodiniphyceae) are peridinioid dinoflagellates. Environmental Microbiology 9: 791–802.
8. Roy S, Morse D (2012) A Full Suite of Histone and Histone Modifying Genes Are Transcribed in the dinoflagellate Lingulodinium. PlosOne 7: e34340.
9. Okamoto N, Horák A, Keeling PJ (2012) Description of Two Species of Early Branching Dinoflagellates, Paramonas pacifica n. g., n. sp. and P. atlantica n. sp. PloS One 7: e49800.
10. Fukuda Y, Endoh H (2006) New details from the complete life cycle of the red-tide dinoflagellate Noctiluca scintillans (Ehrenberg) McCarthy. European Journal of Protistology 42: 209–219.
11. Fukuda Y, Endoh H (2008) Phylogenetic analyses of the dinoflagellate Noctiluca scintillans based on beta-tubulin and Hsp90 genes. European Journal of Protistology 44: 27–33.
12. Ki JS (2010) Nuclear 28S rDNA phylogeny supports the basal placement of Noctiluca scintillans (Dinophyceae; Noctilucales) in dinoflagellates. European Journal of Protistology 46: 111–120.
13. Skovgaard A, Massana R, Saiz E (2007) Parasitic species of the genus Blastodinium (Blastodiniphyceae) are peridinioid dinoflagellates. Journal of Phycolgy 43: 533–560.
14. Kühn SF, Medlin LK (2005) The systematic position of the parasitoid marine dinoflagellate Paonella vantssiki (Dinophyceae) inferred from nuclear-encoded small subunit ribosomal DNA. Protist 156: 393–396.
15. Gast RJ (2006) Molecular phylogeny of a potentially parasitic dinoflagellate isolated from the solitary rhodolith, Thalassia cunctata. Journal of Eukaryotic Microbiology 53: 43–45.
16. Hoppernath M, Leander BS (2010) Dinoflagellate Phylogeny as Inferred from Heat Shock Protein 90 and Ribosomal Gene Sequences. PlosOne 5: e13220.
17. Daugbjerg N, Hansen G, Larsen J, Moestrup O (2000) Phylogeny of some of the major genera of dinoflagellates based on ultrastructure and partial LSU rDNA sequence data, including the erection of three new genera of unarmoured dinoflagellates. Phycologia 39: 302–317.
18. Takayama H (1985) Apical grooves of unarmoured dinoflagellates. Bulletin of Plankton Society of Japan 32: 129–140.
19. Jorgensen MF, Murray S, Daugbjerg N (2004) A new genus of athecate interstitial dinoflagellates, Togula gen. nov., previously encompassed within Amphidinium xenax xenax: Inferred from light and electron microscopy and phylogenetic analyses of partial large subunit ribosomal DNA sequences. Phycological Research 52: 204–299.
20. Dodge JD, Crassford RM (1968) Fine structure of the dinoflagellate Amphidinium carteri. Buffalo. Protistologica 4: 231–242.
21. Leander BS, Keeling PJ (2004) Early evolutionary history of dinoflagellates and apicoplasts (Alveolata) as inferred from hsp96 and actin phylogenies. Journal of Phycolgy 40: 341–350.

22. Saldarriaga JF, Keeling PJ, Earle FM, Taylor FJR (2005) Multiple protein phylogenies show that Oxyrrhis and Peridinium marinus are early branches of the dinoflagellate lineage. International Journal of Systematic and Evolutionary Microbiology 55: 335–365.

23. Zhang H, Bhatchariya D, Lin S (2007) A three-gene dinoflagellate phylogeny suggests monophyly of procentrales and a basal position for Amphidinium and Peridinium. Journal of Systematic and Evolutionary Microbiology 57: 465–474.

24. Shalchian-Tabrizi K, Minge MA, Silberman JD, Orr RJS, Cavalier-Smith T, Shalchian-Tabrizi K, et al. (2009) Evolutionary position of breviate amoebeae and the primary eukaryote divergence. Proceedings of the Royal Society B: Biological Sciences 276: 597–604.

