OPEN ACCESS Freely available online

PLOS ONE

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

Russell J. S. Orr[1], Shauna A. Murray[2,3], Anke Stüken[1], Lesley Rhodes[4], Kjetill S. Jakobsen[1,5]*

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 Sciences, 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 within 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 descendent species of *Alexandrium*.

Citation: Orr RJS, 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/V40 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 amphiesmal 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 Oxyrrhinaceae 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 amphiesmal vesicles is an important character in distinguishing clades of dinoflagellates [16]. The thecate (armored) orders (Dinophysiales, Gonyaulacales, Peridiniales, Prorocentrales and Suessiales) have comparatively fewer, large amphiesmal 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 dinoflagellates 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 dinoflagellate phylogeny based on more than two concatenated genetic markers [16,23]. Presently, sequence data is only available for approximately 10% or less of the known dinoflagellate 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 dinoflagellate lineage [43], are difficult or impossible to culture.

A well-resolved dinoflagellate phylogeny is essential to understanding the evolution of toxin synthesis in this phylum. Approximately 100 known species of dinoflagellates 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 dinoflagellates and freshwater cyanobacteria [47,48]. The toxins appear to be synthesized by similar processes in cyanobacteria and dinoflagellates [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 dinoflagellates *Gymnodinium catenatum* and multiple species within the genus *Alexandrium* [56]. The origin of this gene cluster within the dinoflagellates may have occurred by way of a horizontal gene transfer (HGT) event between an ancestral STX-producing bacterium and the dinoflagellates, before *Alexandrium* 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 within *Alexandrium*, an independent acquisition of STX from a dinoflagellate-dinoflagellate transfer has been postulated [56]. As the phylogenetic relationship between dinoflagellate 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 dinoflagellate 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://dinoflaj.smu.ca/) [6,57]. However, the following amendments are used: "core" dinoflagellates for dinokaryota [58], family Kareniaceae [17,59], the inclusion of *Brachidinium* in Kareniaceae [60], genus *Thoracosphaera* [61], genus *Adenoides* [62], *Biecheleria baltica* (syn. *Woloszynskia halophila*) [63], *Pelagodinium beii* (syn. *Gymnodinium beii*) [64], and *Protodinium simplex* (syn. *Gymnodinum* simplex) [58].

Culturing

The dinoflagellate 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 NSF83 - 1528R and 18sF8 - ITSR01 [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	UIO081	n.d.	n.d.
Amphidinium massartii	CS-259	n.d. [56]	n.d. [56]
Amphidinium mootonorum	CAWD161	n.d.	n.d.
Azadinium spinosum	RCC2538	n.d.	n.d.
Ceratium longipes	CCMP1770	n.d.	n.d.
Coolia monotis	CAWD98	n.d.	n.d.
Gambierdiscus australes	CAWD148	n.d.	n.d.
Gymnodinium aureolum	SCCAP K-1561	n.d.	n.d.
Heterocapsa triquetra	RCC2540	n.d.	n.d.
Karlodinium veneficum	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	UIO292	n.d.	n.d.
Prorocentrum minimum	UIO085	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 RNeasy 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 (*cox1*), 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 *cox1,* 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 HotStarTaq *Plus* polymerase (Qiagen), Bioline Mytag polymerase (Bioline) or BD Advantage 2 polymerase (Clonetech) 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 *sxtA*1 and *sxtA*4 genes were tested for all dinoflagellate strains following the protocol described in [56]. The *sxtA*1 fragment was amplified with primers sxt001 & sxt002 (~550 bp) and the *sxtA*4 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 *sxtA* 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, *cox1* 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 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*+*cox1*+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. perezi*, *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 MODELTEST [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

Table 2. Primers designed specifically for this study: Annealing site is an approximation and can vary slightly between species; *Prorocentrum minimum* was used as a reference.

Primer name	Primer direction	Primer sequence 5′-3′	Annealing site 5′-3′
DinoActinF1	F	GAYGARGCDCAGAGCAAGC	169–187
DinoActinF2	F	ATCATGGTSGGCATGGAC	130–147
DinoActinR2	R	TTGGAGATCCACATCTGCTG	1060–1079
Actin943F	F	ATGAAGATCAAGGTNGTNGC	976–995
DinoBtubF1	F	GGHGCNAARTTYTGGGAGG	49–67
DinoBtubF2	F	GDGCMAAGTTCTGGGARGT	50–68
DinoBtubR1	R	AGGTGGTTCAGGTCHCCGTA	665–683
DinoBtubR2	R	YTCWCCDGTGTACCARTGCAA	1183–1203
Btub305F	F	TSCAGGGBTTCCAGATGT	389–406
Btub305R	R	ACATCTGGAAVCCCTGSA	389–406
DinoCYTbF1	F	WCHGGWATCTTCTTAGCTTTACATTA	73–98
DinoCYTbF2	F	TTRTCACWGGAATCTTMTTAGSTTT	68–92
CYTB343F	F	GGACAAATGAGTTTMTGGGG	343–362
DinoCOXF2	F	CCATTAAGCACKTCTTTYMTGAGTT	349–373
COX211F	F	ATCTTTCAAGGRTCTCCWGAAGTG	211–234
COX631F	F	TTTGGAGGAGATCCTRTWCTCTAT	631–654
COX1021R	R	CCAAGAATTACTCCTGTTGASCC	1021–1043
DinoRhsp90F2	F	ATCCGSTAYGAGTCVATCAC	46–65
DinoRhsp90R2	R	ACCTTGTCKCCSARVACCT	1577–1595

For a full list of primers, primer pairs and annealing temperatures used see Table. S1.
doi:10.1371/journal.pone.0050004.t002

using the l_{cong} index: http://max2.ese.u-psud.fr/bases/upresa/pages/devienne/index.html [89].

Noctiluca 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 [85,86] 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 Bioportal [93] at the University of Oslo (http://www.bioportal.uio.no/).

Results

Sequence Amplification and Assembly

Fifty-five dinoflagellate sequences from 13 species and four orders were successfully amplified for 18S, 5.8S, 28S, actin, beta-tubulin, *cob*, *cox1*, and *hsp90*. All sequences generated in this study have been deposited in Genbank under the accession numbers, 18S: JX262491 and JX262492, 5.8S: JX262493-JX262497, 28S: JX262498, actin: JX262499-JX262509, beta-tubulin: JX262510-JX262519, *cox1*: JX262520-JX262529, *cob*: JX262530-JX262538, and *hsp90*: JX262539-JX262545.

Alignments generated in this study represent the broadest taxon sampling and character number inferred for the dinoflagellates to date. They have been made freely available at TreeBASE (http://www.treebase.org/treebase-web/home.html) under the accession URL: http://purl.org/phylo/treebase/phylows/study/TB2:S12493.

Amphidinium carterae (UIO081), *Ceratium longipes* (CCMP1770), *Coolia monotis* CAWD98, *Heterocapsa triquetra* (RCC2540), *Karlodinium veneficum* (RCC2539), *Lingulodinium polyedrum* (CCMP1931), *Prorocentrum micans* (UIO292), *Prorocentrum minimum* (UIO085), *Pyrocystis noctiluca* (CCMP732) and *Protoceratium reticulatum* CAWD99 were only used for *sxtA* detection via PCR.

