THE UNIVERSITY OF QUEENSLAND

AUSTRALIA

Modelling Sponge-Symbiont Metabolism

Jabin Reid Watson
Bachelor of Science (Hons)

A thesis submitted for the degree of Doctor of Philosophy at
The University of Queensland in 2016
School of Biological Science

Abstract

Marine sponges are increasingly being recognised for their nutrient cycling ecosystem services, linking pelagic nutrients with the benthic ecosystem. Furthermore, sponges are the most prolific producers of bioactive secondary metabolites in the marine environment. Despite their ecological and commercial value, little is know about the metabolic processes that are responsible. The inability to produce sufficient sponge biomass, and thus specific bioactive compounds, has been repeatedly identified as a major bottleneck towards commercialising sponge holobiont derived drugs. Likewise, nutrient cycling by sponges has been a relatively recent advancement in marine ecology and the scale of this process is unknown. This thesis takes a systems approach to understanding sponge-symbiont metabolic processes by utilising the genomic resources available for the demosponge *Amphimedon queenslandica* and its bacterial symbiont *AqS1*. Specifically, I undertake genome-scale modelling and metabolic flux analyses to investigate the metabolic networks present in this sponge as well as its associated vertically-transmitted microbial symbiont. The goal of this thesis is to generate the data required for, and subsequently develop, a dual-species genome-scale metabolic model for *A. queenslandica* and *AqS1*. This will provide a framework to further our understanding of how sponges produce their biomass while living in an oligotrophic, tropical reef environment.

To develop a genome-scale metabolic model, the biochemical composition of the organism must first be determined. Knowledge of the relative quantities of macromolecules and their respective building blocks (e.g. protein and amino acids) is vital, as each requires different substrates, enzymes and cofactors for their synthesis. I adapted and developed methods to characterise the composition and abundance of DNA, RNA, protein, lipids and carbohydrates in a marine sponge. These methods are described in detail that allows them to be easily transferred to other, non-model, large marine organisms. This is followed by a detailed analysis of *A. queenslandica's* macromolecular, amino acid, fatty acid and sterol composition. The biochemical data from this chapter was used to generate a biomass equation that represent the average composition of adult *A. queenslandica* in the metabolic model.

To understand how the metabolic network may work towards producing more biomass, it must be considered in the context of the environmental conditions that naturally constrain growth. The dominant environmental constraint on growth on an oligotrophic reef system is nutrient availability. The abundance of key elements, such as carbon and nitrogen, were quantified throughout the course of a year. To investigate effects on the sponge of any changes in nutrient availability, I concurrently sampled sponges and analysed their biochemical composition to the macromolecular level. Chapter 4 presents these data and discusses a number of trends and correlations in the biochemical composition of the sponges with changing nutrient availability through four seasons of the year. For instance, there was a significant increase in particulate to dissolved organic carbon at the end of summer with a corresponding rise in carbohydrates in *A. queenslandica*. Of particular importance for the subsequent metabolic modelling, was the separation of carbon into the dissolved and particulate fractions.

Chapter 5 presents the metabolic models and their analysis. Initially, I manually constructed a genome-scale metabolic model for the sponge *A. queenslandica*. The model construction identified 10 amino acids, 4 vitamins and a plant-derived phytosterol for which *A. queenslandica* is auxotrophic. These represent nutrients that are essential for sponge growth and may in future form the basis of a defined cell culture medium. The microbial community within *A. queenslandica* is relatively simple and dominated by a species of sulfur oxidising bacteria, called *Aqs1*. I generated a genome-scale model for *Aqs1*, which is able to synthesise all 20 proteinogenic amino acids, in addition to having a diverse range of carbohydrate-specific transporters and enzymes.

To investigate the interactions between the host sponge and *Aqs1*, the two models were joined using a shared compartment. This was called the extracellular matrix, and represents the area of interaction within the sponge body where metabolite transfers can occur. I measured the pumping rate of *A. queenslandica* and calculated the average volume of water pumped per hour, standardised to gram dry weight. This was used to constrain the nutrient uptake rate of the model. To investigate how the metabolic network may respond to different nutrient conditions, low and high nutrient conditions were defined using the ratios of particulate and dissolved carbon from the seasonal environmental profiling.

This work represents the first genome-scale model for a sponge-symbiont system, and marine invertebrates in general. The genome-scale metabolic models resulting from this work are an important resource that will guide future work into the metabolic processes of both *A. queenslandica* and its symbiont, *Aqs1*.

Declaration by author

This thesis is composed of my original work, and contains no material previously published or written by another person except where due reference has been made in the text. I have clearly stated the contribution by others to jointly-authored works that I have included in my thesis.

I have clearly stated the contribution of others to my thesis as a whole, including statistical assistance, survey design, data analysis, significant technical procedures, professional editorial advice, and any other original research work used or reported in my thesis. The content of my thesis is the result of work I have carried out since the commencement of my research higher degree candidature and does not include a substantial part of work that has been submitted to qualify for the award of any other degree or diploma in any university or other tertiary institution. I have clearly stated which parts of my thesis, if any, have been submitted to qualify for another award.

I acknowledge that an electronic copy of my thesis must be lodged with the University Library and, subject to the policy and procedures of The University of Queensland, the thesis be made available for research and study in accordance with the Copyright Act 1968 unless a period of embargo has been approved by the Dean of the Graduate School.

I acknowledge that copyright of all material contained in my thesis resides with the copyright holder(s) of that material. Where appropriate I have obtained copyright permission from the copyright holder to reproduce material in this thesis.

Publications during candidature

Peer reviewed publications

Watson, J., Degnan, B., Degnan, S. & Krömer, J. O. (2014). Determining the biomass composition of a sponge holobiont for flux analysis. In *Metabolic Flux Analysis: Methods and Protocols*, Methods in Molecular Biology. Vol 1191. DOI:10.1007/978-1-4939-1170-7_7

Watson, J. R., Brennan, T. C. R., Degnan, B. M., Degnan, S. M. & Krömer, J. O. (2014). Analysis of the biomass composition of the Demosponge *Amphimedon queenslandica* on Heron Island Reef, Australia. *Marine Drugs*. 12, 3733-3753; DOI:10.3390/md12063733

Peer review publications not included in thesis

Gauthier, M-E. A., **Watson, J. R.** & Degnan, S. M. (2016) Draft genomes shed light on the dual bacterial symbiosis that dominates the microbiome of the coral reef sponge *Amphimedon queenslandica*. Frontiers in Marine Science, http://dx.doi.org/10.3389/fmars.2016.00196

Conference presentations

Watson, J. R., Brennan, T. C. R., Degnan, B. M., Degnan, S. M. & Krömer, J. O. Analysis of the biomass composition of the demosponge *Amphimedon queenslandica*. Oral presentation. *International Marine Biotechnology Conference*, November 2013, Brisbane, Australia.

Watson, J. R., Gauthier, M., Palfreyman, R., Fernandez-Velverde, S., Brennan, T. C. R., Nielsen, L. K., Degnan, B. M., Degnan, S. M. & Krömer, J. O. Genome-scale metabolic modeling promises new insights into biomass production in the marine sponge *Amphimedon queenslandica*. Poster presentation. *Metabolic Engineering - X*, June 2014, Vancouver, Canada

Publications included in this thesis

Watson, J., Degnan, B., Degnan, S. & Krömer, J. O. 2014. Determining the biomass composition of a sponge holobiont for flux analysis. In *Metabolic Flux Analysis: Methods and Protocols*, Methods in Molecular Biology. Vol 1191. DOI:10.1007/978-1-4939-1170-7_7.

This publication was included as Chapter 2

Contributor	Statement of contribution
Jabin R. Watson	Designed study (60%) Conducted analysis (70%) Wrote and edited the chapter (70%)
Bernie Degnan	Designed study (10%) Conducted analysis (10%) Wrote and edited the chapter (10%)
Sandie Degnan	Designed study (10%) Conducted analysis (10%) Wrote and edited the chapter (10%)
Jens Kromer	Designed study (20%) Conducted analysis (10%) Wrote and edited the chapter (10%)

Watson, J. R., Brennan, T. C. R., Degnan, B. M., Degnan, S. M. & Krömer, J. O. 2014. Analysis of the biomass composition of the Demosponge *Amphimedon queenslandica* on Heron Island Reef, Australia. *Marine Drugs*. 12, 3733-3753; DOI:10.3390/md12063733.

This publication was included as Chapter 3

Contributor	Statement of contribution
Jabin R. Watson	Designed study (60%) Conducted analysis (70%) Wrote and edited the chapter (60%)
Bernie Degnan	Designed study (10%) Conducted analysis (5%) Wrote and edited the chapter (20%)
Sandie Degnan	Designed study (10%) Conducted analysis (5%) Wrote and edited the chapter (10%)

Jens Kromer	Designed study (20%) Conducted analysis (20%) Wrote and edited the chapter (10%)
Tim Brennan	Performed the fatty acid methyl ester (FAME) analysis and provided the GC-MS data (100%)

Contributions by others to the thesis

Bernie Degnan, Sandie Degnan and Jens Kromer all contributed to the design off all experiemental designs and manuscript revisions.

Tim Brennan conducted the fatty acid methyl ester analysis and supplied the data in Chapter 3.

Robin Palfreyman provided the Java tool that was used to convert excel to SBML format in Chapter 5 and assisted in generating the *Amphimedon queenslandica* Pathway Tools database.

Statement of parts of the thesis submitted to qualify for the award of another degree

None.

Acknowledgements

First I would like to thank my family. I can't thank enough my parents, Helen and John, who have supported me with out question, through out this process and in everything that I have done before it. It goes without saying I wouldnt be where I am with out you both. To my sister Taigen, brother Nathan, Rosie and Oliver, a big thank you for everything you did for me (and Khan) through out my candidature.

Thank you Bernie and Sandie for taking me on as a student with absolutely no experience in any of the many fields this thesis covered. It was a risk and I hope it was worth it. More so than the molecular and analysis methods I have learnt, it is your accountability for maintaining the high standards that will remain with me. The greatest outcomes for me have not been the facts, rather the mentality towards producing top quality work, while enjoying doing it. I can say I have gained the confidence under your supervision to approach any question with a rationally.

A big should out to Jens. I had no idea about metabolic engineering when you took me on as a student and always made time for my endless questions. Its been a great 4 years and I consider you a friend in addition to a supervisor. Hopefully I made the cut as an "analyst" not just a "biologist".

To the lab members who were here when I joined, or came and left before me. Especially Jarrod, Alyssa, Jake, Nobuo, and Tim. Not the mention the current group, Laura, Will, Katia, Bec, Eunice and Claire. You have all contributed to my development as a scientist and the contents of this thesis in one way or another. A special should out goes to Carmel, Kerry and Simone, you guys know what you mean to me. Andrew, Fede and Shun, it goes without saying that we have had some unreal experiences together and I'm sure we will stay mates long after we have left the lab.

Thank you to Lars Nielson, Andreas Schmid, Robin Palfreyman and Esteban Marcellin for discussions and advice regarding the construction of the metabolic model.

Finally I would like thank boat club. Adro, Pros, Captain Dan, Ads, Ross, Pete and Jake, I just don't know where I would be without all those emails about boats. And Harps, you always come up with solid advice on alot of things. Thanks guys.

Keywords

porifera, microbiome, symbiosis, genome-scale model, biomass composition, constraint-based modelling, flux balance analysis

Australian and New Zealand Standard Research Classifications (ANZSRC)

ANZSRC code: 060114, Systems Biology, 35%
ANZSRC code: 060408, Genomics, 30%
ANZSRC code: 060104, Cell Metabolism, 35%

Fields of Research (FoR) Classification

FoR code: 0601, Biochemistry and Cell Biology, 70%
FoR code: 0604, Genetics, 30%

Table of Contents

FIGURES

TABLES

Appendices

List of Abbreviations

Abbreviation	Definition
HMA	Hight microbial abundance (sponge species)
LMA	Low microbial abundance (sponge species)
OTU	Operational taxonomic unit
EM	Electron microscopy
GEM	Genome-scale metabolic model
DNA	Deoxyribonucleic acid
RNA	Ribonucleic acid
g/gDW	grams per dry weight
FBA	Flux balance analysis
ATP	Adenosine triphosphate
ECM	Extracellular matrix
MiMB	Methods in Molecular Biology
SDS	Sodium dodecyl sulfate
Tris	hydroxylmethyl aminomethane
HCL	Hydrochloric acid
EDTA	Diaminoethane tetraacetic acid
NaOH	Sodium hydroxide
NaCl	Sodium chloride
NA3PO4	Sodium phosphate
HClO4	Perchloric acid
KOH	Potassium hydroxide
BSA	Bovine serum albumin
NASO4	Sodium sulfate
mL	Millilitre
µm	Micron
RO	Reverse osmosis
nm	Nanometer

Abbreviation	Definition
°C	Degrees celcius
µL	Microlitre
mg	Milligram
M	Molar
µg	Microgram
ex	Excitation wavelength
em	Emmission wavelength
FAME	Fatty acid methyl ester
ISTD	Internal standard
GC/MS	Gas-chromatography mass-spectrometry
MS	Mass spectrometry
NIST	National Institute of Standards and Technology
FA	Fatty acid
mol / gDW	mol per gram dry weight
h	Hour/s
µM	Micromolar
mm	Millimeter
ID	Individual
std. dev.	Standard deviation
U# or UU#	Unidentified fatty acid or sterol
mmol / gDW	millimoles per gram dry weight
MW	Molecular weight
DMEM F12	Dulbecco's Modified Eagle Medium
AIMS	Australian Institute for Marine Science
GBROOS	Great Barrier Reef Ocean Observing System
m	Meter
NATA	National Association of Testing Authorities
TOC	Total organic carbon

Abbreviation	Definition
DOC	Dissolved organic carbon
ANOVA	Analysis of variance
Tukey HSD	Tukey's honest significant difference test
mg / L	Milligrams per litre
mmol / h / gDW	Millimoles per hour per gram dry weight
g	Gram
µW	Microwatt
kW	Kilowatt

Chapter 1 - General Introduction

1.1 The marine sponge holobiont

Marine sponges inhabit all benthic marine environments, where they perform vital ecosystem services by providing physical structure and functioning as a trophic link between pelagic nutrients and benthic ecosystems (de Goeij et al., 2013; Hentschel et al., 2012; Rix et al., 2016). In addition to being important members of benthic ecosystems, marine sponges are a rich source of novel secondary metabolites, often with unique structural properties and commercial potential (Blunt et al., 2009, 2010, 2011; Genta-Jouve & Thomas, 2012; Laport et al., 2009; Leal et al., 2012; Munro et al., 1999; Thomas et al., 2010). A key aspect of sponge biology that underlies both the ecosystem services and commercial potential is the symbiotic relationship that sponges maintain with a diverse suite of microorganisms. Together, the host sponge and its associated microbial community are referred to as the sponge holobiont (Hentschel et al., 2012).

The importance of the microbiome to the sponge holobiont becomes evident with the classification of a sponge species as either a high-microbial abundance (HMA) or low-microbial abundance (LMA) species. This distinction is seen as a defining characteristic of sponges (Hentschel et al., 2003, 2006; Weisz et al., 2007b). HMA sponges are characterised as having dense tissue that contains an abundant and diverse microbial community. Additionally, HMA sponges exhibit slower rates of pumping when compared to LMA sponges (Hentschel et al., 2003, 2006; Weisz et al., 2007b). The longer residence time that seawater has within HMA sponges, compared to LMA sponges, has been linked to a greater reliance on the microbial community by the host sponge (Weisz et al., 2007a,b). In LMA sponges, the faster pumping rate gives the sponge access to more environmentally-sourced nutrients, and presumably the host sponge has a reduced reliance on the microbial community to synthesise organic molecules from inorganic sources and recycle waste products (Freeman et al., 2014).

What is known about the composition of the microbial community and potential metabolic interactions with the host sponge comes largely from genomic and transcriptomic approaches that have been applied to different species from a variety of marine environments (Fan et al., 2012; Hentschel et al., 2012; Kamke et al., 2013; Moitinho-Silva et al., 2013; Rua et al., 2015; Schleper et al., 1998; Wilson et al., 2014). Early work identified abundant sponge-specific operational taxonomic units (OTUs) (Hentschel et al., 2002), many of which are vertically transmitted from parent to embryo (Webster et al., 2009). More recently, deeper 16S rRNA gene pyrotag sequencing has revealed that many of the sponge-specific OTUs occur in the rare biosphere of the surrounding seawater and sediments, indicating horizontal transmission could also occur (Moitinho-Silva et al., 2013; Sipkema et al., 2015; Taylor et al., 2013; Webster et al., 2009). Interestingly, sponge-specific OTUs recovered from seawater samples appear to be metabolically inactive (Moitinho-Silva et al., 2013), suggesting that they rely on metabolic or regulatory aspects of the host sponge. This dependency may underlie the limited success of in vitro cultivation of sponge-specific microbes (Abdelmohsen et al., 2014).

The different sequencing techniques have repeatedly shown that sponge microbiomes contain members with specialised metabolic capacities (Hentschel et al., 2012; Taylor et al., 2007), such as the recycling of nitrogenous compounds within the holobiont (Fiore et al., 2015; Hoffmann et al., 2009). For instance, nitrification was shown to occur in four species of cold water sponges (Radax et al., 2011), while the Mediterranean sponge *Aplysinia aerophoba* holobiont exhibited seasonal changes in ammonium uptake and excretion (Bayer et al., 2008). The ability of symbionts to store phosphorous as polyphosphate has also been experimentally shown (Zhang et al., 2015). The genomes of bacteria within the sponge-specific candidate phylum, Poribacteria, encode the capacity for complex carbohydrate degradation (Kamke et al., 2013). Similarly, carbohydrate transporters have been found in the genomes of sulfur-oxidising symbionts in two geographically separate sponges, *Haliclona cymaeformis* collected in Hong Kong, China (Tian et al., 2014), and *Amphimedon queenslandica,* which is the focus of this study, from the Great Barrier Reef, Australia (Gauthier et al., 2016). The ability of members of the sponge microbiome to cycle carbon, nitrogen, phosphorous and sulfur within the host suggests that the central metabolic networks may be intertwined in both

the production of energy and biomass synthesis. These few examples outlined above (for detailed reviews see (Hentschel et al., 2012; Taylor et al., 2007)) highlight the genomic capability that exists for metabolic exchange and dependency, but the inherent complexity of the microbiome has made empirical validation of the exchanges challenging and infrequently reported (Webster & Blackall, 2008).

A notable example of direct metabolite exchange between sponge and microbiome occurs between several tropical sponge species on the Great Barrier Reef and their phototrophic cyanobacteria (Wilkinson, 1982). Feeding the holobiont ^{13}C labelled carbon dioxide revealed that glycogen and glycerol contained ^{13}C that was fixed by the symbiotic cyanobacteria and transferred to the host sponge (Wilkinson, 1983). In another example ^{13}C carbon dioxide was incorporated by zooxanthellar symbionts and exchanged with the host, a coral-boring sponge, *Cliona varians* (Weisz et al., 2010). In these two examples photosynthesis and the fixation of inorganic carbon presented a point of difference between the host and symbiont metabolism that allowed labelling experiments to trace the metabolite transfer. However, the broader application of the labelling-trace experimental approach to quantify sponge-microbe interactions is much more difficult in sponges lacking photosymbionts where the inherent complexity of the microbial communities often precludes assignment of particular physiological or metabolic process to a specific OTU. For this approach to be applied to a wider taxonomic range of sponges, detailed knowledge of the genomic capabilities of both the host and microbiome is essential, a task that is simpler in LMA sponges.

1.2 The metabolism of the coral reef sponge holobiont

Sponges are the second most abundant taxa on most coral reefs after reef-building hermatypic corals (Birkeland, 1997a). Anthropogenic and natural disturbances are causing shifts in the composition of benthic communities, exemplified by the transition of Caribbean reefs that are now dominated by sponge biomass (Loh & Pawlik, 2014; McMurray et al., 2015). Tropical coral reefs are oligotrophic environments that are low in available nutrients; however, they are not lacking in energy sources due to

the high rates of primary productivity (photosynthesis) in the photic zone by macro- and micro-algae, by zooxanthellae symbionts of scleractinian corals, and by cyanobacteria (Hatcher, 2003). Nitrogen is required for the high rate of carbon fixation to be translated into a high rate of biomass synthesis by these primary producers, and this high demand creates a shortage of nitrogenous compounds in the water column (Rix et al., 2016). Primary producers on coral reefs release up to 50% of the carbon they fix, the majority of which enters the dissolved fraction of the water column (de Goeij et al., 2013). The dissolved fraction is mainly, but not exclusively, utilised by prokaryotes; minor uptake also occurs in some unicellular eukaryotes and invertebrate larvae (Carlson & Hansell, 2015). It is in this context that marine sponges, with the help of their microbial communities, perform their crucial trophic role of retaining precious dissolved organic material within the coral reef ecosystem, and linking the pelagic and benthic ecosystems (de Goeij et al., 2013; Rix et al., 2016).

The crucial ecological role that sponges perform on coral reefs by acting as a trophic link between nutrients in the water column and the benthic community (de Goeij et al., 2013; Rix et al., 2016) is facilitated by the vast volumes of water that are actively pumped and from which both particulate and dissolved material is efficiently removed (Yahel et al., 2003b). To maintain a clean and efficient pumping and filtration system, the specialised cells involved, the choanocytes, are shed in large quantities (Alexander et al., 2014; de Goeij et al., 2013; 2009; Shore, 1971). A study of five coral reef species of sponge revealed a turnover rate of 16-19% of the choanocyte cell population every six hours (Alexander et al., 2014). This cellular detritus is expelled with the exhalent water current and becomes food for benthic organisms, ranging from bacteria to invertebrate detritivores. When this turnover of biomass is combined with the large energy demand required for pumping water (Riisgård et al., 1993), it suggests that the central metabolism of the sponge holobiont must be extremely active and efficient, given the oligotrophic status of coral reef ecosystems.

The metabolic rates of diverse sponge species have been investigated through the quantification of oxygen consumption rates and calculating carbon balances (Koopmans et al., 2009; Reiswig, 1981). In *Haliclona oculata,* only 34% of the available carbon is required to meet the metabolic demands

of this temperate species, 10% of which is incorporated into the sponge biomass and 90% of which is required to produce energy (Koopmans et al., 2009). This energy is used for growth-associated maintenance (GAM), such as macromolecule polymerisation, as well as for non-growth associated maintenance (NGAM), including pumping water (Stephanopoulos et al., 1998). The metabolism of *H. oculata* reacts quickly to increases in food availability, as demonstrated by the rapid removal of ^{13}C labelled diatoms. The labelled carbon could be detected in sponge fatty acids for three weeks, indicating that the sponge either stored the excess carbon or incorporated it into cell membranes (Koopmans et al., 2010). Within this sponge, metabolic and growth rates within an individual are not homogenous, with faster growth at the tips (Koopmans et al., 2010). In the HMA sponge *Theonella swinhoei,* twice as much carbon is acquired from the dissolved fraction than the particulate (Yahel et al., 2003b), suggesting a reliance on the microbiome for the uptake of dissolved organics. Likewise, the HMA sponge *Verongia fistularis* is only able to obtain 14% of the carbon required to meet the demands of respiration, with the balance presumably being supplied by the microbiome (Reiswig, 1981).

Quantification of the naturally-occurring carbon and nitrogen isotope ratios in sponges has shown that the carbon and nitrogen interactions within and between both HMA and LMA sponges to be highly variable (Freeman et al., 2014; Freeman & Thacker, 2011; Freeman et al., 2015). HMA sponges are able to inhabit a wider isotopic niche than LMA sponges, presumably due to increased metabolic capacity of their larger microbial community, particularly HMA sponges with photosynthetic symbionts (Freeman et al., 2014). The ratios of $\delta^{13}C$ and $\delta^{15}N$ can be used to generate bivariate plots that place an organism within an isotopic niche; different $\delta^{13}C$ and $\delta^{15}N$ ratios indicate different food sources (Freeman et al., 2014). When applied to a Caribbean sponge population involving 25 species, it was revealed that LMA sponges inhabited a smaller isotopic niche when compared to HMA sponges (Freeman et al., 2014). This further suggests that HMA may have developed a closer dependence on specific symbionts for their metabolic abilities. For instance, HMA sponges of Caribbean reefs, but not LMA sponges, are able to derive nitrogen from microbial sources (Weisz et al., 2007a). Coral reef cavity sponges source 48-73% of their carbon requirements from coral mucus in the dissolved fraction of the water column, as demonstrated by comparing the stable carbon isotope ratio of potential

carbon sources (van Duyl et al., 2011). Although these examples support that metabolic exchange is occurring between host and microbiome, the extent of this exchange is poorly understood, and is likely to be flexible and dependent on the environmental conditions.

1.3 Metabolic products of the sponge holobiont

In addition to being important members of marine benthic ecosystems, marine sponges are a rich source of novel secondary metabolites that often have unique structural properties and bioactivity (Blunt et al., 2009; 2010; 2011; 2014; Leal et al., 2012). The production of secondary metabolites by the sponge holobiont is reliant upon the central metabolic pathways to produce the required precursor building blocks, which is likely to include the exchange of intermediate metabolites between host and symbiont (Haygood et al., 1999; Mueller et al., 2004; Radjasa et al., 2011; Schirmer et al., 2005; G. Wang, 2006; Wilson et al., 2014). This biotechnological potential of marine sponges has stimulated an extensive, and continuing, search for novel sponge-derived compounds and microorganisms. It has yielded success stories such as the compound eribulin mesylate, a synthetic analogue of halichondrin B that is currently approved for clinical application to treat metastatic breast cancer (Huyck et al., 2011). Also of note is the sponge-associated candidate bacterial genus *'Entotheonella'*, which rivals actinomycetes with its large genome size (larger than 9 Mb) and extensive repertoire of secondary metabolite pathways (Wilson et al., 2014).

Still, despite these examples and the considerable effort expended, the commercialisation rate of sponge-derived compounds through to clinical application is extremely low. This is largely due to the difficulty of supplying the drug development phases, let alone the eventual clinical demands, with enough sponge biomass for compound supply. Research into multiple avenues of sponge biomass supply has attempted to solve this problem, all with inconsistent results. In particular, variability in rates of biomass synthesis and in secondary metabolite concentrations are widely reported (for reviews see (Belarbi, 2003; Duckworth, 2009; Mueller et al., 2004; Osinga et al., 1999; Schippers et al., 2012; Sipkema et al., 2005)). Despite this variability, field aquaculture whole sponges and the

culture of explants (sponge pieces) in controlled conditions have revealed the potential for very fast growth rates that have not been reported under natural conditions. Trials conducted in New Zealand, culturing *Lissodendoryx* sp. explants in rope cages suspended in the ocean, reported a biomass increase of 5000% during a single month (Munro et al., 1999; Schippers et al., 2012). This rate, however, was not observed across the multiple study sites or seasons sampled, demonstrating a clear effect of the environment on the central metabolism and its production of biomass.

1.4 A systems approach towards understanding sponge metabolism

In this PhD project, I take a systems approach to increase our understanding of sponge holobiont metabolism by applying genome-scale metabolic modelling to the LMA demosponge *Amphimedon queenslandica*, and one of its bacterial symbionts, *AqS1* (Gauthier et al., 2016). The microbiome of *A. queenslandica* is dominated by *AqS1*, a sulfur-oxidising autotroph (Gauthier et al., 2016). The predictive power of genome-scale metabolic models (GEMs) has made them an indispensable tool utilised in commercial metabolic engineering (Averesch & Krömer, 2014; Oberhardt et al., 2009; Sengupta et al., 2013; Wang et al., 2014). GEMs are able to simulate how metabolic networks may function to achieve a defined objective, such as biomass production, under different nutrient conditions. The application of GEMs has rapidly expanded beyond standard microbial and yeast models to include modelling the metabolic interactions between multiple organisms, including the community dynamics and electron transfer between syntrophic bacteria (Nagarajan et al., 2013; Stolyar et al., 2007). Examples of prokaryote-eukaryote metabolic models include the host pea aphid (*Acyrthosiphon pisum*) and its intracellular symbiont *Buchnera aphidicola* (Macdonald et al., 2012), the response of human alveolar macrophage cells to the pathogenic *Mycobacterium tuberculosis* (Bordbar et al., 2010), and notably the metabolic exchange between host and gut microbiomes of both mouse (Heinken et al., 2013) and human systems (Shoaie et al., 2013).

How metabolic models are constructed and analysed is explained in detail in Chapter 5 but briefly, a GEM is a mathematical representation of the metabolic network of an organism based on the

enzymatic, transport and other biochemical capabilities encoded within the genome of that organism (O'Brien et al., 2015). The metabolic network is represented by the metabolite stoichiometry of the biochemical reactions identified in the genome (for detailed protocols see Feist & Palsson, 2010; Santos et al., 2011; Schellenberger et al., 2011; Thiele & Palsson, 2010). Constraints are placed on the model which define conditions, such as nutrient availability, of each simulation (Lewis et al., 2012). The objective of the simulation is also defined and the possible solutions can then be calculated and further analysed (Feist & Palsson, 2010; Orth et al., 2010). The model predictions can then be tested experimentally, with the results used to refine and improve the model (e.g. Heinken et al., 2014; Orth et al., 2011; O'Brien et al., 2015).

The genomic resources available for the *A. queenslandica* holobiont place it as a model sponge species for the development of a metabolic model. The draft genome for *A. queenslandica* was the first publicly available representative of the phylum *Porifera* (Srivastava et al., 2010). Deep transcriptomic sequencing of 82 developmental stages is available (Anavy et al., 2014) and was recently used to enhance the original gene identification and annotation (Fernandez-Valverde et al., 2015). *A. queenslandica* was identified as a LMA sponge by 16S amplicon sequencing and EM imaging (Feith et al., 2016). The LMA status of *A. queenslandica* is an advantage for metabolic modelling, as it presents a system where the identification of who is able to do what is simplified when compared to the complex microbiome of a HMA sponge. The microbial community associated with the adult *A. queenslandica* is dominated by *AqS1*, a sulfur-oxidising gammaproteobacterium for which a draft genome has recently been assembled and annotated (Gauthier et al., 2016). This is currently the only sponge species for which both host and symbiont draft genomes are available. The process of developing an *A. queenslandica-AqS1* multi-species GEM will increase our understanding of sponge-symbiont metabolism. Ecological questions, including the effects of large and small scale changes in environmental conditions can be simulated, and how the metabolic networks responds can be analysed at the metabolite and reaction level. The results of these simulations can be used to inform the design of validation experiments. Simulating, and thus better understanding, how sponge holobiont biomass can be synthesised will have significant application for advancing the accessibility of sponge holobiont derived commercial metabolites.

1.5 Thesis aims and chapter synopsis

1.5.1 Aim 1 – To develop a guideline for the generation of a biomass equation for a non-model organism.

An essential component of a genome-scale metabolic model is the knowing the biochemical composition of the study organism. For model organisms this information can be readily sourced from the scientific literature. The application of metabolic modelling to non-model organisms and the scarcity of available biochemical data necessitates the generation of organism specific data. Methods for the quantification of the different macromolecule types can be found throughout the literature but a convenient, single reference was not available. Chapter 2 of this thesis is a methodological chapter that presents detailed methods for the quantification of protein, carbohydrates, lipids, DNA and RNA in *Amphimedon queenslandica*, which can be applied to most, if not all sponges and potentially other marine invertebrates.

The experimental validation of predicted metabolic phenotypes is generally done by comparing simulated and experimental rates of biomass synthesis (growth) or product secretion (Wittmann, 2007; Zamboni et al., 2009). Quantifying growth rates in a sponge is not as straightforward as the cultivable microorganisms for which metabolic modelling was originally developed (Zamboni, 2011). Chapter 2 validates using displacement volume as a method to relate a piece of wet, live sponge to the quantified biochemical composition, expressed as grams per gram of dry weight (g/gDW). This will be critical for future validation experiments on live sponge explants, where the change in biomass amount during an experiment must be quantified.

Finally, components of a biomass equation are expressed in molar values. This is because the flow of metabolites through a network must be expressed in a type of unit that is tractable and different metabolites have different molecular weights. The calculation of a biomass equation involves

distributing the relative amount of each macromolecule type (g/gdw) across the relative abundance of its respective building block molecules, using the average molecular weight of the building block molecules. Chapter 2 contains a detailed example of how this conversion is performed.

1.5.2 Aim 2 – Characterise the biochemical composition of *A. queenslandica.*

Chapter 3 applied the methods developed in Chapter 2 to quantify the biochemical composition of *A. queenslandica* in one of its native habitats, Shark Bay on Heron Island Reef (-23.44, 151.92). The biomass composition analysis includes the skeleton, lipid, protein, carbohydrate, RNA and DNA. The composition of amino acids, the building blocks of proteins, was quantified for both the total holobiont biomass and for the skeleton. The abundant fatty acids and sterols in the lipid fraction were identified and their abundance quantified. The variability in macromolecular composition between and within individual sponges from the Shark Bay population is also investigated and discussed.

1.5.3 Aim 3 – To characterisation of the seasonal nutrient availability in the natural habitat of Shark Bay, Heron Island Reef.

To investigate how the metabolic networks of *A. queenslandica* and *AqS1* may function in the context of their natural Shark Bay environment, sampling of nutrient availability was conducted at seasonal time points over 13 months. Concurrent sampling of the skeleton and macromolecular composition of the holobiont was undertaken and potential correlations investigated. For most of the year low nutrient conditions prevail, with early autumn showing a peak in particulate, nutrient rich, material. Chapter 4 presents these results within the context of ecological processes known to occur on Heron Island Reef.

1.5.4 Aim 4 – To construct a metabolic model for *A. queenslandica* and its symbiotic microbe *AqS1* and simulate the flow of metabolites under low and high nutrient conditions.

The overall aim of this thesis is to produce a genome-scale metabolic model for *A. queenslandica* and *AqS1*, and relate this model to biomass synthesis. Chapter 5 presents a detailed description of the central metabolic pathways of a sponge for the first time. This is followed by a description of the reconstructed metabolic network of *AqS1*. These two networks were joined by the addition of a shared compartment, the extracellular matrix within a sponge, where the hypothetical exchange of metabolites between organisms was simulated.

Flux balance analysis (FBA) is a widely used method for analysing the flow of metabolites through a metabolic network (Orth et al., 2010). FBA was applied to validate the model by comparing the maximum ATP synthesis rates of the model, against the ATP demand for pumping and growth. Following this, the maximum biomass synthesis rates were simulated under low and high nutrient conditions and the predicted flow of metabolites through the network inspected.

CHAPTER 2 - DETERMINING THE BIOMASS COMPOSITION OF A SPONGE HOLOBIONT FOR FLUX ANALYSIS

2.1 Abstract

The first step on the path of flux analysis of a new organism with little available literature is the determination of the biomass composition. Once the content of the macromolecular components (protein, RNA, DNA, carbohydrates, lipids) and their composition is known, this composition can be converted into a biomass equation. The biomass equation is an important part of metabolic flux analysis. This equation provides the information about the precursor and energy needs for growth. In many experiments the determination of the growth rate is the simplest flux to be determined, yet this rate determines the net fluxes of a whole range of anabolic pathways in the system and often is used as the objective function in FBA analysis. The challenge for the scientist is to create a biomass equation that represents the organisms of choice under the conditions studied. This chapter outlines basic protocols that can be applied to the quantification of the macromolecular components, using the marine demosponge sponge *Amphimedon queenslandica* as a case study. As is true for all other sponges and indeed marine animals, *A. queenslandica* is a holobiont, comprised of an animal host plus symbiotic and other associated microbial cells. I show how this complexity can be overcome by developing a fast, yet robust, method for biomass quantification of sponges using the displacement volume. The analytical protocols I describe herein are widely applicable not only to other organisms sampled from complex environments, but also to cell cultures. The second part of the chapter highlights the procedures needed to convert a macromolecular composition into a biomass equation.

2.2 Introduction

Genome-scale metabolic reconstructions are powerful tools for understanding metabolic processes, and have become the standard for studying metabolic networks at a system level (Palsson, 2009). To convert a reconstruction into a functioning model for metabolic flux analysis, the biochemical composition of the organism must be determined (Feist et al., 2008). Quantifying the macromolecules of an organism allows the formulation of a biomass equation, which in turn enables the testing of a metabolic model by predicting growth rates under specific conditions. These predictions can be then tested in vivo.

The major macromolecules in a cell are DNA, RNA, proteins, carbohydrates and lipids. By quantifying the macromolecule composition, the precursor molecules that are required for biomass synthesis can be identified. This also allows calculation of the energy requirements for cell maintenance and growth. Although flux analysis has been extensively used on single species cultures (Blank et al., 2005; Holms, 1996; Quek et al., 2010), its application to more complex systems including two and more species is just beginning (Zomorrodi & Maranas, 2012). This chapter provides basic protocols that can be used to determine the macrocomposition of biomass in a complex biological system, exemplified here by a sponge holobiont. The protocols that I describe are applicable to any such characterization, from organisms in complex open environmental systems, to cell cultures.

The organism I chose to analyze as our case study was the marine sponge *Amphimedon queenslandica* (Niphatidae, Haplosclerida, Demospongiae, Porifera), a sessile invertebrate that encrusts complex substrata in tropical coral reef lagoons. At the time of writing, this is the only sponge to have a fully sequenced, assembled and annotated genome (Srivastava et al., 2010). As is the case for nearly all sponges, the *A. queenslandica* body plan consists of an irregular canal system through which water is pumped by groups of ciliated cells (Figure 2.1). The form of an individual sponge is shaped by various environmental factors, including prevailing currents and availability and shape of the substrate (Hooper & Van Soest, 2006; Palumbi, 1986). At the cellular level, the sponge

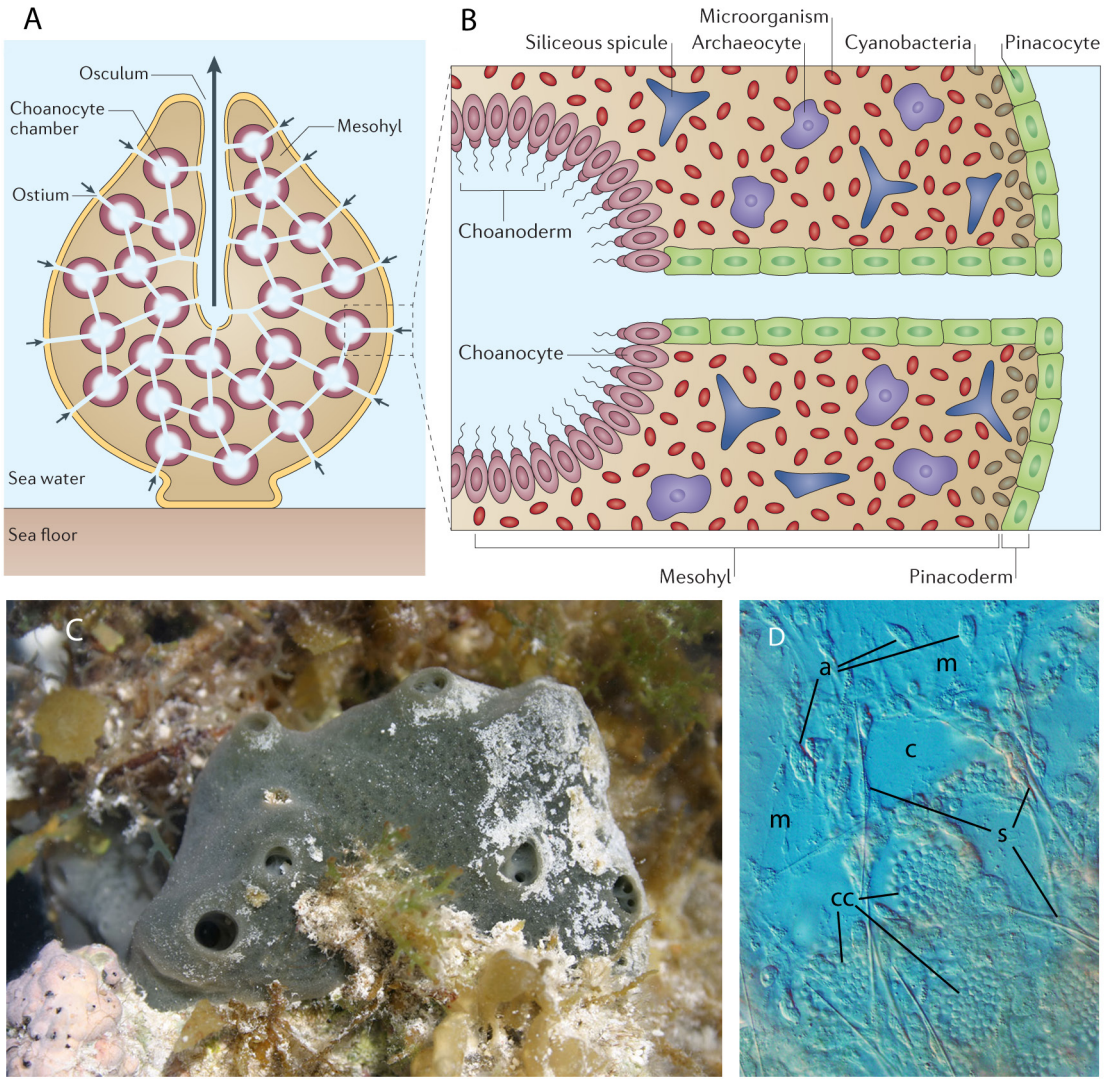


Figure 2.1 Body plan and cellular arrangement of *Amphimedon queenslandica*.

(A) Generalized sponge body plan showing directional water flow mediated by chambers of flagellated choanocyte cells. Connectivity between the inhalant ostia and exhalent oscula is irregular. (B) Cellular arrangement showing interacting microbial and sponge cell populations in the mesohyl. Spicules form part of the supporting skeleton. (C) *A. queenslandica* has an irregular body plan and incorporates foreign material as it grows. (D) This light microscopy image of a live *A. queenslandica* juvenile demonstrates the irregular biomass arrangement at the cellular level. The densely packed cells at the bottom of the image are dominated by choanocyte chambers (cc) and contrast to the low cell density of the extra cellular matrix (m). Archeocytes (a) and sclerocytes (s) are distributed throughout the sponge. Panel A and B reproduced from Hentschel et al (2012), with permission from Nature Publishing Group.

synthesizes a skeleton of silica spicules and a flexible extracellular matrix (ECM) that includes a diversity of proteins and carbohydrates and a sponge specific collagen, spongin (Schutze et al., 2001). The cellular biomass of the *A. queenslandica* holobiont primarily includes both eukaryotic sponge cells and prokaryotic bacterial cells that appear to maintain a symbiotic relationship. Both cell types are easily dislodged from the ECM with minimal physical force.

Long-term cell cultures for sponges have not been established (but see Mueller et al., 2004 and de Caralt et al., 2007 as examples of sponge cell culture methods). This is also the case for other marine invertebrates. Thus, the biomass equation must be generated from, and the subsequent flux analysis is performed on, whole or partial samples of living sponge. Living material that has been directly collected from the field will be inherently variable with respect to cell and ECM composition and density. In organisms that host microbial symbionts, this variability includes the ratio of sponge to non-sponge cells. To address this natural variation, analyses need to be performed on multiple samples from multiple individuals. In this case study, I set a minimum sample size of 20 pieces of sponge per assay; 5 subsamples were collected from 4 different sponges for each of the biochemical assays. To undertake flux analyses on living *A. queenslandica*, I need to be able to relate measurements made on dried samples (i.e. dry weight values) to the live sponge. Given the complexity of the internal aquiferous canal system in sponges, I chose to relate dry weight to displacement volume. This would allow me to measure the starting biomass with minimal disturbance to the metabolic state. Further, to ensure that values obtained from small subsamples are applicable to larger fragments or whole sponges, I determined whether the relationship between dry weight and displacement volume is linear.

This chapter was published in the Methods in Molecular Biology (MiMB) series of textbooks and reflects the methodological focus of the series. The aim is to describe the methods in detail that is not usually included in a journal publication. The required materials (section 2.3) and subsequent methods (section 2.4) are detailed separately. The notes section (2.5) contains hints and tips aimed to assist researchers unfamiliar with the described methods.

2.3 Materials

2.3.1 Materials common to multiple assays

- Pipettes and tips
- 0.22 μm filters
- 2.0 mL tubes
- Glass test tubes
- Glass pipettes
- Analytical balance
- Spectrophotometer
- 1.0 mL cuvettes, disposable and quartz
- Water bath
- Lyophilizer
- Shaker table

2.3.2 Biomass Sample Displacement volume

- *Amphimedon queenslandica*
- Scalpel blades
- Deep petri dishes
- 50 mL measuring cylinder

2.3.3 Skeleton weight

- 2% SDS (sodium dodecyl sulfate). Dissolve 1.0 g of SDS in 50 ml of water

2.3.4 DNA

- Fluorescent plate reader

- 96 well plates
- Calf thymus DNA (Sigma D1501).
- Hoechst 33258 concentrated stock solution, 10 mg/mL (Invitrogen, Life Technologies catalog number H3569).
- 1 M Tris (hydroxymethyl aminomethane) (pH 8.0). The molecular weight of Tris is 121.4 g/mol. Make 500 mL of 1 M Tris (pH 8.0) by adding 60.57 g to 400 mL of MilliQ water. Adjust the pH to 8.0 with HCl (hydrochloric acid). Make to 500 mL with MilliQ water.
- 0.5 M EDTA (diaminoethane tetraacetic acid). The molecular weight of EDTA is 372.2 g/mol. Make 100 mL of 0.5 M EDTA by adding 18.61 g of EDTA to 90 mL of MilliQ water. Adjust the pH to 8 using NaOH (sodium hydroxide) pellets (see note 1). Make to 100 mL with MilliQ water.
- TE (pH 8.0). To make 100.0 mL, combine 1.0 mL of 1 M Tris (pH 8.0) and 0.2 mL of 0.5 M EDTA and make up to 100.0 mL with MilliQ water.
- Fluorometry buffer. 2 M NaCl (sodium chloride) and 50 mM Na_3PO_4 (sodium phosphate). The molecular weight of NaCl is 58.44. To make 500 mL of fluorometry buffer, add 58.44 g to 450 mL of MilliQ water and stir to dissolve. The molecular weight of Na_3PO_4 is 163.94 g/mol. Add 4.099 g and stir to dissolve. Make to 500 mL with MilliQ water. Adjust the pH to 7.4.
- Rehydration buffer. Dissolve 0.5 g of lysozyme in 50 mL TE (pH 8.0). (see note 2)

2.3.5 RNA

- 1. $HClO_4$ (perchloric acid) solutions made from 70% stock (see note 3).
- 2. 3.0 M $HClO_4$. To make 25.0 mL, slowly add 6.468 mL of stock $HClO_4$ to 6.5 mL of MilliQ water. Adjust the final volume to 25 mL with MilliQ water.
- 3. 0.5 M $HClO_4$. Slowly add 8.625 mL of stock $HClO_4$ to 50 mL of MilliQ water. Adjust the final volume to 200 mL with MilliQ water.
- 0.3 M KOH (potassium hydroxide). Add 1.683 g of KOH to 100 mL of water and mix until dissolved.

2.3.6 Carbohydrate

- Concentrated sulfuric acid.
- Phenol (80% w/v).
- Iced water.
- Water bath set to 25°C.
- Analytical grade glucose.
- Glucose standard stock solution. Add 0.01 g of glucose to a 15 mL tube and record the weight of glucose using an analytical balance. Add 10 mL of MilliQ water, recording the weight of water added. Glucose has a molecular weight of 180.16 g/mol. Calculate the concentration of glucose in the stock solution.

2.3.7 Protein

- 1 M NaOH (sodium hydroxide). The molecular weight of NaOH is 39.996 g/mol. To make 500 mL add 19.9985 g to 450 mL of MilliQ water. Mix to dissolve the pellets. Make the total volume to 500 mL with MilliQ water.
- Bovine serum albumin (BSA) 2 mg/mL (Thermo Scientific product number 23209).
- Coomassie Plus™ (Bradford) Assay Kit (Thermo Scientific product number 23236).
- Water bath set to 95°C
- Iced water

2.3.8 Lipid

- 15 mL glass tubes with lids
- Hexane
- Isopropanol
- Make a 3:2 hexane:isopropanol solution. Each sample will require 5 mL.
- Make a 7:2 hexane:isopropanol solution. Each sample will require 5 mL.

- 0.47 M NaSO4 (sodium sulfate). To make 100 mL of 0.47 M $NaSO_4$, add 6.676 g to 100 mL of MilliQ water.
- Nitrogen gas supply, with manifold and flow control valves.

2.4 Methods

2.4.1 Displacement volume

Displacement volume was used to normalize live sponge biomass to dry weight composition. This measurement can be taken on live samples before flux analysis.

1. Cut sponge into similar size pieces using a scalpel blade (see note 4).
2. Filter seawater through a 0.22 µm filter
3. Place 5 mL of filtered seawater into the measuring cylinder, ensuring that the meniscus is in line with the 5 mL mark.
4. Place the piece of sponge into the measuring cylinder (see note 5).
5. Place a petri dish on the analytical balance and zero it.
6. With a pipette, transfer the displaced volume to the petri dish on the analytical balance, by carefully returning the meniscus of the measuring cylinder to the 5.0 mL line.
7. Record the displacement weight.
8. Convert to a volume, correcting for the specific density of the seawater.

2.4.2 Dry weight

Dry weights were obtained by freeze-drying (lyophilization). The application of a vacuum to frozen samples allows the water molecules to sublimate from the solid to the gas phase. This is a gentle process that leaves molecules in their conformational states. Samples that are to be analysed for RNA or lipid content should be lyophilized in a 15 mL glass tube.

1. Sponge pieces were cut and the displacement volume was recorded.
2. Label and weigh a tube and lid for each sample (see note 6).
3. Transfer each sponge piece into a tube and weigh.
4. Lyophilize each sample overnight.
5. Measure and record the weight of the lyophilized sample and tube.
6. Store the lyophilized samples in a -80°C freezer.
7. To calculate the dry weight, subtract the weight of the empty tube and lid from the weight of the tube, lid and dry sample.
8. Correct the dry weight for salt content. First calculate the amount of water that was evaporated by subtracting the weight of the tube, lid and dry sponge piece from the weight of the tube, lid and wet sponge piece. Multiply the amount of water evaporated by the specific gravity of the seawater to estimate the weight of salt left behind (see note 7). Subtract the weight of the salt from the weight of the dry sponge piece.
9. Normalize the salt-corrected dry weight to the displacement of the starting biomass.

2.4.3 Dry weight and displacement volume relationship

To analyze the relationship between dry weight and displacement volume, 50 different-sized pieces of sponge were collected. The displacement volume ranged from 0.0865 mL to 3.404 mL. The above methods were used to analyze ten pieces from five different sponges. The linearity of the relationship was plotted and resulted in an r^2 value of 0.91 (Figure 2.2).

