

***‘Through the looking glass’: Diversity
and its functional significance in
marine benthic microbial eukaryotes***

Arjun Verma

February 2018

Supervisor: Shauna Murray

Co-supervisor: Peter Ralph



A thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy; Climate Change Cluster, School of Life Sciences, University of Technology Sydney

CERTIFICATE OF ORIGINAL AUTHORSHIP

I certify that the work in this thesis has not previously been submitted for a degree nor has it been submitted as part of requirements for a degree except as part of the collaborative doctoral degree and/or fully acknowledged within the text.

I also certify that the thesis has been written by me. Any help that I have received in my research work and the preparation of the thesis itself has been acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This research is supported by an Australian Government Research Training Program Scholarship.

Production Note:

Signature of Student: Signature removed prior to publication.

Date: 23/01/2018

ACKNOWLEDGEMENTS

This thesis would not be possible without the wondrous and enigmatic nature of my study organism, *Ostreopsis* Schmidt, whose beauty and mysteries are endless. I would like to sincerely thank my supervisor, A./Prof. Shauna Murray for her guidance and vision that steered me through my PhD candidature, and also my co-supervisor, Prof. Peter Ralph for his guidance and critical input in certain sections. I would also like to thank the past and present members of the Seafood Safety Group at the Climate Change Cluster for their valuable inputs and friendship. The various facets of my project were completed with the input of various collaborating scientists; Dr. Mona Hoppenrath for her contributions with the SEM images, Dr. Tim Harwood for his contributions with the toxicity analyses, Dr. Steve Brett for his contributions with the monitoring data, Dr. Juan Dorantes Aranda for his contribution with the cell line bioassays, Dr. Gurjeet Singh Kohli for his contribution with the transcriptomic analyses, Dr. Unnikrishnan Kuzhiumparambil for his contribution with metabolomic analyses and David Hughes for his contribution with photophysiology analyses of the *Ostreopsis* strains.

I would also like to thank Drs. Lesley Rhodes and Kirsty Smith for being a mentor and helping me with various projects over the candidature. I would like to thank Dr. Uwe John from the Alfred Wegener Institute for his hospitality during my short lab visit in Germany and for his valuable insight into my project. I would like to thank Dr. Rex Munday for the mouse bioassay data in chapter 2 and Dr. Jennifer Clark, Michaela Larsson, Risa Fujise, Dr. Hazel Farrell, Dr. Gurjeet S. Kohli, Dr. Katrina Petrou and Varunan Balaraju for aid in macroalgal sample collection in chapters 3 and 4. I would like to thank the staff at the Ramaciotti Centre of Genomics, University of New South Wales for their service in analysing RNA quality, preparing and sequencing RNA-Seq libraries and Mike Lake and Anna Liza Kretzschmar for support with the use of high performance computing for data analysis in Chapter 5.

I would also like to thank the technical staff at the University of Technology Sydney for their support with culturing and incubators especially Paul Brooks. I would like to thank John Moore for his assistance with administrative paperwork and overseas travel

formalities. I would also like to thank the UTS travel awards, International Society for the study of Harmful algae (ISSHA) travel award, Gordon and Betty Moore Foundation travel award, Australian Biological Resources Study (ABRS) Taxonomy Forum Travel Grant for the financial support to present my PhD work at various international conferences. I would also like to extend my warmest thanks to Dr. Leo Hardkte, Dr. Buddhi Dayanda and Nasim Shah Mohammadi for their aid in figure generation, volumetric and statistical analyses.

I would like to thank my friends and loved ones who stood by my side during my candidature and my family for their support that gave me the strength to complete this massive endeavour. Big thanks to my mother, Anita Verma who have been my pillar of support.

In the end, I would like to dedicate my PhD work to my grandfather, Mr. Satya Narain, the silent comrade, whose teachings and directions have motivated me to pursue the mysteries of nature and have led me to the endless pursuit of truth, whatever its shape and form. Thank you.

The woods are lovely, dark and deep,

But I have promises to keep,

And miles to go before I sleep,

And miles to go before I sleep.

-Stopping by Woods on a Snowy Evening, Robert Frost

I shall be telling this with a sigh

Somewhere ages and ages hence:

Two roads diverged in a wood, and I—

I took the one less travelled by,

And that has made all the difference.

-The Road Not Taken, Robert Frost

ORIGINAL PUBLICATIONS

Publications included in this thesis:

1. **Verma, A.**, Hoppenrath, M., Harwood, T., Brett, S., Rhodes, L., Murray, S., 2016, Molecular phylogeny, morphology and toxigenicity of *Ostreopsis* cf. *siamensis* (Dinophyceae) from temperate south-east Australia, *Phycological Research* 64(3), 146-59
2. **Verma, A.**, Hoppenrath, M., Dorantes-Aranda, J.J., Harwood, D.T. and Murray, S.A., 2016. Molecular and phylogenetic characterization of *Ostreopsis* (Dinophyceae) and the description of a new species, *Ostreopsis rhodesae* sp. nov., from a subtropical Australian lagoon. *Harmful Algae*, 60, 116-130.
3. **Verma, A.**, Kohli, G. S., Hoppenrath, M., Harwood, T., Kuzhiumparambil, U., Ralph, P. J., Murray, S. A., 2017 Systematics and diversity of the genus *Ostreopsis* in the East Australian Current region. In *Proceedings of the 17th International Conference on Harmful Algae*. International Society for the Study of Harmful Algae 2017.