The Phylogeny of Dinoflagellates

All inferred dinoflagellate phylogenies show good topological congruence with an l_{cong} *P*-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 topological impact (data not shown). In addition, the inference of the corresponding translated supermatrices had minimal topological

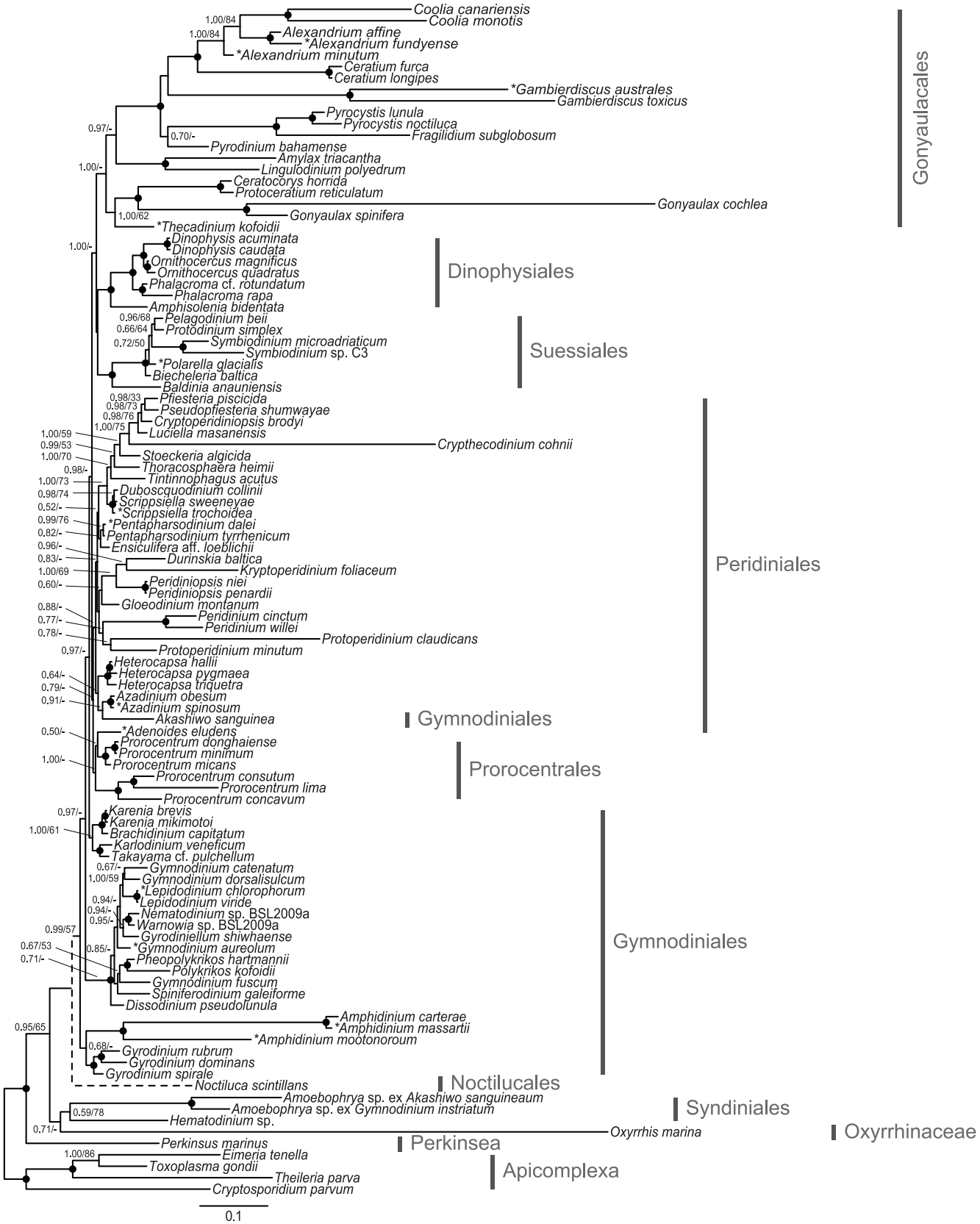

Figure 1. Phylogenetic tree of dinoflagellates inferred from rDNA. Concatenated phylogeny, inferred from 18S+5.8S+28S rDNA (2900 characters). The tree is reconstructed with Bayesian inference (MrBayes). Numbers on the internal nodes represent posterior probability and bootstrap values (>50%) for MrBayes 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

was determined from a parallel Bayesian analysis. * Denotes taxa sequences generated from this study. See Table S2 for a full listing of accessions used.
doi:10.1371/journal.pone.0050004.g001

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 Oxyrrhinaceae (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" Oxyrrhinaceae 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 Gymnodiniales (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 Gymnodiniales, 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 Kareniaceae [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 Gymnodiniales clades were basal to *Gymnodinium sensu stricto* (Fig. 2–3), excluded with moderate support (1.00/60) in the rDNA+mitochondrial+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 *Akashiwo* was unstable, placing within the Peridiniales with few inferred characters (Fig. 1–2). Increasing character number resulted in a position within the Gymnodiniales (Fig. 3–4). The Noctilucales showed an affinity to *Akashiwo*, and thus Gymnodiniales in the largest dataset (Fig. 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 (Gymnodiniales, Noctilucales, Oxyrrhinaceae and Syndiniales) were excluded from the monophyletic thecate (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 *Akashiwo* within the Peridiniales resulted in Gymnodiniales 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, *Adenoides* [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 Peridiniales clade in all phylogenies excluding the cytochrome genes for this lineage (Fig. 1, 2, 3, 4). The inclusion of the mitochondrial genes resulted in *Heterocapsa* being recovered as the most basal "core" dinoflagellate (Fig. S1). *Heterocapsa* and *Adenoides* were constantly recovered paraphyletic to main Peridiniales clade (Fig. 1, 2, 3, 4). Though, *Azadinium* (Fig. 1) and *Kryptoperidinium* were also recovered paraphyletic to the main Peridiniales 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 a 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 50^{th} percentile range of 0.14360–0.73360, and a 25^{th}–75^{th} percentile range of 0.04607–0.77480. The nuclear protein genes diverged approximately 2.38 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 50^{th} percentile of 3.58400–3.64000 and a 25^{th}–75^{th} 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.