2.4.4 Skeleton weight

I defined the sponge skeleton as the silica spicules and the ECM combined. This skeleton plays a structural role but is here assumed not to be turned over metabolically. Instead it can be treated as a metabolic sink during the modeling. Treating the sponge pieces with a detergent to lyse the cellular

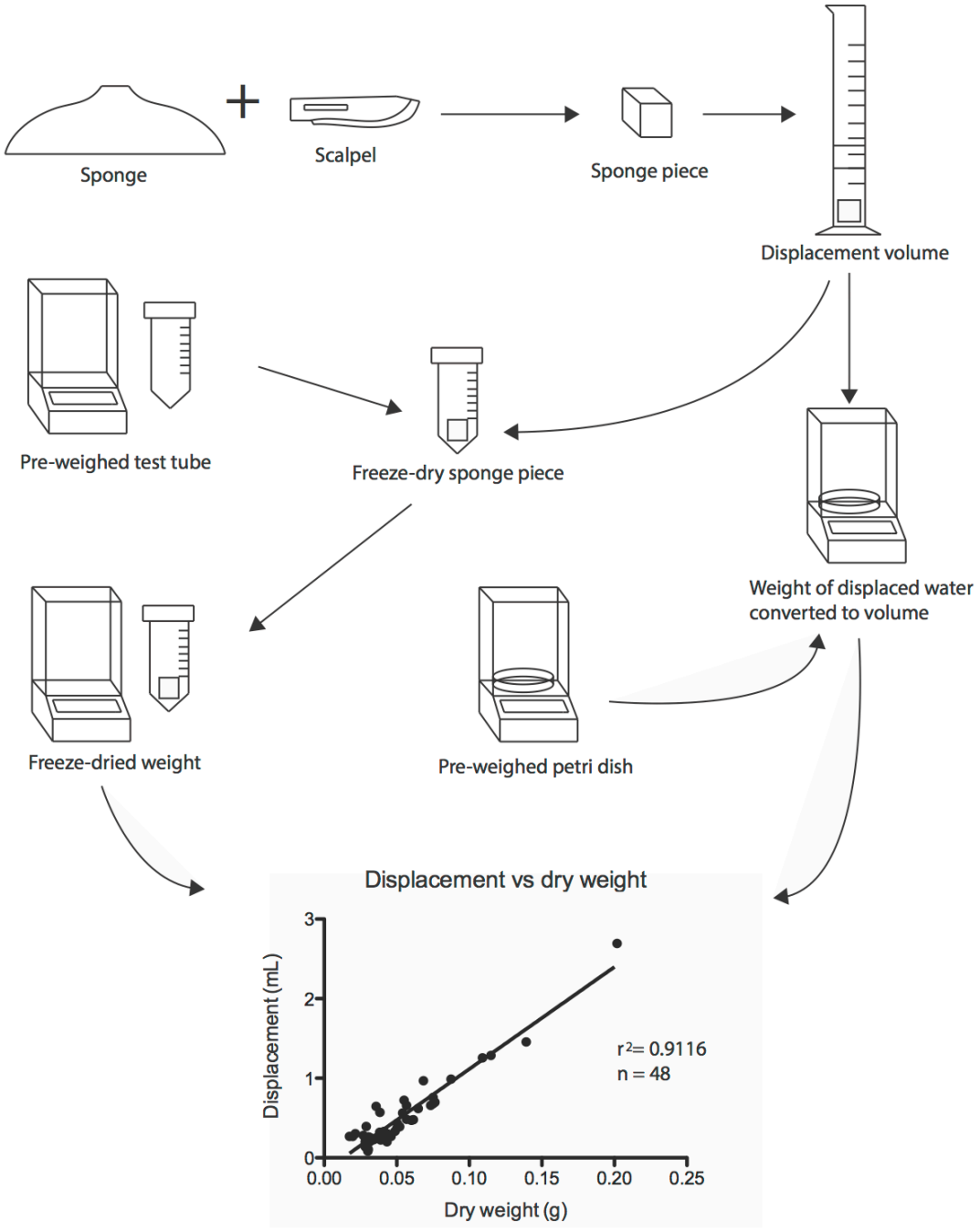


Figure 2.2 This graphical representation outlines the workflow used to generate the relationship between the displacement volume and the dry weight of pieces of sponge.

component allowed me to weigh the remaining skeleton and calculate the amount of biomass that was present.

1. Record the displacement and dry weight of each sample using the method described above.
2. Remove the freeze-dried samples from the -80°C freezer and allow to return to room temperature (see note 8).
3. Weigh and record the weight of each tube with the sample in it (see note 9).
4. Add 3.0 mL of 2% SDS to each tube.
5. Place on a shaker table for 2 hours.
6. Thoroughly wash each piece and tube with RO water to remove all SDS and cellular debris.
7. Examine each piece with a microscope to ensure that only the skeleton remains.
8. Return each piece to its tube and lyophilize overnight.
9. Weigh each piece of lyophilized skeleton.

2.4.5 DNA quantification

To quantify the DNA content, I used the fluorescent dye Hoechst 33258. Hoechst 33258 is a bis-benzimidazole compound that binds non-intercalatively to the minor groove of double stranded DNA, with high specificity to the adenine and thymine regions (Green et al., 2012). It has peak excitation at 350 nm and emission at 461 nm (Cesarone et al., 1979). Samples of unknown concentration are assayed against a standard curve that is generated using DNA of known concentration (see notes 10 & 11). High molecular weight DNA is required for making the standard curve because Hoechst 33258 does not bind efficiently to small DNA fragments (Green et al., 2012). Hoechst 33258 can be used with whole cell lysates as it does not bind to protein, or significantly to RNA (Green et al., 2012).

1. Reconstitute standard stock DNA by adding 100 μg of calf thymus DNA to 1.0 mL of TE buffer (pH 8.0). Leave overnight at 2-8°C to completely dissolve the DNA.

2. Measure the concentration of the standard stock DNA using a spectrophotometer at 260 nm. Multiply the reading by 50. This gives a concentration in μg/mL.
3. Make the standards with a series of 1:1 dilutions of the standard DNA and TE (pH 8.0). Perform 5 serial dilutions, recording the weight at each step for accurate calculation of the DNA concentration. This will give six known concentrations for the standard curve, in addition to a blank.
4. Remove the freeze-dried samples from the -80°C freezer and allow to return to room temperature (see note 8).
5. Weigh and record the weight of each tube with the sample in it (see note 9).
6. The freeze-dried samples need to be rehydrated for analysis. Add 3.0 ml of rehydration buffer and record the weight.
7. Incubate at 37°C for 1 hour.
8. Make up a Hoechst working solution. Add 50 μl of Hoechst 33258 (10 mg/mL) to 50 mL of the fluorometry buffer. Keep solution covered with aluminum foil.
9. Prepare the 96 well plate. For each standard and sample, prepare 3 wells. Transfer 100 μL of the Hoechst working solution into each allocated well. Include a single well for the blank. Weight the plate after each addition and record the exact weight of working solution added.
10. Mix each sample and standard before adding 10 μL to an allocated well. Weigh and record the weight of the plate after each addition.
11. Mix the plate and place in the plate reader.
12. Measure at 350 nm excitation and 461 nm emission.
13. Subtract the blank value from all measurements.
14. Calculate the exact concentrations of the standards using the recorded weights at each step.
15. Make the concentration standard curve.
16. Calculate the concentrations of each replicate before calculating the average concentration of the sample (see note 12).
17. Calculate the concentration of the initial rehydration reaction using the recorded weights (see note 11).
18. Calculate the amount of DNA using the concentration and rehydration reaction volume.
19. Normalize the amount of DNA to the displacement volume of the initial biomass.

2.4.6 RNA quantification

In this assay, an alkali is used to hydrolyze the sample before the cellular components are separated by their solubility in an acidic solution. The UV absorbance is then used to quantify the amount of RNA in solution. The addition of potassium hydroxide hydrolyses the RNA into monomers. These RNA monomers are acid soluble. The hydrolysis of carbohydrates does not produce compounds that affect the absorbance at 260 nm, while the hydrolysis of protein does produce compounds that interfere with the absorbance reading. This production of interfering compounds is restricted by the use of low molarity alkali and a reaction time of 1 hour (Fleck & Munro, 1962). The main proportion of protein is then precipitated when the reaction is made acidic. This acidification also precipitates the DNA fraction that remained stable under the initial alkaline conditions. The precipitated protein and DNA is then removed by centrifugation and the amount of RNA quantified by measuring absorbance at 260 nm.

1. Remove the freeze-dried samples from the -80°C freezer and allow to return to room temperature (see note 8). These are the samples that were prepared and lyophilized in glass tubes (see note 6).
2. Weigh and record the weight of each tube with the sample in it (see note 9).
3. Add 3 mL 0.3 M KOH to each tube and incubate for 60 minutes in a water bath at 37°C to hydrolyze each sample. Mix the solution every 15 minutes.
4. Cool the solution to room temperature.
5. Add 1 mL of 3 M $HClO_4$ and mix before recording the weight.
6. Centrifuge at 8000 g for 10 minutes at 4°C.
7. Transfer the supernatant to a labeled glass tube that can hold at least 15 mL.
8. Resuspend the precipitate in 4 ml cold 0.5 M $HClO_4$. Record the weight added.
9. Centrifuge at 8000 g for 10 minutes at 4°C.
10. Remove the supernatant and add it to the test tube that contains the supernatant from step 7.
11. Repeat steps 8 to 10, combining the $HClO_4$ supernatants.

12. Add 0.5 M $HClO_4$ to the tube containing the combined supernatants, to a total volume of 15 ml.
13. Centrifuge the tube at 8000 g for 10 minutes at 4°C.
14. Transfer to a quartz cuvette.
15. Measure the absorbance at 260 nm.
16. Multiply the absorbance by 40 to quantify the amount of RNA in μg/mL.
17. Calculate the total volume of the cumulative amounts added, taking into account the specific gravity of each solution.
18. Calculate the total amount of RNA using the concentration and total reaction volume.
19. Normalize the amount of RNA in each sample to the displacement volume of the initial biomass.

2.4.7 Carbohydrate quantification

Quantification of the total carbohydrate content is done using the phenol method first described by DuBois et al. (1956), which has since had several modifications to improve accuracy (Dubois et al., 1956; Rao & Pattabiraman, 1989; Taylor, 1995). The addition of concentrated sulfuric acid dehydrates the carbohydrates into their furfural derivatives (Rao & Pattabiraman, 1989). Once the reaction has been cooled, phenol is added that condenses with the furfurals to produce a yellow color (Dubois et al., 1956). Absorbance is measured at 488 nm and carbohydrate concentration is calculated using a standard curve made with glucose.

1. Make a standard series by diluting the glucose standard stock solution with MilliQ water. Add 0, 1, 5, 10, 25, 50, 75, 100, 125 and 150 μL of the glucose standard stock solution to labeled test tubes. Make the total volume of each standard up to 1.0 mL with MilliQ water. Weigh and record the weights of both glucose stock solution and water as they are added. Calculate the concentration of each standard.
2. Remove the freeze-dried samples from the -80°C freezer and allow to return to room temperature (see note 8).

3. Weigh and record the weight of each tube with the sample in it (see note 9).
4. Add 1.0 mL of MilliQ water and record the weight (see note 13).
5. Add 5.0 mL of concentrated H_2SO_4 to each of the sample and standard tubes, and invert several times to mix. Record the weight of H_2SO_4 added.
6. Let the reaction stand for ten minutes.
7. Rapidly cool the tubes by placing in iced water for 1 minute.
8. Add 50 μL of phenol to each of the sample and standard tubes, and mix.
9. Record the weight of phenol added.
10. Incubate the tubes for 30 minutes in a water bath set to 25°C.
11. Measure the absorbance at 488 nm (see note 14 & 15). Take 3 replicate readings for each standard and sample.
12. Average the 3 readings for the replicate standard solutions.
13. Calculate the absorbance verse concentration standard curve.
14. Average the 3 replicate readings per sample.
15. Compare the average absorbance of each sample to the standard curve to calculate the concentration of carbohydrates in each sample.
16. Calculate the total volume of the cumulative amounts added, taking into account the specific gravity of each solution.
17. Calculate the total amount of carbohydrates using the concentration and total reaction volume.
18. Normalize the amount of carbohydrates of each sample to the displacement volume of the initial biomass.

2.4.8 Protein quantification

The Bradford assay is used to quantify the protein content. Sodium hydroxide is used to digest the cellular biomass, as the high concentrations of detergents normally required for cell lysis are not compatible with the Bradford assay (Bradford, 1976). The Bradford assay exploits the change from red to blue of the dye, Coomassie Brilliant Blue 250, upon binding to protein. The reaction is complete in 2 minutes, and is sensitive (Bradford, 1976).

1. Make a set of standard solutions by serially diluting the 2 mg/mL BSA with MilliQ water. Weigh and record the volumes added. Calculate the concentration of each standard.
2. Remove the freeze-dried samples from the -80°C freezer and allow to return to room temperature (see note 8).
3. Weigh and record the weight of each tube with the sample in it (see note 9).
4. Add 3.0 mL of 1 M NaOH to each standard and sample tube and record the weight.
5. Incubate in a water bath for 5 minutes at 95°C.
6. Cool the reaction in iced water.
7. Vortex the tube and centrifuge for 5 minutes at 2000 g to pellet any particulates.
8. For each standard and sample, label 3 tubes.
9. Invert the Coomassie working reagent bottle several times before adding 1.5 mL of to each tube. Record the weight added.
10. Add 50 μL of BSA standard or sample supernatant to the Coomassie working reagent and mix. Weigh and record the volume added (see note 10). A dilution step may be performed here by substituting 1 M NaOH for a fraction of the 50 μL.
11. Incubate for 10 minutes (see note 16).
12. Measure the absorbance at 595 nm, using MilliQ water as a blank.
13. Subtract the absorbance value of the blank from the absorbance of the standards and samples.
14. Calculate the absorbance verse concentration standard curve.
15. Compare the absorbance readings to the standard curve to calculate the concentration of protein in each replicate.
16. For each replicate, calculate the total volume of the cumulative amounts added, taking into account the specific gravity of each solution.
17. Calculate the total amount of protein in each replicate using the concentration and total reaction volume.
18. Calculate the amount of protein in each sample by averaging the 3 replicates.
19. Normalize the amount of protein to the amount of starting biomass used in the assay.

2.4.9 Lipid quantification

Quantification of lipids is done gravimetrically. The lipids are extracted with a 3:2 hexane:isopropanol solution overnight (Carnicer et al., 2009). Cellular components other than lipids are removed by the addition of an aqueous phase. The phases are then separated by centrifugation and the upper hexane layer is transferred. The aqueous layer is washed with 7:2 hexane:isopropanol to recover the total lipid fraction (Hara & Radin, 1978). The hexane is then evaporated under nitrogen and the remaining lipids quantified by weighing.

1. Remove the freeze-dried samples from the -80°C freezer and allow to return to room temperature (see note 8). These are the samples that were prepared and lyophilized in glass tubes (see note 6).
2. Weigh and record the weight of each tube with the sample in it (see note 9).
3. Add 1.0 mL of MilliQ water.
4. Add 5.0 mL of 3:2 hexane:isopropanol.
5. Place the tubes on a slow rotation wheel and extract overnight.
6. Add 5.0 mL of 0.47 M $NaSO_4$ and centrifuge at 4000 g for 10 minutes.
7. For each sample, label and weigh a glass tube with lid included.
8. Pipette the top layer of hexane into the glass tube.
9. Add 5 mL of the 7:2 hexane:isopropanol solution to the $NaSO_4$ solution in the initial tube and mix by inverting for 5 minutes.
10. Centrifuge at 4000g for 10 minutes.
11. Pipette off the hexane supernatant and add to the hexane supernatant in the glass tube from step 9.
12. Evaporate the hexane under a slow flow of nitrogen.
13. Immediately cap the tube while still under nitrogen and weigh it. The increase in weight of the tube is the weight of the lipids present in the sample.
14. Normalize the amount of lipid in the sample to the displacement volume of the starting biomass.

2.4.10 Converting the composition abundance into a biomass equation

Now that I have presented essential protocols to measure the macromolecular composition of the biomass, this information needs to be converted into a biomass equation, or a set of molar drain fluxes that allow the growth rate to be used to estimate the molar flux through a pathway. This is a challenging task that requires a couple of organism specific assumptions and it is beyond the scope of this chapter to address all of them needed for *A. queenslandica*, but here I highlight the procedure using DNA as an example. Generally, to complete the whole biomass equation will require the average amino acid composition of the proteins, the sugar composition of the carbohydrate, the fatty acid composition of the lipids, the nucleotide composition of DNA and RNA. Furthermore some assumptions about the localization of the components must be taken (e.g. distribution of carbohydrates in storage compounds vs cell walls). Nevertheless it is useful to run through the procedure with one component (DNA) as an example as it should highlight the most important steps, especially the conversion of weights and percentages into molar coefficients:

1. Obtain the GC content of the DNA (see note 17) in *A. queenslandica* it is 31.14 %.
2. Translate the GC content into 15.57 % Guanine, 15.57 % Cytosine, 34.43 % Adenine, and 34.43 % Thymine.
3. Calculate molecular weights of the monophosphate-deoxyribonucleotides dxMP (see note 18) M_{dAMP} = 313 g / mol; M_{dTMP} = 304 g / mol; M_{dGMP} = 329 g / mol; M_{dCMP} = 289 g / mol;
4. Calculate the average molecular weight of one DNA nucleotide: (0.3443 x 313) + (0.3443 x 304) + (0.1557 x 329) + (0.1557 x 289) = 308.65 g / mol.
5. Convert the measured DNA content from g_{DNA} / g_{CDW} into mol_{dxMP} / g_{CDW}. In our sponge example the measured DNA content (Table 2.1) was determined to be 0.031 mg_{DNA} / $ml_{displacement}$. A further conversion is required (see note 19): 1.0 ml displacement equates to 0.091g dry sponge. 64 % of that is skeleton (assumed DNA free), while 36 % is non-skeleton tissue or cell dry weight (CDW). Therefore: 0.36 x 0.091 = 0.03276 g_{CDW} / $ml_{displacement}$. This brings the DNA content to 0.946 mg_{DNA} / g_{CDW} and 3.065 μmol_{dxMP} / g_{CDW}, respectively.

Table 2.1 The macromolecular composition of *Amphimedon queenslandica*.

Component	Displacement (mg / ml)	Standard deviation	Relative error	n
Dry weight	92.49	10.22	11.05	42
Skeleton	62.68	10.93	17.44	21
Lipid	10.90	1.79	16.42	19
Protein	8.28	1.14	13.76	20
Carbohydrate	1.66	0.43	25.96	20
RNA	0.198	0.037	18.78	20
DNA	0.031	0.006	19.81	20

6. Using the ratios derived from the GC content from above I can calculate the individual nucleotide yield coefficients for DNA synthesis to be: 0.477 $\mu mol_{dGTP} / g_{CDW}$; 0.477 $\mu mol_{dCTP} / g_{CDW}$; 1.055 $\mu mol_{dATP} / g_{CDW}$; 1.055 $\mu mol_{dTTP} / g_{CDW}$ (see note 20).
7. If a comprehensive genome-scale model is present, one can stop here and use the yield coefficients provided above. If the model does not contain the nucleotide biosynthetic pathways, these yields can be further fragmented into the drains of central metabolites as outlined in the next step.
8. dATP: Ribose-5P + 7 ATP + 2 GLN + CO_2 + GTP + NADPH + 2 ASP + 2 Formyl-THF + GLY → dATP + 6 ADP + AMP + 2 GLU + $NADP^+$ + 2 FUM + GDP + 2 THF + PPi + 4 Pi (see note 21).
9. dGTP: Ribose-5P + 8 ATP + 3 GLN + CO2 + NADPH + ASP + 2 Formyl-THF + GLY + NAD^+ → dGTP + 6 ADP + 3 GLU + 2 AMP + NADP + FUM + 2 THF + NADH + 2 PPi + 3 Pi.
10. dCTP: HCO_3^- + 2 GLN + 6 ATP + ASP + UQ + Ribose-5P + NADPH → dCTP + 2 GLU + 5 ADP + AMP + UQH_2 + CO_2 + 3 Pi + PPi + $NADP^+$.
11. dTTP: HCO_3^- + 2 GLN + 8 ATP + ASP + UQ + Ribose-5P + NADPH + 5,10 methylene-THF → dTTP + 2 GLU + 7 ADP + AMP + UQH_2 + CO_2 + 3 Pi + PPi + $NADP^+$ + DHF + NH_3.

12. Multiply the stoichiometric coefficients in above equations with the individual dxTP yield coefficient outlined above and you can calculate the total drain of precursors that DNA synthesis brings about (in μmol / g_{CDW}). In our example this results in:
13. 3.06 Ribose-5P + 22.5 ATP + 6.605 GLN + 1.1 GTP + 3.06 NADPH + 4.12 ASP + 3.06 Formyl-THF + 1.53 GLY + 1.05 5,10-methylene-THF + 1.5 UQ + 1.53 HCO_3^- + 0.5 NAD^+→ DNA + 19 ADP + 3.5 AMP + 6.6 GLU + 3.06 $NADP^+$ + 3 FUM + 1.05 GDP + 3.06 THF + 3.54 PPi + 10.25 Pi + 1.05 NH3 + 1.05 DHF.
14. In analogy to step 7, one could stop on this level. If a model, however, is simpler than this equation, the equation can be further condensed: For instance, using the following reactions:
15. GDP + ATP→ GTP + ADP; GLU + NH3 + ATP → GLN + ADP; AMP + ATP → 2 ADP the above term becomes:
16. 3.06 Ribose-5P + 33.7 ATP + 5.55 NH3 + 3.06 NADPH + 4.12 ASP + 3.06 Formyl-THF + 1.53 GLY + 1.05 5,10-methylene-THF + 1.5 UQ + 1.53 HCO_3^- + 0.5 NAD^+→ DNA + 33.7 ADP + 3.06 $NADP^+$ + 3 FUM + 1.05 GDP + 3.06 THF + 3.54 PPi + 10.25 Pi + 1.05 DHF.
17. Multiplication of the equation with the growth rate gives fluxes in (in μmol / (g_{CDW} x h))

This section hopefully highlights essential steps in assembling the appropriate equations with yield coefficients and how to get from mass measurements to molar yield coefficients. It also highlights that this is a complex task that need to be done for each biomass component separately using the approach described above. The level on which the fragmentation of the reactions stops will always depend on the model. As such, the more comprehensive the model, the less work needs to be put into the condensation of reactions into common precursors. The drawback: More complicated models may only be used for certain approaches, for instance a genome-scale model is well suited for flux balance analysis, while ^{13}C fluxomics with a full genome-scale isotopomer model is currently difficult to analyse. Therefore the user will have to condense reactions manually as described above to a point that is suitable for the flux analysis that is being conducted.

2.5 Notes

1. EDTA will not dissolve until the pH reaches 8. Stirring with a magnetic stirrer and applying gentle heat (50°C) will increase the rate of dissolution.
2. The freezing and hydration process was enough to lyse the sponge cells but not the bacterial cells. I added lysozyme to the TE used to rehydrate the sample to enzymatically lyse the bacterial fraction. Note that lysozyme does not lyse Archaea because of different cell wall composition to bacteria. *A. queenslandica* has a low abundance of Archaea.
3. Perchloric acid is highly reactive with metals and organic matter. Only use glass pipettes and tubes.
4. Change scalpel blade very frequently. The skeleton of *A. queenslandica* contains silica (spicules). This causes the scalpel to rapidly lose its edge. Avoid placing pressure on the blade to prevent the loss of cellular material that is easily dislodged. Be careful to avoid foreign material such as macroalgae or inorganic debris, which is often incorporated into the sponge biomass as it grows.
5. Excess water must be removed when transferring the sponge piece from the deep petri dish to the measuring cylinder. Briefly touch the bottom of the sponge piece to the edge of the petri dish. Water should be left in the aquiferous system to prevent trapped air affecting the displacement value.
6. Use glass tube for RNA and lipid samples. Hexane will react with plastic tubes.
7. The evaporation of seawater leaves salt that significantly contributes to the dry weight. I assumed that the intracellular salt content was the same as seawater for our calculations.
8. To avoid condensation, do not open the tubes until the sample is at room temperature.
9. Remove any water that has condensed on the outside of the tube while it was warming to room temperature.
10. Each standard curve has a linear range within which it is accurate. Sample concentrations that fall outside this range must be diluted as appropriate to ensure an accurate analysis. Use samples that are uniform in size and assay test samples prior to analyzing the sample set. Published data for a closely related organism may serve as a guide for expected concentrations.

11. Small differences in the amounts of solutions used during an assay can produce cumulative error that is significant. To ensure accuracy, all weights of solutions added to an assay are determined gravimetrically on an analytical balance. The exact volume added can then be calculated using the specific gravity of the solution.
12. Do not average the readings and then calculate concentration. The replicate wells will have different concentrations due to differences in the amount of stock solution and diluent. Rather average the calculated concentrations of the replicates.
13. Water is important for the generation of heat essential for the reaction to be completed.
14. A dilution may be required, depending on the organism being analysed. The reaction with the sponge tissue resulted in an almost black solution. I made a dilution of 1 in 50 with concentrated sulfuric acid. Record weights when performing a dilution step for accurate concentration calculations.
15. Dependent on the reaction, filtration of the acid solution may be required to remove particulate material that may interfere with spectrophotometer readings.
16. It is important to incubate all reactions for the same amount of time, as the absorbance continues to change.
17. The GC content for the genome should be available from the genome sequence. If not sequenced, use the GC content of a closely related species.
18. dxMP stands for dAMP, dTMP, dGMP or dCMP, respectively. Remember that the dxMP's in DNA are polymerized, which happens through the loss of PPi from dxTP's. Using the molecular weight of dxTP's I need to subtract the molecular weight of PPi ($P_2O_7H_4$; M = 178 g/mol) to obtain the molecular weight of one dxMP in the chain. E.g. M_{dATP} = 491.18 g/mol (fully protonated); incorporated in the chain: M_{dAMP} = 313.18 g/mol ($C_{10}H_{12}N_5O_5P$).
19. For single cell organisms it should usually be possible to obtain $mg_{macrocomponent}$ per g_{CDW} straight from the described assays.
20. For the yield coefficients the triphosphates of the nucleotides are given, since they need to be synthesized for DNA polymerization and will be the drain points in the network.
21. The stoichiometric coefficients are derived from Chapter 3.6 (Michal & Schomburg, 2013)

Chapter 3 - Analysis of the biomass composition of the demosponge *Amphimedon queenslandica* on Heron Island Reef, Australia

3.1 Abstract

Marine sponges are a potential source of important pharmaceutical drugs, the commercialisation of which is restricted by the difficulties of obtaining a sufficient and regular supply of biomass. One way to optimize commercial cell lines for production is the in-depth characterization and target identification through genome-scale metabolic modelling and flux analysis. By applying these tools to a sponge, I hope to gain insights into how biomass is formed. I chose *Amphimedon queenslandica* as it has an assembled and annotated genome, a prerequisite for genome-scale modelling. The first stepping-stone on the way to metabolic flux analysis in a sponge holobiont, is the characterisation of its biomass composition. In this study we quantified the macromolecular composition and investigated the variation between and within sponges of a single population. I found lipids and protein to be the most abundant macromolecules, while carbohydrates were the most variable. I also analysed the composition and abundance of the fatty acids and amino acids, the important building blocks required to synthesise the abundant macromolecule types, lipids, and protein. These data complement the extensive genomic information available for *A. queenslandica* and lay the basis for genome-scale modelling and flux analysis.

3.2 Introduction

Marine sponges and their associated bacteria are prolific producers of secondary metabolites that are an important source of bioactive compounds with pharmacological potential. More than 5000 compounds have been identified to date, including 355 new compounds reported in 2012 alone

(Blunt et al., 2014). However, despite this enormous potential, approval for use as a therapeutic compound is scarce (Huyck et al., 2011). The lack of commercial development of sponge-derived compounds is attributed to both a biomass supply problem—no reliable method to culture either whole sponges or cells exists—and a lack of understanding of sponge-bacterial interactions. Together, these constraints limit the applicability of metabolic engineering approaches to microbial isolates or sponge cells for over-production of secondary metabolites. To move forward in the development of sponge-derived drugs requires extending the research activities beyond the current major focus on identifying secondary compounds and, to a limited extent, characterising the biosynthetic pathways that produce these secondary metabolites (Genta-Jouve & Thomas, 2012). We need to understand how the sponge holobiont (the animal and its resident microbes) is able to grow; understanding how the central metabolic pathways contribute to biomass production is crucial to overcoming the biomass supply problem.

The metabolic processes underlying sponge growth can be comprehensively investigated by the application of genome-scale modelling and flux analyses (Oberhardt et al., 2009). An essential step of such analyses is characterisation of the biochemical composition of the sponge holobiont, with a focus on the components that contribute significantly to overall biomass production (Thiele & Palsson, 2010). Characterising the biochemical composition allows the formation of a biomass equation, essential for genome-scale modelling and flux analysis as it constrains hundreds of fluxes through the system (Feist & Palsson, 2010). The focus on secondary metabolic compounds has left the products of central metabolism comparatively overlooked, including analysis of the major cellular macromolecules—protein, lipid, carbohydrate, DNA, and RNA—and their constituent building blocks—amino acids, fatty acids, sugars and nucleotides. Although protein, lipid, and carbohydrate content has been quantitated in a number of marine sponges (Chanas & Pawlik, 1995; Elvin, 1979; Koigoora et al., 2013; McClintock, 1987; Pallela et al., 2011; Stone, 1969), a complete analysis of the main biomass components has not been reported.

In addition to quantifying the components that contribute to biomass formation, information about how variable the composition is across the population and within an individual sponge is important for guiding the design of experiments associated with flux analysis. Although sponges appear to be morphologically simple and homogenous, it has been shown that production of secondary metabolites can be controlled and restricted to specific areas of an individual (Freeman & Gleason, 2010). This apparent heterogeneity may be related to physiological differences in the sponge itself or in the associated microbial community. Marine sponges typically host large and diverse communities of microbes (Hentschel et al., 2012), often in a symbiotic relationship (Thomas et al., 2010). These microbes have been shown to be the source of certain secondary metabolites (Wilson et al., 2014), and also to interact with central metabolic pathways in the host sponge (Radax et al., 2011). Thus the distribution of particular bacterial species and communities within an individual sponge has the potential to influence local metabolic activity. Nonetheless, subsampling the main biomass components within an individual sponge typically has not been performed, as most biochemical analyses performed in sponges have been restricted to a single sample from an individual or small number of individuals within a population (Chanas & Pawlik, 1995; Elvin, 1979; Koigoora et al., 2013; McClintock, 1987; Pallela et al., 2011; Stone, 1969). Analysing multiple samples within individuals will give insights into the variation in the macromolecule composition within an individual sponge.

I report here the first analysis of the products of central metabolism of the marine demosponge, *Amphimedon queenslandica* from the Southern Great Barrier Reef. I have quantified the protein, lipid, carbohydrate, DNA and RNA content of this sponge. These analyses are supplemented with a detailed analysis of amino acid and fatty acid composition. By analysing multiple areas within multiple individuals within one *A. queenslandica* population I reveal marked variation within and between sponges for most of these macromolecular classes. These results, along with the extensive genomic resources available for this demosponge (Srivastava et al., 2010), provide a strong foundation for future detailed understanding of the metabolic pathways that contribute to biomass production.

3.3 Methods

3.3.1 Sample Collection and Preparation

A. queenslandica were collected from Shark Bay, Heron Island, Australia (23°26′37.92″ S, 151°55′8.81″ E). Individual sponges and their coral rubble substrate were collected with a hammer and chisel (Leys et al., 2008) and transported submerged to the Heron Island Research Station. Sponges were held in a flow through aquaria system and processed within 8 h of collection. Biopsy samples were collected from a total of 29 sponges. The samples were cut with a sterile scalpel and contained a representative cross section of sponge biomass that consistently included the pinacoderm layer through to biomass just above the substratum. A thin layer of biomass was removed from the bottom of the sponge to avoid contamination by the substrata. Pieces of sponge biomass that had foreign material incorporated were avoided. To investigate the variation in composition within individual sponges, biopsies were collected from each sponge in multiples of 5. This resulted in 20 sponge samples per analysis, comprising 5 biopsy samples each of 4 individual adult sponges. Samples were snap frozen in liquid nitrogen then stored at −80 °C and transported at on dry ice. Lyophilization was performed overnight. Dry weights were recorded on an analytical balance and corrected for salt content by multiplying the amount of water lost through lyophilization by the specific gravity of seawater. We assumed that the salt content of the cells was the same as the seawater. Throughout the study, the weight of each reagent in each reaction was recorded. The specific gravity of each reagent was then used to ensure that accurate calculations were made.

3.3.2 Skeleton Dry Weight

Cellular material was removed from the *A. queenslandica* biopsies by re-hydrating with a 2 % SDS solution made with MilliQ water and incubating for 2 h at room temperature on a shaker table. The biopsies were then washed thoroughly for 5 min with running RO water. The SDS incubation and washing times were optimized by inspecting with a Nikon Eclipse Ti-E inverted microscope

(Nikon, Tokyo, Japan) to ensure all cellular material was removed. The pieces were then lyophilized overnight and dry weights recorded. These samples were stored at −80 °C.

3.3.3 Nucleic Acids

DNA was quantified by Hoechst fluorescence assay. The lyophilized samples from section 3.3.1 were re-hydrated with 3.0 mL of TE buffer (pH 8.0) with 0.01 g / mL lysozyme and incubated at 37 °C for 1 h on a shaker table set to 100 rpm. Re-hydrating the lyophilized biopsies caused complete lysis of the *A. queenslandica* cells, but not the bacterial fraction. Lysozyme was added to ensure lysis of the bacterial fraction. After incubation, sterile tweezers were used to remove the sponge skeleton from the tube. The piece of skeleton was squeezed and rolled against the inside wall to ensure that the lysis buffer was retained. Samples were then vortexed and centrifuged for 5 min at 4000 *g* to pellet cellular debris. Aliquots of the supernatant were used in the Hoechst fluorescence assay. Calf thymus DNA was used to generate a standard curve (Green et al., 2012). All samples and standards were measured in triplicate.

RNA content was quantified by digesting the lyophilized samples from section 3.3.1 with 3.0 mL of 0.3 M KOH for 1 h at 37 °C (Benthin et al., 1991). The reaction was then acidified with the addition of 1.0 mL of 3 M perchloric acid (Fleck & Munro, 1962). Following centrifugation to remove DNA and protein, the RNA-containing supernatant was quantified in triplicate by UV absorbance at 260 nm (Benthin et al., 1991).

The nucleotide composition of both the DNA and RNA was calculated using the GC content of the genome and an adult transcriptome. The calculations assumed that the DNA and RNA is all sponge derived as the bacterial genomic and transcriptomic sequence data, in addition to bacterial diversity and abundance data, is not yet available. The GC content of the *A. queenslandica* genome and adult transcriptome are 31.14 % (Srivastava et al., 2010) and 39.76 % (Fernandez-Valverde et al.,

2015), respectively. The percentage contribution of each nucleotide was calculated, assuming an even guanine-cytosine and thymine-adenine/thymine-uracil ratios. Next the average molecular weight of one nucleotide was calculated. This was done by multiplying the abundance of each nucleotide by its molecular weight. These values were then added together to give the average molecular weight. This was multiplied by the mean total amount of DNA or RNA (g / gDW) to give the number of mols of average nucleotide per gram of dry weight. The amount of each nucleotide (mols) was then calculated using this value and the guanine-cytosine to thymine-adenine/thymine-uracil ratios. Values were converted to micromoles for reporting.

3.3.4 Protein

Sponge biopsies from section 3.3.1 were digested with 3.0 mL of 1.0 M NaOH incubated at 95 °C for 5 min. The reaction was cooled in iced water and then centrifuged at 4000 *g* for 5 min to pellet suspended skeleton solids. The supernatant was assayed in triplicate using the Coomassie Plus™ (Bradford) Assay Kit (Thermo Scientific, Rockford, MI, USA). Bovine serum albumin was used as the standard (Bradford, 1976).

3.3.5 Carbohydrate

Carbohydrate content was quantified using the phenol-sulfuric acid method (Dubois et al., 1956). The sponge biopsies from section 3.3.1 were rehydrated with 1.0 mL of MilliQ water. 5.0 mL of concentrated sulfuric acid was added and the reaction let proceed for 10 min. The reaction was rapidly cooled by placing the tube in iced water for 1 min (Rao & Pattabiraman, 1989). 50 µL of phenol was added and incubated for 30 min in a water bath set to 25 °C. The reaction was diluted 1:50 using concentrated sulfuric acid and the absorbance read at 488 nm. All measurements were taken in triplicate and glucose was used as the standard (Rao & Pattabiraman, 1989).

3.3.6 Lipid

Biopsy samples from section 3.3.1 were rehydrated with 1.0 mL MilliQ water and extracted overnight with 3:2 hexane:isopropanol (Hara & Radin, 1978). 5.0 mL of 0.47 M $NaSO_4$ was added, mixed and then separated by centrifugation at 4000 *g* for 5 min. The hydrophobic upper hexane phase contains the lipid fraction and was transferred to a clean glass tube. The initial solution was washed with 7:2 hexane:isopropanol to ensure that all of the lipids were recovered (Hara & Radin, 1978). This second hexane phase was again recovered with centrifugation and combined with the first extract. The combined hexane extracts were evaporated under a constant flow of nitrogen to prevent oxidation. The remaining solid lipids were quantified gravimetrically. The lipid solid was then stored under nitrogen at −80 °C for fatty acid analysis.

3.3.7 Fatty Acids and Sterols

A fatty acid methyl ester (FAME) analysis was performed on the extracted lipids from section 3.3.6 as previously described (Brennan et al., 2013). The lipid solid was dissolved in hexane and 5.1 μg of nondecanoic acid (Sigma-Aldrich, St. Louis, MO, USA) was used as an internal standard (ISTD) to a final concentration of 3.6 μg / mL. Saponification, methylation and GC/MS analysis of the lipid mass fraction followed the protocol described in (Brennan et al., 2013). The software program AMDIS (version 2.64, National Institute of Standards and Technology (NIST), Gaithersburg, MD, USA) was used for the analysis of the fatty acid and sterols. A minimum match factor of 80 and signal threshold of 700 was set for identifying compounds. A FAME library was built using the target ions and retention times of thirty-two known fatty acid standards (Sigma-Aldrich, St. Louis, MO, USA). An alkane ($C_{7\text{-}30}$) standard was used as a retention time index. To define compounds important for biomass formation, we used a cut-off of 0.05% of the largest peak in the MS analysis, which means that peaks below those signal intensities would not be reported. For unknown compounds (*i.e.*, that were not identified using the standard library), their fragmentation patterns were compared to the NIST database library (version 2.0, NIST, Gaithersburg, MD, USA). For a range of peaks both the target library as well as the NIST library did not provide accurate hits. Based on manual

interpretation of fragmentation patterns, unknowns that were found consistently in the majority (> 75 %) of the samples were classified into sterols or fatty acids. From the quantified lipid fraction (g / gDW) the estimated weight of glycerol and phosphate groups was subtracted assuming a glycerolipid to phospholipid ratio of 20:80. It was then assumed that the reminder of the lipid fraction would generate a signal in GC-MS and that the MS response and ionization efficiency for all FA and sterol ions was comparable (due to the unknown character of many FA's this cannot be tested, but individual errors for individual compounds will not impact dramatically on the later modelling). All peak areas were normalized to the total signal generated for that sample. Using the molecular weight of the knowns and an average molecular weight (calculated from all identified FA's and sterols) for the unknowns the individual concentrations were calculated in mol / gDW. A total of 23 separate fractions of biomass were analysed and the standard deviation was calculated.

3.3.8 Amino Acids

Amino acid analysis was performed by reversed phase high performance liquid chromatography. 20 samples containing cellular and skeleton components were analysed in addition to 20 samples of skeleton, prepared as section 3.3.1. 5 mL of 6 M HCl was added to each sample and incubated at 105 °C. Hydrolysis times were first optimized and found to be 54 h for the complete biomass samples and 48 h for skeleton samples. Optimization monitored the concentration increase of all amino acids over the time course and 54/48 h was found to be a time at which most amino acids reached the peak concentration. Longer times may favour some compounds but will lead to degradation of others. This compromise was necessary since otherwise individual experiments need to be conducted for all 20 amino acids. Samples were allowed to cool, mixed, and then a 500 µL aliquot was taken and filtered with a 0.22 µm Ultrafree-MC GV centrifugal filter (Merck Millipore, city, country). The HCl was evaporated by heating to 40 °C under a constant flow of nitrogen. The solid pellet was stored at −80 °C. Derivatisation was performed in a high-performance autosampler (Agilent HiP-ALS SL, G1367C, Santa Clara, CA, USA). 0.5 µL of sample containing 250 uM of internal standards, sarcosine and 2-aminobutyric acid, was added into 2.5 µL of borate buffer (0.4 N, pH 10.2, Agilent PN: 5061-3339, Santa Clara, CA, USA), mixed and incubated for 20 s at 4 °C. 1 µL of OPA reagent (10 mg

o-pthalaldehyde/mL in 3-mercaptopropionic acid, Agilent PN: 5061-3335, Santa Clara, CA, USA) was then added to initially derivatise primary amino acids. The reaction was mixed and incubated for 20 s at 4 °C. Then 0.4 µL of FMOC reagent (2.5 mg 9-fluorenylmethyl chloroformate/mL in acetonitrile, Agilent PN:5061-3337, Santa Clara, CA, USA) was added, mixed and incubated for 20 s at 4 °C to derivatise the secondary amines proline and sarcosine. 45.6 uL of Buffer A (40 mM Na_2HPO_4, 0.02 % NaN_3, pH 7.8) was added to lower the pH of the reaction prior to injecting the 50 µL reaction onto an Agilent Zorbax Extend C-18 column (3.5 µm, 4.6 × 150 mm, Agilent PN: 763953-902, Santa Clara, CA, USA) with a guard column (SecurityGuard Gemini C18, Phenomenex PN: AJO-7597, Santa Clara, CA, USA). Column temperature was kept at 39 °C in a thermostatted column compartment (Agilent TCC, G1316B, Santa Clara, CA, USA). Chromatography was performed using an Agilent 1200-SL HPLC system, equipped with an active seal wash and a degasser (Agilent Degasser, G1379B, Santa Clara, CA, USA). The HPLC gradient was 2 % – 45 % buffer B (45 % acetonitrile, 45 % methanol and 10 % water) from 0 to 18 min, 50 % – 60 % buffer B from 18.1 to 20 min. 100 % buffer B from 20.1 to 24 min, and 2 % buffer B from 24.1 to 27 min—using a binary pump (Agilent Bin Pump SL, G1312B, Santa Clara, CA, USA). Flow rate was 2 mL / min. Derivatised amino acids were monitored using a fluorescence detector (Agilent FLD, G1321A, Santa Clara, CA, USA). OPA-derivatised amino acids were detected at 340_{ex} and 450_{em} nm from 1 to 18 min, and FMOC-derivatised amino acids at 266_{ex} and 305_{em} nm from 18 to 27 min. Chromatograms were integrated using ChemStation (Rev B.03.02, Agilent, Santa Clara, CA, USA).

3.3.9 Statistical Analysis

To analyse how the variation between and within individual sponges contributed to the overall variation for each macromolecule separately, the data were fit to a linear mixed effects model using the *nlme* package (Pinheiro et al., 2013) in the R environment (Ihaka & Gentleman, 1996).

The amount of macromolecule (g / gDW) was the response variable and Sponge ID (individual) was the random effect. The among-individual variance was divided by the total variance and expressed as the percentage of the overall variability that was due to variation between individuals. To address whether this between-individual variation was significant, we used a likelihood ratio x^2 test to compare two models: one with the random effect and one without the random effect.

3.4 Results

3.4.1 Displacement to Dry Weight Composition Conversion Factor

Metabolic flux analysis requires the ability to quantify the biomass involved in the turnover of the substrate. In order to relate the amount of live sponge to the dry weight composition, I established a correlation between displacement volume of a sponge piece and its dry weight. The displacement volume and dry weight from 97 samples were used to estimate a conversion factor of 0.091 (std. dev. = 0.012) grams of dry sponge weight (gDW) per mL of displacement.

3.4.2 Overall Macromolecular Composition

For each biochemical analysis, I analysed five biopsies from several parts of four *A. queenslandica* individuals, all of which were collected from a single wild population in Shark Bay on Heron Island Reef, Australia (for detailed methods see Section 3.3.1). This sampling regime allowed us to document biochemical differences both within and between individual sponges. Analysis of mean values of the macro-components shows the skeleton (0.6342 g / gDW) to be the dominant component of the dry weight of *A. queenslandica* (Figure 3.1 and Table 3.1). The most abundant macromolecule type was lipid with 0.1251 g / gDW, followed by protein (0.0881 g / gDW), carbohydrate (0.0197 g / gDW), RNA (0.0021 g / gDW) and DNA (0.0003 g / gDW). The coefficient of variation revealed that carbohydrates were the most variable, followed by DNA, RNA, and lipids (Table 3.1). Protein and skeleton had the least overall variability (Table 3.1).

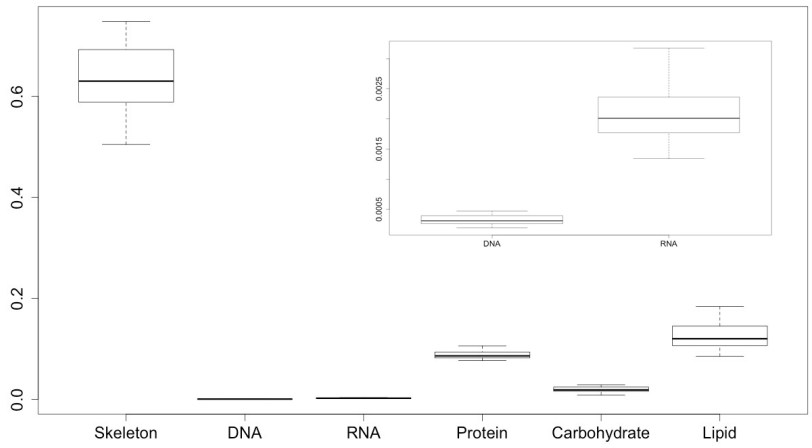


Figure 3.1 Mean skeleton and macromolecular composition of *A. queenslandica* calculated from all biomass samples summarised in Table 3.1.

The box represents the upper and lower quartile, split by the mean value. The upper and lower whiskers denote the minimum and maximum values. Units are grams per gram of dry weight.

Table 3.1 Mean skeletal and macromolecular composition of *A. queenslandica* calculated from all biomass samples.

Macro-Component	Mean (g/gDW)	Standard Deviation	Standard Error	Coefficient of Variation (%)	n
Skeleton	0.6343	0.0647	0.0152	9.81	18
Lipid	0.1252	0.0251	0.0058	20.09	19
Protein	0.0881	0.0082	0.0018	9.31	20
Carbohydrate	0.0197	0.0055	0.0012	27.53	20
RNA	0.0021	0.0005	0.0001	22.24	20
DNA	0.0003	0.0001	0.00002	25.78	19

3.4.3 The Source of Variation in the Overall Composition

I used a linear mixed effects model to investigate how much of the variation observed in the mean values (Table 3.2, Coefficient of variation column) was attributable to the between-individual variation. Over half of the variation seen in the skeleton (58 %) and lipids (54 %) was due to differences between individuals. The significance of this relationship was tested using a x^2 test. The variation of skeleton ($p = 0.009$) and lipids ($p = 0.016$) between individuals had significant effects on the overall variation of both macromolecule types (Table 3.2). Figure 3.2 shows the variation seen within each individual for the respective macro-components.

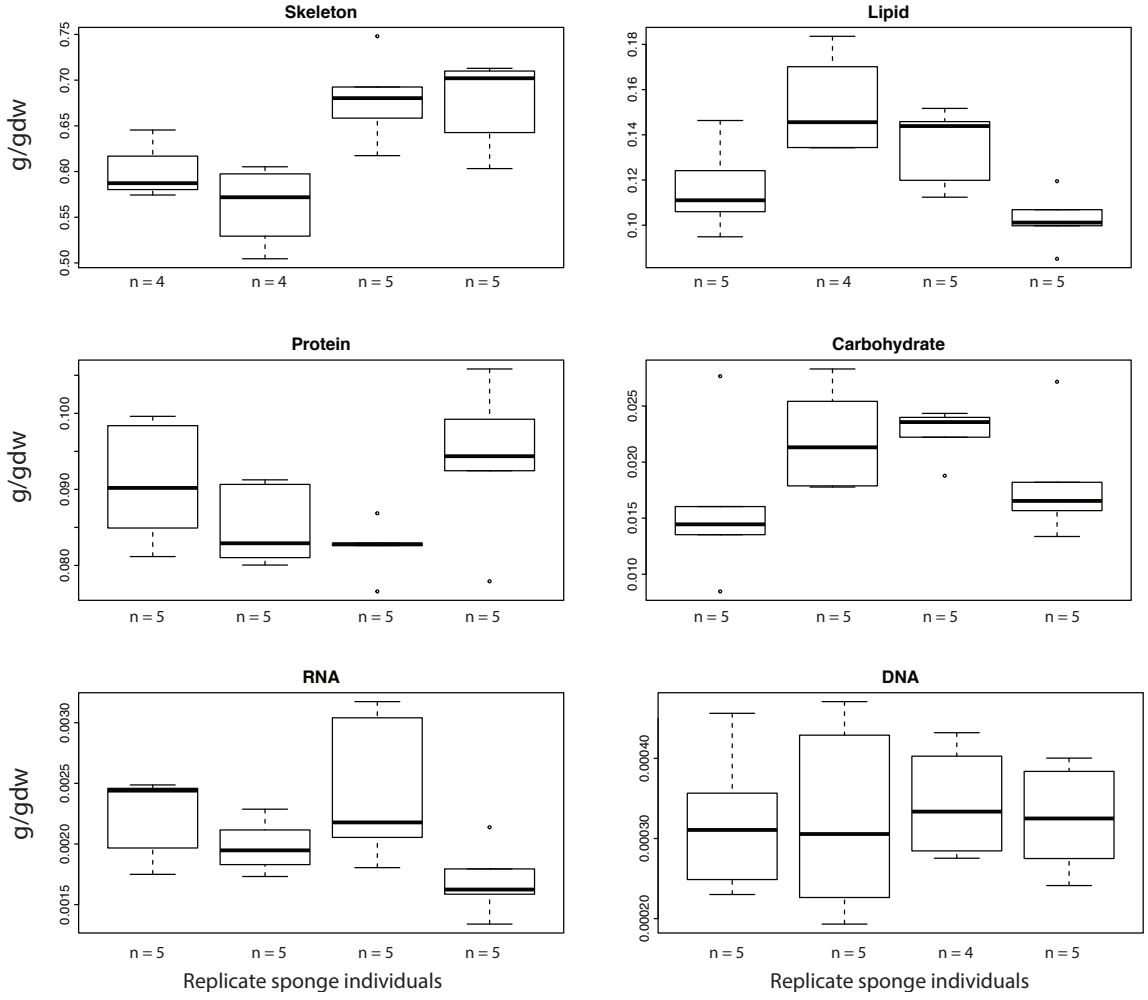


Figure 3.2 Variation of the composition of macro-components within and between individuals of *A. queenslandica*.

