Molecular Phylogenetics and Evolution 92 (2015) 165–180

Contents lists available at ScienceDirect

Molecular Phylogenetics and Evolution

journal homepage: www.elsevier.com/locate/ympev

Gene duplication, loss and selection in the evolution of saxitoxin biosynthesis in alveolates ☆

Shauna A. Murray [a,b,*], Rutuja Diwan [a,b], Russell J.S. Orr [c], Gurjeet S. Kohli [a], Uwe John [d]

[a] Plant Functional Biology and Climate Change Cluster, University of Technology Sydney, NSW 2007, Australia
[b] Sydney Institute of Marine Science, Chowder Bay Rd, Mosman, NSW, Australia
[c] Section for Genetics and Evolutionary Biology, Department of Biosciences, University of Oslo, Norway
[d] Section Ecological Chemistry, Alfred-Wegener-Institut, Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Germany

ARTICLE INFO

Article history:
Received 9 March 2015
Revised 1 June 2015
Accepted 25 June 2015
Available online 30 June 2015

Keywords:
Alexandrium
Saxitoxin
sxtA
sxtG
Paralytic shellfish poisoning
Dinoflagellate

ABSTRACT

A group of marine dinoflagellates (Alveolata, Eukaryota), consisting of ~10 species of the genus *Alexandrium, Gymnodinium catenatum* and *Pyrodinium bahamense*, produce the toxin saxitoxin and its analogues (STX), which can accumulate in shellfish, leading to ecosystem and human health impacts. The genes, *sxt*, putatively involved in STX biosynthesis, have recently been identified, however, the evolution of these genes within dinoflagellates is not clear. There are two reasons for this: uncertainty over the phylogeny of dinoflagellates; and that the *sxt* genes of many species of *Alexandrium* and other dinoflagellate genera are not known. Here, we determined the phylogeny of STX-producing and other dinoflagellates based on a concatenated eight-gene alignment. We determined the presence, diversity and phylogeny of *sxtA*, domains A1 and A4 and *sxtG* in 52 strains of *Alexandrium*, and a further 43 species of dinoflagellates and thirteen other alveolates. We confirmed the presence and high sequence conservation of *sxtA*, domain *A4*, in 40 strains (35 *Alexandrium*, 1 *Pyrodinium*, 4 *Gymnodinium*) of 8 species of STX-producing dinoflagellates, and absence from non-producing species. We found three paralogs of *sxtA*, domain A1, and a widespread distribution of *sxtA1* in non-STX producing dinoflagellates, indicating duplication events in the evolution of this gene. One paralog, clade 2, of *sxtA1* may be particularly related to STX biosynthesis. Similarly, *sxtG* appears to be generally restricted to STX-producing species, while three amidinotransferase gene paralogs were found in dinoflagellates. We investigated the role of positive (diversifying) selection following duplication in *sxtA1* and *sxtG*, and found negative selection in clades of *sxtG* and *sxtA1*, clade 2, suggesting they were functionally constrained. Significant episodic diversifying selection was found in some strains in clade 3 of *sxtA1*, a clade that may not be involved in STX biosynthesis, indicating pressure for diversification of function.

1. Introduction

Secondary metabolites, such as toxins, can act as important models for the study of evolutionary processes, as they often consist of a measurable phenotype with known ecological impacts, produced by a specific group of genes (Fischbach et al., 2008). Processes such as gene duplication, selection and lateral transfer have been found to play a role in the evolution of ecologically significant traits in protists and in cyanobacteria (Murray et al., 2011b; Slamovits and Keeling, 2008; Waller et al., 2006). Saxitoxin and its analogs (STXs), one such secondary metabolite, are alkaloid compounds with major impacts on marine environments, as they can accumulate in organisms including crustaceans, molluscs, echinoderms, cephalopods, fish, turtles, marine mammals and birds (Llewellyn et al., 2006). As they can reach high concentrations in commercial shellfish species, they have major economic impacts on aquaculture industries worldwide. STXs are produced by freshwater prokaryotic cyanobacteria and a group of marine eukaryotic protists. The genes putatively involved in STX biosynthesis (*sxt*), involving ~30 biosynthetic steps, are now known from ~6 cyanobacterial genera (Kellmann et al., 2008; Mihali et al., 2009, 2011; Moustafa et al., 2009; Stucken et al., 2010).

Putative genes involved in the biosynthesis of STX in dinoflagellates have been found, despite the large genome sizes (~1–100 Gb) of dinoflagellates hampering research (Stüken et al., 2011; Murray et al., 2011a; Hii et al., 2012; Hackett et al., 2013; Orr et al., 2013;

☆ This paper was edited by the Associate Editor Dr. J.B. Dacks.

* Corresponding author at: Plant Functional Biology and Climate Change Cluster, University of Technology Sydney, NSW 2007, Australia. Fax: +61 2 9514 4079.
E-mail address: Shauna.Murray@uts.edu.au (S.A. Murray).

http://dx.doi.org/10.1016/j.ympev.2015.06.017
1055-7903/

Suikkanen et al., 2013). They have molecular genetic characteristics typical of dinoflagellates: a conserved 22 bp spliced leader sequence at the 5' end of the transcribed mRNA; eukaryotic poly A tails; a GC content of 62–69%, similar to that reported for *Alexandrium* genomes, compared to 43% for cyanobacterial *sxtA* genes; dinoflagellate-related signal peptides; and a high genomic copy number of $\sim 10^2$ copies cell^{-1}, a feature typical of dinoflagellates (Stüken et al., 2011; Murray et al., 2011a; Orr et al., 2013). A complement of 14 'core' genes (*sxtA–sxtI*, *sxtP–sxtR*, *sxtS*, and *sxtU*) is common between the *sxt* clusters of cyanobacterial STX-producing strains (Murray et al., 2011b), of which eight (*sxtA*, *sxtB*, *sxtD*, *sxtG*, *sxtH* or *sxtT*, *sxtI*, *sxtS*, and *sxtU*) may be directly implicated in STX synthesis (Kellmann et al., 2008). In cyanobacteria, the gene cluster appears to have an ancient origin, and both gene duplication and positive selection have played a role in its evolution (Murray et al., 2011b). Homologs of all genes putatively necessary for STX biosynthesis have been found in dinoflagellates, in the species *Alexandrium fundyense*, *A. minutum*, *Gymnodinium catenatum* and *Pyrodinium bahamense* (Stüken et al., 2011; Hackett et al., 2013).

Dinoflagellate producers of STX that have been studied possess the genes *sxtA*, putatively the initial gene in the STX synthesis pathway (Stüken et al., 2011), and the second gene *sxtG* (Orr et al., 2013). *sxtA* has four catalytic domains, with predicted activities similar to a SAM-dependent methyltransferase (*sxtA1*), GCN-5 related N-acetyltransferase (*sxtA2*), acyl carrier protein (*sxtA3*) and an amidinotransferase (*sxtA4*) (Kellmann et al., 2008). *sxtA* in dinoflagellates shows two isoforms: one of which comprises four domains, *sxtA1–sxtA4*, while the other one encompasses only the domains *sxtA1–sxtA3* (Stüken et al., 2011), and may represent a duplicated, paralogs copy of *sxtA1*. The domain *sxtA4* appears to be necessary for STX biosynthesis, and therefore it is likely that only the isoform with this domain is active in catalysing STX biosynthesis (Murray et al., 2011a). Part or all of *sxtA* appears to be absent from dinoflagellate genera and families that cannot produce STX (Stüken et al., 2011; Orr et al., 2013; Murray et al., 2014; John et al., 2014). Strains of *Alexandrium ostenfeldii,* which do not produce STX appear to lack the domain *sxtA4*, while those that can produce STX have this domain (Suikkanen et al., 2013). *sxtG* was present in all STX-producing dinoflagellates, however it was also present in three other species within the genus *Alexandrium* (Orr et al., 2013).

Only limited information is available regarding the expression and function of *sxtA* copies (Perini et al., 2014; Zhang et al., 2014). It appears that *sxtA* may not be strongly regulated at the transcriptional level in *A. minutum*, as toxin quantities were not well correlated with transcript abundance (Perini et al., 2014). Nevertheless, *sxtA* was found to be significantly down regulated in a non-toxic mutant strain of *A. pacificum* (Zhang et al., 2014). There are several environmental factors that appear to influence the amount of STX found in the cell (reviewed in Murray et al., 2015). Certain growth phases have been correlated with higher rates of cellular toxin quota, including mid-exponential growth phase (Cembella and Destombe, 1996), or stationary phase (Parkhill and Cembella, 1999, Lim and Ogata, 2005). The cellular toxin quota appears to increase upon phosphorus limitation and decrease upon nitrogen limitation (van de Waal et al., 2014). One factor that has been found to lead to a several fold increase in cellular toxin quota is the presence of copepods and their waterborne cues (Selander et al., 2006; Wohlrab et al., 2010; Yang et al., 2011).

Almost all species of dinoflagellates that produce STX belong to the genus *Alexandrium,* of which ~32 species have been described, and are present worldwide in comparable habitats (Balech, 1995; Anderson et al., 2012; Farrell et al., 2013). In order to determine the evolution of *sxt* within dinoflagellates, a resolved phylogeny of the group, including the genus *Alexandrium* and closely related *Pyrodinium bahamense* is necessary. *Alexandrium* has often appeared to form a monophyletic clade in phylogenetic analyses based on regions of the ribosomal DNA array, with a selected outgroup (John et al., 2003; Leaw et al., 2005; Orr et al., 2011; Anderson et al., 2012). Questions remain about the evolution of *Alexandrium*, as different studies have shown that the most basal clade is either: *A. taylori* (John et al., 2003; Rogers et al., 2006), *A. leei* (Leaw et al., 2005), a clade composed of the species *A. satoanum, A. monilatum, A. taylori, A. hiranoi, A. pseudogonyaulax* (Murray et al., 2014), and a clade including most of those species and *A. leei* (Anderson et al., 2012), or a clade including *A. minutum, A. tamutum* and *A. ostenfeldii* (Orr et al., 2011). A recent phylogeny of the group with a comprehensive species coverage, based on a concatenated alignment of the three rRNA regions, found that clades within *Alexandrium* were generally not well supported (Murray et al., 2014). The closest sister group to *Alexandrium* is not yet clear, as most studies of dinoflagellates have either focused on only *Alexandrium* with few outgroups; or covered all dinoflagellates, with only few *Alexandrium* species. In order to determine a resolved phylogeny of the genus, alignments of larger numbers of concatenated genes are necessary, and this has been shown to offer more resolution also for other dinoflagellate groups (Orr et al., 2012).

Many microbial eukaryotes appear to have genomes that are to some extent mosaic, with a significant number of genes laterally acquired, from organisms such as cyanobacteria, proteobacteria, and euglenoids (Keeling and Palmer, 2008; Waller et al., 2006; Wisecaver and Hackett, 2014). The genes *sxtA* and *sxtG* appear likely to have been acquired via an ancient transfer from a prokaryote into a dinoflagellate ancestor (Stüken et al., 2011; Orr et al., 2013). However, the extent to which lateral transfers have occurred between dinoflagellates species is not yet clear. Few species of *Alexandrium* have been examined for the presence of *sxtA,* and representatives of only some dinoflagellate families have been investigated using deep sequencing techniques. Similarly, non-dinoflagellate alveolates have not been screened for the presence of these genes. The potential role of gene duplication in the evolution of this domain is not known. The aim of this study was to examine the potential roles of horizontal transfer, gene duplication and selection in the evolution of *sxtA* and *sxtG* within the dinoflagellates and closely related alveolates, and to examine the diversity of copies of *sxtA4* in species of the genus *Alexandrium*.

2. Material and methods

2.1. Culture maintenance

Twenty-one strains of *Alexandrium* species (Table 1) were obtained from the culture collections: the University of Tasmania culture collection, The National Centre for Culture of Marine Phytoplankton (CCMP strain numbers), Australian National Algae Culture Collection (CS strain numbers) and the Cawthron Institute culture collection of microalgae (CAWD strain numbers) (Table 1, strains in bold). *Alexandrium fundyense* strains A8 and E4 were isolated from the north-east coast of Scotland (N57°54′97.22″, E-1°57′38.89″) in 2010 by Yameng Lu. Species names for the *Alexandrium tamarense* species complex are in accordance with the most recent taxonomic treatment (John et al., 2014), in which Group I ribotype are referred to *Alexandrium fundyense*, Group II are referred to *A. mediterraneum*, Group III are referred to *A. tamarense*, Group IV are referred to *A. pacificum* and Group V to *A. australiense*.

Strains were grown in K, GP or GSe-medium (Keller et al., 1987; Doblin et al., 1999) prepared from filter-sterilized (0.2 μm VacuCap 90 filter units, Pall Corporation, Port Washington) seawater

Table 1
Species that have been investigated for the production of STXs, and the presence of domains A1 and A4 of the *sxtA* gene, either by PCR and sequencing, or by searching EST libraries. n/a – not applicable as sequencing was not attempted, n.d. – not detected. Identification numbers of EST libraries beginning with MMETSP were part of the Moore Foundation Marine Microbial Eukaryote Transcriptome Project. Others are taken from GenBank. The total number of contigs in a library is given in order to determine the comprehensiveness of the library. The total number of sequences of the relevant *sxt* gene identified in the library is shown in brackets (no seqs). Strains in bold were obtained, sequenced or searched in this study.

