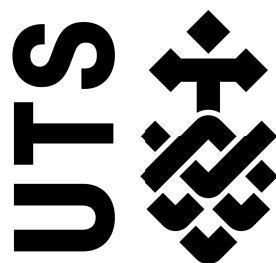


Delving into the Genetic Code of *Gambierdiscus* - the Devil is in the Detail

Thesis by

Anna Liza Kretzschmar

In Partial Fulfillment of the Requirements
for the Degree of
Doctor of Philosophy



University of Technology Sydney
Sydney, NSW, Australia

2019

© 2019

Anna Liza Kretschmar

All Rights Reserved

*Little one better find your way now,
Devil is fine,
Little one better find your way out,
Devil is kind,
Little one better run for your life,
Devil is fine,
Little one where you going with that knife?*

- Zeal & Ardor

Certificate of Original Authorship

I, Anna Liza Kretschmar, declare that this thesis is submitted in fulfillment of the requirements for the award of Doctor of Philosophy, in the Faculty of Science at the University of Technology Sydney. This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution. This research is supported by the Australian Government Research Training Program.

Production Note:

Signature removed prior to publication.

Date: 2019-07-15

Anna Liza Kretschmar

Acknowledgments

This thesis has been a harrowing journey through supervisory change, chronic mental and physical illness, and encountering road blocks rather than the mentorship which one would hope for at crucial junctions for becoming a scientist adherent to the scientific principles of transparency and objectivity.

The fact that I am still here, and smiling, is due to the love and support I was lucky enough to receive. A monumental thank you to my supervisor Aaron Darling for staying with me through this journey through the hard times, and doing so in ever present good humor.

Thank you to my partners, past and present, Kurtis Crisafulli and Leroy Jeffrey, for your love, hugs and making me remember there is more to life.

Thank you to Samantha Goyen, for having an excellent time full of laughing fits and picnics while on day and night long writing benders.

While unusual, it has been requested that I include the following section:

Acknowledgment to Shauna Murray for scholarship funding which amounted to 17.3 % of my candidature. Without your input I would not be the robust person and scientist I am today.

All other funding sources will be acknowledged with gratitude within the relevant chapters.

A. Liza Kretzschmar
Sydney, Australia, 2019.

Thesis Format

A thesis by compilation.

List of Publications

Peer reviewed

1. Murray, S. A., Ajani, P.; Kretzschmar, A. L. & Verma, A. (2017) *Response to " More surprises in the global greenhouse: Human health impacts from recent toxic marine aerosol formulations, due to centennial alterations or world-wide coastal food web"s.* Mar. pol. bulletin. 123(1-2):415-417.
2. Kohli, G.; Haslauer, K.; Chowdhury, S.; Boulter, M.; Kretzschmar, A. L.; Harwood, T.; Laczka, O. & Murray, S. (2017) *Qualitative and Quantitative Assessment of the Presence of Ciguatoxin, P-CTX-1B, in Spanish Mackerel (Scomberomorus commerson) from waters in NSW (Australia).* Toxicol. reports 4:328-334.
3. Kretzschmar, A. L.; Verma, A.; Harwood, T.; Hoppenrath, M. & Murray, S. (2017) *Characterisation of Gambierdiscus lapillus sp. nov. (Gonyaulacales, Dinophyceae): A new toxic dinoflagellate from the Great Barrier Reef (Australia).* J Phycol 53(2):283-297. **Awarded cover image.**

Under review

4. Hoppenrath, M.; Kretzschmar, A. L.; Kaufmann, M.; & Murray, S. (2018) *Morphological and molecular phylogenetic identification and record verification of Gambierdiscus excentricus (Dinophyceae) from Madeira Island (NE Atlantic Ocean).* In review, Mar. Biodiv. Records (2018).
5. Kretzschmar, A. L.; Larsson, M.; Doblin, M.; & Murray, S. *Characterisation of two toxic Gambierdiscus spp. (Gonyaulacales, Dinophyceae) from the Great Barrier Reef (Australia): G. lewisii sp. nov. and G. holmesii sp. nov.* In review, Protist (2018).
6. Kretzschmar, A. L.; Verma, A.; Kohli, G.; & Murray, S. *Development of a quantitative PCR assay for the detection and enumeration of a potentially ciguatoxin-producing dinoflagellate, Gambierdiscus lapillus (Gonyaulacales, Dinophyceae).* In review, PLoS ONE (2019).
7. Kretzschmar, A. L.; Verma, A.; Murray, S.; Kahlke, A; Fourment, M. & Darling, A. *Trial by phylogenetics - Evaluating the Multi-Species Coalescent for evolutionary inference on taxa*

with high levels of paralogy (Gonyaulacales, Dinophyceae). In review, PeerJ (2019).