Four individual sponges were sampled for each macromolecule. Each box and whisker plot represents samples collected from an individual sponge. The box represents the upper and lower quartile, split by the mean value. The upper and lower whiskers denote the minimum and maximum values. All of the components apart from DNA showed marked variation between and within individual sponges.

Table 3.2 The source and degree of variation in macro-components.

Macro-Component	Percentage of Overall Variation Caused by between Individual Variation (%)	x^2	P-value
Skeleton	58	6.691	0.009
Lipid	54	5.763	0.016
Protein	24	1.393	0.237
Carbohydrate	15	0.597	0.439
RNA	1	2.309	0.128
DNA	32	3.53 × 10−8	0.999

3.4.4 Fatty Acid and Sterol Analysis

The lipid extractions were analysed using a fatty acid methyl ester (FAME) analysis using gas chromatography-mass spectrometry (GC-MS). I was interested in the main components that contribute to forming sponge lipids. Compounds were considered essential if they were found in most of the 23 replicate pieces of sponge. Eleven fatty acid (FA) compounds were identified and confirmed with standards. Palmitic acid was the most abundant FA, contributing 98.14 μmol / gDW. Octadecanoic acid (62.21 μmol / gDW) and an unknown FA (UU4-FA, 71.749 in Table 3.3) were other significant FAs. Three additional compounds matched compounds in the NIST database but were not confirmed with a standard (Table 3.3), this included cholesterol. Based on the fragmentation patterns, eleven compounds were classified as unknown FAs and one as an unknown sterol, highlighting the diversity and complexity of the lipid metabolism in *A. queenslandica*. If the total area under all peaks in the chromatogram is assumed to represent the total lipid fraction, then the main FA and sterols reported here cover 82.5% of the lipid fraction.

Table 3.3 Main compounds that contribute to the lipid biomass.

Compound	Lipid Number	μmol/gDW	Standard Deviation
Fatty Acids			
Docosanoic acid FAME	C22:0	9.769	4.299
Eicosanoic acid FAME	C20:1	1.947	0.579
Erucic acid FAME	C22:1ω9	1.721	0.668
Heptadecanoic acid FAME	C17:0	0.922	0.234
Myristic acid FAME	C14:0	3.957	1.455
Nervonic acid FAME	C24:1ω9	10.090	3.451
Octadecanoic acid FAME	C18:0	62.210	19.095
Palmitic acid FAME	C16:0	98.140	22.426
Pentadecanoic acid FAME	C15:0	7.065	2.468
Tetracosanoic acid FAME	C24:0	6.769	3.546
Tricosanoic acid FAME	C23:0	1.147	0.323
11-Eicosenoic acid FAME *	C20:1ω9	5.342	1.737
U6-FA **		1.005	0.580
U7-FA **		5.973	3.845
U8-FA **		2.279	0.722
U10-FA **		3.288	3.014
UU2-FA **		7.732	2.805
UU3-FA **		5.669	1.897
UU4-FA **		71.749	17.820
UU7-FA **		0.690	0.604
UU8-FA **		4.215	2.155
UU9-FA **		29.843	10.551
UU11-FA **		4.460	1.846
UU1 ***		1.728	1.766
UU6 ***		9.617	4.268

Compound	Lipid Number	µmol/gDW	Standard Deviation
Sterols			
Cholesterol *		4.706	2.477
Brassicasterol *		7.581	4.806
UU13-sterol **		6.264	3.375

* Compounds that did not match a compound in the standard mix but matched a compound in a database; ** interpreted from the fragmentation pattern; *** compounds that are neither FA nor sterols.

3.4.5 Amino Acid Analysis

I analysed the amino acid composition of 20 samples of complete sponge biomass (Table 3.4). Samples of biomass were lyophilized and the hydrolysis incubation time was optimised before quantifying the amino acids with high performance liquid chromatography (HPLC). I found glycine to be the most abundant amino acid with 4.582 mmol / gDW, followed by alanine (1.538 mmol / gDW). Due to the oxidation of asparagine and glutamine to aspartate and glutamate respectively during acid hydrolysis, a 50:50 abundance ratio was assumed between asparagine:aspartate and glutamine:glutamate. Tryptophan (0.089 mmol / gDW), tyrosine (0.157 mmol / gDW), and histidine (0.110 mmol / gDW) were the least abundant amino acids. Tryptophan and methionine are known to be degraded during the hydrolysis. While methionine was not detected tryptophan was measurable but *in vivo* concentrations may be higher than what was observed. Neither cysteine, nor cystine was measured with the methods used and thus not reported.

The skeleton is the most abundant component in *A. queenslandica* and is comprised of inorganic silica spicules and a collagen based extracellular matrix. Sponges produce a specific collagen, called spongin (Gross et al., 1956). As the skeleton contributes significantly to the overall sponge composition, I analysed the amino acid composition of the skeleton (Table 3.5). The hydrolysis incubation time was optimised before the analysis was performed on 20 samples. As above, a 50:50

Table 3.4 Mean amino acid composition of complete sponge biomass.

Amino Acid	mmol/gDW	Standard Deviation (mmol/gDW)	Amino Acid Contribution (mol/mol)
ALA	1.538	0.565	0.097
ARG	0.694	0.258	0.044
ASP	0.841	0.312	0.053
ASN	0.841	0.312	0.053
GLN	0.767	0.287	0.048
GLU	0.767	0.287	0.048
GLY	4.582	1.702	0.289
HIS	0.110	0.042	0.007
ILE	0.402	0.149	0.025
LEU	0.597	0.222	0.038
LYS	0.688	0.268	0.043
PHE	0.369	0.137	0.023
PRO	1.050	0.387	0.066
SER	0.806	0.318	0.051
THR	0.826	0.304	0.052
TRP *	0.089	0.053	0.006
TYR	0.157	0.060	0.010
VAL	0.753	0.281	0.047

* Tryptophan is largely destroyed during the hydrolysis, so concentrations in vivo may be higher than observed.

Table 3.5 Mean amino acid composition of the skeleton.

Amino Acid	mmol/gDW	Standard Deviation (mmol/gDW)	Amino Acid Contribution (mol/mol)
ALA	1.257	0.537	0.103
ARG	0.545	0.226	0.045
ASP	0.594	0.250	0.049
ASN	0.594	0.250	0.049
GLN	0.573	0.243	0.047
GLU	0.573	0.243	0.047
GLY	4.248	1.807	0.349
HIS	0.034	0.016	0.003
ILE	0.176	0.080	0.014
LEU	0.290	0.121	0.024
LYS	0.442	0.173	0.036
PHE	0.239	0.102	0.020
PRO	0.916	0.351	0.075
SER	0.534	0.211	0.044
THR	0.559	0.226	0.046
TRP	0.100	0.041	0.008
TYR	0.066	0.028	0.005
VAL	0.440	0.177	0.036

* Tryptophan is largely destroyed during the hydrolysis, so concentrations in vivo may be higher than observed.

ratio between aspartate and asparagine and between glutamine and glutamate was assumed. The most abundant amino acids in the skeleton were glycine with 4.248 mmol / gDW and alanine with 1.257 mmol / gDW. The least abundant amino acids were tyrosine (0.066 mmol / gDW) and histidine (0.034 mmol / gDW).

A comparison of the relative amino acid composition of the skeleton and the cellular component revealed that glycine was the most abundant amino acid for both components (Figure 3.3). The skeleton had relatively more glycine, proline, and alanine than the cellular fraction. The cellular fraction had a higher relative abundance of isoleucine, leucine, and valine than the skeleton. Histidine and tyrosine were the least abundant of the amino acids measured for both the skeleton and cellular components

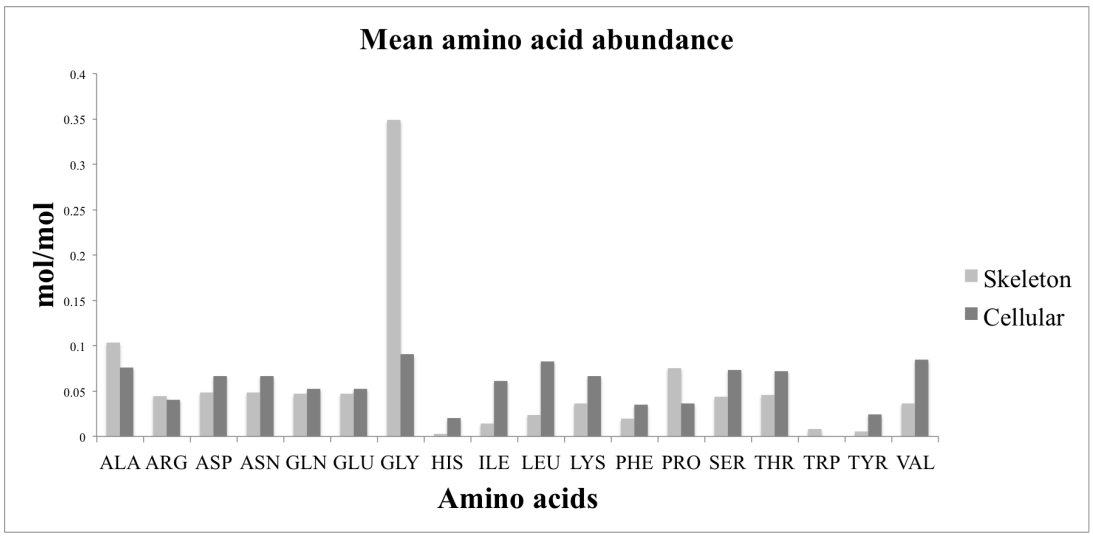


Figure 3.3 A comparison of the relative amino acid composition of the cellular biomass and the skeleton.

The amino acid composition of the cellular components was calculated by subtracting the quantified amino acid values of the skeleton (Table 3.5) from the values for the complete sponge biomass (Table 3.4). The skeletal and estimated cellular amino acid values were then normalized by converting to percentages.

3.4.6 Nucleotide Composition

The molar contribution of nucleotides to the DNA and RNA composition was estimated using available genome and transcriptome data. The genomic GC content was estimated to be 31.14 % using the publicly available genome (Srivastava et al., 2010). The RNA GC content was estimated to be 39.76 %, using an adult transcriptome (Fernandez-Valverde et al., 2015). These percentages were then converted into micromoles per gram of dry weight, using the molecular weight of a polymerized nucleotide (Tables 3.6 and 3.7).

Table 3.6 Deoxyribonucleotide composition.

	MW (g/mol)	mol/ mol DNA	μmol/gDW
dAMP	313	0.3443	0.37
dCMP	289	0.1557	0.17
dGMP	329	0.1557	0.17
dTMP	304	0.3443	0.37

Table 3.7 Ribonucleotide composition.

	MW (g/mol)	mol/ mol DNA	μmol/gDW
AMP	313	0.3012	2.03
CMP	289	0.1988	1.34
GMP	329	0.1988	1.34
UMP	308	0.3012	2.03

3.5 Discussion

3.5.1 Macromolecule Composition

To gain insights into how sponges produce biomass, the complex metabolic processes must be investigated as a complete functioning system. The systems biology approach of genome-scale metabolic modelling requires the biochemical composition of the organism to be known. The biochemical composition identifies the products of central metabolism that contribute to growth. With the products of central metabolism known, the synthesis pathways can be identified and the fluxes of metabolites modelled. As a first step towards genome-scale modelling, I analysed the macromolecular abundance and composition of the marine sponge *A. queenslandica* and investigated the variation seen across the population and within individuals.

The skeleton was the dominant component of the biomass, contributing 63 % of the dry weight. Lipids and protein were the most abundant macromolecules, while carbohydrates were the most variable (Table 3.1). Variation in the amount of skeleton and lipids between individual sponges had significant effects on the overall variability of the respective macro-components (Table 3.2). The apparent variation in the abundance of macromolecules within an individual (Figure 3.2) was not significant and had limited effect on the overall population variability. This implies that not only do the central metabolic pathways have an overall response to the individual's microenvironment, but localised stimulus could also cause compositional differences across an individual. This highlights the need for compositional and metabolic studies to sample appropriately to ensure accurate representation.

Overall, the macromolecular composition of *A. queenslandica* in the current study fits within the ranges reported for protein, lipids, and carbohydrates for 71 tropical marine sponge species (Chanas & Pawlik, 1995). *A. queenslandica* notably contained more lipids than protein (Table 3.1), a characteristic that has been reported previously for two marine demosponges, *Aplysinia archeri* and *Plaktortis halichondroides* (Chanas & Pawlik, 1995). The high lipid content in *A. queenslandica* could

indicate the accumulation of energy storage products in response to increased food availability and, thus, variably depending on food abundance. Alternatively, Elvin (1979) found that, for *Haliclona permollis*, an increase in lipid content was associated with oogenesis and the formation of mesenchymal amoebocytes. Although *A. queenslandica* is reproductively active year round, larval production peaks in the warmer months (Leys & Degnan, 2001). It is not clear if oogenesis is the reason for the high lipid content as the sampling time of this study corresponds with the end of summer when reproductive output is starting to decline. In addition, *A. queenslandica* concentrates its oocytes and embryos in localised brood chambers (Leys & Degnan, 2001), unlike many sponges, suggesting that these levels should vary markedly within an individual sponge.

In *A. queenslandica,* the cellular biomass accounts for less than half of the dry weight. I defined the skeleton as the solid, inorganic spicules and the flexible, organic collagen. The skeleton, being a product of cellular metabolism, is a considerable sink of metabolic resources, accounting for 64 % of the total dry weight (Table 3.1). The amount of skeleton in *A. queenslandica* is within the ranges of both Antarctic (McClintock, 1987) and tropical (Chanas & Pawlik, 1995) sponges. The high skeletal requirement by all sponges, greater than 50 % dry weight, is likely due to the lack of defined tissues, including connective tissues, that are required to support complex body plans. Without the significant support and framework that the skeleton provides, sponges would not be able to achieve the sizes observed across the phylum (Hooper, 2002).

Although sponges do not have defined tissues, the distribution of cells in *A. queenslandica* does not appear to be even if DNA is used as a proxy for cell number. This could be in part due to the large size range of sponge cells, as smaller cells are able to occupy less space yet all cells contain the same amount of DNA. Cells typically contain more RNA than DNA (Nelson & Cox, 2012), and this is also the case for *A. queenslandica*, in which RNA was more abundant than DNA.

3.5.2 The Variation in Macromolecule Composition

Metabolic networks are dynamic, plastic, and redundant (Nelson & Cox, 2012). These characteristics have evolved to give organisms the flexibility to rapidly adapt to variations in environment or external stimulus (Nelson & Cox, 2012). As the aim of this study is to gather baseline data for use in metabolic modelling, we sampled different individuals and took sub-samples within each individual to investigate the source of potential variation. This is important information that will help guide future work. There was marked variation both between and within individual sponges in the abundance of macromolecules (Figure 3.2).

The skeleton was the most abundant component, and the variation in between-individual abundance was shown to have a significant effect on the overall population variation. The skeleton is the structural support for the cellular biomass and it guides the irregular canal network through which water is pumped to supply food and oxygen (Hentschel et al., 2012). The variability seen in the amount of skeleton is likely a result of *A. queenslandica*'s encrusting-lobate growth form (Hooper & Van Soest, 2006), combined with the irregular substrate and different structural requirements of specific internal features, such as canals and brood chambers.

It was previously mentioned that the high lipid content could be due to the reproductive state of the population. The reproductive state would also explain the variability in the lipid content within an individual. *A. queenslandica* forms localised brood chambers where embryos are matured before release (Leys et al., 2008) and not all of the biopsies would have contained brood chambers.

The variation in macromolecule composition may underlie the metabolic responses of the sponge to an unpredictable and competitive environment. Thus, at the scale of within an individual, differences in microhabitat such as food abundance, predatory pressures and conflicts for space can influence the macromolecular composition. This is potentially exacerbated by the nature of the biochemical

assays and the error associated with correcting for salt content. These results reveal significant effects of between-individual variation and a lesser effect of within-individual variation. This indicates that future work on the composition or metabolic processes of *A. queenslandica* should maintain a minimum of four representative individuals, supplemented with at least some subsampling within each individual. Care should still be taken with experimental design as this dataset is representative of a single metabolic state, and assumptions relating these results to different seasonal or ontogenetic stages should be made with caution. These results highlight the importance of biomass composition analysis, rather than the use of literature values.

3.5.3 Fatty Acid and Sterol Composition

The lipidome consists of a mixture of fatty acids and sterols, of which sponges are renowned for their diversity and often unique compounds (Erpenbeck & Soest, 2006; Lawson et al., 1984; Leys et al., 2008; Rodkina, 2005). Despite this diversity among FA structures in sponges, there are no clear or reliable taxonomic trends (Lawson et al., 1984; Rodríguez et al., 2010); this can in part be attributed to the microbial communities associated with sponges (Rodkina, 2005). I did not want to focus on identifying each compound in the lipid fraction as this is not necessary for modelling the sponge holobiont and performing flux balance analysis. As the precursor molecule for all lipid synthesis is acetyl coenzyme A, flux balance analysis may be performed if the contribution of lipids towards the overall biomass is known. I identified what compounds matched our databases as their inclusion will increase the models accuracy. Unidentified compounds were classified as FAs, sterols or neither based on their fragmentation pattern.

The two most abundant FAs in *A. queenslandica* were palmitic acid and octadecanoic acid. Both of these FAs are common in nature and have been reported in sponges previously (Barnathan et al., 1996; Carballeira & Alicea, 2000; Carballeira & Lopez, 1989; Christie et al., 1992; Nechev et al., 2004). Palmitic acid is a precursor with chain elongation, desaturation and addition of side chain modifications possible (Michal & Schomburg, 2013). Apart from pentadecanoic acid (15:0)

and heptadecanoic acid (17:0), the remaining identified FAs are long chain FA. FA with carbon chain length 20 (eocosanoic acid), 22 (docosanoic acid) and 24 (tetracosanoic acid), along with desaturated FAs of the same length (11-eicosanoic acid, erucic acid and nervonic acid, respectively) were identified as contributing to biomass formation.

All of these FAs have been reported in sponges previously (see references above). Eleven out of twenty-three FAs considered important for biomass production were unknown. Sponges, and their associated bacteria, have been shown to produce extensive modifications including methylation (Rodkina, 2005) and hydroxylation (Carballeira & Lopez, 1989) that would prevent the matching of novel compounds to commercial FA databases. This diversity and high number of unknown compounds highlights the complexity of the *A. queenslandica* lipidome and the potential for novel compound discovery.

The FAME analysis also identified the sterol fraction of the extracted lipids. The three sterols that fit my criteria of contributing to the overall biomass were brassicasterol, cholesterol, and an unknown sterol (UU13). Surprisingly, the most abundant of these was the brassicasterol, a plant steroid hormone (Michal & Schomburg, 2013) that is synthesized by terrestrial and marine plants, including macro and unicellular algae. Brassicasterol has been previously reported in a range of filter feeding and grazing marine invertebrates including sponges (Erdman & Thomson, 1972), gorgonian corals (Cnidaria; Anthozoa; Octocorallia; Gorgonacea) (Block, 1974) and sea urchins (Echinodermata; Echinozoa; Echinoidea) (Voogt, 1972). In these other classes of animals, brassicasterol is a minor sterol and its presence could be a result of dietary uptake. This is likely the case for the *A. queenslandica* holobiont, as neither animals nor bacteria have been shown to synthesize brassicasterol (Block, 1974; Erdman & Thomson, 1972; Voogt, 1972). Nevertheless, the high abundance implies an important biological function that this study is unable to discern. While brassicasterol is an unusual, dominant sterol for an animal, cholesterol is considered an important component of all cellular membranes in animals (Nelson & Cox, 2012). The concentration of cholesterol in a membrane performs two important functions. The overall fluidity of a membrane is regulated by the amount of cholesterol.

Increases in cholesterol content reduce the fluidity of the membrane and permability of neutral ions (Haines, 2001). Membrane bound cholesterol can also act as a promoter or inhibitor of membrane bound protein reactions (Yeagle, 1991). Additionally, cholesterol functions as an important substrate molecule for bioactive compounds and other sterols, including brassicasterol (Michal & Schomburg, 2013). Without identifying the second most abundant sterol, UU13, I was unable to infer its biological role, or synthesis pathway.

3.5.4 Amino Acid Composition

Analysing the amino acid composition identifies the relative amounts of precursor molecules that are required to synthesis the protein component of the biomass. When complemented with information from genome-scale modelling of the synthetic pathways that are present in an organism, these data allow for the identification of essential amino acids that cannot be produced by the organism (Thiele & Palsson, 2010). Subsequent flux analysis can identify bottlenecks where the high requirement of a compound for biomass synthesis slows overall growth. The protein component of *A. queenslandica* comprises of two distinct products of metabolism. The overall sponge biomass consists of soluble, cellular derived protein, and an insoluble, extracellular protein matrix. This extracellular matrix is dominated by a sponge specific collagen called spongin (Gross et al., 1956), which is the main component of the sponge skeleton.

I defined the skeleton of *A. queenslandica* as the inorganic silica spicules, together with the organic collagen matrix. Collagen is an important structural component of metazoans. I found that the spongin in *A. queenslandica* contained 34.9 % glycine (Table 3.6). This is close to the values reported for *Chondria reniformis* (30.5 %) (Imhoff & Garrone, 1983), *Scypha graminae* (31.5 % – 32.3 %) (Plez & Gross, 1958) and *Ircinia* sp. (32.3 %) (Junqua et al., 1974). Alanine (10.3 %) and proline (7.5 %) were the next two most abundant amino acids (Table 3.6) and are very close in abundance to the values of these other sponges. Although the amino acid composition of sponge collagens is well documented, the amino acid composition of the complete sponge biomass has not been reported.

The complete amino acid composition may give insights into the many primary cell culture attempts that so far have had limited success (Schippers et al., 2012). Many studies have utilised mammalian cell culture media, with and without serum (Schippers et al., 2012). A comparison of the amino acid composition of the complete *A. queenslandica* biomass (Table 3.4) with a hybridoma cell line (Quek et al., 2010) revealed nine amino acids that are at least twice as abundant in *A. queenslandica*. Proline and glycine stand out at 3.3 and 8.5 times more abundant in *A. queenslandica* than the hybridoma line. If we use DMEM F12 medium as an example of mammalian cell culture media, both proline and glycine are minor contributors to the amino acid fraction (Freshney, 2015) and could present significant bottlenecks for the production of biomass. In addition, due to the different amino acid bias in serum compared to sponge biomass, in the serum experiments excess waste amino acids could lead to growth limitations and may point to necessary changes in the approach to sponge cell culture.

I analysed the amino acid composition of the complete sponge biomass. From this I was able to estimate the average amino acid composition of the cellular fraction by subtracting the average composition of the skeleton. Values were then normalized for comparison by converting to a percentage. This showed that the cellular composition was distinct from the skeleton. The skeleton proportionally contained more glycine, alanine and proline than the cellular biomass. These three amino acids are key in giving collagen its triple helix shape and physical properties (Nelson & Cox, 2012). Amino acids that are essential in higher animals—in particular isoleucine, leucine, and valine (Nelson & Cox, 2012)—were proportionally higher in abundance in the cellular biomass than in the skeleton. Whether these are also essential amino acids for *A. queenslandica* is yet to be shown.

3.6 Conclusion

As a prerequisite for a systems scale investigation of the metabolic processes responsible for sponge biomass production, I analysed the composition and abundance of the macro-components of

A. queenslandica. I found the composition to be within the reported ranges for other tropical marine sponges and that there is significant variation in the composition at the population level and notable variation within the biomass of an individual sponge. The skeleton was the most abundant component of the overall biomass, while lipids and protein were the most abundant macromolecules. Further analysis of the fatty acid and amino acids was undertaken. The FAME analysis revealed a diverse and complex fatty acid composition. The composition of the sterols revealed that brassicasterol and an unknown sterol had a higher mean abundance than cholesterol. There was a distinct difference in amino acid composition between the soluble, cellular fraction and the insoluble skeletal component. This highlights the different protein products of metabolism that share the same pool of precursor molecules. This complete analysis of the biochemical composition of *A. queenslandica* provides baseline information that complements the extensive genomic resources available for *A. queenslandica*, a key species for understanding sponge biochemistry and metabolism.

Chapter 4 - Seasonal changes in environmental nutrient availability and biomass composition in a coral reef sponge

4.1 Abstract

Sponges are crucial ecosystem engineers in most marine habitats, playing a critical role in cycling elements such as carbon, nitrogen, sulfur and phosphorus between the water column and the sea floor. Despite this, it is unclear how the nutritional status of the seawater surrounding a sponge influences its biochemistry and physiology. Here we investigate the seasonal availability of these major nutrients in the water surrounding *Amphimedon queenslandica*, a coral reef demosponge inhabiting low-energy reef flat environments on the southern Great Barrier Reef, and ask how nutrient availability might influence the biomass composition of this sponge. Five replicated, seasonal water and sponge samples were collected over 13 months. Of the 11 environmental parameters that we assessed, all showed significant differences between at least two seasons. Some of the changes in nutrient availability were consistent with established variations in the ecology of Heron Island and its surrounding waters. For instance, the availability and sources of carbon and nitrogen changed throughout the year, with both being strongly influenced by nutrients emanating from an adjacent summer seabird rookery. Several environmental parameters were strongly correlated, such as total and dissolved organic carbon with each other and with nitrate and sulfate, respectively. Amongst biomass components, skeletal content was significantly correlated with low temperature, DNA with low total organic carbon, and the skeleton and lipid biomass with orthophosphate concentration, which was significantly higher in lower water temperatures. Based on these analyses there appears to be limited compelling correlations between biomass composition in *A. queenslandica* and the nutritional status of the surrounding seawater on a seasonal time scale.

4.2 Introduction

Sponges (poriferans) are crucial ecosystem engineers in most marine habitats largely because they are highly efficient biofilters, removing microbes and organic matter from huge volumes of water daily (Bell, 2008; de Goeij et al., 2008a,b; Pile et al., 1996b; Ribes et al., 1999a; Wulff, 2001). They contribute to maintaining water clarity and to cycling nutrients between the water column and the sea floor (Bayer et al., 2008; Hallam et al., 2006; Hoffmann et al., 2009; Liu et al., 2010). These capabilities appear to be particularly important on tropical coral reefs, which typically exist in clear, nutrient-poor seas and yet, paradoxically, are among the most productive ecosystems on Earth (Hatcher, 2003; Sammarco, et al., 1999). Specifically, marine sponges play a critical role in the efficient retention and recycling of nutrients through multiple trophic levels on coral reefs by making dissolved organic matter (DOM) – the most abundant resource produced on many coral reefs – available to higher trophic levels (de Goeij et al., 2013; Rix et al., 2016). DOM turnover by sponges can be close to the gross primary production of the entire reef ecosystem, giving sponges an ecological significance on coral reefs that is rivalled only by the reef-building corals themselves (de Goeij et al., 2013). Sponges procure nutrition from additional sources to varying and often undetermined extent that is contingent upon an array of biotic and abiotic factors, including the species' functional morphology, the type of reef ecosystem, and the diversity and function of associated microbial communities (Bell 2008; Pawlik 2011; Wulff 2012; Pawlik et al. 2015, 2016). For instance, Indo-Pacific and Caribbean coral reef ecosystem vary markedly and the relative abundance and ecology of sponges on these reefs appears to also differ (Wilkinson & Cheshire, 1990; Roff & Mumby, 2012; McMurray et al., 2015).

Despite the central role of sponges in shaping coral reefs and other marine ecosystems and the importance of particulate organic matter (POM) and DOM as being key sources of nutrition for sponges (e.g. Koopmans et al. 2010, 2011, 2015; Rix et al. 2016; de Goeij et al. 2013; Pawlik et al. 2016), there is little understanding about how other water nutrients influence the biochemistry and physiology and thereby the growth and reproduction of these animals. Seasonal changes of physicochemical properties of sea water impact on the growth, physiology and reproduction of other suspension feeding marine invertebrates (e.g. Bayne & Newell, 1983; Kang et al., 2011), suggesting

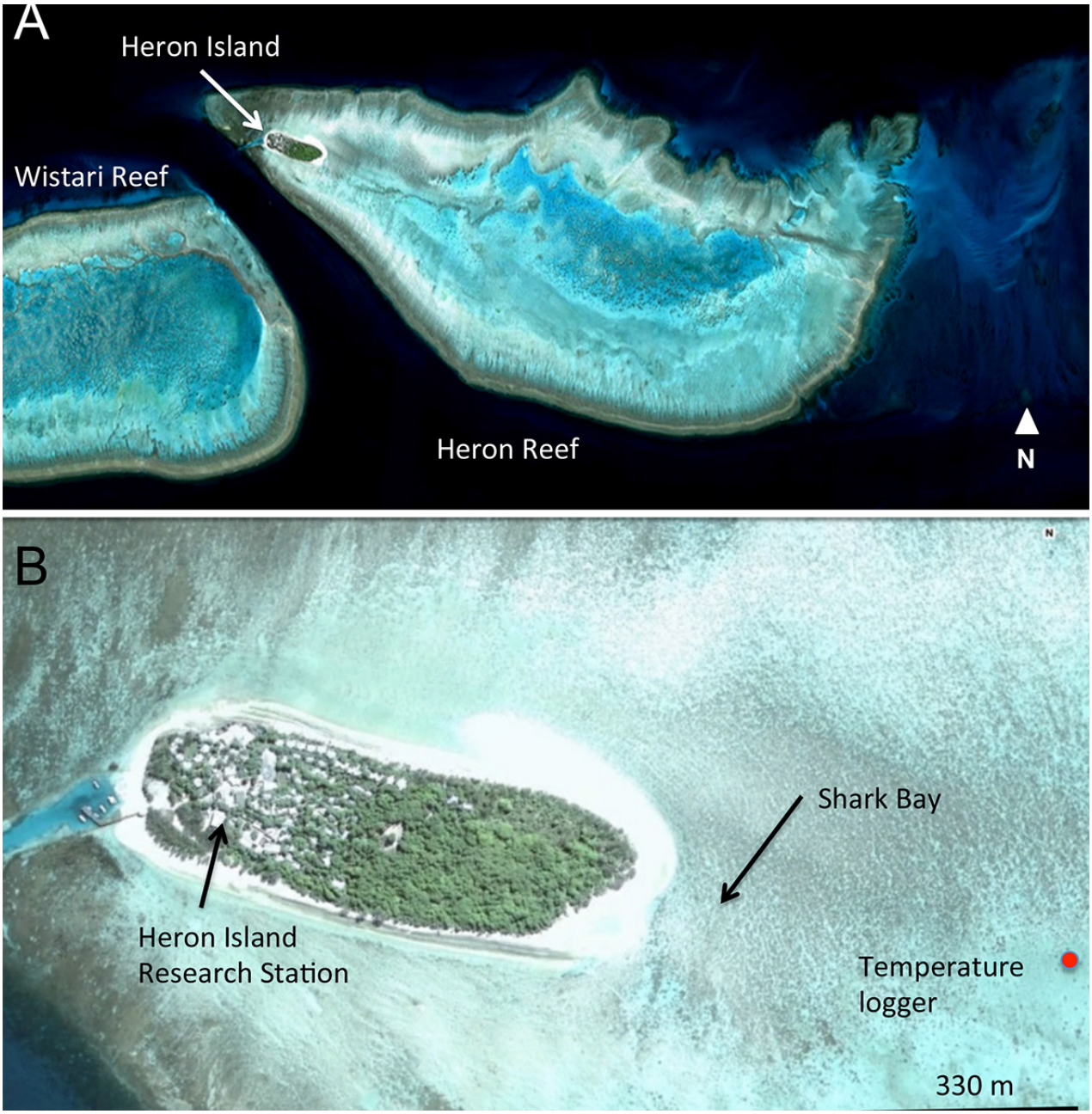


Figure 4.1 A satellite view of Heron Island Reef.

(A) Heron Island Reef and adjacent Wistari Reef. Heron Island is situated at the eastern, leeward end of the lagoon. (B) Heron Island and Shark Bay (-23.44, 151.92). *Amphimedon queenslandica* inhabits coral bommies and rubble in Shark Bay. The GBROOS water-monitoring site (Heron Island SF1) that provided the water temperature data is indicated with the red mark. Scale bar, 330 m.

that sponges will be influenced by changes in nutrient concentration and availability in the surrounding water column.

As a step towards understanding the interplay between sponges and their surrounding coral reef environment, we have characterised the seasonal changes in abundance of nutrients and the corresponding changes in the biomass composition of the demosponge *Amphimedon queenslandica* from Heron Island Reef on the southern Great Barrier Reef (Hooper & Van Soest, 2006). This Indo-Pacific tropical sponge inhabits the shallow, low energy habitats found in coral reef lagoons and reef flats (Degnan et al., 2008; 2015). On Heron Island Reef, it lives on the leeward side of the coral cay that supports a large seabird rookery during the summer breeding season (Figure 4.1). This results in a predictable increase in nurient run-off in the autumn (Smith & Johnson, 1995). This seasonal variation in nutrient loading in the waters that bathe *A. queenslandica*, along with continuous source of DOM from carbohydrate-rich coral mucus that is released on this platform reef daily, affords us with opportunity to determine if the biochemistry of this sponge is influenced by seasonal changes in the concentration of nutrients in the surrounding sea water. As the biomass of *A. queenslandica* has been previously characterised (Watson et al. 2014a,b) and this species has an extensive suite of genomic and transcriptomic resources (e.g. Srivastava et al. 2010; Levin et al. 2015), changes in the environment in the future can be correlated with gene expression, metabolic activity, reproductive output and growth. We find that many nutrients in the surrounding water column varied significantly with water temperature and thus season, and correlated with known changes in nutrient levels emanating from the adjacent coral cay (e.g. guano-derived nutrients), however few unexpectedly correlated with seasonal changes in the biomass composition of *A. queenslandica.*

4.3 Materials and Methods

Both the *Amphimedon queenslandica* biomass samples and the environmental seawater samples were collected from Shark Bay, Heron Island Reef, Australia (-23.44, 151.92; Figure 4.1) in September (early spring) and December (summer) of 2013, and in March (early autumn), July (winter) and October (mid-spring) of 2014. Samples were procured around full moon spring tides.

4.3.1 Analysis of seawater

Water temperature data were sourced from the Great Barrier Reef Ocean Observing System (GBROOS) Project (http://data.aims.gov.au/gbroos) monitoring system maintained by the Australian Institute for Marine Science (AIMS). The sampling station is situated approximately 500 m from the low tide mark at Shark Bay at 0.5 m depth (Figure 4.1B). For this analysis, temperature data was extracted encompassing the four days of water sampling as described below. The first temperature time point began at 00:05 the morning that the first water sample was taken and ended at midnight of the day that the final water sample was taken. The temperature was sampled every five minutes, resulting with each month having 1152 temperature data points.

Eight seawater samples were procured in each of the months (seasons) that were analysed. Four samples were taken two days before the full moon, and four on the day of the full moon. On each of these days, samples were collected at semi-diurnal high and low tides, resulting in four seawater samples per 24-hour period. Seawater samples were collected by opening a submerged 2.0 L acid-washed and sterilised Schott bottle directly adjacent to an *A. queenslandica* individual. This process was repeated three times, with the first two collections discarded to flush the bottle. The Schott bottle was then immediately sealed and transported on ice to Heron Island Research Station, approximately 700 m from the collection site (Figure 4.1B), and treated according to the preparation requirements of the different analysis as detailed below. Once prepared, the samples were frozen at -80°C and transported on dry ice to Brisbane for analysis. All water analyses were conducted by Symbio Alliance,

an independent National Association of Testing Authorities Australia (NATA) accredited laboratory. The methods used to analyse each environmental parameter were derived from the Standard Methods for the Examination of Water and Wastewater, American Public Health Association (APHA) (Rice et al., 2012). For detailed methods of each analysis, refer to the APHA reference number contained in brackets after each parameter.

Ammonia (45000-NH3-G), nitrite (4500-NO2-B), nitrate (4500-NO3-F) and orthophosphate (4500-P-F) were quantified colourimetrically using water samples that had been 0.22 µm filtered prior to freezing at -80°C. Ammonia was quantified with a reaction of alkaline phenol and hypochlorite to produce indophenol blue (Rice et al., 2012), with the intensity of the blue signal being proportional to the amount of ammonia in the sample. Nitrite was quantified by coupling diazotized sulfanilamide with *N*-(1-naphthyl)-ethylene-diamine-dihydrochloride at pH 2-2.5 (Rice et al., 2012) to produce a red-purple azo dye that was quantified by absorbance against a standard curve. Nitrate was quantified by first reducing to nitrite by the addition of cadmium; this total amount of nitrite was then quantified as described above. The amount of nitrate was then calculated by subtracting the amount of nitrite measured previously, without the reduction step (Rice et al., 2012). Orthophosphate was quantified by reacting antimony potassium tartrate with ammonium molybdate and then adding ascorbic acid to produce a blue coloured complex that was proportional to the amount of orthophosphate present (Rice et al., 2012). The detection limit of ammonia, nitrite and nitrate and orthophosphate was 0.005 mg / L.

Total phosphorous (4500-P-I) and total nitrogen (4500-N-B) were quantified from unfiltered seawater samples. To determine total nitrogen, the sample was treated with heated alkaline persulfate and ultraviolet radiation to digest and oxidize all nitrogenous compounds to nitrate. The nitrate was then reduced to nitrite with the addition of cadmium. Under acidic conditions, a diazonium ion was formed from nitrite, diazotization and sulfanilamide. This diazonium ion was subsequently coupled to *N*-(1-naphthyl)-ethylene-diamine-dihydrochloride to produce a pink dye that absorbs at 540 nm (Rice et al., 2012), with a detection limit of 0.05 mg / L. The quantification of total phosphorous

also required an initial digestion with heated persulfate and ultraviolet radiation to convert organic phosphates to orthophosphate (Rice et al., 2012). The amount of orthophosphate was then quantified as described above, but this time with a detection limit of 0.01 mg / L. This is proportional to the total amount of phosphorous.

Silica (silicon dioxide; 3120-B) was quantified using Inductively Coupled Plasma (ICP) optical emission. The ICP is maintained at temperature of 6000 to 8000° K. Injection of the seawater sample as an aerosol caused atomic separation and excitation. Measurements were taken at specific wavelengths and quantified against a known concentration. Silica was measured at 212.41 nm and has a detection limit of 0.01 mg / L.

Total organic carbon (TOC)(5310-B) was quantified by combustion of the sample and infrared detection of formed CO_2. Specifically, the sample was injected into a heated chamber where the seawater vaporized and the carbon was converted into CO_2, the amount of which was then measured with an infrared detector. Dissolved organic carbon (DOC)(5310-B) was quantified using the same procedure, except that the sample was first pre-filtered (0.45 μm) to remove particulates.

Sulfide (4500-S-F) was measured by titrating 0.025 *N* sodium thiosulfate ($Na_2S_2O_3$) against an iodine solution, containing 6 *N* hydrochloric acid and the sample (Rice et al., 2012). To quantify sulfate (4500-SO4-E), it was first precipitated by the addition of an acetic acid solution containing barium chloride ($BaCl_2$). Absorbance of the barium sulfate crystals in suspension was then calculated using a standard curve (Rice et al., 2012).

4.3.2 Sponge biomass composition

At each sampling time point, four adult sponges were collected from the reef using established protocols (Leys et al., 2008) and maintained for < 12 hours in the aquarium system at Heron Island Research Station with flow-through ambient sea water from the same reef flat from where the sponges were collected before being processed. The sponges collected were large enough to allow all replicate pieces of all macromolecule types to be sampled from each individual. The size limitation of large *A. queenslandica* individuals and the need for biopsies free of substrate and foreign material meant that, at most, 18 replicate pieces could be biopsied from a single large sponge. This allowed 3 replicate pieces from each sponge to be allocated for quantitative analysis of each macromolecule type as previously described (Watson et al., 2014a,b). The macromolecules measured were DNA, RNA, protein, lipid, carbohydrate and skeleton.

4.3.3 Statistical analyses

All statistical analysis were performed using R version 3.1.1 within RStudio (Team, 2015). One-way ANOVAs were used to test for significant differences in environmental parameters or biomass composition between the months sampled. A post-hoc Tukey HSD tests were then used to identify months between which significant differences existed. The effect of tide (high or low) or time of day (day or night) on the concentrations of available nutrients was tested with a nested ANOVA. Either tide or time was nested as an interaction with the main effect being month. A Tukey HSD post-hoc test was performed to investigate individual comparisons between months. The *ggplot2* (Ginestet, 2011) package was used to create the box plot figures.

The correlation function in R was used to apply a Pearson correlation test to investigate potential correlations within and between the biomass components and the environmental parameters. Results were visualised using the *corrplot* package, and probability values were estimated with the corr.test function. Due to the uneven sample sizes – 12 biomass replicates and 8 environmental replicates

– statistical analysis of correlations between these were made by randomly removing one biomass sample from each sponge.

4.4 Results

4.4.1 Seasonal changes in environmental parameters

Heron Island Reef is located in the Southern Hemisphere where summer begins in December. Analyses were conducted over multiple seasons during a 13 month period from 2013 to 2014, to capture changes in the environment throughout a year. All data were taken as representative of the season in which the samples were collected and are presented independent of year; the four seasons of the year are represented as follows –summer (December), early autumn (March), winter (July), early spring (September) and mid-spring (October). Complete water analysis results are presented in Appendix 4.1.

In all months encompassing all four seasons, water temperature increased through the day and decreased at night (Figure 4.2). As all samples were collected during spring tides associated with a full moon, each sampling period covered a similar semi-diurnal tidal pattern, with early morning and afternoon low tides, and late morning and evening high tides. Heron Island Reef is a platform reef on which lagoonal water pools above the surrounding sea level during the low tides, which contributed to the daily heating experienced within the lagoon, particularly during the day. Sudden drops in temperature were experienced in early spring (September) and early autumn (March), when the cooler water of the incoming tides coincided with nightfall (Figure 4.2). The winter samples (July) showed a more irregular temperature pattern than the other seasons (Figure 4.2). A one-way ANOVA ($F(4, 5755) = 4713$, $p < 2e\text{-}16$) showed that all months were significantly different from each other (Tukey post-hoc HSD test $p < 0.01$ for all pairwise comparisons).

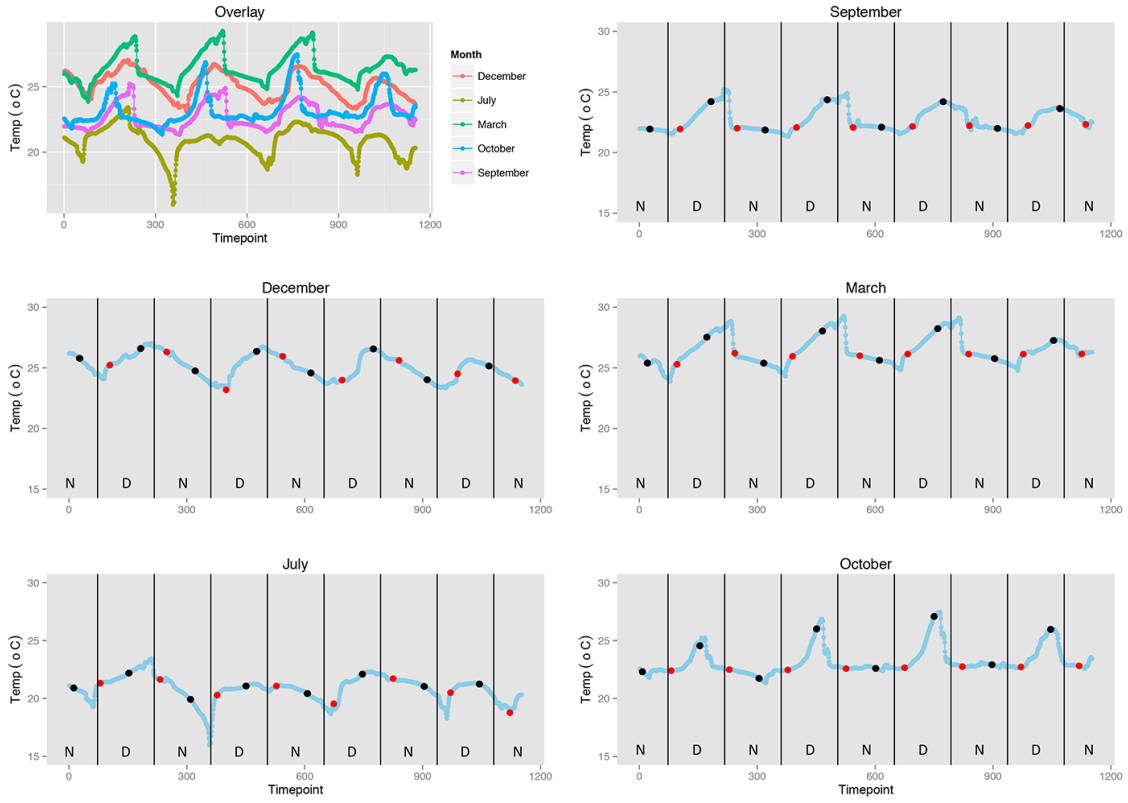


Figure 4.2 Daily fluctuations in water temperature on Heron Island Reef from September 2013 to October 2014.

The top left panel shows the five seasonal samples (September, early spring; December, summer; March, early autumn; July, winter; October, mid-spring), shown in the other five panels, overlayed on each other. The time points highlighted with black and red dots indicate the low and high tidal peaks, respectively. Night and day are indicated with N and D. The black vertical bars are spaced 12 hours apart, indicating 6 am and 6 pm.

Overall, carbon and sulfates were consistently the most abundant nutrients in the water column (Table 4.1, Fig 4.3). Total organic carbon (TOC) levels were highest in March (early autumn) and July (winter), and were significantly different between seasons (one-way ANOVA, $F(4,35) = 6.122$, $p = 7.6e\text{-}4$). Post-hoc Tukey HSD tests identified the significant pairwise comparisons as Dec – July ($p = 0.02$), Mar – Oct ($p = 0.019$), Sept – July ($p = 0.026$), and Oct – July ($p = 0.002$). Dissolved organic carbon (DOC) abundance followed the same trend as the amount of sulfate in the water, with both highest in July (winter) and October (mid-spring) (Table 4.1, Fig. 3). For DOC, the one-way ANOVA, $F(4, 35) = 4.034$, $p = 0.008$ and subsequent post-hoc Tukey HSD identified the following significant

pairwise comparisons: July – Dec ($p = 0.01$); July – Mar ($p = 0.04$); and July – Sept ($p = 0.01$). Sulfate differed significantly (one-way ANOVA, $F(4, 35) = 900.1$, $p = <2e\text{-}16$) between all months (post-hoc Tukey HSD tests, $p < 0.01$) except between September and December. Sulfide was below detection at all time points sampled (data not shown).

Table 4.1 Summary of significant differences in water parameters between seasons.

Parameter	one-way ANOVA	Seasons identified as being significantly different from each other
Total organic carbon	F(4,35) = 6.122, p = 7.6e-4	Dec – July (p = 0.02); Mar – Oct (p = 0.019); Sept – July (p = 0.026); Oct – July (p = 0.002)
Dissolved organic carbon	F(4, 35) = 4.034, p = 0.008	July – Dec (p = 0.01); July – Mar (p = 0.04); July – Sept (p = 0.01)
Sulfate	F(4, 35) = 900.1, p = <2e-16	All pairwise comparison significant (p < 0.01), except for Sept-Dec
Total nitrogen	F(4, 35) = 18.68, p = 2.6e-8	Mar – Dec (p = 2.0e-7), Mar – July (p = 1.0e-7), Mar – Oct (p = 7.5e-6)
Ammonia	F(4, 35) = 3.406, p = 0.01	Mar – Sept (p = 0.047), Mar – July (p = 0.047), Mar – Oct (p = 0.047)
Nitrite	F(4, 35) = 173.6, p < 2e-16	For all pairwise comparisons (p < 0.01)
Nitrate	F(4, 35) = 26.46, p = 3.71e-10	Mar – Dec (p = 5.4e-6), Mar – Sept (p = 6.5e-6), Mar – Oct (p = 3.2e-5), July – Dec (p =1.0e-7), July – Sept (p = 1.0e-7), July – Oct (p = 5.0e-7)
Total phosphate	F(4, 35) = 28.81, p = 1.2e-10	Mar – Dec (p = 2.2e-6), Mar – July (p = 0), Mar –Sept (p = 0), Oct – July (p = 5.5e-5), Oct – Sept (p = 1.5e-4)
Orthophosphate	F(4, 35) = 15.36, p = 2.4e-7	July – Dec (p = 0.02), July – Oct (p = 0.008), Sept – Dec (p = 1.7e-5), Sept – Mar (p = 3.3e-6), Sept – Oct (p = 5.3e-6)

The nitrogenous compounds were consistently the next most abundant nutrients (Table 4.1, Fig 4.3); specifically, we measured total nitrogen, ammonia, nitrite and nitrate, all key components of the nitrogen cycle. The amount of total nitrogen differed significantly between months (one-way ANOVA, $F(4, 35) = 18.68$, $p = 2.6e\text{-}8$), with post-hoc Tukey HSD tests identifying the significant comparisons as Mar – Dec ($p = 2.0e\text{-}7$), Mar – July ($p = 1.0e\text{-}7$), and Mar – Oct ($p = 7.5e\text{-}6$). For ammonia, only March contained detectable levels of ammonia in all samples, with an average concentration of 0.012 +/- 0.002 mg/L; the only other time that ammonia was detected was in one of the eight samples from December (0.038 mg/L). This resulted in March values being significantly different (one-way ANOVA, $F(4, 35) = 3.406$, $p = 0.01$) from September (Tukey HSD, $p = 0.047$), October ($p = 0.047$) and July ($p = 0.047$). Nitrite was only detectable during July, with an average concentration of 0.019 +/- 0.003mg/L, which resulted in July being significantly different from all other months sampled (one-way ANOVA, $F(4, 35) = 173.6$, $p < 2e\text{-}16$ and Tukey $p < 0.01$ for all July comparisons). Nitrate was detected in more samples and in higher concentrations than either ammonia or nitrite (Fig. 4.3), peaking in abundance in March and July. There were significant differences in nitrate between months (one-way ANOVA, $F(4, 35) = 26.46$, $p = 3.71e\text{-}10$), with post-hoc Tukey HSD identifying differences between Mar – Dec ($p = 5.4e\text{-}6$), Mar – Sept ($p = 6.5e\text{-}6$), Mar – Oct ($p = 3.2e\text{-}5$), July – Dec ($p = 1.0e\text{-}7$), July – Sept ($p = 1.0e\text{-}7$), and July – Oct ($p = 5.0e\text{-}7$). The values for nitrate are higher than the values for total nitrogen in the month of July, which is an apparent anomaly that is difficult to explain.

