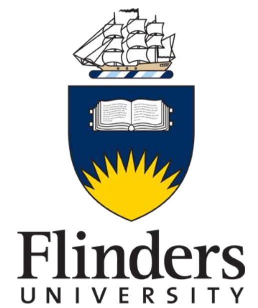


Identifying untapped microbial resources in the marine sponge microbiome

A thesis submitted for the award of

Doctor of Philosophy

at School of Medicine, Faculty of Medicine, Nursing and Health Sciences, Flinders University

Qi Yang

December 2016

TABLE OF CONTENTS

LIST OF TABLES .... **vi**

LIST OF FIGURES .... **viii**

LIST OF ABBREVIATIONS .... **x**

SUMMARY .... **xii**

DECLARATION .... **xiv**

ACKNOWLEDGEMENTS .... **xv**

Chapter 1: Introduction **1**

1.1 Sponge molecular taxonomy **1**

1.1.1 Sponge taxonomy: from morphology to molecular 3

1.1.2 Various DNA markers for sponge molecular taxonomy 5

1.1.3 Phylogenetic classification: update of Demospongiae (class) sponge 9

1.2 Abundant and valuable sponge-associated microorganisms **10**

1.2.1 Sponge-associated microorganisms 10

1.2.2 Sponge-associated actinobacteria: diverse and pharmaceutically valuable bacteria 14

1.2.3 Sponge-associated microorganisms as promising source of natural bioactive compounds 19

1.3 Sponge microbiome revealed by Next Generation Sequencing **21**

1.3.1 Next Generation Sequencing technology 21

1.3.2 Sponge microbiome revealed by 16S rRNA gene based metagenomic sequencing. 23

1.3.3 Sponge specific microbial community 30

1.4 Research plan **33**

1.4.1 Hypotheses 33

1.4.2 Aims of the project 34

1.4.3 Applied methods and techniques 35

1.4.4 Project framework 36

1.4.5 Thesis structure and contributions 36

References **38**

Chapter 2 Development of a multilocus-based approach for sponge (Porifera) identification: refinement and limitations **62**

2.1 Introduction **62**

2.2 Materials and methods **64**

2.2.1 Sponge collection and morphological classification 64

2.2.2 DNA extraction 67
2.2.3 PCR amplification and sequencing 67
2.2.4 Data processing 68
2.2.5 Development of a Sponge Identification Protocol (SIP) 70
2.2.6 Phylogenetic analysis 71
2.2.7 Morphological re-examination and final identification 71
2.3 Results 72
2.3.1 Initial morphological classification 72
2.3.2 Valid sequences 73
2.3.3 Identification using the proposed Sponge Identification Protocol (SIP) 75
2.3.4 SIP reliability validated by phylogenetic analysis 77
2.3.5 Re-examination of morphological identifications 80
2.3.6 Approach to the final identity and the improved discrimination 83
2.3.7 SIP reliability evaluated by morphological identification 84
2.3.8 Suitability and capacity of three DNA markers for sponge identification 85
2.3.9 Limited entries in database 86
2.4 Discussion 88
2.5 Conclusion 90
References 90
Chapter 3 Sponge-associated actinobacterial diversity: Validation of the methods of actinobacterial DNA extraction and optimisation of 16S rRNA gene amplification 96
3.1 Introduction 96
3.2 Materials and methods 98
3.2.1 Specimens 98
3.2.2 Sample preparation 100
3.2.3 DNA extraction 100
3.2.4 PCR amplification of 16S rRNA gene 100
3.2.5 PCR optimisation for 16S rRNA gene 100
3.2.6 PCR amplification of COI mtDNA 101
3.3 Results 102
3.3.1 DNA extraction efficiency of sponge spiked with actinobacteria 102
3.3.2 Highly-diluted DNA required for successful PCR amplification 102
3.3.3 Primer selection and PCR optimisation required for successful PCR amplification of individual sponges 106
3.4 Discussion 108
3.4.1 Comparable DNA recovery from sponge spiked with actinobacterial spores or mycelia

.... 108
3.4.2 Identification and elimination of interferences in the sponge DNA 109
3.4.3 Individual optimisation of PCR conditions required for successful 16S rRNA gene amplification 110
3.5 Conclusion **110**
References **111**
Chapter 4 Combination of different region-specific primers of 16S rRNA gene: essential for sponge microbiome analysis using Illumina platform **116**
4.1 Introduction **117**
4.2 Materials and methods **119**
4.2.1 Sponge collection and community DNA extraction 119
4.2.2 454 GS FLX amplicon library and pyrosequencing 119
4.2.3 Illumina MiSeq amplicon library and sequencing 120
4.2.4 16S rRNA sequencing data processing 121
4.2.5 Sponge microbiome analysis 122
4.3 Results **122**
4.3.1 Classification resolution and taxa richness highly dependent on sequencing techniques 122
4.3.2 Efficacy of microbial taxa identification highly dependent on region-specific primers 124
4.3.3 Biased differentiation performance of various region-specific primers for sponge microbiome 129
4.3.4 Distinct unaffiliated OTU profiles revealed by different various primer sets 130
4.3.5 Power to uncover the likely missed microbial OTUs 137
4.4 Discussion **138**
4.4.1 Illumina MiSeq: a superior method than 454 GS FLX Titanium sequencing 138
4.4.2 Validated combination of different region-specific primers essential for sponge microbiome study 138
4.4.3 Deeper uncovering of microbial taxa by primers specific to the V5 to V8 regions ... 139
4.4.4 Better performance of V5V8 in differentiating sponge-specific microbiomes........... 140
4.4.5 Vast untapped microbial resources revealed by a combination of multiple region-specific primer sets 141
4.5 Conclusion **142**
References **142**
Chapter 5: Combination of Multiple 16S rRNA gene region-specific primer sets reveals comprehensive sponge microbiome and untapped Microbial resource **153**
5.1 Introduction **154**

5.2 Materials and methods **155**
5.2.1 Sponge collection and community DNA extraction 155
5.2.2 Illumina MiSeq amplicon library and sequencing 157
5.2.3 Sequencing data processing 158
5.2.4 Performing Procrustes Analysis 159
5.2.5 Phylogenetic analysis of sponges 160
5.3 Results **160**
5.3.1 Powerful multiple-region based sequencing approach 160
5.3.2 Sponge microbial community structure revealed by a combination of multiple primer sets 163
5.3.3 OTU richness of the sponge microbiome at the order level 164
5.3.4 OTU richness of the sponge microbiome at the family level 165
5.3.5 Comparison of microbial community at the sponge order level 167
5.3.6 Comparison of microbial community at the sponge family level 168
5.3.7 Dominant microbial OTUs 169
5.3.8 Specific microbial OTUs 171
5.3.9 Performing procrustes analysis 181
5.3.10 Phylogenetic relationship of studied sponges 184
5.4 Discussion **186**
5.4.1 A combination of multiple region-specific primer sets required for comprehensive sponge microbiome analysis 186
5.4.2 Structure-specificity rather than diversity-specificity 187
5.4.3 Order-specificity and family-specificity of sponge microbial community 188
5.4.4 Unique microbial taxa as signature for a sponge order or family 189
5.5 Conclusion **189**
References **190**

Chapter 6 Overall conclusion and future directions **194**
6.1 Summary of the project key findings **194**
6.2 Limitations and challenges of this study **196**
6.3 Future research directions **196**
References **198**

APPENDICES **203**

LIST OF TABLES

Table 1 - 1 Various DNA markers applied on sponge molecular taxonomy (2010-2015)............ 8
Table 1 - 2 Demospongiae species (additional to the AFD list) documented from South Australia.......... 14
Table 1 - 3 Actinobacterial genera associated with marine sponges based on the RDP classification of 16S rRNA gene in GenBank 16
Table 1 - 4 Sponge microbiome studies using various 16S rRNA gene region-specific primer sets on different NGS platforms 28
Table 2 - 1 Sponge morphological classification including sampling locations and collection dates.......... 72
Table 2 - 2 Summary of molecular identification of 37 potential sponge species using multilocus approach.......... 74
Table 2 - 3 Trial results of alternative primers for the 28S rRNA gene amplification on the failed sponges and the sampling locations 75
Table 2 - 4 Summary of molecular identification of 34 sponges using three DNA markers and following the sponge order in Table 2-1 76
Table 2 - 5 Reliability of SIP validated by phylogenetic analysis 78
Table 2 - 6 Re-examination of morphological classification and comparison with SIP identification.......... 80
Table 2 - 7 Comparison of molecular identifications by selected DNA marker(s) with or without morphological classifications.......... 85
Table 2 - 8 The submissions of the three DNA loci based on different Porifera orders in the NCBI database.......... 86
Table 2 - 9 The nucleotide submissions from different sponge families, genera and species... 87
Table 3 - 1 The sponge and actinobacteria specimens.......... 99
Table 3 - 2 The applied primer sets for validation 101
Table 3 - 3 The summary of sample combination types and the working concentration of the DNA templates in the 16S rRNA PCR amplification 104
Table 3 - 4 Primer selection for the amplification of the 16S rRNA gene from six sponge samples and individual strains, spores or mycelia of four actinobacterial species 107
Table 3 - 5 The optimised PCR reaction conditions with different primer sets.......... 108
Table 4 - 1 Capacity of different primer sets in revealing the affiliated (known) microbial OTUs for four sponge microbial communities.......... 125
Table 4 - 2 The number of unaffiliated and affiliated OTUs at the phylum/ class level derived from four sponge species revealed by four primer sets 132
Table 4 - 3 Sequence abundance (%) of phylum and class level OTUs revealed by four primer

sets 133
Table 4 - 4 Increased sequences (%) of combined data of four region-specific primer sets based on the total number of the OTUs 134
Table 4 - 5 Comparison of the affiliated (known) and unaffiliated phylum-level OTUs between the microbial profiles of four sponge species in this study and 81 species revealed by Thomas et al. (2016) 135
Table 4 - 6 Likely missed phylum-level OTUs by the V4 region primer set could be revealed by four 16S rRNA primer sets specific to V1V3, V4, V4V5, and V5V8 regions. The likely missed phylum-level OTU list refer to Eloe-Fadrosh et al. (2016) 137
Table 5 - 1 Details of sponges collected from Rapid Bay in South Australia for this study 156
Table 5 - 2 Comparison of the affiliated (known) and unaffiliated phylum-level OTUs for the microbial profiles of 33 sponge species revealed by a combination of multiple 16S rRNA gene regions (V1V3, V4V5, V5V8) in this study and 81 species revealed by V4 region only in the study of Thomas et al. (2016) 161
Table 5 - 3 Comparison for shared and unique dominant OTUs (>1%) between different sponge orders 173
Table 5 - 4 Comparison for shared and unique dominant OTUs (>1%) between different sponge families in one order 175
Table 5 - 5 Comparison for shared and unique OTUs between different sponge orders 177
Table 5 - 6 Comparison for shared and unique OTUs between different sponge families in one order 179

LIST OF FIGURES

Figure 1 - 1 Overview and internal structure of a typical demosponge 11
Figure 1 - 2 South Australian waters include eight Integrated Marine and Coastal Regionalisation of Australia (IMCRA) meso-scale bioregions 13
Figure 1 - 3 Taxonomic distribution within the class Demospongiae collected in South Australia 13
Figure 1 - 4 Development of the culture-independent technologies applied in sponge microbiome study 27
Figure 1 - 5 PhD project framework 36
Figure 2 - 1 The sampling locations and the photos of the 37 sponge species 66
Figure 2 - 2 Sequence data processing flowchart 69
Figure 2 - 3 Sponge Identification Protocol (SIP) flow diagram used in this study.................... 70
Figure 2 - 4 Phylogenetic relationship of three representative sponge species using the Maximum Likelihood method based on 28S rRNA gene and COI mtDNA to validate the sponge identification by SIP 79
Figure 4 - 1 Positions of primer sequences and tandem regions used in this study for 454 pyrosequencing and Illumina sequencing.. 120
Figure 4 - 2 Proportion of affiliated (known) OTUs reaching a specified taxonomic level on 454 pyrosequencing and Illumina MiSeq platforms. 123
Figure 4 - 3 Comparison of the richness of affiliated (known) OTUs between 454 pyrosequencing and Illumina MiSeq platforms based on phylum, class, order, and family levels of the microbial taxa 124
Figure 4 - 4 Number of unique genera and total genera revealed by each region-specific primer set on Illumina MiSeq platform. 127
Figure 4 - 5 Distribution of affiliated (known) microbial OTUs revealed by four region-specific primer sets V1V3, V4, V4V5 and V5V8 on Illumina MiSeq platform for the same sponge samples 128
Figure 4 - 6 Microbial community comparison between four sponge species belonging to four orders based on the affiliated (known) OTUs. 130
Figure 5 - 1 Microbial community composition and structure of 19 sponge species belonging to four taxonomic orders based on the phylum-level microbial OTUs 163
Figure 5 - 2 Microbial community composition and structure of 14 sponge species belonging to five taxonomic families within one order based on the phylum-level microbial OTUs.............. 164
Figure 5 - 3 Rarefaction curves of 16S rRNA gene diversity showing the microbial communities of the sponge species in four taxonomic orders 165
Figure 5 - 4 Rarefaction curves of 16S rRNA gene diversity showing the microbial communities

of the sponge species in five taxonomic families within one order. 166
Figure 5 - 5 Principal Coordinates Analysis (PCoA) plots showing the similarity of microbial communities of the 19 sponge species in four taxonomic orders. 168
Figure 5 - 6 Principal Coordinates Analysis (PCoA) plots showing the similarity of microbial communities of the 14 sponge species in five taxonomic families within one order. 169
Figure 5 - 7 PCoA plot showing the five most abundant class-level microbial OTUs among the 19 sponge species in four taxonomic orders. 170
Figure 5 - 8 PCoA plot showing the five most abundant class-level microbial OTUs among the 14 sponge species in five taxonomic families within one order. 171
Figure 5 - 9 Procrustes analysis of the 19 sponge species belonging to four orders using the unweighted UniFrac metric generated by three 16S rRNA gene region-specific primer sets .. 181
Figure 5 - 10 Procrustes analysis of the 19 sponge species belonging to four orders using the weighted UniFrac metric generated by three 16S rRNA gene region-specific primer sets 182
Figure 5 - 11 Procrustes analysis of the 14 sponge species belonging to five families in one order using the unweighted UniFrac metric generated by three 16S rRNA gene region-specific primer sets 183
Figure 5 - 12 Procrustes analysis of the 14 sponge species belonging to five families in one order using the weighted UniFrac metric generated by three 16S rRNA gene region-specific primer sets 184
Figure 5 - 13 Phylogenetic relationship of the studied sponge species belonging to four taxonomic orders based on the 28S rRNA gene using Neighbor Joining method 185
Figure 5 - 14 Phylogenetic relationship of the studied sponge species belonging to five taxonomic families in one order based on the 28S rRNA gene using Neighbor Joining method 186

LIST OF ABBREVIATIONS

AFD	Australian Faunal Directory
*A*RBS	*Aplysina* Red Band Syndrome
ARISA	Automated Ribosomal Intergenic Spacer Analysis
BLAST	Basic Local Alignment Search Tool
BSA	bovine serum albumin
COI	cytochrome *c* oxidase subunit I
CTAB	cetyltrimethylammonium bromide
DGGE	denaturing gradient gel electrophoresis
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
dNTPs	nucleoside triphosphates containing deoxyribose
EMP	Earth Microbiome Project
emPCR	emulsion polymerase chain reaction
FISH	fluorescent *in situ* hybridization
GS	Genome Sequencer
HIV	human immunodeficiency virus
HMA	high microbial abundance
IGPs	intra-genomic polymorphisms
IMCRA	Integrated Marine and Coastal Regionalisation of Australia
ITS	internal transcribed spacer
LMA	low microbial abundance
LSU	large subunit
ML	Maximum Likelihood
mtDNA	mitochondrial deoxyribonucleic acid
NCBI	National Center for Biotechnology Information
NGS	Next Generation Sequencing
NJ	Neighbor-Joining
NRPSs	nonribosomal peptide synthetases
OTU	Operational Taxonomic Unit
PBDE	polybrominated diphenyl ether
PCoA	principal coordinates analysis
PCR	polymerase chain reaction
PGM	Personal Genome Machine
PKSs	polyketide synthases
PVP	polyvinylpyrrolidone

QIIME	quantitative insights into microbial ecology
RDP	Ribosomal Database Project
RNA	ribonucleic acid
rRNA	ribosomal ribonucleic acid
SARDI	South Australian Research Development Institute
SBL	sequencing by ligation
SBS	sequencing by synthesis
SIP	Sponge Identification Protocol
SMP	Sponge Microbiome Project
ssDNA	single-stranded deoxyribonucleic acid
T-RFLP	terminal restriction fragment length polymorphism
UPGMA	unweighted pair group method with arithmetic mean
WPD	World Porifera Database

SUMMARY

diversity (the composition of the microbial OTUs) within a phylogenetic unit (e.g. order and family). Each order and family has specific microbial OTUs, which are valuable in guiding the exploration of the target microbial groups, particularly for the untapped resources. In conclusion, this project developed a pipeline for an unprecedented complete characterisation of the sponge microbiome, which includes reliable identification of sponge samples, efficient extraction of community DNA, PCR optimisation, evaluation of region-specific primer sets for 16S rRNA gene based amplicon sequencing, and bioinformatics analysis, for a rationale guided discovery of untapped marine sponge associated microbial resources.

DECLARATION

I certify that this thesis does not incorporate without acknowledgment any material previously submitted for a degree or diploma in any university; and that to the best of my knowledge and belief it does not contain any material previously published or written by another person except where due reference is made in the text.

Signed:

Date: 9th December 2016

ACKNOWLEDGEMENTS

Undertaking this Ph.D. has been a truly life-changing experience for me and it would not have been possible without the support and guidance that I received from many people.

Firstly, I would like to express my sincere gratitude to my principal supervisor Professor Chris M.M. Franco for the continuous support of my Ph.D. study and related research, for his patience, motivation, and immense knowledge. He guided me throughout my research and writing of this thesis.

I would like to express my special appreciation and thanks to my co-supervisor Professor Wei Zhang, who is a tremendous mentor for me. I appreciate his contributions of time, ideas, and funding to make my Ph.D. experience productive and stimulating. The joy and enthusiasm he has for his research was contagious and inspirational for me, especially during the tough times in the Ph.D. journey. I would like to thank him for encouraging my research and for allowing me to grow as a research scientist. His advices on both research as well as on my career have been priceless.

I would also like to say a very big thank you to Ms Shirley Sorokin for all the support and encouragement she gave me. She is not only a good advisor but also a great friend for me to help me solve the problems. My thanks also go out to the support I received from Dr Laura S. Weyrich, Australian Center for Ancient DNA of University of Adelaide. I greatly appreciate her guidance on the analysis of the amplicon sequencing data.

My deep appreciation goes out to the whole research team members in Centre for Marine Bioproducts Development (CMBD). Without the endeavours contributed by all of them, I could not have had a good working atmosphere to conduct my experiment and focused on my research. I would also like to acknowledge everyone in the department of Medical Biotechnology for their support and friendship. Especially, I would like to acknowledge Ms Jane Keane and Dr. Hanna Krysinska for their efficient and excellent managements on the laboratory running, event organisation, and appointment schedule.

I gratefully acknowledge the funding received towards my Ph.D. from the Chinese Scholarship Council and Flinders University.

Last but not the least, I would like to say a heartfelt thank you to my family for all their love and encouragement, for my parents who raised me with a love of science and supported me in all my pursuits, and for my partner Edward Xituo Meng who always believed in me and encouraged me to follow my dreams.

CHAPTER 1: INTRODUCTION

1.1 Sponge molecular taxonomy

Sponges (phylum Porifera), sessile and benthic filter feeders, are considered the oldest multicellular animals (Hentschel et al. 2012). Important roles of sponges in marine ecosystem include biogeochemical cycling (Wulff 2012), spatial structuring of the seafloor (Gutt 2007), and benthic-pelagic coupling of nutrient transfer (Bell 2008). Sponges are also commercially important to the pharmaceutical and biomaterial industries as they participate in complex biotic interactions with diverse macrobiotic taxa (Bell 2008) and microbiological communities (Webster & Taylor 2012) to produce highly potent secondary metabolites (Leal et al. 2012).

According to World Porifera Database (Van Soest et al. 2016), there are more than 8,700 valid species, 7,300 of which belong to the class Demospongiae. Species identification of Porifera traditionally relying on morphological feature is extremely challenging (Andreakis, Luter & Webster 2012). These characters are their organic and inorganic skeletons, including skeletal size, shape, structure and composition (Hooper & Van Soest 2002). However, the arrangement of these skeletal elements can be inconsistent, and our understanding of the evolution of skeletal traits is incomplete (Pöppe et al. 2010). Indeed, traditional morphological identification methods often lead to erroneous classification (Xavier et al. 2010). The cryptic species belonging to Porifera are therefore frequently reported (Erwin & Thacker 2007; Reveillaud et al. 2010), which underestimated the existing species diversity and distribution of Porifera (Blanquer & Uriz 2007; Hooper et al. 2013).

Molecular approaches, such as DNA barcoding, provide the potential solution (Erpenbeck et al. 2012; Fontaneto, Flot & Tang 2015; Kress et al. 2005; Wörheide & Erpenbeck 2007). The mitochondrial DNA (mtDNA) exists in all eukaryotic cells and often provides good markers for species identification due to its clonal (maternal) mode of inheritance and clock-like evolutionary rate (Galtier et al. 2009; Shearer et al. 2002; Vargas et al. 2012; Voigt, Eichmann & Wörheide 2012). It has been used to study species identification, sponge diversification patterns (Pöppe et al. 2010) as well as phylogenetic relationships (Erpenbeck et al. 2012) with varying degrees of success (Dohrmann et al. 2012; Lavrov, Wang & Kelly 2008). The COI mtDNA locus is the most commonly used mitochondrial marker of approximately 700 bp at the 5'end of the cytochrome *c* oxidase subunit I (COI) gene - a locus that is relatively easy to amplify as it is conserved across multicellular animals (Folmer et al. 1994) and abundant in eukaryotic DNA (Heim, Nickel & Brümmer 2007b). Blanquer & Uriz (2007) reported that COI mtDNA successfully discriminated between species in the genera *Tethya* and *Scopalina*. However,

some species that can be clearly distinguished on the basis of morphology show similar COI sequences (Heim, Nickel & Brümmer 2007a; Pöppe et al. 2010). Studies on the COI intraspecific variation have been used more regularly to classify other metazoans at the species level, but less so for sponges (Duran, Pascual & Turon 2004; Wörheide 2006).

Slow mitochondrial evolution is a problem for the resolution of phylogenies at the species and genus levels using standard mitochondrial markers. However, the question remains as to whether faster evolving gene regions can be identified for use in conjunction with the standard COI mtDNA barcode (Pöppe et al. 2010). The nuclear ribosomal genes of eukaryotes, such as the 28S (large subunit, LSU) rRNA genes (Gerbi 1985), are arranged in tandemly repeated clusters, where transcribed units alternate with non-transcribed units called spacers, such as the internal transcribed spacer 1 (ITS1) and 2 (ITS2) (Borchiellini et al. 2000; Hillis & Davis 1988). The 28S rRNA gene has regions that are sufficiently heterogeneous to address phylogeny at many different levels (Cárdenas et al. 2011; Erpenbeck et al. 2012). Various regions have been evaluated to be the DNA markers for sponge taxonomy (Cárdenas, Pérez & Boury-Esnault 2012; Redmond et al. 2011; Thacker et al. 2013). The more rapidly evolving ITS regions are commonly used as “high resolution” markers. In general, ITS regions are used to reconstruct relationships ranging from those between populations to those between the taxonomic “families”. They have been used for phylogenetic and phylogeographic analyses of non-bilaterian metazoans such as corals (Pillay et al. 2006) and sponges (Erpenbeck et al. 2011; Wahab et al. 2014).

For sponges, there is no reliable molecular protocol available for species identification. Based on the current knowledge about the resolution of different DNA markers, no single ideal marker for all sponge species exists as each marker has its own strengths and limitations (Duran, Pascual & Turon 2004; Szitenberg et al. 2013; Voigt, Eichmann & Wörheide 2012; Wörheide 2006). The incomplete sequence entries in the gene database limit the application of the phylogeny-based molecular taxonomic approach for species identification. In the NCBI database, the sponge derived gene submissions only cover a few hundreds (Benson et al. 2009) out of the known 8,700 sponge species. It is essential to establish an effective and practical molecular approach for sponge identification to respond to these issues. On the other hand, marine sponge-associated microbial community is highly relevant to the host sponge identity (Rodríguez-Marconi et al. 2015). The communities show the species-specificity regardless of the variations of the living environment (Rodríguez-Marconi et al. 2015). Importantly, these associated microorganisms have unsurpassed capacity for the production of natural chemical compounds with diverse bioactivities (Abdelmohsen, Bayer & Hentschel 2014; Hentschel et al. 2012). Therefore, the baseline study of the special relationship between the

host sponges and their symbiotic microbes is the rationale-based guidance to explore the target secondary metabolites for new drug discovery. To better answer the question of how the phylogeny of the host sponges relate to their associated microbial community, a reliable and effective sponge identification is required to conduct the sample collection and selection. The development of the sponge identification protocol is showed in Chapter II.

1.1.1 Sponge taxonomy: from morphology to molecular

Sponge classification traditionally relies on their morphological characters. The most important element is their spicule. Most of the types are siliceous structures, which are often needle-shaped, form a distinct skeleton, but occasionally they are randomly distributed throughout the sponge body. Moreover, sponges offer only a few different types of the characteristic spicules, which increases the difficulty of morphological identification. Another issue is that no diverse morphological character is available for most of the sponges. The cytological features have been used in sponge systematics (Boury-Esnault et al. 1994), but the phylogenetic information content has not been fully analysed and might be insufficient for higher sponge taxonomy. These limitations could lead to homoplasies and erroneous classification (Mikkelsen & Cracraft 2001).

A group of sponges Keratosa (subclass), including the species in the two orders Dendroceratida Minchin, 1900 (Bergquist & Cook 2002) and Dictyoceratida Minchin, 1900 (Cook & Bergquist 2002), only have the organic skeleton made of spongin fibres instead of mineral spicules. It was found that the spongein fibres are much less diverse than the spicules, which makes it markably difficult to discrinminate this group of sponges. However, an efficient and reliable classificaiton for the sponges in subclass Keratosa is required to establish due to their diverse host-specific natrual products, which have the potent value for new drug discovery (Blunt & Munro 2003). Again, the sponge species in the family Irciniidae Gray 1867 (order Dictyoceratida), have no spicules but only collagen filaments in the mesohyl. The collagen filaments are their key morphological features to distinguish with other families in Keratosa (subclass) (Van Soest et al. 2016). The discrinimation and the identificaiton for the sponge species in the three genera *Ircinia*, *Sarcotragus* Schmidt 1862, and *Psammocinia* Lendenfeld 1889 belonging to family Irciniidae also rely on the structures of the primary filaments or fibres. However, the high diversity of the species and the unstandardised morphological classification increase the possibility of the ambiguous sponge species (Cook & Bergquist 1998; Cook & Bergquist 2002).

In fact, Porifera show considerable morphological variations (polymorphisms); the variations have been suggested to be impacted by environmental factors (Meroz-Fine, Shefer &

Ilan 2005). Polymorphisms might also be responsible for the misidentification and/or phylogenetic misplacement of Porifera species. Different factors, such as bathymetry, water energy, sedimentation, light conditions, and symbionts as well as the availability of nutrients and substrates, are considered to be important parameters for the distribution of sponge species and for their growth morphologies. Thus, growth forms of sponges are commonly used as indicators in palaeoecological studies (Mehl-Janussen 1999). However, most studies on the zonation and distribution of sponges have concentrated on biocoenoses in tropical environments (Duran & Rützler 2006) and only a few studies have concentrated on temperate seas (Erpenbeck et al. 2004). Thus far, the existing field studies on sponge morphology in temperate areas indicated that variation occurred only within a genetically fixed frame (Bell, Barnes & Tuner 2002). However, all the main criteria for morphological investigations are subject to variation caused by environmental effects (Maldonado et al. 1999), which could lead to misidentification and/or phylogenetic misinterpretation.

Few useful morphological characteristics exist for the classification of many of the taxa within the Porifera species (Lévi 1973). Among the sponge taxa with a high number of useful diagnostic features, most species have a relatively restricted geographical distribution, whereas among those with only a few useful taxonomic characters, many species are considered to be very widespread or even cosmopolitan. Therefore, many sponge scientists suspect that the apparent global distribution of various sponge species may merely be a consequence of inadequate taxonomic studies (Solé-Cava et al. 1992). Again, morphological homogeneity may easily lead to the misidentification of species. Cryptic divergence and large genetic variation within species or even species groups have been observed in the recent genetic analyses of species that were considered to be cosmopolitan or have a wide distribution range. The misidentification has led to a series of nomenclature correction of many sponge species and adjustment among families and sometimes, orders within the Demospongiae (Hooper & Van Soest 2002). Application of the advanced technologies such as electron microscopy or chemistry techniques, alongside the traditional morphological tool for species discrimination, has demonstrated that some species are actually comprised of more than one subspecies, which could be the potential new species (Klautau, Solé-Cava & Borojevic 1994; Solé-Cava et al. 1992).

Since the sponge-produced bioactive and pharmaceutically valuable compounds were discovered (Bergmann & Freeney 1950), new phylogenetic hypotheses have been brought up by biochemical data, which offers an alternative means to morphological characteristics of Porifera (Van Soest & Braekman 1999). Presence and absence of particular biochemical compounds or pathways could help on the scant morphological characters for sponge

classification. For example, Bergquist (1978) concluded that family Agelasidae has closer phylogenetic relationship with Axinellidae. Moreover, Braekman et al. (1992) noted an exclusive occurrence of pyrrole-2-imidazole derivates in the species of both orders Agelasida and Axinellida and provided additional evidence for a potential relationship between these two groups. In the same year, Lafay et al. (1992) published one of the first molecular phylogenies on sponges. Their 28S rRNA fragment analysis also favoured a close relationship between *Agelas oroides* (family Agelasidae) and *Axinella damicornis* (family Axinellidae) species. The finding was further supported by the 28S rRNA gene data from Alvarez et al. (2000).

With the development of the molecular technique, DNA barcoding was becoming an effective tool for sponge classification. DNA barcoding can be used to identify known sponge species and also to discriminate the ambiguous species (Hebert et al. 2003; Witt, Threloff & Hebert 2006). Sponges Barcoding Project has been established (www.spongebarcoding.org) (Wörheide, Erpenbeck & Menke 2008). The COI mtDNA (~ 640 bp length) was used as the standard barcoding region. Further downstream the standard barcoding fragment contains a region with a higher substitution rate, when used together with the standard fragment, offers more variable sites (Erpenbeck, Hooper & Wörheide 2006). It is yet to demonstrate the effectiveness of this COI fragment for its use in sponge species, given the fact that substitution rates for mitochondria in Porifera are reduced and hence diminish its resolution power at the species and genus levels (Huang et al. 2008). Although the COI was observed to be too conserved for population studies (Wörheide 2006), it could still be a good marker for species discrimination, at least within some invertebrate groups (Heim, Nickel & Brümmer 2007a, 2007b).

Some study comprehensively considered the morphological classification, in conjunction with molecular and biochemical analyses to further infer the identity. For example, the presence of the characteristic spicules in most genera in combination with molecular 28S sequence data (Nichols 2005) and biochemical evidence (Wörheide 1998) clearly demonstrated a close relationship between family Astroscleridae and Agelasidae despite their fundamentally different skeleton types.

1.1.2 Various DNA markers for sponge molecular taxonomy

The molecular approach has provided much insight into the evolutionary relationships of some invertebrates, particularly where morphological characteristics are few. The molecular approach, however, has yet to be applied as efficiently as in other metazoan animals (Faulkner 2000).

The application of mitochondrial gene in sponge systematics started from the late 20th century (Wörheide, Degnan & Hooper 2000), and it was not until the early 21st century that the complete sponge mitochondrial genomes were published – making it one of the last published among all the major metazoan phyla (Lavrov et al. 2005). The mitochondrial DNA exists in all multicellular animals (Folmer et al. 1994) and often provides good markers for species identification because the COI mtDNA is a conservative region but with highly variable sequences (Galtier et al. 2009; Voigt, Eichmann & Wörheide 2012), and this gene is relatively easy to amplify owning to its highly abundant in eukaryotic DNA (Heim, Nickel & Brümmer 2007b). It has been used to study species identification, sponge diversification patterns (Pöppe et al. 2010) as well as phylogenetic relationships (Erpenbeck et al. 2012) with varying degrees of success (Dohrmann et al. 2012; Lavrov, Wang & Kelly 2008). For example, COI mtDNA successfully discriminated the species in the genera *Tethya* and *Scopalina* (Blanquer & Uriz 2007). More recently, Vargas et al. (2015) collected sponges from the Rose Sea region and sequenced the standard COI region. DNA barcodes of 53 species were produced.

However, some species that can be clearly distinguished based on morphology show highly similar COI sequences (Heim, Nickel & Brümmer 2007a; Pöppe et al. 2010). The mtDNA genes were proven with better resolution for the phylogenetic analysis at lower taxonomic levels for the Metazoan organisms at a higher rank (Duran, Pascual & Turon 2004; Wörheide 2006). COI mtDNA has been reported to be suitable for the sponge phylogenetic analysis at the family level (Erpenbeck et al. 2002; Nichols 2005), but have insufficient variability below this level (Duran, Pascual & Turon 2004; Schröder et al. 2003; Wörheide et al. 2002).

The question was pointed out whether the faster evolving gene regions can be identified for use in conjunction with the standard COI mtDNA barcode (Pöppe et al. 2010) since slow mitochondrial evolution is a problem for the resolution of phylogenies at the species and genus levels. The main progress in recent sponge systematics is due to the gain of DNA sequence data. The 28S rRNA gene, subunit of the cytoplasmatic ribosome, has been used to reconstruct phylogenetic trees (Borchiellini et al. 2004; Manuel et al. 2003). Some 28S rRNA gene analyses yielded evidence against a monophyletic relationship of the family Halichondriidae. An analysis of the 28S C1–D2 region found the family Halichondriidae in a shared clade with family Suberitidae (Chombard & Boury-Esnault 1999). The resulting taxon, called 'Suberitina' still lacked acceptance as it failed to show unambiguous morphological synapomorphies notably regarding the spicule geometry. However, all these molecular analyses comprised only a subset of halichondrids and made a comprehensive study of all major families necessary to elucidate their composition and relationships to non-halichondrids. Erpenbeck et al. (2004) observed that the 28S rRNA gene structure and evolutionary rate in sponges underlie significant differences

on order level. Based on those findings, the taxon set could be narrowed down to a homogeneous and comparable character set without long branches, which mask phylogenetic information and lead to an erroneous signal.

The 18S rRNA gene is approximately 1,800 nucleotides long in eukaryotes (Hillis & Dixon 1991), but unusually long sequences have been reported for a variety of distantly related organisms including protists, plathyhelminthes, and arthropods (Busse & Preisfeld 2002; Gillespie et al. 2005; Giribet & Wheeler 2001). The increase in length is due to large insertions in the variable regions of the gene, with the V4 and V7 regions being the most commonly affected. Variable regions (named 'V' regions) of the 18S rRNA may not be functionally important and as a result have few structural constraints. Therefore, they have a higher substitution rate and greater variation between species. In the tertiary structure of the 18S rRNA gene, conserved regions ('C' regions) are found in the centre of the ribosome and the 'V' regions are positioned on the surface (Doudna & Rath 2002; Wuyts, Van de Peer & De Watcher 2001). The presence of repeated copies and the different rates of evolution among the various regions make the 18S rRNA gene quite versatile in animal systematics (Borchiellini et al. 2004; Borchiellini et al. 2001; Redmond et al. 2007). It is believed that these indels are strong synapomorphies for a monophyletic group of haplosclerids (hereafter referred to as Clade X) not previously described in morphological classifications, and are able to help resolve the phylogeny of the order in conjunction with other phylogenetically informative characters. The use of new molecular synapomorphies will help provide answers to the many questions in both haplosclerid and sponge systematics as a whole.

Ribosomal internal transcribed spacer (ITS) (Schröder et al. 2003) sequence data have been successfully employed to study some sponges at the population level, e.g. *Axinella corrugata* (Lopez et al. 2002), *Leucetta 'chagosensis'* (Wörheide, Hooper & Degnan 2002) and *Crambe crambe* (Duran, Pascual & Turon 2004). However, Wörheide, Nichols & Goldberg (2004) showed that there were heterogeneous levels of intra-genomic polymorphisms (IGPs) in a diverse range of marine sponge taxa. They concluded that the ITS regions should only be used in phylogenetic studies below the family level as multiple sequence alignments above this level would be too variable to confidently establish homologies. More importantly, their work also suggested that care should be taken when using ITS regions due to the different levels of IGPs detected across sponge taxa. They recommended that each taxon should be screened to identify and analyse the level of IGPs present. Alvarez, Krishnan & K. (2007) investigated the intra-genomic variation in the ITS regions in three halichondrid sponges and found at least three different sequence types in any single individual and suggested that the ITS regions are not appropriate markers for fine-scale population studies in species with levels of IGPs greater than

1%.

All of the DNA markers have their suitability and capacity for sponge classification and identification at different taxonomic levels. They showed the various successes on different sponge groups. Table 1-1 summarises the utilisation of the DNA markers as the barcodes to conduct sponge (Porifera) molecular taxonomy.

Table 1 - 1 Various DNA markers applied on sponge molecular taxonomy (2010-2015)

DNA markers	Sponge group	References
COI mtDNA	Demosponge	Vargas et al. 2015
COI mtDNA	~ 7,400 sponge specimens	Vargas et al. 2012
COI mtDNA	Keratose sponges	Erpenbeck, Sutcliffe et al. 2012
COI mtDNA	Halichondrid demosponges	Erpenbeck, Hall et al. 2012
COI mtDNA	Hexactinellida (class)	Dohrmann et al. 2012
COI mtDNA	Irciniidae (family)	Pöppe et al. 2010
COI mtDNA, 28S rRNA gene (C1-D2), 18S rRNA gene	Tetillidae (family): 88 specimens belonging to 28 species	Szitenberg et al. 2013
COI mtDNA, 28S rRNA gene	*Abyssocladia* (genus) *Phelloderma* (genus)	Vargas et al. 2012
COI mtDNA, 28S rRNA gene	Demospongiae (class)	Morrow et al. 2012
COI mtDNA, 28S rRNA gene (C1-D2)	Astrophorida (revised to suborder Astrophorina)	Cárdenas et al. 2011
COI mtDNA, 28S rRNA gene (C1-D2)	Geodiidae (family)	Cárdenas et al. 2010
COI mtDNA, 28S rRNA gene, nadI mtDNA	Haplosclerida (order)	Redmond et al. 2011
COI mtDNA, 28S rRNA gene, Atp8 mtDNA	*Cliona celata*	Xavier et al. 2010
COI mtDNA, ITS2 marker	*Ianthella basta*	Freckelton et al. 2012
COI mtDNA, ITS marker	*Ianthella basta*	Andreakis, Luter & Webster 2012
COI mtDNA, ITS marker	*Spongillina* (genus)	Erpenbeck et al. 2011
28S rRNA gene (nearly complete)	4 classes of sponges, 4 major clades of Demospongiae (class)	Thacker et al. 2013
18S, 28S rRNA genes	*Mycale laevis*	Loh et al. 2012
18S, 28S rRNA genes	Homoscleromorpha (class)	Gazave, Lapébie, et al. 2010

18S rRNA gene (total), 28S rRNA gene (partial)	Axinellidae (family)	Gazave, Carteron, et al. 2010
ITS2 marker	Keratose sponges	Wahab et al. 2014
ITS marker	Clathrinida (order): 50 species	Klautau et al. 2013
ITS marker	*Ephydatia fluviatilis* (freshwater sponge species)	Karlep, Reintamm & Kelve 2013
Large subunit (LSU) rDNA	Calcarea (class)	Voigt, Wülfing & Wörheide 2012
Mitochondrial markers	*Leucetta chagosensis*	Voigt, Eichmann & Wörheide 2012

1.1.3 Phylogenetic classification: update of Demospongiae (class) sponge

Demospongiae is the largest class among the phylum Porifera, contributing 81% of all living sponges worldwide (Morrow & Cárdenas 2015). *Systema* Porifera (Hooper & Van Soest 2002) documented a large global sponge collection to focus on the higher taxa classification of Demospongiae (class), primarily based on morphological characters. The development of the molecular phylogenetic analyses has improved the sponge classification. New clades have been gradually revealed and confirmed. Consequently, a few taxonomical changes were required to revised for the classification in *Systema* Porifera. Recently, an official revision of the Demospongiae classification was proposed by Morrow & Cárdenas (2015) based on molecular data of the last ten years. In this revision, three subclasses and seven of the 13 orders from *Systema* Porifera are retained. The names for five orders (Halichondrida, Halisarcida, Hadromerida, Verticillitida, and lithistids) are abandoned and six order names (Spongillida, Suberitida, Axinellida, Tetractinellida, Merliida, and Sphaerocladina) are upgraded from families. Finally, seven new orders (Clionaida, Tethyida, Trachycladida, Polymastiida, Scopalinida, Bubarida, and Desmacellida) are created. Another two orders (Chondrillida and Biemnida) were recently created. Therefore, there are 22 orders in total for class Demospongiae.

Due to the classification updates, the number of the species in each of the revised group also changed. The revised distribution is reviewed by Morrow & Cárdenas (2015). Based on the revised classification, the order Poecilosclerida is still the largest with 2,209 species, although about 421 species in this order were reduced. The order Haplosclerida is the second largest with 1,073 species, and followed by order Tetractinellida with 1,064 species.

Moving to the lower taxonomic levels, more new genus and species have been identified and uncovered, for example, Alvarez, De Voogd & Van Soest (2016); Garcia-Santos et al. (2016); and Göcke et al. (2016). Phylogenetic relationship for Antarctic sponges Tetillidae (Demospongiae, Tetractinellida) was reassessed to identify new genera and genetic similarity

among morphologically distinct species (Carella et al. 2016). Only a few genera and species of the family Tetillidae have been described from the Antarctic. Classical genera such as *Craniella* recovered their traditional identity by moving the Antarctic *Tetilla* from *Craniella*. The morphological re-examination of specimens used in the previous phylogeny and their comparison to the type material revealed misidentifications. Remarkably, species within the Antarctic genera *Cinachyra* (*C. barbata* and *C. antarctica*) and *Antarctotetilla* (*A. leptoderma*, *A. grandis*, and *A. sagitta*), clearly distinguishable morphologically, were not genetically differentiated with any of the markers assayed. As it has been reported for other Antarctic sponges, both the mitochondrial and nuclear partitions used did not differentiate species that were well characterised morphologically. Antarctic Tetillidae offers a rare example of genetically cryptic (with the traditional markers used for sponges), morphologically distinct species.

1.2 Abundant and valuable sponge-associated microorganisms

Marine sponges (phylum Porifera) are the oldest multicellular animals with a history of more than 600 million years (Hentschel et al. 2012). Contemporarily, they still maintain large populations globally, crossing the offshore and deep regions of the marine benthic ecosystem. For some particular areas, as much as 80% of available surfaces were covered by sponge community (Taylor et al. 2007). One of the striking characteristics of sponge is their highly diverse and abundant symbiotic microorganisms (Hentschel et al. 2012).

With the rapid development of the Next Generation Sequencing technology, a molecular approach is the current choice of technology to conduct the experiments and answer the research questions. The baseline study for the validation of the community DNA extraction and PCR optimisation are essential for downstream studies of sponge microbiome to achieve a precise and comprehensive profiling of the microbial community. Therefore, a validation test was designed in this study to examine the community DNA recovery efficiency of sponge samples and amplification efficiency of the target gene, such as partial of the 16S rRNA gene. The key findings are showed in Chapter III.

1.2.1 Sponge-associated microorganisms

Sponges could inhabit in both seawater and freshwater with large numbers, due largely to their adaptability to dramatic environmental changes (Gatti 2002; Gutt 2007). Sponges organise a simple body structure (Figure 1-1a): the specialised cells serve a variety of functions (Hentschel et al. 2012). One such example are pinacocytes, which cover microscopic chambers, internal canal system, and the outer body layer of sponges (van Soest et al. 2012). A matrix made with collagen fills the space between the chambers and canals, making an ideal

environment for bearing various type of cells, inorganic skeleton, and fibres (Van Soest et al. 2016). Inside the chambers line a layer of chanocytes – flagella-bearing cells that enable to create the water currents to conduct the unique filtering activity (Webster 2012).

Thanks to the effective water pumping capacity, sponges could collect the microscopic-size particles (Bell 2008) and permanently host highly dense and diverse symbionts (bacteria, archaea, and unicellular eukaryotes) in their tissues (Taylor et al. 2007). The microbes can contribute up to 40%-60% of the total sponge biomass (De Voogd et al. 2015; Hentschel et al. 2003; Vacelet & Donadey 1977) and are found at densities exceeding 10^9 microbial cells per cubic centimetre of sponge tissue (3-4 folds of magnitude greater than the density in the surrounding sea water) (Taylor et al. 2007). Most sponge-associated microbes inhabit extracellularly in the mesohyl tissue (Fig. 1-1). However, some symbionts may also exist intracellularly, as evidenced by early electron microscopy studies (Vacelet & Donadey 1977). In the mesohyl, microbes could be digested to use as food sources and most of them could live as sponge symbiotic community (Fig. 1-1b).

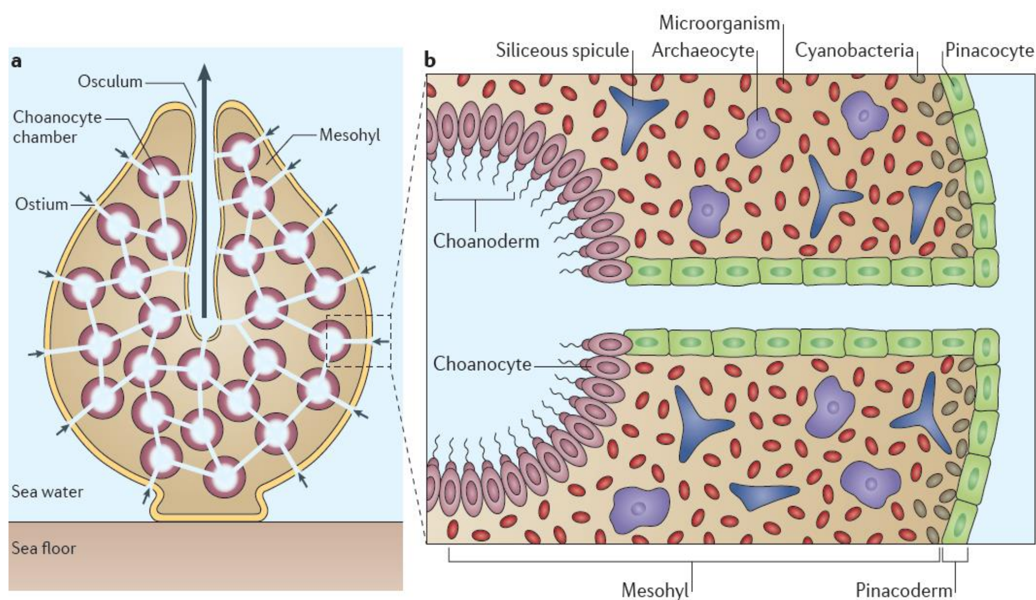


Figure 1 - 1 Overview and internal structure of a typical demosponge. (reproduced from Hentschel et al. 2012). a: Outline of sponge body; b. Internal structure.

An interesting question comes up that how these microbes avoid being digested. The studies showed that the potential symbionts ingested by their hosts could pass through the host unscathed. In contrast, the non-symbiotic bacteria were generally digested (Wehrl, Steinert & Hentschel 2007). It suggested that specific symbionts could have the ability to either be

recognised then ignored by the host sponge or not be recognised at all (Wilkinson, Garrone & Vacelet 1984). One possible reason is that sponges have a surprisingly well-developed innate immune system (Gauthier, Du Pasquier & Degnan 2010; Srivastava et al. 2010) and produce a wide range of antimicrobial compounds (Blunt et al. 2011). Conversely, microbial symbionts of sponges also enjoy some benefits. The efficient filter-feeding activities of the host provide a stable nutrient supply. The ammonia as a metabolic end product can be extracted by sponge to offer a supply of the scarce element nitrogen. The unique way that the sponge body is organised, together with its endurance of symbiotic microbes, make sponges great suitable organisms adapting to the various environmental challenges.

To date, our knowledge of sponge diversity is still far from complete. Approximately 8,700 species are considered valid and have been formally described (Van Soest et al. 2016), however, twice as many are believed to exist. Currently, sponges are divided into 31 taxonomic orders belonging to four classes (Morrow & Cárdenas 2015; Van Soest et al. 2016), but the evaluation of many taxa with higher rank are still limited by insufficient gene sequencing data and inadequate morphological characteristics (Knowlton 1993). Most sponges have regional or local distributions, because of occasional asexual propagation and the limited mobility of their larvae. A review of the distribution patterns of species and higher taxa over the global seas and oceans is summarised by van Soest et al. (2012). It is worth mentioning that the presented species are regarded as a bare minimum estimation of the actual distribution. Based on this review, most of the Australian coastline is covered by a yellow region with 1-100 species and in some parts 201- 461 species.

Based on the review of Sorokin et al. (2013), South Australia has a range of conditions for hosting a variety of sponge species with a south-facing coastline that is the longest in the Southern Hemisphere and a large continental shelf with low nutrients in water. In addition, it has two gulfs that contain sheltered estuaries (Fig. 1-2). Due to the Leeuwin Currently, the waters are usually warm in the west. In the southeast, however, as a result of upwelling and easterly coastal currents they are much cooler. The gulfs become warm enough in summer (>30°C) to support species usually encountered in the tropics.

Ninety named species of demosponge from South Australia, are listed in the Australian Faunal Directory (AFD), though this number falls short of the real diversity that exists in the State. A recent update reported by Sorokin et al. (2013) puts the number of demosponges documented from South Australia at 239 species. The putative species or operational taxonomic units (OTUs), which are not in the AFD list, are shown in Figures 1-3, and the additional Demospongiae species are listed in Table 1-2.

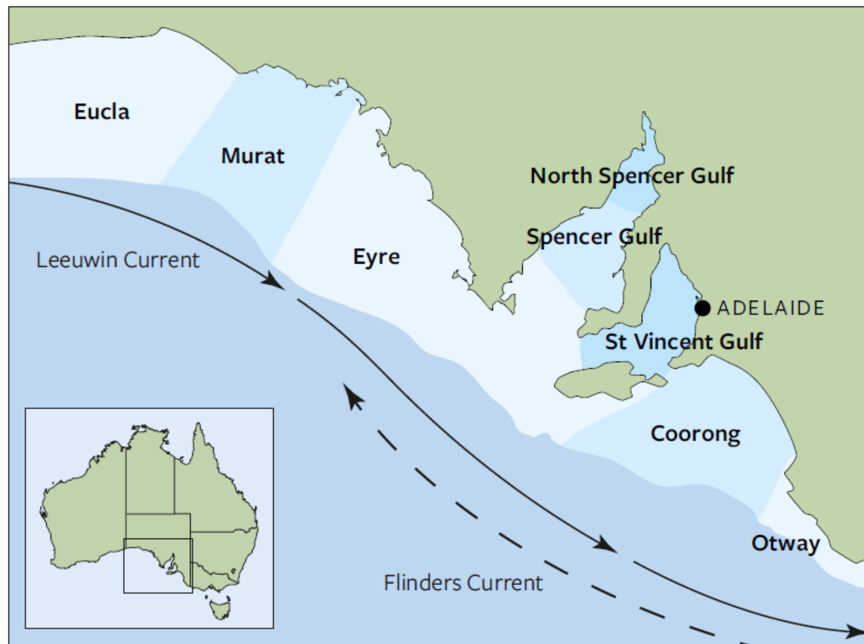


Figure 1 - 2 South Australian waters include eight Integrated Marine and Coastal Regionalisation of Australia (IMCRA) meso-scale bioregions (Sorokin et al. 2013).

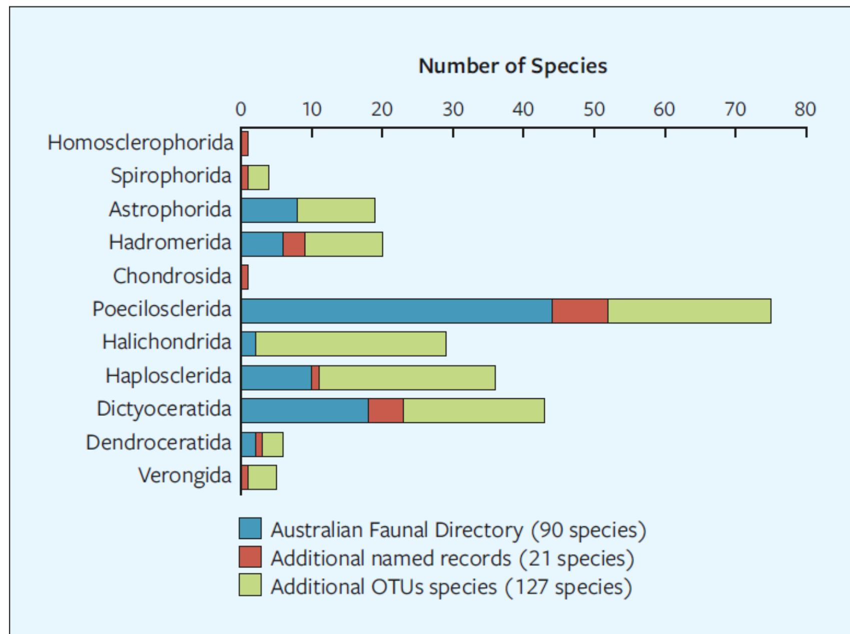


Figure 1 - 3 Taxonomic distribution within the class Demospongiae collected in South Australia (Sorokin et al. 2013).

Table 1 - 2 Demospongiae species (additional to the AFD list) documented from South Australia (Sorokin et al. 2013)

Order	Species
Homosclerophorida	*Corticium candelabrum* Schmidt, 1862
Spirophorida	*Amphitethya microsigma* Lendenfeld, 1907
Hadromerida	*Spheciospongia papillosa* Ridley & Dendy, 1886
	Spheciospongia purpurea Lamarck, 1815
	Tethya ingalli Bowerbank, 1858
Chondrosida	*Chondrilla austaliensis* Carter, 1873
Poecilosclerida	*Acarnus guentheri* Dendy, 1896
	Clathria (*Clathria*) *striata* Whitelegge, 1907
	Clathria (*Isociella*) *macropora* Lendenfeld, 1886
	Echinoclathria axinelloides Dendy, 1896
	Echinoclathria chalinoides Carter, 1885
	Raspailia (*Clathriodendron*) *cacticutis* Carter, 1885
	Strongylacidon stelliderma Carter,1886
	Tedania anhelans Lieberkuhn,1859
Haplosclerida	*Oceanapia ramsayi* Lendenfeld, 1888
Dictyoceratida	*Fenestraspongia intertexta* Carter, 1885
	Psammocinia cf. *bulbosa* Bergquist, 1995
	Psammocinia vesiculifera Poléjaeff, 1884
	Strepichordaia caliciformis Carter, 1885
	Thorectandra typica Carter, 1886
Dendroceratida	*Dictyodendrilla* cf. *dendyi* Bergquist, 1996
Verongida	*Aplysina lendenfeldi* Bergquist 1980

1.2.2 Sponge-associated actinobacteria: diverse and pharmaceutically valuable bacteria

Studies using culture-independent molecular approaches have demonstrated that novel, abundant actinobacterial assemblages are associated with sponges (Abdelmohsen, Bayer & Hentschel 2014; Montalvo et al. 2005; Webster et al. 2001; Xin et al. 2008; Zhang et al. 2006). One sixth of the sponge-associated microbial sequences in public databases belongs to the phylum Actinobacteria, indicating that actinobacteria constitute an important phylum among the

sponge-associated microorganisms (Xi, Ruan & Huang 2012). In order to assess the diversity of marine sponge-associated Actinobacteria (class), the review study of Abdelmohsen et al. (2014) constructed a maximum-likelihood phylogenetic tree of all actinobacterial 16S rRNA gene sequences with a length of >1300 bp and that were available in the NCBI database (Benson et al. 2009). The suborder Micrococcineae represents almost half of the genera isolated from marine sponges, among them *Micrococcus*, *Microbacterium*, and *Arthrobacter*, which are readily isolated because of their fast-growing nature. However, their potential for secondary metabolism appears to be limited to few reports (Bultel-Ponce et al. 1998; Lang et al. 2004). On the contrary, the single genus *Streptomyces* is represented by hundreds of sequence entries that were obtained from many different sponge species, and many of which display novel chemistry. Marine sponges are not only a rich source for diverse actinobacteria but also an impressive habitat for new and rare actinobacterial genera. Rare genera that have been recovered from sponges include *Actinokineospora*, *Actinomadura*, *Amycolatopsis*, *Knoellia*, *Nonomuraea*, *Pseudonocardia*, *Saccharomonospora*, *Saccharopolyspora,* and *Verrucosispora*, which could provide novel lead compounds in the future.

With the application of the sequence data, the diversity of sponge-associated Actinobacteria (phylum) has been understood better, 1,885 16S rRNA gene sequences including both culturable and unculturable sponge-derived actinobacteria were obtained from the GenBank database (Valliappan, Sun & Li 2014). Phylogenetic classification using Ribosomal Database Project (RDP) classifier indicated that the sponge-associated actionbacteria consisted of four classes, namely Acidimicrobiia, Actinobacteria, Nitriliruptoria, and Rubrobacteria, which together encompassed 112 genera. The class Actinobacteria had the abundant members of the sponge-associated actinobacterial communities with 102 genera (Table 1-3). Apart from the routine genera, several rare genera (for example *Ferrimicrobium*, *Ferrithrix*, *Humicoccus*, *Rubrobacter*, and *Serinicoccus*) were also associated with the sponges. Thus, the phylogenetic classification proved that the association between sponges and actinobacteria was diverse and it supports that the marine sponges are a rich source for isolating actinobacteria.

