Metabolic Profiling of Environmental Stress in Scleractinian Corals

Benjamin R. Gordon
B.Chem.Sci. (hons)

For the degree of Doctor of Philosophy
College of Public Health, Medical and Veterinary Science
James Cook University
August 2020

ACKNOWLEDGEMENTS

I would like to begin by expressing my deepest gratitude to my advisory team. Firstly, to Dr. Cherie Motti for her patience, encouragement and support. Her attention to detail, immense knowledge and willingness to challenge my way of thinking have improved this research immensely. I could not have imagined having a better supervisor and mentor. To A/Prof Bill Leggat for having the courage to support a novel metabolomics project and providing me with the assistance and freedom to explore new lines of enquiry – this research would not have been possible without him. And finally, to Prof Bruce Bowden, who willingly gave up time in his retirement to support me and this research.

A special thank you to all who were instrumental to the success of this project: Sarah Gierz, Lubna Ukani and Daisie Ogawa, for their hard work, help and assistance in both the lab and the field; Adrian Lutz, for his support, friendship and assistance with all things coral and biochemical; Teressa Bobeszko, for her support and friendship when things got difficult and; Marnie Freckelton, who gave up valuable time to assist with round-the-clock experiments at the Australian Synchrotron.

I would also like to thank all the wonderful people at Metabolomics Australia, the Australian Synchrotron and the Australian Institute of Marine Science who took the time to impart their knowledge and assist with all the analytical chemistry that was conducted during this research.

This work would not have been realised without the patience and endless support of my family, to them, especially my wife Danielle, goes the biggest thank you of all.

PUBLICATION LIST

Peer reviewed articles

1. Gordon, B., Martin, D., Bambery, K. & Motti, C. (2018) Chemical imaging of a *Symbiodinium* sp. cell using synchrotron infrared microspectroscopy: a feasibility study. *Journal of microscopy,* **270,** 83-91
2. Gierz, S. L., Gordon, B. R. & Leggat, W. (2016) Integral light-harvesting complex expression in *Symbiodinium* within the coral *Acropora aspera* under thermal stress. *Scientific reports,* **6,** 25081
3. Andreakis, N., Høj, L., Kearns, P., Hall, M. R., Ericson, G., Cobb, R. E., Gordon, B. R. & Evans-Illidge, E. (2015) Diversity of marine-derived fungal cultures exposed by DNA barcodes: The algorithm matters. *PloS one,* **10,** e0136130
4. Ovenden, S. P. B., Gordon, B. R., Bagas, C. K., Muir, B., Rochfort, S. & Bourne, D. J. (2010) A Study of the Metabolome of *Ricinus communis* for Forensic Applications. *Australian Journal of Chemistry,* **63,** 8-21
5. Gordon, B. R. & Leggat, W. (2010) *Symbiodinium*-Invertebrate Symbioses and the Role of Metabolomics. *Marine Drugs,* **8,** 2546-2568

Book chapters

1. Gordon, B. R., Leggat, W. & Motti, C. A. (2013) Extraction Protocol for Nontargeted NMR and LC-MS Metabolomics-Based Analysis of Hard Coral and Their Algal Symbionts. In: *Metabolomics Tools for Natural Product Discovery: Methods and Protocols* (eds. U. Roessner & D. A. Dias). Humana Press, Totowa, NJ

Data and software repositories

1. Gordon, B. R. & Motti, C. A. (2020) Revealing Taxonomic and Diel Variation Through Untargeted Metabolomics. v0.0.1 ed. Zenodo
2. Gordon, B. R. & Motti, C. A. (2020) coralmz: Coral Mass Spectral Features. v1.0.3 ed. Zenodo
3. Gordon, B. R. (2020) Predicting the Photosynthetic Efficiency of *Acropora aspera*. v1.0.2 ed. Zenodo

4. Gordon, B. R. (2020) coralclass: Classification of Coral Phenotypes. v1.0.1 ed. Zenodo

Reports

1. Ovenden, S. P. B., Gordon, B. R., Bagas, C. K., Muir, B., Rochfort, S. & Bourne, D. J. (2009) Cultivar determination of *Ricinus communis* via the metabolome: a proof-of-concept investigation. *Pandora Electronic Collection*.

Conference presentations

1. Gordon, B. R. (2012) Metabolomics: The Metabolome of *Symbiodinium* and their Coral Hosts. In: *International Coral Reef Society Conference (oral presentation)*. Cairns, Australia
2. Gordon, B. R. (2011) Climate Change and the Metabolome of Reef-Building Coral. In: *7th International Conference of the Metabolomics Society (poster presentation)*. Cairns, Australia
3. Gordon, B. R. (2010) The Metabolome of *Symbiodinium* Phylotypes and their Coral Hosts: Responses to Environmental Stressors. In: *2nd Australian Symposium on Metabolomics (poster presentation)*. Melbourne

STATEMENT ON THE CONTRIBUTION OF OTHERS

Nature of Assistance	Contribution	Name and Affiliation
Project and Thesis		
Intellectual support	Intellectual, design and editorial support	Dr. Cherie A. Motti[1] and A/Prof. William Leggat[2]
	Assistance with data interpretation and analysis	Dr. Cherie A. Motti[1]
Financial Support	Australian Post graduate Award	Australian Government
	Stipend	School of Public Health Medicine and Veterinary Sciences JCU
		AIMS@JCU
	Science for Management Grant	Great Barrier Reef Marine Park Authority
	Graduate Research Scheme	School of Public Health Medicine and Veterinary Science JCU
Chapter 3 – Methods		
Intellectual Support	Intellectual, design and editorial support	Dr. Cherie A. Motti[1] and A/Prof. William Leggat[2]
	Assistance with data interpretation and analysis	Dr. Cherie A. Motti[1]
Financial Support	Research	A/Prof. William Leggat lab
		Australian Institute of Marine Science
Data Collection	Field assistance	A/Prof. William Leggat[2] and Dr. Daisie Ogawa[3]
	Lab Assistance	Dr. Cherie A. Motti[1]
	Analytical Assistance	Dr. Cherie A. Motti[1]
Chapter 4 – CO_2		
Intellectual Support	Intellectual and design support	Dr. Cherie A. Motti[1], A/Prof. William Leggat[2] and Prof. Ute Roesnner[4]
	Editorial support	Dr. Cherie A. Motti[1] and A/Prof. William Leggat[2]
	Assistance with data interpretation and analysis	Dr. Cherie A. Motti[1] and Dr. Berin Boughton[4]
Financial Support	Research	A/Prof. William Leggat lab
		Australian Institute of Marine Science
		Great Barrier Reef Foundation
Data Collection	Field assistance	A/Prof. William Leggat[2] and Ms Daisie Ogawa[3]
	Lab Assistance	Dr. Cherie A. Motti[1]
	Analytical Assistance	Dr. Cherie A. Motti[1] and Dr. Berin Boughton[4]
Chapter 5 – Bleaching		

Nature of Assistance	Contribution	Name and Affiliation
Intellectual support	Intellectual and design support	Dr. Cherie A. Motti[1], A/Prof. William Leggat[2] and Prof. Ute Roesnner[4]
	Editorial support	Dr. Cherie A. Motti[1]
	Assistance with data interpretation and analysis	Dr. Cherie A. Motti[1]
Financial Support	Research	A/Prof. William Leggat lab
		Great Barrier Reef Foundation
Data Collection	Field assistance	Dr. Sarah Gierz[3]
	Lab Assistance	Dr. Sarah Gierz[3]
	Analytical Assistance	Dr Adrian Lutz[4]
Chapter 6 – Diel variation		
Intellectual support	Intellectual and design support	Dr. Cherie A. Motti[1], A/Prof. William Leggat[2] and Prof. Ute Roesnner[4]
	Editorial support	Dr. Cherie A. Motti[1]
	Assistance with data interpretation and analysis	Dr. Cherie A. Motti[1]
Financial Support	Research	A/Prof. William Leggat lab
		Great Barrier Reef Foundation
Data Collection	Field assistance	Mrs. Danielle Gordon[5]
	Lab Assistance	Mrs. Danielle Gordon[5]
	Analytical Assistance	Dr Adrian Lutz[4]

[1] *Australian Institute of Marine Science. Townsville, Queensland, Australia*
[2] *School of Environmental and Life Sciences, The University of Newcastle, Newcastle, Australia*
[3] *School of Public Health, Medicine and Veterinary Sciences, James Cook University, Townsville, Australia*
[4] *Metabolomics Australia, The University of Melbourne, Melbourne, Australia*
[5] *Faculty of Education, Queensland University of Technology, Brisbane, Australia*

ABSTRACT

The Great Barrier Reef (GBR), encompassing ~3000 coral reefs and spanning an area of 345,000 km^2 along the northeast coast of Australia, is the largest of the world's coral reefs. Globally, it provides a diverse range of essential goods and services such as, food, coastal protection, income, employment, and the discovery of new drugs and biochemicals. Unfortunately, coral reefs worldwide have been degraded by centuries of overfishing and pollution and are now facing further pressure from ocean warming and acidification driven by increasing atmospheric CO_2 concentrations. How this important bioresource is managed under future climate change scenarios is crucial to maintaining the goods and services it provides.

The traditional approach to reef management involves monitoring and assessing the state of the reef ecosystem, as well as impacts and environmental drivers of change; however, due to a lack of detailed indicators, this approach does not fully reflect the health and functional state of corals, nor many of the ecological and biochemical processes that underpin the entire system. Metabolic biomonitoring, which incorporates metabolic profiling and modern chemometrics to monitor the biochemical response of an organism to changes in its environment, may provide a higher resolution solution to fill the current gaps in coral reef monitoring.

This research explores the potential of metabolic profiling and machine learning to search for chemical bioindicators of the health and functional state of coral reefs. This will further our understanding of the complex biological processes that underpin the biological stress response by characterising the molecular features driving the prediction of specific stress responses. The grey and published literature is reviewed to establish *Symbiodiniaceae*-invertebrate symbioses and their metabolic interactions and identify potential knowledge gaps. Initially published in 2010, chapter 2 has been updated to include the current (2010-2019) coral metabolomics literature.

Metabolic biomonitoring requires a thorough and reproducible snapshot of the coral metabolome. Snapshots are affected by the sample preparation and handling procedures that are common to many studies; however, standard metabolic profiling protocols may be unsuitable for the analysis of corals due to the biological complexity of the holobiont.

Chapter 3 presents a protocol for the sample collection, extraction and measurement of hard coral holobiont metabolites using proton Nuclear Magnetic Resonance (^{1}H NMR) spectroscopy and Liquid Chromatography coupled with Mass Spectrometry (LC-MS). It provides all the details required to replicate the method, including the hazards and risks that were discovered during its development. Recently reviewed, the 70% methanol extraction method developed here is now a commonly used method in coral metabolomics.

Discrete descriptors of coral health, such as *stressed* or *healthy*, provide reef managers with coral phenotypes associated with their health for guiding interventions; they can also be supplied to machine learning algorithms as class labels. With this in mind, Chapter 4 explores the capability of two machine learning models – Partial Least Squares Discriminant Analysis and Random Forests (RF) – to classify the ^{1}H-NMR and LC-MS metabolic profiles of *Acropora aspera* during exposure to a simulated bleaching event and pCO_2 levels consistent with the mid-century projections of the representative concentration pathway. Elevated temperature was the major contributor to coral stress while elevated pCO_2 somewhat ameliorated temperature stress and increased symbiont photosynthetic capacity. LC-MS, combined with RF modelling was the most accurate combination of analytical platform and modelling algorithm, correctly classifying corals at different stages of exposure. This approach identified perturbations in the metabolome of corals exposed to realistic mid-century levels of pCO_2; surpassing the sensitivity of more traditional methods such as symbiont cell density and Pulse-amplitude modulation (PAM) fluorometry. This work clearly demonstrates the potential utility of metabolic profiling for precise monitoring purposes.

In contrast to the discrete descriptors predicted by classification models, regression models predict a continuous descriptor. A continuous descriptor based on a quantifiable measure, such as the maximum quantum yield of photosystem II (F_v/F_m), may provide more precise indicators of coral health. In Chapter 5, RF predicted the F_v/F_m during cross-validation to within ± 0.0371 of the true values, while unseen test samples were predicted to within ± 0.0406 of their true values. Twenty spectral features that were important predictors of F_v/F_m were revealed, three of which were provided with a putative compound identification after a thorough search of the coral research literature. LC-MS-based profiling predicted a quantitative measure of coral health with meaningful precision, while simultaneously providing valuable information about the coral metabolome.

Natural variation in coral metabolomes has the potential to introduce unwanted bias in machine learning models, which may undermine the confidence of model predictions and/or lead to false assessments of coral health. Furthermore, metabolic biomonitoring, like many monitoring programs, will rely on a few representative species as sentinels of coral health. Therefore, the natural metabolic variation and homeostatic potential of five coral species over a diel cycle was explored using LC-MS and RF models (Chapter 6). The need for a simple and effective collection technique was also realised during the design of the in-field experiment presented here; consequently, the metabolomes of wild corals quenched using liquid nitrogen were compared with those quenched using a less hazardous, and more user-friendly, methanol-based technique. RF modelling revealed low levels of diel variation in each coral species that was most apparent in low abundant spectral features. *A. aspera* had the lowest intrasample variation, supporting its adoption as a sentinel species for metabolic biomonitoring. Finally, the more user-friendly methanol quenching technique induced similar levels of metabolome variation as the more commonly used liquid nitrogen quenching, supporting its use for in-field sampling of wild corals.

The research in this thesis improves the current understanding of the coral metabolism by providing several putative biomarkers linked to coral health and function, and metabolic phenotypes associated with realistic, mid-century levels of atmospheric CO_2. It furthers the progress of coral metabolomics by testing and validating a variety of sampling, extraction, analysis and chemometric methods. The results and discussion presented here provide a solid foundation for the future development and advancement of metabolic biomonitoring programs of corals and coral reefs.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

GLOSSARY OF TERMS

Bioindicator (or Biological indicator) An organism (or parts of an organism or community of organisms) that contains information on the quality of the environment.

Biomarker (or Biological marker) A verified biological response that is objectively measured and evaluated as an indicator of a particular biological process.

Biomonitoring (or Biological monitoring) The quantitative measurement of an organism (or parts of an organism or community of organisms) to assess the condition of the environment.

List of Abbreviations

AIMS	Australian Institute of Marine Science
ANOVA	Analysis of Variance
DNA	deoxyribonucleic acid
ERETIC	Electronic REference To access In vivo Concentrations
ESI	Electrospray Ionisation
FAA	Free Amino Acids
FTIR	Fourier Transform Infrared
GBR	Great Barrier Reef
GBRMPA	Great Barrier Reef Marine Park Authority
GCMS	Gas Chromatography coupled with Mass Spectrometry
HF	Host Factors
HRF	Host Release Factors
IMOS	Integrated Marine Observing System
IPCC	Intergovernmental Panel on Climate Change
JCU	James Cook University
LC	Liquid Chromatography
LC-MS	Liquid Chromatography coupled with Mass Spectrometry
LOD	Limit Of Detection
MAA	Mycosporine-like Amino Acids
MAE	Mean Absolute Error
MS	Mass Spectrometry
^{1}H NMR	Proton Nuclear Magnetic Resonance
OPLS	Orthogonal projections to latent structures
PAM	Pulse Amplitude Modulated (fluorometry)
PBQC	Pooled Biological Quality Control
PCA	Principal Components Analysis
PDA	photodiode array detection
PLS-DA	Partial Least Squares - Discriminant Analysis
PSII	Photosystem 2
QTOF	Quadrupole Time of Flight (mass spectrometer)
RCP	Representative Concentration Pathway
RF	Random Forests
RLA	Relative Log Abundance
RMANOVA	Repeated measures ANOVA
RMSE	Root Mean Square Error
ROS	Reactive Oxygen Species
SD	Standard Deviation

SR-IRMS	Synchrotron Radiation Infrared Microspectroscopy
UV	Ultraviolet
XML	Extensible Markup Language

Chapter 1
Introduction

The Great Barrier Reef (GBR), encompassing ~3000 coral reefs and spanning an area of 345,000 km^2 along the northeast coast of Australia, is the largest of the world's coral reefs (De'ath *et al.*, 2009). Globally, it provides a diverse range of essential goods and services such as, food, coastal protection, income, employment and the discovery of new drugs and biochemicals (Hoegh-Guldberg, 1999). The GBR is valued at $56 billion as an Australian economic, social and iconic asset; in 2015-16 alone, it contributed $6.4 billion and currently supports over 64,000 jobs (O'Mahoney *et al.*, 2017). Unfortunately, coral reefs worldwide have been degraded by centuries of overfishing and pollution and are now under further pressure from ocean warming and acidification driven by increasing atmospheric CO_2 concentrations (Hughes *et al.*, 2010). On the GBR, the combined effects of increased nutrients and sediments, localised flooding, temperature extremes, predation by Crown-of-Thorns starfish (CoTS) and cyclones have had a major impact on the reef (GBRMPA, 2014a, Hughes *et al.*, 2017b). The cumulative effects of these disturbances have diminished the GBR's ability to recover (GBRMPA, 2014b) and they are predicted to worsen into the future. How this important resource is managed under future climate change scenarios is crucial to maintain the goods and services it provides.

Management of the GBR is a delicate balance between the needs of users, Traditional Owners and stakeholders while ensuring the long-term protection of the reef, its biodiversity and heritage values – it is a nuanced and complex task that involves multiple disciplines and strategies. Traditional approaches to reef management have involved monitoring and assessing the state of the reef ecosystem, as well as impacts and environmental drivers of change (GBRMPA, 2014c, Nichols & Williams, 2006, Christensen *et al.*, 1996). State-based indicators such as, % coral cover, coral colour, number of new coral recruits, etc, are monitored in keystone species to assess the state of the reef as it responds to environmental stressors such as, ocean warming, storm activity, fishing, dredging, water quality, damage, sedimentation and CoTS predation (Cooper & Fabricius,

2007, GBRMPA, 2014c). This traditional approach, with its emphasis on reef state and the abundance of keystone species, does not adequately reflect the health and functional state of corals, nor many of the ecological and biochemical processes that underpin the entire system.

Molecular biomarkers of coral health and functional state offer great potential as a new source of resilience-based metrics. They have the capacity to describe a variety of ecological processes and drivers with greater precision than many of the traditional bioindicators and proxies. Growth and mortality of corals for example, could be more accurately described using molecular biomarkers as they may potentially describe the health of corals prior to the manifestation of visual symptoms. The maximum quantum yield of photosystem II (F_v/F_m), for example, is known to correlate with reduced growth rates and mortality in *Acropora hyacinthus* for up to two years post bleaching (Okamoto *et al.*, 2005). To date, the use of molecular biomarkers such as symbiont F_v/F_m, alteration in gene expression or metabolic dysfunction in reef resilience assessments remains low (Warner *et al.*, 2006, Weis & Allemand, 2009, Weis, 2019). This may be due, in part, to the perceived complexity of emerging techniques and the cost and logistics of employing them at regional scales. In-depth research into the applicability and reliability of molecular biomarkers is required, as the benefits returned from their use in biomonitoring may far outweigh the costs of implementation.

The measured metabolome is the ultimate expression of an organism's phenotype, representing the end product of cellular and genetic processes in response to the environment (Fiehn, 2002). Metabolic profiles are unique fingerprints of an organism's biochemical state and their comprehensive analysis falls within the broad field of metabolomics (see Chapter 2 for a more thorough review of the field). Metabolic profiling is particularly well suited as a biomonitoring tool as the metabolome reflects the response of the genotype to (environmental) disturbance and precedes phenotypic change; in essence, its response is faster than most other biological processes (Caldana *et al.*, 2011).

Metabolic biomonitoring is an emerging approach to ecosystem management that incorporates metabolic profiling and modern chemometrics to monitor the biochemical response of an organism to changes in its environment. Its potential has been considered for a number of environmental management applications, such as: measuring the bioavailability of soil contaminants by monitoring the metabolome of earthworms (Simpson & McKelvie, 2009); monitoring soils, crops, livestock health and water quality in

developing countries using low-cost infrared spectroscopy (Shepherd & Walsh, 2007); utilising sentinel species to monitor pollution and its effects (Moore *et al.*, 2004); managing biotic interactions in agro-ecosystems (Thrall *et al.*, 2011); conducting *in situ* monitoring of fish exposed to impacted surface waters (Skelton *et al.*, 2014), monitoring of bivalves to assess harmful marine compounds (Suárez-Ulloa *et al.*, 2013) and as a biological effects tool to define good marine environmental status (Lyons *et al.*, 2010).

Metabolic biomonitoring has been validated in several non-environmental fields. In health and medicine, for example, metabolic profiling has been used to monitor: the treatment of patients with rheumatoid arthritis (Zabek *et al.*, 2016); the outcomes of gastric bypass surgery (Lopes *et al.*, 2016); the health effects of iron storage disease in the Sumatran Rhinoceros (Watanabe *et al.*, 2016) and cancer treatment toxicity (Hajduk *et al.*, 2016). In agriculture, metabolic biomonitoring has been adopted to monitor biochemical changes in grape berries (Ali *et al.*, 2011); the milk composition and metabolic status of dairy cows (Aernouts *et al.*, 2011); plant responses to abiotic stresses (Obata & Fernie, 2012); the effect of production practices on potatoes (Shepherd *et al.*, 2014) and the quality, authenticity and provenance of a variety of foods (Cevallos-Cevallos *et al.*, 2009). More recently, validated approaches have been applied to environmental biomonitoring programs (reviewed by Bedia *et al.*, 2018), including detection of changes in water quality parameters using *Daphnia magna* (Jeong *et al.*, 2019), and presence of the widely prescribed non-selective non-steroidal anti-inflammatory drug diclofenac based on changes in the metabolome profile of the mussel *Mytilus galloprovincialis* (Bonnefille *et al.*, 2018), both with potential to be an early warning system.

In the coral reef research field, a number of metabolomics studies have highlighted the potential of this technique to provide high-resolution phenotypes of a variety of corals and their symbionts (see Gordon & Leggat, 2010 and Chapter 2 for a more detailed review of the coral metabolomics literature). For example, phenotypes associated with ocean warming (Hillyer *et al.*, 2016, Hillyer *et al.*, 2017a, Hillyer *et al.*, 2017b, Hillyer *et al.*, 2018, Petrou *et al.*, 2018), ocean acidification (Putnam *et al.*, 2016) and combined ocean warming and acidification (Sogin *et al.*, 2016, Farag *et al.*, 2018) have been identified. Phenotypes associated with unique environmental conditions, such as water quality (Januar *et al.*, 2012), provenance (He *et al.*, 2014, Farag *et al.*, 2016) and light conditions (Klueter *et al.*, 2015), have also been identified in a variety of coral species. Finally, phenotypes associated with unique attributes such as, species variation (Sogin *et al.*, 2014), competitive interactions between coral and algae (Quinn *et al.*, 2016), chemical and physical elicitors (Farag *et al.*,

2017a, Farag *et al.*, 2017b), symbiont type (Matthews *et al.*, 2017) and microbiome composition (Sogin *et al.*, 2017) have been characterised.

There is a lack of thorough accounting of the array of coral biomarkers available for targeted monitoring. On this basis, this research explores the potential of metabolic profiling and machine learning methods to establish the presence of molecular biomarkers of coral reef health and functional state that may potentially offer greater precision than traditional bioindicators. In doing so, it aims to further our understanding of the complex biological processes that underpin the biological stress response of corals by characterising influential features driving the prediction of specific stress responses.

To begin with, the relevant grey and published literature was reviewed to establish the state of the field and identify the potential knowledge gaps. This review, which is focussed on *Symbiodiniaceae*-invertebrate symbioses and their metabolic interactions (see Gordon & Leggat, 2010 and Chapter 2) was first published in 2010 and has since been updated to include the current coral metabolomics literature.

Metabolic biomonitoring of coral health relies on a thorough and reproducible snapshot of the coral metabolome. Snapshots are affected by the sample preparation and handling procedures that are common to many studies; however, traditional metabolomics protocols, which have been developed primarily for medical and industrial applications and rely on comprehensive primary metabolite databases, may be unsuitable for the analysis of coral due to the biological complexity of the holobiont and our limited knowledge of coral secondary metabolite chemistry. Prior to 2013, there was no published research on sample preparation for coral metabolomics (see Gordon *et al.*, 2013 and Chapter 3). This protocol describes the sample collection, extraction and measurement of hard coral holobiont metabolites using both proton Nuclear Magnetic Resonance (^{1}H NMR) spectroscopy and Liquid Chromatography coupled with Mass Spectrometry (LC-MS) (Gordon *et al.*, 2013 and Chapter 3). It provides all the details required to replicate the method, including many of the costly and time consuming "pitfalls" or "traps" that were discovered during its development. Other than Gordon *et al.* (2013), there remains a lack of standardised methods for coral metabolomics. Anderson *et al.* (2019) recently reviewed the methanol extraction method, finding it still to be applicable to coral metabolomics. This is discussed in greater detail in Chapter 3, section 3.5.

To identify health-related coral phenotypes from metabolic profiles, robust statistical and machine learning models are required to classify profiles under specific conditions while

monitoring coral health and identifying potential biomarkers. Discrete descriptors of coral phenotype, such as *stressed* or *healthy*, provide reef managers with easy to interpret terms for guiding interventions. With this in mind, Chapter 4 explores the capability of two machine learning models (Partial Least Squares – Discriminant Analysis and Random Forests) to classify the ^{1}H-NMR and LC-MS metabolic profiles of stressed and healthy *Acropora aspera* during exposure to a simulated bleaching event and at pCO_2 levels consistent with mid-century relative concentration pathways (IPCC RCP 4.5, 2014).

In contrast to the discrete classification approach, regression models can predict continuous descriptors that may provide more precise indicators of coral health based on quantifiable measures such as F_v/F_m, widely accepted as a proxy for coral health (Jones *et al.*, 1999). Therefore, the capability of Random Forests (RF) modelling to predict the F_v/F_m of healthy and thermally stressed corals from their LC-MS metabolic profiles is explored in detail in Chapter 5. Mass spectral features driving the prediction of F_v/F_m are discussed in the context of coral health and metabolic biomonitoring.

Unfortunately, natural variation in coral metabolomes has the potential to introduce unwanted bias in machine learning models, which unduly undermines the confidence of model predictions and/or lead to false assessments of coral health. Furthermore, metabolic biomonitoring programs, like many other monitoring programs, will rely on a few representative species as sentinels of coral and overall ecosystem health. In Chapter 6, therefore, the natural metabolic variation of five coral species over a diel cycle is determined using LC-MS and RF modelling. Further data interrogation identified influential spectral features that explain differences in the homeostatic potential as a function of species and time. These are discussed in the context of identifying a suitable sentinel coral species for metabolomics biomonitoring of coral reefs.

Over the course of this research, it became evident that coral biomonitoring introduces an additional level of technical and logistical complexity not experienced in aquaria- or laboratory-based experiments. Field sample collection, for example, is a particularly complex task given the dynamic nature of the coral metabolome and the challenging conditions encountered in the marine environment. The use of traditional quenching agents, such as liquid nitrogen, is problematic in the field due to its high evaporation rate and elevated safety risk, requiring strict handling procedures. In many field studies, it may not be possible (immediately or at all) to immerse freshly collected corals nubbins in liquid nitrogen, where any delay in quenching possibly induces unwanted variation into the coral

metabolome. Therefore, impacts of an alternative, less hazardous and more user-friendly methanol-based quenching technique on the metabolomes of wild corals were assessed.

This thesis concludes with a synthesis of the results adding to the current pool of metabolic biomonitoring literature while advancing our understanding of coral metabolism and coral metabolomics. Additional gaps in the coral metabolomics field are identified and discussed, signalling further research directions that might lower the barrier to implementing future metabolomics research and biomonitoring programs.

Chapter 2
Literature Review

A portion of this chapter is published as:

> Gordon, B. R. & Leggat, W. (2010) *Symbiodinium*-Invertebrate Symbioses and the Role of Metabolomics. *Marine Drugs,* **8,** 2546-2568.

Benjamin Gordon wrote this chapter, designed and conducted the research, and analysed the data. Co-authors provided intellectual and editorial support.

2.1 Foreword

This section was initially prepared and published in 2010 (Gordon & Leggat, 2010). At the time, metabolomics was in its infancy and yet to be applied to the coral animal or its dinoflagellate symbiont. There was a wealth of knowledge on coral and *Symbiodiniaceae* metabolites and natural products going back decades, and one could argue that some of those studies fell within the definition of metabolomics. Given the lack of coral metabolomics studies at the time, the original brief for this review was broad: explore the *potential* of applying metabolomics techniques to corals and their symbionts. With that in mind, the literature was examined and the state of metabolite research on coral and *Symbiodiniaceae* assessed – the genesis of this review. Now, Gordon *et al.* 2010 has been cited more than 120 times, highlighting its relevance to those who have delved into the coral metabolomics field. Nonetheless, more than ten years have passed since this chapter's publication and an update of the coral metabolomics literature was warranted. This chapter offers an updated version of the review with the most recent literature discussed in section 2.5. The literature is discussed within the context of six common themes that best represent the current state of the coral metabolomics field.

2.2 Background

Metabolomics is the newest of the 'omics'-based sciences and is fast becoming a popular tool in many fields of research. The first metabolomics analyses were performed at the turn of the century and, in the decades since, the literature available has gone from a handful published in 2000 (Fiehn *et al.*, 2000, Raamsdonk *et al.*, 2001, Oliver *et al.*, 1998) to more than 600 published in 2009. Now, in 2020 a Google Scholar search for "metabolomics" returns more than 100,000 articles.

Metabolomics was first defined by Oliver *et al* (1998) as:

> "the quantitative complement of all of the low molecular weight molecules present in cells in a particular physiological or developmental state"

In the years before the advent of metabolomics, coral metabolite studies were more focused on specific metabolites, or groups of metabolites, and how they changed according to a given stimuli or from the effects of an alteration (e.g. a genetic mutation). These types of studies were often classified as metabolite target analysis, metabolic profiling or metabolic fingerprinting (Fiehn, 2001). Although definitions can be vague and open to interpretation, the distinguishing factor in all three definitions is that none attempted to study the global suite of metabolites in an unbiased fashion. Metabolomics has evolved over the last twenty years and, for the purpose of clarity, it is now more commonly defined as the study of endogenous, low molecular weight (<1500 Da), global metabolite profiles in a system (cell, tissue, or biofluid) under a given set of conditions (Goodacre *et al.*, 2004, Rochfort, 2005, Viant, 2007).

Metabolomics has become a popular approach for studying the interaction of living organisms with their environment. Part of its appeal, and value, is that it complements other omics methods such as genomics, transcriptomics and proteomics (Viant *et al.*, 2009). Metabolomics can report on the functional status of an organism, which can then be related to its phenotype, and is one reason it has become a popular tool in the environmental sciences. Another contributing factor is its ability to raise questions about the study organism and thereby uncover unexpected metabolite responses and relationships, which can lead to further hypothesis generation and investigation (Bundy *et al.*, 2009, Kell & Oliver, 2004). The study of how organisms interact with their environment is a broad field, encompassing ecology, agriculture (Dixon *et al.*, 2006), viticulture (Rochfort *et al.*, 2010) and forensics (Ovenden *et al.,* 2010).

Unicellular, photosynthetic dinoflagellates of the family Symbiodiniaceae, which were commonly referred to as zooxanthellae, were first described in (Cienkowski *et al.*, 1871). They live freely in the water column and are predominantly found in clear tropical waters with low nutrient concentration and plankton densities (Muscatine & Porter, 1977, Oldeman, 1987). In this low nutrient environment, some invertebrates have formed a symbiotic relationship with *Symbiodiniaceae* to gain a competitive advantage through increased fitness (Trench, 1979), allowing the bilateral exchange of metabolites, including the production of metabolites that are not formed by either organism separately (Lewis & Smith, 1971). Of particular interest to researchers over the past few decades are the relationships of *Symbiodiniaceae* with hermatypic corals. The importance of this relationship cannot be understated given the fundamental role that hermatypic corals have played in the formation and maintenance of coral reef habitats. These habitats provide a livelihood for local communities, with tourism and fishing industries relying heavily upon them. Unfortunately, due to anthropogenic pollution and global climate change, this ecosystem is under increasing threat (Hughes *et al.*, 2003, Yellowlees *et al.*, 2008, Morice *et al.*, 2012, Hughes *et al.*, 2017a, Hoegh-Guldberg *et al.*, 2017).

This chapter analyses the roles that metabolomics and targeted metabolite analysis had in marine science with a specific focus on *Symbiodiniaceae*-invertebrate symbioses (see section 2.3). An in-depth historical analysis of targeted metabolite research of *Symbiodiniaceae* and invertebrate hosts (see section 2.4) sets the tone for a thematic analysis of the coral metabolomics literature (see section 2.5).

2.3 *Symbiodiniaceae* symbiosis

Symbiosis was originally defined by Anton deBary as "the living together of differently named organisms" (Yellowlees *et al.*, 2008). *Symbiodiniaceae* have established intracellular symbioses with corals, anemones, jellyfish, nudibranchs, Ciliophora, Foraminifera, zoanthids and sponges and have extracellular symbioses with giant clams (for a review see Stat *et al.*, 2006). The majority of the estimated 800 coral species have established symbiosis with *Symbiodiniaceae*; in this symbiotic relationship the algae are found within a host-derived vacuole (symbiosome membrane) within the gastrodermal cell layer. The cell membrane is derived during the acquisition and division of the algal symbionts and is analogous to the symbiosome in legumes where the plant membrane encloses the symbiotic rhizobium cells (Yellowlees *et al.*, 2008). Roth *et al* (1988) defined this cell

membrane as the host derived symbiosome. In 1993 research revealed that the cell membrane of the anemone host *Anemonia viridis* controls the availability of phosphorous compounds and other nutrients to zooxanthellae (Rands *et al.*, 1993). Since the transportation of all nutrients must proceed via this cell membrane, it is critical to the metabolic interaction between the symbiont and host.

Symbiodiniaceae cell numbers can range from one symbiont per host cell (Muscatine *et al.*, 1998) to over 60 in some hydroids (Fitt, 2000), with *Symbiodiniaceae* normal cell densities being in the order of 1-2 million cells cm^{-2} of coral tissue. Until recently, *Symbiodiniaceae* were characterised within a single genus (*Symbiodinium*) consisting of several clades, sub-clades, types and strains (LaJeunesse, 2001). However, a recent systematic revision of the genus has brought the formal taxonomy in line with the current understanding of the evolutionary relationships among these dinoflagellate symbionts (LaJeunesse *et al.*, 2018). The genus *Symbiondinium* has now been split into multiple genera and species belonging to the *Symbiodiniaceae* family.

While some *Symbiodiniaceae* are considered generalists, establishing symbioses with more than one species of host, others are restricted in the symbiotic association they can form. Given the diversity in both coral host species and algal genotypes there is potential for a large range of metabolic relationships.

2.4 Targeted metabolite analysis of *Symbiodiniaceae*.

2.4.1 Nutritional roles of *Symbiodiniaceae*.

Much of the metabolite analysis of *Symbiodiniaceae* began in the mid 1950s after Zahl and McLaughlin (1957) reported a method for isolating and cultivating these dinoflagellates. From that point onwards, the study of *Symbiodiniaceae* metabolites accelerated. Muscatine and Hand (1958) were the first to suggest, based on experimental evidence, that *Symbiodiniaceae* provided nutrition to their anemone host. In that study, sea anemones with symbiotic dinoflagellates were exposed to seawater containing ^{14}C-labelled CO_2 for 18 and 48 hours. Autoradiography of dissected host tissue after 18 hours of exposure showed that algae incorporated enough labelled carbon to produce a readable autoradiograph; however, no transfer of labelled metabolites from the algae to the anemone were observed. Transfer of labelled carbon from the algae to the host was observed after one week, and each subsequent week for four weeks. Muscatine *et al.* (1967) subsequently demonstrated that

symbiotic algae from a variety of freshwater hosts produce relatively large amounts of soluble carbohydrates (mostly maltose and glucose) in contrast to free living algae, which produced only small amounts of glycolic acid; speculating that the soluble carbohydrates were a source of carbon utilised by the host to enhance growth and survivability.

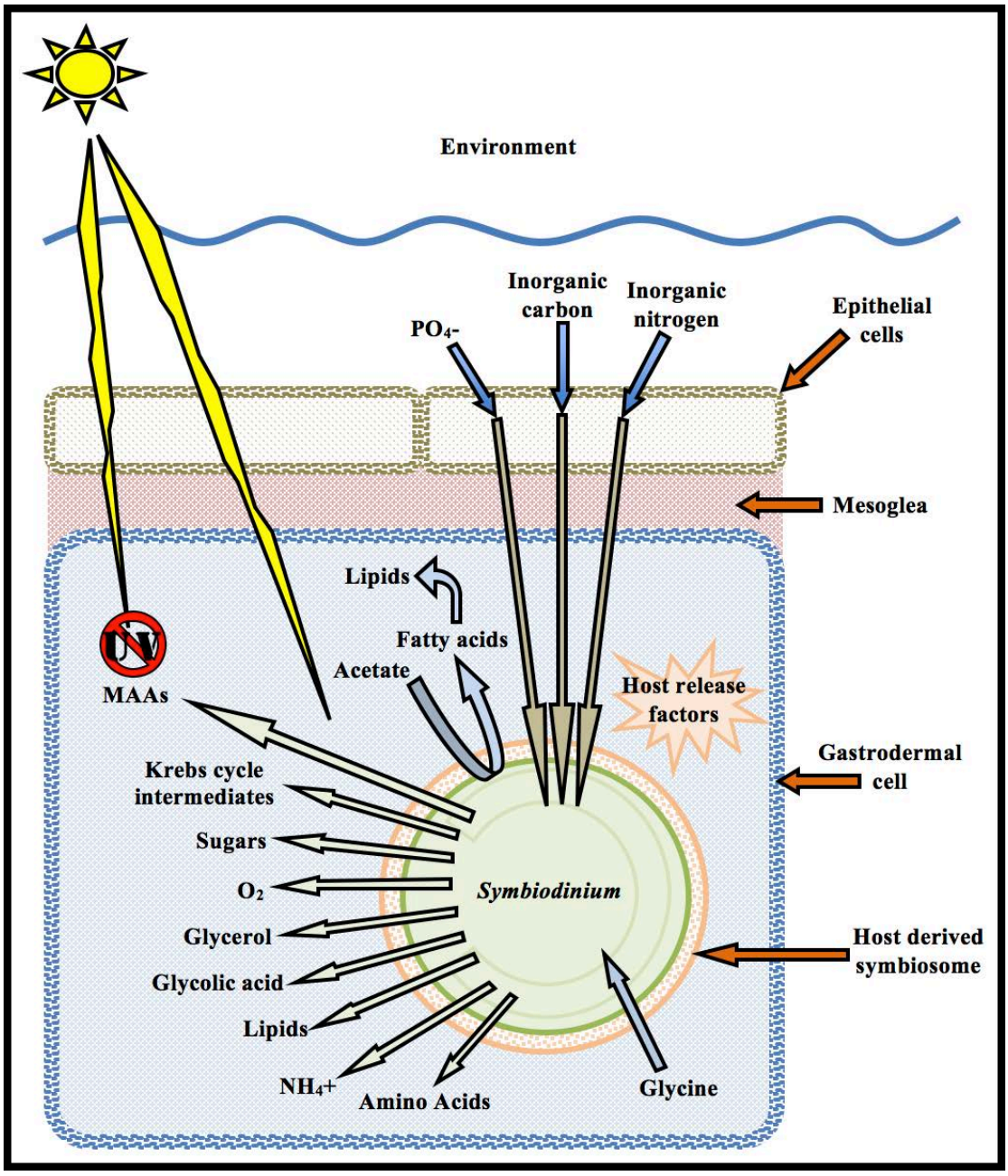


Figure 2.1 Schematic diagram of the invertebrate-*Symbiodiniaceae* relationship

By the 1970's researchers had firmly established that *Symbiodiniaceae* release soluble products of photosynthesis, including sugars, amino acids, esters, alcohols and lipids; all of which were utilised by their hosts (see Table 2.1 and Figure 2.1) (Goreau *et al.*, 1973, Von Holt & Von Holt, 1968b, Cernichiari *et al.*, 1969). Most research during this period emphasised the

nutritional interactions between the host and symbiont. Of particular interest was the role of *Symbiodiniaceae* in translocating photosynthetic products and the recycling of host metabolic products, such as ammonium and phosphate. Once again, most research made use of labelled carbon to identify the metabolic products produced and recycled by *Symbiodiniaceae*.

Table 2.1 Some of the more common metabolites of the invertebrate-*Symbiodiniaceae* relationship[1]

Metabolite	Transport Direction		*Symbiodiniaceae* Specific	Function	Reference
	Host to symbiont	Symbiont to Host			
Maltose		✓		Nutrition	(Muscatine, 1965)
Glycerol		✓		Nutrition	(Muscatine, 1967)
Glucose and other hexoses'		✓		Nutrition	(Muscatine *et al.*, 1967, Bil *et al.*, 1991)
Glycolic acid		✓		Nutrition	(Muscatine *et al.*, 1967)
Glycine	✓			Nucleotide synthesis and nutrition	(Von Holt, 1968)
Alanine		✓		Protein formation	(Lewis & Smith, 1971)
Acetate	✓			Fatty acid synthesis	(Von Holt & Von Holt, 1968a, Patton *et al.*, 1977)
Fatty acids		✓		Lipid synthesis	(Patton *et al.*, 1977)
Lipids		✓		Energy exchange	(Kellogg & Patton, 1983)
Lactate		✓		Metabolite	(Von Holt & Von Holt, 1968a)
Succinate		✓		Krebs cycle	(Von Holt & Von Holt, 1968a)
Citrate		✓		Krebs cycle	(Von Holt & Von Holt, 1968a)
Ketoglutarate		✓		Krebs cycle	(Von Holt & Von Holt, 1968a)
Malate		✓		Krebs cycle	(Von Holt & Von Holt, 1968a)
Pyruvate		✓		Glycolysis product	(Von Holt & Von Holt, 1968a)
PO_4^{3-}	✓			Nutrition	(Cates & McLaughlin, 1979)
NO_2^-, NO_3^-, NH_4^+	✓			Amino acid synthesis	(Cates & McLaughlin, 1979)
Inorganic carbon (i.e. CO_2, HCO_3^-)	✓			Photosynthesis	(Muscatine & Lenhoff, 1963, Patton *et al.*, 1983)
O_2		✓		Photosynthesis	(Dykens & Shick, 1982)
Zooxanthellamide-A and B			✓	Unknown	(Onodera *et al.*, 2003, Onodera *et al.*, 2004)
Mycosporine-like amino acids			✓	UV light and free radical protection	(Dunlap & Yamamoto, 1995)
Zooxanthellatoxins			✓	Unknown	(Nakamura *et al.*, 1993)
Symbioramide-C16			✓	Unknown	(Nakamura *et al.*, 1998)
Zooxanthellabetaines			✓	Unknown	(Nakamura *et al.*, 1998)
Zooxanthellamine			✓	Unknown	(Nakamura *et al.*, 1998)

Despite direct evidence that photosynthetic products were chemically incorporated into the tissue of the host, little was understood about the quantity of metabolites produced by the symbiont, nor how the host utilised them. It was in this context that Robert K. Trench published three papers (Trench, 1971a, Trench, 1971b, Trench, 1971c). The first examined the movement of labelled carbon from symbiont to host in the sea anemone, *Anthopleura*

[1] A more comprehensive list of metabolites can be found in the coralmz R package at GitHub and Zenodo (Gordon & Motti, 2020).

elegantissima and the zoanthid *Palythoa townsleyi*. Chemical extraction and paper chromatography methods were used to fractionate the water-soluble, low molecular weight compounds from the lipids, proteins and nucleic acids. The photosynthetically fixed carbon (predominant as glycerol) was incorporated into the animal tissue primarily as lipids and proteins (Trench, 1971a).

The second paper (Trench, 1971b), focussed on the type of compounds that were produced by *Symbiodiniaceae*. Using similar chemical and chromatographic methods, several different compounds were identified. Glycerol was identified as the major extracellular product and that other labelled carbon-based compounds were produced, including alanine, glucose, fumaric acid, succinic acid, glycolic acid and two other unidentified organic acids. One of the more noteworthy observations in this paper was that different *Symbiodiniaceae* isolates produced similar compounds in significantly differing proportions. Trench (1971b) concluded that although *Symbiodiniaceae* from different hosts were morphologically the same, they did differ biochemically. This was the first observation that *Symbiodiniaceae* is a diverse group of dinoflagellates; it is also consistent with the dynamic response of the metabolome to its environment.

The third and final paper in the series considered the effect of host tissues on the excretion of photosynthetic products *in vitro* by *Symbiodiniaceae* (Trench, 1971c). It was observed that *Symbiodiniaceae* excreted a greater number of metabolites when incubated in a homogenate of host tissue than those incubated in seawater alone. When incubated in a homogenate of host tissue from algal-free host animals, there was no observable increase in the production of extracellular metabolites; however, when algal-free animals were infected with *Symbiodiniaceae*, the host tissue increased production of metabolites, suggesting that the host plays an important role in regulating the type and quantity of metabolites produced by *Symbiodiniaceae*. This series of papers highlighted the metabolic diversity and plasticity of *Symbiodiniaceae* while providing *a priori* knowledge of complex biosynthetic pathways.

The concept of the host controlling the release of metabolites from the symbiont has been formalised in the concept of host factors (HF) or host release factors (HRF) (Muscatine, 1967, Trench, 1971c). Although there are still some questions as to their exact identity, Gates *et al.* (1995) proposed that host-derived free amino acids (FAAs) served as HRFs. They found that a variety of FAAs stimulated *Symbiodiniaceae* carbon fixation up to two-fold, and carbon release up to four-fold. These compounds were tentatively identified as glucose, glycerol, alanine, glycine, serine, glutamine, valine, phenylalanine, leucine citrate,

glycerate, glycolate, lactate and succinate. In contrast, research by others has been unable to reproduce similar results (Withers *et al.*, 1998, Cook & Davy, 2001), proposing instead that small (<1000 Da) unknown host compounds must act as HRFs (Grant *et al.*, 1998). In recent years, inositol has been proposed as a potential host-controlled signalling molecule after a recent study identified increases of this well-known signalling molecule in the metabolome of *Aiptasia* exposed to thermal stress (Hillyer *et al.*, 2017a). Corals may also produce a low molecular weight (<1000 Da) peptide that can inhibit *Symbiodiniaceae* photosynthesis (Grant *et al.*, 2001).

Although coral tissue can consist of up to one-third lipid by dry-weight, it was not until 1977 that Patton *et al.* (1977) provided evidence that *Symbiodiniaceae* primarily performed the role of lipid synthesis in corals from host-derived acetate. They proposed that energy transfer from symbiont to host via acetate recycling might be the key to the ecological success of corals in nutrient poor waters. Their work provided evidence that the host acquires, via digestive and degradative metabolism, carbon from exogenous food sources, where it is subsequently absorbed by the *Symbiodiniaceae*, most likely as acetate. The acetate is then transferred to the chloroplast where, in reaction with excess adenosine triphosphate, it is reduced to triglyceride fatty acids. These triglyceride fatty acids are transferred back to the host via lipid bodies that were formed on the outer surface of the *Symbiodiniaceae* cell wall. They went on to propose that the surface membrane of these lipid bodies were the site of wax ester and triglyceride synthesis.

Later research by Patton *et al.* (1983) showed that lipid bodies exist within both the host and symbiont cells. They found that wax esters and triglycerides comprised 75% of the intact coral lipids whereas the symbiont comprised of only about 8% of these neutral lipids. Conversely, structural lipids such as sterols, phospholipids and galactolipids made up approximately 67% of the symbiont lipids but only 16% of the host lipids. Labelled fatty acids derived from acetate had a higher degree of unsaturation than endogenous fatty acids, which implied that the host either modified the fatty acids or the transfer process was selective for saturated fatty acids. Their results showed that the majority of the coral's reserved energy (wax esters and triglycerides) was located within its own tissue and that *Symbiodiniaceae* were the primary producers of these lipids, biosynthesised from host-derived acetate.

Research on lipids delved deeper into the function and formation of lipid bodies in mammalian cells during the early 2000s. Reviewed by Martin and Parton, (2006) these lipid

bodies were described as pivotal cellular organelles with specific structural and functional characteristics. They consist of a core of neutral lipids predominantly comprised of triacylglycerols or cholesteryl esters, which are surrounded by a monolayer of phospholipids and associated proteins. Functionally, these mammalian lipid bodies are now regarded as complex organelles involved in a number of cellular processes such as, cell signalling (Umlauf *et al.*, 2004) and visual chromophore regeneration (Imanishi *et al.*, 2004) and, as previously discussed, the lipid metabolism in *Symbiodiniaceae*.

Luo *et al* (2009) tested the hypothesis that lipid bodies in *Symbiodiniaceae* and the host gastrodermal cells of *Euphyllia glabrescens* were unique organelles that reflected the dynamic nature of their endosymbiotic status. Using dual emission ratiometric imaging with a solvatochromic fluorescent probe, they found the ratio of polar versus neutral lipids in lipid bodies of the host increased upon bleaching, indicating a decrease in neutral lipid accumulation within the gastrodermal cell. Conversely, neutral lipid accumulation increased in the lipid bodies of the symbiont when bleached. This demonstrated that the composition and morphology of lipid bodies was positively correlated to the endosymbiotic status and hence, implicated the lipid bodies in lipid trafficking between the host and the symbiont for the purpose of regulating the endosymbiosis.

While lipids are essential for energy storage and nutrition, they also play an important role as structural components of symbiont cells; as such, differences in lipid composition can have distinct effects on the ability of different *Symbiodiniaceae* species to adapt to changes in their environment. In 2004, Tchernov *et al* (2004) demonstrated that thermal stress sensitivity of *Symbiodiniaceae* could be categorised by the level of lipid saturation and lipid stacking patterns in the thylakoid membrane. Gas Chromatography Mass Spectrometry (GC-MS) analysis of several *Symbiodiniaceae* isolates revealed a striking contrast between thermally tolerant and sensitive isolates. The thermally tolerant *Symbiodiniaceae* had much lower levels of the major polyunsaturated fatty acid, Δ6,9,12,15-cis-octadecatetraenoic acid (18:4), in relation to Δ9-cis-octadecatetraenoic (18:1). Furthermore, transmission electron micrographs of thermally sensitive *Symbiodiniaceae* exposed to higher temperatures showed a significant disruption in the organised stacking pattern of the thylakoid membrane, which is essential for efficient photochemical energy transduction. This finding demonstrated that thylakoid membrane lipid composition was a key determinate of thermal stress sensitivity in the symbiotic algae of cnidarians.

Research by Chen *et al.* (2015) identified six major lipid classes in host cytoplasmic lipid bodies: wax esters, sterol esters, triacylglycerols, cholesterols, free fatty acids, and phospholipids. The significant differences in concentration of these lipids between the host, symbiont and lipid bodies suggests a high degree of compartmental regulation. Chen *et al.* (2015) also found that wax esters were only present in the host cells and lipid bodies, further evidence that lipid bodies are host derived. In another study by Chen *et al.* (2017), lipid profiling of *Euphyllia glabrescens* specimens collected across the diel cycle revealed significant changes in host gastrodermal cells, lipid bodies and the symbiont; demonstrating the temporally variable nature of lipids in corals. Each of the three compartments possessed a unique lipidome, consisting of both symbiont- and host-derived lipids. The existence of uniquely derived lipids across all three compartments is evidence of the active lipid trafficking that occurs in the holobiont. Given the high degree of temporal lipid exchange from the host to lipid body, and symbiont to lipid body, Chen et al (2017) argued that lipid bodies may act as a relay centre of lipid exchange between the host and symbiont.

Given the nutrient limitation reported in tropical waters, the translocation of organic nitrogen, in the form of amino acids, from *Symbiodiniaceae* to the coral is extremely important. *Symbiodiniaceae* recycle waste nitrogen produced by the host to synthesise several essential amino acids. Until 1999, substantial evidence for the transport of essential amino acids from *Symbiodiniaceae* to the host had yet to be observed. Wang and Douglas (Wang & Douglas, 1999) reasoned that previous experimental designs exploring labelled photosynthate release would fail to detect amino acid transfer to the host tissue if the amino acids were released many hours, or even days, after carbon fixation took place, or that amino acids were not synthesised by photosynthetically-derived carbon. Extended pulse chase experiments over 2 days established the metabolic fate of labelled carbon compounds in the symbiosis and provided direct evidence for the transport of several essential amino acids from *Symbiodiniaceae* to the host.

Both the coral and algae are capable of assimilating ammonium from their environment in addition to producing it metabolically via glutamate dehydrogenase and glutamine synthetase. This is thought to be a major reason for the success of the *Symbiodiniaceae* - invertebrate relationship in nutrient poor environments. Uptake, retention and release of ammonium in reef corals was investigated, providing "evidence to support the hypothesis that combined nitrogen is recycled within the coral-algae symbiosis" (Muscatine & D'Elia, 1978). These experiments found symbiotic corals uptake and retain ammonium and that the process was enhanced by light; in contrast, non-symbiotic corals released ammonium

into the medium under all conditions. It was tentatively proposed that the uptake and retention of ammonium was due to the activity of *Symbiodiniaceae*. This process was considered an adaptation to a deficiency of environmental nitrogen since virtually all of the ammonium excreted by the coral host was taken up by the symbiont. These results were further supported by D'Elia *et al.* (1983) who proposed a depletion-diffusion mechanism for the uptake of ammonium, whereby the symbiont was responsible for the majority of ammonium assimilation. However, contrasting studies have since supported the theory that the host may be limiting the supply of ammonium as a method of controlling the population of *Symbiodiniaceae*. A study of the assimilation of ammonium by the host under dark conditions, for example, found that ammonium assimilation, along with protein and free amino acid pool sizes, were the same in dark conditions as illuminated conditions after supplementing the medium with organic carbon compounds (Wang & Douglas, 1998). This suggests that the host controls the concentration of ammonium via enhanced ammonium assimilation. It also contradicted previous findings that the symbiont was primarily responsible for ammonium assimilation.

Unlike its host, *Symbiodiniaceae* is capable of utilising nitrate as a nitrogen source. The first evidence of this appeared when reef corals were found to have a light independent mechanism for the uptake and reduction of dissolved nitrate from seawater (Franzisket, 1973) and that nitrite was present in both the coral and symbiont tissue. The discovery of expressed sequence tags for transporter enzymes of both nitrate and nitrite provided further evidence for the uptake of nitrates by *Symbiodiniaceae* (Leggat *et al.*, 2007). Further, an increase in *Symbiodiniaceae* cell densities was observed upon increases in nitrate concentration and uptake (Marubini & Davies, 1996).

2.4.2 Diel cycles

Given the importance of photosynthesis to coral symbiosis, it is not surprising there are several biomolecules, proteins and metabolites that exhibit large diel variations. One of the major diel changes seen in symbiotic invertebrates is intracellular oxygen tension, a consequence of *Symbiodiniaceae* photosynthesis. Using oxygen microelectrodes, Dykens and Shick (1982) measured the partial pressure of pure oxygen bubbles on the surface of symbiotic sea anemones and found they were subjected to a continuous flux of hyperbaric oxygen. Similar studies have estimated that oxygen concentrations at the tissue surface vary from less than 2% of air saturation during dark periods to over 250% saturation in the light, while tissue pH varied from 8.5 in the light to 7.3 in the dark (Kuhl *et al.*, 1995).

These high oxygen concentrations have profound implications for the coral host given molecular oxygen undergoes reductions to form oxygen radicals, which are particularly destructive to cells.

An important enzyme involved in reactive oxygen detoxification is superoxide dismutase (SOD), which catalyses the dismutation of two oxygen radicals to produce hydrogen peroxide and molecular oxygen. Dykens and Shick (1982) found that anemones with higher chlorophyll content had higher rates of SOD activity, indicating that the host was altering its protein expression in response to *Symbiodiniaceae* physiology or metabolism. Exposing anemones to exogenous hyperbaric oxygen caused a 62% increase in SOD activity, while anemones kept under dark, and photosynthetically poor conditions, had reduced levels of SOD activity. Levy *et al* (2006) examined the wavelength dependence of two free radical scavenger enzymes, SOD and catalase, finding that the host's enzyme response to the spectral distribution of light was higher than that of the zooxanthellae; probably due to accumulation of free radicals within the host tissue. Furthermore, the activity of the enzymes was affected by the length of the day and night cycles, and in the laboratory, by the duration of the illumination. The activity of scavenger enzymes, such as SOD, are vital for the acclimatisation and survival of corals in shallow water environments with high light radiation, as they reduce the effects of oxidative damage to cells by free radicals.

Recent research has provided evidence of diel-associated metabolite transfer in corals. This research was focussed on the transfer of lipids between the host and symbiont via lipid bodies, which are suspected of acting as a relay centre between the host and symbiont (see Chen *et al.* 2017 and the previous discussion in section 2.4.1). In the giant clam-*Symbiodiniaceae* symbiosis, where sampling the haemolymph (blood) can be easily performed, there is ample evidence that glucose is the major carbohydrate transported to the host with concentrations varying from less than 100 μM in the dark to almost 400 μM at noon (Rees *et al.*, 1993, Leggat *et al.*, 2002).

In addition to driving photosynthesis, high light levels expose the host and symbiont to damaging UV radiation. One particular defence against UV exposure is the production of mycosporine-like amino acids (MAAs). These compounds (**1**-**12** in Figure 2.2) are characterised by a cyclohexanone or cyclohexenimine chromophore conjugated with a nitrogen substituent of an amino acid and absorb UV radiation without any further photochemical reactions (Nakamura *et al.*, 1982). The source (host vs. symbiont) of these compounds is not yet clear, and concentrations in coral tissue can vary according to

fluctuations in light levels. Yakovleva & Hidaka (2004) showed that freshly isolated *Symbiodiniaceae* contained no MAAs, yet they were distributed throughout the host tissue. In contrast, Banaszak *et al* (2000) showed that cultured *Symbiodiniaceae* cells did produce MAAs, while cells from *Breviolum spp.* did not. These discrepancies may indicate that MAAs are produced by the coral's microbial community.

Increases of up to two-fold at midday have been demonstrated for discrete MAA species. In particular, the concentration of imino-MAA species varied in response to light while mycosporine-glycine **1** did not (Yakovleva & Hidaka, 2004). Some MAAs also have the potential to act as free radical scavengers (Dunlap & Yamamoto, 1995) with six common MAAs (five cyclohexenimine MAAs: shinorine **9**, porphyra-334 **10**, palythine **3**, asterina-330 **5** and palythinol **6**; and a single cyclohexanone MAA, mycosporine-glycine **1**) being found in four different marine species. These five imino-MAAs have the oxidative robustness, which is consistent with their sunscreen properties, but no definitive antioxidant activity; mycosporine-glycine showed moderate antioxidant activity.

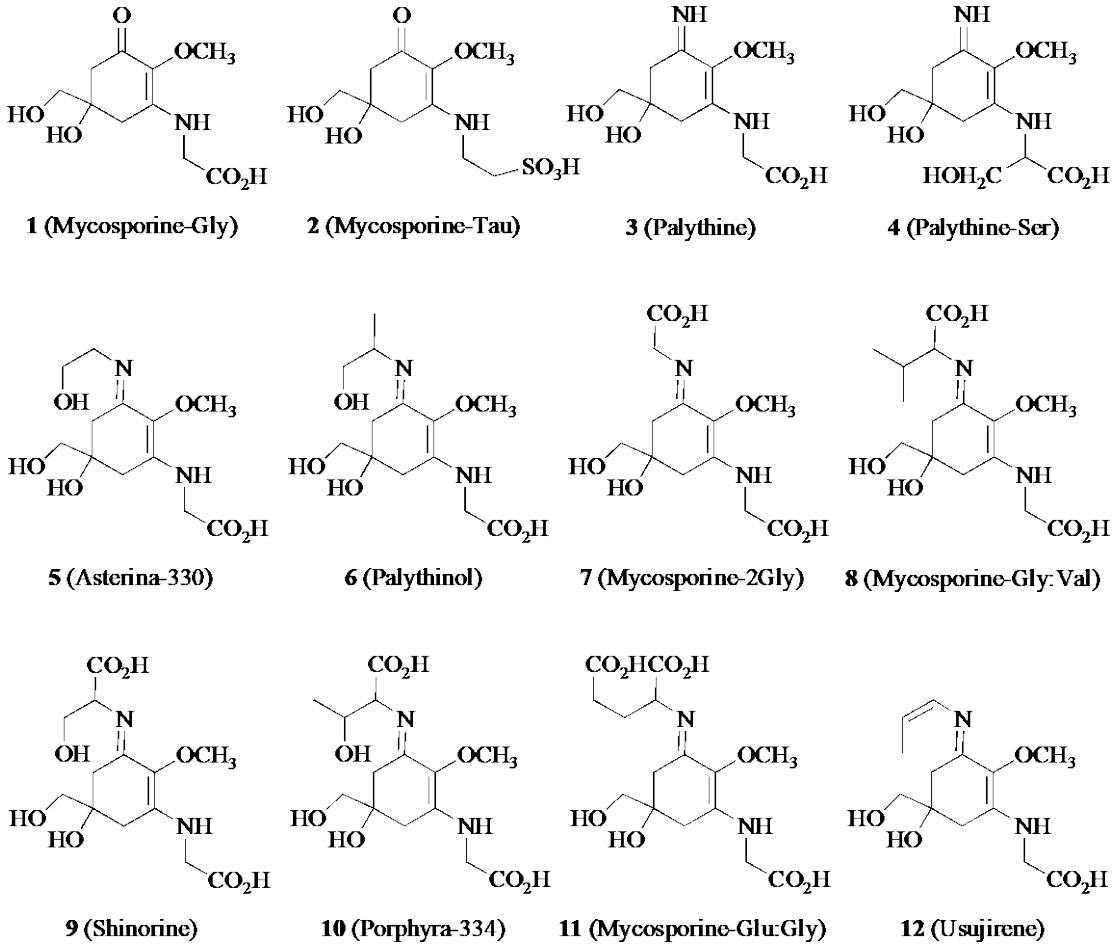


Figure 2.2 Molecular structure of some common mycosporine-like amino acids (MAAs) in marine organisms

The production of MAAs correlates with the availability of ammonium in the red algae *Porphyra leucosticta* and *Porphyra umbilicalis* (Korbee *et al.*, 2005). Both species were cultured at three ammonium concentrations (0, 100 and 300 μM) under artificial light and the effect of ammonium availability on photosynthesis determined by measuring the fluorescence of chlorophyll-a, photosynthetic pigments and MAA content. Photosynthetic activity decreased under culture conditions due to UV radiation and increased ammonium availability. Conversely, four MAAs were identified in both species (shinorine **9**, porphyra-334 **10**, palythine **3** and asterina-330 **5**) that were found to increase in concentration at high ammonium concentrations and exposure to UV radiation.