Genus	Species	Strain	Documented to produce STXs?	Method for identification of *sxtA* domains	EST library		Genes identified or sequenced			Reference
					ID no. of library	Total no. of contigs in library	*sxtA1* (no. seqs)	sxtA4 (no. seqs)	*sxtG* (No. seqs)	
Gonyaulacales										
Alexandrium	*fundyense*	CCMP1719	Yes (Orr et al., 2011)	PCR	–	–	Yes	Yes	Yes	Stüken et al. (2011), Orr et al. (2013)
		CCMP1979	Yes (Orr et al., 2011)	PCR	–	–	Yes	Yes		Stüken et al. (2011)
		OF101	Yes (Scholin et al., 1994)	PCR	–	–	n/a	Yes		John et al. (2014)
		SPE10-10-3	Yes (Hackett et al., 2013)	PCR	–	–	n/a	Yes		John et al. (2014)
		C8	Yes	PCR	–	–	n/a	Yes		This study
		E4	Yes	PCR	–	–	n/a	Yes		This study
		ACQH01	Yes (Scholin et al., 1994)	PCR	–	–	n/a	Yes		John et al. (2014)
		AtGTM253, 38-3, SPE10-03	Yes (?)	EST	169259	92,394	Yes (2)	Yes (1)	Yes (3)	Hackett et al. (2013)
Alexandrium	*mediterraneum*	SZN01	No	PCR			n/a	n.d.		John et al. (2014)
		SZN08	No	PCR			n/a	n.d.		John et al. (2014)
Alexandrium	*tamarense*	ATSW01-1	No	PCR			n/a	n.d.		John et al. (2014)
		CCMP1771	No	PCR			Yes	(?) yes		Stüken et al. (2011)
		CCMP1771	No	EST	MMETSP0384 MMETSP0382	106,664 and 98,253	Yes (2)	n.d.	n.d.	This study
		ATSP1-B	No	EST	prjna169250	94,313	Yes (1)	n.d.	n.d.	Hackett et al. (2013)
Alexandrium	*pacificum*	CCMP1493	Yes (Dantzer and Levin, 1997)	PCR	–	–	Yes	Yes	Yes	Stüken et al. (2011), Orr et al. (2013)
		CS-300/01	Yes uncharacterised	PCR	–	–	Yes	Yes		This study
		CS-315	Yes uncharacterised	PCR	–	–	Yes	Yes		This study
		CS-313/01	Yes uncharacterised	PCR	–	–	Yes	Yes		This study
		CS-798	Yes uncharacterised	PCR	–	–	Yes	Yes		This study
		ACTRA02	Yes (Murray et al., 2011a,b)	PCR	–	–	Yes	Yes		This study; Stüken et al., 2011
		CAWD44	Yes (MacKenzie et al., 2004)	PCR	–	–	Yes	Yes		This study
		CAWD121	Yes (MacKenzie et al., 2004)	PCR	–	–	Yes	Yes		This study
		CCMP1598	Yes	EST	163,063	151,525	Yes (5)	Yes (2)	Yes (4)	This study
		ACCC01	Yes (Murray et al., 2011a,b)	PCR	–	–	Yes	Yes	Yes	This study; Stüken et al., 2011; Orr et al., 2013
		ACSH02	Yes (Murray et al., 2011a,b)	PCR	–	–	Yes	Yes		This study; Stüken et al., 2011
Alexandrium	*australiense*	ATBB01	Yes/no (Orr et al., 2011; Murray et al., 2012a,b)	PCR	–	–	Yes	Yes		Stüken et al. (2011)
		ATCJ33	No (Stüken et al., 2011)	PCR	–	–	Yes	Yes		This study; Stüken et al., 2011
		ATNWB01	Yes (Murray et al., 2012a,b)	PCR	–	–	Yes	Yes		Stüken et al. (2011)
		ATEB01	No (Stüken et al., 2011)	PCR	–	–	Yes	Yes		Stüken et al. (2011)
Alexandrium	*affine*	CCMP112	No	PCR	–	–	n.d.	n.d.	Yes	Stüken et al. (2011), Orr et al. (2013)
Alexandrium	*andersoni*	CCMP2222	No	PCR	–	–	n.d.	n.d.	Yes (1)	Stüken et al. (2011), Orr et al. (2013)
				EST	MMETSP1436	42,240	Yes (1)	n.d.		This study
Alexandrium	*concavum*	**CAWD51**	No	PCR	–	–	Yes	n.d.		This study
Alexandrium	*diversaporum*	**AAKT01**	No	PCR	–	–	Yes	n.d.		This study, Murray et al., 2014
Alexandrium	*fraterculus*	**CAWD52**	No	PCR	–	–	Yes	n.d.		This study

(*continued on next page*)

Table 1 (*continued*)

Genus	Species	Strain	Documented to produce STXs?	Method for identification of *sxtA* domains	EST library		Genes identified or sequenced			Reference
					ID no. of library	Total no. of contigs in library	*sxtA1* (no. seqs)	sxtA4 (no. seqs)	*sxtG* (No. seqs)	
		CAWD97	No	PCR	–	–	Yes	n.d.		This study
Alexandrium	*insuetum*	CCMP2082	No	PCR	–	–	n.d.	n.d.	Yes	Orr et al. (2013)
Alexandrium	*margalefi*	CS-322	No	PCR	–	–	Yes	n.d.		This study
			No	EST	CLC bio assembly	109943	Yes	n.d.		This study
Alexandrium	*minutum*	CCMP113	Yes (Orr et al., 2011)	PCR	–	–	Yes	Yes	Yes	Stüken et al. (2011), Orr et al. (2013)
		ALSP01	Yes (Orr et al., 2011)	PCR	–	–	Yes	Yes		Stüken et al. (2011)
		ALSP02	Yes (Orr et al., 2011)	PCR	–	–	n/a	Yes		Stüken et al. (2011)
		AMAD16	Yes (Orr et al., 2011)	PCR	–	–	n/a	Yes		Stüken et al. (2011)
		CCMP1888	Yes (Orr et al., 2011)	PCR	–	–	Yes	Yes		Stüken et al. (2011)
		CAWD11	Yes (MacKenzie and Berkett, 1997)	PCR	–	–	Yes	Yes		This study
		CAWD12	Yes (MacKenzie and Berkett, 1997)	PCR	–	–	Yes	Yes		This study
		CS-324	Yes (Negri et al., 2003)	PCR	–	–	Yes	Yes	Yes	This study; Stüken et al., 2011; Orr et al., 2013
		CS-320/01	Yes (Orr et al., 2011)	PCR	–	–	Yes	Yes	Yes	Stüken et al. (2011), Orr et al. (2013)
Alexandrium	*monilatum*	JR08	No	EST	MMETSP0093	95,897	Yes (1)	n.d.	n.d.	This study
Alexandrium	*ostenfeldii*	AOVA30	Yes (Suikkanen et al., 2013)	PCR	–	–	Yes	Yes		Suikkanen et al. (2013)
		AOF0927	Yes (Suikkanen et al., 2013)	PCR	–	–	Yes	Yes		Suikkanen et al. (2013)
		NCH85	Yes (Suikkanen et al., 2013)	PCR	–	–	Yes	n.d.		Suikkanen et al. (2013)
		S06/013/01	Yes (Brown et al., 2010)	PCR	–	–	Yes	n.d.		Suikkanen et al. (2013)
Alexandrium	*pseudogonyaulax*	**CAWD54**	No	PCR	–	–	Yes	n.d.		This study
Alexandrium	*tamiyavanichi*		Yes (Hii et al., 2012)	PCR	–	–	Yes	Yes		Hii et al. (2012)
Ceratium	*longipes*	CCMP1770	No	PCR	–	–	n.d.	n.d.		Orr et al., 2012
Ceratium	*fusus*		No	EST	MMETSP1074	82,514	n.d.	n.d.		This study
					MMETSP1075	87,514				
Coolia	*monotis*	CAWD98	No	PCR	–	–	n.d.	n.d.		Orr et al., 2012
Crypthecodinium	*cohnii*	**Seligo**	No	EST	MMETSP0325	70,674	n.d.	n.d.		This study
					MMETSP0324	68,757	n.d.	n.d.		This study
Gambierdiscus	*australes*	**CAWD149**	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
			No	EST	–	106,797	Yes (4)	n.d.	n.d.	This study
Gambierdiscus	*belizeanus*	**CCMP401**	No	EST	–	114,510	Yes (5)	n.d.	n.d.	This study
Gonyaulax	*spinifera*	**CCMP409**	No	EST	MMETSP1439	42,076	Yes (1)	n.d.	n.d.	This study
Lingulodinium	*polyedrum*	**CCMP1931**	No	PCR	–	–	n.d.	n.d.		Orr et al., 2012
				EST	Genbank	74,892	Yes (2)	n.d.	n.d.	This study
				EST	MMETSP1035	88,741	Yes (3)	n.d.	n.d.	This study
Protoceratium	*reticulatum*	**CAWD99**	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
		CCMP1889	No	EST	MMETSP0228	78,281	Yes (1)	n.d.	n.d.	This study
Pyrocystis	*noctiluca*	CCMP732	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
Pyrodinium	*bahamense*	**pbaha01**	Yes (Harada et al., 1982)	EST	MMETSP0796	105,175	Yes (1)	Yes (1)	Yes (2)	This study
Thecadinium	*kofoidii*	SCCAP K-1504	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
Dinophysiales										
Dinophysis	*acuminata*	**DAEP01**	No	EST	MMETSP0797	89,951	Yes (2)	n.d.	n.d.	This study
Gymnodiniales										
Amphidinium	*carterae*	UIO081	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
		CCMP1314	No	EST	MMETSP0258	44,378	n.d.	n.d.	n.d.	This study
					MMETSP0259	45,656	n.d.	n.d.	n.d.	This study
Amphidinium	*massartii*	**CS-259**	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
				EST	CLC assembly	54,729	n.d.	n.d.	n.d.	This study
Amphidinium	*mootonorum*	CAWD161	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
Gymnodinium	*aureolum*	SCCAP K-1561	No	PCR	–		n.d.	n.d.		Orr et al. (2012)

Gymnodinium	*catenatum*	CCMP1937	Yes (Oshima et al., 1987)	PCR	–	–	Yes	Yes	Yes	Stüken et al. (2011), Orr et al. (2013)
		GC744	Yes	EST	MMETSP0784	88,811	Yes (3)	Yes (5)	Yes (2)	This study
		GCTRA01	Yes (Stüken et al., 2011)	PCR	–	–	Yes	Yes		Stüken et al. (2011)
		CS-395	Yes (Stüken et al., 2011)	PCR	–	–	Yes	Yes		Stüken et al. (2011)
Karenia	*brevis*	**Wilson**	No	EST	MMETSP0201	89,216	Yes (2)	n.d.	n.d.	This study
		Wilson			MMETSP0640	83,137	Yes (1)	n.d.	n.d.	This study
		Sp1			MMETSP0574	99,942	Yes (3)	n.d.	n.d.	This study
		Sp3			MMETSP0528	82,936	Yes (2)	n.d.	n.d.	This study
Karlodinium	*veneficum*	RCC2539	No	PCR	–	–	n.d.	n.d.		Orr et al., 2012
Karlodinium	*micrum*	**CCMP2283**	No	EST	MMETSP1016	70,296	Yes (1)	n.d.	n.d.	This study
Lepidodinium	*chlorophorum*	RCC2537	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
Togula	*jolla*	CCCM725	No	EST	MMETSP0224	47,727	n.d.	n.d.	n.d.	This study
Noctilucales										
Noctiluca	*scintillans*	Unknown	No	EST	MMETSP0253	45,249	n.d.	n.d.	n.d.	This study
Oxyrrhinaceae										
Oxyrrhis	*marina*	Unknown	No	EST	MMETSP0469	51,092	n.d.	n.d.	n.d.	This study
Peridiniales										
Adenoides	*eludens*	CCMP1891	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
Azadinium	*spinosum*	RCC2538	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
		3D9	No	EST	MMETSP1036	75,458	Yes (1)	n.d.	n.d.	This study
Brandtodinium	*nutriculum*	**RCC3387**	No	EST	MMETSP1462	68,123	Yes (2)	n.d.	n.d.	This study
Durinskia	*baltica*	**CSIROCS-38**	No	EST	MMETSP0116	78,444	Yes (2)	n.d.	n.d.	This study
Glenodinium	*foliaceum*	**CCAP 116/3**	No	EST	MMETSP0118	80,537	n.d.	n.d.	n.d.	This study
Heterocapsa	*triquetra*	RCC2540	No	PCR	–	–	n.d.	n.d.		Orr et al. (2012)
		CCMP448	No	EST	MMETSP0448	62,735	Yes (1)	n.d.	n.d.	This study
Heterocapsa	*rotundata*	**SCCAP K-0483**	No	EST	MMETSP0503	46,210	Yes (1)	n.d.	n.d.	This study
Heterocapsa	*arctica*	**CCMP445**	No	EST	MMETSP1441	50,206	Yes (2)	n.d.	n.d.	This study
Kryptoperidinium	*foliaceum*	**CCMP1326**	No	EST	MMETSP0120	93,725	Yes (1)	n.d.	n.d.	This study
Peridinium	*aciculiferum*	**PAER-2**	No	EST	MMETSP0370	64,738	Yes (1)	n.d.	n.d.	This study
Pentapharsodinium	*dalei*	SCCAP K-1100	No	PCR	–	–	n.d.	n.d.	n.d.	Orr et al. (2012)
Scrippsiella	*trochoideae*	BS-46	No	PCR	–	–	n.d.	n.d.	n.d.	Orr et al., 2012
		CCMP3099		EST	MMETSP0270	106,989	Yes (1)	n.d.		This study
Prorocentrales										
Prorocentrum	*lima*	CS-869	No	PCR	–	–	n.d.	n.d.		Orr et al., 2012
Prorocentrum	*micans*	UIO292	No	PCR	–	–	n.d.	n.d.		Orr et al., 2012
Prorocentrum	*minimum*	UIO085	No	PCR	–	–	n.d.	n.d.	n.d.	Orr et al., 2012
		CCMP1329		EST	MMETSP0053	80,069	Yes (1)	n.d.		This study
Suessiales										
Pelagodinium	*beii*	**RCC1491**	No	EST	MMETSP1338	55,559	Yes (1)	n.d.	n.d.	This study
Polarella	*glacialis*	**CCMP2088**	No	PCR	–	–	n.d.	n.d.	n.d.	Orr et al., 2012
				EST	MMETSP1440	45,295	Yes (1)	n.d.		This study
Polarella	*glacialis*	**CCMP1383**	No	EST	MMETSP2227	74,437	Yes (1)	n.d.	n.d.	This study
Symbiodinium	sp.	**CCMP2430**	No	EST	MMETSP0115	47,757	Yes (1)	n.d.	n.d.	This study
			No	EST	MMETSP1122	47,710	Yes (1)	n.d.	n.d.	This study