Table 1 - 3 Actinobacterial genera associated with marine sponges based on the RDP classification of 16S rRNA gene in GenBank (reproduced from Valliappan et al. 2014)

Acidimicrobiia (Class)	Actinobacteria (Class)		Nitriliruptoria (Class)	Rubrobacteria (Class)
Acidimicrobium	*Acaricomes*	*Microbacterium*	*Nitriliruptor*	*Conexibacter*
Aciditerrimonas	*Acidothermus*	*Microcella*		*Rubrobacter*
Ferrimicrobium	*Actinaurispora*	*Micrococcus*		*Thermoleophilum*
Ferrithrix	*Actinoalloteichus*	*Microlunatus*		
Iamia	*Actinokineospora*	*Micromonospora*		
Ilumatobacter	*Actinomadura*	*Modestobacter*		
	Actinomyces	*Mycetocola*		
	Actinotalea	*Mycobacterium*		
	Aeromicrobium	*Nocardia*		
	Agrococcus	*Nocardioides*		
	Agromyces	*Nocardiopsis*		
	Alloactinosynnema	*Nonomuraea*		
	Allokutzneria	*Okibacterium*		
	Amycolatopsis	*Ornithinibacter*		
	Angustibacter	*Ornithinimicrobium*		
	Arthrobacter	*Oryzihumus*		
	Auritidibacter	*Phycicoccus*		
	Bogoriella	*Phytohabitans*		
	Brachybacterium	*Piscicoccus*		
	Brevibacterium	*Planosporangium*		
	Brooklawnia	*Plantactinospora*		
	Catenuloplanes	*Plantibacter*		
	Cellulomonas	*Polymorphospora*		
	Cellulosimicrobium	*Propionibacterium*		
	Corynebacterium	*Pseudoclavibacter*		
	Cryobacterium	*Pseudonocardia*		
	Curtobacterium	*Quadrisphaera*		
	Demequina	*Rhodococcus*		
	Dermacoccus	*Rothia*		

Dermatophilus	*Saccharomonospora*
Devriesea	*Saccharopolyspora*
Dietzia	*Salinibacterium*
Frondihabitans	*Salinispora*
Geodermatophilus	*Sanguibacter*
Georgenia	*Serinicoccus*
Glaciibacter	*Sphaerisporangium*
Humicoccus	*Sporichthya*
Blastococcus	*Streptacidiphilus*
Isoptericola	*Streptomonospora*
Janibacter	*Streptomyces*
Jishengella	*Streptosporangium*
Klugiella	*Terracoccus*
Knoellia	*Thermasporomyces*
Kocuria	*Thermocrispum*
Krasilnikovia	*Thermomonospora*
Kribbia	*Tsukamurella*
Kytococcus	*Verrucosispora*
Leifsonia	*Ruania*
Leucobacter	*Xylanibacterium*
Marihabitans	*Yimella*
Marinactinospora	*Yonghaparkia*

Omnipresent in the marine setting, phylum actinobacteria is pivotal to the marine ecology. It recycles refractory biomaterials and utilise them to produce novel natural products, many of which with pharmaceutical applications. Oftentimes, actinobacteria produce bioactive compounds to battle its host against harmful pathogens. Similarly, actinobacteria presented in sponge could offer an important source for future development of marine drugs. With the culture-dependent methods, the number of descriptions of new actinobacteria species and even genera from sponge sources is continuously rising (Hameş-Kocabaş & Uzel 2012; Kwon et al. 2006; Supong, Suriyachadkun, Pittayakhajonwut, et al. 2013; Supong, Suriyachadkun, Suwanborirux, et al. 2013). One example is the obligate marine genus *Salinispora* represented by *S. arenicola*, *S. tropica* and *S. pacifica* which were discovered originally in sediments (Maldonado et al. 2005) but were since then also found in sponges such as *Pseudoceratina*

clavata from the Great Barrier Reef (Kim, Garson & Fuerst 2005). Another recent example is *Verrucosispora andamanensis* sp. nov., isolated from *Xestospongia* sp. collected from the Andaman Sea, Thailand (Supong, Suriyachadkun, Pittayakhajonwut, et al. 2013). More recently, 12 genera of actinobacteria including *Saccharopolyspora*, *Salinispora*, *Prauserella*, *Streptomyces*, *Micromonospora*, *Kocuria*, *Nocardiopsis*, *Serinicoccus*, *Micrococcus*, *Mycobacterium*, *Rhodococcus*, and *Actinoalloteichus* were isolated from the sponges collected from the Yongxing Island in the South China Sea (Valliappan, Sun & Li 2014). So far, 44 genera of the sponge-derived actinobacteria have been identified based on culture-dependent methods.

The distribution of actinobacteria in host sponges does not reveal any patterns that would point to a specific host-symbiont relationship. Instead, the actinobacteria appear to be distributed randomly in the host sponges investigated. The culturable actinobacterial genera from different sponge species have been reviewed by Abdelmohsen, Bayer & Hentschel (2014). This review reported 58 identified actinobacterial genera, which are isolated from 44 sponge species. The sponge species *Hymeniacidon perleve* was repeatedly examined from the two locations offshore China using different media formulations (Sun et al. 2010; Xi, Ruan & Huang 2012; Zhang et al. 2006). With the exception of *Streptomyces*, most other actinobacterial genera were variably present in the three sponges investigated. Similarly, there is no consistent pattern among the three closely related *Aplysina* sponges and neither among the three *Dysidea* sponge species. However, since a systematic study with replicate sampling over space and time aiming to resolve patterns of host specificity is still lacking, it is too early to draw any conclusions. A comprehensive study by Vicente and co-workers also revealed no evidence for a specific relationship between actinomycetes and the host sponges from which they were isolated. Rather sedimentation rate was identified as a determining factor in that sedimentation rich habitats provided more actinobacteria diversity and higher numbers of isolation than pristine waters (Vicente et al. 2013). In contrast to the sponge-specific microbial consortia consisting of other phyla that are vertically transmitted through the reproductive stages and permanently associated with their host sponge (Hentschel et al. 2012), the actinobacteria are very likely taken up from the environment by filtration and appear to persist in the mesohyl matrix.

Among the sponge associated bacteria, actinobacteria are often sponge-specific and a dominant producer of bioactive natural products (Abdelmohsen, Bayer & Hentschel 2014). Abundant bioactive compounds produced by marine actinobacteria includes peptides, polyketides, phenazines, sterols, isoprenoids, indolocarbazoles and others. They have been proven with antibacterial, anticancer, antioxidant, and anti-HIV activities. A comprehensive

review of the natural products produced by marine organism associated actinobacteria was reported by Valliappan, Sun & Li (2014). The studies on the gene levels also confirmed the functional values of the sponge-associated actinobacteria, such as, the presence of polyketide synthases (PKSs) and nonribosomal peptide synthetases (NRPSs) genes (Schneemann et al. 2010).

1.2.3 Sponge-associated microorganisms as promising source of natural bioactive compounds

One of the most valuable properties of marine sponges is their capacity to produce diverse types of natural bioactive compounds. Sponges are among the richest natural sources of secondary metabolites with biological actives and have produced much more compounds than any other marine organisms (Blunt et al. 2011). Many of these substances performed as protective agents against predators (Pawlik 2011), which are believed to contribute to the successful evolutionary development of sponges. Numerous sponge-derived natural products exhibit promising activities against various diseases. For example, eribulin, a synthetic analogue of halichondrin B from sponges *Halichondria* and *Lissodendoryx* spp., was recently approved as a drug for the treatment of metastatic breast cancer (Huyck et al. 2011). However, the challenge of providing long-term supplies for structurally complex metabolites can represent a crucial bottleneck in sponge-based drug discovery. Driven by these challenges, the possibility that the bioactive compounds could be produced by sponge-associated bacteria provides a promising opportunity to release sustainable natural products supply. Two strategies are available: cultivating the bacteria (producing the bioactive compounds) or transferring the biosynthetic genes into a culturable host to scale up the production of the secondary metabolites (Bewley & Faulkner 1998).

Sponge-associated bacteria and fungi are the two main groups of microbes producing antimicrobial compounds (Indraningrat, Hauke & Detmer 2016). Based on the review of Indraningrat, Hauke & Detmer (2016), 90% of the antimicrobial compounds were found in sponge-associated bacteria, while fungi contribute about 10% of the reported compounds. The antimicrobial compounds derived from sponge-associated bacteria can be found from 35 genera, which belong to four phyla of Actinobacteria (48.8%), Proteobacteria (36.6%), Firmicutes (11.4%), and Cyanobacteria (0.4%). In contrast, only one sponge-associated phylum Ascomycota that belongs to fungi has been reported to produce antimicrobials.

Genus *Streptomyces* contributes 30% of the sponge bacteria-derived compounds (Seipke, Kaltenpoth & Hutchings 2012; Traxler & Kolter 2015), particularly for strains of *Streptomyces* sp. HB202 and *Streptomyces* sp. RV15. Three antibacterial substances

(mayamycin, streptophenazine G, and K) have been reported to produce by *Streptomyces* sp. HB202, which is isolated from the sponge *Halichondria panacea*. The compound naphthacene glycoside extracted from *Streptomyces* sp. RV15 is the only anti-Chlamydia reported from sponge-associated microbes so far (Reimer et al. 2015). In addition, isolation of the anti-Trypanosoma and anti-Leishmania compounds valinomycin, staurosporine and butenolide from three strains of *Streptomyces* sp. 43, 21 and 11, respectively (Pimentel-Elardo et al. 2010), further confirm *Streptomyces* as the currently most prominent producer of antimicrobial substances from sponges. Genus *Pseudovibrio* is the second most prolific bacterial genus isolated from sponges (20%) with antimicrobial activities (Harrington et al. 2014). Additionally, sponge-associated *Bacillus* spp. was reported to produce 9% of the currently known bioactives.

For sponge-associated fungi, phylum Ascomycota has been found to produce antimicrobials, including 12 genera (Indraningrat, Hauke & Detmer 2016). Of these 12 fungal genera, *Aspergillus* (30%) and *Penicillium* (23%) are currently the two most prominent groups. *Aspergillus versicolor* (Lee et al. 2010) and an unidentified *Aspergillus* sp. isolated from the sponge *Xestospongia testudinaria* (Li et al. 2012) showed a strong antibacterial activity. The antimicrobial activities found from sponge-associated *Penicillium* spp. are particularly remarkable as it is the only fungal genus that is found to produce antivirals, antibacterials antifungals and antiprotozoals. *Penicillium chrysogenum* (Bringmann et al. 2003) and *Penicillium* sp. FF01 (Subramani et al. 2013) are to date the most promising sponge-associated *Penicillium* isolates with anti-HIV activity (sorbicillactone) and antibacterial activity (citrinin), respectively. Sponge-associated fungi should be considered as an important source of antimicrobial compounds, though the number of produced antimicrobials is still much less than that produced from the sponge-associated bacteria.

Based on the chemical structures of these compounds, six categories are concluded: peptides, terpenoids, phenazines, indoles, phenols, and polyketides (Indraningrat, Hauke & Detmer 2016). Peptide derivatives constitute 19% of the total identified antibacterial substances and about 12.5% from the total antimicrobial compounds based on the examination of the study Indraningrat et al. (2016). Phenazine derivatives are the second most frequently isolated class of antibacterial compounds from sponge-associated microbes (15%) (Choi et al. 2009; Jayatilake et al. 1996; Kunz et al. 2014). In addition, recent studies have begun to investigate the synthetic pathways of the natural product chemistry at the gene levels to demonstrate the clear link between the host sponge and its symbionts (Fisch et al. 2009; Hentschel et al. 2012; Piel et al. 2004). The bacterial symbionts are the main producer of the compounds and can grow on a large scale - a potential solution for the supplying issue of many pharmacologically valuable sponge natural products (Indraningrat, Hauke & Detmer 2016).

1.3 Sponge microbiome revealed by Next Generation Sequencing

Sponge associated microbial community have been revealed by the amplicon (16S rRNA gene) based metagenomic sequencing. Many efforts have been done to enhance the understanding of the sponge microbiome for approaching a complete microbial community. Different sequencing platforms and various sequencing primers specific to the 16S rRNA gene regions have been utilised to analyse sponge species globally. However, no universal primer set or the best 16S rRNA gene region, such as region V4 proposed by Earth Microbiome Project, had the capacity to uncover the complete sponge associated microbial community. Many studies have reported the data about the comparison and evaluation of the performance of different 16S rRNA gene region-specific primer sets, a critical issue yet to be solved is the impact of the region-specific primers selection on its performance in uncovering and differentiating the sponge microbiome. A validation is required to select a proper combination of the region-specific primer sets to reveal a comprehensive microbial community, which was conducted in this study and demonstrated in Chapter IV. In addition, the studies have reported the comparison of the sponge associated microbial communities between different sponge samples, or between the sponges and seawaters. All the studies focused on the sponge-specific or sponge species-specific microbial community using the single-primer-set based 16S rRNA gene metagenomic sequencing approach. However, it is still unclear whether the symbiotic microbial communities have the unique profile at other taxonomic level(s) of the host sponges. To better understand the sponge microbiome, a comprehensive analysis is essential to compare the sponge-associated microbial communities based on different phylogenetic ranks (e.g. family and order), which was designed in this study and reported in Chapter V.

1.3.1 Next Generation Sequencing technology

The first Next Generation Sequencing (NGS) technology was the pyrosequencing method by 454 Life Sciences (Margulies et al. 2005). The 454 Genome Sequencer generated about 200 000 reads (~20 Mb) of 110 base-pairs (bp). Following the 454 platform, the Solexa/Illumina sequencing system was commercialised. The third technology was based on Sequencing by Oligo Ligation Detection (SOLiD) designed by Applied Biosystems (Valouev et al. 2008). The Illumina and SOLiD platforms generated much larger numbers of reads than 454 but the reads produced were only 35 bp long. In 2010, Ion Torrent released the Personal Genome Machine (PGM), which was developed by Jonathan Rothberg, the founder of 454, and resembles the 454 system. An important improvement is that the PGM uses semiconductor technology and does not rely on the optical detection of incorporated nucleotides using fluorescence and camera scanning. This technical improvement allowed for the sequencing with higher speed, lower cost, and smaller instrument size. Meanwhile, other NGS methods have also been

developed, such as Qiagen intelligent biosystems sequencing-by-synthesis (Ju et al. 2006), Polony sequencing (Shendure et al. 2005), and a single molecule detection system (Pushkarev, Neff & Quake 2009).

Sequencing a short read then assembly is the most common approach of NGS. Based on the sequencing technologies, there are two broad categories: sequencing by ligation (SBL) and sequencing by synthesis (SBS) (Goodwin, McPherson & McCombie 2016). The generation of the clonal template populations applied different strategies: bead-based, solid-state and DNA nanoball generation. The DNA template generation starts from fragmentation of the sample DNA, then ligation to a common adaptor set for clonal amplification and sequencing. For bead-based method, one adaptor is complementary to an oligonucleotide fragment that is immobilised on a bead. Using emulsion PCR (emPCR) (Dressman et al. 2003), the DNA template is amplified such that as many as one million clonal DNA fragments are immobilised on a single bead. These beads can be distributed onto a glass surface or arrayed on a PicoTiterPlate (Roche Diagnostics) (Fedurco et al. 2006). For solid-state method, the amplification eschews the use of emPCR in favour of amplification directly on a slide (Harris et al. 2008). The forward and reverse primers are covalently bound to the slide surface, either randomly or on a patterned slide. These primers provide complementary ends to which single-stranded DNA (ssDNA) templates can bind. Precise control over template concentration enables the amplification of templates into localised, non-overlapping clonal clusters, thus maintaining spatial integrity. Recently, several NGS platforms have utilised patterned flow cells. By defining precisely where primers are bound to the slide, more DNA templates can be spatially resolved, enabling higher densities of reaction centre clusters and increasing sequencing throughput. The Complete Genomics technology is currently the only approach that achieves template enrichment in solution. DNA undergoes an iterative ligation, circularisation, and cleavage process to create a circular template, with four distinct adaptor regions. Through the process of rolling circle amplification (Fedurco et al.), up to 20 billion discrete DNA nanoballs are generated. The nanoball mixture is then distributed onto a patterned slide surface containing features that allow a single nanoball to associate with each location (Drmanac et al. 2010).

Different NGS platforms vary on their sequencing throughput, cost, error rate and read structure (Mardis 2013). 454 system offers superior read lengths compared to other short-read sequencers with reads up to an average of 700 bp, providing some advantages for applications that focus on repetitive or complex DNA. However, the insertion and deletion (indel) errors are dominate. Homopolymer regions are problematic for 454 platform, which lacks single-base accuracy in measuring homopolymers larger than 6-8 bp (Forgetta et al. 2013; Loman et al.

2012). Therefore, the 454 platform has been unable to complete with other platforms in terms of yield or cost. This limitation has led Roche to discontinue the platform in 2016 (GenomeWeb 2015). In contrast, Illumina platforms are getting popular for the current researches (Timmerman 2015). Illumina dominates the NGS industry due to its maturity as a technology, a high level of cross-platform compatibility and its wide range of platforms. The suite of instruments available ranges from the low-throughput MiniSeq to the ultra-high-throughput HiSeq X, which supports the sequencing ~1,800 human genomes to 30× coverage per year. Further diversification is derived from the many options available for runtime, read structure and read length (up to 300 bp). In addition, Illumina platform is much less susceptible to the homopolymer errors. Although it achieves an overall accuracy rate of >99.5% (Bentley et al. 2008), the platform sometimes under-represents AT-rich (Dohm et al. 2008; Harismendy et al. 2009) and GC-rich regions (Harismendy et al. 2009; Nakamura et al. 2011), and increases ricks of substitution errors (Minoche, Dohm & Himmelbauer 2011). In 2008, the two-colour labelling system used by the NextSeq and MiniSeq platforms increases speed and reduces costs by reducing scanning to two colour channels and reducing fluorophore usage. However, the two-channel system results in a slightly higher error profile and underperformance for low-diversity samples owing to more ambiguous base discrimination (Wang et al. 2015). HiSeq X is currently the highest-throughput instrument available; however, as a consequence of its optimisation, it is limited to just a few applications, such as whole-genome bisulfite sequencing. HiSeq X is further limited as an all-purpose instrument due to a required initial purchase of five or ten instruments (additional single instruments can be purchased after the initial commitment), placing this system out of reach of most facilities.

1.3.2 Sponge microbiome revealed by 16S rRNA gene based metagenomic sequencing

16S ribosomal RNA (rRNA) genes have been considered as a gold standard for phylogenetic analysis of microbial community and the taxonomic identification of microbes (Woese 1987). Two important characters support these suitability and applicability of the 16S rRNA genes. Firstly, the 16S rRNA genes commonly occur in the genome of prokaryotes (Acinas et al. 2004). Secondly, conservative but complex secondary structure of the 16S rRNA genes (Noller et al. 1985) provide the specific features to a particular taxonomic unit (Jonasson, Olofsson & Monstein 2002; Van de Peer, Chapelle & De Wachter 1996a).

The 16S rRNA gene analysis has been extensively utilised to identify bacterial species thanks to its highly-conserved sequences among diverse bacteria (Choi, Wyss & Gobel 1996; Clarridge 2004; Petti, Polage & Schreckenberger 2005; Schmalenberger, Schwieger & Tebbe 2001). Nine hypervariable regions located at the 16S rRNA genes (Van de Peer, Chapelle & De

Wachter 1996b) are flanked by conserved sequences, which can be used to design the universal primers to amplify the target regions (Baker, Smith & Cowan 2003; Lu et al. 2000). Many successful applications have been reported (Becker et al. 2004; Bertilsson, Cavanaugh & Polz 2002; Maynard et al. 2005; Rothman et al. 2002; Yang et al. 2002). In addition, rather than the amplification of the hypervariable regions, some studies focused on the species-specific sequences within a single hypervariable region to achieve a rapid identification (Stohr et al. 2005; Varma-Basil et al. 2004). Several online databases are available to access the collections of 16S sequences, such as the Ribosomal Database Project (RDP) (Cole et al. 2009), EZ-Taxon (Chun et al. 2007), GreenGenes (DeSantis et al. 2006), and SILVA (Pruesse et al. 2007). Unfortunately, a large part of the currently known bacterial phyla are only based on the 16S rRNA gene data but have no pure cultures available (Rappe & Giovannoni 2003; Wu et al. 2009).

The effective primers for PCR amplification are essential because a failed match between the primer sequences and the target region will not produce the amplicon (Lane et al. 1985). The first set of 16S rRNA gene primers was designed based on the conservative regions of different bacterial species and was named based on the positions of *Escherichia coli* 16S rRNA gene (Lane et al. 1985). With the number of known 16S rRNA sequences increasing, more primers have been designed, such as ARB project (from Latin *arbor*, tree) (Ludwig et al. 2004). Moreover, the bacterial phylum-specific primer set have been designed as well (Nubel, Garcia-Pichel & Muyzer 1997). However, the accumulated polymorphisms in the conservative regions of the 16S rRNA gene (Cole et al. 2009) could influence the efficiency of the commonly used primer sets for some particular bacteria (Baker, Smith & Cowan 2003; Huws et al. 2007; Teske & Sorensen 2007).

The selection of primers specific to different 16S rRNA gene regions is becoming increasingly important driven by the development of the amplicon sequencing in the NGS platforms. Roche 454 Genome Sequencer (GS) FLX Titanium machine allows a high throughput and parallel sequencing to reveal the microbial community of the environmental samples (Armougom & Raoult 2009; Huse et al. 2008; Liu et al. 2007; Rothberg & Leamon 2008). However, using widely accepted primers could lead to the decline of bacterial OTU coverage rates, as a result of the accumulation of known polymorphisms in the regions to design the primers (Jonasson, Olofsson & Monstein 2002). Therefore, for 16S rRNA gene based metagenomic sequencing, the inappropriate primers might omit some bacterial species and consequently lead to incomplete microbial profile, as described by Baker, Smith & Cowan (2003) and Huws et al. (2007). Recently, several known primers were evaluated in terms of the bacterial OTU coverage using metagenomic data and the ones with better efficacy were

proposed (Liu et al. 2008). However, there is no optimal or universal primers for bacterial 16S rRNA genes applied on environmental microbiome studies.

To profile the comprehensive microbial community is one of the most important tasks for microbiologists to explore various ecosystems. However, our understanding of the kingdom Bacteria remains limited because most bacteria can not be cultured or isolated under laboratory conditions (Rinke et al. 2013). In the past few decades, the culture-dependent technologies have been developed and applied. DGGE (Denaturing gradient gel electrophoresis) (Muyzer, de Waal & Uitterlinden 1993), T-RFLP (Terminal restriction fragment length polymorphism) (Liu et al. 1997), FISH (fluorescent in situ hybridization) (Wagner et al. 1998) and Genechips (He, Van Nostrand & Zhou 2012) were used as mainstream methods in studies of bacterial communities and diversity until the development of high-throughput sequencing technology. However, such methods lack details and resolution when they focus on certain groups of bacteria and profile communities. Recently, metagenomic methods provided by next-generation sequencing (Seipke, Kaltenpoth & Hutchings 2012) technology such as Roche 454 (Claesson et al. 2010; Tamaki et al. 2011) and Illumina (Bennett 2004) have shown a better role in profiling microbial communities (Hamady, Lozupone & Knight 2010; Zhang, Shao & Ye 2012).

The principles of these technologies bring with them a few limitations. For example, the study reported that the use of different primers might result in different DGGE patterns (Yu & Morrison 2004). Recent studies utilising high throughput technology have also demonstrated that suboptimal primer pairs would lead to the uneven amplification of certain species, causing either an under- or over-estimation of some species in a microbial community (Baker, Smith & Cowan 2003; Hamady & Knight 2009; Tringe & P. 2008; Wang & Qian 2009). Several studies have focused on optimal primer pairs or, equivalently, optimal variable regions for the study of bacterial communities (Kim, Morrison & Yu 2011; Klindworth et al. 2012; Yu et al. 2008). They utilised synthetic microbial communities and indicated that the taxa chosen in the experiments would largely influence the result. Consequently, the use of different sequencing technologies and targeting of different sub-regions of 16S rRNA genes will result in a distinct composition of a given microbial community.

Moreover, the studies particularly focused on NGS platforms to compare the phylogenetic sensitivity between the 16S rRNA sub-regions to find the best or optimal region (Beckers et al. 2016; Ghyselinck et al. 2013; Group 2012; Kim, Morrison & Yu 2011; Klindworth et al. 2012; Nelson et al. 2014; Tremblay et al. 2015). This methodology is inherently flawed on three major issues, which would unavoidably lead to biased validation and less convincing conclusion. Firstly, the sequencing length has limited coverage – with a few single regions of 16S rRNA

gene and only less than four consecutive hypervariable regions. (Bartram et al. 2011; Kircher & Kelso 2010). Secondly, the chosen primer sets – ones targeting partial rRNA gene – suffer from biased bacterial taxa coverage as they have varying capacities of PCR amplification (Kumar et al. 2011; Wang & Qian 2009). Thirdly, uncertainty dramatically increases in taxonomic assignments at lower levels with the read length being too short (Wang et al. 2007). Why does it matter for strictly correct classification? For one thing, most functional microbes need a classification at the genus/species level for their functional study within a bacterial community. It is therefore important to develop a methodology on validating the optimal combination of the regions instead of the best selecting a single one to reveal the bacterial community.

Sponge microbiome, a biologically complex community, was revealed by culture-dependent methods before the year of 1999. For the last two decades, the development of the sponge microbiome study can be divided into three stages (Fig. 1-4). Fluorescence in situ hybridization (FISH) was applied for the first time in 1999, which revealed the microbial community associated with sponge *Aplysina cavernicola* (Friedrich et al. 1999). Schmidt et al. (2000) subsequently used denaturing gradient-gel electrophoresis (DGGE) in sponge microbiome study. In 2003, terminal restriction fragment length polymorphism (T-RFLP) was applied in conjunction with DGGE to reveal the microbial community associated with sponge *Aplysina aerophoba* (Ahn et al. 2003). Till 2005, after the first Next Generation Sequencing instrument (454 pyrosequencing platform) was released, researchers started uncovering more comprehensive microbiomes using the high throughput sequencing platform (van Dijk et al. 2014). The year after, a sequencing platform (Illumina system) with higher throughput, higher quality, and shorter reading length was developed (van Dijk et al. 2014). The first sponge microbiome study using 454 pyrosequencing platform was reported by Webster et al. (2010). The Illumina platform was successfully applied in 2014 to profile the community structure of host-specific microbiome of tropical marine sponges (Easson & Thacker 2014).

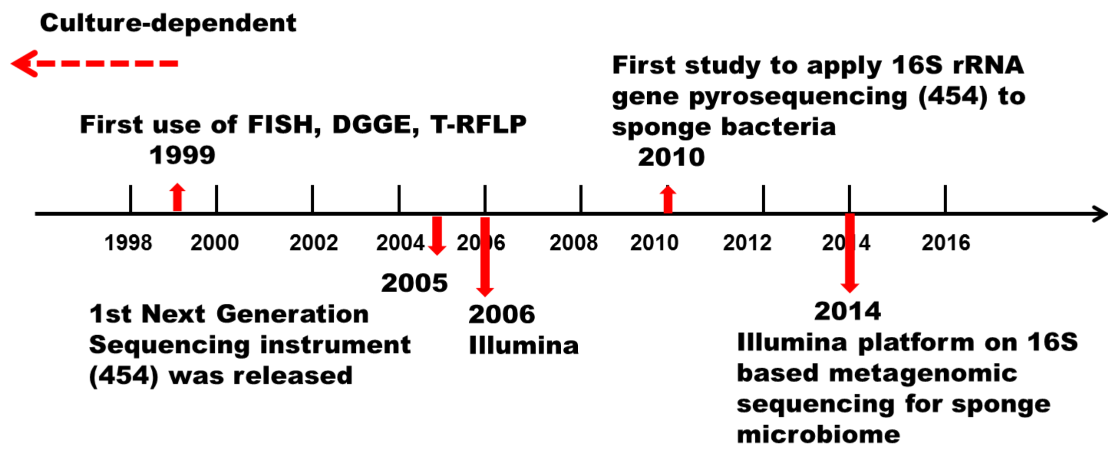


Figure 1 - 4 Development of the culture-independent technologies applied in sponge microbiome study

Analyses using metagenomics provide the depth of results required for a more comprehensive characterisation and understanding of temporal and spatial bacterial community composition. Given the technological advancements in computing power and statistical modelling capabilities for metagenomic data, the characterisation of sponge microbial communities using metagenomics is now becoming more cost-effective and widely available. Although the diversity of sponge symbionts has been extensively addressed using molecular tools, comparative work is hindered due to methodological differences in sampling, sample processing and data analyses (Simister et al. 2012; Taylor et al. 2007; Webster et al. 2010). A few global projects - the Earth Microbiome Project (Gilbert, Jansson & Knight 2014), for instance - have standardised these technical aspects to reliably and consistently describe patterns of microbial diversity and composition. These efforts have generated a large knowledge base for host-associated microbiomes, but equivalent data sets for sponges are still limited. To get a better understanding of the evolution and complexity of microbial symbiotic interactions with sponge, many studies have been conducted using different NGS platforms with various 16S rRNA gene region-specific primer sets (Table 1-4). However, a critical issue remains to be solved: the impact of the region-specific primers selection on its performance in uncovering and differentiating sponge microbiomes.

Table 1 - 4 Sponge microbiome studies using various 16S rRNA gene region-specific primer sets on different NGS platforms

Sponge species	Sequencing Platform	Sequencing region(s) (primer set)	Reference
81 sponge species	Illumina HiSeq	V4 (515F/806R)	Thomas et al. 2016
Hymeniacidon heliophila	Illumina MiSeq	V4 (515F/806R)	Weigela & Erwin 2016
Myxilla (Burtonanchora) sp. *Clathria* sp. *Kirkpatrickia variolosa* *Hymeniacidon torquata* *Leucetta antarctica* *Haliclona (Gellius)* sp. *Megaciella annectens*	Illumina MiSeq	V4 (515F/806R)	Rodríguez-Marconi et al. 2015
Agelas oroides *Chondrosia reniformis* *Petrosia ficiformis* *Dysidea avara* *Axinella damicornis* *Spirastrella cunctatrix*	Illumina MiSeq	V4 (515F/806R)	Ribes et al. 2015
Carteriospongia foliascens	Illumina	V4 (515F/806R)	Luter et al. 2015
Aiolochroia crassa *Amphimedon compressa* *Amphimedon erina* *Aplysina cauliformis* *Aplysina fulva* *Chalinula molitba* *Chondrilla caribensis* *Dysidea etheria* *Ectyoplasia ferox* *Erylus formosus* *Haliclona tubifera* *Haliclona vansoesti* *Iotrochota birotulata* *Lissodendoryx colombiensis* *Mycale laevis* *Mycale laxissima* *Niphates erecta* *Placospongia intermedia* *Tedania ignis* *Xestospongia bocatorensis*	Illumina	V4 (515F/806R)	Easson & Thacker 2014
Hexadella spp. *Mycale* spp.	Illumina HiSeq	V6 (967F/1046R)	Reveillaud et al. 2014

Xestospongia spp.	454	V1V2 (27F/338R)	Montalvo et al. 2014
Raspailia ramose *Stelligera stuposa*	454	V1V3 (63F/518R)	Jackson, S et al. 2012
Axinella corrugata	454	V1V3 (27F/533R)	White et al. 2012
Aplysina aerophoba *Aplysina cavernicola* *Ircinia variabilis* *Petrosia ficiformis* *Pseudocorticium jarrei*	454	V3 (338F/533R)	Schmitt, Hentschel & Taylor 2012
Stylissa massa *Xestospongia testudinaria*	454	V3V4	De Voogd et al. 2015
Suberites diversicolor *Cinachyrella australiensis*	454	V3V4	Cleary et al. 2013
Ancorina alata *Ancorina* sp. *Aplysina aerophoba* *Aplysina archeri* *Aplysina cavernicola* *Biemna ehrenbergi* *Chondrilla australiensis* *Cymbastela coralliophila* *Hippospongia* sp. *Hyrtios altum* *Hyrtios erectus* *Hyrtios* sp. *Ircinia felix* *Ircinia gigantea* *Ircinia* sp. *Ircinia variabilis* *Petrosia ficiformis* *Plakina trilopha* *Polymastia* sp. *Pseudoceratina crassa* *Pseudoceratina* sp. *Pseudocorticium jarrei* *Rhabdastrella globostellata* *Stelletta aremaria* *Stelletta maori* *Stylissa massa* *Theonella swinhoei* *Xestospongia* aff. *carbonaria* *Xestospongia muta* *Xestospongia* sp. *Xestospongia testudinaria*	454	V3 (338F/533R)	Schmitt et al. 2012

Cinachyrella spp.	454	V4 (515F/806R)	Cuvelier et al. 2014
Stylissa carteri *Xestospongia testudinaria*	454	V4V5 (533F/907R)	Moitinho-Silva et al. 2014
Hyrtios erectus *Stylissa carteri* *Xestospongia testudinaria*	454	V6 (789F/106R)	Lee et al. 2011
Ianthella basta *Ircinia ramose* *Rhopaloeides odorabile*	454	V6 (967F/1046R)	Webster et al. 2010
Carteriospongia foliascens	454	V6V9 (905F/1492R)	Gao et al. 2014
Petrosia ficiformi	454	V6V8 (926F/1392R)	Burgsdorf et al. 2014

1.3.3 Sponge specific microbial community

Sponges (Porifera), the oldest multicellular animals, permanently host abundant and diverse symbiotic microorganisms. Interactions between sponges and their associated microorganisms could be mutualism, commensalism, or parasitism (Hentschel, Usher & Taylor 2006). A large number of the studies have focused on these special associations between sponges and their symbiotic microbes, and some general opinions have been concluded: Firstly, the sponge associated microbial community is of host specificity. Secondly, sponge microbiome has apparent difference from the communities of the surrounding water or sediment samples. Thirdly, the sponge host and its associated microbial community enjoy the benefits each other by sharing the nutrients or functional elements (Fan et al. 2012; Jackson et al. 2012; Lee et al. 2011; Schmitt et al. 2012; Webster et al. 2010). Specifically, about 40 bacterial phyla, including the candidate phyla, have been reported from the sponges in the temperate and tropical regions. Five out of the 40 bacterial phyla are generally dominant and common: Actinobacteria, Bacteroidetes, Chloroflexi, Cyanobacteria, and Proteobacteria (Alex et al. 2013; Hentschel et al. 2002; Kennedy et al. 2014; Lee et al. 2011; Webster & Hill 2001). Based on the molecular data, more than 100 sponge-specific 16S rRNA gene clusters were proposed, such as candidate phylum Poribacteria (Fieseler et al. 2004; Taylor et al. 2013) and Candidatus Synechococcus spongiarum (Simister et al. 2012). Moreover, the high throughput sequencing data also demonstrated that marine sponges host highly diverse microorganisms from the Archaea and Eukarya. Phylum Thaumarchaeota is one of the two major described Archaeal phyla, which has been reported from several sponges with dominant percentage (Jackson et al. 2013; Pape et al. 2006). In addition, 11 eukaryotic phyla belonging to fungi and protists have been revealed from sponges (Cerrano et al. 2004; He et al. 2014; Jin et al. 2014; Webster et al. 2004).

Sponges are generally divided into two groups based on the density of their associated microorganisms. High microbial abundance (HMA) sponges are defined as the sponges with highly diverse and abundant microbes (10^8–10^{10} bacteria/g or ml of tissue) (Hentschel et al. 2003; Hentschel, Usher & Taylor 2006). In contrast, low microbial abundance (LMA) sponges are the ones with only 10^5–10^6 bacteria per gram or per ml of the sponge biomass. Particularly, microbial phyla Proteobacteria and Bacteroidetes were found to be dominant in the communities, which was more similar with the microbial community of the seawater samples (Erwin, Olson & Thacker 2011; Giles et al. 2013; Hentschel, Usher & Taylor 2006). Notably, the biggest differences between the HMA and LMA sponges are their water pumping capacity (HMA: lower pumping efficiency; LMA: higher pumping efficiency) and the way how their associated microbes make use of the nutrients, e.g. HMA: photosynthetic (Weisz et al. 2007), LMA: heterotrophic (Freeman & Thacker 2011; Schöttner et al. 2013; Weisz, Lindquist & Martens 2008). Transmission electron microscopy images have shown that the mesohyl of LMA sponges has limited microorganisms (Giles et al. 2013; Kamke, Taylor & Schmitt 2010; Vacelet & Donadey 1977; Weisz et al. 2007). Their larvae are also largely bacteria free (Schmitt et al. 2007). The microbial diversity of LMA sponges was investigated far less than HMA sponges (Schmitt et al. 2012).

On the other hand, the concepts emphasising the presence of "core" microbial taxa, "variable" microbial taxa, and "host-specific" microbial taxa were pointed out (Schmitt et al. 2012). Different types of the techniques have been applied for the microbiome studies, such as clone library sequencing, terminal restriction fragment length polymorphisms (T-RFLPs), denaturing gradient gel electrophoresis (DGGE), and automated ribosomal intergenic spacer analysis (ARISA). Based on the studies of HMA and LMA sponges to date, their microbial communities were found to be highly specific to the host phylogeny identities (Anderson, Northcote & Page 2010; Erwin, Olson & Thacker 2011; Olson, Thacker & Gochfeld 2014; Pita et al. 2013; Schöttner et al. 2013). NGS technologies have greatly improved the throughput and quantity of the sequencing data to reveal sponge microbiome (Reveillaud et al. 2014; Schmitt et al. 2012; Webster et al. 2004). As a result, sponge microbiome was further demonstrated to be largely host-specific, though it was also influenced by the seasonal, environmental, and geographic factors (Cleary et al. 2013; Cuvelier et al. 2014; Hardoim et al. 2012; Jeong, Kim & Park 2013; Kennedy et al. 2014; Luter et al. 2015; Montalvo et al. 2014; White et al. 2012).

To clearly define the specificity of the host-microbe association, the term sponge-specific cluster was applied. It refers to a group of 16S rRNA gene sequences that have been revealed from the sponge species in same or different geographical locations with close relationship to each other (Hentschel et al. 2002). The sponge-specific 16S rRNA gene clusters could suggest

the co-evolved host sponges and their associated microbial communities (Thacker & Freeman 2012). However, the sponge-specific clusters are rarely found in LMA sponge species (Erwin, Olson & Thacker 2011; Giles et al. 2013). In consequence, the microbiome composition of LMA sponges and its variation over space, time and with host phylogeny are currently far from being understood.

Notably, the application of the high throughput NGS datasets allowed deep analyse of the sponge microbiome, although the massive volumes of the data generated from NGS platform greatly increased the challenge of computing the analysis (Huse et al. 2010; Sogin et al. 2006). The most important advantage of NGS technology is to reveal the low abundant microbial taxa, for example, Reveillaud et al. (2014) reported the extremely rare microbes associated with sponge and their host-specificity. Not only the low abundant microbial taxa could be recovered, but the low abundant microbial community of the LMA sponge species also was revealed and confirmed to be of host species specificity (Giles et al. 2013). Moreover, a recent study compared the microbial community composition of an LMA sponge (*Stylissa carteri*), an HMA sponge (*Xestospongia testudinaria*) with seawater by 454 amplicon sequencing and proposed a new term of "sponge-enriched" to improve the original "sponge-specific" (Moitinho-Silva et al. 2014).

When considering the environmental influences, it is found that the most of the sponge microbiome studies have been focused on the temperate and tropical environments. In terms of other environmental types, such as the polar areas and the deep sea, only few studies have been reported (McClintock et al. 2005). For example, Webster et al. (2004) selected five sponge species living in the Antarctic regions with extreme environmental conditions (Hollibaugh, Lovejoy & Murray 2007), and demonstrated the host specific sponge associated microbial communities. In addition, the microbial communities of Antarctic sponges were also revealed by other studies, which primarily focused on the culturable microbial community and microscopic observations (Cerrano et al. 2004; Henríquez et al. 2014; Mangano et al. 2011).

The most recent research reported a comprehensive analysis of microbial communities associated with 81 sponge species (Thomas et al. 2016). A total of 804 sponge samples were collected from the shallow waters globally. 133 seawater and 36 sediment samples were also collected to compare their microbial communities with those of sponges. Microbial community composition for each sample was revealed using 16S rRNA gene-sequencing protocols established by the Earth Microbiome Project (EMP) (Gilbert, Jansson & Knight 2014). The concepts of the diversity, variability, and specificity of sponge associated microbial community were proposed.

1.4 Research plan

1.4.1 Hypotheses

Marine sponges are the most ancient living Metazoa and generally form symbiotic relationships with complex communities of microorganisms (Hentschel et al. 2012). Sponges can maintain highly diverse, yet specific symbiont communities, despite the constant influx of microorganisms in seawater resulting from the filter-feeding activities (Thomas et al. 2016). These symbioses are known to be at least partially underpinned by metabolic exchange between symbiont and host (Taylor et al. 2007). In this respect, the sponge species-specific symbionts perform analogous functions to the symbionts found in mammalian guts and plants. Therefore, sponge-microbe symbioses represent an ecologically relevant example of host-microbe interactions in an early-diverging metazoan clade. The species-specific microorganisms and their natural bioactive metabolites have significant value for novel compounds discovery for pharmaceutical and agrichemical applications (Abdelmohsen, Bayer & Hentschel 2014; Hentschel et al. 2012; Indraningrat, Hauke & Detmer 2016).

A comprehensive understanding of sponge microbiomes will help us explore these natural bioactive metabolites. First of all, the sponge identification is a basic requirement when studying the sponge host-specific microbial community. Although the application of DNA markers could improve the accuracy and efficiency of the traditional morphological classification, there is no reliable molecular protocol available currently.

Successful DNA extraction to obtain both high abundant and low-abundant sponge-associated actinobacterial DNA from the whole sponge tissue is also a primary requirement. In addition, it was not clear whether there were inhibitors in the sponge DNA samples to hinder the amplification of actinobacterial DNA when they are in such a complex biological mix. PCR amplification is always limited by the primer selection and the reaction conditions. However, no validation or optimisation had been done.

The diversity of the sponge-associated microbial community has been extensively addressed using molecular tools, though comparative work has been hindered due to methodological differences in sampling, sample processing and data analyses (Thomas et al. 2016). Large-scale efforts, such as the Human Microbiome Project (Human Microbiome Project Consortium 2012) and the Earth Microbiome Project (Gilbert, Jansson & Knight 2014), have standardised these technical aspects to describe patterns of microbial diversity and composition. These efforts have generated a large knowledge base for host-associated microbiomes of vertebrates, and especially humans, but equivalent data sets for invertebrates

are missing. Notably, most of the studies utilised a single primer set for particular region(s) to reveal what was considered to be an almost complete microbial profile and the validation work done was limited only on finding the best or universal primer set. Therefore, a critical issue needed to be solved: The impact of the region-specific primer selection on its performance in uncovering and differentiating the sponge microbiome. An optimised sequencing approach needed to be developed to reveal a more comprehensive sponge microbiome.

To gain critical insights into the complexity of symbiotic interactions, we require a greater understanding of both structure and composition of the microbial community. To date, microbiome research has focused on species and within-species comparisons or the comparative analysis of microbiomes of disparate host organisms (for example, sponges versus seawater/ sediments) primarily based on the known OTUs. However, to further define the host-specificity, the specificity needs to be tested at different taxonomic levels within the phylum Porifera (for example, between different orders in one class or between different families in one order). The unaffiliated microbial OTUs also need to be considered to reveal the microbiome.

Driven by these issues and gaps, the following hypotheses were set out to test:

I. A multilocus-based molecular identification protocol can be effective and reliable for sponge (phylum Porifera) classification.

II. The extraction method selected is effective for the recovery of microbial community DNA as demonstrated with actinobacteria (spores and mycelia) within the sponge microbial community.

III. The influence of the inhibitor(s) if existing in DNA preparation to PCR amplification can be relieved by optimising the conditions for microbial 16S rRNA gene PCR.

IV. A more comprehensive and reliable sponge microbiome can be revealed through the use of an optimum number of primer-sets targeting different hypervariable regions of the 16S rRNA gene.

V. The structure and the composition of the microbial community of the sponges collected within the same location and the same season are highly specific to the host phylogenetic status (at order and family levels).

1.4.2 Aims of the project

The aims of the project are:

I. To develop a multilocus-based sponge identification protocol to achieve an effective and reliable molecular identification by integrating the various accuracies of the different DNA markers;

II. To validate the community DNA extraction method and optimise the PCR amplification conditions to provide high quality and valid DNA templates for microbial 16S rRNA gene based amplicon sequencing;

III. To evaluate different 16S rRNA gene region-specific primer sets or their combinations using Next Generation Sequencing platform -Illumina MiSeq- to demonstrate the impact of the sequencing region selection on the integrity of the microbial community and to propose an improved amplicon sequencing approach by applying the validated primer sets.

IV. To reveal the composition and the structure of the host-specific microbial communities associated with the sponges at the order and family levels collected in the same location and season using the newly developed approach.

1.4.3 Applied methods and techniques

Multilocus-based sponge molecular identification protocol was developed to infer the identities of the specimens to select the representative sponge species belonging to different families within one order as well as the species belonging to different orders within one class. DNA barcoding was applied in the sponge identification protocol (SIP). The phylogenetic analysis was utilised to validate the reliability of the developed SIP. The morphological classification was compared with the molecular identification for the efficiency and accuracy assessment of the different DNA markers with different evolutionary rates (the COI mtDNA, the 28S rRNA gene, and the nuclear ITS region). To successfully conduct the 16S rRNA gene based metagenomic sequencing, the DNA extraction protocol optimised from the commonly used cetyltrimethylammonium bromide (CTAB) method and the commercial DNA extraction KIT were compared to validate the actinobacterial DNA recovery efficiency by artificially adding actinobacterial spores or mycelia powder into the a given freeze-dried sponge sample. The different combinations of the PCR reaction conditions and programs were tested to optimise PCR amplification protocol to achieve the qualified amplicons for the following metagenomic sequencing. The various 16S rRNA gene region specific primer sets were evaluated to propose a suitable primer set or their combination to reveal a comprehensive sponge-associated microbial community. High throughput sequencing (16S rRNA gene based amplicon sequencing) data were generated by the Next Generation Sequencing (NGS) - Illumina MiSeq platform. The sequencing data were processed and analysed by bioinformatics program. The

statistical analysis were also utilised to compare and analyse the composition and the structure of the sponge-associated microbial communities at both family and order levels to test the hypotheses.

1.4.4 Project framework

The project was conducted according the framework in Figure 1-5.

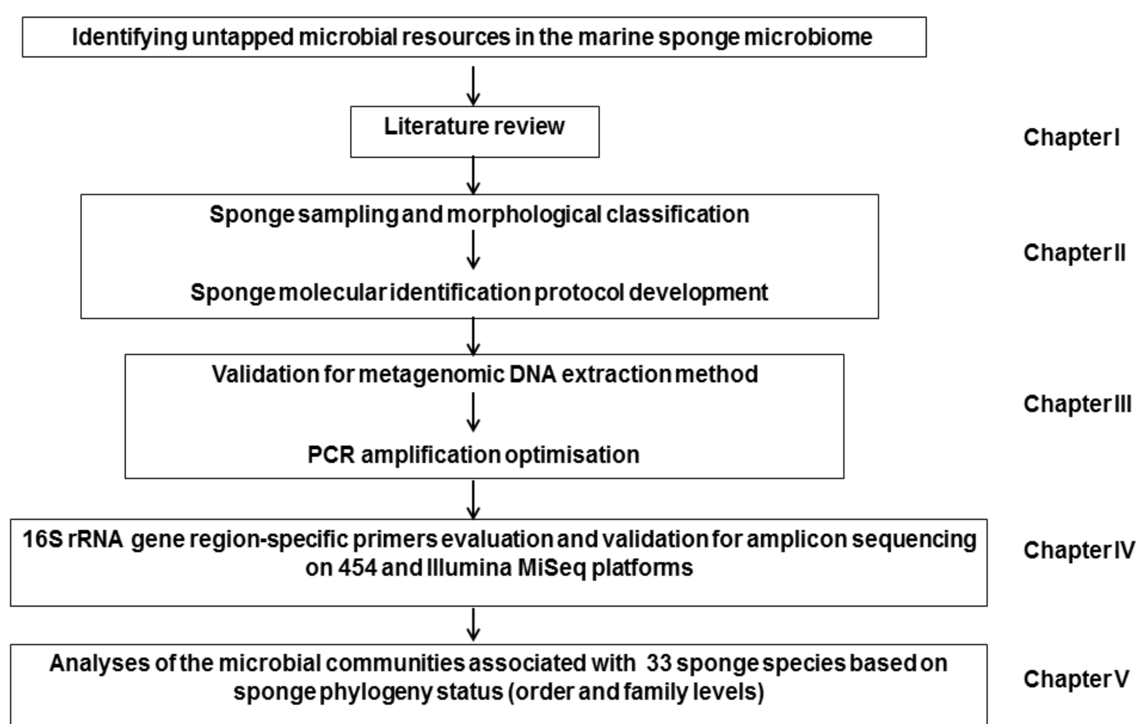


Figure 1 - 5 PhD project framework

1.4.5 Thesis structure and contributions

The PhD project mainly focuses on the methodology development and validation to reveal comprehensive microbial communities associated with sponges at higher taxonomic ranks (order and family levels). The whole thesis includes six chapters.

Chapter I introduces the research background of this thesis. It starts with the sponge classification. The limitation of the morphological characterisation and application of the molecular taxonomy are discussed to highlight the necessity of an effective and reliable sponge identification protocol. The current understanding of sponge-associated microbial community and the values for pharmaceutical application are reviewed. The current advanced technologies

and the main findings on the study of sponge microbiome are summarised and discussed. Based on the background understanding of the sponge microbiome, the hypotheses and aims of this project are developed. The technologies and the framework are constructed.

Chapter II illustrates the development of a multilocus-based approach for sponge (Porifera) identification. This study revealed the issues and the limitations of the sponge molecular identification as well as provided the reasons and solutions to improve the protocol. This part of work has been submitted to the peer-reviewed journal Scientific Reports, and requires minor corrections. The contribution statement of the publication: The research idea was provided by Qi Yang (QY) and Wei Zhang (WZ). The experiments and the data analyses were conducted by QY. The sponge morphological classification was contributed by Shirley Sorokin (SS) and QY. The manuscript was written by QY and revised extensively by CF, WZ and SS.

Chapter III discusses the validation of the modified DNA extraction methods and optimisation of 16S rRNA gene amplification for sponge-associated microbial community study. This experiment was designed to test whether the modified metagenomic DNA extraction method enables to pick up the actinobacterial spores and mycelia, which are the key components to cause a suboptimal DNA product. On the other hand, the inhibitor(s) during the PCR amplification was identified and reduced to optimise the protocol for the following sponge metagenomic DNA study. The findings of this study have been published in the peer-reviewed journal Applied Microbiology and Biotechnology (see the cover page of the published format in the Appendices Page 247. The contribution statement of the publication: the experimental design was provided by QY and CF, the experiments, the data analyses, and manuscript writing were conducted by QY. CF and WZ provided the contributions on data analysis and manuscript revision.

Chapter IV compares and evaluates different 16S rRNA gene region-specific primer sets to reveal the microbial communities of the same sponge DNA samples on 454 pyrosequencing and Illumina MiSeq sequencing platforms. Based on the bioinformatics analyses, an improved approach was proposed to utilise multiple region-specific primer sets to reveal comprehensive sponge microbiomes.

Chapter V reveals the microbial communities associated with 33 different sponge species and discusses the communities based on the sponge phylogenetic identities at the order and family levels. The specific microbial groups are uncovered for the particular sponge hosts.

Chapter VI concludes the key findings of this project and highlights the significance of this

study. The limitations and challenge of this study was evaluated. Additionally, the further research directions are discussed. Since the environmental factor influence on the microbial communities is the other important aspect of sponge microbiome study, the community shifts with the environmental changes, including physical factors, chemical factors, and biological factors, such as diseased hosts, are reviewed in this chapter.

References

Abdelmohsen, UR, Bayer, K & Hentschel, U 2014, 'Diversity, abundance and natural products of marine sponge-associated actinomycetes', *Natural Product Reports*, vol. 31, pp. 381-99.

Abdelmohsen, UR, Pimentel-Elardo, SM, Hanora, A, Radwan, M, Abou-El-Ela, SH, Ahmed, S & Hentschel, U 2010, 'Isolation, phylogenetic analysis and anti-infective activity screening of marine sponge-associated actinomycetes', *Marine Drugs*, vol. 8, no. 3, pp. 399-412.

Acinas, SG, Marcelino, LA, Klepac-Ceraj, V & Polz, MF 2004, 'Divergence and redundancy of 16S rRNA sequences in genomes with multiple rrn operons', *Journal of Bacteriology*, vol. 186, pp. 2629-35.

Ahn, Y-B, Rhee, S-K, Fennell, DE, Kerkhof, J, Hentschel, U & Häggblom, MM 2003, 'Reductive dehalogenation of brominated phenolic compounds by microorganisms associated with the marine sponge *Aplysina aerophoba*', *Applied and Environmental Microbiology*, vol. 69, pp. 4159-66.

Alex, A, Silva, V, Vasconcelos, V & Antunes, A 2013, 'Evidence of unique and generalist microbes in distantly related sympatric intertidal marine sponges (Porifera: Demospongiae)', *PLoS ONE*, vol. 8, no. 11, p. e80653.

Alvarez, B, Crisp, MD, Driver, F, Hooper, JNA & Van Soest, RWM 2000, 'Phylogenetic relationships of the family Axinellidae (Porifera: Demospongiae) using morphological and molecular data', *Zoologica Scripta*, vol. 29, no. 2, pp. 169-98.

Alvarez, B, De Voogd, NJ & Van Soest, RWM 2016, 'Sponges of the family Axinellidae (Porifera: Demospongiae) in Indonesia', *Zootaxa*, vol. 4137, no. 4, pp. 451-77.

Alvarez, B, Krishnan, M & Gibb, K 2007, 'Analysis of intragenomic variation of the rDNA internal transcribed spacers (ITS) in Halichondrida (Porifera: Demospongiae)', *Journal of the Marine Biological Association of the United Kingdom*, vol. 87, pp. 1599-605.

Anderson, SA, Northcote, PT & Page, MJ 2010, 'Spatialand temporal variabilityof the bacterial community in different chemotypes of the New Zealand marine sponge *Mycale hentscheli*', *FEMS Microbiology Ecology*, vol. 72, pp. 328–42.

Andreakis, N, Luter, HM & Webster, NS 2012, 'Cryptic speciation and phylogeographic relationships in the elephant ear sponge *Ianthella basta* (Porifera, Ianthellidae) from northern Australia', *Zoological Journal of the Linnean Society*, vol. 166, pp. 225-35.

Armougom, F & Raoult, D 2009, 'Exploring microbial diversity using 16S rRNA high-throughput methods', *Journal of Computer Science & Systems Biology*, vol. 2, pp. 74-92.

Australian Faunal Directory Australian Biological Resources Study, Canberra.,

http://www.environment.gov.au/biodiversity/abrs/online-resources/fauna/afd/index.html.

Baker, GC, Smith, JJ & Cowan, DA 2003, 'Review and re-analysis of domain-specific 16S primers', *Journal of Microbiological Methods*, vol. 55, pp. 541-55.

Bartram, AK, Lynch, MDJ, Sterns, JC, Moreno-Hagelsieb, G & Neufeld, JD 2011, 'Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end Illumina reads', *Applied and Environmental Microbiology*, vol. 77, no. 11, pp. 3846-52.

Becker, K, Harmsen, D, Mellmann, A, Meier, C, Schumann, P, Peters, G & von Eiff, C 2004, 'Development and evaluation of a quality-controlled ribosomal sequence database for 16S ribosomal DNA-based identification of *Staphylococcus* species', *Journal of Clinical Microbiology*, vol. 42, pp. 4988-95.

Beckers, B, Op De Beeck, M, Thijs, S, Truyens, S, Weyens, N, Boerjan, W & Vangronsveld, J 2016, 'Performance of 16s rdna primer pairs in the study of rhizosphere and endosphere bacterial microbiomes in metabarcoding studies', *Frontiers in Microbiology*, vol. 7, no. 605, pp. 1-15.

Bell, J, Barnes, DKA & Tuner, J 2002, 'The importance of micro and macro morphological variation in the adaptation of a sublittoral demosponge to current extremes', *Marine Biology*, vol. 140, no. 1, pp. 75-81.

Bell, JJ 2008, 'The functional roles of marine sponges', *Estuarine, Coastal and Shelf Science*, vol. 79, pp. 341-53.

Bennett, S 2004, 'Solexa Ltd', *Pharmacogenomics*, vol. 5, no. 4, pp. 433-8.

Benson, DA, Karsch-Mizrachi, I, Lipman, DJ, Ostell, J & Sayers, EW 2009, 'GenBank', *Nucleic Acids Research*, vol. 37, no. Database issue, pp. D26-31.

Bentley, DR, Balasubramanian, S, Swerdlow, HP, Smith, GP, Milton, J, Brown, CG, Hall, KP, Evers, DJ, Barnes, CL & Bignell, HR 2008, 'Accurate whole human genome sequencing using reversible terminator chemistry', *Nature*, vol. 456, pp. 53-9.

Bergmann, W & Freeney, RJ 1950, 'The isolation of a new thymine pentoside from sponges', *Journal of the American Chemical Society*, vol. 72, p. 2805.

Bergquist, PR 1978, *Sponges*, Hutchinson University Library, London.

Bergquist, PR & Cook, SDC 2002, 'Order Dendroceratida Minchin, 1900', in Hooper JNA & Van Soest RWM (eds), *Systema Porifera: A guide to the classification of sponges*, pp. 1067. Kluwer Academic/Plenum Publishers, New York, Boston, Dordrecht, London, Moscow.

Bertilsson, S, Cavanaugh, CM & Polz, MF 2002, 'Sequencing-independent method to generate oligonucleotide probes targeting a variable region in bacterial 16S rRNA by PCR with detachable primers', *Applied and Environmental Microbiology*, vol. 68, pp. 6077-86.

Bewley, CA & Faulkner, DJ 1998, 'Lithistid sponges: star performers or hosts to the stars', *Angewandte Chemie International Edition in English*, vol. 37, pp. 2163-78.

Blanquer, A & Uriz, MJ 2007, 'Cryptic speciation in marine sponges evidenced by mitochondrial and nuclear genes: A phylogenetic approach', *Molecular Phylogenetics and Evolution*, vol. 45,

pp. 392-7.