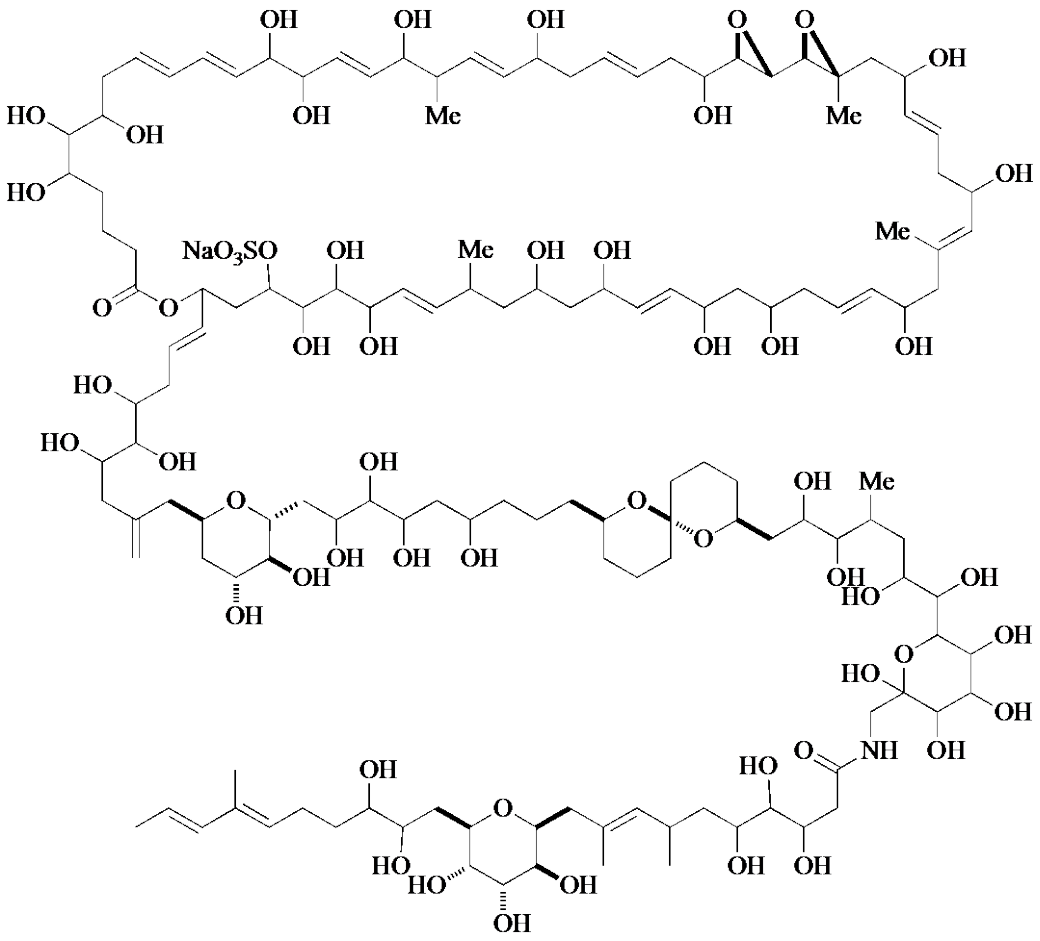
2.4.3 Biomarkers and natural products

The concept of using chemical fingerprints to classify different coral phenotypes was first explored in a 1982 study of secondary metabolites from soft corals and sponges from the Red Sea (Kashman *et al.*, 1982). These organisms are chemically protected from predation by fish and bacterial infection, and a potential source of bioactive pharmaceuticals. Given the difficulty in species identification of soft corals, chemical fingerprints of sesquiterpenes and other volatiles obtained through gas chromatography (GC) and liquid chromatography (LC) were investigated as a complementary tool for phenotype identification. Unique chemical differences in the chromatograms of six *Sinularia* species were observed and were subsequently used to assist traditional taxonomists in defining different species, especially where doubt existed due to different growth forms.

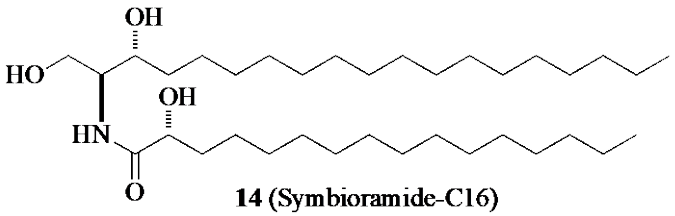
The use of molecular biomarkers has been examined in *Symbiodiniaceae* symbioses (Cuif *et al.*, 1999). Derivatisation of >3 kDa proteins and hydrolysis of polysaccharides, isolated from coral skeletons of symbiotic and non-symbiotic corals, yielded amino acids and monosaccharides. Univariate statistical methods were applied to identify key amino acids and monosaccharides associated with soluble matrices and those which distinguished between symbiotic and non-symbiotic coral types. The amino acids identified were glutamic acid, alanine, serine, threonine and histidine. Monosaccharides included: galactose, mannose, galactosamine, glucosamine and arabinose. The discriminant amino acids and sugars (a total of nine variables) were combined into a single data set and interrogated by principal component analysis (PCA) to expose the variation within. The first three principal components (PCs) explained 80% of the total variance of the data, with PC 1 variation, at 50.1%, driven by galactosamine and serine. Further interrogation of the PCA established non-symbiotic corals contained high levels of serine and galactosamine as well as alanine

and glucosamine while symbiotic corals were high in mannose, galactose, arabinose, glutamic acid and threonine. These results demonstrated that characteristic ratios of certain amino acids and sugars in soluble skeletal matrices could be used to differentiate symbiotic and non-symbiotic corals.

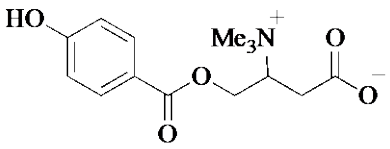
Investigation of zooxanthellae isolated from the flatworm *Amphiscolops* sp. identified the first examples of *Symbiodiniaceae* -derived water-soluble large molecules, zooxanthellatoxins (ZTs) A (**13;** Figure 2.3) and the congener ZT-B (Nakamura *et al.*, 1993). Initial characterisation of these toxins, which induce prominent contraction of rabbit aorta, showed they contain more double bonds and fewer ethereal rings compared to similar toxins isolated from other species of dinoflagellates. Structural similarities between ZTs and palytoxins (Boroujerdi Arezue *et al.*, 2009) suggested that there exist common biogenetic processes, such as the polyketide pathway, that utilise a glycine starting unit and tetrahydropyran ring formation, indicating that *Symbiodiniaceae* may be the source of various toxins isolated from symbiotic marine invertebrates, thus making them an attractive source of novel bioactive materials.



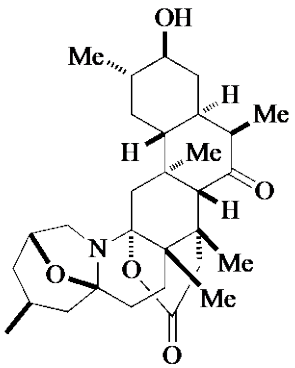
13 (Zooxanthellatoxin-A)



14 (Symbioramide-C16)



15 (Zooxanthellabetaine-A)



16 (Zooxanthellamine)

Figure 2.3 Molecular structures of zooxanthellatoxin-A (13), symbioramide-C16 (14), zooxanthellabetaine (15) and zooxanthellamine (16).

Nakamura *et al* (1998) elucidated four new compounds from ethanol extracts of *Symbiodiniaceae* cells cultured under varied conditions. Two were the betaines, zooxanthellabetaine-A and -B (**15**), another was the C30 alkaloid, zooxanthellamine (**16**), and the last a new ceramide, symbioramide-C16 (**14**) (Figure 2.3). Zooxanthellamine has a

very similar alkaloid structure to a zoanthamine obtained from zoanthids, suggesting this compound was algal rather than host derived (Nakamura *et al.*, 1998). This highlighted that metabolites and the bioactive molecules produced by *Symbiodiniaceae*, could theoretically be manipulated by culture conditions and host factors to produce a variety of different metabolites of interest to natural products researchers.

Compounds related to ZTs have also been isolated from free-living zooxanthellae found in a Hawaiian tidal pool (Onodera *et al.*, 2003). Zooxanthellamide A (**17**, Figure 2.4) has a smaller molecular weight than that of ZTs; however, unlike ZTs, it does not possess bisepoxide and exomethylene moieties. In addition, it contains a pair of amide and sulfate groups, whereas these only exist as lone groups in ZTs. Structural similarities suggest that ZTs and zooxanthellamide A arise from similar biosynthetic pathways; however, subsequent isolation of the δ-lactone derivative zooxanthellamide B led to the proposal that zooxanthellamides were in fact a novel family of large polyhydroxy metabolites (Onodera *et al.*, 2004). This showed that the polyhydroxy metabolism of zooxanthellae was quite diverse.

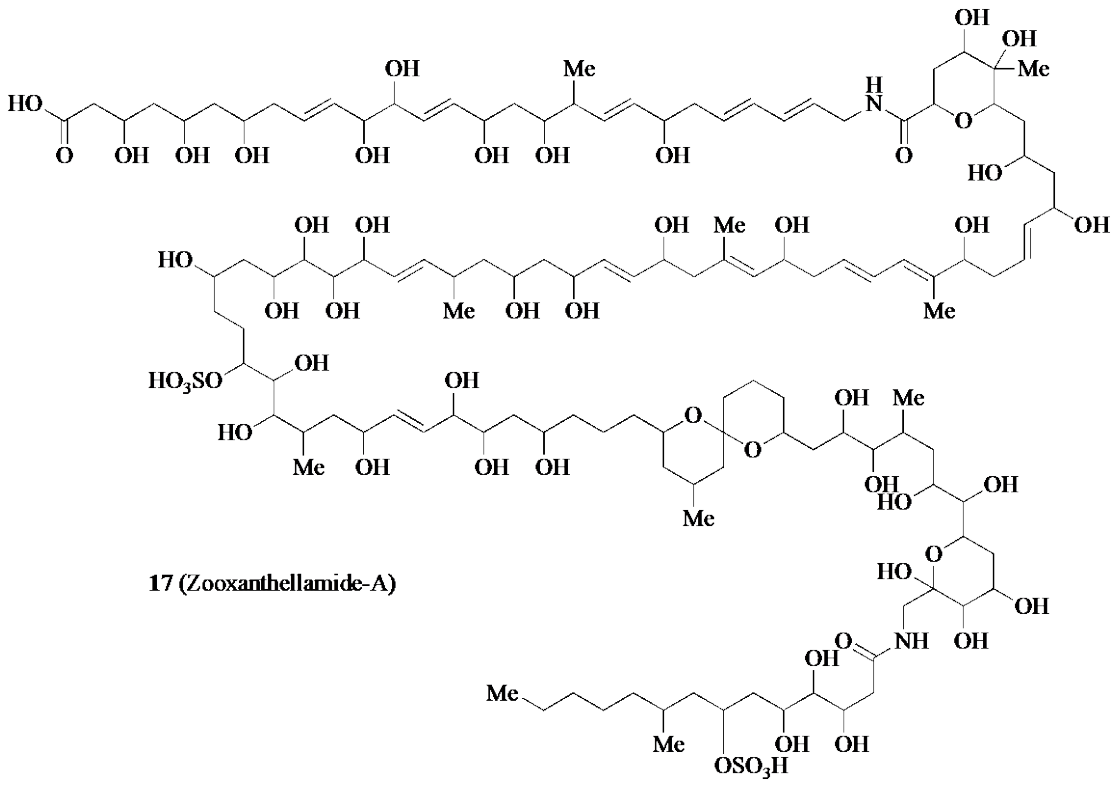


Figure 2.4 Molecular structure of zooxanthellamide-A (17)

2.5 Coral metabolomics

Phenotypic diversity – also referred to as intraspecific variation – in coral species is driven by polymorphism (i.e. genetic differentiation) and encompasses environment-induced variation (i.e. phenotypic plasticity) (Todd, 2008). Morphological assessment, particularly of the coral skeleton, and genotyping have been widely used to characterise phenotypes of the coral holobiont. The measured metabolome, however, is the ultimate expression of an organism's phenotype (Fiehn, 2002). Metabolomics, therefore, may be a more powerful tool for studying phenotypic diversity of both partners at the cellular or molecular level.

In the context of this review, a coral metabolomics study is defined as any published study that reported a metabolomics and/or lipidomics approach (including those using a wider systems biology approach) to examine cnidarians and/or their dinoflagellate symbionts. Since the first named coral metabolomics study in 2009, 35 studies fit these criteria. While they have been varied in their approach and subject, some common themes have emerged; for example, *Aiptasia spp.* has been a popular subject choice due to its informal acceptance and adoption as a model organism of cnidarian physiology (Weis *et al.*, 2008), while temperature, being very relevant to coral bleaching, has been the predominant stressor investigated. What follows is a discussion of the coral metabolomics literature produced since this review was first published in 2010. Key examples are discussed with a focus on metabolite variation, coral-algal-microbial symbiosis maintenance, impacts of climate change, and impacts of anthropogenic agents.

2.5.1 The dawn of coral metabolomics

The first report applying metabolomics to any species of coral was published in 2009 by Boroujerdi *et al* (Boroujerdi Arezue *et al.*, 2009). While not specifically investigating the metabolome of corals nor *Symbiodiniaceae*, they did explore the metabolome of the temperature-dependant coral pathogen *Vibrio coralliilyticus*. This has been linked to coral disease worldwide, and like many other *Vibrio* species, exhibits a temperature-dependent pathogenicity. The effect of temperature on the metabolome was monitored using Nuclear Magnetic Resonance (NMR) spectroscopy and assessed by PCA. Distinct metabolic differences were observed between the virulent (high temperature) and threshold-of-virulency (low temperature) forms of *V. coralliilyticus*, with betaine, succinate and glutamate identified as the metabolites causing the greatest temperature-based separations in the PC scores. With increasing temperature, production of betaine was down-regulated, while

succinate and glutamate were up-regulated; however, significant inter-batch variability was observed. Upon re-extraction, instrumental and statistical reanalysis, systematic errors that may have caused the inter-batch variability were excluded, confirming there must have been significant biological variability during the growth of *V. coralliilyticus*.

Metabolome data describing variation in *Symbiodiniaceae* metabolite profiles was non-existent prior to 2010; however, in 2009 the importance of these daily variations was clearly demonstrated in a toxicity analysis of the unicellular freshwater alga *Scenedesmus vacuolatus*. Kluender *et al.* (2009) used GC-MS to analyse low molecular weight hydrophilic and lipophilic metabolites from synchronized algal populations exposed for 14 hours to the photosystem-II inhibiting herbicide prometryn. Subsequent multivariate analysis identified temporal trends in metabolite levels along with distinct statistical separations between control algal populations and those exposed to the toxin. Results were compared to traditional observational parameters used in phytotoxicity assessment where metabolite levels were found to respond more rapidly to toxin exposure than growth. This was one of the first papers to highlight the potential of metabolomics for the study of unicellular aquatic algae.

2.5.2 Naturally occurring metabolite variation

The metabolic modulation of free-living versus endosymbiotic (within the anemone *Aiptasia pulchella*) *Symbiodiniaceae* over the diel cycle was investigated using synchrotron radiation-based infrared microspectroscopy (SR-IRMS) (Peng *et al.*, 2011). Significant differences in the levels of lipids, nitrogenous compounds, polysaccharides and cell wall components in *Symbiodiniaceae* were detected, revealing that nitrogen limitation was exploited by the host cell to induce lipid and polysaccharide production in endosymbionts. SR-IRMS was also used to examine a single *Symbiodiniaceae* cell (see Gordon *et al.*, 2018 and Chapter 7) and three distinct regions characterised by their unique chemical signatures were established; further demonstrating the potential of this technique for metabolomics approaches.

LC-MS-based metabolomic and genomic approaches have been used in combination to investigate the chemical diversity of three *Symbiodiniaceae* clades (Beedessee *et al.*, 2019). These analyses revealed polyketide synthase genes that were more diversified than non-ribosomal peptide synthetase genes and that evolutionary processes contributed to the diversification. Further, a few of the *Symbiodiniaceae* metabolic pathways were found to be conserved among the three clades, indicating that *Symbiodiniaceae* genomes were well

equipped to generate chemical diversity of secondary metabolites and subsequently adapt to unique host environments.

An LC-MS metabolomics approach was employed to examine the indirect effects of genes of symbiotic and aposymbiotic *Astrangia poculata* and *Symbiodinium psygmophilum* on the cellular environment, independent of the individual bodies in which they reside (the extended phenotype) (Parkinson & Baums, 2014). In this study, PCA clustered the symbiont state of *A. poculata* more strongly than host genotype and a platelet-activating factor was observed at higher levels in aposymbiotic polyps. 13*E*-docosenamide was found to be unique to a *Symbiodiniaceae* sample, highlighting the importance of accounting for intraspecific diversity when designing an experiment. Importantly, these results also highlight the potential of *A. poculata* as a model system for investigating the coral-dinoflagellate symbiosis due to its unique characteristic of harbouring both symbiotic and aposymbiotic polyps.

Natural spatial and taxonomic variation of deep-sea coral species was explored using LC-MS metabolomics (Vohsen *et al.*, 2019). Across three deep-sea sites *Callogorgia delta* exhibited distinct metabolic profiles confirming spatial variation within the taxa, although no unique metabolites were found. Comparison of *C. delta* metabolite diversity with three other deep-sea corals *Stichopathes* sp., *Leiopathes glaberrims* and *Lophelia pertusa*, and one shallow-water species, *A. palmata*, found it to be the least diverse in terms of metabolic richness.

To facilitate analysis and interpretation of such rich and complex chemical data a novel approach to organising mass spectral data of complex mixtures was developed (Garg *et al.*, 2015). The approach involves using a molecular networking technique, which exploits the fact that molecules with similar chemistry and structure have similar fragmentation spectra. Comparison of the fragmentation spectra of cystic fibrosis-afflicted human lung and hard corals, systems that represent complex mixtures arising from diverse microbial communities, identified an inflammatory lipid that was previously unknown in corals. This molecular networking technique has application in building libraries of complex mass spectral data that can assist in the automated annotation of molecular features in a variety of complex mixtures.

The utility of ^{1}H-NMR metabolic profiling to assess natural sample variation, the recovery and quantification of spiked compounds and to discriminate different species of reef-building corals was investigated by Sogin *et al* (2014). ^{1}H-NMR metabolomic profiles

correctly phenotyped four unique reef-building coral species while also highlighting the reproducibility and sensitivity of the technique. A similar NMR-based metabolomics approach was used to investigate spatial variation in soft corals (He *et al.*, 2014). Terpenes, sterols and nitrogenous compounds were identified as important variables in classifying soft coral samples from two different geographical locations and further fractionation of the original coral extracts was found to improve discrimination.

Individually, both LC-MS and ^{1}H-NMR have provided insight into the chemical repertoire of the coral host and algal symbiont. More recently, these techniques have been combined to enable a broader investigation of coral metabolites. A comparative LC-MS and ^{1}H-NMR approach was used to explore the genetic diversity and provenance of 16 *Sarcophyton* soft coral species from the Red Sea (Farag *et al.*, 2016). A total of 120 metabolites, including 65 diterpenes, 8 sesquiterpenes, 18 sterols, and 15 oxylipids were identified, with the cembranoids and oxylipids driving species differences. Cembranoids were also found to drive discrimination of origin. Notably, aquarium reared soft corals were less enriched with cembranoids compared to wild corals. In this study, PCA was as effective as Orthogonal Partial Least Squares in predicting species origin, which highlights the distinguishing nature of metabolic profiles given the unsupervised nature of PCA.

Lohr *et al* (2019) applied a similar approach to unravel the differential responses of genetically different individuals of the same species (i.e. different genotypes) to the same environmental conditions. Distinct profiles for each of three genotypes of *A. cervicornis* from a common ecosystem were obtained, revealing the extent of their intraspecific variation while identifying several metabolites driving separation between genotypes. These findings highlight the metabolic complexity of the coral holobiont and the need to understand the naturally occurring metabolite variation of the study population, an aspect that is particularly relevant if metabolomics is to be included in biomonitoring programs at regional scales.

2.5.3 Maintenance of coral symbioses

The establishment of a symbiotic association involves changes in the biochemistry and metabolic network of all the partners involved; maintenance of this symbiosis is largely mediated by the transfer of metabolites between partners. Metabolomics can explore this biochemistry holistically, and in detail, establishing the physiological state that maintains the partnership between host and symbiont.

The mutualistic translocation model, as applied to coral-algal interactions, suggests that symbiont-derived metabolites, including lipids and fatty acids, are responsible for 90-99% of the hosts energy requirements (Gordon & Leggat, 2010, Yellowlees *et al.*, 2008, Falkowski *et al.*, 1993, Latyshev *et al.*, 1991, Patton *et al.*, 1977, Kellogg & Patton, 1983). With this in mind, ^{13}C stable isotope incorporation into fatty acids was compared between symbiotic and aposymbiotic *Aiptasia pulchella*, *Acropora millepora* and free-living dinoflagellate cultures using LC-MS lipid profiling (i.e. lipidomics) (Dunn *et al.*, 2012). The fatty acid synthesis rates were attributed to only two algal lipogenesis pathways with no evidence of symbiont-derived fatty acids, fatty acid derivatives or dissolved inorganic carbon being utilised in the host long chain (late n-6 pathway) fatty acid synthesis. These findings did not align with the mutualistic translocation model. Indeed, a similar study compared the mass features from symbiotic and aposymbiotic *Aiptasia pallida* (Garrett *et al.*, 2013) and found 149 differential features that differed between the two symbiotic states. Several lipid classes were identified including glycerophospholipids and phosphonosphingolipids, further contradicting the mutual translocation model and highlighting the need for more research in this area.

^{13}C stable isotope labelling and GC-MS was employed in a third study to track the fate of carbon in *Aiptasia* colonised with either the native symbiont, *Breviolum minutum* or the non-native symbiont, *Durusdinium trenchii* (Matthews *et al.*, 2018). *Aiptasia* harbouring non-native symbionts exhibited reductions in the abundance and diversity of carbohydrates and lipogenesis precursors. Additionally, significant alterations to the host molecular signalling pathways, as well as differential activity in the antioxidant- and ammonium-producing pathways, were observed. It was concluded that such significant metabolic differences between native and non-native symbionts may limit the success of the host switching to more thermally tolerant, but non-native, partners.

2.5.4 Coral-microbial interactions

In recent years, extracellular coral metabolites (i.e. the exometabolome) have received increased attention because of the need to link microbial community dynamics, host physiology and ecosystem biogeochemistry for a better understanding of the ecological roles of coral-associated microbes (Kelly *et al.*, 2018). The significant role of microbial communities in reefs has led to the proposal of "ecosystem microbiology", to link the vast amounts of data arising from the omics sciences to establish holistic monitoring of coral

reefs and provide management with microbial indicators of coral health and ecosystem processes.

Metabolomics was employed to examine competition between corals and non-self holobionts such as algae and microbes (Quinn *et al.*, 2016). This study found that algae and microbes induced tissue damage in corals during competitive interactions, promoting the production of platelet activating factor (PAF); a single fatty acid chain phospholipid that is a known pro-inflammatory signal molecule in humans. This finding was backed by rigorous compound identification. Complementary transcriptome data from the coral confirmed that the expression of the gene encoding the protein that converts Lyso-PAF to PAF increased during this competitive interaction.

In a novel approach to study the exometabolome, Ochsenkühn *et al.* (2018) investigated the molecular environment surrounding two species of coral. Metabolomics and meta-barcoding were employed to analyse seawater samples collected from the surface of corals and at distances of 5 and 50 cm. Each species was found to harbour unique bacteria and metabolites at their surface that were distinctly different from those in the surrounding seawater. The molecules at the surface of the corals were identified as chemo-attractants, antibacterials and signalling molecules that may play a significant role in structuring the surface-associated microbial community. Metabolites were also discovered that distinguished diseased corals from their healthy counterparts, likely produced by pathogenic or opportunistic bacteria. This discovery is significant in the context of coral biomonitoring, as it presents a non-destructive method for assessing the chemical state of corals and indicators of coral disease.

Another study compared coral metabolome profiles (^{1}H-NMR) with microbial community composition (Sogin *et al.*, 2017). This untargeted metabolomics study assessed coral holobiont functional status in the presence of various microbial partners. The relative abundance of different *Symbiodiniaceae* sub-clades and microbial communities produced positive and negative metabolomic signatures, with respect to nutrition, in the holobiont. Positive correlations, for example, were observed between cyanobacteria and nitrogen containing branched-chain amino acids and acetate, supporting the hypothesis that cyanobacteria produce nitrogen compounds to support coral nutrition. On the other hand, opportunistic gamma-proteobacteria were negatively correlated, indicating they were consuming the compounds. This work set a foundation for future studies of microbial

coral symbionts and their effect on coral functional status while emphasising the need for more targeted studies.

Meta-mass shift chemical profiling, a method analogous to molecular networking techniques (Garg *et al.*, 2015), utilises the mass differences arising from ion fragmentation and knowledge of molecular networks to determine how molecules are related. The technique was applied to seven coral, algal and fungal mat holobiont metabolomic datasets (Hartmann *et al.*, 2017). Distinct mass shift profiles were identified for all holobionts, suggesting they modified the same molecules in different ways, even within the same genera and species, and despite high genomic similarity. This approach to the analysis of tandem mass spectral profiles has again highlighted the chemical diversity and metabolic complexity of the coral holobiont.

2.5.5 Impacts of climate change-related stressors

To date, much of the scientific literature reporting on climate change related impacts to coral reefs assumes a business as usual emissions scenario of 600-1000 ppm of atmospheric CO_2 (Hughes *et al.*, 2017a). With greater global concern for climate change mitigation, such as the development of new 'green' technologies, this scenario is less likely to eventuate. Nevertheless, even under the most optimistic greenhouse gas emissions scenario considered by the Intergovernmental Panel on Climate Change (IPCC RCP 2.6 2014), the longer-term warming trend is still expected to be in the range of 0.30 ℃ to 0.68 ℃ (Hughes *et al.*, 2017a, Morice *et al.*, 2012). Modelling predicts this level of warming will have severe consequences for coral reefs, especially during summer heatwaves when sea surface temperatures spike above the summer maximum.

Mounting scientific evidence of the longer-term impacts of climate change on coral reefs has highlighted the need to better understand coral-stressor interactions under more realistic emissions scenarios (Hughes *et al.*, 2017a, Hughes *et al.*, 2010, Pandolfi *et al.*, 2011, Steffen *et al.*, 2015). This is a difficult task because the biological response to smaller scale fluctuations in water temperature and CO_2 are not easily detected, require more time to present physically, and are more difficult to quantify. Metabolomics, with its superior sensitivity and resolution, can detect changes in the metabolome from minor disturbances and even natural variation (see section 2.5.2), offering a solution to the problem of detection and quantification. However, long-term studies will still be required to detect those changes that take more time to present – a gap that could be filled by a metabolomics-based biomonitoring program.

Temperature

Elevated temperatures have been implicated in the breakdown of the coral-algae symbiosis leading to coral bleaching and, if protracted, coral mortality. While the mechanism of action of this collapse remains somewhat unclear (Obura, 2009, Bieri *et al.*, 2016), metabolomics has provided an insight into the heat-stress responses of both the host and symbiont.

The metabolome profiles of four symbiosis-forming species of *Symbiodiniaceae* cultured under different light and temperature conditions was examined utilising an untargeted GC-MS metabolomics approach designed to generate hypotheses (Klueter *et al.*, 2015). GC-MS and Random Forests modelling was able to detect significant differences in two major free sterols produced by the four species of *Symbiodiniaceae*, with *Symbiodinium psygmophilum* being the most unique. In addition, changes were detected in the production of inositol, selected sterols and glycerol in *Symbiodiniaceae* exposed to varying temperature and light conditions, although these were considered unlikely to cause a disruption of the symbiosis. This study demonstrated the utility of untargeted metabolomics to observe previously unknown biochemical processes, leading to refined hypotheses.

A ^{13}C stable isotope tracer experiment in combination with GC-MS metabolomics was employed to profile polar and semi-polar compounds in the *Acropora aspera* host and their dinoflagellate symbionts during thermal bleaching (Hillyer *et al.*, 2017b). Changes in GC-MS profiles of both partners were detected, concomitant with bleaching severity. In the dinoflagellate symbiont, accumulations of free fatty acids and α-tocopherol, and a decline in carbohydrates, amino acids and ascorbate, were observed while in the coral host. Undoubtedly, this study furthered the understanding of metabolome changes in the coral holobiont. This was also the first reported attempt to measure the metabolome of the endogenous symbionts and of the coral host after their separation post sampling.

In a follow up to their 2017 study, Hillyer *et al.* (2018) investigated the changes in autotrophic carbon fate during thermal bleaching of *Acropora aspera*. Partner-specific changes in carbon fate were observed in response to heat stress and were found to worsen with prolonged exposure. *Symbiodiniaceae* continued to produce and translocate organic products to the host despite significant photodamage. As bleaching progressed, however, symbiont production of long chain fatty acids decreased to a minimal state, which was also evident in the host fatty acid pools. This important finding, which shows how the quality

of translocated products is reduced during thermal bleaching, will no doubt encourage more research into the mechanisms that trigger the breakdown of symbiosis.

A GC-MS-based metabolomics investigation of thermally induced changes in the metabolome of the *Aiptasia-Symbiodiniaceae* holobiont (Hillyer *et al.*, 2016) found polyunsaturated fatty acids levels increased in the symbiont and decreased in the host, while glycolysis intermediates, amino acids and their intermediates accumulated in both partners. Pathway activity analysis associated these changes with alterations to central metabolism, oxidative state, cell structure, biosynthesis and signalling. This study provided valuable information of the biochemical interplay between host and symbiont under thermal stress and highlighted the importance of hypothesis-generating metabolomics studies.

Subsequently, Hillyer *et al.* (2017a) mapped carbon fate during bleaching of the *Aiptasia-Symbiodiniaceae* symbiosis. Substantial increases in non-labelled lipid and starch stores were observed in the dinoflagellate, while ^{13}C enrichment of compounds associated with ongoing carbon fixation were maintained. In the host, some downstream pathways showed altered carbon states; however, minimal change was observed in ^{13}C-enriched glucose, a major symbiont-derived compound. They concluded that symbionts continue to provide compounds for the host at significant cost to themselves, requiring the symbiont to draw on its own energy stores during thermal bleaching.

The *Aiptasia-Symbiodiniaceae* model system was investigated using a comparative metabolomics and transcriptomics approach to characterise the molecular interplay between host and symbiont (Matthews *et al.*, 2017). Driving this study was the hypothesis that reef corals may adapt to climate related stressors by changing their dominant symbiont type. This led to the examination of gene expression and metabolome profiles of *Aiptasia* sp. when colonised by two distinct symbionts: the thermally tolerant and opportunistic *Symbiodinium trenchii* and the regular or normal symbiont *Symbiodinium minutum*. In hosts colonised by *S. trenchii*, immunity processes, oxidative stress responses, G protein-associated stress signalling and lipid signalling were all significantly activated, while a strong shift toward carbohydrate and lipid transport and storage was also observed. Notably, of the 89 metabolites identified in *S. trenchii*-colonised hosts, all but eight were present in higher abundance. The host immunoresistance and heterotrophy observed in this study has provided further evidence casting doubt on the ability of thermally tolerant *Symbiodiniaceae* species to assist corals in adapting to climate change.

Synchrotron radiation Fourier transform infrared microscopy metabolomics provided insight into the biochemical changes induced by elevated temperatures in individual *in hospite* and expelled *Symbiodiniaceae* cells from thermally stressed *Acropora millepora* (Petrou *et al.*, 2018). *In hospite* symbionts showed increases in free amino acids and molecule phosphorylation along with a decline in protein content. Conversely, expelled symbionts showed decreases in amino acids and molecule phosphorylation and an increase in protein and lipid content. Surprisingly, this suggests that expelled symbionts are more functional after expulsion and that the metabolic profiles of heat stressed, and expelled symbionts are distinctly opposite. This metabolomics approach revealed alterations in the symbiont lipid metabolism that corroborates many of the findings reported by Hillyer *et al.* (Hillyer *et al.*, 2017a, Hillyer *et al.*, 2017b).

Ocean acidification

Ocean acidification adversely affects the calcification process of corals and increases the passive dissolution of coral skeletons (van Woesik *et al.*, 2013). Under a "business as usual" emissions scenario, where atmospheric CO_2 concentrations are expected to reach 800 ppm, the rate of dissolution is predicted to outstrip the average rate of coral reef growth by 2100 (Hughes *et al.*, 2017a, van Woesik *et al.*, 2013). However, due to a lack of research, it is currently difficult to predict the coral response to the more likely end-of-century atmospheric CO_2 concentrations of 450-550 ppm that would maintain a pH of 7.9-8.1 in most tropical waters (Hughes *et al.*, 2017a, IPPC RCP 4.5, 2014, ICCP RCP 2.6, 2014). Unlike temperature, experimentally manipulating and controlling small changes in CO_2 concentration for any length of time is technically challenging, identifying and quantifying its effect on coral even more so.

The capacity for metabolomics to detect CO_2-induced variations in coral metabolomes has been clearly demonstrated. Putnam *et al.* (2016) tested the hypothesis that ocean acidification-induced DNA methylation (i.e. epigenetics) is linked to phenotypic plasticity in two scleractinian corals. In addition to examining host DNA methylation and calcification rates, ^{1}H NMR metabolomics was used to measure phenotypes of *Pocillopora damicornis* and *Montipora capitata* clonal fragments after exposure to high pCO_2 (~1320-2360 μatm; ΔpH ~0.3). *P. damicornis*, the more sensitive of the two species, exhibited DNA methylation, reduced calcification rates and changes in the metabolome, clearly linking phenotypic plasticity with DNA methylation. This was an important finding that supports the use of assisted evolution to aid corals in acclimating to rapid changes in their environment (van Oppen *et al.*, 2015, Huey *et al.*, 1999, Reusch, 2014). However, the pCO_2

and ΔpH used here is consistent with a doubling of atmospheric CO_2 concentrations of ~800 ppm, an unlikely scenario for the average reef this century (Hughes *et al.*, 2017a). Consequently, the suitability and applicability of metabolomics approaches in monitoring the long-term coral response to realistic, end-of-century, ocean acidification scenarios warrant further assessment.

Synergistic effects

While it is imperative to develop approaches and databases to establish the induced changes in the coral metabolome in response to individual stressors, the synergistic effects of ocean warming and acidification on corals must be considered to truly understand the impact of climate change. Comparison of the physiology of *P. damicornis* exposed to simulated future ocean temperatures and elevated pCO_2 levels with GC-MS and LC-MS-based metabolomics offers insight into these stress-induced changes (Sogin *et al.*, 2016). Significant variables were identified and functionally analysed, demonstrating that corals alter carbohydrates, cell structural lipids and signalling compounds in response to these elevated stressors. For the first time, coral metabolic profiles were modelled to predict net photosynthetic rate (measured as oxygen flux) and corals exposed to synergistic temperature and pCO_2 treatments. Understanding this link between the metabolome and photosynthetic performance is critical to identifying early biomarkers of coral stress with potential application in coral reef biomonitoring.

A large-scale multi-platform approach using GC-MS and LC-MS metabolomics profiled the metabolome of the soft corals *Sarcophyton ehrenbergi* and *S. glaucum* and their dinoflagellate symbionts under simulated ocean warming and acidification (Farag *et al.*, 2018). Soft coral response to thermal and high CO_2 stress increased the free amino acid pools, particularly alanine, which is an early indicator of acute anaerobiosis, and proline, which impacts the osmotic balance. The disturbance of osmotic balance was proposed as a biomarker of environmental stress in marine invertebrates. This study also attempted to modulate the bleaching response using four chemical treatments known to mitigate heat stress in plants (alanine, GABA, nicotinic acid and proline); unfortunately, these chemical treatments were unsuccessful. Regardless, identifying molecular biomarkers of coral stress and exploring novel solutions remains the goal for many metabolomics studies.

2.5.6 Impacts of anthropogenic agents

A key priority of managers is to halt and reverse the decline in water quality to maintain the health and resilience of coral reefs (Cooper & Fabricius, 2007, Cooper *et al.*, 2009,

D'Angelo & Wiedenmann, 2014). Water quality is affected by a number of factors such as nutrient loading, sedimentation, pesticides, herbicides and even sunscreen. Changes in water quality impact coral physiology and induce a wide range of responses (Roder *et al.*, 2013, Browne *et al.*, 2015); however, little is understood about the impacts on coral metabolism and health, nor how changes in water quality present in the coral metabolome. A better understanding of how water quality factors alter the coral metabolome will further our knowledge of coral health while also providing highly sensitive and dynamic indicators of water quality itself.

Metabolomics biomonitoring of coral health and water quality was recently reported by Tang *et al.* (2018). Orthogonal projections to latent structures (OPLS) modelled the LC-MS-derived lipidome profiles and photosynthetic efficiency of *Seriatopora caliendrum* in response to Irgarol 1051 exposure; a common photosystem II herbicide used in marine antifouling applications. Irgarol 1051 exposure caused photoinhibition and increased production of reactive oxygen species (ROS) in *S. caliendrum*, consistent with previous studies (Downs & Downs, 2007, Maxwell & Johnson, 2000, Müller *et al.*, 2001). Changes in glycerophosphocholines (a major membrane lipid) were revealed in the lipidome after 4 days of exposure to Irgarol 1051, which was attributed to photosynthetic shortages and cell membrane accommodation of ROS. The models, built using a large suite of metabolites, offer superior calibration and resilience to background noise associated with natural variation – a common issue when using only a handful of molecular biomarkers.

Acrylic acid accumulation in hard and soft corals (*Acropora* sp. and *Lobophytum* sp., respectively) was investigated as a potential molecular biomarker of deteriorating water quality (Westmoreland *et al.*, 2017). NMR metabolomics was employed to analyse changes in acrylic acid, and the polar metabolome, in response to increased ammonia and phosphate concentration and decreased calcium concentration. Deteriorating water quality resulted in nearly two-fold increases in acrylic acid and acetate concentration in both coral species. Increases in trigonelline concentration were also observed in both coral species while lactate and thymine decreased in concentration. The increase in acrylic acid concentration in corals exposed to reduced water quality contrasted with decreases observed in corals exposed to thermal stress. The authors proposed that acrylic acid accumulation could be a protective response to oxidative stress caused by excess ammonia.

Changes in metabolite diversity of the soft coral, *Nephthea spp.* as a response to water quality was explored by LC-MS metabolomics (Januar *et al.*, 2012). Traditional inorganic

nitrogen and phosphate, salinity and pH measurements were used as indicators for water quality index. In a unique data-driven environmental metabolomics approach, an ecological index was generated from the LC-MS data. The index included indicators such as metabolite richness, metabolite diversity and metabolite evenness. A correlation analysis of the water quality index and the LC-MS ecology index found a significant relationship between metabolite richness, particularly amongst terpenoids, and water quality. Water quality and environmental stress affected the metabolite richness of soft corals, with potential consequences for pharmaceutical studies and the capability of soft corals to defend against predation.

With a focus on natural products and drug development, LC-MS metabolomics has been used to assess the potential of abiotic and biotic elicitors to increase secondary metabolite production in soft corals (Farag *et al.*, 2017a, Farag *et al.*, 2017b). *Sarcophyton ehrenbergi* was subjected to one physical and five chemical plant elicitors: physical wounding, methyl jasmonate, salicylic acid, $ZnCl_2$, glutathione and β-glucan. Methyl jasmonate inhibited photosynthetic efficiency (Farag *et al.*, 2017a) and salicylic acid and $ZnCl_2$ altered metabolite production, eliciting a significant increase in the production of cholesteryl acetate and sarcophytonolide I, respectively. In a second study by Farag *et al.* (2017b), *S. glaucum* and *Lobophyton pauciflorum* were exposed to three oxylipin analogues, prostaglandin, methyl jasmonate, and arachidonic acid; the diterpene precursor, geranylpyrophosphate, and physical wounding. Multivariate data analysis revealed changes in the secondary metabolism of *S. glaucum* exposed to prostaglandin and methyl jasmonate. Terpene and sterol metabolites, such as campestene-triol and cembranoid were upregulated, with prostaglandin eliciting the strongest response. The effects of these chemical elicitors were less pronounced in *L. pauciflorum*, which pointed to a differential oxylipin response in soft corals.

Metabolic profiling has also been used to evaluate the toxicity and bioaccumulation of octocrylene in *Pocillopora damicornis*, a common ingredient in many sunscreens and cosmetics (Stien *et al.*, 2018). Metabolic profiles showed that octocrylene was transformed into fatty acid conjugates yielding highly lipophilic octocrylene analogues, which are more likely to accumulate in the coral tissue than octocrylene itself. Exposed corals also had higher levels of acylcarnitines, suggesting a disturbance in the fatty acid metabolism related to mitochondrial dysfunction. The discovery that octocrylene forms lipophilic analogues suggests environmental monitoring of octocrylene has been underestimating octocrylene

bioaccumulation for some time. The authors rightly argue for an in-depth re-evaluation of octocrylene toxicity and its bioaccumulation rate in the ocean's food chain.

2.6 Concluding Remarks

It is clear from the literature available that metabolite analysis of *Symbiodiniaceae* symbioses peaked in the 1970s and 1980s with a keen interest in the nutritional roles that symbionts provided their hosts and how they interacted. The pioneers in these early days of metabolite analysis made a significant contribution to the knowledge that marine science benefits from today. Little was understood in those early years about the number or diversity of metabolites produced by the symbiont, nor how the host utilised them, yet they proved the symbiont was providing nutrition to the host and that free-living zooxanthellae behaved differently to their symbiotic counterparts.

Growing concerns over climate change in the 1990s resulted in a greater focus on monitoring the photophysiological responses of corals and the loss of coral diversity at the expense of metabolism research. This trend has reversed to some extent in recent times with improvements in analytical platforms and the increased popularity of metabolomics. Nevertheless, the fundamental cellular processes and metabolic interplay that establishes and sustains the coral holobiont remain somewhat neglected to this day. For example, coral holobiont phenotype identification, signalling pathways between host and symbiont, and connecting gene and protein expression with metabolite regulation are just a few research areas that have yet to be investigated in any real detail. The specific and synergistic effects of anthropogenic stressors, including ocean warming and acidification (under the most likely emissions scenarios for this century) and water quality on corals remain poorly understood. With coral reefs continuing to experience accelerating rates of change and poor recovery despite decades of research, metabolomics methods provide a solid foundation towards the development of a metabolomics-based biomonitoring program that may address many of these knowledge gaps.

Chapter 3

Extraction protocol for non-targeted NMR and LC-MS metabolomics-based analysis of hard coral and their algal symbionts

The content of this chapter has been published as:

> Gordon, B. R., Leggat, W. & Motti, C. A. (2013) Extraction Protocol for Nontargeted NMR and LC-MS Metabolomics-Based Analysis of Hard Coral and Their Algal Symbionts. In: *Metabolomics Tools for Natural Product Discovery: Methods and Protocols* (eds. U. Roessner & D. A. Dias). Humana Press, Totowa, NJ.

Benjamin Gordon wrote this chapter, designed and conducted the research, and analysed the data. Co-authors provided intellectual, experimental and editorial support.

3.1 Introduction

The coral holobiont is a complex symbiotic organism that consists of the animal host and a plethora of intracellular and extracellular macro- and microbiota, such as fish, crustaceans, polychaetes, dinoflagellates, prokaryotes, viruses, fungi, archaea and endolithic algae (Rohwer *et al.*, 2002, Stella *et al.*, 2011, Bourne *et al.*, 2009). While many of the these have mutualistic associations with the coral host, the most well-known and studied of them is that of the photosynthetic dinoflagellate, *Symbiodiniaceae*. Members of this diverse genus can live freely in the water column, or sediment (Adams *et al.*, 2009), and also in symbiosis with a number of marine invertebrates, such as corals, giant clams and anemones (Stat *et al.*, 2006). This symbiosis allows the animal host to meet part or all of its carbon requirements through autotrophy, thereby gaining a competitive advantage through increased fitness (Muscatine & Porter, 1977, Trench, 1979). The relationship enables the bilateral exchange of metabolites, including metabolites that are not produced solely by either organism (Lewis & Smith, 1971, Gordon & Leggat, 2010).

The success of this symbiosis is based upon the invertebrate host supplying inorganic nutrients to the algal symbiont, which are returned as organic compounds and used to supplement the host's energy and nutritional demands (Yellowlees *et al.*, 2008). In the coral symbiosis, *Symbiodiniaceae* are found in a host-derived vacuole (symbiosome membrane) within the gastrodermal cell layer, which forms during the acquisition of the algal symbiont. The symbiosome membrane closely resembles those in legumes where the plant membrane encloses the symbiotic rhizobium cells (Roth *et al.*, 1988, Rands *et al.*, 1993). Consequently, the exchange of all organic and inorganic nutrients must proceed through this cell membrane and thus it is critical to the metabolic interaction between the symbiont and host.

Of the metabolites involved in the algal-invertebrate symbiosis, not all are involved in nutritional roles. Compounds such as mycosporine-like amino acids (MAAs) play an important role in the protection of coral from ultraviolet light, in addition to acting as free radical scavengers (Dunlap & Yamamoto, 1995). Likewise, dimethylsulfoniopropionate and its breakdown products, dimethylsulfide and acrylate, are involved in free radical scavenging (Sunda *et al.*, 2002), osmotic- and cryoprotection (Trevena *et al.*, 2000) and the structuring of microbial communities (Miller *et al.*, 2004, Sjoblad & Mitchell, 1979, Raina *et al.*, 2010). While the biological role of some of these metabolites have either been identified or closely examined, there are still many metabolites of which the exact function remains unknown. For example, free amino acids and other small compounds may act as host release factors that stimulate the production and release of other metabolites from *Symbiodiniaceae* to the host (Muscatine, 1967, Trench, 1971c, Gates *et al.*, 1995).

What's more, compounds such as zooxanthellatoxins (Nakamura *et al.*, 1993), zooxanthellamides (Onodera *et al.*, 2003), betaines, alkaloids and ceramides (Nakamura *et al.*, 1998) have all been isolated from either *Symbiodiniaceae* or invertebrates in symbiosis with *Symbiodiniaceae*; however, their biological role still remains unclear. On that note, the diversity of metabolites involved in the algal-invertebrate symbiosis, along with the theoretical prospect of changing growth conditions and host factors to manipulate metabolite production, highlights the potential role that metabolomics will have in elucidating metabolites of interest from both hard coral and their algal symbionts for not only the biochemical sciences, but also for natural products research.

The extraction protocol presented here was developed with the following criteria in mind. Firstly, the protocol had to be user friendly and as such, should not require any further handling or input from the analyst than was absolutely necessary. The benefits of such an approach were two-fold; in addition to reducing the working time, it also ensured that any interference from the analyst that may affect the integrity of a sample was kept to a minimum. The next criterion, maximising the number of features detected by ^{1}H NMR and LC-MS, was based on the non-targeted metabolomics approach taken. As such, the choice of extraction solvent and analysis conditions played an integral role in fulfilling that requirement and are discussed later in more detail. The last criterion, but by no means the least, was reproducibility. Reproducibility is the key to a successful metabolomics study with the results of statistical analyses, and hence the conclusions that can be drawn from an experiment, being reliant on reproducible data. Consequently, not only were experimental conditions kept constant and similar between samples, sample collection and preparation were also regulated by maintaining a constant handling time and temperature. As such, aspects of the protocol that could utilise multiple techniques (i.e. sample concentration) for achieving the same outcome were thoroughly tested to ensure that the least disruptive technique was employed.

The choice of extraction solvent played a crucial role in the development of this protocol. For metabolomics, this choice depends largely on the metabolites of interest, the analytical platform used, and the hypothesis put forward by the researcher. If, for example, one was interested in employing a targeted metabolomics approach, such as the analysis of free amino acids, then in most cases an appropriate solvent system that targets those compounds is deemed most suitable. However, it is important to keep in mind that in symbiotic coral, the organism is comprised of two distinctly different cell types, animal and algal cells, of which the latter have more robust cell walls than that of their host. Consequently, many high polarity solvent systems (i.e. those high in water content and suitable for free amino acid extraction) will not easily permeate the cell walls of the algal symbionts and would therefore require the employment of more time-consuming

mechanical disruption methods. With that in mind, a biphasic methanol-chloroform extraction that is capable of permeating the algal cell walls (Viant, 2008), while partitioning polar and non-polar metabolites into two separate phases, would in most cases, be deemed most suitable for the targeted metabolomics analysis of symbiotic coral.

Considering the two distinct cell types of symbiotic coral, and the fact that this protocol was developed for a non-targeted metabolomics approach, the effectiveness of five commonly used solvents employed in the extraction of compounds from both plant and animal material were examined. Each extraction was carried out in the same manner as described in the methods using the symbiotic hard coral, *Acropora aspera*. The solvents tested for this protocol were; 100% methanol, 70% aqueous methanol, 90% aqueous acetone, 100% water and a methanol:dichloromethane:water (MeOH:DCM:H_2O) biphasic solvent system. Of the five solvents, the 100% methanol and 90% aqueous acetone solvents produced the greatest number of features in both the ^{1}H NMR and LC-MS analyses (see Figure 3.1). However, they also extracted large amounts of lipid, which is detrimental to reversed-phased LC-MS analysis and as such, were eliminated as viable extraction solvents. The 100% water solvent was eliminated after microscopy analysis of the residual biological matter showed that the algal cells were not effectively lysed. In 100% water, enzymatic and metabolic activity was not halted, affecting the integrity of the sample. While the biphasic MeOH:DCM:H_2O extraction may be suitable for a targeted metabolomics analysis, there was a distinct disadvantage of this technique when applied to the non-targeted metabolomics analysis of symbiotic hard coral. Essentially, too many of the less polar metabolites partitioned into the organic phase of the solvent system, along with the lipids. This resulted in an aqueous phase that had very few features for both NMR and LC-MS analyses (see Figure 3.1). Of the five extraction solvents examined, the 70% aqueous methanol solvent displayed excellent reproducibility, extracted minimal amounts of lipid, effectively lysed the algal cells and complied with the proposed criteria. Hence, the 70% aqueous methanol solvent was considered the most suitable for a non-targeted metabolomics study of hard coral and their algal symbionts.

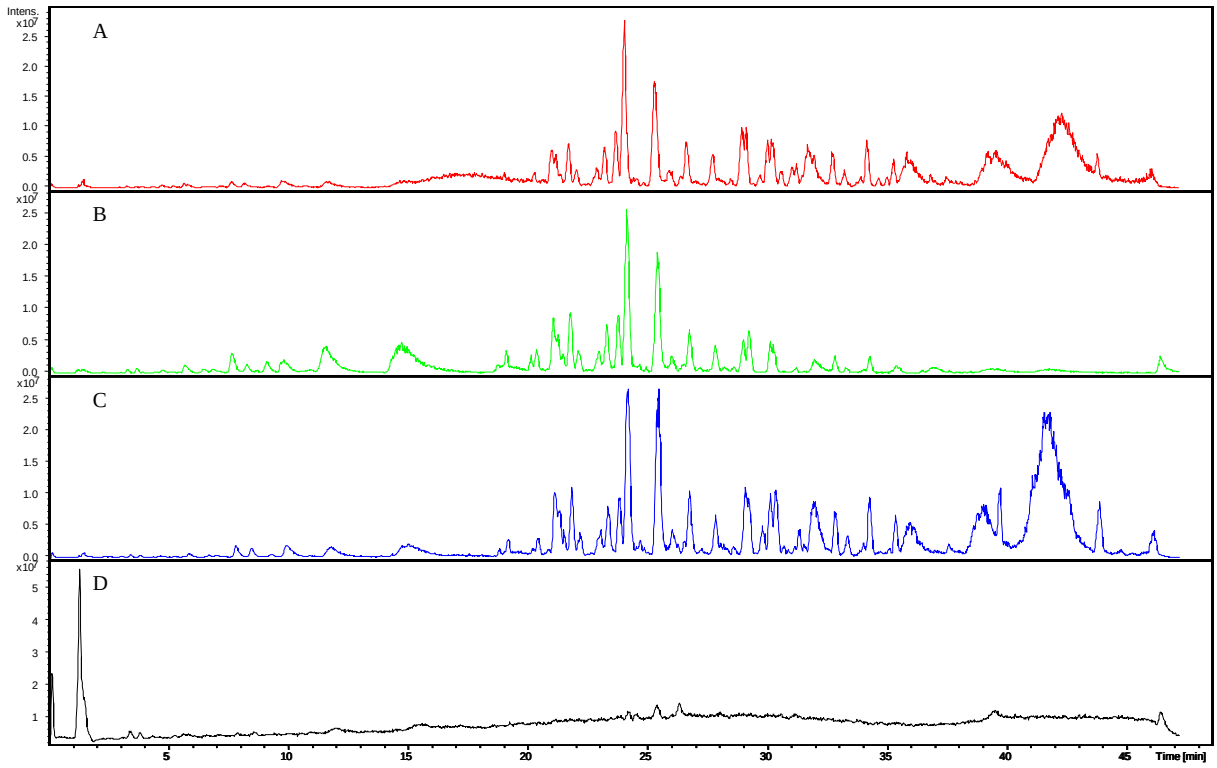


Figure 3.1 LCMS comparison of four different solvents. Total ion chromatograms of 100% methanol (A), 70% aqueous methanol (B), 90% aqueous acetone (C) and the aqueous phase of the MeOH:DCM:H_2O extraction (D)

Metabolomics has become a popular technique for the study of biological samples, which can be attributed to the high number of quality studies published in a diverse number of fields such as environmental science, medicine and pharmaceuticals. A quality metabolomics study depends largely on the hypotheses put forward by the researcher and a thorough understanding of the scientific principals involved. Consequently, each study needs to be considered in the context of the hypothesis. Presented here is an extraction protocol for a non-targeted metabolomics analysis of symbiotic hard coral using LC-MS and ^{1}H NMR. The protocol requires minimal user input and provides reproducible and reliable results using readily available labware and reagents. Every effort has been made to provide the reader with all the details required to perform the technique, including many of the costly and time consuming "pitfalls" or "traps" that were discovered during its development. This protocol can be confidently accomplished by those with less experience in the extraction and analysis of symbiotic hard coral and their algal symbionts.

3.2 Materials

Extraction solvents should be prepared in ultra-clean glassware (see **Note 1**) at 25 °C using Mass Spectrometry (MS) grade solvents (see **Note 2**) and ultrapure (MilliQ) water with a minimum

resistance of 18 MΩ cm^{-1}. Extractions should be performed at approximately 4 °C or cooler (see Note 3) in ultra-clean glassware or disposable scintillation vials of known purity. High quality centrifuge tubes are recommended for removing particulates from liquid extracts (see Note 4). To avoid unwanted contamination, all aspects of the extraction process should be performed while wearing appropriate gloves washed with MilliQ water.

3.2.1 Coral collection and quenching materials

1. Liquid nitrogen
2. Disposable foam dewar
3. 50 mL plastic centrifuge tubes (Nunc/Thermo Fisher Scientific, Scoresby, VIC, Australia) (see Note 4)
4. Double-action, Stille-Liston bone cutters for cutting coral into nubbins
5. Stainless steel forceps or clamps
6. Liquid nitrogen or dry ice for temporary storage
7. Minus 80 °C freezer for long-term storage
8. Nally bin with a tether attached

3.2.2 Sample extraction, concentration and clarification components

1. 70% aqueous methanol extraction solvent: Using a clean glass measuring cylinder, measure and combine 70 parts of MS grade methanol to 30 parts of MilliQ water up to the required volume and store in an acid-washed glass Schott bottle or similar at minus 20 °C until required for use
2. 15 mL centrifuge tubes (Nunc/Thermo Fisher Scientific, Scoresby, VIC, Australia) (see Note 4)
3. 20 mL glass scintillation vials (Sigma) (see Note 5)
4. Dry ice or liquid nitrogen contained in two disposable foam dewars
5. Double-action, Stille-Liston style bone cutters
6. Glass Pasteur pipettes and silicone pipette bulb
7. Freeze dryer (Dynavac FD12)
8. Speed vacuum evaporator (Savant/Thermo Fisher Scientific, Scoresby, VIC, Australia)
9. Sonic water bath maintained at 0 - 4 °C
10. An appropriate labelling system (see Note 6)
11. An appropriate rack for upright handling and storage of glass vials in a minus 80 °C freezer
12. Centrifuge with a carousel suitable for 15 mL centrifuge tubes

3.2.3 LC-MS materials and conditions

1. MS certified sample vials (see **Note 7**) (Verex™, Phenomenex, Lane Cove, NSW, Australia)
2. 99% formic acid in 1 mL glass ampules (see **Note 8**) (Sigma)
3. Superficially porous C_{18} or XB-C_{18} HPLC column and matching guard column (Kinetex™, Phenomenex, Lane Cove, NSW, Australia)
4. MS grade acetonitrile (see **Note 2**) (Fisher Optima®, Thermo Fisher Scientific, Scoresby, VIC, Australia)
5. 1000 µL auto pipette
6. LC-MS platform: Low resolution mass spectral data (see **Note 29**) were measured on a Bruker Daltonics Esquire 3000 plus mass spectrometer (ESI MS) with an Apollo source connected to an Agilent 1100 HPLC system comprising degasser, binary pump, autosampler and PDA. All LC-MS data was collected using Bruker Daltonics Esquire Control v5.3 and Hystar v3.1 operating on Windows XP Professional.

3.2.4 NMR materials

1. NMR sample tubes, 509-UP-7 (see **Note 9**) (Norell Inc., Landisville, NJ, USA)
2. Deuterated methanol (CD_3OD, D 99.8%) (see **Note 10**) (Cambridge Isotope Laboratories, Andover, MA, USA)
3. 1 mL graduated glass syringe (Hamilton Company, Reno, NV, USA)
4. Lint-free tissue (Kimwipes, Kimberley-Clark, Milsons Point, NSW, Australia)
5. 6 mL glass scintillation vials (see **Note 5**)
6. 1000 µL auto pipette
7. NMR spectrometer (^1H NMR spectra in this study were collected on a Bruker Avance 600 MHz NMR spectrometer complete with TXI cryoprobe operating at 600 MHz for ^1H in CD_3OD, δ_H 3.31 ppm)

3.3 Methods

3.3.1 Coral collection and metabolism quenching

Harvesting coral and the quenching method used is dependent on the conditions at the collection site. It is recommended to harvest coral from a sheltered lagoon or reef flat where the collection can be undertaken safely in calm, shallow water that is approximately knee deep. As this method requires small amounts of liquid nitrogen to be taken to the site of collection, care should be

taken to always ensure the safety of personnel. Generally, 2-3 L of liquid nitrogen in a foam dewar is sufficient to harvest coral nubbins over a 30-minute period. The foam dewar containing the liquid nitrogen can be placed in a Nally bin and floated on the water's surface so that nubbins can be placed into the liquid nitrogen immediately after they are excised from the colony. This timing is critical as it is well documented in coral that physical interference can result in an immediate chemical response (Tapiolas *et al.*, 2010).

Collection of hard coral from deeper waters, which will often involve the use of SCUBA and a boat, requires a different sampling technique. For this method, it is recommended to use a hammer and a cold chisel with a 1-inch cutting surface to remove all or part of a colony from the substrate. After which, the whole colonies can be dissected into smaller nubbins while still submerged just below the water's surface and immediately placed into liquid nitrogen that is kept on the boat. Care should be taken to ensure that dive profiles do not involve multiple ascents and descents and as such, all colonies should be brought close to the surface in only one attempt per dive.