(salinity 33–35) at 15–18 °C, with a 14:10 or 12:12 Light:Dark cycle, and an irradiance of 60–150 μmol photons m^{-2} s^{-1}.

2.2. DNA extraction, PCR amplification, cloning and DNA sequencing

Alexandrium cells were harvested by centrifugation for 15 min at 4 °C at 3220*g* (Eppendorf 5415R, Germany). DNA was extracted using a modified CTAB- phenol- chloroform method (Jørgensen et al., 2004). The DNA quality and concentration was determined using a Nanodrop (Thermoscientific, Australia), amplifying rRNA genes and/or visualising on an agarose gel.

Partial sequences of the rRNA genes LSU and SSU and complete 5.8s/ITS genes were amplified using previously published primers: SS3, SS5 (Medlin et al., 1988), D1R, D3b (Scholin et al., 1994), and ITSfor, ITSrev (Murray et al., 2012a,b). For the amplification of *sxtA* genes, *sxtA1* and *sxtA4* domains were amplified using the primers sxt001, sxt002 (A1) and sxt007 and sxt008 (A4) (Stüken et al., 2011). PCR reactions (final volume 25 μL) consisted of 20–100 ng genomic DNA, 0.5 μM of forward and reverse primer, BSA at a final concentration of 1 μg/μL and 12.5 μL of Immomix™ (Bioline, Australia), which contains Reaction Buffer, 3 mM MgCl2, polymerase and additives needed to stabilize the polymerase in the reaction mix. The reactions for SSU, LSU and ITS rDNA were carried out using the following programme: 95 °C for 10 min, 31× (95 °C for 40 s, 55 °C for 40 s, 72 °C for 1 min), 72 °C for 7 min with the primers described in Table 2, using the Verity® Thermal Cycler (Life Technologies, Australia).

The PCR reactions to amplify the *sxtA1* and *sxtA4* genes were carried out in a final volume of 25 μL. The reaction mix consisted of 20–100 ng of DNA, and 0.2 mM dNTPs, 0.5 μM of each forward and reverse primer, 0.25 units of 50× GC2 Advantage Polymerase Mix (BD Biosciences) consisting of 0.25% Glycerol, 0.075 mM Tris–Glycerol (ph-8.0), 0.375 mM KCl, 0.25 μM EDTA and 1 unit of Advantage GC Buffer (BD Biosciences, Australia), which includes 40 mM Tricine–KOH, 5 mM KOAc1, 3.5 mM Mg(OAc)2, 5% DMSO, 3.75 μg/mL BSA, 0.005%, Nonidet P-40, 0.005% Tween-20, and 1 unit of 5 M GC Melt solution, which aids in the successful amplification of GC-rich dinoflagellate sequences. Alternatively, 20 μL PCR reactions were carried out using HotMasterTaq® (5Prime, Hamburg, Germany) buffer 1×, 0.1 mM of dNTPs, 0.1 mM of each forward and reverse primer, 0.75 units of Taq polymerase, and 10–30 ng of genomic DNA. *sxtA1* was amplified with the following programme: 94 °C for 3 min, 31× (94 °C for 30 s, 60 °C for 30 s, 68 °C for 1 min), 68 °C for 3 min. *sxtA4* was amplified as follows: 94 °C for 3 min, 31× (94 °C for 30 s, 57 °C for 30 s, 68 °C for 1 min), 68 °C for 3 min. DNA was visualised on an agarose gel with GelRedTM (Biotium, Australia). Fragments were excised and purified (Zymogen Gel extraction Kit).

Purified PCR products were ligated into the pCR®4-TOPO® vector with the TOPO®TA cloning kit (Invitrogen, Darmstadt, Germany) according to the manufacturer's instructions. 4 μL of the purified PCR product was mixed with 2 μL 5× Advantage GC Buffer (BD Biosciences), 0.2 mM dNTPs and 0.1 μL of 50 × GC2 Advantage Polymerase Mix (BD Biosciences) and incubated at 68 °C for 5 min to carry out poly-A tailing of the PCR product. Poly-A tailed PCR products were stored at −20 °C till required for transformation. For the transformation, 2 μL of pCR®4-TOPO® vector ligation mix was mixed with 25 μL of One Shot® TOP10 chemically *E. coli* cells (Invitrogen, Darmstadt, Germany). Cells were incubated at 45 °C then mixed with 250 μL Super Optimal Broth (SOB, Invitrogen, Darmstadt, Germany). Cells were incubated at 37 °C with shaking at 250 rpm for an hour, plated on pre-warmed Lysogeny Broth agar plates and incubated at 37 °C overnight. Colonies were exised and M13 PCR was performed according to the manufacturer's instructions. PCR reactions were cleaned using SureClean Plus (Bioline, Australia) and checked on a 1% agarose gels. Reactions were sequenced by Sanger sequencing at Macrogen, Korea, AWI or at Ramaciotti Centre for Genomics, Australia. All sequences were submitted to GenBank (Supplementary Table 1).

2.3. Assembly and analyses of EST libraries

EST libraries from thirty-six species (forty-six libraries) of dinoflagellates were investigated for the presence or absence of *sxtA1*, *sxtA4* or *sxtG* contigs. These consisted of every publicly available library of a species of *Alexandrium,* in particular those that were part of the Marine Microbial Eukaryote Transcriptome Sequencing Program (MMETSP) (http://marinemicroeukaryotes.org/), and representative libraries from other dinoflagellate families. We limited our searches to only the most comprehensive publicly available transcriptome libraries, which we judged as those with a sequence depth of a least 40,000 unique contigs after assembly, as we wanted the highest possible likelihood that an inability to find an *sxtA* domain was not due to limitations in EST sequencing methods or library completeness (Table 1).

Custom BLAST searches were performed in Geneious software (Kearse et al., 2012). *Alexandrium* transcriptomes were queried against a local database of *sxtA1*, *sxtA4* (JA343240-JF343356) (Stüken et al., 2011) and *sxtG* (JX995111-JX995130) (Orr et al., 2013) sequences with an e-value cutoff of 10^{-3}. A similarity of 90% with a minimum hit length of 250 bp or more was taken as a positive hit for both *sxtA1* or *sxtA4*. A conserved domain database search (Marchler-Bauer et al., 2011) on NCBI website was conducted on extracted positive sequences, to confirm that the sequences have a predicted functionality of a SAM-dependent methyltransferase (*sxtA1*) or class II aminotransferase (*sxtA4*). Due to high similarity between *sxtG* and other amidinotransferases in EST libraries, phylogenetic analysis was performed to differentiate between *sxtG* and other amidinotransferases.

In addition, thirteen EST libraries of other alveolates: nine species of Ciliata, two Perkinozoa, one Chromerida and one Apicomplexa, were investigated for the presence or absence of *sxtA1*, *sxtA4* or *sxtG* using the method described above (Supplementary Table 2).

2.4. Phylogenetic analyses

The three-rDNA genes (18S, 5.8S, and 28S) were separately aligned using the MAFFTv7 Q-INS-I model (Katoh and Toh, 2008), considering secondary RNA structure (default parameters used). The five protein coding datasets (actin, beta-tubulin, cob, cox1, and hsp90) were separately aligned at the nucleotide level based on the corresponding amino acid alignment, as to maintain codon integrity, inferred with MAFFTv7 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 were established from previous dinoflagellate and protist phylogenies (Murray et al., 2005; John et al., 2003; Shalchian-Tabrizi et al., 2006; Orr et al., 2013). All single gene alignments were checked manually using Mesquite v3.0 (Maddison and Maddison, 2011). The eight separate alignments were then checked with Gblocks v0.91b (Castresana, 2000), 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 a supermatrix. For the supermatrix (concat) in-group taxa (dinoflagellates) required rDNA sequence data in addition to two of the five protein coding genes. The only exception to this was species of the *Alexandrium* genus, whose inclusion was considered paramount to the studies' aim, and which were

Table 2
Percent similarity within a strain of the multiple copies of *sxtA1* and *sxtA4* domains, based on cDNA and gDNA. Strains marked in bold were sequenced in this study, other data were obtained from GenBank.

Species	Strain	*sxtA1*				*sxtA4*				Reference
		cDNA (no. sequences)	% similarity	gDNA (no. sequences)	% similarity	cDNA (no. sequences)	% similarity	gDNA (no. sequences)	% similarity	
A. fundyense	CCMP1719	19	0.983–1.000	10	0.668–0.994	12	0.962–0.994	10	0.917–1.000	Stüken et al. (2011)
	CCMP1979	–	–	2	0.974	–	–	5	0.926–0.979	Stüken et al. (2011)
	A8	–	–	–	–	7	0.935–1.000	19	0.775–1.000	This study
	E4	–	–	–	–	19	0.970–1.000	40	0.768–1.000	This study
	ACQH01	–	–	–	–	–	–	5	0.798–0.990	This study
A. pacificum	CCMP1493	–	–	5	0.711–0.994	–	–	1	–	Stüken et al. (2011)
	ACCC01	–	–	4	0.701–0.990	4	0.960–1.000	5	0.910–1.000	Stüken et al., 2011, This study
	ACTRA02	–	–	4	0.698–0.992	–	–	1	–	Stüken et al., 2011, This study
A. australiense	ATBB01	–	–	5	0.700–0.996	–	–	5	0.983–1.000	Stüken et al. (2011)
	ATCJ33	–	–	2	0.712	–	–	1	–	Stüken et al., 2011, This study
A. minutum	CCMP113	–	–	7	0.983–0.998	–	–	7	0.982–0.996	Stüken et al. (2011)
	ALSP02	–	–	–	–	–	–	6	0.958–1.000	Stüken et al. (2011)
	AMAD16	–	–	5	0.990–0.998	–	–	5	0.981–1.000	Stüken et al. (2011)
	CCMP1888	–	–	5	0.968–0.990	–	–	5	0.984–0.996	Stüken et al. (2011)
	CAWD11	–	–	3	0.707–0.977	–	–	1	–	This study
	CAWD12	–	–	3	0.739–0.930	–	–	1	–	This study

included even when only the rDNA sequence information was present (Wiens, 2005). Maximum Likelihood (ML) analyses were performed with RAxML-VI-HPCv8.0.0, GTRCAT model with 25 rate categories (Stamatakis, 2006). Bayesian inference was carried out using Phylobayes v3.3f (Lartillot et al., 2009) under the GTRCAT substitution model with a free number of mixing categories and a discrete across site variation under 4 categories. Trees were inferred when the largest maximum difference between the bipartitions (chains) was <0.1. The mitochondrial cytochrome genes, cob and cox1 were excluded for *Heterocapsa* species; the mitochondrial genes for this genus have been shown to have a high divergence rate resulting in an artificial evolutionary placement (Orr et al., 2012).

