Scientific Reports

https://doi.org/10.1038/s41598-026-40315-x

Article in Press

Transcriptomic insights into polyketides and toxin biosynthesis genes in freshwater dinoflagellates

Received: 14 August 2025

Accepted: 12 February 2026

Published online: 17 February 2026

Cite this article as: Muhammad B.L., Bui Q.T.N., Kim H. *et al.* Transcriptomic insights into polyketides and toxin biosynthesis genes in freshwater dinoflagellates. *Sci Rep* (2026). https://doi.org/10.1038/s41598-026-40315-x

Buhari Lawan Muhammad, Quynh Thi Nhu Bui, Han-Sol Kim & Jang-Seu Ki

We are providing an unedited version of this manuscript to give early access to its findings. Before final publication, the manuscript will undergo further editing. Please note there may be errors present which affect the content, and all legal disclaimers apply.

If this paper is publishing under a Transparent Peer Review model then Peer Review reports will publish with the final article.

ARTICLE IN PRESS

© The Author(s) 2026. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

> Research Article

Transcriptomic insights into polyketides and toxin biosynthesis genes in freshwater dinoflagellates

Buhari Lawan Muhammad[1,2], Quynh Thi Nhu Bui[1,3], Han-Sol Kim[1,2], and Jang-Seu Ki[1,2*]

[1]Department of Life Science, Sangmyung University, Seoul 03016, South Korea

[2]Institute of Natural Science, Sangmyung University, Seoul 03016, South Korea

[3]Faculty of Biology, Agriculture & Environmental Science, The University of Danang – University of Science and Education, Danang 550000, Vietnam

---

*Corresponding author

Tel: 82-2-2287-5449, Fax: 82-2-2287-0070

E-mail: kijs@smu.ac.kr (J.-S. Ki)

Abstract

Freshwater dinoflagellates are typically considered non-toxic, and their polyketides and toxin biosynthesis genes are largely unexplored. Here, we generated and analyzed the transcriptome of freshwater dinoflagellate *Palatinus apiculatus* and compared it with the transcriptome data from *Peridinium bipes* and *Ceratium furcoides* to investigate the presence and diversity of polyketide synthases (PKS), fatty acid synthases (FAS), and saxitoxin (STX) biosynthesis genes (*sxt*). We identified 95, 117, and 39 PKS-related transcripts in *P. apiculatus*, *P. bipes*, and *C. furcoides*, respectively, which include single-domain PKS, multi-domain PKS, and hybrid NRPS/PKS. Phylogenetic analysis revealed a novel clade of ketosynthase (KS) domains unique to freshwater dinoflagellates, suggesting species-specific diversification. Conserved catalytic residues were found in type II FAS genes across both freshwater and marine taxa. Although core STX biosynthesis genes were absent in all analyzed species, several STX-associated transcripts, including *sxtA4*, *sxtU, sxtS, sxtD, sxtH/T*, and *sxtI,* were identified. Phylogenetic analysis of the sxtA4 domain revealed that freshwater dinoflagellate sequences form a distinct clade from those of toxic marine dinoflagellates and cyanobacteria while retaining conserved active sites, suggesting potential functional variation. These findings reveal unique PKS and STX gene features in freshwater dinoflagellates, highlighting their previously unrecognized biosynthetic diversity, ecological roles, and biotechnological potential.

Keywords: Polyketide synthase, Fatty acids synthase (FAS), Saxitoxins (STXs), Dinoflagellate toxins, Freshwater ecosystems.

Introduction

Dinoflagellates are a diverse group of microeukaryotic protists, renowned for their ecological significance in aquatic environments[1]. They are widely distributed in marine and freshwater ecosystems, with approximately 2,000 marines and 350 freshwater dinoflagellates identified morphologically to date[2]. Dinoflagellates are notable for producing bioactive secondary metabolites, some of which exhibit therapeutic potential[3-5]. However, many of these compounds are potent toxins, such as saxitoxins, karlotoxins, amphidinolides, okadaic acid, and brevetoxins, posing risks to aquatic organisms and human health[6].

Among these secondary metabolites, polyketides represent a major and chemically diverse class of compounds that contribute to both toxic and non-toxic bioactivities in dinoflagellates. To date, at least 25 dinoflagellate species have been reported to produce polyketides, many of which function as marine biotoxins[6-8]. For example*, Prorocentrum lima* produces okadaic acid (OA), responsible for diarrhetic shellfish poisoning (DSP); *Karenia brevis* synthesizes brevetoxins (BTXs), causing neurotoxic shellfish poisoning (NSP); and *Gambierdiscus polynesiensis* produces ciguatoxins (CTXs) associated with ciguatera fish poisoning[9-12]. In addition, polyketides have also been commonly identified in saxitoxin (STX)-producing dinoflagellates, including *Alexandrium* spp., *Gymnodinium catenatum*, and *Pyrodinium bahamense,* which are responsible for paralytic shellfish poisoning (PSP)[13-15]. Most of these toxins are produced via polyketide synthesis pathways, with STXs being the most widely distributed, making them one of the most ecologically and economically significant marine biotoxins.

The biosynthesis of STXs involves several saxitoxin genes (*sxt)*, with the unique initial key gene *sxtA* having a polyketide synthase (PKS)-like structure[16,17]. This highlights the fundamental role of PKSs in toxin biosynthesis, as they facilitate the formation of complex chemical structures. However, PKS-derived compounds are not limited to potent toxins but also include a wide range of bioactive molecules with therapeutic applications, like antibiotics, insecticides, and immunosuppressive and anti-tumor agents[18-21]. Thus, PKSs have been detected in non-toxic species[14,22].

Most biotoxins share a polyketide backbone, which is synthesized through sequential condensations of carboxylic acids, a process catalyzed by polyketide synthases (PKSs)[6]. These enzymes share structural and functional similarities with fatty acid synthases (FASs) across various organisms, suggesting a common evolutionary origin[14,23]. Both PKSs and FASs rely on a core enzymatic framework comprising acyl transferase (AT), ketosynthase (KS), and acyl carrier protein (ACP) for acyl unit condensation, with additional modifying domains selectively present depending on the enzyme type[6]. PKSs are classified into three main types, while FASs are categorized into two groups based on their catalytic domain organization[24,25]. Among dinoflagellates, type I and II PKS/FAS systems have been identified, alongside hybrid non-ribosomal peptide synthase (NRPS)/PKS genes, highlighting the complexity of polyketide biosynthesis in these organisms[26].

Research on dinoflagellate PKS/FAS has gained attention only in the last decade, with most studies focusing on marine species due to their extreme toxicity and significant ecological impacts[3,10,12,27-30]. In contrast, freshwater dinoflagellates have been generally regarded as non-toxic and less ecologically concerning, leading to their relative neglect in PKS research. However, various freshwater microorganisms, including fungi, cyanobacteria, and bacteria, can produce polyketides with diverse ecological functions, including biotoxins such as STX, which is also synthesized by freshwater cyanobacteria[16,31-34]. This suggests that PKS and STX biosynthesis may be more widespread in freshwater environments than previously recognized. Notably, some freshwater dinoflagellates form blooms and may produce compounds linked to polyketide biosynthesis. For instance, *Naiadinium polonicum* has caused red tides in freshwater systems and was suspected in fish mortality events in Japan and Spain[35,36]. In addition, *Peridinium aciculiferum* has also been associated with harmful algal blooms in lakes[37], while *Peridinium bipes* exhibits algicidal activity against *Microcystis* *aeruginosa*[38]. A major bloom of *Palatinus apiculatus* was recently documented in Verbne Lake, Ukraine[39], and Paldang Reservoir in Korea[40], and *Ceratium furcoides* has emerged as an invasive species of concern due to its rapid spread and impact on water quality[41]. These cases suggest that freshwater dinoflagellate blooms may be more widespread than previously recognized, with potential involvement of PKS pathways, warranting further investigation.

The identification of PKS genes in dinoflagellates has been challenging due to their large genomes[42,43]. Advancements in RNA sequencing, however, have facilitated the detection of PKS/FAS genes in marine dinoflagellates, such as *Gambierdiscus* spp., *Azadinium spinosum*, *K. brevis*, *Symbiodinium* spp., and *P. lima*[12,14,44]. Transcriptomic surveys have also identified *sxt* genes in both toxic and non-toxic species, enhancing our understanding of toxin biosynthesis and its potential applications[15,45,46]. Despite these advances, PKS and STX biosynthesis in dinoflagellates remains poorly understood. It is likely that dinoflagellates produce many polyketide compounds that are yet undetected and uncharacterized[14]. While marine dinoflagellates have been extensively studied, few studies have been conducted on freshwater dinoflagellates. At present, only three freshwater dinoflagellates (*Peridinium bipes*, *Peridinium aciculiferum*, and *Ceratium furcoides*) have publicly available transcriptomes. Given the distinct environmental conditions of freshwater ecosystems, understanding the PKS, FAS, and STX-related genes in freshwater dinoflagellates is crucial for a comprehensive view of dinoflagellate evolution and the potential ecological roles of these organisms in freshwater ecosystems.

In the present study, we generated and analyzed the transcriptome of the freshwater dinoflagellate *Palatinus apiculatus* and compared it with available transcriptome data from *Peridinium bipes* and *Ceratium furcoides*. Our aim was to unravel the presence of PKS, FAS, and STX biosynthesis genes in the freshwater dinoflagellates. By identifying gene homologs, analyzing their structural features, and reconstructing their phylogeny, we tried to uncover their diversity and evolutionary trajectories. Finally, we discussed the potential ecological roles of PKS, FAS, and STX proteins in freshwater ecosystems, shedding light on their possible functions in adaptation, chemical interactions, and toxin biosynthesis.

Material and methods

Palatinus apiculatus strains and morphology observation

The freshwater dinoflagellate *P. apiculatus* (FD-02) was isolated at Paldang Reservoir in Korea on 12 December 2021. The samples collected were placed in an icebox and transported to a laboratory. The strain was cultured and maintained in URO medium, without silicates[47,48], at a temperature of 16°C, pH 7.5, and under a cool-white fluorescent illumination (65 µmol photons $m^{-2} \cdot s^{-1}$) on a 12 h:12 h light-dark cycle. The strain was identified as *P. apiculatus* using morphological and molecular methods based on[40].

RNA extraction, sequencing, and library construction

P. apiculatus cells were exposed to different temperature conditions, including normal (20°C), cold shock (4°C for 6 h and 16°C for 6 and 24 h), and thermal shock (24°C for 6 and 24 h and 37°C for 30 min). In addition, cells were exposed to diverse toxic contaminants ($CdSO_4$ 1.0 mg L^{-1}, $NaAsO_2$ 1.0 mg L^{-1}, and triclosan 1.0 mg L^{-1}), for 6 and 24 h and at different growth stages. Cells were harvested at stationary phase and centrifuged at 2000 × *g* for 5 min, and the pellets were mixed with RiboEX reagent (Invitrogen, Carlsbad, CA).

Total RNA was extracted using the GeneAll Hybrid-R RNA Kit (GeneAll Biotechnology Co., Seoul, Korea) according to the manufacturer’s instructions, and genomic DNA was removed with DNase I treatment (TaKaRa, Shiga, Japan). RNA integrity and concentration were further assesed using a Bioanalyzer RNA Pico 6000 chip (Agilent Technologies, Santa Clara, CA) to determine the RNA integrity number (RIN) and rRNA ratio. To generate a comprehensive transcriptome, RNA samples from all environmental conditions were pooled together for sequencing. cDNA libraries were prepared with Illumina TruSeq Stranded mRNA LT Sample Prep Kit (Illumina Inc., San Diego, CA) and sequencing was performed using 150 bp paired-end reads on the Illumina NovaSeq 6000 platform (Illumina Inc., San Diego, CA).

Transcriptome assembly and functional annotation

The quality of Illumina paired-end raw reads from the *P. apiculatus* (FD-02) transcriptome was assessed using FastQC v0.12.1[49]. To enhance data quality, raw sequences underwent adaptor removal and trimming of low-quality reads with Trimmomatic v0.38[50](Bolger et al., 2014). The cleaned reads were then subjected to *de novo* assembly using Trinity with default settings[51]. Assembly statistics were evaluated using the Trinity Stats module[51]. To generate a non-redundant set of transcripts (unigenes), the longest contigs from the assembled data were clustered using CD-HIT-EST[52]. Open reading frames (ORFs) within the unigenes were identified using TransDecoder[53].

For functional annotation, the Trinotate pipeline[51] was employed, integrating various functional annotation tools. ORF prediction was carried out with TransDecoder, while SignalP was used for signal peptide detection and TMHMM for transmembrane domain identification. Protein domain searches were conducted with HMMER against the Pfam database. Additionally, homology searches were performed using DIAMOND-based BlastP and BlastX against the NCBI non-redundant (NR) and UniProtKB/Swiss-Prot databases. Functional insights were further enriched by assigning Gene Ontology (GO) terms, mapping pathways to the Kyoto Encyclopedia of Genes and Genomes (KEGG)[54], and classifying orthologous groups through EggNOG. An E-value threshold of 1e-5 was applied for these analyses.