Contents

Certificate of Authorship	iv
Acknowledgments	v
Thesis Format	vi
List of Publications	vii
Abstract	1
1 General introduction.	2
1.1 Ciguatera fish poisoning	3
1.1.1 Symptoms of ciguatera	3
1.1.2 Polyketide toxins produced by <i>Gambierdiscus</i> spp.	4
1.1.2.1 Ciguatoxins	4
1.1.2.2 Maitotoxins	7
1.1.2.3 Other toxins	8
1.1.2.4 Utilization of bioassays	9
1.2 The genus <i>Gambierdiscus</i>	10
1.2.1 Taxonomy	10
1.2.1.1 Morphology	10
1.2.2 Distribution of <i>Gambierdiscus</i> spp.	14
1.2.2.1 Geographic differences in <i>Gambierdiscus</i> distribution	14
1.2.2.2 Depth dependent variation of <i>Gambierdiscus</i> species	14
1.2.2.3 Local variation <i>Gambierdiscus</i> abundances	14
1.2.2.4 <i>Gambierdiscus</i> in Australia	15
1.2.2.5 Changes in distribution due to climate change	16
1.3 Dinoflagellate genetics	17
1.3.1 Polyketide synthase genes	18

1.3.1.1	Polyketide synthase genes in <i>Gambierdiscus</i>	19
1.4	Thesis aims	21
2	Characterization of <i>Gambierdiscus lapillus</i> sp. nov. (Gonyaulacales, Dinophyceae): A new toxic dinoflagellate from the Great Barrier Reef (Australia).	22
2.1	Authorship Declaration	22
2.2	Abstract	24
2.3	Introduction	24
2.4	Materials and methods	26
2.4.1	Specimen collection and culture conditions	26
2.4.2	Morphological analyses	28
2.4.3	Genomic DNA extraction	29
2.4.4	PCR amplification and sequencing	29
2.4.5	Sequence alignment and phylogenetic analysis	31
2.4.6	Toxin production	31
2.4.6.1	Toxicity via mouse bioassay	31
2.4.6.2	Toxin profile screening by liquid chromatography-mass spectrometry	32
2.5	Results	32
2.5.1	Morphology and taxonomical treatment	32
2.5.2	Phylogenetics	40
2.5.3	Toxicology	43
2.5.3.1	Toxicity via intraperitoneal injection	43
2.5.3.2	Toxicity via gavage	44
2.5.3.3	Toxin profile	44
2.6	Discussion	44
2.7	Acknowledgements	48
2.8	Conflict of interest disclosure	49
2.9	Supplementary Material	49
3	Development of a quantitative PCR assay for the detection and enumeration of a potentially ciguatoxin-producing dinoflagellate, <i>Gambierdiscus lapillus</i> (Gonyaulacales, Dinophyceae).	52
3.1	Authorship Declaration	52
3.2	Abstract	54
3.3	Introduction	54
3.4	Materials and methods	59
3.4.1	Clonal strains and culturing conditions	59
3.4.2	DNA extraction and species specific primer design	60