Three dinoflagellate genes *sxtA1, sxtA4* and *sxtG* were aligned with orthologous amino acid cyanobacterial sequences, apart from *sxtA4*, for which we also aligned nucleotide sequences, and a selection of closely related NCBInr BLASTP hits using MAFFTv7 L-INS-I model under the default settings. The resulting alignments were checked manually, and poorly aligned positions were excluded using Mesquite v3.0. ProtTest v2.4 (Abascal et al., 2005) determined LG as the optimal evolutionary model for all alignments. Maximum likelihood (ML) analyses were performed with the RAxML-VI-HPC v8.0.0, PROTCATLG model with 25 rate categories. The most likely topology was established from 100 separate searches, and bootstrap analyses were performed with 1000 pseudoreplicates. Bayesian inference was carried out using Phylobayes v3.3f (Lartillot et al., 2009) under the LGCAT substitution model, as previously described. All model estimation and phylogenetic analyses were done using either Lifeportal (https://lifeportal.uio.no) or the abel server at the University of Oslo.

2.5. Selection analyses

Codon-based alignments for dinoflagellates including representatives of all clades of *sxtG* and *sxtA1* were generated. Tests for branches and sites under selection were conducted using Branch-site REL (Pond et al., 2011), as implemented using the datamonkey web interface (www.datamonkey.org). Branch-site REL performs likelihood ratio tests to identify which lineages in a phylogeny have a proportion of sites that evolved with d*N*/d*S* significantly greater than 1, which would indicate instances of episodic diversifying selection at some sites in the alignment. Branch-site REL does not require assumptions about which lineages those are, and what happens to the rest of the lineages, as these assumptions could lead to false positives and loss of statistical power (Pond et al., 2011). Probability values are corrected for multiple testing using the Holm Bonferroni method (Pond et al., 2011).

3. Results and discussion

3.1. Phylogeny of dinoflagellates, including the genus Alexandrium

The phylogenetic analysis of dinoflagellates (Fig. 1) including the new gDNA sequences of SSU generated in this study, ITS1/5.8s/ITS2, partial LSU rDNA and the newly identified sequences from transcriptome libraries of actin, beta-tubulin, cytochrome b and cox1 showed that the genus *Alexandrium* was fully supported as a monophyletic clade (100%/1.00, BS/PP). Our phylogeny showed moderate support for the divergence between armoured and unarmoured dinoflagellates (57/0.82), a similar level to that found in a study, which was the first to demonstrate support for this split (Orr et al., 2012). There was support for the monophyly of the major orders of dinoflagellates, including Gonyaulacales (100/1.00), Peridiniales excluding *Heterocapsa* (58/0.86), Suessiales (100/1.00), Prorocentrales (82/0.93), while

Gymnodiniales were paraphyletic, and were sister group to the armoured dinoflagellates (Fig. 1). The most basal lineage of "core" dinoflagellates was shown to be the genus *Amphidinium*, as has been found previously (Jørgensen et al., 2004; Zhang et al., 2007; Orr et al., 2012). The species *Gymnodinium catenatum* was clearly part of the *Gymnodinium* sensu stricto clade (100/1.00).

The fully supported *Alexandrium* clade was within the order Gonyaulacales, and was sister group to a clade consisting of *Gambierdiscus* spp. and *Pyrocystis* spp, with *Pyrodinium bahamense* as sister group to that (Fig. 1). This topology was the same in the analysis on the reduced alignment, including only those species for which all genes were present (data not shown). Within *Alexandrium*, two main clades were identified, although each was only moderately supported (61/0.78, 64/0.94), which consisted of the group of species: *A. margalefi, A. leei, A. diversaporum, A. minutum, A. tamutum, A. andersoni, A. ostenfeldii,* and the group of species *A. monilatum, A. satoanum, A. taylori, A. pseudogonyaulax, A. hiranoi, A. affine, A. fraterculus, A. tropicale, A. tamiyavanichi*, as well as the *A. tamarense* species complex group (*A. fundyense, A. mediterraneum, A. tamarense, A. pacificum*, and *A. australiense* (John et al., 2014)). The *Alexandrium* species producing STXs (highlighted in Fig. 1) included some representatives from both major *Alexandrium* clades (Fig. 1), as well as the species *Pyrodinium bahamense* and *Gymnodinium catenatum*.

3.2. Presence of sxtA domains A1 and A4, and sxtG

The species of *Alexandrium* that are established to produce STXs are: *Alexandrium minutum, A. ostenfeldii, A. australiense* (these three species also include some non-STX producing strains (Yang et al., 2010; Touzet et al., 2007; Murray et al., 2012a,b; Kremp et al., 2014), *A. pacificum, A. fundyense,* and *A. tamiyavanichi.* We searched for *sxtA* domains in the following strains which were known to produce STXs: fifteen strains of species of *Alexandrium*, one strain of *Gymnodinium catenatum* and one strain of *Pyrodinium bahamense*, using PCR (14 strains) and by searching transcriptome libraries (3 strains) (Table 1). An analysis of the presence of *sxtA* in dinoflagellates from our new data and all previous reports (Table 1, data from PCRs in current study in bold), shows that all species which produced STX possessed the domain *sxtA4.* In detail, we have confirmed the presence of *sxtA*, domain *A4*, in 40 strains of 8 species of STX-producing dinoflagellates, and did not identify it in 67 strains of 54 species of non STX-producing species representing different dinoflagellate families (Table 1). These results confirmed the hypothesis that *sxtA4* is an essential domain for saxitoxin synthesis (Kellmann et al., 2008; Stüken et al., 2011; Murray et al., 2011a,b). Conversely, only those species which could produce this toxin possessed this domain, apart from one interesting anomaly: all four strains of *Alexandrium australiense* (formerly *A. tamarense*, Group V), only two of which have been found to produce STXs, possess the domain *sxtA4* (Orr et al., 2011; Murray et al., 2012a,b). The finding of *sxtA4* in one strain (CCMP1771) of *A. tamarense*, (former *A. tamarense*, Group III, John et al., 2014), in a previous study (Stüken et al., 2011) has not been confirmed in this study, or a previous study (John et al., 2014). A search of two comprehensive transcriptomic libraries of CCMP1771, failed to find a sequence of *sxtA*, domain 4 (Table 1). Similarly, a study of a different strain of *Alexandrium tamarense* (formerly Group III) failed to find this sequence (Hackett et al., 2013). *Alexandrium tamarense* (*sensu* John et al., 2014) is not known to produce STXs, and so this finding is in line with the pattern we have observed.

Eight strains of STX-producing *Alexandrium fundyense* (formerly, Group 1) possess both the domain *sxtA1* and the domain *sxtA4*. Similarly, these two domains were found in eleven strains of STX producing *Alexandrium pacificum* (formerly, *A. catenella*, Group IV, John et al., 2014); nine strains of STX-producing *Alexandrium minutum*, four strains of *Alexandrium australiense*; four strains of STX producing *Gymnodinium catenatum*, two strains of *Alexandrium ostenfeldii*, and a single strain each of *Alexandrium tamiyavanichi* and *Pyrodinium bahamense* (Table 1, Stüken et al., 2011; Murray et al., 2012a,b; John et al., 2014; Hackett et al., 2013; Suikkanen et al., 2013).

The species of *Alexandrium* examined, which lacked the domain *sxtA4*, were: *Alexandrium mediterraneum* (formerly *A. tamarense* Group II), *A. tamarense* (formerly Group III), *A. affine, A. andersoni* (CCMP2222), *A. concavum, A. diversaporum, A. fraterculus, A. insuetum, A. margalefi, A. monilatum, A. pseudogonyaulax* (Table 1). This indicates that these strains may lack the ability to synthesise STXs. It has been reported that a strain of *A. affine* can produce low levels of STX (Nguyen Ngoc, 2004), while other studies have found that strains do not produce STXs (Hallegraeff et al., 1991; Band-Schmidt et al., 2003; Stüken et al., 2011; Wang et al., 2006). A strain of *Alexandrium andersoni* (CCMP2222) was reported to produce low levels of STXs (Ciminiello et al., 2000; Frangopulos et al., 2004), while other studies based on CCMP2222 and other strains of *A. andersoni* have not detected STXs (Sampedro et al., 2013; Stüken et al., 2011; Orr et al., 2011; Touzet et al., 2008). It is possible that *sxtA* related genes are no longer expressed in CCMP2222, or that early reports of STX production were in error.

We searched for *sxtA* domains in EST libraries from 28 additional dinoflagellates, which were not species of the genus *Alexandrium* and non-STX producers. We found that *sxtA*, domain *A4*, was not present in the transcriptomes of these species (Table 1). However, we found a paralog of the domain *sxtA1* to be present in the majority of the non-toxic dinoflagellate species, including in non-STX producing species of *Alexandrium* (Table 1).

For *sxtG*, we searched the all EST libraries of all dinoflagellate strains available (Table 1). We found *sxtG* was present in the transcriptome of the three strains of STX producing dinoflagellates: *A. fundyense, A. pacificum* CCMP1598, *Pyrodinium bahamense* (Table 1). It was not found in the transcriptomes of the non-STX producing species of *Alexandrium* or other dinoflagellates (Table 1). This result is in contrast with Orr et al., 2013, in which *sxtG* was found in 3 non-STX producing *Alexandrium* species, as well as 7 species of STX-producing *Alexandrium*.

We searched EST libraries from nine ciliates, two Perkinozoa, one Chromerida and one Apicomplexa and found that *sxtA*, domain A1 and A4 and *sxtG* were not present in the transcriptome of these alveolate species (Supplementary Table 2).

3.3. Intraspecific variation in sxtA1 and sxtA4 domains

Approximately 100–378 genomic copies of *sxtA4* gene cell^{-1} have been reported for three strains of *Alexandrium pacificum* (Stüken et al., 2011; Murray et al., 2011b), which appears to be an average number of copies for a moderately expressed gene in a dinoflagellate (i.e. Le et al., 1997; Bachvaroff and Place, 2008). Expressed transcripts of *sxtA4* have been reported to differ from one another in *A. fundyense* and *A. minutum* (Stüken et al., 2011We examined intragenomic diversity differences within the sequences of cDNA as compared to gDNA for three strains of *A. fundyense* and one strain of *A. pacificum* for *sxtA4* and one strain of *A. fundyense* for *sxtA1* (Table 2). Transcription appears to limit the diversity of copies of the *sxtA4* and *sxtA1* domains (Table 2), as the overall% similarity was consistently higher for cDNA copies than gDNA copies, indicating that only certain gDNA copies are transcribed. It is likely that some of the gDNA copies are pseudogenes, as a deletion of 63 bases was found in some of the gDNA sequences from *A. fundyense* strains E4 and A8, whereas no gap was observed in the corresponding cDNA copies from the same strains. No deletion was found in sequences from CCMP1719, of

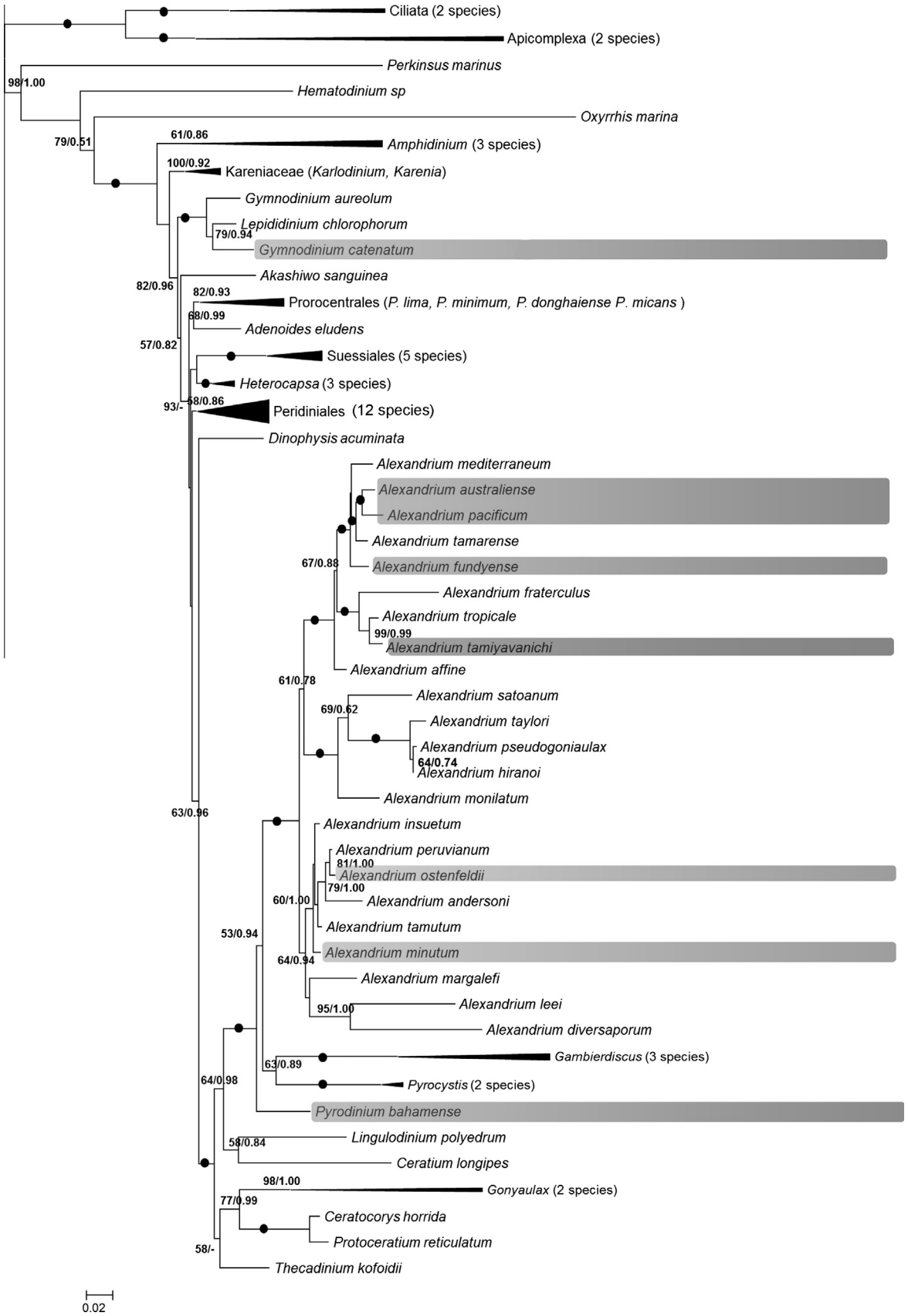

Fig. 1. Phylogeny of dinoflagellates, including the genus *Alexandrium*, based on a concatenated alignment of 8 genes (7308 bp). Support values at nodes represent ML bootstrap values based on 1000 replicates and Bayesian Posterior probabilities (BS/PP). Boxes highlighted show species which are known to produce PSTs. Scale bar indicates the number of substitutions per site.