For metabolomics, diligent work and a reliable method will produce sound results that, in some cases, will not require the commonly used techniques of data scaling (i.e. normalisation and transformation). In this protocol, for example, normalising samples to a common extract concentration by drying and resuspending the sample in the appropriate volume of solvent practically eliminated the need to normalise at the data processing stage. As a means of verifying the robustness of this method and the importance of quickly and effectively quenching the coral metabolism, five stressed coral nubbins (nubbins collected and snap frozen after 30 minutes of agitation in a bucket of seawater) were compared with five non-stressed coral nubbins (nubbins snap frozen immediately according to this method) using ^{1}H NMR and Principal Components Analysis (PCA). As expected, the results of the PCA show that the stressed nubbins displayed considerably greater variability (Figure 3.1) Moreover, scaling of the data matrix rows and/or columns reduced the amount of variability explained by the PCA from 97.96% in 3 PCs without scaling, to 97.16% in 4 PCs with scaling. The fact that scaling reduced the ability of PCA to explain the variability associated with our data adds considerable weight to the reproducibility of this method. Methods reported below conform to the Metabolomics Standards Initiative (Sumner *et al.*, 2007).

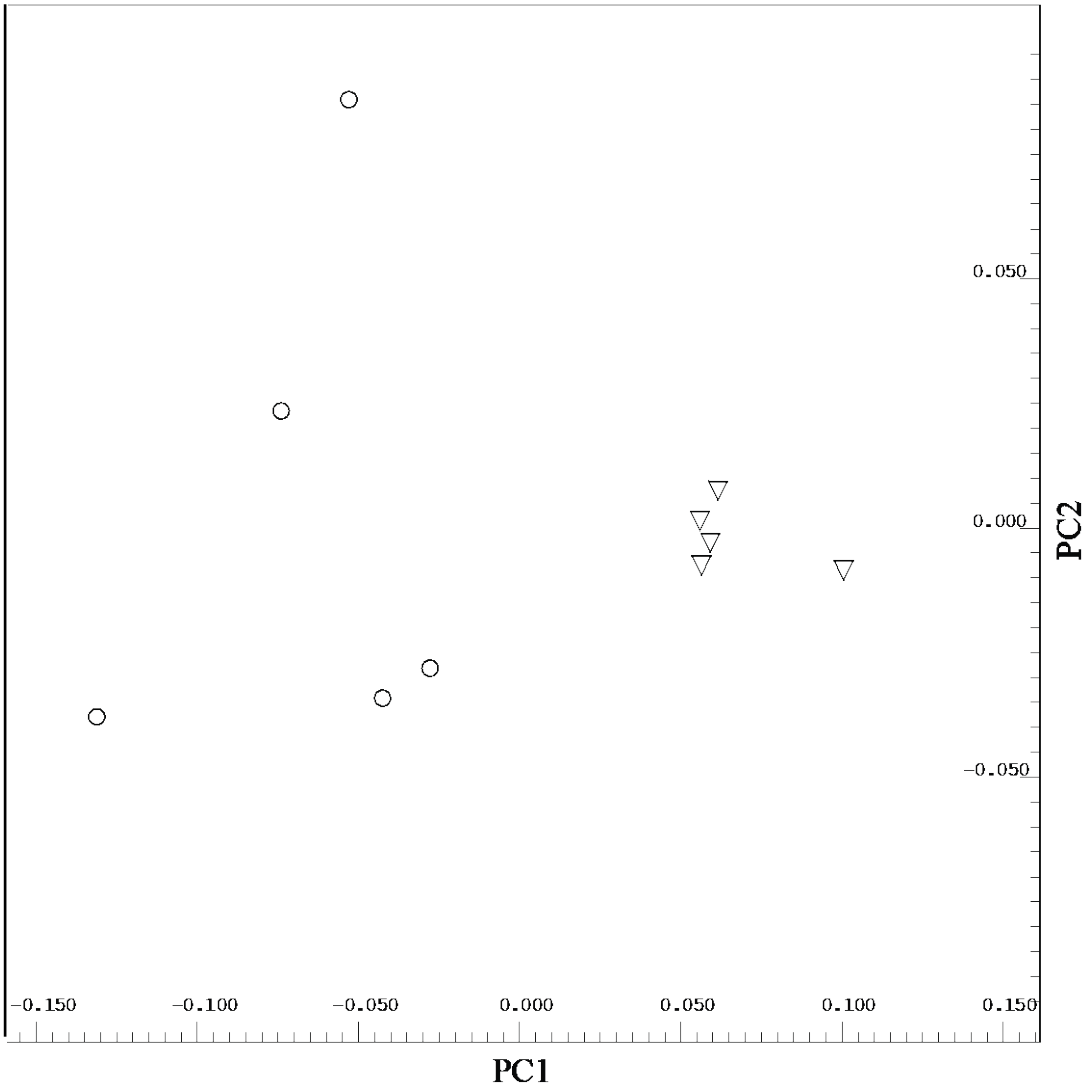


Figure 3.2 PCA scores plot of 70 % methanol extracts of stressed (*open circles*) versus non-stressed (*open triangles*) *A. aspera* analyzed by ^{1}H NMR. The much tighter grouping of the non-stressed samples highlights the reproducibility obtained by this protocol and the importance of fast and effective quenching of the metabolism. Binned NMR data (0.02 ppm widths) was mean-centred and Pareto -scaled prior to PCA.

Method

1. Fill a disposable foam dewar with liquid nitrogen, store, transport and handle with appropriate care
2. Place the foam dewar containing the liquid nitrogen, along with the bone cutters into a tethered Nally bin that is suitable for use as a stable, floating work platform
3. Use the double-hinged bone cutters to cut the coral nubbins into approximately 5 cm pieces (see **Note 11**)
4. Immediately after cutting, place each nubbin into the foam dewar containing liquid nitrogen to snap freeze. Nubbins can remain in the liquid nitrogen until ready for transfer into 50 mL centrifuge tubes

5. Upon returning to shore, remove the coral nubbins from the liquid nitrogen using steel forceps and place immediately into labelled 50 mL centrifuge tubes (see **Note 12**)
6. Place the filled tubes back in liquid nitrogen or onto dry ice to keep frozen until ready for transfer into a minus 80 °C freezer for long-term storage

3.3.2 Sample extraction, concentration and clarification

Given the desire to extract and analyse as many metabolites as is feasible whilst excluding as many lipid classes as possible for reasons mentioned previously, it is recommended to start with the 70% methanol extraction and then store the sample at minus 80 °C in the dark. Subsequently, the lipid, chlorophyll and other less polar classes of compounds can be examined later by re-extracting the biological material in a less polar solvent. As such, long-term storage conditions are an important consideration if samples are to be analysed later and it is widely accepted that storing samples at or below minus 80 °C and in the dark is adequate for halting enzyme or metabolite activity and preserving the integrity of samples (Fiehn, 2002). It should be noted, however, that storing samples above minus 25 °C is not adequate for performing the same task (Lauridsen *et al.*, 2007, Ettinger-Epstein *et al.*, 2007, Tapiolas *et al.*, 2010).

While the reproducibility of a method is influenced by the choice of extraction solvent, other factors such as sample concentration, extraction temperature, sample state before extraction (i.e. wet or dried sample), collection methods, sample storage and sample handling (i.e. the amount of time kept at room temperature for analysis) can also have a significant effect on the outcome of a particular experiment. With regards to the extraction temperature, sample storage and handling, it is largely a case of maintaining constant and similar conditions for all samples while reducing exposure to elevated temperatures that may cause sample degradation. However, where several choices are available, as is the case for sample concentration (speed vacuum, lyophilisation or nitrogen stream) and sample state prior to extraction (wet or lyophilised sample), it is appropriate to consider each option with respect to its potential to minimise sample degradation and data variability. With that in mind, ^{1}H NMR and PCA were used to examine the variability associated with three different solvent removal techniques (speed vacuum, lyophilisation and nitrogen stream at 25 °C) and the extraction of either lyophilised or wet coral nubbins. Of the three sample concentration techniques, drying under a stream of nitrogen gas at 25 °C introduced the greatest variability, while lyophilisation and speed vacuum introduced similar levels of variability (see Appendix Figure 2). As such, drying under a nitrogen stream at 25 °C was not considered an appropriate sample concentration technique. Furthermore, the analysis of extractions performed on wet and lyophilised samples showed that the extraction of wet sample was less reproducible,

most likely due to the greater differences in water and salt composition between wet and dried samples. Consequently, extractions were performed on lyophilised coral and that a combination of lyophilisation and speed vacuum was used for sample concentration.

Method

1. Fill two appropriately sized disposable foam dewars with dry ice or liquid nitrogen. Place the bone cutters and labelled 20 mL scintillation vials in the first container to cool and place the 50 mL centrifuge tubes containing the coral nubbins collected in step 3.1.4 into the second dewar
2. Using the chilled bone cutters cut each coral nubbin into several 1 cm^3 pieces so that each nubbin fits into the 20 mL glass scintillation vials (see **Note 13**). Remove any ice that may have dislodged from the nubbin when cutting (see **Note 14**). Return vials to the dry ice to keep frozen until lyophilisation
3. Undo each lid of the cold vials containing the frozen coral nubbins a quarter of a turn to ensure the atmosphere within each vial can escape when subjected to the vacuum associated with the lyophilisation. Vials should be placed upright in a suitably sized rack or tray and placed into a freeze drier for 24 hours or until completely dry (see **Note 15**)
4. After lyophilisation, fill each scintillation vial containing the coral nubbin pieces with cold (0 - 4 °C) 70% methanol to approximately 75% of its volume, ensuring that the coral nubbin is completely submersed in the extraction solvent
5. Sonicate each vial for 5 mins in a sonication bath chilled to 0 - 4 °C (see **Note 16**)
6. Decant the extract into 15 mL centrifuge tubes (see **Note 4**) and centrifuge at 5800 rcf for 5 mins to settle any particulates still present in the extract solvent
7. Using a glass Pasteur pipette remove the supernatant and transfer into clean, pre-weighed, 20 mL glass scintillation vials (see **Note 17**)
8. Remove the lids of each vial and place the vials containing the 70% methanol extracts into a speed vacuum and dry at ambient temperature for ~ 1 hour (see **Note 18**)
9. Remove the samples from the speed vacuum and freeze the remaining sample in liquid nitrogen or by placing in a minus 80 °C freezer for 1 hour
10. Undo the lids of each vial one quarter of turn and place the frozen samples into a freeze drier and lyophilise to complete dryness
11. Remove the samples from the freezer drier and tighten caps immediately to avoid any absorption of water
12. Weigh each vial and calculate the extract weight by subtracting the empty weight of each vial from the final weight (see **Note 19**)

13. Resuspend each of the dried extracts in an appropriate amount of 70% methanol, ensuring that each extract is made up to a uniform concentration (see **Note 20**)
14. Store samples at minus 80 °C until ready for analysis

3.3.3 LC-MS Analysis

It is important to consider the optimisation of the analytical platforms, which must be done on a case-by-case basis. While it is not pertinent to discuss in detail the optimisation of our own LC-MS instrument and its associated peripheries, the relevant details are provided. In particular, it is worth mentioning some of the details related to our choice of column and mobile phase, along with the parameters and conditions used for our own LC-MS analysis. With regards to the chromatography, the ability of three superficially porous columns to separate the 70% methanol extract (20 mg/mL) of symbiotic hard coral using two different acidified organic mobile phases, 0.1% formic acid in methanol and 0.1% formic acid in acetonitrile, was tested. Columns were rated based on peak separation, resolution and the number of features within each chromatogram. Of the three different columns and two different mobile phases, the XB-C_{18} column and the 0.1% formic acid in acetonitrile mobile phase gave reproducible chromatograms with the best resolution, peak separation and the greatest number of features (see Figure 3.3 and Appendix Figure 1).

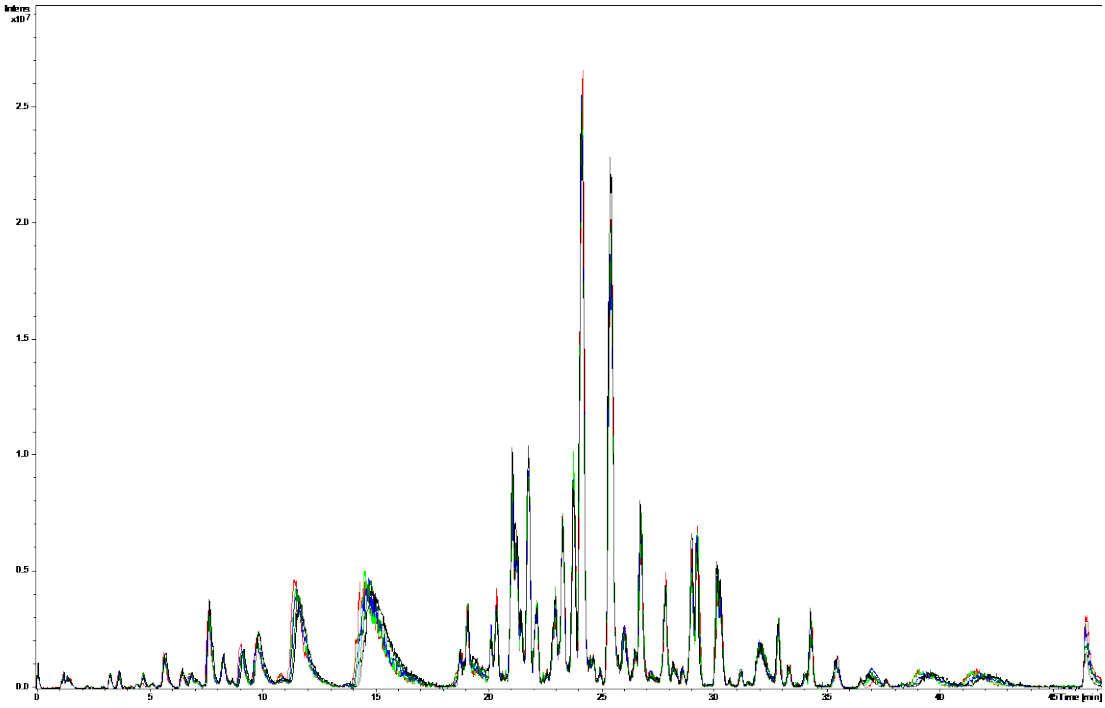


Figure 3.3 Overlaid LC-MS ion trap base peak chromatograms of five replicate 70 % methanol extracts of *A. aspera*. An XB-C_{18} column and a 0.1% formic acid in aqueous acetonitrile mobile phase gradient was used to obtain the data shown

Method

1. Remove the samples prepared according to section 3.2 from the minus 80 °C storage and store at 4 °C for 1 hour to defrost
2. Using a 1000 µL auto pipette and clean pipette tips, place 1 mL of each sample into new 1.5 mL centrifuge tubes and centrifuge each sample at 4 °C for 5 mins at 21,900 rcf to ensure that no particulates remain in the sample (see **Note 21**)
3. Transfer 100 µL of each sample into a new reduced volume HPLC sample vial or one containing a reduced volume insert (see **Note 22**). Return the remaining sample back to minus 80 °C storage
4. Store HPLC vials in the dark at minus 80 °C until ready for analysis
5. Perform the LC-MS analysis. Conditions described in steps 6 and 7 can be used as an initial starting point or guide (see **Note 23**)
6. Chromatography conditions: An XB-C_{18} superficially porous, (Phenomenex, Kinetex 3 × 100 mm, 2.6 µm) column was mated with a 0.5 µm stainless steel filter and guard cartridge of the same stationary phase; 2 µL injection volume; mobile phase A (0.1% formic acid in water); mobile phase B (0.1% formic acid in acetonitrile); gradient elution from 50% A to 100% B at a flow rate of 350 µL/min for 35 mins. After 35 mins 100% B was maintained for 7.1 mins at the same flow rate. At 42.1 mins the mixture was changed to its initial setting (50% A, 50%B) and the column equilibrated until 47.1 mins
7. MS Acquisition parameters: ESI ion source; positive ion polarity; scan range from 50 to 1000 m/z; capillary exit of 120 V; accumulation time of ~ 8000 µs; Drying temperature of 350 °C; Nebuliser pressure of 32 psi; Drying gas flow of 8 L/min
8. Agilent ES tuning mix ACN solution (G2421A) was used as the calibration sample for the performance test. This solution contains several calibration ions at mass values from 117 Da to 2,722 Da.
9. MS processing parameters: Base peak chromatograms were calculated with background removal. Raw data was mean-centred prior to PCA.

3.3.4 ^{1}H NMR analysis

It is not within the scope of this protocol to provide in detail the methods used for the ^{1}H NMR analysis of symbiotic hard coral. As such, using the well-described protocol written by Mark Viant as a solid foundation for the development of a high-quality metabolomics NMR analysis is recommended (Viant, 2008). Although our NMR analysis is not described in detail, there were aspects worth mentioning. For example, and of particular importance, is that ^{1}H NMR was proven to be effective at identifying differences between coral that were either stressed or healthy

at the time of snap freezing (Figure 3.1) (see **Note 24**). However, upon examination of the ^{1}H NMR spectra there were some noticeable, yet minor, shifts in signal position, which were attributed to the effects of high salt concentration and differences in the pH of samples. While most NMR metabolomics experiments employ the use of pH buffers to avoid such signal shifts, in this case, the shifts were small and within the confines of the binned data; hence the PCA was still very effective. As such, buffering of the sample pH using commonly employed buffers, such as sodium phosphate, were not included as part of this protocol (see **Note 25**) nor was a chemical shift standard (see **Note 26**).

Method

1. Remove the samples prepared according to section 3.2 from the minus 80 °C storage and store at room temperature (20 – 25 °C) for 30 minutes to thaw
2. Using a 1000 µL auto pipette and a clean tip for each sample, place 800 µL of each 70% methanol extract into new, individually labelled, 6 mL glass scintillation vials
3. Remove the lids of each vial and place the vials containing the extracts into a speed vacuum and dry at ambient temperature for ~ 10 mins (see **Note 18**)
4. Remove the samples from the speed vacuum and freeze the remaining sample in liquid nitrogen or by placing in a minus 80 °C freezer for 1 hour
5. Undo the lids of each vial one quarter of turn and place the frozen samples into a freeze drier and lyophilise to complete dryness
6. Remove the samples from the freezer drier and tighten caps immediately to avoid any absorption of water
7. Using a 1 mL graduated glass syringe, reconstitute each sample to its original 20 mg/mL concentration by adding 800 µL of deuterated methanol, (CD_3OD, D 99.8%)
8. Place each vial containing the extracts in CD_3OD into a sonic bath for 5 mins to completely dissolve the extracts
9. Using a 1000 µL auto pipette and a clean tip for each sample, transfer the 800 µL of each sample into clean, individual and appropriately labelled 5 mm NMR tubes
10. Cap each tube, seal with parafilm and store upright at minus 80 °C until ready for analysis
11. Prior to analysis, thermally equilibrate samples by storing the tubes at room temperature (~ 25 °C) for 30 minutes (see **Note 27**)
12. Perform the NMR analyses using the acquisition and processing parameters in steps 13 and 14 as a guide
13. Acquisition parameters: Bruker pulse sequence zgesgp (1D excitation sculpting using 180 water-selective pulses) comprising [relaxation delay–180°–acquire]; 8.3 kHz spectral

width; 3 s relaxation delay; typically 4 dummy scans followed by 32 transients (ns) are collected into 32K data points; receiver gain was constant for all samples; temperature set at 298K

14. Processing parameters: Zero-filling was not applied; exponential line broadening of 0.3 Hz; Fourier transformation; manual phase correction (zero- and first-order corrections); baseline correction was not performed (see **Note 28**); calibrate the spectrum by setting CD_3OD peak to 3.31 ppm. Raw data was mean-centred prior to PCA.

3.4 Notes

1. It is critical to ensure that all reusable glassware is free from contaminants that may affect the analysis. LC-MS using superficially porous columns is particularly susceptible to contaminants such as polyethylene glycol, slip agents, biocides and plasticisers. They can be concentrated on reverse phase LC columns and eluted during a gradient. These contaminants are commonly found in detergents used for washing glassware and any plastic items they come in contact with (i.e. drying racks and gloves). They are often identifiable by broad peaks in the chromatogram with repeating mass units in the corresponding mass spectrum. As such, care should be taken to avoid any unnecessary exposure of glassware to plastic items. Gloves should be washed in ultrapure water to remove any slip agents before handling and all glassware of unknown purity should be acid-washed in a solution of 50% nitric acid (HNO_3) for 30 mins then rinsed at least 3 times with MilliQ water. Glass Schott bottles, beakers and measuring cylinders can be filled to the brim with 50% HNO_3 and left for 30 mins and then triple-rinsed with MilliQ water. It is not necessary to wash the outside of any glassware.
2. The use of high-quality MS grade solvents is highly recommended to eliminate the possibility of inferior solvents introducing spurious features to an analysis. This method recommends the use of Fisher Optima® LC-MS solvents (Thermo Fisher Scientific, Scoresby, VIC, Australia).
3. While the 70% methanol extraction solvent effectively halts enzyme and metabolic activity, it is considered good practice to ensure that the temperature is kept as low as feasibly possible throughout the extraction and analysis to avoid unwanted changes in chemistry. Also, as temperature is not as critical for extraction efficiency as solvent choice (Beltran *et al.*, 2012), a temperature range of 0 °C to 4 °C is recommended for this protocol.

4. All plastic centrifuge tubes should be made of virgin polypropylene and free from slip-agents, biocides and plasticisers.
5. Glass scintillation vials should be of known purity and utilise an inert, contaminant and extractable free Teflon or foil-lined gasket. Vials having a plastic lid without such a gasket should be avoided.
6. This protocol recommends labels be made of plastic film and be solvent resistant. Labels made of paper or absorbent material can affect the calculation of extract yields as they contribute to a loss in mass during the solvent removal steps. Alternatively, permanent markers can be used to label vials, however, care should be taken to ensure that spilt solvents do not remove the markings.
7. During the development of this protocol, the silicone/Teflon septa of standard auto-sampler vials was found to contribute significant levels of plasticiser contamination to each analysis. As such, sample vials, lids and septa used in LC-MS auto-samplers should be certified as MS grade. The use of certified Verex™ vials, lids and septa (Phenomenex, Lane Cove, NSW, Australia) is highly recommended. For cases where reduced volume glass inserts are used, then it is acceptable to use standard vials with MS certified lids and septa as the glass inserts are of sufficient purity.
8. The cost of formic acid supplied in 1 mL glass ampules can be very expensive and as such, it is not considered essential for this protocol, however, it does buy peace of mind for the analyst and is preferable. If using formic acid supplied in larger volumes, it is highly recommended to ensure the ongoing integrity of the reagent by exerting extra care when storing and taking aliquots from the storage container. As such, a number of 1 mL aliquots should be prepared in individual clean glass vials (2 mL MS certified HPLC vials and lids are ideal) and stored at 4 °C for later use. In this way, it is a straightforward process of adding the pre-measured 1 mL of formic acid to 1 L of mobile phase to give a final formic acid concentration of 0.1 %.
9. NMR sample tubes should be compatible with the field frequency of the user's NMR and the volume of the sample must cover the entire active region of the shim coils (700 μL in a 5 mm NMR tube). Dirty tubes can be cleaned by filling them with 50% HNO_3 and soaking for 30 minutes. The acid solution can then be removed, and the tubes rinsed twice with ultra-pure water and again with acetone. Tubes can be dried by either air-drying or placing them under vacuum and should never be dried at high temperature. Always follow the manufacturer's instructions.

10. Volatile NMR solvents should be stored in a sealed container with a suitable desiccant to minimise the possibility of solvents absorbing water from the atmosphere.
11. When cutting the coral nubbins, it is preferable to cut the nubbin and allow it to fall into a waiting gloved hand or container. The nubbin can then be quickly transferred directly into the liquid nitrogen and snap-frozen with minimal handling. Care should be taken to avoid handling each nubbin more than required as this will damage the coral tissue, invoking the rapid production of coral mucous resulting in a change of chemistry.
12. It is acceptable to place replicate nubbins in a single 50 mL centrifuge tube after the nubbins have been snap frozen.
13. Once cut into smaller pieces, a nubbin should occupy approximately 30% to 50% of the volume of each scintillation vial. The nubbin pieces should be completely submersed once the vial is filled to 75% of its volume with the extraction solvent.
14. Most of the residual seawater that was frozen with the nubbin during the collection process will separate from the nubbins when cutting them into smaller pieces. It is important that this ice be discarded to reduce the amount of undesirable salts within the sample.
15. In any metabolomics study it is important to recognise the effects that certain techniques will have upon the sample. In the case of lyophilisation, the reduced atmospheric pressure will remove many of the volatile compounds. If these volatile compounds are of interest to the analyst, then drying under such conditions may need to be avoided and a more suitable technique examined.
16. At this point extracts should be orange in colour if effective lysis of *Symbiodinium* cells has occurred.
17. New glass scintillation vials are weighed at this point in order to acquire the empty weight of each vial, which will be used to calculate the weight of the crude, dry extract.
18. The aim of this step is to remove ~ 90% to 95% of the extraction solvent. Therefore, it is important to remove the vials from the speed vacuum concentrator before all the liquid has been evaporated so that the extracts remain cooled by the evaporation process, which commonly takes from 10 – 60 minutes depending on the volume of solvent. The speed vacuum concentrator should not be used with sample heating. Drying alcohol and water mixtures under vacuum can be time consuming without sample heating due to the non-ideal interaction of water and methanol molecules at low concentrations (Wakisaka *et al.*, 1998). As such, this protocol utilises lyophilisation as a second drying step. In this way,

the water-methanol interaction is overcome, the sample does not need to be heated and thus the sample integrity is maintained.

19. By recording the weight of each extract, samples can be prepared at the same concentration for analysis. This has the effect of normalising each sample.
20. For this method, a final concentration of 20 mg/mL is appropriate and will provide sufficient sample concentration for both ^{1}H NMR and LC-MS analyses. However, factors such as LC column loading and NMR sensitivity, need to be accounted for when deciding on the appropriate concentration.
21. Syringe filtering was not employed in this method due to the contamination that the rubber syringe plungers introduce into the sample. Using glass syringes as an alternative was ruled out because cleaning them between samples was deemed too time consuming, in addition to the negative impact that additional sample handling could have on the sample. Alternatively, centrifugal solvent resistant filters could be used; however, they were not tested as part of this protocol.
22. In this instance, a 100 µL aliquot of the 20 mg/mL extract was transferred into each HPLC vial, hence the use of reduced volume inserts. This small amount is sufficient for multiple 2 µL injections and avoids the risk of compromising the entire sample.
23. While the conditions described in steps 6 and 7 will provide a good starting point, LC-MS conditions may still need to be optimised depending on the type of instrument and its associated peripheries.
24. In addition to the typical 1D ^{1}H NMR experiment, the effectiveness of the 2D JRES experiment (Aue *et al.*, 1976) was examined using a double spin-echo water suppression method (Thrippleton *et al.*, 2005). The primary advantage of the JRES experiment for metabolomics lies in its ability to reduce the complexity of the typical 1D ^{1}H NMR experiment by shifting the J-coupling into a second dimension. However, upon comparing the two experiments, it was found that there was not a sufficient gain in spectral decongestion to warrant using the more time consuming JRES experiment.
25. The decision not to use a buffer was based on a number of reasons. Firstly, the observed signal shifts were small and within the tolerance of the 0.04 ppm bin widths used in the PCA. For the small number of instances where shifts in NMR peaks did occur across the borders of a bucket, the loadings of the PCA reflected these shifts whereby multiple and consecutive bins contributed to the explanation of variability in the data. The final reason for avoiding the use of a pH buffer was salt concentration. It was reasoned that adding a buffer to a marine sample, which already had a naturally high salt content, would

contribute to the ionic strength of the sample, exacerbating the detrimental effects of high salt concentration (Weljie *et al.*, 2006). That being the case, it is critical that the analyst thoroughly examines and compares all NMR spectra for shifts in peak position to ascertain the need for pH buffering. When the use of a buffer is required, we recommend using a buffer with low ion mobility to preserve the sensitivity of the NMR probe, particularly when analysing samples high in salt concentration using cryogenically cooled probes (Kelly *et al.*, 2002).

26. Chemical shift standards were not used in this protocol as all spectra were calibrated to the residual methanol signal at 3.31 ppm, thus avoiding the need to manipulate the sample. While chemical shift standards often perform the secondary function of an internal standard used for quantification, we recommend the use of the ERETIC experiment for this purpose (Akoka *et al.*, 1999). The advantage of the ERETIC experiment lies in its ability to add an electronic signal, at a chemical shift of one's choice, into previously acquired spectra. The electronic signal is calibrated and integrated to a known concentration of any analyte of choice and inserted into the previously acquired spectra where it can be used for the quantification of all other signals.
27. Warming the sample to 25 °C is necessary for cases where the NMR experiment is performed at 25 °C (298 K), thereby reducing the time of thermal equilibration within the instrument.
28. Manual baseline correction was not performed. In our case, all baselines were observed to be flat within the signal region, yet distortions did occur at the extremities (visible as upturned edges of the baseline). These distortions could not be removed using baseline correction without being detrimental to nearby regions. To rectify this, our sweep width was increased to 14 ppm (8.3 kHz) and the spectrum offset changed to 6 ppm. In this way, the distortion at the extremities of the baseline was shifted to regions having no signals (approximately -1 ppm and 13 ppm) and were then removed from the subsequent PCA by only integrating the signal region from 0 ppm to 10 ppm.
29. A low-resolution instrument, such as that used to assess the different extractions here, while suited to exploratory investigations of metabolic profiles during method development, is not suited for most metabolomics workflows.

3.5 Discussion and concluding remarks

Despite the rise in popularity of coral metabolomics over the last decade, there have only been two attempts to provide a thoroughly investigated and validated sample preparation procedure for coral metabolomics (Andersson *et al.*, 2019, Gordon *et al.*, 2013). There are a variety of reasons for this but mostly, researchers are tailoring their sample preparation to the specific goals of each experiment, where factors such as the type of analytical platform being employed, the metabolite classes being targeted, or the specific hypotheses formulated (e.g. targeting a specific partner in the holobiont) drive the choice of sample preparation methods. Nevertheless, certain aspects of sample preparation and extraction have been universally adopted in coral metabolomics. For example, quenching is almost always performed using liquid nitrogen followed by storage at -80 ℃ and metabolite extraction is commonly achieved using 70% aqueous methanol (Andersson *et al.*, 2019).

The protocol reported in this chapter (published in 2013) has had a measurable impact on coral metabolomics. For example, extraction using 70% aqueous methanol has been successfully employed in a number of studies (Sogin *et al.*, 2014, Quinn *et al.*, 2016, Putnam *et al.*, 2016, Sogin *et al.*, 2017, Hartmann *et al.*, 2017), revealing measurable changes in the production and utilisation of key endogenous molecules linked to coral functional status. The protocol is suitable for a variety of analytical approaches, including fragmentation mass spectrometry and negative ionisation with little to no modification.

The performance of the 70% aqueous methanol extraction protocol was recently evaluated using ^{1}H-NMR-based metabolomics (Andersson *et al.*, 2019). Processing steps related to metabolome (cryo)preservation, metabolite extraction and subsampling were evaluated for the variation induced, and for the intensity and number of metabolite features detected. Comparative analysis confirmed that while a biphasic extraction solvent of chloroform:methanol:water (2:2:1.8), was superior in terms of reproducibility compared with either methyl *tert*-butyl ether (developed for lipid extraction) or 70% aqueous methanol, extraction with 70% aqueous methanol was superior for feature detection, validating the original intent of the protocol developed here (Gordon *et al.*, 2013).

The 70% aqueous methanol extraction protocol provides a solid foundation for coral metabolomics and while its adoption has already been realised, there is still a fundamental need for a standardised protocol if a metabolomics-based biomonitoring program is to be realised. Therefore, a more rigorous experimental and statistical evaluation of its application was

undertaken and is presented in the following chapters. The suitability of the 70% aqueous methanol extraction protocol for LC-MS and ^{1}H NMR profiling of the metabolome of corals exposed to ocean warming and acidification is assessed in Chapter 4. Its effectiveness in extracting metabolites to predict the maximum quantum yield of photosystem II in corals exposed to ocean warming is assessed in Chapter 5.

In the process of developing a method to classify temporal metabolome changes in wild coral species, it became evident that a more user-friendly and less hazardous method to quench coral samples in the field was needed. Therefore, a modification of this protocol, forgoing cryopreservation with liquid nitrogen for full immersion in 100% methanol, is investigated in Chapter 6. It is anticipated that these investigations will provide a solid foundation towards developing a standardised biomonitoring protocol with the potential to facilitate large-scale and non-expert sampling and data sharing. Ultimately, the findings will improve the prediction of coral functional status to guide management of reef ecosystems facing functional collapse due to cumulative anthropogenic stresses.

Chapter 4

Classification of Discrete *Acropora aspera* Phenotypes Associated with a Simulated Bleaching Event and Elevated Carbon Dioxide

This chapter is inserted as in preparation for submission: Benjamin R. Gordon, Daisie Ogawa, William Leggat, Ute Roessner, Berin Boughton and Cherie A. Motti

Benjamin Gordon wrote this chapter, designed and conducted the research, and analysed the data. Co-authors provided intellectual, experimental and editorial support.

4.1 Introduction

The success of coral reefs is negatively impacted from a variety of anthropogenic pressures such as, overfishing, climate change, nutrient runoff and pollution (Hughes *et al.*, 2017a, Hughes *et al.*, 2017b, Hoegh-Guldberg, 2011). Anthropogenic activity on both local and global scales continues to change the configuration of coral reefs at accelerating rates and in previously unseen ways, despite our current knowledge of these impacts and our efforts to manage them (Hughes *et al.*, 2017a, Hughes *et al.*, 2010, Pandolfi *et al.*, 2011, Hoegh-Guldberg *et al.*, 2017). To conserve and manage the coral reef ecosystem and maintain the service it provides now, and for future generations, it is imperative we have a better understanding of the biological processes that underpins its health and productivity.

Stable seawater temperatures and aragonite concentrations are crucial factors affecting many of the biological processes that maintain healthy coral reefs. In scleractinian corals, the major reef framework builders, prolonged elevated temperatures 1-3 ℃ above the baseline summer maximum are sufficient to cause coral bleaching (Hughes *et al.*, 2017a, Hoegh-Guldberg, 2011), while decreases in carbonate ion concentrations of ~200 μmol kg^{-1} seawater will reduce the rate of reef calcification processes to a state favouring erosion (Hoegh-Guldberg *et al.*, 2007).

Ocean warming and acidification also impacts the broader reef community; for example, coral fitness is reduced, and disease becomes more prevalent (Maynard *et al.*, 2015), resulting in higher rates of coral mortality over larger regions. Algal abundance increases and can degrade or reduce substrate availability, impacting coral recruitment and the ability of coral reefs to recover from other major disturbances such as cyclones and Crown of Thorns Starfish (CoTS) outbreaks (Ledlie *et al.*, 2007, Pandolfi *et al.*, 2011). In turn, a decline in abundance of animals that rely on coral for food or habitat may occur, resulting in a shift in community structure (Stella *et al.*, 2011).

To date, coral metabolic phenotypes associated with ocean warming (Hillyer *et al.*, 2016, Hillyer *et al.*, 2017b, Hillyer *et al.*, 2017a, Hillyer *et al.*, 2018, Petrou *et al.*, 2018), ocean acidification (Putnam *et al.*, 2016) and combined ocean warming and acidification (Sogin *et al.*, 2016, Farag *et al.*, 2018) have been identified; however, our understanding of the coral response to synergistic ocean warming and acidification is underpinned by research assuming a business-as-usual emissions scenario of 600 – 1000 ppm of atmospheric carbon dioxide (CO_2) (Hughes *et al.*, 2017a). These atmospheric CO_2 concentrations are now less likely as countries do more to limit the production of greenhouse gases (as agreed in the Paris Agreement, 2015), and the world sees a shift to a lower emissions scenario. Unfortunately, we lack an understanding of how reefs will change

under more realistic atmospheric CO_2 concentrations of 450 – 550 ppm (Hughes *et al.*, 2017a) that is projected by the Intergovernmental Panel on Climate Change (IPCC) Representative Concentration Pathway 4.5 (IPCC, RCP 4.5, 2014).

Our knowledge of the cumulative and chronic impacts of ocean warming up to 1 ℃ and changes in ocean pH of up to 0.1 units is poorly understood; indeed, no coral metabolic phenotypes have been identified for atmospheric CO_2 concentrations of 450 – 550 ppm projected under RCP 4.5 (Hughes *et al.*, 2017a). This gap in the coral research is due, in part, to the difficulty associated with manipulating and recording temperature and CO_2 variables at such fine scales, and to the potential masking of acidification-induced symptoms by more influential factors such as temperature and irradiance.

Metabolomics biomonitoring is an emerging approach to ecosystem management that incorporates metabolic profiling and modern chemometrics to monitor the biochemical response of an organism to changes in its environment. Given the fast response of the metabolome to external stressors (Caldana *et al.*, 2011) and the superior sensitivity of the analytical platforms used to measure it, metabolomics and chemometric techniques, offer an ideal solution for monitoring an individual coral's response to small changes in seawater temperature and pH (associated with atmospheric CO_2 concentrations) expected this century.

To be an effective biomonitoring tool metabolic profiling must not only provide detailed and informative data, but also have sufficient resolving power to predict the functional state of corals subjected to realistic changes in temperature and CO_2. In particular, the technique must be sufficiently sensitive to describe coral phenotypes associated with atmospheric CO_2 concentrations less than 550 ppm, even when subjected to the more influential factors. To this end, the potential of two commonly used analytical platforms – liquid chromatography coupled with mass spectrometry (LC-MS) and proton nuclear magnetic resonance spectroscopy (^{1}H-NMR) – as well as two popular machine learning algorithms – partial least squares discriminant analysis (PLS-DA) and Random Forests (RF) (Mendez *et al.*, 2019) – to classify *Acropora aspera* phenotypes (supplied as discrete class labels) associated with a simulated bleaching event and pCO_2 levels consistent with the projections of RCP 4.5 were assessed.

4.2 Methods

4.2.1 Experimental conditions

This experiment was designed to compare the responses of *A. aspera* fragments under current day conditions with those subjected to either a simulated bleaching event, elevated pCO_2 or combined bleaching and elevated pCO_2 scenarios. Coral nubbins (n = 96) were collected (Great Barrier Marine Park Authority permit G09/32575.1 and G08/26873.1) from three colonies of *A. aspera* (tan morph, approximately 32 nubbins from each colony) from the Heron Island reef flat, Queensland, Australia. The nubbins, supported in racks, were acclimated in aquaria at the Heron Island research station with a continuous flow of fresh, sand filtered seawater pumped from the Heron Island reef flat. The incised base of the coral nubbins were monitored for tissue regrowth for 14 days to ensure they were in good health.

Acclimated nubbins were randomly assigned to one of four treatments (Table 4.1): **control** (ambient temperature and ambient CO_2), **eT** (elevated temperature at ambient CO_2), **eCO_2** (elevated CO_2 at ambient temperature) and **eCO_2eT** (elevated CO_2 and elevated temperature). Each semi-closed treatment system consisted of a 250 L flow-through sump tank supplying two 65 L replicate aquaria. In heated systems, temperature was controlled using 300W aquarium heaters (Eheim Jager, Deisizou, Germany) and monitored with red spirit thermometers and HOBO temperature loggers (OneTemp, Brisbane, Australia). In acidified systems, a CO_2/air mixing system based on the design of Munday *et al.* (2009) was used to control CO_2 enrichment. Aquaria pH was monitored daily with a YSI 600QS Sonde (YSI, OH, USA).

Table 4.1 Summary of the physical and chemical conditions (mean ± SD) of the semi-closed aquaria during the experiment. Value ranges reflect the diurnal fluctuations of the incoming reef water. Changes in mean temperature and CO_2 are relative to controlled conditions.

Treatment	T (℃)	Peak Mean ΔT	pH	Mean ΔpH	pCO_2 (µatm)	Mean ΔpCO_2
Control	25.4-30.8	-	8.1-8.4	-	142-435	-
eCO_2	26.7-35.2	4.24 ± 0.91	8.1-8.4	-0.04 ± 0.05	141-350	-11 ± 26
eT	26.7-35.2	4.24 ± 0.91	8.1-8.4	-0.04 ± 0.05	141-350	-11 ± 26
eCO_2eT	25.9-35.3	3.77 ± 1.54	7.9-8.4	0.08 ± 0.05	204-827	59 ± 28

For heated treatments (eT and eCO_2eT), aquaria were maintained at 34 °C daily maximum (3–5 °C above the ambient midday maxima) for the first five days, then decreased at a rate of 1 °C per day to a mild heat stress of approximately 2 °C above ambient for the remainder of the experiment. Acidified treatments (eCO_2 and eCO_2eT) were maintained at a pH that was consistent with IPCC projections under RCP4.5 (approximately 0.1 pH units lower than present levels or ~450 ppm $[CO_2]_{atm}$) for the duration of the experiment (IPCC, RCP 4.5, 2014). Due to

the nature of the semi-closed system, the ambient temperature and pH was monitored in every experimental aquarium and baseline fluctuations established (25.4–30.8 °C and 8.1–8.4 pH, respectively). Aquaria were covered with 70% shade cloth (maximum 250 μmol photons m^{-2} s^{-1}) to provide light levels similar to those experienced by the corals in their natural reefal environment. Water samples were collected and preserved with a 0.2% saturated $HgCl_2$ solution then stored at 4 °C for subsequent analysis of alkalinity and pCO_2 as described by Ogawa *et al.* (Ogawa *et al.*, 2013).

4.2.2 Sample collection

On days 1, 4, 6 and 14, three replicate coral fragments were taken from each of the replicate aquaria (six fragments per treatment), snap-frozen immediately in liquid nitrogen then stored at -80 °C for metabolomic analyses. On days 5, 10, and 14, three nubbins from each of the replicate aquaria (6 fragments per treatment) were snap-frozen immediately in liquid nitrogen and stored at -80 °C for pigment quantification and symbiont cell densities.

4.2.3 Quantification of thermal and pH stress

The maximum quantum yield of photosystem II (F_v/F_m) in *A. aspera* was measured using imaging Pulse Amplitude Modulation (PAM) fluorometry (MAXI imaging PAM, Waltz, Effeltrich, Germany) to monitor symbiont photosystem health during the experiment. PAM measurements of the same five replicate nubbins from each aquarium were conducted every second day, 30 min after sunset. The minimum fluorescence was measured with a weak pulse of light, followed by a saturating pulse of 2,700 μmol quanta m^{-2} s^{-1} of photosynthetically active radiation (PAR) for 800 ms to determine the maximal fluorescence and dark-adapted yield. Nubbins were returned to aquaria immediately following measurement.

Coral fragments collected for symbiont cell density analysis had their tissue removed with a dental irrigator filled with 0.22 μm filtered seawater. Tissue samples were homogenised using an immersion blender for 20 s, the blastate volume recorded, and then centrifuged at 3220 *g* for 5 min to pelletize the algal cells. Symbiont cell densities were quantified using replicate (n = 6–10) haemocytometer counts per sample and normalised to the fragment surface area using the 3D modelling method of Jones *et al.* (Jones *et al.*, 2008).

4.2.4 LC-MS and NMR sample preparation

Frozen coral nubbins were placed into 20 mL glass scintillation vials and lyophilised for 24 h after which they were extracted in 70 % aqueous methanol (MeOH) as per Chapter 3 and Gordon et al (Gordon *et al.*, 2013). Coral extracts were centrifuged at 5,800 *g* to remove any

undissolved cellular debris and the supernatant transferred to clean scintillation vials. The extract was concentrated under centrifugal vacuum followed by lyophilisation for 24 h. Dry coral extracts were normalised to total dry extract weight by weighing each dried extract before reconstituting to an extract concentration of 20 mg mL^{-1} in 70 % MeOH for LC-MS analyses and 100% deuterated methanol (CD_3OD; Cambridge Isotope Laboratories Inc sourced from Novachem Australia) for ^{1}H-NMR analysis.

Those extracts earmarked for LC-MS were centrifuged at 21,900 *g* before transferring an aliquot of the supernatant to LC-MS grade amber vials with reduced-volume glass inserts (Phenomenex, Lane Cove, Australia) and then stored at -80 °C for analysis. Pooled biological quality controls (PBQCs) were used to monitor analytical variation. PBQCs were prepared for LC-MS analysis by combining 2 µL of each sample into a single PBQC sample that was analysed after every 10 injections.

A 700 µL aliquot of each extract prepared in CD_3OD was transferred into a 5 mm NMR tube and samples analysed immediately.

4.2.5 LC-MS data acquisition

Reverse Phase chromatography was performed using an Agilent 1200 HPLC system (Santa Clara, CA, USA) consisting of a vacuum degasser, binary pump, thermostatic auto sampler and column compartment. Chromatography was conducted using the following conditions: a Zorbax Eclipse XDB-C18, 2.1 mm x 100 mm, 1.8 µm (Agilent, Australia) column; solvent (A) 0.1% formic acid in Milli-Q water and solvent (B) 0.1% formic acid in acetonitrile (ACN); a flow rate of 0.4 mL min^{-1} and; a column temperature of 40 °C. Samples were injected (10 µL) and eluted with a 10 min linear gradient from 5% (B) to 100% (B), followed by a 2 min hold at 100% (B), then returned to 5% (B) and re-equilibrated for 5 min (total time of 17 min). Samples were randomised to ensure analytical variation did not correlate with biological variation and a pooled biological quality control was run every tenth sample.

The mass spectrometer was an Agilent 6520 ESI-QTOF-MS (Santa Clara, CA, USA) with a dual spray ESI source operated in positive ion mode. The source conditions were: nebuliser pressure of 45 psi; gas temperature of 300 °C; drying gas flow of 10 L min^{-1}; capillary voltage of 4000 V and skimmer 65 V; fast polarity enabled. Measurements were performed in the extended dynamic range mode (m/z range of 100-3200; sampling rate 2 GHz); scan rate of 2.03 spectra/s collected as MS1 centroid data. The mass spectrometer was calibrated using Agilent's ES Tuning Mix ACN Solution, containing several calibration ions at mass values from 117 Da to 2,722 Da. PBQCs

were used to monitor mass deviations of some ubiquitous features and analytical variation such as retention time shifts (typically < 0.1 min) and ionisation efficiency.

4.2.6 LC-MS data pre-processing

LC-MS data was converted to mzXML format using ProteoWizard's MSConvert tool (v3.0.6585; Chambers *et al.*, 2012, Kessner *et al.*, 2008) with an absolute intensity threshold of 1000. LC-MS chromatograms were integrated and aligned in the R environment (v3.6.1; http://www.r-project.org/) using the XCMS package (v1.42.0; Scripps Institute for Metabolomics; Smith *et al.*, 2006). Feature detection was performed using the *centWave* method (Tautenhahn *et al.*, 2008) with the following optimised parameters: ppm=30, peakwidth=c(20,60), mzdiff=0.01, integrate=2 and, prefilter=c(3,1100). Features were matched across samples using the following optimised parameters: bw=5, mzwid=0.05 and max=100. Retention time correction utilised the *obiwarp* method and profStep=0.5. *FillPeaks* was employed using default values. Isotopes and adducts were annotated using the CAMERA package (v1.40.0) with the following optimised parameters: perfwhm=0.7, cor_eic_th=0.75, ppm=10 and polarity='positive'. All M+1, M+2, M+3 and M+4 isotopes identified by the CAMERA package were removed from the resulting peak list. Missing values were imputed using *k*-nearest neighbour averaging implemented by the impute package (v1.58.0) with the following parameters: k=10, rowmax=0.5, colmax=0.8 and maxp=3000. The final, pre-processed, dataset had a total of 1641 features.

4.2.7 ^{1}H-NMR Data Acquisition

^{1}H-NMR spectra were acquired at 600.13 MHz using a Bruker Avance NMR spectrometer fitted with a cryoprobe. Acquisition parameters: Bruker pulse sequence zgesgp (1D excitation sculpting using 180° water-selective pulses) comprising [relaxation delay—180°—acquire]; 8.3 kHz spectral width; 3 s relaxation delay; four dummy scans followed by 32487 transients (ns) were collected into 32 K data points; receiver gain (161) was constant for all samples; temperature set at 298 K. Processing parameters: Zero filling was not applied; exponential line broadening of 0.3 Hz; Fourier transformation; manual phase correction (zero- and first-order corrections); baseline correction was not performed; spectra were calibrated to the CD_3OD peak at 3.31 ppm.

4.2.8 ^{1}H-NMR data pre-processing

^{1}H-NMR data was converted to rectangular bins with a 0.02 ppm width using AMIX Version 3.7 (Bruker BioSpin GmbH, Rheinstetten, Germany) and exported as .csv files. The bin containing the residual CH_3OH/CD_3OD solvent peak (3.30 – 3.32 ppm) was removed. The final, pre-processed, dataset had a total of 499 features.

4.2.9 Principal components analysis (PCA)

Principal components analysis (PCA) was performed to visualise the internal structure of the LC-MS and ^{1}H-NMR data in relation to its variance. PCA was performed in the R environment (v3.5.0; R Core Team, 2018) by singular value decomposition of the mean-centred and scaled data using the base *prcomp* function.

4.2.10 Model training and validation

PLS-DA and RF models were trained in the R environment (version 3.5.0; R Core Team, 2018) to classify coral samples exposed to one of four treatment conditions: control, eT, eCO_2 and eCO_2eT. The temporal variable was not included as an outcome or predictor variable; however, temporal information was used to ensure balance of the training and test sets and to examine its relationship with model predictions. The model training process is outlined in Figure 4.1. The initial step randomly partitioned the data into training and test sets using the *createDataPartition* function within the *caret* package (v6.0-84). A balanced 80:20 (train:test) split was created based on the treatment and exposure time. Both the PLS-DA and RF data were mean-centred and scaled prior to fitting each model. For PLS-DA, the model was trained to the optimal number of components, or latent variables (LVs); while for RF, the model was trained to the optimal number of variables randomly sampled as candidates at each split (known as the mtry value). In both cases, the training process utilised repeated, stratified, *k*-fold cross validation (three repeats; *k*=10) for choosing the optimal model. The optimal model was evaluated based on the overall accuracy. For all models, the optimal model was the most accurate model within one standard error of the empirically best model (Kohavi, 1995, Breiman *et al.*, 1984). Test set predictions were used to evaluate the final model.

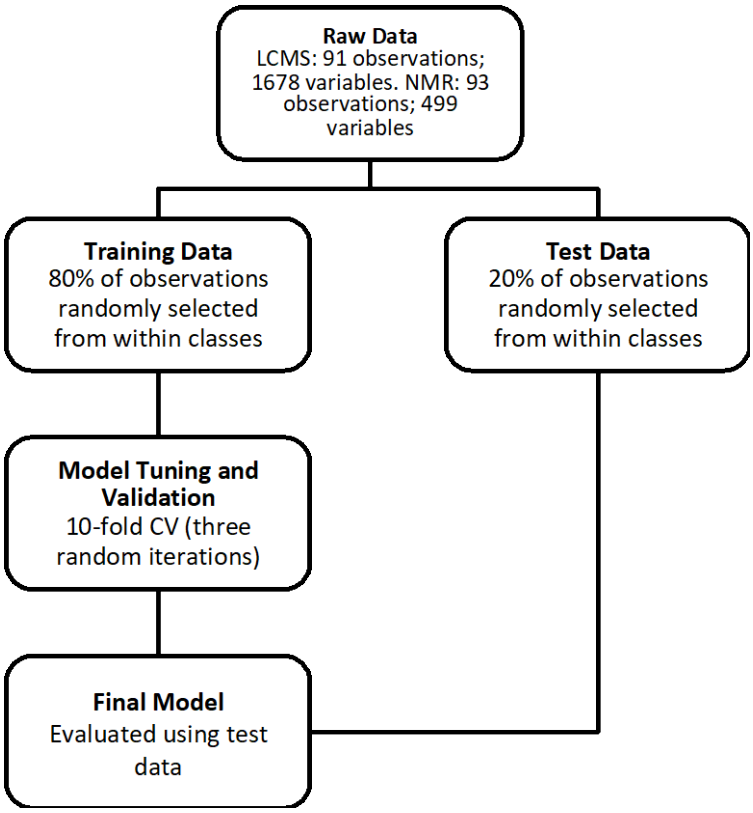


Figure 4.1 Model tuning and validation flow diagram for the analysis of LC-MS and ^{1}H-NMR data.

4.2.11 Variable Selection

Spectral features driving classification were identified for both LC-MS models. For the PLS-DA model, variable selection utilised the Variable Importance in Projection (VIP) coefficient, which is based on the weighted sums of the absolute regression coefficients. For the RF models, variable selection utilised the mean decrease in accuracy as described in the *randomForest* package (Liaw & Wiener, 2002).

An in-house thematic database of known coral and algal mass spectral features was constructed from data in the MarinLit[2] database and coral mass spectral features reported in the coral research literature (Gordon & Motti, 2020a). Important spectral features identified in this research were cross-referenced with this database using a mass error of Δ50 ppm, providing level 2 putative annotations where possible (Sumner *et al.*, 2007). Restricting our search to known coral and algal compounds which have been observed using mass spectrometry provided a greater level of confidence in the level 2 putative annotations and allowed for a wider search window as described by Van Assche et al (2015).

[2] https://pubs.rsc.org/marinlit

4.2.12 Data

The data analysis and R scripts for this experiment are available from GitHub at https://github.com/brgordon17/coralclass (Gordon, 2020a). The full list of metabolites found in the coral research literature and the MarinLit database are available as an R package at GitHub at https://github.com/brgordon17/coralmz (Gordon & Motti, 2020a). Total ion chromatograms for duplicate samples in positive and negative ion modes respectively are illustrated in Appendix Figure 3.

4.3 Results

4.3.1 Thermal and pH stress indicators

Treatment had a significant effect on *Symbiodiniaceae* cell density at day 5 (one-way ANOVA, $F_{3, 20} = 7.98$, $P = 0.001$), day 10 (one-way ANOVA, $F_{3, 20} = 84.08$, $P = < 0.001$) and day 14 (one-way ANOVA, $F_{3, 20} = 209.2$, $P = < 0.001$). Compared to the controls, corals subjected to both elevated temperature treatments (eT and eCO_2eT; Figure 4.2) experienced a reduction in symbiont cell density on all days measured (Bonferroni post hoc, $P < 0.05$), while corals exposed to elevated CO_2 at ambient temperature (eCO_2) did not (Bonferroni post hoc, $P > 0.05$).

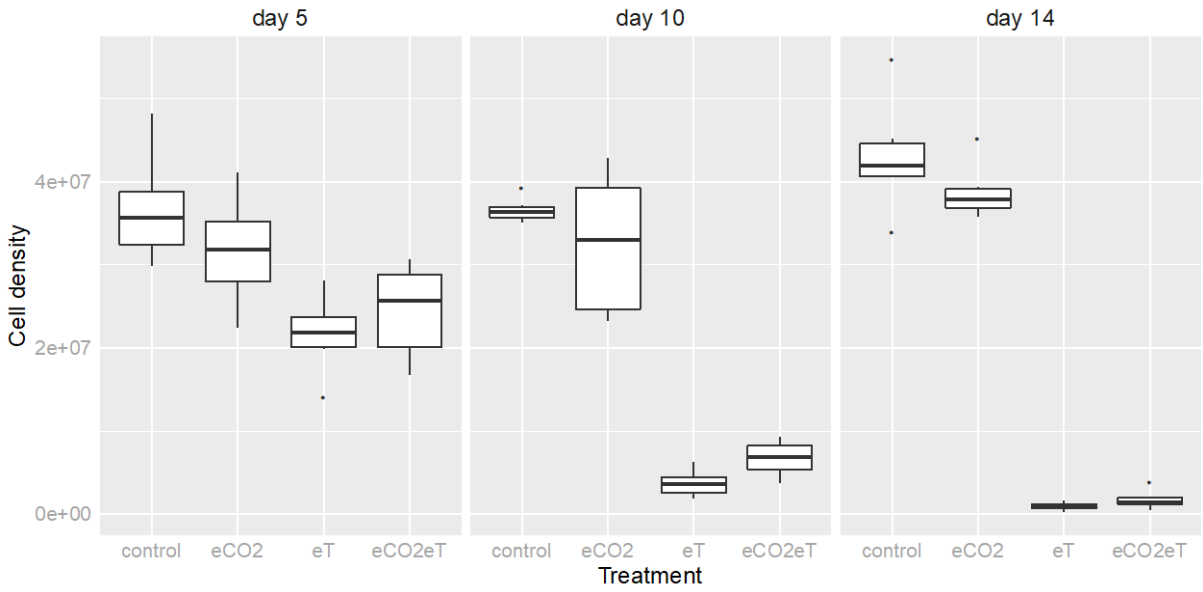


Figure 4.2 Boxplots of symbiont cell density for each treatment at day 5, 10 and 14.

Corals exposed to elevated temperature treatments (eT and eCO_2eT) exhibited significant ($P < 0.05$, Bonferroni post hoc) declines in F_v/F_m (Figure 4.3) on all days measured except day 5. On the other hand, corals exposed to elevated CO_2 at ambient temperatures displayed no significant declines. Compared to control corals, those exposed to the eCO_2 treatment had marginally higher

F_v/F_m. Likewise, corals exposed to the eCO_2eT treatment displayed a similar effect compared to corals exposed to the eT treatment.

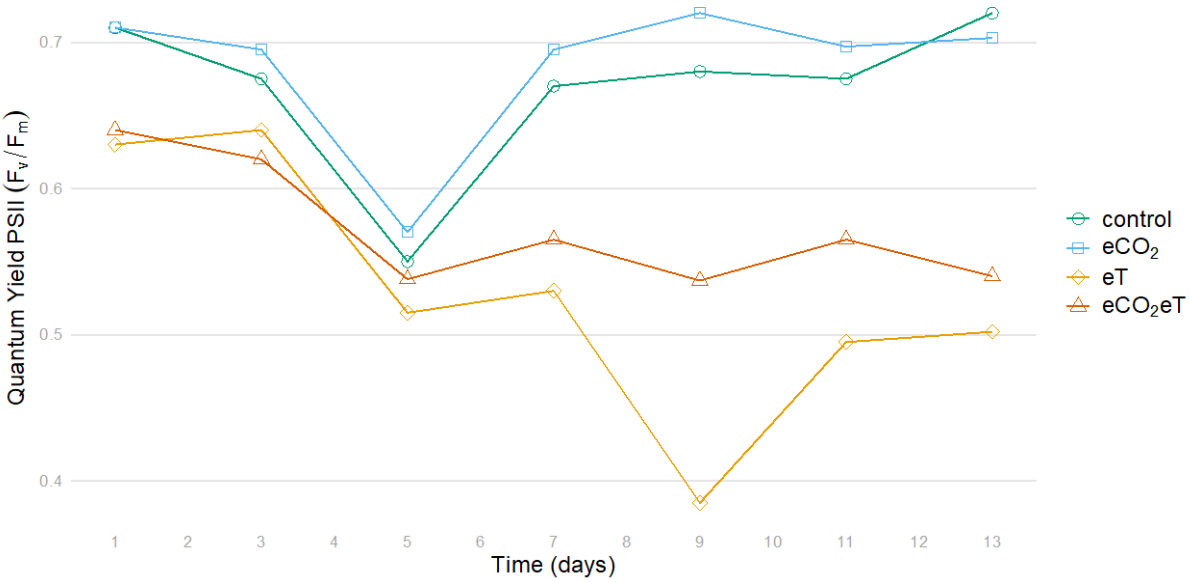


Figure 4.3 The photosynthetic efficiency of *Acropora aspera* for each treatment. Photosynthetic efficiency is defined as the maximum quantum yield of photosystem II (F_v/F_m) for each day it was measured.

4.3.2 Raw data structure

PCA was used to visualise the underlying structure, or variability, in the raw data. The scores plot for the LC-MS data (Figure 4.4a) shows the first two principal components (PCs) captured approximately 46% of the variability. There was some separation in the second PC between control samples and those exposed to elevated temperature (eT and eCO_2eT), while there was no discernible separation between control samples and those exposed to elevated CO_2. For the ^{1}H-NMR data (Figure 4.4b), the first two principal components captured approximately 88% of the variability. The captured ^{1}H-NMR variability was less clearly defined in the PCA scores plot; however, PCA scores for both analytical platforms revealed more clearly defined separation between individual treatments at specific timepoints. For example, corals exposed to elevated temperature treatments for prolonged periods (days 6 and onwards) occupied scores space that was furthest from all other samples (Appendix Figure 4). The Total Ion Chromatograms of PBQCs, which were employed to monitor LC-MS analytical variance, were centred amongst all LC-MS samples, i.e., their variation was adequately represented by the "average" of all variation (see Figure 4.5). Their even separation across both PCs 1 and 2 suggests the analytical variation was unlikely to correlate with any biological variation.