For other freshwater dinoflagellates (*Peridinium bipes* and *Ceratium furcoides*), we retrieved raw transcriptome datasets of *P. bipes* (DOP) and *C. furcoides* (P10K-1) from the NCBI Sequence Read Archive (SRA; SRR11148504) and National Genomics Data Center (https://ngdc.cncb.ac.cn/; SAMC2821193), respectively. Specifically, *Peridinium bipes* was isolated from a tributary of Three Gorges Reservoir, China and cultured under dissolved inorganic phosphorus. The *C. furcoides* dataset was generated by the Protist 10,000 Genomes (P10K) Project using single-cell whole-transcriptome sequencing of a field-isolated cell, for which no standardized culture conditions or growth-phase information are available. The raw transcriptome datasets of the two species were individually subjected to transcriptome assembly and functional annotation as described in *P. apiculatus* above.

Identification of PKS, FAS, and *sxt* genes

PKS genes were identified by the presence of KS domains due to the high sequence conservation within active sites. Therefore, the PKS candidate genes in the three freshwater dinoflagellates were identified by searching the transcriptome databases for characteristics of the KS domain and also by searching annotations for keywords containing conserved PKS domains. Free-standing PKS domains, including ketoacyl synthase (KS), ketoreductase (KR), phosphopantetheine (PP), dehydratase (DH), acyl transferase (AT), enoyl reductase (ER), and thioesterase (TE), were filtered out and classified as single-domain PKS. Additionally, multi-domain PKS proteins were identified and categorized as modular PKS. Furthermore, hybrid non-ribosomal peptide synthase (NRPS)/PKS genes were also identified, which were characterized by having additional adenylation (A) and condensation (C) domains. For type II fatty acid genes (FASII), identifications were performed by searching the gene ID names and annotations for keywords containing conserved FASII genes, for instance, 3-ketoacyl-ACP synthase I, II, and III enzymes for *FabB*, *FabF*, and *FabH*, respectively. 3-ketoacyl-ACP reductase for *FabG*, enoyl-ACP reductase I for *FabI*, malonyl-CoA: ACP transacylase for *FabD*, and 3-hydroxydecanoyl-ACP dehydratase/isomerase for *FabZ*. For *sxt* genes, we used functional keywords and Pfam IDs associated with each *sxt* gene. For example, using PF02274 (amidinotransferase domain) for *sxtG*, using PF08211 (cytidine deaminase domain) for *sxtB*, and using PF04116 (sterol desaturase domain) for *sxtD* among others.

In addition, local tblastn was performed using known sequences from other dinoflagellates and cyanobacteria (in the case of *sxt* genes) against our transcriptome dataset to identify unannotated conserved domains of PKS, FAS, and *sxt* genes, applying an E-value threshold of 1e-5. All the identified candidate transcripts were validated by functional domain analysis using the Conserved Domain Database (CDD; https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi) in NCBI and the InterPro online tool (http://www.ebi.ac.uk/interpro/), which also facilitated the identification of active and binding sites. Furthermore, the full-length open reading frame (ORF) of the candidate genes was determined using dinoflagellate spliced leader (dinoSL) sequence (ranging from 40 to 160 bp upstream of the start codon) and poly(A) tails. The active protein domains were aligned with Clustal X 1.83[55], and the conserved amino acids were identified via a sequence logo tool (http://weblogo.berkeley.edu/logo.cgi).

Phylogenetic analysis

To confirm the taxonomic placement of the freshwater dinoflagellate species analyzed in this study, a ribosomal DNA (rDNA)–based phylogenetic analysis was conducted. 28S rDNA sequences for *Peridinium bipes* and *Ceratium* *furcoides* were retrieved from the NCBI GenBank database, while the *Palatinus apiculatus* rDNA sequence was obtained from the previous study[40]. Additional representative dinoflagellate rDNA sequences covering major orders (Suessiales, Gymnodiniales, Peridiniales, Prorocentrales, and Gonyaulacales) were included for comparative context. *Plasmodium falciparum* was used as an outgroup to root the tree. Sequences were aligned using MAFFT with default parameters, and poorly aligned or ambiguous regions were trimmed prior to phylogenetic reconstruction. Maximum likelihood (ML) analysis was performed using MEGA, applying the Tamura-Nei 93 with Gamma distribution (TN93+G) model determined by model selection. Branch support was assessed using 1,000 bootstrap replicates, and bootstrap values ≥50% are shown on the tree. The resulting phylogeny was visualized and annotated to highlight taxonomic groupings and the positions of the freshwater species examined in this study.

The phylogenetic relationships of predicated KS domains within PKS, type II FAS, and *sxt* sequences were inferred. The protein sequences of the identified candidate genes from *P. apiculatus*, *P. bipes*, and *C. furkoides* were used to construct phylogenetic trees through comparison with sequences from dinoflagellates, cyanobacteria, bacteria, haptophytes, chlorophytes, fungi, metazoans, and apicomplexans available in GenBank and previous studies. Protein sequences were aligned using MAFFT, trimmed manually[56], and phylogenetic trees were constructed using the maximum-likelihood (ML) method. The ML analysis employed the GTR model selected by MEGA and included 1,000 bootstrap replications to ensure robustness. For each type II FAS and *sxt* genes, one representative sequence with the highest identity to a known homolog from each species was selected and shown in the final tree to ensure clarity and ease of interpretation. However, for the KS domain within the PKS, all the identified transcripts were used in the phylogenetic analysis. The aligned amino acid sequences were further refined in BioEdit[57], and sequence motifs were analyzed using the WebLogo tool[58]. Phylogenetic trees were visualized using MEGA X Tree Explorer and finalized in Adobe Illustrator CS6 (Adobe Systems, San Jose, CA).

Results

Transcriptome sequence libraries of *P. apiculatus, P. bipes*, and *C. furcoides*

RNA sequencing generated 31 million raw reads from *Palatinus apiculatus* (Peridiniales, Peridiniaceae; this study). In addition, we obtained 27 and 16 million raw reads from *Peridinium bipes* (Peridiniales, Peridiniaceae; SRR11148504) and *Ceratium furcoides* (Gonyaulacales, Ceratiaceae; SAMC2821193) RNA libraries in public databases. After trimming and assembling the raw reads, a total of 111,352, 179,865, and 82,155 transcripts were obtained from *P. apiculatus*, *P. bipes*, and *C. furcoides*, respectively (**Table 1**). In addition, the number of unigenes in each species corresponds to 75,385, 107,791, and 68,218, respectively. To obtain comprehensive gene function annotations, the unigenes of the three species were separately blasted to NR, Swiss-Prot, Pfam, and GO. Functional annotation across these databases showed varying levels of gene identification. The NR database identified the highest number of matches for *P. bipes* (93,901), followed by *P. apiculatus* (90,922) and *C. furcoides* (35,618). Annotation in the Swiss-Prot, Pfam, and Gene Ontology (GO) databases also revealed higher gene identification for *P. apiculatus* and *P. bipes* compared to *C. furcoides*, which had notably fewer annotated genes across all databases.

Ribosomal DNA (rDNA)-based phylogenetic relationships of freshwater and marine dinoflagellates

The 28S rDNA-based phylogenetic analysis resolved the freshwater dinoflagellates examined in this study within well-supported and taxonomically consistent clades (**Supplementary Fig. S1**). *Palatinus apiculatus* and *Peridinium bipes* clustered within Peridiniales, forming a distinct freshwater lineage with other *Peridinium* species and receiving strong bootstrap support (≥100), clearly separated from marine taxa belonging to Gonyaulacales, Gymnodiniales, Prorocentrales, and Suessiales. In contrast, *Ceratium furcoides* grouped within Gonyaulacales alongside *Alexandrium* species, including known saxitoxin producers, but remained distinct from the freshwater Peridiniales clade. This phylogenetic framework provides the evolutionary context for subsequent comparisons of PKS, FAS, and *sxt* gene diversity among freshwater and marine dinoflagellates.

Modular PKSs and single-domain PKSs

The numbers of transcripts identified as modular PKS were 12 in *P. apiculatus*, 8 in *P. bipes*, and 4 in *C. furcoides*. These modular PKSs encoded one to three modules in at least one contig, with both *cis*- and *trans*-AT architectures present. In the modular PKS chain, all PKS domains were found in at least one contig of *P. apiculatus* and *P. bipes*; however, *C. furcoides* contained relatively few domains. No single complete modules were found in *C. furcoides*. In addition, transcripts identified as the hybrid NRPS-PKS proteins were found in all three dinoflagellates: *P. apiculatus*, *P. bipes*, and *C. furcoides* (**Table 2**). The sequence characteristics and domain structures of modular PKS and NRPS/PKS hybrids found in the three species are listed in **Supplementary Table S1**.

To identify free-standing PKS domains (singular PKS), transcriptomic libraries of the three species were analyzed using conserved domain searches to detect single PKS types, including KS, KR, ACP, AT, DH, ER, and TE domains. Several PKS domains detected in the modular type were also found in the single type (**Table 2**). Among these, the KS domain was the most abundant, with the highest number of contigs identified across all three species (30 in *P. apiculatus*, 58 in *P. bipes*, and 14 in *C. furcoides*). The numbers of contigs identified for the KR domain were 23 in *P. apiculatus*, 28 in *P. bipes*, and 2 in *C. furcoides*, while the ACP domain had 21, 9, and 10 contigs, respectively. The AT domain had the lowest number of contigs, with 4 in *P. apiculatus*, 2 in *P. bipes*, and 3 in *C. furcoides*. On the other hand, DH, ER, and TE domains could not be confidently identified in any of the species, suggesting either their low abundance or absence in the transcriptomes analyzed. Details about the singular PKS genes are provided in **Supplementary Table S2-4**.

KS domain of freshwater and marine dinoflagellates with other organisms

The diversity and evolutionary relationships of freshwater KS domains were explored by constructing phylogenetic trees using singular KS domains and KS domains extracted from modular PKS sequences. We used 145 sequences, including sequences from the freshwater dinoflagellates determined in this study and the sequences from other dinoflagellates and diverse organisms from previous studies (**Supplementary Table S5**). The present ML phylogenetic tree revealed that dinoflagellate KS sequences formed multiple distinct clades, clearly separating from those of bacteria, fungi, haptophytes, apicomplexans, and chlorophytes (**Fig. 1**). Within dinoflagellates, the single and modular KS domains were further resolved into separate clusters. The single and modular KS domains were grouped into five clades each. The single KS clades A and B have all the active sites present in the KS domain (highlighted in orange in **Fig. 1**); for example, the conserved active site cysteine (C) required for decarboxylative condensation, the conserved histidine (H) within the conserved HGTGT sequence required for transacylation, and the second H required for KS activity. However, in the single KS clade C, the KS domains in all the species in this clade do not have the active sites or the conserved residues (DTACSS, HGTGT, and N (I, H) G (S, A) H). In the single KS clade D and E, there is no active site cysteine required for decarboxylative condensation, and rather the active site is replaced by the consensus presented as DT (E/D) X (A/G) (A/S) S(A). The conserved H in the expected HGTGT is present in single KS clade E and some sequences in this clade D, with a consensus sequence of H (D/N) G (G/A) X. The second conserved H is replaced in most sequences with N. Furthermore, the modular KS clades A, B, and C contained all the active sites and the conserved motif residues. However, in the modular KS clade D, the conserved Hs (underlined) in the expected HGTGT and NIGH are replaced with Q and G, with consensus sequences of QCNGS and KGHG**,** respectively. Finally, in the modular clade E, the sequences were irregular, and no conserved sequences were found between the species; however, all KS sequences have a variable active site.

From the phylogenetic analysis of the KS domains, we found that most sequences identified as single KS domains in the freshwater species were grouped within single KS clade A (**Fig. 2**). However, one *C. furcoides* sequence clustered in single KS clade B alongside marine dinoflagellates, such as *Azadinium spinosum, Alexandrium pacificum, Alexandrium ostenfeldii*, *Heterocapsa triquetra, Gambierdiscus excentricus,* and *Karenia brevis*. In addition, single KS clade C contained *P. apiculatus* and *C. furcoides* sequences clustering with *A. pacificum*, while single KS clade D included only *P. bipes* sequences alongside *A. pacificum*, *Symbiodinium* spp., and *G. polynesiensis*. Similarly, *P. apiculatus* clustered separately in single KS clade E with *G. polynesiensis*, *K. brevis*, *Cocklodinium polykrikoides*, *Polarella glacialis*, *Durusdinium trenchii*, and *S. natans*. For modular PKS, all the freshwater dinoflagellates were clustered together with other marine dinoflagellates in modular clades B, C, and D. However, clade A was only comprised of *P. bipes*.