TABLE OF CONTENTS

Certificate of original authorship	i
Acknowledgements	ii
Publications	v
Publications included in this thesis	v
Table of contents	vi
List of Figures	x
List of Tables	xv
List of supplementary data	xvii
Preface	xxi
Thesis abstract	xxii
Chapter 1: General Introduction	1
1.1 Dinoflagellates: ‘Through the looking glass’	3
1.1.1 Contemporary issues in dinoflagellate studies	4
1.1.1.1 Cryptic species	4
1.1.1.2 Population structure	7
1.1.1.3 Functional trait diversity	9
1.1.1.4 Secondary metabolites and their biosynthesis	10
1.1.1.5 Dinoflagellate genetics	11
1.2 The East Australian Current region: a climate change hotspot	13
1.3 Genus <i>Ostreopsis</i>	16
1.3.1 Cryptic morphology	16
1.3.2 Phylogenetic studies	18
1.3.3 Biogeography	20
1.3.4 Toxin producers: Palytoxin	22
1.3.4.1 Chemical structure and mode of action	22
1.3.4.2 Toxin analogues and variability	23
1.3.4.3 Human health and ecological impact	25
1.4 Research objectives and thesis outline	26
Chapter 2: Molecular phylogeny, morphology and toxigenicity of <i>Ostreopsis</i> cf. <i>siamensis</i> (Dinophyceae) from temperate south-east Australia	29
2.1 Abstract	30
2.2 Introduction	31
2.3 Materials and methods	33
2.3.1 Site description	33
2.3.2 Sample collection and culture establishment	34
2.3.3 Light microscopy	35
2.3.4 Scanning electron microscopy (SEM)	36
2.3.5 DNA extraction, PCR amplification and sequencing	36
2.3.6 Sequence alignment and phylogenetic analysis	37
2.3.7 Growth rate and cell size	37
2.3.8 Toxin analysis via LC-MS/MS and bioassays	38

2.4 Results	41
2.4.1 Morphology	41
2.4.2 Phylogeny	42
2.4.3 Culturing and growth rates	42
2.4.4 Distribution and abundance	43
2.4.5 Toxin analysis and toxicity	43
2.5 Discussion	47
2.6 Author Contributions	55
Chapter 3: Molecular and phylogenetic characterization of <i>Ostreopsis</i> (Dinophyceae) and the description of a new species, <i>Ostreopsis rhodesae</i> sp. nov., from a subtropical Australian lagoon	57
3.1 Abstract	58
3.2 Introduction	59
3.3 Materials and methods	61
3.3.1 Sample collection and culture establishment	61
3.3.2 Microscopy	63
3.3.3 DNA extraction and PCR amplification	64
3.3.4 Sequence analysis and phylogenetic reconstruction	64
3.3.5 Modelling ITS2 secondary structure	65
3.3.6 Toxin analysis via LC–MS/MS	66
3.3.7 Fish gill cell line assay for toxicity	67
3.4 Results	67
3.4.1 <i>Ostreopsis rhodesae</i> Verma, Hoppenrath et Murray sp. nov.	67
3.4.1.1 Morphological description	67
3.4.1.2 Holotype	72
3.4.1.3 Isotype	73
3.4.1.4 Type locality	73
3.4.1.5 Etymology	73
3.4.1.6 Accession numbers	74
3.4.2 Molecular analyses and phylogeny	74
3.4.3 ITS2 secondary structure	75
3.4.4 Toxin presence	77
3.4.5 Fish gill cell assays	77
3.5 Discussion	78
3.5.1 Morphological comparison among <i>Ostreopsis</i> species	78
3.5.2 Phylogeny and biogeography of genus <i>Ostreopsis</i>	79
3.5.3 Toxicity	85
3.6 Author contributions	87
Chapter 4: Functional significance of phylogeographic structure in a toxic marine protist (<i>Ostreopsis</i>, Dinophyceae) along a 1500 km of north-south gradient in the East Australian Current	89
4.1 Abstract	90
4.2 Introduction	91
4.3 Materials and methods	92

4.3.1 Site descriptions, sample collection and strain establishment	92
4.3.2 DNA extraction, PCR and sequencing	94
4.3.3 Phylogenetic analyses	96
4.3.4 Growth rates estimates	97
4.3.5 Cell volume analysis	97
4.3.6 PLTX toxin determination	98
4.3.7 FRRf - Dark-acclimated photophysiology	98
4.3.8 FRRf – Photosynthetic-Irradiance (PE) Response	99
4.3.9 Pigments and Photosynthetic Unit (PSU) Size	100
4.3.10 Statistical analysis	101
4.4 Results	102
4.4.1 Sampling and species identification	102
4.4.2 Phylogeographic structures and genetic diversity	102
4.4.3 Trait variability among isolates	107
4.4.3.1 Growth rates and cell volume	107
4.4.3.2 Toxin production	107
4.4.3.3 Photobiological parameters	109
4.5 Discussion	112
4.5.1 Population divergence	112
4.5.2 Phenotypic variation	114
4.5.2.1 Growth rates and cell volume	114
4.5.2.2 Toxin content and composition	115
4.5.2.3 Photophysiological strategies	117
4.6 Conclusion and significance	118
4.7 Author contributions	119
Chapter 5: Transcriptomic and metabolomic insights into polyketide toxin production in species of <i>Ostreopsis</i> (Dinophyceae)	121
5.1 Abstract	122
5.2 Introduction	123
5.3 Materials and methods	125
5.3.1 Cultures	125
5.3.2 PLTX analysis via LC-MS/MS	126
5.3.3 RNA isolation and sequencing	126
5.3.4 Transcriptome Assemblies and Annotation	126
5.3.5 Non-targeted metabolomics using UPLC	128
5.4 Results	129
5.4.1 Toxin analysis	129
5.4.2 <i>De novo</i> assembly and annotation	129
5.4.3 Polyketide biosynthesis	132
5.4.4 Fatty acid synthesis	136
5.4.5 Putative metabolomic profiles of dinoflagellate species	139
5.5 Discussion	140
5.5.1 Polyketide biosynthesis	141
5.5.2 Fatty acid biosynthesis	143