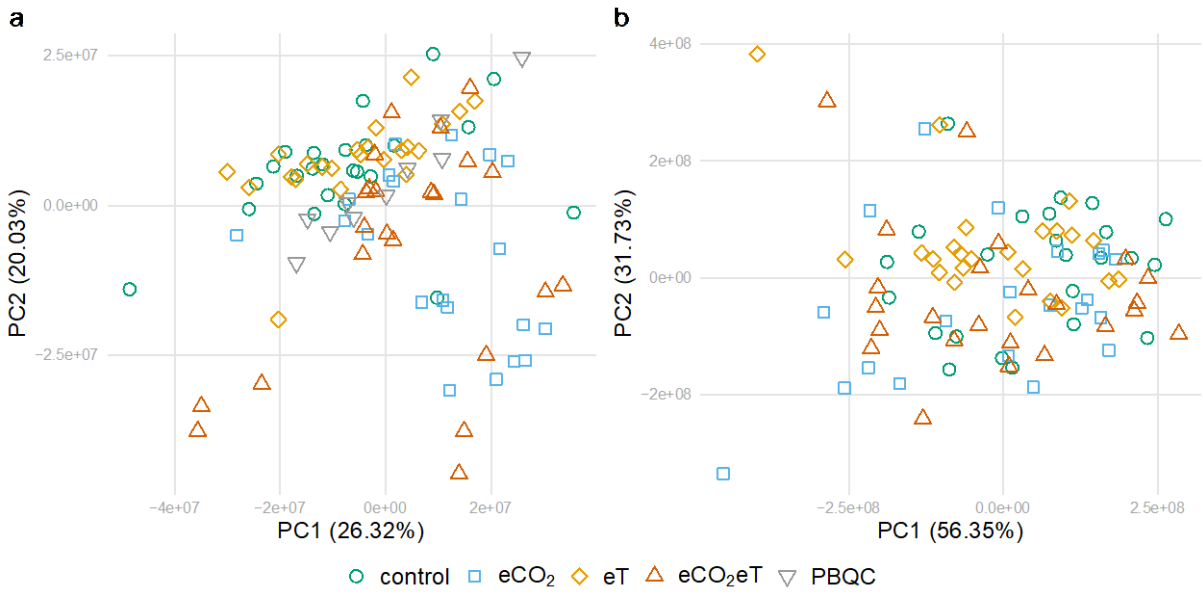


Figure 4.4 PCA scores plot of (a) the mean-centred LCMS data and (b) the mean-centred ^{1}H-NMR data. Corals exposed to elevated temperatures (eT; blue squares and eCO2eT; red triangles) occupy unique regions of the LCMS scores plot in panel a.

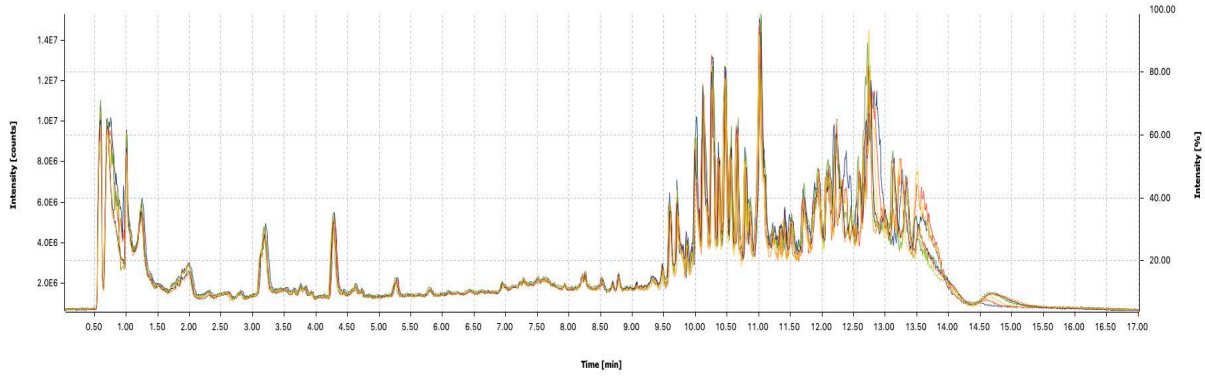


Figure 4.5 Overlaid Total Ion Chromatograms of PBQC samples.

4.3.3 Model performance and parameter optimisation

Model performance was evaluated using three random iterations of a stratified, 10-fold, cross-validation. For both the LC-MS and ^{1}H-NMR datasets, the PLS-DA models benefitted the most from model training (Figure 4.6a-d). The LC-MS PLS-DA and LC-MS RF models performed equally well, achieving a classification accuracy of approximately 84% in both cases. This was not the case for models built from ^{1}H-NMR data, where the PLS-DA model was 15% more accurate at classifying treatment class than the RF model. The LC-MS models provided more accurate predictions of treatment class than the ^{1}H-NMR models, being approximately 7% and 23% more accurate for the PLS-DA and RF models, respectively. The LC-MS PLS-DA model was optimal at 16 latent variables (LVs) with a model accuracy of 84% (Figure 4.5a), while the ^{1}H-NMR PLS-DA model was optimal at 21 LVs with a model accuracy of 77% (Figure 4.5b). As shown in

figure 4.6a-b, increasing the number of LVs beyond that to be considered optimal would not produce any further gain in model accuracy and resulted in overly complex models. The LC-MS RF model was optimal at *mtry* = 350 giving an accuracy of 84% (Figure 4.6c), while the ^{1}H-NMR RF model was optimal at *mtry* = 25 with a model accuracy of 61% (Figure 4.6d).

Figure 4.6 Model tuning results for: (a) the LCMS PLS-DA model, (b) the ^{1}H-NMR PLS-DA model, (c) the LCMS RF model and (d) the ^{1}H-NMR RF model. The best tune indicates the optimal model within one standard error of the empirically best model.

The prediction results from cross-validation were examined to assess the predictive power of the final models in greater detail. For each model, the accuracy, sensitivity (the proportion of samples correctly identified) and specificity (the proportion of samples correctly rejected) for each treatment class were calculated (Table 4.2). The specificity of both LC-MS models and the ^{1}H-NMR PLS-DA model were similar, having true negative rates between 89% and 98% across all treatment classes. The specificity of the ^{1}H-NMR RF model was somewhat lower than the other models, having true negative rates between 86% and 88%. All models for both datasets had slightly, though distinctly higher, specificity for control and eT classes than eCO_2 and eCO_2eT classes.

Sensitivity was slightly more varied between models and treatment class. The LC-MS RF model was particularly sensitive to the control and eCO_2 classes whereas the LC-MS PLS-DA model was more sensitive to the eCO_2 and eCO_2eT classes. Both ^{1}H-NMR models had, in general, lower sensitivity than both LC-MS models; however, the ^{1}H-NMR PLS-DA model and LC-MS PLS-DA model had comparable sensitivity to control and eCO_2 treatment classes.

Table 4.2 Comparison of the cross validation predictions for Partial Least Squares Discriminant Analysis (PLS-DA) and Random Forest (RF) models of the LC-MS and ^{1}H-NMR datasets. Elevated temperature (eT), elevated CO_2 (eCO_2) and combined elevated CO_2 and temperature (eCO_2eT).

		Sensitivity				Specificity			
Method	**Accuracy**	**Control**	**eT**	**eCO_2**	**eCO_2eT**	**Control**	**eT**	**eCO_2**	**eCO_2eT**
LC-MS PLS-DA	0.84	0.8	0.81	0.86	0.89	0.97	0.97	0.93	0.92
LC-MS RF	0.84	0.93	0.72	0.93	0.75	0.98	0.94	0.94	0.93
1H-NMR PLS-DA	0.77	0.82	0.69	0.8	0.77	0.96	0.92	0.93	0.89
1H-NMR RF	0.61	0.65	0.61	0.58	0.6	0.87	0.88	0.86	0.86

4.3.4 Model validation

The validity of all final models was evaluated by predicting an external validation dataset (or test set). However, due to the small number of samples per treatment class in the test set (three for eT and four for each of the other treatments), the test set's sensitivity and specificity results are not sufficiently robust to make explicit comparisons with the respective cross validation results, where a single misclassification would present as a 0.25 change (0.33 for eT). The accuracy measure, being calculated from all 15 test samples is, however, somewhat more robust. The accuracy of the test set predictions (Table 4.3) was either similar, or greater than, the accuracy of predictions performed during cross validation (Table 4.2). The RF model of the LC-MS data most accurately predicted the treatment class of test samples (0.93).

Table 4.3 Comparison of the test set predictions for Partial Least Squares Discriminant Analysis (PLS-DA) and Random Forest (RF) models of the LC-MS and ^{1}H-NMR datasets. Elevated temperature (eT), elevated CO_2 (eCO_2) and combined elevated CO_2 and temperature (eCO_2eT).

		Sensitivity				Specificity			
Method	**Accuracy**	**Control**	**eT**	**eCO_2**	**eCO_2eT**	**Control**	**eT**	**eCO_2**	**eCO_2eT**
LC-MS PLS-DA	0.8	0.75	0.67	0.75	1	0.91	0.92	1	0.91
LC-MS RF	0.93	1	1	1	0.75	1	0.92	1	1
^{1}H-NMR PLS-DA	0.8	0.75	1	0.5	1	0.82	0.92	1	1
^{1}H-NMR RF	0.8	0.75	1	0.75	0.75	0.91	0.92	0.91	1

4.3.5 Temporal and class (treatment) effects on classification

Temporal information was not included in the models; however, its relationship with classification was examined. For both PLS-DA models (Figure 4.7) the greatest proportion of misclassified samples were those exposed to either one or 14 days of treatment. A similar effect

was seen in the LC-MS RF model; however, a greater proportion of samples misclassified as belonging to the eCO_2 treatment class were exposed to six days of treatment. The greater proportion of samples misclassified by the ^{1}H-NMR RF model were, again, exposed to either one or 14 days of treatment; however, there was a greater proportion of misclassified samples exposed to four and six days of treatment than there was with the other models.

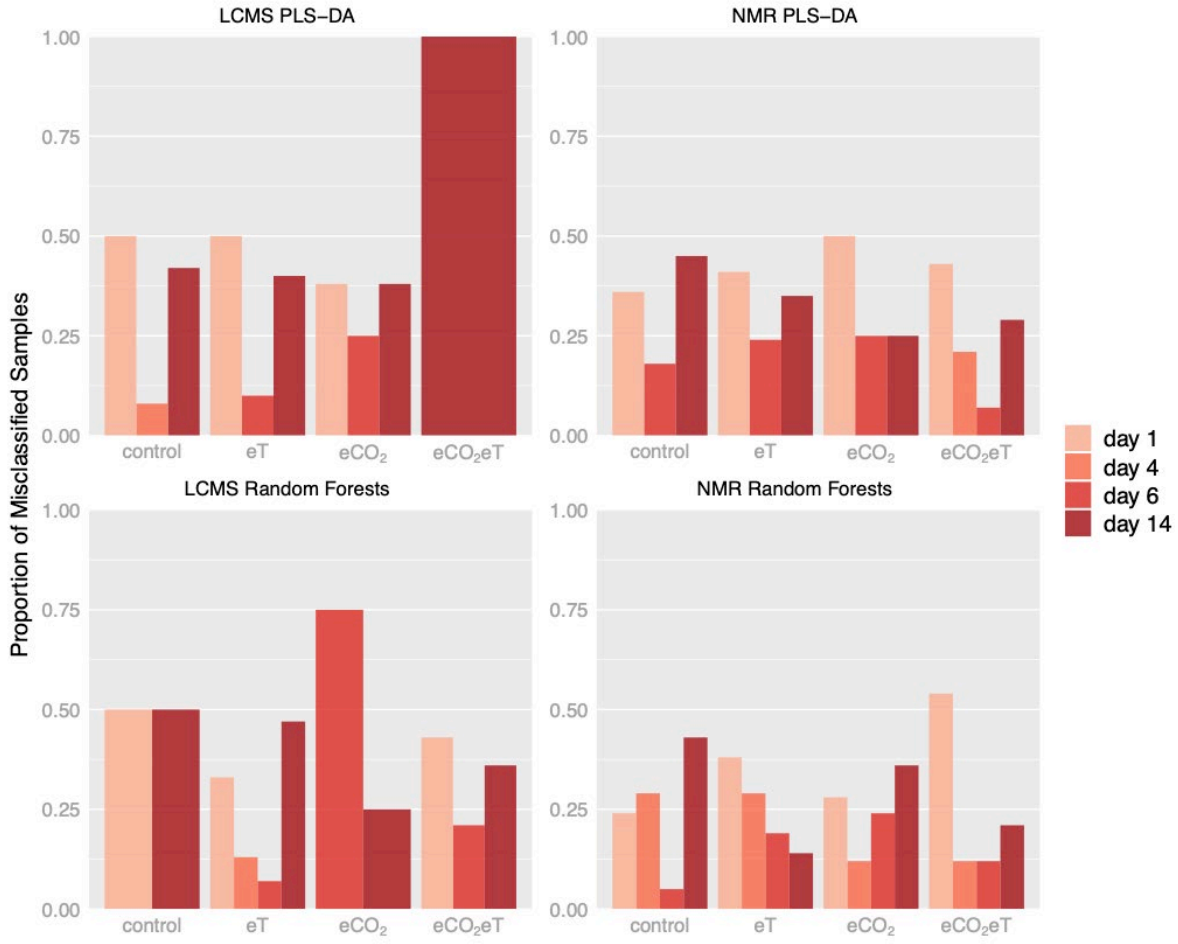


Figure 4.7 The proportion of misclassified samples according to length of exposure as predicted by Partial Least Squares Discriminant Analysis (PLS-DA) and Random Forest (RF) models of the LC-MS and ^{1}H-NMR datasets. Elevated temperature (eT), elevated CO_2 (eCO_2) and combined elevated CO_2 and temperature (eCO_2eT). For example, 50% of Control samples misclassified by the LCMS Random Forests algorithm (bottom left panel) were exposed for 1 day of treatment and the other 50% were exposed for 14 days.

All four models varied in their ability to predict treatment classes, as shown in the confusion matrices (Table 4.4). The LC-MS PLS-DA model, for example, was better at predicting samples exposed to eCO_2 than control and eT samples. The LC-MS RF model performed particularly well in discriminating control and eCO_2 samples having only four incorrect predictions each, while performing less admirably with samples exposed to eT. Models built on ^{1}H-NMR data were less accurate and classification differences between classes were less obvious. The ^{1}H-NMR PLS-

DA model was less effective at classifying eT samples while the ^{1}H-NMR RF model was less effective with samples exposed to elevated CO_2.

Table 4.4 Cross validation confusion matrices for Partial Least Squares Discriminant Analysis (PLS-DA) and Random Forest (RF) models of the LC-MS and ^{1}H-NMR datasets. Elevated temperature (eT), elevated CO_2 (eCO_2) and combined elevated CO_2 and temperature (eCO_2eT).

Predicted Class	LC-MS PLS-DA Observed Class				^{1}H-NMR PLS-DA Observed Class			
	Control	eT	eCO_2	$eTeCO_2$	Control	eT	eCO_2	$eTeCO_2$
Control	48	0	5	0	49	1	6	0
eT	0	44	0	6	0	37	2	12
eCO_2	10	2	49	0	10	1	48	2
$eTeCO_2$	2	8	3	51	1	15	4	46

	LC-MS RF Observed Class				^{1}H-NMR RF Observed Class			
	Control	eT	eCO_2	$eTeCO_2$	Control	eT	eCO_2	$eTeCO_2$
Control	56	0	4	0	39	5	14	3
eT	2	39	0	8	3	33	3	15
eCO_2	2	3	53	6	18	0	35	6
$eTeCO_2$	0	12	0	43	0	16	8	36

4.3.6 Important LC-MS variables

To identify potential compounds involved in the coral stress response, the 20 highest ranked m/z variables driving classification in both LC-MS models were determined (Figure 4.8). The two most important of these, extracted from both the PLS-DA and RF models were m/z 256.1166 (χ^2 (4) = 54.795, p < 0.001) and m/z 274.1992 (χ^2 (4) = 54.271, p < 0.001). Other important variables common to both models were m/z 302.2308 (χ^2 (4) = 44.706, p < 0.001), m/z 567.5809 (χ^2 (4) = 23.533, p < 0.001) and m/z 623.6350 (χ^2 (4) = 18.903, p < 0.001).

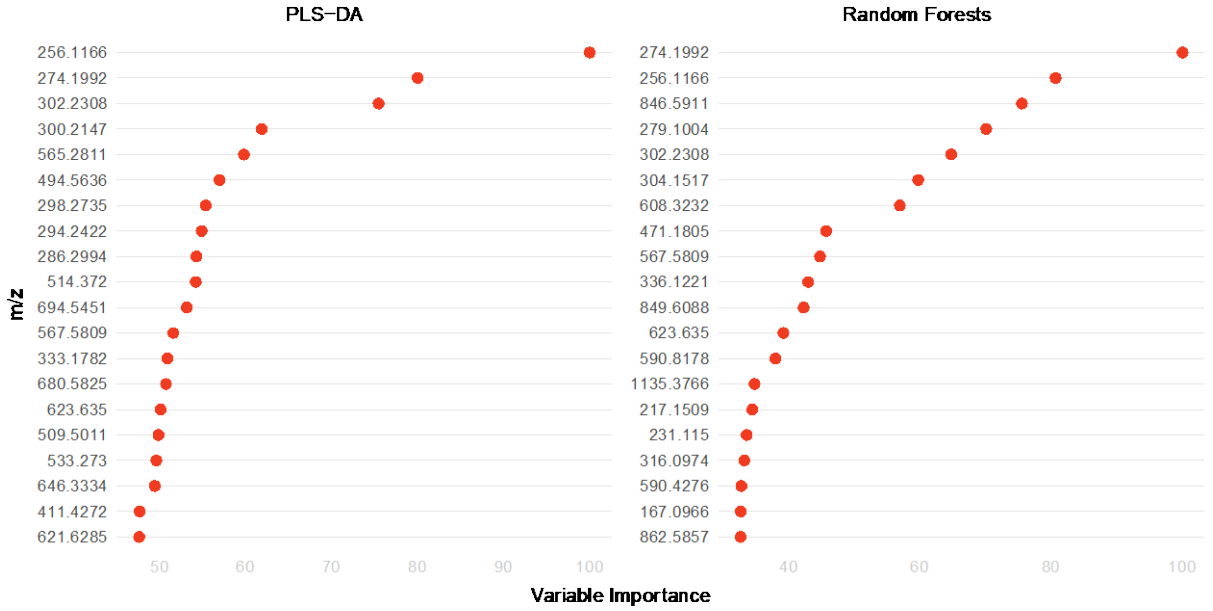


Figure 4.8 The top 20 important LC-MS variables from Partial Least Squares Discriminant Analysis (PLS-DA) and Random Forest (RF) models of the LC-MS dataset.

The top 20 ranked spectral features (Figure 4.8), and their corresponding adducts, were compared with the coral research literature (https://github.com/brgordon17/coralmz) to provide putative identifications. Of the 20 top ranked variables and their corresponding adducts, two spectral features had a similar mass (Δppm < 50 ppm) to three compounds identified in the literature (Table 4.5). One of the most important variables, m/z 256; χ^2 (4) = 54.795, p < 0.001, had a monoisotopic mass within 50 ppm of 3'-deimino-3'-oxoaplysinopsin identified in the *Tubastraea* genus of corals (Guella *et al.*, 1988). The other variable, m/z 217; χ^2 (4) = 36.612, p < 0.001, matched with montiporyne A and montiporyne B isolated from the *Montipora* genus of corals (Bae *et al.*).

An expanded list of the 50 highest ranked m/z variables driving classification in both LC-MS models is provided in Appendix Table 2 and Appendix Table 3.

Table 4.5 Putative assignment of important mass spectral features determined by Partial Least Squares Discriminant Analysis (PLS-DA) and Random Forest (RF) models

Detected *m/z*	Detected Ion	Importance %	Literature *m/z*	Literature Exact Mass	Literature Ion	Molecular Formula	Mass error (ppm)	Compound	Taxon	Reference
PLS—DA										
256.1166	$[M+H]^+$	100	255	255.100784	$[M]^{\bullet+}$	$C_{14}H_{13}N_3O_2$	33	3'-deimino-3'-oxoaplysinopsin	Tubastraea	(Guella *et al.*)
RF										
256.1166	$[M+H]^+$	80.67	255	255.100784	$[M]^{\bullet+}$	$C_{14}H_{13}N_3O_2$	33	3'-deimino-3'-oxoaplysinopsin	Tubastraea	(Guella *et al.*, 1988)
217.1509	$[M+H]^+$	34.46	239	216.1514	$[M+Na]^+$	$C_{15}H_{20}O$	36	Montiporyne A	Montipora	(Bae *et al.*, 2000)
217.1509	$[M+H]^+$	34.46	217	216.1514	$[M+H]^+$	$C_{15}H_{20}O$	36	Montiporyne A	Montipora	(Bae *et al.*, 2000)
217.1509	$[M+H]^+$	34.46	217	216.1514	$[M+H]^+$	$C_{15}H_{20}O$	36	Montiporyne B	Montipora	(Bae *et al.*, 2000)

4.4 Discussion

Temporal ^{1}H-NMR and LC-MS metabolite profiles of the *A. aspera* holobiont subjected to a simulated bleaching event and projected mid-century CO_2 conditions were examined using two common statistical models, PLS-DA and RF, and their prediction accuracy compared. Elevated temperature was the major contributor to coral stress while elevated CO_2 somewhat ameliorated temperature stress and increased symbiont photosynthetic capacity. Of the two analytical techniques and two algorithms examined, LC-MS combined with RF was undoubtedly the most informative, and better at predicting the test data. The two statistical models performed equally well at classifying treatment classes for LC-MS data; however, the RF model proved more sensitive to control samples and those exposed to elevated CO_2, while the PLS-DA model was equally sensitive to all treatment classes. Despite the large temporal variation, and the associated degree of stress, LC-MS, combined with either RF or PLS-DA modelling, proved very capable of correctly classifying corals at different stages of stress exposure. Of note was the ability of each model to classify corals exposed to elevated CO_2 that could not be identified using more traditional methods such as symbiont cell density and PAM fluorometry. Finally, a comparison of the important spectral features with the coral literature provided three putative annotations. These spectral features deserve closer examination with regards to their role in coral health.

4.4.1 Physiological responses to elevated temperature and CO_2

This study compared two commonly used physiological measurements—symbiont cell density and PAM fluorometry—to characterise the level of stress experienced by corals exposed to elevated CO_2 and temperature. While there was no statistically significant response to samples exposed to elevated CO_2, marginally higher F_v/F_m was observed. This is contrary to several other studies that have observed negative responses to photosynthesis and respiration rates under elevated CO_2 conditions. For example, under high irradiance, high CO_2 has been shown to act as a bleaching agent by lowering thermal bleaching thresholds over an 8-week period (Anthony *et al.*, 2008). In *A. millepora*, high CO_2 induced major changes in F_v/F_m, gene expression and respiration for up to 28 days and before changes in mineralisation were apparent (Kaniewska *et al.*, 2012). The lack of a negative PSII response observed in the present study might be due to a commensurate increase in bicarbonate concentration. This has been observed in previous studies over short timescales (~1 week) where the increase in bicarbonate concentrations dominated the subtle effects of

pH and had a positive effect on photosynthesis rates (Weis, 1993, Marubini *et al.*, 2008). Alternatively, atmospheric CO_2 may not act as a bleaching agent at the levels observed in this study; the average elevated ΔpCO_2 from control corals was ~60-70 ppm. The CO_2 induced bleaching of corals observed by Anthony et al (2008) and Kaniewska et al (2012), occurred between 520 ppm and 1300 ppm.

4.4.2 Analytical platform

There is no universally accepted single analytical platform for metabolomics due wholly to the complex and diverse nature of metabolites. MS and NMR are the two most popular detection platforms, and each has its advantages and disadvantages that must be considered prior to analysis or determined experimentally. MS-based methods, for example, have superior sensitivity and lower limits of detection, while NMR is purely quantitative, highly reproducible, and unmatched for determining chemical structures. Signal resolution is particularly important in metabolic profiling experiments and LC-MS is superior to NMR in this regard. Resolution of the NMR signal can be improved with a variety of two-dimensional experiments; however, these are often time consuming and cannot compete with the higher throughput and automation of LC-MS. On the other hand, LC-MS suffers from matrix effects that may mask the detection of biologically important compounds and reduce its quantitative capability (Taylor, 2005). Ideally, many analytical platforms should be employed, preferably simultaneously, to detect and characterise as much chemistry as possible; however, due to cost, time, and logistics, the analyst must often choose one. This is particularly true in the context of biomonitoring, where simplicity and cost are major considerations of large-scale and long-term monitoring.

LC-MS was the most informative analytical platform for classifying corals exposed to elevated CO_2 and temperature. This was visually evident in the PCA scores, where some biological variation within the LC-MS data was observed, while very little variation was observed within the raw ^{1}H-NMR data (Figure 4.4). This was largely due to the increased resolution resulting from the chromatographic separation in LC-MS; however; the complexity of the experiment, which had four treatment classes and four sampling points, also affected the visual separation of samples in the PCA scores plots. Indeed, after reducing the complexity to two treatment classes at a single sampling point (day 14 controls vs day 14 eT, Appendix Figure 5) clear separation was observed in the NMR data.

Superior class predictions further support LC-MS as the preferred analytical platform for biomonitoring of coral. In this experiment, the prediction accuracies of both PLS-DA and

RF models built using LC-MS data were higher than those from ^{1}H-NMR data for both PLS-DA and RF models (Table 4.2). Moreover, the optimal LC-MS PLS-DA model was more parsimonious, requiring less LVs, than the optimal ^{1}H-NMR PLS-DA model. Although model complexity is not an accurate measure of model performance, other studies have observed increases in complexity when differences between classes are smaller and/or model performance deteriorates (Szymańska *et al.*, 2012). From this perspective, the ability of LC-MS to provide higher resolution metabolic profiles and superior predictions than ^{1}H-NMR, naturally places it as the preferred analytical platform for coral metabolic profiling.

4.4.3 Machine learning algorithms

Establishing the best machine learning model for predicting sample classes based on LC-MS data was less straightforward since the PLS-DA and RF models were equally accurate in their overall classification (84%; Table 4.2). However, the PLS-DA model was equally sensitive to all classes, whereas the RF model was less sensitive to both elevated temperature classes (eT and eCO_2eT) and more sensitive to control and eCO_2 classes. This observation was likely due to a stronger interaction between the control and eCO_2 classes compared to the eT and eCO_2eT classes, where tree-based models such as RF are thought to be more accurate (Cutler *et al.*, 2007). Both models were equally specific (specificity > 0.9 across all classes) and so false positive classifications are equally unlikely with each model. At first glance, the PLS-DA with its equal sensitivity to all classes, appears more suitable for predicting samples exposed to ocean warming and acidification from LC-MS data; however, the RF model may be the better choice given it was superior at classifying both the test data and the least influential of the stressors (i.e. CO_2). This is important if synergistic ocean warming and acidification is to be effectively monitored long-term. Also, when the RF model incorrectly predicted eT corals, it classified them as belonging to the other elevated temperature class (eCO_2eT), and vice versa (Table 4.4). PLS-DA, on the other hand, was more likely to falsely classify healthy corals (control and eCO_2 samples) as physiologically stressed corals (eT and eCO_2eT); a far more concerning misclassification that erodes trust in the model. Considering the greater sensitivity of RF models to healthy corals and corals subjected to the low-impact CO_2 treatment, it is the preferred algorithm for metabolomics biomonitoring of synergistic ocean warming and acidification.

4.4.4 Temporal effects on statistical model performance

For all four models, the majority of misclassified samples were those that were exposed to either one or 14 days of treatment (Figure 4.7). Metabolites represent the composite output

of the genomic, transcriptomic and proteomic variability of the organism at any given time (Wang *et al.*, 2010), in effect describing the interaction between an organism and its environment. (i.e. the phenotype). As such, changes in the metabolome are relative to changes in these upstream cellular processes. It is not unsurprising then that the metabolic profiles of those samples exposed to a single day of stress, even 34 °C, did not differ significantly from the controls. A similar phenomenon could be expected for samples exposed to 14 days of elevated temperature, where metabolic profiles would likely converge to a state of complete stress. Samples exposed to four days of treatment were the least likely to be misclassified, where the average metabolic profile of each class is most likely to diverge and where prediction accuracy is maximised.

To explore the temporal effects on coral metabolic fingerprints in more detail, future work should apply a linear temperature gradient along with an increased number of sample replicates and time points. This would facilitate a regression analysis to model changes in temperature (as a continuous variable) as a function of coral biochemistry. That is, instead of classifying corals to binary classes of stressed/healthy, a regression model may offer accurate predictions of ocean temperature at the time of sampling while revealing the biochemical changes that underpin those predictions.

4.4.5 Important variables

Two of the 20 top ranked LC-MS variables were putatively matched to coral metabolites identified in the literature (Gordon & Motti, 2020a). One of these, 3'-deimino-3'-oxoaplysinopsin (m/z 256.1166 $[M+H]^+$), is an alkaloid that was first identified in a scleractinian coral in 1988 (Guella *et al.*, 1988). It has since been identified in marine sponges, sea anemones and molluscs (Bialonska & Zjawiony, 2009). There is little research on its function in corals or its significance to climate related stressors; however, the photoisomerisation of aplysinopsins may act as a protection against UV radiation by entrapping radiant energy (Bialonska & Zjawiony, 2009, Guella *et al.*, 1988). In corals, aplysinopsins have been proposed as defensive compounds such as: antimicrobials (Koh & Sweatman, 2000); antipredation compounds (Okuda *et al.*, 1982); coral-produced inhibitors of cell cleavage in fertilised sea urchin eggs (Fusetani *et al.*, 1986) and as inhibitors of competing coral larval growth (Koh & Sweatman, 2000). Six dihydroaplysinopsins, isolated from sea anemones, have also been identified as chemical attractants for anemone fish, and are important for maintaining the host-guess symbiosis (Murata *et al.*, 1986). Given the somewhat ubiquitous nature of aplysinopsins, a more detailed analysis of these compounds is

warranted; in particular, future work should focus on formally identifying the presence of the compound in *A. aspera* while quantifying any potential correlation with climate related stressors such as temperature, irradiance and CO_2.

The second annotated variable was m/z 217.1509 $[M+H]^+$ and was putatively matched to the polyacetylenes montiporyne A and montiporyne B. These two compounds were first isolated in 2000 from *Montipora* sp. (Bae *et al.*, 2000). To date, much of the research on these compounds has focused on their antitumor, antibacterial, antimicrobial and antifungal properties (Negri, 2015, Legrave *et al.*, 2015). They are cytotoxic compounds that are extremely effective at inhibiting bacterial strains, including known coral pathogens in the surrounding seawater (Gochfeld & Aeby, 2008). Ecologically, they have been shown to act as sperm attractants in *Montipora digitata* (Coll *et al.*, 1994), which is a hermaphroditic coral that releases eggs and sperms in a single bundle. In a certain ratio, these acetylenic alcohols were shown to be more effective in attracting sperm from *M. digitata* than from other species of *Montipora*, reducing the chances of hybridisation between different *Montipora* species. They are also known to irreversibly inhibit photosynthesis in *Symbiodiniaceae* within minutes of exposure (Hagedorn *et al.*, 2015). If montiporynes are indeed a component of the *A. aspera* metabolome then its role in the *A. aspera* stress response may, along with its known cytotoxicity and detriment to *Symbiodiniaceae*, suggest a role in the bleaching process. Nevertheless, future research must formally confirm its presence in *A. aspera* and fully investigate its potential role in the stress response.

None of the remaining 18 top-ranked spectral features driving the classification of stressors could be putatively annotated. Compound identification requires considerable time and resources that were not within the scope of this research and unfortunately, it remains the major bottleneck in coral metabolic profiling. This is especially true for secondary metabolites, as reference standards are not readily available for comparison. In this regard, a multi-platform approach that utilises both NMR and MS platforms to elucidate chemical structures and provide conclusive compound identifications is essential. To advance the field of coral metabolomics and our understanding of coral physiology, future research must focus on isolating and identifying a much greater proportion of the coral metabolome, with the same sense of urgency and effort that has been invested into decoding the coral and symbiont genome.

4.5 Conclusion

This study highlights the potential of LC-MS-based metabolic fingerprinting to predict discrete outcomes of coral exposure to varying degrees of ocean warming and acidification with a high level of accuracy. Of note is the ability of this approach to classify corals exposed to mid-century levels of ocean pH with a high degree of accuracy. It's an important finding given few studies have observed CO_2 induced changes in coral physiology at levels that are likely to persist, and gradually increase, for the remainder of the century: most ocean acidification studies to date have focussed on CO_2 levels representing worst-case emissions scenarios (Hughes *et al.*, 2017a). Additionally, this study has demonstrated the ability of this technique to distinguish corals exposed to a single stressor from those exposed to a combination of stressors. This research provides potential leads, in the form of important variables, which, with further examination, may enlighten our understanding of coral physiology and the biochemical interplay between host and symbiont. This work clearly demonstrates the potential utility of metabolic profiling for biomonitoring of coral reefs. With that in mind, future research should focus on improving model accuracy and resolution via an increase in sampling frequency and replication.

Chapter 5
Predicting the Photosynthetic Efficiency of *Acropora aspera* Exposed to Elevated Temperature using LC-MS Metabolomics and a Random Forests Machine Learning Model

This chapter is in preparation for submission: Benjamin R. Gordon, Sarah Gierz, William Leggat, Ute Roessner, Adrian Lutz and Cherie A. Motti

Benjamin Gordon wrote this chapter, designed and conducted the research, and analysed the data. Co-authors provided intellectual, experimental and editorial support.

5.1 Introduction

Effective management of coral reef ecosystems requires knowledge of coral functional status and health. To date, this knowledge is acquired by monitoring indicators of reef state such as: coral abundance or the presence and prevalence of coral bleaching and/or disease. In many cases, these indicators are symptoms resulting from external stressors such as ocean warming and nutrient runoff, which are monitored by reef managers to predict changes in reef state. Given the complexity of ecological systems, this somewhat reactive approach to monitoring provides managers with only limited descriptions of coral functional status, hindering their ability to predict the resilience and future responses of coral reefs (Lam *et al.*, 2017, Hughes *et al.*, 2017a); as such, the ability to prioritise interventions and avoid undesirable changes in the reef ecosystem is diminished. To manage reefs effectively, more precise indicators of coral functional status and health are required.

Molecular biomarkers offer a more informative overview of coral functional status; however, their inclusion in monitoring programs has been hindered by the cost and logistics associated with instrumentation and their analysis. These hurdles may be overcome as technology and our understanding of coral physiology matures. Indeed, analytical platforms used to study coral functional status, such as Pulse Amplitude Modulation (PAM) fluorometry, are now commonplace in the field and are routinely deployed in regional scale monitoring programs.

PAM is a rapid and non-destructive technique to assess the photosynthetic efficiency of the coral symbiont. Photosynthetic efficiency is determined by the ratio of variable fluorescence to maximal fluorescence (F_v/F_m), otherwise known as the maximum quantum yield of photosystem II (PSII). A stress-induced reduction in F_v/F_m indicates a breakdown in PSII function resulting in a loss of oxygen-evolving capacity, disruption of the PSII reaction centre and the disassociation of its light-harvesting complex that often precedes coral bleaching (Warner *et al.*, 1996, Jones *et al.*, 1998).

Photosynthetic efficiency has been explored as an indicator of coral health in regional scale monitoring because of its sensitivity to a variety of stressors. In a 2007 report to the Australian Government Marine and Tropical Sciences Research Facility (Cooper & Fabricius, 2007), managers advocated for the use of PAM fluorometry techniques based, in part, on evidence that the F_v/F_m of coral has a positive correlation with improvements in water quality (Cooper & Ulstrup, 2009). In a regional monitoring program in the Arabian Gulf, PAM- and visual-based monitoring of coral health was employed over a one-and-a-half-year

period, where it was found that corals experiencing reduced F_v/F_m, despite having no visual signs of stress, were more likely to experience subsequent necrosis (Febbo *et al.*, 2012). In the Sekisei lagoon in Okanawa, Japan, the F_v/F_m of 28 coral genera (c. 68 species) were monitored both before and after a natural bleaching event using PAM fluorometry (Okamoto *et al.*, 2005), providing an important baseline F_v/F_m of healthy and stressed coral populations.

What constitutes a "healthy" or "unhealthy" F_v/F_m of wild coral populations across a variety of reef landscapes is of particular importance for monitoring and management of coral reefs. The current consensus for a healthy coral, as measured by PAM fluorometry, is when photosynthesis is working at peak efficiency. Typically, such corals have F_v/F_m values between 0.50-0.70, depending on the species of coral and irradiance (Fitt *et al.*, 2001, Warner *et al.*, 1996, Jones *et al.*, 1998, Okamoto *et al.*, 2005). While it is well-known that coral F_v/F_m is reduced when exposed to stress, classification of corals as unhealthy based on sub-optimal F_v/F_m below ~0.50 is less straightforward. Indeed, sub-optimal F_v/F_m values have been recorded in otherwise healthy corals for a variety of reasons, such as: the dark acclimation time employed prior to measurement (Fitt & Cook, 2001), seasonal fluctuations (Warner *et al.*, 2002), host and/or symbiont identity (Warner *et al.*, 2006), depth and irradiance (Winters *et al.*, 2006, Okamoto *et al.*, 2005), and diurnal variation (Jones & Hoegh-Guldberg, 2001, Belshe *et al.*, 2007).

Precise interpretation of suboptimal F_v/F_m in corals – to link photoinhibition with thermal stress for example – requires complimentary physiological measurements (Fitt *et al.*, 2001). *Symbiodiniaceae* cell density and chlorophyll (chl) concentration are typically utilised in coral bleaching experiments; however, with these parameters alone, the focus remains solely on the response of the algal symbiont to the stressor(s) and not of the holobiont per se. Linking a stress response with photoinhibition is not trivial given the complexity of the holobiont, which comprises the animal host and algal endosymbiont along with a complex microbiome and numerous macrobiota (Rohwer *et al.*, 2002, Stella *et al.*, 2011, Glasl *et al.*, 2019). To effectively link photoinhibition with the functional state of coral, a more holistic analysis, which better captures and describes the phenotype of all partners in the holobiont, is required.

> "Metabolites are the end products of cellular regulatory processes, and their levels can be regarded as the ultimate response of biological systems to genetic or environmental changes" (Fiehn, 2002).

The coral metabolome at any one time is a unique chemical fingerprint with the potential to provide detailed descriptions of the functional state of the coral holobiont. While the technique has yet to be employed as a coral reef monitoring tool, recent research has identified a variety of coral phenotypes associated with a number of stressors, such as: anthropogenic climate change (Hillyer *et al.*, 2016, Hillyer *et al.*, 2017b, Hillyer *et al.*, 2017a, Hillyer *et al.*, 2018, Petrou *et al.*, 2018, Putnam *et al.*, 2016, Sogin *et al.*, 2016, Farag *et al.*, 2018), unique environmental conditions (Januar *et al.*, 2012, He *et al.*, 2014, Farag *et al.*, 2016, Klueter *et al.*, 2015) and unique physical attributes (Sogin *et al.*, 2014, Quinn *et al.*, 2016, Farag *et al.*, 2017a, Farag *et al.*, 2017b, Matthews *et al.*, 2017, Sogin *et al.*, 2017). Interpretation of these chemical fingerprints, however, is time consuming and difficult, due in part to the relative infancy of metabolomics in coral research and the difficulties associated with metabolite identification. Nevertheless, as the field matures and new discoveries come to light, coral metabolomics will play a defining role in the discovery of molecular biomarkers and bioindicators describing coral health.

This research seeks to establish a molecular phenotype associated with temperature-induced photoinhibition of corals by investigating the hypothesis that Liquid Chromatography Mass Spectrometry (LC-MS) metabolic profiling and machine learning using a Random Forests (RF) algorithm can predict the photosynthetic efficiency of *Acropora aspera* exposed to a linear increase in daytime seawater temperatures over 15 days. Given that long-term monitoring of coral metabolic profiles will inevitably be hampered by batch specific sources of variation, this study analysed samples in two separate batches and investigated the resulting batch effects to determine their extent and how these may be managed under a biomonitoring scenario. Chapter 4 suffered from a lack of sample replication, which hindered model testing; consequently, sample replication was increased here to provide a more representative dataset and to adequately test the model. Model performance and the metabolites driving prediction are discussed in detail, as is the concept of F_v/F_m as an indicator of coral health.

5.2 Methods

5.2.1 Experimental conditions

Coral nubbins (n = 312) were collected (under Great Barrier Reef Marine Park Authority permit G13/36402.1) from four colonies of *Acropora aspera* (tan morph, approximately 78 from each colony) on the reef flat of Heron Island at low tide in May 2013. Submerged coral nubbins were randomly placed in holding racks before being transferred to eight 65 L

experimental tanks, where they were allowed to acclimate at ambient seawater temperature for five days: complete tissue regrowth at the site of excision confirmed acclimation of the nubbins. Each tank was supplied with a continuous flow of sand-filtered seawater pumped from the reef flat via two ~200 L sump tanks (total system volume ~460 L).

Four tanks were subjected to ambient temperature (~21-24 ℃; control) conditions and the other four to elevated temperature (treatment) conditions (Figure 5.1). The control tanks were maintained at ambient seawater temperature for the duration of the experiment. The temperature of the treatment tanks was increased by approximately 0.7 °C per day, during the day, for eleven days (25–32.3 °C), then held at 33 °C for two days and then at 34 °C for the final two days. Treatment tanks were heated with a 300 Watt Eheim Jager heater (Eheim, Deisizou, Germany) and four 25 W Thermosafe™ aquarium heaters (Aqua One, Pet HQ, Townsville, QLD). Diurnal fluctuations in the control tanks were as per the ambient seawater temperature at Heron Island. To allow for some recovery from elevated daytime temperatures, and to maintain diurnal temperature variation, the Eheim heater was turned off at sunset and back on at sunrise. Temperatures in each tank were recorded every 10 minutes with temperature data loggers (HOBO®, OneTemp, Brisbane, Australia). Reef flat temperatures were monitored by three Integrated Marine Observing System (IMOS) sensor floats (AIMS data centre, Heron Island sensor floats 1, 2 and 4). Light levels were monitored over the course of the experiment every 10 min with Photosynthetically Active Radiation recorders (Odyssey, Dataflow Systems Limited, New Zealand).

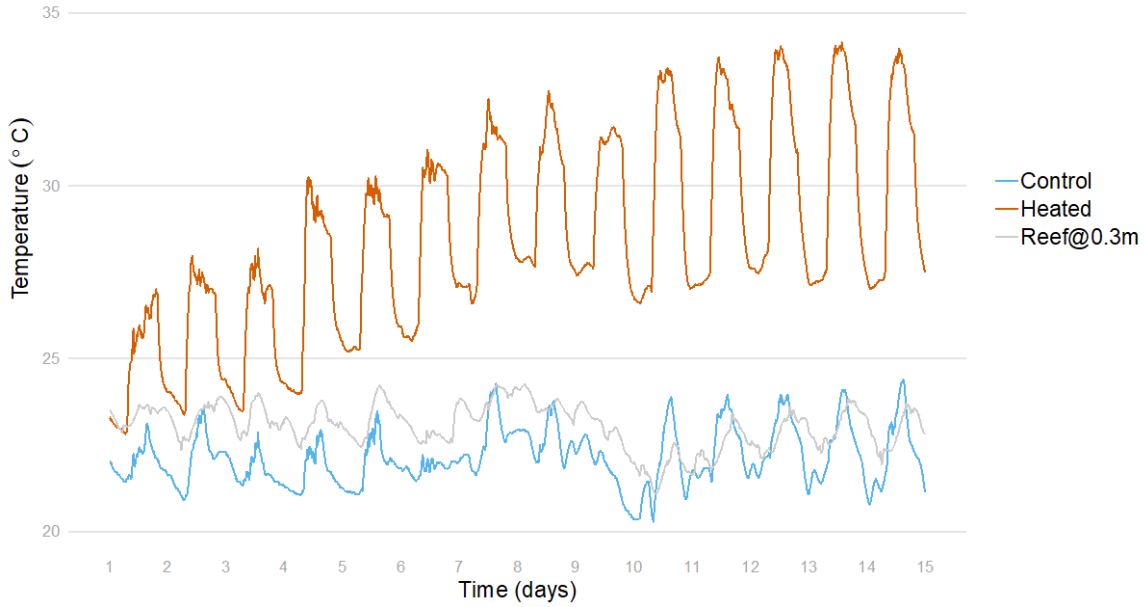


Figure 5.1 Temperature profiles of control (blue) and elevated temperature (red) treatments for the duration of the experiment. The grey data series represents the average temperature of three IMOS sensor floats (at 0.3 m depth) located on the Heron Island reef flats over the same period.

5.2.2 Sample Collection

Three coral nubbins were collected from each of the four control and heated tanks on days 1, 5, 8, 10, 12 and 15 (n = 144; 12 nubbins × 2 treatments × 6 time-points) and snap-frozen in liquid nitrogen. All samples were stored at -80 °C for metabolomic analyses. On days 1, 5, 8, 10, 12 and 15, three coral nubbins were collected from each of the control and heated tanks (n = 144) for chlorophyll quantification and *Symbiodiniaceae* cell count. Three coral nubbins from each of the control and heated tanks (n = 24) were designated for repeated daily PAM measurements.

5.2.3 Quantification of thermal stress

Chlorophyll (chl) fluorescence of *A. aspera* was measured using imaging PAM fluorometry (MAXI imaging PAM, Waltz, Effeltrich, Germany) to monitor the dark-adapted F_v/F_m response to exposure. Repeated PAM measurements of the same three designated coral nubbins from each aquarium were conducted every day ~30 min after sunset. Chl fluorescence was measured with a weak pulse of light, followed by a saturating pulse of 2,700 µmol quanta m^{-2} s^{-1} of photosynthetically active radiation (PAR) for 800 ms. Coral nubbins were returned to aquaria immediately following measurement.

To quantify *Symbiodiniaceae* cell density and chl levels coral tissue was removed from nubbins with a dental irrigator filled with filtered (0.22 μm) seawater. The blastate was homogenised with an immersion blender for 5 s before being centrifuged at 3,076 g for 3 min to pelletise algal cells. Due to a procedural oversight, the blastate was not homogenized with an immersion blender prior to centrifugation on days 1 and 5 and the results were found to be unreliable. Pelletised cells were suspended in 50 mL of seawater and a 1 mL aliquot removed for cell density approximation. Cell number was determined using a Neubauer haemocytometer, with replicate cell counts performed (n = 5). The remaining cells (in 49 mL) were centrifuged at 3,076 g for 3 min to pelletise cells and stored at −80 °C for chl-*a* and chl-*c* quantification. Samples were extracted in 90% acetone for 20 h in the dark at 4 °C and chl quantified using the equations of Jeffrey and Humphrey (1975). The surface area of nubbins was determined using the wax dipping method (Stimson & Kinzie Iii, 1991).

5.2.4 LC-MS sample preparation

Frozen coral nubbins were placed into 20 mL glass scintillation vials and lyophilised for 24 h then extracted in 70 % aqueous methanol as per Chapter 3 and Gordon *et al.* (2013). Coral extracts were centrifuged at 5,800 g to remove any undissolved cellular debris and the supernatant transferred to LC-MS grade, amber vials (Phenomenex, Lane Cove, Australia). All samples were stored at -80 °C prior to analysis. PBQCs were prepared according to the methods described in section 4.2.4 and analysed after every tenth injection to monitor analytical variation.

5.2.5 LC-MS data acquisition

Reverse Phase chromatography was performed using an Agilent 1200 HPLC system (Santa Clara, CA, USA) consisting of a vacuum degasser, binary pump, thermostatic auto sampler and column compartment. Chromatography was conducted using the following conditions: a Zorbax Eclipse XDB-C18, 2.1 mm x 100 mm, 1.8 μm (Agilent, Australia) column; solvent (A) 0.1% formic acid in Milli-Q water and solvent (B) 0.1% formic acid in acetonitrile (ACN); a flow rate of 0.4 mL min^{-1} and; a column temperature of 40 °C. Samples were injected (10 μL) and eluted with a 10 min linear gradient from 5% (B) to 100% (B), followed by a 2 min hold at 100% (B), then returned to 5% (B) and re-equilibrated for 5 min (total time of 17 min). Samples were randomised to ensure analytical variation did not correlate with biological variation and a pooled biological quality control was run every tenth sample.

The mass spectrometer was an Agilent 6520 ESI-QTOF-MS (Santa Clara, CA, USA) with a dual spray ESI source operated in positive ion mode (a small number of samples were analysed in negative ionisation mode to determine the most suitable method; see Appendix Figure 3). The source conditions were: nebuliser pressure of 45 psi; gas temperature of 300 °C; drying gas flow of 10 L min^{-1}; capillary voltage of 4000 V and skimmer 65 V; fast polarity enabled. Measurements were performed in the extended dynamic range mode (m/z range of 100-3200; sampling rate 2 GHz); scan rate of 2.03 spectra/s collected as MS1 centroid data. The mass spectrometer was calibrated using Agilent's ES Tuning Mix ACN Solution, containing several calibration ions at mass values from 117 Da to 2,722 Da. PBQCs were used to monitor mass deviations of some ubiquitous features and analytical variation such as retention time shifts (typically < 0.1 min) and ionisation efficiency.

To assess batch variation samples were split into two batches, where samples from days 5 and 12 were analysed as batch one and all other samples analysed independently, seven days later, as batch two.

5.2.6 LC-MS data pre-processing

LC-MS data was converted to mzXML format using ProteoWizard's MSConvert tool (v3.0.6585; Chambers *et al.*, 2012, Kessner *et al.*, 2008) with an absolute intensity threshold of 1000. LC-MS chromatograms were integrated and aligned in the R environment (v3.5.0; R Core Team, 2018) using the XCMS package (v3.6.2; Smith *et al.*, 2006, Tautenhahn *et al.*, 2008, Benton *et al.*, 2010). Feature detection was performed using the centWave method (Tautenhahn *et al.*, 2008) with the following modified parameters: ppm = 30, peakwidth = c(10, 60), mzdiff = -0.001, integrate = 1 and, prefilter = c(3, 1100). Retention time correction utilised the obiwarp method with binSize = 0.5 (Prince & Marcotte, 2006). Chromatographic peaks were grouped across samples using the peak density parameter with: bw = 5, binSize = 0.025 and minFraction = 0.5. Peak Filling was employed using the *fillChromPeaks()* function with default values.

Isotopes and adducts were annotated using the *annotate()* function within the CAMERA package (v1.38.1; Kuhl *et al.*, 2011) with the following parameters: perfwhm = 0.7, cor_eic_th = 0.75, ppm = 10 and polarity = 'positive'. All M^{+1}, M^{+2}, M^{+3} and M^{+4} isotopes identified by the CAMERA package were removed from the resulting peak list.

Unreliable spectral features with 80% or more missing values across all classes were deleted entirely. Where a class had 65% or more missing values within any spectral feature, a random

value between zero and the minimum intensity was imputed to represent a non-detect specific to that class. The remaining missing values were imputed using a Random Forest trained on the observed values (ntree = 100, mtry = √x), as described by Stekhoven and Buehimann (2011).

Batch effects were removed using a Principal Components Analysis (PCA) and constrained optimisation technique implemented by the Harman package (v1.12.0; Oytam *et al.*, 2016) with limit = 0.99, numrepeats = 100000, and two class levels (control and heat-stressed). Finally, boxplots of the relative log abundance for each sample and a PCA were utilised to assess the data for unwanted variation. The final, pre-processed, dataset had a total of 6334 features.

5.2.7 Model tuning and validation

The experiment utilised an optimised Random Forest model (RF) to predict the F_v/F_m of corals exposed to simulated ocean warming. Model parameters were optimised using repeated k-fold cross-validation (k = 10) as outlined in Figure 5.2. The data was partitioned into training and test sets and the test set excluded from the model construction procedure. A balanced 90:10 (train:test) split was created by grouping the descriptor (F_v/F_m) using percentiles and sampling evenly from within each group. The root mean square error (RMSE) of prediction was used to tune the optimal *mtry* value during model construction. The final, optimal model was the most accurate model within one standard error of the empirically best model (Kohavi, 1995, Breiman *et al.*, 1984).

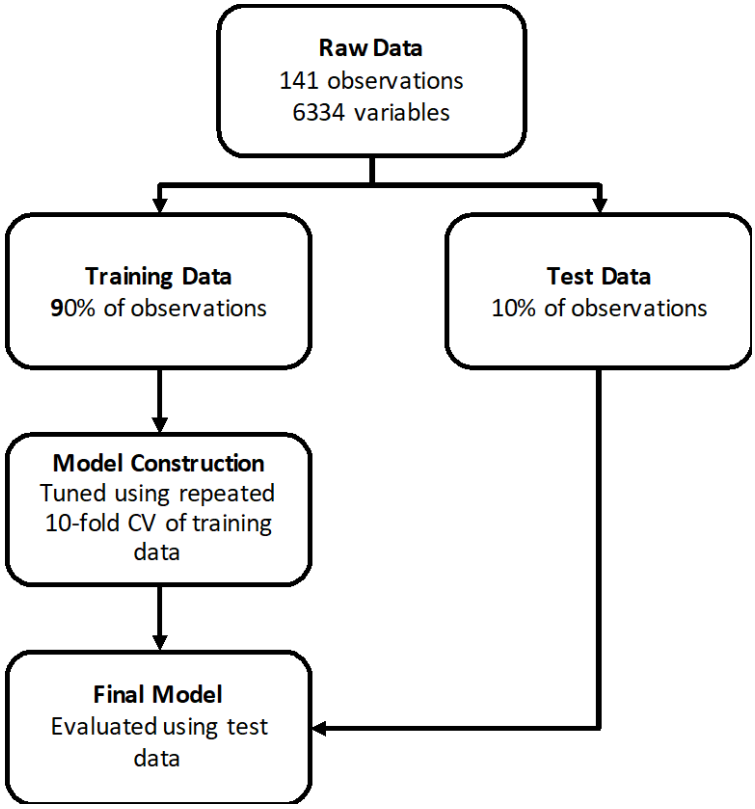


Figure 5.2 Workflow of the model tuning and validation procedure for the analysis of LC-MS data

5.2.8 Variable selection

Variables driving prediction were identified by permuting each spectral feature then comparing the RMSE of prediction on the out-of-bag portion of data both before and after permutation as described in the *randomForest* package (v4.6-14) and by Breiman (Breiman, 2001). Important spectral features were cross-referenced with the MarinLit database and metabolites reported in the coral research literature (Gordon & Motti, 2020a) using a mass error of Δ50 ppm as per Van Assche et al (2015), providing level 2 putative annotations where possible (Sumner *et al.*, 2007).

5.2.9 Data

The R scripts for this experiment are available from GitHub at https://github.com/brgordon17/fvfm-prediction (Gordon, 2020b). The full list of metabolites retrieved from the coral research literature and the MarinLit database is available as an R package on GitHub at https://github.com/brgordon17/coralmz (Gordon & Motti, 2020a).

5.3 Results

5.3.1 Quantification of thermal stress

The functional state of *A. aspera* nubbins was affected by elevated temperature by day 10 of exposure. At that point, coral nubbins had experienced at least six days of seawater temperatures above 30 °C with days 8 to 10 experiencing maximum midday temperatures of approximately 32 °C. By day 10 of exposure, a 22.3% decline in *Symbiodiniaceae* cell density and a 46.6% decline in F_v/F_m was observed.

Exposure of *A. aspera* to the linear temperature gradient as described in section 5.2.1, resulted in statistically significant declines of F_v/F_m between treatments over time (Figure 5.3; RMANOVA, $F_{14, 330} = 208.4$, $p < 0.001$). Tukey's HSD post hoc tests were performed to compare treated samples with controls coinciding with each of the six collection points. There was no statistically significant decline in F_v/F_m for treated samples between days 1 and 6. A statistically significant decline ($p < 0.01$) in F_v/F_m was observed for all treated samples from day 7 onwards. The average decline in F_v/F_m for samples collected after day 5 were: 0.158 ($p < 0.001$) for day 8; 0.296 ($p < 0.001$) for day 10; 0.475 ($p < 0.001$) for day 12 and; 0.660 ($p < 0.001$) for day 15.

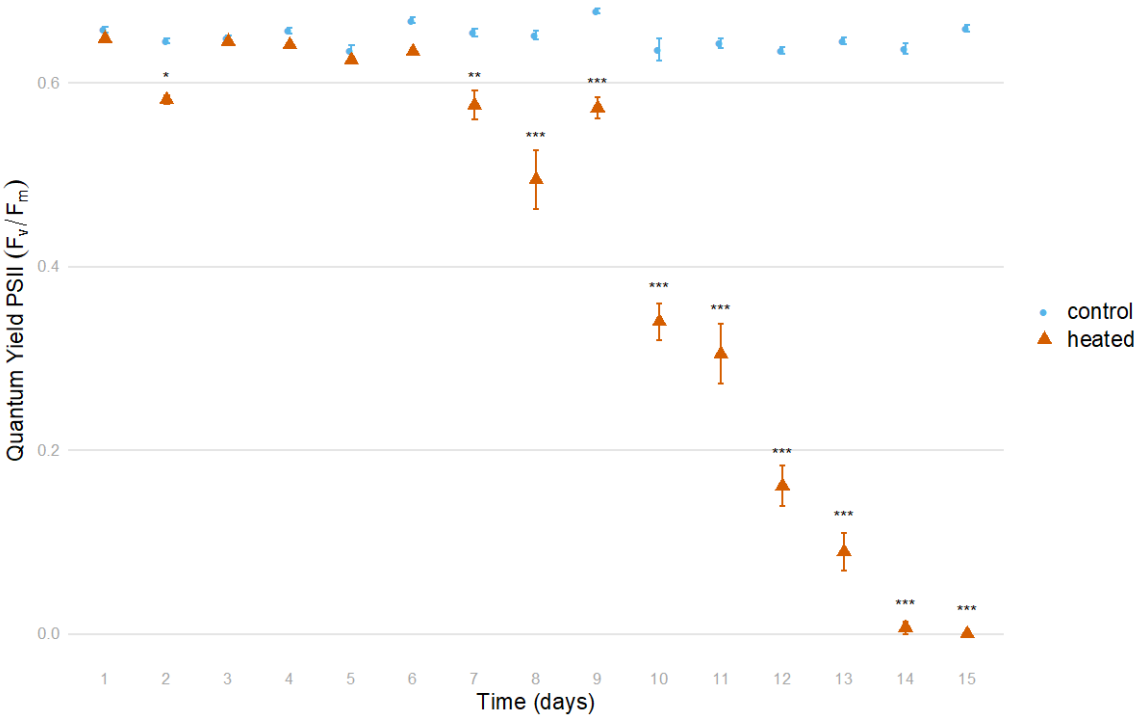


Figure 5.3 The maximum quantum yield of photosystem II (PSII, ratio of variable fluorescence to maximal fluorescence, F_v/F_m mean ± s.e.) over time for control samples (blue circles) and heat-stressed samples (red triangles). Asterisks denote significance at *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

Elevated temperatures correlated with statistically significant decreases in symbiont cell density from days 10 to 15 (Figure 5.4a). By day 15, the mean symbiont cell density of heat-stressed samples was 6×10^5 cells cm^{-2}, representing a decline of 37.5% compared to the control samples (1.6×10^6 cells cm^{-2}, Bonferroni post hoc, $p < 0.001$). In contrast to the declines in cell density and F_v/F_m over the course of the experiment, mean chl-*a* increased over the same temporal scale (Figure 5.4b). By day 15, the mean chl-*a* concentration of heat-stressed samples (0.93 pg $cell^{-1}$) was 102% greater than the control samples (0.46 pg $cell^{-1}$).

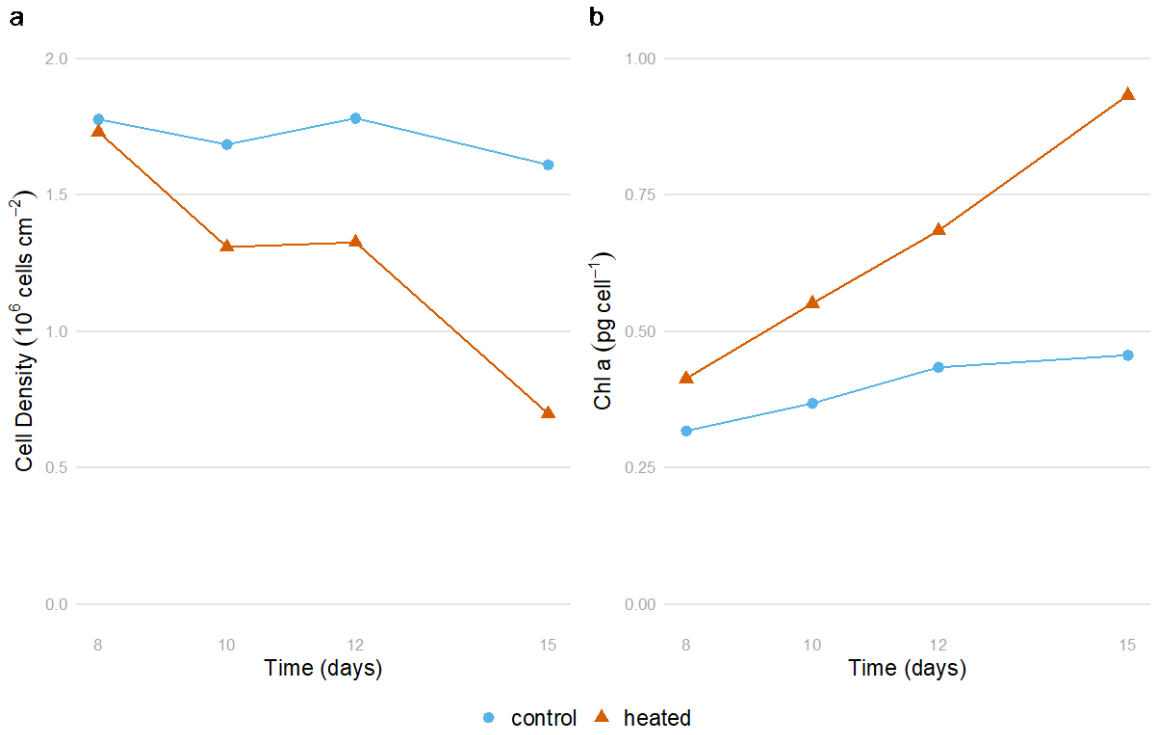


Figure 5.4 Symbiont cell density (a) and chlorophyll-a (chl-a) concentration (b) for heat-stressed (red triangles) and control treatments (blue circles) on days 8, 10, 12, and 15.