Modular PKS architecture

Analysis of the modular PKS architecture revealed that clade A of the modular PKS comprises exclusively *P. bipes* sequences and is characterized by a cis-AT architecture, with a unique fatty acid desaturase (FADS) and p-hydroxybenzoic acid (pHBA) domain embedded within the PKS modules (**Fig. 3A**). Modular Clade B includes both freshwater (colored green) and marine dinoflagellate species and consists of both *cis*- and *trans*-AT architectures (**Fig. 3B**). Notably, all three freshwater dinoflagellates in this clade exhibited a *trans*-AT architecture. This clade contains the highest number of modules, with *P. apiculatus* and *P. bipes* each possessing at least three modules. The largest PKS modular systems were observed in *G. polynesiensis* and *P. glacialis*, with 7 modules each. In modular KS Clade C, all species exhibit a *trans*-AT architecture, except for *P. apiculatus*, which is the only species with *cis*-AT with a domain organization pattern of AT-DH-KR-PP-KS-KR. Clades D and E display unique modular arrangements (**Fig. 3C**). Clade D includes *G. polynesiensis* and *G. pacificus*, featuring a KS-PP domain organization. Clade E contains *P. bipes* and *G. pacificus* and is distinguished by *trans*-AT modules (**Fig. 3D**).

Type II fatty acid synthases (FAS) genes in *P. apiculatus, P. bipes*, and *C. furcoides*

We identified several key genes involved in the type II fatty acid synthase (FAS) pathway, which plays essential roles in fatty acid biosynthesis. The 3-ketoacyl-ACP synthase I, II, and III enzymes (*FabB, FabF*, and *FabH*, respectively), which catalyze elongation steps, showed varying copy numbers of candidate genes across species (**Table 3**). *P. apiculatus,* for example, has 1 contig for *FabF* and 20 for *FabB*, while *P. bipes* has 25 *FabF* and 8 *FabB* contigs. *C. furcoides* has the lowest number of contigs, with zero *FabB* and 5 *FabF* contigs. The 3-ketoacyl-ACP reductase (*FabG*) gene, responsible for reducing ketoacyl intermediates, is found in P. apiculatus (6 contigs) and *P. bipes* (8 contigs) but also absent in *C. furcoides*. Enoyl-ACP reductase I (*FabI*), which is crucial for double bond reduction, is present in all three species, with *P. apiculatus*, *P. bipes*, and C. furcoides containing 1, 2, and 1 contigs, respectively. The malonyl-CoA:ACP transacylase (*FabD*), essential for initiating fatty acid elongation, is found in *P. apiculatus* (4 copies), *P. bipes* (3 copies), and C. furcoides (4 copies). The 3-hydroxydecanoyl-ACP dehydratase/isomerase (*FabZ*), which is involved in dehydration and isomerization, is present in all three species, with *P. bipes* having the highest number (11 contigs), followed by *C. furcoides* (7) and *P. apiculatus* (3). *FabK* and *FabL*, alternative enoyl-ACP reductases, are absent in all three species. All the FASII candidate genes identified in this study are presented in **Supplementary Table S6-8**.

In addition, our phylogenetic results and examination of amino acid residues revealed that the dinoflagellate type II FAS enzymes are distinct from those of other organisms, including diatoms, cyanobacteria, haptophytes, and plants (**Fig. 4**). Within dinoflagellates, the freshwater dinoflagellates *P. apiculatus* and *P. bipes* were clustered together within the dinoflagellate clade; however, they formed a distinct lineage separate from marine species across all the FAS enzymes. In particular, despite this differentiation, the active sites of these enzymes were highly conserved not only among freshwater and marine dinoflagellates but also across diverse taxa, such as diatoms, apicomplexans, chlorophytes, and plants, indicating strong evolutionary conservation of functional residues. All type II FAS sequences of *C. furcoides* were partial and, therefore, excluded from the analysis. All the sequences used for the phylogenetic analysis of the type II FAS are presented in **Supplementary Table S9**.

Saxitoxin (*sxt*) gene homologs in *P. apiculatus, P. bipes*, and *C. furcoides*

Putative *sxt* candidate genes were searched and obtained from the *P. apiculatus, P. bipes, and C. furcoides* transcriptome library (**Supplementary Table S10-12**). Here, we identified 266, 185, and 110 putative homologs of 16 dinoflagellates and cyanobacterial STX genes in *P. apiculatus*, *P. bipes*, and *C. furcoides*, respectively (**Table 4**). Based on the STX biosynthetic pathway (**Fig. 5A**) and in comparison with toxic cyanobacteria (**Fig. 5B**) and toxic (red) and non-toxic (blue) marine dinoflagellates (**Fig. 5C**), fewer *sxt* genes were found in the freshwater dinoflagellates among the eight *sxt* core genes (*sxtA, sxtB, sxtD*, *sxtG, sxtH/T, sxtI, sxtS*, and *sxtU*) (**Fig. 5D**). Only *sxtI, sxtS*, and *sxtH/T* were found in all three freshwater species. *sxtA4* domain and *sxtD* were only found in *P. apiculatus* and *P. bipes*. In addition, *sxtU* was found only in *P. apiculatus*. Moreover, *sxtA1, sxtA2, sxtA3*, *sxtB,* and *sxtG* were not found in any of the freshwater species. Other *sxt* genes proposed to be involved in STX synthesis, tailoring (*sxtC, sxtL, sxtO*, and *sxtX*), regulatory (*sxtZ*), transport (*sxtF/M* and *sxtP*), and others (*sxtV* and *sxtW*) were found in one or more species. *sxtN* and *sxtQ* were not found in any of the species (**Table 4**). Detailed information on all the *sxt* homologs identified in the three species is provided in **Supplementary Table S10-12**.

In addition, phylogenetic relationships of the sxtA4 domain of the freshwater dinoflagellates were examined alongside sxtA4 of toxic dinoflagellates and toxic and non-toxic cyanobacteria (**Fig. 5E**). The maximum-likelihood (ML) phylogenetic tree revealed four well-supported clades: toxic dinoflagellates, toxic cyanobacteria, freshwater dinoflagellates (non-toxic), and non-toxic cyanobacteria. The toxic dinoflagellate clade comprising *A. fundyense, A. catenella, A. affine,* and *G. catenatum* clustered closely together with the toxic cyanobacteria *H. crispum*, *D. circinale*, and *C. raciborskii*. However, the freshwater dinoflagellates analyzed here were obviously differentiated from the toxic dinoflagellates and cyanobacteria. In addition, the non-toxic cyanobacteria were also differentiated from the three clades. Moreover, the protein alignment of the sxtA4 revealed significant variability between freshwater dinoflagellates and other organisms; however, the active sites (black box) and 3 binding sites (red box) were conserved between the freshwater dinoflagellates and the toxic species. However, one binding site residue varied among the species. The protein alignment of the sxtA4 domain are provided in **Supplementary Fig S2.**

Discussion

Dinoflagellates are commonly found in marine and freshwater environments and are capable of producing bioactive compounds, including some biotoxins by marine dinoflagellates[1,2]. However, unlike marine species, the genetic characteristics of bioactive compounds in freshwater dinoflagellates remain poorly understood due to the limited number of known freshwater species and challenges in culturing them under laboratory conditions. As a result, key biosynthetic pathways and evolutionary traits in freshwater dinoflagellates remain largely unexplored. In the present study, we successfully cultured the freshwater dinoflagellate *Palatinus apiculatus*, determined its transcriptome sequences, and compared them with available data from previously reported freshwater and marine dinoflagellates.

Widespread presence of PKS genes in freshwater dinoflagellates

In the present study, we identified a large number and diversity of genes putatively involved in polyketide biosynthesis, including single-domain PKS (KS, ACP, AT, KR), multi-domain PKS, and hybrid NRPS/PKS in *P. apiculatus*, *P. bipes*, and *C. furcoides*. Conserved domain analysis and phylogenetic reconstruction consistently supported the dinoflagellate origin of these PKS genes. Most domains showed the highest BLASTx similarity to sequences from other protists, including dinoflagellates, further confirming their taxonomic identity. Although the cultures were non-axenic and associated bacterial transcripts cannot be completely excluded, multiple lines of evidence, including dinoflagellate-specific transcript features (e.g., spliced leader sequences) and phylogenetic clustering of conserved domains within dinoflagellate rather than bacterial clades, support a predominantly eukaryotic origin of the PKS sequences analyzed. These findings align with previous results that reported a diversity of polyketide synthase genes in several marine dinoflagellates, including both toxin- and non-toxin-producing species[3,59]. While a variable number of PKS-related genes and domains were recovered among the species, with *C. furcoides* showing fewer detected sequences, the transcriptomes analyzed here were generated using different experimental designs, sequencing depths, and assembly strategies. Accordingly, comparisons are focused on qualitative patterns of gene presence and domain diversity rather than quantitative differences among species. The present study reports the presence and diversity of PKS genes in freshwater dinoflagellates, indicating their potential to synthesize multiple polyketide compounds, some of which may possess unknown bioactivities. The detection of PKS genes in these freshwater taxa highlights the biochemical potential of underexplored dinoflagellate lineages and provides a foundation for future studies on their ecological and pharmacological relevance.

A novel freshwater-specific KS clade

Most studies investigating PKS biosynthesis genes have targeted the KS domains, as they are the most conserved and essential among the PKS genes. Commonly, KS domains have very conserved active site residues, Cysteine-Histidine-Histidine (C–H-H), important for their functionality[3,14,12,59]. Thus, KS domains are particularly valuable for identifying divergent homologs, making them key to understanding the evolutionary history of PKS genes in dinoflagellates and protists in general[14,60]. In the present study, our comprehensive phylogenetic analysis revealed the KS domains from both freshwater and marine dinoflagellates, alongside other organisms, showing multiple divergent homologs distributed across distinct phylogenetic clades, including a novel freshwater-specific clade. Previous researchers have identified three major single KS domain clades in dinoflagellates, two of which contained conserved C-H-H active sites, while the third exhibited altered active sites[12,14,29,28]. However, in all phylogenetic analyses, there was no clear clustering based on species lineage or the types of compounds they produced, suggesting these genes may be involved in synthesizing common polyketides produced by all species[12,14,30,61].

Interestingly, this study identified five distinct single KS domain clades, including previously reported clades and one additional novel clade that was exclusive to the freshwater dinoflagellates *P. apiculatus, P. bipes*, and *C. furcoides*. The sequences within this freshwater-specific clade retained the conserved C-H-H active site motifs, suggesting functional conservation. Additionally, some sequences from freshwater dinoflagellates clustered alongside marine dinoflagellates in other clades. This highlights the evolutionary diversity of KS domains in freshwater dinoflagellates, and the diversity may be driven by the evolution of PKS genes under selective pressures, as proposed by[62]. For instance, multiple gene duplication events and domain shuffling have been reported in KS domains, with the gain or loss of domains likely shaped by their functional relevance to the organism[63]. Consistent with this, the rDNA-based phylogeny shows that the freshwater dinoflagellates nested within marine species but form a well-supported distinct lineage supporting the interpretation that differences in PKS gene features may reflect ecological adaptation to freshwater environments rather than broad phylogenetic divergence. To our knowledge, this is the first report of a potential species-specific clade within the dinoflagellate KS domain that is exclusive to freshwater dinoflagellates. This novel clade may indicate a divergence in polyketide biosynthetic pathways or the production of unique

PKS-derived compounds, potentially shaped by ecological pressures distinct from those in marine environments[62].

Modular PKS Architecture in freshwater dinoflagellates

With advancements in sequencing technology, modular PKSs (multi-domain) have been identified in various dinoflagellates[3,12,38,59,64]. These modular PKSs are classified into two types: *cis*-AT PKS and *trans*-AT PKS, distinguished by the presence or absence of an acyltransferase (AT) domain within the multi-domain structure, respectively. In the present study, both *cis*-AT and *trans*-AT PKS transcripts were found in the transcriptome of freshwater dinoflagellates, with most sequences having *trans*-AT PKS. The majority of modular PKS identified are *trans*-AT[12,65], which was postulated that during the evolutionary history of PKS in algae, integrated AT domains were replaced to yield *trans*-AT PKS[65]. The phylogenetic tree showed the multi-domain PKSs also constituted various subclades that consisted of both freshwater and marine dinoflagellates. However, a novel clade with 'unusual' PKS domains, which consisted of only *P. bipes,* was found. Similar 'unusual' PKS domains have previously been reported in marine dinoflagellates[10,29], which highlights the complexity of dinoflagellate-derived polyketides and the involvement of domains with variable effects on the final product. In modular PKS systems, the number of modules typically correlates with the structural complexity and diversity of the synthesized metabolites[60]. The largest modular PKS contig in this study was found in *P. bipes* (DN5219_c0_g3_i2) with a 13,796 nt long transcript encoding for a 3-module PKS. *P. apiculatus* also has three modules with shorter contigs, while *C. furkoides* has only one module. These modular PKS sequences are relatively shorter compared to those reported in marine dinoflagellates known for producing diverse bioactive compounds, such as *Gambierdiscus* spp., which can encode up to seven PKS modules[10]. Therefore, the presence of fewer modules in freshwater dinoflagellates may indicate a reduced complexity in their polyketide biosynthetic pathways and could explain the non-toxic nature of freshwater species. On the other hand, PKS loci are often very large (>10 kb) and highly modular, whereas transcriptome assemblies are constrained by read length, sequencing depth, and assembly strategy. Consequently, long PKS transcripts may be fragmented or incompletely recovered, and the observed module architectures in the present study likely represent detectable transcript fragments rather than complete loci. These technical constraints are intrinsic to short-read transcriptomic approaches and underscore the need for future studies employing deeper sequencing, standardized experimental designs, long-read transcriptomics, and genome-resolved analyses to fully reconstruct PKS loci and validate inferred biosynthetic potential.