5.5.3 Metabolomic insights	144
5.6 Conclusion and significance	146
5.7 Author contributions	147
Chapter 6: General Discussion	149
6.1 Overview	150
6.2 Significance and future of findings	151
6.3 Thesis conclusion	157
Bibliography	159
Supplementary data	189
Original publications from thesis	317

LIST OF FIGURES

Figure 1.1 A: The schematic of the major lineages in the eukaryotic tree of life as represented in Burki et al. (2014). B: Phylogenetic tree of dinoflagellates inferred from rDNA modified from Orr et al. (2012).	6
Figure 1.2 Unrooted Bayesian phylogeny showing cryptic population structure interred from 135,035 polymorphic sites in a fungal species <i>Neurospora crassa</i> as represented in Ellison et al. (2011).	8
Figure 1.3 Number of protein coding genes compared to genome size (in log scale) of various organisms highlighting the large genome sizes of dinoflagellates as represented in Murray et al. (2016).	12
Figure 1.4 The warm EAC jet flows along the shelf-edge off eastern Australia as represented in Ridgway and Hill (2009).	14
Figure 1.5 Seasonal occurrence of <i>Ostreopsis</i> spp. in aquaculture plankton samples collected from shellfish producing estuaries along the New South Wales coastline between July 2005 and December 2013. Cell count represents cells identified microscopically in one litre of water sample.	15
Figure 1.6 Original drawings of <i>Ostreopsis siamensis</i> as represented in Schmidt (1902). The epithecal view is presented on the left, the hypothecal view on the right (Parsons et al., 2010).	17
Figure 1.7 <i>Ostreopsis</i> species drawings of the plate patterns as represented in Hoppenrath et al. (2014).	19
Figure 1.8 Global distribution of <i>Ostreopsis</i> species modified from Rhodes (2011). Black dots indicate locations of molecular and morphological reports.	21
Figure 1.9 Structure of PLTX as represented in Ramos and Vasconcelos (2010)	23
Figure 2.1 Map showing Merimbula lake inlet, south-east New South Wales, Australia.	35
Figure 2.2 <i>Ostreopsis</i> cf. <i>siamensis</i> CAWD203 from Merimbula taken using light microscopy. A, C–E: Differential interference contrast showing the general morphology; and B: epifluorescence demonstrating the autofluorescence of the chloroplasts. A: Typical very wide cell, note the colourless ventral area and the dorsal nucleus (n). B: Chloroplast fluorescence of the cell shown in A. C: Smaller and narrower cell with one pusule (p) visible in the ventral cell half connected to the ventral area. D, E: Same cell in different focus. D: Note the nucleus (n) in the dorsal area and the two pusules (p). E: The apical pore complex (arrow) in the dorsal area. F, G: Line drawings illustrate the thecal plate pattern of the epitheca (F); and hypotheca (G) including cingular plates. The scale bars represent 10 µm. '?' represent verification of cingular plate borders.	44

Figure 2.3 Scanning electron micrographs (SEM) of <i>Ostreopsis</i> cf. <i>siamensis</i> CAWD203 from Merimbula. A: Epitheca; and B: Hypotheca. C: Inside view of the apical pore complex and surrounding plates. D: Plate detail showing two size classes of pores, small (arrowhead) and large (arrow) ones. Scale bars represent 10 μ m in A, B; 5 μ m in C, D.	45
Figure 2.4 SEM of <i>Ostreopsis</i> cf. <i>siamensis</i> CAWD203 from Merimbula. Details of the sulcal area. A: Hypothecal view of the ventral area. B–F: Inside views of broken cells. sa = anterior sulcal plate, ssa = anterior left sulcal plate, ssp = posterior left sulcal plate, sda = anterior right sulcal plate, sdp = posterior right sulcal plate, sp = posterior sulcal plate, 1'''' = first antapical plate. Scale bars represent 5 μ m.	46
Figure 2.5 Maximum Likelihood (ML) phylogenetic trees of various <i>Ostreopsis</i> strains using A: D8/D10; and B: D1/D3 LSU rDNA regions. Merimbula strain CAWD203 shown in bold letters in <i>Ostreopsis</i> cf. <i>siamensis</i> clade shaded grey. External black vertical bars show each distinct <i>Ostreopsis</i> clade and internal vertical bars show each <i>Ostreopsis</i> sub-clade. Med, Atl, Pac and Ind represent Mediterranean Sea, Atlantic, Pacific and Indian Oceans sub-clades, respectively. South China Sea and Thailand are the <i>Ostreopsis</i> cf. <i>ovata</i> South China Sea and Gulf of Thailand sub-clades respectively. Numbers at nodes represent posterior probabilities from BI and bootstrap support values from ML based on 1,000 pseudo-replicates. Robust branches (BI=1.00 and ML=100) are indicated by asterisks.	48
Figure 2.6 ML phylogenetic trees of various <i>Ostreopsis</i> strains using A: ITS1/5.8S/ITS2; and B: SSU rDNA regions. See the caption in Figure 2.5 for the detailed information.	49
Figure 2.7 Growth pattern of <i>Ostreopsis</i> cf. <i>siamensis</i> from Merimbula in f/2 and f/10 batch cultures. Each point represents the mean \pm SE of three experiments of three replicates.	51
Figure 2.8 Seasonal occurrence of <i>Ostreopsis</i> spp. in aquaculture plankton samples collected from Merimbula Lake Inlet (two sites within the inlet) between October 2005 and May 2015, indicating year around presence with higher abundance in austral spring-summer.	52
Figure 2.9 Extracted ion chromatograms from the solid phase extraction and on-column oxidation of A: Palytoxin (PLTX) standard (50 ng mL ⁻¹); and B: <i>Ostreopsis</i> cf. <i>siamensis</i> CAWD203 from Merimbula, Australia.	54
Figure 3.1 A: Map of the north-eastern coastline of Australia, showing Heron Island. B: Map of Heron reef lagoon, showing sampling site during June 2014 and February 2015 (shown as black dot).	62
Figure 3.2 Light micrographs of <i>Ostreopsis rhodesae</i> sp. nov. HER32 showing the cell shape and general features. A, B: Same cell in different focal planes. Ovate cell ventrally tapering and ventral area devoid of chloroplasts. A: Note the APC (short arrow) and the transverse flagellum (long arrow). B: The nucleus (n) is located in the right dorsal area.	68