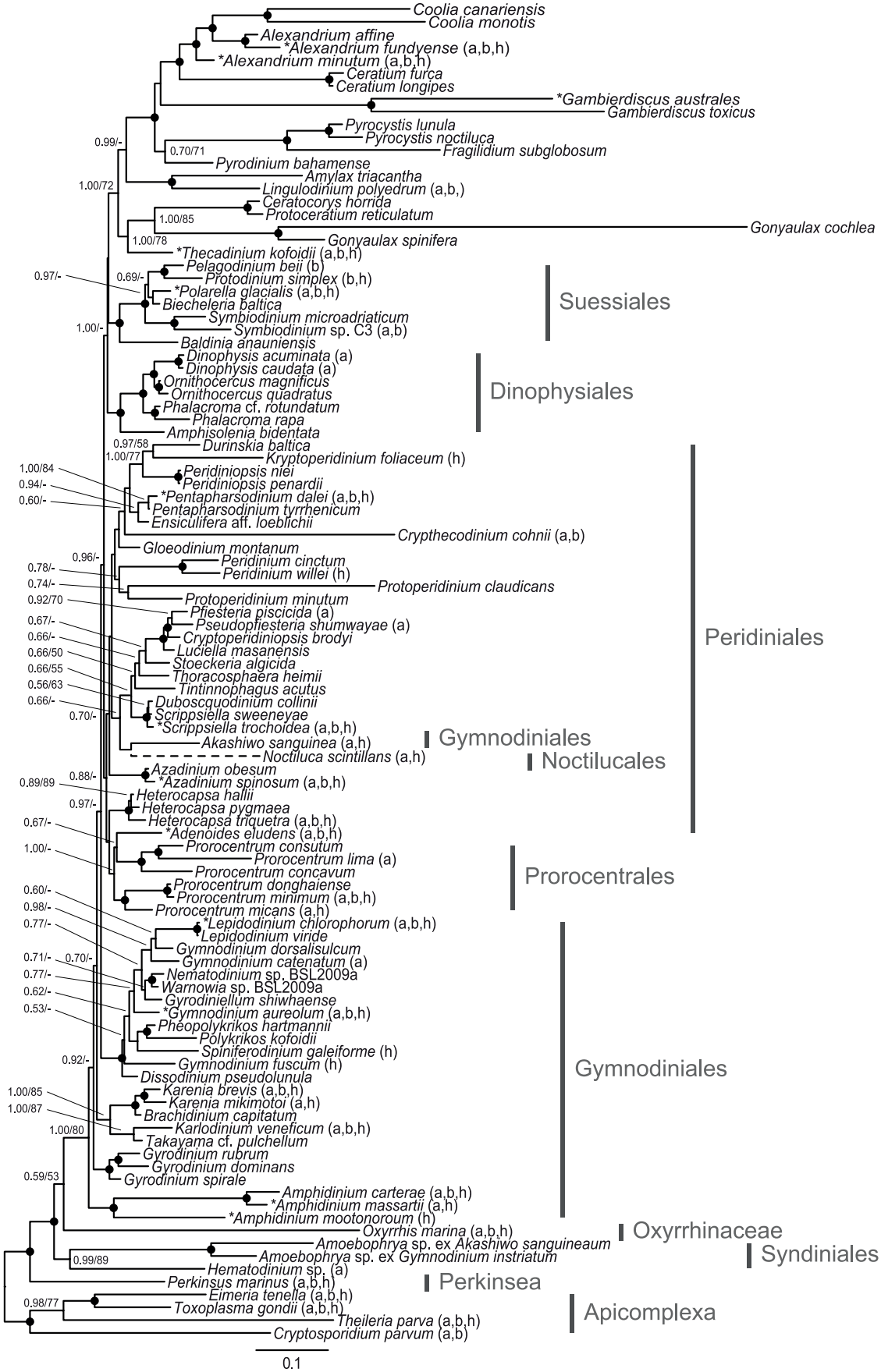
Coolia canariensis
Coolia monotis
Alexandrium affine
*Alexandrium fundyense (a,b,h)
*Alexandrium minutum (a,b,h)
Ceratium furca
Ceratium longipes
*Gambierdiscus australes
Gambierdiscus toxicus
Pyrocystis lunula
Pyrocystis noctiluca
Fragilidium subglobosum
Pyrodinium bahamense
Amylax triacantha
Lingulodinium polyedrum (a,b,)
Ceratocorys horrida
Protoceratium reticulatum
Gonyaulax cochlea
Gonyaulax spinifera
*Thecadinium kofoidii (a,b,h)
Gonyaulacales
Pelagodinium beii (b)
Protodinium simplex (b,h)
*Polarella glacialis (a,b,h)
Biecheleria baltica
Symbiodinium microadriaticum
Symbiodinium sp. C3 (a,b)
Baldinia anauniensis
Suessiales
Dinophysis acuminata (a)
Dinophysis caudata (a)
Ornithocercus magnificus
Ornithocercus quadratus
Phalacroma cf. rotundatum
Phalacroma rapa
Amphisolenia bidentata
Dinophysiales
Durinskia baltica
Kryptoperidinium foliaceum (h)
Peridiniopsis niei
Peridiniopsis penardii
*Pentapharsodinium dalei (a,b,h)
Pentapharsodinium tyrrhenicum
Ensiculifera aff. loeblichii
Crypthecodinium cohnii (a,b)
Gloeodinium montanum
Peridinium cinctum
Peridinium willei (h)
Protoperidinium claudicans
Protoperidinium minutum
Pfiesteria piscicida (a)
Pseudopfiesteria shumwayae (a)
Cryptoperidiniopsis brodyi
Luciella masanensis
Stoeckeria algicida
Thoracosphaera heimii
Tintinnophagus acutus
Duboscquodinium collinii
Scrippsiella sweeneyae
*Scrippsiella trochoidea (a,b,h)
Akashiwo sanguinea (a,h)
Noctiluca scintillans (a,h)
Gymnodiniales
Noctilucales
Peridiniales
Azadinium obesum
*Azadinium spinosum (a,b,h)
Heterocapsa hallii
Heterocapsa pygmaea
Heterocapsa triquetra (a,b,h)
*Adenoides eludens (a,b,h)
Prorocentrum consutum
Prorocentrum lima (a)
Prorocentrum concavum
Prorocentrum donghaiense
Prorocentrum minimum (a,b,h)
Prorocentrum micans (a,h)
Prorocentrales
*Lepidodinium chlorophorum (a,b,h)
Lepidodinium viride
Gymnodinium dorsalisulcum
Gymnodinium catenatum (a)
Nematodinium sp. BSL2009a
Warnowia sp. BSL2009a
Gyrodiniellum shiwhaense
*Gymnodinium aureolum (a,b,h)
Pheopolykrikos hartmannii
Polykrikos kofoidii
Spiniferodinium galeiforme (h)
Gymnodinium fuscum (h)
Dissodinium pseudolunula
Karenia brevis (a,b,h)
Karenia mikimotoi (a,h)
Brachidinium capitatum
Karlodinium veneficum (a,b,h)
Takayama cf. pulchellum
Gyrodinium rubrum
Gyrodinium dominans
Gyrodinium spirale
Amphidinium carterae (a,b,h)
*Amphidinium massartii (a,h)
*Amphidinium mootonoroum (h)
Gymnodiniales
Oxyrrhis marina (a,b,h)
Oxyrrhinaceae
Amoebophrya sp. ex Akashiwo sanguineaum
Amoebophrya sp. ex Gymnodinium instriatum
Syndiniales
Hematodinium sp. (a)
Perkinsus marinus (a,b,h)
Perkinsea
Eimeria tenella (a,b,h)
Toxoplasma gondii (a,b,h)
Theileria parva (a,b,h)
Cryptosporidium parvum (a,b)
Apicomplexa
0.1

Figure 2. Phylogenetic tree of dinoflagellates inferred from rDNA and nuclear protein genes. Concatenated phylogeny, inferred from 18S+5.8S+28S+actin+beta-tubulin+*hsp90* (5626 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 was determined from a parallel Bayesian analysis. * Denotes taxa sequences generated from this study. See Table S2 for a full listing of accessions used. Non-ribosomal gene presence for each taxon is represented in brackets behind each species name (a: actin, b: beta-tubulin, h: *hsp90*).
doi:10.1371/journal.pone.0050004.g002