Blunt, JW, Copp, BR, Munro, MHG, Northcote, PT & Prinsep, MR 2011, 'Marine natural products', *Natural Product Reports*, vol. 28, pp. 196-268.

Blunt, JW & Munro, MHG 2003, *MarinLit. A database of the literature on marine natural products for use on a macintosh computer prepared and maintained by the Marine Chemistry Group*, Department of Chemistry, University of Canterbury, Canterbury, New Zealand.

Borchiellini, C, Chombard, C, Lafay, B & Boury-Esnault, N 2000, 'Molecular systematics of sponges (Porifera)', *Hydrobiologia*, vol. 420, pp. 15-27.

Borchiellini, C, Chombard, C, Manuel, M, Alivon, E, Vacelet, J & Boury-Esnault, N 2004, 'Molecular phylogeny of Demospongiae: implications for classification and scenarios of character evolution', *Molecular Phylogenetics and Evolution*, vol. 32, pp. 823-37.

Borchiellini, C, Manuel, M, Alivon, E, Boury-Esnault, N, Vacelet, J & Le Parco, Y 2001, 'Sponge paraphyly and the origin of Metazoa', *Journal of Evolutionary Biology*, vol. 14, pp. 171-9.

Boury-Esnault, N, Hajdu, E, Klautau, M, Custodio, M & Borojevic, R 1994, 'The value of cytological criteria in distinguishing sponges at the species level – the example of the Genus Polymastia', *Canadian Journal of Zoology*, vol. 72, pp. 795-804.

Braekman, JC, Daloze, D, Stoller, C & Van Soest, RWM 1992, 'Chemotaxonomy of *Agelas* (Porifera, Demospongiae)', *Biochemical Systematics and Ecology*, vol. 20, pp. 417-31.

Bringmann, G, Lang, G, Muhlbacher, J, Schaumann, K, Steffens, S, Rytik, PG, Hentschel, U, Morschhauser, J & Müller, WEG 2003, 'Sorbicillactone A: A structurally unprecedented bioactive novel-type alkaloid from a sponge-derived fungus', *Progress in Molecular and Subcellular Biology*, vol. 37, pp. 231-53.

Bultel-Ponce, VV, Debitus, C, Berge, JP, Cerceau, C & Guyot, M 1998, 'Metabolites from the sponge-associated bacterium *Micrococcus luteus*', *Journal of Marine Biotechnology*, vol. 6, pp. 233-6.

Burgsdorf, I, Erwin, PM, López-Legentil, S, Cerrano, C, Haber, M, Frenk, S & Steindler, L 2014, 'Biogeography rather than association with cyanobacteria structures symbiotic microbial communities in the marine sponge *Petrosia ficiformis*', *Frontiers in Microbiology*, vol. 10, no. 529, pp. 1-11.

Busse, I & Preisfeld, A 2002, 'Unusually expanded SSU ribosomal DNA of primary osmotrophic euglenids: Molecular evolution and phylogenetic inference', *Journal of Molecular Evolution*, vol. 55, pp. 757-67.

Cárdenas, P, Pérez, T & Boury-Esnault, N 2012, 'Sponge systematics facing new challenges', *Advances in Marine Biology*, vol. 61, pp. 79-209.

Cárdenas, P, Rapp, HT, Schander, C & Tendal, OS 2010, 'Molecular taxonomy and phylogeny of the Geodiidae (Porifera, *Demospongiae*, Astrophorida)-combining phylogenetic and Linnaean classification', *Zoologica Scripta*, vol. 39, pp. 89-106.

Cárdenas, P, Xavier, JR, Reveillaud, J, Schander, C & Rapp, HT 2011, 'Molecular phylogeny of the Astrophorida (Porifera, *Demospongiae*[p]) reveals an unexpected high level of spicule homoplasy', *PLoS ONE*, vol. 6, p. e18318.

Carella, M, Agell, G, Cárdenas, P & Uriz, MJ 2016, 'Phylogenetic Reassessment of Antarctic Tetillidae (Demospongiae, Tetractinellida) Reveals New Genera and Genetic Similarity among Morphologically Distinct Species ', *PLoS ONE*, vol. 11, no. 8, p. e0160718.

Cerrano, C, Calcinai, B, Cucchiari, E, Di Camillo, C, Totti, C & Bavestrello, G 2004, 'The diversity of relationships between Antarctic sponges and diatoms: the case of *Mycale acerata* Kirkpatrick, 1907 (Porifera, Demospongiae)', *Polar Biology* , vol. 27, no. 4, pp. 231-7.

Choi, BK, Wyss, C & Gobel, UB 1996, 'Phylogenetic analysis of pathogen-related oral spirochetes', *Journal of Clinical Microbiology*, vol. 34, pp. 1922-5.

Choi, EJ, Kwon, HC, Ham, J & Yang, HO 2009, '6-Hydroxymethyl-1-phenazine-carboxamide and 1,6-phenazinedimethanol from a marine bacterium, *Brevibacterium* sp. KMD 003, associated with marine purple vase sponge', *Journal of Antibiotics*, vol. 62, pp. 621-4.

Chombard, C & Boury-Esnault, N 1999, 'Good congruence between morphology and molecular phylogeny of Hadromerida or how to bother sponge taxonomists', *Memoirs of the Queensland Museum*, vol. 44, p. 100.

Chombard, C, Boury-Esnault, N, Tillier, A & Vacelet, J 1997, 'Polyphyly of “sclerosponges” (Porifera,Demospongiae) supported by 28S ribosomal sequences', *Biological Bulletin*, vol. 193, pp. 359-67.

Chun, J, Lee, JH, Jung, Y, Kim, M, Kim, S, Kim, BK & Lim, YW 2007, 'EzTaxon: A web-based tool for the identification of prokaryotes based on 16S ribosomal RNA gene sequences', *International Journal of Systematic and Evolutionary Microbiology*, vol. 57, pp. 2259-61.

Claesson, MJ, Wang, Q, O'Sullivan, O, Greene-Diniz, R, Cole, JR, Ross, RP & O'Toole, PW 2010, 'Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions', *Nucleic Acids Research*, vol. 38, no. 22, p. e200.

Clarridge, JE 2004, '3rd Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases', *Clinical Microbiology Reviews*, vol. 17, pp. 840-62.

Cleary, DFR, Becking, LE, De Voogd, NJ, Pires, ACC, Polónia, ARM, Egas, C & Gomes, NCM 2013, 'Habitat- and host-related variation in sponge bacterial symbiont communities in Indonesian waters', *FEMS Microbiology Ecology*, vol. 85, no. 3, pp. 465-82.

Cole, JR, Wang, Q, Cardenas, E, Fish, J, Chai, B, Farris, RJ, Kulam-Syed-Mohideen, AS, McGarrell, DM, Marsh, T & Garrity, GM 2009, 'The Ribosomal Database Project: Improved alignments and new tools for rRNA analysis', *Nucleic Acids Research*, vol. 37, pp. D141-D5.

Cook, SDC & Bergquist, PR 1998, 'Revision of the genus *Psammocinia* (Porifera: Demospongiae: Dictyoceratida), with six new species from New Zealand', *New Zealand Journal of Marine and Freshwater Research*, vol. 32, pp. 399-426.

Cook, SDC & Bergquist, PR 2002, 'Family Irciniidae Gray, 1867', in Hooper JNA & Van Soest RWM (eds), *Systema Porifera: A guide to the classification of sponges*, pp. 1022-7, Kluwer Academic/Plenum Publishers, New York, Boston, Dordrecht, London, Moscow.

Cook, SDC & Bergquist, PR 2002, 'Order Dictyoceratida Minchin, 1900', in Hooper JNA & Van

Soest RWM (eds), *Systema Porifera: A guide to the classification of sponges*, p. 1021, Kluwer Academic/Plenum Publishers, New York, Boston, Dordrecht, London, Moscow.

Cuvelier, ML, Blake, E, Mulheron, R, McCarthy, PJ, Blackwelder, P, Vega Thurber, RL & Lopez, JV 2014, 'Two distinct microbial communities revealed in the sponge Cinachyrella', *Frontiers in Microbiology*, vol. 5, no. 581, pp. 1-14.

De Voogd, NJ, Cleary, DFR, Polónia, ARM & Gomes, NCM 2015, 'Bacterial community composition and predicted functional ecology of sponges, sediment and seawater from the thousand islands reef complex, West Java, Indonesia', *FEMS Microbiology Ecology*, vol. 91, no. 4.

DeSantis, TZ, Hugenholtz, P, Larsen, N, Rojas, M, Brodie, EL, Keller, K, Huber, T, Dalevi, D, Hu, P & Andersen, GL 2006, 'Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB', *Applied and Environmental Microbiology*, vol. 72, pp. 5069-72.

Dohm, JC, Lottaz, C, Borodina, T & Himmelbauer, H 2008, 'Substantial biases in ultra-short read data sets from high-throughput DNA sequencing', *Nucleic Acids Research*, vol. 36, p. e105.

Dohrmann, M, Haen, KM, Lavrov, DV & Wörheide, G 2012, 'Molecular phylogeny of glass sponges (Porifera, Hexactinellida): increased taxon sampling and inclusion of the mitochondrial protein-coding gene, cytochrome oxidase subunit I', *Hydrobiologia*, vol. 687, pp. 11-20.

Doudna, JA & Rath, VL 2002, 'Structure and function of the eukaryotic ribosome: the next frontier', *Cell*, vol. 109, pp. 153–6.

Dressman, D, Yan, H, Traverso, G, Kinzler, KW & Vogelstein, B 2003, 'Transforming single DNA molecules into fluorescent magnetic particles for detection and enumeration of genetic variations', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 100, pp. 8817-22.

Drmanac, R, Sparks, AB, Callow, MJ, Halpern, AL, Burns, NL, Kermani, BG, Carnevali, P, Nazarenko, I, Nilsen, GB, Yeung, G, Dahl, F, Fernandez, A, Staker, B, Pant, KP, Baccash, J, Borcherding, AP, Brownley, A, Cedeno, R, Chen, L, Chernikoff, D, Cheung, A, Chirita, R, Curson, B, Ebert, JC, Hacker, CR, Hartlage, R, Hauser, B, Huang, S, Jiang, Y, Karpinchyk, V, Koenig, M, Kong, C, Landers, T, Le, C, Liu, J, McBride, CE, Morenzoni, M, Morey, RE, Mutch, K, Perazich, H, Perry, K, Peters, BA, Peterson, J, Pethiyagoda, CL, Pothuraju, K, Richter, C, Rosenbaum, AM, Roy, S, Shafto, J, Sharanhovich, U, Shannon, KW, Sheppy, CG, Sun, M, Thakuria, JV, Tran, A, Vu, D, Zaranek, AW, Wu, X, Drmanac, S, Oliphant, AR, Banyai, WC, Martin, B, Ballinger, DG, Church, GM & Reid, CA 2010, 'Human genome sequencing using unchained base reads on self-assembling DNA nanoarrays', *Science*, vol. 327, no. 5961, pp. 78-81.

Duran, S, Pascual, M & Turon, X 2004, 'Low levels of genetic variation in mtDNA sequences over the western Mediterranean and Atlantic range of the sponge *Crambe crambe* (Poecilosclerida)', *Marine Biology*, vol. 144, pp. 31-5.

Duran, S & Rützler, K 2006, 'Ecological speciation in a Caribbean marine sponge', *Molecular Phylogenetics and Evolution*, vol. 40, no. 1, pp. 292-7.

Easson, CG & Thacker, RW 2014, 'Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges', *Frontiers in Microbiology*, vol. 5, no. 532, pp. 1-11.

Erpenbeck, D, Breeuwer, JAJ, van der Velde, H & Van Soest, RWM 2002, 'Unravelling host and symbiont phylogenies of halichondrid sponges (Demospongiae, Porifera) using a mitochondrial marker', *Marine Biology*, vol. 141, pp. 377-86.

Erpenbeck, D, Hall, K, Alvarez, B, Büttner, G, Sacher, K, Schätzle, S, Schuster, A, Vargas, S, Hooper, JNA & Wörheide, G 2012, 'The phylogeny of halichondrid demosponges: past and present re-visited with DNA-barcoding data', *Organisms Diversity and Evolution*, vol. 12, pp. 57-70.

Erpenbeck, D, Hooper, JNA & Wörheide, G 2006, *Molecular Ecology Notes*, vol. 6, pp. 550-553.

Erpenbeck, D, Knowlton, AL, Talbot, SL, Highsmith, RC & Van Soest, RWM 2004, 'A molecular comparison of Alaskan and North East Atlantic *Halichondria panicea* (Pallas 1766) (Porifera: Demospongiae) populations', *Bollettino die Musei e degli Istituti Biologici dell'Università di Genova*, vol. 68, pp. 319-25.

Erpenbeck, D, McCormack, GP, Breeuwer, JAJ & Van Soest, RWM 2004, 'Order level differences in the demosponge LSU structure - New insights in an old taxon', *Molecular Phylogenetics and Evolution*, vol. 32, no. 1, pp. 388-95.

Erpenbeck, D, Sutcliffe, P, Cook, S, Dietzel, A, Maldonado, M, Soest, RWMv, Hooper, JNA & Wörheide, G 2012, 'Horny sponges and their affairs: on the phylogenetic relationships of keratose sponges', *Molecular Phylogenetics and Evolution*, vol. 63, pp. 809-16.

Erpenbeck, D, Weier, T, De Voogd, NJ, Wörheide, G, Sutcliffe, P, Todd, JA & Michel, E 2011, 'Insights into the evolution of freshwater sponges (Porifera: Demospongiae: Spongillina): Barcoding and phylogenetic data from Lake Tanganyika endemics indicate multiple invasions and unsettle existing taxonomy', *Molecular Phylogenetics and Evolution*, vol. 61, pp. 231-6.

Erwin, PM, Olson, JB & Thacker, RW 2011, 'Phylogenetic diversity, host-specificity and community profiling of sponge-associated bacteria in the northern Gulf of Mexico', *PLoS ONE*, vol. 6, p. e26806.

Erwin, PM & Thacker, RW 2007, 'Phylogenetic analyses of marine sponges within the order Verongida: a comparison of morphological and molecular data', *Invertebrate Biology*, vol. 126, pp. 220–34.

Fan, L, Reynolds, D, Liu, M, Stark, M, Kjelleberg, S, Webster, NS & Thomas, T 2012, 'Functional equivalence and evolutionary convergence in complex communities of microbial sponge symbionts', *Proceedings of the National Academy of Sciences of the United States of America*, pp. E1878-E87.

Faulkner, DJ 2000, 'Highlights of marine natural products chemistry (1972–1999)', *Natual Product Reports*, vol. 17, pp. 1-6.

Fedurco, M, Romieu, A, Williams, S, Lawrence, I & Turcatti, G 2006, 'BTA, a novel reagent for DNA attachment on glass and efficient generation of solid-phase amplified DNA colonies', *Nucleic Acids Research.*, vol. 34, p. e22.

Fieseler, L, Horn, M, Wagner, M & Hentschel, U 2004, 'Discovery of the novel candidate phylum "Poribacteria" in marine sponges', *Applied and Environmental Microbiology*, vol. 70, no. 6, pp. 3724-32.

Fisch, KM, Gurgui, C, Heycke, N, van der Sar, SA, Anderson, SA, Webb, VL, Taudien, S,

Platzer, M, Rubio, BK, Robinson, SJ, Crews, P & Piel, J 2009, 'Polyketide assembly lines of uncultivated sponge symbionts from structure-based gene targeting', *Nature Chemical Biology*, vol. 5, no. 7, pp. 494-501.

Folmer, O, Black, M, Hoeh, W, Lutz, R & Vrijenhoek, R 1994, 'DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates', *Molecular Marine Biology and Biotechnology*, vol. 3, no. 5, pp. 294-9.

Fontaneto, D, Flot, J-F & Tang, CQ 2015, 'Guidelines for DNA taxonomy, with a focus on the meiofauna', *Marine Biodivercity*, vol. 45, no. 3, pp. 433-51.

Forgetta, V, Leveque, G, Dias, J, Grove, D, Lyons Jr., R, Genik, S, Wright, C, Singh, S, Peterson, N, Zianni, M, Kieleczawa, J, Steen, R, Perera, A, Bintzler, D, Adams, S, Hintz, W, Jacobi, V, Bernier, L, Levesque, R & Dewar, K 2013, 'Sequencing of the dutch elm disease fungus genome using the Roche/454 GS-FLX titanium system in a comparison of multiple genomics core facilities', *Journal of Biomolecular Techniques*, vol. 24, no. 1, pp. 39-49.

Freckelton, ML, Luter, HM, Andreakis, N, Webster, NS & Motti, CA 2012, 'Qualitative variation in colour morphotypes of *Ianthella basta* (Porifera: Verongida)', *Hydrobiologia*, vol. 687, pp. 191-203.

Freeman, CJ & Thacker, RW 2011, 'Complex interactions between marine sponges and their symbiotic microbial communities', *Limnology and Oceanography,* vol. 56, pp. 1577-86.

Friedrich, AB, Merkert, H, Fendert, T, Hacker, J, Proksch, P & Hentschel, U 1999, 'Microbial diversity in the marine sponge *Aplysina cavernicola* (formerly Verongia cavernicola) analyzed by fluorescence in situ hybridization (FISH)', *Marine Biology*, vol. 134, no. 3, pp. 461-70.

Fromont, J & Bergquist, PR 1990, 'Structural characters and their use in sponge taxonomy: When is a sigma not a sigma? ', in Rützler K (ed.), pp. 273-8, *New perspectives in Sponge Biology*, Smithonian Institution ress., Washington DC.

Galtier, N, Nabholz, B, Glémin, S & Hurst, GDD 2009, 'Mitochondrial DNA as a marker of molecular diversity: a reappraisal', *Molecular Ecology*, vol. 18, pp. 4541-50.

Gao, ZM, Wang, Y, Lee, OO, Tian, RM, Wong, YH, Bougouffa, S, Batang, Z, Al-Suwailem, A, Lafi, FF, BajicVladimir, B & Qian, PY 2014, 'Pyrosequencing reveals the microbial communities in the red sea sponge *Carteriospongia foliascens* and their impressive shifts in abnormal tissues', *Microbial Ecology*, vol. 68, pp. 621-32.

Garcia-Santos, G, Pinheiro, U, Hajdu, E & Van Soest, RWM 2016, 'New Genus and species of Heteroxyidae from Brazil (Axinellida: Demospongiae: Porifera), with a revised identification key for the family', *Zootaxa*, vol. 4158 no. 1, pp. 105-16.

Gatti, S 2002, 'The role of sponges in high-antarctic carbon and silicon cycling - a modelling approach ', *Bericht zur Polar-und Meeresforschung*, vol. 434, pp. 1-124.

Gauthier, EA, Du Pasquier, L & Degnan, BM 2010, 'The genome of the sponge Amphimedon queenslandica provides new perspectives into the origin of Toll-like and interleukin 1 receptor pathways', *Evolution & Development*, vol. 12, pp. 519-33

Gazave, E, Carteron, S, Chenuil, A, Richelle-Maurer, E, Boury-Esnault, N & Borchiellini, C 2010, 'Polyphyly of the genus *Axinella* and of the family Axinellidae (Porifera: *Demospongiae*[p])', *Molelular Phylogenetics and Evolution*, vol. 57, pp. 35-47.

Gazave, E, Lapébie, P, Renard, E, Vacelet, J, Rocher, C, Ereskovsky, AV, Lavrov, DV & Borchiellini, C 2010, 'Molecular phylogeny restores the supra-generic subdivision of Homoscleromorpha sponges (Porifera, Homoscleromorpha)', *PLoS ONE*, vol. 5, p. e14290.

GenomeWeb 2015, *Roche shutting down 454 sequencing business*, GenomeWeb viewed 15 Oct. 2015.

Gerbi, SA 1985, *Evolution of ribosomal DNA*, New York/London.

Ghyselinck, J, Pfeiffer, S, Heylen, K, Sessitsch, A & De Vos, P 2013, 'The effect of primer choice and short read sequences on the outcome of 16s rrna gene based diversity studies', *PLoS ONE*, vol. 8, no. 8, p. e71360.

Gilbert, JA, Jansson, JK & Knight, R 2014, 'The Earth Microbiome project: successes and aspirations', *BMC Biology*, vol. 12, no. 69, pp. 1-4.

Giles, EC, Kamke, J, Moitinho-Silva, L, Taylor, MW, Hentschel, U, Ravasi, T & Schmitt, S 2013a, 'Bacterial community profiles in low microbial abundance sponges', *FEMS Microbiology Ecology.*, vol. 83, pp. 232-41.

Gillespie, JJ, McKenna, CH, Yoder, MJ, Gutell, RR, Johnston, JS, Kathirithamby, J & Cognato, AI 2005, 'Assessing the odd secondary structural properties of nuclear small subunit ribosomal RNA sequences (18S) of the twisted-wing parasites (Insecta: Strepsiptera)', *Insect Molecular Biology*, vol. 14, no. 6, pp. 625-43.

Giribet, G & Wheeler, WC 2001, 'Some unusual small-subunit ribosomal RNA sequences of metazoans', *American Museum Novitates*, vol. 3337, p. 14.

Göcke, C, Hestetun, JT, Uhlir, C, Freiwald, A, Beuck, L & Janussen, D 2016, '*Cladorhiza corallophila* sp. nov., a new carnivorous sponge (Cladorhizidae, Demospongiae) living in close association with Lophelia pertusa and Madrepora oculata (Scleractinia)', *Zootaxa*, vol. 4168 no. 3, pp. 512-24.

Goodwin, S, McPherson, JD & McCombie, WR 2016, 'Coming of age: ten years of nextgeneration sequencing technologies', *Nature*, vol. 17, pp. 333-51.

Jumpstart Consortium Human Microbiome Project Data Generation Working Group 2012, 'Evaluation of 16S rDNA-Based Community Profiling for Human Microbiome Research', *PLoS ONE*, vol. 7, no. 6, p. e39315.

Gutt, J 2007, 'Antarctic macro-zoobenthic communities: a review and an ecological classification', *Antarctic Science*, vol. 19, no. 2, pp. 165-82.

Hamady, M & Knight, R 2009, 'Microbial community profiling for human microbiome projects: Tools, techniques, and challenges', *Genome Research*, vol. 19, no. 7, pp. 1141-52.

Hamady, M, Lozupone, C & Knight, R 2010, 'Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data', *ISME Journal*, vol. 4, no. 1, pp. 17-27.

Hameş-Kocabaş, EE & Uzel, A 2012, 'Isolation strategies of marine-derived actinomycetes from sponge and sediment samples', *Journal of Microbiological Methods*, vol. 88, no. 3, pp. 342-7.

Hardoim, CC, Esteves, AI, Pires, FR, Gonçalves, JM, Cox, CJ, Xavier, JR & Costa, R 2012, 'Phylogenetically and spatially close marine sponges harbour divergent bacterial communities', *PLoS ONE*, vol. 7, p. e53029.

Harismendy, O, Ng, PC, Strausberg, RL, Wang, X, Stockwell, TB, Beeson, KY, Schork, NJ, Murray, SS, Topol, EJ, Levy, S & Frazer, KA 2009, 'Evaluation of next generation sequencing platforms for population targeted sequencing studies', *Genome Biology*, vol. 10, no. 3, p. R32.

Harrington, C, Reen, F, Mooij, M, Stewart, F, Chabot, J-B, Guerra, A, Glöckner, F, Nielsen, K, Gram, L & Dobson, A 2014, 'Characterisation of non-autoinducing tropodithietic acid (TDA) production from marine sponge *Pseudovibrio* species', *Marine Drugs*, vol. 12, pp. 5960-78.

Harris, TD, Buzby, PR, Babcock, H, Beer, E, Bowers, J, Braslavsky, I, Causey, M, Colonell, J, Dimeo, J, Efcavitch, JW, Giladi, E, Gill, J, Healy, J, Jarosz, M, Lapen, D, Moulton, K, Quake, SR, Steinmann, K, Thayer, E, Tyurina, A, Ward, R, Weiss, H & Xie, Z 2008, 'Single-molecule DNA sequencing of a viral genome', *Science*, vol. 320, no. 5872 pp. 106-9.

Hartman, WD 1980, 'Systematics of the Porifera', in Hartman WD, Wendt JW & Wiedenmayer F (eds), pp. 24-52, *Living and fossil sponges*, Sedimenta viii rosenstiel school of marine and atmospheric science, Miami.

He, L, Liu, F, Karuppiah, V, Ren, Y & Li, Z 2014, 'Comparisons of the fungal and protistan communities among different marine sponge holobionts by pyrosequencing', *Microbial Ecology*, vol. 67, no. 4, pp. 951-61.

He, ZL, Van Nostrand, JD & Zhou, JZ 2012, 'Applications of functional gene microarrays for profiling microbial communities', *Current Opinion in Biotechnology*, vol. 23, no. 3, pp. 460-6.

Hebert, PDN, Cywinska, A, Ball, SL & de Waard, JR 2003, 'Biological identifications through DNA barcodes', *Proceedings of the royal society of London. Series B: Biological sciences*, vol. 270, pp. 313–32.

Heim, I, Nickel, M & Brümmer, F 2007a, 'Molecular markers for species discrimination in poriferans: a case study on species of the genus *Aplysina*', *Porifera Research: Biodiversity, Innovation and Sustainability*, vol. 1, pp. 361-71.

Heim, I, Nickel, M & Brümmer, F 2007b, 'Phylogeny of the genus Tethya (Tethyidae: Hadromerida: Porifera): molecular and morphological aspects', *Journal of the Marine Biological Association of the United Kingdom*, vol. 87, pp. 1615-27.

Henríquez, M, Vergara, K, Norambuena, J, Beiza, A, Maza, F, Ubilla, P, Araya, I, Chávez, R, San-Martín, A, Darias, J, Darias, MJ & Vaca, I 2014, 'Diversity of cultivable fungi associated with Antarctic marine sponges and screening for their antimicrobial, antitumoral and antioxidant potential', *World Journal of Microbiology and Biotechnology*, vol. 30, no. 1, pp. 65-76.

Hentschel, U, Fieseler, L, Wehrl, M, Gernert, C, Steinert, M & Hacker, J 2003, 'Microbial diversity of marine sponges', in Müller WEG (ed.), pp. 59-83, *Sponges (Porifera)*, Springer, Heidelberg, Germany.

Hentschel, U, Hopke, J, Horn, M, Friedrich, AB, Wagner, M, Hacker, J & Moore, BS 2002, 'Molecular evidence for a uniform microbial community in sponges from different oceans', *Applied and Environmental Microbiology*, vol. 68, no. 9, pp. 4431-40.

Hentschel, U, Piel, J, Degnan, SM & Taylor, MW 2012, 'Genomic insights into the marine

sponge microbiome', *Nature Reviews Microbiology*, vol. 10, pp. 641-54.

Hentschel, U, Usher, KM & Taylor, MW 2006, 'Marine sponges as microbial fermenters', *FEMS Microbiology Ecology*, vol. 55, pp. 167-77.

Hillis, DM & Davis, SK 1988, 'Ribosomal DNA: Intraspecific Polymorphism, Concerted Evolution, and Phylogeny Reconstruction', *Systematic Zoology*, vol. 37, no. 1, pp. 63-6.

Hillis, DM & Dixon, MT 1991, 'Ribosomal RNA: molecular evolution and phylogenetic inference', *Quarterly Review of Biology*, vol. 66, pp. 411-53.

Hollibaugh, J, Lovejoy, C & Murray, A 2007, 'Microbiology in polar oceans', *Oceanography*, vol. 20, no. 2, pp. 138-43.

Hooper, JNA, Hall, KA, Ekins, M, Erpenbeck, D, Wörheide, G & Jolley-Rogers, G 2013, 'Managing and sharing the escalating number of sponge "unknowns": the SpongeMaps project', *Integrative and Comparative Biology*, vol. 33, no. 3, pp. 473-81.

Hooper, JNA & Van Soest, RWM 2002, *Systema Porifera. A guide to the classification of sponges*, 2 vols., Kluwer Academic/Plenum Publishers, New York.

Huang, D, Meier, R, Todd, PA & Chou, LM 2008, 'Slow mitochondrial COI sequence evolution at the base of the metazoan tree and its implications for DNA barcoding', *Journal of Molecular Evolution*, vol. 66, pp. 167-74.

Human Microbiome Project Consortium 2012, 'Structure, function and diversity of the healthy human microbiome', *Nature*, vol. 486, pp. 207-14

Huse, SM, Dethlefsen, L, Huber, JA, Mark, WD, Relman, DA & Sogin, ML 2008, 'Exploring microbial diversity and taxonomy using SSU rRNA hypervariable tag sequencing', *PLOS Genetics*, vol. 4, no. 11, p. e1000255.

Huse, SM, Welch, DM, Morrison, HG & Sogin, ML 2010, 'Ironing out the wrinkles in the rare biosphere through improved OTU clustering', *Environmental Microbiology*, vol. 12, pp. 1889-98.

Huws, SA, Edwards, JE, Kim, EJ & Scollan, ND 2007, 'Specificity and sensitivity of eubacterial primers utilised for molecular profiling of bacteria within complex microbial ecosystems', *Journal of Microbiological Methods*, vol. 70, pp. 565-9.

Huyck, TK, Gradishar, W, Manuguid, F & Kirkpatrick, P 2011, 'Eribulin mesylate', *Nature Reviews Drug Discovery*, no. 10, pp. 173-4.

Indraningrat, AAG, Hauke, S & Detmer, S 2016, 'Bioprospecting sponge-associated microbes for antimicrobial compounds', *Marine Drugs*, vol. 14, no. 5, pp. 87.

Jackson, S, Kennedy, J, Morrissey, J, O'Gara, F & Dobson, AW 2012, 'Pyrosequencing Reveals Diverse and Distinct Sponge-Specific Microbial Communities in Sponges from a Single Geographical Location in Irish Waters', *Microbial Ecology*, vol. 64, no. 1, pp. 105-16.

Jackson, SA, Flemer, B, McCann, A, Kennedy, J, Morrissey, JP, O'Gara, F & Dobson, AD 2013, 'Archaea appear to dominate the microbiome of *Inflatella pellicula* deep sea sponges', *PLoS ONE*, vol. 8, no. 12, p. e84438.

Jayatilake, GS, Thornton, MP, Leonard, AC, Grimwade, JE & Baker, BJ 1996, 'Metabolites from

an Antarctic sponge-associated bacterium, *Pseudomonas aeruginosa*', *Journal of Natural Products*, vol. 59, pp. 293-6.

Jeong, I-H, Kim, K-H & Park, J-S 2013, 'Analysis of bacterial diversity in sponges collected off Chujado, an Island in Korea, using barcoded 454 pyrosequencing: analysis of a distinctive sponge group containing Chloroflexi', *Journal of Microbiology*, vol. 51 pp. 570-7.

Jin, L, Liu, F, Sun, W, Zhang, F, Karuppiah, V & Li, Z 2014, '*Pezizomycotina* dominates the fungal communities of South China Sea sponges *Theonella swinhoei* and *Xestospongia testudinaria*', *FEMS Microbiology Ecology*, vol. 90, no. 3, pp. 935-45.

Jonasson, J, Olofsson, M & Monstein, HJ 2002, 'Classification, identification and subtyping of bacteria based on pyrosequencing and signature matching of 16s rDNA fragments', *Acta Pathologica, Microbiologica et Immunologica Scandinavica*, vol. 115, pp. 668-77.

Ju, J, Kim, DH, Bi, L, Meng, Q, Bai, X, Li, Z, Li, X, Marma, MS, Shi, S, Wu, J, Edwards, JR, Romu, A & Turro, NJ 2006, 'Four-color DNA sequencing by synthesis using cleavable fluorescent nucleotide reversible terminators', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 52, pp. 19635-40.

Kamke, J, Taylor, MW & Schmitt, S 2010, 'Activity profiles for marine sponge-associated bacteria obtained by 16S rRNA vs 16S rRNA gene comparisons', *ISME Journal*, vol. 4, pp. 498-508.

Karlep, L, Reintamm, T & Kelve, M 2013, 'Intragenomic profiling using multicopy genes: The rDNA internal transcribed spacer sequences of the freshwater sponge *Ephydatia fluviatilis*', *PLoS ONE*, vol. 8, no. 6, p. e66601.

Kennedy, J, Flemer, B, Jackson, SA, Morrissey, JP, O'Gara, F & Dobson, AD 2014, 'Evidence of a putative deep sea specific microbiome in marine sponges', *PLoS ONE*, vol. 9, no. 3, p. e91092.

Khan, ST, Takagi, M & Shin-ya, K 2012, 'Actinobacteria associated with the marine sponges *Cinachyra* sp., *Petrosia* sp., and *Ulosa* sp. and their culturability', *Microbes and Environments*, vol. 27, no. 1, pp. 99-104.

Kim, M, Morrison, M & Yu, Z 2011, 'Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes', *Journal of microbiological methods*, vol. 84, pp. 81-7.

Kim, TK, Garson, MJ & Fuerst, JA 2005, 'Marine actinomycetes related to the "*Salinospora*" group from the Great Barrier Reef sponge *Pseudoceratina clavata*', *Environmental Microbiology*, vol. 7, no. 4, pp. 509-18.

Kircher, M & Kelso, J 2010, 'High-throughput DNA sequencing-concepts and limitations', *Bioessays*, vol. 32, no. 6, pp. 524-36.

Klautau, M, Azevedo, F, Cóndor-Luján, B, Rapp, HT, Collins, A & Russo, CAdM 2013, 'A molecular phylogeny for the order Clathrinida rekindles and refines haeckel's taxonomic proposal for calcareous sponges', *Integrative and Comparative Biology*, vol. 53, no. 3, pp. 447-61.

Klautau, M, Solé-Cava, AM & Borojevic, R 1994, 'Biochemical systematics of silbing sympatric species of *Clathrina* (Porifera: Calcarea)', *Biochemical Systematics and Ecology*, vol. 22, no. 4,

pp. 367-75.

Klindworth, A, Pruesse, E, Schweer, T, Peplies, J, Quast, C, Horn, M & Glöckner, FO 2012, 'Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies', *Nucleic Acids Research*, pp. 1-11.

Knowlton, N 1993, 'Sibling species in the sea', *Annual review of ecology system*, vol. 24, pp. 189-214.

Kress, WJ, Wurdack, KJ, Zimmer, EA, Weigt, LA & Janzen, DH 2005, 'Use of DNA barcodes to identify flowering plants', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 23, pp. 8369-74.

Kumar, PS, Brooker, MR, Dowd, SE & Camerlengo, T 2011, 'Target region selection is a critical determinant of community fingerprints generated by 16S Pyrosequencing', *PLoS ONE*, vol. 6, no. 6, p. e20956.

Kunz, AL, Labes, A, Wiese, J, Bruhn, T, Bringmann, G & Imhoff, JF 2014, 'Nature's lab for derivatization: New and revised structures of a variety of streptophenazines produced by a sponge-derived *Streptomyces* strain', *Marine Drugs*, vol. 12, pp. 1699-714.

Kwon, HC, Kauffman, CA, Jensen, PR & Fenical, W 2006, 'Marinomycins A-D, antitumor-antibiotics of a new structure class from a marine actinomycete of the recently discovered genus "*marinispora*"', *Journal of the American Chemical Society*, vol. 128, no. 5, pp. 1622-32.

Lafay, B, Boury-Esnault, N, Vacelet, J & Christen, R 1992, 'An analysis of partial 28S ribosomal RNA sequences suggests early radiation of sponges', *BioSystems*, vol. 28, pp. 139–51.

Lane, D, Pace, B, Olsen, GJ, Stahl, DA, Sogin, ML & Pace, NR 1985, 'Rapid determination of 16S ribosomal RNA sequences for phylogenetic analyses', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 82, pp. 6955-9.

Lang, S, Beli, W, Tokuda, H, Wicke, C & Lurtz, V 2004, 'Improved production of bioactive glycosylmannosyl-glycerolipid by sponge-associated *Microbacterium* species', *Marine Biotechnology*, vol. 6, pp. 152-6.

Lavrov, DV, Forget, L, Kelly, M & Lang, BF 2005, 'Mitochondrial genomes of two demosponges provide insights into an early stage of animal evolution', *Molecular Biology and Evolution*, vol. 22, no. 5, pp. 1231-9.

Lavrov, DV, Wang, X & Kelly, M 2008, 'Reconstructing ordinal relationships in the Demospongiae using mitochondrial genomic data', *Molecular Phylogenetics and Evolution*, vol. 49, pp. 111-24.

Leal, MC, Puga, J, Serôdio, J, Gomes, NCM & 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*, vol. 7, p. e30580.

Lee, OO, Wang, Y, Yang, J, Lafi, FF, Al-Suwailem, A & Qian, PY 2011, 'Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea', *ISME Journal*, vol. 5, no. 4, pp. 650-64.

Lee, Y, Li, H, Hong, J, Cho, H, Bae, K, Kim, M, Kim, D-K & Jung, J 2010, 'Bioactive metabolites from the sponge-derived fungus Aspergillus versicolor', *Archives of Pharmacal Research*, vol.

33, pp. 231-5.

Lévi, C 1973, 'Systématique de la classe des Demospongiaria (Démosponges)', in Grassé P (ed.), vol. 3, pp. 1-716, *Traité de Zoologie. Anatomie, Systématique, Biologie.*

Li, CQ, Liu, WC, Zhu, P, Yang, JL & Cheng, KD 2011, 'Phylogenetic diversity of bacteria associated with the marine sponge *Gelliodes carnosa* collected from the Hainan Island coastal waters of the South China Sea', *Microbial Ecology*, vol. 62, pp. 800-12.

Li, D, Xu, Y, Shao, C-L, Yang, R-Y, Zheng, C-J, Chen, Y-Y, Fu, X-M, Qian, P-Y, She, Z-G & De Voogd, NJ 2012, 'Antibacterial bisabolane-type sesquiterpenoids from the sponge-derived fungus *Aspergillus* sp. ', *Marine Drugs*, vol. 10, pp. 234-41.

Liu, WT, Marsh, TL, Cheng, H & Forney, LJ 1997, 'Characterization of microbial diversity by determining terminal restriction fragment length polymorphisms of genes encoding 16S rRNA', *Applied and Environmental Microbiology*, vol. 63, no. 11, pp. 4516-22.

Liu, Z, DeSantis, TZ, Andersen, GL & Knight, R 2008, 'Accurate taxonomy assignments from 16S rRNA sequences produced by highly parallel pyrosequencers', *Nucleic Acids Research*, vol. 36, p. e120.

Liu, Z, Lozupone, C, Hamady, M, Bushman, FD & Knight, R 2007, 'Short pyrosequencing reads suffice for accurate microbial community analysis', *Nucleic Acids Research*, vol. 35, p. e120.

Loh, T-L, López-Legentil, S, Song, B & Pawlik, JR 2012, 'Phenotypic variability in the Caribbean Orange Icing sponge *Mycale laevis* (Demospongiae: Poecilosclerida)', *Hydrobiologia*, vol. 687, pp. 205-17.

Loman, NJ, Misra, RV, Dallman, TJ, Constantinidou, C, Gharbia, SE, Wain, J & Pallen, MJ 2012, 'Performance comparison of benchtop high-throughput sequencing platforms', *Nature Biotechnology*, vol. 30, no. 5, pp. 434-9.

Lopez, JV, Peterson, CL, Willoughby, R, Wright, AE, Enright, E, Zoladz, S, Reed, JK & Pomponi, SA 2002, 'Characterization of genetic markers for in vitro cell line identification of the marine sponge *Axinella corrugata*', *The American Genetic Association*, vol. 93, pp. 27-36.

Lu, JJ, Perng, CL, Lee, SY & Wan, CC 2000, 'Use of PCR with universal primers and restriction endonuclease digestions for detection and identification of common bacterial pathogens in cerebrospinal fluid', *Journal of Clinical Microbiology*, vol. 38, pp. 2076-80.

Ludwig, W, Strunk, O, Westram, R, Richter, L, Meier, H, Yadhukumar, Buchner, A, Lai, T, Steppi, S, Jobb, G, Förster, W, Brettske, I, Gerber, S, Ginhart, AW, Gross, O, Grumann, S, Hermann, S, Jost, R, König, A, Liss, T, Lüssmann, R, May, M, Nonhoff, B, Reichel, B, Strehlow, R, Stamatakis, A, Stuckmann, N, Vilbig, A, Lenke, M, Ludwig, T, Bode, A & Schleifer, KH 2004, 'ARB: a software environment for sequence data', *Nucleic Acids Research*, vol. 32, no. 4, pp. 1363-71.

Luter, HM, Widder, S, Botté, ES, Wahab, MA, Whalan, S, Moitinho-Silva, L, Thomas, T & Webster, NS 2015, 'Biogeographic variation in the microbiome of the ecologically important sponge, *Carteriospongia foliascens*', *PeerJ*, vol. 3, p. e1435.

Maldonado, LA, Fenical, W, Jensen, PR, Kauffman, CA, Mincer, TJ, Ward, AC, Bull, AT & Goodfellow, M 2005, '*Salinispora arenicola* gen. nov., sp. nov. and *Salinispora tropica* sp. nov., obligate marine actinomycetes belonging to the family Micromonosporaceae', *International*

Journal of Systematic and Evolutionary Microbiology, vol. 55, no. 5, pp. 1759-66.

Maldonado, M, Carmona, MC, Uriz, MJ & Cruzado, A 1999, 'Decline in Mesozoic reef-building sponges explained by silicon limitation', *Nature*, vol. 401, pp. 785–8.

Mangano, S, Caruso, C, Michaud, L, Giudice, AL & Bruni, V 2011, 'Incidence of plasmid and antibiotic resistance in psychrotrophic bacteria isolated from Antarctic sponges', *Accademia Peloritana dei Pericolanti. Classe di Scienze, Fisiche, Matematiche e Naturali*, vol. 89, no. 1, p. C1A8901003.

Manuel, M, Borchiellini, C, Alivon, E, Parco, YL, Vacelet, J & Esnault, NB- 2003, 'Phylogeny and evolution of calcareous sponges: monophyly of Calcinea and Calcaronea, high level of morphological homoplasy, and the primitive nature of axial symmetry', *Systematic Biology*, vol. 52, no. 3, pp. 311-33.

Mardis, ER 2013, 'Next-generation sequencing platforms', *Annual Review of Analytical Chemistry (Palo Alto Calif.)*, vol. 6, pp. 287-303.

Margulies, M, Egholm, M, Altman, WE, Attiya, S, Bader, JS, Bemben, LA, Berka, J, Braverman, MS, Chen, Y-J, Chen, Z, Dewell, SB, Du, L, Fierro, JM, Gomes, XV, Godwin, BC, He, W, Helgesen, S, Ho, CH, Irzyk, GP, Jando, SC, Alenquer, MLI, Jarvie, TP, Jirage, KB, Kim, J-B, Knight, JR, Lanza, JR, Leamon, JH, Lefkowitz, SM, Lei, M, Li, J, Lohman, KL, Lu, H, Makhijani, VB, McDade, KE, McKenna, MP, Myers, EW, Nickerson, E, Nobile, JR, Plant, R, Puc, BP, Ronan, MT, Roth, GT, Sarkis, GJ, Simons, JF, Simpson, JW, Srinivasan, M, Tartaro, KR, Tomasz, A, Vogt, KA, Volkmer, GA, Wang, SH, Wang, Y, Weiner, MP, Yu, P, Begley, RF & Rothberg, JM 2005, 'Genome sequencing in microfabricated high-density picolitre reactors', *Nature*, vol. 437, no. 7057, pp. 376-80.

Maynard, C, Berthiaume, F, Lemarchand, K, Harel, J, Payment, P, Bayardelle, P, Masson, L & Brousseau, R 2005, 'Waterborne pathogen detection by use of oligonucleotide-based microarrays', *Applied and Environmental Microbiology*, vol. 71, pp. 8548-57.

McClintock, JB, Amsler, CD, Baker, BJ & Van Soest, RW 2005, 'Ecology of Antarctic marine sponges: an overview', *Integrative and Comparative Biology*, vol. 45, no. 2, pp. 359-68.

Mehl-Janussen, D 1999, 'Die frühe Evolution der Porifera, Münchner Geowissenschaftliche Abhandlungen Reihe A ', *Geologie und Paläontologie*, vol. 37, pp. 1-72.

Meroz-Fine, E, Shefer, S & Ilan, M 2005, 'Changes in morphology and physiology of the East Mediterranean sponge in different habitats', *Marine Biology*, vol. 147, no. 1, pp. 243-50.

Mikkelsen, P & Cracraft, J 2001, 'Marine biodiversity and the need for systematic inventory', *Bulletin of Marine Science*, vol. 69, pp. 525–34.

Minoche, AE, Dohm, JC & Himmelbauer, H 2011, 'Evaluation of genomic high-throughput sequencing data generated on Illumina HiSeq and genome analyzer systems', *Genome Biology*, vol. 12, p. R112.

Moitinho-Silva, L, Bayer, K, Cannistraci, CV, Giles, EC, Ryu, T, Seridi, L, Ravasi, T & Hentschel, U 2014, 'Specificity and transcriptional activity of microbiota associated with low and high microbial abundance sponges from the Red Sea', *Molecular Ecology*, vol. 23, pp. 1384-63.

Montalvo, NF, Davis, J, Vicente, J, Pittiglio, R, Ravel, J & Hill, RT 2014, 'Integration of Culture-Based and Molecular Analysis of a Complex Sponge-Associated Bacterial Community', *PLoS*

ONE, vol. 9, no. 3, p. e90517.

Montalvo, NF, Mohamed, NM, Enticknap, JJ & Hill, RT 2005, 'Novel actinobacteria from marine sponges', *Antonie Van Leeuwenhoek*, vol. 87, pp. 29-36.

Morrow, C & Cárdenas, P 2015, 'Proposal for a revised classification of the Demospongiae (Porifera)', *Frontiers in Zoology*, vol. 12, no. 7, pp. 1-27.

Morrow, CC, Picton, BE, Erpenbeck, D, Boury-Esnault, N, Maggs, CA & Allcock, AL 2012, 'Congruence between nuclear and mitochondrial genes in Demospongiae: a new hypothesis for relationships within the G4 clade (Porifera: Demospongiae)', *Molecular Phylogenetics and Evolution*, vol. 62, no. 1, pp. 174-90.

Muyzer, G, de Waal, EC & Uitterlinden, AG 1993, 'Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA', *Applied and Environmental Microbiology*, vol. 59, no. 3, pp. 695-700.

Nakamura, K, Oshima, T, Morimoto, T, Ikeda, S, Yoshikawa, H, Shiwa, Y, Ishikawa, S, Linak, MC, Hirai, A, Takahashi, H, Altaf-Ul-Amin, M, Ogasawara, N & Kanaya, S 2011, 'Sequence-specific error profile of Illumina sequencers', *Nucleic Acids Research*, vol. 39, no. 13.

Nelson, MC, Morrison, HG, Benjamino, J, Grim, SL & Graf, J 2014, 'Analysis, Optimization and Verification of Illumina-Generated 16S rRNA Gene Amplicon Surveys', *PLoS ONE*, vol. 9, no. 4, p. e94249.

Nichols, SA 2005, 'An evaluation of support for order-level monophyly and interrelationships within the class Demospongiae using partial data from the large subunit rDNA and cytochrome oxidase subunit I', *Molecular Phylogenetics and Evolution*, vol. 34, pp. 81-96.

Noller, HF, Stolk, BJV, Moazed, D, Douthwaite, S & Gutell, RR 1985, 'Studies on the structure and function of 16S ribosomal RNA using structure-specific chemical probes', *Journal of Biosciences*, vol. 8, pp. 747-55.

Nubel, U, Garcia-Pichel, F & Muyzer, G 1997, 'PCR primers to amplify 16S rRNA genes from cyanobacteria', *Applied and Environmental Microbiology*, vol. 63, pp. 3327-32.

Olson, JB, Thacker, RW & Gochfeld, DJ 2014, 'Molecular community profiling reveals impacts of time, space, and disease status on the bacterial community associated with the Caribbean sponge Aplysina cauliformis', *FEMS Microbiology Ecology*, vol. 87, no. 1, pp. 268-79.

Pape, T, Hoffmann, F, Queric, NV, von Juterzenka, K, Reitner, J & Michaelis, W 2006, 'Dense populations of Archaea associated with the demosponge *Tentorium semisuberites* Schmidt, 1870 from Arctic deep-waters', *Polar Biology*, vol. 29, no. 8, pp. 662-7.

Pawlik, JR 2011, 'The chemical ecology of sponges on Caribbean reefs: natural products shape natural systems', *Bioscience*, vol. 61, pp. 888-98.

Petti, CA, Polage, CR & Schreckenberger, P 2005, 'The role of 16S rRNA gene sequencing in identification of microorganisms misidentified by conventional methods', *Journal of Clinical Microbiology*, vol. 43, pp. 6123-5.

Piel, J, Hui, D, Wen, G, Butzke, D, Platzer, M, Fusetani, N & Matsunaga, S 2004, 'Antitumor polyketide biosynthesis by an uncultivated bacterial symbiont of the marine sponge *Theonella swinhoei*', *Proceedings of the National Academy of Sciences of the United States of America*,

vol. 101, no. 46, pp. 16222-7.

Pillay, KRM, Asahida, T, Chen, CA, Terashima, H & Ida, H 2006, 'ITS Ribosomal DNA Distinctions and the Genetic Structures of Populations of Two Sympatric Species of *Pavona* (Cnidaria: Scleractinia) from Mauritius', *Zoological Studies*, vol. 45, no. 1, pp. 132-44.

Pimentel-Elardo, SM, Kozytska, S, Bugni, TS, Ireland, CM, Moll, H & Hentschel, U 2010, 'Anti-parasitic compounds from *Streptomyces* sp. strains isolated from mediterranean sponges', *Marine Drugs*, vol. 8, pp. 373-80.

Pita, L, Turon, X, López-Legentil, S & Erwin, PM 2013, 'Host rules: spatial stability of bacterial communities associated with marine sponges (*Ircinia* spp.) in the Western Mediterranean Sea', *FEMS Microbiology Ecology*, vol. 86, pp. 268-76.

Pöppe, J, Sutcliffe, P, Hooper, JNA, Wörheide, G & Erpenbeck, D 2010, 'COI barcoding reveals new clades and radiation patterns of Indo-Pacific sponges of the family Irciniidae (Demospongiae: Dictyoceratida)', *PLoS ONE*, vol. 5, no. 4, p. e9950.

Pruesse, E, Quast, C, Knittel, K, Fuchs, BM, Ludwig, W, Peplies, J & Glockner, FO 2007, 'SILVA: A comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB', *Nucleic Acids Research*, vol. 35, pp. 7188-96.

Pushkarev, D, Neff, NF & Quake, SR 2009, 'Single-molecule sequencing of an individual human genome', *Nature Biotechnology*, vol. 27, no. 9, pp. 847-50.

Radwan, M, Hanora, A, Zan, J, Mohamed, NM, Abo-Elmatty, DM, Abou-El-Ela, SH & Hill, RT 2010, 'Bacterial community analyses of two Red Sea sponges', *Marine Biotechnology*, vol. 12, no. 3, pp. 350-60.

Rappe, MS & Giovannoni, SJ 2003, 'The uncultured microbial majority', *Annual Review Microbiology*, vol. 57, pp. 369-94.

Redmond, NE, Raleigh, J, Soest, RWMv, Kelly, M, Travers, SAA, Bradshaw, B, Vartia, S, Stephens, KM & McCormack, GP 2011, 'Phylogenetic Relationships of the Marine Haplosclerida (Phylum Porifera) Employing Ribosomal (28S rRNA) and Mitochondrial (*cox1, nad1*) Gene Sequence Data', *PLoS ONE*, vol. 6, no. 9, p. e24344.

Redmond, NE, Van Soest, RWM, Kelly, M, Raleigh, J, Travers, SAA & Raleigh, J 2007, 'Reassessment of the classification of the order Haplosclerida (Class Demospongiae, Phylum Porifera) using 18S rRNA gene sequence data', *Molecular Phylogenetics and Evolution*, vol. 43, pp. 344-52.

Reimer, A, Blohm, A, Quack, T, Grevelding, CG, Kozjak-Pavlovic, V, Rudel, T, Hentschel, U & Abdelmohsen, UR 2015, 'Inhibitory activities of the marine streptomycete-derived compound SF2446A2 against *Chlamydia trachomatis* and *Schistosoma mansoni*', *Journal of Antibiotic,* vol. 68, no. 11, pp. 674-9.

Reveillaud, J, Maignien, L, Eren, AM, Huber, JA, Apprill, A, Sogin, ML & Vanreusel, A 2014, 'Host-specificity among abundant and rare taxa in the sponge microbiome', *ISME Journal*, vol. 8, pp. 1198-209.

Reveillaud, J, Remerie, T, Soest, Rv, Erpenbeck, D, Cárdenas, P, Derycke, S, Xavier, JR, Rigaux, A & Vanreusel, A 2010, 'Species boundaries and phylogenetic relationships between Atlanto-Mediterranean shallow-water and deep-sea coral associated *Hexadella* species

(Porifera, Ianthellidae)', *Molecular Phylogenetics and Evolution*, vol. 56, pp. 104-14.

Ribes, M, Dziallas, C, Coma, R & Riemannb, L 2015, 'Microbial diversity and putative diazotrophy in high- and low- microbial-abundance mediterranean sponges', *Applied and Environmental Microbiology*, vol. 81, no. 17, pp. 5683-93.

Rinke, C, Schwientek, P, Sczyrba, A, Ivanova, NN, Anderson, IJ, Cheng, JF, Darling, A, Malfatti, S, Swan, BK, Gies, EA, Dodsworth, JA, Hedlund, BP, Tsiamis, G, Sievert, SM, Liu, WT, Eisen, JA, Hallam, SJ, Kyrpides, NC, Stepanauskas, R, Rubin, EM, Hugenholtz, P & Woyke, T 2013, 'Insights into the phylogeny and coding potential of microbial dark matter', *Nature*, vol. 499, no. 7459, pp. 431-7.

Rodríguez-Marconi, S, De la Iglesia, R, Díez, B, Fonseca, CA, Hajdu, E & Trefault, N 2015, 'Characterization of bacterial, archaeal and eukaryote symbionts from antarctic sponges reveals a high diversity at a three-domain level and a particular signature for this ecosystem', *PLoS ONE*, vol. 10, no. 9, p. e0138837.

Rothberg, JM & Leamon, JH 2008, 'The development and impact of 454 sequencing', *Nature Biotechnology*, vol. 26, pp. 1117-24.

Rothman, RE, Majmudar, MD, Kelen, GD, Madico, G, Gaydos, CA, Walker, T & Quinn, TC 2002, 'Detection of bacteremia in emergency department patients at risk for infective endocarditis using universal 16S rRNA primers in a decontaminated polymerase chain reaction assay', *Journal of Infectious Diseases*, vol. 186, pp. 1677-81.

Schmalenberger, A, Schwieger, F & Tebbe, CC 2001, 'Effect of primers hybridizing to different evolutionarily conserved regions of the small-subunit rRNA gene in PCR-based microbial community analyses and genetic profiling', *Applied and Environmental Microbiology*, vol. 67, pp. 3557-63.

Schmidt, EW, Obraztsova, AY, Davidson, SK, Faulkner, DJ & Haygood, MG 2000, 'Identification of the antifungal peptide-containing symbiont of the marine sponge *Theonella swinhoei* as a novel delta-proteobacterium, '*Candidatus* Entotheonella palauensis'', *Marine Biology*, vol. 136, pp. 969-77.

Schmitt, S, Hentschel, U & Taylor, M 2012, 'Deep sequencing reveals diversity and community structure of complex microbiota in five Mediterranean sponges', in Maldonado M, Turon X, Becerro M & Jesús Uriz M (eds), vol. 219, pp. 341-51, *Ancient Animals, New Challenges*, Springer Netherlands.

Schmitt, S, Tsai, P, Bell, J, Fromont, J, Ilan, M, Lindquist, N, Perez, T, Rodrigo, A, Schupp, PJ, Vacelet, J, Webster, N, Hentschel, U & Taylor, MW 2012, 'Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges', *ISME Journal*, vol. 6, no. 3, pp. 564-76.

Schmitt, S, Wehrl, M, Lindquist, N, Weisz, JB & Hentschel, U 2007, 'Morphological and molecular analyses of microorganisms in Caribbean reef adult sponges and in corresponding reproductive material', in Custódio MR, Lôbo-Hajdu G, Hajdu E & Muricy G (eds), pp. 561-8, *Porifera Research: Biodiversity, Innovation and Sustainability*, MuseuNacional, Rio de Janeiro.

Schneemann, I, Nagel, K, Kajahn, I, Labes, A, Wiese, J & Imhoff, JF 2010, 'Comprehensive Investigation of Marine Actinobacteria Associated with the Sponge *Halichondria panicea*', *Applied and Environmental Microbiology*, vol. 76, no. 11, pp. 3702–14.

Schönberg, CHL & Barthel, D 1997, 'Unreliability of demosponge skeletal characters: the example of *Halichondria panicea*', in Watanabe Y & Fusetani N (eds), pp. 41-53, *Sponge Sciences: Multidisciplinary Perspectives*, Springer, Tokyo.

Schöttner, S, Hoffmann, F, Cárdenas, P, Rapp, HT, Boetius, A & Ramette, A 2013, 'Relationships between Host Phylogeny, Host Type and Bacterial Community Diversity in Cold-Water Coral Reef Sponges', *PLoS ONE*, vol. 8, no. 2, p. e55505.

Schröder, HC, Efremova, SM, Itskovich, VB, Belikov, S, Masuda, Y, Krasko, A, Müller, IM & Müller, WEG 2003, 'Molecular phylogeny of the freshwater sponges in Lake Baikal', *Journal of Zoological Systematics and Evolutionary Research*, vol. 41, pp. 80-6.

Seipke, RF, Kaltenpoth, M & Hutchings, MI 2012, '*Streptomyces* as symbionts: An emerging and widespread theme?', *FEMS Microbiology Reviews*, vol. 36, pp. 862-76.

Shearer, TL, Oppen, MJHV, Romano, SL & Wörheide, G 2002, 'Slow mitochondrial DNA sequence evolution in the Anthozoa (Cnidaria)', *Molecular Ecology*, vol. 11, pp. 2475-87.

Shendure, J, Porreca, GJ, Reppas, NB, Lin, X, McCutcheon, JP, Rosenbaum, AM, Wang, MD, Zhang, K, Mitra, RD & Church, GM 2005, 'Accurate multiplex polony sequencing of an evolved bacterial genome', *Science*, vol. 309, no. 5741, pp. 1728-32.