Conserved FAS genes across dinoflagellates

In this study, we identified several type II FAS genes, with the freshwater dinoflagellates clustered together in the phylogeny within marine species. Moreover, the active site residues were highly conserved in the freshwater dinoflagellates and other dinoflagellates and protists, suggesting that the type II FAS system genes are fully functional in dinoflagellates. No significant divergence was observed between freshwater and marine dinoflagellates. Previous studies have reported similar conservation[14,15] and suggest that fatty acid synthesis is crucial for survival, leading to the retention of these genes in the nucleus of protists under strong selective pressure[14,59]. Previous transcriptomic studies in *Karenia brevis* reported the apparent absence of Type II FAS, suggesting the presence of a Type I FAS system or the potential involvement of PKS-like domains in fatty acid biosynthesis[66]. However,[14] later demonstrated a clear evolutionary distinction between Type II FAS and PKS genes, reinforcing that fatty acid biosynthesis in dinoflagellates is likely driven by Type II FAS enzymes. While fatty acids and PKS share a similar biosynthetic mechanism, PKS systems are more diverse, modular, and capable of producing structurally complex molecules, while FAS systems generate linear fatty acids with a relatively conserved structure. Amongst dinoflagellates, their evolution broadly follows a similar trend as reported by ribosomal gene phylogenies[14], highlighting the conserved nature of FAS domains amongst the species.

Saxitoxin (STX) biosynthesis genes in the freshwater dinoflagellates

Transcriptomic surveys in both toxic and non-toxic marine dinoflagellates have shown that the absence of key saxitoxin (STX) biosynthesis genes, such as the fourth domain of *sxtA* (sxtA4) and *sxtG*, often correlates with a loss of toxicity[15,45,46,67]. Here, we investigated non-STX-producing freshwater dinoflagellates and detected several *sxt* genes, although key *sxt* genes essential for complete toxin production, namely *sxtG* and *sxtB*, were absent. However, we identified the sxtA4 domain in *P. apiculatus* and *P. bipes*, with conserved catalytic and binding sites despite clear divergence from toxic marine dinoflagellates and cyanobacteria, suggesting functional roles beyond STX biosynthesis. While *sxtA4* has been consistently linked to STX production and is absent in confirmed non-toxic species like *Amphidinium carterae* and *Prorocentrum micans*[15], its presence in freshwater species raises intriguing questions about its evolutionary origin and alternative metabolic functions. The origin of these genes in marine dinoflagellates remains debated; some propose horizontal gene transfer (HGT) from cyanobacteria[68], while others suggest an independent acquisition followed by gene loss in non-toxic lineages[63,69,70]. Nevertheless, our identification of *sxt* genes in freshwater dinoflagellates is critical for understanding the evolution and distribution of toxin biosynthesis pathways beyond well-characterized marine species. The presence of *sxt* genes, even in non-toxic freshwater species, particularly conserved domains such as *sxtA4*, underscores the need for further investigation integrating standardized culture conditions and direct toxin measurements to clarify the functional, ecological, and evolutionary significance of *sxt* genes in freshwater dinoflagellates. These genes may reveal insights into biosynthetic capabilities and adaptive strategies in this diverse and understudied group of organisms.

Potential environmental implications of PKS, FAS and *sxt* genes in the freshwater dinoflagellates

In the present study, we found the widespread presence of PKS, FAS, and *sxt* genes in *P. apiculatus, P. bipes,* and *C. furcoides*. The presence of PKS, FAS, and *sxt* genes in freshwater dinoflagellates may have significant ecological implications for freshwater ecosystems (**Fig. 6**). For example, polyketides synthesized by PKS enzymes are structurally diverse metabolites that play essential roles in chemical defense, interspecies interactions, and adaptation to environmental stressors[6,34,71]. Our results show that the freshwater dinoflagellates have possessed diverse KS domains, most of which are distinct from marine dinoflagellates and other organisms. This finding is particularly intriguing, because substantial PKS gene diversity has been reported in marine dinoflagellates[14]; however, no clear phylogenetic patterns have been established. This suggests that the freshwater dinoflagellates might have a divergent PKS and have evolved unique biosynthetic pathways, possibly to cope with environmental pressures specific to freshwater ecosystems[62]. Indeed, freshwater environments differ significantly from marine systems in terms of ionic composition, nutrient availability, and biological interactions[72], which may have driven the divergence of PKS gene structures and functions in freshwater dinoflagellates[62]. However, further research is needed to include a broader range of species to better understand this distinction and their environmental implications.

From a broader environmental perspective, PKS enzymes are central to the biosynthesis of marine biotoxins such as brevetoxins, ciguatoxins, yessotoxins, and saxitoxins, all of which contribute to harmful algal blooms (HABs)[6,8]. Most of these toxins typically possess polyether backbones synthesized through complex modular PKS pathways[12]. Here, we found that *P. bipes* exhibited the longest PKS contigs (3 modules) among the freshwater species, which were still shorter than those typically found in highly toxic marine dinoflagellates. For instance, toxic strains like *Gambierdiscus* and *Prymnesium* harbor longer modular PKS transcripts (7 and 11 modules, respectively) than non-toxic strains[10,11]. Thus, the shorter PKS modules in freshwater species may indicate simpler biosynthetic systems, possibly due to the absence or loss of key modular PKS genes involved in polyether synthesis. This loss could stem from relaxed selective pressure in freshwater environments, where low ionic strength might destabilize polyether toxins and render their production less beneficial[12,62]. Therefore, freshwater dinoflagellates may not have complex PKS systems for synthesizing complex toxins but may still use polyketides for different ecological and physiological functions such as antimicrobial activity, structural and metabolic roles, and stress response and environmental adaptation[34,35,37]. Importantly, because these conclusions are drawn from transcriptome-based datasets generated under different experimental conditions, observed differences in PKS, FAS, and *sxt* gene architecture are interpreted qualitatively and do not constitute definitive evidence of biological differences among species. Accordingly, future studies should examine a broader range of freshwater dinoflagellate species cultured under standardized conditions to more rigorously assess gene presence, architecture, and expression patterns and to clarify their functional and ecological roles.

Overall, the presence of PKS, FAS, and *sxt* genes in freshwater dinoflagellates expands our understanding of PKS diversity and provide an initial transcriptome-based perspective in a poorly explored freshwater dinoflagellates. These suggest their potential to produce bioactive compounds that influence ecological interactions and may pose environmental and public health risks. In addition, freshwater dinoflagellates could represent an untapped source of novel polyketides with unique structures and biological activities, offering potential applications in biotechnology[73]. Indeed, various novel polyketides, such as Annularins A-H, Tanzawaic acid Z1, and Asperdichrome, have been identified in other freshwater organisms, including fungi, cyanobacteria, and bacteria, with promising pharmaceutical and industrial applications[32,34,74,75]. Therefore, there is a need for further research into the biochemical and ecological functions of PKS and *sxt* genes in non-marine species. Investigating these pathways could provide insights into their evolutionary significance and potential biotechnological applications.

5. Conclusions

In this study, we provide the first comprehensive investigation of PKS, FAS, and STX biosynthetic pathways in freshwater dinoflagellates using transcriptome data from three freshwater dinoflagellates; *P. apiculatus, P. bipes*, and *C. furcoides*. We identified 95, 117, and 39 PKS transcripts in *P. apiculatus, P. bipes*, and *C. furcoides*, respectively, encompassing single-domain PKS, multi-domain PKS, and hybrid NRPS/PKS. Phylogenetic analysis revealed a novel clade of ketosynthase (KS) domains exclusive to freshwater dinoflagellates, suggesting potential adaptations to their environment. In addition, type II FAS genes exhibited conserved residues and active sites across freshwater and marine dinoflagellates and other protists. Several STX-related genes, including the sxtA4 domain, were detected; however, key STX biosynthesis genes such as *sxtG* and *sxtB* were not observed, which may be associated with the apparent non-STX producing phenotype of these species. Overall, our findings demonstrate the widespread presence of PKS, FAS, and *sxt* genes in freshwater dinoflagellates, uncovering unique genetic diversity and structural variations that may reflect evolutionary adaptations to freshwater-specific environmental pressures.

Data Availability

The datasets generated and/or analyzed during the current study are available in the NCBI Sequence Read Archive (SRA) under Bio Project, PRJNA1307768.

References

1. Bi, R. et al. Responses of marine diatom–dinoflagellate competition to multiple environmental drivers: Abundance, elemental, and biochemical aspects. *Front. Microbiol.* **12**, 731786 (2021).
2. Taylor, F. J. R., Hoppenrath, M. & Saldarriaga, J. F. Dinoflagellate diversity and distribution. *Biodivers. Conserv.* **17**, 407–418 (2008).
3. Beedessee, G., Hisata, K., Roy, M. C., Van Dolah, F. M., Satoh, N. & Shoguchi, E. Diversified secondary metabolite biosynthesis gene repertoire revealed in symbiotic dinoflagellates. *Sci. Rep.* **9**, 1204 (2019).
4. Beedessee, G. et al. Integrated omics unveil the secondary metabolic landscape of a basal dinoflagellate. *BMC Biol.* **18**, 16 (2020).
5. Camacho-Muñoz, D., Praptiwi, R. A., Lawton, L. A. & Edwards, C. High value phycotoxins from the dinoflagellate *Prorocentrum*. *Front. Mar. Sci.* **8**, 638739 (2021).
6. Verma, A., Barua, A., Ruvindy, R., Savela, H., Ajani, P. A. & Murray, S. A. The genetic basis of toxin biosynthesis in dinoflagellates. Microorganisms 7, 222 (2019).
7. Rein, K. S. & Borrone, J. Polyketides from dinoflagellates: Origins, pharmacology and biosynthesis. Comp. Biochem. Physiol. B Biochem. *Mol. Biol.* **124**, 117–131 (1999).
8. Kellmann, R., Stüken, A., Orr, R. J., Svendsen, H. M. & Jakobsen, K. S. Biosynthesis and molecular genetics of polyketides in marine dinoflagellates. *Mar. Drugs* **8**, 1011–1048 (2010).
9. Pawlowiez, R., Morey, J. S., Darius, H. T., Chinain, M. & Van Dolah, F. M. Transcriptome sequencing reveals single domain Type I-like polyketide synthases in the toxic dinoflagellate Gambierdiscus polynesiensis. *Harmful Algae* **36**, 29–37 (2014).
10. Kohli, G. S. et al. Polyketide synthesis genes associated with toxin production in two species of Gambierdiscus (Dinophyceae). *BMC Genomics* **16**, 410 (2015).
11. Van Dolah, F. M. et al. Transcriptomic analysis of polyketide synthases in a highly ciguatoxic dinoflagellate, Gambierdiscus polynesiensis and low toxicity *Gambierdiscus pacificus*, from French Polynesia. *PLoS ONE* **15**, e0231400 (2020).
12. Wan, X. et al. Transcriptomic analysis of polyketide synthesis in dinoflagellate, Prorocentrum lima. *Harmful Algae* **123**, 102391 (2023).
13. Snyder, R. V. et al. Polyketide synthase genes from marine dinoflagellates. *Mar. Biotechnol.* **5**, 1–12 (2003).

14. Kohli, G. S., John, U., Van Dolah, F. M. & Murray, S. A. Evolutionary distinctiveness of fatty acid and polyketide synthesis in eukaryotes. *ISME J.* **10**, 1877–1890 (2016).

15. Wang, H., Kim, H. & Ki, J. S. Transcriptome survey and toxin measurements reveal evolutionary modification and loss of saxitoxin biosynthesis genes in the dinoflagellates *Amphidinium carterae* and *Prorocentrum micans*. Ecotoxicol. *Environ. Saf.* **195**, 110474 (2020).

16. Kellmann, R. et al. Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. *Appl. Environ. Microbiol.* **74**, 4044–4053 (2008).

17. Stüken, A., Orr, R. J., Kellmann, R., Murray, S. A., Neilan, B. A. & Jakobsen, K. S. Discovery of nuclear-encoded genes for the neurotoxin saxitoxin in dinoflagellates. *PLoS ONE* **6**, e20096 (2011).

18. Camacho, F. G. *et al.* Biotechnological significance of toxic marine dinoflagellates. *Biotechnol. Adv.* **25**, 176–194 (2007).

19. Assunção, J., Guedes, A. C. & Malcata, F. X. Biotechnological and pharmacological applications of biotoxins and other bioactive molecules from dinoflagellates. *Mar. Drugs* **15**, 393 (2017).

20. Morales-Amador, A., Souto, M. L., Hertweck, C., Fernández, J. J. & García-Altares, M. Rapid screening of polyol polyketides from marine dinoflagellates. *Anal. Chem.* **94**, 14205–14213 (2022).

21. Pradhan, B. & Ki, J. S. Phytoplankton toxins and their potential therapeutic applications: A journey toward the quest for potent pharmaceuticals. *Mar. Drugs* **20**, 271 (2022).