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 *cob* 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 *cob* 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 *cob* 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 [36,98]. 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 *Gyrodinium*. 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 Noctilucales 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 Suessiales. 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 Noctilucales to be a basal "core" dinoflagellate (Fig. 1). However, increasing the number of inferred characters resulted in an affinity

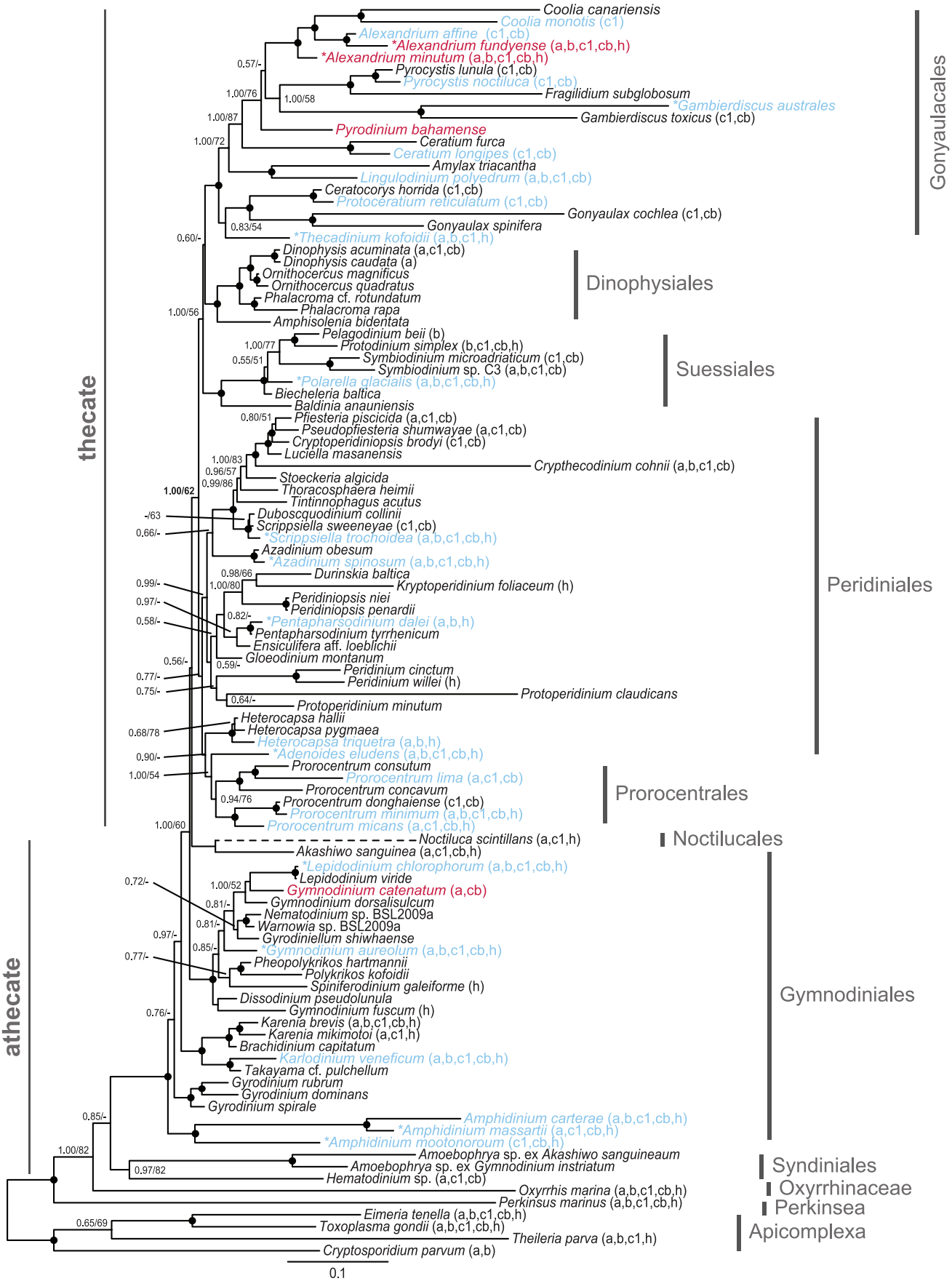
Coolia canariensis
Coolia monotis (c1)
Alexandrium affine (c1,cb)
*Alexandrium fundyense (a,b,c1,cb,h)
*Alexandrium minutum (a,b,c1,cb,h)
Pyrocystis lunula (c1,cb)
Pyrocystis noctiluca (c1,cb)
Fragilidium subglobosum
*Gambierdiscus australes
Gambierdiscus toxicus (c1,cb)
Pyrodinium bahamense
Ceratium furca
Ceratium longipes (c1,cb)
Amylax triacantha
Lingulodinium polyedrum (a,b,c1,cb)
Ceratocorys horrida (c1,cb)
Protoceratium reticulatum (c1,cb)
Gonyaulax cochlea (c1,cb)
Gonyaulax spinifera
*Thecadinium kofoidii (a,b,c1,h)
Dinophysis acuminata (a,c1,cb)
Dinophysis caudata (a)
Ornithocercus magnificus
Ornithocercus quadratus
Phalacroma cf. rotundatum
Phalacroma rapa
Amphisolenia bidentata
Pelagodinium beii (b)
Protodinium simplex (b,c1,cb,h)
Symbiodinium microadriaticum (c1,cb)
Symbiodinium sp. C3 (a,b,c1,cb)
*Polarella glacialis (a,b,c1,cb,h)
Biecheleria baltica
Baldinia anauniensis
Pfiesteria piscicida (a,c1,cb)
Pseudopfiesteria shumwayae (a,c1,cb)
Cryptoperidiniopsis brodyi (c1,cb)
Luciella masanensis
Crypthecodinium cohnii (a,b,c1,cb)
Stoeckeria algicida
Thoracosphaera heimii
Tintinnophagus acutus
Duboscquodinium collinii
Scrippsiella sweeneyae (c1,cb)
*Scrippsiella trochoidea (a,b,c1,cb,h)
Azadinium obesum
*Azadinium spinosum (a,b,c1,cb,h)
Durinskia baltica
Kryptoperidinium foliaceum (h)
Peridiniopsis niei
Peridiniopsis penardii
*Pentapharsodinium dalei (a,b,h)
Pentapharsodinium tyrrhenicum
Ensiculifera aff. loeblichii
Gloeodinium montanum
Peridinium cinctum
Peridinium willei (h)
Protoperidinium claudicans
Protoperidinium minutum
Heterocapsa hallii
Heterocapsa pygmaea
Heterocapsa triquetra (a,b,h)
*Adenoides eludens (a,b,c1,cb,h)
Prorocentrum consutum
Prorocentrum lima (a,c1,cb)
Prorocentrum concavum
Prorocentrum donghaiense (c1,cb)
Prorocentrum minimum (a,b,c1,cb,h)
Prorocentrum micans (a,c1,cb,h)
Noctiluca scintillans (a,c1,h)
Akashiwo sanguinea (a,c1,cb,h)
*Lepidodinium chlorophorum (a,b,c1,cb,h)
Lepidodinium viride
Gymnodinium catenatum (a,cb)
Gymnodinium dorsalisulcum
Nematodinium sp. BSL2009a
Warnowia sp. BSL2009a
Gyrodiniellum shiwhaense
*Gymnodinium aureolum (a,b,c1,cb,h)
Pheopolykrikos hartmannii
Polykrikos kofoidii
Spiniferodinium galeiforme (h)
Dissodinium pseudolunula
Gymnodinium fuscum (h)
Karenia brevis (a,b,c1,cb,h)
Karenia mikimotoi (a,c1,h)
Brachidinium capitatum
Karlodinium veneficum (a,b,c1,cb,h)
Takayama cf. pulchellum
Gyrodinium rubrum
Gyrodinium dominans
Gyrodinium spirale
Amphidinium carterae (a,b,c1,cb,h)
*Amphidinium massartii (a,c1,cb,h)
*Amphidinium mootonoroum (c1,cb,h)
Amoebophrya sp. ex Akashiwo sanguineaum
Amoebophrya sp. ex Gymnodinium instriatum
Hematodinium sp. (a,c1,cb)
Oxyrrhis marina (a,b,c1,cb,h)
Perkinsus marinus (a,b,c1,cb,h)
Eimeria tenella (a,b,c1,cb,h)
Toxoplasma gondii (a,b,c1,cb,h)
Theileria parva (a,b,c1,h)
Cryptosporidium parvum (a,b)
0.1
thecate
athecate
Gonyaulacales
Dinophysiales
Suessiales
Peridiniales
Prorocentrales
Noctilucales
Gymnodiniales
Syndiniales
Oxyrrhinaceae
Perkinsea
Apicomplexa