Simister, RL, Deines, P, Botté, ES, Webster, NS & Taylor, MW 2012, 'Sponge-specific clusters revisited: a comprehensive phylogeny of sponge-associated microorganism', *Environmental Microbiology*, vol. 14, no. 2, pp. 517-24.

Sogin, ML, Morrison, HG, Huber, JA, Welch, DM, Huse, SM, Neal, PR, Arrieta, JM & Herndl, GJ 2006, 'Microbial diversity in the deep sea and the underexplored “rare biosphere” ', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, pp. 12115-20.

Solé-Cava, AM, Boury-Esnault, N, Vacelet, J & Thorpe, JP 1992, 'Biochemical divergence and systematics in sponges of the genera *Corticium* and *Oscarella* (Demospongiae: Homoscleromorpha) in the Mediterranean Sea', *Marine Biology*, vol. 113, no. 2, pp. 299-304.

Sorokin, S, Laperousaz, T, Currie, D, Fromont, J & Goudie, L 2013, 'Biodiversity of demosponge species in shelf waters off South Australia, including Commonwealth waters to the shelf edge', paper presented to The 9th World Sponge Conference Perth, Australia.

Srivastava, M, Simakov, O, Chapman, J, Fahey, B, Gauthier, MEA, Mitros, T, Richards, GS, Conaco, C, Dacre, M, Hellsten, U, Larroux, C, Putnam, NH, Stanke, M, Adamska, M, Darling, A, Degnan, SM, Oakley, TH, Plachetzki, DC, Zhai, Y, Adamski, M, Calcino, A, Cummins, SF, Goodstein, DM, Harris, C, Jackson, DJ, Leys, SP, Shu, S, Woodcroft, BJ, Vervoort, M, Kosik, KS, Manning, G, Degnan, BM & Rokhsar, DS 2010, 'The Amphimedon queenslandica genome and the evolution of animal complexity', *Nature*, vol. 466, pp. 720-6.

Stohr, K, Hafner, B, Nolte, O, Wolfrum, J, Sauer, M & Herten, DP 2005, 'Species-specific identification of mycobacterial 16S rRNA PCR amplicons using smart probes', *Analytical Chemistry*, vol. 77, pp. 7195-203.

Subramani, R, Kumar, R, Prasad, P & Aalbersberg, W 2013, 'Cytotoxic and antibacterial substances against multi-drug resistant pathogens from marine sponge symbiont: Citrinin, a secondary metabolite of *Penicillium* sp.', *Asian Pacific Journal of Tropical Biomedicine*, vol. 3, pp. 291-6.

Sun, W, Dai, S, Jiang, S, Wang, G, Liu, G, Wu, H & Li, X 2010, 'Culture-dependent and culture-independent diversity of Actinobacteria associated with the marine sponge *Hymeniacidon perleve* from the South China Sea', *Antonie Van Leeuwenhoek*, vol. 98, no. 1, pp. 65-75.

Supong, K, Suriyachadkun, C, Pittayakhajonwut, P, Suwanborirux, K & Thawai, C 2013, '*Micromonospora spongicola* sp. nov., an actinomycete isolated from a marine sponge in the Gulf of Thailand', *The Journal of Antibiotics*, vol. 66, pp. 505-9

Supong, K, Suriyachadkun, C, Suwanborirux, K, Pittayakhajonwut, P & Thawai, C 2013, '*Verrucosispora andamanensis* sp. nov., isolated from a marine sponge', *International Journal of Systematic and Evolutionary Microbiology*, vol. 63, pp. 3970-4.

Szitenberg, A, Becking, LE, Vargas, S, Fernandez, JCC, Santodomingo, N, Wörheide, G, Ilan, M, Kelly, M & Huchon, D 2013, 'Phylogeny of Tetillidae (Porifera, Demospongiae, Spirophorida) based on three molecular markers', *Molecular Phylogenetics and Evolution*, vol. 67, pp. 509-19.

Tabares, P, Pimentel-Elardo, SM, Schirmeister, T, Hünig, T & Hentschel, U 2011, 'Anti-protease and immunomodulatory activities of bacteria associated with Caribbean sponges', *Marine Biotechnology (NY)*, vol. 13, no. 5, pp. 883-92.

Tamaki, H, Wright, CL, Li, X, Lin, Q, Hwang, C, Wang, S, Thimmapuram, J, Kamagata, Y & Liu, W-T 2011, 'Analysis of 16S rRNA amplicon sequencing options on the Roche/454 next-generation titanium sequencing platform', *PLoS ONE*, vol. 6, no. 9, p. e25263.

Taylor, MW, Radax, R, Steger, D & Wagner, M 2007, 'Sponge-associated microorganisms: evolution, ecology, and biotechnological potential', *Microbiology and Molecular Biology Reviews*, vol. 71, pp. 295-347.

Taylor, MW, Tsai, P, Simister, RL, Deines, P, Botte, E, Ericson, G, Schmitt, S & Webster, NS 2013, ''Sponge-specific' bacteria are widespread (but rare) in diverse marine environments', *ISME Journal*, vol. 7, pp. 438-43.

Teske, A & Sorensen, KB 2007, 'Uncultured archaea in deep marine subsurface sediments: have we caught them all?', *ISME Journal*, vol. 2, pp. 3-18.

Thacker, RW & Freeman, CJ 2012, 'Sponge-microbe symbioses: recent advances and new directions', in Becerro MA, Uriz MJ, Maldonado M & Turon X (eds), *Advances in Marine Biology*, vol. 62.

Thacker, RW, Hill, AL, Hill, MS, Redmond, NE, Collins, AG, Morrow, CC, Spicer, L, Carmack, CA, Zappe, ME, Pohlmann, D, Hall, C, Diazjj, MC & Bangalore, PV 2013, 'Nearly complete 28S rRNA gene sequences confirm new hypotheses of sponge evolution', *Integrative and Comparative Biology*, vol. 53, no. 3, pp. 373-87.

Thomas, T, Moitinho-Silva, L, Lurgi, M, Björk, JR, Easson, C, Astudillo-García, C, Olson, JB, Erwin, PM, López-Legentil, S, Luter, H, Chaves-Fonnegra, A, Costa, R, Schupp, PJ, Steindler, L, Erpenbeck, D, Gilbert, J, Knight, R, Ackermann, G, Lopez, JV, Taylor, MW, Thacker, RW, Montoya, JM, Hentschel, U & Webster, NS 2016, 'Diversity, structure and convergent evolution of the global sponge microbiome', *Nature Communications*, vol. 7, p. 11870.

Timmerman, L 2015, *DNA sequencing market will exceed $20 billion, says Illumina CEO Jay Flatley* Forbes, viewed 29 Apr 2015.

Traxler, MF & Kolter, R 2015, 'Natural products in soil microbe interactions and evolution',

Natural Product Reports, vol. 32, pp. 956-70.

Tremblay, J, Singh, K, Fern, A, Kirton, ES, He, S, Woyke, T, Lee, J, Chen, F, Dangl, JL & Tringe, SG 2015, 'Primer and platform effects on 16S rRNA tag sequencing', *Frontiers in Microbiology*, vol. 6, pp. 771.

Tringe, SG & P., H 2008, 'A renaissance for the pioneering 16S rRNA gene', *Current Opinion in Microbiology*, vol. 11, no. 5, pp. 442-6.

Uriz, MJ, Turon, X, Becerro, MA & Agell, G 2003, 'Siliceous spicules and skeleton frameworks in sponges: origin, diversity, ultrastructural patterns, and biological functions', *Microscopy Research and Technique*, vol. 62, pp. 279–99.

Vacelet, J & Donadey, C 1977, 'Electron microscope study of the association between some sponges and bacteria', *Journal of Experimental Marine Biology and Ecology*, vol. 30, pp. 301-14.

Valliappan, K, Sun, W & Li, Z 2014, 'Marine actinobacteria associated with marine organisms and their potentials in producing pharmaceutical natural products', *Applied Microbiology and Biotechnology*, vol. 98, pp. 7365-77.

Valouev, A, Ichikawa, J, Tonthat, T, Stuart, J, Ranade, S, Peckham, H, Zeng, K, Malek, JA, Costa, G, McKernan, K, Sidow, A, Fire, A & Johnson, SM 2008, 'A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning', *Genome Research*, vol. 18, no. 7, pp. 1051–63.

Van de Peer, Y, Chapelle, S & De Wachter, R 1996a, 'A quantitative map of nucleotide substitution rates in bacterial rRNA', *Nucleic Acids Research*, vol. 24, pp. 3381-91.

Van de Peer, Y, Chapelle, S & De Wachter, R 1996b, 'A quantitative map of nucleotide substitution rates in bacterial rRNA', *Nucleic Acids Research*, vol. 24, pp. 3381-91.

van Dijk, EL, Auger, H, Jaszczyszyn, Y & Thermes, C 2014, 'Ten years of next-generation sequencing technology', *Trends in Genetics*, vol. 30, no. 9, pp. 418-26.

Van Soest, RWM, Boury-Esnault, N, Hooper, JNA, Rützler, K, De Voogd, NJ, Alvarez de Glasby, B, Hajdu, E, Pisera, AB, Manconi, R, Schoenberg, C, Klautau, M, Picton, B, Kelly, M, Vacelet, J, Dohrmann, M, Díaz, M-C, Cárdenas, P & Carballo, JL 2016, *World Porifera database*, http://www.marinespecies.org/porifera, 2016-11-03.

van Soest, RWM, Boury-Esnault, N, Vacelet, J, Dohrmann, M, Erpenbeck, D, De Voogd, NJ, Santodomingo, N, Vanhoorne, B, Kelly, M & Hooper, JNA 2012, 'Global diversity of sponges (Porifera)', *PLoS ONE*, vol. 7, no. 4, p. e35105.

Van Soest, RWM & Braekman, JC 1999, 'Chemosystematics of Porifera: a review', *Memoirs of the Queensland Museum*, vol. 44, pp. 569–89.

Vargas, S, Erpenbeck, D, Göcke, C, Hall, KA, Hooper, JNA, Janussen, D & Wörheide, G 2012, 'Molecular phylogeny of *Abyssocladia* (Cladorhizidae: Poecilosclerida) and *Phelloderma* (Phellodermidae: Poecilosclerida) suggests a diversification of chelae microscleres in cladorhizid sponges', *Zoologica Scripta*, vol. 00, pp. 000-.

Vargas, S, Kelly, M, Schnabel, K, Mills, S, Bowden, D & Wörheide, G 2015, 'Diversity in a cold hot-spot: DNA-barcoding reveals patterns of evolution among Antarctic Demosponges (class Demospongiae, phylum Porifera)', *PLoS ONE*, vol. 10, no. 6, p. e0127573.

Vargas, S, Schuster, A, Sacher, K, Büttner, G, Schätzle, S, Läuchli, B, Hall, K, Hooper, JNA, Erpenbeck, D & Wörheide, G 2012, 'Barcoding sponges: An overview based on comprehensive sampling', *PLoS ONE*, vol. 7, no. 7, p. e39345.

Varma-Basil, M, El-Hajj, H, Marras, SA, Hazbon, MH, Mann, JM, Connell, ND, Kramer, FR & Alland, D 2004, 'Molecular beacons for multiplex detection of four bacterial bioterrorism agents', *Clinical Chemistry*, vol. 50, pp. 1060-2.

Vicente, J, Stewart, A, Song, B, Hill, RT & Wright, JL 2013, 'Biodiversity of actinomycetes associated with Caribbean sponges and their potential for natural product discovery', *Marine Biotechnoogy*, vol. 15, pp. 413-24.

Voigt, O, Eichmann, V & Wörheide, G 2012, 'First evaluation of mitochondrial DNA as a marker for phylogeographic studies of Calcarea: a case study from *Leucetta chagosensis*', *Hydrobiologia*, vol. 687, pp. 101-6.

Voigt, O, Wülfing, E & Wörheide, G 2012, 'Molecular phylogenetic evaluation of classification and scenarios of character evolution in Calcareous Sponges (Porifera, Class Calcarea)', *PLoS ONE*, vol. 7, p. e33417.

Wagner, M, Noguera, DR, Juretschko, S, Rath, G, Koops, HP & Schleifer, KH 1998, 'Combining fluorescent in situ hybridization (FISH) with cultivation and mathematical modeling to study population structure and function of ammonia-oxidizing bacteria in activated sludge', *Water Science and Technology*, vol. 37, no. 4-5, pp. 441-9.

Wahab, MAA, Fromont, J, Whalan, S, Webster, N & Andreakis, N 2014, 'Combining morphometrics with molecular taxonomy: How different are similar foliose keratose sponges from the Australian tropics?', *Molecular Phylogenetics and Evolution*, vol. 73, pp. 23-9.

Wang, Q, Garrity, GM, Tiedje, JM & Cole, JR 2007, 'Naive Bayesian classification for rapid assignment of rRNA sequences into the new bacterial taxonomy', *Applied and Environmental Microbiology*, vol. 73, no. 16, pp. 5261-7.

Wang, X, Yu, X, Zhu, W, McCombie, WR, Antoniou, E, Powers, RS, Davidson, NO, Li, E & Williams, J 2015, 'A trimming-and-retrieving alignment scheme for reduced representation bisulfite sequencing', *Bioinformatics*, vol. 31, no. 12, pp. 2040-2.

Wang, Y & Qian, P-Y 2009, 'Conservative fragments in bacterial 16S rRNA genes and primer design for 16S ribosomal DNA amplicons in metagenomic studies', *PLoS ONE*, vol. 4, no. 10, p. e7401.

Webster, N, Negri, A, Munro, M & Battershill, C 2004, 'Diverse microbial communities inhabit Antarctic sponges', *Environmental Microbiology*, vol. 6, no. 3, pp. 288-300.

Webster, NS & Hill, RT 2001, 'The culturable microbial community of the Great Barrier Reef sponge *Rhopaloeides odorabile* is dominated by an α-Proteobacterium', *Marine Biology*, vol. 138, no. 4, pp. 843-51.

Webster, NS & Taylor, MW 2012, 'Marine sponges and their microbial symbionts: love and other relationships', *Environmental Microbiology*, vol. 14, no. 2, pp. 335-46.

Webster, NS, Taylor, MW, Behnam, F, Lücker, S, Rattei, T, Whalan, S, Horn, M & Wagner, M 2010, 'Deep sequencing reveals exceptional diversity and modes of transmission for bacterial

sponge symbionts', *Environmental Microbiology*, vol. 12, no. 8, pp. 2070–82.

Webster, NS, Wilson, KJ, Blackall, LL & Hill, RT 2001, 'Phylogenetic diversity of bacteria associated with the marine sponge *Rhopaloeides odorabile* ', *Applied and Environmental Microbiology*, vol. 67, pp. 434-44.

Wehrl, M, Steinert, M & Hentschel, U 2007, 'Bacterial uptake by the marine sponge *Aplysina aerophoba*', *Microbial Ecology*, vol. 53, pp. 355-65.

Weigela, BL & Erwin, PM 2016, 'Intraspecific variation in microbial symbiont communities of the sun sponge, *Hymeniacidon heliophila*, from intertidal and subtidal habitats', *Applied and Enrivonmental Micorbiology*, vol. 82, pp. 650–8.

Weisz, J, Hentschel, U, Lindquist, N & Martens, C 2007, 'Linking abundance and diversity of sponge-associated microbial communities to metabolic differences in host sponges', *Marine Bioogy*, vol. 152, pp. 475-83.

Weisz, JB, Lindquist, N & Martens, CS 2008, 'Do associated microbial abundances impact marine demosponge pumping rates and tissue densities?', *Oecologia*, vol. 155, pp. 367-76.

White, JR, Patel, J, Ottesen, A, Arce, G, Blackwelder, P & Lopez, JV 2012, 'Pyrosequencing of Bacterial Symbionts within *Axinella corrugata* Sponges: Diversity and Seasonal Variability', *PLoS ONE*, vol. 7, no. 6, p. e38204.

Wilkinson, CR, Garrone, R & Vacelet, J 1984, 'Marine sponges discriminate between food bacteria and bacterial symbionts: electron microscope radioautography and in situ evidence', *Proceedings of the Royal Society of London. Series B*, vol. 220, pp. 519-28

Witt, JD, Threloff, DL & Hebert, PDN 2006, 'DNA barcoding reveals extraordinary cryptic diversity in an amphipod genus: implications for desert spring conservation', *Molecular Ecology*, vol. 15, pp. 3073–82.

Woese, CR 1987, 'Bacterial evolution', *Microbiology Review*, vol. 51, pp. 221-71.

Wörheide, G 1998, 'The reef cave dwelling ultraconservative coralline Demosponge *Astrosclera willeyana* LISTER 1900 from the Indo-Pacific-micromorphology, ultrastructure, biocalcification, isotope record, taxonomy, biogeography, phylogeny', *Facies*, vol. 38, pp. 1-88.

Wörheide, G 2006, 'Low variation in partial cytochrome oxidase subunit I (COI) mitochondrial sequences in the coralline demosponge *Astrosclera willeyana* across the Indo-Pacific', *Marine Biology*, vol. 148, pp. 907-12.

Wörheide, G, Degnan, BM & Hooper, JNA 2000, 'Population phylogenetics of the common coral reef sponges *Leucetta* spp. and *Pericharax* spp. (Porifera: Calcarea) from the Great Barrier Reef and Vanuatu', in Hopley D, Hopley P, Tamelander J & Done T (eds), vol. 21, *Ninth International Coral Reef Symposium*, Bali, Indonesia.

Wörheide, G, Degnan, BM, Hooper, JNA & Reitner, J 2002, 'Biogeography and taxonomy of the Indo-Pacific reef cave dwelling coralline demosponge *Astrosclera* '*willeyana*': new data from nuclear internal transcribed spacer sequences', in Moosa KM, Soemodihardjo S & Soegiarto A (eds), pp. 339-45, *The 9th International Coral Reef Symposium*, Bali, Indonesia.

Wörheide, G & Erpenbeck, D 2007, 'DNA taxonomy of sponges-progress and perspectives', *Journal of the Marine Biological Association of the United Kingdom*, vol. 87, pp. 1629-33.

Wörheide, G, Erpenbeck, D & Menke, C 2008, 'The Sponge Barcoding Project -aiding in the identification and description of poriferan taxa', in Custódio MR, Lôbo-Hajdu G, Hajdu E & Muricy G (eds), pp. 123-8, *Porifera research: Biodiversity, innovation & sustainability*, Museu Nacional de Rio de Janiero Book Series, Rio de Janeiro.

Wörheide, G, Hooper, JNA & Degnan, BM 2002, 'Phylogeography of western Pacific *Leucetta 'chagosensis'* (Porifera: Calcarea) from ribosomal DNA sequences: implications for population history and conservation of the Great Barrier Reef World Heritage Area (Australia)', *Molecular Ecology*, vol. 11, pp. 1753-68.

Wörheide, G, Nichols, SA & Goldberg, J 2004, 'Intragenomic variation of the rDNA internal transcribed spacers in sponges (Phylum Porifera): implications for phylogenetic studies', *Molecular Phylogenetics and Evolution*, vol. 33, pp. 816-30.

Wu, D, Hugenholtz, P, Mavromatis, K, Pukall, R, Dalin, E, Ivanova, NN, Kunin, V, Goodwin, L, Wu, M & Tindall, BJ 2009, 'A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea', *Nature*, vol. 462, pp. 1056-60

Wulff, J 2012, 'Ecological interactions and the distribution, abundance, and diversity of sponges', *Advances in Marine Biology*, vol. 61, pp. 273-344.

Wuyts, J, Van de Peer, Y & De Watcher, R 2001, 'Distribution of substitution rates and location of insertion sites in the tertiary structure of ribosomal RNA', *Nucleic Acids Research.*, vol. 29, no. 24, pp. 5017-28.

Xavier, JR, Rachello-Dolmen, PG, Parra-Velandia, F, Schönberg, CHL, Breeuwer, JAJ & Van Soest, RWM 2010, 'Molecular evidence of cryptic speciation in the ''cosmopolitan" excavating sponge Cliona celata (Porifera, Clionaidae)', *Molecular Phylogenetics and Evolution*, vol. 56, pp. 13-20.

Xi, L, Ruan, J & Huang, Y 2012, 'Diversity and biosynthetic potential of culturable actinomycetes associated with marine sponges in the China seas', *International Journal of Molecular Sciences*, vol. 13, pp. 5917-32.

Xin, Y, Huang, J, Deng, M & Zhang, W 2008, 'Culture-independent nested PCR method reveals high diversity of actinobacteria associated with the marine sponges *Hymeniacidon perleve* and *Spongia* sp.', *Antonie Van Leeuwenhoek*, vol. 94, pp. 533-42.

Yang, S, Lin, S, Kelen, GD, Quinn, TC, Dick, JD, Gaydos, CA & Rothman, RE 2002, 'Quantitative multiprobe PCR assay for simultaneous detection and identification to species level of bacterial pathogens', *Journal of Clinical Microbiology*, vol. 40, pp. 3449-54.

Yu, Z, García-González, R, Schanbacher, FL & Morrison, M 2008, 'Evaluations of different hypervariable regions of archaeal 16S rRNA genes in profiling of methanogens by Archaea-specific PCR and denaturing gradient gel electrophoresis', *Applied and Environmental Microbiology*, vol. 74, no. 3, pp. 889-93.

Yu, ZM, M. 2004, 'Comparisons of different hypervariable regions of rrs genes for use in fingerprinting of microbial communities by PCR-denaturing gradient gel electrophoresis', *Applied and Environmental Microbiology*, vol. 70, no. 8, pp. 4800-6.

Zhang, H, Zhang, W, Jin, Y, Jin, M & Yu, X 2008, 'A comparative study on the phylogenetic diversity of culturable actinobacteria isolated from five marine sponge species', *Antonie Van*

Leeuwenhoek., vol. 93, no. 3, pp. 241-8.

Zhang, HT, Lee, YK, Zhang, W & Lee, HK 2006, 'Culturable actinobacteria from the marine sponge *Hymeniacidon perleve*: isolation and phylogenetic diversity by 16S rRNA gene-RFLP analysis', *Antonie Van Leeuwenhoek*, vol. 90, pp. 159-69.

Zhang, T, Shao, MF & Ye, L 2012, '454 Pyrosequencing reveals bacterial diversity of activated sludge from 14 sewage treatment plants', *ISME Journal*, vol. 6, no. 6, pp. 1137-47.

CHAPTER 2 DEVELOPMENT OF A MULTILOCUS-BASED APPROACH FOR SPONGE (PORIFERA) IDENTIFICATION: REFINEMENT AND LIMITATIONS

For sponges (phylum Porifera), there is no reliable molecular protocol available for species identification. To address this gap, a multilocus-based Sponge Identification Protocol (SIP) was developed and validated by a sample of 37 sponge species belonging to 10 orders from South Australia. The universal barcode COI mtDNA, 28S rRNA gene (D3-D5), and the nuclear ITS1-5.8S-ITS2 region were evaluated for their suitability and capacity for sponge identification. The highest Bit Score was applied to infer the identity. The reliability of SIP was validated by phylogenetic analysis. The 28S rRNA gene and COI mtDNA performed better than the ITS region in classifying sponges at various taxonomic levels. A major limitation is that the databases are not well populated and possess low diversity, making it difficult to conduct the molecular identification protocol. The identification is also impacted by the accuracy of the morphological classification of the sponges whose sequences have been submitted to the database. Re-examination of the morphological identification further demonstrated and improved the reliability of sponge identification by SIP. Integrated with morphological identification, the multilocus-based SIP offers an improved protocol for more reliable and effective sponge identification, by coupling the accuracy of different DNA markers.

2.1 Introduction

Sponges (phylum Porifera) are the oldest multicellular animals, and are sessile and benthic filter-feeders (Wulff 2012). In marine habitats, they are highly diverse and play important roles in biogeochemical cycling (Wulff 2012), in the spatial structuring of the seafloor, and in the benthic-pelagic coupling of nutrient transfer within ocean ecosystems (Bell 2008). Sponges are also commercially important for the pharmaceutical and biomaterial industries as they participate in complex biotic interactions with diverse macrobiotic taxa (Bell 2008) and microbiological communities (Webster & Taylor 2012) to produce up to 30% of all active marine metabolites found (Leal et al. 2012).

According to the World Porifera Database (Van Soest et al. 2016), there are more than 8,700 valid species, 7,300 of which belong to the class Demospongiae. Porifera are an important group of Metazoa in which species identification is particularly difficult because the available characters used for classification are limited (Hooper & Soest 2002). For example, the sponge family Polymastiidae possesses a relatively simple spicule assortment providing a rather scant set of taxonomic characters (Plotkin, Rapp & Gerasimova 2012). Some features

are in fact also displayed by some taxa from other families (Plotkin, Rapp & Gerasimova 2012). Generally, these characters are their organic and inorganic skeletons, including skeletal size, shape, structure and composition (Hooper & Soest 2002). However, the arrangement of these skeletal elements can be inconsistent, and our understanding of the evolution of skeletal traits is incomplete (Cárdenas et al. 2011; Morrow et al. 2013; Nichols 2005; Plotkin et al. 2016; Pöppe et al. 2010; Vargas et al. 2015). Indeed, traditional morphological identification methods often lead to erroneous classification (Morrow et al. 2012; Redmond et al. 2013) and the actual species diversity and distribution may be underestimated (Hooper et al. 2013).

Molecular approaches, such as DNA barcoding, provide a potential solution for sponge classification (Erpenbeck et al. 2012; Fontaneto, Flot & Tang 2015; Wörheide & Erpenbeck 2007). This study applied DNA markers to assist in the identification of individuals against already known species, which consists of comparing standardised stretches of DNA (barcodes) to reference databases to identify sponges. The mitochondrial genome (mtDNA) exists in all eukaryotic cells and is a good marker for species identification because of its clonal (maternal) mode of inheritance and clock-like evolutionary rate (Galtier et al. 2009; Voigt, Eichmann & Wörheide 2012). It has been used to study species identification, sponge diversification patterns (Pöppe et al. 2010) as well as phylogenetic relationships (Erpenbeck, et al. 2012) with varying degrees of success (Dohrmann et al. 2012; Lavrov, Wang & Kelly 2008). The COI mtDNA locus is a conservative region but with highly variable sequences. It is the most commonly used mitochondrial marker of approximately 700 bp at the 5'end of the cytochrome *c* oxidase subunit I (COI) gene. This gene is relatively easy to amplify as it is conserved across multicellular animals (Folmer et al. 1994) and abundant in eukaryotic DNA (Heim, Nickel & Brümmer 2007). Many studies reported that COI mtDNA successfully discriminated sponges at different taxonomic levels (Morrow et al. 2012; Morrow et al. 2013; Nichols 2005; Vargas et al. 2015). However, some species that can be clearly distinguished on the basis of morphology show similar COI sequences (Carella et al. 2016; Heim, Nickel & Brümmer 2007; Pöppe et al. 2010). Studies on the COI intraspecific variation has been used more regularly to classify other metazoans at the species level, but less so for sponges (Duran, Pascual & Turon 2004; Wörheide 2006).

Slow mitochondrial evolution is a problem for the resolution of phylogenies at the species and genus levels using standard mitochondrial markers. However, the question remains as to whether faster evolving gene regions can be identified for use in conjunction with the standard COI mtDNA barcode (Pöppe et al. 2010). The nuclear ribosomal genes of eukaryotes, such as the 28S (large subunit, LSU) rRNA genes (Gerbi 1985), are arranged in tandemly repeated clusters, where transcribed units alternate with non-transcribed units called

spacers, such as the internal transcribed spacer 1 (ITS1) and 2 (ITS2) (Borchiellini et al. 2000). The 28S rRNA gene has regions that are sufficiently heterogeneous to address phylogeny at different levels (Cárdenas et al. 2011; Erpenbeck et al. 2012). For example, the regions D3-D5 have been used as the DNA markers for sponge taxonomy (Morrow et al. 2013; Redmond et al. 2011). The more rapidly evolving ITS regions are commonly used as “high resolution” markers. In general, ITS regions are used to reconstruct relationships ranging from those between populations to those between the taxonomic “families”. They have been used for phylogenetic and phylogeographic analyses of non-bilaterian metazoans such as corals (Pillay et al. 2006) and sponges (Wahab et al. 2014).

Based on current knowledge about the resolution of different DNA markers, no single ideal marker for all sponge species exists, as each marker has its own strengths and limitations (Duran, Pascual & Turon 2004; Szitenberg et al. 2013; Voigt, Eichmann & Wörheide 2012; Wörheide 2006). The incomplete sequence entries in the gene database limit the application of the phylogeny-based molecular taxonomic approach for species identification. In the NCBI database, the sponge derived gene submissions only cover a few hundred (Benson et al. 2009) out of the known 8,700 sponge species. Therefore, the aims are to establish an effective and practical multilocus-based molecular approach for sponge identification to respond to these issues. This study, using South Australian sponges, was set up to (1) develop a Sponge Identification Protocol (SIP) using three DNA markers (the mitochondrial COI gene, the nuclear 28S rRNA gene and the nuclear ITS region); (2) validate the reliability of the SIP by phylogenetic analysis; (3) evaluate the efficiency of the proposed SIP in identifying 37 sponge species (three individuals for each); (4) approach a final identity by re-examining the morphological characters and mutual validation with SIP identification to discriminate the sponges whose identities are ambiguous; and (5) demonstrate the resolution and the suitability of different DNA markers in conjunction with morphological characters for sponge classification.

2.2 Materials and methods

2.2.1 Sponge collection and morphological classification

Sponges in this study were sampled under Exemption Permit Number 9902620 by the South Australian Research Development Institute (SARDI), issued by Primary Industries and Regions South Australia. Materials were used by Flinders University under a Material Transfer Agreement with SARDI and did not involve endangered or protected species. Sponges were collected from four different geographic locations in South Australia in 2012 and 2013 (Fig. 2-1a): Rapid Bay (35°31'17.25"S 138°11'15.26"E), Outer Harbour (34°46'26.90"S 138°29'08.54"E), Klein Point (34°55'41.79"S 137°47'19.42"E), and Williams Island (35°01'37.51"S 135°58'2.11"E).

A portable fridge and iceboxes kept the sponge chilled during the transportation.

The sample treatment was conducted immediately after the samples arrived at the laboratory. They were washed by sterilised sea water to remove soil or contaminations. The photos were taken as records. The sponges were cut into small pieces (about 1 cm^3) and were stored in sealed sample bags. They were kept in -80ºC for the following analysis and long term storage. Several pieces covering the surface and internal tissues were stored in a sample jar with 70% ethanol for morphological identification.

Figure 2 - 1 The sampling locations and the photos of the 37 sponge species. a. Four sampling locations (A-D), Map data: Google Earth. b. Photos of the 37 sponge species (underwater photos courtesy of David Wiltshire).

All the 37 potential species (111 sponge individuals) were first classified using morphological features (Hooper & Soest 2002). Preparation for histological sections and spicule preparations followed the methods in ‘sponguide’ (Hooper 2003). The classification followed the revised Demosponge classification (Morrow & Cárdenas 2015). All the species are lodged with the South Australian Museum in Adelaide, South Australia with the museum vouchers SAMA S1960 to SAMA S1996. The photos of the 37 sponge species are presented in Figure 2-1b. Brief descriptions of the morphological characteristics are available in the Appendix Table 2-1.

2.2.2 DNA extraction

Whole genomic DNA was extracted from sponge tissue frozen at -80°C. A conventional cetyltrimethylammonium bromide (CTAB)-based protocol (Taylor et al. 2004) was used for isolating DNA. Briefly, the sponge tissues were ground under liquid nitrogen. The CTAB extraction buffer was applied to lyse tissues and then combined with polyvinylpyrrolidone (PVP) and β-mercaptoethanol to help remove phenolic compounds and tannins in the extract. To separate the proteins and polysaccharides from nucleic acids, phenol: chloroform: isoamyl alcohol (25:24:1) was utilised before DNA was precipitated with chilled isopropanol. DNeasy Blood & Tissue Kit (QIAGEN, Germany) was used for sponges that did not yield high quality DNA with the CTAB method. For each potential species, triplicate DNA extractions from three different individuals were obtained. The purified DNA was resuspended in 35 µl of sterile distilled water and stored at -20°C. The purity and quantity of DNA were determined with a Nanodrop 1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and only high quality DNA was used for subsequent PCR reactions.

2.2.3 PCR amplification and sequencing

The COI locus was amplified using the universal primers LCO1490 (5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3') and HCO2198 (5'-TAA ACT TCA GGG TGA CCA AAA AAT CA-3') (Folmer et al. 1994). The thermocycler was programmed as follows: a 1-min initial denaturation at 94°C; 5 cycles of 94°C for 30 sec, 45°C for 90 sec and 72°C for 1 min; 35 cycles of 94°C for 30 sec, 51°C for 40 sec and 72°C for 1 min; and a final extension step at 72°C for 5 min. For the sponges that could not be amplified using these universal primers, the following two pairs of primers universal for metazoan were applied in a nested-PCR: C1-J2165 (5'-AAG TTT ATA TTT TAA TTT TAC CCC AGT GG-3') and C-Npor 2760 (5'-TCT AGG TAA TCC AGC TAA ACC-3') (Erpenbeck et al. 2002); CO1porF1 (5'-CCN CAN TTN KCN GMN AAA AAA CA-3') and CO1porR1 (5'-AAN TGN TGN GGR AAR AAN G-3') (Erpenbeck et al. 2006). For the second round of the nested-PCR, 5 µl of the amplicon adjusted to a DNA concentration of 50 ng/ µl was used as the template.

One set of primers developed to amplify the partial 28S rRNA gene was NL4F (5'-GAC CCG AAA GAT GGT GAA CTA-3') and NL4R (5'-ACC TTG GAG ACC TGA TGC G-3') (Nichols 2005) for regions D3- D5. The complete 28S rRNA gene alignment demonstrated that this region exhibited suitable levels of variability between sponge taxa allowing for the resolution of relatively deep phylogenetic relationships (Medina et al. 2001). Thermocycler conditions were as follows: a 10-min initial denaturation at 95°C; 35 cycles of 95°C for 1 min, 56°C for 1 min and 72°C for 1 min; and a final extension step at 72°C for 7 min. For the sponges that could not be amplified, alternative primers were employed: RD3A (5'-GAC CCG TCT TGA AAC ACG A-3') and RD5B2 (5'-ACA CAC TCC TTA GCG GA-3') (Erpenbeck et al. 2012).

The ITS was amplified using the following thermocycler program: 94°C for 2 min; 35 cycles of 94°C for 30 s, 45°C for 20 s, 65°C for 60 s; and a final extension step of 72°C for 10 min. The ITS primers were ITSRA2 (5'-GTC CCT GCC CTT TGT ACA CA-3') and ITS2.2 (5'-CCT GGT TAG TTT CTT TTC CTC CGC-3') (Adlard & Lester 1995).

Duplicate PCRs of each locus were prepared for three individuals belonging to one potential species. The amplification products were purified using the Ultra Clean® PCR Clean-Up Kit (MoBio), then Sanger-sequenced at the Institute of Microbiology, Chinese Academy of Sciences, Beijing, China.

2.2.4 Data processing

The following protocol was developed to process the sequence data in order to work through the proposed SIP in this study (Fig. 2-2), including trimming the sequences (forward and reverse), checking the validity, phasing heterozygous sequences, as well as consensus generation, and filtering.

Generally, the raw data of the forward and reverse sequences were trimmed individually by Sequencer 5.3 (Angermeier et al. 2012) under the setting: for the 5' end, trimming no more than 25%, trim until the first 50 bases contain less than 1 ambiguity; and for the 3' end, starting from 100 bases after the 5' trim, trim the first 50 bases containing more than 1 ambiguity. The trimmed forward and reverse sequences were aligned separately against the sequences in NCBI Genbank database (Benson et al. 2009). The sequences not belonging to Porifera and the ones with substandard lengths (<200bp) were excluded.

The heterozygosity was checked to separate the sequence pairs from each sponge species into three groups: (1) the pairs are homozygous (45%); (2) the pairs with their heterozygotes presenting a single double peak in the chromatogram (30%); (3) the pairs with multiple length-variant heterozygotes (25%) on either one or both of the sequence strands. For

the first two groups, the data can be edited easily following the protocol in Fig. 2-2. In group (3), the ambiguous bases were replaced with a mixed base symbol for nuclear marker phasing. The converted sequences were reconstructed to an optimal sequence of the corresponding two alleles (Flot et al. 2006) before running the online Basic Local Alignment Search Tool (BLAST) searches. The reconstruction was accomplished by a web tool (Champuru) available online at http://jfflot.mnhn.fr/champuru/ that automates the process (Flot 2007). The assembled consensus sequences derived from every sponge species were checked directly using BLAST searches (Altschul et al. 1990) against the NCBI GenBank database (Benson et al. 2009). Additionally, the results from the BLAST search, when consisted of less than 50% Coverage Region, were considered unreliable and excluded. In other words, the participating part of the sequence needs to represent more than half of the query.

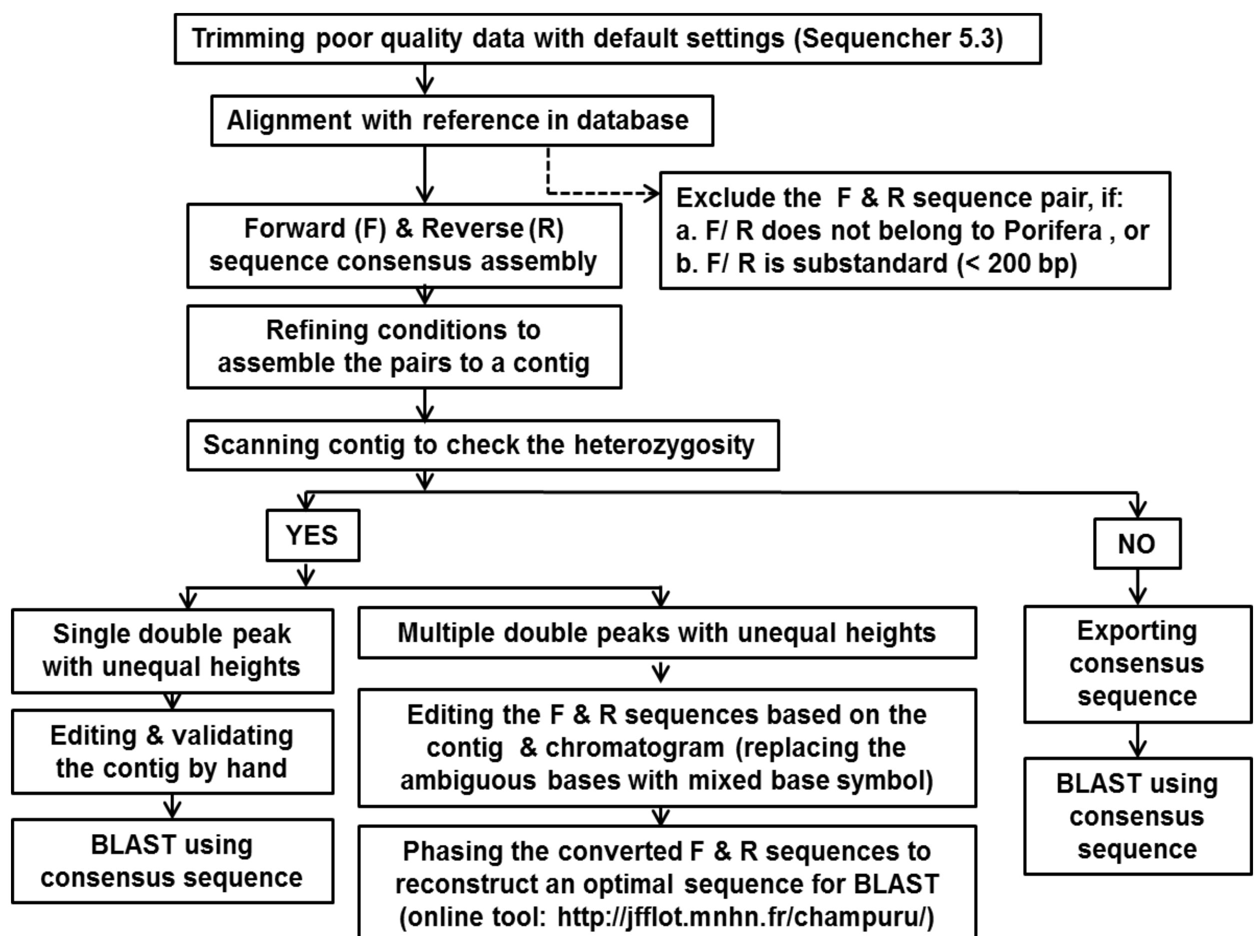


Figure 2 - 2 Sequence data processing flowchart

2.2.5 Development of a Sponge Identification Protocol (SIP)

The proposed Sponge Identification Protocol (SIP) is shown in Fig. 2-3. The molecular identities of each potential sponge species were inferred by SIP, a purely BLAST based non-phylogenetic sponge identification protocol. The application of the multiple locus strategy in SIP was an attempt to mitigate effects of the database limitations existing for single DNA marker based protocols to approach reliable sponge identification. Instead of building a phylogenetic tree separately, the three identities were referred to each other to infer the classification, as using any single DNA locus is restricted by its limited number of database submissions, resulting in inaccurate identifications. However, these limitations could not be fully avoided, due to unreliable morphologically identified submissions in the database.

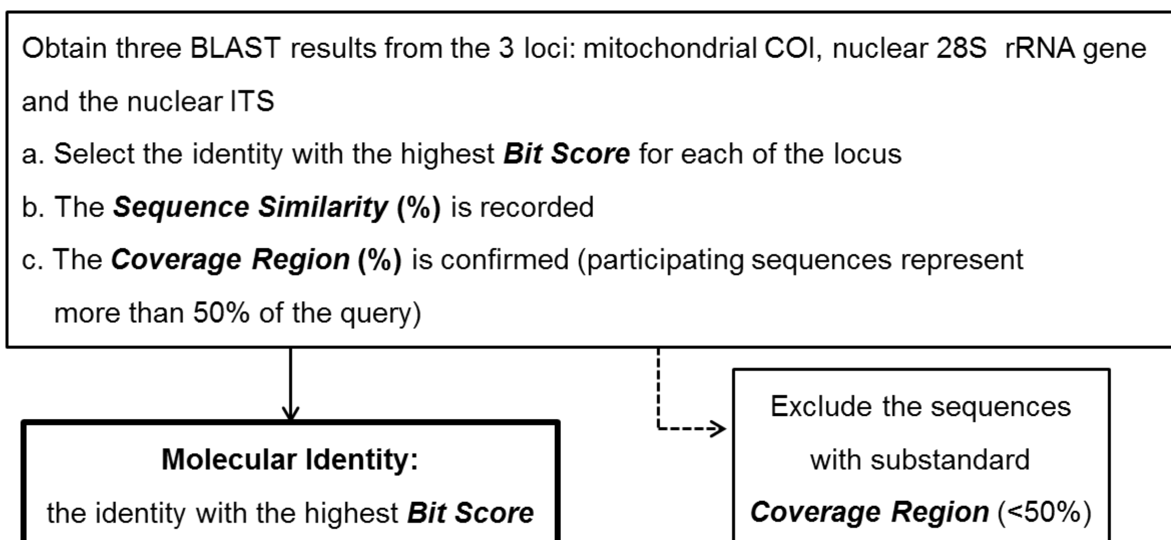


Figure 2 - 3 Sponge Identification Protocol (SIP) flow diagram used in this study

Reliable inferences require reliable statistical estimates. The Bit Score is a prominent statistical indicator used in addition to the E-value in a BLAST output (Pearson 2013). As a raw similarity score, Bit Score and E-value reflect the evolutionary distance of the two aligned sequences, the length of the sequences, and the scoring matrix used for the alignment. A Bit Score is normalised with respect to the scoring system and can be used to compare alignment scores from different searches. The higher the Bit Score, the more highly significant the match is. In contrast, sequence similarity (%) is not as sensitive or reliable. It is a useful approximation for analyses that depend on evolutionary distance, but evolutionary distance is not linear with percent similarity. The evolutionary distance associated with a 10% change in similarity is much

greater at longer distances. Here, the Bit Score and E-value are far more useful. However, the E-values showed 0.0 for most cases, which rules it out as a comparative indicator. Therefore, the Bit Score was selected as the priority identification parameter to identify sponges in this newly proposed SIP: (1) The one with the highest Bit Score in the BLAST result list was selected as the identity for every single DNA locus. It is important to consider the top 20 Blast results in the list for any possible errors, otherwise the one with highest Bit Score accompanied with the best E-value, sequence similarity, and percentage coverage was selected. (2) The one with the highest Bit Score among the three loci was chosen to be the final SIP identity. All of the consensus sequences used for providing the identities of different individuals were submitted into NCBI GenBank under the accession numbers KJ546351-546362, KJ546354-546368, KJ620376-620395, KJ620398-620409, KJ782592, KJ782595, KJ782600, KJ782602, KJ782604, KJ801654, KJ801656 and KJ801658-801661.

2.2.6 Phylogenetic analysis

The valid consensuses for each of the query sequence were submitted to the NCBI database to be searched by online Basic Local Alignment Search Tool (BLAST) (Altschul et al. 1990). The top 20 sequences were downloaded as a FASTA format file. An outgroup sequence belonging to different genus need to be added for the next step alignment. Software BioEdit (Hall 1999) was applied to combine the query sequence and the references in the FASTA file obtained from the previous BLAST search. Software MEGA6 (Galtier, Gouy & Gautier 1996) was used to align the sequences using ClustalW algorithm, and trim the aligned sequences to make them ready for tree construction. The alignment was exported as a MEGA format file. The Maximum Likelihood (ML) and Neighbor-Joining (NJ) methods were utilised to construct the phylogenetic trees for COI mtDNA and 28S rRNA gene separately. Both of the methods were set with 1000 of bootstrap replications. ML trees used Tamura-Nei model and NJ trees applied p-distance model.

2.2.7 Morphological re-examination and final identification

Guided by the molecular identification of SIP (the discrepancy between the SIP and initial morphological identities), a re-examination of the morphological features was conducted. The re-examination followed the protocol mentioned in section 2.2.1. The rules applied in this study were (1) if the re-examined morphological identity matched with the SIP identity at genus level, the SIP identity was chosen to the final identity; otherwise, (2) to check the morphological features of these sponges matched at the order or family level, if the morphological features are highly similar and difficult to discriminate, their SIP identity was selected as the final; (3) if they are not morphologically similar, a threshold of 98% for the sequence similarity was utilised to

determine the final identity. The SIP identity with ≥98% sequence similarity against the database entries was selected as the final identity, the others were assigned the final identity based on the morphological identification.

2.3 Results

2.3.1 Initial morphological classification

The 111 sponge individuals belonging to 37 potential species (three individuals for each species) were classified using morphological features. Table 2-1 shows the classifications, the sampling dates, and the locations. All the 37 sponges could be identified at the order/family level. Thirty-one sponges were identified to genus level and among of them nine were identified to species.

Table 2 - 1 Sponge morphological classification including sampling locations and collection dates

Museum Voucher	Order	Genus/ Species	Location	Date
SAMA S1991	Poecilosclerida	*Chondropsis* sp.	Rapid Bay	25/02/13
SAMA S1994		*Chondropsis* sp.	Rapid Bay	25/02/13
SAMA S1982		*Chondropsis* sp.	Outer Harbour	01/03/13
SAMA S1984		*Chondropsis* sp.	Outer Harbour	01/03/13
SAMA S1978		*Chondropsis* sp.	Klein Point	04/03/13
SAMA S1987		*Mycale* (*Zygomycale*) sp.	Outer Harbour	01/03/13
SAMA S1966		*Mycale* (*Arenochalina*) sp.	Williams Island	19/05/13
SAMA S1975		*Crella* sp. 1	Klein Point	04/03/13
SAMA S1977		*Crella* sp. 1	Klein Point	04/03/13
SAMA S1992		Poecilosclerid sp.	Rapid Bay	25/02/13
SAMA S1993		*Tedania* cf. *anhelans*	Rapid Bay	25/02/13
SAMA S1969		*Clathria* sp.	Williams Island	19/05/13
SAMA S1976	Tetractinellida	*Ecionemia* sp.	Klein Point	04/03/13
SAMA S1962		*Ecionemia* sp.	Williams Island	19/05/13
SAMA S1983		Geodiid sp.	Outer Harbour	01/03/13
SAMA S1963		Ancorinid sp.	Williams Island	19/05/13
SAMA S1968		Astrophorin sp.	Williams Island	19/05/13

SAMA S1996	Dictyoceratida	*Ircinia* sp.	Rapid Bay	25/02/13
SAMA S1974		*Ircinia* sp.	Williams Island	19/05/13
SAMA S1979		*Euryspongia* cf. *arenaria*	Klein Point	04/03/13
SAMA S1970		*Thorectandra* sp.	Williams Island	19/05/13
SAMA S1995	Axinellida	*Echinodictyum mesenterinum*	Rapid Bay	25/02/13
SAMA S1967		*Echinodictyum mesenterinum*	Williams Island	19/05/13
SAMA S1972		*Echinodictyum mesenterinum*	Williams Island	19/05/13
SAMA S1973	Verongiida	*Aplysina lendenfeldi*	Williams Island	19/05/13
SAMA S1985		Aplysinellid sp.	Outer Harbour	01/03/13
SAMA S1988		Verongiid sp.	Outer Harbour	01/03/13
SAMA S1971	Haplosclerida	*Callyspongia* (*Callyspongia*) sp.	Williams Island	19/05/13
SAMA S1986		*Chalinula* sp.	Outer Harbour	01/03/13
SAMA S1980		*Haliclona* sp.	Klein Point	04/03/13
SAMA S1960	Clionaida	*Cliona* sp.	Williams Island	19/05/13
SAMA S1961		*Spheciospongia* sp.	Williams Island	19/05/13
SAMA S1989	Dendroceratida	*Aplysilla rosea*	Rapid Bay	27/06/12
SAMA S1990		*Aplysilla rosea*	Rapid Bay	22/08/12
SAMA S1965	Suberitida	*Caulospongia* sp.	Williams Island	19/05/13
SAMA S1981		*Suberites* sp.	Outer Harbour	01/03/13
SAMA S1964	Tethyida	*Tethya* cf. *bergquistae*	Williams Island	19/05/13

2.3.2 Valid sequences

DNAs from the 111 specimens belonging to 37 species were successfully extracted. For PCR amplification, three duplicates of each species showed consistent performance. Ninety-three COI mtDNA amplicons for 31 sponge species, 75 amplicons of 28S rRNA gene for 25 species, and 96 amplicons of ITS region for 32 species were successfully obtained. All the amplicons were sequenced and eventually offered 93 COI mtDNA sequences derived from 31 species, 66 sequences of 28S rRNA gene for 22 species, and 81 ITS sequences for 27 species. The success rates of sequencing for the three DNA markers were 84%, 59%, and 73%, respectively. The alignment between the three duplicates of each species matched (>99% similarity) so that only one of the three was used for the following analysis.

The COI primers LCO1490 and HCO2198 yielded PCR products from 31 of the 37

potential sponge species (Table 2-2). All the PCR products were sequenced with sizes ranging from 672-699 bp. Two of the six sponges with no PCR products for COI locus had available sequence data from both 28S and ITS loci. The other four with sequence data only from the ITS locus were subjected to the alternative COI primer sets C1-J2165/C-Npor2760 and CO1porF1/CO1porR1 in a nested-PCR. However, no products were obtained. Implementing the data processing protocol in Fig. 2-2, 29 valid sequence pairs (forward and reverse) were obtained from the total of 31 successfully sequenced COI mtDNA amplicons (Table 2-2). Two were excluded as the forward sequence of one did not belong to Porifera, and the other had only 170 bp of the forward sequence (Appendix Table 2-2).

Table 2 - 2 Summary of molecular identification of 37 potential sponge species using multilocus approach

		COI mtDNA	Success rate	28S rRNA gene	Success rate	ITS region	Success rate
Data processing	**DNA preparation**	37	100%	37	100%	37	100%
	PCR products	31	84%	25	68%	32	86%
	Sequencing results	31	100%	22	88%	27	84%
	Valid results	29	94%	20	91%	12	44%
	Belong to Porifera	29	100%	20	100%	11	92%
SIP	**Putative identification**	11	-	19	-	4	-
	Putative different species	8	-	15	-	2	-

For the 28S rRNA gene, 25 PCR products were amplified from the 37 sponge species using primers NL4F/NL4R with sizes of 845-1227 bp, of which 22 were sequenced successfully (Table 2-2). Of the 15 unsuccessfully sequenced sponges, the identities of three had the data from both COI and ITS loci. The rest (12 sponges) were subjected to the second pair of 28S primers, RD3A/RD5B2, which yielded PCR products from four of the 12 sponges (Table 2-3). Two of the 22 successfully sequenced 28S amplicons were excluded as their reverse readings showed identities not belonging to Porifera (Appendix Table 2-2).

Table 2 - 3 Trial results of alternative primers for the 28S rRNA gene amplification on the failed sponges and the sampling locations

	SAMA S1990, S1993 Rapid Bay	SAMA S1985, S1988 Outer Harbour	SAMA S1975, S1977, S1978, S1979, S1980 Klein Point	SAMA S1969, S1970, S1974 Williams Island
Specimen working with alternative primers	SAMA S1990 SAMA S1993	-	-	SAMA S1969 SAMA S1974
Success rate	2/2	0/2	0/5	2/3
Overall successful rate			4/12	

The nuclear ITS amplicons of 32 sponge species were successfully amplified by the primer set, ITSRA2/ITS2.2, giving PCR products ranging from 334-1142 bp (Table 2-2). Of the 32 amplicons, 27 were successfully sequenced. After checking the percentage of the coverage region, 15 were excluded because the coverage percentage (the size of the sequence participating in the cluster analysis/ the size of the whole query sequence) was less than 50%. The other 12 sequence pairs included 11 valid ones and one non-Porifera sequence.

2.3.3 Identification using the proposed Sponge Identification Protocol (SIP)

After the data processing, 29, 20 and 11 valid sequence pairs, for which there was consensus between the sequence of their forward and reverse strands, were obtained, respectively, from the three DNA markers of the mitochondrial COI, the nuclear 28S rRNA gene and the nuclear ITS region (Table 2-2). The three loci derived from the same sponge were given equal weightage.

With the application of the proposed SIP in this study (Fig. 2-3), the closest inferences against the entries in the Genbank database were obtained, and assigned to 34 out of the 37 sponges using the valid sequence pairs as the initial SIP identities (Table 2-4). The sponges with no sequencing results for the COI locus were checked to see whether they had a matching identity from both 28S and ITS nuclear loci (Appendix Table 2-2). From the 29 valid COI sequences, 11 with the highest Bit Score among the three loci were used to infer the identities (Table 2-2). Eight of the 11 sponges belonged to eight distinct species in eight families of six orders, and the other three were identified as duplicate species (Table 2-4). For the 28S rRNA gene, of the 15 unsuccessfully sequenced sponges, the identities of three were inferred from their COI and ITS loci. Based on the 20 valid 28S rRNA gene sequences, 19 with the highest

Bit Score were inferred as 15 different species in 15 families in 10 orders (Table 2-2). Six of the 10 orders were already identified in the results from the COI mtDNA locus analysis whereas four more were identified (Clionaida, Haplosclerida, Dictyoceratida and Tethyida) (Table 2-4). In regard to the ITS locus, 11 valid sequences were obtained (Table 2-2), from which another four sponges were inferred as two different species. Both of the orders (Suberitida and Dictyoceratida) were also identified from other sponges by 28S rRNA gene. Family Halichondriidae in the order Suberitida, however, was only identified by its ITS locus (Table 2-4).

Overall, eighteen sponge species had valid sequencing data for both the COI mtDNA and 28S rRNA gene. Similarly, the COI and ITS loci were available for eight sponge species, and the 28S and ITS loci were available for six sponge species. Only five sponge species had sequencing results from all three DNA loci.