3.4.3	Evaluation of primer specificity	61
3.4.4	Evaluation of primer sensitivity	61
3.4.4.1	Gene based calibration curve	61
3.4.4.2	Cell based calibration curve	62
3.4.5	Determination of gene copies per cell for <i>G. lapillus</i>	62
3.4.6	Screening environmental samples for <i>G. lapillus</i>	62
3.5	Results	64
3.5.1	Evaluation of primer specificity	64
3.5.2	Evaluation of primer sensitivity	66
3.5.3	Quantification of SSU rDNA copy number per cell of <i>G. lapillus</i>	69
3.5.4	Screening environmental samples for <i>G. lapillus</i> abundance	69
3.6	Discussion	72
3.7	Conclusion	74
3.8	Acknowledgements	74
3.9	Supplementary Material	74
4	Using comparative transcriptomics for the pan-transcriptome of <i>Gambierdiscus</i>.	78
4.1	Abstract	79
4.2	Introduction	79
4.3	Methods	82
4.3.1	Transcriptome acquisition	82
4.3.2	Spliced leader search	83
4.3.3	Homolog clustering	83
4.3.4	Comparison of gene ontology annotations	84
4.3.5	Ketosynthase domain search	84
4.4	Results	84
4.4.1	Overview of the transcriptomes	84
4.4.1.1	Level 1 GO annotations between <i>Gambierdiscus</i> species	87
4.4.1.2	Level 2 GO annotations between <i>Gambierdiscus</i> species	91
4.4.2	Pan-transcriptomics of <i>Gambierdiscus</i>	95
4.4.2.1	Core transcriptome	95
4.4.2.2	Softcore transcriptome	95
4.4.2.3	The unique portion of the transcriptome	95
4.4.3	Level 1 GO annotations for pan-transcriptomes	96
4.4.4	Level 2 GO annotations for pan-transcriptomes	100
4.4.5	Ketosynthase active domain search	104
4.5	Discussion	106
4.5.1	Core, softcore and unique clusters	106
4.5.2	Ketosynthase domain detection	107

4.5.3	Areas for possible improvement in this study	107
4.5.3.1	Coverage of taxa	108
4.5.3.2	PKS active domain search	108
4.6	Conclusion	108
5	Trial by phylogenetics - Evaluating the Multi-Species Coalescent for evolutionary inference on taxa with high levels of paralogy (Gonyaulacales, Dinophyceae).	109
5.1	Authorship Declaration	109
5.2	Abstract	111
5.3	Introduction	111
5.3.1	Summary: Statistical nomenclature & errors this study seeks to address . .	113
5.3.2	Summary: Who/what are the Gonyaulacales?	114
5.4	Materials and methods	115
5.4.1	Culture conditions	115
5.4.2	RNA isolation, library preparation and Sequencing	115
5.4.2.1	Publicly available transcriptome libraries	115
5.4.3	Transcriptome processing scripts	115
5.4.3.1	Transcriptome assembly	116
5.4.3.2	Construction of multiple sequence alignments	116
5.4.4	Assembly analysis	116
5.4.5	Phylogenetic inferences	116
5.4.5.1	Ribosomal DNA based inference	116
5.4.5.2	Concatenated single copy gene based inference	117
5.4.5.2.1	Maximum likelihood with concatenated sequences	117
5.4.5.2.2	Bayesian inference with concatenated sequences	117
5.4.5.2.3	Bayesian probability under the MSC	117
5.4.6	Stepping stone analysis	117
5.4.6.1	Generation of figures	117
5.4.6.2	Implementation	117
5.5	Results	119
5.5.1	Transcriptomes overview	119
5.5.2	Single copy gene search with BUSCOv2	120
5.5.3	Phylogenetic inferences	120
5.5.3.1	rDNA based phylogeny	121
5.5.3.2	Concatenated single copy gene based phylogeny inferred with ML	121
5.5.3.3	Concatenated single copy gene based phylogeny inferred with BI	122
5.5.3.4	Single copy gene based phylogeny under MSC	123
5.6	Discussion	124
5.6.1	Considerations for data set selection and pre-processing	125

5.6.1.1	Quantity of taxa in phylogenetic inference	125
5.6.1.2	Quality of transcriptome assemblies	125
5.6.1.2.1	Assembly parameters	126
5.6.1.3	Selection of paralogs to infer species evolution	126
5.6.1.4	Model selection for inference	128
5.6.2	Comparison to commonly employed models and data sets	128
5.6.2.1	Phylogenetic inference using ribosomal genes	128
5.6.2.2	Concatenating selected genes and using ML methods for species inference	131
5.6.2.3	Concatenating selected genes and using BI methods for species inference	133
5.6.3	Areas for possible improvement of this study	135
5.6.3.1	Contamination of other taxa	135
5.6.3.2	No representative genome for comparison	135
5.6.3.3	Different methods for RNA-seq	135
5.6.3.4	Total evidence phylogenetics	135
5.6.3.5	Model comparison methods	135
5.7	Conclusion	136
5.8	Acknowledgments	137
5.9	Supplementary material	137
6	Discussion	141
6.1	General discussion	141
6.1.1	Targeting the species at bottom of the food chain	141
6.1.2	Targeting toxin production potential directly	141
6.2	Significance and future directions	142
6.3	Thesis conclusion	142
7	Bibliography	143
8	Appendix.	170
8.1	Published manuscripts	170
8.1.1	Qualitative and quantitative assessment of the presence of ciguatoxin, P- CTX-1B, in Spanish Mackerel (<i>Scomberomorus commerson</i>) from waters in New South Wales (Australia)	170
8.1.2	Response to “More surprises in the global greenhouse: Human health im- pacts form recent toxic marine aerosol formulations, due to centennial al- terations or world-wide coastal food webs”	170
8.2	Submitted manuscripts	171