Figure 3. 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+*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. scintilans* 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. triquetra* 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

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 athecate 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

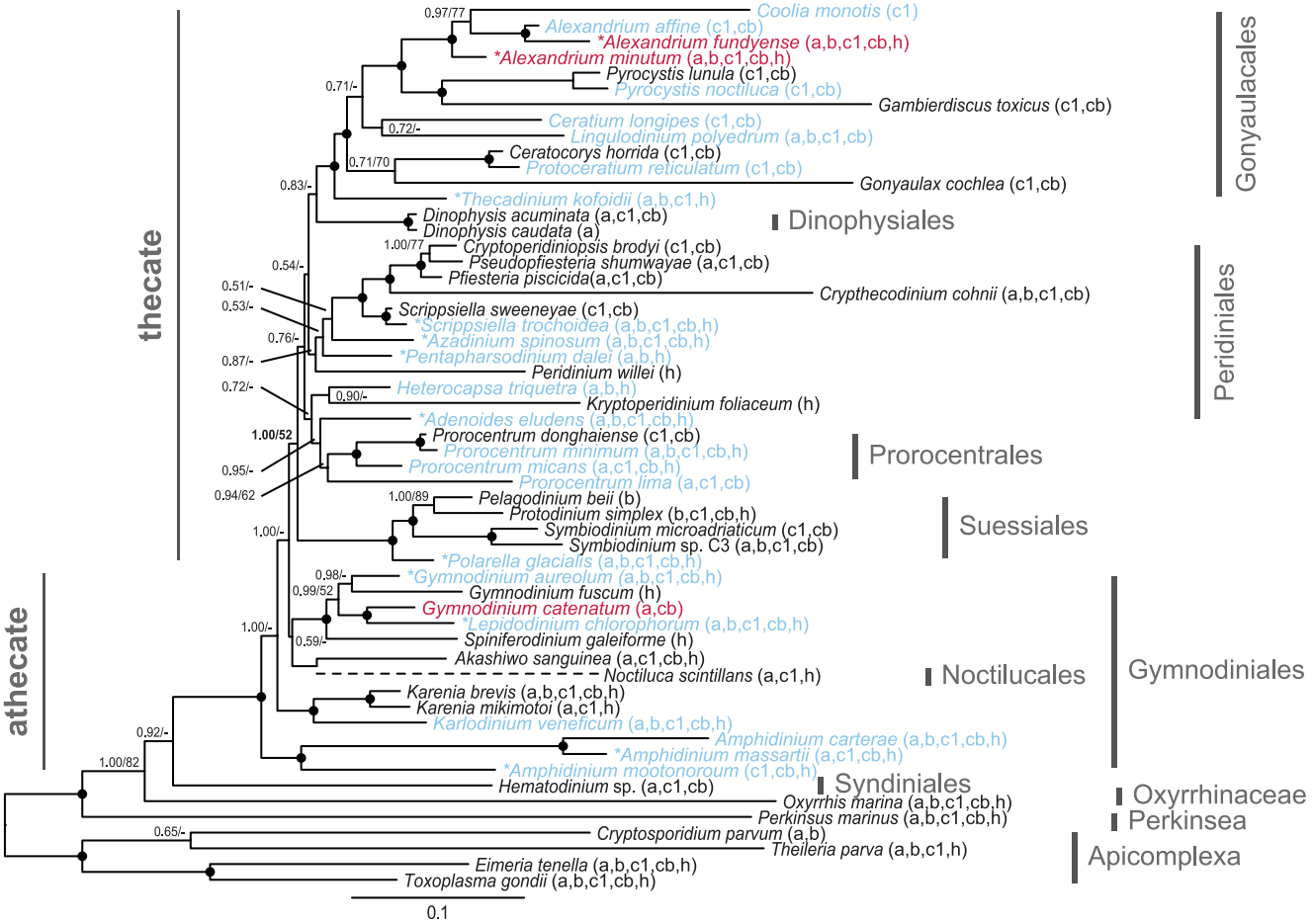

Figure 4. Phylogenetic tree of dinoflagellates inferred from rDNA, mitochondrial and nuclear protein genes (reduced phylogeny). Concatenated phylogeny, inferred from 18S+5.8S+28S+*cob*+*cox1*+actin+beta-tubulin+*hsp90* (7138 characters). This phylogeny was inferred excluding taxa with only rDNA signal; 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. scintilans* 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. triquetra* 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	1	2	3	4
Dataset	**rDNA**	**rDNA +nuclear protein**	**rDNA +mitochondrial +nuclear protein**	**Reduced rDNA +mitochondrial +nuclear protein**
Taxa/Characters	**104/2900**	**104/5626**	**104/7138**	**56/7138**
Order/Lineage/clade				
Dinoflagellates	0.95/65	1.00/95	1.00/82	1.00/82
Syndiniales	0.59/78	0.99/89	0.97/82	–
"Core" dinoflagellates	0.99/57	1.00/80	1.00/98	1.00/100
Gymnodiniales	(polyphyletic)	(polyphyletic)	(paraphyletic)	(paraphyletic)
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	(paraphyletic)	(paraphyletic)	0.94/76	0.94/62
Peridiniales	(paraphyletic)	(paraphyletic)	(paraphyletic)	(paraphyletic)
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/−	1.00/72	1.00/99	1.00/91
Dinophysiales, Gonyaulacales and Suessiales	1.00/−	1.00/−	1.00/56	–

doi:10.1371/journal.pone.0050004.t003

proposed that the Suessiales, with their numerous distinct latitudinal plates, were the most basal thecate dinoflagellate clade, with the Peridiniales giving rise to the Dinophysiales and the Prorocentrales [36]. In general, we did not find support for the Peridiniales being basal to the Dinophysiales and the Prorocentrales, however support was not conclusive (Fig. 2). Interestingly, and in agreement with this hypothesis, we found that the Suessiales was the most basal thecate lineage, with the removal of taxa only possessing rDNA signal (Fig. 4). In contrast to the position for Suessiales observed in Fig. 3, the basal placement is unsupported. The result shows no support for a trend toward sutural loss [36]. Both more genes and more taxa appear to be necessary in order to fully investigate the pattern of thecal plate evolution.