Table 2 - 4 Summary of molecular identification of 34 sponges using three DNA markers and following the sponge order in Table 2-1

Museum Voucher	Identification Result (Order; Family; Genus/Species) (ID locus, Bit score, % similarity, coverage %)	Accession No.
SAMA S1991	Poecilosclerida; Tedaniidae; *Tedania tubulifera* (28S, 992, 99%, 97%)	KJ620377
SAMA S1994	Poecilosclerida; Tedaniidae; *Tedania tubulifera* (28S, 1000, 99%, 96%)	KJ620378
SAMA S1982	Poecilosclerida; Tedaniidae; *Tedania tubulifera* (28S, 979, 98%, 98%)	KJ620381
SAMA S1984	Poecilosclerida; Tedaniidae; *Tedania tubulifera* (28S, 870, 95%, 92%)	KJ620384
SAMA S1978	Poecilosclerida; Desmacididae; *Desmapsamma anchorata* (COI, 581, 97%, 93%)	KJ546367
SAMA S1987	Poecilosclerida; Mycalidae; *Mycale setosa* (28S, 305, 84%, 55%)	KJ620385
SAMA S1966	Poecilosclerida; Mycalidae; *Mycale setosa* (28S, 1008, 99%, 85%)	KJ620392
SAMA S1975	Poecilosclerida; Hymedesmiidae; *Phorbas bihamiger* (COI, 469, 91%, 95%)	KJ546364
SAMA S1977	Poecilosclerida; Hymedesmiidae; *Phorbas bihamiger* (COI, 453, 91%, 90%)	KJ546366
SAMA S1993	Poecilosclerida; Desmacididae; *Desmapsamma anchorata* (COI, 614, 98%, 97%)	KJ546354
SAMA S1969	Dictyoceratida; Irciniidae; *Ircinia felix* f. *felix* (ITS, 358, 87%, 70%)	KJ801659
SAMA S1976	Suberitida; Halichondriidae; *Halichondria okadai* (ITS, 617, 94%, 99%)	KJ801656
SAMA S1962	Tetractinellida; Ancorinidae; *Ecionemia robusta* (COI, 1010, 99%, 99%)	KJ620388
SAMA S1983	Tetractinellida; Ancorinidae; *Tethyopsis mortenseni* (28S, 967, 98%, 97%)	KJ620383
SAMA S1963	Tetractinellida; Ancorinidae; *Stelletta clavosa* (28S, 1023, 99%, 97%)	KJ620389
SAMA S1968	Poecilosclerida; Microcionina; *Clathria rugosa* (COI, 510, 93%, 94%)	KJ620406
SAMA S1996	Dictyoceratida; Irciniidae; *Ircinia strobilina* (28S, 845, 94%, 94%)	KJ620380
SAMA S1974	Dictyoceratida; Irciniidae; *Ircinia felix* f. *felix* (ITS, 398, 83%, 100%)	KJ801661
SAMA S1979	Dendroceratida; Dictyodendrillidae; *Acanthodendrilla australis* (COI, 526, 97%, 84%)	KJ546368
SAMA S1970	Dictyoceratida; Irciniidae; *Ircinia felix* f. *felix* (ITS, 349, 81%, 99%)	KJ801660
SAMA S1995	Axinellida; Raspailiidae; *Raspailia vestigifera* (28S, 822, 96%, 82%)	KJ620379
SAMA S1967	Haplosclerida; Petrosiidae; *Petrosia lignosa* (28S, 826, 93%, 96%)	KJ620393
SAMA S1972	Axinellida; Raspailiidae; *Echinodictyum cancellatum* (COI, 467, 93%, 85%)	KJ620408
SAMA S1973	Verongiida; Aplysinidae; *Aplysina archeri* (28S, 1005, 98%, 97%)	KJ620395
SAMA S1985	Verongiida; Pseudoceratinidae; *Pseudoceratina* sp. (COI, 634, 99%, 95%)	KJ546361

SAMA S1988	Verongiida; Pseudoceratinidae; *Pseudoceratina* sp. (COI, 619, 99%, 93%)	KJ546363
SAMA S1971	Haplosclerida; Chalinidae; *Cladocroce* sp. (28S, 1085, 99%, 98%)	KJ620394
SAMA S1980	Suberitida; Suberitidae; '*Protosuberites*' sp. (COI, 404, 87%, 96%)	KJ620398
SAMA S1960	Clionaida; Spirastrellidae; *Spirastrella hartmani* (28S, 1012, 99%, 98%)	KJ620386
SAMA S1961	Poecilosclerida; Podospongiidae; *Diacarnus spinipoculum* (28S, 953, 99%, 92%)	KJ620387
SAMA S1989	Dendroceratida; Dictyodendrillidae; *Igernella notabilis* (28S, 977, 98%, 98%)	KJ620376
SAMA S1965	Suberitida; Halichondriidae; *Hymeniacidon heliophila* (28S, 1004, 99%, 96%)	KJ620391
SAMA S1981	Suberitida; Suberitidae; *Suberites aurantiacus* (28S, 975, 98%, 97%)	KJ620381
SAMA S1964	Tethyida; Tethyidae; *Tethya* sp. (28S, 973, 98%, 98%)	KJ620390

2.3.4 SIP reliability validated by phylogenetic analysis

In order to validate the reliability of the proposed SIP for sponge identification, the phylogenetic analysis was compared with the BLAST and the SIP identities (Table 2-5). The sponges have a cut-off sequence similarity of 96% as the sequences with lower similarity could not be aligned with other reference sequences to compute the valid phylogenetic relationship for identification. Furthermore, considering the minimum coverage percentage to have a valid sequence alignment, all of the ITS loci were excluded from the phylogenetic analysis due to their lower coverages.

The analysis illustrated that the identities of all sponges concluded from SIP matched with the phylogenetic analysis for both DNA markers, even though the BLAST results of the two loci were different (Appendix Fig. 2-1 and 2-2). For example, sponge SAMA S1981 was inferred to have the closest match with *Suberites aurantiacus* by SIP with the highest Bit Score among the three loci. The identity inferred from the BLAST result of 28S locus was different from the BLAST of COI locus, but they shared the same sequence similarity (98%). Using the phylogenetic analysis, it was demonstrated that the query sequence showed closer relationship with the species *Suberites aurantiacus* in both of the trees for COI and 28S loci based on the Maximum Likelihood (Fig. 2-4a, b) and Neighbor Joining methods (Appendix Fig. 2-1 and 2-2), which validated the identity inferred by SIP. Similarly, sponge SAMA S1962 was inferred to have the closest match with *Ecionemia robusta* from the COI locus by SIP and confirmed by the phylogenetic analysis of COI (Fig. 2-4d) and 28S loci (Fig. 2-4c) at the genus level. The phylogenetic tree for 28S locus showed that the query was closer to *Ecionemia acervus*, in the same genus as the SIP identity, instead of the BLAST result *Stelletta clavosa*. Checking the database, there was no 28S rRNA gene entry for the species *Ecionemia robusta*. For sponges (e.g. sponge SAMA S1960) where the same genus does not exist in the both COI and 28S databases, the trees showed the same phylogenetic status of the query sequence (Fig. 2-4e, f).

Table 2 - 5 Reliability of SIP validated by phylogenetic analysis

	BLAST result		Inference from SIP	Phylogenetic analysis- Maximum Likelihood		Phylogenetic analysis- Neighbor Joining	
	COI identity (Bit Score, % similarity)	**28S identity (Bit Score, % similarity)**		**COI locus**	**28S locus**	**COI locus**	**28S locus**
SAMA S1981	*Rhizaxinella* sp. (608, 98%)	*Suberites aurantiacus* (975, 98%)	*Suberites aurantiacus*	*Suberites aurantiacus*	*Suberites aurantiacus*	*Suberites aurantiacus*	*Suberites aurantiacus*
SAMA S1963	*Ecionemia* sp. (641, 99%)	*Stelletta clavosa* (1023, 99%)	*Stelletta clavosa*	*Stelletta clavosa*	*Stelletta clavosa*	*Stelletta clavosa*	*Stelletta clavosa*
SAMA S1965	*Protosuberites* sp. (598, 97%)	*Hymeniacidon heliophila* (1004, 99%)	*Hymeniacidon heliophila*	*H. perlevis; H. helliophila; H. sinapium*	*Hymeniacidon heliophila*	*H. perlevis; H. helliophila; H. sinapium*	*Hymeniacidon heliophila*
SAMA S1973	*Aplysina lacunose* (611, 98%)	*Aplysina archeri* (1005, 98%)	*Aplysina archeri*	*Aplysina archeri*	*Aplysina archeri*	*Aplysina archeri*	*Aplysina archeri*
SAMA S1989	*Igernella notabilis* (629, 98%)	*Igernella notabilis* (977, 98%)	*Igernella notabilis*	*Igernella notabilis*	*Igernella notabilis*	*Igernella notabilis*	*Igernella notabilis*
SAMA S1961	*Diacarnus spinipoculum* (657, 99%)	*Diacarnus spinipoculum* (953, 99%)	*Diacarnus spinipoculum*	*Diacarnus spinipoculum*	*Diacarnus spinipoculum*	*Diacarnus spinipoculum*	*Diacarnus spinipoculum*
SAMA S1962	*Ecionemia robusta* (1010, 99%)	*Stelletta clavosa* (642, 99%)	*Ecionemia robusta*	*Ecionemia robusta*	*Ecionemia acervus*	*Ecionemia robusta*	*Ecionemia acervus*
SAMA S1991	*Desmapsamma anchorata* (612, 98%)	*Tedania tubulifera* (992, 99%)	*Tedania tubulifera*	*T. ignis; T. klausi*	*Tedania tubulifera*	*T. ignis; T. klausi*	*Tedania tubulifera*
SAMA S1982	*Desmapsamma anchorata* (615, 99%)	*Tedania tubulifera* (979, 98%)	*Tedania tubulifera*	*T. ignis; T. klausi*	*Tedania tubulifera*	*T. ignis; T. klausi*	*Tedania tubulifera*
SAMA S1994	*Tedania ignis* (604, 98%)	*Tedania tubulifera* (1000, 99%)	*Tedania tubulifera*	*Tedania ignis*	*Tedania tubulifera*	*Tedania ignis*	*Tedania tubulifera*
SAMA S1966	*Mycale mirabilis* (633, 99%)	*Mycale setosa* (1008, 99%)	*Mycale setosa*	*Mycale mirabilis*	*Mycale setosa*	*Mycale mirabilis*	*Mycale setosa*
SAMA S1960	*Clionaopsis platei* (605, 97%)	*Spirastrella hartmani* (1012, 99%)	*Spirastrella hartmani*	*Cliona chilensis*	*Spirastrella hartmani*	*Cliona chilensis*	*Spirastrella hartmani*
SAMA S1971	*Callyspongia siphonella* (581, 98%)	*Cladocroce* sp. (1085, 99%)	*Cladocroce* sp.	*Callyspongia siphonella*	*Cladocroce* sp.	*Callyspongia siphonella*	*Cladocroce* sp.
SAMA S1983	*Ancorina* sp. (600, 98%)	*Tethyopsis mortenseni* (967, 98%)	*Tethyopsis mortenseni*	*Pleroma menoui*	*Tethyopsis mortenseni*	*Pleroma menoui*	*Tethyopsis mortenseni*

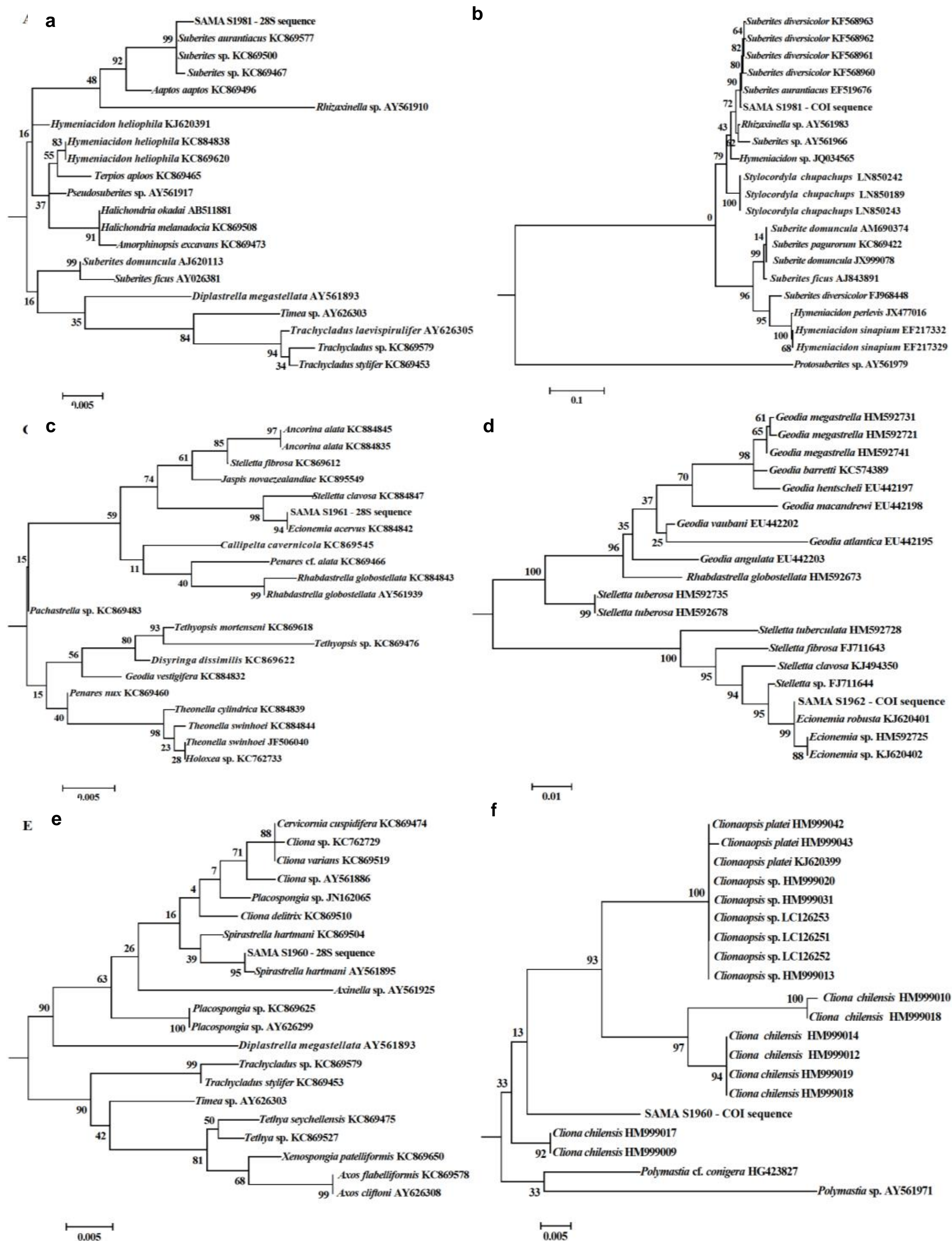


Figure 2 - 4 Phylogenetic relationship of three representative sponge species using the Maximum Likelihood method based on 28S rRNA gene and COI mtDNA to validate the sponge identification by SIP. a. Phylogenetic relationship of sponge SAMA S1981 based on 28S rRNA gene. b. Phylogenetic relationship of sponge SAMA S1981 based on COI mtDNA. c. Phylogenetic relationship of sponge SAMA S1962 based on 28S rRNA gene. d. Phylogenetic relationship of sponge SAMA S1962 based on COI mtDNA. e. Phylogenetic relationship of sponge SAMA S1960 based on 28S rRNA gene. f. Phylogenetic relationship of sponge SAMA S1960 based on COI mtDNA.

2.3.5 Re-examination of morphological identifications

There were only nine sponge species identified to the same genus by the SIP and morphological identifications. To resolve the discrepancy, the 34 sponge species inferred by SIP (Table 2-4) were re-examined based on their morphological features (Table 2-6). Morphological descriptions of the 34 sponge species are documented in Appendix Table 2-1. Twenty-seven sponge species were assigned to the same classifications as the initial ones based on the available morphological characters. Based on both of the morphological and the SIP identification, the 27 sponge species were divided into four categories. Category I includes nine sponge species of which the re-examined identities matched with the SIP identities at genus level. Category II includes seven sponge species which matched only at the order or family level, but where the orders or families have highly similar morphological features. Category III has eight sponge species where the species matched only at the order level and Category IV is for three species where sponge did not match even at the order level.

Importantly, the initial morphological identities of seven sponge species were corrected and revised (category V). Of these, three re-examined morphological identities matched with the SIP identities at the genus level, and another three matched at the order level. One sponge species did not match at the order level.

Table 2 - 6 Re-examination of morphological classification and comparison with SIP identification

Museum Voucher	Initial morphological classification	Re-examination of morphological classification	SIP identification	Notes on difference between morphological and molecular identities
Category I: Genus level match between morphology and SIP				
SAMA S1962	*Ecionemia* sp.	*Ecionemia* sp.	*Ecionemia robusta* (1010, 99%, 99%)	*E. robusta* is now accepted as *Ancorina robusta*, occurs in South Australia (SA). This is possibly a species match.
SAMA S1964	*Tethya* cf. *bergquistae*	*Tethya* cf. *bergquistae*	*Tethya* sp. (973, 98%, 98%)	
SAMA S1966	*Mycale* (*Arenochalina*) sp.	*Mycale* (*Arenochalina*) sp.	*Mycale setosa* (1008, 99%, 85%)	*M. setosa* is a red sea sponge.
SAMA S1972	*Echinodictyum mesenterinum*	*Echinodictyum mesenterinum*	*Echinodictyum cancellatum* (467, 93%, 85%)	Both *E. mesenterinum* and *E. cancellatum* occur in SA, but their morphology is distinct.

SAMA S1973	*Aplysina lendenfeldi*	*Aplysina lendenfeldi*	*Aplysina archeri* (1005, 98%, 97%)	*A. archeri* (Higgin, 1875) is a Caribbean yellow tubular sponge; it is very similar in appearance to Australia's *A. lendenfeldi* (Bergquist, 1980).
SAMA S1974	*Ircinia* sp.	*Ircinia* sp.	*Ircinia felix* f. *felix* (398, 83%, 100%)	*I. felix* is Caribbean/Brazilian sponge.
SAMA S1981	*Suberites* sp.	*Suberites* sp.	*Suberites aurantiacus* (975, 98%, 97%)	*S. aurantiacus* is a Caribbean sponge.
SAMA S1987	*Mycale* (*Zygomycale*) sp.	*Mycale* (*Zygomycale*) sp.	*Mycale setosa* (305, 84%, 55%)	*M. setosa* is a red sea sponge.
SAMA S1996	*Ircinia* sp.	*Ircinia* sp.	*Ircinia strobilina* (845, 94%, 94%)	*I. stobilina* is a Caribbean/Brazilian sponge.
Category II: Order/ family level match between morphology and SIP; highly similar each other				
SAMA S1960	*Cliona* sp.	*Cliona* sp.	*Spirastrella hartmani* (1012, 99%, 98%)	*S. hartmani* (Boury-Esnault, Klautau, Bézac, Wulff & Solé-Cava, 1999) is a Caribbean/ Brazillian sponge. The families Clionaidae and Spirastrellidae are close, which contain *Cliona* and *Spirastrella*, respectively. Historically, some genera moving from one to another (Hooper, J.N.A. & Soest 2002).
SAMA S1971	*Callyspongia* (*Callyspongia*) sp.	*Callyspongia* (*Callyspongia*) sp.	*Cladocroce* sp. (1085, 99%, 98%)	*Callyspongia* (Callyspongiidae) and *Cladocroce* (Chalinidae) are in different families.
SAMA S1975	*Crella* sp. 1	*Crella* sp. 1	*Phorbas bihamiger* (469, 91%, 95%)	*P. bihamiger* (Waller, 1878) is a green encrusting UK/North Atlantic sponge. The families Crellidae (containing *Crella*) and Hymedesmiidae (containing *Phorbas*) are very close differing in only the surface arrangement of spicules (Hooper, J.N.A. & Soest 2002).
SAMA S1977	*Crella* sp. 1	*Crella* sp. 1	*Phorbas bihamiger* (453, 91%, 90%)	As above.
SAMA S1978	*Chondropsis* sp.	*Chondropsis* sp.	*Desmapsamma anchorata* (581, 97%, 93%)	*Desmapsamma* (Desmacididae) and *Chondropsis* (Chondropsidae) are both Poecilosclerid, sand-bearing sponges, but in different families.

SAMA S1989	*Aplysilla rosea*	*Aplysilla rosea*	*Igernella notabilis* (977, 98%, 98%)	*I. notabilis* (Duchassaing & Michelotti, 1864) is a fleshy pink conulose Caribbean sponge. *A. rosea* (Barrois, 1876) is a fleshy pink conulose sponge from NE Atlantic region and the Mediterranean area. Van Soest (2004) stated that other records of *A. rosea* are inaccurate, although it is listed in Australian Faunal Directory and commonly cited.
SAMA S1995	*Echinodictyum mesenterinum*	*Echinodictyum mesenterinum*	*Raspailia vestigifera* (822, 96%, 82%)	*E. mesenterinum* and *R. vestigifera* share the same family.
Category III: Order level match between morphology and SIP				
SAMA S1963	Ancorinid sp.	Ancorinid sp.	*Stelletta clavosa* (1023, 99%, 97%)	*S. clavosa* does occur in SA.
SAMA S1965	*Caulospongia* sp.	*Caulospongia* sp.	*Hymeniacidon heliophila* (1004, 99%, 96%)	
SAMA S1982	*Chondropsis* sp.	*Chondropsis* sp.	*Tedania tubulifera* (979, 98%, 98%)	
SAMA S1983	Geodiid sp.	Geodiid sp.	*Tethyopsis mortenseni* (967, 98%, 97%)	
SAMA S1984	*Chondropsis* sp.	*Chondropsis* sp.	*Tedania tubulifera* (870, 95%, 92%)	
SAMA S1991	*Chondropsis* sp.	*Chondropsis* sp.	*Tedania tubulifera* (992, 99%, 97%)	
SAMA S1992	*Tedania* cf. *anhelans*	*Tedania* cf. *anhelans*	*Desmapsamma anchorata* (614, 98%, 97%)	
SAMA S1994	*Chondropsis* sp.	*Chondropsis* sp.	*Tedania tubulifera* (1000, 99%, 96%)	
Category IV: Not match at order level				
SAMA S1961	*Spheciospongia* sp.	*Spheciospongia* sp.	*Diacarnus spinipoculum* (953, 99%, 92%)	*D. spinipoculum* does occur in SA.
SAMA S1968	Astrophorin sp.	Astrophorin sp.	*Clathria rugosa* (510, 93%, 94%)	Possible contamination by an encrusting sponge.
SAMA S1980	*Haliclona* sp.	*Haliclona* sp.	'*Protosuberites*' sp. (404, 87%, 96%)	

Category V: Incorrect initial morphological classification or classification refined further on re-examination				
SAMA S1985	Aplysinellid sp.	*Pseudoceratina* sp.	*Pseudoceratina* sp. (634, 99%, 95%)	Updated ID match at genus level
SAMA S1988	Verongid sp.	*Pseudoceratina* sp.	*Pseudoceratina* sp. (619, 99%, 93%)	Updated ID match at genus level
SAMA S1979	*Euryspongia* cf. *arenaria*	*Acanthodendrilla* sp.	*Acanthodendrilla australis* (526, 97%, 84%)	Updated ID match at genus level
SAMA S1967	*Echinodictyum mesenterinum*	*Callyspongia bilamellata*	*Petrosia lignosa* (826, 93%, 96%)	Updated ID match at order level
SAMA S1969	*Clathria* sp.	Spongiid sp.	*Ircinia felix* f. *felix* (358, 87%, 70%)	Updated ID match at order level
SAMA S1970	*Thorectandra* sp.	Thorectid sp.	*Ircinia felix* f. *felix* (349, 81%, 99%)	Updated ID match at order level
SAMA S1976	*Ecionemia* sp.	*Chondrosia* sp.	*Halichondria okadai* (617, 94%, 99%)	

2.3.6 Approach to the final identity and the improved discrimination

The differences between morphological classifications and SIP identifications occurred at various taxonomic levels (Table 2-6). The final identities were concluded by comprehensively considering the confidence level for both of the molecular and morphological identifications and followed the rules in section 2.2.7.

In Category I (Table 2-6), the final identities of the nine sponge species was based on SIP as their identities generated by SIP matched the morphological identification at the genus level. In particular, the species *Ecionemia robusta* (SAMA S1962), inferred by SIP, has been reported in South Australia. It was possible to achieve even a species-level identity. In Category II, the final identities of the seven sponge species followed the SIP identification due to their highly similar morphological features being difficult to discriminate. For example, the SIP identity of the sponge SAMA S1989 was *Igernella notabilis* (Duchassaing & Michelotti, 1864), a fleshy pink conulose Caribbean sponge. The morphological ID was *Aplysilla rosea* (Barrois, 1876), a fleshy pink conulose sponge from NE Atlantic region and the Mediterranean area. van Soest (2004) stated that other records of *A. rosea* are inaccurate, although *A. rosea* is listed in Australia (Australian Faunal Directory database) and commonly cited. Due to the need to revise

the genus *Aplysilla*, it is possible for sponge SAMA S1989 to be corrected as *Igernella*.

In terms of the 18 sponges in categories III, IV, and V, their final identities were determined by the same rule by combined consideration of both SIP and morphological identifications. Three sponge species matched identities at the genus level (Category V: SAMA S1979, S1985, and S1988). Therefore, their SIP identities were the final identification. The remaining 15 sponge species had mismatched identities between morphological and SIP identifications at the family or order levels. Of those, eight SIP identities were taken as the final identities based on the threshold of 98% sequence similarity. The other seven were given their final identities based on the morphological identification.

Overall, 27 of the 34 sponge species were finally identified by SIP and seven by morphology. The final identities were assigned to the genes of the three loci amplified from the 34 sponge species and submitted into the NCBI database.

For the morphologically ambiguous sponges that were defined as the ones sharing the same genera/ species, the discrimination could be improved by SIP accompanied with the re-examination of the morphology. For example, five sponges identified as *Chondropsis* sp. in initial morphological examination were distinguished into three species based on re-examination after SIP identification: *Tedania* sp. (SAMA S1982, S1991, and S1994), *Desmapsamma* sp. (SAMA S1978), and *Chondropsis* sp. (SAMA S1984).

2.3.7 SIP reliability evaluated by morphological identification

To mitigate the limitation of the incomplete gene database, the sponge identities inferred from any two of the DNA loci were referred to each other to confirm an identity. The comparison with the re-examined morphological classification was to evaluate the reliability of the identification. Specifically, the sequence information of 18 sponges (Appendix Table 2-2) with both of the COI mtDNA and 28S rRNA gene loci were compared with the revised morphological identifications at the genus, family, and order levels. There were better matches at the higher taxonomic rank with a maximum matching rate of 94% at the order level (Table 2-7). When the locus used to identify the sponges was compared with the morphological classification, the order level matching rate (%) remained the same (94%); however, the matching rates at the family and genus levels were noticeably higher. For sponges with at least two valid sequencing results, the three DNA markers supported each other as evidenced by the high matching rate at the order, family, and genus levels (91%, 73%, and 45%, respectively).

Table 2 - 7 Comparison of molecular identifications by selected DNA marker(s) with or without morphological classifications

	COI and 28S Loci & Morphological classification			One locus used for inferring identity & Morphological classification			Any two loci matching each other		
	Genus match	Family match	Order match	Genus match	Family match	Order match	Genus match	Family match	Order match
Number of sponges	3	6	17	6	8	17	10	16	20
Percentage	3/18	6/18	17/18	6/18	8/18	17/18	10/22	16/22	20/22
%	17%	33%	94%	33%	44%	94%	45%	73%	91%
	Single locus & morphological classification								
	ITS: 11 valid sequences			COI: 29 valid sequences			28S: 20 valid sequences		
	Genus match	Family match	Order match	Genus match	Family match	Order match	Genus match	Family match	Order match
Number of sponges	2	3	10	9	13	27	6	9	19
Percentage	2/11	3/11	10/11	9/29	13/29	27/29	6/20	9/20	19/20
%	18%	27%	90%	31%	44%	93%	30%	45%	95%
	Identity sequence similarity ≥98%								
	ITS			COI			28S		
	Genus match	Family match	Order match	Genus match	Family match	Order match	Genus match	Family match	Order match
Number of sponges	0	0	0	6	8	17	4	6	14
Percentage	-	-	-	6/9	8/13	17/27	4/6	6/9	14/18
%	0%	0&	0%	66%	61%	62%	67%	67%	78%

2.3.8 Suitability and capacity of three DNA markers for sponge identification

The resolution of the three DNA loci to identify sponges was analysed by separately comparing the matching rate between the molecular information of each locus and the re-examined morphological classification for a given sponge species (Table 2-7). The molecular identification results for 19 out of the 20 valid 28S rRNA sequence pairs matched the morphological classification at the order level, nine at the family level, and six at the genus level. Similarly, the COI mtDNA and ITS sequence data using the 29 and 11 valid sequencing results, respectively, were analysed. Twenty-seven among the 29 for COI locus and 10 out of the 11 for ITS locus matched at the order level. The 28S rRNA locus showed the highest matching rate (%) among these three loci at the order and family levels. For the genus level, COI and 28S loci showed similar performance but significantly better than ITS locus. Correlating the morphological classification with the molecular identification with sequence similarities ≥98%, more sponge species inferred by COI locus than 28S locus matched with morphological identity at the order, family, and genus levels. However, 28S locus had higher matching rate than COI

locus at all three taxonomic levels due possibly to its less number of the valid sequences. Notably, the ITS locus showed a significantly lower matching rate than the other two loci in all these cases.

2.3.9 Limited entries in database

The limited number of the ITS submissions in the reference databases often results in a low Bit Score and sequence similarity as well as the potential for misidentification. To validate this hypothesis, we enumerated the total number of submissions in the NCBI database for COI mtDNA, 28S rRNA gene, and ITS belonging to the 10 orders of Porifera identified in this study (Table 2-8). The number of submissions associated with the ITS region was the lowest of the three loci. A total of 964 submissions for the ITS locus included 367 sequences covering 26 different species from the order Clionaida that has the largest contribution of the 10 orders. In contrast, the smallest contribution was for the order Tethyida with only 11 accessions.

Table 2 - 8 The submissions of the three DNA loci based on different Porifera orders in the NCBI database

Porifera Order	COI mtDNA Submissions	28S rRNA gene Submissions	ITS region Submissions
Haplosclerida	717	374	15
Poecilosclerida	280	502	27
Tetractinellida	237	272	27
Dictyoceratida	127	292	159
Suberitida	98	178	159
Verongiida	87	177	147
Tethyida	30	77	11
Axinellida	28	258	37
Clionaida	26	379	367
Dendroceratida	6	31	15
Total	1636	2540	964

For the sponge species belonging to the orders Haplosclerida (e.g. SAMA S1971), Poecilosclerida (e.g. SAMA S1982), and Tetractinellida (e.g. SAMA S1983), the COI mtDNA and 28S rRNA loci had a much higher Bit Score and sequence similarity (≥98%) than the ITS locus (Appendix Table 2-2). Table 2-9 shows some sponge species as examples in the orders Haplosclerida, Poecilosclerida, and Tetractinellida having the varying numbers of the submissions associated with the three DNA loci in the database.

Table 2 - 9 The nucleotide submissions from different sponge families, genera and species [a]

Porifera Order	Family	Genus	Species	COI mtDNA nucleotide No.	28S rRNA gene nucleotide No.	ITS region nucleotide No.
Haplosclerida	Callyspongiidae	*Callyspongia*	*Callyspongia siphonella*	1	1	0
			Other species	561	74	0
		Other genera	-	1	5	0
	Chalinidae	*Cladocroce*	*Cladocroce* sp.	0	2	0
			Other species	1	3	0
		Haliclona	*Haliclona* sp.	4	23	2
			Other species	37	30	5
		Other genera	-	0	26	0
Poecilosclerida	Desmacididae	*Desmapsamma*	*Desmapsamma anchorata*	7	2	0
			Other species	0	0	0
		Other genera	-	0	0	0
	Tedaniidae	*Tedania*	*Tedania tubulifera*	0	1	0
			Tedania ignis	12	3	1
			Other species	6	3	0
		Other genera	-	0	0	0
Tetractinellida	Ancorinidae	*Ancorina*	*Ancorina* sp.	1	1	0
			Other species	0	2	0
		Tethyopsis	*Tethyopsis mortenseni*	0	1	0
			Other species	0	1	0
		Other genera	-	31	45	3
	Geodiidae	*Pachymatisma*	*Pachymatisma johnstonia*	12	9	7
			Other species	8	6	5
		Other genera	-	86	41	2

[a] All the data are from NCBI database searched by May 2016.

2.4 Discussion

Primer selection is a crucial part of PCR amplification to conduct the proposed SIP for sponge identification. Particularly for marine sponges, no single pair of primers could amplify the desired gene from all these sponge species (Tables 2-2 and 2-4). The quality of the sponge specimens is another essential factor for successful amplification. The degradation or contamination of the sponge specimens may result in a failure of gene amplification (Table 2-3: SAMA S1975, S1977, S1978, S1979, and S1980). Consistent with the previous study (Redmond & McCormack 2009), the results here indicate that identification using the ITS locus has much less value than the other loci. The ITS region has a higher evolutionary rate than the other loci (Liao 1999) and insufficient database entries. The advantage of ITS locus is the variety in size due to indels which may be informative at some level. Consequently, using secondary structure to guide the alignment may also help in this situation.

The Sponge Identification Protocol (SIP) developed in this study was tested and evaluated with 111 individual sponges. The outcomes revealed the problems and challenges inherent in the current molecular identification process using a single molecular marker. The reliability of the developed SIP has been validated and well matched by phylogenetic analysis (Table 2-5 and Appendix Fig. 2-1, 2-2). The character-based (Maximum Likelihood) phylogenetic trees demonstrated that the final inferences from SIP matched exactly with the identities inferred from each of the 28S rRNA gene and COI mtDNA phylogenetic trees, if the databases for the two DNA loci have the matching entries. Otherwise, the identity with the higher Bit Score always had a closer phylogenetic relationship with the query sequence based on the trees for the COI and 28S loci. In the distance-based (Neighbor Joining) phylogenetic trees, the results obtained were consistent. All of these analyses implied that the incomplete database entries not only limit the BLAST search but also restrict the application of phylogenetic analysis for sponge identification, unless a multi-locus approach such as SIP, developed in this study, can be utilised.

In the SIP, applying the normalised similarity score (Bit Score) as the identification indicator, in conjunction with the proper multiple DNA markers, was proven to be a highly efficient and reliable sponge identification approach. The combination of the three loci not only offered a broader coverage of identifiable sponges (92%) but also improved the reliability when the identities from two or more loci matched each other. Compared to the identification performance of the single marker in SIP (COI mtDNA: 11; 28S rRNA: 19; ITS region: 4; refer to Supplementary Table 2-2), the multilocus approach substantially increased the identification efficacy (Table 2-4). Thirty-four sponges were classified following the established SIP. Only

three sponges failed to be identified due either to no PCR amplicon, substandard Coverage Region, or a non-Porifera sequence. More than 50% of sponges had a ≥ 98% sequence similarity. When the sequence similarity was ≥ 90%, 85% of the sponges could be classified. Without SIP, the BLAST results from any one of the three loci were: COI mtDNA: 29; 28S rRNA: 20; ITS region: 11 (refer to Table 2-2). However, most of these results for the COI mtDNA (18 out of 29) and ITS region (7 out of 11) were unable to infer the sponge identity due to their lower similarities or lower sequence quality (Appendix Table 2-2).

Mismatches between morphological and molecular identifications are common in sponges and fall across a spectrum of discordance (DeBiasse & Hellberg 2015). Re-examination of the morphological features after the molecular identification is essential to approach a more reliable identity when considering the limitations of both of the molecular and morphological approaches. On the one hand, the Genbank data are limited by the number and the diversity of entries. The reliability of the identities of the genes submitted to the database also greatly impacts the identification. On the other hand, there are many sponge taxa with little or no spicule diversity at all for species level discrimination (de Paula et al. 2012). Some different families even showed high degree of overlap of the spicules (de Paula et al. 2012). Particularly, for some local sponge species (e.g. South Australian sponges), it is difficult to identify them into the species level solely using the morphological features, without the accepted type species available. In our study, a threshold of 98% sequence similarity was selected to determine a more reliable identity to comprehensively consider the accuracy of both of the morphological and molecular identifications. This 98% threshold may sound arbitrary, but this level of confidence is required for a robust SIP identification. In conjunction with the morphological re-examination, the SIP offered more reliable identities.

A multilocus approach has been employed by other researchers in different combinations with mixed results (Blanquer & Uriz 2007; Erpenbeck et al. 2004; Nichols 2005; Redmond et al. 2011). In addition, the amount of congruence observed between molecular and morphological data sets varied among different types of molecular data and sponge taxa (Alvarez et al. 2000; Usher et al. 2004). In this study, when compared with morphological classification (Table 2-7), the molecular identification had a high matching rate at the order level for both of these two categories: the three element matching (COI mtDNA, 28 rRNA, and morphological classification) and the two element matching (the locus used to infer the identity and morphological classification). Notably, the latter category increased at the family and genus levels. Additionally, the matching rates of the identities from any two DNA markers were much higher than the previous two categories at the family and genus levels (Table 2-7). Therefore, the reliability of the developed SIP (molecular loci only) was further validated, as the main

principle of SIP is also to utilise one of the three loci with the highest similarity (Bit Score).

In contrast to the phylogenetic approach, the multilocus-based BLAST protocol greatly improves the quality of inferences by coupling the accuracy present in different DNA markers. The resolution of the DNA markers was jointly demonstrated with the morphological classifications in three aspects: accuracy, reliability and suitability (Table 2-7). In our study, the nuclear 28S rRNA gene was found to have higher resolution than mitochondrial COI gene and nuclear ITS region. However, the COI locus is common for sponges as it offered the largest number of valid sequences. In some cases, such as the sponges in Poecilosclerida, the COI locus showed a better resolution at the order level (Appendix Table 2 - 3).

Notably, limited database submissions can skew the identification accuracy regardless of the BLAST approach or the phylogenetic tree method. Caution should be exercised in inferring a new species or genus when the sequence similarity was lower than 98% using any one locus alone (Table 2-8). The nucleotide database for ITS locus is not only limited by the small number of submissions, but also the lack of species diversity. For the order Clionaida with the largest ITS submissions among the 10 sponge orders in this study, the numbers of nucleotides associated with COI, 28S, and ITS loci are 26, 379 and 367, respectively (Tables 2-8). However, the ITS locus only covered 26 species compared to the 28S loci covering more than 170 species (Benson et al. 2009). The application of multiple loci reflected the limitation of a single locus based approach and highlighted the necessity to consider the diversity of the database entries when identifying the sponges.

2.5 Conclusion

Overall, the results indicate that the application of multiple loci is essential for both of the SIP and phylogenetic approach to achieve a level of confidence for sponge identification due to the limited gene database entry of a single DNA marker. The SIP, in which the various resolutions of three DNA markers complement each other, was completely validated by phylogenetic analysis. SIP is more effective and practical than building different trees for one query sequence identification. Re-examination of the morphological identification guided by SIP identities leads to revised sponge identities, which demonstrates the better reliability of SIP. In conclusion, the multilocus-based SIP integrated with morphological identification offers an improved protocol that is effective for more reliable sponge identification.

References

Adlard, RD & Lester, RJG 1995, 'The Life-Cycle and Biology of Anilocra-Pomacentri (Isopoda, Cymothoidae), an Ectoparasitic Isopod of the Coral-Reef Fish, Chromis-Nitida (Perciformes,

Pomacentridae)', *Australian Journal of Zoology*, vol. 43, no. 3, pp. 271 - 81

Altschul, SF, Gish, W, Miller, W, Myers, EW & Lipman, DJ 1990, 'Basic local alignment search tool', *Jounal of Molecular Biology*, vol. 215, pp. 403-10.

Alvarez, B, Crisp, MD, Driver, F, Hooper, JNA & Soest, RWMv 2000, 'Phylogenetic relationships of the family Axinellidae (Porifera: Demospongiae) using morphological and molecular data', *Zoologica Scripta*, vol. 29, pp. 169-98.

Angermeier, H, Glöckner, V, Pawlik, JR, Lindquist, NL & Hentschel1, U 2012, 'Sponge white patch disease affecting the Caribbean sponge *Amphimedon compressa*', *Diseases of Aquatic Organisms*, vol. 99, pp. 95–102.

Australian Faunal Directory Australian Biological Resources Study, Canberra., http://www.environment.gov.au/biodiversity/abrs/online-resources/fauna/afd/index.html .

Bell, JJ 2008, 'The functional roles of marine sponges', *Estuarine, Coastal and Shelf Science*, vol. 79, pp. 341-53.

Benson, DA, Karsch-Mizrachi, I, Lipman, DJ, Ostell, J & Sayers, EW 2009, 'GenBank', *Nucleic Acids Research*, vol. 37, no. Database issue, pp. D26-31.

Blanquer, A & Uriz, MJ 2007, 'Cryptic speciation in marine sponges evidenced by mitochondrial and nuclear genes: A phylogenetic approach', *Molecular Phylogenetics and Evolution*, vol. 45, pp. 392-7.

Borchiellini, C, Chombard, C, Lafay, B & Boury-Esnault, N 2000, 'Molecular systematics of sponges (Porifera)', *Hydrobiologia*, vol. 420, pp. 15-27.

Cárdenas, P, Xavier, JR, Reveillaud, J, Schander, C & Rapp, HT 2011, 'Molecular phylogeny of the Astrophorida (Porifera, *Demospongiae*[p]) reveals an unexpected high level of spicule homoplasy', *PLoS ONE*, vol. 6, p. e18318.

Carella, M, Agell, G, Cárdenas, P & Uriz, MJ 2016, 'Phylogenetic Reassessment of Antarctic Tetillidae (Demospongiae, Tetractinellida) Reveals New Genera and Genetic Similarity among Morphologically Distinct Species ', *PLoS ONE*, vol. 11, no. 8, p. e0160718.

de Paula, TS, Zilberberg, C, Hajdu, E & Lôbo-Hajdu, G 2012, 'Morphology and molecules on opposite sides of the diversity gradient: Four cryptic species of the *Cliona celata* (Porifera, Demospongiae) complex in South America revealed by mitochondrial and nuclear markers', *Molecular Phylogenetics and Evolution*, vol. 62, pp. 529-41.

DeBiasse, MB & Hellberg, ME 2015, 'Discordance between morphological and molecular species boundaries among Caribbean species of the reef sponge *Callyspongia*', *Ecology and Evolution*, vol. 5, no. 3, pp. 663-75.

Dohrmann, M, Haen, KM, Lavrov, DV & Wörheide, G 2012, 'Molecular phylogeny of glass sponges (Porifera, Hexactinellida): increased taxon sampling and inclusion of the mitochondrial protein-coding gene, cytochrome oxidase subunit I', *Hydrobiologia*, vol. 687, pp. 11-20.

Duran, S, Pascual, M & Turon, X 2004, 'Low levels of genetic variation in mtDNA sequences over the western Mediterranean and Atlantic range of the sponge *Crambe crambe* (Poecilosclerida)', *Marine Biology*, vol. 144, pp. 31-5.

Erpenbeck, D, Breeuwer, JAJ, Parra-Velandia, FJ & Soest, RWMv 2006, 'Speculation with spiculation?-Three independent gene fragments and biochemical characters versus morphology in demosponge higher classification', *Molecular Phylogenetics and Evolution*, vol. 38, pp. 293-305.

Erpenbeck, D, Breeuwer, JAJ, Velde, HCvd & Soest, RWMv 2002, 'Unravelling host and symbiont phylogenies of halichondrid sponges (Demospongiae, Porifera) using a mitochondrial marker', *Marine Biology*, vol. 141, pp. 377-86.

Erpenbeck, D, Hall, K, Alvarez, B, Büttner, G, Sacher, K, Schätzle, S, Schuster, A, Vargas, S, Hooper, JNA & Wörheide, G 2012, 'The phylogeny of halichondrid demosponges: past and present re-visited with DNA-barcoding data', *Organisms Diversity and Evolution*, vol. 12, pp. 57-70.

Erpenbeck, D, McCormack, GP, Breeuwer, JAJ & Van Soest, RWM 2004, 'Order level differences in the demosponge LSU structure-New insights in an old taxon', *Molecular Phylogenetics and Evolution*, vol. 32, no. 1, pp. 388-95.

Erpenbeck, D, Sutcliffe, P, Cook, S, Dietzel, A, Maldonado, M, Soest, RWMv, Hooper, JNA & Wörheide, G 2012, 'Horny sponges and their affairs: on the phylogenetic relationships of keratose sponges', *Molecular Phylogenetics and Evolution*, vol. 63, pp. 809-16.

Flot, J 2007, 'CHAMPURU 1.0: a computer software for unraveling mixtures of two DNA sequences of unequal lengths', *Molecular Ecology Notes*, vol. 7, pp. 974-7.

Flot, J, Tillier, A, Samadi, S & Tillier, S 2006, 'Phase determination from direct sequencing of length-variable DNA regions', *Molecular Ecology Notes*, vol. 6, pp. 627-30.

Folmer, O, Black, M, Hoeh, W, Lutz, R & Vrijenhoek, R 1994, 'DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates', *Molecular Marine Biology and Biotechnology*, vol. 3, no. 5, pp. 294-9.

Fontaneto, D, Flot, J-F & Tang, CQ 2015, 'Guidelines for DNA taxonomy, with a focus on the meiofauna', *Marine Biodivercity*, vol. 45, no. 3, pp. 433-51.

Galtier, N, Gouy, M & Gautier, C 1996, 'SEA VIEW and PHYLO_ WIN: two graphic tools for sequence alignment and molecular phylogeny', *Computer Applications in the Bioscience*, vol. 12, no. 6, pp. 543-8.

Galtier, N, Nabholz, B, Glémin, S & Hurst, GDD 2009, 'Mitochondrial DNA as a marker of molecular diversity: a reappraisal', *Molecular Ecology*, vol. 18, pp. 4541-50.

Gerbi, SA 1985, *Evolution of ribosomal DNA*, New York/London.

Hall, TA 1999, 'BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT', *Nucleic Acids Symposium Series*, vol. 41, pp. 95-8.

Heim, I, Nickel, M & Brümmer, F 2007a, 'Molecular markers for species discrimination in poriferans: a case study on species of the genus *Aplysina*', *Porifera Research: Biodiversity, Innovation and Sustainability*, vol. 1, pp. 361-71.

Heim, I, Nickel, M & Brümmer, F 2007b, 'Phylogeny of the genus Tethya (Tethyidae: Hadromerida: Porifera): molecular and morphological aspects', *Journal of the Marine Biological Association of the United Kingdom*, vol. 87, pp. 1615–27.

Hooper, JNA 2003, *'Sponguide'. Guide to sponge collection and identification* Queensland Museum, http://www.qm.qld.gov.au/.

Hooper, JNA, Hall, KA, Ekins, M, Erpenbeck, D, Wörheide, G & Jolley-Rogers, G 2013, 'Managing and sharing the escalating number of sponge "unknowns": the SpongeMaps project', *Integrative and Comparative Biology*, vol. 33, no. 3, pp. 473-81.

Hooper, JNA & Van Soest, RWM 2002, *Systema Porifera. A guide to the classification of sponges*, 2 vols., Kluwer Academic/Plenum Publishers, New York.

Lavrov, DV, Wang, X & Kelly, M 2008, 'Reconstructing ordinal relationships in the Demospongiae using mitochondrial genomic data', *Molecular Phylogenetics and Evolution*, vol. 49, pp. 111-24.

Leal, MC, Puga, J, Serôdio, J, Gomes, NCM & 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*, vol. 7, p. e30580.

Liao, D 1999, 'Concerted Evolution: Molecular Mechanism and Biological Implications', *The American Journal of Human Genetics*, vol. 64, pp. 24-30.

Medina, M, Collins, AG, Silberman, JD & Sogin, ML 2001, 'Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 98, pp. 9707-12.

Morrow, C & Cárdenas, P 2015, 'Proposal for a revised classification of the Demospongiae (Porifera)', *Frontiers in Zoology*, vol. 12, no. 7, pp. 1-27.

Morrow, CC, Picton, BE, Erpenbeck, D, Boury-Esnault, N, Maggs, CA & Allcock, AL 2012, 'Congruence between nuclear and mitochondrial genes in Demospongiae: a new hypothesis for relationships within the G4 clade (Porifera: Demospongiae)', *Molecular Phylogenetics and Evolution*, vol. 62, no. 1, pp. 174-90.

Morrow, CC, Redmond, NE, Picton, BE, Thacker, RW, Collins, AG, Maggs, CA, Sigwart, JD & Allcock, AL 2013, 'Molecular phylogenies support homoplasy of multiple morphological characters used in the taxonomy of Heteroscleromorpha (Porifera: Demospongiae)', *Integrative and Comparative Biology*, vol. 53, pp. 428–46.

Nichols, SA 2005, 'An evaluation of support for order-level monophyly and interrelationships within the class Demospongiae using partial data from the large subunit rDNA and cytochrome oxidase subunit I', *Molecular Phylogenetics and Evolution*, vol. 34, pp. 81-96.

Pearson, WR 2013, 'An introduction to sequence similarity ("homology") searching', *Current Protocols in Bioinformatics*, vol. 6, no. 3, pp. 1-9.

Pillay, KRM, Asahida, T, Chen, CA, Terashima, H & Ida, H 2006, 'ITS Ribosomal DNA Distinctions and the Genetic Structures of Populations of Two Sympatric Species of *Pavona* (Cnidaria: Scleractinia) from Mauritius', *Zoological Studies*, vol. 45, no. 1, pp. 132-44.

Plotkin, A, Rapp, HT & Gerasimova, E 2012, 'Phylogenetic reconstruction of Polymastiidae (Demospongiae: Hadromerida) based on morphology', *Hydrobiologia*, vol. 687, pp. 21-41.

Plotkin, A, Voigt, O, Willassen, E & Rapp, HT 2016, 'Molecular phylogenies challenge the

classification of Polymastiidae (Porifera, Demospongiae) based on morphology', *Organisms Diversity & Evolution*, pp. 1-22.

Pöppe, J, Sutcliffe, P, Hooper, JNA, Wörheide, G & Erpenbeck, D 2010, 'COI barcoding reveals new clades and radiation patterns of Indo-Pacific sponges of the family Irciniidae (Demospongiae: Dictyoceratida)', *PLoS ONE*, vol. 5, no. 4, p. e9950.

Redmond, NE & McCormack, GP 2009, 'Ribosomal internal transcribed spacer regions are not suitable for intra- or inter-specific phylogeny reconstruction in haplosclerid sponges (Porifera: Demospongiae)', *Journal of the Marine Biological Association of the United Kingdom*, vol. 89, no. 6 pp. 1251-6.

Redmond, NE, Morrow, CC, Thacker, RW, Diaz, MC, Boury-Esnault, N, Cárdenas, P, Hajdu, E, Lôbo-Hajdu, G, Picton, BE, Pomponi, SA, Kayal, E & Collins, AG 2013, 'Phylogeny and systematics of Demospongiae in light of new small subunit ribosomal DNA (18S) sequences.', *Integrative and Comparative Biology*, vol. 53, pp. 388–415.

Redmond, NE, Raleigh, J, Soest, RWMv, Kelly, M, Travers, SAA, Bradshaw, B, Vartia, S, Stephens, KM & McCormack, GP 2011, 'Phylogenetic Relationships of the Marine Haplosclerida (Phylum Porifera) Employing Ribosomal (28S rRNA) and Mitochondrial (*cox1, nad1*) Gene Sequence Data', *PLoS ONE*, vol. 6, no. 9, p. e24344.

Szitenberg, A, Becking, LE, Vargas, S, Fernandez, JCC, Santodomingo, N, Wörheide, G, Ilan, M, Kelly, M & Huchon, D 2013, 'Phylogeny of Tetillidae (Porifera, Demospongiae, Spirophorida) based on three molecular markers', *Molecular Phylogenetics and Evolution*, vol. 67, pp. 509-19.

Taylor, MW, Schupp, PJ, Dahllöf, I, Kjelleberg, S & Steinberg, PD 2004, 'Host specificity in marine sponge-associated bacteria, and potential implications for marine microbial diversity', *Environmental Microbiology*, vol. 6, no. 2, pp. 121-30.

Usher, KM, Sutton, DC, Toze, S, Kuo, J & Fromont, J 2004, 'Biogeography and phylogeny of *Chondrilla* species (Demospongiae) in Australia', *Marine Ecology Progress Series*, vol. 270, pp. 117-27.

Van Soest, RWM, Boury-Esnault, N, Hooper, JNA, Rützler, K, De Voogd, NJ, Alvarez de Glasby, B, Hajdu, E, Pisera, AB, Manconi, R, Schoenberg, C, Klautau, M, Picton, B, Kelly, M, Vacelet, J, Dohrmann, M, Díaz, M-C, Cárdenas, P & Carballo, JL 2016, *World Porifera database*, http://www.marinespecies.org/porifera, 2016-11-03.

Vargas, S, Kelly, M, Schnabel, K, Mills, S, Bowden, D & Wörheide, G 2015, 'Diversity in a cold hot-spot: DNA-barcoding reveals patterns of evolution among Antarctic Demosponges (class Demospongiae, phylum Porifera)', *PLoS ONE*, vol. 10, no. 6, p. e0127573.

Voigt, O, Eichmann, V & Wörheide, G 2012, 'First evaluation of mitochondrial DNA as a marker for phylogeographic studies of Calcarea: a case study from *Leucetta chagosensis*', *Hydrobiologia*, vol. 687, pp. 101-6.

Wahab, MAA, Fromont, J, Whalan, S, Webster, N & Andreakis, N 2014, 'Combining morphometrics with molecular taxonomy: How different are similar foliose keratose sponges from the Australian tropics?', *Molecular Phylogenetics and Evolution*, vol. 73, pp. 23-9.

Webster, NS & Taylor, MW 2012, 'Marine sponges and their microbial symbionts: love and other relationships', *Environmental Microbiology*, vol. 14, no. 2, pp. 335-46.

Wörheide, G 2006, 'Low variation in partial cytochrome oxidase subunit I (COI) mitochondrial sequences in the coralline demosponge *Astrosclera willeyana* across the Indo-Pacific', *Marine Biology*, vol. 148, pp. 907-12.

Wörheide, G & Erpenbeck, D 2007, 'DNA taxonomy of sponges—progress and perspectives', *Journal of the Marine Biological Association of the United Kingdom*, vol. 87, pp. 1629-33.

Wulff, J 2012, 'Ecological interactions and the distribution, abundance, and diversity of sponges', *Advances in Marine Biology*, vol. 61, pp. 273-344.

CHAPTER 3 SPONGE-ASSOCIATED ACTINOBACTERIAL DIVERSITY: VALIDATION OF THE METHODS OF ACTINOBACTERIAL DNA EXTRACTION AND OPTIMISATION OF 16S rRNA GENE AMPLIFICATION

Experiments were designed to validate the two common DNA extraction protocols (CTAB-based method and DNeasy Blood & Tissue Kit) used to effectively recover actinobacterial DNA from sponge samples in order to study the sponge-associated actinobacterial diversity. This was done by artificially spiking sponge samples with actinobacteria (spores, mycelia and a combination of the two). The results demonstrated that both DNA extraction methods were effective in obtaining DNA from the sponge samples as well as the sponge samples spiked with different amounts of actinobacteria. However, it was noted that in the presence of the sponge, the bacterial 16S rRNA gene could not be amplified unless the combined DNA template was diluted. To test the hypothesis that the extracted sponge DNA contained inhibitors, dilutions of the DNA extracts were tested for six sponge species representing five Orders. The results suggested that the inhibitors were co-extracted with the sponge DNA and a high dilution of this DNA was required for the successful PCR amplification for most of the samples. The optimised PCR conditions, including primer selection, PCR reaction system and program optimisation, further improved the PCR performance. However, no single PCR condition was found to be suitable for the diverse sponge samples using various primer sets. These results highlight for the first time that the DNA extraction methods used are effective in obtaining actinobacterial DNA, and that the presence of inhibitors in the sponge DNA requires high dilution coupled with fine tuning of the PCR conditions to achieve success in the study of sponge-associated actinobacterial diversity.

3.1 Introduction

Marine sponges (phylum Porifera) are considered to be the oldest multicellular animals with a history of more than 600 million years. They have attracted substantial research interests because of their ecological importance and their production of a wide range of bioactive compounds for pharmacological use (Ando et al. 2010; Blunt et al. 2010; Leal et al. 2012; Sirirak et al. 2013; Vogel 2008; Waters et al. 2010). One striking characteristic of sponges is their microbial association with a remarkable array of microorganisms, such as archaea (Preston et al. 1996; Radax et al. 2012; Turque et al. 2010), bacteria (Hentschel et al. 2001; Hentschel et al. 2012; Montalvo & Hill 2011; Radwan et al. 2010; Richardson et al. 2012; Schmitt et al. 2011; Webster & Hill 2001) including actinobacteria (Abdelmohsen, Bayer & Hentschel 2014) and cyanobacteria (Alex et al. 2012; Thacker & Starnes 2003), unicellular algae (Annenkova et al. 2011; He et al. 2014; Hentschel et al. 2012; Wecker et al. 2015), and

fungi (Gopi et al. 2012; Maldonado et al. 2005). These microorganisms are reported to comprise between 35% to 40% of the total tissue volume in some sponge species (Hentschel et al. 2012; Taylor et al. 2007; Vacelet & Donadey 1977) and exceed a density of 10^9 microbial cells per ml of sponge tissue (De Voogd et al. 2015; Webster & Hill 2001). The extraordinary high abundance and diversity of microorganisms in sponges has led to ecological questions on their role and how the association is established and maintained.

Actinobacteria, producing novel natural products with pharmaceutical applications, are a promising resource for new drug discovery (Zotchev 2012). Marine sponges have been reported to host a high diversity of actinobacteria and many of them are rare and new (Abdelmohsen, Bayer & Hentschel 2014; Hentschel et al. 2012; Webster & Taylor 2012). These sponge-associated actinobacteria produce novel natural products with a diverse range of bioactivities, such as anticancer, antimicrobial, antiparasitic, neurological, antioxidant and anti-HIV activities (Abdelmohsen, Bayer & Hentschel 2014). The study on the diversity of sponge-associated actinobacteria is an important area to uncover novel strains for drug discovery and development.

Advances in molecular techniques, mainly based on 16S rRNA gene sequences, provide a culture-independent means to characterise the microbial diversity and to make a more accurate assessment of phylogenetic affiliation of microbes in a complex community such as sponges (Hentschel et al. 2012). Deep sequencing with next-generation platforms, such as Illumina (Logares et al. 2014), has revealed unparalleled bacterial diversity in the invertebrate host. In line with the previously described ‘sponge-specific’ clusters (Taylor et al. 2007; Webster et al. 2010), studies using next-generation sequencing have demonstrated sponge-specific bacterial communities in a number of sponge species (Lee et al. 2011; Schmitt et al. 2012).

Notably, for successful application of the next generation sequencing techniques, DNA extraction to obtain both high abundant and low abundant sponge-associated actinobacterial DNA from the whole sponge tissue is a primary requirement. However, there are no reports on the study of the DNA extraction efficiency from the sponge samples and their associated actinobacteria. Moreover, it was noted that with a number of sponge samples PCR products could not be obtained from extracted community DNA. It is not clear whether there are inhibitors in the sponge DNA samples to hinder the amplification of actinobacterial DNA when they are in such a complex biological mix. In addition, PCR amplification was also limited by the primer selection and the reaction conditions. While it is a common practice to optimise the PCR reaction conditions, however there are few reports on what causes the problem of unsuccessful PCR.

The purpose of this study was therefore to determine whether the DNA extraction methods

were effective in obtaining DNA of both sponges and their associated actinobacteria, as well as to determine whether successful PCR amplification of the 16S rRNA gene can be achieved to study the actinobacterial diversity. To do this, two commonly used sponge-community DNA extraction methods were chosen. We tested the DNA extraction efficiency of externally added actinobacteria, as spores and mycelia, in the presence of sponge by quantifying the DNA yield. The sponge species chosen for these studies was one for which there was no successful PCR amplification of the 16S rRNA gene. Therefore, any PCR products with the actinobacteria-spiked sponge samples would confirm the efficacy of the DNA extraction methods for the spiked actinobacteria. Furthermore, we tested the possibility of inhibitors in the DNA extract of the sponge samples which could be responsible for the unsuccessful actinobacterial 16S rRNA gene amplification. Finally, the PCR conditions were optimised by using four pairs of primers combined with five PCR programs.