We quantified both the total amount of phosphorous and the abundance of orthophosphate, which is the biologically-available fraction of phosphate. Overall, the seasonal patterns of abundance differ between the two forms (Figure 4.3). Total phosphorous concentration differs significantly between most months (one-way ANOVA, $F(4, 35) = 28.81$, $p = 1.2e\text{-}10$), with post-hoc Tukey HSD tests identifying significant differences between Mar – Dec ($p = 2.2e\text{-}6$), Mar – July ($p = 0$), Mar – Sept ($p = 0$), Oct – July ($p = 5.5e\text{-}5$), and Oct – Sept ($p = 1.5e\text{-}4$). Orthophosphate concentration also differs between some months (one-way ANOVA, $F(4, 35) = 15.36$, $p = 2.4e\text{-}7$), with post –hoc Tukey HSD tests identifying the significant comparisons as July – Dec ($p = 0.02$), July – Oct ($p = 0.008$), Sept – Dec ($p = 1.7e\text{-}5$), Sept – Mar ($p = 3.3e\text{-}6$), and Sept – Oct ($p = 5.3e\text{-}6$).

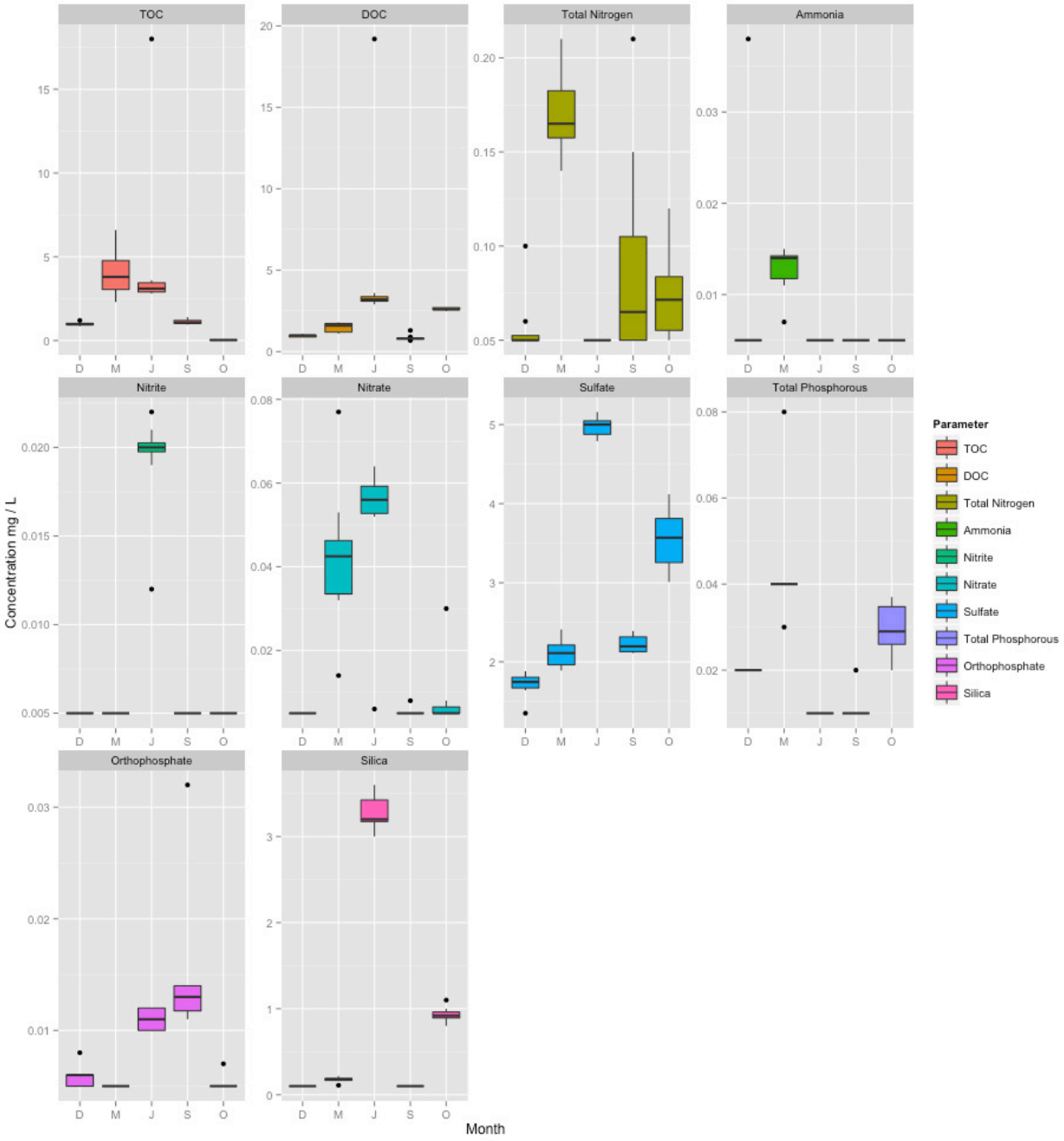


Figure 4.3 Changes in the availability of major nutrients in Shark Bay, Heron Island Reef from September 2013 to October 2014.

Eight replicate water samples are analysed per seasonal sample, at two high and two low tides during the day and night, over two non-successive days. The y-axis scales on each parameter are independent. The black dots represent outlier values that are indicative of the variability that can occur within small time frames in this natural system. The months sampled were December, 2013 (D; summer), March, 2014 (M; early spring), July, 2014 (J; winter), September, 2013 (S; early spring) and October, 2014 (O; mid-spring). DOC, dissolved organic carbon; TOC, total organic carbon.

Silica was in low abundance through most of the year, with the exception of a substantial peak in abundance in July and a smaller peak in October (Figure 4.3), a seasonal pattern similar to DOC and sulfate. As a result, these months were significantly different from all others (one-way ANOVA, $F(4, 35) = 1449$, $p < 2e\text{-}16$, and post-hoc Tukey HSD tests $p < 0.01$).

4.4.2 Effects of time and tide on water parameters

To test whether nutrient availability changes through the day–night cycle (Appendix 4.2), we used a nested ANOVA with time of sampling (day or night) nested as an interaction within the effect month. We found no overall significant effect for any of the water parameters analysed. Further, Tukey post-hoc tests of individual day–night comparisons within months identified no statistically significant comparisons. We performed a separate nested ANOVA with tidal height (high or low) as the nested interaction within month (Appendix 4.3), and again found no statistically significant effect on any of the nutrients in any month.

4.4.3 Changes in sponge biomass composition

Of the macromolecular components that we measured, total protein and RNA were relatively constant across all seasons, while all others were more variable (Figure 4.4; Appendix 4.4). Skeleton levels were highest in the winter (July) and early spring (September), decreasing in the warmer months (Figure 4.4) with significant differences between some months (one-way ANOVA, $F(4, 39) = 5.972$, $p = 0.0005$). Post-hoc Tukey tests identified the significantly differences between Sept – Dec ($p = 0.018$), Sept – Mar ($p = 0.014$), Sept – Oct ($p = 0.002$), and Oct – July ($p = 0.02$). The lipid content followed a similar pattern to the skeletal component, with higher relative abundances in the cooler months (Fig. 4.4), with significant differences (one-way ANOVA, $F(4, 54) = 6.507$, $p = 0.0002$) between Dec – July (Tukey post-hoc test, $p = 0.003$), Dec – Sept ($p = 0.02$), Oct – July ($p = 0.001$), and Oct – Sept ($p = 0.01$). Carbohydrate levels were highest in December (summer) and March (early autumn) (Fig. 4.4; one-way ANOVA, $F(4, 54) = 4.228$, $p = 0.004$) and post-hoc Tukey HSD identified

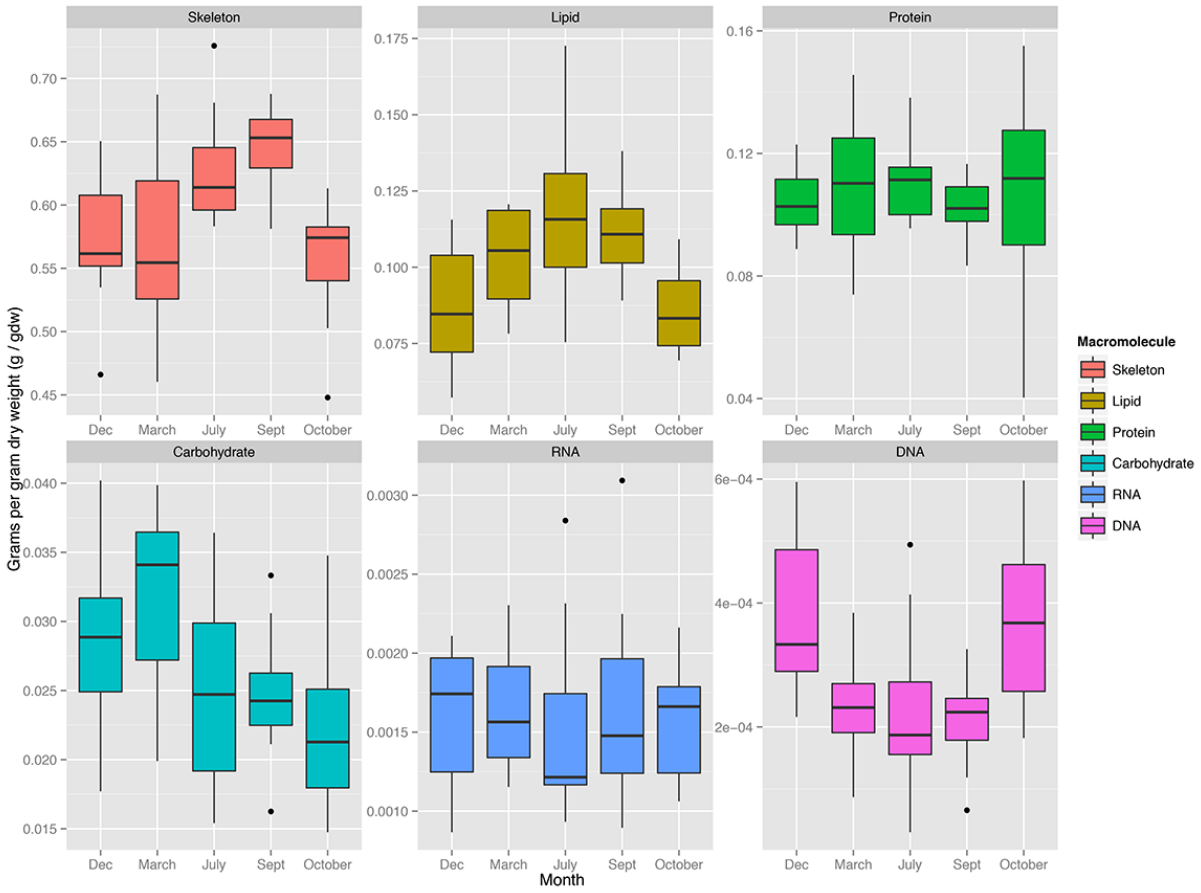


Figure 4.4 Changes in the macromolecular composition of *Amphimedon queenslandica* in Shark Bay, Heron Island Reef from September 2013 to October 2014.

There were 12 biomass samples analysed for each component during each season sample. All values are presented in grams per gram of sponge dry weight (Watson et al., 2014). The y-axis scales are independent due to the large differences in abundance of the different macromolecular components. The dots represent outlier samples that are greater than 1.5 times the interquartile range.

March as significantly different from both spring months of October ($p = 0.005$) and September ($p = 0.04$). The total amount of DNA in the sponges varied significantly (one-way ANOVA, $F(4, 53) = 6.254$, $p = 0.0003$), and was highest in warmest months, December and October, with significant differences observed in Dec – July (Tukey HSD post-hoc test, $p = 0.02$), Dec – Mar ($p = 0.02$), Dec – Sept ($p = 0.008$), Oct – July ($p = 0.02$), Oct – Mar ($p = 0.03$), and Oct – Sept ($p = 0.01$) comparisons.

4.4.4 Correlations within and between biomass components and environmental parameters

We identified all correlations in abundance among the environmental parameters and biomass components. All biomass components, except protein, and all environmental parameters, except ammonia, were significantly ($p < 0.01$) correlated, either positively or negatively, with at least one other component or parameter (Figure 4.5; Appendix 4.5). Temperature correlated with the most factors (1 biomass component and 7 environmental parameters) followed by the nutrients sulfate (6 environmental parameters), silica (5 environmental parameters) nitrite (4 environmental parameters), phosphorous (4 environmental parameters) and orthophosphate (2 biomass components and 2 environmental parameters) (Figure 4.5; Appendix 4.4). Several environmental parameters were significantly correlated ($p < 0.01$, n = 8) with each other, either positively (notably sulfate – silica, sulfate – DOC, TOC – DOC, TOC – nitrate, and total nitrogen – total phosphorous) or negatively (notably temperature – sulfate, temperature – silica, and temperature – orthophosphate). Significant correlations ($p < 0.01$, n = 8) among biomass components were restricted to a positive correlation between carbohydrate – RNA, and negative correlations between DNA – RNA, and DNA – skeleton. Of most interest to us were the significant correlations between biomass components and environmental parameters. Of these, positive correlations were found between total nitrogen – carbohydrate, between orthophosphate – skeleton, and between orthophosphate – lipid. Negative correlations were found between temperature – skeleton, and between TOC – DNA.

The correlations supported a number of shared trends that we observed when plotting seasonal variations, including TOC and DOC with nitrate and sulfate, respectively (Figures 4.3 and 4.5). From a biological perspective, although both skeleton and silica were significantly correlated with low temperature, and the *A. queenslandica* skeleton is partially comprised of siliceous spicules, we found no significant correlation of skeleton with silica (Fig. 4.5; Appendix 4.5). Orthophosphate was significantly correlated with skeletal and lipid components and low temperature, and we did find a positive correlation of skeleton and lipid, but only at $0.01 < p < 0.05$ (Figure 4.5; Appendix 4.5).

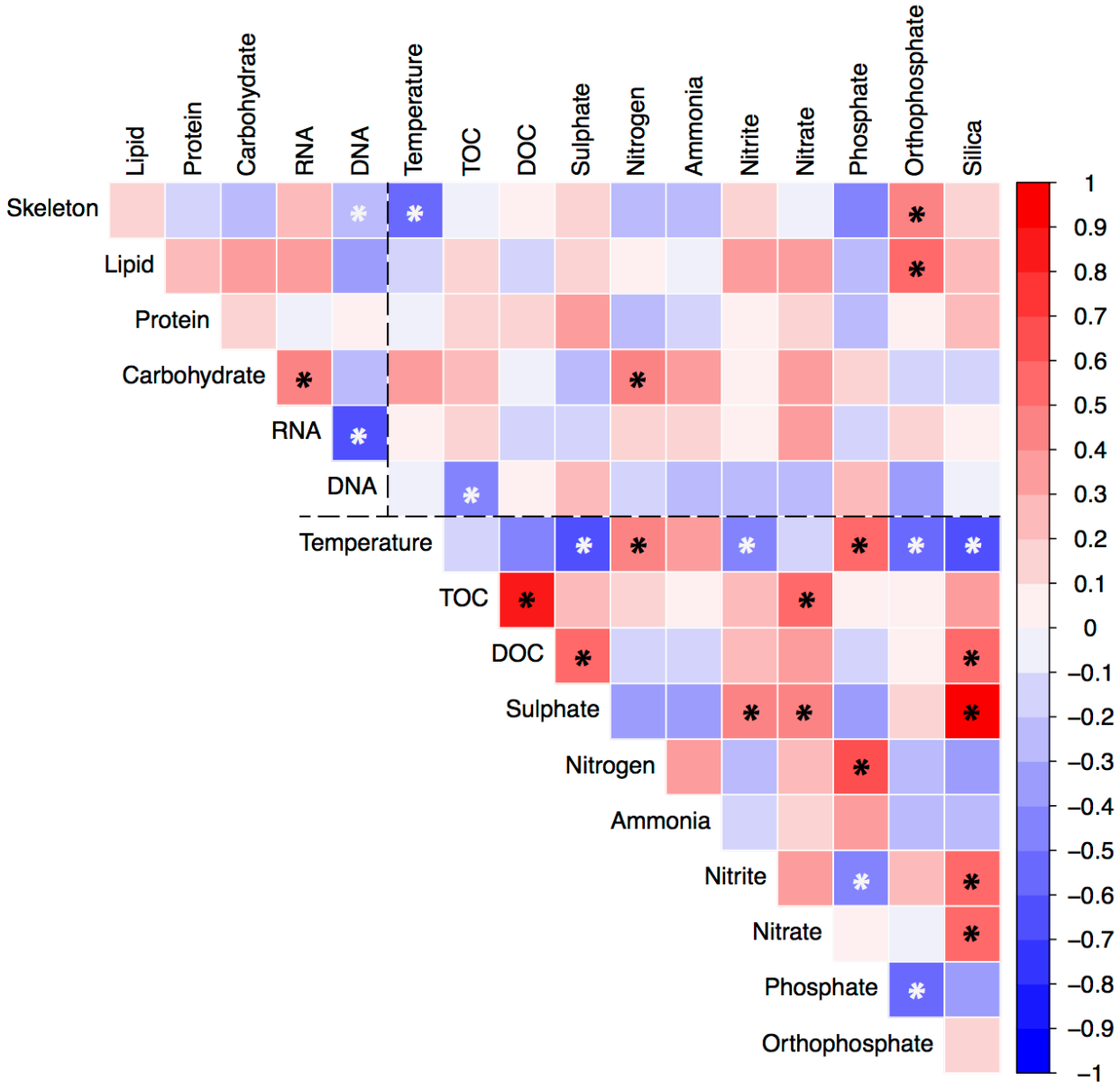


Figure 4.5 A correlation matrix summarising all pairwise comparisons between biomass components and environmental parameters.

The intensity of the colour indicates the strength of the correlation, with red being positive and blue being negative. For biomass vs biomass component, n = 12. For water parameter vs water parameter and water parameter vs biomass, n = 8. Positive and negative correlations with greater than 99% confidence are denoted by black and white asterisks, respectively. Nitrogen is an abbreviation for total nitrogen; phosphorous is an abbreviation for total phosphorous. Dashed line separates biomass components from environmental parameters.

4.5 Discussion

In this study, we characterised the seasonal availability of nutrients in the water column in a natural habitat of the demosponge *Amphimedon queenslandica*, while concurrently monitoring the biomass composition of the sponges themselves to determine if the former directly impacts on the latter. We focussed our sampling on a population of sponges residing in the shallow lagoonal habitat of Shark Bay adjacent to the coral cay of Heron Island on the southern Great Barrier Reef, from September 2013 to October 2014. *A. queenslandica* appears to be restricted to this relatively small area of Heron Island Reef (Degnan et al., 2015); personal observations) that is on the leeward side of the lagoon and is characterised as a low-energy environment (Frith, 1983). Generally, the low energy environments on coral reefs are dominated by heterotrophic processes that include the deposition and degradation of organic material (Birkeland, 1997b). As such, we intentionally did not filter larger particular matter from the water samples, as this would remove strands of nutrient-rich coral mucus and associated particulate matter, which are observed often adhering the the surface of *A. queenslandica* (unpublished observations) and thus potentially providing a source of additional nutrients to heterotrophic bacteria in the vicinity of the sponge.

Organic carbon is most abundant waterborne nutrient available to *A. queenslandica*. By comparing total organic carbon (TOC) and dissolved organic carbon (DOC), it is possible to estimate the amount of particulate carbon in the water column. Given it is likely that sponge cells primarily phagocytise the particulate matter within a given size range, typically between 0.1 -70 μm (e.g. Pile et al., 1996a; Ribes et al., 1999b; Turon et al., 1997), and bacterial cells assimilate the dissolved nutrients (Weisz et al., 2007b; Yahel et al., 2003a), distinguishing particulate and dissolved carbon sources is important to determining metabolic inputs into the holobiont; as stated above, larger (>70 μm) particulate matter was included in this study as this may be an indirect source of nutrients. The putative functional partitioning of nutrient uptake and assimilation is likely to underlie the capacity of sponge holobiont to make dissolved organic matter (DOM) – the most abundant resource produced on a coral reef – available to higher trophic levels (de Goeij et al., 2013; Rix et al., 2016). Our results show that the dissolved fraction is more abundant than the particulate fraction in the water column of

the *A. queenslandica* habitat, with DOC accounting for 75% to 95% of the TOC for all months sampled except in early autumn (March). Coral mucus is a dominant contributor to DOC on coral reefs and transporter of particulates (van Duyl & Jan Gast, 2001; Wild et al., 2010, 2004b; Yahel et al., 2003a) and included in our analyses. This constantly secreted mucus is primarily comprised of carbohydrates that are synthesized by the coral's symbiotic zooxanthellae, with up to 50% of the fixed carbon being released as mucus (de Goeij et al., 2013; Wild et al., 2004b). Although the specific roles of microbial and sponge cells in the *A. queenslandica* holobiont in nutrient assimilation were not determined in this study, the dominant bacterial symbiont of *A. queenslandica* is a gamma proteobacterium, the genome of which is enriched in carbohydrate transporter genes and utilisation enzymes (Gauthier at al., 2016). Together, these observations suggest that, for most of the year, the predominant carbon source for the *A. queenslandica* holobiont is DOC. Although DOC contains energy-rich carbohydrates, it has a low nitrogen content, and may be of low nutritional value compared to the particulate fraction (Wild et al., 2005). During the early autumn month of March, the average amount of carbon in the dissolved fraction was reduced to 41%. At this time, the TOC is at its peak, while the DOC is low, suggesting that there is an increase in particulate organic matter at this time. This may include phytoplankton and bacterial cells that can be directly consumed by sponge cells.

During the period of the study, TOC was found to be more abundant than nitrogen, generally by an order of magnitude. Seasonal changes in total nitrogen concentration tend to follow TOC, as anticipated by the Redfield ratio that describes the ratio of C:N:P found in the particulate fraction and dissolved pools of marine environments (Redfield, 1934). The Redfield ratio of C:N is 6.625. This value is below the average TOC:Total nitrogen ratio for each season in Shark Bay on Heron Island Reef (September 13.9, December 18.0, March 23.5, July 99.5 and October 49.6), identifying this area as a low nutrient environment that is limited in nitrogen (Sterner et al., 2008). These low C:N ratios include both particulate and dissolved fractions of carbon, with the dissolved fraction skewing the ratio. The higher the Redfield ratio is, the greater the relative contribution of DOC to TOC. Nitrogen is essential for the synthesis of amino acids, which in *A. queenslandica* contribute 73% of the total sponge dry weight (Watson et al., 2014a). This large requirement for, but low abundance of, nitrogen in the environment may be a driving factor for the apparent fast pumping rates in *A. queenslandica*,

especially since their most abundant symbiotic bacteria appear unable to fix nitrogen (Gauthier et al, 2016). This is congruent with the low microbial abundance (LMA) status of this sponge (Feith et al., 2016) that implies *A. queenslandica* needs to process large volumes of water to meet its nitrogen requirements from particulate matter (Hentschel et al., 2003; 2006; Weisz et al., 2007b).

Dissolved inorganic nitrogenous compounds were found in low concentrations or, in some months, were undetectable by the methods used. Ammonia, nitrite and nitrate all are important compounds in the nitrogen cycle. This cycle has been studied in detail on Heron Island and the surrounding reef because this coral cay supports a large seabird rookery of up to 80,000 breeding pairs of white-capped noddys (*Anous minutus*) and wedgetail shearwaters (*Puffinus pacificus*) during the summer breeding season (Smith & Johnson, 1995). The amount of guano deposited annually is estimated to be 130 tonnes, which includes 9.4 tonnes of nitrogen (Smith & Johnson, 1995). As a result, the groundwater is very high in nitrate (up to 530 μmol), but low in ammonia, and tidal-driven flushing of the island's permeable sediments mixes seawater with this island groundwater (Santos et al., 2010). The nitrogen-enriched water flows out to the reef as the tides recedes, accumulating in the pooled lagoon, including Shark Bay (Santos et al., 2010). In our data, total nitrogen peaked during the March sampling period that is at the end of the seabird breeding season; this same sampling period was the only time in which we detected ammonia and nitrate, while nitrite remained below detection throughout the entire year. The terrestrially-sourced nitrogen is primarily in the form of nitrate, which neither *A. queenslandica* nor its symbionts appear to be able to directly utilise. However, the increased organic load in the low energy environment of Shark Bay over summer stimulates the conversion of nitrate to ammonia (Erler et al., 2014); nitrate that enters the sediments is reduced, with ammonia being the end product. It has been experimentally shown in Shark Bay that the flux through the dissimilatory nitrate reduction to ammonia (DNRA) pathway is enhanced by the addition of organic material to the sediment (Erler et al., 2014; Eyre et al., 2008). The *A. queenslandica* holobiont possesses the enzymes to utilise ammonium but not nitrogen (Gauthier et al., 2016).

Sulfate and DOC showed a positive correlation that could be linked through metabolic processes occurring in the sediments. Anoxic conditions have been identified as shallow as 5 mm into the sediments of Shark Bay, and DOC that enters this zone stimulates the production of reduced sulfur compounds such as sulfide (Wild et al., 2004b). Rapid oxidation of sulfide to sulfate would subsequently occur when the sulfide diffuses into the water column and reacts with oxygen. The dominant gamma proteobacterial symbiont of *A. queenslandica* can fix carbon from carbon dioxide, using energy gained from the oxidation of reduced sulfur compounds (Gauthier et al., 2016). Sulfide was one of the parameters that we analysed in the present study, but it was below the detection limit for all samples in all months. It is possible that this was due to oxygen exposure during the processing of the water samples themselves, which would provide opportunity for the oxidisation of the sulfide to sulfate. The acquisition of further data on reduced sulfur compounds in the Shark Bay environment is something that we would like to follow up on in future studies, not least because the capacity to oxidise reduced sulfur compounds is a defining characteristic of the dominant *A. queenslandica* bacterial symbiont (*AqS1*; Gauthier et al., 2016). Developing a greater understanding of the availability of reduced sulfur compounds in the sponge habitat will help us to determine whether in fact these are essential for the sponge-bacterial holobiont.

Our concurrent measuring of the sponge biomass composition at the macromolecular level allowed us to identify correlations between sponge biomass and the environmental parameters that were quantified. Although the relative amounts of several biomass components changed over the four seasons of the year, we identified only a few significant positive correlations between a biomass component and an environmental nutrient. These include both skeletal and lipid components being positively correlated with orthophosphate concentration, and intriguingly carbohydrate being positively correlated with total nitrogen. DNA, which was highest in the warmer months, was negatively correlated with TOC, which was highest in the autumn and winter. Most marine animals reproduce seasonally and thus seasonally partition their available resources between growth and reproduction. Data on the growth rates of *A. queenslandica* are not available, but the reproductive cycle is well characterised (Degnan et al., 2008; 2015). *A. queenslandica* is reproductive throughout the year, but there is a substantial increase in larval output in the warmer spring and summer months

of October through to March (Leys et al., 2008). Interestingly, we observed in this study that skeletal and lipid content peaked during the winter (July sampling month), and propose that this may indicate the storage of lipids and extracellular components in preparation for the main reproductive season that follows. In support of this hypothesis, both the skeleton and lipid biomass components decreased in our samples as the water temperature increased, consistent with the deployment of lipid and extracellular stores to the brood chambers for oocyte and embryonic growth. *A. queenslandica* brood chambers that contain eggs and embryos, which are likely to have high lipid contents (Ereskovsky, 2010), were deliberately not surveyed in this study as they are located basally in adult sponges near the substratum. We infer therefore that the decrease in lipids and skeletal content in the somatic regions during times of highest reproduction is because of deployment of these components to eggs and embryos by parental nurse cells. The sponge *Haliclona permollis* also has an increase in lipid content prior to oogenesis in summer (Elvin, 1979) and lipid inclusions are a common feature of sponge oocytes and early embryos (Ereskovsky, 2010).

The lipid content of *A. queenslandica* increased from summer to the next measurement in early autumn (December to March), and this corresponded with an increase in carbohydrate content. Together, these observations suggest there is an increase through the central energy pathways to meet the increase in energy demanded associated with faster metabolism in warmer water, and to provide substrate metabolites for lipid biosynthesis, and oocyte and embryo formation. In the summer there is also an increase in DNA concentration consistent with an increase in cell abundance, although we did not determine in the present study whether this was attributable to cells of the sponge itself or of its associated bacteria.

4.6 Summary

To begin to understand how the biochemical composition of *A. queenslandica* individuals in Shark Bay on Heron Reef might be influenced by the surrounding water column, we characterised the seasonal availability of key environmentally-derived nutrients, namely carbon, nitrogen, phosphorous

and sulfur. For most of the year, nutrients were predominantly available as dissolved material. The month of March, representing the beginning of autumn and following immediately from the end of an intensive seabird nesting season on the adjacent Heron Island coral cay, stood out as the only month with more carbon in the particulate (plankton) fraction than in the dissolved fraction. Nitrogen availability was not directly related to carbon, and its low abundance suggested it could be a limiting nutrient in this environment. This characterisation of nutrients seasonally available in Shark Bay provides a first step in revealing ecological and physicochemical environmental processes that may be influencing *A. queenslandica's* growth and distribution on Heron Island Reef. The extent to which the *A. queenslandica* holobiont actually relies on terrestrial nitrogenous compounds for nitrogen, or reduced sulfur compounds for energy and carbon fixation, can now be explored by genome-scale metabolic modelling of the system.

Chapter 5 - Development of a dual-species genome-scale metabolic model for the marine sponge *Amphimedon queenslandica* and its dominant bacterial symbiont, *AqS1*, with a focus on biomass synthesis.

5.1 Abstract

Genome-scale metabolic models are powerful tools for simulating metabolic and cellular processes such as biomass synthesis and metabolite production at the system scale. Based on genomic content and biochemical data, metabolic modelling allows the properties of complex systems, such as metabolic networks, to be simulated even though the systems are typically underdetermined (not all biochemical and biophysical properties known). Metabolic modelling has become a routine tool for the improvement of commercially important microbe strains and the prediction of cellular functionality from genomic content. Sponges and their associated microbial community are the source of the greatest diversity in natural products in the marine environment. In addition, marine sponges are important ecosystem engineers that act as a link between the planktonic and benthic habitats. Metabolic modelling was applied to the marine sponge *Amphimedon queenslandica* and the dominant member of its symbiotic bacterial community *AqS1*, to further our understanding of host-symbiont metabolism and interactions. Reconstructing the central metabolic pathways of *A. queenslandica* revealed conserved metazoan variants of glycolysis and the TCA cycle. Ten amino acids were found to be essential for growth in *A. queenslandica*, all of which can be synthesised by the *AqS1* metabolic model. A shared compartment representative of the extracellular space in the sponge mesophyll was used to join the two metabolic models; this is where metabolite exchange occurs between the two organisms. Flux balance analysis (FBA) was performed with nutrient constraints defined by low and high nutrient conditions that are experienced in a natural habitat of *A. queenslandica*, the reef flat of Heron Island Reef. The ability of the model to generate enough energy for pumping and growth of the sponge holobiont was used to validate the model. FBA predicted the exchange of metabolites,

including amino acids, between *A. queenslandica* and *AqS1* under both nutrient conditions. This chapter presents the first genome-scale dual-species metabolic model for any marine invertebrate – symbiont association and is able to synthesise biomass under real nutrient conditions.

5.2 Introduction

5.2.1 Sponges are key nutrient cyclers and producers of secondary metabolites in the marine environment

Sponges are key cyclers of carbon, nitrogen, sulfur and phosphorous in the marine environment due in part to their relationship with symbiotic microbial communities (de Goeij et al., 2013; Hentschel et al., 2012; Rix et al., 2016; F. Zhang et al., 2015). They efficiently filter particulate and dissolved material from huge volumes of seawater each day (Reiswig, 1971; Riisgård et al., 1993; Topçu et al., 2010; Turon et al., 1997; Yahel et al., 2003b), making these nutrients available to benthic trophic levels with the release of detritus material (Alexander et al., 2014; de Goeij et al., 2013; de Goeij et al., 2008a; Rix et al., 2016). The efficient cycling of nutrients has long been seen as vital to the high levels of productivity that coral reefs display in low nutrient environments (de Goeij et al., 2013; Hatcher, 1997; 2003; Pawlik et al., 2016). Primary production on coral reefs is possibly restricted by the availability of important inorganic nutrients such as nitrogen, sulfur and phosphorous; this limits carbon fixation by phototrophic organisms (Hatcher, 2003; Rix et al., 2016), which typically release dissolved organic carbon (DOC) back into the water column on coral reefs (van Duyl & Jan Gast, 2001; Yahel et al., 2003b). The majority of coral reef DOC is produced by the photosynthetic zooxanthellae symbionts of scleractinian corals. 50 % of the carbon fixed by zooxanthellae is released as coral mucus (de Goeij et al., 2013; van Duyl & Jan Gast, 2001; Wild et al., 2004a; Yahel et al., 2003b). Sponges are able to use this DOC with the help of their symbiotic microbial communities (de Goeij et al., 2008a; Mueller et al., 2014; Yahel et al., 2003b).

The feeding habits and the use of nutrients in sponges and how sponge-derived nutrients are cycled back to the surrounding environment are varied and complex, with different aspects having been characterised in different species. Nutrient sources can vary between sponge species, for example *Theonella swinhoei* (Yahel et al., 2003b) and *Xestospongia muta* (McMurray et al., 2016) use DOC as their main carbon source, while picoplankton (0.2-2.0 μm) was identified as the major food source for *Spongia officinalis* (Topçu et al., 2010). Further, most of the carbon contributing to *S. officinalis* biomass was sourced from larger nanoeukaryotes (3-20 μm) (Topçu et al., 2010). *X. muta* has also been shown to discriminate between different particulate food sources, preferring to use cyanobacteria when available in the plankton (McMurray et al., 2016). During periods of increased particulate food, *Haliclona oculata* and *Dysidea avara* responded with high clearance rates and incorporation of carbon that was processed over weeks (Koopmans et al., 2010). A carbon balance in *Haliclona oculata* found that only 34% of the particulate carbon pumped through was used for both respiration and growth, of which 10% was used in biomass synthesis and 90% for energy generation (Koopmans et al., 2009). Much of the nutrients that are converted into biomass appears to be returned to the benthic ecosystem by the shedding of choanocytes, specialised cells that pump seawater from which they filter food (Alexander et al., 2014; de Goeij et al., 2009; 2013; Rix et al., 2016; Shore, 1971).

Despite these insights into nutrient uptake and cycling in a variety of sponges, the biochemical processes connecting the uptake of food with biomass formation are poorly understood in these aquatic invertebrates. In this chapter I apply genome-scale modelling to the coral reef sponge *A. queenslandica* and its dominant symbiont, *AqS1*, to develop a resource that can help predict the metabolic processes, based on genomic and environmental data. A sponge genome-scale metabolic model (GEM) will provide a framework for future work into sponge metabolism. This will enable hypothesis driven experiments to improve our understanding of how sponges use their available nutrients to form biomass, improving our understanding of the services they provide to coral reefs.

In addition to their ecosystem services, marine sponges and their associated bacteria are the most prolific producers of secondary metabolites in the marine environment (Blunt et al., 2014; Haygood et

al., 1999; Laport et al., 2009; Leal et al., 2012; Thakur & Müller, 2004; Thomas et al., 2010). Novel bioactive compounds are routinely discovered and identified, but the low concentrations in which they often exist means that sourcing them from wild stocks of sponges would require unsustainable levels of harvesting (Sipkema et al., 2005). The production of these compounds in commercially viable quantities is currently hampered by these low concentrations, coupled with variable growth rates in both natural and controlled systems (Bergman et al., 2011; Duckworth, 2009; Louden et al., 2007; Osinga et al., 1999), making it very difficult to achieve successful and efficient production of sponge biomass.

Commercial production of the bioactive compounds of sponges has three potential alternative solutions. The first is artificial synthesis, which can be complex and expensive, and produces chemical waste products (Morris & Phillips, 2009; 2011). The second alternative is to use recombinant production, but this would require knowledge on the secondary biosynthetic pathways that are poorly characterised in sponges. Finally, is the use of sponge cell culture to biosynthesise the compounds of interest (Otero & Nielsen, 2010). Cell culture would present conditions that could be controlled and combined with the ability to engineer the metabolic network to improve yields. Currently this approach is limited by the difficulties in sustaining long-term cultures of either sponge cells or their associated bacteria (De Rosa et al., 2003; Grasela et al., 2011; Osinga et al., 2003; Rinkevich, 2011; Schippers et al., 2012). We simply do not know enough about the metabolism of sponges, or their bacterial symbionts, to establish and maintain long-term and sustainable cultures. For example, important nutritional information that would be considered common knowledge for model organisms, such as which amino acids the organism is unable to synthesise and therefore must acquire in its diet, is unknown for sponges. This knowledge gap extends to the entire central metabolic network, which is responsible for energy generation and growth. To improve our understanding of sponge-symbiont metabolism, with a view to informing future culture efforts, I applied constraint-based metabolic modelling to *A. queenslandica* and its dominant species of symbiotic bacteria, *AqS1* (Gauthier et al., 2016). *AqS1* falls within the family Ectothiorhodospiraceae and order Chromatiales, which are commonly referred to as sulfur-oxidising bacteria.

A. queenslandica is a marine sponge of tropical coral reefs. It is unparalleled as a target sponge species for metabolic modelling because of the extensive genomic and biochemical data that are available, including a high quality draft genome of the host sponge itself (Fernandez-Valverde et al., 2015; Srivastava et al., 2010) and also for the two dominant bacterial symbionts (Gauthier et al., 2016). These resources currently are not available for any other species of marine sponge. Access to both sponge host and bacterial symbiont genomes allows for the construction of a dual-species metabolic model to investigate the metabolic nature of the symbiotic relationship, a relationship that is widespread and well documented between marine sponges and microorganisms (Hentschel et al., 2012; Kennedy et al., 2008; Webster & Taylor, 2011). Here I have constructed a dual-species metabolic model to simulate biomass synthesis, drawing upon existing knowledge of the biochemical composition of *A. queenslandica* (Chapter 3; Watson et al., 2014a), and seasonal nutrient availability in its natural habitat (Chapter 4; Watson et al., in revision), which allowed me to investigated how the host and symbiont metabolic networks function and interact in real, low and high nutrient conditions. This allowed me to simulate how their metabolic networks function and interact under different environmental conditions.

5.2.2 Constraint-based metabolic modelling

Metabolic modelling centres on reconstructing a mathematical representation of an organism's metabolic network based on the stoichiometry of biochemical reactions. Various approaches to metabolic modelling exist, including constraint-based, thermodynamic and kinetic modelling (Henry et al., 2007; Lewis et al., 2012; Steuer et al., 2006). I have chosen to apply the constrained-based modelling approach (Lewis et al., 2012) to the dual-species model developed in this chapter. Constraint-based modelling allows the prediction of metabolic phenotypes, such as growth rates and the flow of material through the network, under defined conditions (O'Brien et al., 2015). The advantages of being able to simulate the effects of different nutrient conditions, or of gene deletions and additions, have made constraint-based modelling an indispensable tool for commercial metabolic engineering (e.g. Averesch & Krömer, 2014; Oberhardt et al., 2009; Sengupta et al., 2013; Wang et al., 2014). The

application of genome-scale metabolic models (GEMs) has expanded from the traditional microbial models to include examples of complex multi-organism interactions that span the domains of life. Prokaryote-eukaryote metabolic models include the host pea aphid (*Acyrthosiphon pisum*) and its intracellular symbiont *Buchnera aphidicola* (Macdonald et al., 2012), the response of human alveolar macrophage cells to the pathogenic *Mycobacterium tuberculosis* (Bordbar et al., 2010), and notably the metabolic exchange between host and gut microbiomes of both mouse (Heinken et al., 2013) and human systems (Shoaie et al., 2013).

The reconstructed network is an organism-specific database of the biochemical (enzymatic and spontaneous) reactions that are encoded within its genome (Figure 5.1). The initial database of metabolic reactions may be compiled manually (Thiele & Palsson, 2010), or be automatically generated using publicly available bioinformatic resources (Agren et al., 2013; Henry et al., 2010). Automatically-generated reaction databases still require manual curation due to annotation and computational errors, or the presence of novel metabolic capabilities, before being converted into a metabolic model (Henry et al., 2010; Thiele & Palsson, 2010; Thorleifsson & Thiele, 2011). The reaction database contains detailed information for each reaction, including the metabolite stoichiometry of each reaction, the directions in which the reaction can proceed, and the genes that encode an enzyme that can perform each reaction (Thiele & Palsson, 2010). The reaction database is then converted into a mathematical model in the form of a stoichiometric matrix. The stoichiometric matrix is populated using the metabolite stoichiometry of each reaction in the reaction database (Figure 5.1), after which it can be analysed. A review published in 2012 (Lewis et al., 2012) referenced over 100 analysis methods including protocols aimed at uncovering network topology, determining reaction-enzyme-gene importance and essentiality, and the prediction of metabolic and cellular phenotypes. Among these, flux balance analysis (FBA) can be used to simulate growth rates and the relative flow of metabolites through a network (Krömer et al., 2014; Orth et al., 2010; Santos et al., 2011). For FBA, and all metabolic modelling, constraints are placed on the model to define the specific conditions under which the metabolic network is being simulated.

The inherent complexity of a metabolic network means that there are many possible solutions to the same problem; this is termed the solution space (Orth et al., 2010; O'Brien et al., 2015). Constraints are applied that define the conditions to which the network must conform for a particular simulation (Lewis et al., 2012; O'Brien et al., 2015). This allows the user to reduce the solution space and, importantly, to place the model into a relevant biological context. Important constraints for modelling growth are the model inputs, such as nutrient availability, the direction each reaction is allowed to proceed, and the model output specifically being the biochemical composition of the organism (Lewis et al., 2012; Orth et al., 2010; O'Brien et al., 2015). Specifically, the biomass composition is an equation containing all of the metabolites that contribute to the organisms' biomass (Feist & Palsson, 2010). The coefficients of the biomass equation can be used to constrain the model, by setting the specific proportions in which each contributing metabolite must be synthesised so as to meet the observed biomass composition of the organism of interest (Feist & Palsson, 2010; Lewis et al., 2012; O'Brien et al., 2015). A steady state assumption is applied, meaning that the inputs must be mass balanced through the network using the stoichiometry of the reactions to produce the defined biomass composition; that is any metabolites produced by active reactions must be consumed elsewhere in the network or exported as product (Figure 5.1) (Lewis et al., 2012; Orth et al., 2010; O'Brien et al., 2015). A solution to a simulation is essentially a stoichiometrically-balanced biochemical equation at the genome-scale.

Each simulation of the metabolic network results in the prediction of a value for each reaction, these are termed fluxes (Orth et al., 2010), and refer to the relative amount of material that passes through the reaction to achieve the mass-balanced solution. Fluxes are expressed in molar values, as the different molecular weight of individual metabolites is not easily calculated for the flow of mass through a metabolic network. The rate of the fluxes through the network is determined by the units of measurement applied to the input constraints, for example millimol per hour per gram dry weight (mmol / h / gDW).

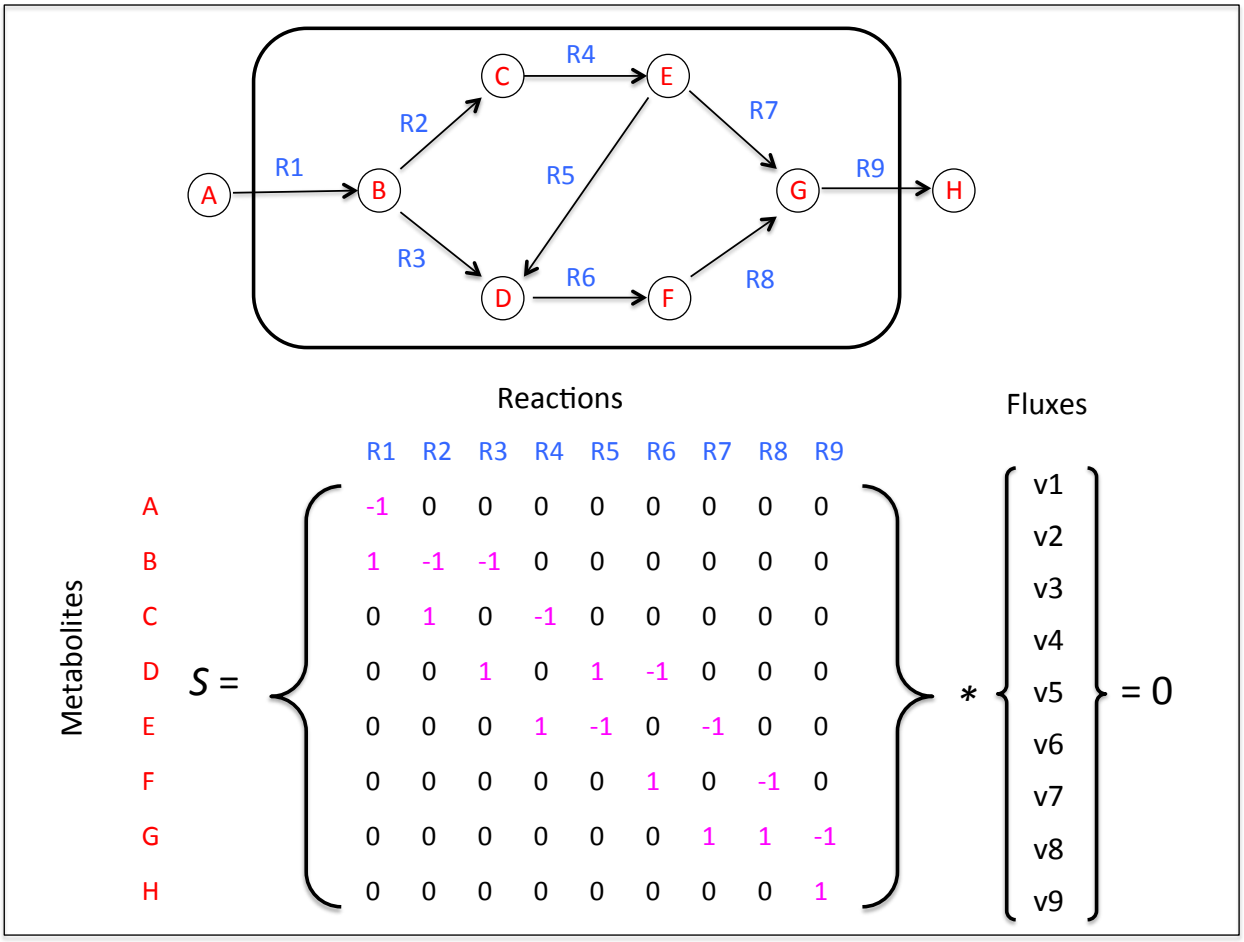


Figure 5.1 A simple reaction network (above) and corresponding stoichiometric matrix (below).

Red letters represent metabolites in the network, while arrows represent the reactions with their corresponding flux in blue. The metabolites A and H represent boundary species. The resulting stoichiometric matrix (S) is sparsely populated with the metabolite stoichiometry of each reaction. Sparsely populated matrix reflects how the metabolites in a metabolic network are loosely coupled. A negative number means that the metabolite is consumed and positive number means that it is a reaction product. This network contains no gaps and the stoichiometric matrix is connected, meaning that mass can flow through the network and a steady state achieved where all of the produced metabolites are consumed.

5.3 Methods

Individual metabolic models were initially constructed for *Amphimedon queenslandica* and *AqS1* before they were joined to create the dual-species model. The *A. queenslandica* model was manually constructed and the draft *AqS1* model was automatically generated as described below. Both models

were individually curated following the steps detailed below for *A. queenslandica*. Generating a functional model and setting realistic constraints relied upon data sourced from *A. queenslandica* and *AqS1*'s genomic resources, general information on sponges from the scientific literature and specific experiments conducted in this study to fill knowledge gaps. As many of the parameters and settings used when performing metabolic simulations are based on particular decisions, I have elected to include explanations of the rationale underpinning these methods in this section. Before constructing these models and simulating their function, I recount decisions for using particular gene models, and a number of parameters based on the literature and prior analyses (i.e. Chapters 3 and 4). If calculations are based on literature values then they are presented here. I also include new experimentation and data that are deemed necessary for accurate modelling.

5.3.1 Estimating the pumping rate of *A. queenslandica*

Ten adult *A. queenslandica* were collected from Shark Bay, Heron Island, as described in (Leys et al., 2008) and transported to the adjacent Heron Island Research Station. The sponges were maintained in individual tanks with a flowthrough supply of ambient seawater from the reef flat. Prior to measuring pumping rates, the flowthrough water was stopped and the water in the aquaria allowed to settle. The flow rate of each osculum was determined by loading approximately 2.0 mL of 5.0 g/L fluorescein dye (Sigma, F6377) in sea water adjacent to the outer sponge surface and recording the exhalent flow on video (Figure 5.2). Three replicate videos were taken of each osculum. Between each video, the water in the holding tank was completely flushed with fresh seawater. The exhalent speed of each osculum was analysed using the Open Source Physics software, Tracker (http://physlets.org/tracker/). The diameter of each osculum was measured and the dry weight of each sponge determined by lyophilisation (Watson et al., 2014b). The average diameter of osculum per gDW was then calculated. This was then converted into an average area of osculum per gDW and subsequently, average volume pumped per hour per gDW of sponge.

5.3.2 Defining the composition of the different nutrient sources

To define the nutrients that can be obtained from the capture and phagocytosis of particulate food, the biochemical composition of *Escherichia coli* and *Chlamydomonas reinhardtii* were used to represent the nutrients obtained from bacterial and phytoplankton cells. The biomass composition for bacterioplankton, as represented by *E. coli,* strain K-12 MG1655, contains 28 metabolites (Appendix 5.1) and was sourced from the metabolic model *i*JO1366 (Orth et al., 2011). The selection of a representative species of phytoplankton was restricted as the two metabolic models that exist for marine phytoplankton (Krumholz et al., 2012) do not contain coefficients for the biomass components. Knowing the relative amount that each metabolite contributes is essential, leading to the choice of a freshwater algae, *C. reinhardtii* (de Oliveira Dal'Molin et al., 2011), to represent phytoplankton. The *C. reinhardtii* biomass input equation contained 23 metabolites (Appendix 5.2).

The literature repeatedly reports that the most significant contributor to the dissolved organic fraction on coral reefs is mucus that is excreted by corals (Ferrier-Pages et al., 1998; Rix et al., 2016; van Duyl & Jan Gast, 2001; Yahel et al., 2003b). To represent the dissolved nutrient fraction in the model, I used an average metabolite composition calculated from known compositions of the mucus excreted by the six most abundant coral species in Shark Bay (Wild et al., 2005). The resulting coral mucus nutrient source was composed entirely of carbohydrates – L-arabinose, L-Rhamnose, L-fucose, xylose, D-mannose, galactose, D-glucose and N-acetyl-D-glucosamine. The coefficients were calculated using the seasonal abundances of dissolved organic carbon (DOC) reported in Chapter 4 (Watson et al., in revision). See Appendix 5.3 for the complete list of metabolites and their coefficients.