As has been found previously [23,104], the monophyly of Prorocentrales was only recovered after the inclusion of the mitochondrial cytochrome 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 peridinioids [36].

The monophyly of Dinophysiales, Gonyaulacales and Suessiales were 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 [5,24,29,30]. 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 descendent 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 descendent 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 [108].

We were unable to detect *sxtA* for any species external to Gonyaulacales and Gymnodiniales. In addition, *sxtA* was undetected for other species of *Gymnodinium sensu stricto* tested. The dissimilar morphology of Gymnodiniales 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. S*xtA* 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 *incerta 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], *Pseudothecadinium* [111] and *Halostylodinium* [112] have unclear family level affinities. *Adenoides* [105], *Plagiodinium* [113], *Pileidinium* [114] and *Tovelliaceae* [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+*hsp90* (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. scintilans* 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+*hsp90* (4238 characters). The cytochrome genes *cob* and *cox1* for *H. triquetra* 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_{cong} *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.
(EPS)

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.
(EPS)

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.
(EPS)

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.
(EPS)

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.
(EPS)

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.
(EPS)

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.
(EPS)

Table S1 Primers specifically designed for this study or used in previous studies. T_M 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+ DinoActinR2, DinoActinF2+AUAP, SL+ DinoBtubR1, SL+ DinoBtubR2, NSF83+1528R, 18F1574+28R691new, 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, 18SF8+ ITSR01. 57°C: Actin943F+AUAP. 64°C: Sxt001+Sxt002, Sxt007+Sxt008.
(DOC)

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

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. 798.
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, Hoppenrath 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 (2004) Molecular data and the evolutionary history of dinoflagellates. European Journal of Protistology 40: 85–111.
6. Fensome RA, Taylor FJR, Norris G, Sarjeant WAS, Wharton DI, et al. (1993) A classification of living and fossil dinoflagellates. Micropaleontology, Special Publication Number 7. Pennsylvania, USA: Sheridan Press. 351.
7. Lin S, Zhang H, Zhuang Y, Tran B, Gill J (2010) Spliced leader–based metatranscriptomic analyses lead to recognition of hidden genomic features in dinoflagellates. Proceedings of the National Academy of Sciences 107: 20033–20038.
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, *Psammosa pacifica* n. g., n. sp. and *P. atlantica* n. sp. PlosOne 7: e34900.
10. Fukuda Y, Endoh H (2006) New details from the complete life cycle of the red-tide dinoflagellate *Noctiluca scintillans* (Ehrenberg) McCartney. 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 Phycology 43: 553–560.
14. Kühn SF, Medlin LK (2005) The systematic position of the parasitoid marine dinoflagellate *Paulsenella vonstoschii* (Dinophyceae) inferred from nuclear-encoded small subunit ribosomal DNA. Protist 156: 393–398.
15. Gast RJ (2006) Molecular phylogeny of a potentially parasitic dinoflagellate isolated from the solitary radiolarian, *Thalassicolla nucleata*. Journal of Eukaryotic Microbiology 53: 43–45.
16. Hoppenrath 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. Jørgensen MF, Murray S, Daugbjerg N (2004) A new genus of athecate interstitial dinoflagellates, *Togula* gen. nov., previously encompassed within *Amphidinium sensu* lato: Inferred from light and electron microscopy and phylogenetic analyses of partial large subunit ribosomal DNA sequences. Phycological Research 52: 284–299.
20. Dodge JD, Crawford RM (1968) Fine structure of the dinoflagellate *Amphidinium carteri* Hulburt. Protistologica 4: 231–242.