25. Minge MA, Silberman JD, Orr RJS, Cavalier-Smith T, Shalchian-Tabrizi K, et al. (2009) Reconciling morphological and molecular based phylogenies. Biology Letters 5: 366–406.

26. Fensome RA, Saldarriaga JF, Taylor FJR (1999) Dinoflagellate phylogeny revisited: reconciling morphological and molecular based phylogenies. Grana 38: 66–80.

27. Burgki F, Shalchian-Tabrizi K, Pawlsowski J (2008) Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. Biology Letters 4: 463–465.

28. Murray SA, Mihali TK, Neilan BA (2011) Extraordinary Conservation, Gene Loss, and Positive Selection in the Evolution of an Ancient Neurotoxin. Molecular Biology and Evolution 28: 1173–1182.

29. Stucken K, John U, Cembella A, Murillo AA, Soto-Liebe K, et al. (2010) The smallest known genomes of multicellular and toxic cyanobacteria: Comparison, Minimal Gene Sets for Linked Traits and the Evolutionary Implications. PLoS One 5: e9252.

30. Tillmann U, Salas R, Gottschling M, Krock B, O'Driscoll D, et al. (2012) Morphological diversity of globally distributed Syndiniales (Dinophyceae) inferred from SSU and LSU rDNA sequences. BMC Evolutionary Biology 9: 237–259.

31. Gaines G, Elbracht M (1987) Heterotrophic nutrition. In: Taylor FJR, editor. Dinoflagellates. PlosOne 6: e20096.

32. Moestrup Ø, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.

33. Saunders GW, Hill DRA, Sexton JP, Andersen RA (1997) Small-subunit ribosomal RNA gene sequences from selected dinoflagellates: testing classical evolutionary hypotheses with molecular systematic methods. Plant Systematics and Evolution 11: 237–259.

34. Saunders GW, Hill DRA, Sexton JP, Andersen RA (1997) Small-subunit ribosomal RNA sequences from selected dinoflagellates: testing classical evolutionary hypotheses with molecular systematic methods. Plant Systematics and Evolution 11: 237–259.

35. Hoppenrath M, Bachvaroff TR, Handy SM, Delwiche CF, Leander BS (2009) The ‘megagroup’ including most photosynthetic eukaryotes. Biology Letters 4: 366–406.

36. Moustafa A, Loram JE, Hackett JD, Anderson DM, Plumley FG, et al. (2009) Origin of Saxitoxin Biosynthetic Genes in Cyanobacteria. PlosOne 4: e5720.

37. Moestrup O, Daugbjerg N, Fernández-Tejedor M (2006) On the Identity of Kaldaluinum species and Description of Kaldaluinum amargae sp. nov., (Dinophyceae), Based on Light and Electron Microscopy, Nuclear-Encoded Genes and functional homology reveal a saxitoxin gene cluster in cyanobacteria. Applied and Environmental Microbiology 74: 444–453.

38. Fensome RA, Saldarriaga JF, Taylor FJR (1999) Dinoflagellate phylogeny revisited: reconciling morphological and molecular based phylogenies. Grana 38: 66–80.

39. John U, Fensome RA, Medlin LK (2003) The application of a molecular clock based on molecular sequences and the fossil record to explain biogeographic distributions within the Alexandrium tamarense ‘species complex’ (Dinophyceae). Molecular Biology and Evolution 20: 1015–1027.

40. Murray SA, Mihali TK, Neilan BA (2011) Extraordinary Conservation, Gene Loss, and Positive Selection in the Evolution of an Ancient Neurotoxin. Molecular Biology and Evolution 28: 1173–1182.

41. Stucken K, John U, Cembella A, Murillo AA, Soto-Liebe K, et al. (2010) The smallest known genomes of multicellular and toxic cyanobacteria: Comparison, Minimal Gene Sets for Linked Traits and the Evolutionary Implications. PLoS One 5: e9252.

42. Murray SA, Mihali TK, Neilan BA (2011) Extraordinary Conservation, Gene Loss, and Positive Selection in the Evolution of an Ancient Neurotoxin. Molecular Biology and Evolution 28: 1173–1182.