Chloroplasts are elongated (arrows). C: Cell with different shape confirming the nucleus (n) position. Scale bars = 10 μ m.

Figure 3.3 Scanning electron micrographs of *Ostreopsis rhodesae* sp. nov. HER32 showing the general thecal tabulation. A, B: Epitheca in apical view. Note the suture between plates 3' and 6'' (arrowhead). C: Epitheca in right lateral view. Note the suture between plates 3' and 6'' (arrowhead). D: Epitheca in apical view with heptagonal 1' plate, with suture between plates 1' and 5'' (arrow). E, F: Hypotheca in antapical view. Scale bars = 10 μ m. 70

Figure 3.4 Scanning electron micrographs of *Ostreopsis rhodesae* sp. nov. HER32 showing details of the second apical plate (2'). A: Inside view of a broken theca. B: Outside view of a broken theca. C: Outside view of an intact cell. Note the 2' plate margins (arrows). Scale bars = 5 μ m. 71

Figure 3.5 Scanning electron micrographs of *Ostreopsis rhodesae* sp. nov. HER32. A: Cell in left lateral view showing the undulated cingulum path. B-F: Sulcal details. B: Ventral hypotheca, outside view. C: Ventral hypotheca, outside view of a broken cell. D: Ventral part of a broken cell. E: Isolated ssa plate in connection with the first cingular plate. F: Sulcal plates separated from the theca, inside view. G-J: Thecal pores. G: Thecal plate detail, outside view. H-J: Details of the inside of thecal pores showing the sieve-like structure. Note the simple small pores in I and J. Scale bars A: 10 μ m, C-F: 5 μ m, G-J: 1 μ m. 72

Figure 3.6 Line drawings showing thecal plate patterns. A, B: *Ostreopsis rhodesae* sp. nov. A: Epitheca; and B: Hypotheca. C, D: *Ostreopsis* heptagona (from Hoppenrath et al., 2014). C: Epitheca; and D: Hypotheca. 73

Figure 3.7 Maximum Likelihood (ML) phylogenetic trees of various *Ostreopsis* strains using primer sets for A: ITS1-5.8S-ITS2; and B: D8-D10 LSU rDNA regions. External Black vertical bars show each distinct *Ostreopsis* clade and internal vertical bars show each *Ostreopsis* subclade. Med, Atl, Pac and Ind are the *Ostreopsis* cf. *ovata* Mediterranean Sea, Atlantic, Pacific and Indian Oceans subclades respectively. Malacca, Celebes, South China Sea and Thailand are the *Ostreopsis* cf. *ovata* Malacca strait, Celebes Sea, South China Sea and Gulf of Thailand subclades respectively. Numbers at nodes represent posterior probabilities from Bayesian Inferences (BI) and bootstrap support values from Maximum Likelihood (ML) based on 1000 pseudo-replicates. * represents 1, 100 support values for BI and ML respectively. 81

Figure 3.8 ML phylogenetic trees of various *Ostreopsis* strains using primer sets for A: D1-D3 LSU rDNA; and B: SSU rDNA regions. See the caption in Figure 3.7 for the detailed information. 82

Figure 3.9 Predicted ITS2 secondary structure of *Ostreopsis* strains. A: *Ostreopsis* cf. *ovata* HER27; B: *Ostreopsis* cf. *siamensis* HER24; and C: *Ostreopsis rhodesae* HER26. 84