An exchange reaction was added to allow the export of each input metabolite from the *A. queenslandica* or *AqS1* compartment into the ECM (see section 5.3.11). This was to prevent the *A. queenslandica* or *AqS1* sub-network from being forced to mass balance the complete utilisation of

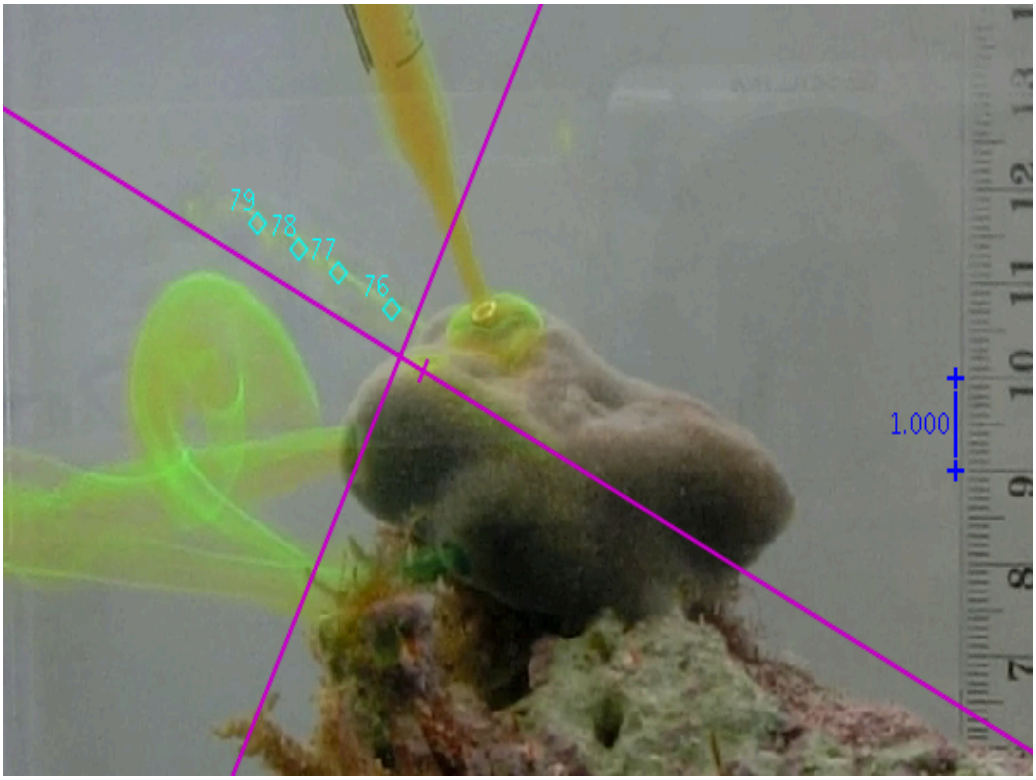


Figure 5.2 The average volume that *A. queenslandica* pumps being estimated by video analysis.

Using fluorescein seawater loaded near each osculum. The purple axes are manually aligned to the flow. The blue scale bar is at the same depth from the camera as the oscula of interest. Successive frames marking the leading edge of the labelled flow are marked with number blue squares; Tracker software calculates the velocity of the exhalent water based on these.

each metabolite in the algal, bacterial and coral mucus food sources. If a metabolite was exported from the sponge compartment, it was available for *AqS1* if a membrane transport reaction was present within the *AqS1* model. Likewise for *A. queenslandica*. Exchange reactions were also added at the system boundary (out of the ECM) to allow each input metabolite to be excreted from the model. These boundary exchange reactions only allow the excretion of metabolites that either the *A. queenslandica* or *AqS1*s metabolic network cannot use or are in excess. Reactions were added to *A. queenslandica*, *AqS1* and at the ECM boundary to allow ethanol and organic acids – pyruvate, formate, propionate, acetate, citrate, fumarate, succinate, lactate – to be exported. Finally, the exchange of both essential compounds, such as water and ions, and of waste products, such as carbon dioxide, was allowed to occur across the model boundary.

5.3.3 Estimating the availability of nutrients

Defining the nutrient constraints was undertaken using the seasonal variation in nutrient availability as reported in Chapter 4 (Watson et al., in revision), combined with the pumping rate (from 5.4.1) and composition of food sources (from 5.3.2). As a starting point for modelling, the October and March time points from Chapter 4 were chosen to represent low and high nutrient conditions experienced by *A. queenslandica*. For most of the year in which measurements were taken, carbon was more abundant in the dissolved fraction than it was in the particulate fraction; the exception was the month of March (Chapter 4, Watson et al., in revision). To calculate the relative amounts of DOC to POC, the quantified carbon values reported in Chapter 4 were broken down into particulate and dissolved concentrations and then converted from mg / L to mmol / h / gDW.

As POC was not directly quantified in Chapter 4 (Watson et al., in revision), the amount of POC was estimated by subtracting the DOC from the total organic carbon (TOC). An assumption was made that the POC was comprised of only bacterioplankton and phytoplankton cells. To estimate the relative amounts of bacteria to algae in the POC, the abundances reported in a study conducted previously at Shark Bay (Sorokin & Sorokin, 2009) were used. An average of 11.53% phytoplankton was calculated from the data reported by Sorokin (2009), whose samples were representative of both day and night abundances. This percentage was applied to the POC to estimate the relative abundance of bacterioplankton and phytoplankton, expressed as mg of carbon / L. Multiplication by the pumping rate gave mg of carbon / h / gDW.

Finally, the mg of carbon / h / gDW values were converted to mmol carbon / h / gDW. An average molecular weight was estimated for both *E. coli* and *C. reinhardtii* from the molecular weights of all metabolites included in their respective biomass equations (see 5.3.2). This average molecular weight was divided by the mg of carbon / h / gDW to produce mmol / h / gDW. Subsequently all flux values are reported as mmol / h / gDW.

5.3.4 Estimating oxygen availability

The availability of oxygen in the water column was quantified during summer (December 2012) using a water quality monitoring probe (YSI EXO2 Multiparamter Sonde) deployed adjacent to an *A. queenslandica* individual in Shark Bay. The probe was deployed for two 24-hour periods, sampling every two minutes. The two sampling periods were separated by 24 h and generated 1492 oxygen concentration readings. The average oxygen concentration was then calculated and converted into mmol / h / gDW using the estimated pumping rate (see 5.4.1).

5.3.5 Estimating metabolic cost of choanocyte cell turnover in *A. queenslandica*

Choanocytes are responsible for pumping water through the sponge body and extracting food particles from this water and are shed constantly into the environment (de Goeij et al., 2009). As the rate of choanocyte turnover has not yet been quantified specifically in *A. queenslandica*, the two published rates were averaged; complete replacement of choanocyte cell population occurs every 5.4 and 20 h for *Halisarca caerula* (tropical reef species) and *Hymeniacidon sinapium* (subtropical species), respectively (de Goeij et al., 2009; Shore, 1971); choanocytes comprise 18% of the tissue volume for *H. caerulea* (de Goeij et al., 2009). Based on these values, an estimated metabolic cost of producing and shedding choanocyte cells was made for *A. queenslandica*, with 18% cellular biomass being replaced every 12.7 h (1.42% per h).

5.3.6 Reconstructing the metabolic network of *A. queenslandica*

The *A. queenslandica* metabolic model was manually created in Excel (Microsoft) following a previously published protocol (see {Thiele:2010fr} for details). The *A. queenslandica* metabolic model was created using the Model SEED metabolite names and IDs to facilitate future work on the *A. queenslandica* holobiont, as microbial and metagenome scale metabolic models can be generated

quickly by the Model SEED online bioinformatics toolkit. The reaction database was constructed using the most current *A. queenslandica* gene models (Aqu2.1; {FernandezValverde:2015bo}. These predicted gene models were annotated with the Uniprot database (www.uniprot.org), and subsequently a Pathway Tools {Karp:2010cy} Aqu2.1 organism pathway database was generated using the Pathologic tool {Karp:2002vi}. The Aqu2.1 organism pathway database was not made public as no further curation was performed after generating the automated database. The Aqu2.1 Pathway Tools database was generated to use as a resource to complement KEGG, Uniprot and the Model SEED pathway databases. The Aqu1 KEGG database (http://www.genome.jp/kegg-bin/show_organism?org=aqu) was used to assist in determining what pathways to include. The *A. queenslandica* KEGG database refers to the original Aqu1 gene models and annotation {Srivastava:2010iea}. If a pathway map was present in the Aqu1 KEGG database, then all of the reactions on that map were manually checked for inclusion in the Aqu2.1 metabolic model. This was done by searching the Aqu2.1 Pathway Tools database for the presence of EC numbers with associated gene ID's.

If a reaction was present in the *A. queenslandica* Pathway Tools database, the reactions, metabolite names and IDs were extracted from the Model SEED website (http://modelseed.org/). Reactions that were not present in the Model SEED database were entered using KEGG metabolite names and IDs. The gene models associated with a reaction were sourced from the *A. queenslandica* Pathway Tools database. The reconstructed network was compiled in Excel (Microsoft) format for ease of curation and converted into SBML (www.sbml.org) as in Dal'Molin et al. (2010). The FBA were performed with the COBRA Toolbox {Schellenberger:2011fs} in the MATLAB environment (Release 2015B, The MathWorks, Inc.). All FBA analyses were performed as mixed integer linear programming (MILP) problems to remove biologically infeasible loops from the solution {Price:2002bx, Schellenberger:2011fs}.

5.3.7 Generating the *A. queenslandica* biomass equation

The biomass equation describes the molecular composition of an organism, in their relative proportions required to synthesise biomass (Feist & Palsson, 2010). Besides external measurements

of substrates and products of the system, the biomass equation is the most important constraint that can be put on internal flux distributions. The biomass equation was created for *A. queenslandica* using the biochemical data reported in Chapters 3 (Watson et al., 2014a) and 4 (Watson et al., in revision). The coefficient calculations were performed using the abundances of building block molecules (fatty acids, sterols and amino acids) detailed in Chapter 3 (Watson et al., 2014a). The mass contributions (g / gDW) of each macromolecule type to the overall biomass were different under the two nutrient (seasonal) conditions and sourced from the relative amounts reported in Chapter 4 (Watson et al., in revision). Separate biomass coefficient calculations were made for the two nutrient conditions - low and high nutrient abundance - that were simulated (see section 5.3.3). This gave different biomass objectives that reflected the differences in biochemical composition in different nutrient conditions (Chapter 4; Watson et al., in revision). As the intracellular pH of both *A. queenslandica* and *AqS1* is unknown, I assumed a neutral pH (Henry et al., 2010). There is a detailed example of calculating nucleotide coefficients from biochemical measurements in Chapter 2 (Watson et al., 2014b). In short, separate calculations were performed for the building block molecules of each macromolecule type. Using amino acids and protein as an example, first an average molecular weight was calculated for the amino acids included in the biomass equation. Next the relative abundance (percentage) of each amino acid to the protein content was calculated. A total amount of amino acid moles was calculated by dividing the amount of protein (g /gDW) by the average molecular mass of the included amino acids. This amount of average amino acid moles was then proportionally distributed among the included amino acids using their relative abundance (percentage of total protein).

The overall biomass of *A. queenslandica* can be broken down into two main components with distinct differences in composition and turnover rates. These two components are the skeleton and the cellular biomass (Chapter 2; Watson et al., 2014b, and Chapter 3; Watson et al., 2014a). Separate biomass equations were created for both of these components so that their relative contributions to the overall biomass were adjustable within the model.

The skeleton is the dominant component of the total sponge biomass, contributing to 57.3 +/- 0.07 % of the total dry weight of *A. queenslandica* during the high nutrient month of March and 58.5 +/- 0.08 % during the low nutrient month of October (Chapter 3; Watson et al., 2014a). The skeleton is largely comprised of spongin, a sponge-specific collagen formed from amino acid building blocks (Clancey, 1926). Chapter 3 details the skeletal amino acid composition, although values for L-cysteine and L-methionine and L-tryptophan are unavailable. L-cysteine was not measured by the methods used, while L-methionine and L-trpytophan were destroyed during the acid hydrolysis sample preparation (Chapter 3; Watson et al., 2014a). Published literature on the composition of sponge collagens supports the assumption taken in generating the skeletal amino acid composition that L-cysteine and L-methionine, the sulfur containing amino acids, contribute little to the overall spongin or collagen composition (Clancey, 1926; Imhoff & Garrone, 1983; Junqua et al., 1974; Lipinski et al., 2001). The analysis of collagen within *Spongia graminea* showed 0.48% and 1.67% sulfur content (Gross et al., 1956), while *Chondrosia reniformis* skeleton did not contain detectable cysteine and only trace amounts of methionine (Imhoff & Garrone, 1983).

The *A. queenslandica* cellular biomass equation includes the nucleotides, fatty acids, sterols and cellular-derived amino acids. The nucleotide coefficients were calculated using the relative abundance reported in Chapter 3 (Watson et al., 2014a). The biomass analysis reported several fatty acids that could not be identified (Chapter 3; Watson et al., 2014a), and while these contribute significantly to the lipid fraction, they were unable to be included in the *A. queenslandica* model. Instead, to calculate the coefficients for the fatty acids that were identified, it was assumed that the total lipid dry weight was comprised of only the identified fatty acids and sterols. This assumption should have only a minor effect on the overall fluxes as fatty acids are all derived from the acetyl-coenzyme-A pool and the cost of fatty acid carbon chain elongation (reducing power in the form of NADPH) is relatively small when compared to amino acid biosynthesis (Michal & Schomburg, 2013). The sterols, cholesterol and brassicasterol, were pooled with the fatty acids when calculating the coefficients as the total lipid dry weight from the biochemical analysis did not discriminate between these fractions (Chapter 3; (Watson et al., 2014a).

The average total carbohydrate content of *A. queenslandica* was quantified during the biochemical analysis reported in Chapters 3 (Watson et al., 2014a) and 4 (Watson et al., in revision), although the composition was not identified. Glycogen is the common carbohydrate storage compound in eukaryotes (Michal & Schomburg, 2013), and complete glycogen synthesis and degradation pathways are encoded by the *A. queenslandica* genome. Thus glycogen was added to the biomass equation of both nutrient conditions, and its coefficient calculated from the data in Chapter 4 (Watson et al., in revision) to represent the total carbohydrate content.

The amino acid analyses previously reported for *A. queenslandica* were conducted on the skeleton by itself, and on samples of complete biomass (Chapter 3; Watson et al., 2014a). I chose not to use these data as the cellular amino acid composition would be an estimate from subtracting the average skeletal amino acid composition from the average amino acid composition of the complete biomass. This would not contain values for L-cysteine, L-methionine and L-tryptophan (see above). As an alternative to the data in Chapter 3 (Watson et al., 2014a), the codon frequency was calculated using all of the *A. queenslandica* adult translated protein sequences (Fernandez-Valverde et al., 2015). It was important to include values for L-cysteine, L-methionine and L-tryptophan due to their involvement with sulfur and niacin (vitamin B3) metabolism. The list of biomass components and respective coefficients for both sponge skeleton (Appendix 5.5) and sponge cellular biomass (Appendix 5.6) under both nutrient conditions is detailed in the Appendices.

5.3.8 Network gaps and gap filling in the *A. queenslandica* model

A metabolic simulation represents a stoichiometrically balanced metabolic steady-state where all of the metabolites that are produced must be consumed (Bordbar et al., 2014; Orth et al., 2010; O'Brien et al., 2015). Gaps in the network that prevent the synthesis or consumption of a metabolite can result in an unfeasible model with no possible solution (Bordbar et al., 2014; Orth et al., 2010; O'Brien et al., 2015). Gaps may reflect a true biological loss of function in the organism, represent

unknown biochemical capabilities, or they may be a result of incorrect gene prediction or annotation (Montero, & Peret, 2013; Orth & Palsson, 2012; O'Brien et al., 2015). For the model to be functional, gaps in the network must be identified. In the case of annotation error, the reactions may be added to the model once the correct gene models are identified in the genome. Reactions may also be added to fill gaps where no respective gene is identified to ensure the network is connected and the model can function to achieve its objective, this can be done manually or with automated bioinformatics tools (Henry et al., 2010; Latendresse et al., 2012; Orth & Palsson, 2012; Reed et al., 2006; Kumar et al., 2007; Thiele & Palsson, 2010).

Several pathways were missing or incomplete in the draft *A. queenslandica* model as it was unable to produce biomass from simple carbon (D-glucose) and nitrogen (ammonia) sources. A gap was filled with the missing reaction if it consisted of a single reaction; if multiple successive reactions were missing, then the gap was assumed to be real. This conservative approach allowed the model to be more connected in the absence of biochemical data. Biomass metabolites that could not be synthesised were added as a boundary species and were considered an essential nutrient that must be supplied to *A. queenslandica* to enable biomass synthesis. Boundary species are metabolites that do not require to be mass balanced to produce a flux solution (Figure 5.1); that is, they are essentially metabolites that can enter and leave the system (this rate can be constrained)(F. Santos et al., 2011; Thiele & Palsson, 2010).

Transport reactions in *A. queenslandica* were identified by the Transport Inference Parser in Pathway Tools (Lee et al., 2008), and were added to the draft model if they contained specific metabolites that were already present; non-specific metabolite names such as "an amino acid" could not be included in the model. Additional transport reactions were added for metabolites that were required for model functionality, such as oxygen and carbon dioxide (Thiele & Palsson, 2010). Where a reaction potentially had different cofactors, such as NAD or NADH, separate reactions were included for all cofactors (Thiele & Palsson, 2010).

Energy in the form of ATP is generated in cells by creating a proton gradient across a membrane. In prokaryotes this occurs between the cellular membrane and the periplasmic space, while in eukaryotes this occurs across the inner mitochondrial membrane (Michal & Schomburg, 2013). The protons involved in oxidative phosphorylation represent a pool that is not available to other reactions. To reflect this, these protons were given a different metabolite ID when they were theoretically transported across their respective membranes. This allowed the protons involved in ATP synthesis to remain separate pools in both *A. queenslandica* and *AqS1*, which were unavailable to other reactions in the model.

5.3.9 Generating the *AqS1* metabolic reconstruction

The draft *AqS1* genome (Gauthier et al, 2016) was annotated using the Rapid Annotations using Subsystems Technology (RAST) server (Aziz et al., 2008). From this annotated genome, a draft metabolic model was automatically generated using the model SEED (Henry et al., 2010). This included automatic gap filling to enable each component of the biomass equation to be hypothetically produced from complete media. In the model SEED, the definition of complete media is organism specific and is determined by the transport reactions identified in the organism's genome; the complete media is comprised of all those metabolites for which a transport reaction was identified. Transport reactions for compounds that are not identified within the genome, but are essential for model functionality (such as water and oxygen transport), are also added automatically in the draft reconstruction. The biomass composition of *AqS1* is unknown, so the default gram-negative biomass composition generated by the model SEED was used as a template, with the following changes applied (see Appendix 5.7 for the complete list of *AqS1* biomass metabolites and their coefficients). First, the amino acid coefficients were changed to match their relative observed abundance in the *AqS1* genome. Second, the overall complexity was simplified to contain only amino acids and nucleotides (Appendix 5.7). The amino acid (protein) fraction accounts for 55 % of the dry weight in bacteria making it the most influential biomass component (Stephanopoulos et al., 1998). DNA and RNA

account for 3 % and 20 % respectively, meaning that nearly 80 % of the dry weight is theoretically adjusted (Stephanopoulos et al., 1998).

An Excel (Microsoft) format of the draft model was exported from the Model SEED and manually curated following established protocols {Henry:2010dga, Thiele:2010fr}. To aid with curation of the draft model in Excel (Microsoft), an *AqS1* specific Pathway Tools database was created {Karp:2010cy}, which was used to complement the ModelSEED and KEGG pathway databases during pathway inspections and gap filling. The *AqS1* Pathway Tools database was generated using a *AqS1* Genbank file exported from the SEED viewer. The initial Pathway Tools curation steps were performed with Pathologic (Karp et al., 2002). The resulting *AqS1* specific Pathway Tools database (Appendix 5.8) did not include reactions that were gap filled by the Model SEED, meaning that the *AqS1* Pathway Tools database could be compared to the pathway maps available within the Model SEED viewer, which are generated from KEGG pathway maps. The draft model exported from The Model SEED contained errors and the *AqS1* Pathway Tools database provided gap-filling solutions and allowed identification of missing reactions and genes. Model curation, such as editing incorrect reactions, reversibility and gap-filling, was performed in Excel (Microsoft) and converted into SBML (www.sbml.org) using an in-house Java Script (Oracle Corporation), before being analysed with the COBRA Toolbox (Schellenberger et al., 2011) in the MATLAB environment (Release 2015B, The MathWorks, Inc.).

5.3.10 Energy demands for growth, maintenance and pumping water

Adenosine triphosphate (ATP) is the energy carrying metabolite of all cells (Michal & Schomburg, 2013) and is required to power a diverse range of cellular functions that must be accounted for in a metabolic model. There are ATP costs associated with making the biomass and then maintenance energy. Some maintenance is non-growth associated (basic life support - NGAM) some scales with growth (growth associated maintenance - GAM). Growth associated is the synthesis of building block metabolites such as amino acids, and their subsequent polymerisation into protein. The synthesis of

building block metabolites is accounted for in the reactions within the model but the ATP cost of polymerisation must be calculated and added to the biomass equations. The different compositions of the *A. queenslandica* skeleton and cellular biomass equations, and for the *AqS1* biomass equation, meant that they had different ATP requirements for polymerisation. This ATP cost was individually calculated and added to each of these equations following the protocol of Thiele et al. (2010).

Non-growth associated maintenance costs include cellular processes such as repairing damage and maintaining ion gradients (for review of cellular ATP costs see Locasale & Cantley, 2011). This energetic cost is unknown for both *A. queenslandica* and *AqS1*. A reported NGAM cost for a hybridoma cell line was 28.6 $mmol_{ATP}$ per 10^9 cells per hour (Xie & Wang, 2003). With an average dry weight of 450 pg per cell (Lake-Ee Quek, personal communication), an NGAM value of 6.36 $mmol_{ATP}$ / h / gDW was assumed for *A. queenslandica*. The literature value of 8.39 $mmol_{ATP}$ / h / gDW was taken from *Escherichia coli* (Thiele & Palsson, 2010) and applied to *AqS1*. Sponges also have the non-growth energetic cost of pumping water to account for.

Metabolic models are validated by comparing the predicted rates of growth and metabolite production with quantified *in situ* measurements. With such data lacking for *A. queenslandica*, the ability of the model to produce enough ATP to meet the requirements for both pumping, growth and cellular maintenance was used to test the model. To estimate the pumping energy cost, the standard sponge pump model proposed by Riisgard et al (1993) was applied to *A. queenslandica*. This standard sponge pump model (Riisgård et al., 1993) was calculated from the pumping rates of two demosponges, *Halichondria panicea* and *Haliclona urceolus*, and is defined as having a dry weight of 0.1 g, pumping 6 mL / min, at a maximum power output of 0.667 μW (Riisgård et al., 1993). To test the suitability of applying this model to *A. queenslandica*, I compared the pumping rates that were measured with the predicted pumping rates of the theoretical standard sponge pump. Riisgard et al (1993) quantified the pumping rates of *H. panicea* and *H. urceolus* at different temperatures, which revealed a linear relationship. This relationship is represented by the equation, $F = 0.77T - 3.2$, where

F is pumping rate and T is temperature. When this formula is applied to the July average temperature of 21.15 °C (Chapter 4; Watson et at., in review), it predicts a pumping rate of 7.85 L / h / gDW.

This compares well to the actual flow rate measured in July of 9.52 L / h / gDW (see 5.3.1 and 5.4.1), when the following points are taken into account. First, the water temperature data is sourced from a monitoring station located in the lagoon adjacent to Shark Bay and represents an average over a 96-hour period (Chapter 4; Watson et al., in revision). Second, the flow rate was quantified during the day, with the holding tanks being supplied by a flow through system where the header supply tank is position in direct sunlight. As a result the incoming water is warmer than the reef water. As a comparison calculation, if the yearly average temperature of 23.75 °C (2.6 °C higher than the lagoon temperature when the flow rate was quantified) is applied to the equation, it predicts a pumping rate of 9.05 L / h / gDW, which is very close to the measured pumping rate (9.52 L / h / gDW). Using the sponge pump model, the ATP demand for pumping was calculated to be 20.82 mmol / ATP / h / gDW using the quantified flow rate (9.52 L / h / gDW). The calculations are explained in detailed below.

To estimate the energetic pumping cost, the power output of the sponge pump model (Riisgård et al., 1993) was converted from μW / 0.1 gDW to mmol ATP. As all units in the GEM are expressed per gram of sponge dry weight, the maximum power output of the sponge pump model was multiplied by 10 to give 6.67 μW / gDW. The sponge pump model has a pumping rate of 3.6 L / h / gDW (Riisgård et al., 1993). This lower pumping rate than *A. queenslandica* is due to the sponge pump model being characterised at 12°C (Riisgård et al., 1993). The maximum power output of the standard sponge pump model (6.67 μW / gDW) was divided by its pumping rate of 3.6 L / h / gDW to give 1.85 μW / L / h / gDW. This value was then multiplied by the measured pumping rate of 9.52 L / h / gDW to give 17.64 μW / h / gDW. For ease of calculation, μW were converted to kW. As 1 kilowatt hour is equal to 3.60E+10 joules, the calculated kW / h / gDW was converted to joules by multiplying by 3.60E+10. Joules were then converted to kilojoules as 30.5 kilojoules of energy is equal to 1 mol of ATP (Michal & Schomburg, 2013). Finally, the amount of kilojoules was divided by 30.5 to give 0.02 mol / h / gDW and converted to 20.82 $mmol_{ATP}$ / h / gDW.

5.3.11 Generating the dual-species model

To join the *A. queenslandica* and *AqS1* metabolic models, I created a third compartment called the extracellular matrix (ECM) (Figure 5.3). This compartment functions to join the *A. queenslandica* and *AqS1* metabolic networks by representing the extracellular space within the sponge mesohyl where the sponge and symbiont cells interact. Sponges actively pump water through their bodies to access nutrients and remove waste products (Simpson, 2012). The nutrients that are available in the water column of a coral reef consist of both particulate and dissolved matter (Birkeland, 1997a; Carlson & Hansell, 2015; de Goeij et al., 2008a; van Duyl & Jan Gast, 2001). The particulate fraction includes planktonic algae and bacteria, while the dissolved fraction consists of smaller organic and inorganic compounds. It is likely that the particulate algae and bacteria are phagocytised by the sponge cells (Pile et al., 1996a; Ribes et al., 1999b; Turon et al., 1997), while the bacterial cells are able to use the dissolved fraction (Weisz et al., 2007b; Yahel et al., 2003b). The dual-species model is setup so that the *A. queenslandica* network uptakes the algal and bacterial food sources, can use what metabolites it requires, and then exports excess to the ECM. If a transport reaction was identified in the *AqS1* genome, it can use metabolites from the ECM. Likewise, *AqS1* is able to use the dissolved nutrient food source before exporting excess metabolites to the ECM where *A. queenslandica* may use them if a specific transport enzyme was identified in its genome.

5.3.12 Analysis of the multi species metabolic model

All FBA were performed using the COBRA Toolbox (Schellenberger et al., 2011) in the MATLAB environment (Release 2015B, The MathWorks, Inc.). An artefact of reconstructed networks is that reactions and metabolites that are not connected in reality can form biologically infeasible loops in the models (Price et al., 2002). Loops can generate artificially high fluxes within themselves and can erroneously influence FBA. To avoid this, all FBA were performed with the option of UNFEASIBLE LOOPS REMOVED. Either the biomass composition or the rate of ATP synthesis, were set as the model objective for individual simulations.

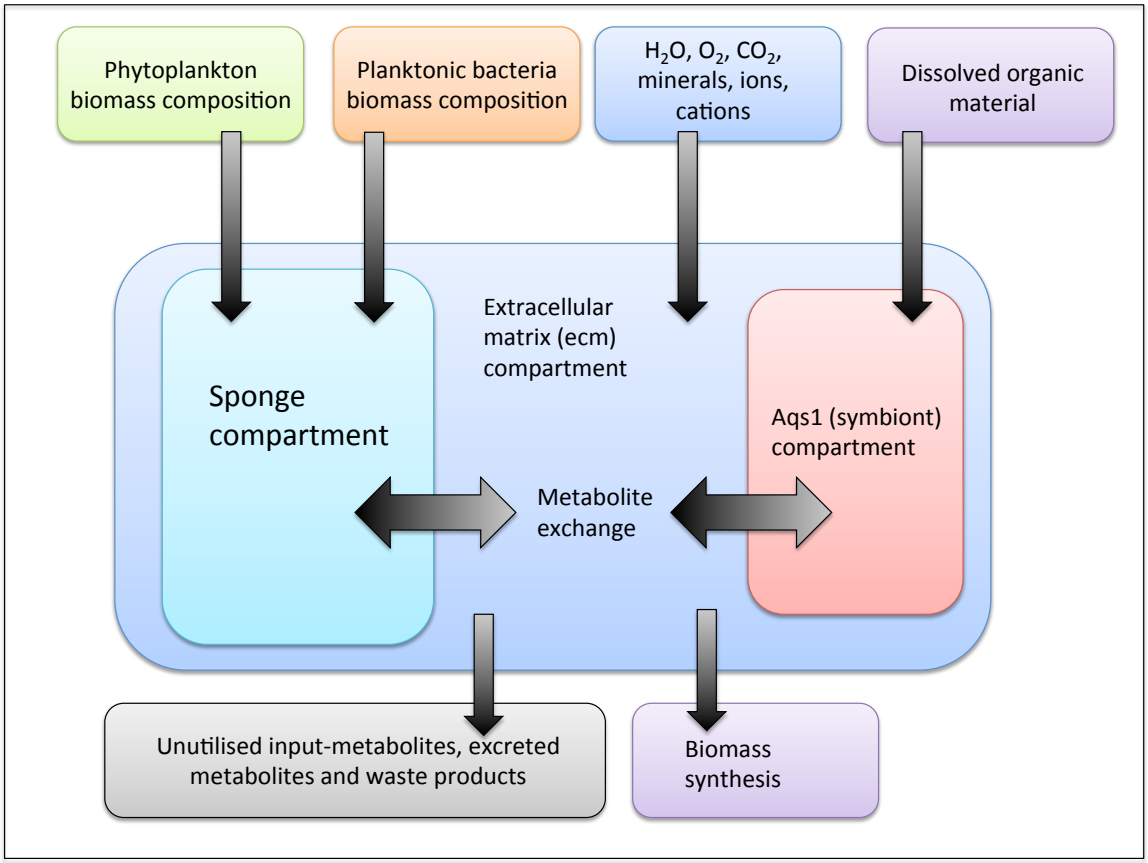


Figure 5.3 A schematic view of compartmentalisation of the dual-species model.

The sponge compartment is constrained to uptake phytoplankton and bacterioplankton cells, while dissolved organic material is utilised by the bacteria, *AqS1*. Metabolites from the different input sources that are not utilised, or waste products from the holobiont metabolism, can then be exported into the extracellular matrix compartment. These compounds can then be utilised by the other cell type, or excreted from the holobiont system.

5.4 Results

5.4.1 The pumping rate of *A. queenslandica*

The pumping rate of *A. queenslandica* was quantified to estimate (i) the ATP requirements for pumping for validation of the model, and (ii) the rate at which food could be obtained from the

water. The latter also depends on the concentration of the different food sources in the water, which is detailed in the next section 5.3.5.3. Triplicate measurements of the flow rate of fluorescein labelled seawater were made on ten newly collected *A. queenslandica* adults (Appendix 5.9), taking into account the number and size of the oscula per sponge as described in the Methods. The average flow rate of adult *A. queenslandica* at 24^{o}C was estimated to be 9.52 +/- 2.09 L / h / gDW.

5.4.2 Constraining oxygen availability

The oxygen concentration ranged from 1.63 – 11.82 mg / L, with an average of 6.24 mg / L (Appendix 5.10). The oxygen constraints were calculated from the average concentration. This was done by converting into mmol / L and then multiplying by the quantified pumping rate of 9.52 L / h / gDW to give 3.7127 mmol O_2 / h / gDW. This value sets the maximum oxygen consumption rate allowed by the model.

5.4.3 The central metabolic network of *A. queenslandica*

Here I have successfully reconstructed the first working metabolic model of a marine invertebrate. Notably, during the reconstruction, I identified ten amino acids, five vitamins and a phytosterol that *A. queenslandica* is unable to synthesise, but that are required for its growth (Table 5.1). These metabolites must be considered essential nutrients for the growth of this sponge species and represent a significant step towards understanding the nutritional requirements of a sponge. Below I describe core components of the central metabolic network of *A. queenslandica*, using Enzyme Commission (EC) numbers following common enzyme names. Corresponding *A. queenslandica* gene IDs are provided in the complete list of reactions (Appendix 5.11).

5.4.3.1 Amino acid synthesis pathways

Amino acids can be grouped according to the precursor molecule from which they are synthesised (Stephanopoulos et al., 1998). The metabolite 2-oxoglutarate is the precursor molecule for L-glutamate, L-glutamine, L-proline and L-arginine (Figure 5.4); of these, complete biosynthesis pathways are present in *A. queenslandica* for L-glutamate, L-glutamine and L-proline, but not for L-arginine. L-glutamate can be synthesised via two different pathways: (i) directly from 2-oxoglutarate and ammonia using the enzyme glutamate dehydrogenase (EC 1.4.1.3); and (ii) indirectly via the degradation of L-ornithine, initially to L-glutamate-5-semialdehyde by ornithine aminotransferase (EC 2.6.1.13) and then to L-glutamate by delta-1-pyrroline-5-decarboxylate dehydrogenase (EC 1.2.1.88). L-glutamine is synthesised by the enzyme glutamine synthase (EC 6.3.1.2), which is present and converting L-glutamate into L-glutamine.

The L-proline biosynthesis pathway is present in *A. queenslandica* (Figure 5.4). L-glutamate is converted to L-glutamate 5-phosphate by the enzyme glutamate 5-kinase (EC 2.7.2.11). L-glutamate 5-phosphate is then converted into L-glutamate-5-semialdehyde by glutamate-5-semialdehyde dehydrogenase (EC 1.2.1.41). L-glutamate-5-semialdehyde spontaneously rearranges to 1-pyrroline-5-carboxylate. This is then converted to L-proline by pyrroline-5-carboxylate reductase (EC 1.5.1.2).

A. queenslandica cannot synthesize L-arginine (Figure 5.4). L-arginine is an intermediate metabolite of the urea cycle, which is not present in *A. queenslandica*. The sponge is missing argininosuccinate lyase (EC 4.3.2.1) that converts L-arginino-succinate into L-arginine. This gap prevents the synthesis of L-arginine. Additionally, *A. queenslandica* does not have a complete L-arginine degradation pathway. L-arginine can be converted to nitric oxide via nitric oxide synthase (EC 1.14.13.39).

Table 5.1 Essential metabolites that must be derived from nutritional inputs or from the symbiotic microbial community.

Metabolite name	Class of metabolite	Biosynthetic pathway present in AqS1
L-arginine	Amino acid	Yes
L-histidine	Amino acid	Yes
L-isoleucine	Amino acid	Yes
L-leucine	Amino acid	Yes
L-lysine	Amino acid	Yes
L-methionine	Amino acid	Yes
L-phenylalanine	Amino acid	Yes
L-threonine	Amino acid	Yes
L-tryptophan	Amino acid	Yes
L-valine	Amino acid	Yes
Thiamine	Vitamin (B1)	Yes
Pantothenate	Vitamin (B5)	No
Pyridoxal	Vitamin (B6)	No
Retinol	Vitamin (A)	No
Folate	Vitamin (B9)	Yes
Brassicasterol	Phytosterol	No

Pyruvate is the precursor metabolite for the synthesis of L-alanine, L-leucine and L-valine. Of these, the sponge can synthesize only L-alanine (Figure 5.4), which it can do via two alternate pathways. One pathway converts L-glutamate and pyruvate to L-alanine, by the enzyme alanine transaminase (EC 2.6.1.2). The other pathway is performed by cysteine desulfurase (EC 2.8.1.7), which acts on L-cysteine to produce L-alanine in addition to a range of sulfur containing compounds.

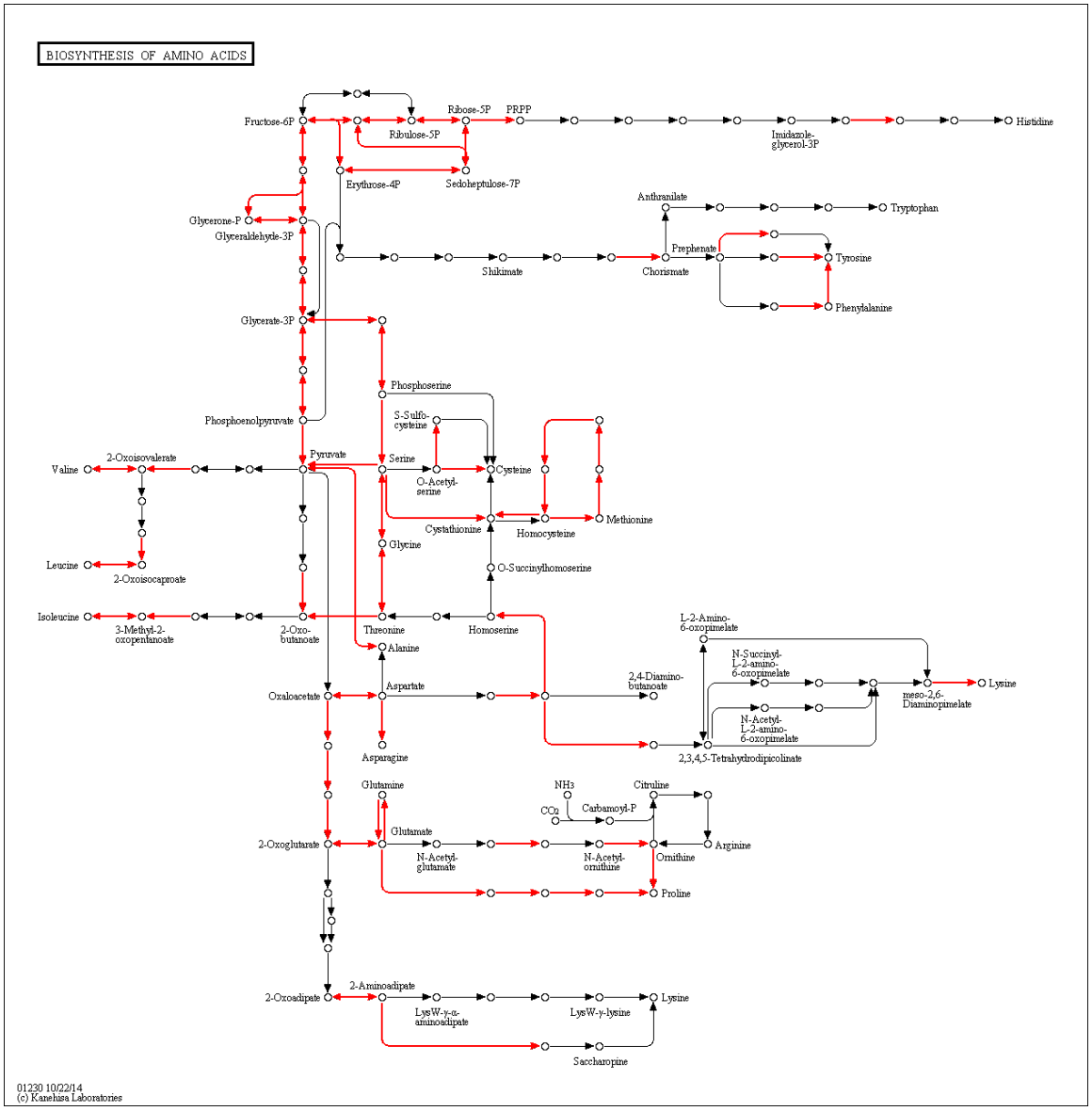


Figure 5.4 Overview of the amino acid synthesis pathways in *A. queenslandica*.

Lines in red represent reactions that are present, while black represents absent reactions. The amino acids that cannot be synthesized by *A. queenslandica* are L-arginine, L-histidine, L-isoleucine, L-leucine, L-lysine, L-methionine, L-phenylalanine, L-threonine, L-tryptophan and L-valine. Not every reaction in a pathway is represented. This figure was created with the online KEGG mapping tool (www.kegg.jp).

The TCA cycle metabolite oxaloacetate is the precursor for L-aspartate and L-asparagine, both of which *A. queenslandica* can produce (Figure 5.4). Oxaloacetate is converted into L-aspartate by aspartate transaminase (EC 2.6.1.1). L-aspartate is then converted into L-asparagine by either aspartate-ammonia ligase (EC 6.3.1.1) or asparagine synthase (EC 6.3.5.4), which respectively use

ammonia and glutamine as the amino group donor. L-aspartate itself is the precursor for L-methionine, L-cysteine, L-threonine, L-isoleucine and L-lysine; of these, *A. queenslandica* can synthesize only L-cysteine, and only then if L-methionine is sourced from the environment or bacterial community (Figure 5.4).

A. queenslandica is missing the Shikimate pathway, which means that it cannot individually synthesize the aromatic amino acids L-phenylalanine, L-tryptophan and L-tyrosine (Figure 5.4). However, the enzyme phenylalanine-4-hydroxylase (EC 1.14.16.1) is present, so *A. queenslandica* can produce L-tyrosine if L-phenylalanine is sourced from the environment or associated bacterial community. L-histidine does not group with any of the other amino acids but, like the shikimate pathway, most of the enzymes required for its synthesis were not identified in the *A. queenslandica* genome.

3-phosphoglycerate is the precursor for the synthesis of L-serine and L-glycine (Figure 5.4). The three enzymes required to synthesize L-serine from 3-phosphoglycerate are present in the *A. queenslandica* genome. Four different enzymes, each of which utilizes different substrates, can synthesise L-glycine. Alanine-glyoxylate transaminase (EC 2.6.1.44) produces L-glycine and pyruvate from glyoxylate and L-alanine. L-threonine aldolase (EC 4.1.2.5/4.1.2.48) converts L-threonine to L-glycine and acetaldehyde. Glycine dehydrogenase (EC 1.4.4.2) produces L-glycine and lipoylprotein from S-aminomethyldihydrolipoylprotein and carbon dioxide. Finally, hydroxymethyltransferase (EC 2.1.2.1) can interchange L-serine and L-glycine.

5.4.3.2 Central carbon metabolism

The central carbon metabolism of *A. queenslandica* contains all of the carbohydrate pathways common to eukaryotes; glycolysis, the TCA cycle, pentose phosphate pathway and C_1 metabolism reactions are all present (Figure 5.5). The glycolysis pathway is complete from glucose to pyruvate,

but *A. queenslandica* is lacking lactate dehydrogenase (EC 1.1.1.27), which converts pyruvate into lactate and NAD+ under anoxic conditions. The TCA cycle in *A. queenslandica* contains two isoforms of succinyl-CoA ligase. The first form uses ATP (EC 6.2.1.5) and is primarily used by both prokaryotes and eukaryotes (Johnson et al., 1998). The second form uses GTP (EC 6.2.1.4) as a cofactor, which is not used by prokaryotes (Johnson et al., 1998).

The pentose phosphate pathway in *A. queenslandica* contains the enzymes to perform both the oxidative and non-oxidative branches of the pathway. The oxidative branch produces reducing power in the form of NADPH with the enzymes glucose 6-phosphate dehydrogenase (EC 1.1.1.49) and phosphogluconate dehydrogenase (decarboxylating) (EC 1.1.1.351 and EC 1.1.1.44). Both of these reactions produce ribulose 5-phosphate, which feeds into the non-oxidative branch. This branch produces two precursor metabolites, ribose 5-phosphate and erythrose 4-phosphate. Ribose 5-phosphate is synthesized from ribulose 5-phosphate via ribose 5-phosphate isomerase (EC 5.3.1.6); this is subsequently converted to erythrose 4-phosphate by transketolase (EC 2.2.1.1).

5.4.3.3 Vitamin metabolism

The vitamins present in the *A. queenslandica* metabolic reconstruction are thiamine (B1), riboflavin (B2), niacin (B3), pantothenate (B5), pyridoxal 5-phosphate (active form of B6 group), folate (B9) and retinol (A). Of these, *A. queenslandica* is able to *de-novo* synthesize riboflavin (B2) and niacin (B3) (Table 5.1).

For model functionality, transport reactions were added to allow the uptake of the essential vitamins that cannot be synthesised by *A. queenslandica* (Appendix 5.11). This included thiamine, pantothenate, retinal and a representative of the B6 group. The vitamin B6 group contains the metabolites pyridoxol, pyridoxamine, pyridoxal and pyridoxal 5-phosphate. Pyridoxal 5-phosphate is the functional molecule that acts as the cofactor for many reactions involving amino acids and

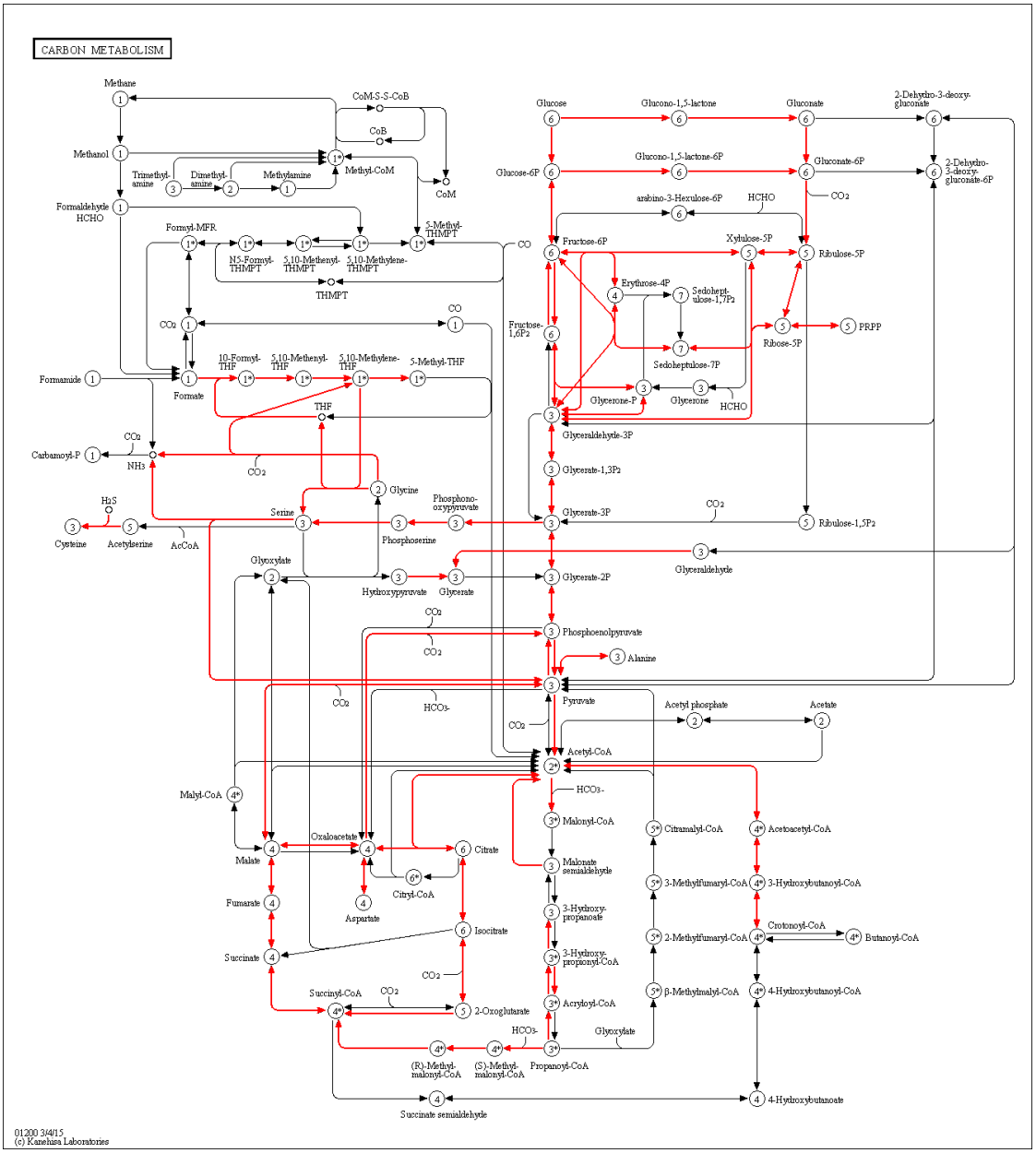


Figure 5.5 Central carbon metabolism of *A. queenslandica*.

This map is a generalized representation of carbon metabolism. The red lines indicate pathway steps that are present. Glycolysis runs vertically down the centre, from the pentose phosphate pathway to the TCA cycle. C1 metabolism is centred on tetrahydrofolate (THF). This figure was created with the online KEGG mapping tool (www.kegg.jp).

glycogen. *A. queenslandica* contains the enzymatic capabilities to convert the three non-functional forms – pyridoxol, pyridoxamine or pyridoxal – into the functional form. Subsequently, only one of the Vitamin B6 group was required as input for growth. This was arbitrarily chosen to be pyridoxol for which a transport reaction was added to *A. queenslandica*.

5.4.3.4 Lipid metabolism

The biochemical analysis of *A. queenslandica* identified twelve fatty acids, three of which are unsaturated (Chapter 3; Watson et al., 2014a). The *A. queenslandica* model can produce all twelve fatty acids, with acetyl-CoA and propionyl-CoA being the respective substrate molecules for the even and uneven chain length fatty acid pathways (Michal & Schomburg, 2013). Biosynthesis of the unsaturated fatty acids, 11-Eicosenoic acid (Icosenoic acid), erucic acid and nervonic acid, was enabled by added individual desaturase reactions acting on FA 20:0, 22:0 and 24:0.

Cholesterol and brassicasterol are the two most abundant sterols in *A. queenslandica* (Chapter 3; Watson et al., 2014a). In the draft model, two gaps were filled in the cholesterol synthesis pathway (Fig. 5.6). The first missing reaction converts 4,4-dimethyl-5alpha-cholesta-8,14,24-trien-3beta-ol to 14-demethyllanosterol (EC 1.3.1.70). This enzyme is possibly missing due to an annotation error as the gene was present in the original genome annotation (Srivastava et al., 2010). Also, the sequence for the corresponding gene in the re-annotation (Aqu2.1.42778_001) matches to lamin-B receptor (delta-14-sterol-reductase) when BLASTed against the NCBI database (ncbi.nlm.nih.gov). No putatively associated gene could be identified for the second gap-filled reaction required by *A. queenslandica* to synthesise cholesterol. This gap-filled reaction converts zymostenol to zymosterol (EC 1.3.1.72). Both of these reactions were manually added to the model.

The *A. queenslandica* metabolic reconstruction is unable to synthesize the phytosterol known as brassicasterol (Fig. 5.6), which was found to be the second most abundant sterol in the biomass composition analysis (Chapter 3, Watson et al., 2014a). There is currently no data to support that *A. queenslandica* can synthesise phytosterols, so a transport reaction was manually added to facilitate its uptake.