43. Gaines G, Elbracht M (1987) Heterotrophic nutrition. In: Taylor FJR, editor. Dinoflagellates. PlosOne 6: e20096.

44. Moestrup O, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.

45. Moestrup O, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.

46. Moestrup O, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.

47. Alam M, Ikawa M, Sasner JJ, Sawyer PJ (1973) Purification of saxitoxin from Aphanizomenon flos aquae- toxin and its chemical and physiological properties. Toxicon 11: 65–72.

48. Shannon J, Lynch JM, Vayvdad G, Matsumot K, Rapport H (1966) Purification and characterization of poison produced by Gonyaulax catenella in axenic culture. Biochemistry 5: 1191–1195.

49. Shimizu Y (1996) Microalgal metabolites: a new perspective. In: Ornston LN, editor. Algal Secondary Metabolites: Biosynthesis, Bioactivity, and Biotechnology. Boca Raton: CRC Press. 11–15.

50. Saiz-Jimenez C, Moestrup O, Bierau M, Plante B, Trimpont J, et al. (2010) Determination of the Thoracosphaeraceae (Dinophyceae), Including the Identity of Alexandrium tamarense. Journal of Phycology 46: 203–220.

51. Moustafa A, Loram JE, Hackett JD, Anderson DM, Plumley FG, et al. (2009) Origin of Saxitoxin Biosynthetic Genes in Cyanobacteria. PlosOne 4: e5720.

52. Moestrup O, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.

53. Moestrup O, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.

54. Moestrup O, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.

55. Murray SA, Mihali TK, Neilan BA (2011) Extraordinary Conservation, Gene Loss, and Positive Selection in the Evolution of an Ancient Neurotoxin. Molecular Biology and Evolution 28: 1173–1182.

56. Stucken K, John U, Cembella A, Murillo AA, Soto-Liebe K, et al. (2010) The smallest known genomes of multicellular and toxic cyanobacteria: Comparison, Minimal Gene Sets for Linked Traits and the Evolutionary Implications. PLoS One 5: e9252.

57. Fensome RA, Saldarriaga JF, Taylor FJR (1999) Dinoflagellate phylogeny revisited: reconciling morphological and molecular based phylogenies. Grana 38: 66–80.

58. Biecheleria TR, Hardy SM, Delwiche CF, Leander BS (2009) Molecular phylogeny of the ocellot-boring dinoflagellates, Protoperidinium, Diptilomya and Protoperidinium (Dinophyceae), inferred from large subunit rDNA. Journal of Phycology 45: 1081–1095.

59. Saldarriaga JF, Taylor FJR, Keeling PJ, Cavalier-Smith T (2001) Dinoflagellate nuclear LSU rDNA phylogeny suggests multiple plastid losses and replacements. Journal of Molecular Evolution 53: 204–213.

60. Moestrup O, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.

61. Gottschling M, Soehner S, Zinsmeister C, John U, Plötzner J, et al. (2012) Determination of the Thoracosphaeraceae (Dinophyceae), Including the Calcareous Dinoflagellates, Based on Large Amounts of Ribosomal RNA Sequence Data. PLoS ONE 163: 15–24.

62. Taylor FJR (1990) Phylum Dinoflagellata. In: Margulis L, Corliss JO, editors. Handbook of Protoctista. Boston: Jones and Bartlett Publishers. 419–437.

63. Moestrup O, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.

64. Moestrup O, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.

65. Moestrup O, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid gene clusters in cyanobacteria. PlosOne 4: e5720.
96. Heath TA, Zwicki DJ, Kim J, Hillis DM (2008) Taxon sampling affects inferences of macroevolutionary processes from phylogenetic trees. Systematic Biology 57: 160–166.

97. Parfrey LW, Grant J, Tekle YL, Lasek-Nesselquist E, Morrison HG, et al. (2010) Broadly Sampled Multigene Analyses Yield a Well-Resolved Eukaryotic Tree of Life. Systematic Biology 59: 518–533.