Figure 3.10 Effect of crude extracts from A: <i>Ostreopsis</i> cf. <i>siamensis</i> HER24; B: <i>Ostreopsis rhodesae</i> HER26; and C: <i>Ostreopsis</i> cf. <i>ovata</i> HER27 on viability of fish gill cells RTgill-W1. Plots are average from quadruplicate wells and bars represent their standard deviation. Symbols (*) indicate significant differences (at $p \leq 0.05$) when comparing the effect of extracts from exponential and stationary growth phase at each extract concentration. Arrows and negative values show the decrease in gill cell viability.	86
Figure 4.1 Map of the south-eastern coastline of Australia showing sampling locations in this study. Macroalgal samples were collected during April-July 2014. MW, BH, FR, LM, PC, GB, KM and MER represent sampling sites, i.e. Minnie Waters, Bonny Hills, Wallis Lake in Forster, Lake Macquarie, Patonga Creek, Gordons Bay, Kiama and Merimbula Lake Inlet respectively. Numbers in the circles represent the number of clonal isolates that were established from the sampling site. Isothermal lines represent the mean sea surface temperature (SST) from 2012-2017 varying north to south from 24-17°C by a gradient of 1°C.	93
Figure 4.2 A: Maximum likelihood (ML) phylogenetic tree based on ITS-5.8S/D1-D3 and D8-D10 LSU rDNA concatenated sequences representing the two sub-clades of <i>Ostreopsis</i> cf. <i>siamensis</i> . B: Phylogram representing the various haplotypes within the two sub-clades. The internal grey line represents haplotype 1 common to all locations. Numbers at nodes represent posterior probabilities from Bayesian Inferences (BI) and bootstrap support values from ML based on 1000 pseudo-replicates. Only bootstrap values > 50% are shown. * represents 1, 100 support values for BI and ML respectively. Colour codes represent origin of strains as represented in Figure 4.1.	103
Figure 4.3 Haplotype network based on 68 concatenated sequences using statistical parsimony. Mutation steps are shown in black dots. Colour codes represent origin of strain as represented in Figure 4.1. Strain codes are representatives of haplotypes as described in Table 4.5. Numbers in the pie chart represent the number of isolates from each location that belong to haplotype 1.	104
Figure 4.4 Phenotypic variation of 53 <i>Ostreopsis</i> cf. <i>siamensis</i> strains (represented on the x axes). Colour codes represent origin of strain as represented in Figure 4.1. A: Mean growth rates. Error bars represent standard error of three replicate measurements. B: Cell volume. Error bars represent standard error of twenty measurements. C: Cellular toxin content. Error bars represent 8-10% relative standard deviation of repeatability for LC-MS measurements. NA represents toxin amount below the limit of detection	108
Figure 4.5 Toxin profile variation amongst <i>Ostreopsis</i> cf. <i>siamensis</i> strains in this study. A: Both amino and amide aldehyde fragments observed; B: only amino aldehyde fragment observed; C: Only amide aldehyde fragment observed; and D: No fragments observed (Below the limit of detection)	109

Figure 4.6 Whisker plots of A: maximum photosynthetic rate (ETR_{max}); B: light utilisation efficiency (α); C: light saturation parameter (E_k); and D: PSU size amongst *Ostreopsis cf. siamensis* isolates based upon sampling sites. Whiskers above and below the boxes indicate the 90/10 percentiles, dots the respective 95/5 percentiles. 110

Figure 4.7 Functional groupings based on phenotypic variability in *Ostreopsis cf. siamensis* clones. Light harvesting (F_v/F_m , σ , Cellular RCII concentration ([RCII])) and light utilization ([1-C] and [1-Q]) along with cell volume, toxin amounts and growth rates were measured across all strains. Cluster analysis and multi-dimensional scaling (MDS) were performed on the average of each variable per strain; similarity is shown at 90 and 95% level and vectors driving the clustering are shown in black. 111

Figure 5.1 BLASTx analysis using BLAST2GO (e-value cut off 10^{-3}) for *Ostreopsis cf. ovata* HER27, *Ostreopsis cf. siamensis* BH1, *Ostreopsis rhodesae* HER26 and *Coolia malayensis* MAB. 131

Figure 5.2 Phylogenetic analysis of ketoacyl synthase (KS) domains from prokaryotic and eukaryotic type I & II polyketide synthases (PKS) and fatty acid synthases. Consensus schematic representation of the multi-KS and NRPS/PKS hybrid domains are displayed on the right. AT: acyl transferase, A: Non-Ribosomal Peptide Synthase, KS: ketosynthase, DH: dehydratase, KR: ketoreductase, TE: thioesterase, ACP: acyl carrier protein, ER: enoyl reductase; SL: 5' splice leader; Poly-A: 3' poly A tail. 135

Figure 5.3 Phylogenetic analysis of ketoacyl reductase (KR) domains from prokaryotic and eukaryotic type I & II polyketide synthases (PKS) and fatty acid synthases. Consensus schematic representation of the multi-KS and NRPS/PKS hybrid domains are displayed similar to Figure 5.2. 136

Figure 5.4 Concatenated phylogeny of five enzymes involved in type II fatty acid synthesis (3-ketoacyl ACP synthase III, s-malonyltransacylase, trans3-ketoacyl ACP reductase, 3-hydroxyacyl-ACP dehydratase and enoyl-ACP reductase) from 22 dinoflagellates and one other alveolate *Chromera velia* which was used as an outgroup. Phylogeny was inferred using RAxML, GAMMA model of rate heterogeneity and 1,000 bootstraps. 138

Figure 5.5 Non-targeted metabolomics on *Ostreopsis* and *Coolia* spp. A: Heat map of differential analysis of the four metabolomic profiles between 500-2400 m/z ratios. B: 3-D principal component plot of the four dinoflagellate metabolomic profiles. 139

Figure 5.6 Venn diagram displaying shared metabolites detected in the non-targeted metabolic analyses of *Ostreopsis* species. The numbers represent the total entities that are significantly different between *Coolia malayensis* and the respective *Ostreopsis* species. 140