22. Beedessee, G., Hisata, K., Roy, M. C., Satoh, N. & Shoguchi, E. Multifunctional polyketide synthase genes identified by genomic survey of the symbiotic dinoflagellate, *Symbiodinium minutum*. *BMC Genomics* **16**, 941 (2015).

23. Jenke-Kodama, H., Sandmann, A., Müller, R. & Dittmann, E. Evolutionary implications of bacterial polyketide synthases. *Mol. Biol. Evol.* **22**, 2027–2039 (2005).

24. Shen, B. *et al.* Prerequisites for combinatorial biosynthesis: Evolution of hybrid NRPS/PKS gene clusters. *Biocombinatorial Approaches for Drug Finding*, 107–126 (2005).

25. Keatinge-Clay, A. T. The structures of type I polyketide synthases. *Nat. Prod. Rep.* **29**, 1050–1073 (2012).

26. Rasmussen, S. A. *et al.* Chemical diversity, origin, and analysis of phycotoxins. *J. Nat. Prod.* **79**, 662–673 (2016).

27. Meyer, J. M. *et al.* Transcriptomic characterisation and genomic glimpse into the toxigenic dinoflagellate *Azadinium spinosum*, with emphasis on polyketide synthase genes. *BMC Genomics* **16**, 27 (2015).

28. Kohli, G. S. *et al.* Role of modular polyketide synthases in the production of polyether ladder compounds in ciguatoxin-producing *Gambierdiscus polynesiensis* and *G. excentricus* (Dinophyceae). *J. Eukaryot. Microbiol.* **64**, 691–706 (2017).

29. Van Dolah, F. M., Kohli, G. S., Morey, J. S. & Murray, S. A. Both modular and single-domain Type I polyketide synthases are expressed in the brevetoxin-producing dinoflagellate, *Karenia brevis* (Dinophyceae). *J. Phycol.* **53**, 1325–1339 (2017).

30. Wang, H., Guo, R., Lim, W. A., Allen, A. E. & Ki, J. S. Comparative transcriptomics of toxin synthesis genes between the non-toxin producing dinoflagellate *Cochlodinium polykrikoides* and toxigenic *Alexandrium pacificum*. *Harmful Algae* **93**, 101777 (2020).

31. Paguigan, N. D., Raja, H. A., El-Elimat, T. & Oberlies, N. H. New polyketides from a freshwater *Lindgomycetaceae* sp. *Planta Med.* **80**, PC33 (2014).

32. Ebada, S. S. & Ebrahim, W. A new antioxidant decalin polyketide from freshwater-sediment-derived fungus *Penicillium* sp. strain S1a1. *Chem. Select* **4**, 9814–9816 (2019).

33. Kaluzhnaya, O. V. & Itskovich, V. B. Features of diversity of polyketide synthase genes in the community of freshwater sponge *Baikalospongia fungiformis*. *Russ. J. Genet.* **58**, 336–346 (2022).

34. Scesa, P. *et al.* Defensive polyketides produced by an abundant gastropod are candidate keystone molecules in estuarine ecology. *Sci. Adv.* **10**, eadp8643 (2024).

35. Roset, J. *et al.* Mortality of rainbow trout (*Oncorynchus mykiss* (Walbaum)) associated with freshwater dinoflagellate bloom (*Peridinium polonicum* (Woloszynska)) in a fish farm. *Aquac. Res.* **33**, 159–164 (2002).

36. Oshima, Y., Minami, H., Takano, Y. & Yasumoto, T. Ichthyotoxins in a freshwater dinoflagellate *Peridinium polonicum*. In *Red Tides: Biology, Environmental Science and Toxicology. Proceedings of the First International Symposium on Red Tides* 375–377 (Elsevier, 1989).

37. Rengefors, K. & Legrand, C. Toxicity in *Peridinium aciculiferum*—an adaptive strategy to outcompete other winter phytoplankton? *Limnol. Oceanogr.* **46**, 1990–1997 (2001).

38. Wu, J. T., Kuo-Huang, L. L. & Lee, J. Algicidal effect of *Peridinium bipes* on *Microcystis aeruginosa*. *Curr. Microbiol.* **37**, 257–261 (1998).

39. Krahmalniy, A. F. Mass development of *Palatinus apiculatus* (Dinoflagellata) in the Verbne Lake (Kyiv, Ukraine). *Hydrobiol. J.* **54**, 41–47 (2018).

40. Kim, T. & Ki, J. S. New record of the cold freshwater dinoflagellate *Palatinus apiculatus* (Dinophyceae) from the Paldang Reservoir, Korea. *J. Species Res.* **11**, 162–168 (2022).

41. Corrêa, R. F. *et al.* First report of the invasive *Ceratium furcoides* (dinoflagellate) in Paracambi Reservoir, Rio de Janeiro: Risks to the world's largest domestic water treatment plant. *Lakes Reserv. Res. Manag.* **27**, e12400 (2022).

42. Lin, S. A decade of dinoflagellate genomics illuminating an enigmatic eukaryote cell. *BMC Genomics* **25**, 932 (2024).

43. Wang, H., Wu, P., Xiong, L., Kim, H. S., Kim, J. H. & Ki, J. S. Nuclear genome of dinoflagellates: size variation and insights into evolutionary mechanisms. *Eur. J. Protistol.* **93**, 126061 (2024).

44. Van Dolah, F. M. *et al.* Transcriptomic analysis of polyketide synthases in a highly ciguatoxic dinoflagellate, *Gambierdiscus polynesiensis* and low toxicity *Gambierdiscus pacificus*, from French Polynesia. *PLoS ONE* **15**, e0231400 (2020).

45. Bui, Q. T. N., Pradhan, B., Kim, H. S. & Ki, J. S. Environmental factors modulate saxitoxins (STXs) production in toxic dinoflagellate *Alexandrium*: An updated review of STXs and synthesis gene aspects. *Toxins* **16**, 210 (2024).

46. Muhammad, B. L., Kim, H. S., Bui, Q. T. N. & Ki, J. S. Transcriptomic comparison unveils saxitoxin biosynthesis genes in the marine dinoflagellate *Gymnodinium catenatum*. *Harmful Algae* **137**, 102872 (2025).

47. Kimura, B. & Ishida, Y. Photophagotrophy in *Uroglena americana*, Chrysophyceae. *Jpn. J. Limnol.* **46**, 315–318 (1985).

48. Guillard, R. R. L. Culture of phytoplankton for feeding marine invertebrates. In *Culture of Marine Invertebrate Animals: Proceedings—1st Conference on Culture of Marine Invertebrate Animals, Greenport* 29–60 (Springer, 1975).

49. Andrews, S. FastQC: A quality control tool for high throughput sequence data. Available at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc (2010).

50. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).

51. Grabherr, M. G. *et al.* Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011).

52. Li, W. & Godzik, A. Cd-hit: A fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* **22**, 1658–1659 (2006).

53. Haas, B. J. *et al.* De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013).

54. Kanehisa, M. Toward pathway engineering: a new database of genetic and molecular pathways. *Science & Technology Japan*, No. 59, pp. 34-38 (1996).

55. Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.* **23**, 403–405 (1998).

56. Katoh, K. & Standley, D. M. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013).

57. Hall, T. A. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98 (1999).

58. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).

59. Verma, A., Kohli, G. S., Harwood, D. T., Ralph, P. J. & Murray, S. A. Transcriptomic investigation into polyketide toxin synthesis in *Ostreopsis* (Dinophyceae) species. *Environ. Microbiol.* **21**, 4196–4211 (2019).

60. John, U. *et al.* Novel insights into evolution of protistan polyketide synthases through phylogenomic analysis. *Protist* **159**, 21–30 (2008).

61. Kimura, K. *et al.* RNA sequencing revealed numerous polyketide synthase genes in the harmful dinoflagellate *Karenia mikimotoi*. *PLoS ONE* **10**, e0142731 (2015).

62. Nivina, A., Yuet, K. P., Hsu, J. & Khosla, C. Evolution and diversity of assembly-line polyketide synthases: focus review. *Chem. Rev.* **119**, 12524–12547 (2019).

63. Murray, S. A., Diwan, R., Orr, R. J., Kohli, G. S. & John, U. Gene duplication, loss and selection in the evolution of saxitoxin biosynthesis in alveolates. *Mol. Phylogenet. Evol.* **92**, 165–180 (2015).

64. Vingiani, G. M. *et al.* De novo transcriptome of the non-saxitoxin producing *Alexandrium tamutum* reveals new insights on harmful dinoflagellates. *Mar. Drugs* **18**, 386 (2020).

65. Shelest, E., Heimerl, N., Fichtner, M. & Sasso, S. Multimodular type I polyketide synthases in algae evolve by module duplications and displacement of AT domains in trans. *BMC Genomics* **16**, 1015 (2015).

66. Van Dolah, F. M. *et al.* Subcellular localization of dinoflagellate polyketide synthases and fatty acid synthase activity. *J. Phycol.* **49**, 1118–1127 (2013).

67. Cho, Y. *et al.* Intracellular abundance, localization, and enzymatic activity of a saxitoxin biosynthesis enzyme, SxtG, in two sister subclones of the dinoflagellate *Alexandrium catenella* with extremely different levels of paralytic shellfish toxins. *Harmful Algae* **139**, 102723 (2024).

68. Orr, R. J., Stüken, A., Murray, S. A. & Jakobsen, K. S. Evolution and distribution of saxitoxin biosynthesis in dinoflagellates. *Mar. Drugs* **11**, 2814–2828 (2013).

69. Hackett, J. D. *et al.* Evolution of saxitoxin synthesis in cyanobacteria and dinoflagellates. *Mol. Biol. Evol.* **30**, 70–78 (2013).

70. Akbar, M. A. *et al.* Biosynthesis of saxitoxin in marine dinoflagellates: An omics perspective. *Mar. Drugs* **18**, 103 (2020).

71. Van Donk, E. & Ianora, A. Induced defences in marine and freshwater phytoplankton: a review. *Hydrobiologia* **668**, 3–19 (2011).

72. Kilham, P. & Hecky, R. E. Comparative ecology of marine and freshwater phytoplankton. *Limnol. Oceanogr.* **33**, 776–795 (1988).

73. Hertweck, C. The biosynthetic logic of polyketide diversity. *Angew. Chem. Int. Ed.* **48**, 4688–4716 (2009).

74. Li, C., Nitka, M. V., Gloer, J. B., Campbell, J. & Shearer, C. A. Annularins A–H: new polyketide metabolites from the freshwater aquatic fungus *Annulatascus triseptatus*. *J. Nat. Prod.* **66**, 1302–1306 (2003).

75. Kaluzhnaya, O. V. & Itskovich, V. B. Features of diversity of polyketide synthase genes in the community of freshwater sponge *Baikalospongia fungiformis*. *Russ. J. Genet.* **58**, 336–346 (2022).

Acknowledgements

We thank Mr. T Kim for cell culture and valuable comments to the early version of our manuscript.

Author Contributions

B.L.M designed the research**,** analyzed the data, wrote the original draft, and reviewed and edited the paper. H.S.K performed the experiment**,** reviewed and edited the paper. Q.T.N.B**.** reviewed and edited the paper. J.S.K**.** designed the research, supervised the research, coordinated with co-authors, and provided extensive feedback on the text.

Funding

This work was supported by Korea Environment Industry & Technology Institute (KEITI) through Aquatic Ecosystem Conservation Research Program funded by Korea Ministry of Environment (MOE) (2022003050002 and RS-2022-KE002119211530) and by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (RS-2024-00354842).

Declarations

Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Additional information

Supplementary Information The online version contains supplementary material available at:

Figures legends

Fig. 1. Maximum likelihood phylogenetic analysis of KS domains extracted from singular and modular PKS including dinoflagellates, Bacteria, Fungi, Metazoan, Chlorophytes, Haptophyte, and Apicomplexans. The numbers above or below the branches are bootstrap values based on 1,000 bootstrap replications. Bootstrap values ≥50% were shown. Sequence logos of the active sites are shown for each major dinoflagellate clade. For modular KS clade E, the sequences were highly divergent in the catalytic regions, preventing robust alignment and consensus logo construction; these sequences were therefore classified based on overall KS domain homology rather than conserved active-site motifs.

Fig. 2. Detailed phylogenetic analysis of KS domains extracted from singular and modular PKS including sequences from the freshwater dinoflagellate analyzed this study (*Palatinus apiculatus*, *Peridinium bipes*, and *Ceratium furcoides*) and other sequences from marine dinoflagellates, and other prokaryotic and eukaryotic taxa. This circular phylogenetic analysis included 140 amino acid sequences (**Supplementary Table S3**). The inner ring color codes the taxonomic classification, including dinoflagellates, chlorophytes, bacteria, fungi, metazoa, haptophytes, and apicomplexans. The outer ring categorizes PKS sequences into different clades of modular and single KS domains. Black dots highlight freshwater-specific clades. The numbers above or below the branches are bootstrap values based on 1,000 bootstrap replications. Bootstrap values ≥50% were shown.