98. Wieser JH, Hackett JD (2011) Dinoflagellate Genome Evolution. Annual Review of Microbiology, Vol 65: 369–387.

99. Taylor FJR (2001) Dinoflagellates. eLS. Chichester: John Wiley & Sons, Ltd.

100. Jørgensen MF, Murray S, Daugbjerg N (2004) Amphidinium Revisited. I. Redefinition of Amphidinium (Dinophyceae) based on cladistic and molecular phylogenetic analyses. Journal of Phycology 40: 351–365.

101. Büják JP, Williams GL (1990) The evolution of dinoflagellates. Canadian Journal of Botany 59: 2077–2087.

102. Loeblich ARD (1976) Dinoflagellate Evolution: speculation and evidence. Journal of Protozoology 23: 13–28.

103. Dorhofer G, Davies EH (1990) Evolution of archeophyte and tabulation in rhizogonyaulacine dinoflagellate cysts. Toronto: Royal Ontario Museum. 91 p.

104. Murray S, Ip CLC, Moore R, Nagahama Y, Fukuyo Y (2009) Are Procentroid Dinoflagellates Monophyletic? A Study of 25 Species Based on Nuclear and Mitochondrial Genes. Protist 160: 245–264.

105. Hoppenrath M, Schwekert M, Elbracht M (2003) Morphological reinvestigation and characterization of the marine, sand-dwelling dinoflagellate Adenoides clados (Dinophyceae). European Journal of Phycology 38: 383–394.

106. Hansen G, Daugbjerg N (2011) Micractinium oblonga gen. & comb. nov. (syn.: Gyrodinium oblongum), a new marine dinoflagellate genus characterized by light and electron microscopy, photoenzymatic pigments and LSU-dsDNA sequence. Phycology 50: 583–599.

107. Qiu DJ, Huang LM, Liu S, Lin SJ (2011) Nuclear, Mitochondrial and Plastid Gene Phylogenies of Dinophyceae: Evidence of Variable Types of Chloroplasts. PlosOne 6: e29398.

108. Hackett JD, Wisecaver JH (2011) Dinoflagellate Genome Evolution. Annual Review of Microbiology, Vol 65: 369–387.

109. Hobday ARD, Moore R, Nagahama Y, Nagahama Y, Fukuyo Y (2009) Are Procentroid Dinoflagellates Monophyletic? A Study of 25 Species Based on Nuclear and Mitochondrial Genes. Protist 160: 245–264.

110. Murray S, Patterson DJ (2004) Cabra manta, gen. nov., sp. nov., a new benthic, heterotrophic dinoflagellate. European Journal of Phycology 39: 229–234.

111. Hoppenrath M, Selina M (2006) Gyrodinium oblongum gen. et sp. nov., a peridinioid, sand-dwelling dinoflagellate (Dinophyceae). Journal of Phycology 42: 934–942.

112. Murray S, Hoppenrath M, Prisfeldt A, Larsen J, Yoshimatsu S, et al. (2006) Phylogenetics of Rhadinodinium brevicolor gen. et sp. nov., a peridinioid, sand-dwelling dinoflagellate (Dinophyceae). Journal of Phycology 42: 934–942.

113. Faust MA, Balech E (1993) A Further SEM Study of Marine Benthic Dinoflagellates from a Mangrove Island, Twin Cays, Belize, Including Pileidinium ciceropse gen. et sp nov (Dinophyceae). Journal of Phycology 39: 229–234.

114. Tamura M, Horiguchi T (2005) Halostylodinium arenarium gen. et sp. nov. (Dinophyceae), a coccoid sand-dwelling dinoflagellate from subtropical Japan. Journal of Phycology 41: 960–971.

115. Lindberg K, Moestrup O, Daugbjerg N (2005) Studies on woloszynskioid dinoflagellates I: Woloszynskia corona re-examined using light and electron microscopy and partial LSU rDNA sequences, with description of Tovellia gen. nov. and Jadwigia gen. nov. (Tovelliaeae fam. nov.). Phycologia 44: 416–410.