8.2.1	Morphological and molecular phylogenetic identification and record verification of <i>Gambierdiscus excentricus</i> (Dinophyceae) from Madeira Island (NE Atlantic Ocean).	171
8.2.2	Characterisation of two toxic <i>Gambierdiscus</i> spp. (Gonyaulacales, Dinophyceae) from the Great Barrier Reef (Australia): <i>G. lewisii</i> sp. nov. and <i>G. holmesii</i> sp. nov.	171
8.2.3	Development of a quantitative PCR assay for the detection and enumeration of a potentially ciguatoxin-producing dinoflagellate, <i>Gambierdiscus lapillus</i> (Gonyaulacales, Dinophyceae).	171
8.2.4	Trial by phylogenetics - Evaluating the Multi-Species Coalescent for evolutionary inference on taxa with high levels of paralogy (Gonyaulacales, Dinophyceae).	171

List of Figures

2.1	Australian sites in which <i>Gambierdiscus</i> have been detected.	27
2.2	Light micrographs of <i>Gambierdiscus lapillus</i> sp. nov.	33
2.3	Scanning electron micrographs of <i>Gambierdiscus lapillus</i> hypotheca	35
2.4	Scanning electron micrographs of the epitheca of <i>Gambierdiscus lapillus</i>	36
2.5	Scanning electron micrographs of <i>Gambierdiscus lapillus</i> showing the reticulate-foveate thecal ornamentation.	37
2.6	Line drawings showing the thecal tabulation pattern of <i>Gambierdiscus lapillus</i> . . .	39
2.7	Maximum likelihood phylogeny of <i>Gambierdiscus</i> species/phylotypes of the LSU D8-D10 rDNA region.	41
2.8	Maximum likelihood phylogeny of <i>Gambierdiscus</i> species/phylotypes of the SSU rDNA region.	42
2.9	Visual representation of the size difference between <i>Gambierdiscus</i> and <i>Fukuyoa</i> species.	46
2.10	Suppl. scanning electron micrographs of <i>Gambierdiscus lapillus</i> (strain HG26). . .	50
2.11	Suppl. scanning electron micrographs of <i>Gambierdiscus lapillus</i> showing sulcal and cingular details.	50
2.12	Suppl. scanning electron micrographs of <i>Gambierdiscus lapillus</i> (strain HG7) of morphological intra-strain variability.	51
3.1	Bioaccumulation of ciguatoxin through the food web.	55
3.2	<i>Gambierdiscus</i> sampling site on Heron Island, Queensland.	63
3.3	Cell based quantitative PCR standard curve.	67
3.4	Gene fragment based quantitative PCR curve.	68
3.5	<i>Gambierdiscus lapillus</i> detection on macroalgae across environmental sampling sites.	70
3.6	<i>Gambierdiscus lapillus</i> enumeration on macroalgae across environmental sampling sites.	71
4.1	Venn diagram of distribution of nucleotide clusters between the five <i>Gambierdiscus</i> species.	86