Fig. 3. Phylogenetic relationships and domain organization of modular PKS clades (a-e) identified in Fig. 1 and 2 across dinoflagellates and other eukaryotic and prokaryotic taxa. (a-d) The phylogenetic trees on the left illustrate the evolutionary relationships among taxa within each modular PKS

clade. On the right, the corresponding domain organizations of each taxon are displayed, showing the arrangement of catalytic modules. Species analyzed in this study are highlighted in green. The dashed lines in *G. polynesiensis* and *P. glacialis* indicate omitted PKS modules for clarity. The domain key (bottom right) provides abbreviations and functional roles for each module. This figure highlights the diversity and conservation of modular PKS architectures across different lineages.

Fig. 4. Phylogenetic relationships and conserved active sites in type II fatty acid synthase enzymes in *Palatinus apiculatus* (Pa), *Peridinium bipes* (Pb) analyzed in this study (in red), and marine dinoflagellates and other species. (a) 3-ketoacyl ACP reductase (FabG); (b) 3-ketoacyl ACP synthase II (FabF); (c) malonyl-CoA: ACP transacylase (FabD); (d) 3-ketoacyl ACP synthase III (FabH); (e) 3-ketoacyl ACP synthase I (FabB); (f) enoyl-ACP reductase I (FabI). Active site residues are highlighted in black boxes. Detail information of the sequences is provided in **Supplementary Table S5.**

Fig. 5. Structure of the saxitoxin (STX) biosynthesis pathway in (a) dinoflagellate (Verma et al., 2019a) and (b) cyanobacteria (Kellmann et al., 2008). (c) *sxt* gene cluster find in the toxic (green) and non-toxic (blue) *Alexandrium* species based on transcriptomic surveys (Bui et al., 2024). (d) *sxt* gene cluster identified in the present study in the three freshwater dinoflagellates (*Palatinus apiculatus*, *Peridinium bipes*, and *Ceratium furcoides*). (e) The ML phylogenetic tree and the conserved active and binding sites of *sxtA4* domain identified in *P. apiculatus* and *P. bipes*, comparing with toxic dinoflagellates and toxic and non-toxic cyanobacteria.

Fig. 6. Conceptual overview of polyketide synthase (PKS), fatty acid synthase (FAS), and saxitoxin (*sxt*) gene diversity and functions in freshwater

dinoflagellates. The diagram illustrates the comparative genomic and functional landscape of secondary metabolism-related gene clusters in freshwater dinoflagellates. The figure is divided into four quadrants representing Modular PKS, Singular PKS, Type II FAS, and Saxitoxins.

Tables

Table 1. Transcriptome assembly statistics and annotations of *Palatinus apiculatus, Peridinium bipes, and Ceratium furcoides* library.

Species	*Palatinus apiculatus*	*Peridinium bipes*	*Ceratium furcoides*
Strain	FD-02	DOP	P10K-1
Data Source	This study	NCBI	NGDC
Sequencing			
Sequencing Platform	NovaSeq 6000	HiSeq 2500	HiSeq 4000
Raw reads	31,358,869	27,008,331	16,717,658
Assembly			
Number of transcripts	111,352	179,865	82,155
Number of unigenes	75,385	107,791	68,218
No of bases (bp)	110,348,805	191,307,715	57,103,880
N50	1485	1,728	838
GC (%)	67.50	67.44	53.91
Annotation			
NCBI NR	90,922	93,901	35,618
Swissprot	19,483		27,832
Pfam	37,220	22,851	8,352
GO	22,324	10,838	4203

Table 2. The total number of polyketide synthase associated domains found for the three freshwater dinoflagellates.

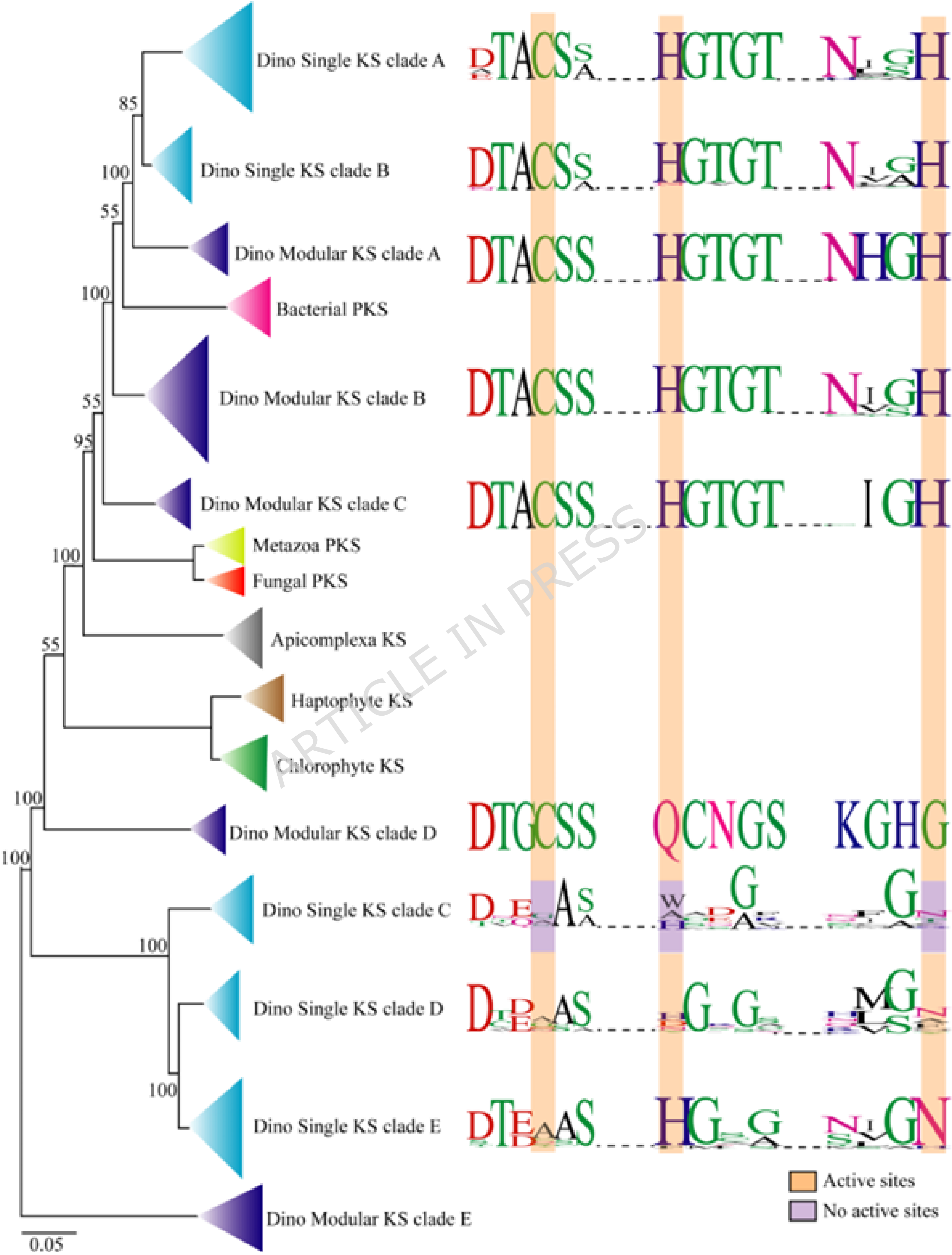
PKS	**Annotation**	***P. apiculatus***	***P. bipes***	***C. furcoides***
Modular PKS	Multi domain	12	8	4
NPRS-PKS Hybrid	Multi domain PKS + NPRS	5	2	6
KS	Ketoacyl synthase	30	58	14
KR	Ketoreductase	23	38	2
ACP (PP)	Acyl Carrier Protein (Phosphopantetheine-binding protein)	21	9	10
AT	Acyl transferase	4	2	3

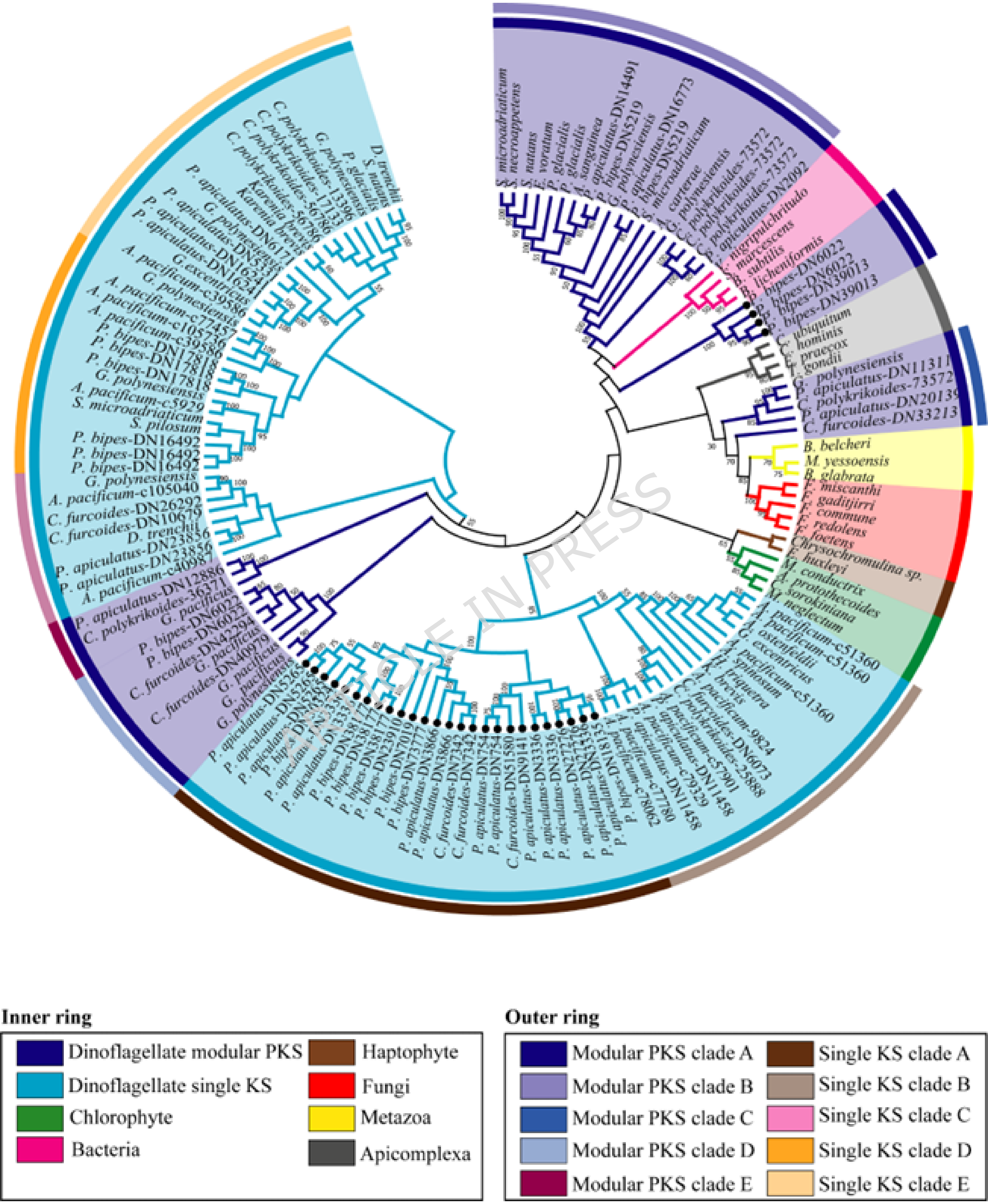
Table 3. List of type II FAS encoding transcripts from the three transcriptomes of freshwater dinoflagellates.