5.3.2 Batch effects and data structure

The relative log abundance (RLA) of each sample's metabolic profile was examined to compare the variation between samples. A pronounced batch effect was visible in box plots of the RLA of the unadjusted data (Figure 5.5a). Samples analysed in batch one had noticeably higher spectral feature abundances and more outliers than samples analysed in batch two. The Harman batch correction algorithm (Oytam *et al.*, 2016) was employed to mitigate and/or remove the unwanted variation while quantifying the loss of biological variation. Using a confidence level of 0.99 and 100,000 simulations, the Harman algorithm identified principal components PC1 and PC159 having a confidence less than 0.99 and thus containing the batch effect. These two PCs were corrected by the algorithm, where the

correction applied to each was: 0.05 and 0.09 respectively. RLA plots of the spectral features both before (Figure 5.5a) and after (Figure 5.5b) the Harman algorithm was applied, clearly shows its effectiveness in removing the bulk of the batch effect.

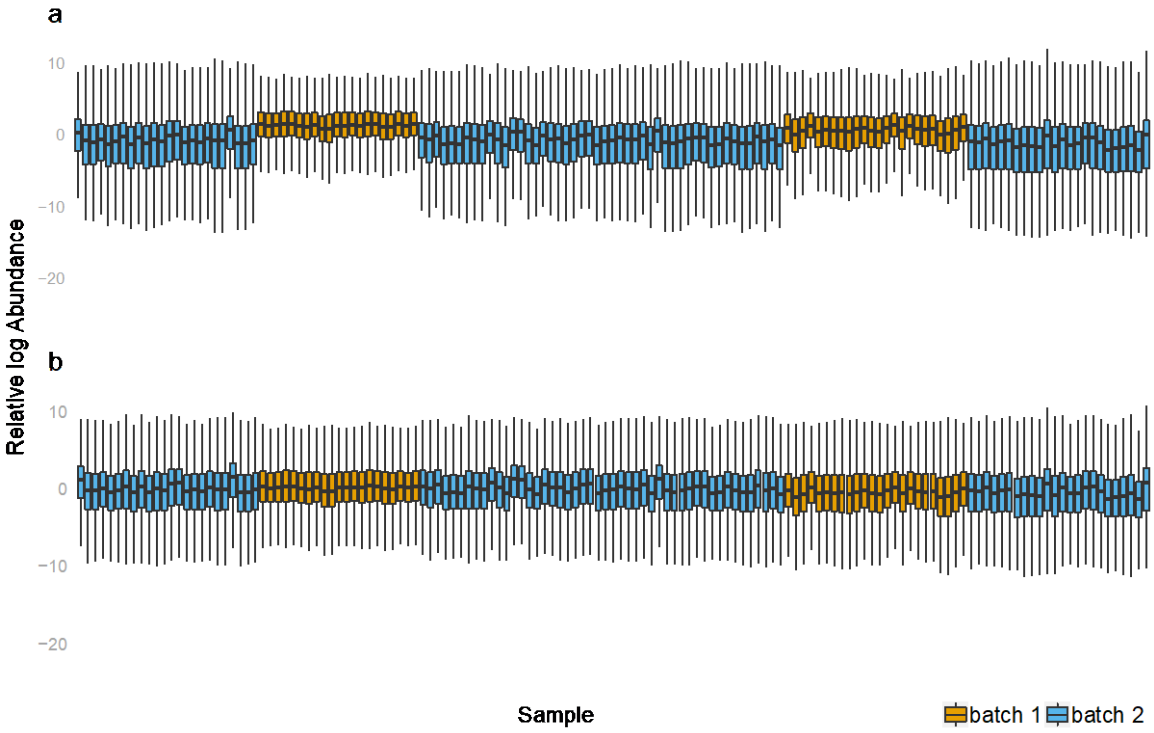


Figure 5.5 Relative log abundance plots of spectral features across all samples in the raw, unadjusted data (a) and in the batch corrected data (b).

A PCA was performed on the batch-corrected heat-stressed samples to examine the remaining metabolome variation and ensure that biological variation was retained. The first five PCs (Figure 5.6) accounted for 66.8% of the variation and pairwise comparisons of these display, in greater detail, the variation that was captured by each PC. The variation captured by the positive component of PC1 (30.9%) was strongly associated with heat-stressed samples collected at day 15, while the negative component of PC1 captured some unknown variation present in one sample from day 1 and one from day 10. The positive component of PC2 captures much of the variation associated with day 12 samples, while the negative component of PC2 captures unknown variation associated with two samples from day 15. Combined, PC1 and PC2 appear to capture biological variation associated with day 12 and day 15 heat-stressed samples along with some unknown variation. The positive component of PC3 captured the variation associated with days 1 and 8, while the negative component captured the variation associated with batch one samples (days 5 and 12). The plot of PC2 (17.6%) vs PC3 (8.4%) explains the variation of samples collected on days 1, 8, and 12 reasonably well. The variation captured by the positive component of PC4 was most strongly

associated with the variation of day 5 samples, while the negative component also captured some variation within day 12 samples. Finally, PC5 (4.8%) appears to have captured unknown variation not associated with treatment.

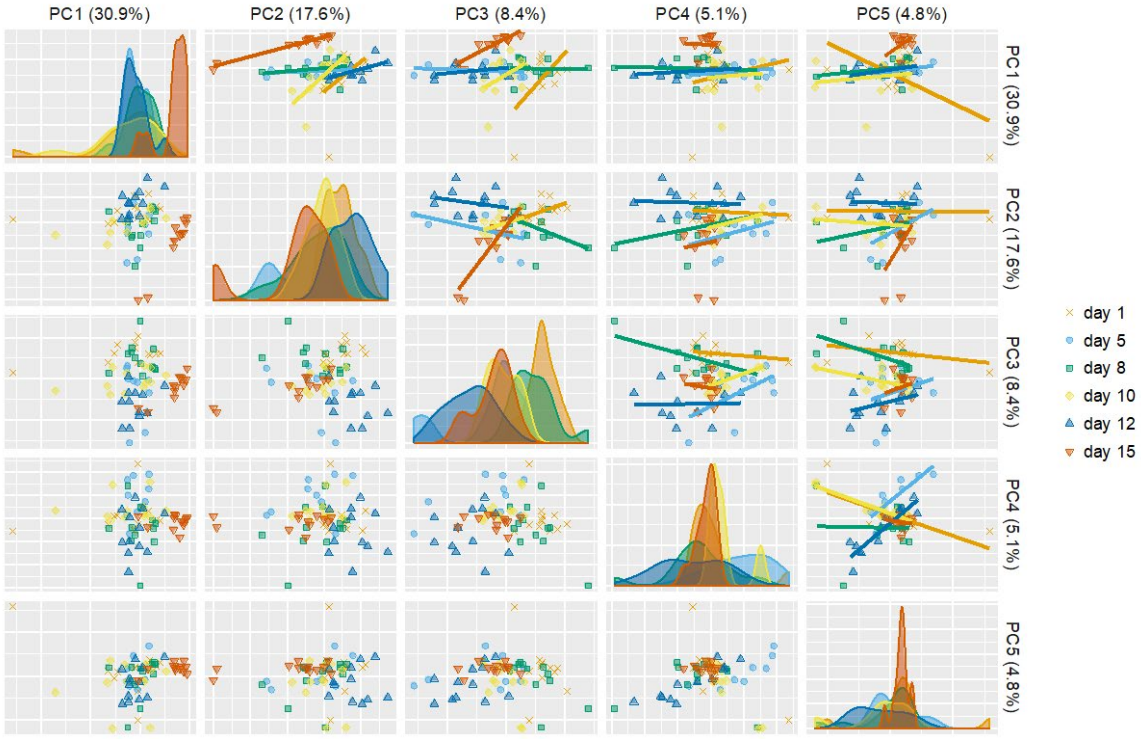


Figure 5.6 Composite pairwise Principal Component Analysis (PCA) scores plot of the corrected spectral data. The bottom left section of the plot shows the PCA scores plots for each pairwise comparison from PC1 to PC5. The plots lying on the diagonal from top left to bottom right are density plots of each PC for each day. The top right half of the plot shows the same PCA scores plots as the bottom left (axes reversed) with the addition of a linear smoother for each day.

5.3.3 Model Accuracy

Model construction utilised repeated, ten-fold cross-validation to tune the optimal number of variables to be considered at each split of the RF tree (the *mtry* value). The optimal model had an *mtry* value of 2113, a RMSE of prediction of 0.0371, a Q^2_{CV} of 0.9721 and a mean absolute error (MAE) of 0.0269. As the RMSE is in the same units as the F_v/F_m, the cross-validation results suggest the model can predict F_v/F_m within ±0.0371 of the measured F_v/F_m value.

The validity of the final model was evaluated by predicting 12 test samples that were not included in the construction of the model (Table 5.1). The model predicted the F_v/F_m of the test samples to ±0.0406 of their mean measured values (RMSE = 0.0406, Q^2 = 0.9684, MAE = 0.0290). The difference between the RMSE obtained during cross-validation and the

RMSE obtained from the prediction of the test samples was 0.0035. The prediction accuracy of the unseen test samples varied according to the observed F_v/F_m. For example, F_v/F_m values of two day 8 samples, having an average measured value of 0.495, were both predicted to be higher at ~0.58; an error that is more than twice the RMSE (Shaded rows in Table 5.1). Removing the prediction results for day 8 samples, the RMSE of prediction for the test samples reduced to 0.0216 and the Q^2 increased to 0.9962.

Table 5.1 Observed and predicted maximum quantum yield of photosystem II (F_v/F_m) values for unseen model validation samples. Shaded rows highlight samples with the greatest prediction error.

Day	Treatment	Observed F_v/F_m	Predicted F_v/F_m
day 1	Control	0.658	0.629
day 1	Control	0.658	0.651
day 1	Heated	0.648	0.635
day 1	Heated	0.648	0.642
day 5	Heated	0.625	0.627
day 8	Heated	0.495	0.583
day 8	Heated	0.495	0.581
day 10	Heated	0.636	0.629
day 10	Heated	0.34	0.356
day 12	Heated	0.636	0.599
day 12	Heated	0.161	0.201
day 15	Heated	0	0.016

Analysis of the 387 predictions made by the model during cross-validation provided insight into the increased prediction error observed for the day 8 test samples (Figure 5.7). Boxplots of the model predictions for each of the mean F_v/F_m values (Figure 5.7) show that day 8 samples (those having an F_v/F_m of 0.495) had the greatest interquartile range with the least distance from optimal F_v/F_m values of ~0.6. In other words, the F_v/F_m of day 8 samples may represent a borderline F_v/F_m that is neither "healthy" nor "unhealthy", somewhat explaining the difficulty in predicting these samples. In contrast, there was a larger distance between the predictions of samples having an observed mean F_v/F_m below ~0.4. Predictions of mean F_v/F_m close to the control values (~0.65) had much tighter distributions, highlighting the bias arising from the greater proportion of samples with an optimal F_v/F_m (more than half, i.e. all control samples and heat-stressed samples prior to day 7, had an F_v/F_m above 0.60; see Figure 5.3).

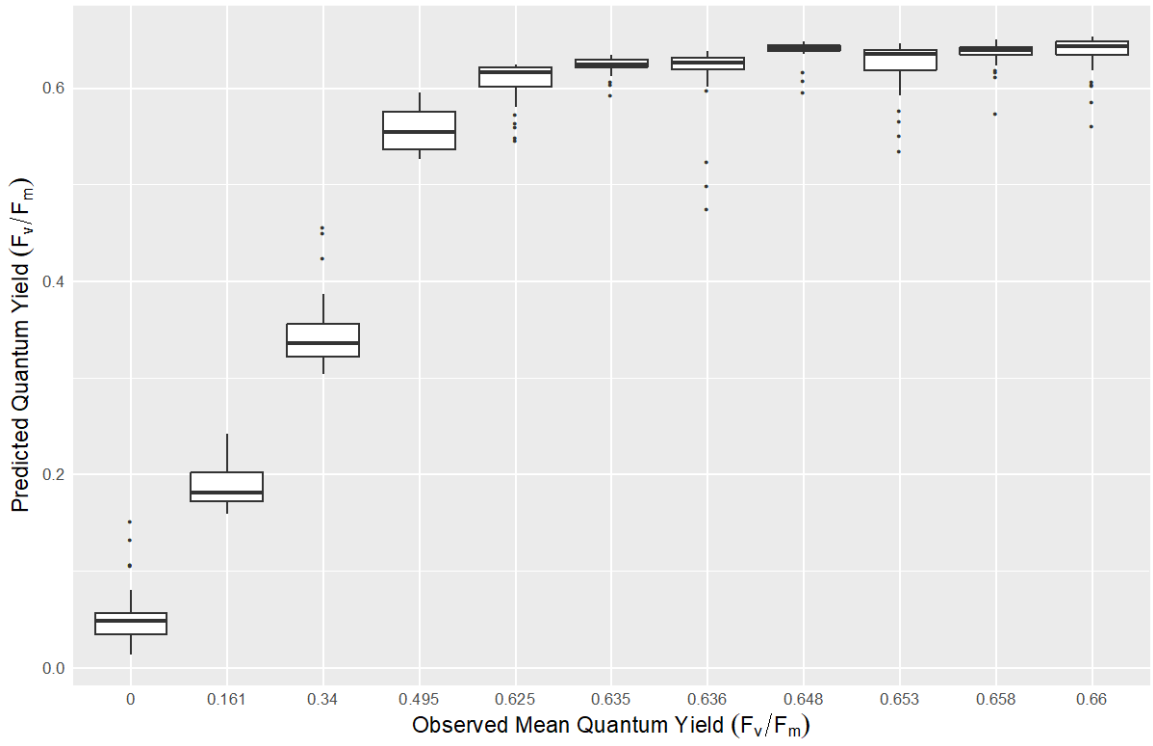


Figure 5.7 Boxplots of predicted maximum quantum yield of photosystem II (F_v/F_m) for each mean measured F_v/F_m made during cross-validation of the final model.

5.3.4 Important variables

To identify potential molecular biomarkers of coral functional status, the twenty most important spectral features driving prediction were identified by feature permutation (Figure 5.8). The most important feature driving prediction of F_v/F_m was m/z 222, which produced a mean decrease in accuracy of 14% after permutation. The next three most important variables, m/z 593, m/z 339 and m/z 817 produced mean decreases in accuracy of 11.4%, 8.3% and 7.6%, respectively. Of the 20 important variables, five were in a narrow m/z range between m/z 208-259.

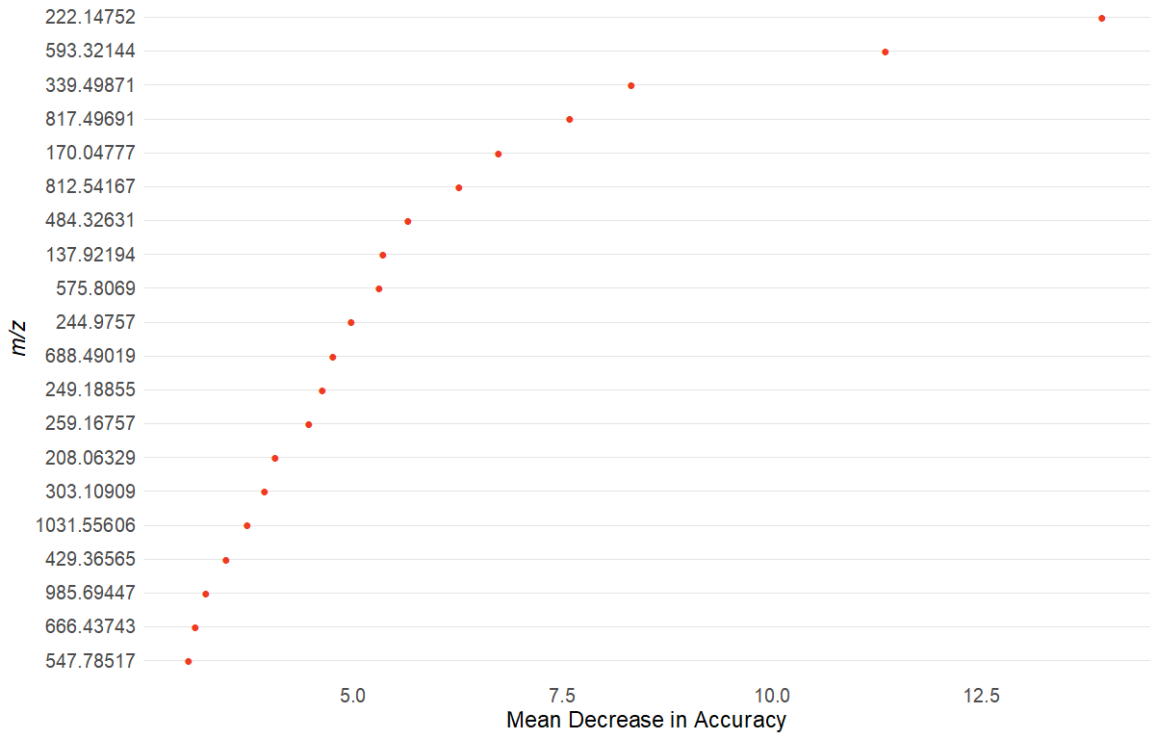


Figure 5.8 Important variables driving the prediction of the maximum quantum yield of photosystem II identified by feature permutation.

The top 20 ranked spectral features (Figure 5.8) and their corresponding adducts or precursor ions were compared with the coral research literature (Gordon & Motti, 2020a) to provide level 2 putative annotations where possible (Sumner *et al.*, 2007). Three spectral features having a similar mass (Δppm ≤ 50 ppm) were matched with four compounds (Table 5.2). The first, m/z 249, had a monoisotopic mass within 50 ppm of dihomomontiporyne H, first isolated from the *Montipora* genus of corals (Alam *et al.*, 2002). At elevated temperatures, the intensity of m/z 249 was lower than controls on days 8 and 15 (Figure 5.9) and statistically significant differences in the mean intensity were observed over time (Krukal-Wallis, χ^2 (5) = 36.985, $P < 0.001$). The second feature, m/z 170, was identified as a potential potassium adduct within 50 ppm of two betaines, β-alanine betaine and alanine betaine (131.095 Da), which have been isolated from ten different species of *Scleractinian* coral (Hill *et al.*, 2010). At elevated temperatures, the measured intensity of m/z 170 was notably higher than controls from day 10 onwards (Figure 5.9) and statistically significant differences in the mean intensity were observed over time (Krukal-Wallis, χ^2 (5) = 24.601, $P < 0.001$). The final feature, m/z 985, was annotated as a potential sodium adduct and dimeric ion within 50 ppm of the C_{16} phosphocoline, (or C_{16} variant of lyso platelet activating factor) lyso-PAF-C_{16} (481.353 Da); its intensity at elevated temperature was higher than days 5 and 8 controls

(Figure 5.9) and statistically significant differences in the mean intensity were observed over time (Krukal-Wallis, χ^2 (5) = 22.256, $P < 0.001$).

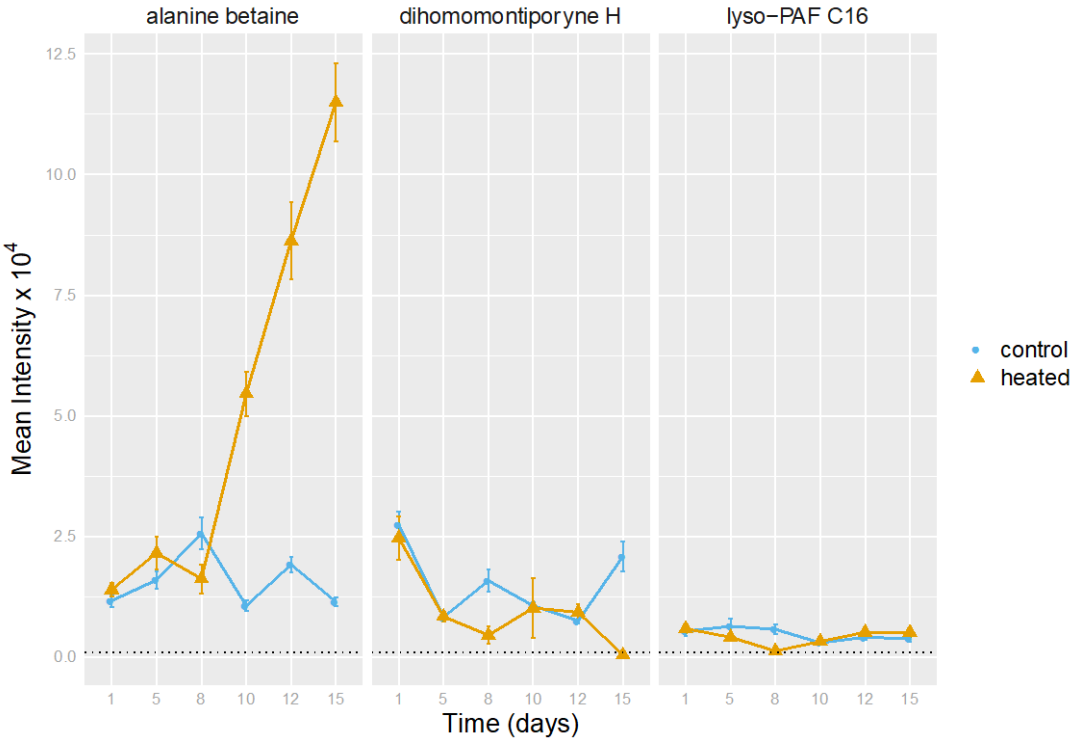


Figure 5.9 The *m/z* ion intensity each day (mean ± s.e) for three putatively annotated spectral features. The dotted horizontal line represents the limit of detection.

Table 5.2 Spectral features having a similar mass (Δppm < 50 ppm) with those reported in the coral research literature

Detected *m/z*	Detected ion	Literature *m/z*	Literature Exact Mass	Literature Ion	Molecular Formula	Mass Error (ppm)	Compound	Taxon	Reference
249.18855	$[M+H]^+$	271	248.17763	$[M+Na]^+$	$C_{16}H_{24}O_2$	13	Dihomomontiporyne H	*Montipora*	(Alam *et al.*, 2002)
170.04777	$[M+K]^+$	NA	131.094635	NA	$C_6H_{13}NO_2$	28	β-alanine betaine	*Scleractinia*	(Hill *et al.*, 2010)
170.04777	$[M+K]^+$	NA	131.094635	NA	$C_6H_{13}NO_2$	28	Alanine betaine	*Scleractinia*	(Hill *et al.*, 2010)
985.69447	$[2M+Na]^+$	482.36	481.35321	$[M+H]^+$	$C_{24}H_{52}NO_6P$	4	Lyso-PAF C_{16}	*Scleractinia*	(Quinn *et al.*, 2016)

5.4 Discussion

F_v/F_m values are continuous and quantitative model descriptors of coral health (Fitt *et al.*, 2001, Warner *et al.*, 1996, Jones *et al.*, 1998, Okamoto *et al.*, 2005). This study applied an LC-MS metabolomics approach and tree-based modelling to predict the F_v/F_m values of *A. aspera* exposed to an elevated temperature regime. Coral nubbins subjected to a ten-day elevated temperature regime experienced significant disruptions to their functional state and a ~50% decline in F_v/F_m. The variation arising from batch effects in the raw data were successfully removed while quantifying the loss of biological variation that is inherent in such normalisation techniques. The final model predicted the F_v/F_m RMSE during cross-validation to within ±0.0371 of the observed values, while the unseen test samples were predicted to within ±0.0406 of their observed values. Twenty spectral features driving the prediction of F_v/F_m were identified, three of which were putatively identified after a thorough search of the coral research literature. This study confirms that an LC-MS-based metabolomics approach can be successfully employed to predict a quantitative measure of coral health with meaningful precision, while simultaneously providing valuable information about temperature induced changes in the *A. aspera* metabolome.

5.4.1 The physiological response of *A. aspera* to elevated temperature

Thermal stress significantly affected symbiont physiology after eight days of incremental increases in seawater temperature. Symbiont cells in thermally stressed corals experienced significant declines in cell density (Fig. 5.4a) and F_v/F_m (Fig 5.3) compared to controls, which is consistent with declines in coral health and symbiosis breakdown established by several previous studies (Lesser, 1996, Lesser, 1997, Warner *et al.*, 1999, Dunn *et al.*, 2007). In contrast, chl-*a* negatively correlated with symbiont cell density (Figure 5.4b), which has been observed previously in *A. aspera* (Gierz *et al.*, 2016, Ogawa *et al.*, 2013), along with other coral species such as *Montastrea annularis*, *Seriatopora hystrix* (Fitt *et al.*, 1993a) and *A. formosa* (Jones, 1997). The negative correlation of chl-*a* with symbiont cell density runs counterintuitive to our understanding of coral bleaching and the consensus that pigment concentrations decline with thermal stress and cell density (Dove *et al.*, 2006, Fitt *et al.*, 2001). It could be that the remaining algal cells increase chlorophyll production to maintain overall photosynthetic output, both requiring and resulting in increased nutrient production per cell. A number of arguments have been put forward

that may explain this increase in symbiont pigment concentration, including that: 1) *Symbiodiniaceae* with a higher thermal tolerance may repopulate the coral after bleaching (Ogawa *et al.*, 2013); 2) repackaging of chl may result in pigment-protein complexes that absorb light more strongly (Gierz *et al.*, 2016, Bissett *et al.*, 1997); 3) the remaining symbionts experience a more nutrient rich intracellular environment than those that occupy corals with higher symbiont cell densities (Fitt *et al.*, 1993b) or; 4) seasonal fluctuations and environmental factors are responsible (Brown *et al.*, 1999b). The cause of elevated chl over the course of this experiment could be explained by any or all of these previous arguments; however, there is significant evidence that high nutrient levels correlate with increased chl concentrations (Hoegh-Guldberg & Jason-Smith, 1989, Muscatine *et al.*, 1989, Stambler *et al.*, 1994) and the argument for elevated nutrient enrichment at lower cell densities first put forward by Fitt et. al. (1993b) deserves more attention.

5.4.2 Batch effects and their implication for long-term metabolomics-based monitoring

Analysing samples in two separate batches resulted in pronounced analytical variation in the metabolic profiles. These batch effects are an unavoidable phenomenon in long-term mass spectrometry experiments involving large numbers of samples because of the high sensitivity of mass spectrometers, and the ions they detect. The major variations between batches one and two were associated with their date of analysis, eluent batch and the column usage time. As far as possible, all other variables were kept constant wherever possible: for example, instrument, column, operator, extraction and handling procedures were consistent across both batches.

The most common methods used to minimise or eliminate batch effects involve the addition of internal standards and/or preparing and analysing PBQC samples at regular intervals and applying chemometric methods to normalise the data (Dunn *et al.*, 2011, Hendriks *et al.*, 2011, Wehrens *et al.*, 2016). Unfortunately, there are several risks involved in removing batch related variation from the data using these techniques. Internal standards, for example, can change the physical characteristics of the sample, mask co-eluting compounds or suppress ionisation; as a result, their inclusion in untargeted metabolomics studies, such as this one, is often avoided (Wehrens *et al.*, 2016). While the use of internal standards is somewhat contentious, the use of PBQCs to remove batch effects is still considered good practice as they provide a means to monitor analytical

variation in long-term studies. As they do not directly affect the sample, their inclusion will, at worst, increase throughput time of each batch and, potentially, the number of batches required. A major consideration with PBQCs, however, is how they are utilised for removing batch effects. Traditionally, the workflow involves modelling the unwanted variation in the PBQCs and applying a correction to the sample data. Unfortunately, this process is not absolute and relies on the probability that batch effects can be discerned from genuine biological signals, inevitably resulting in the removal of some biologically relevant variation (Wehrens *et al.*, 2016, Oytam *et al.*, 2016).

As the first and foremost concern of this study was to maintain the biologically relevant variation while detecting the highest possible number of spectral features, internal standards were avoided and the *Harman* technique: a recently developed risk-conscious approach to batch variation removal described by Oytam et. al. (2016) was employed. The technique, which was originally developed for genome and proteome data, was shown to be superior to the leading genome batch removal technique while preserving much of the biologically relevant information (Oytam *et al.*, 2016). The technique separates data into its principal components and any variance that is associated with the batch information is then removed. The strength of the Harman technique is that the probability of removing variation associated with the biological factors, which are supplied to the algorithm, can be quantified by means of a confidence limit.

In this study, the batch variation was prominent in the RLA plot of the unadjusted data (Figure 5.5a). The Harman technique was employed to remove this variation with a confidence limit of 0.99 (i.e. the probability of removing genuine biological signal was no greater than 1%). Some of the remaining batch variation could be observed in PC2 of the corrected data (Figure 5.6): however, decreasing the confidence limit of the Harman algorithm risked removing too much biological variation, which was naturally low due to the nature of the experiment. Heat-stressed samples collected on days 1 and 5, for example, exhibited "healthy" phenotypes, both having an F_v/F_m similar to that of controls. As a result, the stress-related variation in their metabolomic profiles was also low in comparison with control samples. In contrast, samples collected on days 8 to 15 sustained a linear decline in their F_v/F_m and an increase in the stress-induced variation, as observed in their metabolomic profiles. The degree of stress-related variation is most obvious in the PCA plots, where treated samples collected on days 12 and 15, when seawater temperatures were at their highest, occupied more unique scores space than samples collected at earlier timepoints (Figure 5.6; PC1 and PC2).

The co-removal of biological variation with batch variation was of particular concern for samples collected on days 8 and 10, which had significant, although relatively small, declines in their F_v/F_m and relatively low variation in their metabolic profiles compared to samples exposed to longer periods of elevated temperature. For these samples, increasing the probability of removing biologically relevant variation by decreasing the confidence limit of the Harman algorithm was deemed too risky, especially given the large amount of the batch variation that was removed with the confidence limit set at 0.99.

Batch-related variation is unavoidable if the technique described here is to be scaled up for long-term metabolomics biomonitoring. For current long-term metabolomics studies this issue has largely been resolved with the judicious use and preparation of quality control samples, detailed experimental designs and the use of chemometrics to model the unwanted variation (Dunn *et al.*, 2011, Wehrens *et al.*, 2016, De Livera *et al.*, 2015, Oytam *et al.*, 2016). It should be noted that, in this study, only the treatment labels (control and heat-stressed samples) were supplied to the Harman algorithm for designating biological variation. It is therefore conceivable, that the algorithm may have removed other biologically relevant variation such as the length of exposure. Although the loss of all biologically-relevant variation was not definitively quantified in this study, the Harman technique removed the majority of variation that was not associated with either treatment group while also providing a means to quantify, and thus minimise, the probability of losing variation that was associated with each treatment. Whether this technique will translate to longer-term studies, where seasonal and diurnal variations must also be accounted for, has yet to be verified and should be explored in more detail.

5.4.3 F_v/F_m as an indicator of healthy or stressed coral phenotypes

The difficulty of linking the F_v/F_m value with a coral's phenotype remains a major consideration in the use of F_v/F_m to monitor coral health, especially in the long term. While the connection between the functional state of the coral holobiont and the F_v/F_m of its algal symbionts has been firmly established over the past few decades of coral photophysiology research (Warner *et al.*, 1996, Fitt *et al.*, 2001, Ralph *et al.*, 2015), other measures such as cell density and chl concentration are still utilised to provide context to the F_v/F_m measurements. For the purpose of this study, an F_v/F_m range representing a healthy phenotype was extracted from the vast amount of PAM fluorometry results reported in the coral research literature.

One of the earliest investigations into coral photophysiology reported that peak photosynthetic efficiency would produce an F_v/F_m between 0.5-0.7 in most corals depending on environmental conditions (Fitt *et al.*, 2001), and this generally holds true to this day; however, this window may be even narrower for shallow water corals. Studies of *A. aspera*, for example, have recorded F_v/F_m values in "healthy" control corals between 0.6 and 0.7 units (Hillyer *et al.*, 2018, Hillyer *et al.*, 2017b, Ogawa *et al.*, 2013, Gierz *et al.*, 2016, Rosic *et al.*, 2014, Middlebrook *et al.*, 2008). Indeed, the work by Okamoto et. al. determined the average "healthy" F_v/F_m of 68 different species was 0.664 ($\sigma = 0.029$) (Okamoto *et al.*, 2005). While there is no doubt that a long-term biomonitoring program must have adequate sampling to substantiate a baseline F_v/F_m, current research suggests that a healthy F_v/F_m phenotype of shallow water corals may be more in the region of 0.60–0.70.

Healthy *A. aspera* nubbins (controls) were found to have an average F_v/F_m of ~0.65. By day 10 of heat stress there was an obvious impact on coral health, where F_v/F_m had declined to an average of 0.34, corresponding to a significant decline in *Symbiodiniaceae* cell density. While there was not a significant decline in cell density on day 8 of the temperature regime, knowledge that corals were experiencing above average seawater temperatures (~32 °C), and that cell density declined by a further ~15% over the subsequent two days, supports the a priori assumption that the reduced F_v/F_m readings for day 8 (~0.5) were indicators of dysfunctional corals in this experiment. At the very least, the ~15% reduction in measured F_v/F_m on day 8 would be indicative of at-risk corals that warrant closer inspection. With this in mind, a ~15% reduction in F_v/F_m may represent a threshold where shallow water corals shift from a healthy to stressed phenotype.

5.4.4 Model Performance

RF modelling of coral metabolomic profiles accurately predicted the F_v/F_m of both healthy and heat-stressed corals at varying degrees of suboptimal F_v/F_m values. With the exception of samples collected on day 8, the F_v/F_m of all unseen test samples were predicted to within ±0.0406 of their observed values. This result is well within the window of each of the suboptimal F_v/F_m values recorded, as is evident in the boxplots of the cross-validation predictions, which show minimal overlap between the range of predictions (Figure 5.7). The same boxplots also highlight the bias that was introduced into the model, where samples having an F_v/F_m close to control values (~0.65) were

more accurately predicted. This bias may be responsible for the greater prediction error of the two unseen test samples collected on day 8 (F_v/F_m of 0.495; Table 5.1), which were both predicted to be high and with twice the RMSE. It was suspected that the greater proportion of control samples, along with those less-exposed samples exhibiting a healthy F_v/F_m phenotype, were responsible for this bias; however, further analysis showed only a small increase in the RMSE of test samples when controls were removed from the model. As a result, removing the control samples from the model was deemed unnecessary and more likely to remove information that could be utilised for prediction of healthy coral phenotypes. The RMSE of prediction represents a 6.3% error margin for the control corals, well within the proposed ~15% threshold that delineates healthy corals phenotypes from stressed phenotypes. While more research on wild coral populations across a variety of regions is required, these results provide new and promising evidence that the information contained within metabolic profiles can be used to accurately describe healthy and stressed coral phenotypes as defined by their F_v/F_m.

5.4.5 Putative molecular biomarkers driving the prediction of F_v/F_m

An examination of the coral literature provided level 2 putative annotations (Sumner *et al.*, 2007) for three of the 20 most influential spectral features driving F_v/F_m prediction (Table 5.2) The strongest driver of F_v/F_m was m/z 170.04777 (m/z 170), which was putatively annotated as a potassium adduct of alanine betaine or its isomer, β-alanine betaine. Betaines are abundant metabolites commonly found in marine algae (Kato *et al.*, 1996, Blunden *et al.*, 1992, Blunden *et al.*, 1986), *Symbiodiniaceae* (Nakamura *et al.*, 1998, Yancey *et al.*, 2009, Leblond *et al.*, 2015) and a number of corals and giant clams (Hill *et al.*, 2010, Hill *et al.*, 2017, Yancey *et al.*, 2009). Betaines act as osmolytes in animals (Hill *et al.*, 2010, Yancey, 2005, Anthoni *et al.*, 1991) and are widely known to reduce the impacts of cellular stress by stabilising proteins and cellular membranes (Hill *et al.*, 2017, Rhodes & Hanson, 1993). Of note, is their role in protecting PSII from high irradiance and temperature, which has been firmly established in green algae and higher plants (Papageorgiou & Murata, 1995, Schiller & Dau, 2000).

In corals, betaines have been measured at 30-94% higher concentrations when exposed to high irradiance (Hill *et al.*, 2010). Under very low irradiance, i.e. in corals experiencing dark stress, up-regulation of the *betaine-homocysteine S-methyltransferase 1* (BHMT) gene has also been observed (DeSalvo *et al.*, 2012). Up-regulation of BMHT may be indicative of elevated betaine concentrations and thus further evidence of their defensive role in corals

experiencing PSII stress (DeSalvo *et al.*, 2012). The intensity of the ion at m/z 170 had a positive correlation with F_v/F_m (Figure 5.9), which is consistent with the high irradiance behaviour reported by Hill *et al.* (2010), increasing confidence in the putative annotation of m/z 170 as alanine betaine. While future research is required to provide an unambiguous identification, this study provides tentative evidence of increased coral betaine production corresponding with a concomitant temperature induced reduction in symbiont F_v/F_m, and of the photoprotective role of betaines in corals.

The next most influential spectral feature, m/z 985.69447, was putatively annotated as lyso-PAF-C_{16}; a C_{16} phospholipid involved in the immune response of a number of animals where it acts as a signalling molecule to induce inflammation (Quinn *et al.*, 2016, Camussi & Brentjens, 1987, Yost *et al.*, 2010). Parkinson *et al.* (2014) reported the first observation of lyso-PAF-C_{16} in corals during an investigation of the functional diversity of coral and algal individuals. Examination of the metabolomes of symbiotic and aposymbiotic polyps of *Astrangia poculata* colonies found lyso-PAF-C_{16} to be more abundant in aposymbiotic polyps than symbiotic ones, highlighting the importance of accounting for intraspecific diversity in experimental designs.

A second observation of lyso-PAF-C_{16} in corals came from a methods-based study of molecular networking for metabolite identification in complex metabolite mixtures (Garg *et al.*, 2015 6389). Co-networking of coral (*Montastraea annularis*) extracts with diseased lung extracts, along with data from several spectral libraries, enabled the rapid identification of lyso-PAF-C_{16} and other PAFs in coral extracts, highlighting the utility of molecular networking as an organisational tool of tandem MS data.

PAF-C_{16}, lyso-PAF-C_{16} and lyso-PAF-C_{18} were identified in a 2016 study using a LC-MS-based untargeted metabolomics approach and gene expression analysis to identify major components of the coral lipidome (Quinn *et al.*, 2016). PAF-C_{16} concentration increased in four different coral genera in response to the competitive interactions with algae along with the gene that encodes the protein responsible for converting lyso-PAF-C_{16} to PAF-C_{16}. This was the first evidence that PAF-C_{16} was involved in the inflammation response of stressed corals.

The ion at m/z 985 (putatively annotated as lyso-PAF-C_{16}) was found to be a strong predictor of symbiont F_v/F_m. Initially, its abundance decreased from days 1-8 (Figure 5.9) in response to increases in temperature (Figure 5.1). But, by day 10 its abundance began to increase and continued to do so for the remainder of the experiment, returning

close to its original baseline level at day 1 (Figure 5.9). This behaviour coincides with the known action of the PAF signalling cascade in mammals (Yost *et al.*, 2010, Camussi & Brentjens, 1987), invertebrates (Sugiura *et al.*, 1992) and corals (Quinn *et al.*, 2016), where lyso-PAF is converted to PAF by the lyso-PAF acetyltransferase (*lyso-PAF-AT*) protein, initiating an inflammatory response. Conversely, conversion of PAF back to lyso-PAF by the PAF acetylhydrolase (*PAF-AH*) protein is known to reduce the inflammatory response (Yost *et al.*, 2010) in mammals. This interconversion between lyso-PAF and PAF is tightly controlled; however, the inflammatory response in mammals can be uncontrolled and, in many cases, fatal (sepsis and anaphylaxis for example).

Without monitoring the concentration of both PAF and lyso-PAF, it is difficult to explain with certainty the return of m/z 985 intensity back to baseline levels from day 10 onwards. Logically, *PAF-AH* should be up-regulated, converting PAF to lyso-PAF, resulting in a reduction of the inflammation response; however, this scenario runs counterintuitive to the expected increase of the inflammatory response in organisms under continued physiological stress. One possible explanation may be that lyso-PAF is being produced, via the phospholipid remodelling pathway (Uemura *et al.*, 1991), at a rate that exceeds the *lyso-PAF-AT* mediated conversion of lyso-PAF to PAF. This explanation, however, would also suggest a reduction in the inflammatory response at a time when the coral was most physiologically stressed, since lyso-PAF is known to reduce the inflammatory response in mammals (Yost *et al.*, 2010). The latter would appear more likely given the continued thermal stress and further decreases in F_v/F_m observed in this experiment. Also, the lyso-PAF/PAF mediated inflammatory response in humans and other animals is known to be uncontrolled and often lethal under certain circumstances (e.g. in cases of sepsis and anaphylaxis) (Yost *et al.*, 2010). Thus, the reduction in abundance of m/z 985 from days 1-8 may represent an initial inflammatory response to thermal stress involving the conversion of lyso-PAF-C_{16} to PAF-C_{16}. With continued stress, a critical point may have been reached by day 10, where m/z 985 biosynthesis is increased due to a dysregulation of the PAF signalling cascade—a common feature of lethal inflammatory responses in humans (Yost *et al.*, 2010).

The putative annotation of m/z 985 as lyso-PAF-C_{16}, and its role as a predictor of F_v/F_m in response to thermal stress, provides supporting evidence for the lyso-PAF/PAF inflammatory response first observed in corals by Quinn et. al (2016). However, the $M+H^+$ ion (482.36 Da) observed by Quinn et. al (2016) was not in the top 20 mass spectral features driving model classification, despite being analysed with similar

instruments. This could be due to matrix effects or ion suppression that are likely to be more prevalent when analysing crude coral extracts. Nevertheless, given that PAF-mediated events are known to disrupt defensive mechanisms that predispose the host to further injury and/or infection (Yost *et al.*, 2010), further research into this mechanism in corals should be prioritised, as it presents an opportunity to intervene in the PAF signalling cascade and so reduce the lethality of heat stress in corals.

The final important spectral feature found in the coral research literature was m/z 249.18855, putatively annotated as the acetylenic toxin, dihomomontiporyne H. The original isolation and structure elucidation of dihomomontiporyne H from *Montipora* corals was performed with the aim of finding new cytotoxic natural products that may have potential as chemotherapy drugs (Alam *et al.*, 2002 6322). Against five human tumour cell lines, dihomomontiporyne H was found to have only moderate activity compared to other acetylenes isolated from *Montipora* sp. (Alam *et al.*, 2002 6322, Alam *et al.*, 2001 6321).

In a study of the ecological roles of dihomomontiporyne H, Higa et. al. (1990) isolated four related acetylenes from *Montipora* sp. that exhibited ichthyotoxicity and inhibited the growth of some bacteria and fungi. They may act as sperm attractants in *Montipora digitata*, a hermaphroditic coral that releases eggs and sperms in a single bundle (Coll *et al.*, 1994). In a certain ratio, acetylenic alcohols are more effective in attracting sperm from *M. digitata* than from other species of *Montipora*, which may reduce the chances of hybridisation between different *Montipora* species (Coll *et al.*, 1994). Both the tissue and eggs of Montiporid corals are known to harbour compounds that inhibit photosynthesis in endogenous and exogenous *Symbiodiniaceae* within minutes (Hagedorn *et al.*, 2015, Hagedorn *et al.*, 2010). Crushed eggs of *M. capitata* reduced the F_v/F_m of exogenous *Symbiodiniaceae* by up to 44% within 40 minutes, while the F_v/F_m of endogenous *Symbiodiniaceae* reduced to zero within 5 mins of crushing the *M. capitata* sperm bundle. While the ecological role of these toxins is still unclear, the literature suggests they play a role in defending the coral from other coral species, providing a chemical advantage to corals of the *Montipora* genus (Gunthorpe & Cameron, 1990, Hagedorn *et al.*, 2015).

This research provides tentative evidence of dihomomontiporyne H in a coral outside of the *Montipora* genus. Given their ability to inhibit photosynthesis, acetylenic toxins may play an important role in the breakdown of PSII during thermal stress if the putative annotation provided here proves to be correct. Clearly, future research is essential

investigate this phenomenon further and fully resolve the ecological role of acetylenic toxins in corals.

5.5 Conclusion

This research presents a reliable metabolomics-based method to accurately predict photoinhibition (i.e. F_v/F_m) in corals exposed to elevated temperatures. RF modelling of the LC-MS profiles of *A. aspera* identified a number of spectral features that were strongly associated with the F_v/F_m of *A. aspera* under varying degrees of thermal stress. A search of these spectral features against the coral research literature putatively identified three metabolites, all with potentially important roles in symbiont photophysiology and the breakdown of the coral-algal symbiosis. The successful implementation of the Harman algorithm, and its removal of batch related variation, has overcome a considerable impediment to the future implementation of long-term monitoring of coral metabolomes. While further research is required to investigate the utility of this approach beyond that of controlled laboratory conditions, these findings clearly highlight the potential of metabolic profiling to provide more informative, early indicators of coral health that may enhance the ability of reef managers to prioritise interventions and avoid undesirable changes in the reef ecosystem.

Chapter 6
Revealing Taxonomic and Diel Variation Through Untargeted Metabolomics

This chapter is in preparation for submission: Benjamin R. Gordon, William Leggat, Ute Roessner, Danielle Gordon, Adrian Lutz and Cherie A. Motti

Benjamin Gordon wrote this chapter, designed and conducted the research, and analysed the data. Co-authors provided intellectual, experimental and editorial support.

6.1 Introduction

Metabolic profiles of wild corals are highly variable and impacted by the vast array of environmental conditions to which they are exposed and the large number of organisms that constitute the holobiont. Variation in metabolic profiles can be captured using metabolomics – a spectral description of a sample's measurable biochemistry – and monitored to assess the functional status and/or health of an organism. Indeed, metabolic biomonitoring of agricultural products such as, olive oils (Goodacre *et al.*, 2002), wine (Cuadros-Inostroza *et al.*, 2010, Ali *et al.*, 2011) and beef (Jung *et al.*, 2010) has been in use for well over a decade, resulting in a thorough understanding of the conditions needed to control organismal behaviour and metabolic diversity, enabling the production of goods of specific quality.

Robust biomonitoring of coral health must be representative of the coral population and account for natural or unknown variation. At small scales, a thorough sampling regime of a variety of corals provides an adequate representation of the coral reef population; however, due to the sheer number of samples required across multiple species, geographies and time scales, this approach does not scale well. A feasible solution may be found through the adoption of a sentinel species that is sensitive to changes in the environment while also being robust to localised variations and representative of a variety of species across a variety of landscapes. Restricting a metabolomics-based biomonitoring program to include only one, or even a few, important species would allow for greater sample replication and consequently, more accurate and robust predictions of coral and reef health.

Metabolome research of wild corals is still in its infancy; however, a number of untargeted metabolic profiling studies have provided a glimpse into the metabolic response of key coral species to climate change-related stressors under controlled laboratory environments (see Chapter 2 for a review of the current coral metabolomics literature). Coral metabolic profiles are unique to their geographical location; for example, research of a single species of soft coral has shown that their sterol, terpene and *N*-containing compounds vary according to their geographical location, and by inference, their living environmental conditions (He *et al.*, 2014). A similar discovery was made in the deep-sea coral, *Callogorgia delta*, which was shown to have unique metabolic variation, not only across different geographic locations, but also between individual colonies at the same location (Vohsen *et al.*, 2019). Such intraspecific variation of coral metabolomes is

common, even within local environments. Betaines, for example, are as much as ~90% more abundant in corals that occupy areas of the reef that are high in irradiance (Hill *et al.*, 2010). As well as provenance and reef morphology, intraspecific variation is also caused by the genotypic diversity within a species; for example, untargeted metabolic profiling of three distinct genotypes of *Acropora cervicornis* revealed metabolic profile variation unique to each genotype (Lohr *et al.*, 2019). Interestingly, previous research had shown that the three genotypes had unique growth and stress tolerance phenotypes, which could be related to the observed metabolic variation. Clearly, a better understanding of the natural variation of corals is required to identify a sentinel coral species and instil confidence in the forecasts provided by metabolomics-based biomonitoring. To this end, this chapter explores the metabolic profiles of five different coral species and their variation over a single diel cycle.

In addition to understanding natural metabolome variation, effective and efficient sample collection and quenching will be a major consideration of a metabolomics-based biomonitoring program. This aspect of biomonitoring can be both time consuming and costly in terms of the logistics involved, yet it is crucial to the quality of the data and the decisions that arise from its analysis. In particular, the traditional approach of collecting coral nubbins and quenching them in liquid nitrogen is problematic if coral sampling is to occur over a wide landscape. This is due, in large part, to the logistics and specialist handling required for liquid nitrogen in the marine environment. Nevertheless, there are a number of viable alternatives that could be explored in more detail; for example, methanol has been widely employed as a quenching agent in cell culture metabolomics (de Koning & van Dam, 1992, Sáez & Lagunas, 1976, Mashego *et al.*, 2007, Bolten *et al.*, 2007, Canelas *et al.*, 2008) and examined as an alternative to liquid nitrogen for the in-field collection and quenching of plant samples (Maier *et al.*, 2010). Alternatively, the metabolites present at the corals surface (i.e. 0-5 cm) can also be sampled with minimal to no impact on the coral itself (Ochsenkühn *et al.*, 2018). In an effort to tailor the collection of wild corals to a metabolomics-based biomonitoring program, this chapter also assesses the suitability of methanol as a chemical-based alternative quenching agent of coral metabolomes.

6.2 Materials and methods

6.2.1 Sample collection and processing for quenching assessment

On the 26th March 2014 at the collection site on the reef flat of Heron Island, ten fragments of *A. aspera* were collected (under Great Barrier Marine Park Authority permit G13/36402.1) from a single colony and snap frozen in liquid nitrogen, while another ten from the same colony were placed into 20 mL glass scintillation vials (Sigma-Aldrich, Sydney, NSW, Australia) containing 15 mL of 100% LCMS grade methanol at ambient temperature, ~25 ℃ (Thermo Fisher Scientific, Scoresby, VIC, Australia). After one hour in liquid nitrogen, the snap-frozen samples were transferred to 20 mL glass scintillation vials containing 15 mL of 100% methanol. All twenty samples were transferred to -80 °C storage for approximately one month prior to analysis.

Samples were prepared for LCMS analysis by centrifuging a 1 mL aliquot of the methanol extract from each sample at 21,900 g for 5 mins to settle any particulate matter before transferring the supernatant to glass autosampler vials (Verex™ vials, Phenomenex, Lane Cove, NSW, Australia). PBQCs were prepared by combining 5 µL of each sample extract into a single pooled sample. Three PBQCs were prepared; one for each of the two different quenching treatments and one for all samples.

6.2.2 Sample collection and processing for the assessment of taxonomic and diel variation

To ensure the greatest amount of metabolic diel variation was captured, sample collection was conducted both before, and after, the light/dark transitions at dawn and dusk. Coral samples were collected at 0500, 1000, 1600 and 2100 h (under Great Barrier Marine Park Authority permit G13/36402.1) on the reef flat of Heron Island on the 21st April 2014 from five coral colonies comprising five different, visually identified, species: *Acropora aspera, Montipora digitata, Montipora aequituberculata, Pocillopora damicornis* and *Porities cylindrica* (5 replicates per colony per time point; n = 100). For ease of sampling, collections were timed so that they occurred ~2.5 h either side of the low tides, which were recorded at 0731 and 1905 h. At all collection times, corals were at a collection depth of approximately 30 – 60 cm.

Methanol quenching was employed as it was deemed to be a viable alternative to snap freezing using liquid nitrogen, as discussed in section 6.4.1. Small nubbins approximately

2 cm in length were excised from coral colonies and placed directly into 15 mL of 100% methanol at ~25 ℃ contained in 20 mL glass scintillation vials then stored at -80 °C for approximately one month prior to analysis. Samples and PBQCs were prepared for LCMS analysis as described in section 6.2.1. Six PBQCs were prepared: one for each species and one for all samples.

6.2.3 LCMS data acquisition

Reverse Phase chromatography was performed using an Agilent 1200 HPLC system (Santa Clara, CA, USA) consisting of a vacuum degasser, binary pump, thermostatic auto sampler and column compartment. Chromatography was conducted using the following conditions: a Zorbax Eclipse XDB-C18, 2.1 mm x 100 mm, 1.8 μm (Agilent, Australia) column; solvent (A) 0.1% formic acid in Milli-Q water and solvent (B) 0.1% formic acid in acetonitrile (ACN); a flow rate of 0.4 mL min^{-1} and; a column temperature of 40 °C. Samples were injected (10 μL) and eluted with a 10 min linear gradient from 5% (B) to 100% (B), followed by a 2 min hold at 100% (B), then returned to 5% (B) and re-equilibrated for 5 min (total time of 17 min). Samples were randomised to ensure analytical variation did not correlate with biological variation and a pooled biological quality control was run every tenth sample.

The mass spectrometer was an Agilent 6520 ESI-QTOF-MS (Santa Clara, CA, USA) with a dual spray ESI source operated in positive ion mode. The source conditions were: nebuliser pressure of 45 psi; gas temperature of 300 °C; drying gas flow of 10 L min^{-1}; capillary voltage of 4000 V and skimmer 65 V; fast polarity enabled. Measurements were performed in the extended dynamic range mode (m/z range of 100-3200; sampling rate 2 GHz); scan rate of 2.03 spectra/s collected as MS1 centroid data. The mass spectrometer was calibrated using Agilent's ES Tuning Mix ACN Solution, containing several calibration ions at mass values from 117 Da to 2,722 Da. PBQCs were used to monitor mass deviations of some ubiquitous features and analytical variation such as retention time shifts (typically < 0.1 min) and ionisation efficiency.

6.2.4 LCMS data pre-processing

Raw LCMS data was acquired from the instrument (Agilent MassHunter) then converted to mzXML format using ProteoWizard's MSConvert tool (v3.0.6585; Chambers *et al.*, 2012, Kessner *et al.*, 2008) with an absolute intensity threshold of 1000 counts. LCMS chromatograms were integrated and aligned in the R environment (v3.5.0) (R Core

Team, 2018) using the XCMS package (v3.4.2; Smith *et al.*, 2006, Tautenhahn *et al.*, 2008, Benton *et al.*, 2010). Feature detection was performed using the centWave method (Tautenhahn *et al.*, 2008) with the following modified parameters: ppm = 30, peakwidth = c(10, 60), mzdiff = -0.001, integrate = 1 and, prefilter = c(3, 1100). Retention time correction utilised the obiwarp method with binSize = 0.5 (Prince & Marcotte, 2006). Chromatographic peaks were grouped across samples using the peak density parameter with: bw = 5, binSize = 0.025 and minFraction = 0.5. Peak Filling was employed using the *fillChromPeaks()* function with default values.

Isotopes and adducts were annotated using the *annotate()* function within the CAMERA package (v1.38.1) (Kuhl *et al.*, 2011) with the following parameters: perfwhm = 0.7, cor_eic_th = 0.75, ppm = 10 and polarity = 'positive'. All M^{+1}, M^{+2}, M^{+3} and M^{+4} isotopes identified by the CAMERA package were removed from the resulting peak list.

Unreliable spectral features with 60% or more missing values across all classes were deleted entirely. Where a class had 80% or more missing values within any spectral feature, a random value between zero and the minimum intensity was imputed to represent a non-detect specific to that class. The remaining missing values were imputed using a random forest trained on the observed values (ntree = 100, mtry = $\sqrt{x}$), as described by Stekhoven and Buehimann (Stekhoven & Bühlmann, 2011).

The structure and overall variation of the raw data was assessed using principal components analysis (PCA). PCA was performed in the R environment (v3.5.0; R Core Team, 2018) by a singular value decomposition of the centred and scaled data. Boxplots of the relative log abundance (RLA) of each sample were employed to assess the variation in metabolite abundance between samples. The data was standardised in the R environment (v3.5.0; R Core Team, 2018) by subtracting the median from each log-transformed spectral feature before constructing the boxplots. The final, pre-processed, dataset had a total of 6542 features.

6.2.5 Assessment of variation between liquid nitrogen- and methanol-quenched samples

PCA and boxplots of the RLA of each sample were employed to assess metabolic variation as described in section 6.2.4. A Random Forests (RF) model was constructed in the R environment (v3.5.0; R Core Team, 2018) using the randomForest package (v4.6-14)(Liaw & Wiener, 2002) to more accurately model minor variations in spectral feature

intensity. The optimal number of variables randomly sampled as candidates at each split (the mtry value) was optimised using repeated k-fold cross-validation (six repeats; k = 5). The optimal model was the most accurate model within one standard error of the empirically best model (Kohavi, 1995, Breiman *et al.*, 1984).

6.2.6 Modelling of taxonomic and diel variation

Six RF models were constructed in the R environment (v3.5.0; R Core Team, 2018) using the randomForest package (v4.6-14; Liaw & Wiener, 2002): one to model the taxonomic variation between the five different species, and a further five to model the diel variation within each of the five species. For the taxonomic model, mtry was optimised using repeated k-fold cross-validation (three repeats; k = 10) using the caret package (v6.0-81; Kuhn, 2008). For the five models of diel variation, mtry was optimised using six repeats and k = 5 (see chapter 5, Figure 5.2 for a graphical representation of the cross-validation procedure). Due to the small number of samples (n = 20), the data was not partitioned into test and training sets: instead, all samples were used to optimise mtry and the results of cross validation were utilised for assessing the quality of the model. In each case, the optimal model was the most accurate model within one standard error of the empirically best model (Kohavi, 1995, Breiman *et al.*, 1984).

6.2.7 Variable selection

Important spectral features driving classification were identified for each of the RF models by permuting each spectral feature then comparing the mean decrease in accuracy as described in the randomForest package (Liaw & Wiener, 2002). Important spectral features were cross-referenced with the MarinLit database and metabolites reported in the coral research literature (Gordon & Motti, 2020a) using a mass error of Δ50 ppm as per Van Assche et al (2015), providing level 2 putative annotations where possible (Sumner *et al.*, 2007).

6.2.8 Data

The R scripts for this experiment are available for viewing or download from GitHub at https://github.com/brgordon17/diel-species (Gordon & Motti, 2020b). The full list of metabolites found in the coral research literature and the MarinLit database are available as an R package at GitHub; https://github.com/brgordon17/coralmz (Gordon & Motti, 2020a).

6.3 Results

6.3.1 Metabolome variation as a function of quenching method

PCA scores revealed no observable variations in the metabolome of *A. aspera* samples collected using the two different quenching methods (Figure 6.1a). Principal components (PCs) one and two captured a combined 57.3% of the variation in the data. Samples from both quenching methods occupy similar scores space in the plot, suggesting the variation captured by PCA was either very small or was not induced by either quenching method.

As with the PCA, the RLA plot displayed little variation in metabolite abundance between quenching methods (Figure 6.1b). All samples had very similar interquartile ranges and distributions. Compared to previous observations of biological and analytical variation (see chapter 5, Figure 5.5), overall variation is very low. Nonetheless, the methanol-quenched samples show slightly more variation in their metabolite abundances than samples quenched in liquid nitrogen.

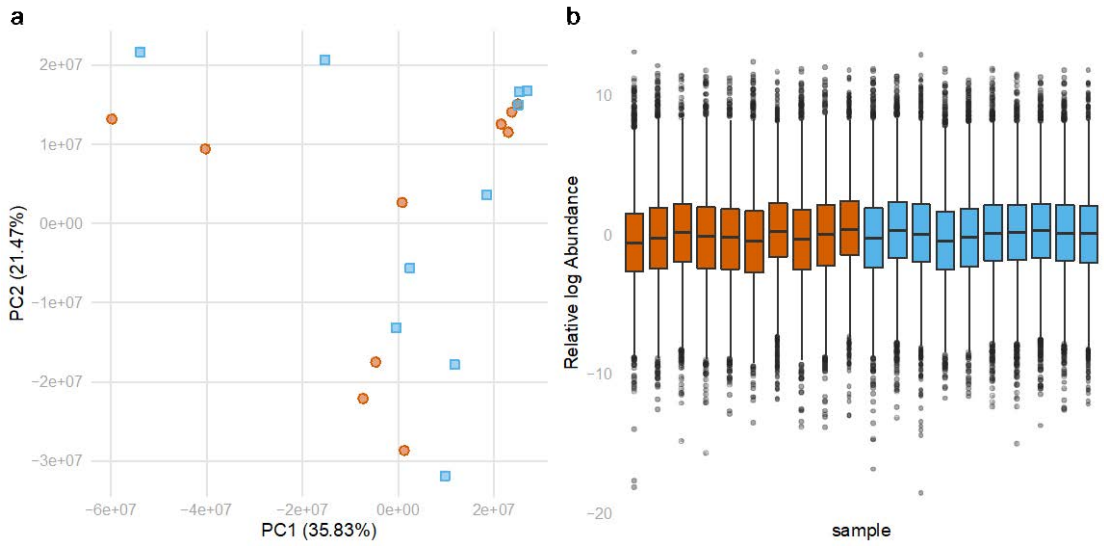


Figure 6.1 PCA scores plot (a) and relative log abundance plot (b) of samples quenched in liquid nitrogen (blue) and samples quenched in methanol (red)

Although variation due to the quenching method was less than the variation that is typically observed from biological or analytical factors, the quenching technique was modelled with RF to determine the true extent of the variation. Six repeats of five-fold cross validation produced a parsimonious model (mtry = 25) that was 100% accurate in predicting the quenching method used for each sample. This is unequivocal evidence that there is a difference in the metabolite content at the point of arrest of cellular

metabolism between the two quenching methods; consequently, a closer examination of the spectral features driving class separation was warranted.