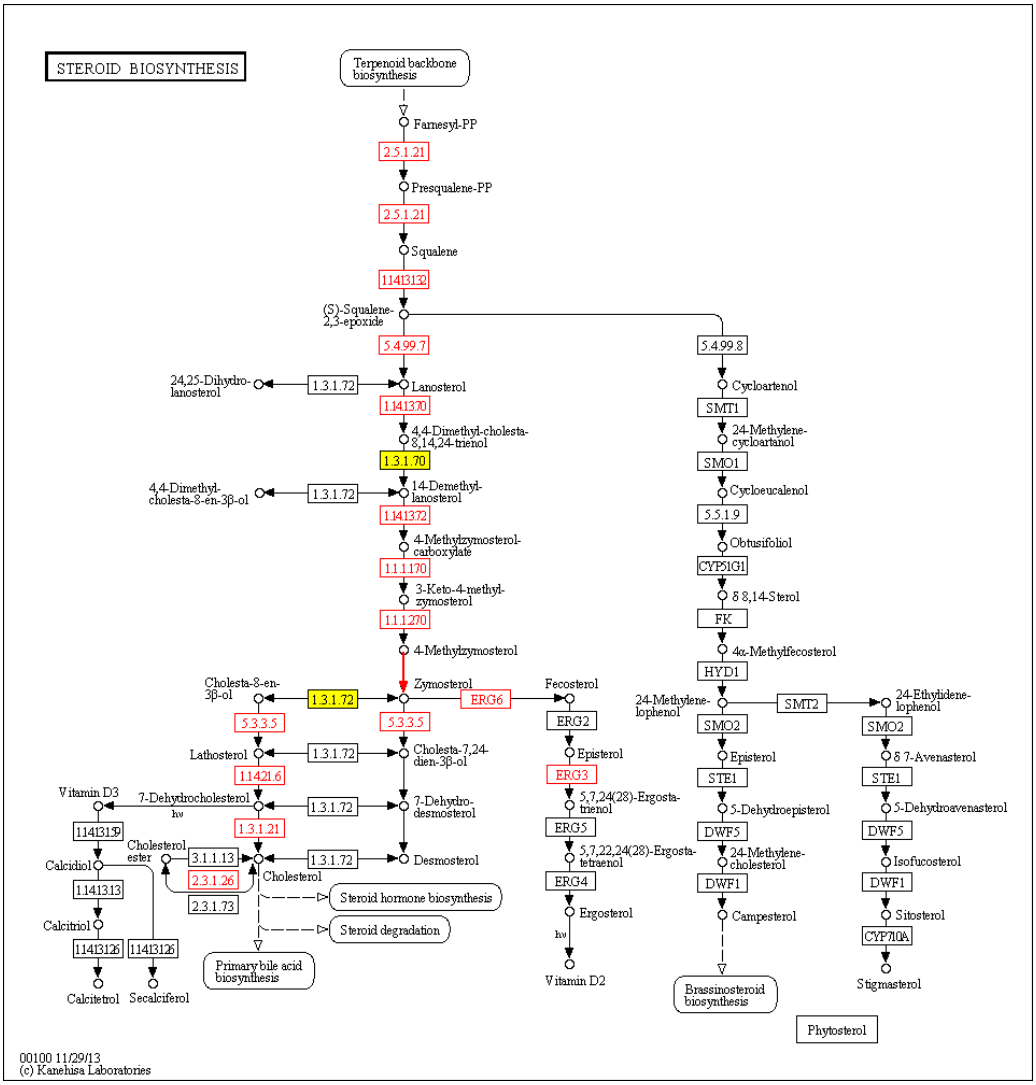


Figure 5.6 Steroid biosynthesis in *A. queenslandica*.

Red boxes and red arrow represent reactions for which enzymes were found in the *A. queenslandica* genome. The yellow boxes were gap filled to allow cholesterol synthesis. *A. queenslandica* cannot synthesize phytosterols, the branch to the right, which include brassicasterol that was abundant in the biomass composition analysis. This figure was created with the online KEGG mapping tool (www.kegg.jp).

5.4.3.5 Gap-filled and spontaneous reactions in the *A. queenslandica* model

This section summarises the reactions that were missing in the *A. queenslandica* metabolic network and were added to the model to allow the model to function. These reactions represent points

in the metabolic network that require further investigation to test their validity as it is possible that they may be real gaps, replaced with metabolic interactions with *AqS1*. Including reactions discussed above, the *A. queenslandica* metabolic reconstruction contains 13 reactions for which putative gene IDs could not be assigned. These gap-filled reactions, in addition to 10 spontaneous reactions (Table 5.2) are currently essential for the *A. queenslandica* metabolic model to produce biomass. As a comparison, the *Arabidopsis* model, AraGEM, contains 75 gap-filled reactions for which no putative gene has been assigned (de Oliveira Dal'Molin et al., 2010).

5.4.4 Characteristics of the *AqS1* metabolic model

AqS1 is a species of sulfur-oxidizing bacteria (SOB) (Gauthier et al., 2016). In general, SOBs are autotrophs, capable of oxidizing hydrogen and reduced sulfur compounds for energy, which is coupled to fixing carbon dioxide to synthesize organic compounds (Kleiner et al., 2012a; Kleiner et al., 2012b). Similar to a previously reported symbiotic SOB from a marine sponge (Tian et al., 2014), *AqS1* has enhanced heterotrophic capabilities when compared to free living SOB (Gauthier et al., 2016). This is evidenced by increased abundance of transporters and metabolic capabilities to use a wide range of carbohydrates that would be available from the environment or from the host sponge biomass (Gauthier et al., 2016). As no *AqS1*-specific biochemical data is available, the *AqS1* metabolic reconstruction was gap-filled using biochemical data from related species of SOBs. Five reactions were gap-filled for which no putative gene could be assigned. For more details on *AqS1* refer to Gauthier et al. (2016). This section focuses on the pathways and gap filling that were relevant to reconstructing a functional metabolic model that can produce all biomass metabolites.

5.4.4.1 Chemoautotrophic metabolism

AqS1 contains an incomplete Sox system that is missing SoxCD. This is the same as free-living, symbiotic (Kleiner et al., 2012a; Muyzer et al., 2011a,b) and phototrophic SOB (Friedrich et al., 2005; Ogawa et al., 2008). The complete Sox system yields 8 electrons per thiosulfate molecule, but the loss

Table 5.2 Gap filled and spontaneous reactions in the *A. queenslandica* metabolic reconstruction.

rxnID	Pathway	Spontaneous	Enzyme	Reaction	Rev
rxn01759	Cysteine and Methionine metabolism	No	3'-Phosphoadenylylsulfate:2-Aminoacrylate sulfotransferase	3-phosphoadenylylsulfate + Dehydroalanine + 2 H+ = Adenosine 3-5-bisphosphate + L-Cysteate	1
R10409	Cysteine and Methionine metabolism	No		S-sulfocysteine + Glutaredoxin = L-Cysteine + Sulfite + Glutaredoxin disulfide	1
rxn01208	Valine, Leucine and Isoleucine biosynthesis	No	2-Oxo-4-methyl-3-carboxypentanoate decarboxylation	4MOP + CO2 = 2-isopropyl-3-oxosuccinate + H+	-1
rxn01662	Lysine degradation	No	N6-(L-1,3-Dicarboxypropyl)-L-lysine:NAD+ oxidoreductase	H2O + NAD + Saccharopine = 2-Aminoadipate 6-semialdehyde + H+ + L-Glutamate + NADH	0
rxn00484	Tryptophan metabolism	No	L-tryptophan decarboxy-lyase	H+ + L-Tryptophan = CO2 + Tryptamine	1
rxn03360	Tryptophan metabolism	No	5-Hydroxykynurenamine decarboxy-lyase	5-Hydroxykynurenamine + CO2 = 5-Hydroxykynurenine + H+	-1
rxn01948	Tryptophan metabolism	No	5-Hydroxy-L-tryptophan decarboxy-lyase	5-Hydroxy-L-tryptophan + H+ = CO2 + Serotonin	1
SP2	NAD synthesis	No	Nicotinate-nucleotide adenylyltransferase	Nicotinate ribonucleotide + ATP + H+ = Deamido-NAD + PPi	1

rxnID	Pathway	Spontaneous	Enzyme	Reaction	Rev
rxn01465	Pyrimidine metabolism	No	(S)-Dihydroorotate amido-hydrolase	H2O + S-Dihydroorotate = H+ + N-Carbamoyl-L-aspartate	0
rxn03911	Steroid biosynthesis	No	4,4-dimethyl-5a-cholesta-8,24-dien-3b-ol:NADP+ D14-oxidoreductase	14-Demethyllanosterol + NADP = 4,4-Dimethyl-5alpha-cholesta-8,14,24-trien-3beta-ol + H+ + NADPH	0
rxn07327	Steroid biosynthesis	No		4-methylzymosterol = Zymosterol	0
rxn03007	Steroid biosynthesis	No		NADP + Zymostenol = H+ + NADPH + Zymosterol	-1
rxn01466	Terpenoid backbone biosynthesis	No	Geranyl-diphosphate:isopentenyl-diphosphate geranyltrans-transferase	Geranyldiphosphate + Isopentenyldiphosphate = Farnesyldiphosphate + PPi	1
rxn02374	Arginine and Proline metabolism	spontaneous	L-glutamate 5-semialdehyde dehydratase	L-Glutamate5-semialdehyde = 1-Pyrroline-5-carboxylate + H+ + H2O	0
rxn05129	Arginine and Proline metabolism	spontaneous	Non-enzymatic	Phosphocreatine = Creatinine + Phosphate	1
rxn03070	Arginine and Proline metabolism	spontaneous	L-1-Pyrroline-3-hydroxy-5-carboxylate spontaneous conversion to L-4-Hydroxyglutamate semialdehyde, mitochondrial	3-Hydroxy-L-1-pyrroline-5-carboxylate + H+ + H2O = L-4-Hydroxyglutamatesemialdehyde	-1

rxnID	Pathway	Spontaneous	Enzyme	Reaction	Rev
R04861	Cysteine and Methionine metabolism	spontaneous	Non-enzymatic	3-Sulfinopyruvate + H2O = Sulfite + Pyruvate	1
rxn11597	Histidine metabolism	spontaneous	Non-enzymatic	4-Imidazolone-5-propanoate + A + 3 H2O = 4-Oxoglutaramate + AH2 + Formate + NH3	0
rxn00201	Histidine metabolism	spontaneous	Non-enzymatic	2-Oxoglutarate + NH3 = 4-Oxoglutaramate + H2O	-1
rxn11972	Tyrosine metabolism	spontaneous	Dopaquinone = H+ + Leucodopachrome	1	
rxn02609	Tyrosine metabolism	spontaneous	L-Dopa + L-Dopachrome = Dopaquinone + Leucodopachrome	1	
rxn02611	Tyrosine metabolism	spontaneous	L-Dopachrome:oxygen oxidoreductase	H+ + L-Dopachrome = CO2 + DHI	1
SP1	NAD synthesis	spontaneous	2-Amino-3-carboxymuconate semialdehyde = Quinolinate + H2O	1	

Rev denotes the direction in which the reaction may proceed, with 1 being forward (left to right as written), 0 being a reversible reaction and -1 indicating the reaction may only proceed in the right to left direction.

of the soxCD enzyme results in just 2 electrons being generated and the formation of sulfur globules (Fazzini et al., 2013). As *AqS1* likely belongs to the family Ectothiorhodospiraceae (Gauthier et al., 2016), it is characteristic that the sulfur globules are exported from the cell (Garrity et al., 2007). In the metabolic reconstruction, the Sox system is represented by a single reaction.

When sulfur compounds are limiting, *AqS1* can import taurine as a sulfur source by oxidizing it to sulfite and aminoacetaldehyde, using the enzyme alpha-ketoglutarate-dependent taurine dioxygenase (EC 1.14.11.17). Taurine is found in high abundance in the tissues of many marine invertebrates (Carr et al., 1996), notably in invertebrates that host symbiotic SOB (Joyner et al., 2003). *A. queenslandica* can produce taurine and is a potential source of sulfur for *AqS1*.

In addition to reduced sulfur compounds, the *AqS1* metabolic model is capable of oxidizing hydrogen (H_2) as a source of electrons. The electrons enter the electron transport chain at the level of quinone (Friedrich & Schwartz, 1993). Formate can be synthesised and exported by the *A. queenslandica* metabolic network, potentially being used by *AqS1*, which encodes the enzyme formate dehydrogenase (EC 1.2.1.2). Formate dehydrogenase oxidizes formate to carbon dioxide, generating reducing power in the form of NADH.

In SOBs, the oxidation of reduced compounds is linked to fixing CO2 via the Calvin-Benson-Bassham Cycle. As predicted by Pathway Tools, *AqS1* may also be able to fix CO2 by the reductive acetyl-coenzyme A pathway. Carbon monoxide dehydrogenase (EC 1.2.7.4), which converts carbon dioxide to carbon monoxide, is present in the *AqS1* genome. Other reactions identified in the *AqS1* genome that are unique to reductive acetyl-coenzyme A pathway are methylenetetrahydrofolate reductase (ferredoxin) (EC 1.5.7.1), 5-methyltetrahydrofolate:corrinoid/iron-sulfur protein (EC 2.1.1.258) and CO-methylating acetyl-CoA synthase (EC 2.3.1.169). Further, a BLAST search of the *AqS1* draft genome on The SEED Viewer server (rast.nmpdr.org/seedviewer.cgi) resulted in methylenetetrahydrofolate reductase (ferredoxin) (EC 1.5.7.1) being identified (peg.2667). The final

two enzymes in the reductive acetyl-coenzyme A pathway, 5-methyltetrahydrofolate:corrinoid/iron-sulfur protein (EC 2.1.1.258) and CO-methylating acetyl-CoA synthase (EC 2.3.1.169), do appear to be missing from the draft genome. With two sequential enzymes missing, the reductive acetyl-coenzyme A pathway was not included in the model. The presence of two reactions unique to this pathway suggest that this could be a novel carbon fixation pathway for SOBs that warrants further investigation that was outside the scope of this study.

5.4.4.2 Amino acid metabolism

The inability of the *AqS1* model to individually synthesise all 20 proteinogenic amino acids from simple carbon (D-glucose) and nitrogen (ammonia) sources revealed several incomplete pathways. Manual inspection revealed ten enzymes that were missing from the amino acid biosynthesis pathways in the draft model, six of which were able to have putative genes assigned to them using The SEED Viewer server (rast.nmpdr.org/seedviewer.cgi). The four reactions that were missing and could not have putative genes assigned were gap filled as they were single missing reactions in a pathway. The gap filled reactions were in the L-arginine, L-asparagine, L-cysteine and L-lysine biosynthesis pathways (Table 5.3).

5.4.4.3 Vitamin and cofactor metabolism

The synthesis pathway for menaquinone (Vitamin K) is nearly complete. This pathway involves two reactions that convert octaprenyl-diphosphate, a product of the terpenoid backbone pathway, to menaquinone. Demethylmenaquinone methyltransferase (EC 2.1.1.163), required for the first reaction, is present. The other reaction involved in the pathway, 1,4-dihydroxy-2-naphthoate polyprenyltransferase (EC 2.5.1.74), was gap-filled.

Table 5.3 Reactions that were gap filled and genes manually assigned in the amino acid pathways of the *AqS1* metabolic reconstruction.

Amino acid pathway	geneID	EnzymeName	EC number	Reaction
L-arginine	Unknown	N2-Acetyl-L-ornithine:2-oxoglutarate aminotransferase	2.6.1.11	2-Oxoglutarate + N-Acetylornithine = 2-Acetamido-5-oxopentanoate + L-Glutamate
L-asparagine	Unknown	L-aspartate:L-glutamine amido-ligase (AMP-forming)	6.3.5.4	H2O_thio + ATP_thio + L-Aspartate_thio + L-Glutamine_thio = PPi_thio + AMP_thio + L-Glutamate_thio + L-Asparagine_thio
L-cysteine	Unknown	serine O-acetyltransferase	2.3.1.30	Acetyl-CoA + L-Serine = CoA + O-Acetyl-L-serine
L-lysine	Unknown	N-Succinyl-L-2,6-diaminoheptanedioate:2-ocoglytarate	2.6.1.17	2-Oxoglutarate + N-Succinyl-L-2,6-diaminopimelate = L-Glutamate + N-Succinyl-L-2-amino-6-oxopimelate
L-lysine	peg.1539 peg.2751 peg.1286	4-hydroxy-tetrahydrodipicolinate synthase	4.3.3.7	Pyruvate + L-Aspartate4-semialdehyde = (2S,4S)-4-Hydroxy-2,3,4,5-tetrahydrodipicolinate + H2O + H+
L-lysine	peg.744	4-hydroxy-tetrahydrodipicolinate reductase	1.17.1.8	(2S,4S)-4-Hydroxy-2,3,4,5-tetrahydrodipicolinate + NADPH + H+ = tetrahydrodipicolinate + NADP + H2O
L-methionine	peg.2875 peg.4009 peg.1983	O-Succinyl-L-homoserine succinate-lyase (adding cysteine)	2.5.1.48	Cystathionine + Succinate = L-Cysteine + O-Succinyl-L-homoserine
L-methionine	peg.909	L-Homoserine:NAD+ oxidoreductase	1.1.1.3	L-Homoserine + NAD = H+ + L-Aspartate4-semialdehyde + NADH
L-methionine	peg.909	L-Homoserine:NADP+ oxidoreductase	1.1.1.3	L-Homoserine + NADP = H+ + L-Aspartate4-semialdehyde + NADPH
L-phenylalanine	peg.3158	Chorismate pyruvatemutase	5.4.99.5	Chorismate = Prephenate

The geneID refers to the *AqS1* database on The SEED Viewer server (rast.nmpdr.org/seedviewer.cgi).

AqS1 is unable to synthesize the essential cofactor NAD *de novo* (Fig. 5.7) because it is lacking the enzymes that convert L-aspartate, the substrate molecule for NAD biosynthesis in prokaryotes, to nicotinate D-ribonucleoside. NAD can, however, be synthesised if the precursor vitamin niacin is supplied; this could represent a reliance on the host sponge *A. queenslandica*, which is able to

synthesize both NAD and niacin. The precursor for NAD biosynthesis in eukaryotes is L-tryptophan, not L-aspartate as in prokaryotes, and *A. queenslandica* is unable to synthesize L-tryptophan, although *AqS1* can do so. This could form a feed back loop where *AqS1* is able to synthesise L-tryptophan for *A. queenslandica*, which in turn synthesises NAD for *AqS1*.

5.4.5 Metabolic simulations of the dual-species model

With the metabolic networks for *A. queenslandica* and *AqS1* both reconstructed, they were joined to form a dual-species GEM. The dual-species metabolic network consists of 2510 reactions, including biomass and transport reactions. There are 1700 unique metabolites across 3 compartments that represent *A. queenslandica*, *AqS1* and a shared compartment for metabolite exchange. This section details the analysis of initial FBA simulations of the dual-species network under realistic, high and low nutrient conditions. First the model is validated by comparing the amount of ATP required for growth and pumping water with the maximum rate of ATP synthesis. The low (October) and high (March) nutrient constraints were calculated using data from the environmental characterisation of Shark Bay that I reported in Chapter 4 (Watson et al., in revison). The constraints, objective function and resulting fluxes for each set of simulations are detailed in Appendix 5.12..

5.4.5.1 Validation of the dual-species metabolic model

The first step towards validating the mixed-species metabolic model was to compare the maximum rate at which ATP could be produced, with the estimated ATP requirements for pumping (see section 5.3.3.5 above), growth and maintenance. This would show if the theoretical model could survive under the realistic nutritional constraints. To calculate the amount of ATP required for growth, the holobiont biomass reaction was set as the objective of the simulation and FBA was performed under both nutrient conditions. This resulted with a flux of 0.07 mmol / h / gDW (low nutrient) and 0.26 mmol / h / gDW (high nutrient) through the ATP synthase reaction in *A. queenslandica* (Table 5.4). Setting the *A. queenslandica* ATP synthase reaction as the objective function revealed the

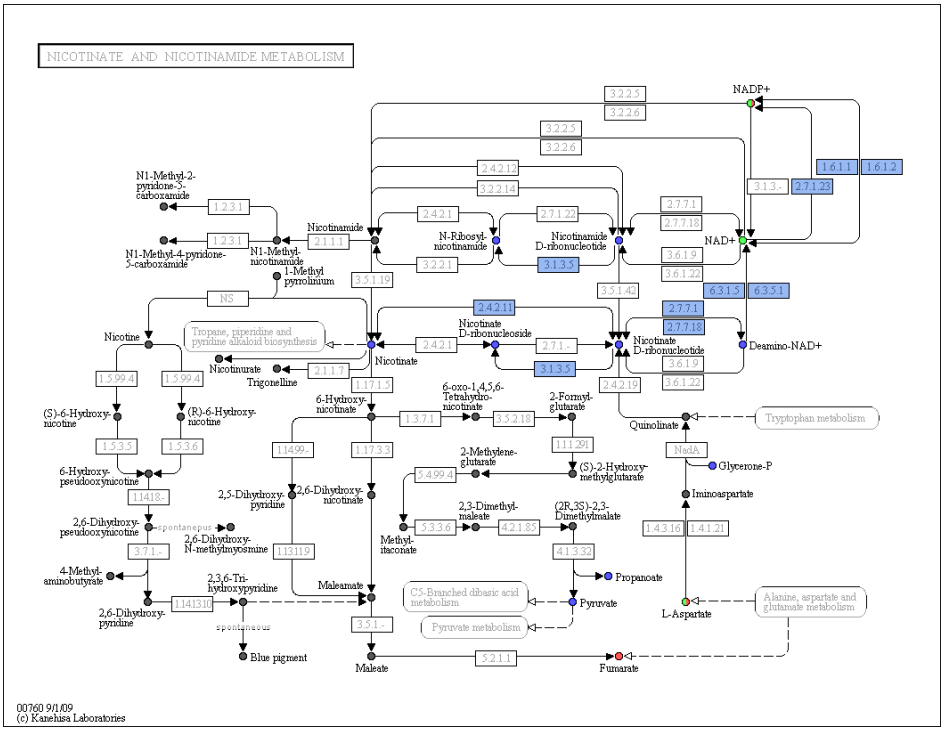


Figure 5.7 Nicotinate and nicotinamine metabolic map of *AqS1*.

De novo synthesis of NAD in bacteria uses L-aspartate as the precursor molecule. Present enzymes are coloured in blue. The pathway from L-aspartate to nicotinate-D-ribonucleotide is missing. NAD can be synthesized if the vitamin niacin (nicotinate) is supplied. This figure was created with the online KEGG mapping tool (www.kegg.jp).

maximum rates of ATP synthesis to be 37.63 (low nutrient) and 46.88 (high nutrient) mmol / h / gDW of ATP. These values translate to the model being able to produce a surplus of ATP, easily meeting the energetic demands for growth, cell maintenance and pumping (Table 5.4).

5.4.5.2 Simulating biomass synthesis under low and high nutrient conditions

How the dual-species GEM is able to synthesise biomass under low and high nutrient conditions was simulated using FBA. The resulting fluxes from FBA are representative of an optimum growth state under the respective nutrient conditions. Flux values represent the theoretical flow of mass through the metabolic network, with high flux values indicating a greater flow of mass and importance. Low fluxes indicate reactions of less importance, and reactions that do not carry any flux are non-

essential in achieving the simulations objective. The complete flux analysis results are detailed in Appendix 5.12.

With the synthesis of biomass defined as the objective, the dual-species network consumed all of the available bacterial and algal food sources in both of the nutrient simulations. Only a small amount of the available coral mucus was taken up, with a decrease from 6 % in the low nutrient to 3 % in the high nutrient simulation. This suggests that overall, the majority of the nutrients and energy that *A. queenslandica* requires can be sourced from the particulate food sources. The flux through the total biomass reaction (the overall growth rate) for the low nutrient simulation was only 25 % of the high nutrient simulation (Table 5.5). The rate of oxygen consumption by the dual-species model was higher in the high nutrient simulation (0.056 mmol /h /gDW) compared to the low nutrient simulation (0.022 mmol /h / gDW), although both of these oxygen consumption rates are small when compared to the maximum allowed rate of oxygen consumption (3.7127 mmol / h /gDW; section 5.4.2).

The increased growth rate predicted in the high nutrient simulation had a corresponding increase in the rate of carbon dioxide export out of the model; the low and high nutrients respectively resulting in 0.034 and 0.066 mmol / h / gDW. In both nutrient simulations the model excreted ammonia, with the export rate being twice as high in the high nutrient simulation than the low (0.021 and 0.044 mmol / h / gDW). In the low nutrient simulation, excess sulfur was exported at the model boundary as sulfate (0.002 mmol / h / gDW). The export of sulfate halved in the high nutrient simulation, but was replaced by an equal amount of sulfite export (both 0.001 mmol / h / gDW). Also, a small amount of hydrogen sulphide was exported from the model in the high nutrient simulation.

Phosphorous was only exchanged at the model boundary in the high nutrient simulation. Phosphate was consumed (0.0003 mmol / h / gDW) and di-phosphate exported (0.00009 mmol / h / gDW). The low nutrient simulation predicted the export of all amino acids apart from L-aspartate, L-glutamate and L-serine. The high nutrient simulation predicted the export of all amino acids apart

Table 5.4 Estimated ATP costs for pumping and growth compared to the maximum rate of ATP synthesis by the model. Nutrient condition data taken from Chapter 4 (Watson et al., in review).

	Low nutrient conditions	High nutrient conditions
Estimated ATP cost of pumping (mmol / h / gDW)	20.82	20.82
Estimated ATP cost for maintenance (mmol / h / gDW)	2.33	2.33
Flux through ATP synthase for maximum growth rate in A. queenslandica compartment (mmol / h / gDW)	0.07	0.26
ATP for max growth, maintenance and pumping (mmol / h / gDW)	23.22	23.41
Max ATP generation by model (mmol / h / gDW)	37.63	46.88
Amount of extra ATP that can be produced (mmol / h / gDW)	14.41	23.47
Percentage of max ATP required for growth, maintenance and pumping (%)	61.70	49.93
Flux through the holobiont biomass reaction	0.0027	0.0106

from L-aspartate, L-glutamine and L-glutamate. Most of the amino acids that were exported from the model under both nutrient simulations showed increased rates in the high nutrient condition, the exception was L-glutamine (Appendix 5.12).

Of the organic acids, pyruvate (0.01 mmol / h / gDW), formate (0.00006 mmol / h / gDW) and acetate (0.002 mmol / h / gDW) were exported from the model in the low nutrient simulation. In the high nutrient simulation the export of pyruvate, formate and acetate was greatly reduced to 0.0001,

4e-13 and 0.0009 mmol / h / gDW. Succinate was also exported in the high nutrient simulation (0.034 mmol / h / gDW), but not in the low nutrient simulation.

The host *A. queenslandica* is initially able to access all of the proteinogenic amino acids and carbon sources from the phytoplankton and bacterioplankton food sources. Even though *AqS1* is able to synthesise all of the proteinogenic amino acids, FBA predicted that L-methionine, L-histidine, L-lysine, L-threonine, L-tryptophan, L-leucine, L-proline and L-valine were exported from *A. queenslandica* and into *AqS1* in both nutrient simulations. L-arginine and L-aspartate were also predicted to be exchanged from *A. queenslandica* to *AqS1*, but only in the low nutrient simulation. L-isoleucine was predicted to be exported by *A. queenslandica* and imported by *AqS1* only in the high nutrient simulation. *A. queenslandica* exported ammonia in both nutrient conditions, some of which was consumed by *AqS1* only in the high nutrient simulation. The organic acid, succinate was predicted to be exchanged from *A. queenslandica* to *AqS1* during in both nutrient simulations. Fumarate was exported by *A. queenslandica* to *AqS1* only during the low nutrient simulation. A small amount of hydrogen sulfide (H2S) was produced by *A. queenslandica* and taken up by *AqS1* in the low nutrient simulation. *AqS1* also imported H2S during the high nutrient simulation, but it was not produced by *A. queenslandica* and instead imported from outside the system boundary (the environment).

In addition to the amino acids exported from *A. queenslandica* and used by *AqS1*, excess L-alanine, L-asparagine, L-cysteine, L-glycine, L-phenylalanine and L-tyrosine were exported by *A. queenslandica* and out of the model in both nutrient simulations (Appendix 5.12). L-glutamine was exported only in the low nutrient simulation and L-serine only in the high nutrient simulation. L-glutamate was not exported from the *A. queenslandica* compartment in either of the nutrient conditions simulated, but can be synthesised from 2-oxoglutarate (TCA cycle intermediate) and ammonia. This suggests that the current model is not limited by the amount of available nitrogenous metabolites. The FBA simulations predicted that *A. queenslandica* exports acetate in low nutrient conditions and pyruvate in high nutrient conditions.

The fluxes through the *A. queenslandica* metabolic network during the low nutrient conditions were largest through ATP synthase, L-alanine – L-aspartate – L-glutamate metabolism, the TCA cycle, fatty acid degradation and elongation (Appendix 5.12). The predicted fluxes of metabolites through these reactions and pathways were smaller in the low nutrient simulation than the high nutrient simulation. Exceptions where the high nutrient simulation had smaller fluxes than the low nutrient simulation include pyrimidine metabolism, pyruvate metabolism and cysteine and methionine metabolism. If a metabolite was a component of *A. queenslandica's* biomass, and was not available in the algal or bacterial food compositions, then the fluxes through the respective biosynthetic pathways were larger in the high nutrient conditions where overall growth was greater. Examples of this were the unsaturated fatty acids and sterols (Appendix 5.12). Nearly all amino acid synthesis pathways showed no flux in either nutrient simulation.

AqS1 required fewer reactions to actively carry flux to synthesise its biomass components. In the low nutrient simulation 238 reactions were active, while only 149 reactions were required in the high nutrient simulation to synthesise biomass. In the low nutrient simulation, the reactions that showed the larger fluxes were associated with ATP synthesis, aromatic amino acid synthesis, glycolysis, C_1 metabolism and the TCA cycle (Appendix 5.12). Similar to the low nutrient conditions, the reactions that carried flux in the high nutrient simulation were associated with ATP synthesis, glycolysis, the TCA cycle and aromatic amino acid synthesis. In the high nutrient simulation *AqS1* was predicted to synthesise and store 3-hydroxybutanoate (PHB), which is a storage carbohydrate that many bacteria synthesise when excess nutrients are present (Mauchline et al., 1999; Uchino et al., 2007).

5.5 Discussion

The overall aim of this chapter was to develop a genome-scale metabolic model that can convert realistic food sources into *Amphimedon queenslandica* and *AqS1* biomass. This was achieved, with the model being able to convert DOC, phytoplankton and bacterioplankton into the metabolites and

energy required to synthesise the biomass components. The development of this resource can now guide future work to improve the understanding coral reef nutrient cycling, as well as to assist the development of sponge and symbiont cell cultures. The development of a genome-scale model is a continual process (O'Brien et al., 2015; Thiele & Palsson, 2010). Novel biochemical pathways can be incorporated, biomass components expanded upon, annotations improved and biochemical data added. As such, the model presented here is the baseline for further development and can serve as a guide for creating GEMs for other sponge species. It is important to note that a simulation is a theoretical representation of a possible situation and not assumed to be perfect; it represents a possible metabolic phenotype that is defined not only by the environmental constraints placed upon it, but the reactions that it contains. Subsequently, the gap-filled and added reactions represent important points to focus rational improvement towards understanding sponge-symbiont metabolism.

5.5.1 The metabolic networks of *A. queenslandica* and *AqS1*

This reconstruction of the central metabolic network of *A. queenslandica* represents the first detailed analysis of the biosynthetic pathways that produce important building block metabolites in a marine sponge. Reconstructing the amino acid synthesis pathways in *A. queenslandica* revealed that it has the same essential dietary amino acid requirements as other metazoans (section 5.4.3) (Payne & Loomis, 2006; Rose et al., 2003). This has implications for sponge cell culture where commercial mammalian cell culture media has been used (De Rosa et al., 2003; GarciaCamacho et al., 2006; Grasela et al., 2011; Zhang et al., 2008; Zhao et al., 2008). The inclusion off the essential amino acids for *A. queenslandica* in mammalian cell culture media indicates that the lack of developing a primary culture is not due to an amino acid deficiency and other metabolites may be essential nutrients.

The process of reconstructing the network and subsequent FBA, identified what metabolites the network could produce from simple carbon and nitrogen sources and in the process identified the downstream gaps that blocked the synthesis of particular metabolites. Some of these gaps, or essential nutrients, could have been missed if a pathway map was simply visually inspected. It was previously

reported that *A. queenslandica* was auxotrophic for only L-lysine, L-histidine and L-tryptophan (Fiore et al., 2015), an assessment made by visually inspecting the *A. queenslandica* KEGG pathway maps (http://www.genome.jp/kegg-bin/show_organism?org=aqu). Gap filling a reaction, to enable network functionality (growth) or a known biochemical process to occur, is common practise in metabolic modelling (Latendresse et al., 2012; Kumar et al., 2007). There were fourteen reactions that were gap filled in the *A. queenslandica* model that could not have putative gene IDs assigned (Table 5.2). These fourteen gap filled reactions present the first hypothesis that should be investigated *in vitro*, which is typically performed using stable carbon isotope labelling (Sauer, 2006; Wittmann, 2007; Zamboni et al., 2009). This will prove challenging as the isolation of sponge cells from their associated prokaryotic community would be required to ensure that any biomass labelled was due to only the sponges metabolism.

As with gap filling the metabolic pathways within each organism's sub-network, the metabolites that may be exchanged between *A. queenslandica* and *AqS1* in the dual-species model is at a preliminary stage of investigation. Metabolite exchange is a key aspect of symbiosis in marine sponges (Hentschel et al., 2012) that is inherently difficult to study experimentally due to the complexity of sponge associated microbial communities (Webster & Blackall, 2008). The limited sequence data available for transport reactions, in particular export reactions, limits a genomic approach to investigating metabolite exchange. Transport reactions are not as well annotated as intracellular enzymes, with many transport reactions being annotated with generic terms such as amino acid transporter. The dual-species model presented here serves as a baseline to simulate potential metabolic interactions between *A. queenslandica* and *AqS1*. Being able to assess the effect of adding or removing reactions to a network, and predicting the change in functionality, is a fundamental aspect of genome-scale metabolic models (Bordbar et al., 2014; Esvelt & Wang, 2013; O'Brien et al., 2015; Sengupta et al., 2013), but was outside the scope of this thesis.

5.5.2 Validation of the dual-species model

The maximum rate of ATP synthesis was compared with the estimated requirements for pumping and growth as a method to validate the model in the absence of growth rates or other quantitative biochemical data; simply put, can the theoretical metabolic model survive under the environmental conditions imposed on it? The dual-species model was able to produce more than the amount of ATP required for pumping and growth under both low and high nutrient conditions that were simulated. This is an important first indication that the predictions of the dual-species model are realistic.

One limitation of the simulations is that the environmental characterisation of Shark Bay (Chapter 4; Watson et al., in revision) did not quantify the availability of reduced sulfur compounds to the *A. queenslandica* holobiont. This could be important, because the dominant bacterial symbiont *AqS1* is able to oxidise reduced sulfur compounds as an energy source, linked to the fixation of carbon from CO_2 (Gauthier et al., 2016). The nutrient availability represented average values that are experienced by *A. queenslandica* during different seasons, while the oxygen availability and pumping rate were derived from a single, different, time-point. Concurrently quantifying these values may reveal when there is a dependence of *A. queenslandica* on *AqS1* for the supply of energy (carbon) rich compounds derived from carbon fixation. As a follow up study, a sensitivity analysis can now be performed by modelling the upper and lower boundaries of oxygen and ATP production. This will enable the limits of the system to be explored.

5.5.3 Flux balance analysis of biomass synthesis

The model predicted a clear difference in the ability of the metabolic network to produce biomass under the low and high nutrient conditions. The greater availability of particulate food sources during the high nutrient simulation resulted in 75% more biomass being synthesised than the low nutrient condition. In general growth rates of sponges under natural conditions are poorly studied, with most of what is known coming from aquaculture attempts to supply secondary metabolites (for reviews see

(Belarbi, 2003; Duckworth, 2009; Mueller et al., 2004; Osinga et al., 1999; Schippers et al., 2012; Sipkema et al., 2005)). Likewise, the growth rate of *A. queenslandica* has not been quantified but the FBA results clearly indicate that much higher growth rates are possible towards the end of summer and periods of increased particulate food. *A. queenslandica* is reproductive through-out the year, with a peak in reproductive output that coincides with this increased capacity to synthesise biomass (Leys et al., 2008). Incorporating into the model the biomass drain of large daily larval releases may help explain the nutrients that were exported from the model, possibly to satisfy the steady state assumption (Fig 5.1).

Under low nutrient conditions, nitrogen for amino acid synthesis is scarce (Hatcher, 2003; Rix et al., 2016), yet the mixed-species model exported ammonia, amino acids and nucleotides in both nutrient simulations (Appendix 5.12). The export of excess nitrogenous compounds from the overall network, while carbon sources (such as coral mucus) and oxygen are not being utilised, indicates that the model may be missing a cofactor or metabolite that is restricting the resulting the metabolic network to use available nitrogen. Alternatively, it might be a nutrient that is limiting the full use of the available nitrogenous compounds. If the export of nitrogenous compounds occurs *in situ*, this could be an important insight where sponges are vital for cycling organic nitrogenous compounds between the particulate and dissolved fractions of tropical coral reef waters. This is an interesting concept that would complement the fixation and release of carbon compounds by photosynthetic organisms (de Goeij et al., 2013; Hatcher, 2003; Rix et al., 2016). To my knowledge, the exhalent water of any sponge has not been analysed specifically at free amino acids or nucleotides, but fourteen species of sponges on Caribbean reefs have been shown to be a source of dissolved nitrogen (ammonia, nitrite and nitrate) under certain environmental conditions (Fiore et al., 2013; Southwell et al., 2008).

A. queenslandica required thiamine, pantothenate, pyridoxal and retinol to produce several biomass metabolites during the gap-filling phase of model development. There was no uptake flux of these vitamins under both nutrient conditions, indicating that the biomass components that require thiamine, pantotheonate, pyridoxal or retinal can be readily sourced from the diet or from *AqS1*.

Data on vitamins in sponges is scarce although *X. muta* possibly obtains the vitamins riboflavin and thiamine from its microbial community. *X. muta* was shown to express catabolic pathway enzymes for riboflavin and thiamine, while the prokaryotic metatranscriptome contained transcripts of the biosynthetic pathways (Fiore et al., 2015).

It is important to note that only a single objective function, either biomass synthesis or ATP production, was defined in each FBA simulation. Realistically, there are multiple objectives towards which an organism must allocate its available resources (O'Brien et al., 2015). These objectives change over the short and long terms, as do the availability of nutrient and oxygen conditions. Running separate FBA and simply changing the objective function of the simulations did not capture this trade off between the cost of pumping and growth. Methods to perform FBA with multiple objectives have been developed (Oh et al., 2009; Sengupta et al., 2013) and can be applied to this model in future work.

The FBA predicted that the algal and bacterioplankton food sources were important for maximum rates of biomass synthesis under both nutrient conditions. Recently *X. muta* has been shown to selectively feed on different food sources to optimise the nutritional gain (McMurray et al., 2016). The model also predicted that *A. queenslandica* sources most of its carbon from the particulate fraction of the water column. This has been shown in *Spongia officinalis*, which derives most of the carbon that is retained in its biomass from nanoeukaryotic-sized cells in the plankton (Topçu et al., 2010). The lipid content of the algal food is key for the growth of *A. queenslandica* as brassicasterol, the essential phytosterol, cannot be synthesised by the sponge.

5.6 Summary and future direction

The model presented here represents a significant step towards understanding the metabolism of sponges and their symbionts. The model can uptake different food sources and converts them into

both energy and biomass. Understanding how sponge biomass is synthesised at the molecular level is essential for the commercialisation of sponge-symbiont derived compounds (Mueller et al., 2004; Osinga et al., 1999; Rinkevich, 2011). The identification of essential nutrients including 10 amino acids, vitamins and a phytosterol provides the basis for the formulation of a defined or minimal media for culturing *A. queenslandica* cells.

The continued future development of the genome-scale metabolic model will involve the incorporation of new data as it becomes available. This in turn will increase the model's predictive power. Quantifying uptake and excretion rates of metabolites will greatly help to both constrain the solution space and improve the predictive power. In particular quantifying the consumption of oxygen could be related to the estimated ATP demands for maintenance and pumping. This model sets an important precedent using *A. queenslandica* towards understanding how the metabolism of sponge holobionts function at the molecular and system level.

Chapter 6 - General discussion

6.1 Overview

The overall aim of this thesis was to generate a dual-species genome-scale metabolic model (GEM) that was able to simulate the biosynthesis of biomass for the marine sponge *Amphimedon queenslandica* and its symbiotic bacteria *AqS1*. The development of a sponge-symbiont GEM will enable omics-scale data to be integrated and interpreted in a metabolic framework that gives resolution at the enzymes and reaction level, at the scale of the complete system. In constructing a dual-species GEM for the *A. queenslandica – AqS1* system, specific methods and data sets were developed and generated. These may serve as a guide for the development of GEMs for other sponges and marine invertebrates.

Insights from the analyses undertaken over the course this PhD research have the potential to inform marine sponge biology, ecology and biotechnology. For instance, marine sponges are valued for their ecosystem services, in particular cycling and retaining nutrients on coral reefs (Alexander et al., 2014; de Goeij et al., 2013; Rix et al., 2016), and as a source of bioactive secondary metabolites with commercial potential (Blunt et al., 2014; 2015; Konstantinidis & Tiedje, 2004; Leal et al., 2012; Mehbub et al., 2014; Thomas et al., 2010). It is accepted that the low commercialisation rate of sponge holobiont derived natural products is due to the inability to reliably produce sponge biomass (Belarbi, 2003; Leal et al., 2014; Osinga et al., 1999; Schippers et al., 2012; Sipkema et al., 2005; Wijffels, 2008). Simplified, growth is the catabolic and anabolic conversion of food by the central metabolic network into biomass. While the potential food sources that sponges can utilise are relatively well understood (de Goeij et al., 2008a,b, 2013; McMurray et al., 2016; Pile et al., 1996a; Rix et al., 2016; Yahel et al., 2003b), only partial analyses of the biochemical composition of sponge biomass has been reported (Chanas & Pawlik, 1995; Elvin, 1979; Koigoora et al., 2013; McClintock, 1987; Pallela et al., 2011; Stone, 1969; Vacelet et al., 1988) and very little is known of the metabolic network.

The generation of the biochemical data and subsequent conversion into a biomass equation is a complex task that previously required protocols for individual macromolecule types to be independently sourced. Chapter 2 resolves this by presenting a single, detailed methodological reference for lipid, protein, carbohydrate, DNA and RNA quantification that was missing from the scientific literature. The methods I chose avoided expensive commercial quantification kits and instead used reagents and equipment that most labs would have or can readily source. The protocols within Chapter 2 are used in both Chapter 3, where the detailed biomass composition is reported, and again in Chapter 4, where seasonal changes in the biomass composition were analysed. These protocols are also essential for future *in vivo* stable isotope labelling experiments which are used in metabolic modelling to test the *in silico* predictions (Sauer, 2006; Wittmann, 2007; Zamboni, 2011; Zamboni et al., 2009). Also essential for in vivo experiments was a method to relate wet sponge biomass to sponge dry weight. I used displacement volume, and showed that despite the heterogeneous nature of *A. queenslandica*'s biomass, this method was accurate and displayed a linear relationship. This method can be applied to all sponge species.

The first biomass composition analysis to include the skeleton, all macromolecule types and most of their respective building block molecules (analysis of carbohydrate composition was not undertaken) was presented as Chapter 3. This analysis presents a significant resource for future GEMs of other sponge-symbiont systems, as biomass compositions of related organisms are frequently used when the specific composition is unknown. To exemplify this point, recently the use of the nucleic acid and amino acid composition of yeast was accepted in the biomass composition of the recent *Caenorhabditis elegans* GEM (Yilmaz & Walhout, 2016). For organisms which there is little biochemical data, such as growth rates or oxygen consumption, knowing the actual biochemical composition and not using proxy data, ensures that the most significant constraint on network functionality is as accurate as possible (Feist & Palsson, 2010).

It was important to investigate the source of variation in the biochemical composition of *A. queenslandica* to ensure that further sampling and the design of future *in vivo* experiments represents the variability observed. By sampling five replicate biomass pieces from four individual sponges, I was able to determine the amount of variability due to differences within and between individuals (Chapter 3). I found that biochemical composition of *A. queenslandica* was heterogeneous, not just between sponges, but also within individuals (Chapter 3). This is significant, as sampling of sponge biochemical components in the past has been restricted to a single representative pieces from an individual (Chanas & Pawlik, 1995; Elvin, 1979; Pallela et al., 2011). The variability observed within individual sponges ensured that the sampling regime for the seasonal changes in composition (Chapter 4) contained replicates biomass samples from each sponge. It is unknown if this variability extends to other species of sponges, in particular if species with regular body plans exhibit the same within individual variation.

The biomass composition revealed that the most abundant macromolecule in the cellular fraction were lipids (Chapter 3). Generally protein is the most abundant macromolecule type found in cells (Michal & Schomburg, 2013; Stephanopoulos et al., 1998), although a higher relative abundance of lipids has been reported in two demosponges before (Chanas & Pawlik, 1995). With the data from Chapter 3, it was unclear if the high proportional lipid content was due to nutrient availability or the reproductive state of the *A. queenslandica* population. The seasonal sampling of both environmental parameters and biochemical composition (Chapter 4; Watson et al., in revision) indicated that the lipid content is likely related to reproductive state, with high lipid accumulations occurring before the increase in reproductive output associated with summer (Chapter 4; Watson et al., in revision; Leys et al., 2008). Sponges have been identified as key nutrient cyclers on coral reefs by linking pelagic nutrients with the benthic community (de Goeij et al., 2013; Rix et al., 2016). This seasonal increase in reproductive output may also contribute significantly to this nutrient cycling, as brooding sponges such as *A. queenslandica* release tens to hundreds of energy and nutrient rich larvae daily.

Sponges are renown for biosynthesising a diverse range of fatty acids and sterols, often with unique structural and bioactive properties (Blunt et al., 2009; 2010; 2011; 2014; Genta-Jouve & Thomas, 2012; Laport et al., 2009; Leal et al., 2012; Munro et al., 1999; Thomas et al., 2010). The lipid analysis revealed that twenty fatty acids contributed to the majority of the lipid biomass, half of these could not be identified using a commercially available compound library (Chapter 3; Watson et al., 2014a). It is anticipated that some of the unknown fatty acids and sterols will be identified in the future. Identification of the unknown fatty acids and sterols may reveal novel compounds, the synthesis pathways of which can be added to the metabolic model to improve its accuracy.

To place the dual-species GEM into its environmental context I monitored changes in the seasonal abundances of key nutrients in a natural habitat of *A. queenslandica* (Chapter 4; Watson et al., in revision). This allowed me to place realistic nutrient constraints when performing flux balance analysis (FBA) on the dual-species GEM (Chapter 5). *A. queenslandica* is a model sponge with a relatively well-characterise genome, development and microbiome, yet the environment where the study population resides had not previously been characterised. This environmental profiling filled an important gap in the literature for *A. queenslandica*. I found that for most of the year the available nutrients were in the dissolved fraction of the water column. Early autumn (March) was the exception with elevated levels of particulate nutrients that coincided with the end of the seabird-breeding season (Smith & Johnson, 1995). Nitrogen was identified as a potential limiting nutrient for *A. queenslandica* in Shark Bay (Chapter 4; Watson et al., in revision). Nitrogen is essential for the synthesis of amino acids, which dominate the overall macromolecule composition of the skeletal and cellular biomass (Watson et al., 2014a). This large requirement for, but low abundance of, nitrogen in the environment may be a driving factor for the apparent fast pumping rates in *A. queenslandica*, especially since their most abundant symbiotic bacteria appear unable to fix nitrogen (Gauthier et al, 2016). This is congruent with the low microbial abundance (LMA) status (Feith et al., 2016) of this sponge that implies *A. queenslandica* needs to process large volumes of water to meet its nitrogen requirements from particulate matter (Hentschel et al., 2003; 2006; Weisz et al., 2007b).

This thesis culminates with the creation of a dual-species genome-scale metabolic model for *A. queenslandica* and *AqS1*. The GEM can successfully create biomass from dissolved nutrients, represented by coral mucus, and particulate phytoplankton and bacterioplankton cells. The reported growth rates – biomass synthesis – in sponges varies significantly throughout the scientific literature (Belarbi, 2003; Duckworth, 2009; Mueller et al., 2004; Osinga et al., 1999; Schippers et al., 2012; Sipkema et al., 2005). This model will enable the maximum growth rates under different environmental conditions to be estimated, which in turn could provide valuable insights as to why there is so much variability in natural systems, and why the potential growth rates are not observed.

The process of reconstructing the metabolic network identified ten proteinogenic amino acids that *A. queenslandica* cannot synthesise, all of which *AqS1* is able to (Chapter 5). An additional five vitamins and a phytosterol were also found to be essential nutrients for *A. queenslandica*. Two of the vitamins can be synthesised and therefore possibly provided by *AqS1* (Chapter 5). The identification of these nutritional requirements is vital for unravelling the metabolic relationship between host and symbiont and for the development of suitable cell culture media for this, and possibly other, species of sponges. The ability of the dual-species GEM to generate enough energy (ATP) to satisfy the estimated requirements for growth and non-growth associated demands, including pumping water, validated that the model can theoretically survive in the simulated low and high nutrient conditions. This served as a proxy for model validation but measurements of oxygen uptake or product secretion should now be taken to further validate the model.

Here I have successfully completed the foundational work and created a genome-scale metabolic model for a marine invertebrate holobiont for the first time. This is a powerful tool that can now be used to investigate nutrient fluxes and biomass synthesis in marine sponges. Such knowledge can then be translated into a raft of biotechnological applications, including development of media and conditions for the cultivation and propagation of cells and whole sponges.

6.2 Future directions

The work presented here provides the foundation for the generation of *A. queenslandica* holobiont metabolic model. Specifically, I developed methods to accurately estimate biomass in this sponge in a range of contexts that provided insight into the macromolecular composition of this sponge and variation in this composition within a sponge, between sponges and between seasons. I also developed a metabolic model for *A. queenslandica* and its primary symbiont *AqS1*. However, this model only accommodates average nutrient constraints, and thus the solutions derived from the model represent average functional states. Applying this model and approach to the range of environmental conditions experienced by the *A. queenslandica* holobiont is the next step. For example, simulating how the system functions at the lower and upper limits of oxygen could have significant impacts on the ability of the sponge to produce enough energy (ATP) to grow and pump water. In low oxygen conditions, the exchange of organic acids may become more pronounced and essential as *AqS1* shifts its metabolism to fermentation.

The microbial community associated with *A. queenslandica* contains a second, abundant betaproteobacterium called *AqS2*. This species has reduced metabolic capabilities when compared to *AqS1*, but like *AqS1*, contains heterotrophic transporters and pathways. The heterotrophic transporters and pathways differ between *AqS1* and *AqS2*, which could translate into less metabolites being exported from the model and more efficient overall growth. The addition of *AqS2* to the model would encompass 90-99 % of the microbial community and be truly representative of a sponge holobiont. To supplement this it would be necessary to quantify the relative abundance of *A. queenslandica*, *AqS1* and *AqS2* cells within the holobiont and to develop algorithm to constrain fluxes accordingly.