4.2	Summary of GO annotations between <i>Gambierdiscus</i> species that mapped to molecular functions at GOSUM level 1.	88
4.3	Summary of GO annotations for cellular components between <i>Gambierdiscus</i> species at GOSUM level 1.	89
4.4	Summary of GO annotations in the biological processes category, between <i>Gambierdiscus</i> species at GOSUM level 1.	90
4.5	Summary of GO annotations for biological processes between <i>Gambierdiscus</i> species at GOSUM level 2.	92
4.6	Summary of GO annotations matching to cellular components between <i>Gambierdiscus</i> species at GOSUM level 2.	93
4.7	Summary of GO annotations for molecular functions between <i>Gambierdiscus</i> species at GOSUM level 2.	94
4.8	Summary of pan-transcriptome clusters matching to cellular component GO annotations at GOSUM level 1.	97
4.9	Summary of biological processes GO annotations of pan-transcriptome at GOSUM level 1.	98
4.10	Summary of level 1 molecular function GO annotations of pan-transcriptome.	99
4.11	Summary of cellular component category GO annotations for the pan-transcriptome at level 2.	101
4.12	Summary of level 2 molecular function GO annotations for pan-transcriptome clusters.	102
4.13	Summary of biological processes annotations for the pan-transcriptome clusters at level 2.	103
4.14	Venn diagram of the ketosynthase domain clusters across five <i>Gambierdiscus</i> spp.	105
5.1	Maximum likelihood phylogenetic inference of concatenated ribosomal DNA genes.	121
5.2	Maximum likelihood phylogenetic inference of concatenated single copy gene set.	122
5.3	Bayesian phylogenetic inference of concatenated single copy gene set.	123
5.4	Bayesian phylogenetic inference of single copy gene set under multispecies coalescent.	124
5.5	Tanglegram of the single copy gene topology comparison between this study and Price et al. (2017).	128
5.6	Tanglegram comparing the topology between concatenated ribosomal DNA sequences and Bayesian inference of single copy genes under the multispecies coalescence.	130
5.7	Tanglegram comparing the topology of single copy gene inference generated with concatenated maximum likelihood vs. Bayesian inference under the multispecies coalescence.	132

5.8 Tanglegram comparing the topologies from the single copy gene set with Bayesian inference concatenated vs. under the multispecies coalescence. 134

List of Tables

1.1	Ciguatoxin congeners identified and characterized to date.	6
1.2	Maitotoxin congeners identified and characterized to date.	8
1.3	Presently characterized <i>Gambierdiscus</i> spp. and their taxonomic identifications. . .	11
2.1	<i>Gambierdiscus</i> spp. identified around Australia.	28
2.2	Primer sets used for phylogenetic analysis of <i>Gambierdiscus</i> spp.	30
2.3	Morphological measurements of <i>G. lapillus</i>	32
2.4	Morphological characteristics of sister species to <i>Gambierdiscus lapillus</i>	45
3.1	Published qPCR assays for <i>Gambierdiscus</i> and <i>Fukoyoa</i> spp.	58
3.2	Ciguatera cases reported in Queensland, 2011 to 2015	58
3.3	List of <i>Gambierdiscus</i> clonal strains used for the qPCR assay.	60
3.4	<i>G. lapillus</i> specific qPCR primer set for SSU rDNA designed in this study.	60
3.5	Quantitative PCR primer set cross reactivity.	65
3.6	Suppl. macroalgagl samples screened for <i>Gambierdiscus lapillus</i>	74
4.1	<i>Gambierdiscus</i> species transcriptomes used in this study along with associated in- formation.	83
4.2	Progression of clusters found in each <i>Gambierdiscus</i> transcriptome during processing.	85
5.1	Summary of transcriptome sequencing and assembly statistics.	120
5.2	Suppl. culturing conditions for species processed for this study.	137
5.3	Transcriptomes and associated information used for study.	138
5.4	Accession numbers for ribosomal DNA sequences.	139

Abstract

Ciguatera is a global problem but predominantly affects communities in the Pacific Island Nations, which is predicted to increase in prevalence and severity as the effects of climate change unfold. Despite the menace posed by ciguatera, very little is known about how to monitor for risk of outbreaks and report rates of the disease are estimated around 20 % at best. To address this, a global ciguatera strategy was called into action by an international body of experts and endorsed by UNESCO's Intergovernmental Oceanographic Commission, which have called for ciguatera to be considered a neglected tropical disease, which needs to be characterised as a matter of urgency.

Element 1 of the global ciguatera strategy revolves around the characterization of *Gambierdiscus* species and their distribution, as well as their toxin production. This thesis contributes to Element 1 by investigating the diversity of *Gambierdiscus* around Heron Island, a region where ciguatera is endemic, developing a monitoring tool for the species discovered there and then characterizing the evolution of the genus in the context of other closely related toxic genera as well as within the species. It contributes to the understanding of *Gambierdiscus* diversity, the species concept with description of a new species, adds information to ameliorate the scarcity of genetic data and molecular markers as well as monitoring tools for *Gambierdiscus*, and explores evolution of the Gonyaulacales.

Dissertation directed by Associate Professor Aaron E. Darling
iThree institute, Faculty of Science
University of Technology Sydney