Boxplots of the 12 most important spectral features driving the classification of the two quenching methods were constructed to understand how the two different quenching methods affected the metabolites captured, i.e. the metabolic profile (Figure 6.2). Non-detects drove separation of the two quenching methods: for example, spectral features *m/z* 349, 389, 399 and 813 were not detected in samples that were quenched with liquid nitrogen, while the remaining eight spectral features were not detected in samples quenched with 100% methanol. Nine of the spectral features from both methanol and nitrogen quenched samples had comparatively low intensities (<1 × 10^4 counts) that were close to the mass spectrometer's limit of detection. Three of the 12 spectral features, *m/z* 383, 462 and 696, had moderate intensities above 1×10^4 counts.

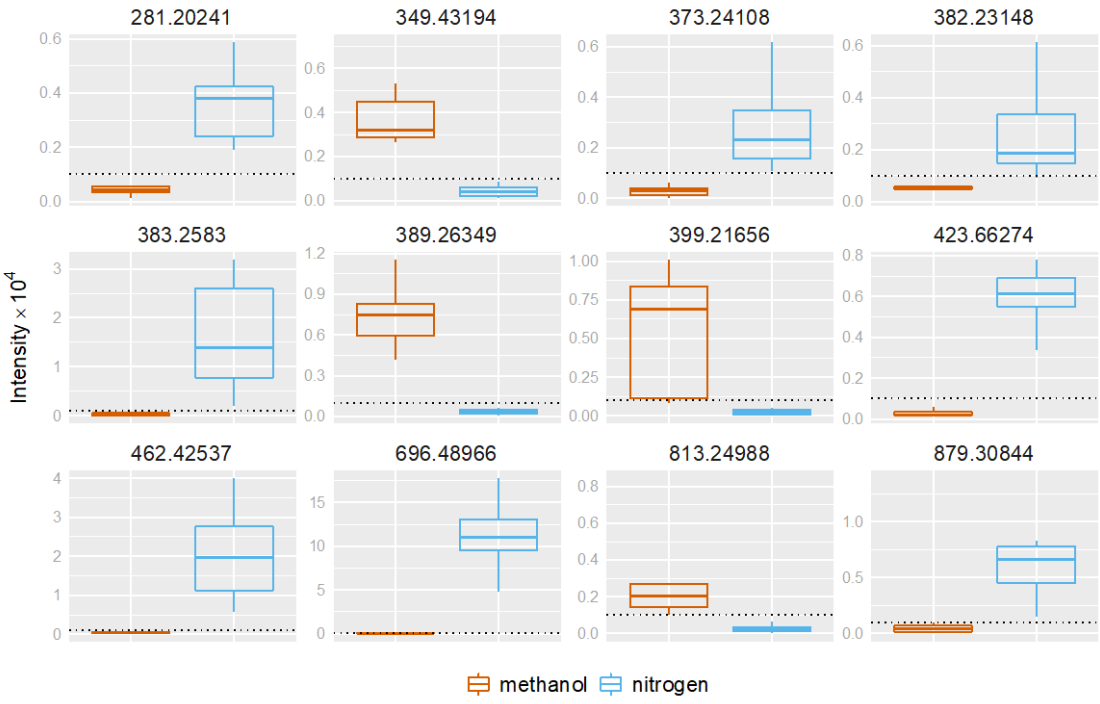


Figure 6.2 Boxplots of the 12 most important spectral features driving the separation of samples quenched by either methanol (red) or liquid nitrogen (blue). Each spectral feature is identified by its *m/z* above each of the nine panels. The dotted horizontal line represents the limit of detection (**0.1 × 10^4**). Boxplots of intensities below the limit of detection are non-detects (i.e. narrow distributions of imputed noise). Comparisons between features should take into account the free y-axis of each of the 12 panels.

6.3.2 Taxonomic and diel metabolome variation

Tight clustering in the PCA of all five species' clearly shows that taxonomic variation was greater than analytical variation or diel variation (Figure 6.3a). The negative component of PC2 and the positive component of PC1 captured variation associated with the two Montiporid corals, while the positive components of PC2 and PC1 captured the variation associated with *P. damicornis*. Metabolic variation of the *A. aspera* samples was best explained by the negative components of PC1 and PC2, while variation in the *P. cylindrica* samples was best explained by the negative component of PC1 and the positive component of PC2.

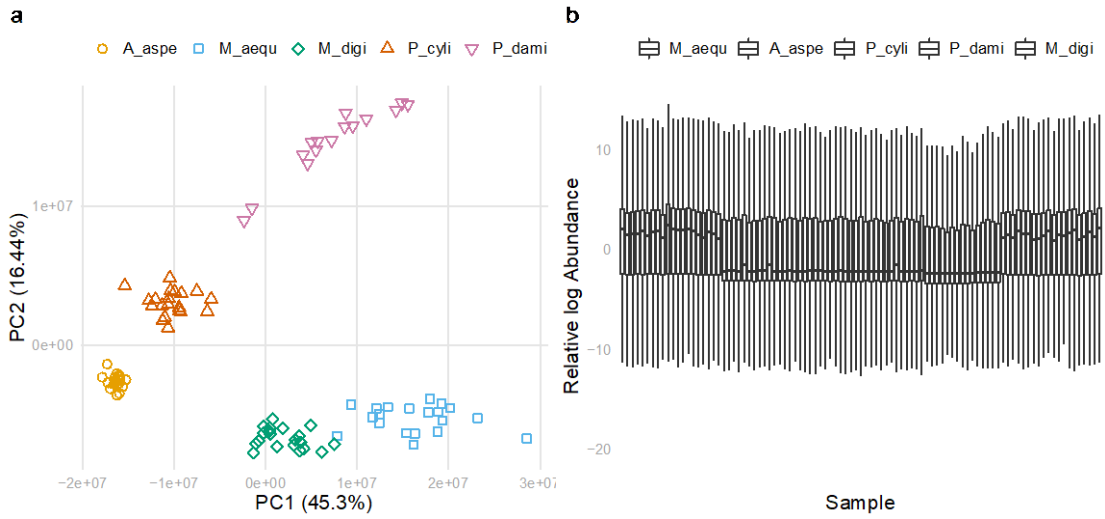


Figure 6.3 PCA scores (a) and relative log abundance (b) plots of the measured metabolomes of *A. aspera* (A_aspe), *M. aequituberculata* (M_aequ), *M. digitata* (M_digi), *P. cylindrica* (P_cyli) and *P. damicornis* (P_dami).

All five species occupied unique regions of the scores space for PC1 and PC2 and were distinctly separated from each other. *M. aequituberculata* and *M. digitata* occupied the most distant region of the scores plot compared to other species. Intraspecific variation was greatest in the *P. damicornis* samples and lowest in the *A. aspera* samples. Two trajectories appear to correlate with increasing intraspecific variation in the PCA scores plot: the first, extends from the *A. aspera* samples, through *P. cylindrica* and ending at the *P. damicornis* samples, while the second, extends from the *A. aspera* samples through to both of the Montiporid samples. All three species belonging to the *Acroporidae* family (*A. aspera*, *M. aequituberculata* and *M. digitata*), occupied similar scores space in the negative component of PC2.

The relative log abundance boxplots of each sample highlight the variation in metabolite abundances between species. Montiporid corals had the highest metabolite abundances

of all species (Figure 6.3b) and the greatest variation of their sample medians. The sample medians and distributions of *A. aspera* and *P. cylindrica* were very similar; suggesting similar metabolite abundances and low sample variations, which is supported by their tight clustering in the PCA. Apart from two samples, *A. aspera* had the lowest variation between the sample medians while *P. damicornis* had the lowest overall metabolite abundance. Montiporid corals had the least number of high-abundant outliers of all species, while *P. damicornis* had the greatest number of high-abundant outliers. The number of low-abundant outliers was similar across all species.

PCA was conducted on the raw data from individual species to further assess the variation within species and, in particular, to assess the level of variation associated with potential diel differences (Figure 6.4). A small amount of diel separation was observed in the scores space of each species; however, there was also a large amount of overlap between samples collected at different times. *P. damicornis* had the greatest separation between any two collection times with all samples collected at 05:00 and 16:00 h occupying unique areas of the scores space, although it should be noted that no data was available for *P. damicornis* at 21:00 h. For all species, diel variation within each class (timepoint) was similar to the variation between each class, indicating that diel variation was low and at levels similar to the noise in the data. Log transformation of the data, which is known to emphasise the standard deviations of low concentration metabolites (Kvalheim *et al.*, 1994, van den Berg *et al.*, 2006), revealed the structure associated with diel variation more clearly, further supporting the hypothesis that diel variation was low in comparison to taxonomic variation (Appendix Figure 3).

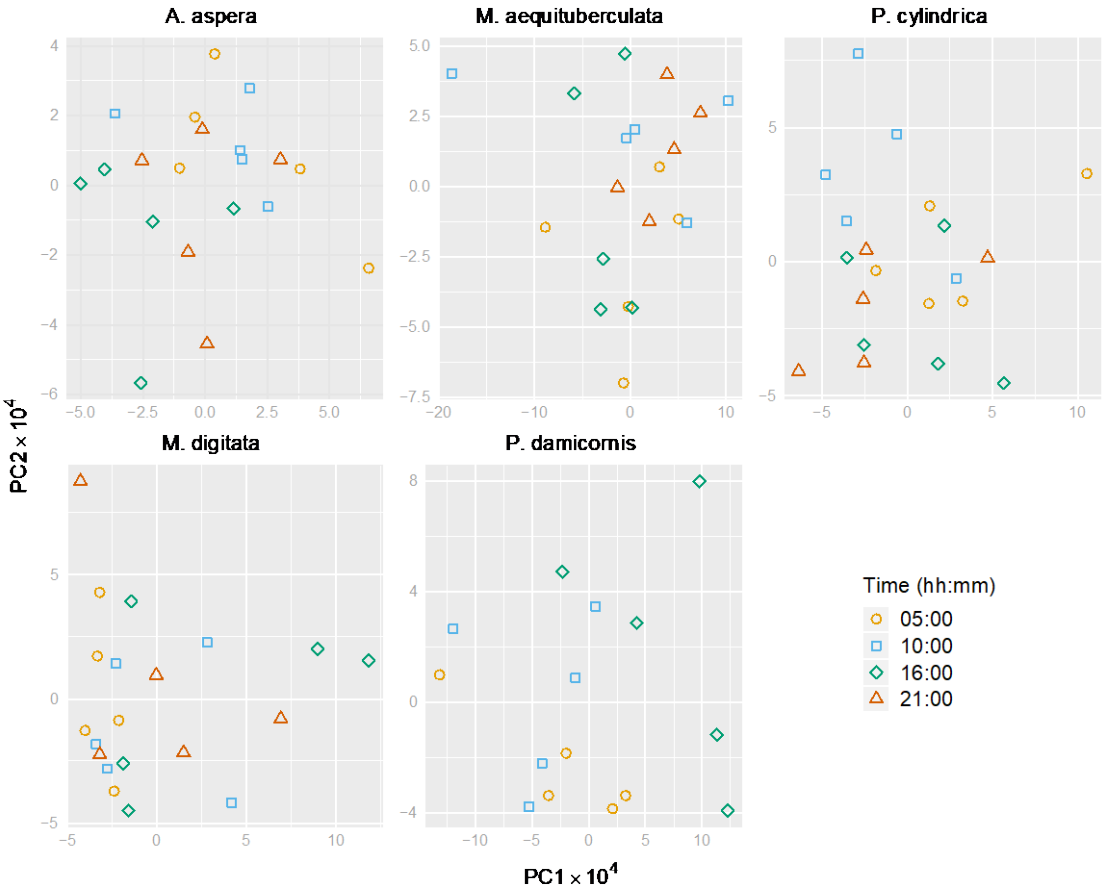


Figure 6.4 Composite PCA scores plot of PC1 and PC2 for each species

6.3.3 Spectral features driving taxonomic variation

RF modelling predicted the taxon of all training and test samples with 100% accuracy. To identify features driving taxonomic classification, the top 20 spectral features with the greatest contribution to the accuracy of the model were identified and labelled with their species-specific bias to model accuracy (Figure 6.5). Each of the top five features was biased to a single species and produced a mean decrease in model accuracy after permutation of 2.0-2.2%. The next six important features were biased to four species and produced a mean decrease in model accuracy after permutation from 1.8-2.0%. The remaining nine metabolites produced a mean decrease in model accuracy after permutation of ~1.7%.

Montiporid corals required fewer important spectral features for effective classification: for example, two features from the top 20 drove classification of *M. aequituberculata* (m/z 173 and m/z 1467; Figure 1.4), while for *M. digitata* three features drove classification (m/z 382, 749 and 404). In contrast, *A. aspera* utilised more of the top 20 features for classification than any other species, requiring six features in total. *P. damicornis* utilised five features and *P. cylindrica* utilised four.

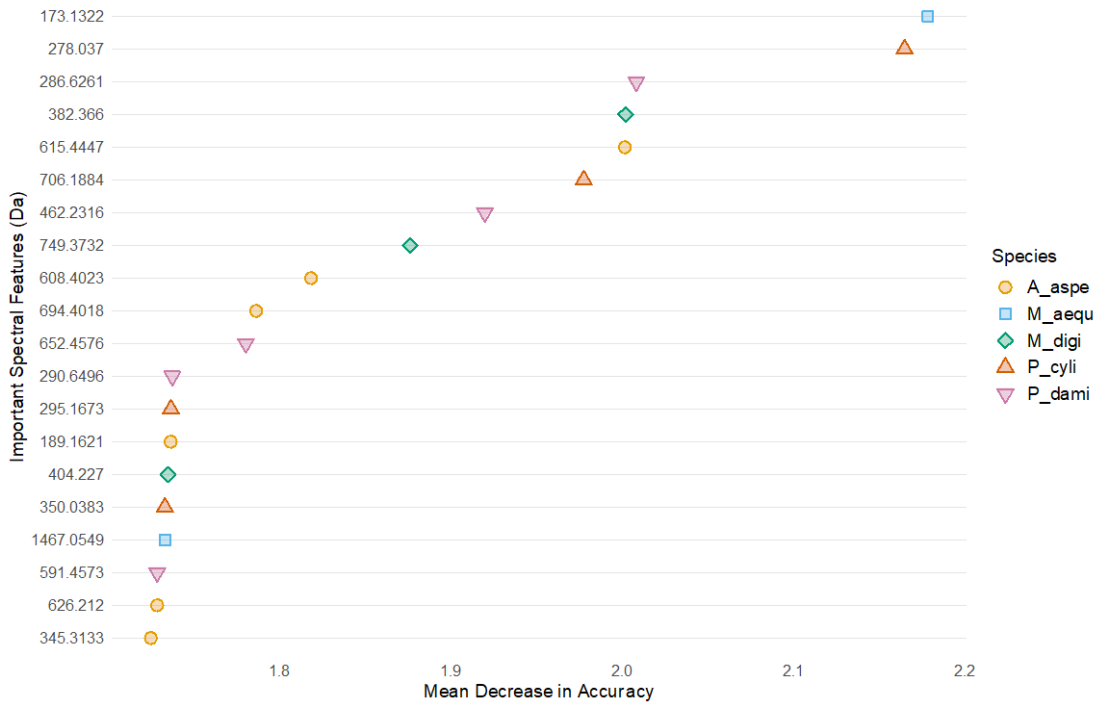


Figure 6.5 Top 20 important spectral features driving taxonomic separation. Symbol shape and colour identifies the species-specific importance of each feature (see plot legend).

Feature intensities of the 12 most important features (Figure 6.5) were examined to better understand their behaviour and influence (Figure 6.6). Many of the features driving taxonomic separation were unique to a single species or genus of coral. The two ions m/z 173 and 462 were detected in Montiporid corals only where m/z 173 was more abundant in *M. aequituberculata* and m/z 462 was more abundant in *M. digitata*. The ion m/z 382 was detected in *M. digitata* only, m/z 278 and 706 were unique to *P. cylindrica*, m/z 290 was unique to *P. damicornis* and m/z 608 was unique to *A. aspera*.

Where features were detected in more than one species, notable differences in abundance drove taxonomic separation, for example, m/z 615 and 749 were the only features of the top 12 detected in all species and drove separation of *A. aspera* and *M. digitata* where they were most abundant. On the other hand, m/z 286 was most abundant in *A. aspera* but drove separation of *P. damicornis* where its abundance was low in comparison. Finally, two of the 12 important features were absent from only a single species: m/z 652 was not detected in *P. cylindrica* and m/z 694 was not identified in *P. damicornis*. None of the 12 important features driving taxonomic variation could be provided with putative annotations from the coral research literature or the MarinLit database.

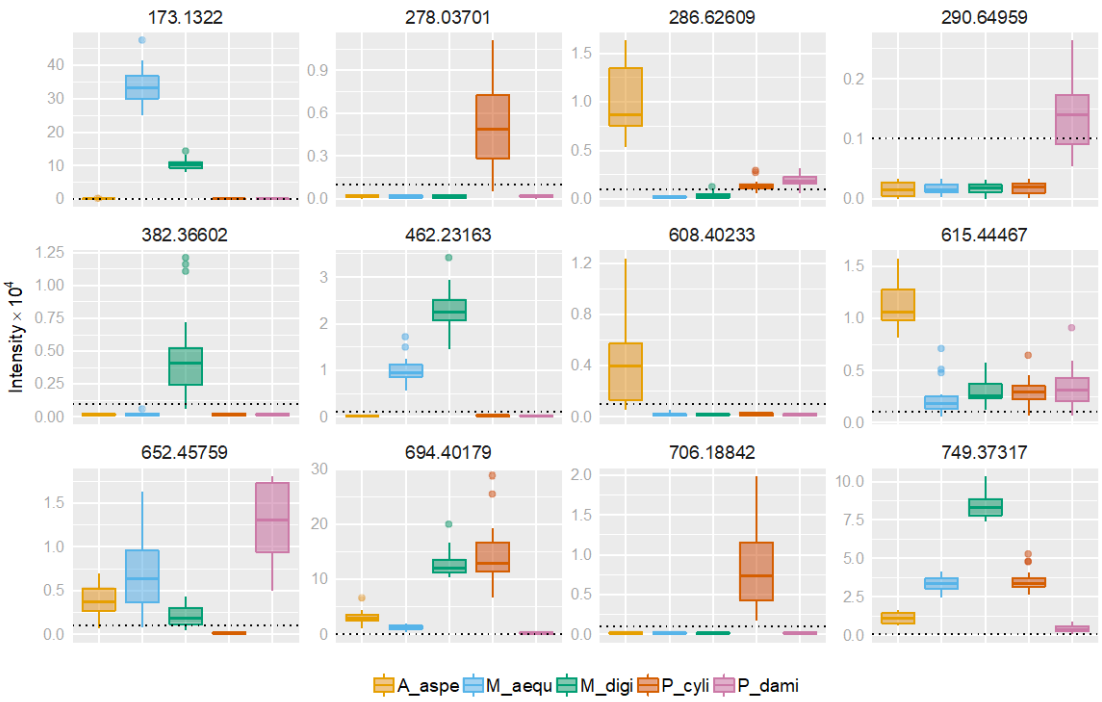


Figure 6.6 Feature intensities for the 12 most important features driving taxonomic separation. Each spectral feature is identified by its *m/z* above each of the 12 panels. The dotted horizontal line represents the limit of detection (0.1×10^4). Boxplots of intensities below the limit of detection are non-detects (i.e. narrow distributions of imputed noise). Comparisons between features should take into account the free y-axis of each of the 12 panels.

6.3.4 Spectral features driving diel variation

Five models were constructed to classify the diel variation within each species and identify the spectral features driving separation. Due to the small number of samples (n = 20 per species), a subsample of the data was not set aside for validation of the final model. Instead, the prediction results obtained from cross-validation were used to assess the quality and fit of each model (Table 6.1). The best performing models were those constructed for *A. aspera* and *P. damicornis*, which had prediction accuracies of 0.98 and 0.97, respectively. The model for *M. digitata* had an accuracy of 0.93, while the models for *P. cylindrica* and *M. aequituberculata* had prediction accuracies of 0.87 and 0.85, respectively. The sensitivity and specificity of all the models were above 0.80, with exception of the *M. aequituberculata* and *M. digitata* models, which had sensitivities of 0.60 for 10:00 h samples and 0.73 for 16:00 h samples, respectively. The sensitivity and specificity of the *A. aspera* model was particularly good, having values of 1.00 across three classes.

Table 6.1 Cross-validation prediction results for each Random Forests model including: the overall prediction accuracy of the model and the sensitivity and specificity of each model for the class descriptors (time).

		Time (hh:mm)			
Statistic	**Accuracy**	**05:00**	**10:00**	**16:00**	**21:00**
A. aspera	0.98				
Sensitivity		1.00	1.00	1.00	0.90
Specificity		1.00	0.97	1.00	1.00
M. aequituberculata	0.85				
Sensitivity		0.80	0.60	1.00	1.00
Specificity		1.00	0.99	0.88	0.93
M. digitata	0.93				
Sensitivity		1.00	0.97	0.73	1.00
Specificity		1.00	0.91	0.99	1.00
P. cylindrica	0.87				
Sensitivity		0.83	0.80	1.00	0.83
Specificity		0.89	0.94	0.99	1.00
P. damicornis	0.97				
Sensitivity		0.97	-	0.93	1.00
Specificity		0.97	-	0.98	1.00

The top 20 features driving the classification of diel variation for each species were identified by their contribution to model accuracy after permutation (Figure 6.7). In each case, features contributed ~1.4 to ~2.0% accuracy to the model. Only a single feature was common to more than one species; m/z 246.3009, which was common to both *A. aspera* and *M. aequituberculata* and drove separation of samples collected at 16:00 h in both species. Features having an m/z ~200-500 featured prominently in all species and their contributions to model accuracy were evenly spread across all four classes (sampling times), except for *P. damicornis*, where none of the top 20 features had a strong contribution to the classification accuracy of samples collected at 21:00 h.

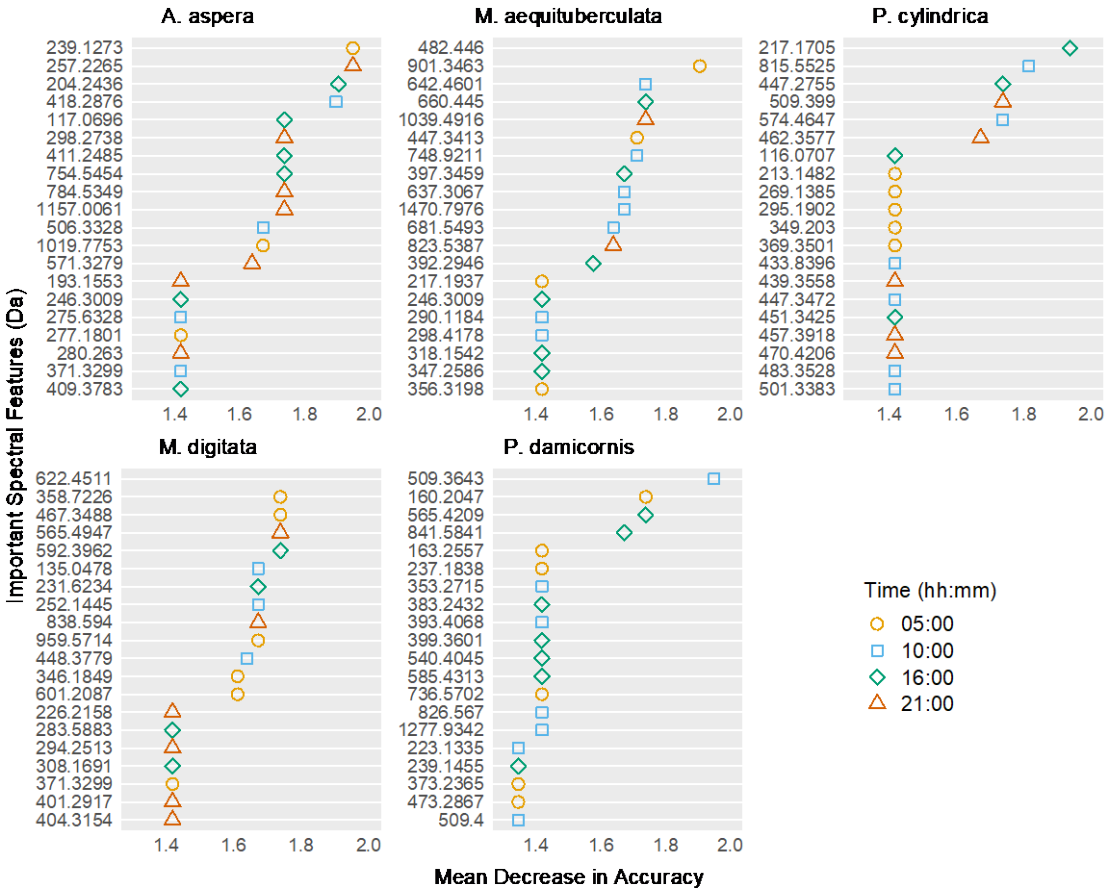


Figure 6.7 Important spectral features driving diel separation of five coral species.

The top 20 spectral features were compared with coral metabolites reported in the coral mass spectrometry research literature and the MarinLit database (Gordon & Motti, 2020a) to provide level 2 putative annotations (Sumner *et al.*, 2007). Five spectral features having a similar mass (Δppm ≤ 50 ppm) matched five compounds in the research literature (Table 6.2). The mean intensity of each feature was plotted across all collection times to examine their behaviour and influence in each species (Figure 6.8).

Feature m/z 784 drove class separation in *A. aspera* and produced a mean decrease in accuracy of ~1.75% after permutation (Figure 6.7). It had a monoisotopic mass within 50 ppm of arachidonoylthio-phosphorylcholine (also known as arachidonoylthio-PC) identified in *Pocillopora* sp. (Sogin *et al.*, 2016). Statistically significant differences in the mean intensity over time were observed in: *A. aspera*, where m/z 784 was detected in only two replicates at 21:00 h (one-way ANOVA, $F_{3,16} = 12.574$, $P < 0.001$); *P. damicornis* where m/z 784 was detected in only three replicates at 21:00 h (one-way ANOVA, $F_{2,16} = 13.598$, $P < 0.001$) and; in *M. digitata*, where m/z 784 was below the LOD at 05:00 h and then increased in abundance during the day before decreasing again at 21:00 h (one-way ANOVA, $F_{3,16} = 5.838$, $P < 0.01$).

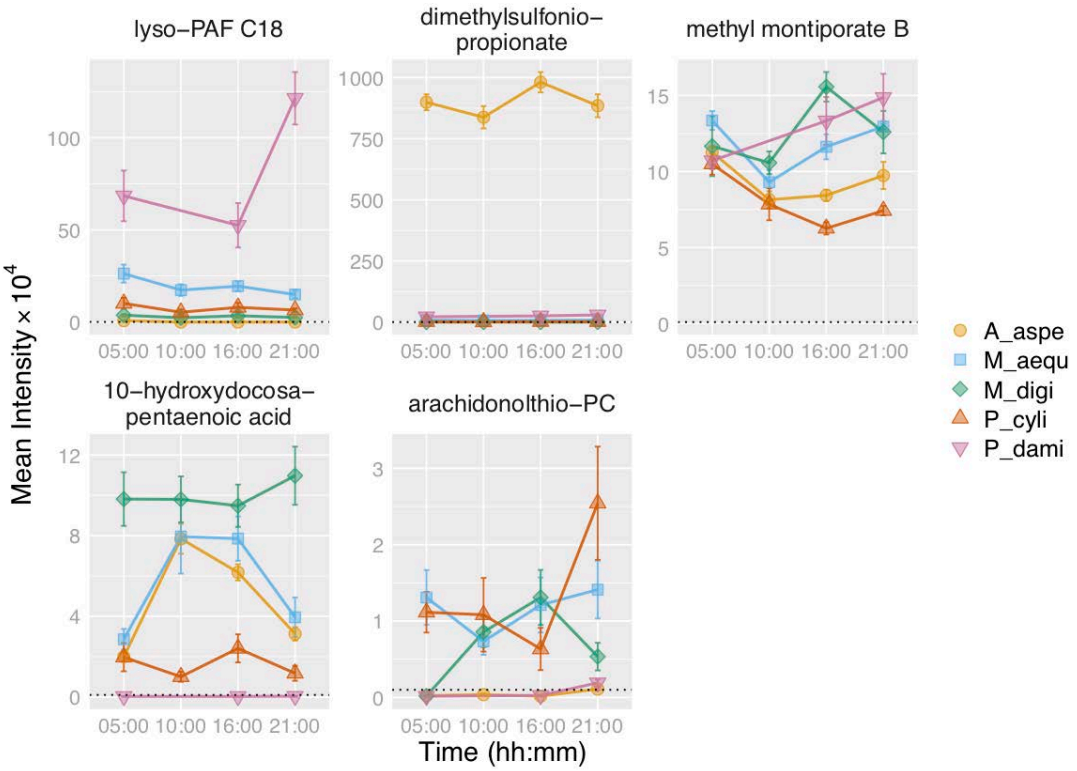


Figure 6.8 Annotated feature intensities for each species over time (mean ± s.e): The dotted horizontal line represents the limit of detection (0.1 x 10^4). Note that the y-axis scale differs for each plot.

The second feature, m/z 277, drove class separation in *A. aspera* and produced a mean decrease in accuracy of ~1.6% after permutation (Figure 6.7). It had a monoisotopic mass within 50 ppm of methyl montiporate B identified in *Montipora* sp. (Alam *et al.*, 2001). Statistically significant differences in the mean intensity over time was observed in *A. aspera* (one-way ANOVA, $F_{3,16} = 5.782$, $P < 0.001$), *M. aequituberculata* (one-way ANOVA, $F_{3,16} = 7.839$, $P < 0.01$) and *P. cylindrica* (one-way ANOVA, $F_{3,16} = 6.836$, $P < 0.01$) where notable declines in m/z 277 abundance was observed from 05:00 to 10:00 h. This initial decline was not observed in *P. damicornis* due to missing 10:00 h samples. In *P. cylindrica*, the initial decline in abundance from 05:00 continued beyond 10:00 h to 16:00 h. The abundance of m/z 277 continued to increase from 16:00 h to 21:00 in all species, except *M. digitata*, where m/z 277 decreased during that time.

The third feature, m/z 1019, drove class separation in *A. aspera* and produced a mean decrease in accuracy of ~1.68% after permutation (Figure 6.7). It was annotated as a dimer having a monoisotopic mass within 50 ppm of 1-octadecylglycero-3-phosphocholine (also known as Lyso-PAF C-18) as identified in *Pocillopora* sp. (Sogin *et al.*, 2016) and a number of other scleractinian corals (Quinn *et al.*, 2016). A statistically significant difference in the mean intensity over time was observed in: *A. aspera*, where

m/z 1019 was detected in all replicates at 05:00 h but not at any other time (one-way ANOVA, $F_{3, 16} = 26.401$, $P < 0.001$) and; in *P. damicornis,* where *m/z* 1019 was notably more abundant at 21:00 h (one-way ANOVA, $F_{2, 16} = 7.268$, $P < 0.01$).

The fourth feature, *m/z* 347, drove class separation in *M. aequituberculata* and produced a mean decrease in accuracy of ~1.4% after permutation (Figure 6.7). It had a monoisotopic mass within 50 ppm of 10-hydroxydocosapentaenoic acid as identified in *Modrepora* sp. (Mancini *et al.*, 1999). In *M. aequituberculata* and *A. aspera*, *m/z* 347 increased during the day before returning to 05:00 levels by 21:00 (one-way ANOVA, $F_{3, 16} = 4.771$, $P < 0.05$ for *M. aequituberculata* and $F_{3, 16} = 33.152$, $P < 0.001$ for *A. aspera*). The feature was most abundant in *M. digitata*; however, there was no statistically significant difference in the mean intensity over time. It was not detected in *P. damicornis* and there was no statistically significant difference in the mean intensity over time in *P. cylindrica.*

The last important spectral feature to be provided with a putative annotation was *m/z* 135, which drove class separation in *M. digitata* and produced a mean decrease in accuracy of ~1.68% after permutation (Figure 6.7). It had a monoisotopic mass within 50 ppm of dimethylsulfoniopropionate (Mancini *et al.*, 1999). In *M. digitata*, the species in which it drove class separation, *m/z* 135 was detected at low abundance (~2.5 x 10^4) in only two of the five replicates at a single time point (10:00 h). These two replicates drove the statistically significant difference in the mean intensity over time for *M. digitata* and the importance of this feature to classification. No statistical difference was observed in the mean intensity over time for any of the remaining four species; however, unlike *M. digitata*, *m/z* 135 was detected in all replicates of the remaining four species at mean intensities ranging from ~1.5 x 10^4 in *P. cylindrica* to ~90 x 10^4 in *A. aspera.*

Table 6.2 Table of features important to the prediction of diel variation matching metabolites identified in the coral mass spectrometry research literature and the MarinLit database. The shaded columns represent the data obtained in this experiment. The remainder of the table represents the data obtained from the literature. Important features (*m/z*) were compared to the monoisotopic masses reported in the research literature. Multiple matches of the same feature occur when another feature was identified as a potential adduct or when a metabolite was reported in two or more references. NA = not available; compound reported in the literature, but no data provided.

Sample Taxon	Detected *m/z*	Detected Ion	Literature *m/z*	Literature Exact Mass	Literature ion	Molecular formula	Mass Error (ppm)	Compound	Isolated from	Reference
A. aspera	784.5349	$[M+H]^+$	NA	783.560059	NA	$C_{44}H_{82}NO_6PS$	41	Arachidonoyl thio-phosphatidylcholine	*Pocillopora*	(Sogin *et al.*, 2016)
A. aspera	277.1801	$[M+H]^+$	299	276.1725	$[M+Na]^+$	$C_{17}H_{24}O_3$	1	Methyl montiporate B	*Montipora*	(Alam *et al.*, 2001)
A. aspera	1019.7753	$[2M+H]^+$	NA	509.384521	NA	$C_{26}H_{56}NO_6P$	1	Lyso-PAF C-18	*Pocillopora*	(Sogin *et al.*, 2016)
A. aspera	1019.7753	$[2M+H]^+$	510.42	509.384521	$[M+H]^+$	$C_26H_{56}NO_6P$	1	Lyso-PAF C-18	Scleractinia	(Quinn *et al.*, 2016)
M. aequituberculata	347.2586	$[M+H]^+$	NA	346.2508	NA	$C_{22}H_{34}O_3$	2	10-hydroxydocosapentaenoic acid	*Madrepora*	(Mancini *et al.*, 1999)
M. digitata	135.0478	$[M+H]^+$	135	134.040146	$[M+H]^+$	$C_5H_{10}O_2S$	3	Dimethylsulfoniopropionate	*Acropora*	(Swan *et al.*, 2017)

6.4 Discussion

Long-term *in situ* biomonitoring provides a means of monitoring and predicting changes in reef communities, such as the effects of temperature and salinity-induced bleaching events, and recovery from disturbances such as Crown-of-Thorns starfish outbreaks and cyclones. This data is vital for assessing the impacts of management programs such as the 2003 Great Barrier Reef Marine Park zoning plan (GBRMPA, 2004). In recent years, reef restoration has gained currency given the cumulative impacts on coral cover as it aims to build resilience of reefs and mitigate further damage to their delicate ecosystems. Metabolomics biomonitoring will be critical to assess the effectiveness of these intervention efforts while also having an increasingly important role as reef management moves further to a resilience-based approach.

This study established solvent-based sample quenching as a feasible alternative to the commonly employed liquid nitrogen method (Vuckovic, 2012), greatly simplifying the sampling and extraction procedure of corals in the field towards developing a robust biomonitoring method. Quenching with methanol induced a similar level of metabolome variation as liquid nitrogen confirming its suitability. Applying the methanol quenching method, the metabolome variation of five coral species over a 24-hour period was examined to ascertain the taxonomic and diel metabolome variation and the spectral features driving differentiation. Within taxa, *A. aspera* had the lowest intrasample variation, supporting its adoption as a sentinel species for metabolomics-based biomonitoring. RF modelling revealed low levels of diel variation in each coral species that was most apparent in low abundant spectral features. Together, these results highlight the importance of understanding the natural inter- and intra-specific variation of coral metabolomes when considering the suitability of a potential sentinel species and for instilling confidence in biomonitoring predictions.

6.4.1 Methanol quenching: a field-friendly and reproducible alternative to using liquid nitrogen.

In-field methanol quenching of the five coral species induced variation that was comparable in magnitude to the variation induced by the commonly used liquid nitrogen method; the defining difference being that liquid nitrogen had a greater selectivity for low-abundant metabolites. Although further work is required to better understand this phenomenon, there are several plausible explanations: 1) liquid nitrogen can capture low-

abundant metabolites with high turnover rates before they're metabolised into product compounds, 2) the low abundant metabolites are volatile and easily degas at the higher temperatures experienced during methanol quenching or, 3) liquid nitrogen is not as effective as methanol and is capturing metabolic processes. Given that liquid nitrogen is cold and fast acting it is likely to outperform methanol quenching while also reducing the volatility and reactivity of metabolites. For this reason, it should remain as the favoured method under laboratory conditions, or wherever it is convenient to employ, until the exact identity and nature of these low-abundant metabolites can be explored in greater detail. Nevertheless, the fact that variation was restricted to very low-abundant metabolites shows that methanol is a viable alternative to liquid nitrogen for the rapid quenching of wild coral metabolomes.

Quenching in methanol has four major advantages over liquid nitrogen for field-based sampling of corals: 1) it is a quick and simple one-step extraction technique; 2) it is a highly effective extraction solvent (see chapter three and Gordon *et al.*, 2013), increasing the throughput of sample preparation; 3) it does not require insulating containers and samples are unlikely to undergo major enzymatic changes that would normally occur at ambient temperatures; and 4) it does not require cumbersome protective equipment and can be easily handled by relatively unskilled personnel.

Methanol quenching is uncommon in coral metabolomics; however, the technique is widely used in microbial and cell culture metabolomics (de Koning & van Dam, 1992, Sáez & Lagunas, 1976, Mashego *et al.*, 2007, Bolten *et al.*, 2007, Canelas *et al.*, 2008). In microbial metabolomics, cold aqueous methanol (60% methanol; -40 °C) is routinely employed to arrest microbial activity while simultaneously washing the cells of any extracellular metabolites. In this case, it is the temperature of the solution that quenches the metabolism and the methanol is added to keep the cell suspension from freezing at sub-zero temperatures (Canelas *et al.*, 2008).

In contrast to its use in microbial metabolomics, methanol was employed here as a quenching agent for its ability to arrest cellular (metabolic) activity and not for its capacity to remain liquid at sub-zero temperatures. Methanol, which is inherently toxic to coral, quickly disrupts the cell walls in both the coral and its symbiont, effectively halting metabolic activity. Indeed, methanol and other popular solvents such as chloroform and acetone are known to quickly disrupt the cells of marine organisms and as such, are widely used in the extraction and analysis of metabolites (Andreae & Klumpp, 1979, Crews *et al.*,

1986). For example, methanol is the solvent of choice to analyse photosynthetic pigments (Daigo *et al.*, 2008, Kim *et al.*, 2013) and other key metabolites in corals (see chapter three and Gordon *et al.*, 2013). In contrast, methanol is far less effective on yeast cells grown in culture where it has been shown to be less disruptive to the cell wall, reducing metabolite leakage (Canelas *et al.*, 2008). This counterintuitive phenomenon could be due to the more hydrophilic yeast cell wall, which has a requirement for active nutrient exchange with its aqueous environment, and the very low temperatures at which the methanol is employed.

Liquid nitrogen is inherently difficult to work with in the field and limited by its high volatility at ambient temperatures. Indeed, this is not the first attempt to find a more field-friendly sampling method. In a similar study, field sampling and quenching of fresh plant material using methanol:dichloromethane (2:1) was shown to be a favourable alternative to liquid nitrogen (Maier *et al.*, 2010). In fact, solvent based quenching of plant material in this case was shown to be slightly more reproducible than liquid nitrogen.

This study showed that methanol quenching compares favourably to liquid nitrogen in terms of its reproducibility —a finding that is supported by the previously mentioned research in plants (Maier *et al.*, 2010). The effects of this one-step quenching and extraction method on the coral metabolome requires further investigation if the technique is to be used in a quantifiable manner; in particular, methanol's nucleophilic nature and its effect on reactive, electrophilic components of the metabolome, as well as extraction efficiency and compound degradation, must be explored in greater detail. Nevertheless, these findings were sufficient to adopt methanol quenching for this study given the inherent difficulties associated with in-field sampling of coral.

6.4.2 The suitability of a metabolomics biomonitoring sentinel species is influenced by the inter- and intra-specific variation of coral metabolomes.

This study measured varying degrees of metabolome variation and metabolite abundance (Figure 6.3a-b) in five species of coral during a single diel cycle. Montiporid corals had the greatest metabolite abundance and metabolome variation of the five species, while *A. aspera* had the least. While these differing degrees of variation are influenced by the genotype of each species (Vohsen *et al.*, 2019), they may also represent the homeostatic potential, or homeostatic range, of each species. Homeostatic control of the metabolome could affect the interpretation of results in two ways: 1) perturbed metabolites may be highly regulated, thus masking their initial response or, 2) perturbed metabolites may be less regulated, thus becoming more distinguishable compared to other, more highly regulated, metabolites.

Nevertheless, metabolic regulation is an important consideration when choosing a sentinel species for a metabolomics-based model of coral health.

The advantage of an untargeted metabolomics approach to model coral health is that it is not restricted to a specific class of metabolites, such as primary metabolites, which are more actively regulated. Many coral metabolomics studies have been untargeted in their approach; however, their scope is often narrowed to include only those metabolites that can be identified using commercially available standards or molecular reference libraries (Hillyer *et al.*, 2017b, Hillyer *et al.*, 2016, Klueter *et al.*, 2015, Sogin *et al.*, 2016). This is less of an issue for metabolomics biomonitoring, which relies on monitoring the changes to the metabolic profile, rather than quantifying changes in known metabolites, to signify a deviation from the pre-defined baseline. Given the desirability for a stable metabolic baseline of "healthy" coral, species with a high homeostatic potential and/or low metabolome variation such as *A. aspera*, represent the better choice as a sentinel species (compared to other species used in this study).

6.4.3 Putatively annotated metabolites associated with natural diel variation

An examination of the coral research literature provided level 2 putative annotations (Sumner *et al.*, 2007) for some of the most important spectral features associated with diel variation in *A. aspera*, *P. cylindrica, M. digitata* and *M. aequituberculata* (Table 6.2) The spectral feature m/z 784.53491 was strongly associated with diel variation in *A. aspera* and putatively annotated as arachidonoylthio phosphatidylcholine (ATPC). The first report of ATPC in corals was published by Sogin et al (2016) where ATPC was shown to be down regulated by ~23 % in *P. damicornis* exposed to thermal stress; prior to this report, ATPC was not known to occur naturally. ATPC is a synthetic compound first synthesised by Reynolds et al (1994) for use as a substrate in microtiter plate assays for the human cytosolic phospholipase A_2 protein. Consequently, the putative annotation provided here is unreliable and further research is required to confirm the identity of m/z 784.

Methyl montiporate B (m/z 277) was also associated with diel variation in *A. aspera*. This acetylenic compound was first isolated from *Montipora* sp. by Alam et al (2001). The exact ecological role of methyl montiporate B remains unknown; however, acetylenes have been found extremely effective at inhibiting bacteria and other coral pathogens in the surrounding seawater (Gochfeld & Aeby, 2008, Higa *et al.*, 1990), as ichthyotoxins (Higa *et al.*, 1990), and as sperm attractants in *Montipora digitata* (Coll *et al.*, 1994). As a sperm attractant, acetylenes in certain ratios were shown to be more effective in attracting sperm

from *M. digitata* than from other species of *Montipora*, which suggests they play a role in reducing the chances of hybridisation between different *Montipora* species. They are also known to irreversibly inhibit photosynthesis in *Symbiodiniaceae* within minutes of exposure (Hagedorn *et al.*, 2015). Natural products researchers have a keen interest in acetylenes due to their significant antitumor, antibacterial, antimicrobial and antifungal properties (Negri, 2015, Nathalie *et al.*, 2015)

Acetylenes may be common to many coral species given that m/z 277 was detected in all five species examined in this experiment. Statistically significant changes in the intensity of m/z 277 throughout the diel cycle in *A. aspera*, *M. aequituberculata* and *P. cylindrica* suggests the role of m/z 277 in coral physiology and/or photophysiology is not trivial. Indeed, acetylenes, and their ecological role in corals, could be worthy of far greater attention. With this in mind, m/z 277 should be formally identified to validate the strong association of acetylenes with F_v/F_m observed here.

A third important spectral feature associated with diel variation was the dimer m/z 1019, which was putatively annotated as lyso-PAF-C18 (Table 6.2). In chapter five, the C-16 form of lyso-PAF was associated with temperature induced changes in F_v/F_m. Both forms of lyso-PAF are very closely related and their role in corals has been discussed extensively in chapter 5; however, the C-18 form was found to be less common, and at much lower concentrations, than the C-16 forms in corals affected by algal and microbial interactions (Quinn *et al.*, 2016). In this study, lyso-PAF-C_{18} was only detected in *A. aspera* at a single timepoint (0500 h), artificially inflating its importance to the model. In contrast, lyso-PAF-C_{18} was far more abundant in *P. damicornis*, the only other species where lyso-PAF-C_{18} was observed to change significantly with the diel cycle. Since lyso-PAF-C_{18} could not be quantified at each timepoint in *A. aspera*, and since it wasn't flagged as important by the *P. damicornis* model, it is difficult to conclude that lyso-PAF-C_{18} changes throughout the diel cycle with any authority. These findings, and those of chapter 5, do, however, bring this compound to the attention of the analyst for closer scrutiny – a primary aim of any untargeted metabolomics study.

In *M. aequituberculata*, m/z 347 drove classification of diel variation. This feature was putatively annotated as 10-hydroxydocosapentaenoic acid, which was first identified in the deep-water corals, *Modrepora oculata* and *Lophelia pertusa* (Mancini *et al.*, 1999). To date, the discovery of 10-hydroxydocosapentaenoic acid by Mancini *et al.* (1999) remains as the only observation of hydroxypolyenoic acids in Scleractinian corals; however, they have been

observed in other marine invertebrates such as red algae (Guerriero *et al.*, 1990a, Lopez & Gerwick, 1988), sponges (Guerriero *et al.*, 1990b) and a starfish (d'Auria *et al.*, 1988). In mammals, hydroxypolyenoic acids are intermediates in the arachidonic acid cascade, an important signalling molecule present in the phospholipids of cell membranes (Shimizu & Wolfe, 1990). The arachidonic acid cascade is a popular biochemical target for the pharmaceutical industry due, in part, to its key role in the inflammatory response (Willenberg *et al.*, 2015).

Here, statistically significant increases in m/z 347 were observed during the day in *M. aequituberculata* and *A. aspera* but not in other species. Diel variation has been observed in other intermediates in the arachidonic acid cascade. In humans, for example, eicosanoids in saliva (Rigas & Levine, 1983) and gastric juices (Tonnesen *et al.*, 1974) have been shown to change in concentration during sleep-wake cycles. This may suggest that hydroxypolyenoic acids could play an important role in the circadian rhythms of corals; however, their biological role in corals remains unclear. For this to be resolved, future work should focus on isolating and unequivocally identifying the existence of 10-hydroxydocosapentaenoic acid in these corals.

The final important spectral feature to be provided with a putative annotation was m/z 135, which was annotated as dimethylsulfoniopropionate (DMSP). DMSP drove prediction of diel variation in *M. digitata*; however, m/z 135's importance for prediction in this case is likely to be overstated as DMSP concentration was below the limit of detection in all but two replicates at a single timepoint. Certainly, it is not feasible to conclude that diel variation of DMSP concentration was observed in *M. digitata*, nor was it observed in any other species. This is consistent with previous research that has found DMSP remains stable in coral tissue throughout the diel cycle (Tapiolas *et al.*, 2013).

6.4.4 Low levels of diel variation may affect the interpretation of metabolomic data and the confidence of biomonitoring predictions.

The feature annotated as lyso platelet activating factor (lyso-PAF C-18) is a prime example of how natural variation may affect the biological interpretation of metabolomic data. Investigation of the photosynthetic efficiency of *A. aspera* exposed to elevated temperature (Chapter 5) found the C-16 variant of lyso-PAF to be a potentially strong predictor of the maximum quantum yield of photosystem II (F_v/F_m), changing in abundance with continued thermal stress. Here, the C-18 variant of lyso-PAF was identified as a potential predictor of diel variation in *A. aspera* (Table 6.2 and Figure 6.8). As F_v/F_m is known to be

influenced by the natural diel cycle and its associated changes in solar irradiance (Lesser & Gorbunov, 2001, Brown *et al.*, 1999a), the C-18 variant is also likely to be a predictor of F_v/F_m. Conversely, the C-16 variant is likely to be a predictor of diel variation.

Minor changes in low-abundant spectral features are more likely to influence predictive models than minor changes in high-abundant spectral features. This happens because the absolute noise of the instrument increases with signal intensity, masking minor changes as metabolite abundance increases (Kvalheim *et al.*, 1994). However, MS is not truly quantitative due to a variety of factors that affect the ionisation process (Gosetti *et al.*, 2010, Ho *et al.*, 2003). Consequently, the measurement of small variations by MS can often be unreliable in low-abundant spectral features, resulting in the incorrect identification of important variables. Here, m/z 135 (annotated as DMSP), identified by RF as driving classification of diel separation in *M. digitata*, was unreliable in this study due to a lack of reproducibility. In this case, two of the five replicates from a single time point had significantly higher intensities, which inflated the importance of m/z 135. The unreliable variation of m/z 135 in *M. digitata* highlights the importance of closely examining the spectral features that drive class separation.

Low-level diel variations are an important consideration to metabolomics-based modelling of coral reef health on a landscape scale. Under lab conditions, diel variations can be accounted for with a considered experimental design that includes an appropriate number of controls and data pre-processing. A biomonitoring approach at landscape scales does not utilise controls in this traditional sense. Instead, many more corals need to be sampled over much larger spatial and time scales to establish a stable baseline of coral health or functional state. Nevertheless, diel variations should be carefully considered where control samples are lacking, especially when deriving conclusions about coral health from the behaviour of low abundant metabolites.

6.5 Conclusion

This research has advanced the prospect of metabolomics biomonitoring of corals by exploring the metabolic variation associated with quenching method, the diel cycle and taxonomy. Methanol is a viable alternative to liquid nitrogen for the rapid quenching of wild coral metabolomes. Apart from its superior ease of handling in the field, methanol quenching induces variation that is comparable in magnitude to the variation induced by

the commonly used liquid nitrogen method. Low levels of diel variation elevated the influence of low-abundant spectral features, highlighting the importance of a robust collection protocol, sound data processing and the close inspection of important features in machine learning. Even so, the reliance of all models on low-abundant spectral features revealed diel variation in the coral metabolome was relatively low; consequently, a shift in coral functional status away from its healthy baseline is unlikely to be masked by diel variation. *A. aspera* had the most stable metabolic baseline throughout the diel cycle, flagging it as a potential sentinel species. This stability is important in a biomonitoring context, as major biochemical changes should reflect the functional state of the coral rather than changes associated with the natural cycles within its environment.

Chapter 7
General Discussion: Predicting Environmental Changes in the Coral Metabolome

Benjamin Gordon wrote this chapter with intellectual and editorial support from Cherie Motti and Bruce Bowden.

By capturing and analysing high-resolution metabolic profiles (i.e. chemical fingerprints) of the coral metabolome, this study has advanced the current understanding of coral metabolomics and associated methods. It has established a link between the metabolic profile and coral health status, thereby enabling prediction of coral health and function. It has also provided a basis for a more thorough chemical investigation of keystone coral species by identifying a number of potential stress-specific biomarkers for further isolation and characterisation. Findings presented here also provide a solid foundation for the future development and advancement of metabolomics-based biomonitoring programs of corals and coral reefs. In summary, analytical methods have been firmly established with the publication of the first coral metabolomics extraction and analysis protocol, described in Chapter 3 (Gordon *et al.*, 2013). These methods were further optimised with the development and validation of a user-friendly chemical quenching method for in-field sampling, as described in Chapter 6. Chemometric analysis of metabolic profiles has been thoroughly explored, validating several machine learning approaches that accurately predict both discrete labels and continuous quantitative descriptors of coral health. Further, metabolic phenotypes associated with realistic, mid-century, levels of atmospheric CO_2 have been established and thoroughly examined. Provided below is a reasoned synthesis of this research and recommendations on future research directions for advancing coral metabolomics and its role in long-term biomonitoring.

The need for standardised analytical and data pre-processing protocols for coral metabolic profiling

This research began at a time when coral metabolomics was just emerging and there was no prior research in this field on which to build. From this modest base, analytical methods were developed (Chapter 3) to address a key gap in the coral metabolomics research; specifically, to determine, with proper validation, a sampling and extraction protocol for an untargeted metabolome analysis of scleractinian corals using LC-MS and ^{1}H-NMR. Since its publication in 2013 (Gordon *et al.*, 2013), this protocol has had a significant impact on the field, being utilised in several studies (Sogin *et al.*, 2014, Quinn *et al.*, 2016, Putnam *et al.*, 2016, Sogin *et al.*, 2017, Hartmann *et al.*, 2017). Seven years later, however, methods for the extraction and analysis of coral for metabolomics remain diverse, reflecting the unique research questions of individual studies.

Without well characterised biomarkers, coral metabolomics biomonitoring will need to use an *untargeted* metabolic profiling approach to monitor changes in the coral metabolome induced by environmental stressors. Analytical platforms such as mass spectrometry coupled with chromatographic separation (LC-MS) provide superior resolution (~1 ppm)

and are widely used in a diverse range of metabolic profiling experiments. Indeed, LC-MS was found to be superior to ^{1}H-NMR for classifying coral exposure to elevated CO_2 and temperature, in large part due to the greater amount of molecular information obtained. Both the Partial Least Squares Discriminant Analysis (PLS-DA) and Random Forest (RF) models, for example, were more parsimonious and had greater prediction accuracies for LC-MS data than those built with ^{1}H-NMR data.

The full repertoire of chemistry produced by corals is yet to be discovered, with isolation efforts focussed primarily on identification of secondary metabolites with specific bioactivity. Unfortunately, this represents a significant limitation to coral metabolomics studies not faced by other well-studied systems (German *et al.*, 2005; Sumner *et al.*, 2015*)*, including the availability of applicable metabolite databases and appropriate internal standards. While metabolomics analysis of keystone coral species over a diurnal cycle and exposed to climate-associated stressors (i.e. elevated temperature and pCO_2) has identified potential metabolites of interest, more research on the use of internal standards and their effects on the coral metabolome is sorely needed. Here, common internal standards were not used because they may potentially interact and react with the suite of metabolites present and alter the overall sample profile, which may affect the prediction outcomes and the interpretation of results (Wehrens *et al.*, 2016, Remane *et al.*, 2010, Chamberlain *et al.*, 2019). In this research, the use of an unadulterated sample was essential to provide confidence in the predictions that were provided. This was non-negotiable in this early-phase, discovery-based, research to ensure that interpretations of the modelled metabolic fingerprint were not based on the assumed behaviour of metabolites in the mass spectrometer. In addition to the above, there a variety of other factors that must be considered when using internal standards in untargeted, non-quantitative, metabolomics studies like that performed here. For example:

1. The availability of stable isotope labelled standards is very limited and they are often expensive;
2. The number of metabolites is often too large and chemically diverse to provide a suite of internal standards with similar chemical characteristics and behaviour as the entire metabolome; that is, one cannot accurately normalise an entire dataset containing thousands of highly variable, and chemically diverse, features based on a few, or even many, internal standards (Wehrens *et al.*, 2016, Chamberlain *et al.*, 2019);

3. Internal standards may coelute with other compounds, which is not ideal when it is not clearly known in advance what metabolites are being detected;
4. Internal standards are known to contribute to ion suppression and/or enhancement;
5. When used as quality controls, i.e. to monitor data quality and the variation induced during sample preparation and handling, internal standards may not be representative of all compounds – this is especially true for symbiotic organisms such as corals, which are comprised of a variety of cell types (animal, algal, bacterial and viral).

Nevertheless, internal standards are widely used in the metabolomics field and the body of evidence suggests their benefits far outweigh any potential downsides. Indeed, internal standards play an important role as quality controls to assess variation from a variety of sources – not to be confused with internal standards that are used for quantification, which must be added just prior to analysis (Broadhurst *et al.*, 2018). While the discovery-based approach taken in this study precluded the use of internal standards, identification of internal standards suited to investigating coral extracts– and their effects on the extracted coral metabolome – must be considered in future research, as they will be essential for monitoring the quality of data. This is particularly important for large datasets collected over extended timeframes, as is required in biomonitoring programs.

Complex pre-processing is required before LC-MS data can be confidently analysed, representing the major downside to using this platform for metabolic profiling. Retention time correction, feature extraction, batch correction and isotope and adduct identification, for example, must be tailored and validated for every taxa, sample matrix, LC-MS instrument, and LC-MS separation and detection protocols. Coral samples and LC-MS conditions were kept consistent across each of the experiments conducted here and effort directed towards validating and testing a variety of data pre-processing approaches. Fortunately, data pre-processing in later experiments only required validation without testing. For most new studies, however, data pre-processing and analytical approaches are an obvious bottleneck in the coral metabolomics pipeline. Future research should, therefore, focus on the development of a standardised set of coral-specific pre-processing, separation, and analysis protocols suitable for a variety of platforms and study conditions. This would not only streamline the coral metabolic profiling workflow, but also encourage the use of a standard set of procedures; improving the insights that can be gained through the increased availability and the utility of standardised datasets.

Machine learning in untargeted metabolic profiling of the coral metabolome

There is a wide variety of machine learning and statistical approaches for analysing metabolome data. Supervised machine learning algorithms predict a predefined outcome by iteratively optimising the model according to a measure of performance, such as accuracy or area under the receiver operating characteristic curve (AUC). Some more commonly used algorithms include: RF, PLS-DA, Support Vector Machines and principal components regression. The most widely employed algorithm in metabolomics is the linear PLS-DA algorithm, as it performs admirably well for binary classification problems and is computationally inexpensive compared to other algorithms (Gromski *et al.*, 2015); however, the performance of the PLS-DA algorithm may suffer in comparison to more complex algorithms – as was the case in Chapter 4 – particularly with larger datasets.

Untargeted metabolic profiling using LC-MS platforms generate far more molecular features than targeted approaches, or even untargeted approaches utilising alternative platforms, such as GC-MS and NMR. The untargeted LC-MS approach, for example, routinely generates metabolic profiles with anywhere from ~1000 to ~6000 molecular features, while those generated using targeted or semi-targeted approaches have far fewer features as a consequence of filtering data to identifiable metabolites only. Therefore, non-linear algorithms, such as RF and neural networks, are more suitable for classifying the untargeted and unfiltered profiles that are encountered in LC-MS-based metabolomics biomonitoring.

Metabolomics biomonitoring has unique chemometric demands for classifying metabolic profiles; however, as the field is still in its infancy a number of questions are yet to be thoroughly explored. For example:

1. How are unwanted variations associated with long-term metabolic profiling handled?
2. How current should a model be, and how is new data integrated, aligned and incorporated with existing data?
3. How do we incorporate data from different platforms, sources or organisms for use by a single model or decision support tool?
4. What model outcomes are important to marine park managers?
5. How do we ensure the quality of predictions and instil confidence in those using them?

Regarding question one, this research shows that modelling unwanted variation can be an effective solution to removing the variation associated with different batches. This was achieved using the Harman algorithm (Chapter 5) by modelling both the batch-associated variation and the known biological variation, then removing the former while monitoring the probability of removing the latter. Low levels of natural diel variation were successfully identified using the RF algorithm and these too, could be successfully accounted for when drawing conclusions from metabolome data (Chapter 6). Modelling variation in pooled biological quality controls (PBQCs) is another commonly used method and, given the ease of preparing and analysing PBQCs, their inclusion in metabolic profiling experiments is considered best practice. However, the long-term stability and availability of PBQCs must be considered before implementing a long-term metabolomics biomonitoring programming, as they must remain consistent over time. Therefore, future research should be conducted to identify a stable, consistent, and widely available PBQC that can adequately and reproducibly represent the chemistry of the coral metabolome.

The second question concerns the incorporation of new data into a predictive model. This is a time-consuming process because of the computationally expensive task of aligning the new LC-MS data with pre-existing data. To date, this is an unavoidable problem that is commonly handled by reprocessing (feature extraction and alignment, etc) then retraining and releasing new models at strategically timed intervals; however, this approach limits the speed that the model could provide feedback on coral health, exacerbating the problem as the volume of data grows. Recent research has highlighted the potential of *mapping* spectral features rather than aligning their retention times through warping functions (Smith *et al.*, 2015). Realising this approach would overcome the computational expense associated with alignment and facilitate the fusion of data from experiments using different chromatographic conditions.