3.2 Materials and methods

3.2.1 Specimens

Six sponge specimens were collected from Klein Point, Outer Harbour and Rapid Bay in South Australia. They were put on ice immediately after collection for transport to the laboratory and transferred into a -80ºC freezer for storage. All the details of the specimens are given in Table 3-1. The six sponge species chosen for this study as representatives of five orders: there were *Halichondria* sp. (*Hs*), *Pseudoceratina* sp. (*Ps*), *Igernella notabilis* (*In*), *Raspailia vestigifera* (*Rv*) from four orders and another two, *Tedania* sp. (*Ts*), *Desmapsamma* sp. (*Ds*), from the fifth order. Sponge species *Hs* was used in the DNA extraction efficiency test as it did not show any PCR products with the microbial 16S rRNA gene primers under this test PCR reaction condition. It was spiked individually with spores and mycelia of four actinobacterial species belonging to four genera. The actinobacterial strains were provided by the Department of Medical Biotechnology, Flinders University (Kaewkla and Franco 2010; Kaewkla and Franco 2013; Shomura et al. 1987; Sveshnikova et al. 1969). The spores were harvested from cultures grown on agar media and the mycelia were harvested from liquid media cultured for two weeks. All the sponges, actinobacterial spores and mycelia were freeze dried before DNA extraction.

Table 3 - 1 The sponge and actinobacteria specimens

Specimen No.		Species name (Species; Family; Order)	GenBank No. /Deposition No.	Location/ Reference
Sponge	*Hs* [a]	*Halichondria* sp. [b]; *Halichondriidae*; *Suberitida* [c]	KJ546365 (COI mtDNA)	Klein Point
	Ps	*Pseudoceratina* sp.; *Pseudoceratinidae*; *Verongida*	KJ546363 (COI mtDNA)	Outer Harbour
	In	*Igernella notabilis*; *Dictyodendrillidae*; *Dendroceratida*	KJ546351 (COI mtDNA)	Rapid Bay
	Rv	*Raspailia vestigifera*; *Raspailiidae*; *Axinellida* [d]	KJ620379 (28S rRNA gene)	Rapid Bay
	Ts	*Tedania* sp. [e]; *Tedaniidae*; *Poecilosclerida*	KJ620377 (28S rRNA gene)	Rapid Bay
	Ds	*Desmapsamma* sp. [f]; *Desmacididae*; *Poecilosclerida*	KJ546354 (COI mtDNA)	Rapid Bay
Actinobacteria	*Ke*	*Kribbella endophytica*; *Nocardioidaceae*; *Actinomycetales*	DSM 23718	Kaewkla and Franco 2013
	Ma	*Micromonospora aurantiaca*; *Micromonosporaceae*; *Actinomycetales*	DSM 43813	Sveshnikova et al. 1969
	Sa	*Streptomyces aculeolatus*; *Streptomycetaceae*; *Actinomycetales*	DSM 41644	Shomura et al. 1987
	Pa	*Pseudonocardia adelaidensis*; *Pseudonocardiaceae*; *Actinomycetales*	DSM 45352	Kaewkla and Franco 2010

[a] The abbreviations represent the species of sponge and actinobacteria samples. The abbreviations are valid in the whole chapter; [b] The species *Halichondria* sp. was identified as *Halichondria okadai* with Accession Number of KJ546365 in GenBank, but it has not previously been found in Australia based on World Porifera Database (WPD); [c] The recent classification of sponges has transferred the genus *Halichondria* to the order Suberitida based on WPD; [d] The recent classification of sponges has transferred the genus *Raspailia* to the order Axinellida; [e] The species *Tedania* sp. was identified as *Tedania tubulifera* with Accession Number of KJ620377 in GenBank, but not reported previously from Australia based on the World Porifera Database (WPD); [f] The species *Desmapsamma* sp. was identified as *Desmapsamma anchorata* with Accession Number of KJ546354 GenBank, but not previously reported in Australia based on World Porifera Database (WPD).

3.2.2 Sample preparation

To test the DNA extraction efficiency of sponge and their associated actinobacteria, artificially spiked samples were prepared. Spores and mycelial samples of four actinobacteria *Kribbella endophytica* (*Ke*), *Micromonospora aurantiaca* (*Ma*), *Streptomyces aculeolatus* (*Sa*), *Pseudonocardia adelaidensis* (*Pa*) were combined individually with sponge species *Hs* using three different mixing ratios (3: 7; 5: 5; 7: 3) to a total 10 mg. Another four samples combined 3.5 mg of spores, 3.5 mg mycelium and 3 mg sponge (*Hs*) for four actinobacteria strains. In addition, four DNA combinations were prepared by mixing the spore DNA of the four actinobacterial strains individually with the DNA of sponge *Hs* at a ratio of 7: 3 (volume) to 10 µl. The control group consists of 13 samples: four pure actinobacterial strains, four pure spore samples, four pure mycelium samples and one sponge sample.

3.2.3 DNA extraction

The DNA extraction methods utilised in this study: 1) CTAB-based method (Doyle & Dickson 1987). Briefly, the CTAB extraction buffer was combined with polyvinylpyrrolidone (PVP) and β-mercaptoethanol to lyse tissues. For sponge samples, the freeze dried sponge tissues were ground and suspended in the water for injection for one hour, and the tissue deposited in the bottom of the tube was collected. A bead-beating step using 1.0 mm dia. silica beads (Biospec Products) was applied to increase the DNA release (Simister et al. 2011). 2) DNeasy Blood & Tissue Kit (QIAGEN, Germany), following the manufacturer's protocol. The Kit was employed only when the first method failed in this study. The purified DNA was resuspended in 35 µl of sterile distilled water and stored at - 20°C. The samples were extracted at least in duplicate.

3.2.4 PCR amplification of 16S rRNA gene

The actinobacteria-biased primers used for the 16S rRNA gene are 27F (5'-GAG AGT TTG ATC CTG GCT CAG-3') and 765R (5'-CTG TTT GCT CCC CAC GCT TTC-3') (Coombs and Franco 2003). The amplification reaction was performed in a 20 µl volume containing 2 µl of 10× ThermoPol reaction buffer; 0.4 µl of dNTPs (10 mM); 0.8 µl of primers each (10 µM); 2 U *Taq* polymerase and 4 µl purified DNA template. If the undiluted DNA samples could not be amplified the DNA template was diluted by 10^1, 10^2, 10^3, and 10^4 fold for the PCR reaction. The thermocycler program was as follows: a 5-mins initial denaturation at 95°C; 25 cycles of 94°C for 1 min, 53°C for 1 min and 72°C for 1 min; and a final extension step at 72°C for 10 min. The samples were done in triplicate.

3.2.5 PCR optimisation for 16S rRNA gene

If the actinobacteria-biased PCR primers did not work, an additional four pairs of primers were employed to select the optimal primer set (Table 3-2). Five reaction programs were

compared to improve the PCR amplification. The modifications of the reaction system components and the PCR program were based on the PCR principles of Mullis et al. (1994) to improve the performance of the amplification: I. replacing the 10×standard *Taq* reaction buffer with the 10×ThermoPol reaction buffer which contains 2.0 mM $MgSO_4$ (1×); II. using the One*Taq* GC reaction buffer that contains 10-20% High GC enhancer (1×) and 5% (v/v) DMSO (1×), which are suitable for the high GC DNA samples; III. adding DMSO at a final concentration of 3% (v/v) (Mammedov et al. 2008). In addition, the reaction system accompanied with the PCR program together were optimised: IV. using One*Taq* GC reaction buffer, reducing *Taq* polymerase to a final concentration of 0.01 U/μl as the DMSO needs a low *Taq* working condition and decreasing denaturation temperature to 90°C (Mammedov et al. 2008); and V. using the standard buffer, adding DMSO to 8% (v/v) and reducing the *Taq* polymerase to 0.01 U/μl as well as decreasing the denaturation temperature to 90°C. Actinobacterial DNA, four spore DNA and four mycelium DNA samples derived from four strains (*Ke*, *Ma*, *Sa* and *Pa*) were amplified by these additional four pairs of primers under the basic PCR program (see Section 3.2.4) as a control

.

Table 3 - 2 The applied primer sets for validation

Primers	Sequences	Reference
S-C-Act-235-a-S-20 [a] S-C-Act-878-a-A-19	5'-CGC GGC CTA TCA GCT TGT TG-3' 5'-CCG TAC TCC CCA GGC GGG G-3'	Stach et al. 2003
8F-926R [b]	5'-AGA GTT TGA TCC TGG CTC AG-3' 5'-CCG TCA ATT CCT TTR AGT TT-3'	Olson et al. 2013; Olson and Gao 2013; Lee et al. 2009; 2011
27F-1492R [b]	5'-AGA GTT TGA TCC TGG CTC AG-3' 5'-GGT TAC CTT GTT ACG ACT T-3'	De Voogd et al. 2015; Montalvo & Hill 2011
28F-519R [b]	5'-GAG TTT GAT CNT GGC TCA G-3' 5'-GTN TTA CNG CGG CKG CTG-3'	Croué et al. 2013

[a] Actinobacteria-specific primers; [b] Universal bacteria primers, which are used for sponge-associated microbial community study with positive results.

3.2.6 PCR amplification of COI mtDNA

The PCR amplification of the sponge COI mtDNA was carried out to confirm the quality of the sponge DNA extraction. The universal primers employed were: LCO1490 (5'-GGT CAA CAA ATC ATA AAG ATA TTG G-3') and HCO2198 (5'-TA A ACT TCAG GGT GAC CAA AAA ATC A-3') (Folmer et al. 1994). Amplification reactions were performed in a 20 μl volume containing 2 μl of 10× standard *Taq* reaction buffer; 0.4 μl of dNTPs (10 mM); 0.8 μl of primers each (10 μM); 0.5 U *Taq* polymerase and 4μl purified DNA templates. The thermocycler

program was as follows: a 1-min initial denaturation at 94°C; 5 cycles of 94°C for 30 sec, 45°C for 90 sec and 72°C for 1 min; 35 cycles of 94°C for 30 sec, 51°C for 40 sec and 72°C for 1 min; and a final extension step at 72°C for 5 min.

3.3 Results

3.3.1 DNA extraction efficiency of sponge spiked with actinobacteria

The DNA extracted from the various combinations of samples is summarised in Table 3-3. The A260/A280 values for all the DNA samples are in the range of 1.8-2.0. DNA was extracted from the mixtures, which were composed of either actinobacterial spores or mycelia with the sponge, as well as the spores and the mycelia powders alone in group I. In the cases of the actinobacterial strains *Ma*, *Sa a*nd *Pa*, when the weight of the spores or mycelium was 7 mg, the CTAB extraction method failed to yield measurable amount of DNA from the mixed samples. These samples were extracted successfully by the DNeasy Blood & Tissue Kit (QIAGEN, Germany) instead. In group II, for the mixed sample containing *Ma*, measurable amount of DNA was obtained with the extraction Kit, rather than the CTAB method. The samples in this group showed variable DNA yields.

3.3.2 Highly-diluted DNA required for successful PCR amplification

In group I of Table 3-3, four samples extracted from the combination of sponge species *Hs* with spores of each of the four actinobacterial strains (*Ke*, *Ma*, *Sa*, *Pa*) with a ratio of 7:3 failed to amplify the microbial 16S rRNA gene. Similarly, the other four combinations with the mycelium in place of the spores also did not result in amplification of the 16S rRNA gene. The other two mixing ratios (5:5 and 3:7), resulted in successful PCR amplification after diluting the DNA by 10^1 to 10^4 fold (Table 3-3). The DNA dilution required for successful PCR amplification was lower when the mycelia replaced the spores in the combinations at the same mixing ratio. It was noted however that the DNA yields of the mycelia were similar to that of the spores when they were individually combined with the sponge in the same ratio. In addition, with the amount of sponge powder reduced from 5 mg to 3 mg in the mixed samples, the DNA template dilution reduced by 10 fold when they were working on the 16S rRNA primers whilst the DNA yield was only 10% lower. In this group, the positive control for the DNA of the combined samples and the sponge alone sample was their successful sponge COI mtDNA amplification at the original DNA concentration. The validity of the DNA extraction from actinobacterial spores and mycelia was confirmed by the 16S rRNA gene products, which were amplified using the individual spores or mycelial DNA at the original concentrations and dilutions, even to a maximum 10^4 fold in certain cases. Apart from the combinations with 7 mg sponge powder, the original concentrations of the four actinobacterial spore DNA extracts were between 417.6 to 949.1 ng/ µl, which were all higher than the original DNA concentration of their relevant mixed samples irrespective of the

mixing ratios. This indicates that the concentration of actinobacterial spore DNA used, per se, was not inhibitory. For the actinobacterial mycelia, the DNA concentrations were also all much higher than their combined samples that successfully worked in the PCR reaction. In group II, the 16S rRNA gene amplification using the DNA extracted from the sponge sample combined with both of the spores and mycelia also required templates to be highly diluted.

In group III, the DNA was extracted separately from the sponge samples and the actinobacteria and then mixed prior to the PCR reaction. In this case, PCR products were obtained for all of the combined DNA samples after at least a 10 fold dilution. No dilutions were required for the amplification of the actinobacterial DNA samples. However, the 16S rRNA gene could not be amplified from the sponge samples at any dilutions tested.

Table 3 - 3 The summary of sample combination types and the working concentration of the DNA templates in the 16S rRNA PCR amplification

Group I									
Test Sample (mg) Spores [a]: Sponge		Conc. (ng/μl)[b]	Dilution [c] for 16S rRNA PCR	Dilution [c] for COI PCR	Test Sample (mg) Mycelium [a]: Sponge		Conc. (ng/μl)[b]	Dilution [c] for 16S rRNA PCR	Dilution [c] for COI PCR
	3:7	620.3	No	0-2		3:7	650.9	No	0-2
Ke: *Hs*	5:5	523.9	4	0-2	*Ke*: *Hs*	5:5	540.0	3	0-2
	7:3	515.5	3	0-2		7:3	522.2	2	0-2
	3:7	429.6	No	0-2		3:7	459.0	No	0-2
Ma: *Hs*	5:5	396.8	3	0-1	*Ma*: *Hs*	5:5	398.7	2	0-2
	7:3[d]	375.9	2	0-1		7:3[d]	388.4	1	0-2
	3:7	523.1	No	0-2		3:7	570.7	No	0-2
Sa: *Hs*	5:5	516.8	3-4	0-2	*Sa*: *Hs*	5:5	555.3	2	0-2
	7:3[d]	489.5	2	0-2		7:3[d]	500.8	1	0-2
	3:7	877.4	No	0-3		3:7	771.4	No	0-3
Pa: *Hs*	5:5	800.5	4	0-3	*Pa*: *Hs*	5:5	741.1	3	0-3
	7:3[d]	792.4	3	0-3		7:3[d]	711.8	2	0-3
Ke	10	579.2	0-4	n.d.	*Ke*	10	281.5	0-3	n.d.
Ma	10	417.6	0-3	n.d.	*Ma*	10	352.4	0-2	n.d.
Sa	10	856.2	0-4	n.d.	*Sa*	10	366.3	0-2	n.d.
Pa	10	949.1	0-4	n.d.	*Pa*	10	489.5	0-3	n.d.
Group II									
Test Sample (mg) Spores [a] (3.5): Mycelium (3.5): Sponge (3)					Conc. (ng/μl)[b]			Dilution [c] for 16S rRNA PCR	Dilution [c] for COI PCR

Ke-s: *Ke*-m: *Hs*	510.0	2-3	0-2
Ma-s: *Ma*-m: *Hs* [d]	381.2	2-3	0-2
Sa-s: *Sa*-m: *Hs*	465.7	2-3	0-2
Pa-s: *Pa*-m: *Hs*	756.6	2-4	0-2
Group III			
Test Sample (μl) Spores [a] DNA (7): Sponge DNA (3)	Conc. (ng/μl)[b]	Dilution [c] for 16S rRNA PCR	Dilution [c] for COI PCR
Ke: *Hs*	629.6	1-3	0-1
Ma: *Hs*	511.9	1-2	0-1
Sa: *Hs*	820.4	2-3	0-1
Pa: *Hs*	877.5	2-3	0-1
Ke	433.0	0-2	n.d.
Ma	371.9	0-2	n.d.
Sa	503.8	0-2	n.d.
Pa	644.7	0-2	n.d.
Hs	793.3	n.d.	0-3

[a] Abbreviations as per Table 3 -1; [b] Concentration of DNA template; [c] The number in the Dilution column: 0=10^0 fold dilution (original), 1=10^1 fold dilution, 2=10^2 fold dilution, 3=10^3 fold dilution, 4=10^4 fold dilution; [d] DNA was extracted by DNeasy Blood & Tissue Kit (QIAGEN, Germany).

3.3.3 Primer selection and PCR optimisation required for successful PCR amplification of individual sponges

One pair of actinobacteria-specific primers and three pairs of universal bacterial primers were chosen to amplify microbial 16S rRNA genes in this study. The test results are shown in Table 3-4. All the individual actinobacteria strains, either as spores or mycelium successfully yielded amplification products from the four primer sets using the basic PCR program (see Section 3.2.4). The working concentration of the DNA templates for every spore or mycelium sample alone ranged from the original to 10^2 fold dilution. Sponge-community DNA (which includes DNA from the sponge and all associated microorganisms) from the four different sponge species (*Hs*, *Ps*, *In* and *Ds*) gave PCR products with the primer sets 27F-1492R and 28F-519R using the basic PCR reaction conditions. At least 10 fold dilution of these four sponge-community DNA was required to successfully obtain the PCR products.

The sponge species *Rv* and *Ts*, with no amplification products were then evaluated by the same two pairs of primers 27F-1492R and 28F-519R under another five PCR reaction conditions. The test results showed that only the last PCR condition (V) amplified the 16S rRNA gene successfully from the sponge-community DNA of species *Rv* (821.7 ng/µl, 10^2 to 10^3 fold dilution) using both pairs of primers as well as from the sponge-community DNA of species *Ts* (500.5 ng/µl, 10^2 fold dilution) by the primer set 28F-519R.

Using this improved PCR reaction condition, all the six sponge species were tested to amplify the 16S rRNA gene using the above four pairs of primers (Table 3-5). Four sponge species (*Hs*, *Rv*, *Ts*, *Ds*) were amplified by the actinobacteria-specific primers (S-C-Act-235-a-S-20 & S-C-Act-878-a-A-19) and only one species (Margulies et al.) could work with the primer sets 8F-926R and 27F-1492R. The primer pair 28F-519R gave positive results with all of these six sponge species using 10^1 to 10^3 fold dilutions of the sponge-community DNA.

Table 3 - 4 Primer selection for the amplification of the 16S rRNA gene from six sponge samples and individual strains, spores or mycelia of four actinobacterial species

Primer Set	Sponge a community DNA Working Conc. (ng/µl) [b]	Actinobacteria DNA Effective Dilution [b]	Spores DNA Effective Dilution [b]	Mycelia DNA Effective Dilution [b]
S-C-Act-235-a-S-20 S-C-Act-878-a-A-19	-			
8F-926R	-			
27F-1492R	*Hs*: 793.3; 1-2 *Ps*: 681.4; 1-3 *In*: 513.2; 1-3 *Ds*: 494.6; 1-2	*Ke*: 0-2 *Ma*: 0-2 *Sa*: 0-2 *Pa*: 0-2	*Ke*: 0-2 *Ma*: 0-1 *Sa*: 0-2 *Pa*: 0-2	*Ke*: 0-1 *Ma*: 0-2 *Sa*: 0-2 *Pa*: 0-2
28F-519R	*Hs*: 793.3; 1-2 *Ps*: 681.4; 1-2 *In*: 513.2; 1-2 *Ds*: 494.6; 1-2			

[a] Abbreviations as per Table 3-1; [b] Concentration of DNA template; 0=10^0 fold dilution (original), 1=10^1 fold dilution, 2=10^2 fold dilution, 3=10^3 fold dilution.

Table 3 - 5 The optimised PCR reaction conditions with different primer sets

Primer Set	PCR reaction condition	Results [a, b]
S-C-Act-235-a-S-20 S-C-Act-878-a-A-19	V	*Hs*: 1-2 *Rv*: 2-3 *Ts*: 2 *Ds*: 2
8F-926R	V	*Rv*: 2-3
27F-1492R	V	*Rv*: 3
28F-519R	V	*Hs*: 2-3 *Ps*: 1-2 *In*: 1-2 *Rv*: 2-3 *Ts*: 2 *Ds*: 1-2

[a] Abbreviations as per Table 3 -1; [b] Concentration of DNA template; 1=10^1 fold dilution, 2=10^2 fold dilution, 3=10^3 fold dilution.

3.4 Discussion

3.4.1 Comparable DNA recovery from sponge spiked with actinobacterial spores or mycelia

The DNA extraction efficacy of the actinobacterial population within sponges is one of the critical elements for the successful analysis of sponge-associated actinobacterial diversity (Abdelmohsen et al. 2014). The recovery of high-quality, high-purity representative DNA from environmental samples highlights the importance of the extraction efficacy of the microbial DNA from the initial samples (Fang et al. 2014; Kennedy et al. 2014; Mirsepasi et al. 2014).

In this study, apart from the variation in DNA content of the different actinobacterial strains, the DNA yields correlated with the percentage of the actinobacterial spores or mycelia in the artificially combined samples. When the sponge:actinobacteria was combined in the ratio (3:7, w/w), the CTAB method (Schmitt et al. 2012; Taylor et al 2004) was not effective in releasing actinobacterial DNA. The DNeasy extraction Kit (QIAGEN, Germany) showed better results for these samples. Therefore, a suitable DNA extraction method is the foundation for the effective extraction of the sponge-associated actinobacterial DNA from the sponge samples. With both the extraction methods, the spore samples gave higher yields of DNA than the mycelium samples for the same actinobacterial strain. On the other hand, in combination with sponge, a lower yield of DNA was obtained with the spores compared to when the mycelium was added, though both yields were lower than expected. It demonstrated that the spores have a lower

DNA recovery efficiency, which indicates some interference by the sponge on the efficiency of actinobacterial DNA extraction. This was also noted with mycelium, but to a lesser extent. Nevertheless, the DNA extraction methods employed in this study were able to recover the added actinobacterial DNA from sponge samples.

3.4.2 Identification and elimination of interferences in the sponge DNA

A suitable final concentration of the valid DNA template is a second vital element for successful PCR amplification (Gillings 2014; Gunawardana et al. 2014; Mirsepasi et al. 2014). With the combined samples, amplification of the 16S rRNA gene was possible only after the DNA was diluted to a low concentration in order to reduce the interference from some sponge-derived components. It was found that when the sponge was present at 70% (w/w) in combination with either spores or mycelium, the DNA could not be amplified successfully by the microbial 16S rRNA gene primers, even though the concentration of DNA was highest among the three combinations. In addition, the DNA extracted from the sponge tissue alone did not amplify with the primers for the microbial 16S rRNA gene, even up to a 10^4 fold dilution, which indicated that they did not have sufficient actinobacterial DNA or simply could not overcome the interference present with the sponge DNA extract.

When the ratio of the spores or mycelium increased, the PCR worked successfully at 10^1 to 10^4 fold dilutions of the DNA template. The highest working concentration of the DNA template was approximately 50 ng/µl. The same spore and mycelial DNA could give a PCR product at the original concentration (10^0 fold), at which the concentrations were much higher than those of the highest working DNA concentrations for their relevant combined samples. In addition, the DNA extracted from the combined samples was amplified successfully by the COI mtDNA primers at their original concentrations. Therefore, the results illustrated that the reason for the high dilution was to relieve the interference affected by the sponge and not due to the DNA concentration being too high.

Notably, the DNA templates of the spores-sponge combination needed a 10 times higher dilution than those of the mycelia-sponge group using the same strain at the same ratio whilst they yielded similar amounts of DNA. Furthermore, the dilution of the DNA extracted from the combined samples with all the three components required a minimum of 10^2 fold dilution, whereas the DNA extracted from each sample and then combined, gave a successful PCR amplification with a 10 fold dilution. For all these combined samples, the amount of DNA template with the highest dilution was only about 0.2 ng (Table 3-3 -Group I- *Ke* spores (5) + *Hs* (5): 0.05239 ng/µl, 4 µl in the reaction system). Most of them were lower than the average requirement of the amount of the DNA template (~ 200 ng) in the PCR reaction system.

Therefore, the results proved the existence of an inhibitor(s) for the PCR amplification. In most cases, the dilution of the DNA template resulted in successful amplification of the 16S rRNA gene.

Whilst DNA template dilution is an effective way to eliminate the interference, it could result in revealing only the highly abundant bacteria associated with sponges. This problem has been reflected in many molecular studies of sponge-associated bacteria diversity that do not match even with the culturable bacteria diversity (Montalvo et al. 2014; Zhu et al. 2008). Therefore, the method-dependent bias on the sponge-associated bacterial diversity should be carefully considered in drawing any conclusions of either culture-independent or culture-dependent studies.

3.4.3 Individual optimisation of PCR conditions required for successful 16S rRNA gene amplification

PCR amplification-based techniques are commonly used to study sponge microbial diversity (Ahn et al. 2003; Anderson et al. 2010; Lee et al. 2009; Lee et al. 2011; Olson & Gao 2013; Schmitt et al. 2012; Taylor et al. 2005). A suitable primer set is the third key element to successfully amplify the actinobacterial 16S rRNA gene from the sponge-community DNA. The failure of the actinobacteria-specific primers to amplify the sponge-community DNA is the main reason for the less-than-ideal pyrosequencing results (Gao et al. 2014). After confirming that all the actinobacterial spore and mycelial DNA could be amplified individually by the four pairs of primers for PCR optimisation, we found that four of the six sponge-community DNA samples worked with the two pairs of the primers, 27F-1492R and 28F-519R, after appropriate dilutions were made (Table 3-4).

Sponge species *Rv* could be amplified by the primer pairs of 27F-1492R and 28F-519R, however, sponge species *Ts* could only work with primers 28F-519R. Therefore, one primer set 28F-519R accompanied with one basic and one modified PCR reaction conditions was established for all the representative sponge species in this study, including sponge species *Hs* (Table 3-3). It was noted that this improved PCR condition was not effective for all PCR primers for all six sponge species (Table 3-5).

3.5 Conclusion

In conclusion, the causes of the unsuccessful amplification of the microbial 16S rRNA gene from sponge samples were highlighted. The two common sponge-community DNA extraction methods tested in this study were proven to be effective for actinobacteria in the presence of sponge. The presence of a co-extracted inhibitor from the host sponges required dilution of the

template DNA to successfully obtain the target PCR products. PCR conditions were optimised so that one pair of primers successfully amplified the 16S rRNA gene of all six sponge samples. To improve our understanding of the sponge-associated actinobacterial diversity, especially for low abundant actinobacteria, it is critical to develop and apply methods that can eliminate the interference in sponge-community DNA and optimise individual PCR conditions.

References

Abdelmohsen, UR, Bayer, K & Hentschel, U 2014, 'Diversity, abundance and natural products of marine sponge-associated actinomycetes', *Natural Product Reports*, vol. 31, pp. 381-99.

Abdelmohsen UR, Yang C, Horn H, Hajjar D, Ravasi T, Hentschel U 2014, 'Actinomycetes from Red Sea sponges: Sources for chemical and phylogenetic diversity', *Marine Drugs*, vol. 12, pp. 2771-2789.

Ahn YB, Rhee SK, Fennell DE, Kerkhof LJ, Hentschel U, Häggblom MM 2003, 'Reductive dehalogenation of brominated phenolic compounds by microorganisms associated with the marine sponge *Aplysina aerophoba'*, *Applied and Environmental Microbiology*, vol. 6, pp. 4159-4166.

Alex A, Vasconcelos V, Tamagnini P, Santos A, Antunes A 2012, 'Unusual symbiotic cyanobacteria association in the genetically diverse intertidal marine sponge *Hymeniacidon perlevis* (Demospongiae, Halichondrida)'*, PLoS ONE*, vol. 7, p. e51834.

Anderson SA, Northcote PT, Page MJ 2010, 'Spatial and temporal variability of the bacterial community in different chemotypes of the New Zealand marine sponge *Mycale hentscheli'*, *FEMS Microbiology Ecology*, vol. 72, pp. 328-342.

Ando H, Ueoka R, Okada S, Fujita T, Iwashita T, Imai T, Yokoyama T, Matsumoto Y, van Soest RW, Matsunaga S 2010, 'Penasins A-E, long-chain cytotoxic sphingoid bases, from a marine sponge *Penares* sp.', *Journal of Natural Products*, vol. 73, pp. 1947-1950.

Annenkova NV, Lavrov DV, Belikov SI 2011, 'Dinoflagellates associated with freshwater sponges from the Ancient lake Baikal', *Protist*, vol. 162, pp. 222-236.

Blunt JW, Copp BR, Munro MH, Northcote PT, Prinsep MR 2010, 'Marine natural products', *Natural Products Report*, vol. 27, pp. 165-237.

Coombs JT, Franco CMM 2003, 'Isolation and identification of actinobacteria from surface-sterilized wheat roots', *Applied and Environmental Microbiology,* vol. 69, pp. 5603-5608.

Croué J, West NJ, Escande M, Intertaglia L, Lebaron P, Suzuki MT 2013, 'A single beta-proteobacterium dominates the microbial community of the crambescidine-containing sponge *Crambe crambe'*, *Scientific Report*s, vol. 3, p. 2583.
De Voogd NJ, Cleary DFR, Polónia ARM, Gomes NCM 2015, 'Bacterial community composition and predicted functional ecology of sponges, sediment and seawater from the thousand-islands reef complex, West-Java, Indonesia'. *FEMS Microbiology Ecology*, vol. 4.

Fang Y, Xu M, Chen X, Sun G, Guo J, Wu W, Liu X 2014, 'Modified pretreatment method for total microbial DNA extraction from contaminated river sediment', *Frontiers of Environmental Science & Engineering,* vol. 9, no. 3, pp. 444-452.

Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R 1994, 'DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates'. *Molecular Marine Biology and Biotechnology*, vol. 3, pp. 294-299.

Gao ZM, Wang Y, Lee OO, Tian RM, Wong YH, Bougouffa S, Batang Z, Al-Suwailem A, Lafi FF, Bajic VB, Qian PY 2014, 'Pyrosequencing reveals the microbial communities in the Red Sea sponge *Carteriospongia foliascens* and their impressive shifts in abnormal tissues', *Microbiology Ecology*, vol. 68, pp. 621-632.

Gillings MR 2014, 'Rapid extraction of PCR-competent DNA from recalcitrant environmental samples'. In Paulsen IT, Holmes AJ (eds.), pp. 7-23, *Environmental Microbiology: Methods and Protocols*. Springer: Science+Business Media, UK.

Gopi M, Kumar TTA, Balagurunatham R, Vinoth R, Dhaneesh KV, Rajasekaran R, Balasubramanian T 2012, 'Phylogenetic study of sponge associated bacteria from the Lakshadweep archipelago and the antimicrobial activities if their secondary metabolites'. *World Journal of Microbiology and Biotechnology*, vol. 28, pp. 761-766.

Gunawardana M, Chang S, Jimenez A, Holland-Moritz D, Holland-Moritz H, La Val TP, Lund C, Mullen M, Olsen J, Sztain TA, Yoo J, Moss JA, Baum MM 2014, 'Isolation of PCR quality microbial community DNA from heavily contaminated environments'. *Journal of Microbiology Methods*, vol. 102, pp. 1-7.

He L, Liu F, Karuppiah V, Ren Y, Li Z 2014, 'Comparisons of the fungal and protistan communities among different marine sponge Holobionts by pyrosequencing'. *Microbial Ecology*, vol. 67, pp. 951-961.

Hentschel U, Piel J, Degnan SM, Taylor MW 2012, 'Genomic insights into the marine sponge microbiome'. *Nature Reviews Microbiology*, vol. 10, pp. 641-654.

Hentschel U, Schmid M, Wanger M, Fieseler L, Gernert C, Hacker J 2001, 'Isolation and phylogenetic analysis of bacteria with antimicrobial activities from the Mediterranean sponges *Aplysina aerophoba* and *Aplysina cavernicola'*, *FEMS Microbiology Ecology*, vol. 35, pp. 305-312.

Kaewkla O, Franco CMM 2013, '*Kribbella endophytica* sp. nov., an ehdophytic actinobacterium isolated from the surface-sterilized leaf of a native apricot tree'. *International Journal of Systematic and Evolutionary Microbiology*, vol. 63, pp. 1249-1253.

Kaewkla O, Franco CMM 2010, '*Pseudonocardia adelaidensis* sp. nov., an endophytic actinobacterium isolated from the surface-sterilized stem of a grey box tree (Eucalyptus microcarpa)'. *International Journal of Systematic and Evolutionary Microbiology*, vol. 60, pp. 2818-2822.

Kennedy NA, Walker AW, Berry SH, Duncan SH, Farquarson FM, Louis P, Thomson JM, Satsangi J, Flint HJ, Parkhill J, Lees CW, Hold GL 2014, 'The impact of different DNA extraction kits and laboratories upon the assessment of human gut microbiota composition by 16S rRNA gene sequencing'. *PLoS ONE*, vol. 9, p. e88982.

Leal MC, Puga J, Serodio J, Gomes NC, 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*, vol. 7, p. e30580.

Lee OO, Wong YH, Qian PY 2009, 'Inter- and intraspecific variations of bacterial communities associated with marine sponges from San Juan Island, Washington', *Applied and Environmental Microbiology,* vol. 75, pp. 3513-3521.

Lee OO, Wang Y, Yang J, Lafi FF, Al-Suwailem A, Qian PY 2011, 'Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea'. *ISME Journal*, vol. 5, pp. 650-664.

Logares R, Sunagawa S, Salazar G, Cornejo-Castillo FM, Ferrera I, Sarmento H, Hingamp P, Ogata H, de Vargas C, Lima-Mendez G, Raes J, Poulain J, Jaillon O, Wincker P, Kandels-Lewis S, Karsenti E, Bork P, Acinas SG 2014, 'Metagenomic 16S rDNA Illumina tags are a powerful alternative to amplicon sequencing to explore diversity and structure of microbial communities'. *Environmental Microbiology*, vol. 16, pp. 2659-2671.

Maldonado M, Cortadellas N, Trillas MI, Rützler K 2005, 'Endosymbiotic yeast maternally transmitted in a marine sponge'. *Biological Bulletin*, *vol.* 209, pp. 94-106.

Mammedov TG, Pienaar E, Whitney SE, TerMaat JR, Carvill G, Goliath R, Subramanian A, Viljoen HJ 2008, 'A fundamental study of the PCR amplification of GC-rich DNA templates'. *Computational Biology and Chemistry*, vol. 32, pp. 452-457.

Montalvo NF, Hill RT 2011, 'Sponge-associated bacteria are strictly maintained in two closely related but geographically distant sponge hosts'. Applied and Environmental Microbiology, vol. 77, pp. 7207-7216.

Montalvo NF, Davis J, Vicente J, Pittiglio R, Ravel J, Hill RT 2014, 'Integration of culture-based and molecular analysis of a complex sponge-associated bacterial community'. *PLoS ONE*, vol. 9, p. e90517.

Mirsepasi H, Persson S, Struve C, Andersen LOB, Petersen AM, Krogfelt KA 2014, 'Microbial diversity in fecal samples depends on DNA extraction method: easyMag DNA extraction compared to QIAamp DNA stool mini kit extraction', BMC Research Notes*,* vol. 7, p. 50.

Mullis KB, Ferré F, Gibbs RA 1994, 'The polymerase chain reaction'. Springer: Science + Business Media, New York, USA.

Olson JB, Gao X 2013, 'Characterizing the bacterial associates of three Caribbean sponges along a gradient from shallow to mesophotic depths', *FEMS Microbiology Ecology*, vol. 85, pp. 74-84.

Olson JB, Thacker RW, Gochfeld DJ 2013, 'Molecular community profiling reveals impacts of time, space, and disease status on the bacterial community associated with the Caribbean sponge *Aplysina cauliformis'*, *FEMS Microbiology Ecology*, *vol.* 87, pp. 268-279.

Preston CM, Wu KY, Molinski TF, DeLong EF 1996, 'A psychrophilic crenarchaeon inhabits a marine sponge: *Cenarchaeum symbiosum* gen. nov., sp. nov.', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 93, pp. 6241-6246.

Radax R., Hoffmann F., Rapp T.R., Leninger S., Schleper C 2012, 'Ammonia-oxidising Archaea as main drivers of nitrification in cold-water sponges', *Environmental Microbiology*, vol.14, pp. 909-923.

Radwan M, Hanora A, Zan J, Mohamed NM, Abo-Elmatty DM, Abou-El-Ela SH, Hill RT 2010, 'Bacterial community analyses of two Red Sea sponges'. *Marine Biotechnology*, vol. 12, pp.

350-360.

Richardson C, Hill M, Marks C, Runyen-Janecky L, Hill A 2012, ‘Experimental manipulation of sponge/bacterial symbiont community composition with antibiotics: sponge cell aggregates as a unique tool to study animal/microorganism symbiosis’, *FEMS Microbiology Ecology*, vol. 81, pp. 407-418.

Schmitt S, Hentschel U, Taylor MW 2011, ‘Deep sequencing reveals diversity and community and community structure of complex microbiota in five Mediterranean sponges’, *Hydrobiologia*, vol. 1, pp. 341-351.

Schmitt S, Tsai P, Bell J, Fromont J, Ilan M, Lindquist N, Perez T, Rodrigo A, Schupp PJ, Vacelet J, Webster N, Hentschel U, Taylor MW 2012, ‘Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges’. *ISME Journal*, vol. 6, pp. 564-576.

Shomura T, Gomi S, Ito M, Yoshida J, Tanaka E, Amano S, Watabe H, Ohuchi S, Itoh J, Sezaki M, Takebe H and Uotani K 1987, ‘Studies of new antibiotics SF 2415. I. Taxonomy, fermentation, isolation, physico-chemical properties and biological activities’. *Journal of Antibiotics*, vol. 40, pp. 732-739.

Simister RL, Schmitt S, Taylor MW 2011, ‘Evaluating methods for the preservation and extraction of DNA and RNA for analysis of microbial communities in marine sponges’, *Journal of Experimental Marine Biology and Ecology*, vol. 397, pp. 38-43.

Sirirak T, Brecker L, Plubrukam A 2013, ‘Kabiramide L, a new antiplasmodial trisoxazole macrolide from the sponge *Pachastrissa nux’*, *Natural Products Research*, vol. 27, pp. 1213-1219.

Stach JEM, Maldonado LA, Ward AC, Goodfellow M, Bull AT 2003,'New primers for the class Actinobacteria: application to marine and terrestrial environments', *Environmental Microbiology*, vol. 5, pp. 828-841.

Sveshnikova MA, Maksimova TS and Kudrina ES 1969, ‘The species belonging to the genus Micromonospora Ørskov, 1923 and their taxonomy’, *Mikrobiol*, vol. 38, pp. 754-763. (in Russian).

Taylor MW, Radax R, Steger D, Wagner M 2007, ‘Sponge associated microorganisms: evolution, ecology, and biotechnological potential’. *Microbiology and Molecular Biology Reviews*, vol. 71, pp. 295-347.

Taylor MW, Schupp PJ, de Nys R, Kjelleberg S, Steinberg PD 2005, ‘Biogeography of bacteria associated with the marine sponge *Cymbastela concentric’*, *Environmental Microbiology*, vol. 7, pp. 419-433.

Taylor MW, Schupp PJ, Dahllof I, Kjelleberg S, Steinberg PD 2004, ‘Host specificity in marine sponge-associated bacteria, and potential implications for marine microbial diversity’, *Environmental Microbiology*, vol. 6, pp. 121-130.

Thacker RW, Starnes S 2003, ‘Host specificity of the symbiotic cyanobacterium *Oscillatoria spongeliae* in marine sponge, *Dysides* spp.’, *Marine Biology*, vol. 142, pp. 643-648.

Turque AS, Batista D, Silveira CB, Cardoso AM, Vieira RP, Moraes FC, Clementino MM, Albano RM, Paranhos R, Martins OB, Muricy G 2010, ‘Environmental shaping of sponge

associated archaeal communities', *PLoS ONE*, vol. 5, p. e15774.

Vacelet J, Donadey C 1977, 'Electron microscope study of the association between some sponges and bacteria', *Journal of Experimental Marine Biology and Ecology*, vol. 30, pp. 301-314.

Vogel G 2008, 'The inner lives of sponges', *Science*, vol. 320, pp. 1028-1030.

Waters AL, Hill RT, Place AR, Hamann MT 2010, 'The expanding role of marine microbes in pharmaceutical development', *Current Opinion in Biotechnology*, vol. 21, pp. 1-7.

Webster NS, Hill RT 2001, 'The culturable microbial community of the Great Barrier Reef sponge *Rhopaloeides odorabile* is dominated by α-proteobacterium', *Marine Biology*, vol. 138, pp. 843-851.

Webster NS, Taylor MW 2012, 'Marine sponges and their microbial symbionts: love and other relationships'. *Environmental Microbiology*, vol. 14, pp. 335-346.

Webster NS, Taylor MW, Behnam F, Lücker S, Rattei T, Whalan S, Horn M, Wagner M 2010, 'Deep sequencing reveals exceptional diversity and modes of transmission for bacterial sponge symbionts', *Environmental Microbiology*, vol. 12, pp. 2070-2082.

Wecker P, Fournier A, Bosserelle P , Debitus C, Lecellier G, Berteaux-Lecellier V 2015, 'Dinoflagellate diversity among nudibranchs and sponges from French Polynesia: Insights into associations and transfer', *Comptes Rendus Biologies*, vol. 338, pp. 278-283.

Zhu P, Li Q, Wang G 2008, 'Unique microbial signatures of the alien Hawaiian marine sponge *Suberites zeteki*', *Microbial Ecology*, vol. 55, pp. 406-414.

Zotchev SB 2012, 'Marine actinomycetes as an emerging resource for the drug development pipelines', *Journal of Biotechnology*, vol. 158, pp. 168-175.

CHAPTER 4 COMBINATION OF DIFFERENT REGION-SPECIFIC PRIMERS OF 16S rRNA GENE: ESSENTIAL FOR SPONGE MICROBIOME ANALYSIS USING ILLUMINA PLATFORM

The Illumina platform has outpaced 454 pyrosequencing to be the current technology of choice to characterise sponge microbiomes using 16S rRNA gene based metagenomic analysis. However, there is no quantitative comparison on the difference in revealing the microbial taxa richness. Different region-specific primers could be used on the Illumina platform, however, most of the studies if not all used the single primer set based approach. The study in chapter four aims to evaluate the impact of the region-specific primer set selection on their performance in uncovering and differentiating microbiomes, and as a result, a new paradigm was developed to reveal more comprehensive sponge microbiomes than any of existing studies.

The Illumina MiSeq dataset was proven to be able to cover all the affiliated (known) microbial taxa revealed by 454 using the same primers with about a four-fold increase at the class level richness. Amplification by four region-specific primer sets on Illumina indicated that they have significantly different performance on the affiliated (known) microbial OTU identification. No universal, comprehensive single primer set exists. Among the tested four pairs of the region-specific primers, the V5V8 primer set performed better on affiliated (known) microbial taxa identification and community differentiation between sponges in different orders. This contrasts with the V4 primers commonly used by the Earth Microbiome Project (EMP) and Sponge Microbiome Project (SMP) that only revealed an average 68% of the microbial genera by the V5V8 primers. Similarly, the unaffiliated microbial OTUs, including unassigned and candidate OTUs, showed substantially distinct numbers and profiles revealed by each of the four primer sets. The total number of unaffiliated phylum level OTUs was 53.7% of the total number of phylum level OTUs, and up to 62.3% at class level. It implies that the majority of the sponge microbiome has not been uncovered yet and most of them are novel and undescribed. Applying the four combined primer sets, targeting almost the full length of 16S rRNA gene, a minimum 4.6-12.9% to a maximum 1-11 fold increases were obtained in the sequence abundance (only account the sequence abundance of distinct OTUs and removed the redundant) compared to that by any single primer set. Limited overlap in microbiome structure observed between different sponge species suggested that sponge-associated microbial communities are a significant global source of unique and untapped microbial diversity. This study indicated that each type of environmental sample, such as sponges, requires a validated

combination of multiple region-specific primer sets to analyse their comprehensive microbiome so as to avoid biased conclusions. Importantly, the proposed approach provides an efficient tool to reveal the unaffiliated microbial OTUs, which are the vastly untapped microbial resources.

4.1 Introduction

Complex host-associated microbial communities, such as the marine sponge microbiome, are receiving increased attention thanks to the advanced technologies in culture-independent characterisation methods and the high possibility of discovering new and valuable bacteria for pharmaceutical and industry uses. High-throughput next-generation sequencing (NGS) provides unparalleled insight into the community structure. Until recently this was typically carried out by 454 Genome Sequencer (GS) FLX instruments using pyrosequencing technology (Margulies et al. 2005). The major drawbacks of GS are the relatively high error rates, high operational cost, and significantly lower throughput. With the emergence of the Illumina sequencing platform (Bennett 2004), these drawbacks of 454 could be overcome. Illumina instruments have a lower per-base error rate and are not as susceptible to indel errors (Jünemann et al. 2013; Loman et al. 2012), but with much shorter read lengths. Additionally, read merging strategy could improve the shortcoming of the short length of Illumina sequenced amplicons (Bartram et al. 2011; Eren et al. 2013; Kozich et al. 2013). These advantages have encouraged many researches to focus on 16S rRNA gene amplicons using the Illumina platform (Bartram et al. 2011; Caporaso et al. 2012; Degnan & Ochman 2012; Gloor et al. 2010).

The 16S small ribosomal subunit gene (16S rRNA gene) has been applied as the gold standard for phylogenetic study of bacterial communities and taxonomic identification of bacteria (Riviere et al. 2009). For most bacteria, their 16S rRNA gene has nine different hyper-variable regions, with conserved regions in between (Neefs et al. 1993), which can be used to design the primers to amplify the target regions (Woese 1987). Although relatively lower resolution than the full-length 16S rRNA gene, high throughput sequencing technology focusing on a short region provides a much larger coverage of the microbial profile for a given sample (Claesson et al. 2009) and allows for many more samples running simultaneously (Hamady et al. 2008). Different region-specific primer sets have been applied in various studies. However, the selection of the specific primers was not generally based on the literature or recommendations from the researchers in this field without a systematic validation for the studied samples. The assessments of the region suitability have indicated the PCR bias of varying degrees depending on amplicon generation (Group 2012; He et al. 2013; Kim, Morrison & Yu 2011; Klindworth et al. 2012; Kumar et al. 2011; Lee et al. 2012; Pinto & Raskin 2012; Schmalenberger, Schwieger & Tebbe 2001). Although it has been realised that length of the

sequenced amplicon could affect the resultant microbial community richness and evenness, the selection of hyper-variable region had a much larger impact on the microbial profile (Engelbrektson et al. 2010; Ghyselinck et al. 2013; Tremblay et al. 2015). Therefore, the selection of the specific regions to study the diversity of the microbial community is important and will determine the success of the study. It is however a largely overlooked parameter in most, if not all, studies that use only one region-specific amplicon, such as the Earth Microbiome Project (EMP) using V4 only (Caporaso et al. 2012; Gilbert, Jansson & Knight 2014).

Marine sponges (phylum Porifera) are evolutionary the oldest animals that contributed to over 30% of all marine natural products discovered so far (Blunt et al. 2016). They house enormously dense and diverse communities of symbiotic bacteria, archaea, and unicellular eukaryotes in their tissues, which are considered as contributing to the production of these highly diverse natural products. These microorganisms can make up to 60% of the total biomass in certain sponge species, and are found at densities exceeding 10^9 microbial cells/ cm^3 of sponge tissue (De Voogd et al. 2015). Studying the diversity of sponge-associated microbes is therefore a hot topic and essential to recover the novel strains for pharmaceutical and industry applications. There are a number of studies using 454 pyrosequencing (Hentschel et al. 2012), while more recent studies have used Illumina. However, one critical issue remains to be resolved, that is the impact of the 16S rRNA gene regions-specific primers selection on its performance in uncovering and differentiating sponge-associated microbiomes of different taxonomic groups. This issue is not only critical for the study of sponge microbiome but also for some other sponge-related studies using the Illumina platform, which applied only one amplicon, mainly the V4 region following most of the other studies such as EMP. It also raises question for all other microbiome studies using the Illumina platform without validating the selection of region-specific primers.

Therefore, the purpose of this study was to validate the selection of region-specific primer sets toward revealing the complete sponge microbial community via the Illumina MiSeq platform. Initially, the classification resolution and taxa richness of the affiliated (known) microbial OTUs were compared between 454 and Illumina MiSeq platform using four sponge species from different orders. The efficacy of the identification of affiliated (known) microbial taxa by the selected four primer sets (V1V3, V4, V4V5, and V5V8 regions) was assessed in Illumina MiSeq with the aim of gaining a deeper understanding of sponge-associated microbial communities, and a better differentiation between these communities associated with sponges in different orders. Moreover, the unaffiliated microbial OTUs were jointly analysed and compared to identify if there is an optimum primer set or indeed a combination of various

region-specific primers would be required to reveal comprehensive sponge microbial communities.

4.2 Materials and methods

4.2.1 Sponge collection and community DNA extraction

Sponge specimens were collected via scuba diving at the depths of 4-15 metres at Rapid Bay, Adelaide, South Australia (35°31'16.6"S, 138°11'07.5"E) in February and March of 2013. Each specimen was kept separately using a sterile plastic bag and in ice box during transport. Specimens were washed with 0.22 μm membrane filtered seawater to remove attached microorganisms. A $10cm^3$ sponge tissue for each specimen was cut into small parts with a sterile blade and stored in -80°C freezer for subsequent DNA extraction. The sponge identities are the closed match with the sequences in Gene Database. The identification followed the protocol (SIP) developed in Chapter 2. They are: *Aplysina archeri* with accession number KJ620395 (Yang, Franco & Zhang 2015), *Halichondria okadai* - KJ546365 (Yang, Franco & Zhang 2015), *Igernella notabilis* - KJ546352 (Geer et al. 2010), and *Tedania tubulifera* - KJ620377 (Geer et al. 2010), which belong to four different orders Verongida, Suberitida, Poecilosclerida, and Dendroceratida, respectively.

The DNA extraction method utilised the validated CTAB-based method in Chapter 3, which was modified based on Schmitt et al. (2012) and Taylor et al. (2004). Briefly, the freeze-dried sponge tissues were ground and suspended in the sterile distilled water for one hour, and after the low speed (600 ×g) centrifugation the tissue deposited at the bottom of the tube was collected. The CTAB extraction buffer was applied to lyse tissues, which was combined with polyvinylpyrrolidone (PVP) and β-mercaptoethanol to help remove phenolic compounds and clean tannins in the extract. A bead-beating step using 1.0mm diameter silica beads (Biospec Products) was applied to increase the DNA release (Simister, Schmitt & Taylor 2011). The purified DNA was resuspended in 35 μl of sterile distilled water. Purity and quantity of DNA were determined with a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The samples were extracted in duplicate. The quantified DNA sample for each specimen (A260/280: 1.8-2.0; Con. > 100 ng/μl) was divided into two aliquots of equal volume. They were kept at -20°C for subsequent PCR reactions and sequencing. One aliquot was sequenced by Roche 454 GS FLX Titanium XL+ (mode read length 700 bp), and the other with the Illumina MiSeq (2 × 300 bp paired-end reads).

4.2.2 454 GS FLX amplicon library and pyrosequencing

454 pyrosequencing amplicon libraries for each sponge community DNA were created

using the V1V3 hyper-variable region-specific primer set 28F-519R for 16S rRNA gene (28F: 5'-GAG TTT GAT CNT GGC TCA G-3'; 519R: 5'-GTN TTA CNG CGG CKG CTG-3') (Croué et al. 2013) and primer set 341F-926R for V3V5 region (341F: 5'-CCT ACG GGN GGC WGC AG-3'; 926R: 5'-CCG TCA ATT CNT TTR AGT-3') (Claesson et al. 2010; Wang & Qian 2009). They were selected from the widely used primer sets by the 454 platform for 16S rRNA gene-based metagenomic sequencing. Their positions in the gene are marked in Figure 4-1.

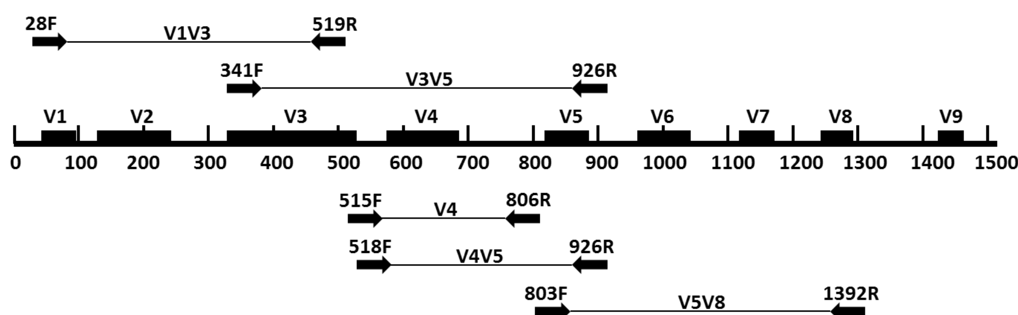


Figure 4 - 1 **Positions of primer sequences and tandem regions used in this study for 454 pyrosequencing and Illumina sequencing.** It is mapped along 16S rRNA gene (the blocks mark the hyper-variable regions V1-V9).

Quantified DNA (100 ng) was amplified in a PCR reaction buffer containing 25 mM $MgCl_2$, 1% Triton, 10 mM dNTPs, and 10 mg/ml BSA. Forward primers were tagged with 10 bp unique barcode sequences along with the 454 GS FLX Titanium sequencing adaptor (Hamady & Knight 2009). PCR was performed under the following conditions: initial denaturing at 95°C for 5 mins followed by 20 cycles of 95°C for 30 s; 56°C for 30 s; 72°C for 90 s. The reaction was terminated after 8 mins extension at 72°C. The control reactions used water to replace the DNA template in the reaction system. Post-PCR products were free of contamination. The amplicons from each DNA sample were amplified in triplicate. The purification of amplicons was conducted by fowlloing a QIAquick Gel Extraction Kit (Qiagen). Purified amplicon DNAs (A260/280: 1.8-2.0) were quantified using a Quant-iT PicoGreen kit (Invitrogen, Carlsbad, CA) and loaded for pyrosequencing using the Titanium method, provided by a Titanium genomic kit (454 Life Sciences, Roche, USA).

4.2.3 Illumina MiSeq amplicon library and sequencing

The V1V3, V4, V4V5, and V5V8 regions were selected for Illumina sequencing in order to cover the full-length 16S rRNA gene (Fig. 4-1). This excludes V9, as the previous analysis

had shown that the V9 region evolves at a rate much slower than the rest of the gene (Schloss 2010). Apart from the V1V3 region mentioned previously, the other primers used for the amplification were 515F-806R for the V4 region (515F: 5'-GTG YCA GCM GCC GCG GTA A-3'; 806R: 5'-GGA CTA CNV GGG TWT CTA AT-3') (Caporaso et al. 2012), 518F-926R for the V4V5 region (518F: 5'-CCA GCA GCY GCG GTA AN-3'; 926R: 5'-CCG TCA ATT CNT TTR AGT-3') (Nelson et al. 2014), and 803F-1392R for the V5V8 region (803F: 5'-TTA GAN ACC CNN GTA GTC-3'; 1392R: 5'-ACG GGC GGT GWG TRC-3') (Engelbrektson et al. 2010). PCR was performed based on the protocol presented in Caporaso et al. (2011). Briefly, both of the forward and reverse primers were added to the 5' and 3' Illumina adapter, respectively. The employed PCR reaction conditions were 2 mM $MgCl_2$, 0.2 μM each primer and 200 μM dNTPs. The PCR conditions were 94°C for 3 mins, followed by 94°C for 45 s, 50°C for 60 s, 72°C for 90 s in 35 cycles, and a final elongation step at 72°C for 10 mins. Negative controls and triplicate amplification were applied. After the purification and quantification, the amplicons (A260/280: 1.8-2.0) were pooled and subsequently sequenced on an Illumina MiSeq Sequencer from both ends of paired-end library preparations (2× 300 bps), using sequencing kit version 3.0 followed by base-calling using the GAPipeline version 1.4.0.

4.2.4 16S rRNA sequencing data processing

4.2.4.1 Dataset quality filtering

The sequences generated by the 454 GS FLX Titanium sequencer were filtered by the quantitative insights into microbial ecology (QIIME) pipeline (Caporaso et al. 2010). A threshold of the sequence length was applied: The sequences with < 200 bp or > 1,000 bp were discarded. In addition, the sequences contained incorrect primer sequences, or contained more than 1 ambiguous base need filter out. Then, the sequences were assigned sample aliquots based on their unique barcodes (Hamady et al. 2008). Chimeric sequences were removed using ChimeraSlayer (Haas et al. 2011). The filtered data in fasta files were used for the following analysis. Similarly, the quality control for the Illumina MiSeq dataset was processed in QIIME pipeline: the first 20 bases of all fastq files were trimmed to remove the primer sequence, and quality trimmed to remove poor quality sequence using a sliding window of 4 bases with an average base quality above 15 using the software Trimmomatic (Version 0.35). All reads were then hand trimmed to 250 bases, and any with less than 250 bases excluded. Fastq files were finally converted to fasta files for next step analysis.

4.2.4.2 Closed reference OTU picking

Reads were assigned to OTUs using a closed-reference OTU picking protocol of the QIIME toolkit (Caporaso et al. 2010), where uclust (Edgar 2010) was applied to search sequences against a subset of the Greengenes database (DeSantis et al. 2006), filtered at 97%

identity. Reads were assigned to OTUs based on their best hit to this database at greater than or equal to 97% sequence identity. Reads that did not match a reference sequence were discarded. The resulting OTU table was filtered to remove any OTU with an abundance of less than 0.05%. Representative OTU sequences were then BLASTed against the reference database (Greengenes). After aligning OTU representative sequences and filtering the alignment, the OTU table was made in QIIME. The table summarised the OTU abundances in each sample with taxonomic identifiers for each OTU.

4.2.5 Sponge microbiome analysis

Phylotype-based OTU grouping was applied to infer the representative OTUs to the genus level with a threshold of 97% similarity (Chen et al. 2013). Classification resolution and taxa richness were analysed to compare the capacity of the two sequencing platforms (454 pyrosequencing and Illumina MiSeq) on revealing the affiliated (known) microbial OTUs generated from four sponge species. The applicability and the performance of the four selected primer sets, as evaluated by Illumina MiSeq platform, were analysed for their recovery efficiency of the affiliated (known) microbial taxa. In addition, the capacity of different primer sets on differentiating the microbial communities between four sponge species was evaluated by the microbial community profiles of the affiliated (known) phylum-level OTUs on a per-primer based bar chart. The unaffiliated (unassigned and candidate) phylum/ class level OTUs were further analysed for the four sponge species revealed by four primer sets. The sequence abundance (%) was compared between the microbial profiles revealed by each single primer set and the combination of four primer sets.