LIST OF TABLES

Table 2.1 Geographic, morphological and molecular reports of <i>Ostreopsis siamensis</i> / <i>Ostreopsis</i> cf. <i>siamensis</i> .	40
Table 3.1 Morphological characteristics (means and standard deviations, ranges) of <i>Ostreopsis</i> strains determined by light microscopy: dorso-ventral diameter (DV), trans-diameter (W) and DV/W ratio. All data were from cultured cells. * represents type strain.	69
Table 3.2 Distance values (pairwise uncorrected p-distances) based on the ITS/5.8S, D1/D2, D8/D10 LSU and 18S rDNA sequences respectively within <i>Ostreopsis rhodesae</i> strains and between <i>Ostreopsis</i> cf. <i>siamensis</i> HER24 and <i>Ostreopsis</i> cf. <i>ovata</i> HER27 from Heron Island (based on Clustal W alignment). Standard error estimate(s) are shown in brackets and were obtained by a bootstrap procedure (1000 replicates).	75
Table 3.3 List of Compensatory base changes (CBCs) and hemi-CBCs between <i>Ostreopsis rhodesae</i> , <i>Ostreopsis</i> cf. <i>ovata</i> strain HER27 and <i>Ostreopsis</i> cf. <i>siamensis</i> strain HER24.	76
Table 4.1 Details of the sampling sites and macroalgal samples used for establishing monoclonal cultures of <i>Ostreopsis</i> cf. <i>siamensis</i> in this study	95
Table 4.2 Primers used for phylogenetic analyses in this study and the annealing temperature (Ta) used for the PCR reactions.	97
Table 4.3 Molecular diversity indexes (uncorrected p-distances) in the ITS1-5.8S-ITS2, D1-D3, D8-D10 rDNA and concatenated sequences within sampling locations. Standard error estimate(s) are shown in brackets and were obtained by a bootstrap procedure (1000 replicates).	104
Table 4.4 Molecular diversity indexes (uncorrected p-distances) in the ITS1-5.8S-ITS2, D1-D3, D8-D10 rDNA and concatenated sequences between sampling locations. Standard error estimate(s) are shown in brackets and were obtained by a bootstrap procedure (1000 replicates).	105
Table 4.5 Haplotype diversity of <i>Ostreopsis</i> cf. <i>siamensis</i> isolates based upon location of origin as obtained from Arlequin. Standard error estimate(s) are shown in brackets and were obtained by a bootstrap procedure (1000 replicates). Strain codes are representatives of the various haplotypes. Number in the bracket represents the number of isolates that belong to the representative haplotype. 1 represent the strain codes and the number of isolates from each location that belong to haplotype 1.	106
Table 5.1 Transcriptome assembly statistics including the total number of polyketide synthase associated domains found for the three <i>Ostreopsis</i> species and <i>Coolia malayensis</i> used in this study.	132

Table 5.2 List of PKS-NRPS found in the four transcriptomes. ACP- Acyl carrier protein; TE-Thioestrane; NRPS(a)- The adenylation domain of non-ribosomal peptide synthetases (NRPS); ER- Enoylreductase; KS- Ketosynthase; KR-ketoreductase; AT-Acyl transferase; DH- dehydratase; NRPS(p)- partial NRPS domain	134
Table 5.3 List of FAS encoding transcripts from the four transcriptomes	137

LIST OF SUPPLEMENTARY DATA

S1 Primers used for amplification and sequencing	191
S2 List of <i>Ostreopsis</i> spp. clones used for phylogenetic reconstruction and for inferring <i>p</i> -distances	192
S3 Distance values (pairwise uncorrected <i>p</i> -distances) based on the SSU rDNA sequences (Clustal W alignment) between and within clades of <i>Ostreopsis</i> .	194
S4 Comparison of acclimated growth rates of <i>Ostreopsis</i> spp. Growth rates were calculated from the exponential phase portion of the growth curve. Values are mean of three replicates.	195
S5 Palytoxin (PLTX)-equivalents quantification in <i>Ostreopsis siamensis</i> isolates from the Australian and New Zealand waters according to Selwood et al., 2012.	196
S6: ITS2 Secondary structures of <i>Ostreopsis</i> cf. <i>ovata</i> subclades. A: <i>Ostreopsis</i> cf. <i>ovata</i> Malacca Sea sub-clade; B: <i>Ostreopsis</i> cf. <i>ovata</i> Celebes Sea sub-clade; C: <i>Ostreopsis</i> cf. <i>ovata</i> South China Sea sub-clade; D: <i>Ostreopsis</i> cf. <i>ovata</i> Thailand sub-clade and; E: <i>Ostreopsis</i> cf. <i>ovata</i> Med/Pac clade. Full arrows represent the Hemi-CBCs between Celebes Sea and other sub-clades. Dashed arrows represent the Hemi-CBCs in Med/Pac sub-clade.	197
S7A Maximum Likelihood (ML) phylogenetic tree of various <i>Ostreopsis</i> cf. <i>siamensis</i> strains isolated along the New South Wales coastline using ITS1-5.8S-ITS2 primer set. Numbers at nodes represent posterior probabilities from Bayesian Inferences (BI) and bootstrap support values from ML based on 1000 pseudo-replicates. Only bootstrap values > 50% are shown. * represents 1, 100 support values for BI and ML respectively.	202
S7B Maximum Likelihood (ML) phylogenetic tree of various <i>Ostreopsis</i> cf. <i>siamensis</i> strains isolated along the New South Wales coastline using LSU D1-D3 rDNA region primer set. Numbers at nodes represent posterior probabilities from Bayesian Inferences (BI) and bootstrap support values from ML based on 1000 pseudo-replicates. Only bootstrap values > 50% are shown. * represents 1, 100 support values for BI and ML respectively.	203