Predictive models are only as good as the data provided. Incorporating data from different laboratories or different platforms into a single model requires extensive data cleaning and pre-processing to ensure the quality and consistency of the data. This process takes considerable time, requires robust quality control and is often too difficult or time consuming to consider; consequently, data fusion is rarely considered in metabolomics. Stacking multiple models, rather than incorporating data into a single model, is one possible solution to the data fusion problem.

Model stacking involves training a new, or top layer, machine learning algorithm to combine the predictions of multiple base layer models. This approach to machine learning offers predictions that typically perform as well as, or better than, the individual base layer models (Polley & Van der Laan, 2010). Model stacking has had very low adoption rates in metabolomics to date due to the perceived complexity of the technique, poor interpretability and limitations in computing power. Data science, however, has advanced significantly in recent years and stacked models, along with new methods to interpret them, are now commonplace in machine learning. Stacked models facilitate the use of pre-existing datasets from different laboratories, reducing the effort and time required to build a diverse and representative dataset. Hence, metabolomics biomonitoring could be implemented quicker, and at lower cost, using such an approach. For this reason, future research should examine the stacked model approach for combining the predictions of multiple unique metabolic profiling experiments into a single model for predicting coral health.

Model outcomes must be underpinned with a sound understanding of the underlying biology or processes that may be affected by the decisions that the model supports. This could be achieved by incorporating scientifically valid data from alternative sources. A model that predicts the susceptibly of coral to bleaching, for example, might incorporate physical and water quality measurements from existing sources such as oceanographic moorings and satellites. This could be approached from a variety of angles; for example, pre-existing data could be incorporated into a new predictive model or, alternatively, metabolic profiling data could be incorporated into pre-existing coral reef monitoring programs. Either way, the success of any metabolomics biomonitoring program will be reliant on the support of reef managers and agencies already invested in coral reef monitoring. Consequently, coral sampling for metabolomics biomonitoring should prioritise areas that are actively monitored by existing programs. Engaging stakeholders involved in these programs is a high priority for the advancement of metabolomics biomonitoring.

Isolating the individual response of host and symbiont

Unfortunately, traditional metabolomics-based approaches are not ideal for the analysis of an intracellular symbiosis as they do not provide the spatial resolution required to investigate the individual responses of the host and symbiont. Quenching the metabolism of the study species, followed by extraction of metabolite classes of interest in a tailored solvent, is the standard workflow used for sample preparation (Gordon *et al.*, 2013);

however, this approach does not distinguish between the host and symbiont, and attempts to differentiate their response through data analysis is extremely challenging. Separating the host tissue from the symbiont before extraction is not practical, as it is difficult to maintain the low temperatures required to suppress enzymatic activity during separation. Hillyer et al (Hillyer *et al.*, 2016, Hillyer *et al.*, 2018) attempted this by grinding and homogenising the coral holobiont under liquid nitrogen then suspending the material in chilled MilliQ water before separating the symbiont from the host by centrifugation (Hillyer *et al.*, 2016). Although every attempt was made to minimise metabolic activity, their method has not been validated in the research literature and future research should verify their method and explore potential host-symbiont separation methods in more detail.

Measuring the cellular level profiles of host and symbiont *in situ* is possible using platforms such as nano scale secondary ion mass spectrometry (nanoSIMS, Marshall *et al.*, 2007), matrix assisted laser desorption ionisation mass spectrometry (MALDI-MS; Kopp *et al.*, 2015) and Synchrotron Radiation Infrared Microspectroscopy (SR-IRM) (Gordon *et al.*, 2018), the latter having the potential to examine live algal cells in culture. Utilising a liquid flow-through cell, media can be manipulated and delivered to live cells and their response can be monitored in near real time (Heraud *et al.*, 2005). Given that our current knowledge of the signal transduction pathways and molecular interactions that govern the initiation and stability of the coral-algal symbiosis remains poor (Weis *et al.*, 2008, Davy *et al.*, 2012), future research should focus on using a combination of platforms to decouple the host and symbiont responses observed by the LC-MS analyses conducted here.

Metabolomics biomonitoring can provide real world data about the coral response to ocean acidification and warming

Recent research has revealed that traditional CO_2 and temperature manipulations in the scientific research literature have been too severe (Hughes *et al.*, 2017a). There is now an urgent need to recalibrate ocean warming and acidification experiments to better understand coral-stressor interactions under more realistic emissions scenarios (Hughes *et al.*, 2017a, Hughes *et al.*, 2010, Pandolfi *et al.*, 2011, Steffen *et al.*, 2015). This is a difficult task because the biological response to smaller scale fluctuations in water temperature and CO_2 are not as easily detected, require more time to present physically, and are more difficult to control experimentally. These barriers could be overcome with the establishment of a long-term metabolomics biomonitoring programming. Experimental manipulations of pCO_2 and temperature, for example, would then be redundant, as a well-designed metabolomics biomonitoring program would capture the biochemical response to

natural and anthropogenic-induced variation in the coral's environment. Indeed, Chapter 4 shows that metabolic profiling can identify molecular phenotypes associated with reductions in pH, elevated temperature, and combinations of the two, at pCO_2 levels consistent with atmospheric CO_2 concentrations under RCP 4.5. To date, however, metabolomics has yet to be employed in the study of coral health long-term. It is imperative, therefore, that future research focus firmly on long-term metabolomics biomonitoring and long-term chemometric approaches.

The functional roles of many secondary metabolites in the coral holobiont are poorly understood

Several secondary metabolites were putatively identified as directing classification of coral health. The application of a 70% aqueous methanol extraction protocol which is more selective for secondary metabolites and the reliance on a database that contained a high proportion of secondary metabolites (Gordon & Motti, 2020a), has likely biased their assignment at the expense of primary metabolites. Nevertheless, secondary metabolites clearly play an important role in distinguishing the coral's response to stress and environmental variation.

Acetylenes, for example, were putatively identified in *A. aspera* in three separate experiments (Chapters 4-6), indicating their potentially important role in *A. aspera* physiology. In chapter 4, m/z 217, putatively annotated as montiporyne A and montiporyne B, was strongly associated with *A. aspera* nubbins exposed to thermal and CO_2 stressors. In chapter 5, m/z 249, putatively annotated as dihomomontiporyne H, was an important feature driving the prediction of the maximum quantum yield of photosystem II (F_v/F_m). In chapter 6, m/z 277, putatively annotated as methyl montiporate B was associated with diel variation. While it is difficult to speculate on the ecological function of acetylenes in *A. aspera* without a definitive identification, significant changes in F_v/F_m were common to all three experiments. In Chapters 4 and 5, for example, F_v/F_m reduced in response to stress, while in Chapter 6, F_v/F_m was expected to change throughout the diel cycle (Lesser & Gorbunov, 2001, Brown *et al.*, 1999a, Sorek *et al.*, 2013), particularly in response to the significant shift in irradiance that occurs at dawn and dusk. Given this common factor, and the fact that Montiporid corals are known to harbour compounds that inhibit photosynthesis in *Symbiodiniaceae* (Hagedorn *et al.*, 2015, Hagedorn *et al.*, 2010), *A. aspera* may be utilising acetylenes or closely related compounds to regulate the photosynthetic performance of their coral symbionts.

Lyso-PAF was another putatively identified metabolite that drove prediction of samples undergoing changes in F_v/F_m in more than one experiment. In Chapter 5, for example, m/z 985, putatively annotated as lyso-PAF-C_{16}, was a strong predictor of symbiont F_v/F_m. Its abundance initially decreased in response to increasing temperature before returning to original baseline levels, consistent with the PAF-induced inflammatory response in mammals (Yost *et al.*, 2010, Camussi & Brentjens, 1987), invertebrates (Sugiura *et al.*, 1992) and corals (Quinn *et al.*, 2016). In Chapter 6, the dimer m/z 1019 was putatively annotated as lyso-PAF-C_{18} and contributed to the classification of corals throughout the diel cycle, which, as mentioned previously, causes changes in symbiont F_v/F_m. Given the well-known role of lyso-PAF in the immune response of many animals, changes in symbiont F_v/F_m may elicit a similar immune response in either partner. To advance the field of coral metabolomics, future research must focus on isolating and identifying these common features from corals. Indeed, coral metabolomics should revisit its roots, incorporating a natural products chemistry approach to formally identify a greater proportion of the coral metabolome with the same sense of urgency and effort that has been invested into decoding genomes.

Considerations for coral metabolomics biomonitoring

Tropical coral reefs are under considerable and increasing pressure from global stressors caused by a rapidly changing climate and their resilience into the future is uncertain. A more robust method to detect subtle variation and decline in coral fitness within the holobiont (i.e. host and symbiont) - before changes in phenotype expression become evident - will better inform reef managers by offering an early warning system and decision support tools that stratify corals reefs according to a variety of risk factors.

Decision support tools and predictive models must have the confidence of marine park managers. To instil this confidence, predictive models must be validated and robust to unforeseen environmental changes, such as abnormal weather events and anthropogenic influences, as well as variations introduced through subtle differences in the collection, handling, processing, and analysis of coral samples.

This research has considered several factors to improve the robustness and quality of predictions made using coral metabolic profiles. The quality of predictions, for example, was shown to be affected by a lack of sample replication in Chapter 4. In Chapter 6, a simplified method for sample collection was evaluated after recognising the complexity and potential for errors associated with the commonly used liquid nitrogen quenching

approach. Nevertheless, there remain numerous gaps in the research that must be explored before the coral metabolic profiling techniques explored here can be utilised for metabolomics biomonitoring at full-scale. Some of the major gaps identified here for consideration include:

1. The development of standardised sample handling, extraction and analysis protocols, with appropriate quality controls and quality assurance, to improve the quality and comparability of data;
2. Improvements in, and standardisation of, mass spectral feature detection, alignment and/or mapping to improve the quality of data and facilitate the sharing and fusion of datasets.;
3. Identification of a stable and common PBQC to facilitate the sharing and fusion of datasets;
4. Research into the stacked model approach for combining the predictions of multiple concurrent biomonitoring programs, improving predictive power and model stability;
5. Prioritising regions that are actively monitored by existing programs to make use of current resources and datasets that may improve model predictions and stability;
6. Engaging stakeholders involved in existing programs to document the business needs and risks;
7. Formally identifying a greater proportion of the coral metabolome, and specifically those putatively assigned features identified here;
8. Furthering the work by Tang *et al.* (2018) to explore the full potential of the lipidome for biomonitoring purposes;
9. Exploring the full potential of affordable technology, such as older, lower resolution mass spectrometers, to reduce the investment needed. Chapter 4 provides evidence that affordable and low-resolution instruments can provide data to categorise the health status of corals.

This thesis has developed protocols and provided insights into the suitability of metabolic profiling to monitor coral health long term. While more rigorous studies, including biomarker identification, are needed to progress metabolomics-based biomonitoring of corals, findings reported here provide a basis to further explore this strategy.

BIBLIOGRAPHY

Adams, L. M., Cumbo, V. & Takabayashi, M. (2009) Exposure to sediment enhances primary acquisition of *Symbiodinium* by asymbiotic coral larvae. *Marine Ecology Progress Series,* **377,** 179-156.

Aernouts, B., Polshin, E., Saeys, W. & Lammertyn, J. (2011) Mid-infrared spectrometry of milk for dairy metabolomics: a comparison of two sampling techniques and effect of homogenization. *Analytica chimica acta,* **705,** 88-97.

Akoka, S., Barantin, L. & Trierweiler, M. (1999) Concentration Measurement by Proton NMR Using the ERETIC Method. *Analytical chemistry,* **71,** 2554-2557.

Alam, N., Bae, B. H., Hong, J., Lee, C. O., Im, K. S. & Jung, J. H. (2001) Cytotoxic diacetylenes from the stony coral *Montipora* species. *Journal of natural products,* **64,** 1059-1063.

Alam, N., Hong, J., Lee, C.-O., Choi, J. S., Im, K. S. & Jung, J. H. (2002) Additional cytotoxic diacetylenes from the stony coral *Montipora* sp. *Chemical and pharmaceutical bulletin,* **50,** 661-662.

Ali, K., Maltese, F., Fortes, A. M., Pais, M. S., Choi, Y. H. & Verpoorte, R. (2011) Monitoring biochemical changes during grape berry development in Portuguese cultivars by NMR spectroscopy. *Food Chem,* **124,** 1760-1769.

Andersson, E. R., Day, R. D., Loewenstein, J. M., Woodley, C. M. & Schock, T. B. (2019) Evaluation of Sample Preparation Methods for the Analysis of Reef-Building Corals Using ^{1}H-NMR-Based Metabolomics. *Metabolites,* **9,** 32-50. doi: 10.3390/metabo9020032

Andreae, M. O. & Klumpp, D. (1979) Biosynthesis and release of organoarsenic compounds by marine algae. *Environmental science & technology,* **13,** 738-741.

Anthoni, U., Christophersen, C., Hougaard, L. & Nielsen, P. H. (1991) Quaternary ammonium compounds in the biosphere—an example of a versatile adaptive strategy. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry,* **99,** 1-18.

Anthony, K. R. N., Kline, D. I., Diaz-Pulido, G., Dove, S. & Hoegh-Guldberg, O. (2008) Ocean acidification causes bleaching and productivity loss in coral reef builders. *Proceedings of the National Academy of Sciences,* **105**(45), 17442-17446. https://doi.org/10.1073/pnas.0804478105

Aue, W. P., Karhan, J. & Ernst, R. R. (1976) Homonuclear broad band decoupling and two-dimensional J-resolved NMR spectroscopy. *The Journal of Chemical Physics,* **64,** 4226-4227.

Bae, B. H., Im, K. S., Choi, W. C., Hong, J., Lee, C.-O., Choi, J. S., Son, B. W., Song, J.-I. & Jung, J. H. (2000) New acetylenic compounds from the stony coral *Montipora* sp. *Journal of natural products,* **63,** 1511-1514.

Banaszak, T., LaJeunesse, T. C. & Trench, R. K. (2000) The synthesis of mycosporine-like amino acids (MAAs) by cultured, symbiotic dinoflagellates. *Journal of Experimental Marine Biology and Ecology,* **249,** 219-233.

Bedia, C., Cardoso, P., Dalmau,N., Garreta-Lara. E., Gómez-Canela, C., Gorrochategui, E., Navarro-Reig, M., Ortiz-Villanueva, E., Puig-Castellví, F. & Tauler, R. (2018) Applications of Metabolomics Analysis in Environmental Research. *Compr. Anal. Chem.*, **82**, 533-582. https://doi.org/10.1016/bs.coac.2018.07.006

Beedessee, G., Hisata, K., Roy, M. C., Van Dolah, F. M., Satoh, N. & Shoguchi, E. (2019) Diversified secondary metabolite biosynthesis gene repertoire revealed in symbiotic dinoflagellates. *Scientific reports,* **9,** 1204.

Belshe, E., Durako, M. J. & Blum, J. E. (2007) Photosynthetic rapid light curves (RLC) of *Thalassia testudinum* exhibit diurnal variation. *Journal of Experimental Marine Biology and Ecology,* **342,** 253-268.

Beltran, A., Suarez, M., Rodr√ ≠guez, M. A., Vinaixa, M., Samino, S., Arola, L., Correig, X. & Yanes, O. (2012) Assessment of compatibility between extraction methods for NMR- and LC/MS-based metabolomics. *Analytical chemistry*.

Benton, H. P., Want, E. J. & Ebbels, T. M. (2010) Correction of mass calibration gaps in liquid chromatography–mass spectrometry metabolomics data. *Bioinformatics,* **26,** 2488-2489.

Bialonska, D. & Zjawiony, J. K. (2009) Aplysinopsins-marine indole alkaloids: Chemistry, bioactivity and ecological significance. *Marine drugs,* **7,** 166-183.

Bieri, T., Onishi, M., Xiang, T., Grossman, A. R. & Pringle, J. R. (2016) Relative contributions of various cellular mechanisms to loss of algae during cnidarian bleaching. *PLoS One,* **11,** e0152693.

Bil, K. Y., Kolmakov, P. V., Parnik, T., Titlyanov, E. A. & Muskatine, L. (1991) Photosynthetic products in zooxanthellae of the symbiotic corals *Stylophora pistillata* and *Seriatopora coliendrum* situated at various depths. *Fiziologiya Rastenii (Moscow),* **38,** 846-54.

Bissett, W. P., Patch, J. S., Carder, K. L. & Lee, Z. P. (1997) Pigment packaging and Chl a-specific absorption in high-light oceanic waters. *Limnology and Oceanography,* **42,** 961-968.

Blunden, G., Gordon, S. M., Crabb, T. A., Roch, O. G., Rowan, M. G. & Wood, B. (1986) NMR spectra of betaines from marine algae. *Magnetic Resonance in Chemistry,* **24,** 965-971.

Blunden, G., Smith, B. E., Irons, M. W., Yang, M.-H., Roch, O. G. & Patel, A. V. (1992) Betaines and tertiary sulphonium compounds from 62 species of marine algae. *Biochemical Systematics and Ecology,* **20,** 373-388.

Bolten, C. J., Kiefer, P., Letisse, F., Portais, J.-C. & Wittmann, C. (2007) Sampling for metabolome analysis of microorganisms. *Analytical chemistry,* **79,** 3843-3849.

Bonnefille, B., Gomez, E., Alali, M., Rosain, D., Fenet, H., & Courant, F. (2018). Metabolomics assessment of the effects of diclofenac exposure on *Mytilus galloprovincialis*: Potential effects on osmoregulation and reproduction. *Science of the Total Environment*, **613**, 611-618. https://doi.org/10.1016/j.scitotenv.2017.09.146

Boroujerdi Arezue, F. B., Vizcaino Maria, I., Meyers, A., Pollock Elizabeth, C., Huynh Sara, L., Schock Tracey, B., Morris Pamela, J. & Bearden Daniel, W. (2009) NMR-based microbial metabolomics and the temperature-dependent coral pathogen *Vibrio coralliilyticus*. *Environmental science & technology,* **43,** 7658-64.

Bourne, D. G., Garren, M., Work, T. M., Rosenberg, E., Smith, G. W. & Harvell, C. D. (2009) Microbial disease and the coral holobiont. *Trends in Microbiology,* **17,** 554-562.

Breiman, L. (2001) Random forests. *Machine learning,* **45,** 5-32.

Breiman, L., Friedman, J., Stone, C. J. & Olshen, R. A. (1984) *Classification and regression trees,* CRC press.

Broadhurst, D., Goodacre, R., Reinke, S. N., Kuligowski, J., Wilson, I. D., Lewis, M. R., & Dunn, W. B. (2018). Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics*, *14*(6), 1-17.

Brown, B. E., Ambarsari, I., Warner, M. E., Fitt, W. K., Dunne, R. P., Gibb, S. W. & Cummings, D. G. (1999a) Diurnal change in photochemical efficiency and xanthophyll concentrations in shallow water reef corals: evidence for photoinhibition and photoprotection. *Coral Reefs,* **18,** 99-105.

Brown, B. E., Dunne, R. P., Ambarsari, I., Le Tissier, M. D. A. & Satapoomin, U. (1999b) Seasonal fluctuations in environmental factors and variations in symbiotic algae and chlorophyll pigments in four Indo-Pacific coral species. *Marine Ecology Progress Series,* **191,** 53-69.

Browne, N. K., Tay, J. K., Low, J., Larson, O. & Todd, P. A. (2015) Fluctuations in coral health of four common inshore reef corals in response to seasonal and anthropogenic changes in water quality. *Marine environmental research,* **105,** 39-52.

Bundy, J., Davey, M. & Viant, M. (2009) Environmental metabolomics: a critical review and future perspectives. *Metabolomics,* **5,** 3-21.

Caldana, C., Degenkolbe, T., Cuadros - Inostroza, A., Klie, S., Sulpice, R., Leisse, A., Steinhauser, D., Fernie, A. R., Willmitzer, L. & Hannah, M. A. (2011) High-density kinetic analysis of the metabolomic and transcriptomic response of *Arabidopsis* to eight environmental conditions. *The Plant Journal,* **67,** 869-884.

Camussi, G. & Brentjens, J. R. (1987) The role of platelet-activating factor in inflammation. In: *Platelet-activating factor and related lipid mediators*. Springer.

Canelas, A. B., Ras, C., Ten Pierick, A., van Dam, J. C., Heijnen, J. J. & Van Gulik, W. M. (2008) Leakage-free rapid quenching technique for yeast metabolomics. *Metabolomics,* **4,** 226-239.

Cates, N. & McLaughlin, J. J. A. (1979) Nutrient availability for zooxanthellae derived from physiological activities of *Condylactis* spp. *Journal of Experimental Marine Biology and Ecology,* **37,** 31-41.

Cernichiari, E., Muscatine, L. & Smith, D. C. (1969) Maltose excretion by the symbiotic algae of *Hydra viridis*. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **173,** 557-576.

Cevallos-Cevallos, J. M., Reyes-De-Corcuera, J. I., Etxeberria, E., Danyluk, M. D. & Rodrick, G. E. (2009) Metabolomic analysis in food science: a review. *Trends Food Sci Tech,* **20,** 557-566.

Chamberlain, C. A., Rubio, V. Y., & Garrett, T. J. (2019). Impact of matrix effects and ionization efficiency in non-quantitative untargeted metabolomics. *Metabolomics*, *15*(10), 1-9.

Chambers, M. C., Maclean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B. & Egertson, J. (2012) A cross-platform toolkit for mass spectrometry and proteomics. *Nature biotechnology,* **30,** 918-920.

Chapin III, F. S., Kofinas, G. P., Folke, C. & Chapin, M. C. (2009) Principles of ecosystem stewardship: resilience-based natural resource management in a changing world, Springer Science & Business Media.

Chen, H.-K., Song, S.-N., Wang, L.-H., Mayfield, A. B., Chen, Y.-J., Chen, W.-N. U. & Chen, C.-S. (2015) A compartmental comparison of major lipid species in a coral-*Symbiodinium* endosymbiosis: evidence that the coral host regulates lipogenesis of its cytosolic lipid bodies. *PLoS one,* **10,** e0132519.

Chen, H.-K., Wang, L.-H., Chen, W.-N. U., Mayfield, A. B., Levy, O., Lin, C.-S. & Chen, C.-S. (2017) Coral lipid bodies as the relay center interconnecting diel-dependent lipidomic changes in different cellular compartments. *Scientific reports,* **7,** 3244.

Christensen, N. L., Bartuska, A. M., Brown, J. H., Carpenter, S., D'Antonio, C., Francis, R., Franklin, J. F., MacMahon, J. A., Noss, R. F. & Parsons, D. J. (1996) The report of the Ecological Society of America committee on the scientific basis for ecosystem management. *Ecological applications,* **6,** 665-691.

Cienkowski, L. (1871) On the production of spores in the radiolaria. *Archiv. Fur. Mikroskop. Anatomie.,* **7,** 396-403.

Coll, J., Bowden, B., Meehan, G., Konig, G., Carroll, A., Tapiolas, D., Alino, P., Heaton, A., De Nys, R. & Leone, P. (1994) Chemical aspects of mass spawning in corals. I. Sperm-attractant molecules in the eggs of the scleractinian coral *Montipora digitata*. *Mar. Biol.,* **118,** 177-182.

Cook, C. B. & Davy, S. K. (2001) Are free amino acids responsible for the `host factor' effects on symbiotic zooxanthellae in extracts of host tissue? *Hydrobiologia,* **461,** 71-78.

Cooper, T. & Fabricius, K. (2007) Coral-based indicators of changes in water quality on nearshore coral reefs of the Great Barrier Reef. *Report to the Australian Government's Marine and Tropical Sciences Research Facility (MTSRF).*

Cooper, T., Gilmour, J. & Fabricius, K. (2009) Bioindicators of changes in water quality on coral reefs: review and recommendations for monitoring programmes. *Coral Reefs,* **28,** 589-606.

Cooper, T. F. & Ulstrup, K. E. (2009) Mesoscale variation in the photophysiology of the reef building coral *Pocillopora damicornis* along an environmental gradient. *Estuarine, Coastal and Shelf Science,* **83,** 186-196.

Crews, P., Manes, L. V. & Boehler, M. (1986) Jasplakinolide, a cyclodepsipeptide from the marine sponge, *Jaspis* sp. *Tetrahedron letters,* **27,** 2797-2800.

Cuadros-Inostroza, A., Giavalisco, P., Hummel, J., Eckardt, A., Willmitzer, L. & Peña-Cortés, H. (2010) Discrimination of wine attributes by metabolome analysis. *Analytical chemistry,* **82,** 3573-3580.

Cuif, J. P., Dauphin, Y., Freiwald, A., Gautret, P. & Zibrowius, H. (1999) Biochemical markers of zooxanthellae symbiosis in soluble matrices of skeleton of 24 Scleractinia species. *Comparative Biochemistry and Physiology,* **123,** 269-278.

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J. & Lawler, J. J. (2007) Random forests for classification in ecology. *Ecology,* **88,** 2783-2792.

D'Angelo, C. & Wiedenmann, J. (2014) Impacts of nutrient enrichment on coral reefs: new perspectives and implications for coastal management and reef survival. *Current Opinion in Environmental Sustainability,* **7,** 82-93.

D'Auria, M., Minale, L., Riccio, R. & Uriarte, E. (1988) Marine eicosanoids: Occurrence of 8-(R)-HETE in the starfish *Patiria miniata*. *Experientia,* **44,** 719-720.

D'Elia, C. F., Domotor, S. L. & Webb, K. L. (1983) Nutrient uptake kinetics of freshly isolated zooxanthellae. *Mar. Biol.,* **75,** 157-167.

Daigo, K., Nakano, Y., Casareto, B., Suzuki, Y. & Shioi, Y. (2008) High-performance liquid chromatographic analysis of photosynthetic pigments in corals: An existence of a variety of epizoic, endozoic and endolithic algae. In: *Proceedings of the 11th International Coral Reef Symposium.*

Davy, S. K., Allemand, D. & Weis, V. M. (2012) Cell biology of cnidarian-dinoflagellate symbiosis. *Microbiology and Molecular Biology Reviews,* **76,** 229-261.

de Koning, W. & van Dam, K. (1992) A Method for the determination of changes of glycolytic metabolites in yeast on a subsecond time scale using extraction at neutral pH. *Analytical biochemistry,* **204,** 118 - 123.

De Livera, A. M., Sysi-Aho, M., Jacob, L., Gagnon-Bartsch, J. A., Castillo, S., Simpson, J. A. & Speed, T. P. (2015) Statistical methods for handling unwanted variation in metabolomics data. *Analytical chemistry*.

De'ath, G., Lough, J. M. & Fabricius, K. E. (2009) Declining coral calcification on the Great Barrier Reef. *Science,* **323,** 116-119.

DeSalvo, M., Estrada, A., Sunagawa, S. & Medina, M. (2012) Transcriptomic responses to darkness stress point to common coral bleaching mechanisms. *Coral Reefs,* **31,** 215-228.

Dixon, R. A., Gang, D. R., Charlton, A. J., Fiehn, O., Kuiper, H. A., Reynolds, T. L., Tjeerdema, R. S., Jeffery, E. H., German, J. B., Ridley, W. P. & Seiber, J. N. (2006) Applications of metabolomics in agriculture. *Journal of Agricultural and Food Chemistry,* **54,** 8984-8994.

Dove, S., Ortiz, J. C., Enriquez, S., Fine, M., Fisher, P., Iglesias-Prieto, R., Thornhill, D. & Hoegh-Guldberg, O. (2006) Response of holosymbiont pigments from the scleractinian coral *Montipora monasteriata* to short-term heat stress. *Limnology and Oceanography,* **51,** 1149-1158.

Downs, C. & Downs, A. (2007) Preliminary examination of short-term cellular toxicological responses of the coral *Madracis mirabilis* to acute Irgarol 1051 exposure. *Archives of environmental contamination and toxicology,* **52,** 47-57. doi: 10.1007/s00244-005-0213-6

Dunlap, W. C. & Yamamoto, Y. (1995) Small-molecular antioxidants in marine organisms: antioxidant activity of mycosporine-glycine. *Comparative Biochemistry and Physiology,* **112,** 106-114.

Dunn, S. R., Schnitzler, C. E. & Weis, V. M. (2007) Apoptosis and autophagy as mechanisms of dinoflagellate symbiont release during cnidarian bleaching: every which way you lose. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **274,** 3079-3085.

Dunn, S. R., Thomas, M. C., Nette, G. W. & Dove, S. G. (2012) A lipidomic approach to understanding free fatty acid lipogenesis derived from dissolved inorganic carbon within cnidarian-dinoflagellate symbiosis. *PloS one,* **7,** e46801.

Dunn, W. B., Broadhurst, D., Begley, P., Zelena, E., Francis-McIntyre, S., Anderson, N., Brown, M., Knowles, J. D., Halsall, A. & Haselden, J. N. (2011) Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry. *Nature protocols,* **6,** 1060-1083.

Dykens, J. A. & Shick, J. M. (1982) Oxygen production by endosymbiotic algae controls superoxide dismutase activity in their animal host. *Nature,* **297,** 579-580.

Ettinger-Epstein, P., Motti, C. A., De Nys, R., Wright, A. D., Battershill, C. N. & Tapiolas, D. M. (2007) Acetylated Sesterterpenes from the Great Barrier Reef sponge *Luffariella variabilis*. *Journal of natural products,* **70,** 648-651.

Falkowski, P. G., Dubinsky, Z., Muscatine, L. & McCloskey, L. (1993) Population control in symbiotic corals. *BioScience,* **43,** 606-611.

Farag, M. A., Al-Mahdy, D. A., Meyer, A., Westphal, H. & Wessjohann, L. A. (2017a) Metabolomics reveals biotic and abiotic elicitor effects on the soft coral *Sarcophyton ehrenbergi* terpenoid content. *Scientific Reports,* **7,** 648. doi: 10.1038/s41598-017-00527-8

Farag, M. A., Meyer, A., Ali, S. E., Salem, M. A., Giavalisco, P., Westphal, H. & Wessjohann, L. A. (2018) Comparative metabolomics approach detects stress-specific responses during coral bleaching in soft corals. *Journal of proteome research,* **17,** 2060-2071.

Farag, M. A., Porzel, A., Al-Hammady, M. A., Hegazy, M.-E. F., Meyer, A., Mohamed, T. A., Westphal, H. & Wessjohann, L. A. (2016) Soft corals biodiversity in the Egyptian Red Sea: a comparative MS and NMR metabolomics approach of wild and aquarium grown species. *Journal of proteome research,* **15,** 1274-1287.

Farag, M. A., Westphal, H., Eissa, T. F., Wessjohann, L. A. & Meyer, A. (2017b) Effect of Oxylipins, Terpenoid Precursors and Wounding on Soft Corals' Secondary Metabolism as Analyzed via UPLC/MS and Chemometrics. *Molecules,* **22**(12), 2195. doi: 10.3390/molecules22122195

Febbo, E., Richard, C., Horlin, E., LeGall, R., Dutrieux, E. & Adenekan, A. (2012) PAM chlorophyll a fluorometry for monitoring health of corals along the coast of Qatar. In: *International Conference on Health, Safety and Environment in Oil and Gas Exploration and Production*. Society of Petroleum Engineers.

Fiehn, O. (2001) Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and Functional Genomics,* **2,** 155-168.

Fiehn, O. (2002) Metabolomics – the link between genotypes and phenotypes. *Plant Molecular Biology,* **48,** 155-171.

Fiehn, O., Kopka, J., Dormann, P., Altmann, T., Trethewey, R. N. & Willmitzer, L. (2000) Metabolite Profiling for plant functional genomics. *Nature Biotechnology,* **18,** 1157-1161.

Fitt, W. K. (2000) Cellular growth of host and symbiont in a cnidarian-zooxanthellar symbiosis. *The Biological Bulletin,* **198,** 110-120.

Fitt, W. K., Brown, B. E., Warner, M. E. & Dunne, R. P. (2001) Coral bleaching: interpretation of thermal tolerance limits and thermal thresholds in tropical corals. *Coral reefs,* **20,** 51-65.

Fitt, W. K. & Cook, C. B. (2001) Photoacclimation and the effect of the symbiotic environment on the photosynthetic response of symbiotic dinoflagellates in the tropical marine hydroid *Myrionema amboinense*. *Journal of Experimental Marine Biology and Ecology,* **256,** 15-31.

Fitt, W. K., Rees, T. A. V., Braley, R. D., Lucas, J. S. & Yellowlees, D. (1993a) Nitrogen flux in giant clams - size-dependency and relationship to zooxanthellae density and clam biomass in the uptake of dissolved inorganic nitrogen. *Mar. Biol.,* **117,** 381-386.

Fitt, W. K., Spero, H. J., Halas, J., White, M. W. & Porter, J. W. (1993b) Recovery of the Coral *Montastrea-Annularis* in the Florida Keys After the 1987 Caribbean Bleaching Event. *Coral Reefs,* **12,** 57-64.

Franzisket, L. (1973) Uptake and accumulation of nitrate and nitrite by reef corals. *Naturwissenschaften,* **60,** 552.

Fusetani, N., Asano, M., Matsunaga, S. & Hashimoto, K. (1986) Bioactive marine metabolites—XV. Isolation of aplysinopsin from the scleractinian coral *Tubastrea aurea* as an inhibitor of development of fertilized sea urchin eggs. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry,* **85,** 845-846.

Garg, N., Kapono, C. A., Lim, Y. W., Koyama, N., Vermeij, M. J., Conrad, D., Rohwer, F. & Dorrestein, P. C. (2015) Mass spectral similarity for untargeted metabolomics data analysis of complex mixtures. *Int J Mass Spectrom,* **377,** 719-727.

Garrett, T. A., Schmeitzel, J. L., Klein, J. A., Hwang, J. J. & Schwarz, J. A. (2013) Comparative lipid profiling of the cnidarian *Aiptasia pallida* and its dinoflagellate symbiont. *PloS one,* **8,** e57975.

Gates, R. D., Hoegh-Guldberg, O., McFall-Ngai, M. J., Bil, K. Y. & Muscatine, L. (1995) Free amino acids exhibit anthozoan "host factor" activity: they induce the release of photosynthate from symbiotic dinoflagellates *in vitro*. *Proceedings of the National Academy of Sciences,* **92,** 7430-7434.

GBRMPA (2004) *Great barrier reef marine park zoning plan 2003,* Great Barrier Reef Marine Park Authority.

GBRMPA (2014a) Great Barrier Reef Outlook Report 2014. Great Barrier Reef Marine Park Authority, Townsville, Australia.

GBRMPA (2014b) Great Barrier Reef Region Strategic Assessment: Program Report. Great Barrier Reef Marine Park Authority, Townsville, QLD, Australia.

GBRMPA (2014c) Great Barrier Reef Region Strategic Assessment: Strategic Assessment Report. Great Barrier Reef Marine Park Authority, Townsville, QLD, Australia.
German, J. B., Hammock, B. D., & Watkins, S. M. (2005). Metabolomics: building on a century of biochemistry to guide human health. Metabolomics, 1(1), 3-9
Gierz, S. L., Gordon, B. R. & Leggat, W. (2016) Integral light-harvesting complex expression in *Symbiodinium* within the coral *Acropora aspera* under thermal stress. *Scientific reports,* **6**.
Glasl, B., Smith, C. E., Bourne, D. G. & Webster, N. S. (2019) Disentangling the effect of host-genotype and environment on the microbiome of the coral *Acropora tenuis*. *PeerJ,* **7,** e6377.
Gochfeld, D. J. & Aeby, G. S. (2008) Antibacterial chemical defenses in Hawaiian corals provide possible protection from disease. *Marine Ecology Progress Series,* **362,** 119-128.
Goodacre, R., Vaidyanathan, S., Bianchi, G. & Kell, D. B. (2002) Metabolic profiling using direct infusion electrospray ionisation mass spectrometry for the characterisation of olive oils. *Analyst,* **127,** 1457-1462.
Goodacre, R., Vaidyanathan, S., Dunn, W. B., Harrigan, G. G. & Kell, D. B. (2004) Metabolomics by numbers: acquiring and understanding global metabolite data. *Trends in Biotechnology,* **22,** 245-252.
Gordon, B., Martin, D., Bambery, K. & Motti, C. (2018) Chemical imaging of a *Symbiodinium* sp. cell using synchrotron infrared microspectroscopy: a feasibility study. *Journal of microscopy,* **270,** 83-91.
Gordon, B. R. (2020a) coralclass: Classification of Coral Phenotypes. v1.0.1 ed. Zenodo.
Gordon, B. R. (2020b) Predicting the Photosynthetic Efficiency of *Acropora aspera*. v1.0.2 ed. Zenodo.
Gordon, B. R. & Leggat, W. (2010) *Symbiodinium*-Invertebrate Symbioses and the Role of Metabolomics. *Marine Drugs,* **8,** 2546-2568.
Gordon, B. R., Leggat, W. & Motti, C. A. (2013) Extraction Protocol for Nontargeted NMR and LC-MS Metabolomics-Based Analysis of Hard Coral and Their Algal Symbionts. In: *Metabolomics Tools for Natural Product Discovery: Methods and Protocols* (eds. U. Roessner & D. A. Dias). Humana Press, Totowa, NJ.
Gordon, B. R. & Motti, C. A. (2020a) coralmz: Coral Mass Spectral Features. v1.0.3 ed. Zenodo.
Gordon, B. R. & Motti, C. A. (2020b) Revealing Taxonomic and Diel Variation Through Untargeted Metabolomics. v0.0.1 ed. Zenodo.
Goreau, T. F., Goreau, N. I. & Yonge, C. M. (1973) On the utilization of photosynthetic products from zooxanthellae and of a dissolved amino acid in *Tridacna maxima f. elongata* (Mollusca: Bivalvia). *Journal of Zoology,* **169,** 417-454.
Gosetti, F., Mazzucco, E., Zampieri, D. & Gennaro, M. C. (2010) Signal suppression/enhancement in high-performance liquid chromatography tandem mass spectrometry. *J Chromatogr A,* **1217,** 3929-3937.
Grant, A. J., Remond, M. & Hinde, R. (1998) Low molecular-weight factor from *Plesiastrea versipora* (Scleractinia) that modifies release and glycerol metabolism of isolated symbiotic algae. *Marine Biology,* **130,** 553-557.
Grant, A. J., Remond, M., Withers, K. J. T. & Hinde, R. (2001) Inhibition of algal photosynthesis by a symbiotic coral. *Hydrobiologia,* **461,** 63-69.
Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L. & Goodacre, R. (2015) A tutorial review: Metabolomics and partial least squares-discriminant analysis–a marriage of convenience or a shotgun wedding. *Analytica chimica acta,* **879,** 10-23.
Guella, G., Mancini, I., Zibrowius, H. & Pietra, F. (1988) Novel Aplysinopsin-type Alkaloids from Scleractinian Corals of the Family Dendrophylliidae of the Mediterranean and the Philippines. Configurational-assignment criteria, stereospecific synthesis, and photoisomerization. *Helvetica chimica acta,* **71,** 773-782.
Guerriero, A., D'Ambrosio, M. & Pietra, F. (1990a) Novel Hydroxyicosatetraenoic and Hydroxyicosapentaenoic Acids and a 13 - Oxo Analog. Isolation from a Mixture of the Calcareous Red Algae *Lithothamnion corallioides* and *Lithothamnion calcareum* of Brittany Waters. *Helvetica chimica acta,* **73,** 2183-2189.
Guerriero, A., D'Ambrosio, M., Pietra, F., Ribes, O. & Duhet, D. (1990b) Hydroxyicosatetraenoic, hydroxyicosapentaenoic, hydroxydocosapentaenoic, and hydroxydocosahexaenoic acids from the sponge *Echinochalina mollis* of the Coral Sea. *Journal of natural products,* **53,** 57-61.

Gunthorpe, L. & Cameron, A. M. (1990) Intracolonial variation in toxicity in scleractinian corals. *Toxicon,* **28,** 1221-1227.
Hagedorn, M., Carter, V., Leong, J. & Kleinhans, F. (2010) Physiology and cryosensitivity of coral endosymbiotic algae (Symbiodinium). *Cryobiology,* **60,** 147-158.
Hagedorn, M., Farrell, A., Carter, V., Zuchowicz, N., Johnston, E., Padilla-Gamiño, J., Gunasekera, S. & Paul, V. (2015) Effects of toxic compounds in *Montipora capitata* on exogenous and endogenous zooxanthellae performance and fertilization success. *PLoS One,* **10,** e0118364.
Hajduk, A., Mrochem-Kwarciak, J., Skorupa, A., Ciszek, M., Heyda, A., Składowski, K. & Sokół, M. (2016) 1 H NMR based metabolomic approach to monitoring of the head and neck cancer treatment toxicity. *Metabolomics,* **12,** 102.
Hartmann, A. C., Petras, D., Quinn, R. A., Protsyuk, I., Archer, F. I., Ransome, E., Williams, G. J., Bailey, B. A., Vermeij, M. J. & Alexandrov, T. (2017) Meta-mass shift chemical profiling of metabolomes from coral reefs. *Proceedings of the National Academy of Sciences,* **114,** 11685-11690. doi.org/10.1073/pnas.1710248114
He, Q., Sun, R., Liu, H., Geng, Z., Chen, D., Li, Y., Han, J., Lin, W., Du, S. & Deng, Z. (2014) NMR-based metabolomic analysis of spatial variation in soft corals. *Marine drugs,* **12,** 1876-1890.
Hendriks, M. M., van Eeuwijk, F. A., Jellema, R. H., Westerhuis, J. A., Reijmers, T. H., Hoefsloot, H. C. & Smilde, A. K. (2011) Data-processing strategies for metabolomics studies. *TrAC Trends in Analytical Chemistry,* **30,** 1685-1698.
Heraud, P., Wood, B. R., Tobin, M. J., Beardall, J. & McNaughton, D. (2005) Mapping of nutrient-induced biochemical changes in living algal cells using synchrotron infrared microspectroscopy. *FEMS Microbiol. Lett.,* **249,** 219-225.
Higa, T., Tanaka, J., Kohagura, T. & Wauke, T. (1990) Bioactive polyacetylenes from stony corals. *Chemistry Letters,* **19,** 145-148.
Hill, R., Li, C., Jones, A., Gunn, J. & Frade, P. (2010) Abundant betaines in reef-building corals and ecological indicators of a photoprotective role. *Coral Reefs,* **29,** 869-880.
Hill, R. W., Armstrong, E. J., Florn, A. M., Li, C., Walquist, R. W. & Edward, A. (2017) Abundant betaines in giant clams (Tridacnidae) and western Pacific reef corals, including study of coral betaine acclimatization. *Marine Ecology Progress Series,* **576,** 27-41.
Hillyer, K. E., Dias, D. A., Lutz, A., Roessner, U. & Davy, S. K. (2017a) Mapping carbon fate during bleaching in a model cnidarian symbiosis: the application of ^{13}C metabolomics. *New Phytologist,* **214,** 1551-1562.
Hillyer, K. E., Dias, D., Lutz, A., Roessner, U. & Davy, S. K. (2018) ^{13}C metabolomics reveals widespread change in carbon fate during coral bleaching. *Metabolomics,* **14,** 12.
Hillyer, K. E., Dias, D. A., Lutz, A., Wilkinson, S. P., Roessner, U. & Davy, S. K. (2017b) Metabolite profiling of symbiont and host during thermal stress and bleaching in the coral *Acropora aspera*. *Coral Reefs,* **36,** 105-118.
Hillyer, K. E., Tumanov, S., Villas-Bôas, S. & Davy, S. K. (2016) Metabolite profiling of symbiont and host during thermal stress and bleaching in a model cnidarian–dinoflagellate symbiosis. *Journal of Experimental Biology,* **219,** 516-527.
Ho, C. S., Lam, C., Chan, M., Cheung, R., Law, L., Lit, L., Ng, K., Suen, M. & Tai, H. (2003) Electrospray ionisation mass spectrometry: principles and clinical applications. *The Clinical Biochemist Reviews,* **24,** 3.
Hoegh-Guldberg, O. (1999) Climate change, coral bleaching and the future of the world's coral reefs. *Marine and Freshwater Research,* **50,** 839-866.
Hoegh-Guldberg, O. (2011) Coral reef ecosystems and anthropogenic climate change. *Regional Environmental Change,* **11,** 215-227.
Hoegh-Guldberg, O. & Jason-Smith, G. (1989) Influence of the population density of zooxanthellae and supply of ammonia on the biomass and metabloic characteristics of the reef corals *Seriatopora hystrix* and *Stylophora pistillata*. *Marine Ecology Progress Series,* **57,** 173-186.
Hoegh-Guldberg, O., Mumby, P. J., Hooten, A. J., Steneck, R. S., Greenfield, P., Gomez, E., Harvell, C. D., Sale, P. F., Edwards, A. J., Caldeira, K., Knowlton, N., Eakin, C. M., Iglesias-Prieto, R., Muthiga, N., Bradbury, R. H., Dubi, A. & Hatziolos, M. E. (2007) Coral Reefs Under Rapid Climate Change and Ocean Acidification. *Science,* **318,** 1737-1742.

Hoegh-Guldberg, O., Poloczanska, E. S., Skirving, W. & Dove, S. (2017) Coral reef ecosystems under climate change and ocean acidification. *Frontiers in Marine Science,* **4,** 158.
Huey, R. B., Berrigan, D., Gilchrist, G. W. & Herron, J. C. (1999) Testing the adaptive significance of acclimation: a strong inference approach. *American Zoologist,* **39,** 323-336.
Hughes, T. P., Baird, A. H., Bellwood, D. R., Card, M., Connolly, S. R., Folke, C., Grosberg, R., Hoegh-Guldberg, O., Jackson, J. B. C., Kleypas, J., Lough, J. M., Marshall, P., Nyström, M., Palumbi, S. R., Pandolfi, J. M., Rosen, B. & Roughgarden, J. (2003) Climate Change, Human Impacts, and the Resilience of Coral Reefs. *Science,* **301,** 929-933.
Hughes, T. P., Barnes, M. L., Bellwood, D. R., Cinner, J. E., Cumming, G. S., Jackson, J. B., Kleypas, J., van de Leemput, I. A., Lough, J. M. & Morrison, T. H. (2017a) Coral reefs in the Anthropocene. *Nature,* **546,** 82-90.
Hughes, T. P., Graham, N. A., Jackson, J. B., Mumby, P. J. & Steneck, R. S. (2010) Rising to the challenge of sustaining coral reef resilience. *Trends in Ecology & Evolution,* **25,** 633-642.
Hughes, T. P., Kerry, J. T., Álvarez-Noriega, M., Álvarez-Romero, J. G., Anderson, K. D., Baird, A. H., Babcock, R. C., Beger, M., Bellwood, D. R. & Berkelmans, R. (2017b) Global warming and recurrent mass bleaching of corals. *Nature,* **543,** 373-377.
Imanishi, Y., Gerke, V. & Palczewski, K. (2004) Retinosomes: new insights into intracellular managing of hydrophobic substances in lipid bodies. *J. Cell Biol.,* **166,** 447-453.
IPCC RCP 4.5, 2014: Representative Concentration Pathway 4.5 in Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. (eds. Core Writing Team, R. K. Pachauri & L. A. Meyer). IPCC, Geneva, Switzerland, 151 pp. https://ar5-syr.ipcc.ch/topic_futurechanges.php
IPCC RCP 2.6, 2014: Representative Concentration Pathway 2.6 in Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. (eds. Core Writing Team, R. K. Pachauri & L. A. Meyer). IPCC, Geneva, Switzerland, 151 pp. https://ar5-syr.ipcc.ch/topic_futurechanges.php
Januar, H., Marraskuranto, E., Patantis, G. & Chasanah, E. (2012) LC-MS metabolomic analysis of environmental stressor impacts on the metabolite diversity in *Nephthea* spp. *Chronicles of Young Scientists,* **3,** 57-57.
Jeffrey, S. W. & Humphrey, G. F. (1975) New spectrophotometric equations for determining chlorophylls a, b, c1 and c2 in higher plants, algae and natural phytoplankton. *Biochemie und physiologie der pflanzen,* **167,** 191-194.
Jeong, T. Y., & Simpson, M. J. (2019). *Daphnia magna* metabolic profiling as a promising water quality parameter for the biological early warning system. *Water research*, **166**, 115033
Jones, A., Cantin, N., Berkelmans, R., Sinclair, B. & Negri, A. (2008) A 3D modeling method to calculate the surface areas of coral branches. *Coral Reefs,* **27,** 521-526.
Jones, R. J. (1997) Changes in zooxanthellar densities and chlorophyll concentrations in corals during and after a bleaching event. *Marine Ecology Progress Series,* **158,** 51-59.
Jones, R. J. & Hoegh-Guldberg, O. (2001) Diurnal changes in the photochemical efficiency of the symbiotic dinoflagellates (Dinophyceae) of corals: photoprotection, photoinactivation and the relationship to coral bleaching. *Plant, Cell and Environment,* **24,** 89-99.
Jones, R. J., Hoegh-Guldberg, O., Larkum, A. W. D. & Schreiber, U. (1998) Temperature-induced bleaching of corals begins with impairment of the CO_2 fixation mechanism in zooxanthellae. *Plant, Cell and Environment,* **21,** 1219-1230.
Jones, R. J., Kildea, T. & Hoegh-Guldberg, O. (1999) PAM chlorophyll fluorometry: a new *in situ* technique for stress assessment in scleractinian corals, used to examine the effects of cyanide from cyanide fishing. *Marine pollution bulletin,* **38,** 864-874.
Jung, Y., Lee, J., Kwon, J., Lee, K.-S., Ryu, D. H. & Hwang, G.-S. (2010) Discrimination of the geographical origin of beef by 1H NMR-based metabolomics. *Journal of Agricultural and Food Chemistry,* **58,** 10458-10466.
Kaniewska, P., Campbell, P. R., Kline, D. I., Rodriguez-Lanetty, M., Miller, D. J., Dove, S. & Hoegh-Guldberg, O. (2012) Major cellular and physiological impacts of ocean acidification on a reef building coral. *PloS one,* **7,** e34659.

Kashman, Y., Groweiss, A., Carmely, S., Kinamoni, Z., Czarkie, D. & Rotem, M. (1982) Recent research in marine natural products from the Red Sea. *Pure and Applied Chemistry,* **54,** 1995-2010.

Kato, M., Sakai, M., Adachi, K., Ikemoto, H. & Sano, H. (1996) Distribution of betaine lipids in marine algae. *Phytochemistry,* **42,** 1341-1345.

Kell, D. B. & Oliver, S. G. (2004) Here is the evidence, now what is the hypothesis? The complementary roles of inductive and hypothesis-driven science in the post-genomic era. *Bioessays,* **26,** 99-105.

Kellogg, R. B. & Patton, J. S. (1983) Lipid droplets, medium of energy exchange in the symbiotic anemone *Condylactis gigantea* : a model coral polyp. *Mar. Biol.,* **75,** 137-149.

Kelly, A. E., Ou, H. D., Withers, R. & D√ ∂ tsch, V. (2002) Low-conductivity buffers for high-sensitivity NMR measurements. *Journal of the American Chemical Society,* **124,** 12013-12019.

Kelly, L. W., Haas, A. F. & Nelson, C. E. (2018) Ecosystem Microbiology of Coral Reefs: Linking Genomic, Metabolomic, and Biogeochemical Dynamics from Animal Symbioses to Reefscape Processes. *MSystems,* **3,** e00162-17.

Kessner, D., Chambers, M., Burke, R., Agus, D. & Mallick, P. (2008) ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics,* **24,** 2534-2536.

Kim, S. M., Kang, S.-W., Lee, E. A., Seo, E.-K., Song, J.-I. & Pan, C.-H. (2013) Analysis of Carotenoids in 25 Indigenous Korean Coral Extracts. *Journal of Applied Biological Chemistry,* **56,** 43-48.

Kluender, C., Sans-Piche, F., Riedl, J., Altenburger, R., Haertig, C., Laue, G. & Schmitt-Jansen, M. (2009) A metabolomics approach to assessing phytotoxic effects on the green alga *Scenedesmus vacuolatus*. *Metabolomics,* **5,** 59-71.

Klueter, A., Crandall, J. B., Archer, F. I., Teece, M. A. & Coffroth, M. A. (2015) Taxonomic and Environmental Variation of Metabolite Profiles in Marine Dinoflagellates of the Genus *Symbiodinium*. *Metabolites,* **5,** 74-99. doi: 10.3390/metabo5010074

Koh, E. G. & Sweatman, H. (2000) Chemical warfare among scleractinians: bioactive natural products from *Tubastraea faulkneri* kill larvae of potential competitors. *Journal of Experimental Marine Biology and Ecology,* **251,** 141-160.

Kohavi, R. (1995) A study of cross-validation and bootstrap for accuracy estimation and model selection. In: *Ijcai*. Stanford, CA.

Kopp, C., Wisztorski, M., Revel, J., Mehiri, M., Dani, V., Capron, L., Carette, D., Fournier, I., Massi, L. & Mouajjah, D. (2015) MALDI-MS and NanoSIMS imaging techniques to study cnidarian–dinoflagellate symbioses. *Zoology,* **118,** 125-131.

Korbee, N., Houvinen, P., Figueroa, F. L., Aguilera, J. & Karsten, U. (2005) Availability of ammonium influences photosynthesis and the accumulation of mycosporine-like amino acids in two *Porphyra* species (*Bangiales, Rhodophyta*). *Mar. Biol.,* **146,** 645-654.

Kuhl, C., Tautenhahn, R., Böttcher, C., Larson, T. R. & Neumann, S. (2011) CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Analytical chemistry,* **84,** 283-289.

Kuhl, M., Cohen, Y., Dalsgaard, T., Jorgensen, B. B. & Revsbech, N. P. (1995) Microenvironment and photosynthesis of zooxanthellae in scleractinian corals studies with microsensors for O_2, pH and light. *Marine Ecology Progress Series,* **117,** 159-172.

Kuhn, M. (2008) Building predictive models in R using the caret package. *Journal of statistical software,* **28,** 1-26.

Kvalheim, O. M., Brakstad, F. & Liang, Y. (1994) Preprocessing of analytical profiles in the presence of homoscedastic or heteroscedastic noise. *Analytical chemistry,* **66,** 43-51.

LaJeunesse, T. C. (2001) Investigating the biodiversity, ecology, and phylogeny of endosymbiotic dinoflagellates in the genus *Symbiodinium* using the ITS region: in search of a "species" level marker. *Journal of Phycology,* **37,** 866-880.

LaJeunesse, T. C., Parkinson, J. E., Gabrielson, P. W., Jeong, H. J., Reimer, J. D., Voolstra, C. R. & Santos, S. R. (2018) Systematic Revision of Symbiodiniaceae Highlights the Antiquity and Diversity of Coral Endosymbionts. *Current Biology*.

Lam, V. Y., Doropoulos, C. & Mumby, P. J. (2017) The influence of resilience-based management on coral reef monitoring: A systematic review. *PloS one,* **12,** e0172064.

Latyshev, N. A., Naumenko, N. V., Svetashev, V. I. & Latypov, Y. Y. (1991) Fatty Acids of Reef-Building Corals. *Mar Ecol Prog Ser,* **76,** 295-301.

Lauridsen, M., Hansen, S. H., Jaroszewski, J. W. & Cornett, C. (2007) Human Urine as Test Material in 1H NMR-Based Metabonomics:‚Äâ Recommendations for Sample Preparation and Storage. *Analytical chemistry,* **79,** 1181-1186.

Leblond, J. D., Khadka, M., Duong, L. & Dahmen, J. L. (2015) Squishy lipids: Temperature effects on the betaine and galactolipid profiles of a C 18/C 18 peridinin-containing dinoflagellate, *Symbiodinium microadriaticum* (Dinophyceae), isolated from the mangrove jellyfish, *Cassiopea xamachana*. *Phycological research,* **63,** 219-230.

Ledlie, M., Graham, N., Bythell, J., Wilson, S., Jennings, S., Polunin, N. & Hardcastle, J. (2007) Phase shifts and the role of herbivory in the resilience of coral reefs. *Coral Reefs,* **26,** 641-653.

Leggat, W., Hoegh-Guldberg, O., Dove, S. & Yellowlees, D. (2007) Analysis of an EST library from the dinoflagellate (*Symbiodinium* sp.) symbiont of reef-building corals. *Journal of Phycology,* **43,** 1010-1021.

Leggat, W., Marendy, E. M., Baillie, B., Whitney, S. M., Ludwig, M., Badger, M. R. & Yellowlees, D. (2002) Dinoflagellate symbioses: strategies and adaptations for the acquisition and fixation of inorganic carbon. *Functional Plant Biology,* **29,** 309-322.

Legrave, N., Fahmi, E. M., Mohamed, M. & Philippe, A. (2015) Marine polyacetylenes: distribution, biological properties, and synthesis. In: *Studies in Natural Products Chemistry*. Elsevier.

Lesser, M. P. (1996) Elevated temperatures and ultraviolet radiation cause oxidative stress and inhibit photosynthesis in symbiotic dinoflagellates. *Limnology and Oceanography,* **41,** 271-283.

Lesser, M. P. (1997) Oxidative stress causes coral bleaching during exposure to elevated temperatures. *Coral Reefs,* **16,** 187-192.

Lesser, M. P. & Gorbunov, M. Y. (2001) Diurnal and bathymetric changes in chlorophyll fluorescence yields of reef corals measured in situ with a fast repetition rate fluorometer. *Marine Ecology Progress Series,* **212,** 69-77.

Levy, O., Achituv, Y., Yacobi, Y. Z., Stambler, N. & Dubinsky, Z. (2006) The impact of spectral composition and light periodicity on the activity of two antioxidant enzymes (SOD and CAT) in the coral *Favia favus*. *Journal of Experimental Marine Biology and Ecology,* **328,** 35-46.

Lewis, D. H. & Smith, D. C. (1971) The Autotrophic Nutrition of Symbiotic Marine Coelenterates with Special Reference to Hermatypic Corals. I. Movement of Photosynthetic Products between the Symbionts. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **178,** 111-129.

Liaw, A. & Wiener, M. (2002) Classification and regression by randomForest. *R news,* **2,** 18-22.

Lohr, K. E., Khattri, R. B., Guingab-Cagmat, J., Camp, E. F., Merritt, M. E., Garrett, T. J. & Patterson, J. T. (2019) Metabolomic profiles differ among unique genotypes of a threatened Caribbean coral. *Scientific reports,* **9,** 6067. https://doi.org/10.1038/s41598-019-42434-0

Lopes, T. I., Geloneze, B., Pareja, J. C., Calixto, A. R., Ferreira, M. M. & Marsaioli, A. J. (2016) "Omics" prospective monitoring of bariatric surgery: Roux-en-Y gastric bypass outcomes using mixed-meal tolerance test and time-resolved 1H NMR-based metabolomics. *Omics: a journal of integrative biology,* **20,** 415-423.

Lopez, A. & Gerwick, W. H. (1988) Ptilodene, a novel icosanoid inhibitor of 5-lifoxygenase and Na+/K+ ATPase from the red marine alga ptilotaFilicinaJ. Agardh. *Tetrahedron letters,* **29,** 1505-1506.

Luo, Y. J., Wang, L. H., Chen, W. N. U., Peng, S. E., Tzen, J. T. C., Hsiao, Y. Y., Huang, H. J., Fang, L. S. & Chen, C. S. (2009) Ratiometric imaging of gastrodermal lipid bodies in coral-dinoflagellate endosymbiosis. *Coral Reefs,* **28,** 289-301.

Lyons, B., Thain, J., Stentiford, G., Hylland, K., Davies, I. & Vethaak, A. (2010) Using biological effects tools to define good environmental status under the European Union Marine Strategy Framework Directive. *Marine pollution bulletin,* **60,** 1647-1651.

Maier, T., Kuhn, J. & Muller, C. (2010) Proposal for field sampling of plants and processing in the lab for environmental metabolic fingerprinting. *Plant Methods,* **6,** 6.

Mancini, I., Guerriero, A., Guella, G., Bakken, T., Zibrowius, H. & Pietra, F. (1999) Novel 10-Hydroxydocosapolyenoic Acids from Deep-Water Scleractinian Corals. *Helvetica chimica acta,* **82,** 677-684.

Marshall, A. T., Clode, P. L., Russell, R., Prince, K. & Stern, R. (2007) Electron and ion microprobe analysis of calcium distribution and transport in coral tissues. *Journal of Experimental Biology,* **210,** 2453-2463.

Martin, S. & Parton, R. (2006) Lipid droplets: a unified view of a dynamic organelle. *Nature Reviews. Molecular Cell Biology,* **7,** 373-378.

Marubini, F. & Davies, P. S. (1996) Nitrate increases zooxanthellae population density and reduces skeletogenesis in corals. *Mar. Biol.,* **127,** 319-328.

Marubini, F., Ferrier-Pages, C., Furla, P. & Allemand, D. (2008) Coral calcification responds to seawater acidification: a working hypothesis towards a physiological mechanism. *Coral Reefs,* **27,** 491-499.

Mashego, M. R., Rumbold, K., De Mey, M., Vandamme, E., Soetaert, W. & Heijnen, J. J. (2007) Microbial metabolomics: past, present and future methodologies. *Biotechnol Lett,* **29,** 1-16.