These *in silico* lines of enquiry must be supplemented with biochemical data that is specific to *A. queenslandica*. Measuring exchange rates of compounds, such as oxygen and carbon dioxide, with the environment will help constrain the solution space and give confidence in the predicted metabolic phenotypes.

Finally, assumptions were made by gap filling reactions in both the *A. queenslandica* and *AqS1* metabolic networks. Investigating these gap filled reactions, and confirming their status as present or absent is important to understanding the metabolic interdependency between host and symbiont. Can *AqS1* actually produce the four amino acids that required the addition of an enzyme for their synthesis pathway to be complete? Can *AqS1* actually fix carbon dioxide into organic metabolites or has it lost this ability? These questions can now be answered by combining *in silico* modelling with *in vivo* labelling experiments where selected ^{13}C tracers are combined with in depth analytics. This will help to unravel who is eating what from the particulate and dissolved nutrient sources, and can extent to possibly feeding whole labelled cells. A brief example that was not reported in the main body of the thesis, I fed *A. queenslandica* ^{13}C labelled glucose and detected the production of poly-hydroxy alkanoates using gas-chromatography mass-spectrometry. Once completed, the model predicted the bacteria would produce one of these metabolites, (R)-3-Hydroxybutanoate, under high nutrient conditions (Chapter 5). Nano-Sims could be employed to detect where in the holobiont the ^{13}C labelling is occurring.

Overall this research has contributed significantly to advancing the understanding of sponge metabolism. The development of a genome-scale metabolic model for *A. queenslandica* and *AqS1* represents a significant step forward, more so because of the potential future applications.

References

Abdelmohsen, U. R., Bayer, K., & Hentschel, U. (2014). Diversity, abundance and natural products of marine sponge-associated actinomycetes. *Natural Product Reports*, *31*(3), 381. http://doi.org/10.1039/c3np70111e

Agren, R., Liu, L., Shoaie, S., Vongsangnak, W., Nookaew, I., & Nielsen, J. (2013). The RAVEN Toolbox and Its Use for Generating a Genome-scale Metabolic Model for *Penicillium chrysogenum*. *PLoS Computational Biology*, *9*(3), e1002980. http://doi.org/10.1371/journal.pcbi.1002980.s014

Alexander, B. E., Liebrand, K., Osinga, R., van der Geest, H. G., Admiraal, W., Cleutjens, J. P. M., Schutte, B., Verheyen, F., Ribes, M.,van Loon, E. & de Goeij, J. M.. (2014). Cell Turnover and Detritus Production in Marine Sponges from Tropical and Temperate Benthic Ecosystems. *PLoS ONE*, *9*(10), e109486. http://doi.org/10.1371/journal.pone.0109486.s002

Anavy, L., Levin, M., Khair, S., Nakanishi, N., Fernandez-Valverde, S. L., Degnan, B. M., & Yanai, I. (2014). BLIND ordering of large-scale transcriptomic developmental timecourses. *Development*, *141*(5), 1161–1166. http://doi.org/10.1242/dev.105288

Averesch, N. J. H., & Krömer, J. O. (2014). Tailoring strain construction strategies for muconic acid production in *S. cerevisiae* and *E. coli*. *Metabolic Engineering Communications*, *1*(C), 19–28. http://doi.org/10.1016/j.meteno.2014.09.001

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., et al. (2008). The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*, *9*(1), 75. http://doi.org/10.1186/1471-2164-9-75

Barnathan, G., Korpnrobst, J.-M., Doumenq, P., & Miralles, J. (1996). New unsaturated long-chain fatty acids in the phospholipids from the axinellida sponges *Trikentrion loeve* and *Pseudaxinella cf. lunaecharta*. *Lipids*, *31*(2), 193–200.

Bayer, K., Schmitt, S., & Hentschel, U. (2008). Physiology, phylogeny and in situ evidence for bacterial and archaeal nitrifiers in the marine sponge *Aplysina aerophoba*. *Environmental Microbiology*, *10*(11), 2942–2955. http://doi.org/10.1111/j.1462-2920.2008.01582.x

Bayne, B. L., & Newell, R. C. (1983). The Mollusca. (K. M. Wilbur & A. S. M. Saleuddin, Eds.) (Vol. 4). New York: Academic Press.

Belarbi, E. (2003). Producing drugs from marine sponges. *Biotechnology Advances*, *21*(7), 585–598. http://doi.org/10.1016/S0734-9750(03)00100-9

Bell, J. J. (2008). The functional roles of marine sponges. *Estuarine, Coastal and Shelf Science*, *79*(3), 341–353. http://doi.org/10.1016/j.ecss.2008.05.002

Benthin, S., Nielsen, J., & Villadsen, J. (1991). A simple and reliable method for the determination of cellular RNA content. *Biotechnology Techniques*, *5*(1), 39–42.

Bergman, O., Mayzel, B., Anderson, M. A., Shpigel, M., Hill, R. T., & Ilan, M. (2011). Examination of Marine-Based Cultivation of Three Demosponges for Acquiring Bioactive Marine Natural Products. *Marine Drugs*, *9*(12), 2201–2219. http://doi.org/10.3390/md9112201

Bernhard Palsson. (2009). Metabolic systems biology. *FEBS Letters*, *583*(24), 3900–3904. http://doi.org/10.1016/j.febslet.2009.09.031

Birkeland, C. (1997a). Life and Death Of Coral Reefs. Springer Science & Business Media, New York : Chapman & Hall.

Blank, L. M., Kuepfer, L., & Sauer, U. (2005). Large-scale 13C-flux analysis reveals mechanistic principles of metabolic network robustness to null mutations in yeast. *Genome Biology*, *6*(6), R49. http://doi.org/10.1186/gb-2005-6-6-r49

Block, J. H. (1974). Marine sterols from some gorgonians. *Steroids*, *23*(3), 421–424.

Blunt, J. W., Copp, B. R., Hu, W.-P., Munro, M. H. G., Northcote, P. T., & Prinsep, M. R. (2009). Marine natural products. *Natural Product Reports*, *26*(2), 170. http://doi.org/10.1039/b805113p

Blunt, J. W., Copp, B. R., Munro, M. H. G., Northcote, P. T., & Prinsep, M. R. (2010). Marine natural products. *Natural Product Reports*, *27*(2), 165. http://doi.org/10.1039/b906091j

Blunt, J. W., Copp, B. R., Munro, M. H. G., Northcote, P. T., & Prinsep, M. R. (2011). Marine natural products. *Natural Product Reports*, *28*(2), 196. http://doi.org/10.1039/c005001f

Blunt, J. W., Copp, B. R., Keyzers, R. A., Munro, M. H. G., & Prinsep, M. R. (2014). Marine natural products. *Natural Product Reports*, *31*(2), 160. http://doi.org/10.1039/c3np70117d

Blunt, J. W., Copp, B. R., Keyzers, R. A., Munro, M. H. G., & Prinsep, M. R. (2015). Marine natural products. *Natural Product Reports*, *32*, 116–211. http://doi.org/10.1039/C4NP00144C

Bobadilla Fazzini, R. A., Cortés, M. P., Padilla, L., Maturana, D., Budinich, M., Maass, A., & Parada, P. (2013). Stoichiometric modeling of oxidation of reduced inorganic sulfur compounds (Riscs) in *Acidithiobacillus thiooxidans*. *Biotechnology and Bioengineering*, *110*(8), 2242–2251. http://doi.org/10.1002/bit.24875

Bordbar, A., Lewis, N. E., Schellenberger, J., Palsson, B. O., & Jamshidi, N. (2010). Insight into human alveolar macrophage and M. tuberculosis interactions via metabolic reconstructions. *Molecular Systems Biology*, *6*, 1–14. http://doi.org/10.1038/msb.2010.68

Bordbar, A., Monk, J. M., King, Z. A., & Palsson, B. O. (2014). Constraint-based models predict metabolic and associated cellular functions. *Nature Reviews Genetics*, *15*(2), 107–120. http://doi.org/10.1038/nrg3643

Bradford, M. M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical Biochemistry*, *72*, 248–254.

Brennan, T. C. R., Kromer, J. O., & Nielsen, L. K. (2013). Physiological and Transcriptional Responses of Saccharomyces cerevisiae to d-Limonene Show Changes to the Cell Wall but Not to the Plasma Membrane. *Applied and Environmental Microbiology*, *79*(12), 3590–3600. http://doi.org/10.1128/AEM.00463-13

Brown, D. Tracker: Video Analysis and Modeling Tool. Open Source Physics. Version 4.90. Retrieved from http://physlets.org/tracker/

Carballeira, N. M., & Alicea, J. (2000). The first naturally occurring alpha-methoxylated branched-chain fatty acids from the phospholipids of *Amphimedon complanata*. *Lipids*, *36*(1), 83–87.

Carballeira, N. M., & Lopez, M. R. (1989). On the isolation of 2-hydroxydocosanoic and 2-hydroxytricosanoic acids from the marine sponge *Amphimedon compressa*. *Lipids*, *24*(1), 89–91.

Carlson, C. A., & Hansell, D. A. (2015). DOM Sources, Sinks, Reactivity, and Budgets. *Biogeochemistry of Marine Dissolved Organic Matter* (Second Edition, pp. 65–126). Elsevier Inc. http://doi.org/10.1016/B978-0-12-405940-5.00003-0

Carnicer, M., Baumann, K., Töplitz, I., Sánchez-Ferrando, F., Mattanovich, D., Ferrer, P., & Albiol, J. (2009). Macromolecular and elemental composition analysis and extracellular metabolite balances of Pichia pastoris growing at different oxygen levels. *Microbial Cell Factories*, *8*(1), 65. http://doi.org/10.1186/1475-2859-8-65

Carr, W. E. S., James C Netherton, S. I. S., Gleeson, R. A., & Derby, C. D. (1996). Stimulants of Feeding Behavior in Fish: Analyses of Tissues of Diverse Marine Organisms. *Biological Bulletin*, *190*(2), 149–160. http://doi.org/10.2307/1542535?ref=no-x-route:1da895bfb9904f4f9b8d45dec1e83baf

Cesarone, C. F., Bolognesi, C., & Santi, L. (1979). Improved microfluorometric DNA determination in biological material using 33258 Hoechst. *Analytical Biochemistry*, *100*(1), 188–197.

Chanas, B., & Pawlik, J. R. (1995). Defenses of Caribbean sponges against predatory reef fish. II. Spicules, tissue toughness, and nutritional quality. *Marine Ecology Progress Series*, *127*(1), 195–211.

Christie, W. W., Brechany, E. Y., Stefanov, K., & Popov, S. (1992). The fatty acids of the sponge *Dysidea fragilis* from the black sea. *Lipids*, *27*(8), 640–644.

Clancey, V. J. (1926). The Constitution of Sponges: The Common Bath Sponge, *Hippospongia equina*. *Biochemical Journal*, *20*(6), 1186–1189.

de Caralt, S., Uriz, M. J., & Wijffels, R. H. (2007). Cell culture from sponges: pluripotency and immortality. *Trends in Biotechnology*, *25*(10), 467–471. http://doi.org/10.1016/j.tibtech.2007.08.006

de Goeij, J. M., De Kluijver, A., van Duyl, F. C., Vacelet, J., Wijffels, R. H., De Goeij, A. F. P. M., et al. (2009). Cell kinetics of the marine sponge *Halisarca caerulea* reveal rapid cell turnover and shedding. *Journal of Experimental Biology*, *212*(23), 3892–3900. http://doi.org/10.1242/jeb.034561

de Goeij, J. M., Moodley, L., Houtekamer, M., Carballeira, N., & van Duyl, F. C. (2008a). Tracing 13C-enriched dissolved and particulate organic carbon in the bacteria-containing coral reef sponge *Halisarca caerulea*: Evidence for DOM-feeding., *53*(4), 1376–1386.

de Goeij, J. M., van den Berg, H., van Oostveen, M. M., Epping, E., & van Duyl, F. C. (2008b). Major bulk dissolved organic carbon (DOC) removal by encrusting coral reef cavity sponges. *Marine Ecology Progress Series*, *357*, 139–151. http://doi.org/10.3354/meps07403

de Goeij, J. M., van Oevelen, D., Vermeij, M. J. A., Osinga, R., Middelburg, J. J., de Goeij, A. F. P. M., & Admiraal, W. (2013). Surviving in a marine desert: the sponge loop retains resources within coral reefs. *Science*, *342*(6154), 108–110. http://doi.org/10.1126/science.1241981

de Oliveira Dal'Molin, C. G., Quek, L. E., Palfreyman, R. W., Brumbley, S. M., & Nielsen, L. K. (2010). AraGEM, a Genome-Scale Reconstruction of the Primary Metabolic Network in Arabidopsis. *Plant Physiology*, *152*(2), 579–589. http://doi.org/10.1104/pp.109.148817

de Oliveira Dal'Molin, C. G., Quek, L.-E., Palfreyman, R. W., & Nielsen, L. K. (2011). AlgaGEM--a genome-scale metabolic reconstruction of algae based on the *Chlamydomonas reinhardtii* genome. *BMC Genomics*, *12 Suppl 4*, S5–S5. http://doi.org/10.1186/1471-2164-12-S4-S5

De Rosa, S., De Caro, S., Iodice, C., Tommonaro, G., Stefanov, K., & Popov, S. (2003). Development in primary cell culture of demosponges. *Journal of Biotechnology*, *100*(2), 119–125.

Degnan, B. M., Adamska, M., Craigie, A., Degnan, S. M., Fahey, B., Gauthier, M., Hooper, J. N., Larroux, C., Leys, S. P., Lovas, E. & Richards, G.S. (2008). The Demosponge *Amphimedon queenslandica*: Reconstructing the Ancestral Metazoan Genome and Deciphering the Origin of Animal Multicellularity., *Cold Spring Harbour Protocols,* http://doi: 10.1101/pdb.emo108

Degnan, B. M., Adamska, M., Richards, G. S., Larroux, C., Leininger, S., Bergum, B., Calcino, C., Taylor, K., Nakanishi, Y. & Degnan, S. (2015). Evolutionary Developmental Biology of Invertebrates 1, Porifera (pp. 65–106), Vienna: Springer Vienna. http://doi.org/10.1007/978-3-7091-1862-7_4

Dubois, M., Gilles, K. A., Hamilton, J. K., Rebers, P. T., & Smith, F. (1956). Colorimetric method for determination of sugars and related substances. *Analytical Chemistry*, *28*(3), 350–356.

Duckworth, A. (2009). Farming Sponges to Supply Bioactive Metabolites and Bath Sponges: A Review. *Marine Biotechnology*, *11*(6), 669–679. http://doi.org/10.1007/s10126-009-9213-2

Elvin, D. W. (1979). The relationship of seasonal changes in the biochemical components to the reproductive behavior of the intertidal sponge, *Haliclona permollis*. *The Biological Bulletin*, *156*(1), 47–61.

Erdman, T. R., & Thomson, R. H. (1972). Sterols from the sponges *Cliona celata* Grant and *Hymeniacidon perleve* Montagu. *Tetrahedron*, *28*(20), 5163–5173.

Ereskovsky, A. V. (2010). The Comparative Embryology of Sponges. Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-90-481-8575-7

Erler, D. V., Santos, I. R., & Eyre, B. D. (2014). Inorganic nitrogen transformations within permeable carbonate sands. *Continental Shelf Research*, *77*(C), 69–80. http://doi.org/10.1016/j.csr.2014.02.002

Erpenbeck, D., & Soest, R. W. M. (2006). Status and Perspective of Sponge Chemosystematics. *Marine Biotechnology*, *9*(1), 2–19. http://doi.org/10.1007/s10126-005-6109-7

Esvelt, K. M., & Wang, H. H. (2013). Genome-scale engineering for systems and synthetic biology. *Molecular Systems Biology*, *9*, 1–17. http://doi.org/10.1038/msb.2012.66

Eyre, B. D., Glud, R. N., & Patten, N. (2008). Mass coral spawning: A natural large-scale nutrient addition experiment. *Limnology and Oceanography*, *53*(3), 997–1013.

Fan, L., Reynolds, D., Liu, M., Stark, M., Kjelleberg, S., Webster, N. S., & Thomas, T. (2012). Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts. *Proceedings of the National Academy of Sciences*, *109*(27), E1878–E1887. http://doi.org/10.1073/pnas.1203287109/-/DCSupplemental

Feist, A. M., & Palsson, B. O. (2010). The biomass objective function. *Current Opinion in Microbiology*, *13*(3), 344–349. http://doi.org/10.1016/j.mib.2010.03.003

Feist, A. M., Herrgård, M. J., Thiele, I., Reed, J. L., & Palsson, B. O. (2008). Reconstruction of biochemical networks in microorganisms. *Nature Reviews Microbiology*, *7*(2), 129–143. http://doi.org/10.1038/nrmicro1949

Feith, R. A., Gauthier, M-E., A., Bayes, J., Green., K. M. & Degnan, S. M. (2016). Ontogenetic changes in the Bacterial Symbiont Community of the Tropical Demosponge *Amphimedon queenslandica*: Metamorphosis Is a New Beginning.Frontiers in Marine Science, doi: 10.3389/fmars.2016.00228

Fernandez-Valverde, S. L., Calcino, A. D., & Degnan, B. M. (2015). Deep developmental transcriptome sequencing uncovers numerous new genes and enhances gene annotation in the sponge *Amphimedon queenslandica*. *BMC Genomics*, *16*(1), 720. http://doi.org/10.1186/s12864-015-1588-z

Ferrier-Pages, C., Gattuso, J. P., Cauwet, G., Jaubert, J., & Allemand, D. (1998). Release of dissolved organic carbon and nitrogen by the zooxanthellate coral. *Marine Ecology Progress Series*, *172*, 265–274.

Fiore, C. L., Jarett, J. K., & Lesser, M. P. (2013). Symbiotic prokaryotic communities from different populations of the giant barrel sponge, *Xestospongia muta*. *MicrobiologyOpen*, *2*(6), 938–952. http://doi.org/10.1002/mbo3.135

Fiore, C. L., Labrie, M., Jarett, J. K., & Lesser, M. P. (2015). Transcriptional activity of the giant barrel sponge, *Xestospongia muta* Holobiont: molecular evidence for metabolic interchange. *Frontiers in Microbiology*, *6*. http://doi.org/10.3389/fmicb.2015.00364

Flack, A., & Munro, H. N. (1962). The precision of ultraviolet absorption measurements in the Schmidt-Thannhauser procedure for nucleic acid estimation. *Biochimica Et Biophysica Acta*, *55*, 571–583.

Freeman, C. J., & Gleason, D. F. (2010). Chemical defenses, nutritional quality, and structural components in three sponge species: *Ircinia felix, I. campana*, and *Aplysina fulva*. *Marine Biology*, *157*(5), 1083–1093. http://doi.org/10.1007/s00227-010-1389-5

Freeman, C. J., & Thacker, R. W. (2011). Complex interactions between marine sponges and their symbiotic microbial communities. *Limnology and Oceanography*, *56*(5), 1577–1586. http://doi.org/10.4319/lo.2011.56.5.1577

Freeman, C. J., Baker, D. M., Easson, C. G., & Thacker, R. W. (2015). Shifts in sponge-microbe mutualisms across an experimental irradiance gradient. *Marine Ecology Progress Series*, *526*, 41–53. http://doi.org/10.3354/meps11249

Freeman, C. J., Easson, C. G., & Baker, D. M. (2014). Metabolic diversity and niche structure in sponges from the Miskito Cays, Honduras. *PeerJ*, *2*, e695. http://doi.org/10.7717/peerj.695/supp-1

Freshney, R. I. (2015). Culture of Animal Cells. John Wiley & Sons.

Friedrich, B., & Schwartz, E. (1993). Molecular biology of hydrogen utilization in aerobic chemolithotrophs. *Annual Reviews in Microbiology*. 47, 351-383

Friedrich, C. G., Bardischewsky, F., Rother, D., Quentmeier, A., & Fischer, J. (2005). Prokaryotic sulfur oxidation. *Current Opinion in Microbiology*, *8*(3), 253–259. http://doi.org/10.1016/j.mib.2005.04.005

Frith, C. A. (1983). Some aspects of lagoon sedimentation and circulation at One Tree Reef, southern Great Barrier Reef. *BMR Journal of Australian Geology Geophysics*, *8*, 211–221.

Garcia Camacho, F., Chileh, T., Ceron Garcia, M. C., Sanchez Miron, A., Belarbi, E. H., Contreras Gomez, A., & Molina Grima, E. (2006). Sustained Growth of Explants from Mediterranean Sponge *Crambe crambe* Cultured In Vitro with Enriched RPMI 1640. *Biotechnology Progress*, *22*(3), 781–790. http://doi.org/10.1021/bp050341m

Garrity, G., Brenner, D. J., Krieg, N. R., & Staley, J. R. (2007). Bergey's Manual of Systematic Bacteriology. Springer Science & Business Media.

Gauthier, M-E. A., Watson, J. R. & Degnan, S. M. (2016) Draft genomes shed light on the dual bacterial symbiosis that dominates the microbiome of the coral reef sponge *Amphimedon queenslandica*. Frontiers in Marine Science, http://dx.doi.org/10.3389/fmars.2016.00196

Genta-Jouve, G., & Thomas, O. P. (2012). Sponge Chemical Diversity: From Biosynthetic Pathways to Ecological Roles. *Advances in Sponge Science: Physiology, Chemical and Microbial Diversity, Biotechnology* (1st ed., Vol. 62, pp. 183–230). Elsevier Ltd. http://doi.org/10.1016/B978-0-12-394283-8.00004-7

Grasela, J. J., Pomponi, S. A., Rinkevich, B., & Grima, J. (2011). Efforts to develop a cultured sponge cell line: revisiting an intractable problem. *In Vitro Cellular & Developmental Biology - Animal*, *48*(1), 12–20. http://doi.org/10.1007/s11626-011-9469-5

Green, M. R., Sambrook, J., & Sambrook, J. (2012). Molecular cloning: a laboratory manual 4 edition Cold Spring Harbor Laboratory Press. New York.

Gross, J., Sokal, Z., & Rougvie, M. (1956). Structural and chemical studies on the connective tissue of marine sponges. *Journal of Histochemistry & Cytochemistry*, *4*(3), 227–246. http://doi.org/10.1177/4.3.227

Haines, T. H. (2001). Do sterols reduce proton and sodium leaks through lipid bilayers? *Progress in Lipid Research*, *40*(4), 299–324.

Hallam, S. J., Mincer, T. J., Schleper, C., Preston, C. M., Roberts, K., Richardson, P. M., & DeLong, E. F. (2006). Pathways of Carbon Assimilation and Ammonia Oxidation Suggested by Environmental Genomic Analyses of Marine Crenarchaeota, *4*(4), e95–17. http://doi.org/10.1371/journal.pbio.0040095

Hara, A., & Radin, N. S. (1978). Lipid extraction of tissues with a low-toxicity solvent. *Analytical Biochemistry*, *90*(1), 420–426.

Hatcher, B. G. (1997). Coral reef ecosystems: how much greater is the whole than the sum of the parts? *Coral Reefs*. 16, S77-S91

Hatcher, B. G. (2003). Coral reef primary productivity. A hierarchy of pattern and process. *Trends in Ecology & Evolution*, 1–7.

Haygood, M. G., Schmidt, E. W., Davidson, S. K., & Faulkner, D. J. (1999). Microbial symbionts of marine invertebrates: opportunities for microbial biotechnology. *Journal of Molecular Microbiology and Biotechnology*, *1*(1), 33–43.

Heinken, A., Khan, M. T., Paglia, G., Rodionov, D. A., Harmsen, H. J. M., & Thiele, I. (2014). Functional Metabolic Map of *Faecalibacterium prausnitzii*, a Beneficial Human Gut Microbe. *Journal of Bacteriology*, *196*(18), 3289–3302. http://doi.org/10.1128/JB.01780-14

Heinken, A., Sahoo, S., Fleming, R. M. T., & Thiele, I. (2013). Systems-level characterization of a host-microbe metabolic symbiosis in the mammalian gut. *Gut Microbes*, *4*(1), 28–40. http://doi.org/10.4161/gmic.22370

Henry, C. S., Broadbelt, L. J., & Hatzimanikatis, V. (2007). Thermodynamics-Based Metabolic Flux Analysis. *Biophysical Journal*, *92*(5), 1792–1805. http://doi.org/10.1529/biophysj.106.093138

Henry, C. S., DeJongh, M., Best, A. A., Frybarger, P. M., Ben Linsay, & Stevens, R. L. (2010). High-throughput generation, optimization and analysis of genomes-scale metabolic models. *Nature Biotechnology*, *28*(9), 969–974. http://doi.org/10.1038/nbt.1672

Hentschel, U., Fieseler, L., Wehrl, M., & Gernert, C. (2003). Microbial diversity of marine sponges. Sponges (Porifera), *Progress in Molecular and Subcellular Biology,* Springer Publishing

Hentschel, U., Hopke, J., Horn, M., Friedrich, A. B., Wagner, M., Hacker, J., & Moore, B. S. (2002). Molecular Evidence for a Uniform Microbial Community in Sponges from Different Oceans. *Applied and Environmental Microbiology*, *68*(9), 4431–4440. http://doi.org/10.1128/AEM.68.9.4431-4440.2002

Hentschel, U., Piel, J., Degnan, S. M., & Taylor, M. W. (2012). Genomic insights into the marine sponge microbiome. *Nature Publishing Group*, *10*(9), 641–654. http://doi.org/10.1038/nrmicro2839

Hentschel, U., Usher, K. M., & Taylor, M. W. (2006). Marine sponges as microbial fermenters. *FEMS Microbiology Ecology*, *55*(2), 167–177. http://doi.org/10.1111/j.1574-6941.2005.00046.x

Hoffmann, F., Radax, R., Woebken, D., Holtappels, M., Lavik, G., Rapp, H. T., et al. (2009). Complex nitrogen cycling in the sponge *Geodia barretti*. *Environmental Microbiology*, *11*(9), 2228–2243. http://doi.org/10.1111/j.1462-2920.2009.01944.x

Holms, H. (1996). Flux analysis and control of the central metabolic pathways in *Escherichia coli*. *FEMS Microbiology Reviews*, *19*(2), 85–116.

Hooper, J. (2002). Systema Porifera. Taylor & Francis US.

Hooper, J. N., & Van Soest, R. W. (2006). A new species of Amphimedon (Porifera, Demospongiae, Haplosclerida, Niphatidae) from the Capricorn-Bunker Group of Islands, Great Barrier Reef, Australia: target species for the "sponge genome project." *Zootaxa*, *1314*, 31–39.

Huyck, T. K., Gradishar, W., Manuguid, F., & Kirkpatrick, P. (2011). Eribulin mesylate, *Nature Reviews Drug Discovery*. 10;173-174. http://doi.org/10.1038/nrd3389

Ihaka, R., & Gentleman, R. (1996). R: a language for data analysis and graphics. *Journal of Computational and Graphical Statistics*, 5(3):299-314.

Imhoff, J. M., & Garrone, R. (1983). Solubilization and characterization of *Chondrosia reniformis* sponge collagen. *Connective Tissue Research*, *11*(2-3), 193–197.

James, B. W., Mauchline, W. S., Dennis, P. J., Keevil, C. W., & Wait, R. (1999). Poly-3-hydroxybutyrate in *Legionella pneumophila*, an energy source for survival in low-nutrient environments. *Applied and Environmental Microbiology*, *65*(2), 822–827.

Johnson, J. D., Mehus, J. G., Tews, K., Milavetz, B. I., & Lambeth, D. O. (1998). Genetic evidence for the expression of ATP- and GTP-specific succinyl-CoA synthetases in multicellular eucaryotes. *The Journal of Biological Chemistry*, *273*(42), 27580–27586.

Joyner, J. L., Peyer, S. M., & Lee, R. W. (2003). Possible Roles of Sulfur-Containing Amino Acids in a Chemoautotrophic Bacterium-Mollusc Symbiosis. *Biological Bulletin*, *205*(3), 331–338. http://doi.org/10.2307/1543296?ref=no-x-route:f67dadb9981a6c254e12a6461c3fc8b8

Junqua, S., Robert, L., Garrone, R., Pavans de Ceccatty, M., & Vacelet, J. (1974). Biochemical and morphological studies on collagens of horny sponges. *Ircinia* filaments compared to spongines. *Connective Tissue Research*, *2*(3), 193–203.

Kamke, J., Sczyrba, A., Ivanova, N., Schwientek, P., Rinke, C., Mavromatis, K., et al. (2013). Single-cell genomics reveals complex carbohydrate degradation patterns in poribacterial symbionts of marine sponges, *7*(12), 2287–2300. http://doi.org/10.1038/ismej.2013.111

Kang, C.-K., Choy, E. J., Lee, W. C., Kim, N. J., Park, H.-J., & Choi, K.-S. (2011). Physiological energetics and gross biochemical composition of the ascidian *Stylea clava* cultured in suspension in a temperate bay of Korea. *Aquaculture*, *319*(1-2), 168–177. http://doi.org/10.1016/j.aquaculture.2011.06.016

Karp, P. D., Paley, S. M., Krummenacker, M., Latendresse, M., Dale, J. M., Lee, T. J., et al. (2010). Pathway Tools version 13.0: integrated software for pathway/genome informatics and systems biology. *Briefings in Bioinformatics*, *11*(1), 40–79. http://doi.org/10.1093/bib/bbp043

Karp, P. D., Paley, S., & Romero, P. (2002). The pathway tools software. *Bioinformatics*, *18*(1);S225-S232. doi: 10.1093/bioinformatics/18.suppl_1.S

Kennedy, J., Codling, C. E., Jones, B. V., Dobson, A. D. W., & Marchesi, J. R. (2008). Diversity of microbes associated with the marine sponge, *Haliclona simulans*, isolated from Irish waters and identification of polyketide synthase genes from the sponge metagenome. *Environmental Microbiology*, *10*(7), 1888–1902. http://doi.org/10.1111/j.1462-2920.2008.01614.x

Kleiner, M., Petersen, J. M., & Dubilier, N. (2012a). Convergent and divergent evolution of metabolism in sulfur-oxidizing symbionts and the role of horizontal gene transfer. *Current Opinion in Microbiology*, *15*(5), 621–631. http://doi.org/10.1016/j.mib.2012.09.003

Kleiner, M., Wentrup, C., Lott, C., Teeling, H., Wetzel, S., Young, J., et al. (2012b). Metaproteomics of a gutless marine worm and its symbiotic microbial community reveal unusual pathways for carbon and energy use. *Proceedings of the National Academy of Sciences*, *109*(19), E1173–E1182. http://doi.org/10.1073/pnas.1121198109

Koigoora, S., Pallela, R., Nelakurti, C. R., Sunkara, M. S., & Janapala, V. R. (2013). Morphometric, elemental and biochemical characterisation of three Haliclona species (Demospongiae) from the Gulf of Mannar, India. *Chemical Speciation and Bioavailability*, *25*(1), 15–23. http://doi.org/10.3184/095422912X13496990950939

Konstantinidis, K. T., & Tiedje, J. M. (2004). Trends between gene content and genome size in prokaryotic species with larger genomes. *Proceedings of the National Academy of Sciences*, *101*(9), 3160–3165. http://doi.org/10.1073/pnas.0308653100

Koopmans, M., Martens, D., & Wijffels, R. H. (2009). Growth Efficiency and Carbon Balance for the Sponge *Haliclona oculata*. *Marine Biotechnology*, *12*(3), 340–349. http://doi.org/10.1007/s10126-009-9228-8

Koopmans, M., van Rijswijk, P., Martens, D., Egorova-Zachernyuk, T. A., Middelburg, J. J., & Wijffels, R. H. (2010). Carbon conversion and metabolic rate in two marine sponges. *Marine Biology*, *158*(1), 9–20. http://doi.org/10.1007/s00227-010-1538-x

Krömer, J. O., Nielsen, L., & Blank, L. M. (2014). Metabolic Flux Analysis. Humana Press.

Krumholz, E. W., Yang, H., Weisenhorn, P., Henry, C. S., & Libourel, I. G. L. (2012). Genome-wide metabolic network reconstruction of the picoalga Ostreococcus. *Journal of Experimental Botany*, *63*(6), 2353–2362. http://doi.org/10.1093/jxb/err407

Laport, M. S., Santos, O. C. S., & Muricy, G. (2009). Marine sponges: potential sources of new antimicrobial drugs. *Current Pharmaceutical Biotechnology*, *10*(1), 86–105.

Latendresse, M., Krummenacker, M., Trupp, M., & Karp, P. D. (2012). Construction and completion of flux balance models from pathway databases. *Bioinformatics*, *28*(3), 388–396. http://doi.org/10.1093/bioinformatics/btr681

Lawson, M. P., Bergquist, P. R., & Cambie, R. C. (1984). Fatty acid composition and the classification of the Porifera. *Biochemical Systematics and Ecology*, *12*(4), 375–393. http://doi.org/10.1016/j.chemphyslip.2006.06.001

Leal, M. C., Puga, J., Serôdio, J., Gomes, N. C. M., & Calado, R. (2012). Trends in the Discovery of New Marine Natural Products from Invertebrates over the Last Two Decades – Where and What Are We Bioprospecting? *PLoS ONE*, *7*(1), e30580. http://doi.org/10.1371/journal.pone.0030580.t003

Leal, M., Sheridan, C., Osinga, R., Dionísio, G., Rocha, R., Silva, B., et al. (2014). Marine Microorganism-Invertebrate Assemblages: Perspectives to Solve the "Supply Problem" in the Initial Steps of Drug Discovery. *Marine Drugs*, *12*(7), 3929–3952. http://doi.org/10.3390/md12073929

Lee, T. J., Paulsen, I., & Karp, P. (2008). Annotation-based inference of transporter function. *Bioinformatics*, *24*(13), i259–i267. http://doi.org/10.1093/bioinformatics/btn180

Lewis, N. E., Nagarajan, H., & Palsson, B. O. (2012). Constraining the metabolic genotype–phenotype relationship using a phylogeny of in silico methods. *Nature Publishing Group*. http://doi.org/10.1038/nrmicro2737

Leys, S. P., & Degnan, B. M. (2001). Cytological basis of photoresponsive behavior in a sponge larva. *The Biological Bulletin*, *201*(3), 323–338.

Leys, S. P., Larroux, C., Gauthier, M., Adamska, M., Fahey, B., Richards, G. S., et al. (2008). Isolation of Amphimedon Developmental Material. *Cold Spring Harbor Protocols*, *2008*(13), pdb.prot5095–pdb.prot5095. http://doi.org/10.1101/pdb.prot5095

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (2001). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Advanced Drug Delivery Reviews*, *46*(1-3), 3–26.

Liu, M. Y., Kjelleberg, S., & Thomas, T. (2010). Functional genomic analysis of an uncultured & delta-proteobacterium in the sponge *Cymbastela concentrica*, *5*(3), 427–435. http://doi.org/10.1038/ismej.2010.139

Locasale, J. W., & Cantley, L. C. (2011). Metabolic Flux and the Regulation of Mammalian Cell Growth. *Cell Metabolism*, *14*(4), 443–451. http://doi.org/10.1016/j.cmet.2011.07.014

Loh, T. L., & Pawlik, J. R. (2014). Chemical defenses and resource trade-offs structure sponge communities on Caribbean coral reefs. *Proceedings of the National Academy of Sciences*, *111*(11), 4151–4156. http://doi.org/10.1073/pnas.1321626111

Louden, D., Whalan, S., Evans-Illidge, E., Wolff, C., & de Nys, R. (2007). An assessment of the aquaculture potential of the tropical sponges *Rhopaloeides odorabile* and *Coscinoderma sp*. *Aquaculture*, *270*(1-4), 57–67. http://doi.org/10.1016/j.aquaculture.2007.02.038

Mueller, W. G., Grebenjuk, V., Le Pennec, G. L., Schr der, H.-C., Bremmer, F., Hentschel, U., et al. (2004). Sustainable Production of Bioactive Compounds by Sponges? Cell Culture and Gene Cluster Approach: A Review. *Marine Biotechnology*, *6*(2). http://doi.org/10.1007/s10126-002-0098-6

Macdonald, S. J., Lin, G. G., Russell, C. W., Thomas, G. H., & Douglas, A. E. (2012). The central role of the host cell in symbiotic nitrogen metabolism. *Proceedings of the Royal Society B: Biological Sciences*, *279*(1740), 2965–2973. http://doi.org/10.1098/rspb.2012.0414

McClintock, J. B. (1987). Investigation of the relationship between invertebrate predation and biochemical composition, energy content, spicule armament and toxicity of benthic sponges at McMurdo Sound, Antarctica. *Marine Biology*, *94*(3), 479–487.

McMurray, S. E., Finelli, C. M., & Pawlik, J. R. (2015). Journal of Experimental Marine Biology and Ecology. *Journal of Experimental Marine Biology and Ecology*, *473*(C), 73–80. http://doi.org/10.1016/j.jembe.2015.08.007

McMurray, S. E., Johnson, Z. I., Hunt, D. E., Pawlik, J. R., & Finelli, C. M. (2016). Selective feeding by the giant barrel sponge enhances foraging efficiency. *Limnology and Oceanography*, n/a–n/a. http://doi.org/10.1002/lno.10287

Mehbub, M., Lei, J., Franco, C., & Zhang, W. (2014). Marine Sponge Derived Natural Products between 2001 and 2010: Trends and Opportunities for Discovery of Bioactives. *Marine Drugs*, *12*(8), 4539–4577. http://doi.org/10.3390/md12084539

Michal, G., & Schomburg, D. (2013). Biochemical Pathways. John Wiley & Sons.

Moitinho-Silva, L., Bayer, K., Cannistraci, C. V., Giles, E. C., Ryu, T., Seridi, L., et al. (2013). Specificity and transcriptional activity of microbiota associated with low and high microbial abundance sponges from the Red Sea. *Molecular Ecology*, *23*(6), 1348–1363. http://doi.org/10.1111/mec.12365

Morris, J. C., & Phillips, A. J. (2009). Marine natural products: synthetic aspects. *Natural Product Reports*, *26*(2), 245. http://doi.org/10.1039/b805111a

Morris, J. C., & Phillips, A. J. (2011). Marine natural products: Synthetic aspects. *Natural Product Reports*, *28*(2), 269. http://doi.org/10.1039/c0np00066c

Mueller, B., de Goeij, J. M., Vermeij, M. J. A., Mulders, Y., van der Ent, E., Ribes, M., & van Duyl, F. C. (2014). Natural Diet of Coral-Excavating Sponges Consists Mainly of Dissolved Organic Carbon (DOC). *PLoS ONE*, *9*(2), e90152. http://doi.org/10.1371/journal.pone.0090152.s002

Munro, M. H., Blunt, J. W., Dumdei, E. J., Hickford, S. J., Lill, R. E., Li, S., et al. (1999). The discovery and development of marine compounds with pharmaceutical potential. *Journal of Biotechnology*, *70*(1-3), 15–25.

Muyzer, G., Sorokin, D. Y., Mavromatis, K., Lapidus, A., Clum, A., Ivanova, N., et al. (2011a). Complete genome sequence of *Thioalkalivibrio "sulfidophilus" HL-EbGr7*. *Standards in Genomic Sciences*, *4*(1), 23–35. http://doi.org/10.4056/sigs.1483693

Muyzer, G., Sorokin, D. Y., Mavromatis, K., Lapidus, A., Foster, B., Sun, H., et al. (2011b). Complete genome sequence of *Thioalkalivibrio sp. K90mix*. *Standards in Genomic Sciences*, *5*(3), 341–355. http://doi.org/10.4056/sigs.2315092

Ponce-de-Leon, M., Montero, F., & Peret, J. (2013). Solving gap metabolites and blocked reactions in genome-scale models: application to themetabolic network of *Blattabacterium cuenoti*. *BMC Systems Biology*, *7*(1), 1–1. http://doi.org/10.1186/1752-0509-7-114

Nagarajan, H., Embree, M., Rotaru, A.-E., Shrestha, P. M., Feist, A. M., Palsson, B. O., et al. (2013). Characterization and modelling of interspecies electron transfer mechanisms and microbial community dynamics of a syntrophic association. *Nature Communications*, *4*, 1–10. http://doi.org/10.1038/ncomms3809

Nechev, J., Christie, W. W., Robaina, R., de Diego, F., Popov, S., & Stefanov, K. (2004). Chemical composition of the sponge *Hymeniacidon sanguinea* from the Canary Islands. *Comparative Biochemistry and Physiology Part a: Molecular & Integrative Physiology*, *137*(2), 365–374. http://doi.org/10.1016/j.cbpb.2003.10.016

Nelson, D. L., & Cox, M. M. (2012). Lehninger Principles of Biochemistry. W. H. Freeman.

Oberhardt, M. A., Palsson, B. O., & Papin, J. A. (2009). Applications of genome-scale metabolic reconstructions. *Molecular Systems Biology*, *5*, 1–15. http://doi.org/10.1038/msb.2009.77

Ogawa, T., Furusawa, T., Nomura, R., Seo, D., Hosoya-Matsuda, N., Sakurai, H., & Inoue, K. (2008). SoxAX Binding Protein, a Novel Component of the Thiosulfate-Oxidizing Multienzyme System in the Green Sulfur Bacterium *Chlorobium tepidum*. *Journal of Bacteriology*, *190*(18), 6097–6110. http://doi.org/10.1128/JB.00634-08

Oh, Y.-G., Lee, D.-Y., Lee, S. Y., & Park, S. (2009). Multiobjective flux balancing using the NISE method for metabolic network analysis. *Biotechnology Progress*, *25*(4), 999–1008. http://doi.org/10.1002/btpr.193

Orth, J. D., & Palsson, B. (2012). Gap-filling analysis of the iJO1366 *Escherichia coli* metabolic network reconstruction for discovery of metabolic functions. *BMC Systems Biology*, *6*, 1–15. http://doi.org/10.1186/1752-0509-6-30

Orth, J. D., Conrad, T. M., Na, J., Lerman, J. A., Nam, H., Feist, A. M., & Palsson, B. O. (2011). A comprehensive genome-scale reconstruction of *Escherichia coli* metabolism—2011. *Molecular Systems Biology*, *7*, 1–9. http://doi.org/10.1038/msb.2011.65

Orth, J. D., Thiele, I., & Palsson, B. O. (2010). What is flux balance analysis? *Nature Biotechnology*, *28*(3), 245–248. http://doi.org/10.1038/nbt.1614

Osinga, R., Hassan Belarbi, El, Grima, E. M., Tramper, J., & Wijffels, R. H. (2003). Progress towards a controlled culture of the marine sponge *Pseudosuberites andrewsi* in a bioreactor. *Journal of Biotechnology*, *100*(2), 141–146.

Osinga, R., Tramper, J., & Wijffels, R. (1999). Cultivation of Marine Sponges. *Marine Biotechnology*, *1*(6), 509–532.

Otero, J. M., & Nielsen, J. (2010). Industrial systems biology. *Biotechnology and Bioengineering*, *105*(3), 439–460. http://doi.org/10.1002/bit.22592

O'Brien, E. J., Monk, J. M., & Palsson, B. O. (2015). Using Genome-scale Models to Predict Biological Capabilities. *Cell*, *161*(5), 971–987. http://doi.org/10.1016/j.cell.2015.05.019

Pallela, R., Koigoora, S., Gunda, V. G., Sunkara, M. S., & Janapala, V. R. (2011). Comparative morphometry, biochemical and elemental composition of three marine sponges (Petrosiidae) from Gulf of Mannar, India. *Chemical Speciation and Bioavailability*, *23*(1), 16–23. http://doi.org/10.3184/095422911X12966340771966

Palumbi, S. R. (1986). How body plans limit acclimation: responses of a demosponge to wave force. *Ecology*, 208–214.

Pawlik, J. R., Burkepile, D. E., & Thurber, R. V. (2016). A Vicious Circle? Altered Carbon and Nutrient Cycling May Explain the Low Resilience of Caribbean Coral Reefs. *Bioscience*, *66*(6), 470–476. http://doi.org/10.1093/biosci/biw047

Payne, S. H., & Loomis, W. F. (2006). Retention and Loss of Amino Acid Biosynthetic Pathways Based on Analysis of Whole-Genome Sequences. *Eukaryotic Cell*, *5*(2), 272–276. http://doi.org/10.1128/EC.5.2.272-276.2006

Pile, A. J., Patterson, M. R., & Witman, J. D. (1996a). In situ grazing on plankton <10 μm by the boreal sponge *Mycale lingua*. *Marine Ecology Progress Series*. 141, 95-102.

Pinheiro, J., Bates, D., DebRoy, S., & Sarkar, D. (2013). nlme: Linear and Nonlinear Mixed Effects Models. R Package Version 2013. Available online at http://CRAN.R-project.org/package=nlme (accessed 26 October 2013).

Plez, K. A., & Gross, J. (1958). The amino acid composition and morphology of some invertebrate and vertebrate collagens. *Biochimica Et Biophysica Acta*, *34*, 24–39. http://doi.org/10.1016/0006-3002(59)90229-X

Price, N. D., Famili, I., Beard, D. A., & Palsson, B. O. (2002). Letter to the Editor. *Biophysical Journal*, *83*(5), 2879–2882. http://doi.org/10.1016/S0006-3495(02)75297-1

Quek, L.-E., Dietmair, S., Krömer, J. O., & Nielsen, L. K. (2010). Metabolic flux analysis in mammalian cell culture. *Metabolic Engineering*, *12*(2), 161–171. http://doi.org/10.1016/j.ymben.2009.09.002

Radax, R., Hoffmann, F., Rapp, H. T., Leininger, S., & Schleper, C. (2011). Ammonia-oxidizing archaea as main drivers of nitrification in cold-water sponges. *Environmental Microbiology*, *14*(4), 909–923. http://doi.org/10.1111/j.1462-2920.2011.02661.x

Radjasa, O. K., Vaske, Y. M., Navarro, G., Vervoort, H. C., Tenney, K., Linington, R. G., & Crews, P. (2011). Bioorganic & Medicinal Chemistry. *Bioorganic & Medicinal Chemistry*, *19*(22), 6658–6674. http://doi.org/10.1016/j.bmc.2011.07.017

Rao, P., & Pattabiraman, T. N. (1989). Reevaluation of the phenol-sulfuric acid reaction for the estimation of hexoses and pentoses. *Analytical Biochemistry*, *181*(1), 18–22.

Redfield, A. C. (1934). On the proportions of organic derivations in sea water and their relation to the composition of plankton. In R. J. Daniel (Ed.), *James Johnstone Memorial Volume* (pp. 176–192).

Reed, J. L., Patel, T. R., Chen, K. H., Joyce, A. R., Applebee, M. K., Herring, C. D., et al. (2006). Systems approach to refining genome annotation. *Proceedings of the National Academy of Sciences*, *103*(46), 17480–17484. http://doi.org/10.1073/pnas.0603364103

Reiswig, H. M. (1971). In situ pumping activities of tropical Demospongiae. *Marine Biology*, *9*(1), 38–50.

Reiswig, H. M. (1981). Partial carbon and energy budgets of the bacteriosponge *Verohgia fistularis* (Porifera: Demospongiae) in Barbados. *Marine Ecology*, *2*(4), 273–293.

Ribes, M., Coma, R., & Gili, J.-M. (1999). Natural diet and grazing rate of the temperate sponge *Dysidea avara* (Demospongiae, Dendroceratida) throughout an annual cycle. *Marine Ecology Progress Series*. 176;179-190.

Rice, E. W., Baird, R. B., Eaton, A. D., & Clesceri, L. S. (2012). Standard Methods for the Examination of Water and Wastewater: American Public Health Association, AW (22nd ed.). American Water Works Association.

Riisgård, H. U., Thomassen, S., Jakobsen, H., Weeks, J. M., & Larsen, P. S. (1993). Suspension-feeding in marine sponges *Halichondria panicea* and *Haliclona urceolus* effects of temperature on filtration-rate and energy-cost of pumping. *Marine Ecology Progress Series*, *96*(2), 177–188.

Rinkevich, Baruch. (2011). Cell Cultures from Marine Invertebrates: New Insights for Capturing Endless Stemness. *Marine Biotechnology*, *13*(3), 345–354. http://doi.org/10.1007/s10126-010-9354-3

Rix, L., de Goeij, J. M., Mueller, C. E., Struck, U., Middelburg, J. J., van Duyl, F. C., et al. (2016). Coral mucus fuels the sponge loopin warm- and cold-water coral reefecosystems. *Scientific Reports*, 1–11. http://doi.org/10.1038/srep18715

Rodkina, S. (2005). Fatty Acids and Other Lipids of Marine Sponges. *Russian Journal of Marine Biology/Biologiya Morya*, *31*(0), S49–S60. http://doi.org/10.1007/s11179-006-0015-3

Rodríguez, W., Osorno, O., Ramos, F. A., Duque, C., & Zea, S. (2010). Biochemical Systematics and Ecology. *Biochemical Systematics and Ecology*, *38*(4), 774–783. http://doi.org/10.1016/j.bse.2010.07.011

Roff, G., & Mumby, P. J. (2012). Global disparity in the resilience of coral reefs. *Trends in Ecology & Evolution*, *27*(7), 404–413. http://doi.org/10.1016/j.tree.2012.04.007

Rose, W. C., Oesterling, M. J., & Womack, M. (1948). Comparative growth on diets containing ten and nineteen ammo acids, with further observations upon the role of glutamic and aspartic acids. *Journal of Biological Chemistry,* 176;753-762.