21. Leander BS, Keeling PJ (2004) Early evolutionary history of dinoflagellates and apicomplexans (Alveolata) as inferred from *hsp90* and actin phylogenies. Journal of Phycology 40: 341–350.
22. Saldarriaga JF, McEwan ML, Fast NM, Taylor FJ, Keeling PJ (2003) Multiple protein phylogenies show that *Oxyrrhis marina* and *Perkinsus marinus* are early branches of the dinoflagellate lineage. International Journal of Systematic and Evolutionary Microbiology 53: 355–365.
23. Zhang H, Bhattacharya D, Lin S (2007) A three-gene dinoflagellate phylogeny suggests monophyly of prorocentrales and a basal position for *Amphidinium* and *Heterocapsa*. Journal of Molecular Evolution 65: 463–474.
24. Shalchian-Tabrizi K, Minge MA, Cavalier-Smith T, Nedreklepp JM, Klaveness D, et al. (2006) Combined heat shock protein 90 and ribosomal RNA sequence phylogeny supports multiple replacements of dinoflagellate plastids. Journal of Eukaryotic Microbiology 53: 217–224.
25. Minge MA, Silberman JD, Orr RJS, Cavalier-Smith T, Shalchian-Tabrizi K, et al. (2009) Evolutionary position of breviate amoebae and the primary eukaryote divergence. Proceedings of the Royal Society B-Biological Sciences 276: 597–604.
26. Burki F, Shalchian-Tabrizi K, Pawlowski J (2008) Phylogenomics reveals a new 'megagroup' including most photosynthetic eukaryotes. Biology Letters 4: 366–369.
27. Fensome RA, Saldarriaga JF, Taylor FJR (1999) Dinoflagellate phylogeny revisited: reconciling morphological and molecular based phylogenies. Grana 38: 66–80.
28. 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.
29. Murray S, Flø Jørgensen M, Ho SY, Patterson DJ, Jermiin LS (2005) Improving the analysis of dinoflagellate phylogeny based on rDNA. Protist 156: 269–286.
30. Tillmann U, Salas R, Gottschling M, Krock B, O'Driscoll D, et al. (2012) Amphidoma languida sp. nov. (Dinophyceae) Reveals a Close Relationship between Amphidoma and Azadinium. Protist 163: 701–719.
31. Gribble KE, Anderson DM (2006) Molecular phylogeny of the heterotrophic dinoflagellates, *Protoperidinium*, *Diplopsalis* and *Preperidinium* (Dinophyceae), inferred from large subunit rDNA. Journal of Phycology 42: 1081–1095.
32. Saldarriaga JF, Taylor FJ, Keeling PJ, Cavalier-Smith T (2001) Dinoflagellate nuclear SSU rRNA phylogeny suggests multiple plastid losses and replacements. Journal of Molecular Evolution 53: 204–213.
33. 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.
34. Yamaguchi A, Kawamura H, Horiguchi T (2006) A further phylogenetic study of the heterotrophic dinoflagellate genus, *Protoperidinium* (Dinophyceae) based on small and large subunit ribosomal RNA gene sequences. Phycological Research 54: 317–329.
35. Hoppenrath M, Bachvaroff TR, Handy SM, Delwiche CF, Leander BS (2009) Molecular phylogeny of ocelloid-bearing dinoflagellates (Warnowiaceae) as inferred from SSU and LSU rDNA sequences. BMC Evolutionary Biology 9: 116.
36. Taylor FJR (2004) Illumination or confusion? Dinoflagellate molecular phylogenetic data viewed from a primarily morphological standpoint. Phycological Research 52: 308–324.
37. Lopez-Garcia P, Rodriguez-Valera F, Pedros-Alio C, Moreira D (2001) Unexpected diversity of small eukaryotes in deep-sea Antarctic plankton. Nature 409: 603–607.
38. Moon-van der Staay SY, De Wachter R, Vaulot D (2001) Oceanic 18S rDNA sequences from picoplankton reveal unsuspected eukaryotic diversity. Nature 409: 607–610.
39. Skovgaard A, Meneses I, Angelico MM (2009) Identifying the lethal fish egg parasite *Ichthyodinium chabelardi* as a member of Marine Alveolate Group I. Environmental Microbiology 11: 2030–2041.
40. Brate J, Krabberod AK, Dolven JK, Ose RF, Kristensen T, et al. (2012) Radiolaria associated with large diversity of marine alveolates. Protist 163: 767–777.
41. Harada A, Ohtsuka S, Horiguchi T (2007) Species of the parasitic genus *Duboscquella* are members of the enigmatic Marine Alveolate Group I. Protist 158: 337–347.
42. Guillou L, Viprey M, Chambouvet A, Welsh RM, Kirkham AR, et al. (2008) Widespread occurrence and genetic diversity of marine parasitoids belonging to *Syndiniales* (Alveolata). Environmental Microbiology 10: 3349–3365.
43. Gaines G, Elbrächter M (1987) Heterotrophic nutrition. In: Taylor FJR, editor. The biology of dinoflagellates. Oxford: Blackwell Scientific Publications. 224–268.
44. Hallegraeff GM (1993) A review of harmful algal blooms and their apparent global increase. Phycologia 32: 79–99.
45. Deeds JR, Landsberg JH, Etheridge SM, Pitcher GC, Longan SW (2008) Non-traditional vectors for paralytic shellfish poisoning. Marine Drugs 6: 308–348.
46. Wiese M, D'Agostino PM, Mihali TK, Moffitt MC, Neilan BA (2010) Neurotoxic Alkaloids: Saxitoxin and Its Analogs. Marine Drugs 8: 2185–2211.
47. Alam M, Ikawa M, Sasner JJ, Sawyer PJ (1973) Purification of *Aphanizomenon flos- aquae* toxin and its chemical and physiological properties. Toxicon 11: 65–72.
48. Schantz EJ, Lynch JM, Vayvada G, Matsumot K, Rapoport 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. Annual Review of Microbiology. 431–465.
50. Mihali T, Kellmann R, Neilan B (2009) Characterisation of the paralytic shellfish toxin biosynthesis gene clusters in *Anabaena circinalis* AWQC131C and *Aphanizomenon* sp. NH-5. BMC Biochemistry 10: 8.
51. Moustafa A, Loram JE, Hackett JD, Anderson DM, Plumley FG, et al. (2009) Origin of Saxitoxin Biosynthetic Genes in Cyanobacteria. PlosOne 4: e5758.
52. Kellmann R, Mihali TK, Jeon YJ, Pickford R, Pomati F, et al. (2008) Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. Applied and Environmental Microbiology 74: 4044–4053.
53. 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. PlosOne 5: e9235.
54. Mihali TK, Carmichael WW, Neilan BA (2011) A Putative Gene Cluster from a *Lyngbya wollei* Bloom that Encodes Paralytic Shellfish Toxin Biosynthesis. PlosOne 6: e14657.
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. Stüken A, Orr RJS, Kellmann R, Murray SA, Neilan BA, et al. (2011) Discovery of Nuclear-Encoded Genes for the Neurotoxin Saxitoxin in Dinoflagellates. PlosOne 6: e20096.
57. Fensome RA, MacRae RA, Williams GL (2008) DINOFLAJ2, Version 1. American Association of Stratigraphic Palynologists, Data Series no 1.
58. Moestrup Ø, Daugbjerg N (2007) On dinoflagellate phylogeny and classification. In: Brodie J, Lewis JM, editors. Unraveling the Algae: The Past, Present, and Future of Algal Systematics. New York: CRC Press. 215–230.
59. Bergholtz T, Daugbjerg N, Moestrup Ø, Fernández-Tejedor M (2006) On the Identity of *Karlodinium veneficum* and Description of *Karlodinum armiger* sp. nov., (Dinophyceae), Based on Light and Electron Microscopy, Nuclear-Encoded LSU rDNA, and Pigment Composition. Journal of Phycology 42: 170–193.
60. Henrichs DW, Sosik HM, Olson RJ, Campbell L (2011) Phylogentic Analysis of *Brachidinium capitatum* (Dinophyceae) from the Gulf of Mexico Indicates Membership in the Kareniaceae. Journal of Phycology 47: 366–374.
61. Gottschling M, Soehner S, Zinssmeister C, John U, Plötner J, et al. (2012) Delimitation of the Thoracosphaeraceae (Dinophyceae), Including the Calcareous Dinoflagellates, Based on Large Amounts of Ribosomal RNA Sequence Data. Protist 163: 15–24.
62. Taylor FJR (1990) Phylum Dinoflagellata. In: Margulis L, Corliss JO, Melkonian M, Chapman DJ, editors. Handbook of Protoctista. Boston: Jones and Bartlett Publishers. 419–437.
63. Moestrup Ø, Lindberg K, Daugbjerg N (2009) Studies on woloszynskioid dinoflagellates IV: The genus *Biecheleria* gen. nov. Phycological Research 57: 203–220.
64. Siano R, Montresor M, Probert I, Not F, de Vargas C (2010) *Pelagodinium* gen. nov. and *P. beii* comb. nov., a dinoflagellate symbiont of planktonic foraminifera. Protist 161: 385–399.
65. Guillard RRL, Hargraves PE (1993) *Stichochrysis immobilis* is a diatom, not a chrysophyte. Phycologia 32: 234–236.
66. Blackburn SI, Bolch CJS, Haskard KA, Hallegraeff GM (2001) Reproductive compatibility among four global populations of the toxic dinoflagellate *Gymnodinium catenatum* (Dinophyceae). Phycologia 40: 78–87.
67. Hendriks L, Goris A, Neefs JM, Vandepeer Y, Hennebert G, et al. (1989) The Nucleotide-Sequence of the Small Ribosomal-Subunit RNA of the Yeast *Candida albicans* and the Evolutionary Position of the Fungi among the Eukaryotes. Systematic and Applied Microbiology 12: 223–229.
68. Medlin L, Elwood HJ, Stickel S, Sogin ML (1988) The Characterization of Enzymatically Amplified Eukaryotic 16s-Like Rrna-Coding Regions. Gene 71: 491–499.
69. Orr RJS, Stuken A, Rundberget T, Eikrem W, Jakobsen KS (2011) Improved phylogenetic resolution of toxic and non-toxic *Alexandrium* strains using a concatenated rDNA approach. Harmful Algae 10: 676–688.
70. Doyle JJ, Doyle JL (1987) A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemistry Bulletin, Botanical Society of America 19: 11–15.
71. Kim S, Bachvaroff TR, Handy SM, Delwiche CF (2011) Dynamics of Actin Evolution in Dinoflagellates. Molecular Biology and Evolution 28: 1469–1480.
72. Gadberry MD, Malcomber ST, Doust AN, Kellogg EA (2005) Primaclade - a flexible tool to find conserved PCR primers across multiple species. Bioinformatics 21: 1263–1264.
73. Kibbe WA (2007) OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Research 35: W43–W46.
74. Gordon D, Abajian C, Green P (1998) Consed: A graphical tool for sequence finishing. Genome Research 8: 195–202.
75. Maddison WP, Maddison DR (2000) MacClade. 4 ed. Sunderland Massachusetts: Sinauer Associates.