4.3 Results

4.3.1 Classification resolution and taxa richness highly dependent on sequencing techniques

For the affiliated (known) microbial OTUs, when comparing the classification resolution, defined as the proportion of the OTUs confidently classified (using a threshold of 97% similarity) to a certain taxonomic level, both sequencing platforms achieved a high resolution at the phylum and class levels (Fig. 4-2). The 454 platform provided a better resolution at the order, family, and genus levels, though both techniques showed a declining trend. Notably, the Illumina technique had a far worse performance (<20%) than 454 (50%-70%) for all the four sequenced amplicons to classify the microbial community at the genus level.

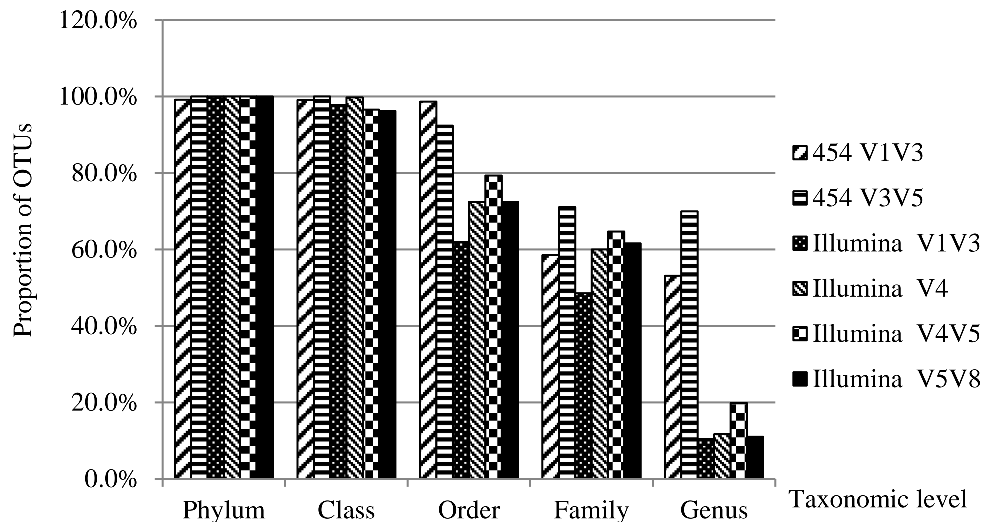


Figure 4 - 2 Proportion of affiliated (known) OTUs reaching a specified taxonomic level on 454 pyrosequencing and Illumina MiSeq platforms. Six columns in each taxonomic level represent the dataset revealed by 454 and Illumina MiSeq platforms using the different region-specific primer sets. The percentage % = the number of the OTUs identified at phylum (class/ order/ family/ genus) level/ the total number of the OTUs. The V1V3 and V3V5 region-specific primer sets were applied in 454 platform. The V1V3, V4, V4V5, and V5V8 primer sets were applied in Illumina MiSeq platform.

In order to compare the taxa richness uncovered by these two techniques, the proportion of affiliated (known) OTUs shared among these two datasets was normalised to evaluate whether the microbial community generated by Illumina could cover the one revealed by 454 when analysing the same sponge community DNA. The 454 dataset was completely covered by the Illumina dataset using the same primers for the V1V3 region regardless of the taxonomic levels (phylum, class, order or family) of the classified microbial OTUs (Fig. 4-3a). The combined 454 dataset including all the affiliated (known) taxa revealed by two primer sets were also covered by the combined Illumina MiSeq dataset (Fig. 4-3b). Importantly, the OTUs shared between those combined 454 and Illumina datasets was only a small percentage of the Illumina dataset, which were 25% and 14% for the classification at the phylum and class levels, 22% for the order and family levels, respectively.

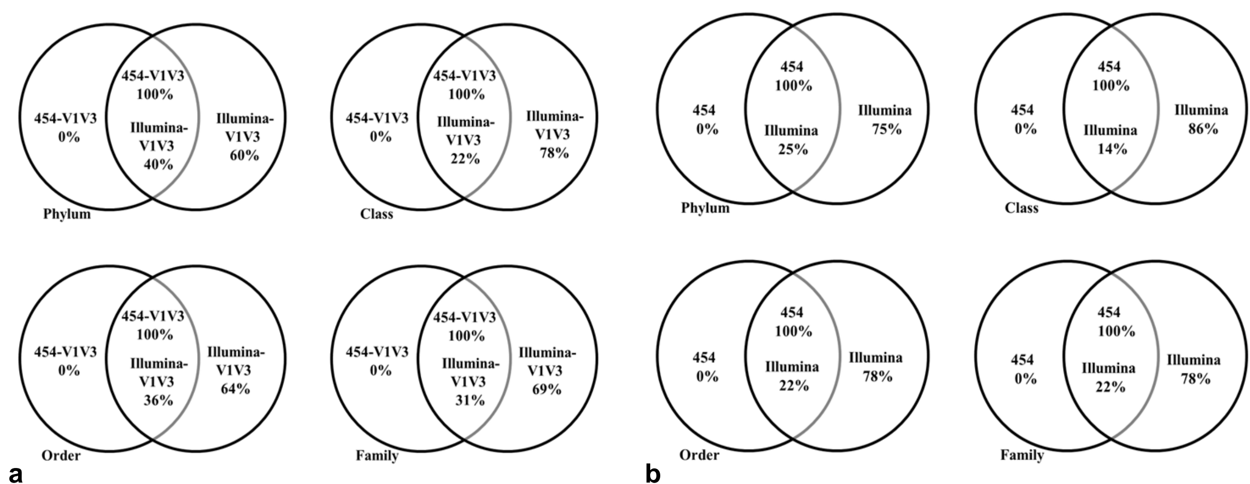


Figure 4 - 3 Comparison of the richness of affiliated (known) OTUs between 454 pyrosequencing and Illumina MiSeq platforms based on phylum, class, order, and family levels of the microbial taxa. a. comparison of the taxa revealed by same region-specific primer set V1V3 in 454 and Illumina platforms; b. comparison of the taxa revealed by 454 platform using two primer sets V1V3 and V3V5 together and the taxa revealed by Illumina MiSeq platform using all four primer sets V1V3, V4, V4V5 and V5V8.

4.3.2 Efficacy of microbial taxa identification highly dependent on region-specific primers

As shown in Table 4-1, the efficacy of affiliated (known) microbial taxa identification was assessed by the coverage of observed microbial taxa associated with the host sponge species, typically including the number and the diversity of the affiliated (known) OTUs.

It is found that the number of microbial OTUs varied significantly when revealed by the different region-specific primer sets of the 16S rRNA gene for the same sponge species. Using the 454 platform, the V1V3 region-specific primers revealed a larger number of OTUs than V3V5 when identifying the microbial community at different taxonomic levels (phylum, class, order, family, and genus). Using the Illumina MiSeq platform, the V5V8 region-specific primers showed a significantly better performance than the other three primer sets for the microbial identification at the order, family, and genus levels, though there were less differences at the phylum and class levels to reveal the same sponge species. Particularly, in the case of sponge *I. notabilis*, 50%- 58% more OTUs were generated at the genus level. For the other three sponges, a range of 19%- 50% more genus level OTUs were obtained with the V5V8 primers.

Table 4 - 1 Capacity of different primer sets in revealing the affiliated (known) microbial OTUs for four sponge microbial communities

Sponge species	*Aplysina archeri*		*Halichondria okadai*		*Igernella notabilis*		*Tedania tubulifera*	
Sequencing Platform	**454**	**Illumina MiSeq**	**454**	**Illumina MiSeq**	**454**	**Illumina MiSeq**	**454**	**Illumina MiSeq**
Primer set	P1[a], P2	P1, P3, P4, P5	P1, P2	P1, P3, P4, P5	P1, P2	P1, P3, P4, P5	P1, P2	P1, P3, P4, P5
No. of OTU /Genus [b]	20, 7	54, 49, 47, 70	13, 8	49, 46, 58, 69	21, 5	62, 60, 59, 93	10, 6	50, 49, 55, 69
No. of OTU /Family	16, 5	45, 42, 42, 54	10, 6	47, 42, 46, 64	14, 3	50, 44, 45, 67	8, 6	43, 42, 45, 54
No. of OTU /Order	14, 4	36, 32, 32, 42	9, 2	43, 35, 37, 57	11, 3	41, 35, 38, 48	6, 2	34, 32, 32, 43
No. of OTU /Class	8, 3	24, 20, 24, 26	4, 1	39, 32, 34, 51	5, 2	33, 28, 33, 34	3, 1	28, 23, 27, 34
No. of OTU /Phylum	6, 3	15, 14, 15, 16	2, 1	20, 15, 18, 24	4, 2	19, 14, 18, 19	3, 1	14, 12, 16, 18

[a] primer set and its relevant 16S rRNA gene region. P1: primer set 28F-519R for the V1V3 region; P2: primer set 341F-926R for the V3V5 region; P3: primer set 515F-806R for the V4 region; P4: primer set 518F-926R for the V4V5 region; P5: primer set 803F-1392R for the V5V8 region. [b] the number of OTUs at different taxonomic levels.

Particularly, for the diversity of the affiliated (known) microbial OTUs at the genus level, we found that the V5V8 region-specific primers achieved a significantly better efficacy as they revealed the largest number of genera and the majority of them were unique, which was defined as those genera observed by one specific primer set only (Fig. 4-4, Appendix Table 4-1). In each of the microbial profiles belonging to four sponge species, the V5V8 primer set consistently provided nearly 50% of the genera out of the combined profile that was revealed by all the four primer sets. Additionally, more than 50% of the unique genera were observed by V5V8. Corresponding with its superior performance in the sponge *I. notabilis* on the number of the identified OTUs, the highest proportion of 61.5% of the unique genera were uncovered by V5V8 in this sponge.

It is important to note that the combination of V1V3 and V5V8 datasets could cover more than 90% of the unique genera in any of these four microbial profiles (Fig. 4-5). Similarly, if applying these two primer sets, 75%- 80% of the genera could be covered (Appendix Table 4-1). In contrast, the V4 and V4V5 primer sets performed at a much lower efficacy on the identification of the unique and total number of genera. Comparing the genus diversity revealed by all four primer sets for one sponge (Appendix Table 4-1), for example *A. archeri*, we found two genera were jointly identified, which represented only 1.4% of the total genera derived from this sponge species using a combination of four primer sets. In terms of sponges *H. okadai*, *I. notabilis*, and *T. tubulifera*, there were 1.3%, 1.6%, and 2.0%, respectively.

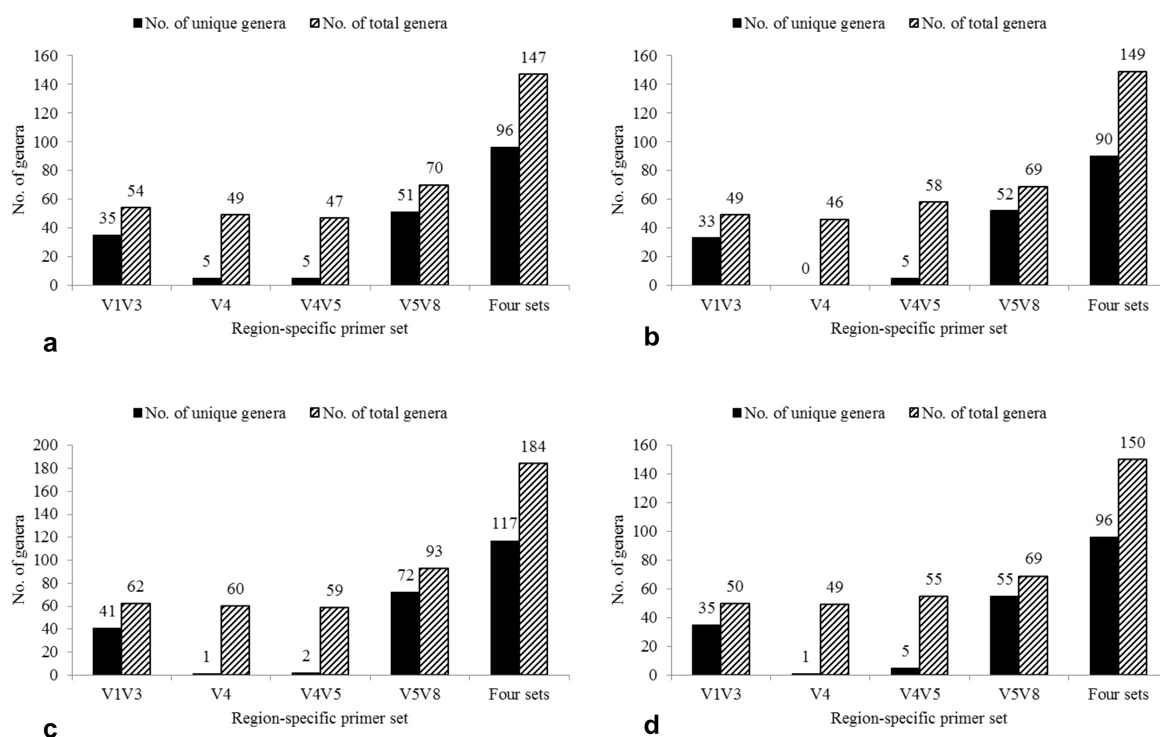


Figure 4 - 4 Number of unique genera and total genera revealed by each region-specific primer set on Illumina MiSeq platform. The specific primers are for the V1V3, V4, V4V5 and V5V8 regions of 16S rRNA gene. The number of unique genera and the total number of genera using all four primer sets were showed as the last group in each chart. a. Sponge *Aplysina archeri*; b. Sponge *Halichondria okadai*; c. Sponge *Igernella notabilis*; d. Sponge *Tedania tubulifera*.

To further evaluate the efficacy of affiliated (known) microbial taxa identification in terms of the diversity of microbial community, data for all the four sponge species were combined (Fig. 4-5) to analyse the distribution of OTUs at the phylum, class, order, and family levels revealed by four primer sets specific to regions V1V3, V4, V4V5, and V5V8. It was found that the shared OTUs revealed by four primer sets took the smallest proportion in the V5V8 profile at any taxonomic levels. The performance of V5V8 primers consistently surpassed the other three primer sets, as it uncovered a significantly higher percentage of unique OTUs with 20%, 38%, 44%, and 51% at the phylum, class, order, and family levels, respectively. Based on the different OTU throughput between these four profiles, which are noted in each of the interaction blocks, the percentage of the unique OTUs revealed by the V1V3 primer set was normalised to compare them with V5V8. There were 12%, 21%, and 21% at the class, order, and family levels, respectively, without any unique OTUs at the phylum level. With regards to the V4 primer

set, all the OTUs were covered by the other three profiles, and similarly, the V4V5 primer set revealed only a mere <2% of unique OTUs at the family level. Consequently, the V5V8 primer set is more powerful in detecting the unique OTUs than any of the other three primer sets.

In addition, evaluating the performance of one primer set to different the microbial communities of different sponge species, there were 25, 27, 36, and 27 affiliated (known) microbial genera shared among the four sponge species when using the V1V3, V4, V4V5, and V5V8 region-specific primer sets, respectively (Appendix Table 4-2). Among these shared genera, only *Spirochaeta* was common, which implied that the inherent biases due to the region-specific primers selection can greatly affect the data interpretation or even offer a misleading conclusion in microbiome studies.

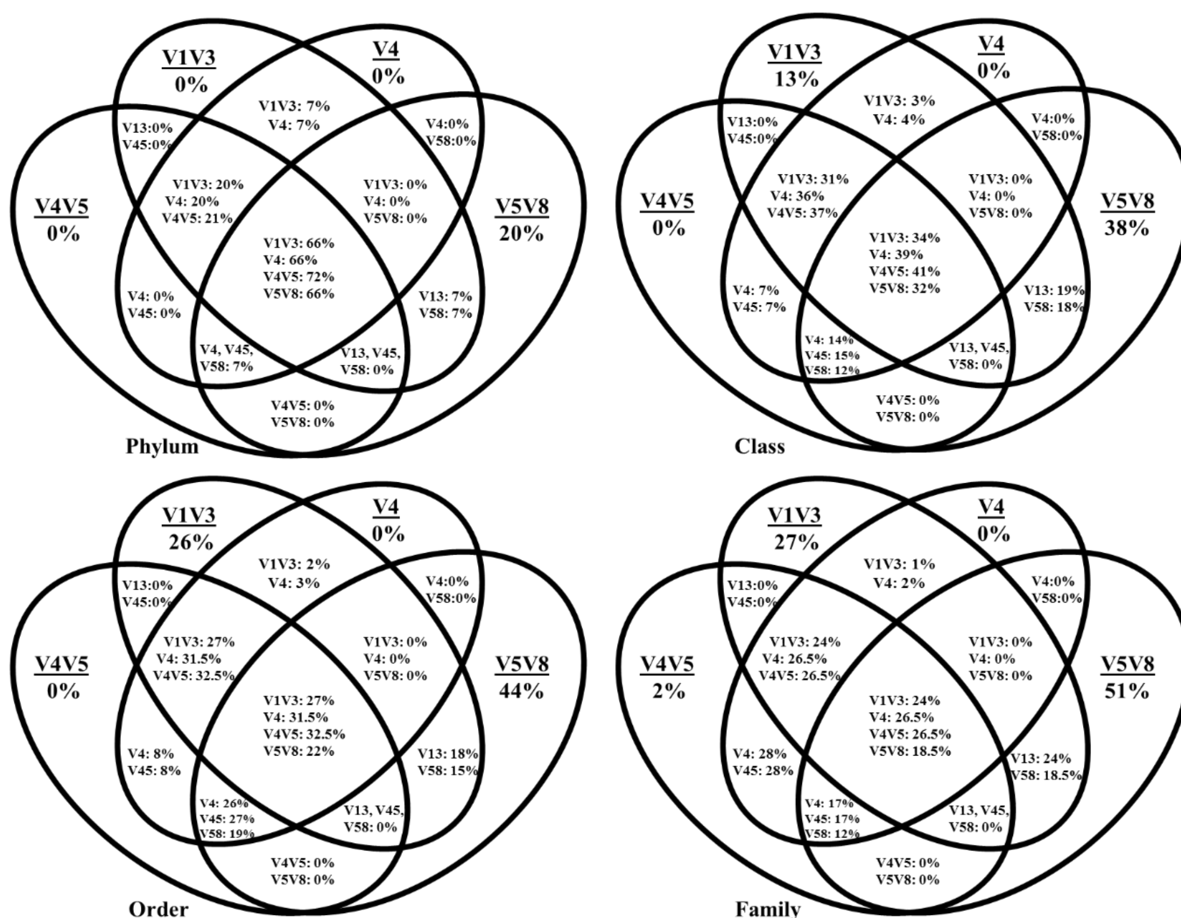


Figure 4 - 5 Distribution of affiliated (known) microbial OTUs revealed by four region-specific primer sets V1V3, V4, V4V5 and V5V8 on Illumina MiSeq platform for the same sponge samples. The comparison is based on phylum, class, order and family levels of the microbial communities. The proportion % in the interactions of the Venn diagram= the number

of the shared taxa/ the total number of the taxa revealed by each of the primer sets. Each circle represents the whole microbial OTUs revealed by one region-specific primer set with a sum of 100%.

4.3.3 Biased differentiation performance of various region-specific primers for sponge microbiome

There are 24 affiliated (known) microbial phyla revealed from sponge *H. okadai* using all the four primer sets on the Illumina MiSeq platform. For the sponges *A. archeri*, *I. notabilis*, and *T. tubulifera*, 15, 21, and 17 affiliated (known) phyla were revealed, respectively (Fig. 4-6). It was found that these four microbial communities associated with different sponge species could be easily distinguished between each other based on a single primer set, when accounting for the relative abundance of each microbial phylum. The V5V8 region-specific primers showed the best differentiation power as it is the only one that could reveal different two dominant phyla for each of the four sponge species in different orders (Fig. 4-6). Importantly, the structure of these four sponge microbial communities varied significantly when using the different four primer sets. Focusing on one sponge species, the microbial phylum richness and the proportion of each phylum also varied when changing the primer sets.

Apart from the differentiation of the microbial communities at the phylum level, the comparison of the affiliated (known) genus-level OTUs further indicated the substantial advantage of the V5V8 region-specific primers to distinguish the sponge microbial communities. As shown in Appendix Table 4-2, the V5V8 primer set revealed the smallest number of the shared microbial genera between different sponges, with 23% out of the total number of the affiliated (known) genera revealed by V5V8 for four sponges. In terms of the other three primer sets, there were 30%, 41%, and 55% for V1V3, V4, and V4V5, respectively.

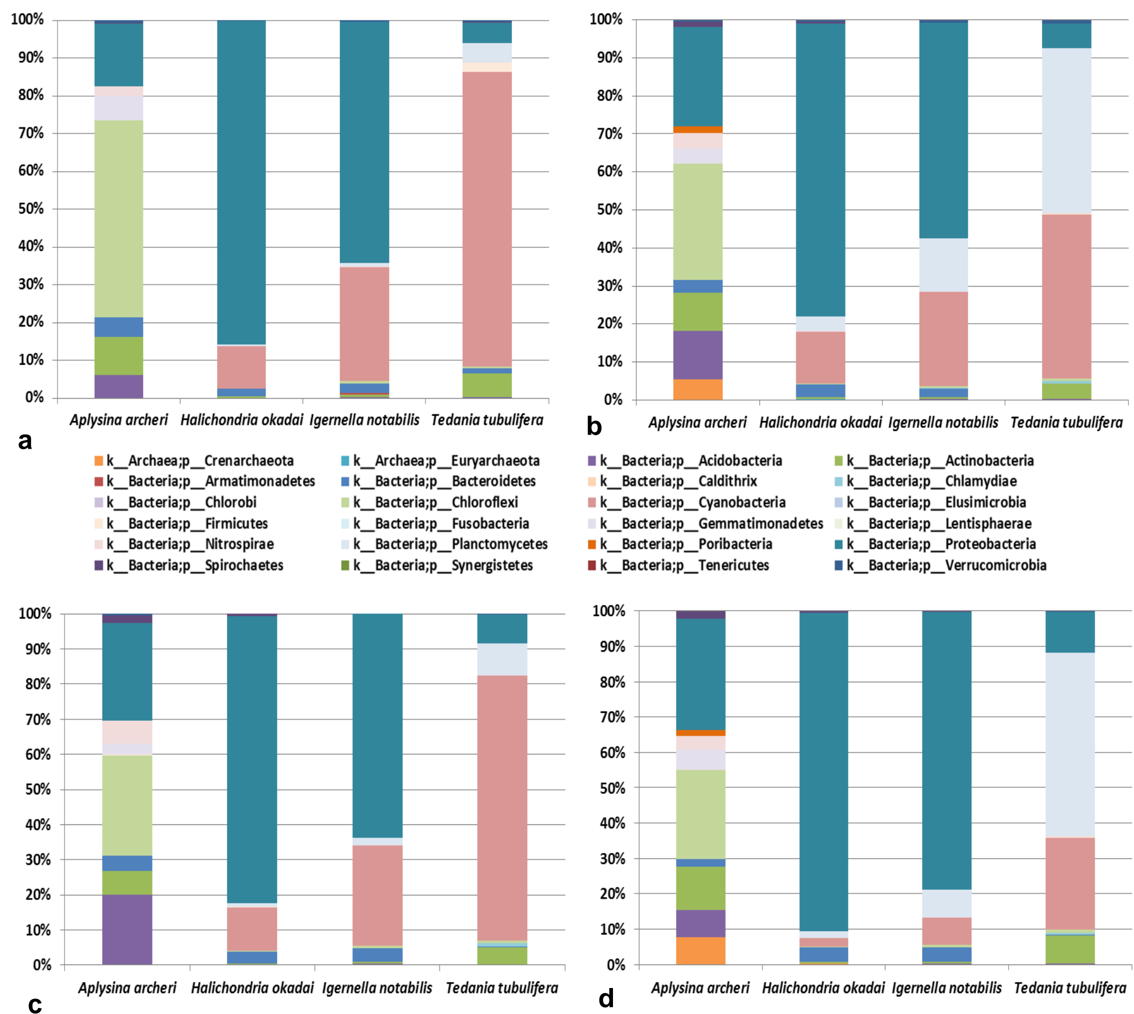


Figure 4 - 6 Microbial community comparison between four sponge species belonging to four orders based on the affiliated (known) OTUs. a. V1V3 region-specific primer set; b. V4 region-specific primer set; c. V4V5 region-specific primer set; d. V5V8 region-specific primer set.

4.3.4 Distinct unaffiliated OTU profiles revealed by different various primer sets

The number of the unaffiliated OTUs, including unassigned and candidate OTUs, showed a substantial discrepancy among the microbial community profiles revealed by different primer sets for a given sponge sample (Table 4-2). For phylum-level microbial OTUs, more than half (53.7%) are unaffiliated using the combined data for four sponge species revealed by four primer sets. For class-level OTUs, a higher percentage of 62.3% among the total number of the OTUs are the unaffiliated. Comparing the unaffiliated OTU profiles generated from four primer sets for one sponge species, we found that they only shared a very small part. If considering the affiliated (known) OTUs together, 46.3% shared OTUs were obtained at the phylum level out of the whole profile for the four sponge species. Moreover, for class-level, even less OTUs

(36.5%) shared among the profiles of four primer sets.

On the other hand, comparing the microbial profiles associated with four sponge species revealed by one primer set, only a small percentage of the unaffiliated OTUs shared between the sponge species. In addition, the proportion of the shared unaffiliated OTUs was much less than the shared affiliated (known) OTUs regardless of the primer selection and the microbial taxonomic levels (refer to the phylum or class levels in this study in Table 4-2).

Importantly, the better capacity of region-specific primer set V5V8 in revealing microbial communities is only applicable for the affiliated (known) microbial OTUs. The four primer sets specific to region V1V3, V4, V4V5, and V5V8 performed at varying degrees of capacity to reveal the unaffiliated OTUs. No single primer set can be considered universal and optimal for all sponge species.

Not only the diverse and unique unaffiliated OTUs were revealed by the four region-specific primer sets, the relative abundance of the unaffiliated OTUs (unassigned and candidate) showed significantly various between the primer sets (Table 4-3). Applying the combination of the four primer sets, the sequence throughput improved greatly (Table 4-4). The total sequence number was calculated by adding the sequences numbers for each distinct OTU and taking the highest sequence number for the shared OTUs. A range of 82.8% to 5.7 times increase in total sequence number was achieved comparing to the performance of the widely used region-specific primer V4. The maximum increase was evaluated by comparing the total combined sequence number and the smallest one among the four datasets (four primer sets), with up to 11.1 times increase for sponge *Halichondria okadai.*

Considering the most comprehensive sponge microbiome study so far as the reference, this study analysed and compared the microbial profiles of the four sponge species using four region-specific primer sets (V1V3, V4, V4V5, V5V8) with the profile generated from 81 species in the study of Thomas et al. (2016) using region V4 (Table 4-5). Using the V4 region only, a total of 18 affiliated (known) OTUs and 26 unaffiliated OTUs at the phylum level were revealed for four sponge species in this study, in comparison with a total of 25 affiliated (known) OTUs and 16 unaffiliated OTUs revealed for 81 sponge species by Thomas et al. (2016). The region-specific primer set V4 in this study had a distinct advantage on revealing the unaffiliated microbial OTUs. Moreover, the combined four primer sets further improved the coverage of the diverse unaffiliated microbial OTUs as 13 more unaffiliated phylum-level OTUs were revealed from four sponge species than those for 81 sponge species (Thomas et al. 2016), with an 80% increase. By focusing the phylum Proteobacteria as an example, V4 primer set failed to reveal microbial Class Zetaproteobacteria that could be covered by the combined dataset (four primer sets) (Table 4-5).

Table 4 - 2 The number of unaffiliated and affiliated OTUs at the phylum/ class level derived from four sponge species revealed by four primer sets

Phylum level OTUs

	Aplysina archeri			*Halichondria okadai*			*Igernella notabilis*			*Tedania tubulifera*			Total	
Primer set	Unaffiliated [a]	Affiliated (known)	Unaffiliated/ total [b] (%)	Unaffiliated	Affiliated (known)	Unaffiliated/ total (%)	Unaffiliated	Affiliated (known)	Unaffiliated/ total (%)	Unaffiliated	Affiliated (known)	Unaffiliated/ total (%)	Unaffiliated	Affiliated (known)
V1V3	4	15	13.8%	9	20	17.6%	9	19	24.3%	6	14	19.4%	11 (4) [c]	21 (12) [c]
V4	6	14	20.7%	24	15	47.1%	12	14	32.4%	11	12	35.5%	26 (4) [c]	18 (11) [c]
V4V5	6	15	31.6%	11	18	21.6%	8	18	21.6%	8	16	25.8%	14 (4) [c]	22 (13) [c]
V5V8	12	16	41.4%	10	24	19.6%	10	19	27.0%	10	18	32.3%	21 (5) [c]	24 (15) [c]
Total	13 (4) [d]	16 (14) [d]	44.8%	26 (4) [d]	25 (12) [d]	51.0%	16 (4) [d]	21 (14) [d]	43.2%	13 (5) [d]	18 (12) [d]	45.2%	29 (7) [c] 29 (9) [d]	25 (15) [c] 25 (16) [d]
Shared OTUs (%)	18/ 29= 62.1%			16/ 51=31.4%			18/ 37 = 48.6%			17/ 31=54.8%			25/ 54=46.3%	

Total No. of unaffiliated phylum level OTUs/ total No. of phylum level OTUs=29/ (29+25) =53.7%

Class level OTUs

	Aplysina archeri			*Halichondria okadai*			*Igernella notabilis*			*Tedania tubulifera*			Total	
Primer set	Unaffiliated	Affiliated (known)	Unaffiliated/ total (%)	Unaffiliated	Affiliated (known)	Unaffiliated/ total (%)	Unaffiliated	Affiliated (known)	Unaffiliated/ total (%)	Unaffiliated	Affiliated (known)	Unaffiliated/ total (%)	Unaffiliated	Affiliated (known)
V1V3	17	24	22.4%	27	39	20.4%	19	33	22.4%	20	28	25.0%	38 (10) [c]	41 (18) [c]
V4	24	20	31.6%	76	32	57.6%	36	28	42.4%	36	23	45.0%	86 (15) [c]	38 (13) [c]
V4V5	25	24	32.9%	36	34	27.3%	28	33	32.9%	22	27	27.5%	53 (10) [c]	42 (18) [c]
V5V8	35	26	46.1%	31	51	23.5%	23	34	27.1%	24	34	30.0%	57 (12) [c]	55 (21) [c]
Total	43 (15) [d]	33 (14) [d]	56.6%	79 (15) [d]	53 (25) [d]	59.8%	46 (12) [d]	39 (25) [d]	54.1%	42 (15) [d]	38 (17) [d]	52.5%	99 (21) [C] 99 (27) [d]	60 (27) [c] 60 (31) [d]
Shared OTUs (%)	29/ 76=38.2%			40/ 132=30.3%			37/ 85=43.5%			32/ 80=40%			58/ 159=36.5%	

Total No. of unaffiliated class level OTUs/ total No. of class level OTUs=99/ (99+60) =62.3%

[a] The unaffiliated OTUs include the candidate phyla and the unassigned OTUs; [b] The total refers to the total No. of the unaffiliated OTUs and the affiliated (known) OTUs revealed by four primer sets for one sponge species; [c] The number in the () refers to the No. of the shared phylum/ class level OTUs among four sponge species; [d] The number in the () refers to the No. of the shared phylum/ class level OTUs revealed by four primer sets.

Table 4 - 3 Sequence abundance (%) of phylum and class level OTUs revealed by four primer sets

Phylum level OTUs

	Aplysina archeri			*Halichondria okadai*			*Igernella notabilis*			*Tedania tubulifera*			Combined		
Primer set	Affiliated (known)	Candidate	Unassigned	Affiliated (known)	Candidate	Unassigned	Affiliated (known)	Candidate	Unassigned	Affiliated (known)	Candidate	Unassigned	Affiliated (known)	Candidate	Unassigned
V1V3	71.82	5.10	23.08	87.16	0.07	12.77	95.14	0.26	4.60	88.75	1.02	10.23	85.72	1.61	12.67
V4	77.16	8.83	14.01	89.87	0.11	10.02	97.07	0.27	2.66	91.67	0.68	7.65	88.94	2.47	8.59
V4V5	76.17	10.31	13.52	87.79	0.08	12.13	97.79	0.36	1.85	96.44	0.75	2.81	89.54	2.88	7.58
V5V8	67.56	10.17	22.27	90.04	0.10	9.86	92.78	0.31	6.91	96.27	0.94	2.79	86.66	2.88	10.46
Combined	73.18	8.60	18.22	88.71	0.09	11.20	95.69	0.30	4.01	93.28	0.85	5.87	87.71	2.46	9.83

Class level OTUs

	Aplysina archeri			*Halichondria okadai*			*Igernella notabilis*			*Tedania tubulifera*			Combined		
Primer set	Affiliated (known)	Candidate	Unassigned	Affiliated (known)	Candidate	Unassigned	Affiliated (known)	Candidate	Unassigned	Affiliated (known)	Candidate	Unassigned	Affiliated (known)	Candidate	Unassigned
V1V3	32.38	40.82	26.80	86.86	0.28	12.86	94.05	1.14	4.81	87.95	1.35	10.70	75.31	10.90	13.79
V4	46.45	30.43	23.12	89.45	0.43	10.12	96.23	0.89	2.88	90.65	1.45	7.90	80.69	8.30	11.01
V4V5	46.08	30.45	23.47	87.40	0.41	12.19	96.88	1.01	2.11	95.24	1.65	3.11	81.40	8.38	10.22
V5V8	44.86	22.17	32.97	89.40	0.30	10.30	90.89	0.78	8.33	95.17	1.75	3.08	80.08	6.25	13.67
Combined	42.44	30.97	26.59	88.28	0.36	11.37	94.51	0.96	4.53	92.25	1.55	6.20	79.37	8.46	12.17

Table 4 - 4 Increased sequences (%) of combined data of four region-specific primer sets based on the total number of the OTUs

	V1V3	V4	V4V5	V5V8	Combination	Minimum increased %	Maximum increased %	Increased compared with V4 %
Aplysina archeri	96338	100584	172013	298920	312792	4.6	224.7	211.0
Halichondria okadai	69573	100280	55452	641381	670862	4.6	1109.8	569.0
Igernella notabilis	80920	100166	96798	162256	183152	12.9	126.3	82.8
Tedania tubulifera	49392	42069	44587	129829	138611	6.8	229.5	229.5

Table 4 - 5 Comparison of the affiliated (known) and unaffiliated phylum-level OTUs between the microbial profiles of four sponge species in this study and 81 species revealed by Thomas et al. (2016)

Affiliated (known) OTUs (phylum)	4 species V4	4 species Combined primer sets	81 species V4	Unaffiliated OTUs (phylum)	4 species V4	4 species Combined primer sets	81 species V4
Crenarchaeota	-	+	+	Bacteria; AncK6	+	+	+
Euryarchaeota	-	+	+	Bacteria; BHI80-139	+	+	-
Acidobacteria	+	+	+	Bacteria; BRC1	+	+	-
Actinobacteria	+	+	+	Bacteria; FBP	+	+	-
Armatimonadetes	+	+	+	Bacteria; GN02	+	+	+
Bacteroidetes	+	+	+	Bacteria; GN04	+	+	+
Caldithrix	-	+	+	Bacteria; GOUTA4	-	+	-
Chlamydiae	-	+	+	Bacteria; KSB3	+	+	-
Chlorobi	-	+	+	Bacteria; LD1	+	+	-
Chloroflexi	+	+	+	Bacteria; MVS-104	-	+	-
Cyanobacteria	+	+	+	Bacteria; NC10	+	+	-
Elusimicrobia	-	+	+	Bacteria; NKB19	+	+	-
Firmicutes	+	+	+	Bacteria; OD1	+	+	+
Fusobacteria	+	+	+	Bacteria; OP1	-	-	+
Gemmatimonadetes	+	+	+	Bacteria; OP11	+	+	-
Lentisphaerae	+	+	+	Bacteria; OP3	+	+	-
Nitrospirae	+	+	+	Bacteria; OP8	-	-	+
Parvarchaeota	-	-	+	Bacteria; Other	+	+	-
Planctomycetes	+	+	+	Bacteria; PAUC34f	+	+	+
Poribacteria	+	+	+	Bacteria; SAR406	+	+	+
Proteobacteria	+	+	+	Bacteria; SBR1093	+	+	+

Spirochaetes	+	+	+	Bacteria; SR1	+	+	-
Synergistetes	+	+	-	Bacteria; TM6	+	+	+
Tenericutes	-	+	+	Bacteria; TM7	+	+	-
Thermi	+	+	+	Bacteria; unassigned	+	+	-
Verrucomicrobia	+	+	+	Bacteria; WPS-2	+	+	+
				Bacteria; WS2	+	+	+
				Bacteria; WS3	+	+	+
				Bacteria; WS5	-	-	+
				Bacteria; WWE1	-	+	-
				Bacteria; ZB3	+	+	+
				Unassigned; Other	+	+	+
Affiliated (known) OTUs (class in phylum Proteobacteria)	**4 species V4**	**4 species Combined primer sets**	**81 species V4**	**Unaffiliated OTUs (class in phylum Proteobacteria)**	**4 species V4**	**4 species Combined primer sets**	**81 species V4**
Alphaproteobacteria	+	+	+	Proteobacteria; Other	+	+	-
Betaproteobacteria	+	+	+	Proteobacteria; TA18	+	+	+
Deltaproteobacteria	+	+	+	Proteobacteria; unassigned	+	+	+
Epsilonproteobacteria	+	+	+				
Gammaproteobacteria	+	+	+				
Zetaproteobacteria	-	+	-				

+/-: Presence/ absence of the OTUs.

4.3.5 Power to uncover the likely missed microbial OTUs

Based on the comparative study for the amplicon sequencing and the shotgun sequencing, some likely missed phylum-level OTUs by the widely used environmental primer set (V4 region: 515F-806R) were reported by Eloe-Fadrosh et al. (2016). In this study, the combined four 16S rRNA primer sets specific to the regions V1V3, V4, V4V5, and V5V8 revealed 30 of the total 52 likely missed phylum-level OTUs as reported in the study of Eloe-Fadrosh et al. (Table 4-6). This improved coverage is only for four sponge samples, proving the powerful capacity of the multiple primer sets combined approach developed in this study.

Table 4 - 6 Likely missed phylum-level OTUs by the V4 region primer set could be revealed by four 16S rRNA primer sets specific to V1V3, V4, V4V5, and V5V8 regions. The likely missed phylum-level OTU list refer to Eloe-Fadrosh et al. (2016)

Likely missed phylum-level OTUs	V1V3	V4	V4V5	V5V8
k__Archaea;p__Crenarchaeota	-	+	+	+
k__Archaea;p__Euryarchaeota	-	+	-	+
k__Bacteria;p__Acidobacteria	+	+	+	+
k__Bacteria;p__Actinobacteria	+	+	+	+
k__Bacteria;p__Armatimonadetes	+	+	+	-
k__Bacteria;p__Bacteroidetes	+	+	+	+
k__Bacteria;p__BRC1	-	+	-	-
k__Bacteria;p__Chlamydiae	-	+	+	+
k__Bacteria;p__Chlorobi	-	+	+	+
k__Bacteria;p__Chloroflexi	+	+	+	+
k__Bacteria;p__Cyanobacteria	+	+	+	+
k__Bacteria;p__Firmicutes	+	+	+	+
k__Bacteria;p__Fusobacteria	+	+	+	+
k__Bacteria;p__Gemmatimonadetes	-	-	-	+
k__Bacteria;p__GN02	+	-	-	+
k__Bacteria;p__Lentisphaerae	+	+	+	+
k__Bacteria;p__Nitrospirae	+	+	+	+
k__Bacteria;p__OD1	+	+	+	+
k__Bacteria;p__OP11	+	+	+	+
k__Bacteria;p__OP3	-	+	+	+
k__Bacteria;p__Planctomycetes	+	+	+	+
k__Bacteria;p__Proteobacteria	+	+	+	+
k__Bacteria;p__Spirochaetes	+	+	+	+
k__Bacteria;p__SR1	-	+	-	-
k__Bacteria;p__Synergistetes	+	-	-	-
k__Bacteria;p__Tenericutes	-	+	+	+
k__Bacteria;p__TM7	+	+	+	+
k__Bacteria;p__Verrucomicrobia	+	+	+	+
k__Bacteria;p__WS3	-	-	-	+
k__Bacteria;p__WWE1	-	-	-	+

+/-: presence/ absence of the OTUs.

4.4 Discussion

4.4.1 Illumina MiSeq: a superior method than 454 GS FLX Titanium sequencing

The first NGS technology- 454 sequencing platform has been utilised increasingly in studying the sponge microbiome since 2010 (De Voogd et al. 2015; Lee et al. 2011; Schmitt et al. 2012). In the last two years, Illumina with higher throughput outpaced 454 to study the sponge microbial community using 16S rRNA gene-based metagenomic sequencing (Easson & Thacker 2014; Luter et al. 2015; Ribes et al. 2015; Rodríguez-Marconi et al. 2015). However, a trade-off between the length and quantity of reads always needs to be considered.

The results in this study indicated that the Illumina MiSeq could readily supersede 454 GS FLX Titanium sequencing (Fig. 4-3). The present analysis first indicated clearly that all the affiliated (known) OTUs of bacteria at the phylum, class, order and family levels revealed by 454 could be re-captured by Illumina MiSeq when using the same primer set. Additionally, other comparative studies also concluded that Illumina MiSeq had remarkable superiority generally in terms of the throughput, though lacking the detailed comparison on the microbial OTUs coverage (Dees et al. 2014; Koskey et al. 2014; Szafranski et al. 2014; Van Treuren et al. 2015). In this study, the direct comparison demonstrated that around four-fold higher OTUs could be revealed by Illumina MiSeq in comparison with 454 platform.

The length and quantity issue also exists in the selection of different Illumina instrument types. It is worth noting that many more researchers have switched to the Illumina MiSeq instead of HiSeq due to its higher reliability for validation, though the HiSeq is more suitable for high-throughput screening (Grimmond, Taft & Miller online resource; Loman et al. 2012).

4.4.2 Validated combination of different region-specific primers essential for sponge microbiome study

The results of this study indicated that the validation is essential to select the effective region-specific primers to reveal a comprehensive sponge microbial community. The V5V8 region-specific primer set had more power to identify the unique microbial OTUs (affiliated) than the other three (Figs 4-4 & 4-5), though about 50% of the total microbial community were left out (Fig. 4-5). Therefore, a combination of the V1V3 (28F-519R) and V5V8 (803F-1392R) region-specific primer sets is advised as a minimum to study sponge microbiome in Illumina MiSeq platform as they could effectively uncover the majority (80%) of the microbial community offered by four primer sets covering the V1- V8 regions of the 16S rRNA gene. Moreover, using another primer set (518F-926R) specific to the V4V5 region in conjunction with the V1V3 and V5V8 region-specific primer sets, at least 90% of the microbial community could be revealed when compared to the combined dataset from four primer sets. Most importantly, this study

showed that using a single prime set alone for one region to profile the microbial community of the complex environmental samples lacks reliability and fidelity. In fact, the widely used V4 region-specific primer set had a much poor performance when identifying affiliated (known) microbial OTUs in this study, though it has been used in most if not all the 16S rRNA gene based metagenomic sequencing, including the EMP (Caporaso et al. 2012; Gilbert, Jansson & Knight 2014), and particularly the sponge microbiome project (Thomas et al. 2016).

Furthermore, using any single amplicon targeting the specific 16S rRNA gene region(s) to analyse and compare sponge microbiomes will probably offer a misleading conclusion as reported in almost all the published literature. As shown in Appendix Table 4-1, less than 2% of the genera revealed by four primer sets were the same when analysing the same sponge DNA sample. Based on the significantly different performances of these primer sets, no single sequencing amplicon could be applied to compare the sponge microbial communities correctly.

Alternatively, one could also propose the hypothesis that regardless of the region, longer reads will improve the efficacy of microbial taxa identification when applying 16S based metagenomic sequencing (Schloss 2010). To address this question, the combined microbial profile revealed by V1V3 and V4 primer sets covering the length of 778 bps, which amplified four hyper-variable regions, was compared with the one revealed by V5V8 (589 bps), also covering four hyper-variable regions. Contrary to the expectation, the V5V8 profile with shorter sequencing reads was found to be superior as it revealed many more unique microbial taxa compared to the longer read (Fig. 4-5). It implies that the length of the sequenced regions may not be the critical factor in revealing microbiomes.

4.4.3 Deeper uncovering of microbial taxa by primers specific to the V5 to V8 regions

The Illumina platform is a developing technique, and reports of the sponge microbiome studies using it are still rare; most of them only utilised the V4 region specific primer set, as was done by the EMP. For the latter, Easson & Thacker (2014) reported the host-specific microbiomes of 20 tropical marine sponges using the V4 region specific primer set (515F-806R). Among the 20 species, *Aplysina cauliformis*, *Aplysina fulva*, and *Tedania ignis*, which belong to the same genera of the sponges in this study, only 10, 10, and four microbial phyla (affiliated and unaffiliated) were revealed compared to 13, 13, and 13 affiliated phyla using V5V8 region in this study, respectively. It suggested that an effective primer set will greatly increase the sequencing depth and the throughput.

Within sponge species, it was found that the second half of the 16S rRNA gene including four hyper-viable regions (V5 to V8) had a better performance on the efficacy of

microbial taxa identification based on both affiliated and unaffiliated OTUs (Reveillaud et al. 2014), which is consistent with the results of this study. An analysis of the microbial diversity of the sponge *X. testudinaria* in relation to the primer set showed V1V2 (27F-338R) revealed the presence of 9-11 microbial phyla (Montalvo et al. 2014); V3V4 revealed 10 microbial phyla (De Voogd et al. 2015). Furthermore, sponge *X. bocatorensis* belonging to the same genus was analysed in the Illumina platform by EMP using V4 (515F-806R) and revealed 11 microbial phyla including two unclassified phyla (Easson & Thacker 2014). However, using the V5V6 region-specific primer set (U789F-U1068R) significantly improved the microbial taxa identification efficacy and revealed 19 microbial phyla from sponge *X. testudinaria* (Lee et al. 2011). In the case of the *Pestrosia ficiformis*, the V3 region (338F-533R) only revealed 13-14 microbial phyla (Schmitt, Hentschel & Taylor 2012), while it was significantly improved by utilising the V6V8 region (926F-1392R) with 19 microbial phyla (Burgsdorf et al. 2014).

Given these results published so far, it is clear that, apart from the differences due to diverse sponge species, the selection of the 16S rRNA gene hyper-variable regions significantly influenced the efficacy of microbial taxa identification and the fidelity of the sponge microbiome. There is currently no accepted consensus of which hyper-variable region offers the least biased view of a microbial community of marine sponges, as clearly no universal hyper-variable region exists. Applying the V5V8 region-specific primer set provided deeper insight into the microbial diversity (affiliated OTUs) of these four sponge species belonging to different orders, and the care must be taken to ensure that the validation is done to select the most effective primer sets for diverse sponge species. Importantly, the validation in this study shows that it would not be reliable to use any single region-specific primer set with the purpose of uncovering the comprehensive microbial community of sponges. The primer set targeting the regions V5 to V8 is strongly recommended for this purpose if a single prime set has to be used.

4.4.4 Better performance of V5V8 in differentiating sponge-specific microbiomes

For microbiome studies, one of the most basic questions is how to differentiate the microbial communities between different samples. For marine sponges, there is no evidence as to best practice for selecting the effective region-specific primers to distinguish the microbial communities belonging to different sponge species. The results in this study showed that the V5V8 region-specific primer set had the better differentiation power as it revealed a higher percentage (77% vs. V1V3- 70%, V4- 59%, and V4V5- 45%) of the sponge species-specific OTUs (affiliated) at genus level (Appendix Table 4-2).

The different proportions of each microbial OTUs rather than the simple presence or absence in the community represent the divergence of the microbial diversity (Fig. 4-6). There

were 147, 149, 184, and 150 affiliated (known) genera recovered from the four sponges *A. archeri*, *H. okadai*, *I. notabilis*, and *T. tubulifera*, respectively, by all four primer sets combined (Fig. 4-4). It implied that a much less differences exist on the number of the microbial OTUs among the sponge species than compared to the significant differences when using any single primer set. Additionally, there were 115 affiliated (known) genera shared between these four sponges when using all four primer sets (Appendix Table 4-1), which represented 78%, 77%, 63%, and 77% of the total number of the affiliated (known) genera revealed from each of these four sponges.

4.4.5 Vast untapped microbial resources revealed by a combination of multiple region-specific primer sets

Focusing on the unaffiliated microbial OTUs, a large proportion of unaffiliated OTUs could be revealed by applying a combination of multiple 16S rRNA gene region-specific primer sets on Illumina MiSeq platform (Table 4-2). These microbial resources are possibly ignored by the singe-primer-set based sequencing approach. The previous studies could only reveal a part of the microbial community and provided an overly underestimated assessment of the highly diverse and abundant sponge microbiomes. In conjunction with the relative abundance of the sequences shown in Table 4-3, it was found that a small number of the sequences could represent a large number of the distinct microbial OTUs. Particularly, compared to a small number of the sequences representing a large number of the OTUs in the category ‘Candidate’, the larger number of the unassigned sequences could be believed to contain many novel OTUs as in this study the unassigned OTUs was only referred to one. Additionally, the sequencing OTU-based throughput was significantly improved using the combined four primer sets (Table 4-4), which offers the possibility to reveal the low abundant microbial OTUs in the complex sponge-microbe association. The discovery of these unaffiliated OTUs will greatly enhance our understanding of the true and complete sponge microbiome to explore the vast untapped marine microbial resources.

Compared to the latest study of sponge microbiome by Thomas et al (2016) focusing on 804 specimens belonging to 81 species globally, our approach provided even better coverage on the unaffiliated OTUs from four sponge species only (Table 4-5). For the likely missed OTUs by the amplicon sequencing reported by Eloe-Fadrosh et al. (2016), the combination of the four region-specific primer sets covering almost the full length of the 16S rRNA gene was proven to be able to detect 58% of them (Table 4-6), which further indicate the outstanding capacity of the proposed approach for amplicon sequencing in this study.

4.5 Conclusion

This study demonstrated all the affiliated (known) taxa revealed by 454 platform could be covered by the Illumina dataset using same primers. The Illumina had at least four-fold higher OTU richness at various microbial taxonomic levels. The selected four 16S rRNA gene region-specific primer sets covering V1-V8 regions have significantly different performances in identifying affiliated (known) microbial OTUs on the Illumina MiSeq platform. It is not reliable to apply a single region-specific primer set to uncover the comprehensive microbial diversity of the complex sponge microbial community. The V5V8 region-specific primers (803F-1392R) performed significantly better on the efficacy of affiliated (known) microbial taxa identification and the differentiation capacity of different sponge microbiomes, though still missed about 4% of the affiliated (known) phylum-level OTUs and 8% affiliated (known) class-level OTUs. The selection of the hyper-viable regions is the dominant factor to determine the efficacy of microbial taxa identification rather than the length or the number of the regions of the sequenced amplicon. Importantly, more than half of the phylum-level OTUs and class-level OTUs are unaffiliated (unassigned and candidate OTUs) using the multiple primer sets. The performances of the primer sets specific to different 16S rRNA gene regions showed significant discrepancy when revealing the unaffiliated microbial OTUs of different sponge species. No universal or optimal primer set exists. The multiple primer sets based amplicon sequencing, covering almost the full length of 16S rRNA gene, has the equivalent capacity with the metagenomic sequencing and could be the future recommended strategy to reveal comprehensive sponge microbiome. An efficient validation is of paramount value to select Illumina sequencing primers and an optimal combination of the specific primer sets for different 16S rRNA gene regions is essential for sponge microbiome analysis and any environmental samples. With a significantly underestimated sponge microbiome in the past literature, the unaffiliated microbial OTUs are the untapped microbial resources with immeasurable value for many discussions and applications. This new paradigm will reveal more comprehensive sponge microbiomes.

References

Adlard, RD & Lester, RJG 1995, 'The Life-Cycle and Biology of Anilocra-Pomacentri (Isopoda, Cymothoidae), an Ectoparasitic Isopod of the Coral-Reef Fish, Chromis-Nitida (Perciformes, Pomacentridae)', *Australian Journal of Zoology*, vol. 43, no. 3, pp. 271-81.

Altschul, SF, Gish, W, Miller, W, Myers, EW & Lipman, DJ 1990, 'Basic local alignment search tool', *Jounal of Molecular Biology*, vol. 215, pp. 403-10.

Alvarez, B, Crisp, MD, Driver, F, Hooper, JNA & Van Soest, RWM 2000, 'Phylogenetic relationships of the family Axinellidae (Porifera: Demospongiae) using morphological and molecular data', *Zoologica Scripta*, vol. 29, pp. 169-98.

Andreakis, N, Luter, HM & Webster, NS 2012, 'Cryptic speciation and phylogeographic relationships in the elephant ear sponge *Ianthella basta* (Porifera, Ianthellidae) from northern Australia', *Zoological Journal of the Linnean Society*, vol. 166, pp. 225-35.

Angermeier, H, Glöckner, V, Pawlik, JR, Lindquist, NL & Hentschel1, U 2012, 'Sponge white patch disease affecting the Caribbean sponge *Amphimedon compressa*', *Diseases of Aquatic Organisms*, vol. 99, pp. 95–102.

Australian Faunal Directory, Australian Biological Resources Study, Canberra., http://www.environment.gov.au/biodiversity/abrs/online-resources/fauna/afd/index.html.

Bartram, AK, Lynch, MDJ, Stearns, JC, Moreno-Hagelsieb, G & Neufeld, JD 2011, 'Generation of Multimillion-Sequence 16S rRNA Gene Libraries from Complex Microbial Communities by Assembling Paired-End Illumina Reads', *Applied and Environmental Microbiology*, vol. 77, pp. 3846-52.

Bell, JJ 2008, 'The functional roles of marine sponges', *Estuarine, Coastal and Shelf Science*, vol. 79, pp. 341-53.

Bennett, S 2004, 'Solexa Ltd', *Pharmacogenomics*, vol. 5, pp. 433-8.

Benson, DA, Karsch-Mizrachi, I, Lipman, DJ, Ostell, J & Sayers, EW 2009, 'GenBank', *Nucleic Acids Research*, vol. 37, no. Database issue, pp. D26-31.

Blanquer, A & Uriz, MJ 2007, 'Cryptic speciation in marine sponges evidenced by mitochondrial and nuclear genes: A phylogenetic approach', *Molecular Phylogenetics and Evolution*, vol. 45, pp. 392-7.

Blunt, JW, Copp, BR, Keyzers, RA, Munro, MHG & Prinsep, MR 2016, 'Marine natural products', *Natural Product Reports*, vol. 33, pp. 382–431.

Borchiellini, C, Chombard, C, Lafay, B & Boury-Esnault, N 2000, 'Molecular systematics of sponges (Porifera)', *Hydrobiologia*, vol. 420, pp. 15-27.

Burgsdorf, I, Erwin, PM, López-Legentil, S, Cerrano, C, Haber, M, Frenk, S & Steindler, L 2014, 'Biogeography rather than association with cyanobacteria structures symbiotic microbial communities in the marine sponge *Petrosia ficiformis*', *Frontiers in Microbiology*, vol. 10, no. 529, pp. 1-11.

Caporaso, JG, Kuczynski, J, Stombaugh, J, Bittinger, k, Bushman, FD, Costello, EK, Fierer, N, Peña, AG, J.K., G, Gordon, JI, Huttley, GA, Kelley, ST, Knights, D, Koenig, JE, Ley, RE, Lozupone, CA, McDonald, D, Muegge, BD, Pirrung, M, Reeder, J, Sevinsky, JR, Turnbaugh, PJ, Walters, WA, Widmann, J, Yatsunenko, T, Zaneveld, J & Knight, R 2010, 'QIIME allows analysis of high-throughput community sequencing data', *Nature Methods*, vol. 7, no. 5, pp. 335–6.

Caporaso, JG, Lauber, CL, Walters, WA, Berg-Lyons, D, Huntley, J, Fierer, N, Owens, SM, Betley, J, Fraser, L, Bauer, M, Gormley, N, Gilbert, JA, Smith, G & Knight, R 2012, 'Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms', *ISME Journal*, vol. 6, no. 8, pp. 1621-4.

Caporaso, JG, Lauber, CL, Walters, WA, Berg-Lyons, D, Lozupone, CA, Turnbaugh, PJ, Fierer, N & Knight, R 2011, 'Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample', *Proceedings of the National Academy of Sciences,* vol. 108, pp. 4516–22.

Cárdenas, P, Pérez, T & Boury-Esnault, N 2012, 'Sponge systematics facing new challenges', *Advances in Marine Biology*, vol. 61, pp. 79-209.

Cárdenas, P, Xavier, JR, Reveillaud, J, Schander, C & Rapp, HT 2011, 'Molecular phylogeny of the Astrophorida (Porifera, *Demospongiae*p) reveals an unexpected high level of spicule homoplasy', *PLoS ONE*, vol. 6, p. e18318.

Carella, M, Agell, G, Cárdenas, P & Uriz, MJ 2016, 'Phylogenetic Reassessment of Antarctic Tetillidae (Demospongiae, Tetractinellida) Reveals New Genera and Genetic Similarity among Morphologically Distinct Species ', *PLoS ONE*, vol. 11, no. 8, p. e0160718.

Chen, W, Zhang, CK, Cheng, Y, Zhang, S & Zhao, H 2013, 'A comparison of methods for clustering 16s rRNA sequences into OTUs', *PLoS ONE*, vol. 8, no. 8, p. e70837.

Claesson, MJ, O'Sullivan, O, Wang, Q, Nikkilä, J, Marchesi, JR, Hauke Smidt, de vos, WM, Ross, RP & O'Toole, PW 2009, 'Comparative Analysis of Pyrosequencing and a Phylogenetic Microarray for Exploring Microbial Community Structures in the Human Distal Intestine', *PLoS ONE*, vol. 4, no. 8, p. e6669.