S7C Maximum Likelihood (ML) phylogenetic tree of various <i>Ostreopsis</i> cf. <i>siamensis</i> strains isolated along the New South Wales coastline using LSU D8-D10 rDNA region primer set. Numbers at nodes represent posterior probabilities from Bayesian Inferences (BI) and bootstrap support values from ML based on 1000 pseudo-replicates. Only bootstrap values > 50% are shown. * represents 1, 100 support values for BI and ML respectively.	204
S8A List of all strains isolated, genotyped and used for physiological assessment from the New South Wales sampling sites.	205
S8B Mean growth rate (in div/day) (n=3), cell volume (in μM^3) (n=20) and PLTX-amine fragment equivalents (in pg cell ⁻¹) for <i>Ostreopsis</i> cf. <i>siamensis</i> strains isolated along the New South Wales coastline. SE represents the standard error.	207
S8C Steady-state growth and light-harvesting characteristics across <i>Ostreopsis</i> cf. <i>siamensis</i> isolates of photosystem II (PSII) maximum photochemical efficiency (Fv/Fm, dimensionless), PSII absorption cross-section (σ , nm ²), PSII reaction center content ([RCII], mol RCII m ⁻³ x 10 ⁻¹⁸ cell ⁻¹), Chlorophyll –a content (in pg cell ⁻¹) and PSU size (moles of chl/ moles of RCII). SE represents the standard error.	209
S8D Rapid light curve derived parameters for <i>Ostreopsis</i> cf. <i>siamensis</i> isolates representing maximum photosynthetic rate (ETR _{max}), the light saturation parameter (E _k), the light utilisation efficiency (α), photochemical (1-C) and non-photochemical quenching (1-Q). SE represents the standard error of n=3.	211
S9A Analysis of variance (one-way ANOVA) on growth rates, FvFm, σ , ETR _{max} , α , E _k and cellular biovolume for strains isolated from Minnie Water	213
S9B Analysis of variance (one-way ANOVA) on growth rates, FvFm, σ , ETR _{max} , α , E _k and cellular biovolume for strains isolated from Bonny Hills	214
S9C Analysis of variance (one-way ANOVA) on growth rates, FvFm, σ , ETR _{max} , α , E _k and cellular biovolume for strains isolated from Wallis Lake, Forster	215
S9D Analysis of variance (one-way ANOVA) on growth rates, FvFm, σ , ETR _{max} , α , E _k and cellular biovolume for strains isolated from Lake Macquarie	216
S9E Analysis of variance (one-way ANOVA) on growth rates, FvFm, σ , ETR _{max} , α , E _k and cellular biovolume for strains isolated from Patonga Creek	217
S9F Analysis of variance (one-way ANOVA) on growth rates, FvFm, σ , ETR _{max} , α , E _k and cellular biovolume for strains isolated from Gordons Bay	218

S9G Analysis of variance (one-way ANOVA) on growth rates, FvFm, σ , ETRmax, α , Ek and cellular biovolume for strains isolated from Kiama	219
S9H Analysis of variance (one-way ANOVA) on growth rates, FvFm, σ , ETRmax, α , Ek and cellular biovolume for strains isolated from Merimbula Lake inlet.	220
S10 Functional groupings based on all factors regulating the ETR (σ , Fv/Fm, cellular RCII concentration ([RCII]), [1 – C] and [1 – Q] across all types. Cluster analysis and multidimensional scaling (MDS) were performed on the average of each variable per variant; similarity is shown at the 90 and 95% levels and vectors driving the clustering are shown in black.	221
S11A Two-dimensional principal component analysis of phenotypic variables in <i>Ostreopsis</i> cf. <i>siamensis</i> clones. Light harvesting (Fv/Fm, σ , Cellular RCII concentration ([RCII])) and light utilization ([1-C] and [1-Q]) along with cell volume, toxin amounts and growth rates were normalised across all strains and indicated in black bars.	221
S11B Percentage variations in <i>Ostreopsis</i> cf. <i>siamensis</i> strains being explained by the factors making up the principal component axes.	222
S11C Eigenvectors: Coefficient in the linear combination of variables making up the Principal component axes	222
S12A Description of sequences from <i>Ostreopsis</i> cf. <i>ovata</i> HER27 encoding essential enzymes of various metabolic pathways. NA= enzymes not present in the transcriptome	223
S12B Description of sequences from <i>Ostreopsis</i> cf. <i>siamensis</i> BH1 encoding essential enzymes of various metabolic pathways. NA= enzymes not present in the transcriptome	234
S12C Description of sequences from <i>Ostreopsis rhodesae</i> HER26 encoding essential enzymes of various metabolic pathways. NA= enzymes not present in the transcriptome	244
S12D Description of sequences from <i>Coolia malayensis</i> MAB encoding essential enzymes of various metabolic pathways. NA= enzymes not present in the transcriptome	254
S13 Distributions of second level cellular component, biological processes and molecular function GO annotations in the annotated transcriptomes.	265
S14A List of transcripts encoding full ketoacyl synthase domain identified in <i>Coolia malayensis</i> MAB, <i>Ostreopsis</i> cf. <i>ovata</i> HER27, <i>Ostreopsis</i> cf. <i>siamensis</i> BH1 and <i>Ostreopsis rhodesae</i> HER26.	266