A. fundyense, suggesting a further strain-specific variability within this species.

In a comparison of the cDNA sequences with the gDNA sequences of each strain, and there were certain base positions for which the bases were consistently different for the cDNA sequences, as compared to the gDNA sequences. In strain E4, there were 6 sites at which different bases were found in almost all of the transcribed sequences as compared to the gDNA, including a gap, which resulted in a frame shift in the protein (data not shown). In *A. fundyense* strain CCMP1719, three sites were identified which differed consistently, apart from in a single gDNA. It has been reported that in *A. pacificum* strain ACCC01, there were 19 sites at which differences were apparent between almost all cDNA clones, and these correlated with one of the gDNA clones (Wiese, 2012). This is similar to the finding of an analysis of gDNA and cDNA copies of actin in *Amphidinium carterae* Hulburt, in which some pseudogenes were found, and transcriptional and post-transcriptional processes appeared to reduce diversity in actin copies (Bachvaroff and Place, 2008). Therefore it is likely that transcriptional and post-transcriptional processes also serve to limit the diversity of *sxtA* sequences in *Alexandrium*. However, this pattern was not universal, as no consistent differences were observed in *A. fundyense* strain A8 between the gDNA and cDNA clones. The presence of multiple, slightly different genomic copies of *sxtA4* may provide adaptive plasticity of strains and may partly explain the broad phenotypic diversity in terms of STX profile and toxicity within *Alexandrium* populations (Alpermann et al., 2010; Kremp et al., 2012). Additional information on the *sxtA* genes in different species of STX producing *Alexandrium* species may lead to a greater clarification of the relationship between toxicity and *sxtA* gene diversity, copy number and expression in *Alexandrium*.

3.4. Phylogeny of sxtA1, sxtA4, and sxtG

Paralogs of *sxtA4* in every dinoflagellate STX producing species were found to form a monophyletic clade with full support (100/1.00, Fig. 2A). The dinoflagellate clade was highly conserved, and the sister group to cyanobacterial *sxtA4* (Fig. 2A). In the analysis based on the nucleotide alignment, (Fig. 2B) the species *Alexandrium minutum*, *A. australiense*, *A. ostenfeldii*, *Gymnodinium catenatum,* and *Pyrodinium bahamense* appeared to have largely distinctive *sxtA4* copies, which did not group with copies from any other species (Fig. 2B). There was little evidence of widespread horizontal transfer amongst all species of *Alexandrium* (Fig. 2B). Within closely related species, for example those of the *A. tamarense* species complex, *A. pacificum* and *A. fundyense*, sequences from different clonal strains from these two species formed a single clade, albeit with little or no support, indicating either a common evolutionary history or similar functional constraints (Fig. 2B). One possibility is that, as these two species have diverged comparatively recently (Fig. 1), their *sxtA* copies may not be sufficiently distinct as yet. One sequence apparently from *A. minutum* was included in this clade, which, if confirmed, might suggest that there may be some genetic exchange amongst species. We have no models how this might functionally have been happened amongst species within the genus. Therefore, a more likely evolutionary scenario may argue that different mutation rates within and amongst species has resulted in this unusual placement (Fig. 2B).

The phylogeny of *sxtA1* (Fig. 3) shows that three paralogs with moderate to high support were present in dinoflagellates. One paralog, clade 2, formed a highly supported clade (98/1.00) that appeared to be made up of species that produce STXs: *A. minutum, A. pacificum, A. australiense*, *A. fundyense*, *Pyrodinium bahamense* and *Gymnodinium catenatum*, as well as two genomic sequences from the non-STX producing species *A. margalefi* and *A. pseudogonyaulax* (Fig. 3). In the EST library of the same strain of *A. margalefi* investigated in this study (Table 1), the corresponding cDNA sequence similar to this *sxtA1* gDNA copy was not found, suggesting that it might be a pseudogene or may not be highly expressed, and therefore not detected.

A further paralog, clade 3, was made up of sequences from almost every dinoflagellate family and order, including Gymnodiniales (*Karenia*), Peridiniales (*Scrippsiella, Heterocapsa*), Suessiales (*Symbiodinium, Pelagodinium*), and Gonyaulacales (*Lingulodinium, Gonyaulax, Protoceratium, Gambierdiscus*). The structure of this group overall was not well supported (–/0.84). However some clades showed high support, such as the paralogs from *Heterocapsa* species, which grouped together with high support (99/1.00), and the second clade from *Alexandrium* species, which also grouped together with high support (96/1.00; Fig. 3).

Finally, clade 1 of *sxtA1* (Fig. 3) was composed of sequences from *Alexandrium* species, which are either known to not produce STXs, or are known to have toxic and non-toxic strains (*A. minutum*, e.g. Yang et al., 2010; Touzet et al., 2007). As sequences of this clade were not recovered from almost all STX producing strains, it is unlikely that these sequences are necessary for STX production, suggesting that this clade may be comprised of a pseudogene or paralog with a separate function to STX production.

The phylogeny of *sxtG* (Fig. 4) shows that one clade (*sxtG*) was highly conserved, and well supported, and included all of the species known to produce STX, including *Gymnodinium catenatum, Pyrodinium bahamense, A. minutum, A. pacificum, A. fundyense* and *A. australiense*. This clade includes sequences from gDNA for *A. andersoni, A. insuetum* and *A. affine* from a previous study (Fig. 4, Orr et al., 2013). In contrast, in this study, *sxtG* was not found in any of the transcriptomes of species not known to produce STX (Table 1). This gene appears to be related to a bacterial clade including the cyanobacterial *sxtG*. This gene is consequently interpreted to be the clade involved in STX biosynthesis. However, as expected, many similar amidinotransferase enzymes, a common functional enzyme domain, were also found throughout the dinoflagellates (clades 1, 2, 3). These were in separate, generally well-supported clades (84/1.00, 100/1.00, 93/0.99, for clades 1, 2, 3 respectively), which included enzymes from many bacterial species as sister groups, and the amidinotransferase enzymes from cylindrospermopsin producing cyanobacteria, *cyrA*. These paralogs were widely presented in dinoflagellates examined, including species of every dinoflagellate order, as well as closely related or basal species such as *Oxyrrhis marina* and *Perkinsus marinus,* pointing to a more common presented function for these genes within the dinoflagellates.

3.5. The evolution of sxt genes within the dinoflagellates

The phylogenies of *sxtA1* and *sxtG* did not show a clear similarity in topology with the species phylogeny of those dinoflagellates possessing these paralogs, indicating that the evolution of these genes has been complex. The possession of multiple paralogs of *sxtA1* and *sxtG* adds to the complexity of the interpretation of these phylogenies. However, within the phylogeny based on nucleotides of *sxtA4*, clade 1 of *sxtA1*, as well as in *sxtG*, the copies present in *Pyrodinium bahamense* and *Gymnodinium catenatum* appear distinctive when compared with the copies in *Alexandrium* species, and occur in highly supported clades (Figs. 2 and 4). In particular, although *sxtA4* was highly conserved, the phylogeny of *sxtA4* based on the nucleotide alignment (Fig. 2) appears most similar to the species phylogeny of these taxa, with *Gymnodinium catenatum* as the fully supported (100/1.00) outgroup. This indicates that the gene may have evolved in a common dinoflagellate ancestor of *Gymnodinium catenatum, Pyrodinium bahamense* and *Alexandrium* spp. the radiation within the *Alexandrium* species occurring later, $\geqslant$77 MYA (John et al., 2003). This would suggest that a secondary

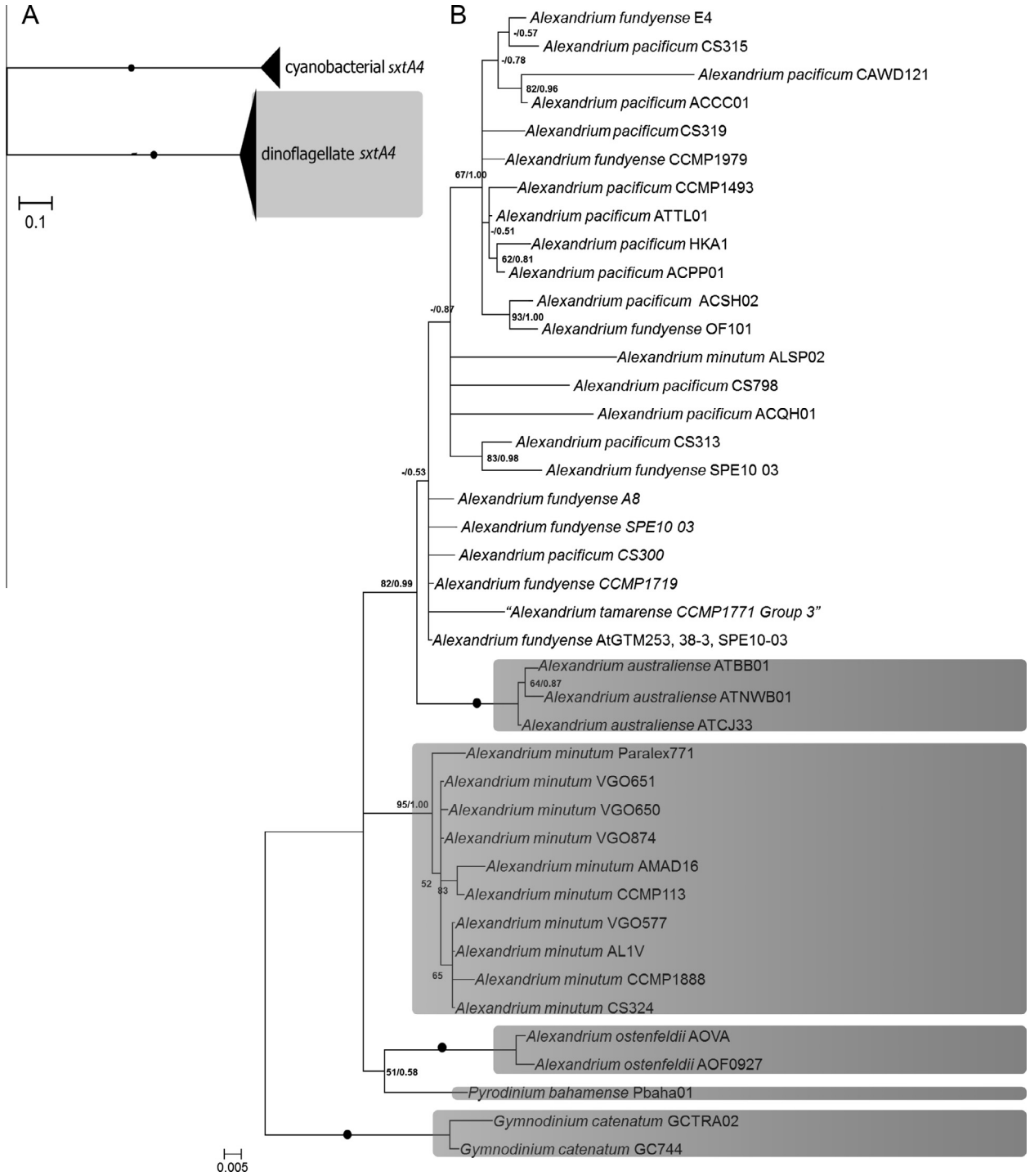

Fig. 2. Phylogeny of *sxtA4*. (A) Based on an alignment of proteins (221 amino acids). The highlighted box shows the single clade found in dinoflagellates, which was highly conserved in all species. (B) Based on an alignment of nucleotides (663 bp). The highlighted boxes show that *sxtA4* copies in multiple strains of species tended to be more closely related to each other. Support values at nodes represent ML bootstrap values based on 1000 replicates and Bayesian Posterior probabilities. Scale bar indicates the number of substitutions per site.

loss of those genes has occurred in the non-toxic *Alexandrium* species. The most parsimonious explanation for the origin of STX in dinoflagellates is that a single horizontal transfer event occurred early in dinoflagellate evolution. The orders Peridiniales and Gonyaulacales have an origin as far back as $\geqslant$190–180 MYA (John et al., 2003). If the origin of these genes is this ancient, it may explain the current patchy distribution of the genes within the genome, the high number of gene duplication events in *sxtA1*, and the high number of paralogs and pseudogenes. Similarly, an ancient origin of the horizontal transfer of these genes from bacterial hosts may explain why multiple horizontal transfers appear to have occurred in the evolution of amidinotransferase genes and *sxtG* in dinoflagellates (Fig. 4).