Matthews, J. L., Crowder, C. M., Oakley, C. A., Lutz, A., Roessner, U., Meyer, E., Grossman, A. R., Weis, V. M. & Davy, S. K. (2017) Optimal nutrient exchange and immune responses operate in partner specificity in the cnidarian-dinoflagellate symbiosis. *Proceedings of the National Academy of Sciences,* **114,** 13194-13199. https://doi.org/10.1073/pnas.1710733114

Matthews, J. L., Oakley, C. A., Lutz, A., Hillyer, K. E., Roessner, U., Grossman, A. R., Weis, V. M. & Davy, S. K. (2018) Partner switching and metabolic flux in a model cnidarian–dinoflagellate symbiosis. *Proceedings of the Royal Society B,* **285,** 20182336. doi: 10.1098/rspb.2018.2336

Maxwell, K. & Johnson, G. N. (2000) Chlorophyll fluorescence—a practical guide. *J. Exp. Bot.,* **51,** 659-668.

Maynard, J., Van Hooidonk, R., Eakin, C. M., Puotinen, M., Garren, M., Williams, G., Heron, S. F., Lamb, J., Weil, E. & Willis, B. (2015) Projections of climate conditions that increase coral disease susceptibility and pathogen abundance and virulence. *Nature Climate Change,* **5,** 688.

Middlebrook, R., Hoegh-Guldberg, O. & Leggat, W. (2008) The effect of thermal history on the susceptibility of reef-building corals to thermal stress. *Journal of Experimental Biology,* **211,** 1050-1056.

Miller, T. R., Hnilicka, K., Dziedzic, A., Desplats, P. & Belas, R. (2004) Chemotaxis of *Silicibacter* sp. strain TM1040 toward dinoflagellate products. *Applied and Environmental Microbiology,* **70,** 4692-4701.

Moore, M. N., Depledge, M. H., Readman, J. W. & Leonard, D. P. (2004) An integrated biomarker-based strategy for ecotoxicological evaluation of risk in environmental management. *Mutation Research: Fundamental and Molecular Mechanisms of Mutagenesis,* **552,** 247-268.

Morice, C. P., Kennedy, J. J., Rayner, N. A. & Jones, P. D. (2012) Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research: Atmospheres,* **117,** D08101. doi:10.1029/2011JD017187

Müller, P., Li, X.-P. & Niyogi, K. K. (2001) Non-photochemical quenching. A response to excess light energy. *Plant Physiology,* **125,** 1558-1566.

Munday, P. L., Dixson, D. L., Donelson, J. M., Jones, G. P., Pratchett, M. S., Devitsina, G. V. & Døving, K. B. (2009) Ocean acidification impairs olfactory discrimination and homing ability of a marine fish. *Proceedings of the National Academy of Sciences,* **106,** 1848-1852.

Murata, M., Miyagawa-Kohshima, K., Nakanishi, K. & Naya, Y. (1986) Characterization of compounds that induce symbiosis between sea anemone and anemone fish. *Science,* **234,** 585-587.

Muscatine, L. (1965) Symbiosis of hydra and algae - III. Extracellular products of the algae. *Comparative Biochemistry and Physiology,* **16,** 77-92.

Muscatine, L. (1967) Glycerol excretion by symbiotic algae from corals and *Tridacna* and its control by the host. *Science,* **156,** 516-519.

Muscatine, L., Blackburn, A., Gates, R. D., Baghdasarian, G. & Allemand, D. (1998) Cell-specific density of symbiotic dinoflagellates in tropical anthozoans. *Coral Reefs,* **17,** 329-337.

Muscatine, L. & D'Elia, C. F. (1978) The uptake, retention, and release of ammonium by reef corals. *Limnology and Oceanography,* **23,** 725-34.
Muscatine, L., Falkowski, P. G., Dubinsky, Z., Cook, P. A. & McCloskey, L. R. (1989) The Effect of External Nutrient Resources on the Population-Dynamics of Zooxanthellae in A Reef Coral. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **236,** 311-324.
Muscatine, L. & Hand, C. (1958) Direct Evidence for the Transfer of Materials from Symbiotic Algae to the Tissues of a Coelenterate. *Proceedings of the National Academy of Sciences,* **44,** 1259-1263.
Muscatine, L., Karakashian, S. J. & Karakashian, M. W. (1967) Soluble extracellular products of algae symbiotic with a ciliate, a sponge and a mutant hydra. *Comparative Biochemistry and Physiology,* **20,** 1-6.
Muscatine, L. & Lenhoff, H. M. (1963) Symbiosis: On the role of algae symbiotic with *Hydra*. *Science,* **142,** 956-958.
Muscatine, L. & Porter, J. W. (1977) Reef corals: mutualistic symbioses adapted to nutrient-poor environments. *BioScience,* **27,** 454-460.
Nakamura, H., Asari, T., Ohizumi, Y., Kobayashi, J., Yamasu, T. & Murai, A. (1993) Isolation of Zooxanthellatoxins, Novel Vasoconstrictive Substances from the Zooxanthella *Symbiodinium* sp. *Toxicon,* **31,** 371-376.
Nakamura, H., Kawase, Y., Maruyama, K. & Murai, A. (1998) Studies on polyketide metabolites of a symbiotic dinoflagellate, *Symbiodinium* sp.: a new C30 marine alkaloid, zooxanthellamine, a plausible precursor for *zoanthid* alkaloids. *Bulletin of the Chemical Society of Japan,* **71,** 781-787.
Nakamura, H., Kobayashi, J. & Hirata, Y. (1982) Separation Of Mycosporine-Like Amino-Acids In Marine Organisms Using Reversed-Phase High-Performance Liquid-Chromatography. *Journal of Chromatography,* **250,** 113-118.
Negri, R. (2015) Polyacetylenes from terrestrial plants and fungi: recent phytochemical and biological advances. *Fitoterapia,* **106,** 92-109.
Nichols, J. D. & Williams, B. K. (2006) Monitoring for conservation. *Trends in ecology & evolution,* **21,** 668-673.
O'Mahoney, J., Simes, R., Redhill, D., Heaton, K., Atkinson, C., Hayward, E. & Nguyen, M. (2017) At what price? The economic, social and icon value of the Great Barrier Reef. Deloitte Access Economics, http://hdl.handle.net/11017/3205
Obata, T. & Fernie, A. R. (2012) The use of metabolomics to dissect plant responses to abiotic stresses. *Cellular and Molecular Life Sciences,* **69,** 3225-3243.
Obura, D. O. (2009) Reef corals bleach to resist stress. *Marine pollution bulletin,* **58,** 206-212.
Ochsenkühn, M. A., Schmitt-Kopplin, P., Harir, M. & Amin, S. A. (2018) Coral metabolite gradients affect microbial community structures and act as a disease cue. *Communications biology,* **1,** 184.
Ogawa, D., Bobeszko, T., Ainsworth, T. & Leggat, W. (2013) The combined effects of temperature and CO_2 lead to altered gene expression in *Acropora aspera*. *Coral reefs,* **32,** 895-907.
Okamoto, M., Nojima, S., Furushima, Y. & Nojima, H. (2005) Evaluation of coral bleaching condition in situ using an underwater pulse amplitude modulated fluorometer. *Fisheries Science,* **71,** 847-854.
Okuda, R. K., Klein, D., Kinnel, R. B., Li, M. & Scheuer, P. (1982) Marine natural products: the past twenty years and beyond. *Pure and Applied Chemistry,* **54,** 1907-1914.
Oldeman, R. A. A. (1987) *Nederlands Bosbouwtijdschrift*, **59**(9/10), 349-353.
Oliver, S. G., Winson, M. K., Kell, D. B. & Baganz, F. (1998) Systematic functional analysis of the yeast genome. *Trends in Biotechnology,* **16,** 373-378.
Onodera, K.-i., Nakamura, H., Oba, Y. & Ojika, M. (2003) Zooxanthellamide A, a novel polyhydroxy metabolite from a marine dinoflagellate of *Symbiodinium* sp. *Tetrahedron,* **59,** 1067-1071.
Onodera, K.-i., Nakamura, H., Oba, Y. & Ojika, M. (2004) Zooxanthellamide B, a novel large polyhydroxy metabolite from a marine dinoflagellate of *Symbiodinium* sp. *Bioscience, Biotechnology, and Biochemistry,* **68,** 955-958.
Ovenden, S. P. B., Gordon, B. R., Bagas, C. K., Muir, B., Rochfort, S. & Bourne, D. J. (2010) A Study of the Metabolome of *Ricinus communis* for Forensic Applications. *Australian Journal of Chemistry,* **63,** 8-21.

Oytam, Y., Sobhanmanesh, F., Duesing, K., Bowden, J. C., Osmond-McLeod, M. & Ross, J. (2016) Risk-conscious correction of batch effects: maximising information extraction from high-throughput genomic datasets. *BMC bioinformatics,* **17,** 332.

Pandolfi, J. M., Connolly, S. R., Marshall, D. J. & Cohen, A. L. (2011) Projecting coral reef futures under global warming and ocean acidification. *Science,* **333,** 418-422.

Papageorgiou, G. C. & Murata, N. (1995) The unusually strong stabilizing effects of glycine betaine on the structure and function of the oxygen-evolving photosystem II complex. *Photosynthesis Research,* **44,** 243-252.

Parkinson, J. E. & Baums, I. B. (2014) The extended phenotypes of marine symbioses: ecological and evolutionary consequences of intraspecific genetic diversity in coral–algal associations. *Frontiers in microbiology,* **5,** 445.

Patton, J. S., Abraham, S. & Benson, A. A. (1977) Lipogenesis in the intact coral *Pocillopora capitata* and its isolated zooxanthellae: Evidence for a light-driven carbon cycle between symbiont and host. *Mar. Biol.,* **44,** 235-247.

Patton, J. S., Battey, J. F., Rigler, M. W., Porter, J. W., Black, C. C. & Burris, J. E. (1983) A comparison of the metabolism of bicarbonate ^{14}C and acetate 1-^{14}C and the variability of species lipid composition in reef corals. *Mar. Biol.,* **75,** 121-130.

Peng, S.-E., Chen, C.-S., Song, Y.-F., Huang, H.-T., Jiang, P.-L., Chen, W.-N. U., Fang, L.-S. & Lee, Y.-C. (2011) Assessment of metabolic modulation in free-living versus endosymbiotic *Symbiodinium* using synchrotron radiation-based infrared microspectroscopy. *Biol. Lett.,* **8,** 434-437.

Petrou, K., Nielsen, D. A. & Heraud, P. (2018) Single-Cell Biomolecular Analysis of Coral Algal Symbionts Reveals Opposing Metabolic Responses to Heat Stress and Expulsion. *Frontiers in Marine Science,* **5,** 110.

Polley, E. C. & Van der Laan, M. J. (2010) Super learner in prediction.

Prince, J. T. & Marcotte, E. M. (2006) Chromatographic alignment of ESI-LC-MS proteomics data sets by ordered bijective interpolated warping. *Analytical chemistry,* **78,** 6140-6152.

Putnam, H. M., Davidson, J. M. & Gates, R. D. (2016) Ocean acidification influences host DNA methylation and phenotypic plasticity in environmentally susceptible corals. *Evolutionary applications,* **9,** 1165-1178.

Quinn, R. A., Vermeij, M. J., Hartmann, A. C., d'Auriac, I. G., Benler, S., Haas, A., Quistad, S. D., Lim, Y. W., Little, M. & Sandin, S. (2016) Metabolomics of reef benthic interactions reveals a bioactive lipid involved in coral defence. *Proc. R. Soc. B.* **283**, 20160469. https://royalsocietypublishing.org/doi/10.1098/rspb.2016.0469

R Core Team (2018) R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Raamsdonk, L. M., Teusink, B., Broadhurst, D., Zhang, N. S., Hayes, A., Walsh, M. C., Berden, J. A., Brindle, K. M., Kell, D. B., Rowland, J. J., Westerhoff, H. V., van Dam, K. & Oliver, S. G. (2001) A functional genomics strategy that uses metabolome data to reveal the phenotype of silent mutations. *Nature Biotechnology,* **19,** 45-50.

Raina, J.-B., Dinsdale, E. A., Willis, B. L. & Bourne, D. G. (2010) Do the organic sulfur compounds DMSP and DMS drive coral microbial associations? *Trends Microbiol.,* **18,** 101-108.

Ralph, P. J., Hill, R., Doblin, M. A. & Davy, S. K. (2015) Theory and Application of Pulse Amplitude Modulated Chlorophyll Fluorometry in Coral Health Assessment. *Diseases of Coral***,** 506-523.

Rands, M. L., Loughman, B. C. & Douglas, A. E. (1993) The Symbiotic interface in an alga invertebrate symbiosis. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **253,** 161-165.

Rees, T. A. V., Fitt, W. K., Baillie, B. & Yellowlees, D. (1993) A method for temporal measurement of haemolymph composition in the giant clam symbiosis and its application to glucose and glycerol levels during a diel cycle. *Limnology and Oceanography,* **38,** 213-217.

Remane, D., Wissenbach, D. K., Meyer, M. R., & Maurer, H. H. (2010). Systematic investigation of ion suppression and enhancement effects of fourteen stable-isotope-labelled internal standards by their native analogues using atmospheric-pressure chemical ionization and

electrospray ionization and the relevance for multi-analyte liquid chromatographic/mass spectrometric procedures. *Rapid Communications in Mass Spectrometry*, **24(7)**, 859-867
Reusch, T. B. (2014) Climate change in the oceans: evolutionary versus phenotypically plastic responses of marine animals and plants. *Evolutionary Applications,* **7,** 104-122.
Reynolds, L. J., Hughes, L. L., Yu, L. & Dennis, E. A. (1994) 1-Hexadecyl-2-arachidonoylthio-2-deoxy-sn-glycero-3-phosphorylcholine as a substrate for the microtiterplate assay of human cytosolic phospholipase A2. *Analytical biochemistry,* **217,** 25-32.
Rhodes, D. & Hanson, A. (1993) Quaternary ammonium and tertiary sulfonium compounds in higher plants. *Annual review of plant biology,* **44,** 357-384.
Rigas, B. & Levine, L. (1983) Human salivary eicosanoids: circadian variation. *Biochemical and biophysical research communications,* **115,** 201-205.
Rochfort, S. (2005) Metabolomics reviewed: A new "omics" platform technology for systems biology and implications for natural products research. *Journal of Natural Products,* **68,** 1813-1820.
Rochfort, S., Ezernieks, V., Bastian, S. E. P. & Downey, M. O. (2010) Sensory attributes of wine influenced by variety and berry shading discriminated by NMR metabolomics. *Food Chem,* **121,** 1296-1304.
Roder, C., Wu, Z., Richter, C. & Zhang, J. (2013) Coral reef degradation and metabolic performance of the scleractinian coral *Porites lutea* under anthropogenic impact along the NE coast of Hainan Island, South China Sea. *Cont Shelf Res,* **57,** 123-131.
Rohwer, F., Seguritan, V., Azam, F. & Knowlton, N. (2002) Diversity and distribution of coral-associated bacteria. *Marine Ecology Progress Series,* **243,** 1-10.
Rosic, N., Kaniewska, P., Chan, C.-K. K., Ling, E. Y. S., Edwards, D., Dove, S. & Hoegh-Guldberg, O. (2014) Early transcriptional changes in the reef-building coral *Acropora aspera* in response to thermal and nutrient stress. *BMC genomics,* **15,** 1052.
Roth, E., Jeon, K. & Stacey, G. (1988) Homology in endosymbiotic systems: the term 'symbiosome' In Molecular Genetics of Plant-Microbe Interactions, ed. R Palacios, DPS Verma, pp. 220-25. SI. Paul, MN: *Am. Phytopathol. Soc.*
Sáez, M. J. & Lagunas, R. (1976) Determination of intermediary metabolites in yeast. Critical examination of the effect of sampling conditions and recommendations for obtaining true levels. *Mol. Cell. Biochem.,* **13,** 73-78.
Schiller, H. & Dau, H. (2000) Preparation protocols for high-activity photosystem II membrane particles of green algae and higher plants, pH dependence of oxygen evolution and comparison of the S2-state multiline signal by X-band EPR spectroscopy. *Journal of Photochemistry and Photobiology B: Biology,* **55,** 138-144.
Shepherd, K. D. & Walsh, M. G. (2007) Infrared spectroscopy—enabling an evidence-based diagnostic surveillance approach to agricultural and environmental management in developing countries. *Journal of Near Infrared Spectroscopy,* **15,** 1-19.
Shepherd, L. V. T., Hackett, C. A., Alexander, C. J., Sungurtas, J. A., Pont, S. D. A., Stewart, D., McNicol, J. W., Wilcockson, S. J., Leifert, C. & Davies, H. V. (2014) Effect of agricultural production systems on the potato metabolome. *Metabolomics,* **10,** 212-224.
Shimizu, T. & Wolfe, L. S. (1990) Arachidonic acid cascade and signal transduction. *Journal of neurochemistry,* **55,** 1-15.
Simpson, M. J. & McKelvie, J. R. (2009) Environmental metabolomics: new insights into earthworm ecotoxicity and contaminant bioavailability in soil. *Analytical and Bioanalytical Chemistry,* **394,** 137-149.
Sjoblad, R. D. & Mitchell, R. (1979) Chemotactic responses of *Vibrio alginolyticus* to algal extracellular products. *Canadian Journal of Microbiology,* **25,** 964-967.
Skelton, D., Ekman, D., Martinovic-Weigelt, D., Ankley, G., Villeneuve, D., Teng, Q. & Collette, T. (2014) Metabolomics for in situ environmental monitoring of surface waters impacted by contaminants from both point and nonpoint sources. *Environmental science & technology,* **48,** 2395-2403.
Smith, C. A., Want, E. J., O'Maille, G., Abagyan, R. & Siuzdak, G. (2006) XCMS: Processing mass spectrometry data for metabolite profiling using Nonlinear peak alignment, matching, and identification. *Analytical chemistry,* **78,** 779 - 787.

Smith, R., Ventura, D. & Prince, J. T. (2015) LC-MS alignment in theory and practice: a comprehensive algorithmic review. *Briefings in bioinformatics,* **16,** 104-117.

Sogin, E. M., Anderson, P., Williams, P., Chen, C.-S. & Gates, R. D. (2014) Application of ^{1}H-NMR Metabolomic Profiling for Reef-Building Corals. *PloS one,* **9,** e111274.

Sogin, E. M., Putnam, H. M., Anderson, P. E. & Gates, R. D. (2016) Metabolomic signatures of increases in temperature and ocean acidification from the reef-building coral, *Pocillopora damicornis*. *Metabolomics,* **12,** 71. https://doi.org/10.1007/s11306-016-0987-8

Sogin, E. M., Putnam, H. M., Nelson, C. E., Anderson, P. & Gates, R. D. (2017) Correspondence of coral holobiont metabolome with symbiotic bacteria, archaea and *Symbiodinium* communities. *Environmental microbiology reports,* **9,** 310-315. https://doi.org/10.1111/1758-2229.12541

Sorek, M., Yacobi, Y. Z., Roopin, M., Berman-Frank, I. & Levy, O. (2013) Photosynthetic circadian rhythmicity patterns of Symbiodium, the coral endosymbiotic algae. *Proceedings of the Royal Society B: Biological Sciences,* **280,** 20122942.

Stambler, N., Cox, E. F. & Vago, R. (1994) Effect of ammonium enrichment on respiration, zooxanthellar densities, and pigment concentrations in two species of Hawaiian corals. *Pacific Science,* **48,** 284-290.

Stat, M., Carter, D. & Hoegh-Guldberg, O. (2006) The evolutionary history of *Symbiodinium* and scleractinian hosts - Symbiosis, diversity, and the effect of climate change. *Perspectives in Plant Ecology, Evolution and Systematics,* **8,** 23-43.

Steffen, W., Broadgate, W., Deutsch, L., Gaffney, O. & Ludwig, C. (2015) The trajectory of the Anthropocene: the great acceleration. *The Anthropocene Review,* **2,** 81-98.

Stekhoven, D. J. & Bühlmann, P. (2011) MissForest—non-parametric missing value imputation for mixed-type data. *Bioinformatics,* **28,** 112-118.

Stella, J. S., Pratchett, M. S., Hutchings, P. A. & Jones, G. P. (2011) Coral-associated invertebrates: diversity, ecological importance and vulnerability to disturbance. *Oceanography and Marine Biology: An Annual Review,* **49,** 43-104.

Stien, D., Clergeaud, F., Rodrigues, A. M., Lebaron, K., Pillot, R., Romans, P., Fagervold, S. & Lebaron, P. (2018) Metabolomics reveal that octocrylene accumulates in *Pocillopora damicornis* tissues as fatty acid conjugates and triggers coral cell mitochondrial dysfunction. *Analytical chemistry,* **91,** 990-995.

Stimson, J. & Kinzie Iii, R. A. (1991) The temporal pattern and rate of release of zooxanthellae from the reef coral *Pocillopora damicornis* (Linnaeus) under nitrogen-enrichment and control conditions. *Journal of Experimental Marine Biology and Ecology,* **153,** 63-74.

Suárez-Ulloa, V., Fernández-Tajes, J., Manfrin, C., Gerdol, M., Venier, P. & Eirín-López, J. (2013) Bivalve omics: state of the art and potential applications for the biomonitoring of harmful marine compounds. *Marine drugs,* **11,** 4370-4389.

Sugiura, T., Fukuda, T., Miyamoto, T. & Waku, K. (1992) Distribution of alkyl and alkenyl ether-linked phospholipids and platelet-activating factor-like lipid in various species of invertebrates. *Biochimica et Biophysica Acta (BBA)-Lipids and Lipid Metabolism,* **1126,** 298-308.

Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W. M., Fiehn, O., Goodacre, R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J. C., Marriott, P., Nicholls, A. W., Reily, M. D., Thaden, J. J. & Viant, M. R. (2007) Proposed minimum reporting standards for chemical analysis. *Metabolomics,* **3,** 211-221.

Sumner, L. W., Lei, Z., Nikolau, B. J., & Saito, K. (2015). Modern plant metabolomics: advanced natural product gene discoveries, improved technologies, and future prospects. Natural product reports, 32(2), 212-229.

Sunda, W., Kieber, D., Kiene, R. & Huntsman, S. (2002) An antioxidant function for DMSP and DMS in marine algae. *Nature,* **418,** 317-320.

Swan, H. B., Deschaseaux, E. S., Jones, G. B. & Eyre, B. D. (2017) Quantification of dimethylsulfoniopropionate (DMSP) in *Acropora* spp. of reef-building coral using mass spectrometry with deuterated internal standard. *Analytical and Bioanalytical Chemistry,* **409,** 1929-1942.

Szymańska, E., Saccenti, E., Smilde, A. K. & Westerhuis, J. A. (2012) Double-check: validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics,* **8,** 3-16.

Tang, C.-H., Shi, S.-H., Lin, C.-Y., Li, H.-H. & Wang, W.-H. (2018) Using lipidomic methodology to characterize coral response to herbicide contamination and develop an early biomonitoring model. *Science of The Total Environment*, 648, 1275-1283.

Tapiolas, D., Motti, C., Holloway, P. & Boyle, S. (2010) High levels of acrylate in the Great Barrier Reef coral *Acropora millepora*. *Coral Reefs,* **29,** 621-625.

Tapiolas, D. M., Raina, J.-B., Lutz, A., Willis, B. L. & Motti, C. A. (2013) Direct measurement of dimethylsulfoniopropionate (DMSP) in reef-building corals using quantitative nuclear magnetic resonance (qNMR) spectroscopy. *Journal of Experimental Marine Biology and Ecology,* **443,** 85-89.

Tautenhahn, R., Böttcher, C. & Neumann, S. (2008) Highly sensitive feature detection for high resolution LC/MS. *BMC bioinformatics,* **9,** 504.

Taylor, P. J. (2005) Matrix effects: the Achilles heel of quantitative high-performance liquid chromatography–electrospray–tandem mass spectrometry. *Clin. Biochem.,* **38,** 328-334.

Tchernov, D., Gorbunov, M. Y., de Vargas, C., Yadav, S. N., Milligan, A. J., Haggblom, M. & Falkowski, P. G. (2004) Membrane lipids of symbiotic algae are diagnostic of sensitivity to thermal bleaching in corals. *Proc. Natl. Acad. Sci.,* **101,** 13531-13535.

Thrall, P. H., Oakeshott, J. G., Fitt, G., Southerton, S., Burdon, J. J., Sheppard, A., Russell, R. J., Zalucki, M., Heino, M. & Ford Denison, R. (2011) Evolution in agriculture: the application of evolutionary approaches to the management of biotic interactions in agro - ecosystems. *Evolutionary Applications,* **4,** 200-215.

Thrippleton, M. J., Edden, R. A. E. & Keeler, J. (2005) Suppression of strong coupling artefacts in J-spectra. *J Magn Reson,* **174,** 97-109.

Todd, P. A. (2008) Morphological plasticity in scleractinian corals. *Biological Reviews,* **83,** 315-337.

Tonnesen, M. G., Jubiz, W., Moore, J. G. & Frailey, J. (1974) Circadian variation of prostaglandin E (PGE) production in human gastric juice. *The American journal of digestive diseases,* **19,** 644-648.

Trench, R. K. (1971a) The Physiology and Biochemistry of Zooxanthellae Symbiotic with Marine Coelenterates. I. The Assimilation of Photosynthetic Products of Zooxanthellae by Two Marine Coelenterates. *Proceedings of the Royal Society of London. Series B. Biological Sciences,* **177,** 225-235.

Trench, R. K. (1971b) The Physiology and Biochemistry of Zooxanthellae Symbiotic with Marine Coelenterates. II. Liberation of Fixed ^{14}C by Zooxanthellae *in vitro*. *Proceedings of the Royal Society of London. Series B. Biological Sciences,* **177,** 237-250.

Trench, R. K. (1971c) The Physiology and Biochemistry of Zooxanthellae Symbiotic with Marine Coelenterates. III. The Effect of Homogenates of Host Tissues on the Excretion of Photosynthetic Products *in vitro* by Zooxanthellae from Two Marine Coelenterates. *Proceedings of the Royal Society of London. Series B. Biological Sciences,* **177,** 251-264.

Trench, R. K. (1979) The cell biology of plant-animal symbiosis. *Annual Review of Plant Physiology and Plant Molecular Biology,* **30,** 485-531.

Trevena, A. J., Jones, G. B., Wright, S. W. & van den Enden, R. L. (2000) Profiles of DMSP, algal pigments, nutrients and salinity in pack ice from eastern Antarctica. *Journal of Sea Research,* **43,** 265-273.

Uemura, Y., Lee, T. & Snyder, F. (1991) A coenzyme A-independent transacylase is linked to the formation of platelet-activating factor (PAF) by generating the lyso-PAF intermediate in the remodeling pathway. *Journal of Biological Chemistry,* **266,** 8268-8272.

Umlauf, E., Csaszar, E., Moertelmaier, M., Schuetz, G. J., Parton, R. G. & Prohaska, R. (2004) Association of stomatin with lipid bodies. *Journal of Biological Chemistry,* **279,** 23699-23709.

Van Assche, R., Temmerman, L., Dias, D. A., Boughton, B., Boonen, K., Braeckman, B. P., Schoofs, L. & Roessner, U. (2015) Metabolic profiling of a transgenic *Caenorhabditis elegans* Alzheimer model. *Metabolomics,* **11,** 477-486.

van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K. & van der Werf, M. J. (2006) Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC genomics,* **7,** 142.

van Oppen, M. J., Oliver, J. K., Putnam, H. M. & Gates, R. D. (2015) Building coral reef resilience through assisted evolution. *Proceedings of the National Academy of Sciences,* **112,** 2307-2313. doi.org/10.1073/pnas.1422301112

van Woesik, R., Van Woesik, K., Van Woesik, L. & Van Woesik, S. (2013) Effects of ocean acidification on the dissolution rates of reef-coral skeletons. *PeerJ,* **1,** e208. 10.7717/peerj.208

Venn, A. A., Wilson, M. A., Trapido-Rosenthal, H. G., Keely, B. J. & Douglas, A. E. (2006) The impact of coral bleaching on the pigment profile of the symbiotic alga, Symbiodinium. *Plant, cell & environment,* **29,** 2133-2142.

Viant, M. R. (2007) Metabolomics of aquatic organisms: the new 'omics' on the block. *Marine Ecology: Progress Series,* **332,** 301-306.

Viant, M. R. (2008) Environmental metabolomics using 1H-NMR spectroscopy. *Methods in Molecular Biology (Totowa, NJ, United States),* **410,** 137-150.

Viant, M. R., Bearden, D. W., Bundy, J. G., Burton, I. W., Collette, T. W., Ekman, D. R., Ezernieks, V., Karakach, T. K., Lin, C. Y., Rochfort, S., de Ropp, J. S., Teng, Q., Tjeerdema, R. S., Walter, J. A. & Wu, H. (2009) International NMR-based environmental metabolomics intercomparison exercise. *Environmental Science & Technology,* **43,** 219-225.

Viant, M. R. & Sommer, U. (2013) Mass spectrometry based environmental metabolomics: a primer and review. *Metabolomics,* **9,** 144-158.

Vohsen, S. A., Fisher, C. R. & Baums, I. B. (2019) Metabolomic richness and fingerprints of deep-sea coral species and populations. *Metabolomics,* **15,** 34.

Von Holt, C. (1968) Uptake of glycine and release of nucleoside polyphosphates by zooxanthellae. *Comparative Biochemistry and Physiology,* **26,** 1071-9.

Von Holt, C. & Von Holt, M. (1968a) Secretion of organic compounds by zooxanthellae isolated from various types of *Zoanthus*. *Comparative Biochemistry and Physiology,* **24,** 83-92.

Von Holt, C. & Von Holt, M. (1968b) Transfer of photosynthetic products from zooxanthellae to coelenterate hosts. *Comparative Biochemistry and Physiology,* **24,** 73-81.

Vuckovic, D. (2012) Current trends and challenges in sample preparation for global metabolomics using liquid chromatography–mass spectrometry. *Analytical and Bioanalytical Chemistry,* **403,** 1523-1548.

Wakisaka, A., Abdoul-Carime, H., Yamamoto, Y. & Kiyozumi, Y. (1998) Non-ideality of binary mixtures Water-methanol and water-acetonitrile from the viewpoint of clustering structure. *Journal of the Chemical Society, Faraday Transactions,* **94,** 369-374.

Wang, J. H., Byun, J. & Pennathur, S. (2010) Analytical approaches to metabolomics and applications to systems biology. In: *Semin. Nephrol.* Elsevier.

Wang, J. T. & Douglas, A. E. (1998) Nitrogen recycling or nitrogen conservation in an alga-invertebrate symbiosis? *Journal of Experimental Biology,* **201,** 2445-2453.

Wang, J. T. & Douglas, A. E. (1999) Essential amino acid synthesis and nitrogen recycling in an alga-invertebrate symbiosis. *Mar. Biol.,* **135,** 219-222.

Warner, M., Chilcoat, G., McFarland, F. & Fitt, W. (2002) Seasonal fluctuations in the photosynthetic capacity of photosystem II in symbiotic dinoflagellates in the Caribbean reef-building coral Montastraea. *Mar. Biol.,* **141,** 31-38.

Warner, M., Fitt, W. & Schmidt, G. (1996) The effects of elevated temperature on the photosynthetic efficiency of zooxanthellae in hospite from four different species of reef coral: a novel approach. *Plant, Cell & Environment,* **19,** 291-299.

Warner, M. E., Fitt, W. K. & Schmidt, G. W. (1999) Damage to photosystem II in symbiotic dinoflagellates: a determinant of coral bleaching. *Proceedings of the National Academy of Sciences,* **96,** 8007-8012.

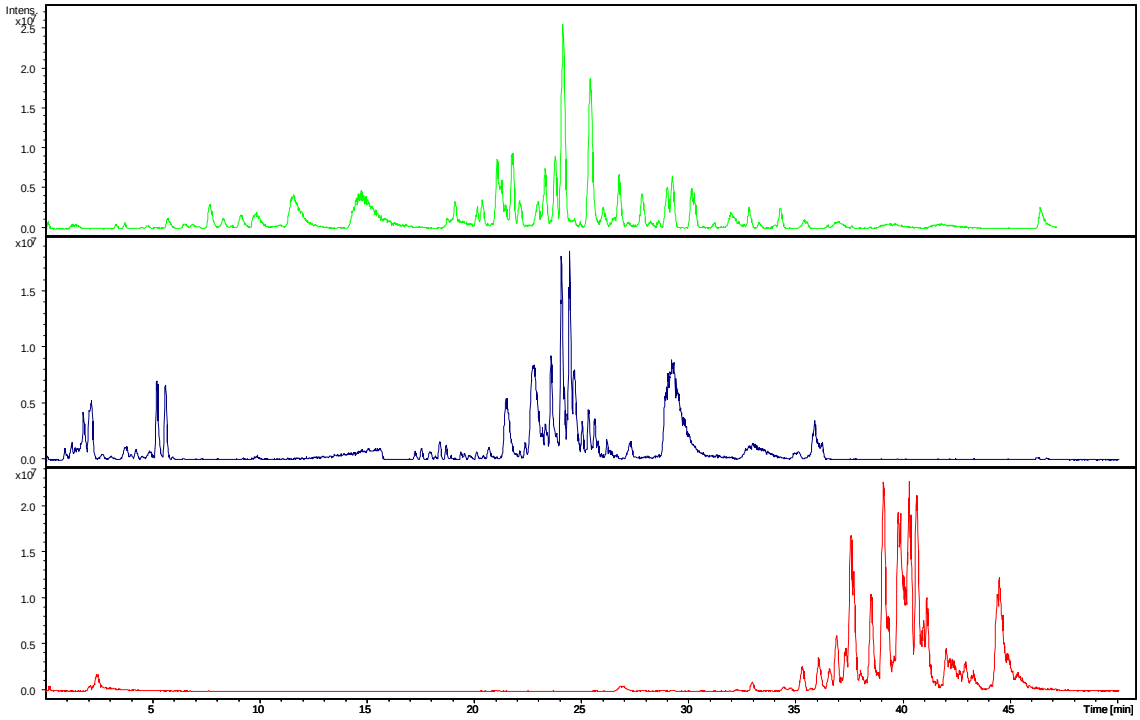
Warner, M. E., LaJeunesse, T. C., Robison, J. D. & Thur, R. M. (2006) The ecological distribution and comparative photobiology of symbiotic dinoflagellates from reef corals in Belize: potential implications for coral bleaching. *Limnology and Oceanography,* **51,** 1887-1897.

Watanabe, M., Roth, T. L., Bauer, S. J., Lane, A. & Romick-Rosendale, L. E. (2016) Feasibility Study of NMR based serum metabolomic profiling to animal health monitoring: A case study on Iron Storage Disease in captive sumatran rhinoceros (*Dicerorhinus sumatrensis*). *PloS one,* **11,** e0156318.

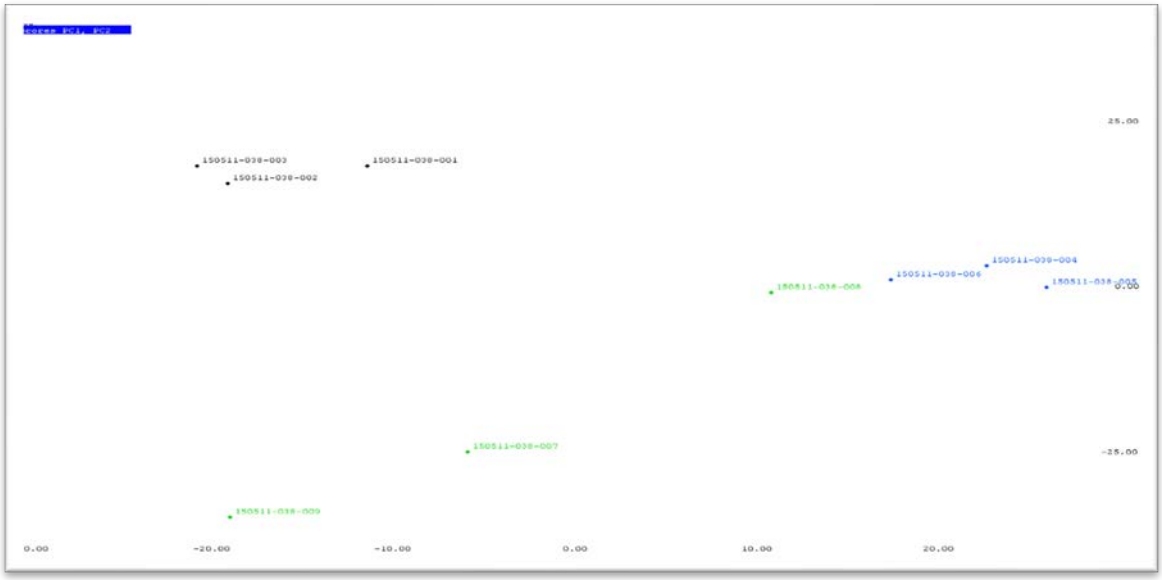
Wehrens, R., Hageman, J. A., van Eeuwijk, F., Kooke, R., Flood, P. J., Wijnker, E., Keurentjes, J. J., Lommen, A., van Eekelen, H. D. & Hall, R. D. (2016) Improved batch correction in untargeted MS-based metabolomics. *Metabolomics,* **12,** 88.

Weis, V. M. (1993) Effect of dissolved inorganic carbon concentration on the photosynthesis of the symbiotic sea anemone *Aiptasia pulchella Carlgre*n: role of carbonic anhydrase. *Journal of Experimental Marine Biology and Ecology,* **174,** 209-225.
Weis, V. M. (2019) Cell biology of coral symbiosis: Foundational study can inform solutions to the coral reef crisis. *Integrative and Comparative Biology*, **59**(4), 845–855.
Weis, V. M. & Allemand, D. (2009) What determines coral health? *Science,* **324,** 1153-1155.
Weis, V. M., Davy, S. K., Hoegh-Guldberg, O., Rodriguez-Lanetty, M. & Pringle, J. R. (2008) Cell biology in model systems as the key to understanding corals. *Trends Ecol. Evol.,* **23,** 369-376.
Weljie, A. M., Newton, J., Mercier, P., Carlson, E. & Slupsky, C. M. (2006) Targeted Profiling: Quantitative Analysis of 1H NMR Metabolomics Data. *Analytical chemistry,* **78,** 4430-4442.
Westmoreland, L. S., Niemuth, J. N., Gracz, H. S. & Stoskopf, M. K. (2017) Altered acrylic acid concentrations in hard and soft corals exposed to deteriorating water conditions. *Facets,* **2,** 531-544. https://doi.org/10.1139/facets-2016-0064
Willenberg, I., Ostermann, A. I. & Schebb, N. H. (2015) Targeted metabolomics of the arachidonic acid cascade: current state and challenges of LC–MS analysis of oxylipins. *Analytical and Bioanalytical Chemistry,* **407,** 2675-2683.
Winters, G., Loya, Y. & Beer, S. (2006) In situ measured seasonal variations in Fv/Fm of two common Red Sea corals. *Coral Reefs,* **25,** 593-598.
Withers, K. J. T., Grant, A. J. & Hinde, R. (1998) Effects of free amino acids on the isolated symbiotic algae of the coral Plesiastrea versipora (Lamarck): absence of a host release factor response. *Comparative Biochemistry and Physiology A-Molecular & Integrative Physiology,* **120,** 599-607.
Yakovleva, I. & Hidaka, M. (2004) Diel fluctuations of mycosporine-like amino acids in shallow-water scleractinian corals. *Mar. Biol.,* **145,** 863-873.
Yancey, P. H. (2005) Organic osmolytes as compatible, metabolic and counteracting cytoprotectants in high osmolarity and other stresses. *Journal of Experimental Biology,* **208,** 2819-2830.
Yancey, P. H., Heppenstall, M., Ly, S., Andrell, R. M., Gates, R. D., Carter, V. L. & Hagedorn, M. (2009) Betaines and dimethylsulfoniopropionate as major osmolytes in cnidaria with endosymbiotic dinoflagellates. *Physiological and Biochemical Zoology,* **83,** 167-173.
Yellowlees, D., Rees, T. A. V. & Leggat, W. (2008) Metabolic interactions between algal symbionts and invertebrate hosts. *Plant Cell Environ.,* **31,** 679-694.
Yost, C. C., Weyrich, A. S. & Zimmerman, G. A. (2010) The platelet activating factor (PAF) signaling cascade in systemic inflammatory responses. *Biochimie,* **92,** 692-697.
Zabek, A., Swierkot, J., Malak, A., Zawadzka, I., Deja, S., Bogunia-Kubik, K. & Mlynarz, P. (2016) Application of 1H NMR-based serum metabolomic studies for monitoring female patients with rheumatoid arthritis. *Journal of pharmaceutical and biomedical analysis,* **117,** 544-550.
Zahl, P. A. & McLaughlin, J. J. A. (1957) Isolation and Cultivation of Zooxanthellae. *Nature,* **180,** 199-200.

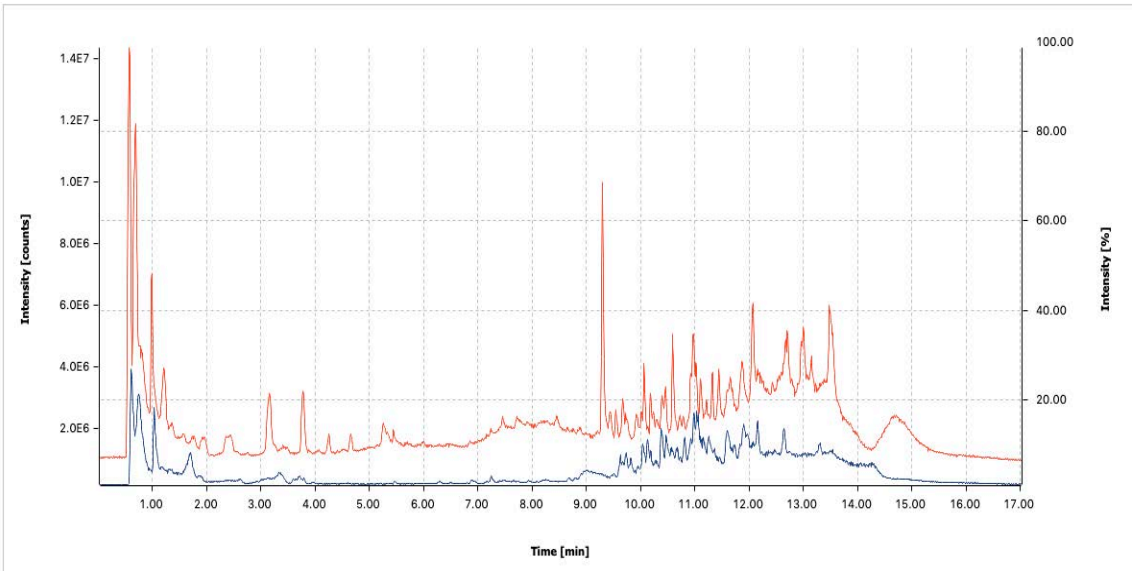
APPENDIX



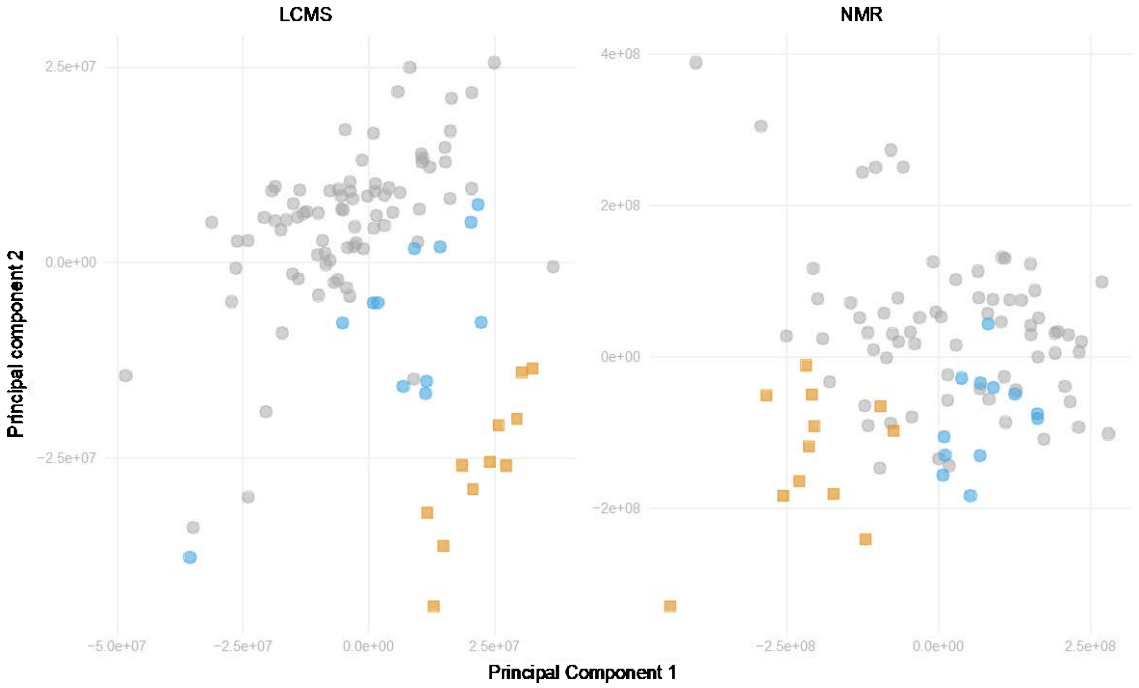
Appendix Figure 1 Comparison of an identical sample on 3 different reversed phase chromatography columns in positive ion detection mode. XB-C_{18} (top), PFP (middle) and Phenyl-Hexyl (bottom).



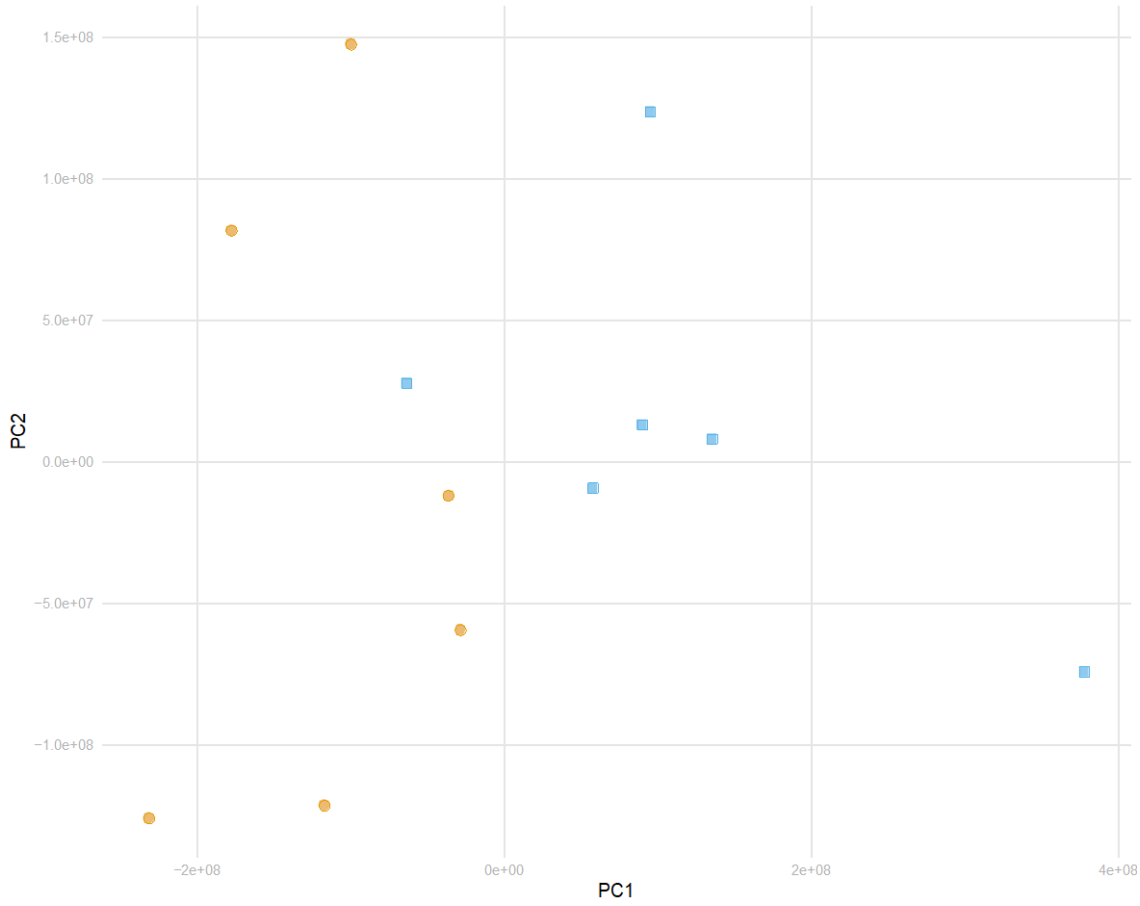
Appendix Figure 2 Principal Components Analysis of coral extracts subjected to three different solvent removal techniques. Lyophilisation (black), speed vacuum (blue) and nitrogen stream (green).



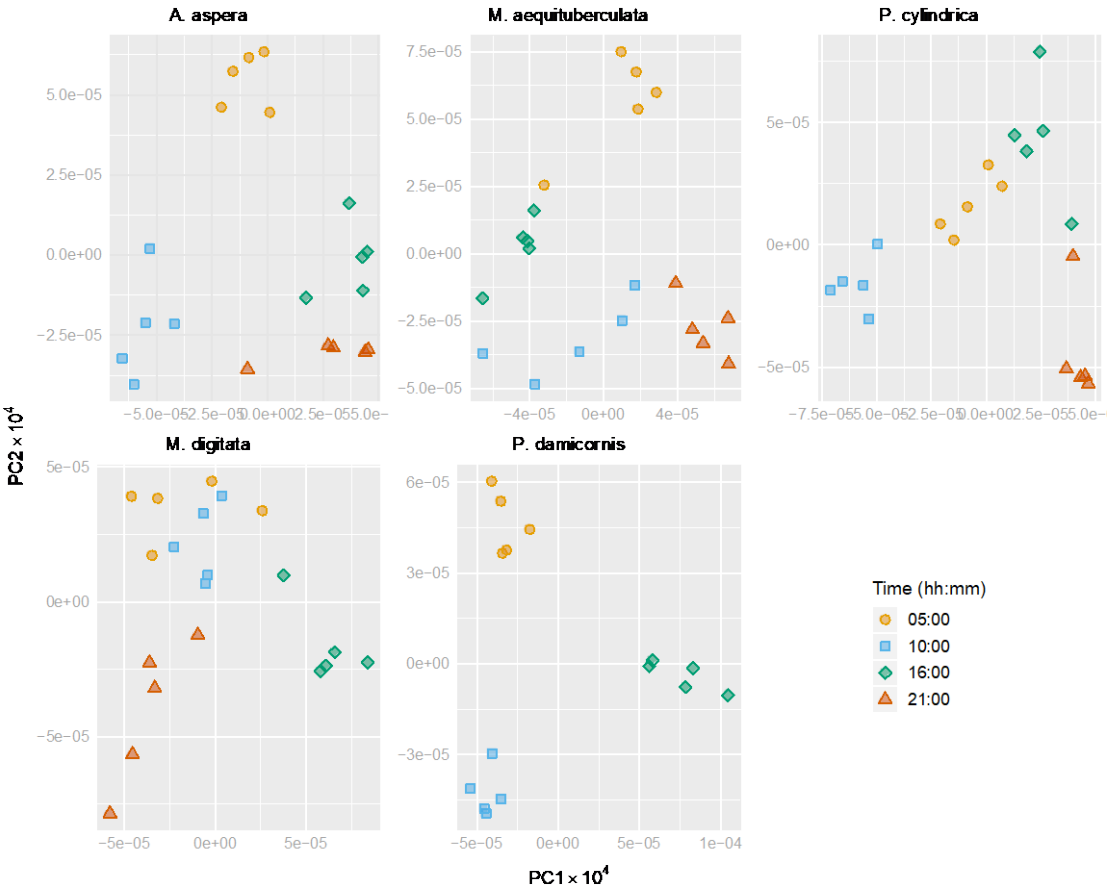
Appendix Figure 3 Total ion chromatograms of identical samples analysed in positive ionisation mode (red) and negative ionisation mode (blue)



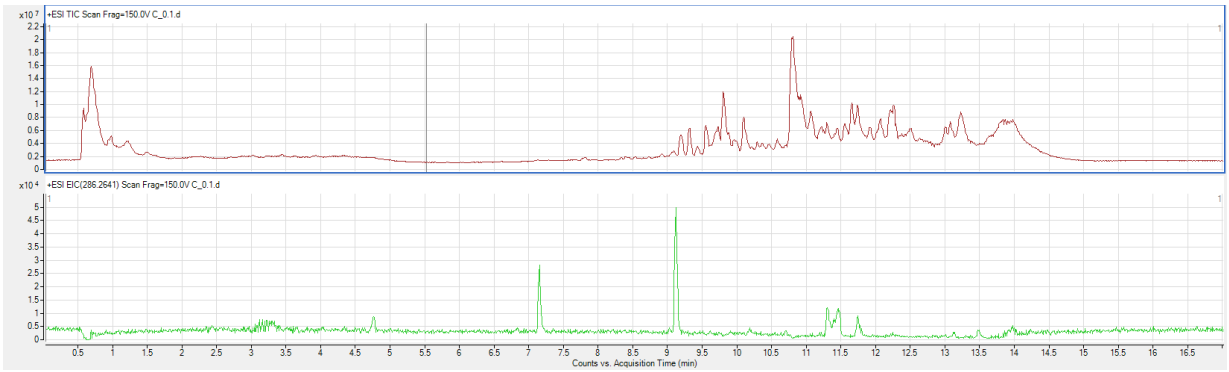
Appendix Figure 4 LC-MS and ^{1}H-NMR PCA analysis of all samples. Samples exposed to elevated temperature (eT and eCO_2eT) at day 6 (blue circles) and day 14 (orange squares) are highlighted.



Appendix Figure 5 PCA of ^{1}H-NMR data for elevated temperature (eT, blue squares) and control (orange circles) classes at day 14 of exposure



Appendix Figure 6 PCA of each species after log transformation of the data: log transformation reveals the structural variation of low concentration metabolites by reducing large values in the dataset more than smaller values while emphasising the larger relative standard deviation of low abundant metabolites.



Appendix Figure 7 An example total ion chromatogram (top) and extracted ion chromatogram (bottom; m/z 286.2641) of *A. aspera*

Appendix Table 1 Springer Nature license details for Chapter 3

This Agreement between Mr. Benjamin Gordon ("You") and Springer Nature ("Springer Nature") consists of your license details and the terms and conditions provided by Springer Nature and Copyright Clearance Center.

License Number	4883070218703
License date	Aug 06, 2020
Licensed Content Publisher	Springer Nature
Licensed Content Publication	Springer eBook
Licensed Content Title	Extraction Protocol for Nontargeted NMR and LC-MS Metabolomics-Based Analysis of Hard Coral and Their Algal Symbionts
Licensed Content Author	Benjamin R. Gordon, William Leggat, Cherie A. Motti
Licensed Content Date	Jan 1, 2013
Type of Use	Thesis/Dissertation

Appendix Table 2 The top 50 most influential metabolites to the PLS-DA model in Chapter 4.

mz	vip	control	eT	eCO2	eCO2eT
256.1166	100.00	45.12	93.69	44.88	100.00
274.1992	80.00	79.83	68.98	56.32	80.00
302.2308	75.44	75.44	55.46	72.80	73.87
300.2147	61.85	60.76	44.38	59.97	61.85
565.2811	59.78	12.80	41.98	12.67	59.78
494.5636	56.94	16.25	43.63	20.63	56.94
298.2735	55.33	18.42	44.48	17.43	55.33
294.2422	54.86	15.63	46.58	13.96	54.86
286.2994	54.25	11.55	40.10	20.01	54.25
514.3720	54.18	18.72	24.14	24.65	54.18
694.5451	53.13	19.52	32.90	19.47	53.13
567.5809	51.55	11.12	38.68	18.65	51.55
333.1782	50.89	18.86	29.08	31.40	50.89
680.5825	50.71	12.29	42.29	13.80	50.71
623.6350	50.08	11.73	43.55	16.07	50.08
509.5011	49.84	48.07	27.25	49.84	40.75
533.2730	49.59	26.44	38.70	19.05	49.59
646.3334	49.41	22.01	30.54	23.77	49.41
411.4272	47.66	10.09	35.23	15.69	47.66
621.6285	47.60	16.52	41.23	20.60	47.60
340.1156	47.54	47.54	17.76	26.47	18.14
247.0418	46.33	18.77	37.32	13.48	46.33
296.2571	46.25	11.35	38.62	12.90	46.25
99.9799	46.04	11.31	41.95	15.47	46.04
284.2934	45.56	4.27	33.84	11.56	45.56
686.4859	45.48	17.51	33.82	24.36	45.48
426.2220	44.95	9.82	26.50	10.25	44.95
560.3340	44.74	27.68	28.14	33.93	44.74

299.2934	44.53	17.77	28.20	22.31	44.53
292.0441	44.23	29.59	38.53	17.16	44.23
457.3533	44.04	19.83	35.85	18.86	44.04
687.5141	43.92	25.92	34.56	30.70	43.92
794.6061	43.75	18.55	27.69	19.28	43.75
246.0963	43.71	17.21	34.52	22.49	43.71
246.9871	43.67	43.67	33.55	30.29	38.04
280.6700	43.53	23.91	28.36	32.94	43.53
742.5740	43.30	43.30	32.38	42.45	37.21
139.0300	42.91	13.09	38.58	13.83	42.91
590.4276	42.69	14.64	37.27	11.76	42.69
487.2822	42.63	21.45	33.70	31.92	42.63
372.3842	42.21	20.10	32.03	17.25	42.21
228.0860	42.00	42.00	13.71	40.40	22.69
810.5997	41.91	8.89	27.43	9.06	41.91
290.0853	41.86	41.86	18.16	22.52	15.67
826.6058	41.79	21.99	23.98	19.65	41.79
692.5772	41.71	18.66	34.41	19.57	41.71
678.4905	41.56	9.86	35.56	10.24	41.56
574.3158	41.43	16.77	29.53	28.03	41.43
170.0794	41.23	24.30	32.16	26.66	41.23
173.1083	41.10	19.60	35.92	22.80	41.10

Appendix Table 3 The top 50 most influential metabolites for the RF model in Chapter 4.

mz	vip	control	eT	eCO2	eCO2eT
274.1992	100.00	100.00	46.55	86.94	33.28
256.1166	80.67	20.32	80.67	39.07	47.59
846.5911	75.53	59.70	60.08	75.53	67.69
279.1004	70.11	58.99	69.14	70.11	58.35
302.2308	64.76	63.90	21.50	64.76	15.56
304.1517	59.73	59.73	55.14	55.08	59.38
608.3232	56.93	56.93	48.55	34.75	53.55
471.1805	45.70	37.28	37.72	44.48	45.70
567.5809	44.78	20.46	44.78	21.14	42.42
336.1221	42.97	35.13	39.39	42.97	40.42
849.6088	42.27	26.30	35.98	42.27	40.02
623.6350	39.19	21.14	36.92	12.48	39.19
590.8178	37.99	33.21	37.99	28.78	8.32
1135.3766	34.79	33.03	34.35	34.79	28.32
217.1509	34.46	34.46	27.15	14.71	23.14
231.1150	33.60	33.60	18.02	24.45	15.65
316.0974	33.25	33.25	28.35	23.40	18.34

590.4276	32.80	14.48	32.80	14.48	13.59
167.0966	32.70	25.33	27.61	32.68	32.70
862.5857	32.67	26.68	32.67	24.59	26.82
340.1156	32.16	32.16	21.99	18.44	30.25
333.1553	32.15	14.48	29.38	32.15	30.93
247.0418	31.88	9.39	31.88	11.95	6.83
464.1640	31.73	31.73	19.11	20.05	23.56
661.1742	31.65	27.43	26.47	31.65	20.95
300.1996	31.40	14.48	23.52	13.93	31.40
677.6813	31.27	18.62	31.27	21.14	5.41
347.1031	31.15	31.15	9.80	16.97	8.80
161.0415	30.62	28.87	19.42	4.42	30.62
292.5880	30.52	28.71	25.57	30.52	0.00
268.1028	30.48	30.48	21.14	21.14	14.48
509.5011	30.43	30.43	21.14	29.01	8.25
742.5740	30.19	14.48	30.19	21.14	23.41
872.6058	30.00	15.81	28.68	30.00	23.37
165.0574	29.81	21.14	29.81	18.13	7.41
103.0518	28.59	20.66	28.59	23.72	25.57
251.0160	28.42	11.24	28.42	21.14	16.53
178.0924	27.77	27.77	16.98	21.14	14.48
192.0242	27.75	15.67	27.75	22.91	18.34
472.1842	27.63	19.34	27.63	25.82	17.80
551.3206	27.42	27.42	16.03	22.86	21.14
985.5799	27.42	14.48	21.95	21.14	27.42
766.5766	27.23	7.82	13.61	14.48	27.23
376.2978	27.09	23.00	27.09	14.05	6.33
211.1894	26.99	6.37	14.48	26.99	14.48
746.5665	26.99	26.99	23.91	25.75	8.60
417.2315	26.92	26.92	21.14	9.91	14.48
292.1733	26.65	26.65	23.89	25.93	14.48
469.3107	26.04	26.04	14.48	21.14	7.82
131.1283	26.02	23.06	21.14	26.02	23.72