S14B List of transcripts encoding partial ketoacyl synthase domain identified in <i>Coolia malayensis</i> MAB, <i>Ostreopsis</i> cf. <i>ovata</i> HER27, <i>Ostreopsis</i> cf. <i>siamensis</i> BH1 and <i>Ostreopsis rhodesae</i> HER26.	273
S14C List of transcripts encoding full ketoacyl reductase domain identified in <i>Coolia malayensis</i> MAB, <i>Ostreopsis</i> cf. <i>ovata</i> HER27, <i>Ostreopsis</i> cf. <i>siamensis</i> BH1 and <i>Ostreopsis rhodesae</i> HER26.	276
S14D List of transcripts encoding partial ketoacyl reductase domain identified in <i>Coolia malayensis</i> MAB, <i>Ostreopsis</i> cf. <i>ovata</i> HER27, <i>Ostreopsis</i> cf. <i>siamensis</i> BH1 and <i>Ostreopsis rhodesae</i> HER26.	277
S15 List of multi-domain PKSs found in <i>Coolia malayensis</i> MAB, <i>Ostreopsis</i> cf. <i>ovata</i> HER27, <i>Ostreopsis</i> cf. <i>siamensis</i> BH1 and <i>Ostreopsis rhodesae</i> HER26	280
S16 Top BLAST hits of polyketide synthase genes found in <i>Coolia malayensis</i> MAB, <i>Ostreopsis</i> cf. <i>ovata</i> HER27, <i>Ostreopsis</i> cf. <i>siamensis</i> BH1 and <i>Ostreopsis rhodesae</i> HER26 transcriptomes (492 contigs out of 557)	282
S17 List of putatively annotated entities featured in the non-targeted metabolite analyses of <i>Ostreopsis</i> cf. <i>ovata</i> HER27, <i>Ostreopsis</i> cf. <i>siamensis</i> BH1, <i>Ostreopsis rhodesae</i> HER26 and <i>Coolia malayensis</i> MAB.	315

PREFACE

Lewis Carroll's 'Alice's Adventures in Wonderland' (1865) and 'Through the Looking-Glass and What Alice Found There' (1871) have fascinated readers for generations and have had a considerable impact on popular culture. Characters and references from these books have been used by scientists to explain the intricate phenomenon in microbial ecology, and particularly in marine microbial eukaryotes. Van Valen's 'Red Queen' hypothesis as a metaphor for an evolutionary 'arms race' and the 'Cheshire Cat' as a symbol of the complex phenomenon of sexual reproduction in *Emiliana huxleyi* are a few popular examples.

'Through the looking glass' depicts mirrors as a gateway to the wonderland and reflective of how mirrors are often illusionary. The book's finale centred around a game of chess where Alice finds herself as a pawn in the bigger game, highlights the importance of strategy to survive. The 'mirrors' and 'chess' from the storyline are symbolic of cryptic diversity and functional traits that exist in marine microbial eukaryotes, at a species, population, genetic and metabolic levels, enabling them to survive in the changing oceanic conditions. Cryptic diversity and strategic functional traits in *Ostreopsis* species are the fundamental questions that I have pursued in my dissertation and hence used the book title as a reference to symbolise the details of my aims and findings. The reference also highlights the enigmatic 'wonderland' of marine microbial eukaryotes that we witness through the lenses of a microscope (looking glass).

***"There is a place like no other on earth. A land full of wonder, mystery and danger!
Some say, to survive it, you need to be as mad as a hatter."***

-Lewis Carroll, *Alice's Adventures in Wonderland*

THESIS ABSTRACT

Marine microbial eukaryotes are of immense ecological and evolutionary significance in marine ecosystems. Understanding their biodiversity and functional evolutionary traits are key to improving our understanding of marine ecosystem functioning. The East Australian Current (EAC) is a global climate change hotspot, and yet we lack in our understanding of its impact on phytoplankton distribution and dynamics. *Ostreopsis* species have been reported to cause severe blooms and produce palytoxin (PLTX) - like compounds all around the globe but we do not have basic information on the distribution and dynamics of *Ostreopsis* species in Australia.

In this dissertation, I established the first comprehensive report of *Ostreopsis* species from Australian waters and explored cryptic diversity and functional traits in this genus. Extensive sampling along a north-south gradient of 1800 km from sub-tropical to temperate waters yielded the identification of three species, including a novel pseudo-cryptic *Ostreopsis rhodesae* from the Great Barrier Reef, along with *Ostreopsis* cf. *ovata*. *Ostreopsis* cf. *siamensis* was identified at all locations and its eco-physiological traits and genetic population structure were investigated. The genetic diversity in the northern sub-tropical locations was greater compared to the more southern locations, reflecting a long-standing divergence and local radiations originating from the ancestral population and a potential southward range expansion, which may be related to the intensification of the EAC over the past century.

Intra- and inter-population variations in physiological traits were investigated to understand its range expansion and functional trade-offs. This is the first study to our knowledge to report growth rates, cell size, cellular toxic concentrations and photobiological parameters on fifty-three clones of a marine protist, in order to investigate intra-specific diversity in key functional traits. The toxin biosynthesis pathway in the three species was investigated using *de novo* transcriptomics and compared to *Coolia malayensis*. All essential domains needed to synthesize a PLTX-like carbon backbone were identified in the three *Ostreopsis* species and were also found in the non-PLTX producing *C. malayensis*. Putative molecules with potential polyketide-like backbone structures were reported in this investigation using non-targeted metabolomics,

suggesting a greater diversity of polyketide compounds amongst these species than previously anticipated.

Results from this dissertation add to the knowledge of species biodiversity, population structures, eco-physiological traits and toxin biosynthesis mechanisms in marine microbial eukaryotes, and *Ostreopsis* species in particular.