The peculiarities of the dinoflagellate genomes might also partly explain the high number of copies of each gene within a species, and the complexity of their evolutionary histories (Lin, 2011;

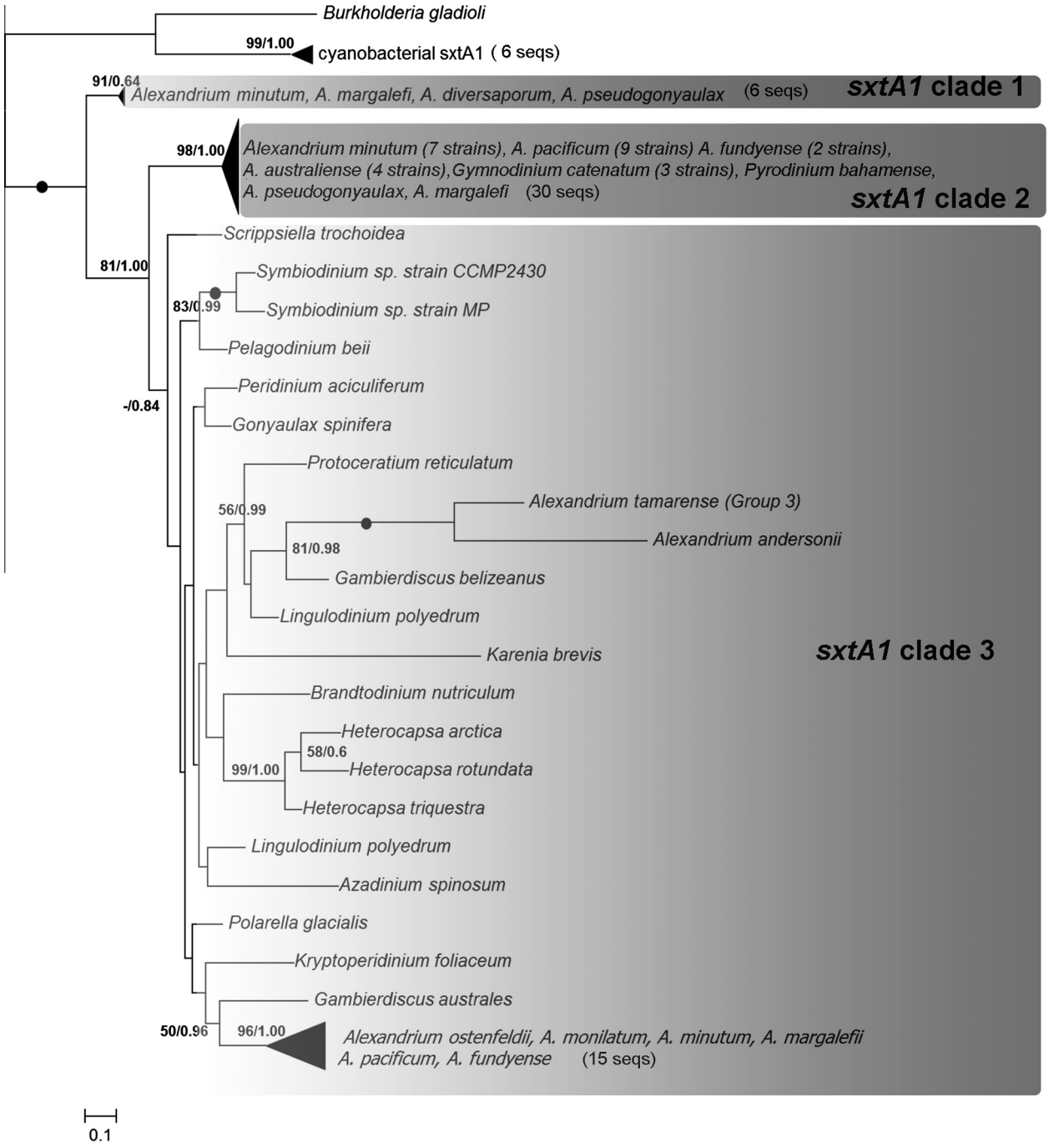

Fig. 3. Phylogeny of *sxtA1*. Based on an alignment of proteins (179 amino acids).Support values at nodes represent ML bootstrap values based on 1000 replicates and Bayesian Posterior probabilities. Boxes highlighted show species. Scale bar indicates the number of substitutions per site.

Jaeckisch et al., 2011). For example, as we have shown here, gene duplication and loss appears to be very common, and genes are presented in many copies in the genome, including pseudogenes, due to processes such as the "recycling" of processed cDNAs, in which they are reverse transcribed into the genome (Zhang et al., 2007; Slamovits and Keeling, 2008). The selfish operon model (Lawrence, 1999), postulated that gene clusters encoding a particular function were likely to stay intact over time, due to common selective processes operating on their function. However, this does not appear to fit the evolution of dinoflagellate genomes. Although all genes of the *sxt* cluster putatively necessary for STX biosynthesis have now been found in dinoflagellates (Stüken et al., 2011; Hackett et al., 2013), the potential signatures for an ancient horizontal gene transfer from cyanobacteria to dinoflagellates have only been detected for *sxtA* and *sxtG* so far (Stüken et al., 2011; Orr et al., 2013; Hackett et al., 2013). The other orthologs of the STX gene cluster appear to have lost their original signature over time, which was indicated by their closest sister clades being other bacterial and eukaryotic organisms, not necessarily the cyanobacterial *sxt* analogs, in analyses of those genes (Stüken et al., 2011; Hackett et al., 2013). The original genes may have become exchanged with other, functionally equivalent dinoflagellate genes over evolutionary time, and the genes subsequently lost. Another example in dinoflagellates in which a previously large gene cluster

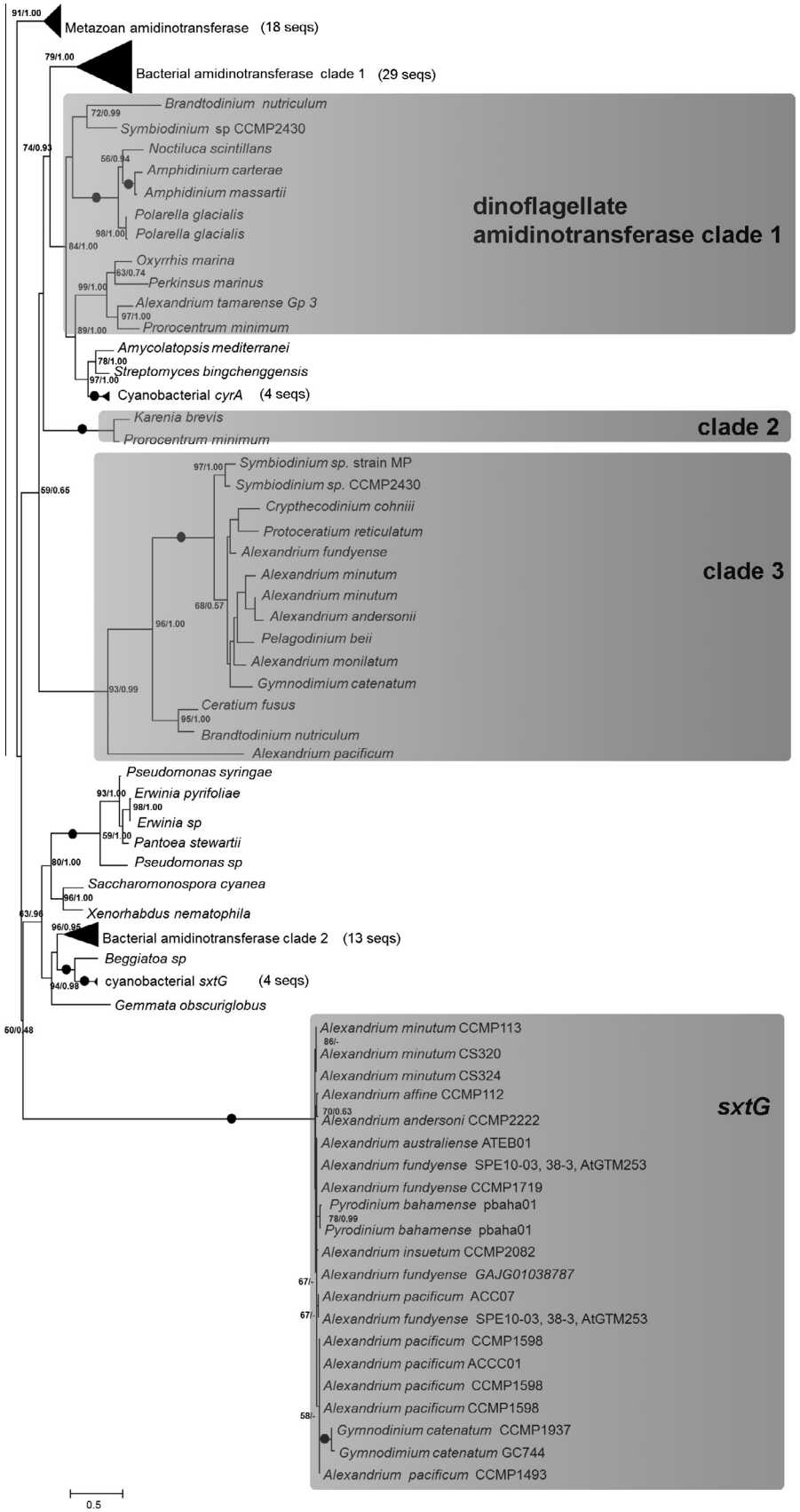

Fig. 4. Phylogeny of *sxtG*. Based on an alignment of proteins (232 amino acids).Support values at nodes represent ML bootstrap values based on 1000 replicates and Bayesian Posterior probabilities. Boxes highlighted show species. Scale bar indicates the number of substitutions per site.

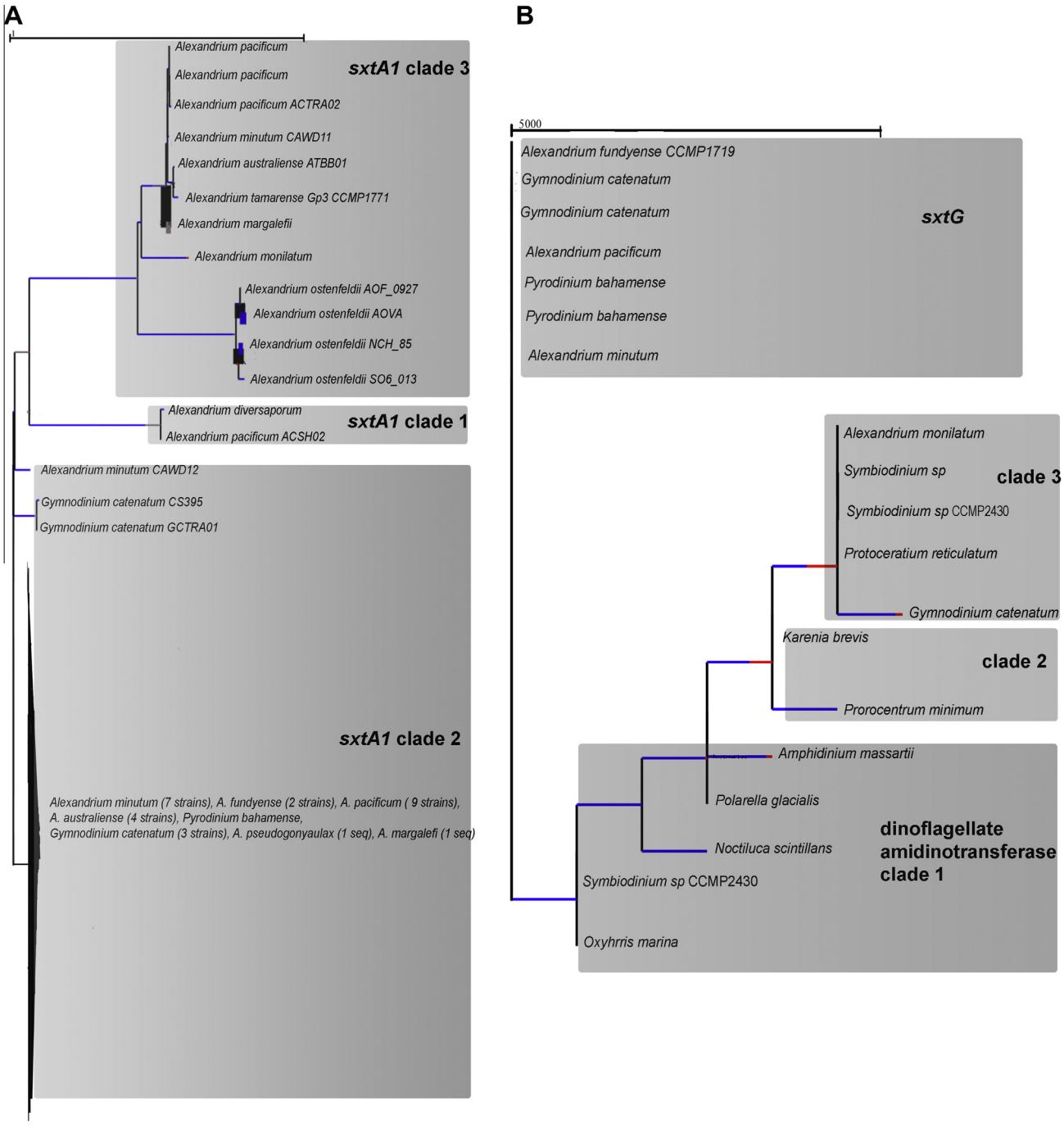

Fig. 5. Analyses of selection on branches in the phylogenies which had more than one paralog of an *sxt* gene; (A) *sxtA1*, and (B) *sxtG*. The colour of the branch indicates the strength of selection, with red corresponding to d*N*/d*S* > 5 (positive selection), blue to d*N*/d*S* = 0 and grey to d*N*/d*S* = 1 (neutral selection). The width of each colour component represents the proportion of sites in the corresponding class. Thicker branches have been classified as undergoing episodic diversifying selection by the sequential likelihood ratio test at corrected $p \leqslant 0.05$.