Rua, C. P. J., Gregoracci, G. B., Santos, E. O., Soares, A. C., Francini-Filho, R. B., & Thompson, F. (2015). Potential metabolic strategies of widely distributed holobionts in the oceanic archipelago of St Peter and St Paul (Brazil). *FEMS Microbiology Ecology*, *91*(6), http://doi.org/10.1093/femsec/fiv043

Sammarco, P. W., Risk, M. J., Schwarcz, H. P., & Heikoop, J. M. (1999). Cross-continental shelf trends in coral δ15N on the Great Barrier Reef: further consideration of the reef nutrient paradox, *180*, 131–138, http://cat.inist.fr/?aModele=afficheN&cpsidt=1850818

Santos, F., Boele, J., & Teusink, B. (2011). A Practical Guide to Genome-Scale Metabolic Models and Their Analysis. *Methods in Systems Biology* (1st ed., Vol. 500, pp. 509–532). Elsevier Inc. http://doi.org/10.1016/B978-0-12-385118-5.00024-4

Santos, I. R., Erler, D., Tait, D., & Eyre, B. D. (2010). Breathing of a coral cay: Tracing tidally driven seawater recirculation in permeable coral reef sediments. *Journal of Geophysical Research*, *115*(C12), http://doi.org/10.1029/2010JC006510

Satish Kumar, V., Dasika, M. S., & Maranas, C. D. (2007). Optimization based automated curation of metabolic reconstructions. *BMC Bioinformatics*, *8*(1), 212. http://doi.org/10.1186/1471-2105-8-212

Sauer, U. (2006). Metabolic networks in motion: 13C-based flux analysis. *Molecular Systems Biology*, *2*. http://doi.org/10.1038/msb4100109

Schutze, J., Skorokhod, A., Muller, I. M., & Muller, W. E. G. (2001). Molecular Evolution of the Metazoan Extracellular Matrix: Cloning and Expression of Structural Proteins from the Demosponges *Suberites domuncula* and *Geodia cydonium*. *Journal of Molecular Evolution*, *53*(4-5), 402–415. http://doi.org/10.1007/s002390010230

Schellenberger, J., Que, R., Fleming, R. M. T., Thiele, I., Orth, J. D., Feist, A. M., et al. (2011). Quantitative prediction of cellular metabolism with constraint-based models: the COBRA Toolbox v2.0. *Nature Protocols*, *6*(9), 1290–1307. http://doi.org/10.1038/nprot.2011.308

Schippers, K. J., Sipkema, D., Osinga, R., Smidt, H., Pomponi, S. A., Martens, D. E., & Wijffels, R. H. (2012). Cultivation of Sponges, Sponge Cells and Symbionts: Achievements and Future Prospects. *Advances in Sponge Science: Physiology, Chemical and Microbial Diversity, Biotechnology* (1st ed., Vol. 62, pp. 273–337). Elsevier Ltd. http://doi.org/10.1016/B978-0-12-394283-8.00006-0

Schirmer, A., Gadkari, R., Reeves, C. D., Ibrahim, F., DeLong, E. F., & Hutchinson, C. R. (2005). Metagenomic Analysis Reveals Diverse Polyketide Synthase Gene Clusters in Microorganisms

Associated with the Marine Sponge *Discodermia dissoluta*. *Applied and Environmental Microbiology*, *71*(8), 4840–4849. http://doi.org/10.1128/AEM.71.8.4840-4849.2005

Schleper, C., DeLong, E. F., Preston, C. M., Feldman, R. A., Wu, K. Y., & Swanson, R. V. (1998). Genomic analysis reveals chromosomal variation in natural populations of the uncultured psychrophilic archaeon *Cenarchaeum symbiosum*. *Journal of Bacteriology*, *180*(19), 5003–5009.

Sengupta, T., Bhushan, M., & Wangikar, P. P. (2013). Metabolic modeling for multi-objective optimization of ethanol production in a Synechocystis mutant. *Photosynthesis Research*, *118*(1-2), 155–165. http://doi.org/10.1007/s11120-013-9935-x

Shoaie, S., Karlsson, F., Mardinoglu, A., Nookaew, I., Bordel, S., & Nielsen, J. (2013). Understanding the interactions between bacteria in the human gut through metabolic modeling. *Scientific Reports*, http://doi.org/10.1038/srep02532

Shore, R. E. (1971). Growth and renewal studies of the choanocyte population in *Hymeniacidon sinapium* (Porifera: Demospongiae) using colcemid and 3 - H thymidine. *Journal of Experimental Zoology*.

Simpson, T. L. (2012). The Cell Biology of Sponges. Springer Science & Business Media.

Sipkema, D., de Caralt, S., Morillo, J. A., Al-Soud, W. A., Sørensen, S. J., Smidt, H., & Uriz, M. J. (2015). Similar sponge-associated bacteria can be acquired via both vertical and horizontal transmission. *Environmental Microbiology*, *17*(10), 3807–3821. http://doi.org/10.1111/1462-2920.12827

Sipkema, D., Osinga, R., Schatton, W., Mendola, D., Tramper, J., & Wijffels, R. H. (2005). Large-scale production of pharmaceuticals by marine sponges: Sea, cell, or synthesis? *Biotechnology and Bioengineering*, *90*(2), 201–222. http://doi.org/10.1002/bit.20404

Smith, J. S., & Johnson, C. R. (1995). Nutrient inputs from seabirds and humans on a populated coral cay. *Marine Ecology Progress Series*, *124*, 189–200.

Sorokin, Y. I., & Sorokin, P. Y. (2009). Analysis of plankton in the southern Great Barrier Reef: abundance and roles in throphodynamics. *Journal of the Marine Biological Association of the UK*, *89*(02), 235. http://doi.org/10.1017/S0025315409003063

Southwell, M. W., Weisz, J. B., & Martens, C. S. (2008). In situ fluxes of dissolved inorganic nitrogen from the sponge community on Conch Reef, Key Largo, Florida. *Limnology and Oceanography,* 53(3), 986-996. hppt://dx.doi.org/10.4319/1o2008.53.3.0986

Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M. E. A., Mitros, T., Richards, G. S., Conaco, C., Dacre, M., Hellsten, U., Larroux, C., Putnam, N. H., Stanke, M., Adamska, M., Darling,

A., Degnan, S., Oaklet, T. H., Plachetzki, D. C., Zhai, Y., Adamski, M., Calcino, A., Cummins, S. F., Goodstein, D. M., Harris, C., Jackson, D. J., Leys, S. P., Shu, S., Woodcroft, B. J., Vervoot, M., Kosik, K. S., Manning, G., Degnan, B & Rokhsar, D. S. (2010). The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature*, *466*(7307), 720–726. http://doi.org/10.1038/nature09201

Stephanopoulos, G., Aristidou, A. A., & Nielsen, J. (1998). Metabolic Engineering. Academic Press.

Sterner, R. W., Andersen, T., Elser, J. J., Hessen, D. O., Hood, J. M., McCauley, E., & Urabe, J. (2008). Scale-dependent carbon : nitrogen : phosphorus seston stoichiometry in marine and freshwaters. *Limnology and Oceanography*, *53*(3), 1169–1180.

Steuer, R., Gross, T., Selbig, J., & Blasius, B. (2006). Structural kinetic modeling of metabolic networks, 103;11868–11873. Presented at the Proceedings of the National Academy of Sciences. http://doi.org/10.1073/pnas.0600013103

Stolyar, S., Van Dien, S., Hillesland, K. L., Pinel, N., Lie, T. J., Leigh, J. A., & Stahl, D. A. (2007). Metabolic modeling of a mutualistic microbial community. *Molecular Systems Biology*, *3*. http://doi.org/10.1038/msb4100131

Stone, A. R. (1969). Seasonal variation in the gross biochemical composition of *Hymeniacidon perleve* (Montagu). *Journal of Experimental Marine Biology and Ecology*, *5*(3), 265–271. http://doi.org/10.1016/0022-0981(70)90005-5

Taylor, K. A. (1995). A modification of the phenol/sulfuric acid assay for total carbohydrates giving more comparable absorbances. *Applied Biochemistry and Biotechnology*, *53*(3), 207–214.

Taylor, M. W., Radax, R., Steger, D., & Wagner, M. (2007). Sponge-Associated Microorganisms: Evolution, Ecology, and Biotechnological Potential. *Microbiology and Molecular Biology Reviews*, *71*(2), 295–347. http://doi.org/10.1128/MMBR.00040-06

Taylor, M. W., Tsai, P., Simister, R. L., Deines, P., Botte, E., Ericson, G., et al. (2013). "Sponge-specific" bacteria are widespread (but rare) in diverse marine environments. *The ISME Journal*, *7*(2), 438–443. http://doi.org/10.1038/ismej.2012.111

Team, R. S. (2015). RStudio: Integrated Development for R. RStudio, Inc., Boston, MA Retrieved from http://www.rstudio.com/

Thakur, N. L., & Müller, W. E. (2004). Biotechnological potential of marine sponges. *Current Science*, *86*(11), 1506–1512.

Thiele, I., & Palsson, B. O. (2010). A protocol for generating a high-quality genome-scale metabolic reconstruction. *Nature Protocols*, *5*(1), 93–121. http://doi.org/10.1038/nprot.2009.203

Thomas, T. R. A., Kavlekar, D. P., & Loka Bharathi, P. A. (2010). Marine Drugs from Sponge-Microbe Association—A Review. *Marine Drugs*, *8*(4), 1417–1468. http://doi.org/10.3390/md8041417

Thorleifsson, S. G., & Thiele, I. (2011). rBioNet: A COBRA toolbox extension for reconstructing high-quality biochemical networks. *Bioinformatics*, *27*(14), 2009–2010. http://doi.org/10.1093/bioinformatics/btr308

Tian, R.-M., Wang, Y., Bougouffa, S., Gao, Z.-M., Cai, L., Bajic, V., & Qian, P.-Y. (2014). Genomic analysis reveals versatile heterotrophic capacity of a potentially symbiotic sulfur-oxidizing bacterium in sponge. *Environmental Microbiology*, *16*(11), 3548–3561. http://doi.org/10.1111/1462-2920.12586

Topçu, N. E., Pérez, T., Grégori, G., & Harmelin-Vivien, M. (2010). Journal of Experimental Marine Biology and Ecology. *Journal of Experimental Marine Biology and Ecology*, *389*(1-2), 61–69. http://doi.org/10.1016/j.jembe.2010.03.017

Turon, X., Galera, J., & Uriz, M. J. (1997). Clearance rates and aquiferous systems in two sponges with contrasting life - history strategies. *Journal of Experimental Zoology*. 278(1), 22-36.

Uchino, K., Saito, T., Gebauer, B., & Jendrossek, D. (2007). Isolated Poly(3-Hydroxybutyrate) (PHB) Granules Are Complex Bacterial Organelles Catalyzing Formation of PHB from Acetyl Coenzyme A (CoA) and Degradation of PHB to Acetyl-CoA. *Journal of Bacteriology*, *189*(22), 8250–8256. http://doi.org/10.1128/JB.00752-07

Vacelet, J., Verdenal, B., & Perinet, G. (1988). The iron mineralization of *Spongia officinalis* L.(Porifera, Dictyoceratida) and its relationships with the collagen skeleton. *Biology of the Cell*, *62*(2), 189–198.

van Duyl, F. C., Moodley, L., Nieuwland, G., van Ijzerloo, L., Van Soest, R. W. M., Houtekamer, M., et al. (2011). Coral cavity sponges depend on reef-derived food resources: stable isotope and fatty acid constraints. *Marine Biology*, *158*(7), 1653–1666. http://doi.org/10.1007/s00227-011-1681-z

van Duyl, F., & Jan Gast, G. (2001). Linkage of small-scale spatial variations in DOC, inorganic nutrients and bacterioplankton growth with different coral reef water types. *Aquatic Microbial Ecology*, *24*, 17–26. http://doi.org/10.3354/ame024017

Voogt, P. A. (1972). Sterols of some echinoids. *Archives of Physiology and Biochemistry*, *80*(5), 883–891.

Wang, G. (2006). Diversity and biotechnological potential of the sponge-associated microbial consortia. *Journal of Industrial Microbiology & Biotechnology*, *33*(7), 545–551. http://doi.org/10.1007/s10295-006-0123-2

Wang, X., Zhang, C., Wang, M., & Lu, W. (2014). Genome-scale metabolic network reconstruction of *Saccharopolyspora spinosa* for Spinosad Production improvement. *Microbial Cell Factories*, *13*(1), 1–7. http://doi.org/10.1186/1475-2859-13-41

Watson, J., Brennan, T., Degnan, B., Degnan, S., & Krömer, J. (2014a). Analysis of the Biomass Composition of the Demosponge *Amphimedon queenslandica* on Heron Island Reef, Australia. *Marine Drugs*, *12*(6), 3733–3753. http://doi.org/10.3390/md12063733

Watson, J., Degnan, B., Degnan, S., & Krömer, J. O. (2014b). Determining the Biomass Composition of a Sponge Holobiont for Flux Analysis. In *Methods in Molecular Biology* (Vol. 1191, pp. 107–125). New York, NY: Springer New York. http://doi.org/10.1007/978-1-4939-1170-7_7

Webster, N. S., & Blackall, L. L. (2008). What do we really know about sponge-microbial symbioses? *The ISME Journal*, *3*(1), 1–3. http://doi.org/10.1038/ismej.2008.102

Webster, N. S., & Taylor, M. W. (2011). Marine sponges and their microbial symbionts: love and other relationships. *Environmental Microbiology*, *14*(2), 335–346. http://doi.org/10.1111/j.1462-2920.2011.02460.x

Webster, N. S., Taylor, M. W., Behnam, F., Lücker, S., Rattei, T., Whalan, S., et al. (2009). Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts. *Environmental Microbiology*. http://doi.org/10.1111/j.1462-2920.2009.02065.x

Weisz, J. B., Hentschel, U., Lindquist, N., & Martens, C. S. (2007a). Linking abundance and diversity of sponge-associated microbial communities to metabolic differences in host sponges. *Marine Biology*, *152*(2), 475–483. http://doi.org/10.1007/s00227-007-0708-y

Weisz, J. B., Lindquist, N., & Martens, C. S. (2007b). Do associated microbial abundances impact marine demosponge pumping rates and tissue densities? *Oecologia*, *155*(2), 367–376. http://doi.org/10.1007/s00442-007-0910-0

Weisz, J. B., Massaro, A. J., Ramsby, B. D., & Hill, M. S. (2010). Zooxanthellar symbionts shape host sponge trophic status through translocation of carbon. *The Biological Bulletin*, *219*(3), 189–197.

Wijffels, R. H. (2008). Potential of sponges and microalgae for marine biotechnology. *Trends in Biotechnology*, *26*(1), 26–31. http://doi.org/10.1016/j.tibtech.2007.10.002

Wild, C., Huettel, M., Klueter, A., Kremb, S. G., Rasheed, M. Y. M., & Jørgensen, B. B. (2004a). Coral mucus functions as an energy carrier and particle trap in the reef ecosystem. *Nature*, *428*(6978), 66–70. http://doi.org/10.1038/nature02344

Wild, C., Naumann, M., Niggl, W., & Haas, A. (2010). Carbohydrate composition of mucus released by scleractinian warm- and cold-water reef corals. *Aquatic Biology*, *10*(1), 41–45. http://doi.org/10.3354/ab00269

Wild, C., Rasheed, M., Werner, U., Franke, U., & Johnstone, R. (2004b). Degradation and mineralization of coral mucus in reef environments. *Marine Ecology Progress Series*, *267*, 159–171.

Wild, C., Woyt, H., & Heuttel, M. (2005). Influence of coral mucus on nutrient fluxes in carbonate sands. *Marine Ecology Progress Series*, *287*, 87–98.

Wilkinson, C. (1982). Significance of sponges with cyanobacterial symbionts on Davies Reef, Great Barrier Reef (pp. 1–8). Proceedings of the 4th International Coral Reef Symposium, Manilla, Phillipines.

Wilkinson, C. R. (1983). Net primary productivity in coral reef sponges. *Science*, *219*(4583), 410–412. http://doi.org/10.1126/science.219.4583.410

Wilson, M. C., Mori, T., Rückert, C., Uria, A. R., Helf, M. J., Takada, K., et al. (2014). An environmental bacterial taxon with a large and distinct metabolic repertoire. *Nature*, 1–18. http://doi.org/10.1038/nature12959

Wittmann, C. (2007). Fluxome analysis using GC-MS. *Microbial Cell Factories*, *6*, 6. http://doi.org/10.1186/1475-2859-6-6

Wulff, J. (2001). Assessing and monitoring coral reef sponges: why and how?, *69*(2), 831–846.

Xie, L., & Wang, D. I. (2003). Energy metabolism and ATP balance in animal cell cultivation using a stoichiometrically based reaction network. *Biotechnology and Bioengineering*, *52*(5), 591–601. http://doi.org/10.1002/(SICI)1097-0290(19961205)52:5<591::AID-BIT6>3.0.CO;2-E

Yahel, G., Sharp, J. H., Marie, D., & Hase, C. (2003a). In situ feeding and element removal in the symbiont - bearing sponge *Theonella swinhoei*: Bulk DOC is the major source for carbon. *Limnology and Oceanography*, *48*(1), 141–149. Retrieved from http://onlinelibrary.wiley.com/doi/10.4319/lo.2003.48.1.0141/abstract

Yahel, G., Sharp, J. H., Marie, D., Hase, C., & Genin, A. (2003b). In situ feeding and element removal in the symbiont-bearing sponge *Theonella swinhoei*: Bulk DOC is the major source for carbon. *Limnology and Oceanography*, *48*(1), 141–149.

Yeagle, P. L. (1991). Modulation of membrane function by cholesterol. *Biochimie*, *73*(10), 1303–1310.

Yilmaz, L. S., & Walhout, A. J. M. (2016). A *Caenorhabditis elegans* Genome-Scale Metabolic Network Model. *Cell Systems*, *2*(5), 297–311. http://doi.org/10.1016/j.cels.2016.04.012

Zamboni, N. (2011). C metabolic flux analysis in complex systems. *Current Opinion in Biotechnology*, *22*(1), 103–108. http://doi.org/10.1016/j.copbio.2010.08.009

Zamboni, N., Fendt, S.-M., Rühl, M., & Sauer, U. (2009). 13C-based metabolic flux analysis. *Nature Protocols*, *4*(6), 878–892. http://doi.org/10.1038/nprot.2009.58

Zhang, F., Blasiak, L. C., Karolin, J. O., Powell, R. J., Geddes, C. D., & Hill, R. T. (2015). Phosphorus sequestration in the form of polyphosphate by microbial symbionts in marine sponges. *Proceedings of the National Academy of Sciences*, *112*(14), 201423768. http://doi.org/10.1073/pnas.1423768112

Zhang, X., Pennec, G. L., Steffen, R., Müller, W. E. G., & Zhang, W. (2008). Application of a MTT Assay for Screening Nutritional Factors in Growth Media of Primary Sponge Cell Culture. *Biotechnology Progress*, *20*(1), 151–155. http://doi.org/10.1021/bp0341601

Zhao, Q., Zhang, W., Jin, M., Yu, X., & Deng, M. (2008). Formulation of a Basal Medium for Primary Cell Culture of the Marine Sponge *Hymeniacidon perleve*. *Biotechnology Progress*, *21*(3), 1008–1012. http://doi.org/10.1021/bp050029c

Zomorrodi, A. R., & Maranas, C. D. (2012). OptCom: A Multi-Level Optimization Framework for the Metabolic Modeling and Analysis of Microbial Communities. *PLoS Computational Biology*, *8*(2), e1002363. http://doi.org/10.1371/journal.pcbi.1002363.s002

APPENDICES

Several appendices are impractiacally large for print format. These files are available to download via Cloudstor+, a cloud storage service run by the Australian Academic and Research Network (AARNet). These appendices are referenced in-text and listed below in order.

Download information for these files is as follows:

URL: https://cloudstor.aarnet.edu.au/plus/index.php/s/AP5EYw4hANDXd1S

Password: spongemodel

Link expiry date: none

Appendix 4.1 Complete dataset for the seasonal analysis of water parameters in Shark Bay, Heron Island Reef from September 2013 to October 2014

December, 2013	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8
Date sampled	17/12/13	17/12/13	17/12/13	18/12/13	19/12/13	19/12/13	19/12/13	20/12/13
Tide height	High	Low	High	Low	High	Low	High	Low
TOC (mg/L)	0.86	0.96	1	1.1	0.96	1.2	0.94	0.97
DOC (mg/L)	0.9	0.9	1	1	1	1.1	0.9	1
Total nitrogen (mg/L)	<0.05	<0.05	<0.05	0.06	<0.05	0.1	<0.05	<0.05
Ammonia (mg/L)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	0.038
Nitrite (mg/L)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Nitrate (mg/L)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Total phosphorous (mg/L)	0.02	0.02	0.02	0.02	0.02	0.02	0.02	0.02
Orthophosphate (mg/L)	0.008	0.006	0.006	0.006	0.005	0.005	<0.005	0.006
Sulfate (mg/L)	1.88	1.64	1.82	1.35	1.75	1.68	1.80	1.74
Sulfide (mg/L)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Silicon (mg/L)	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1
Samples where the concentration was below the level of detection are represented by at < sign, followed by the detection limit.								

October, 2014	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8
Date sampled	6/10/14	6/10/14	6/10/14	7/10/14	8/10/14	8/10/14	8/10/14	9/10/14
Tide height	High	Low	High	Low	High	Low	High	Low
TOC (mg/L)	3.3	3.8	3.6	3.6	3.3	2.9	3.4	3.4
DOC (mg/L)	2.5	2.6	2.7	2.7	2.6	2.5	2.7	2.6
Total nitrogen (mg/L)	<0.05	<0.05	0.07	0.073	0.075	0.057	0.12	0.11
Ammonia (mg/L)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Nitrite (mg/L)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Nitrate (mg/L)	<0.005	<0.005	0.005	0.006	<0.005	0.005	0.008	0.03
Total phosphorous (mg/L)	0.03	0.02	0.034	0.028	0.037	0.026	0.037	0.026
Orthophosphate (mg/L)	0.007	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Sulfate (mg/L)	3.76	4.12	3.97	3.03	3.56	3.33	3.58	3.01
Sulfide (mg/L)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Silicon (mg/L)	0.8	0.91	0.87	0.9	0.95	1	1.1	0.93
Samples where the concentration was below the level of detection are represented by at < sign, followed by the detection limit.								

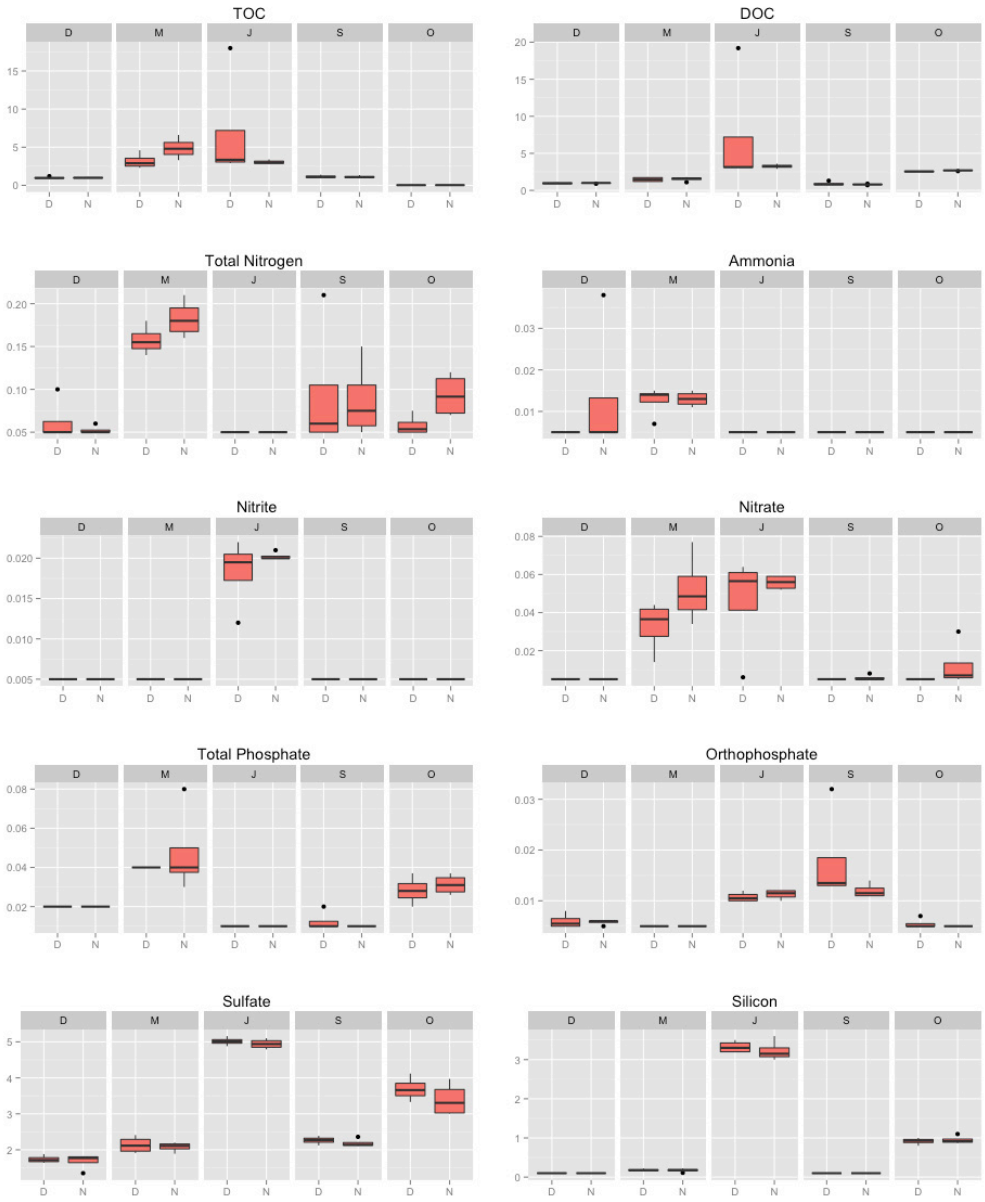
March, 2014	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	Sample 6	Sample 7	Sample 8
Date sampled	15/03/14	15/03/14	15/03/14	16/03/14	17/03/14	17/03/14	17/03/14	18/03/14
Tide height	High	Low	High	Low	High	Low	High	Low
TOC (mg/L)	4.6	3.2	5.3	4.3	2.3	2.6	6.6	3.3
DOC (mg/L)	1.2	1.8	1.6	1.6	1.7	1.2	1.8	1.1
Total nitrogen (mg/L)	0.14	0.18	0.19	0.16	0.16	0.15	0.21	0.17
Ammonia (mg/L)	0.007	0.015	0.015	0.011	0.014	0.014	0.014	0.012
Nitrite (mg/L)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Nitrate (mg/L)	0.032	0.041	0.034	0.044	0.044	0.014	0.053	0.077
Total phosphorous (mg/L)	0.04	0.04	0.04	0.08	0.04	0.04	0.04	0.03
Orthophosphate (mg/L)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Sulfate (mg/L)	1.98	2.41	2.15	2.07	2.25	1.91	2.20	1.89
Sulfide (mg/L)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Silicon (mg/L)	0.17	0.18	0.22	0.18	0.16	0.22	0.18	0.11
Samples where the concentration was below the level of detection are represented by at < sign, followed by the detection limit.								

July, 2014	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 4**	**Sample 5**	**Sample 6**	**Sample 7**	**Sample 8**
Date sampled	10/07/14	10/07/14	10/07/14	11/07/14	12/07/14	12/07/14	12/07/14	13/07/14
Tide height	High	Low	High	Low	High	Low	High	Low
TOC (mg/L)	3.1	3.6	2.9	2.8	18	2.9	3.4	3.1
DOC (mg/L)	3.1	3.1	2.9	3.6	19.2	3.2	3.2	3.3
Total nitrogen (mg/L)	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05
Ammonia (mg/L)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Nitrite (mg/L)	0.022	0.012	0.02	0.02	0.019	0.02	0.02	0.021
Nitrate (mg/L)	0.053	0.006	0.059	0.052	0.06	0.064	0.053	0.059
Total phosphorous (mg/L)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Orthophosphate (mg/L)	0.012	0.01	0.011	0.01	0.011	0.01	0.012	0.012
Sulfate (mg/L)	5.03	5.16	4.79	5.01	4.99	4.88	5.10	4.87
Sulfide (mg/L)	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1
Silicon (mg/L)	3.2	3.5	3.1	3	3.4	3.2	3.6	3.2
Samples where the concentration was below the level of detection are represented by at < sign, followed by the detection limit.								

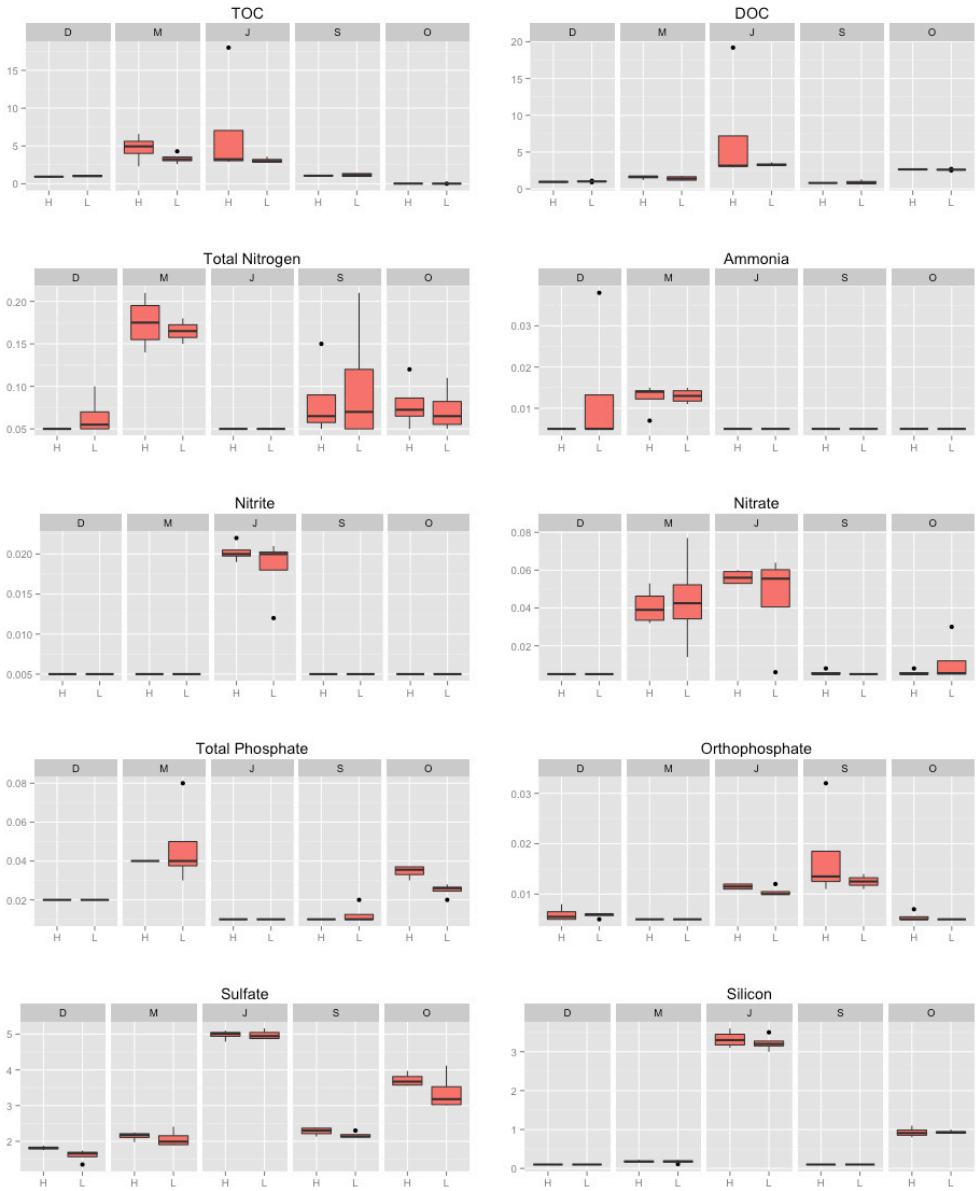
September, 2013	**Sample 1**	**Sample 2**	**Sample 3**	**Sample 4**	**Sample 5**	**Sample 6**	**Sample 7**	**Sample 8**
Date sampled	19/09/13	19/09/13	19/09/13	20/09/13	21/09/13	21/09/13	21/09/13	22/09/13
Tide height	High	Low	High	Low	High	Low	High	Low
TOC (mg/L)	1.15	1	1.1	0.95	1.04	1.4	1.04	1.35
DOC (mg/L)	0.8	0.7	0.8	0.7	0.8	1.3	0.8	0.9
Total nitrogen (mg/L)	<0.05	<0.05	0.06	<0.05	0.07	0.21	0.15	0.09
Ammonia (mg/L)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Nitrite (mg/L)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Nitrate (mg/L)	<0.005	<0.005	0.008	<0.005	<0.005	<0.005	<0.005	<0.005
Total phosphorous (mg/L)	<0.01	<0.01	<0.01	<0.01	<0.01	0.02	<0.01	<0.01
Orthophosphate (mg/L)	0.032	0.013	0.014	0.011	0.013	0.014	0.011	0.012
Sulfate (mg/L)	2.30	2.13	2.36	2.11	2.24	2.30	2.13	2.15
Sulfide (mg/L)	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01	<0.01
Silicon (mg/L)	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1	<0.1
Samples where the concentration was below the level of detection are represented by at < sign, followed by the detection limit.								

Appendix 4.2 Comparison of day and night time availability of major nutrients in Shark Bay, Heron Island Reef from September 2013 to October 2014.

Eight replicate water samples are analysed per seasonal sample, four during the day and four during night, over two non-successive days. The y-axis scales on each parameter are independent. The black dots represent outlier values that are indicative of the variability that can occur within small time frames in this natural system. The months sampled were December, 2013 (D; summer), March, 2014 (M; early autumn), July, 2014 (J; winter), September, 2013 (S; early spring) and October, 2014 (O; mid-spring). DOC, dissolved organic carbon; TOC, total organic carbon.



Appendix 4.3 Comparison of high and low tide availability of major nutrients in Shark Bay, Heron Island Reef from September 2013 to October 2014.
Eight replicate water samples are analysed per seasonal sample, four during high tide and four during low tide, over two non-successive days. The y-axis scales on each parameter are independent. The black dots represent outlier values that are indicative of the variability that can occur within small time frames in this natural system. The months sampled were December, 2013 (D; summer), March, 2014 (M; early autumn), July, 2014 (J; winter), September, 2013 (S; early spring) and October, 2014 (O; mid-spring). DOC, dissolved organic carbon; TOC, total organic carbon.



Appendix 4.4 Complete dataset for the seasonal analysis of biomass composition in adult Amphimedon queenslandica collected from Shark Bay, Heron Island Reef from September 2013 to October 2014.

Three replicate samples from four sponges – 12 samples in total – were analysed. The number after "S" corresponds to the individual sponge. The number after "R" corresponds to the replicate sample for each sponge.

December, 2013	S1R1	S1R2	S1R3	S2R1	S2R2	S2R3	S3R1	S3R2	S3R3	S4R1	S4R2	S4R3
Skeleton (g/gdw)	0.55181	0.59749	0.55458	0.72224	0.40656	0.61115	0.65045	0.62620	0.46604	0.55175	0.53494	0.56862
Lipid (g/gdw)	0.11462	0.08464	0.10488	0.08175	0.10294	-0.0173	0.06651	0.05727	0.07780	0.06345	0.08941	0.11560
Protein (g/gdw)	0.10294	0.09243	0.11533	0.10245	0.08882	0.11226	0.11046	0.10032	0.12289	0.08984	0.09818	0.11134
Carbohydrate (g/gdw)	0.02493	0.03120	0.01999	0.02486	0.01771	0.02758	0.03231	0.02602	0.04020	0.03459	0.03014	0.03149
RNA (g/gdw)	0.00087	0.00172	0.00122	0.00200	0.00161	0.00197	0.00124	0.00125	0.00198	0.00180	0.00177	0.00211
DNA (g/gdw)	0.00034	0.00060	0.00027	0.00032	0.00030	0.00025	0.00051	0.00022	0.00030	0.00049	0.00042	0.00048

March, 2013	S1R1	S1R2	S1R3	S2R1	S2R2	S2R3	S3R1	S3R2	S3R3	S4R1	S4R2	S4R3
Skeleton (g/gdw)	0.46033	0.51448	0.52949	0.54985	0.60903	0.55908	0.49745	0.59679	0.54558	0.64941	0.68724	0.67553
Lipid (g/gdw)	0.10335	0.11854	0.11892	0.10458	0.08691	0.08051	0.12062	0.12059	0.11501	0.09049	0.07823	0.10635
Protein (g/gdw)	0.12710	0.10626	0.10989	0.09585	0.08939	0.11172	0.13013	0.14444	0.10963	0.16936	0.17031	0.14425
Carbohydrate (g/gdw)	0.02262	0.03441	0.03961	0.02562	0.02773	0.02971	0.03768	0.03986	0.03575	0.03606	0.03379	0.01990
RNA (g/gdw)	0.00134	0.00137	0.00230	0.00170	0.00199	0.00134	0.00126	0.00178	0.00189	0.00115	0.00143	0.00204
DNA (g/gdw)	0.00035	0.00020	0.00027	0.00023	0.00038	0.00023	0.00024	0.00022	0.00028	0.00017	0.00009	0.00009

July, 2013	S1R1	S1R2	S1R3	S2R1	S2R2	S2R3	S3R1	S3R2	S3R3	S4R1	S4R2	S4R3
Skeleton (g/gdw)	0.46948	0.58322	0.59986	0.64638	0.59554	0.62797	0.58347	0.59739	0.45591	0.64229	0.72581	0.68083
Lipid (g/gdw)	0.12070	0.10425	0.12754	0.08787	0.07543	0.10184	0.13704	0.14788	0.17261	0.11071	0.09449	0.12859
Protein (g/gdw)	0.11227	0.11074	0.14375	0.11294	0.11805	0.11546	0.15578	0.12350	0.06942	0.21547	0.11042	0.11505
Carbohydrate (g/gdw)	0.02541	0.01964	0.01757	0.02026	0.02403	0.02695	0.03196	0.03641	0.02919	0.03350	0.01780	0.01541
RNA (g/gdw)	0.00108	0.00136	0.00115	0.00127	0.00093	0.00231	0.00284	0.00213	0.00435	0.00118	0.00119	0.00121
DNA (g/gdw)	0.00031	0.00018	0.00023	0.00096	0.00015	0.00005	0.00023	0.00041	0.00049	0.00003	0.00016	0.00019

September, 2013	S1R1	S1R2	S1R3	S2R1	S2R2	S2R3	S3R1	S3R2	S3R3	S4R1	S4R2	S4R3
Skeleton (g/gdw)	0.63094	0.68783	0.67083	0.65988	0.66657	0.68158	0.58430	0.58121	0.65009	0.62423	0.65606	0.63608
Lipid (g/gdw)	0.13808	0.11301	0.11560	0.08911	0.10229	0.09994	0.10188	0.12988	0.11427	0.09025	0.10862	0.13372
Protein (g/gdw)	0.10497	0.10927	0.09719	0.08867	0.09233	0.08664	0.07611	0.08437	0.10077	0.11045	0.08940	0.10962
Carbohydrate (g/gdw)	0.02160	0.02477	0.02325	0.01625	0.02278	0.02385	0.03333	0.02807	0.02566	0.03060	0.02466	0.02112
RNA (g/gdw)	0.00120	0.00141	0.00148	0.00309	0.00089	0.00128	0.00451	0.00225	0.00100	0.00159	0.00188	0.00205
DNA (g/gdw)	0.00019	0.00012	0.00007	0.00017	0.00025	0.00020	0.00022	0.00031	0.00033	0.00024	0.00022	0.00075

October, 2013	S1R1	S1R2	S1R3	S2R1	S2R2	S2R3	S3R1	S3R2	S3R3	S4R1	S4R2	S4R3
Skeleton (g/gdw)	0.57331	0.53909	0.58477	0.61317	0.57504	0.44791	0.75335	0.70900	0.60482	0.54349	0.50270	0.57637
Lipid (g/gdw)	0.07351	0.09523	0.07300	0.08002	0.07453	0.10411	0.08447	0.06945	0.08482	0.10917	0.09668	0.08204
Protein (g/gdw)	0.09365	0.15513	0.11597	0.10774	0.10277	0.12732	0.07359	0.07959	0.14192	0.12142	0.04034	0.12826
Carbohydrate (g/gdw)	0.02078	0.02484	0.02128	0.02503	0.01931	0.02517	0.01475	0.01619	0.01661	0.02954	0.04038	0.03477
RNA (g/gdw)	0.00167	0.00296	0.00174	0.00213	0.00106	0.00115	0.00123	0.00166	0.00126	0.00216	0.00140	0.00184
DNA (g/gdw)	0.00040	0.00047	0.00037	0.00027	0.00021	0.00023	0.00018	0.00037	0.00060	0.00036	0.00049	0.00046

Appendix 4.5 Correlation in seasonal changes in all biomass components and environmental parameters.
Numerical correlation matrix below the diagonal and significance above. 95 and 99% significance are shaded lighter and darker grey, respectively. Dashed line separates biomass components from environmental parameters.

	Skele.	Lipid	Prot.	Carbo.	RNA	DNA	Temp.	TOC	DOC	Sulf.	Nitro.	Amm.	Nitrite	Nitrat.	Phos.	Ortho.	Silica
Skeleton		0.043	0.773	0.296	0.097	0.006	0.001	0.876	0.635	0.526	0.148	0.095	0.452	0.699	0.002	0.001	0.411
Lipid	0.142		0.452	0.367	0.078	0.013	0.491	0.492	0.340	0.335	0.606	0.749	0.011	0.014	0.212	0.001	0.194
Protein	-0.143	0.283		0.485	0.828	0.834	0.945	0.417	0.451	0.056	0.079	0.478	0.751	0.349	0.074	0.587	0.082
Carbohydrate	-0.288	0.335	0.190		0.007	0.060	0.012	0.124	0.602	0.210	0.003	0.035	0.725	0.013	0.345	0.399	0.487
RNA	0.245	0.318	-0.049	0.415		0.000	0.747	0.285	0.315	0.334	0.437	0.498	0.580	0.053	0.304	0.334	0.785
DNA	-0.248	-0.397	0.089	-0.295	-0.639		0.847	0.010	0.985	0.116	0.362	0.204	0.216	0.062	0.131	0.027	0.983
Temperature	-0.506	-0.112	-0.011	0.392	0.053	-0.032		0.216	0.004	0.000	0.002	0.028	0.001	0.303	0.000	0.001	0.000
TOC	-0.025	0.112	0.132	0.247	0.173	-0.401	-0.200		0.000	0.078	0.307	0.562	0.115	0.000	0.864	0.746	0.011
DOC	0.077	-0.155	0.123	-0.085	-0.163	0.003	-0.450	0.824		0.000	0.284	0.454	0.087	0.013	0.379	0.800	0.000
Sulfate	0.103	0.156	0.305	-0.203	-0.157	0.253	-0.647	0.282	0.542		0.016	0.059	0.003	0.004	0.050	0.291	0.000
Nitrogen	-0.233	0.084	-0.281	0.463	0.126	-0.146	0.480	0.166	-0.174	-0.379		0.040	0.137	0.131	0.000	0.097	0.011
Ammonium	-0.268	-0.052	-0.115	0.334	0.110	-0.205	0.347	0.095	-0.122	-0.301	0.326		0.261	0.462	0.036	0.120	0.118
Nitrite	0.122	0.396	0.052	0.057	0.090	-0.200	-0.500	0.253	0.274	0.466	-0.239	-0.182		0.028	0.009	0.095	0.000
Nitrate	-0.063	0.387	0.152	0.388	0.308	-0.297	-0.167	0.587	0.390	0.451	0.243	0.120	0.348		0.551	0.871	0.000
Phosphorous	-0.469	-0.202	-0.286	0.153	-0.167	0.243	0.548	0.028	-0.143	-0.312	0.620	0.333	-0.408	0.097		0.000	0.015
Orthophosphate	0.498	0.502	0.089	-0.137	0.157	-0.349	-0.502	0.053	0.041	0.171	-0.266	-0.250	0.268	-0.027	-0.553		0.283
Silica	0.134	0.210	0.279	-0.113	0.045	-0.004	-0.617	0.400	0.567	0.941	-0.397	-0.251	0.535	0.555	-0.382	0.174	

Appendix 5.1 Metabolite composition and coefficients for the bacterial food source.

Metabolite	Coefficient (mmol/gdw)
L-Alanine	0.513689
L-Arginine	0.295792
L-Asparagine	0.241055
L-Aspartate	0.241055
CTP	0.133508
L-Cysteine	0.09158
dATP	0.026166
dCTP	0.027017
dGTP	0.027017
TTP	0.026166
L-Glutamine	0.26316
L-Glutamate	0.26316
Glycine	0.612638
GTP	0.215096
L-Histidine	0.094738
L-Isoleucine	0.290529
L-Leucine	0.450531
L-Lysine	0.343161
L-Methionine	0.153686
NH3	0.013013
L-Phenylalanine	0.185265
L-Proline	0.221055
L-Serine	0.215792
L-Threonine	0.253687
L-Tryptophan	0.056843
L-Tyrosine	0.137896
UTP	0.144104
L-Valine	0.423162

Appendix 5.2 Metabolite composition and coefficients for the phytoplankton (algal) food source.

Metabolite	Coefficient (mmol / gdw)
Palmitate	0.298
Starch	0.189
Sucrose	0.378
D-Fructose	0.378
D-Glucose	0.378
Glycine	0.217
L-Alanine	0.531
L-Asparagine	0.132
L-Aspartate	0.132
L-Cysteine	0.023
L-Glutamate	0.175
L-Glutamine	0.175
L-Isoleucine	0.072
L-Leucine	0.177
L-Lysine	0.01
L-Methionine	0.024
L-Phenylalanine	0.076
L-Proline	0.099
L-Serine	0.04
L-Threonine	0.066
L-Tryptophan	0.003
L-Tyrosine	0.003
L-Valine	0.122

Appendix 5.3 Metabolite composition and coefficients for the coral mucus (DOC) food source.

Metabolite	Coefficient (mmol / gdw)
L-Arabinose	1.949
L-Rhamnose	0.294
L-Fucose	0.332
Xylose	0.336
D-Mannose	0.644
Galactose	0.281
D-Glucose	1.296
N-Acetyl-D-glucosamine	0.694

Appendix 5.4 Link to the *A. queenslandica* Pathway Tools database

This requires a local installation of Pathway Tools. This database has not been curated.

URL: https://cloudstor.aarnet.edu.au/plus/index.php/s/AP5EYw4hANDXd1S

Password: spongemodel

Appendix 5.5 Biomass composition of *A. queenslandica*'s skeletal component for both nutrient conditions.

Amino acid	Low nutrient (October) coefficient (mmol/gdw)	High nutrient (March) coefficient (mmol/gdw)
L-alanine	0.658593662	0.645084048
L-arginine	0.287735095	0.281832837
L-aspartate	0.313311548	0.306884644
L-asparagine	0.313311548	0.306884644
L-cysteine	Not included	Not included
L-glutamine	0.300523321	0.29435874
L-glutamate	0.300523321	0.29435874
L-glycine	2.231545515	2.185770222
L-histidine	0.01918234	0.018788856
L-isoleucine	0.089517585	0.087681327
L-leucine	0.153458717	0.150310846
L-lysine	0.230188076	0.225466269

Amino acid	Low nutrient (October) coefficient (mmol/gdw)	High nutrient (March) coefficient (mmol/gdw)
L-methionine	Not included	Not included
L-phenylalanine	0.127882264	0.125259039
L-proline	0.479558492	0.469721394
L-serine	0.281340982	0.275569885
L-threonine	0.294129208	0.288095789
L-tryptophan	0.051152906	0.050103615
L-tyrosine	0.031970566	0.03131476
L-valine	0.230188076	0.225466269

Appendix 5.6 Biomass composition of *A. queenslandica*'s cellular component for both nutrient conditions.

Metabolite	Low nutrient (October) coefficient (mmol/gdw)	High nutrient (March) coefficient (mmol/gdw)
L-alanine	0.245437684	0.233571417
L-arginine	0.2022128	0.192436343
L-aspartate	0.246413859	0.234500397
L-asparagine	0.220868267	0.210189867
L-cysteine	0.100737914	0.095867501
L-glutamine	0.173833112	0.165428738
L-glutamate	0.282438518	0.268783358
L-glycine	0.257986171	0.245513218
L-histidine	0.112270309	0.106842334
L-isoleucine	0.284166262	0.27042757
L-leucine	0.430320911	0.409516027
L-lysine	0.266538505	0.253652069
L-methionine	0.094134356	0.089583208
L-phenylalanine	0.169518518	0.161322743
L-proline	0.214447612	0.204079634
L-serine	0.407377064	0.387681454
L-threonine	0.282685466	0.269018367
L-tryptophan	0.048667874	0.04631491
L-tyrosine	0.154279814	0.14682079
L-valine	0.288308975	0.274369995
dAMP	0.001832964	0.000972448
dCMP	0.000828906	0.000439762
dGMP	0.000828906	0.000439762

Metabolite	Low nutrient (October) coefficient (mmol/gdw)	High nutrient (March) coefficient (mmol/gdw)
dTMP	0.001832964	0.000972448
AMP	0.007769451	0.00589419
CMP	0.005128044	0.003890322
GMP	0.005128044	0.003890322
UMP	0.007769451	0.00589419
Docosanoic acid (C22:0)	0.056007253	0.057805014
Eicosanoic acid (C20:0)	0.013025808	0.01344392
Erucic acid (C22:1)	0.010058354	0.010381214
Myristic acid (C14:0)	0.024433788	0.025218082
Nervonic acid (C24:1)	0.057669998	0.059521131
Octadecanoic (C18:0)	0.355691549	0.367108792
Palmitic acid (C16:0)	0.559162208	0.577110598
Pentadecanoic acid (C15:0)	0.042785952	0.044159327
Tetracosanoic acid (C24:0)	0.039965885	0.041248739
11-Eicosenoic acid (20:1ω9)	0.032489598	0.033532472
Heptadecanoic acid (C17:0)	0.007801235	0.008051645
Tricosanoic acid (C23:0)	0.00638522	0.006590177
Cholesterol	0.062685301	0.064697419
Brassicasterol	0.045522017	0.046983216
Glycogen	0.159899989	0.181958992

Appendix 5.7 Biomass composition of *AqS1*.

Metabolite	Coefficient (mmol/gdw)
L-Alanine	0.464526325
L-Arginine	0.372598856
L-Asparagine	0.119232513
L-Aspartate	0.28848643
L-Cysteine	0.053570035
L-Glutamate	0.287728364
L-Glutamine	0.12048497
L-Glycine	0.366003311
L-Histidine	0.096567422
L-Isoleucine	0.276328068

Metabolite	Coefficient (mmol/gdw)
L-Leucine	0.408667442
L-Lysine	0.159435674
L-Methionine	0.10430556
L-Phenylalanine	0.167192123
L-Proline	0.236714514
L-Serine	0.311997479
L-Threonine	0.20519799
L-Tryptophan	0.054701641
L-Tyrosine	0.092388901
L-Valine	0.273266505
ATP (includes growth associated maintenance)	40.11017574
GTP	0.135406821
CTP	0.084103616
UTP	0.090831905
dATP	0.016020672
dGTP	0.016020672
dCTP	0.016020672
TTP	0.016020672

Appendix 5.8 Link to *AqS1* Pathway Tools database.

This requires a local installation of Pathway Tools. This database has not been curated.

URL: https://cloudstor.aarnet.edu.au/plus/index.php/s/AP5EYw4hANDXd1S

Password: spongemodel

Appendix 5.9 Link to the movies of *A. queenslandica* pumping water.

URL: https://cloudstor.aarnet.edu.au/plus/index.php/s/AP5EYw4hANDXd1S

Password: spongemodel

Appendix 5.10 Link to the water temperature data

URL: https://cloudstor.aarnet.edu.au/plus/index.php/s/AP5EYw4hANDXd1S

Password: spongemodel

Appendix 5.11 Link to the dual-species genome-scale metabolic model files.

Files are in SMBL format. A respective SBML file is included for each of the nutrient (low and high) and objective (biomass synthesis and ATP generation) simulations.

URL: https://cloudstor.aarnet.edu.au/plus/index.php/s/AP5EYw4hANDXd1S

Password: spongemodel

Appendix 5.12 Link to excel file containing flux data.

URL: https://cloudstor.aarnet.edu.au/plus/index.php/s/AP5EYw4hANDXd1S

Password: spongemodel