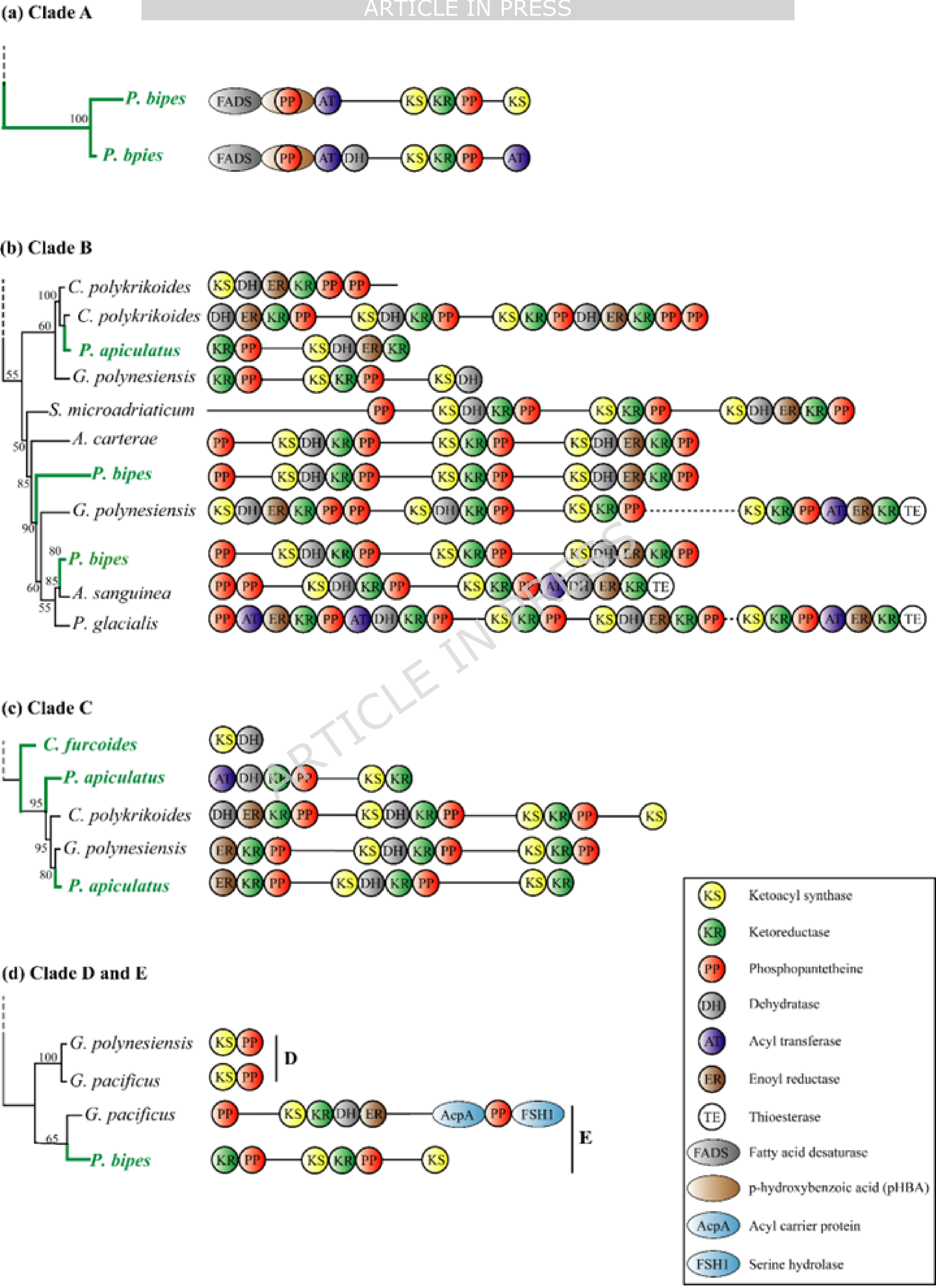
FAS II genes	**Annotation**	*Palatinus apiculatus*	*Peridinium bipes*	*Ceratium furcoides*
FabB	3-ketoacyl-ACP synthase I	20	8	0
FabD	malonyl-CoA:ACP transacylase	4	3	25
FabF	3-ketoacyl-ACP synthase II	1	25	5
FabG	3-ketoacyl-ACP reductase	6	8	0
FabH	3-ketoacyl-ACP synthase III	5	5	6
FabI	enoyl-ACP reductase I	1	2	1
FabK	enoyl-ACP reductase II	0	0	0
FabZ	3-hydroxydecanoyl-ACP dehydratase/isomerase	3	11	7

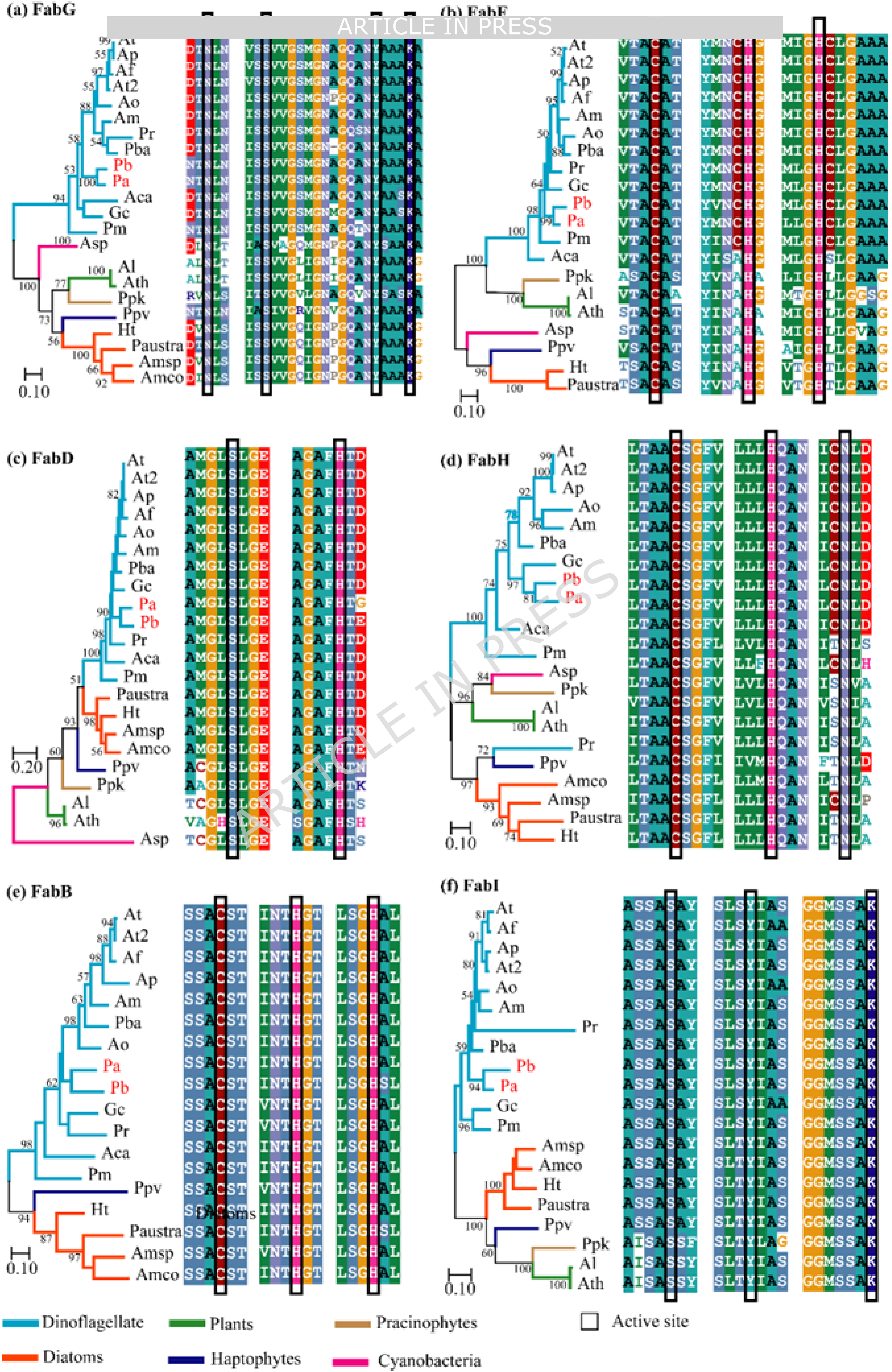
Table 4. Numbers of candidate *sxt* genes identified in the freshwater dinoflagellate *Palatinus apiculatus, Peridinium bipes,* and *Ceratium furcoides*

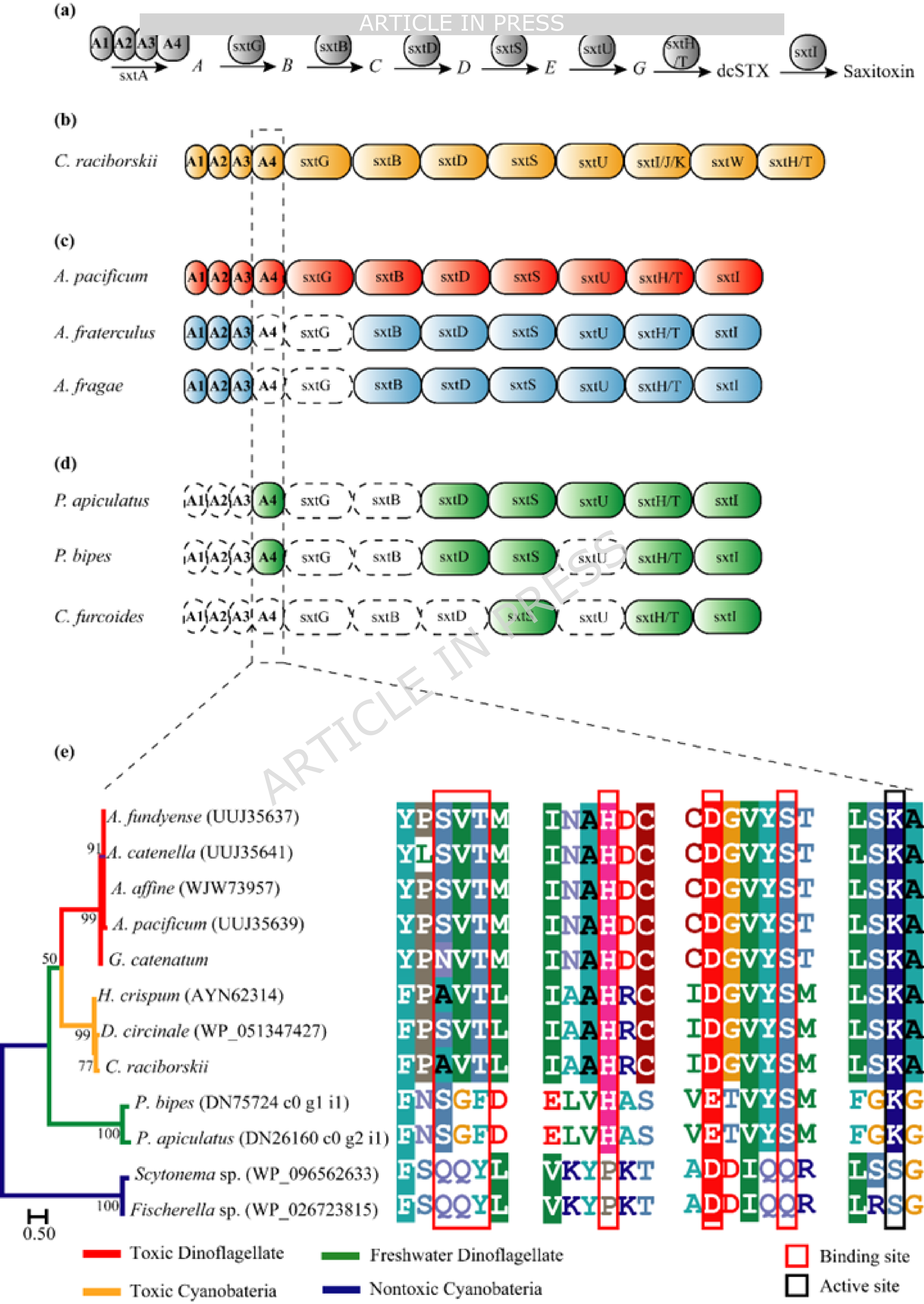
sxt genes	**Putative Function**	*Palatinus apiculatus*	*Peridinium bipes*	*Ceratium furcoides*
sxtA4	Aspartate aminotransferase	1	1	0
sxtB	Cytidine deaminase	0	0	0
sxtG	Amidinotransferase	0	0	0
sxtD	Sterol desaturase	3	2	0
sxtI	Carbomoyltransferase	2	3	2
sxtS	Phytanoyl-CoA dioxygenase	98	22	27
sxtU	Short-chain alcohol dehydrogenase/oxidoreductase	23	17	4
sxtH/T	Phenylpropionate dioxygenase/Rieske [2Fe-2S]	38	29	8
sxtO	Adenylyl sulfate kinase	3	3	0
sxtX	Cephalosporin hydroxylase	0	7	0
sxtF/M	Multidrug/MATE	44	56	25
sxtP	Binding of PSTs	2	0	0
sxtV	Succinate dehydrogenase/Fumarate reductase	7	15	25
sxtN	Sulfotransferase	0	0	0
sxtC	Amidohydrolase	8	0	0
sxtL	SGNH hydrolase/GDSL-Lipase	18	7	0
sxtW	Ferrodoxin [4Fe-4S]	16	21	19
sxtZ	Two-component sensor histidine kinase	3	2	0
sxtQ	Yqcl/YcgG family	0	0	0