encoding a single multidomain enzyme has become separated into domains with many copies, rather than remaining as an operon, are polyketide synthesis (PKS) type I genes, which apparently have also experienced independent gene duplication and shuffling processes (John et al., 2008; Eichholz et al., 2012; Meyer et al., 2015).

3.6. Selection analyses of sxtG and sxtA1

In order to better understand the evolutionary constraints responsible for the demonstrated tree topologies, we examined the potential selective forces operating on the different clades. The two gene families for which more than one paralog was identified in dinoflagellates, *sxtA1* and *sxtG*, were examined to determine the role of selection following gene duplication, in the case of *sxtA1*, or following potential different horizontal transfer events, using the branch-site REL method. The aim of this method was to identify which lineages in the phylogenies had a proportion of sites that evolved with d*N*/d*S* significantly greater than 1, which would indicate instances of episodic diversifying selection at some sites in the alignment. Four branches (Fig. 5A, thick lines, $p < 0.0001$), which represented lineages of *sxtA1*, clade 3, in *Alexandrium ostenfeldii,* and *Alexandrium margalefi,* showed evidence of episodic

positive selection. The clade putatively involved in STX biosynthesis, *sxtA1*, clade 2 (Fig. 5A, Fig. 3), did not show branches with significant positive selection, but the branches were instead under negative (purifying) selection, indicating pressure to constrain function. In the analysis of *sxtG*, no branches showed significant episodic positive selection (Fig. 5B). The areas in the tree in which d*N*/d*S* were highest were in dinoflagellate amidinotransferase clades 2 and 3 (Fig. 5B, red lines, indicating d*N*/d*S* > 5). Again, the clade putatively involved in STX biosynthesis, *sxtG*, did not show branches with significant positive selection, but the branches were instead under negative selection, indicating functional constraint also for this gene. Therefore, for *stxA1* and similarly *stxG*, we have shown the presence of several paralogs in different, well-supported clades, indicating that they may perform separate functions, and supported by instances of positive selection, which may have enhanced the development of differing functions. In contrast, those clades putatively involved in STX biosynthesis were experiencing negative selection, and therefore appeared to be functional conserved.

4. Conclusions

An understanding of the evolution of a marine biotoxin with major ecological impacts can yield important information on eukaryotic evolutionary processes. We have investigated the vertical and/or lateral inheritance of genes involved in STX biosynthesis in the marine dinoflagellates, in comparison to the evolution of the species producing STX. We found that gene duplication followed by loss and stronger functional constraints has played a major role in the evolution of *sxtA*, domain A1, and *sxtG*. The domain *sxtA4*, however, was highly conserved, had no paralogs, and was restricted to those species which can produce STX.

Acknowledgments

This study was funded by the Australian Research Council Grant DP120103199 to SM and UJ. This is contribution number xx from the Sydney Institute of Marine Science. Additional financial support to U.J. was provided by the PACES research program of the Alfred-Wegener-Institute Helmholtz-Zentrum für Polar- und Meeresforschung and to UJ and SM via the PT-DLR IB (Australian Academy of Science German Australian mobility call project 01DR14006). We thank the Cawthron Institute and Gustaaf Hallegraeff for the provision of cultures.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.ympev.2015.06.017.

References

Abascal, F., Zardoya, R., Posada, D., 2005. ProtTest: selection of best-fit models of protein evolution. Bioinformatics 21, 2104–2105.
Alpermann, T.J., Tillmann, U., Beszteri, B., Cembella, A.D., John, U., 2010. Phenotypic variation and genotypic diversity in a planktonic population of the toxigenic marine dinoflagellate *Alexandrium tamarense* (Dinophyceae). J. Phycol. 46, 18–32.
Anderson, D.M., Alpermann, T.J., Cembella, A.D., Collos, Y., Masseret, E., Montresor, M., 2012. The globally distributed genus *Alexandrium*: multifaceted roles in marine ecosystems and impacts on human health. Harmful Algae 14, 10–35.
Bachvaroff, T.R., Place, A.R., 2008. From stop to start: tandem gene arrangement, copy number and trans-splicing sites in the dinoflagellate *Amphidinium carterae*. PLoS ONE 3 (8), e2929.
Balech, E., 1995. The Genus *Alexandrium* Halim (Dinoflagellata). Shenkin Island Press.
Band-Schmidt, C.J., Lilly, E.L., Anderson, D.M., 2003. Identification of *Alexandrium affine* and *A. margalefi* (Dinophyceae) using DNA sequencing and LSU rDNA-based RFLP-PCR assays. Phycologia 42, 261–268.
Brown, L., Bresnan, E., Graham, J., Lacaze, J.-P., Turrell, E., Collins, C., 2010. Distribution, diversity and toxin composition of the genus *Alexandrium* (Dinophyceae) in Scottish waters. Eur. J. Phycol. 45, 375–393.
Castresana, J., 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17 (4), 540–552.
Cembella, A., Destombe, C., 1996. Genetic differentiation among *Alexandrium* populations from eastern Canada. In: Harmful and Toxic Algal Blooms. I. O. C. UNESCO, Paris, pp. 447–450.
Ciminiello, P., Fattorusso, E., Forino, M., Montresor, M., 2000. Saxitoxin and neosaxitoxin as toxic principles of *Alexandrium andersonii*(Dinophyceae) from the Gulf of Naples, Italy. Toxicon 38 (12), 1871–1877.
Dantzer, W., Levin, R., 1997. Bacterial influence on the production of paralytic shellfish toxins by dinoflagellated algae. J. Appl. Microbiol. 83, 464–469.
Doblin, M.A., Blackburn, S.I., Hallegraeff, G.M., 1999. Comparative study of selenium requirements of three phytoplankton species: *Gymnodinium catenatum*, *Alexandrium minutum* (Dinophyta) and *Chaetoceros cf. tenuissimus* (Bacillariophyta). J. Plankton Res. 21, 1153–1169.
Eichholz, K., Beszteri, B., John, U., 2012. Putative monofunctional type I polyketide synthase units: a dinoflagellate-specific feature? PLoS ONE 7 (11), e48624.
Farrell, H., Brett, S., Ajani, P., Murray, S., 2013. Distribution of the genus *Alexandrium* and paralytic shellfish toxins along the coastline of New South Wales, Australia. Mar. Poll. Bull. 72, 133–145.
Fischbach, M.A., Walsh, C.T., Clardy, J., 2008. The evolution of gene collectives: how natural selection drives chemical innovation. Proc. Nat. Acad. Sci. USA 105, 4601–4608.
Frangopulos, M., Guisande, C., deBlas, E., Maneiro, I., 2004. Toxin production and competitive abilities under phosphorus limitation of *Alexandrium* species. Harmful Algae 3 (2), 131–139.
Hackett, J.D., Wisecaver, J.H., Brosnahan, M.L., Kulis, D.M., Anderson, D.M., Bhattacharya, D., Plumley, F.G., Erdner, D.L., 2013. Evolution of saxitoxin synthesis in cyanobacteria and dinoflagellates. Mol. Biol. Evol. 30, 70–78.
Hallegraeff, G., Bolch, C., Blackburn, S., Oshima, Y., 1991. Species of the toxigenic dinoflagellate genus *Alexandrium* in southeastern Australian waters. Bot. Mar. 34 (6), 575–588.
Harada, T., Oshima, Y., Yasumoto, T., 1982. Structures of two paralytic shellfish toxins, gonyautoxins V and VI, isolated from a tropical dinoflagellate, *Pyrodinium bahamense* var. compressa. Agric. Biol. Chem. 46, 1861–1864.
Hii, K.S., Lim, P.-T., Tan, T.-H., Leaw, C.-P., 2012. Characterization of the saxitoxin biosynthetic starting gene, sxtA in the toxic dinoflagellate *Alexandrium tamiyavanichi*. In: The 12th Symposium of the Malaysian Society of Applied Biology: Solutions to Global Challenges and Issues. p. 196–202.
Jaeckisch, N., Yang, I., Wohlrab, S., Gloeckner, G., Kroymann, J., Vogel, H., Cembella, A., John, U., 2011. Comparative genomic and transcriptomic characterization of the toxigenic marine dinoflagellate *Alexandrium ostenfeldii*. PLoS ONE 6, e28012.
John, U., Fensome, R.A., Medlin, L.K., 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). Mol. Biol. Evol. 20 (7), 1015–1027.
John, U., Beszteri, B., Derelle, E., Van de Peer, Y., Read, B., Moreau, H., Cembella, A.D., 2008. Novel insights into evolution of protistan polyketide synthases through phylogenomic analysis. Protist 159, 21–30.
John, U., Litaker, W., Montresor, M., Murray, S., Broshanan, M., Anderson, D., 2014. Formal revision of the *Alexandrium tamarense* species complex (Dinophyceae) taxonomy: The introduction of five species with emphasis on molecular-based (rDNA) classification. Protist 165, 779–804.
Jørgensen, M.F., Murray, S.A., Daugbjerg, N., 2004. *Amphidinium* revisited. I. Redefinition of *Amphidinium* (Dinophyceae) based on cladistic and molecular phylogenetic analyses. J. Phycol. 40, 351–365.
Katoh, K., Toh, H., 2008. Recent developments in the MAFFT multiple sequence alignment program. Briefings Bioinf. 9 (4), 286–298.
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28 (12), 1647–1649.
Keeling, P.J., Palmer, J.D., 2008. Horizontal gene transfer in eukaryotic evolution. Nat. Rev. Genet. 9, 605–618.
Keller, M.D., Selvin, R.C., Claus, W., Guillard, R.R.L., 1987. Media for the culture of oceanic ultraphytoplankton. J. Phycol. 23, 633–638.
Kellmann, R., Mihali, T.K., Jeon, Y.J., Pickford, R., Pomati, F., Neilan, B., 2008. Biosynthetic intermediate analysis and functional homology reveal a saxitoxin gene cluster in cyanobacteria. Appl. Environ. Microbiol. 74, 4044–4053.
Kremp, A., Godhe, A., Egardt, J., Dupont, S., Suikkanen, S., Casabianca, S., Penna, A., 2012. Intra-specific variability in the response of bloom forming marine microalgae to changing climatic conditions. Ecol. Evol. 2, 1195–1207.
Kremp, A., Tahvanainen, P., Litaker, W., Krock, B., Suikkanen, S., Leaw, C.-P., Tomas, C., 2014. Phylogenetic relationships, morphological variation and toxin patterns in the *Alexandrium ostenfeldii* (Dinophyceae) complex: implications for species boundaries and identities. J. Phycol. 50, 81–100.
Lartillot, N., Lepage, T., Blanquart, S., 2009. PhyloBayes 3: a Bayesian software package for phylogenetic reconstruction and molecular dating. Bioinf. App. Note 25 (17), 2286–2288.
Lawrence, J., 1999. Selfish operons: the evolutionary impact of gene clustering in prokaryotes and eukaryotes. Curr. Opinion Genet. Dev. 9, 642–648.

Le, Q.H., Markovic, P., Hastings, J.W., Jovine, R.V., Morse, D., 1997. Structure and organization of the peridinin chlorophyll a-binding protein gene in *Gonyaulax polyedra*. Mol. Gen. Genet. 255, 595–604.

Leaw, C.P., Lim, P.T., Ng, B.K., Cheah, M.Y., Ahmad, A., Usup, G., 2005. Phylogenetic analysis of *Alexandrium* species and *Pyrodinium bahamense* (Dinophyceae) based on theca morphology and nuclear ribosomal gene sequence. J. Phycol. 44 (5), 550–565.