76. Kiryu H, Kin T, Asai K (2007) Robust prediction of consensus secondary structures using averaged base pairing probability matrices. Bioinformatics 23: 434–441.
77. Hofacker IL, Fekete M, Stadler PF (2002) Secondary structure prediction for aligned RNA sequences. Journal of Molecular Biology 319: 1059–1066.
78. Katoh K, Toh H (2008) Improved accuracy of multiple ncRNA alignment by incorporating structural information into a MAFFT-based framework. BMC Bioinformatics 9: 212.
79. Castresana J (2000) Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Molecular Biology and Evolution 17: 540–552.
80. Johnson KA, Holland BR, Heslewood MM, Crayn DM (2012) Supermatrices, supertrees and serendipitous scaffolding: Inferring a well-resolved, genus-level phylogeny of Styphelioideae (Ericaceae) despite missing data. Molecular Phylogenetics and Evolution 62: 146–158.
81. Wiens JJ (2006) Missing data and the design of phylogenetic analyses. Journal of Biomedical Informatics 39: 34–42.
82. Wiens JJ (2003) Missing data, incomplete taxa, and phylogenetic accuracy. Systematic Biology 52: 528–538.
83. Wiens JJ (2005) Can incomplete taxa rescue phylogenetic analyses from long-branch attraction? Systematic Biology 54: 731–742.
84. Posada D, Crandall KA (1998) MODELTEST: testing the model of DNA substitution. Bioinformatics 14: 817–818.
85. Stamatakis A (2006) RAxML-VI-HPC: Maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22: 2688–2690.
86. Stamatakis A (2006) Phylogenetic models of rate heterogeneity: a high performance computing perspective; 25–29 April 2006. 8 pp.
87. Huelsenbeck JP, Ronquist F, Nielsen R, Bollback JP (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. Science 294: 2310–2314.
88. Ronquist F, Huelsenbeck JP (2003) MrBayes 3: Bayesian phylogenetic inference under mixed models. Bioinformatics 19: 1572–1574.
89. de Vienne DM, Giraud T, Martin OC (2007) A congruence index for testing topological similarity between trees. Bioinformatics 23: 3119–3124.
90. Zhang H, Lin S (2008) mRNA editing and spliced-leader RNA trans-splicing groups *Oxyrrhis*, *Noctiluca*, *Heterocapsa*, and *Amphidinium* as basal lineages of dinoflagellates. Journal of Phycology 44: 703–711.
91. Bachvaroff TR, Handy SM, Place AR, Delwiche CF (2011) Alveolate Phylogeny Inferred using Concatenated Ribosomal Proteins. Journal of Eukaryotic Microbiology 58: 223–233.
92. Bachvaroff TR, Sanchez-Puerta MV, Delwiche CF (2006) Rate variation as a function of gene origin in plastid-derived genes of peridinin-containing dinoflagellates. Journal of Molecular Evolution 62: 42–52.
93. Kumar S, Skjaeveland A, Orr RJS, Enger P, Ruden T, et al. (2009) AIR: A batch-oriented web program package for construction of supermatrices ready for phylogenomic analyses. BMC Bioinformatics 10: 357.
94. Taylor FJ (1980) On dinoflagellate evolution. Biosystems 13: 65–108.
95. Dunn CW, Hejnol A, Matus DQ, Pang K, Browne WE, et al. (2008) Broad phylogenomic sampling improves resolution of the animal tree of life. Nature 452: 745–749.
96. Heath TA, Zwickl 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 YI, 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. Wisecaver JH, Hackett JD (2011) Dinoflagellate Genome Evolution. Annual Review of Microbiology, Vol 65 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. Bujak JP, Williams GL (1980) 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. Dörhöfer G, Davies EH (1980) Evolution of archeophyle and tabulation in rhaetogonyaulacinean dinoflagellate cysts. Toronto: Royal Ontario Museum. 91 p.
104. Murray S, Ip CLC, Moore R, Nagahama Y, Fukuyo Y (2009) Are Prorocentroid Dinoflagellates Monophyletic? A Study of 25 Species Based on Nuclear and Mitochondrial Genes. Protist 160: 245–264.
105. Hoppenrath M, Schweikert M, Elbrachter M (2003) Morphological reinvestigation and characterization of the marine, sand-dwelling dinoflagellate *Adenoides eludens* (Dinophyceae). European Journal of Phycology 38: 385–394.
106. Hansen G, Daugbjerg N (2011) *Moestrupia oblonga* gen. & comb. nov (syn.: *Gyrodinium oblongum*), a new marine dinoflagellate genus characterized by light and electron microscopy, photosynthetic pigments and LSU rDNA sequence. Phycologia 50: 583–599.
107. Qiu DJ, Huang LM, Liu S, Lin SJ (2011) Nuclear, Mitochondrial and Plastid Gene Phylogenies of *Dinophysis miles* (Dinophyceae): Evidence of Variable Types of Chloroplasts. PlosOne 6: e29398.
108. Hackett JD, Wisecaver JH, Brosnahan ML, Kulis DM, Anderson DM, et al. (2012) Evolution of saxitoxin synthesis in cyanobacteria and dinoflagellates. Molecular Biology and Evolution.
109. Murray S, Hoppenrath M, Preisfeld A, Larsen J, Yoshimatsu S, et al. (2006) Phylogenetics of *Rhinodinium broomeense* gen. et sp nov., a peridinioid, sand-dwelling dinoflagellate (Dinophyceae). Journal of Phycology 42: 934–942.
110. Murray S, Patterson DJ (2004) *Cabra matta*, gen. nov., sp nov., a new benthic, heterotrophic dinoflagellate. European Journal of Phycology 39: 229–234.
111. Hoppenrath M, Selina M (2006) *Pseudothecadinium campbellii* gen. nov. et sp. nov. (Dinophyceae), a phototrophic, thecate, marine planktonic species found in the Sea of Okhotsk, Russia. Phycologia 45: 260–269.
112. Horiguchi T, Yoshizawa-Ebata J, Nakayama T (2000) *Halostylodinium arenarium*, gen et sp nov (dinophyceae), a coccoid sand-dwelling dinoflagellate from subtropical Japan. Journal of Phycology 36: 960–971.
113. Faust MA, Balech E (1993) A Further SEM Study of Marine Benthic Dinoflagellates from a Mangrove Island, Twin Cays, Belize, Including *Plagiodinium belizeanum* gen. et sp. nov. Journal of Phycology 29: 826–832.
114. Tamura M, Horiguchi T (2005) *Pileidinium ciceropse* gen. et sp nov (Dinophyceae), a sand-dwelling dinoflagellate from Palau. European Journal of Phycology 40: 281–291.
115. Lindberg K, Moestrup Ø, Daugbjerg N (2005) Studies on woloszynskioid dinoflagellates I: *Woloszynskia coronata* re-examined using light and electron microscopy and partial LSU rDNA sequences, with description of *Tovellia* gen. nov. and *Jadwigia* gen. nov. (Tovelliaceae fam. nov.). Phycologia 44: 416–440.