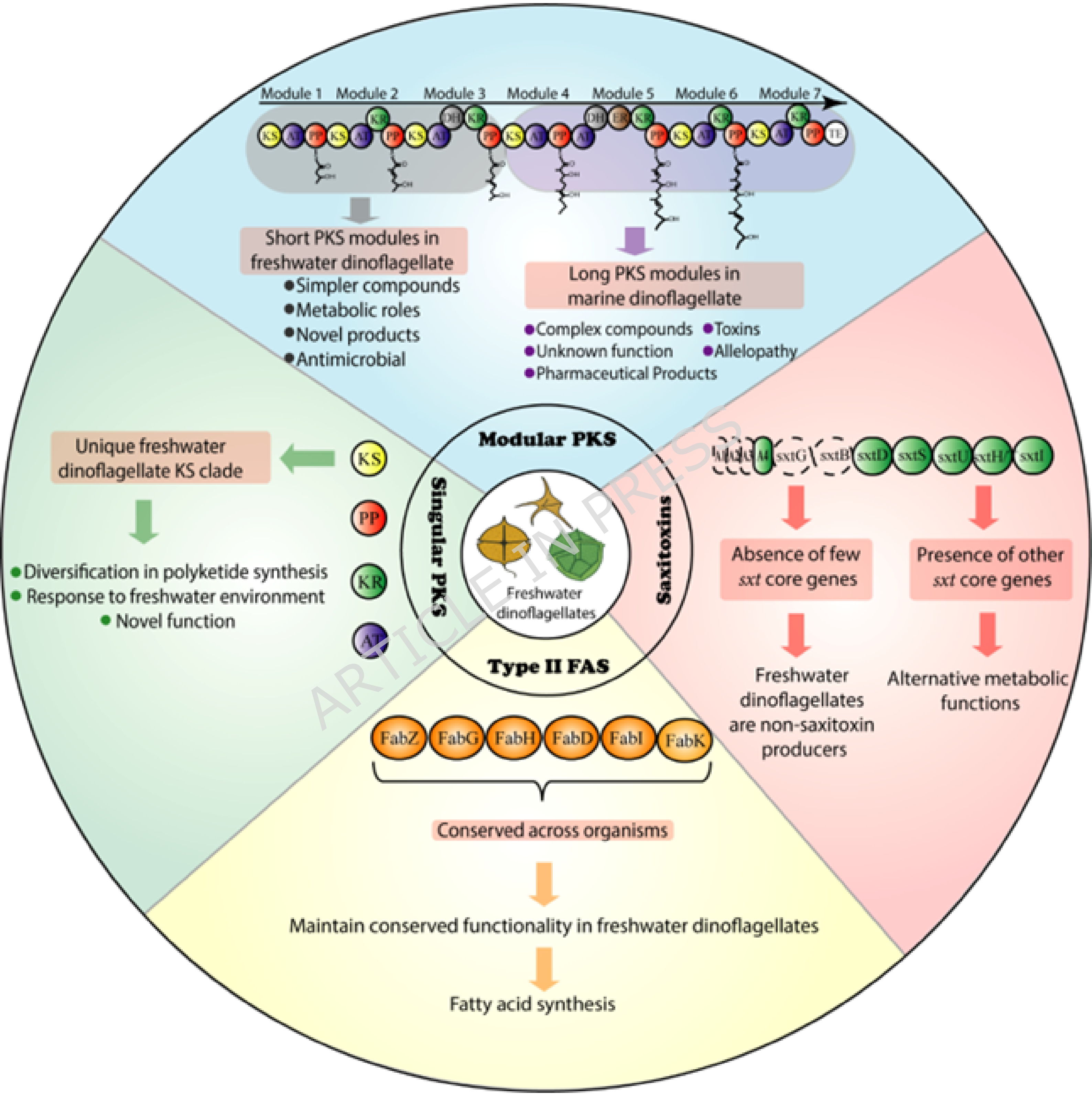

Dino Single KS clade A
85
Dino Single KS clade B
100
55
Dino Modular KS clade A
100
Bacterial PKS
Dino Modular KS clade B
55
95
Dino Modular KS clade C
100
Metazoa PKS
Fungal PKS
Apicomplexa KS
55
Haptophyte KS
Chlorophyte KS
100
Dino Modular KS clade D
100
Dino Single KS clade C
100
Dino Single KS clade D
100
Dino Single KS clade E
Dino Modular KS clade E
0.05
DTACSS
HGTGT
NHGH
DTGCSS
QCNGS
KGHG
Active sites
No active sites


S. microadriaticum
S. necroappetens
S. natans
S. voratum
P. glacialis
P. glacialis
A. sanguinea
P. apiculatus-DN14491
P. bipes-DN5219
G. polynesiensis
P. apiculatus-DN16773
P. bipes-DN5219
S. microadriaticum
A. carterae
G. polynesiensis
C. polykrikoides-73572
C. polykrikoides-73572
C. polykrikoides-73572
P. apiculatus-DN2092
S. nigripulchritudo
S. marcescens
B. subtilis
B. licheniformis
P. bipes-DN6022
P. bipes-DN6022
P. bipes-DN39013
P. bipes-DN39013
C. ubiquitum
C. hominis
E. praecox
T. gondii
G. polynesiensis
P. apiculatus-DN11311
C. polykrikoides-73572
P. apiculatus-DN20139
C. furcoides-DN33213
B. belcheri
M. yessoensis
B. glabrata
F. miscanthi
F. gaditjirri
F. commune
F. redolens
F. foetens
Chrysochromulina sp.
E. huxleyi
M. conductrix
A. protothecoides
C. sorokiniana
M. neglectum
A. pacificum-c51360
A. pacificum-c51360
G. excentricus
A. ostenfeldii
A. pacificum-c51360
S. spinosum
H. triquetra
K. brevis
C. furcoides-DN6073
A. pacificum-9824
C. polykrikoides-25888
A. pacificum-c57901
P. bipes-DN11458
P. apiculatus-DN11458
A. pacificum-c79339
A. pacificum-c77780
P. apiculatus-DN18135
P. apiculatus-DN2270
P. apiculatus-DN3336
P. apiculatus-DN3336
P. apiculatus-DN9141
C. furcoides-DN51580
P. apiculatus-DN754
P. apiculatus-DN754
C. furcoides-DN7345
C. furcoides-DN7345
P. apiculatus-DN3866
P. apiculatus-DN73777
P. bipes-DN73777
P. bipes-DN7016
P. bipes-DN23910
P. bipes-DN38177
P. bipes-DN38177
P. apiculatus-DN333
G. polynesiensis
G. pacificus
G. pacificus
C. furcoides-DN40979
G. pacificus
C. furcoides-DN42294
P. bipes-DN6022
P. bipes-DN6022
G. pacificus
C. polykrikoides-36371
P. apiculatus-DN12886
A. pacificum-c40987
P. apiculatus-DN23856
P. apiculatus-DN23856
D. trenchii
C. furcoides-DN10679
C. furcoides-DN26292
A. pacificum-c105040
G. polynesiensis
P. bipes-DN16492
P. bipes-DN16492
P. bipes-DN16492
S. pilosum
S. microadriaticum
A. pacificum-c5929
G. polynesiensis
P. bipes-DN17818
P. bipes-DN17818
P. bipes-c39586
A. pacificum-c105745
A. pacificum-c77457
G. polynesiensis
A. pacificum-c39586
G. excentricus
P. apiculatus-DN1654
P. apiculatus-DN1654
P. apiculatus-DN533
G. polynesiensis
Karenia brevis
C. polykrikoides-DN617
C. polykrikoides-56789
C. polykrikoides-17133
C. polykrikoides-13398
G. polynesiensis
P. glacialis
S. natans
D. trenchii
Inner ring
Dinoflagellate modular PKS
Dinoflagellate single KS
Chlorophyte
Bacteria
Haptophyte
Fungi
Metazoa
Apicomplexa
Outer ring
Modular PKS clade A
Modular PKS clade B
Modular PKS clade C
Modular PKS clade D
Modular PKS clade E
Single KS clade A
Single KS clade B
Single KS clade C
Single KS clade D
Single KS clade E


(a) Clade A
P. bipes
100
P. bpies
FADS PP AT KS KR PP KS
FADS PP AT DH KS KR PP AT
(b) Clade B
C. polykrikoides
C. polykrikoides
P. apiculatus
G. polynesiensis
S. microadriaticum
A. carterae
P. bipes
G. polynesiensis
P. bipes
A. sanguinea
P. glacialis
100
60
55
50
85
90
80
85
60
55
KS DH ER KR PP PP
DH ER KR PP KS DH KR PP KS KR PP DH ER KR PP PP
KR PP KS DH ER KR
KR PP KS KR PP KS DH
PP KS DH KR PP KS KR PP KS DH ER KR PP
PP KS DH KR PP KS KR PP KS DH ER KR PP
PP KS DH KR PP KS KR PP KS DH ER KR PP
KS DH ER KR PP PP KS DH KR PP KS KR PP KS KR PP AT ER KR TE
PP KS DH KR PP KS KR PP KS DH ER KR PP
PP PP KS DH KR PP KS KR PP AT DH ER KR TE
PP AT ER KR PP AT DH KR PP KS KR PP KS DH ER KR PP KS KR PP AT ER KR TE
(c) Clade C
C. furcoides
P. apiculatus
C. polykrikoides
G. polynesiensis
P. apiculatus
95
95
80
KS DH
AT DH KR PP KS KR
DH ER KR PP KS DH KR PP KS KR PP KS
ER KR PP KS DH KR PP KS KR PP
ER KR PP KS DH KR PP KS KR
(d) Clade D and E
G. polynesiensis
G. pacificus
G. pacificus
P. bipes
100
65
KS PP
KS PP
D
PP KS KR DH ER AcpA PP FSH1
KR PP KS KR PP KS
E
KS Ketoacyl synthase
KR Ketoreductase
PP Phosphopantetheine
DH Dehydratase
AT Acyl transferase
ER Enoyl reductase
TE Thioesterase
FADS Fatty acid desaturase
p-hydroxybenzoic acid (pHBA)
AcpA Acyl carrier protein
FSH1 Serine hydrolase


(a) FabG
(b) FabF
(c) FabD
(d) FabH
(e) FabB
(f) FabI
At
At2
Ap
Af
Ao
Am
Pr
Pba
Pb
Pa
Aca
Gc
Pm
Asp
Al
Ath
Ppk
Ppv
Ht
Paustra
Amsp
Amco
Diatoms
0.10
0.20
Dinoflagellate
Plants
Pracinophytes
Active site
Diatoms
Haptophytes
Cyanobacteria


(a)
A1 A2 A3 A4
sxtA
A
sxtG
B
sxtB
C
sxtD
D
sxtS
E
sxtU
G
sxtH/T
dcSTX
sxtI
Saxitoxin
(b)
C. raciborskii
A1 A2 A3 A4 sxtG sxtB sxtD sxtS sxtU sxtI/J/K sxtW sxtH/T
(c)
A. pacificum
A1 A2 A3 A4 sxtG sxtB sxtD sxtS sxtU sxtH/T sxtI
A. fraterculus
A1 A2 A3 A4 sxtG sxtB sxtD sxtS sxtU sxtH/T sxtI
A. fragae
A1 A2 A3 A4 sxtG sxtB sxtD sxtS sxtU sxtH/T sxtI
(d)
P. apiculatus
A1 A2 A3 A4 sxtG sxtB sxtD sxtS sxtU sxtH/T sxtI
P. bipes
A1 A2 A3 A4 sxtG sxtB sxtD sxtS sxtU sxtH/T sxtI
C. furcoides
A1 A2 A3 A4 sxtG sxtB sxtD sxtS sxtU sxtH/T sxtI
(e)
A. fundyense (UUJ35637)
91 A. catenella (UUJ35641)
A. affine (WJW73957)
99 A. pacificum (UUJ35639)
50 G. catenatum
H. crispum (AYN62314)
99 D. circinale (WP_051347427)
77 C. raciborskii
P. bipes (DN75724 c0 g1 i1)
100 P. apiculatus (DN26160 c0 g2 i1)
Scytonema sp. (WP_096562633)
100 Fischerella sp. (WP_026723815)
0.50
YPSVTM INAHDC CDGVYST LSKA
YLSVTM INAHDC CDGVYST LSKA
YPSVTM INAHDC CDGVYST LSKA
YPSVTM INAHDC CDGVYST LSKA
YPNVTM INAHDC CDGVYST LSKA
FPAVTL IAAHRC IDGVYSM LSKA
FPSVTL IAAHRC IDGVYSM LSKA
FPAVTL IAAHRC IDGVYSM LSKA
FNSGFD ELVHAS VETVYSM FGKG
FNSGFD ELVHAS VETVYSM FGKG
FSQQYL VKYPKT ADDIQQR LSSG
FSQQYL VKYPKT ADDIQQR LRSG
Toxic Dinoflagellate
Freshwater Dinoflagellate
Toxic Cyanobateria
Nontoxic Cyanobateria
Binding site
Active site


Module 1
Module 2
Module 3
Module 4
Module 5
Module 6
Module 7
KS
AT
PP
KR
DH
ER
TE
Short PKS modules in freshwater dinoflagellate
●Simpler compounds
●Metabolic roles
●Novel products
●Antimicrobial
Long PKS modules in marine dinoflagellate
●Complex compounds
●Toxins
●Unknown function
●Allelopathy
●Pharmaceutical Products
Modular PKS
Singular PKS
Saxitoxins
Type II FAS
Freshwater dinoflagellates
Unique freshwater dinoflagellate KS clade
● Diversification in polyketide synthesis
● Response to freshwater environment
● Novel function
sxtG
sxtB
sxtD
sxtS
sxtU
sxtH/
sxtI
Absence of few *sxt* core genes
Presence of other *sxt* core genes
Freshwater dinoflagellates are non-saxitoxin producers
Alternative metabolic functions
FabZ
FabG
FabH
FabD
FabI
FabK
Conserved across organisms
Maintain conserved functionality in freshwater dinoflagellates
Fatty acid synthesis