Lim, P.-T., Ogata, T., 2005. Salinity effect on growth and toxin production of four tropical *Alexandrium* species (Dinophyceae). Toxicon 45, 699–710.

Lin, S., 2011. Genomic understanding of dinoflagellates. Res. Microbiol. 166, 551–569.

Llewellyn, L., Negri, A., Robertson, A., 2006. Paralytic shellfish toxins in tropical oceans. Toxin Rev. 25, 159–196.

MacKenzie, L., Berkett, N., 1997. Cell morphology and PSP-toxin profiles of *Alexandrium minutum* in the Marlborough Sounds, New Zealand. N. Z. J. Mar. Freshwat. Res. 31, 403–409.

MacKenzie, L., de Salas, M., Adamson, J., Beuzenberg, V., 2004. The dinoflagellate genus *Alexandrium* (Halim) in New Zealand coastal waters: comparative morphology, toxicity and molecular genetics. Harmful Algae 3 (1), 71–92.

Maddison, W.P., Maddison, D.R., 2011. Mesquite: a Modular System for Evolutionary Analysis. Version 2.75. 2011. URL <http://mesquiteproject.org>.

Marchler-Bauer, A., Lu, S., Anderson, J.B., Chitsaz, F., Derbyshire, M.K., DeWeese-Scott, C., Fong, J.H., Geer, L.Y., Geer, R.C., Gonzales, N.R., Gwadz, M., Hurwitz, D.I., Jackson, J.D., Ke, Z., Lanczycki, C.J., Lu, F., Marchler, G.H., Mullokandov, M., Omelchenko, M.V., Robertson, C.L., Song, J.S., Thanki, N., Yamashita, R.A., Zhang, D., Zhang, N., Zheng, C., Bryant, S.H., 2011. CDD: a conserved domain database for the functional annotation of proteins. Nucleic Acids Res. 39 (Suppl. 1), D225–D229.

Medlin, L., Elwood, H.J., Stickel, S., Sogin, M.L., 1988. The characterization of enzymatically amplified eukaryotic 16S-like rRNA-coding regions. Gene 71 (2), 491–499.

Meyer, J.M., Rödelsberger, C., Eichholz, K., Tillmann, U., Cembella, A., McGaughran, A., John, U., 2015. Transcriptomic characterisation and genomic glimpse into the toxigenic dinoflagellate *Azadinium spinosum* with emphasis on polyketide synthase genes. BMC Genom. http://dx.doi.org/10.1186/s12864-014-1205-6.

Mihali, T.K., Kellmann, R., Neilan, B.A., 2009. Characterisation of the paralytic shellfish toxin biosynthesis gene clusters in *Anabaena circinalis* AWQC131C and *Aphanizomenon* sp. NH-5. BMC Biochem. 10, 8.

Mihali, T.K., Carmichael, W.W., Neilan, B.A., 2011. A putative gene cluster from a *Lyngbya wollei* bloom that encodes paralytic shellfish toxin biosynthesis. PLoS ONE 6, e14657.

Moustafa, A., Loram, J.E., Hackett, J.D., Anderson, D.M., Plumley, F.G., 2009. Origin of saxitoxin biosynthetic genes in cyanobacteria. PLoS ONE 4, e5758.

Murray, S.A., Flø Jørgensen, M., Ho, S.Y., Patterson, D.J., Jermiin, L.S., 2005. Improving the analysis of dinoflagellate phylogeny based on rDNA. Protist 156, 269–286.

Murray, S.A., Wiese, M., Stüken, A., Brett, S., Kellmann, R., Hallegraeff, G.M., Neilan, B.A., 2011a. *SxtA*-based quantitative molecular assay to identify saxitoxin-producing harmful algal blooms in marine waters. Appl. Environ. Microbiol. 77, 7050–7057.

Murray, S.A., Mihali, T.K., Neilan, B.A., 2011b. Extraordinary conservation, gene loss, and positive selection in the evolution of an ancient neurotoxin. Mol. Biol. Evol. 28, 1173–1182.

Murray, S.A., Garby, T., Hoppenrath, M., Neilan, B.A., 2012a. Genetic diversity, morphological uniformity and polyketide production in dinoflagellates (*Amphidinium*, Dinoflagellata). PLoS ONE 7 (6), e38253.

Murray, S.A., Wiese, M., Brett, S., Orr, R., Neilan, B.A., Hallegraeff, G., 2012b. A reinvestigation of saxitoxin production and *sxtA* in the 'non-toxic' *Alexandrium tamarense* Group V clade. Harmful Algae 18, 96–104.

Murray, S.A., Hoppenrath, M., Orr, R.J.S., Bolch, C., John, U., Diwan, R., Yauwenas, R., Harwood, T., de Salas, M., Neilan, B., Hallegraeff, G., 2014. *Alexandrium diversaporum* sp. nov., a new non-saxitoxin producing species: Phylogeny, morphology, and *sxtA* genes. Harmful Algae 31, 54–65.

Murray, S.A., John, U., Kremp, A., 2015. *Alexandrium* spp: genetic and ecological factors influencing saxitoxin production and proliferation. In: Botana, L.M., Louzao, M.C., Vilarino, N. (Eds.), Climate Change and Marine and Freshwater Toxins. DE Gruyter, Berlin, pp. 123–139.

Negri, A., Llewellyn, L., Doyle, J., Webster, N., Frampton, D., Blackburn, S., 2003. Paralytic shellfish toxins are restricted to few species among Australia's taxonomic diversity of cultured microalgae. J. Phycol. 39 (4), 663–667.

Nguyen Ngoc, L., 2004. An autecological study of the potentially toxic dinoflagellate *Alexandrium affine* isolated from Vietnamese waters. Harmful Algae 3, 117–129.

Orr, R.J.S., Stüken, A., Rundberget, T., Eikrem, W., Jakobsen, K.S., 2011. Improved phylogenetic resolution of toxic and non-toxic *Alexandrium* strains using a concatenated rDNA approach. Harmful Algae 10 (6), 676–688.

Orr, R.J.S., Murray, S., Stüken, A., Rhodes, L., Jakobsen, K.S., 2012. When naked became armored: an eight-gene phylogeny reveals monophyletic origin of theca in dinoflagellates. PLoS ONE 7 (11), e50004.

Orr, R.J.S., Stüken, A., Murray, S., Jakobsen, K.S., 2013. Evolutionary acquisition and loss of saxitoxin biosynthesis in Dinoflagellates: lessons from the second "core" gene – *sxtG*. Appl. Environ. Microbiol. 79, 2128–2136.

Oshima, Y., Hasegawa, M., Yasumoto, T., Hallegraeff, G., Blackburn, S., 1987. Dinoflagellate *Gymnodinium catenatum* as the source of paralytic shellfish toxins in Tasmanian shellfish. Toxicon 25, 1105–1111.

Parkhill, J.-P., Cembella, A.D., 1999. Effects of salinity, light and inorganic nitrogen on growth and toxigenicity of the marine dinoflagellate *Alexandrium tamarense* from northeastern Canada. J. Plankton Res. 21, 939–955.

Perini, F., Galluzzi, L., Dell'Aversano, C., Iacovo, E.D., Tartaglione, L., Ricci, F., Forino, M., Ciminiello, P., Penna, A., 2014. *SxtA* and *sxtG* gene expression and toxin production in the Mediterranean *Alexandrium minutum* (Dinophyceae). Mar. Drugs 12 (10), 5258–5276.

Pond, S.L.K., Murrell, B., Fourment, M., Frost, S.D., Delport, W., Scheffler, K., 2011. A random effects branch-site model for detecting episodic diversifying selection. Mol. Biol. Evol. 28 (11), 3033–3043.

Rogers, J.E., Leblond, J.D., Moncreiff, C.A., 2006. Phylogenetic relationship of *Alexandrium monilatum* (Dinophyceae) to other *Alexandrium* species based on 18S ribosomal RNA gene sequences. Harmful Algae 5 (3), 275–280.

Sampedro, N., Franco, J.M., Zapata, M., Riobó, P., Garcés, E., Penna, A., Caillaud, A.., Diogène, J., Cacho, E., Camp, J., 2013. The toxicity and intraspecific variability of *Alexandrium andersonii* Balech. Harmful Algae 25, 26–38.

Scholin, C.A., Herzog, M., Sogin, M., Anderson, D.M., 1994. Identification of group and strain specific genetic markers for globally distributed *Alexandrium* (Dinophyceae). II Sequence analysis of a fragment of the LSU rRNA gene. J. Phycol. 30 (6), 999–1011.

Selander, E., Thor, P., Toth, G., Pavia, H., 2006. Copepods induce paralytic shellfish toxin production in marine dinoflagellates. Proc. R. Soc. London B 273, 1673–1680.

Shalchian-Tabrizi, K., Minge, M.A., Cavalier-Smith, T., Nedreklepp, J.M., Klaveness, D., 2006. Combined heat shock protein 90 and ribosomal RNA sequence phylogeny supports multiple replacements of dinoflagellate plastids. J. Euk. Microb. 53, 217–224.

Slamovits, C.H., Keeling, P.J., 2008. Widespread recycling of processed cDNAs in dinoflagellates. Curr. Biol. 18 (13), R550–R552.

Stamatakis, A., 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. Bioinformatics 22 (21), 2688.

Stucken, K., John, U., Cembella, A., Murillo, A.A., Soto-Liebe, K., 2010. The smallest known genomes of multicellular and toxic cyanobacteria: comparison, minimal gene sets for linked traits and the evolutionary implications. PLoS ONE 5, e9235.

Stüken, A., Orr, R.J.S., Kellmann, R., Murray, S.A., Neilan, B.A., Jakobsen, K.S., 2011. Discovery of nuclear-encoded genes for the neurotoxin saxitoxin in dinoflagellates. PLoS ONE 6 (5), e20096.

Suikkanen, S., Kremp, A., Hautala, H., Krock, B., 2013. Effects of salinity on growth rate and PST/spirolide production of the dinoflagellate *Alexandrium ostenfeldii*. Harmful Algae 26, 52–59.

Touzet, N., Franco, J.M., Raine, R., 2007. Characterization of nontoxic and toxin producing strains of *Alexandrium minutum* (Dinophyceae) in Irish coastal waters. Appl. Environ. Microbiol. 73 (10), 3333–3342.

Touzet, N., Franco, J.M., Raine, R., 2008. Morphogenetic diversity and biotoxin composition of *Alexandrium* (Dinophyceae) in Irish coastal waters. Harmful Algae 7, 782–797.

Van de Waal, D.B., Smith., V.H., Declerck, S.A.J., Stam, E.C.M., Elser, J.J., 2014. Stoichiometric regulation of phytoplankton toxins. Ecol. Lett., 1–7

Waller, R.F., Slamovits, C.H., Keeling, P.J., 2006. Lateral gene transfer of a multigene region from cyanobacteria to dinoflagellates resulting in a novel plastid-targeted fusion protein. Mol. Biol. Evol. 23 (7), 1437–1443.

Wang, D.-Z., Zhang, S.G., Gu, H.-F., Chan, L.L., Hong, H.-S., 2006. Paralytic shellfish toxin profiles and toxin variability of the genus *Alexandrium* (Dinophyceae) isolated from the Southeast China Sea. Toxicon 48, 138–151.

Wiens, J.J., 2005. Can incomplete taxa rescue phylogenetic analyses from long-branch attraction? Syst. Biol. 54, 731–742.

Wiese, M., 2012. Investigations into Abiotic and Biotic Factors Regulating Saxitoxin Synthesis in the Dinoflagellate Genus *Alexandrium*. PhD thesis, University of New South Wales, Australia.

Wisecaver, J., Hackett, J., 2014. The impact of automated filtering of BLAST-determined homologs in the phylogenetic detection of horizontal gene transfer from a transcriptome assembly. Mol. Phylogenet. Evol. 71, 184–192.

Wohlrab, S., Iversen, M., John, U., 2010. A molecular and co-evolutionary context for grazer induced toxin production in *Alexandrium tamarense*. PLoS ONE 5 (11), e15039.

Yang, I., John, U., Beszteri, S., Glockner, G., Krock, B., Goesmann, A., Cembella, A.D., 2010. Comparative gene expression in toxic versus non-toxic strains of the marine dinoflagellate *Alexandrium minutum*. BMC Genom. 11, 248.

Yang, I., Selander, E., Pavia, H., John, U., 2011. Grazer-induced toxin formation in dinoflagellates: a transcriptomic model study. Eur. J. Phycol. 46 (1), 66–73.

Zhang, H., Hou, Y., Miranda, L., Campbell, D.A., Sturm, N.R., Gaasterland, T., Lin, S., 2007. Spliced leader RNA trans-splicing in dinoflagellates. Proc. Nat. Acad. Sci. USA 104, 4618–4623.

Zhang, Y., Zhang, S.F., Lin, L., Wang, D.Z., 2014. Comparative transcriptome analysis of a toxin-producing dinoflagellate *Alexandrium catenella* and its non-toxic mutant. Mar. Drugs 12, 5698–5718.