Claesson, MJ, Wang, Q, O'Sullivan, O, Greene-Diniz, R, Cole, JR, Ross, RP & O'Toole, PW 2010, 'Comparison of two next-generation sequencing technologies for resolving highly complex microbiota composition using tandem variable 16S rRNA gene regions', *Nucleic Acids Research*, vol. 38, no. 22, p. e200.

Croué, J, West, NJ, Escande, M-L, Intertaglia, L, Lebaron, P & Suzuki, MT 2013, 'A single betaproteobacterium dominates the microbial community of the crambescidine-containing sponge Crambe crambe', *Scientific Reports*, vol. 3, p. 2583.

de Paula, TS, Zilberberg, C, Hajdu, E & Lôbo-Hajdu, G 2012, 'Morphology and molecules on opposite sides of the diversity gradient: Four cryptic species of the *Cliona celata* (Porifera, Demospongiae) complex in South America revealed by mitochondrial and nuclear markers', *Molecular Phylogenetics and Evolution*, vol. 62, pp. 529-41.

De Voogd, NJ, Cleary, DFR, Polónia, ARM & Gomes, NCM 2015, 'Bacterial community composition and predicted functional ecology of sponges, sediment and seawater from the thousand islands reef complex, West Java, Indonesia', *FEMS Microbiology Ecology*, vol. 91, no. 4.

DeBiasse, MB & Hellberg, ME 2015, 'Discordance between morphological and molecular species boundaries among Caribbean species of the reef sponge *Callyspongia*', *Ecology and Evolution*, vol. 5, no. 3, pp. 663-75.

Dees, MW, Lysøe, E, Nordskog, B & Brurberg, MB 2014, 'Bacterial communities associated with surfaces of leafy greens: Shift in composition and decrease in richness over time', *Applied and Environmental Microbiology*, vol. 81, no. 4, pp. 1530 –9.

Degnan, PH & Ochman, H 2012, 'Illumina-based analysis of microbial community diversity', *ISME Journal*, vol. 6, no. 1, pp. 183-94.

De Santis, TZ, Hugenholtz, P, Larsen, N, Rojas, M, Brodie, EL, K. Keller, Huber, T, Dalevi, D, Hu, P & G.L., A 2006, 'Greengenes, a chimera-checked 16s rrna gene database and workbench compatible with ARB', *Applied and Environmental Microbiology*, vol. 72, no. 7, pp. 5069-72.

Dohrmann, M, Haen, KM, Lavrov, DV & Wörheide, G 2012, 'Molecular phylogeny of glass sponges (Porifera, Hexactinellida): increased taxon sampling and inclusion of the mitochondrial protein-coding gene, cytochrome oxidase subunit I', *Hydrobiologia*, vol. 687, pp. 11-20.

Doyle, JJ & Dickson, E 1987, 'Preservation of plant samples for DNA restriction endonuclease analysis', *Taxon*, vol. 36, pp. 715-22.

Duran, S, Pascual, M & Turon, X 2004, 'Low levels of genetic variation in mtDNA sequences over the western Mediterranean and Atlantic range of the sponge *Crambe crambe* (Poecilosclerida)', *Marine Biology*, vol. 144, pp. 31-5.

Easson, CG & Thacker, RW 2014, 'Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges', *Frontiers in Microbiology*, vol. 5, no. 532, pp. 1-11.

Edgar, RC 2010, 'Search and clustering orders of magnitude faster than BLAST', *Bioinformatics*, vol. 26, no. 19, pp. 2460-1.

Engelbrektson, A, Kunin, V, Wrighton, KC, Zvenigorodsky, N, Chen, F, Ochman, H & Hugenholtz, P 2010, 'Experimental factors affecting PCR-based estimates of microbial species richness and evenness', *ISME Journal*, vol. 4, no. 5, pp. 642-7.

Eren, AM, Vineis, JH, Morrison, HG & Sogin, ML 2013, 'A Filtering Method to Generate High Quality Short Reads Using Illumina Paired-End Technology', *PLoS ONE*, vol. 8, no. 6, p. e66643.

Erpenbeck, D, Breeuwer, JAJ, Parra-Velandia, FJ & Soest, RWMv 2006, 'Speculation with spiculation?-Three independent gene fragments and biochemical characters versus morphology in demosponge higher classification', *Molecular Phylogenetics and Evolution*, vol. 38, pp. 293-305.

Erpenbeck, D, Breeuwer, JAJ, Velde, HCvd & Soest, RWMv 2002, 'Unravelling host and symbiont phylogenies of halichondrid sponges (Demospongiae, Porifera) using a mitochondrial marker', *Marine Biology*, vol. 141, pp. 377-86.

Erpenbeck, D, Hall, K, Alvarez, B, Büttner, G, Sacher, K, Schätzle, S, Schuster, A, Vargas, S, Hooper, JNA & Wörheide, G 2012, 'The phylogeny of halichondrid demosponges: past and present re-visited with DNA-barcoding data', *Organisms Diversity and Evolution*, vol. 12, pp. 57-70.

Erpenbeck, D, McCormack, GP, Breeuwer, JAJ & Van Soest, RWM 2004, 'Order level differences in the demosponge LSU structure - New insights in an old taxon', *Molecular Phylogenetics and Evolution*, vol. 32, no. 1, pp. 388-95.

Erpenbeck, D, Sutcliffe, P, Cook, S, Dietzel, A, Maldonado, M, Van Soest, RWM, Hooper, JNA & Wörheide, G 2012, 'Horny sponges and their affairs: on the phylogenetic relationships of keratose sponges', *Molecular Phylogenetics and Evolution*, vol. 63, pp. 809-16.

Erpenbeck, D, Weier, T, De Voogd, NJ, Wörheide, G, Sutcliffe, P, Todd, JA & Michel, E 2011, 'Insights into the evolution of freshwater sponges (Porifera: Demospongiae: Spongillina): Barcoding and phylogenetic data from Lake Tanganyika endemics indicate multiple invasions and unsettle existing taxonomy', *Molecular Phylogenetics and Evolution*, vol. 61, pp. 231-6.

Erwin, PM & Thacker, RW 2007, 'Phylogenetic analyses of marine sponges within the order Verongida: a comparison of morphological and molecular data', *Invertebrate Biology*, vol. 126, pp. 220–34.

Flot, J 2007, 'CHAMPURU 1.0: a computer software for unraveling mixtures of two DNA sequences of unequal lengths', *Molecular Ecology Notes*, vol. 7, pp. 974-7.

Flot, J, Tillier, A, Samadi, S & Tillier, S 2006, 'Phase determination from direct sequencing of length-variable DNA regions', *Molecular Ecology Notes*, vol. 6, pp. 627-30.

Folmer, O, Black, M, Hoeh, W, Lutz, R & Vrijenhoek, R 1994, 'DNA primers for amplification of mitochondrial cytochrome *c* oxidase subunit I from diverse metazoan invertebrates', *Molecular Marine Biology and Biotechnology*, vol. 3, no. 5, pp. 294-9.

Fontaneto, D, Flot, J-F & Tang, CQ 2015, 'Guidelines for DNA taxonomy, with a focus on the meiofauna', *Marine Biodivercity*, vol. 45, no. 3, pp. 433-51.

Galtier, N, Gouy, M & Gautier, C 1996, 'SEA VIEW and PHYLO_ WIN: two graphic tools for sequence alignment and molecular phylogeny', *Computer Applications in the Bioscience*, vol. 12, no. 6, pp. 543-8.

Galtier, N, Nabholz, B, Glémin, S & Hurst, GDD 2009, 'Mitochondrial DNA as a marker of molecular diversity: a reappraisal', *Molecular Ecology*, vol. 18, pp. 4541-50.

Geer, LY, Marchler-Bauer, A, Geer, RC, Han, L, He, J, He, S, Liu, C, Shi, W & Bryant, SH 2010, 'The NCBI BioSystems database', *Nucleic Acids Research*, vol. 38, no. suppl 1, pp. D492-D6.

Gerbi, SA 1985, *Evolution of ribosomal DNA*, New York/London.

Ghyselinck, J, Pfeiffer, S, Heylen, K, Sessitsch, A & De Vos, P 2013, 'The effect of primer choice and short read sequences on the outcome of 16s rrna gene based diversity studies', *PLoS ONE*, vol. 8, no. 8, p. e71360.

Gilbert, JA, Jansson, JK & Knight, R 2014, 'The Earth Microbiome project: successes and aspirations', *BMC Biology*, vol. 12, no. 69, pp. 1-4.

Gloor, GB, Hummelen, R, Macklaim, JM, Dickson, RJ, Fernandes, AD, MacPhee, R & Reid, G 2010, 'Microbiome Profiling by Illumina Sequencing of Combinatorial Sequence-Tagged PCR Products', *PLoS ONE*, vol. 5, no. 10, p. e15406.

Grimmond, S, Taft, R & Miller, D '*The switch is on',* Institute for Molecular Bioscience The University of Queensland in Australia, http://www.illumina.com/landing/iswitched.html

Jumpstart Consortium Human Microbiome Project Data Generation Working Group 2012, 'Evaluation of 16S rDNA-Based Community Profiling for Human Microbiome Research', *PLoS ONE*, vol. 7, no. 6, p. e39315.

Gutt, J 2007, 'Antarctic macro-zoobenthic communities: a review and an ecological classification', *Antarctic Science*, vol. 19, no. 2, pp. 165-82.

Haas, BJ, Gevers, D, Earl, AM, Feldgarden, M, Ward, DV, Giannoukos, G, Ciulla, D, D., T, Highlander, SK, Sodergren, E, Methe, B, DeSantis, TZ, Consortium, THM, Petrosino, JF, Knight, R & Birren, BW 2011, 'Chimeric 16s rrna sequence formation and detection in sanger and 454-pyrosequenced pcr amplicons', *Genome Research*, vol. 21, pp. 494-504.

Hall, TA 1999, 'BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT', *Nucleic Acids Symposium Series*, vol. 41, pp. 95-8.

Hamady, M & Knight, R 2009, 'Microbial community profiling for human microbiome projects: Tools, techniques, and challenges', *Genome Research*, vol. 19, pp. 1141–52.

Hamady, M, Walker, JJ, Harris, KJ, Gold, NJ & Knight, R 2008, 'Error-correcting barcoded primers allow hundreds of samples to be pyrosequenced in multiplex', *Nature methods*, vol. 5, no. 3, pp. 235-7.

He, Y, Zhou, B-J, Deng, G-H, Jiang, X-T, Zhang, H & Zhou, H-W 2013, 'Comparison of microbial diversity determined with the same variable tag sequence extracted from two different PCR amplicons', *BMC Microbiology*, vol. 13, p. 208.

Heim, I, Nickel, M & Brümmer, F 2007, 'Molecular markers for species discrimination in poriferans: a case study on species of the genus *Aplysina*', *Porifera Research: Biodiversity, Innovation and Sustainability*, pp. 361-71.

Heim, I, Nickel, M & Brümmer, F 2007, 'Phylogeny of the genus Tethya (Tethyidae: Hadromerida: Porifera): molecular and morphological aspects', *Journal of the Marine Biological Association of the United Kingdom*, vol. 87, pp. 1615–27.

Hentschel, U, Piel, J, Degnan, SM & Taylor, MW 2012, 'Genomic insights into the marine sponge microbiome', *Nature Reviews Microbiology*, vol. 10, pp. 641-54.

Hillis, DM & Davis, SK 1988, 'Ribosomal DNA: Intraspecific Polymorphism, Concerted Evolution, and Phylogeny Reconstruction', *Systematic Zoology*, vol. 37, no. 1, pp. 63-6.

Hooper, JNA 2003, *'Sponguide'. Guide to sponge collection and identification* Queensland Museum, http://www.qm.qld.gov.au.

Hooper, JNA, Hall, KA, Ekins, M, Erpenbeck, D, Wörheide, G & Jolley-Rogers, G 2013, 'Managing and sharing the escalating number of sponge "unknowns": the SpongeMaps project', *Integrative and Comparative Biology*, vol. 33, no. 3, pp. 473-81.

Hooper, JNA & Van Soest, RWM 2002, *Systema Porifera. A guide to the classification of sponges*, 2 vols., Kluwer Academic/Plenum Publishers, New York.

Jünemann, S, Sedlazeck, FJ, Prior, K, Albersmeier, A, John, U, Kalinowski, J, Mellmann, A, Goesmann, A, Haeseler, Av, Stoye, J & Harmsen, D 2013, 'Updating benchtop sequencing performance comparison', *Nature Biotechnology*, vol. 31, no. 4, pp. 294-6.

Kim, M, Morrison, M & Yu, Z 2011, 'Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes', *Journal of microbiological methods*, vol. 84, pp. 81-7.

Klindworth, A, Pruesse, E, Schweer, T, Peplies, J, Quast, C, Horn, M & Glöckner, FO 2012, 'Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies', *Nucleic Acids Research*, pp. 1-11.

Koskey, AM, Fisher, JC, Traudt, MF, Newton, RJ & S.L., M 2014, 'Analysis of the gull fecal microbial community reveals the dominance of *Catellicoccus marimammalium* in relation to culturable enterococci', *Applied and Environmental Microbiology*, vol. 80, no. 2.

Kozich, JJ, Westcott, SL, Baxter, NT, Highlander, SK & Schloss, PD 2013, 'Development of a Dual-Index Sequencing Strategy and Curation Pipeline for Analyzing Amplicon Sequence Data on the MiSeq Illumina Sequencing Platform', *Applied and Enrivonmental Micorbiology*, vol. 79, pp. 5112-20.

Kress, WJ, Wurdack, KJ, Zimmer, EA, Weigt, LA & Janzen, DH 2005, 'Use of DNA barcodes to identify flowering plants', *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 23, pp. 8369-74.

Kumar, P, Brooker, M, Dowd, S & Camerlengo, T 2011, 'Target Region Selection Is a Critical Determinant of Community Fingerprints Generated by 16S Pyrosequencing', *PLoS ONE*, vol. 6, no. 6, p. e20956.

Lavrov, DV, Wang, X & Kelly, M 2008, 'Reconstructing ordinal relationships in the Demospongiae using mitochondrial genomic data', *Molecular Phylogenetics and Evolution*, vol. 49, pp. 111-24.

Leal, MC, Puga, J, Serôdio, J, Gomes, NCM & 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*, vol. 7, p. e30580.

Lee, CK, Herbold, CW, Polson, SW, Wommack, KE, Williamson, SJ, McDonald, IR & Cary, SC 2012, 'Groundtruthing Next-Gen Sequencing for Microbial Ecology–Biases and Errors in Community Structure Estimates from PCR Amplicon Pyrosequencing', *PLoS ONE*, vol. 7, no. 9, p. e44224.

Lee, OO, Wang, Y, Yang, J, Lafi, FF, Al-Suwailem, A & Qian, PY 2011, 'Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea', *ISME Journal*, vol. 5, no. 4, pp. 650-64.

Liao, D 1999, 'Concerted Evolution: Molecular Mechanism and Biological Implications', *The American Journal of Human Genetics*, vol. 64, pp. 24-30.

Loman, NJ, Misra, RV, Dallman, TJ, Constantinidou, C, Gharbia, SE, Wain, J & Pallen, MJ 2012, 'Performance comparison of benchtop high-throughput sequencing platforms', *Nature Biotechnology*, vol. 30, no. 5, pp. 434-9.

Luter, HM, Widder, S, Botté, ES, Wahab, MA, Whalan, S, Moitinho-Silva, L, Thomas, T & Webster, NS 2015, 'Biogeographic variation in the microbiome of the ecologically important sponge, *Carteriospongia foliascens*', *PeerJ*, vol. 3, p. e1435.

Margulies, M, Egholm, M, Altman, WE, Attiya, S, Bader, JS, Bemben, LA, Berka, J, Braverman, MS, Chen, Y-J, Chen, Z, Dewell, SB, Du, L, Fierro, JM, Gomes, XV, Godwin, BC, He, W, Helgesen, S, Ho, CH, Irzyk, GP, Jando, SC, Alenquer, MLI, Jarvie, TP, Jirage, KB, Kim, J-B, Knight, JR, Lanza, JR, Leamon, JH, Lefkowitz, SM, Lei, M, Li, J, Lohman, KL, Lu, H, Makhijani, VB, McDade, KE, McKenna, MP, Myers, EW, Nickerson, E, Nobile, JR, Plant, R, Puc, BP, Ronan, MT, Roth, GT, Sarkis, GJ, Simons, JF, Simpson, JW, Srinivasan, M, Tartaro, KR, Tomasz, A, Vogt, KA, Volkmer, GA, Wang, SH, Wang, Y, Weiner, MP, Yu, P, Begley, RF & Rothberg, JM 2005, 'Genome sequencing in microfabricated high-density picolitre reactors', *Nature*, vol. 437, no. 7057, pp. 376-80.

Medina, M, Collins, AG, Silberman, JD & Sogin, ML 2001, 'Evaluating hypotheses of basal animal phylogeny using complete sequences of large and small subunit rRNA', *Proceedings of*

the National Academy of Sciences of the United States of America, vol. 98, pp. 9707-12.

Montalvo, NF, Davis, J, Vicente, J, Pittiglio, R, Ravel, J & Hill, RT 2014, 'Integration of Culture-Based and Molecular Analysis of a Complex Sponge-Associated Bacterial Community', *PLoS ONE*, vol. 9, no. 3, p. e90517.

Morrow, C & Cárdenas, P 2015, 'Proposal for a revised classification of the Demospongiae (Porifera)', *Frontiers in Zoology*, vol. 12, no. 7, pp. 1-27.

Morrow, CC, Picton, BE, Erpenbeck, D, Boury-Esnault, N, Maggs, CA & Allcock, AL 2012, 'Congruence between nuclear and mitochondrial genes in Demospongiae: a new hypothesis for relationships within the G4 clade (Porifera: Demospongiae)', *Molecular Phylogenetics and Evolution*, vol. 62, no. 1, pp. 174-90.

Morrow, CC, Redmond, NE, Picton, BE, Thacker, RW, Collins, AG, Maggs, CA, Sigwart, JD & Allcock, AL 2013, 'Molecular phylogenies support homoplasy of multiple morphological characters used in the taxonomy of Heteroscleromorpha (Porifera: Demospongiae)', *Integrative and Comparative Biology*, vol. 53, pp. 428–46.

Neefs, JM, Peer, YVd, Rijk, PD, Chapelle, S & Wachter, RD 1993, 'Compilation of small ribosomal subunit RNA structures', *Nucleic Acids Research*, vol. 21, no. 13, pp. 3025-49.

Nelson, MC, Morrison, HG, Benjamino, J, Grim, SL & Graf, J 2014, 'Analysis, Optimization and Verification of Illumina-Generated 16S rRNA Gene Amplicon Surveys', *PLoS ONE*, vol. 9, no. 4, p. e94249.

Nichols, SA 2005, 'An evaluation of support for order-level monophyly and interrelationships within the class Demospongiae using partial data from the large subunit rDNA and cytochrome oxidase subunit I', *Molecular Phylogenetics and Evolution*, vol. 34, pp. 81-96.

Pearson, WR 2013, 'An introduction to sequence similarity (“homology”) searching', *Current Protocols in Bioinformatics*, vol. 6, no. 3, pp. 1-9.

Pillay, KRM, Asahida, T, Chen, CA, Terashima, H & Ida, H 2006, 'ITS Ribosomal DNA Distinctions and the Genetic Structures of Populations of Two Sympatric Species of *Pavona* (Cnidaria: Scleractinia) from Mauritius', *Zoological Studies*, vol. 45, no. 1, pp. 132-44.

Pinto, AJ & Raskin, L 2012, 'PCR biases distort bacterial and archaeal community structure in pyrosequencing datasets', *PLoS ONE*, vol. 7, no. 8, p. e43093.

Plotkin, A, Rapp, HT & Gerasimova, E 2012, 'Phylogenetic reconstruction of Polymastiidae (Demospongiae: Hadromerida) based on morphology', *Hydrobiologia*, vol. 687, pp. 21-41.

Plotkin, A, Voigt, O, Willassen, E & Rapp, HT 2016, 'Molecular phylogenies challenge the classification of Polymastiidae (Porifera, Demospongiae) based on morphology', *Organisms Diversity & Evolution*, pp. 1-22.

Pöppe, J, Sutcliffe, P, Hooper, JNA, Wörheide, G & Erpenbeck, D 2010, 'COI barcoding reveals new clades and radiation patterns of Indo-Pacific sponges of the family Irciniidae (Demospongiae: Dictyoceratida)', *PLoS ONE*, vol. 5, no. 4, p. e9950.

Redmond, NE & McCormack, GP 2009, 'Ribosomal internal transcribed spacer regions are not suitable for intra- or inter-specific phylogeny reconstruction in haplosclerid sponges (Porifera: Demospongiae)', *Journal of the Marine Biological Association of the United Kingdom*, vol. 89,

no. 6 pp. 1251-6.

Redmond, NE, Morrow, CC, Thacker, RW, Diaz, MC, Boury-Esnault, N, Cárdenas, P, Hajdu, E, Lôbo-Hajdu, G, Picton, BE, Pomponi, SA, Kayal, E & Collins, AG 2013, 'Phylogeny and systematics of Demospongiae in light of new small subunit ribosomal DNA (18S) sequences.', *Integrative and Comparative Biology*, vol. 53, pp. 388–415.

Redmond, NE, Raleigh, J, Soest, RWMv, Kelly, M, Travers, SAA, Bradshaw, B, Vartia, S, Stephens, KM & McCormack, GP 2011, 'Phylogenetic Relationships of the Marine Haplosclerida (Phylum Porifera) Employing Ribosomal (28S rRNA) and Mitochondrial (*cox1, nad1*) Gene Sequence Data', *PLoS ONE*, vol. 6, no. 9, p. e24344.

Reveillaud, J, Maignien, L, Eren, AM, Huber, JA, Apprill, A, Sogin, ML & Vanreusel, A 2014, 'Host-specificity among abundant and rare taxa in the sponge microbiome', *ISME Journal*, vol. 8, pp. 1198-209.

Reveillaud, J, Remerie, T, Soest, Rv, Erpenbeck, D, Cárdenas, P, Derycke, S, Xavier, JR, Rigaux, A & Vanreusel, A 2010, 'Species boundaries and phylogenetic relationships between Atlanto-Mediterranean shallow-water and deep-sea coral associated *Hexadella* species (Porifera, Ianthellidae)', *Molecular Phylogenetics and Evolution*, vol. 56, pp. 104-14.

Ribes, M, Dziallas, C, Coma, R & Riemannb, L 2015, 'Microbial diversity and putative diazotrophy in high- and low- microbial-abundance mediterranean sponges', *Applied and Environmental Microbiology*, vol. 81, no. 17, pp. 5683-93.

Riviere, D, Desvignes, V, Pelletier, E, Chaussonnerie, S, Guermazi, S, Weissenbach, J, Li, T, Camacho, P & Sghir, A 2009, 'Towards the definition of a core of microorganisms involved in anaerobic digestion of sludge', *ISME Journal*, vol. 3, no. 6, pp. 700-14.

Rodríguez-Marconi, S, De la Iglesia, R, Díez, B, Fonseca, CA, Hajdu, E & Trefault, N 2015, 'Characterization of bacterial, archaeal and eukaryote symbionts from antarctic sponges reveals a high diversity at a three-domain level and a particular signature for this ecosystem', *PLoS ONE*, vol. 10, no. 9, p. e0138837.

Schloss, PD 2010, 'The Effects of Alignment Quality, Distance Calculation Method, Sequence Filtering, and Region on the Analysis of 16S rRNA Gene-Based Studies', *PLoS Computational Biology*, vol. 6, no. 7, p. e1000844.

Schmalenberger, A, Schwieger, F & Tebbe, CC 2001, 'Effect of Primers Hybridizing to Different Evolutionarily Conserved Regions of the Small-Subunit rRNA Gene in PCR-Based Microbial Community Analyses and Genetic Profiling', *Applied and Environmental Microbiology*, vol. 67, no. 8, pp. 3557-63.

Schmitt, S, Hentschel, U & Taylor, M 2012, 'Deep sequencing reveals diversity and community structure of complex microbiota in five Mediterranean sponges', in M Maldonado, X Turon, M Becerro & M Jesús Uriz (eds), *Ancient Animals, New Challenges*, Springer Netherlands, vol. 219, pp. 341-51.

Schmitt, S, Tsai, P, Bell, J, Fromont, J, Ilan, M, Lindquist, N, Perez, T, Rodrigo, A, Schupp, PJ, Vacelet, J, Webster, N, Hentschel, U & Taylor, MW 2012, 'Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges', *ISME Journal*, vol. 6, no. 3, pp. 564-76.

Shearer, TL, Oppen, MJHV, Romano, SL & Wörheide, G 2002, 'Slow mitochondrial DNA

sequence evolution in the Anthozoa (Cnidaria)', *Molecular Ecology*, vol. 11, pp. 2475-87.

Simister, RL, Schmitt, S & Taylor, MW 2011, 'Evaluating methods for the preservation and extraction of DNA and RNA for analysis of microbial communities in marine sponges', *Journal of Experimental Marine Biology and Ecology*, vol. 397, no. 1, pp. 38-43.

Szafranski, SP, Wos-Oxley, ML, Vilchez-Vargas, R, Jáuregui, R, Plumeier, I, Klawonn, F, Tomasch, J, Meisinger, C, Kühnisch, J, Sztajer, H, Pieper, DH & Wagner-Döblera, I 2014, 'High-resolution taxonomic profiling of the subgingival microbiome for biomarker discovery and periodontitis diagnosis', *Applied and Environmental Microbiology*, vol. 81, no. 3, pp. 1047-58.

Szitenberg, A, Becking, LE, Vargas, S, Fernandez, JCC, Santodomingo, N, Wörheide, G, Ilan, M, Kelly, M & Huchon, D 2013, 'Phylogeny of Tetillidae (Porifera, Demospongiae, Spirophorida) based on three molecular markers', *Molecular Phylogenetics and Evolution*, vol. 67, pp. 509-19.

Taylor, MW, Schupp, PJ, Dahllöf, I, Kjelleberg, S & Steinberg, PD 2004, 'Host specificity in marine sponge-associated bacteria, and potential implications for marine microbial diversity', *Environmental Microbiology*, vol. 6, no. 2, pp. 121-30.

Thacker, RW, Hill, AL, Hill, MS, Redmond, NE, Collins, AG, Morrow, CC, Spicer, L, Carmack, CA, Zappe, ME, Pohlmann, D, Hall, C, Diazjj, MC & Bangalore, PV 2013, 'Nearly complete 28S rRNA gene sequences confirm new hypotheses of sponge evolution', *Integrative and Comparative Biology*, vol. 53, no. 3, pp. 373-87.

Tremblay, J, Singh, K, Fern, A, Kirton, ES, He, S, Woyke, T, Lee, J, Chen, F, Dangl, JL & Tringe, SG 2015, 'Primer and platform effects on 16S rRNA tag sequencing', *Frontiers in Microbiology*, vol. 6.

Usher, KM, Sutton, DC, Toze, S, Kuo, J & Fromont, J 2004, 'Biogeography and phylogeny of *Chondrilla* species (Demospongiae) in Australia', *Marine Ecology Progress Series*, vol. 270, pp. 117-27.

Van Soest, RWM, Boury-Esnault, N, Hooper, JNA, Rützler, K, De Voogd, NJ, Alvarez de Glasby, B, Hajdu, E, Pisera, AB, Manconi, R, Schoenberg, C, Klautau, M, Picton, B, Kelly, M, Vacelet, J, Dohrmann, M, Díaz, M-C, Cárdenas, P & Carballo, JL 2016, *World Porifera database*, http://www.marinespecies.org/porifera, 2016-11-03.

Van Treuren, W, Ponnusamy, L, Brinkerhoff, RJ, Gonzalez, A, Parobek, CM, Juliano, JJ, Andreadis, TG, Falco, RC, Ziegler, LB, Hathaway, N, Keeler, C, Emch, M, Bailey, JA, Roe, RM, Apperson, CS, Knight, R & Meshnickn, SR 2015, 'Variation in the microbiota of ixodes ticks with regard to geography, species, and sex', *Applied and Environmental Microbiology*, vol. 81, no. 18, pp. 6200–9.

Vargas, S, Kelly, M, Schnabel, K, Mills, S, Bowden, D & Wörheide, G 2015, 'Diversity in a cold hot-spot: DNA-barcoding reveals patterns of evolution among Antarctic Demosponges (class Demospongiae, phylum Porifera)', *PLoS ONE*, vol. 10, no. 6, p. e0127573.

Vargas, S, Schuster, A, Sacher, K, Büttner, G, Schätzle, S, Läuchli, B, Hall, K, Hooper, JNA, Erpenbeck, D & Wörheide, G 2012, 'Barcoding sponges: An overview based on comprehensive sampling', *PLoS ONE*, vol. 7, no. 7, p. e39345.

Voigt, O, Eichmann, V & Wörheide, G 2012, 'First evaluation of mitochondrial DNA as a marker for phylogeographic studies of Calcarea: a case study from *Leucetta chagosensis*', *Hydrobiologia*, vol. 687, pp. 101-6.

Wahab, MAA, Fromont, J, Whalan, S, Webster, N & Andreakis, N 2014, 'Combining morphometrics with molecular taxonomy: How different are similar foliose keratose sponges from the Australian tropics?', *Molecular Phylogenetics and Evolution*, vol. 73, pp. 23-9.

Wang, Y & Qian, P-Y 2009, 'Conservative Fragments in Bacterial 16S rRNA Genes and Primer Design for 16S Ribosomal DNA Amplicons in Metagenomic Studies', *PLoS ONE*, vol. 4, no. 10, p. e7401.

Webster, NS & Taylor, MW 2012, 'Marine sponges and their microbial symbionts: love and other relationships', *Environmental Microbiology*, vol. 14, no. 2, pp. 335-46.

Woese, CR 1987, 'Bacterial evolution', *Microbiological Reviews*, vol. 51, no. 2, pp. 221-71.

Wörheide, G 2006, 'Low variation in partial cytochrome oxidase subunit I (COI) mitochondrial sequences in the coralline demosponge *Astrosclera willeyana* across the Indo-Pacific', *Marine Biology*, vol. 148, pp. 907-12.

Wörheide, G & Erpenbeck, D 2007, 'DNA taxonomy of sponges—progress and perspectives', *Journal of the Marine Biological Association of the United Kingdom*, vol. 87, pp. 1629-33.

Wulff, J 2012, 'Ecological interactions and the distribution, abundance, and diversity of sponges', in *Advances in Marine Biology*, vol. 61, pp. 273-344.

Xavier, JR, Rachello-Dolmen, PG, Parra-Velandia, F, Schönberg, CHL, Breeuwer, JAJ & van-Soest, RWM 2010, 'Molecular evidence of cryptic speciation in the ‘‘cosmopolitan” excavating sponge Cliona celata (Porifera, Clionaidae)', *Molecular Phylogenetics and Evolution*, vol. 56, pp. 13-20.

Yang, Q, Franco, CMM & Zhang, W 2015, 'Sponge-associated actinobacterial diversity: validation of the methods of actinobacterial DNA extraction and optimization of 16S rRNA gene amplification', *Applied Microbiology and Biotechnology*, vol. 99, no. 20, pp. 8731-40.

CHAPTER 5: COMBINATION OF MULTIPLE 16S rRNA GENE REGION-SPECIFIC PRIMER SETS REVEALS COMPREHENSIVE SPONGE MICROBIOME AND UNTAPPED MICROBIAL RESOURCE

Marine sponges (phylum Porifera) are enriched by abundant and diverse microorganisms that make up to 60% of the mesohyl volume. These microbial communities present a species-specific relationship to their host. However, to define the host-specificity, a deeper comparison and analysis of the sponge microbial community on their structure and composition are required. Moreover, the co-evolution needs to be tested at higher taxonomic levels. In this study, the microbial communities of 19 sponge species belonging to four orders (Dendroceratida, Poecilosclerida, Verongida and Suberitida) were investigated to compare the structure and the diversity at the sponge order level, and to test if there is a sponge order-specific community. Another 14 sponge species belonging to five families (Dysideidae, Irciniidae, Spongiidae, Thorectidae, and Verricillitidae) of the fifth order Dictyoceratida were analysed to test if there is a sponge specific microbial community at the family level. A multiple primer sets specific to the regions V1V3, V4V5, and V5V8 of the 16S rRNA gene were utilised for amplicon sequencing by the Illumina MiSeq platform. The results confirmed the previous conclusion in Chapter 4 and further demonstrated that the biased single region-specific primer set based amplicon sequencing greatly influenced the reliability of the microbial profiles. The proposed combination of multiple 16S rRNA gene region-specific primer sets provided the unsurpassed capacity to reveal the comprehensive sponge microbiome. Seventy-one phylum-level microbial OTUs were revealed, including 32 affiliated (known) OTUs and 39 unaffiliated OTUs (one unassigned and 38 candidate OTUs). This contrasts with a total number of 41 phylum-level OTUs reported from all sponges studied to date. Based on the integrated sequencing data, the sponge microbial community showed the specificity at both the order level and family level. The specificity was more about the structure (relative abundance of each microbial OTU) than the composition of the microbial taxa within a given community. Different sponge orders and families have specific dominant microbial OTUs (relative abundance >1%), which are commonly shared between the sponge species within the same order or family. These microbial OTUs are specific to the sponge order and family. Importantly, each sponge order or family has unique microbial OTUs, though some of them are not dominant. These unique OTUs are the signature taxa for a particular order or family. Most unique microbial OTUs are unaffiliated, which are a promising resource of the untapped microorganisms.

5.1 Introduction

Sponges (phylum Porifera) represent a significant component of benthic communities existing globally, not only for their biomass but also for their potential influence to benthic or pelagic ecosystems (Hentschel et al. 2012). More than 8,700 described species to date have been reported from marine and freshwater systems and are found across tropical, temperate, and polar regions (Van Soest et al. 2016). Sponges are among the most ancient living Metazoa and generally form symbiotic relationships with complex communities of microorganisms (Hentschel et al. 2012; Taylor et al. 2007; Thomas et al. 2016). Up to 60% of the tissue volume of certain sponge species consists of microorganisms with a density exceeding 10^9 microbial cells per ml of sponge tissue, orders of magnitude greater than that found in surrounding seawater or sediment (De Voogd et al. 2015). Sponges can maintain highly diverse, yet specific symbiont communities, despite the constant influx of microorganisms in seawater resulting from their filter-feeding activities (Hentschel et al. 2002; Taylor et al. 2013). The associated microbial communities can provide nutrients and secondary metabolites that help sponge defence against predation or diseases (Reveillaud et al. 2014). Collectively, marine sponges and their microbiomes synthesize an impressive number of metabolic products that not only contribute to their nutritional ecology but also have elicited the interest of the pharmaceutical industry owing to their production of an unprecedented number of biologically active compounds.

The sponge microbiome has been extensively characterised by 16S rRNA gene based metagenomic analysis by various molecular tools, such as 454 pyrosequencing, Illumina MiSeq and HiSeq sequencing platforms (Reveillaud et al. 2014; Rodríguez-Marconi et al. 2015; Schmitt et al. 2012). The comparative work was hindered due to methodological differences in sampling, sample processing and data analyses (Thomas et al. 2016). As discussed earlier the Chapter 4, the V4 region of the 16S rRNA gene was applied in Earth Microbiome Project (Gilbert, Jansson & Knight 2014), though we demonstrated that it provides limited coverage of the microbial community. Different regions revealed significantly distinct profiles of the OTU richness (Beckers et al. 2016; Ghyselinck et al. 2013; Kim, Morrison & Yu 2011; Tremblay et al. 2015). Therefore, the combination of multiple primer sets targeting different specific regions of 16S rRNA gene is essential to gain reliable and comprehensive microbiome.

The main conclusion on the sponge microbiome study is that the communities present host-species specificity, e.g. Easson & Thacker 2014; Hentschel et al. 2012; Lee et al. 2011; Thomas et al. 2016). The structural diversity of the microbial communities was found to be highly correlated with the phylogeny of the sponge species (Cuvelier et al. 2014; Reveillaud et al. 2014). Moreover, the microbial communities show the host species-specific structure and

remarkable stability under the large variations of temperature and irradiance (Erwin et al. 2012). Thus, the concept of 'core', 'variable', and 'species-specific' bacterial communities was established to describe sponge microbiomes (Schmitt et al. 2012). In a global effort toward sponge microbiome survey, Thomas et al. (2016) further demonstrated that the core sponge microbiomes are stable and characterised by generalist symbionts. Symbionts that are phylogenetically unique to sponges do not disproportionally contribute to the core microbiome, and host phylogeny impacts complexity rather than composition of the symbiont community. On the other hand, environmental factors, such as biogeographic location and season, could impact the microbial communities for the same species (Burgsdorf et al. 2014; Luter et al. 2015; Weigela & Erwin 2016; White et al. 2012). Several studies also discussed the impact of the environmental factors, such as seawater temperatures (Cebrian et al. 2011; Webster et al. 2008), heavy metals (Selvin et al. 2009; Webster et al. 2001), on the microbial communities. However, to gain critical insights into the complexity and completeness of sponge microbiomes, a greater understanding of the structure and the diversity of the microbial communities in closely related host species at higher taxonomic ranks (such as, family and order levels) is required.

Here we utilised the newly developed approach in Chapter 4 of combining multiple primer sets targeting various 16S rRNA gene regions (V1V3, V4V5, and V5V8) on Illumina MiSeq platform, and provide a comprehensive analysis of the microbial communities. The 33 sponge species, belonging to 32 genera of 19 families, were selected to represent five different taxonomic orders. They were collected from the same biogeographic location of South Australian and the same season to minimise the environmental influences. The aims are to compare the diversity and the relative abundance of the associated microbial OTUs between different sponge orders as well as between different families in one order to uncover the structural and compositional specificity of microbial community at sponge order and family levels.

5.2 Materials and methods

5.2.1 Sponge collection and community DNA extraction

Thirty-three sponge species analysed in this chapter are listed in Table 5-1, following a strict experimental design to test the sponge-microbe specificity at the order and family levels. The specimens were collected via scuba diving at the depths of 4-15 metres at Rapid Bay, Adelaide, South Australia (35°31'16.6"S, 138°11'07.5"E) in February and March, 2015. Each specimen was kept separately using a sterile plastic bag in an ice box during transport. Specimens were flushed with 0.22 µm membrane-filtered seawater to remove loosely attached

microbes and debris. A10cm^3 sponge tissue for each specimen was cut into small parts with a sterile blade and stored in -80ºC freezer for subsequent DNA extraction.

The DNA extraction method utilised in this study is the CTAB-based method (Schmitt et al. 2012; Taylor et al. 2004) with modificaiton described in Chapter 3. Briefly, the freeze dried sponge tissues were ground and suspended in the sterile distilled water for one hour, and after the low speed (600 × g) centrifugation the tissue deposited at the bottom of the tube was collected. The CTAB extraction buffer was applied to lyse tissues, which was combined with polyvinylpyrrolidone (PVP) and β-mercaptoethanol to help remove phenolic compounds and clean tannins in the extract. A bead-beating step using 1.0mm diameter silica beads (Biospec Products) was applied to increase the DNA release (Simister, Schmitt & Taylor 2011). The purified DNA was resuspended in 35 µl of sterile distilled water. Purity and quantity of DNA were determined with a NanoDrop ND-1000 Spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The samples were extracted in duplicate. The quantified DNA sample for each specimen (A260/280: 1.8-2.0; Con. > 100 ng/µl) was divided into two aliquots of equal volume. They were kept at -20ºC for subsequent PCR reactions and sequencing on Illumina MiSeq (2 × 300 bp paired-end reads).

Table 5 - 1 Details of sponges collected from Rapid Bay in South Australia for this study

Sponge identity				Sampling
Order	**Family**	**Putative Species/ genus [a]**	**Accession No.**	**Depths (m)**
Poecilosclerida	Tedaniidae	*Tedania tubulifera*	KJ620377	4-5
	Crellidae	*Crella incrustans*	KC869608	4-5
	Hymedesmiidae	*Phorbas bihamiger*	KJ546366	7-8
	Microcionidae	*Clathria prolifera*	KJ546353	6-7
	Mycalidae	*Mycale setosa*	KJ620392	6-7
	Myxillidae	*Myxilla* cf. *rosacea*	KC883686	7-8
Suberitida	Hachondriidae	*Halichondria okadai*	KJ546365	8-9
	Stylocordylidae	*Stylocordyla chupachups*	LN850243	8-9
	Suberitidae	*Pseudosuberites nudus*	LN850224	11-12
		Rhizaxinella sp.	KJ546357	11-12
		Suberites sp.	KJ620381	4-5
Verongida	Aplysinidae	*Aplysina archeri*	KJ620395	4-5
		Aplysina sp.	KC869638	4-5
	Aplysinellidae	*Aplysinella rhax*	KP026315	14-15

		Suberea creba	KC869606	14-15
Dendroceratida	Dictyodendrillidae	*Igernella notabilis*	KJ620376	5-6
	Darwinellidae	*Aplysilla sulfurea*	EF646837	5-6
		Dendrilla sp.	KU533858	5-6
	Halisarcidae	*Halisarca* sp.	KC869621	14-15
Dictyoceratida	Dysideidae	*Euryspongia lobata*	KC869651	5-10
		Dysidea sp.	KC706752	5-10
		Lamellodysidea herbacea	KC869535	5-10
	Irciniidae	*Ircinia felix*	KJ801661	4-8
		Psammocinia halmiformis	JQ082837	4-8
		Sarcotragus sp.	EF646841	4-8
	Spongiidae	*Spongia* sp.	KU060616	4-8
		Rhopaloeides odorabile	EU644447	4-8
		Hyattella intestinalis	KC869547	4-8
	Thorectidae	*Aplysinopsis* sp.	KC869644	7-9
		Carteriospongia foliascens	KC869574	7-9
		Hyrtios altus	KC869646	7-9
		Thorectandra excavatus	JQ082845	7-9
	Verricillitidae	*Vaceletia* sp.	AM900018	7-9

[a] The closest match in the NCBI database. When the similarity (98%) and E-value (0.0) are confirmed, the species name was assigned to the sequenced gene and then submitted to the NCBI database. If the similarity and E-value failed to meet the threshold, the sponge affiliation was confirmed at the genus level and submitted with the sequence in the NCBI. The Accession No. of each sequence is shown in the next column.

5.2.2 Illumina MiSeq amplicon library and sequencing

The V1V3, V4V5, and V5V8 regions were applied in Illumina MiSeq following the newly developed amplicon sequencing approach for sponge microbiome study, which was proposed and validated in Chapter 4. The primers for 16S rRNA gene region V1V3 are 28F-519R (28F: 5'-GAG TTT GAT CNT GGC TCA G-3'; 519R: 5'-GTN TTA CNG CGG CKG CTG-3') (Croué et al. 2013), 518F-926R for the V4V5 region (518F: 5'-CCA GCA GCY GCG GTAAN-3'; 926R: 5'-CCG TCA ATT CNT TTR AGT-3') (Nelson et al. 2014), and 803F-1392R for the V5V8 region (803F: 5'-TTA GAN ACC CNN GTA GTC-3'; 1392R: 5'-ACG GGC GGT GWG TRC-3') (Engelbrektson et al. 2010). PCR was performed based on the protocol presented in Caporaso et al. (2011). Briefly, both forward and reverse primers were added to the 5' and 3' Illumina adapter, respectively. The employed PCR reaction conditions were 2 mM $MgCl_2$, 0.2 μM each primer and 200 μM dNTPs. The PCR conditions were 94°C for 3 mins, followed by 94°C for 45 s, 50°C for 60 s, 72°C for 90 s in 35 cycles, and a final elongation step at 72°C for 10 mins.

Negative controls and triplicate amplification were applied. After the purification and quantification, the amplicons (A260/280: 1.8-2.0) were pooled and subsequently sequenced on an Illumina MiSeq Sequencer from both ends of paired-end library preparations (2× 300 bps), using sequencing kit version 3.0 followed by base-calling using the GAPipeline version 1.4.0.

5.2.3 Sequencing data processing

5.2.3.1 Demultiplex and quality filter reads

The sequences generated by the Illumina MiSeq dataset was processed in quantitative insights into microbial ecology (QIIME) pipeline (Caporaso et al. 2010). The R1 and R2 files belonging to each sample were merged before processing demultiplex. Under the command *split_libraries.py*, the multiplexed reads were assigned to samples based on their nucleotide barcode (demultiplexing). This step also performed quality filtering based on the characteristics of each sequence, and removed any low quality or ambiguous reads. Specially, the first 20 bases of all fastq files were trimmed to remove the primer sequence, and quality trimmed to remove poor quality sequence using a sliding window of 4 bases with an average base quality above 15 using the software Trimmomatic (Version 0.35). All reads were then hand trimmed to 250 bases, and any with less than 250 bases excluded. Fastq files were finally converted to fasta files for next step analysis. The filtered data in fasta files were used for the following analysis.

5.2.3.2 Closed reference OTU picking

Using the output file (*.fna) from the command of split_libraries.py run the command of pick_closed_reference_otus.py. In closed-reference OTU picking, input sequences were aligned to pre-defined cluster centroids in a reference database. The input sequence that did not match any reference sequences at a pre-defined identity threshold (%) was excluded. The advantages of closed-reference are that closed-reference OTUs give accurate taxonomy assignment and can be used to compare different regions of the same gene (Rideout et al. 2014).

Reads were assigned to OTUs using a closed-reference OTU picking protocol of the QIIME toolkit (Caporaso et al. 2010), where uclust (Edgar 2010) was applied to search sequences against a subset of the Greengenes database (DeSantis et al. 2006), filtered at 97% identity. Reads were assigned to OTUs based on their best hit to this database at greater than or equal to 97% sequence identity. Reads that did not match a reference sequence were discarded. The resulting OTU table was filtered to remove any OTU with an abundance of less than 0.05%. Representative OTU sequences were then BLASTed against the reference database (Greengenes). After aligning OTU representative sequences and filtering the

alignment, the OTU table was made in QIIME. The table (*. biom) summarised the OTU abundances in each sample with taxonomic identifiers for each OTU. It was visualised by running the command summarize_taxa_through_plots.py. The script will generate new tables at various taxonomic levels.

5.2.3.3 Alpha diversity and rarefaction plots

Workflow alpha_rarefaction.py was applied to compute he alpha diversity to evaluate the microbial community within the given sample (Caporaso et al. 2010). This analysis performed the following four steps: generate rarefied OTU tables (multiple_rarefactions.py); compute measures of alpha diversity for each rarefied OTU table (alpha_diversity.py); collate alpha diversity results (collate_alpha.py); and generate alpha rarefaction plots (make_rarefaction_plots.py). The metrics were selected as following three: PD_whole_tree, chao1, observed_otus.

5.2.3.4 Beta diversity and ordination plots

Beta diversity was computed for the analysis between all pairs of samples in the study following the QIIME pipeline (Caporaso et al. 2010). It represents the explicit comparison of microbial communities based on their composition. Beta diversity metrics thus assess the differences between microbial communities. The fundamental output of these comparisons was visualised by Principal Coordinates Analysis (PCoA). The metrics of weighted and unweighted UniFrac were applied. To perform this analysis, the command beta_diversity_through_plots.py was run, which performed the following four steps: rarefy OTU table to remove sampling depth heterogeneity (single_rarefaction.py); compute beta diversity (beta_diversity.py); run Principal Coordinates Analysis (principal_coordinates.py); and generate Emperor PCoA plots (make_emperor.py).

5.2.3.5 3-D biplots

Add microbial taxa to a 3-D PCoA plot using Emperor's make_emperor.py (Caporaso et al. 2010). The coordinates of a given taxon were plotted as a weighted average of the coordinates of all samples, where the weights were the relative abundances of the taxon in the samples. The size of the sphere representing a taxon was proportional to the mean relative abundance of the taxon across all samples.

5.2.4 Performing Procrustes Analysis

The data generated from each primer set was analysed following the pipeline mentioned in the section 5.2.3. The three output files (unweighted_unifrac_pc.txt, generated from beta_diversity.py) for three primer sets were applied for the Procrustes analysis (Gower 1975).

Transform the second and the third coordinates set by rotating, scaling, and then translate it to minimize the distances between corresponding points in these three shapes. This is done with transform_coordinate_matrices.py. The results were visualised using QIIME by running make_emperor.py. The three sets of the coordinates were plotted in the resulting figure, with bars connecting the corresponding points from each data set. Similarly, the three output files (weighted_unifrac_pc.txt, generated from beta_diversity.py) were analysed as above.

5.2.5 Phylogenetic analysis of sponges

Phylogenetic analysis of the 33 sponge species followed the protocol developed in Chapter 2. The 28S rRNA gene used for identification were submitted to NCBI Genbank (Benson et al. 2009). Software BioEdit (Hall 1999) was applied to combine the 28S rRNA gene sequences. Software MEGA6 (Galtier, Gouy & Gautier 1996) was used to align the sequences using ClustalW algorithm, and trim the aligned sequences to make them ready for tree construction. The alignment was exported as a MEGA format file. Neighbor Joining method was utilised to construct the phylogenetic tree using 1000 bootstrap replications and p-distance model.

5.3 Results

5.3.1 Powerful multiple-region based sequencing approach

To test if applying the same primer set to duplicates of the same sponge samples shows consistent microbial profiles, four sponge species belonging to different orders (two replicates each) were separately analysed by two primer sets for regions V1V3 and V5V8. The reproducible results in the beta diversity analysis (Appendix Figure 5-1 and 5-2) demonstrated the reliability of the applied approach in this chapter.

Using three pairs of 16S rRNA gene region-specific primers (V1V3, V4V5, V5V8), 71 phylum-level OTUs were revealed from 33 sponge species. The comparison with the single V4 region based study of 81 sponge species (Thomas et al. 2016) is shown in Table 5-2. Thirty-two affiliated (known) phylum-level OTUs and 39 unaffiliated OTUs (one unassigned and 38 candidate phylum-level OTUs) were revealed from this study using a combination of multiple primer sets targeting V1-V8 regions. This unprecedented capacity of revealing sponge microbiome is in contrast with 25 affiliated and 16 unaffiliated OTUs (phylum level) revealed by V4 primer set only from 81 species of a few hundred samples in the global Sponge Microbiome Project (Thomas et al. 2016). For the classes in phylum Proteobacteria, there were nine OTUs revealed from the 33 sponge species using the new approach, but only seven OTUs revealed from 81 sponge species using a single V4 region-specific primer set (Thomas et al. 2016).

Table 5 - 2 Comparison of the affiliated (known) and unaffiliated phylum-level OTUs for the microbial profiles of 33 sponge species revealed by a combination of multiple 16S rRNA gene regions (V1V3, V4V5, V5V8) in this study and 81 species revealed by V4 region only in the study of Thomas et al. (2016)

Affiliated (known) OTUs (phylum)	33 species (V13,V45,V58)	81 species V4	Unaffiliated OTUs (phylum)	33 species (V13,V45,V58)	81 species V4
Crenarchaeota	+	+	Bacteria; AD3	+	-
Euryarchaeota	+	+	Bacteria; AncK6	+	+
Nanoarchaeota	+	-	Bacteria; BHI80-139	+	-
Parvarchaeota	+	+	Bacteria; BRC1	+	-
Acidobacteria	+	+	Bacteria; FBP	+	-
Actinobacteria	+	+	Bacteria; GN02	+	+
Aquificae	+	-	Bacteria; GN04	+	+
Armatimonadetes	+	+	Bacteria; GOUTA4	+	-
Bacteroidetes	+	+	Bacteria; Kazan-3B-28	+	-
Caldithrix	+	+	Bacteria; KSB3	+	-
Caldithrix-2	+	-	Bacteria; LCP-89	+	-
Chlamydiae	+	+	Bacteria; LD1	+	-
Chlorobi	+	+	Bacteria; MVS-104	+	-
Chloroflexi	+	+	Bacteria; NC10	+	-
Cyanobacteria	+	+	Bacteria; NKB19	+	-
Deferribacteres	+	-	Bacteria; OD1	+	+
Elusimicrobia	+	+	Bacteria; OP1	+	+
Fibrobacteres	+	-	Bacteria; OP11	+	-
Firmicutes	+	+	Bacteria; OP3	+	-
Fusobacteria	+	+	Bacteria; OP8	+	+
Gemmatimonadetes	+	+	Bacteria; OP9	+	-
Lentisphaerae	+	+	Bacteria; Other	+	-
Nitrospirae	+	+	Bacteria; PAUC34f	+	+

Planctomycetes	+	+	Bacteria; SAR406	+	+
Poribacteria	+	+	Bacteria; SBR1093	+	+
Proteobacteria	+	+	Bacteria; SR1	+	-
Spirochaetes	+	+	Bacteria; TM6	+	+
Synergistetes	+	-	Bacteria; TM7	+	-
Tenericutes	+	+	Bacteria; TPD-58	+	-
Thermi	+	+	Bacteria; unassigned	+	-
Thermotogae	+	-	Bacteria; WPS-2	+	+
Verrucomicrobia	+	+	Bacteria; WS1	+	-
			Bacteria; WS2	+	+
			Bacteria; WS3	+	+
			Bacteria; WS4	+	-
			Bacteria; WS5	+	+
			Bacteria; WWE1	+	-
			Bacteria; ZB3	+	+
			Unassigned; Other	+	+
	32	25		39	16

Affiliated (known) OTUs (class in phylum Proteobacteria)	**33 species (V13,V45,V58)**	**81 species V4**	**Unaffiliated OTUs (class in phylum Proteobacteria)**	**33 species (V13,V45,V58)**	**81 species V4**
Alphaproteobacteria	+	+	Proteobacteria; Other	+	-
Betaproteobacteria	+	+	Proteobacteria; TA18	+	+
Deltaproteobacteria	+	+	Proteobacteria; unassigned	+	+
Epsilonproteobacteria	+	+			
Gammaproteobacteria	+	+			
Zetaproteobacteria	+	-			
	6	5		3	2

+/-: presence/ absence of the OTUs.

5.3.2 Sponge microbial community structure revealed by a combination of multiple primer sets

The microbial diversity and relative abundance of each phylum-level OTU were revealed by the combination of three primer sets (Fig. 5-1). Sixty-two phylum-level microbial OTUs were revealed in total for all the 19 sponge species studied, among which 31 were the unaffiliated OTUs (one unassigned and 30 candidate phyla). The microbial communities revealed from different host species within the same sponge order demonstrated high similarity but did show significant differences with other orders. The shared dominant phylum-level OTUs of the microbial community were distinct for each sponge order and easy to distinguish.

Similarly, for the five sponge families in the order Dictyoceratida, within the same sponge family, the microbial communities presented consistently dominant OTUs for various host species (Fig. 5-2). The differences between different sponge families were obvious. Sixty-three phylum-level microbial OTUs were uncovered for the studied 14 sponge species, and 33 of them were unaffiliated.

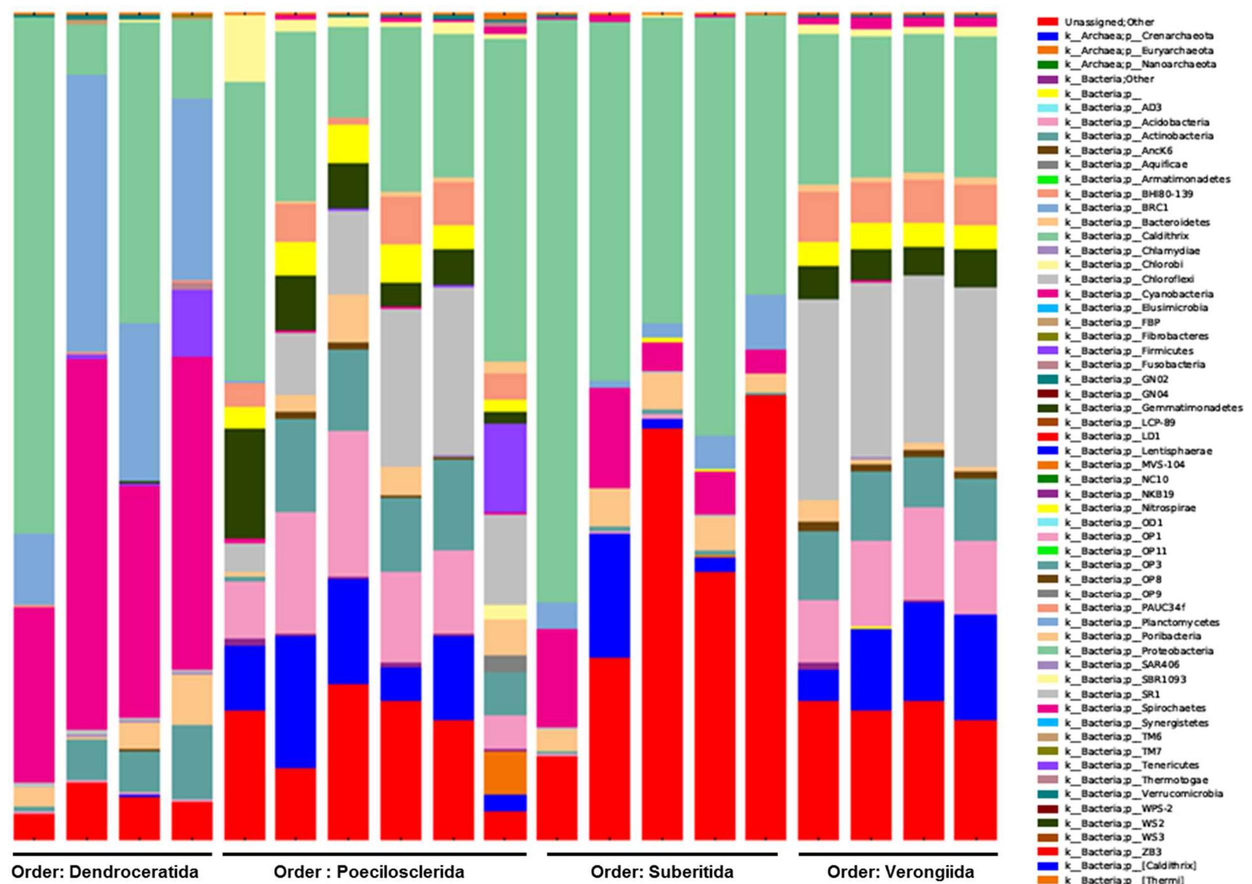


Figure 5 - 1 Microbial community composition and structure of 19 sponge species belonging to four taxonomic orders based on the phylum-level microbial OTUs

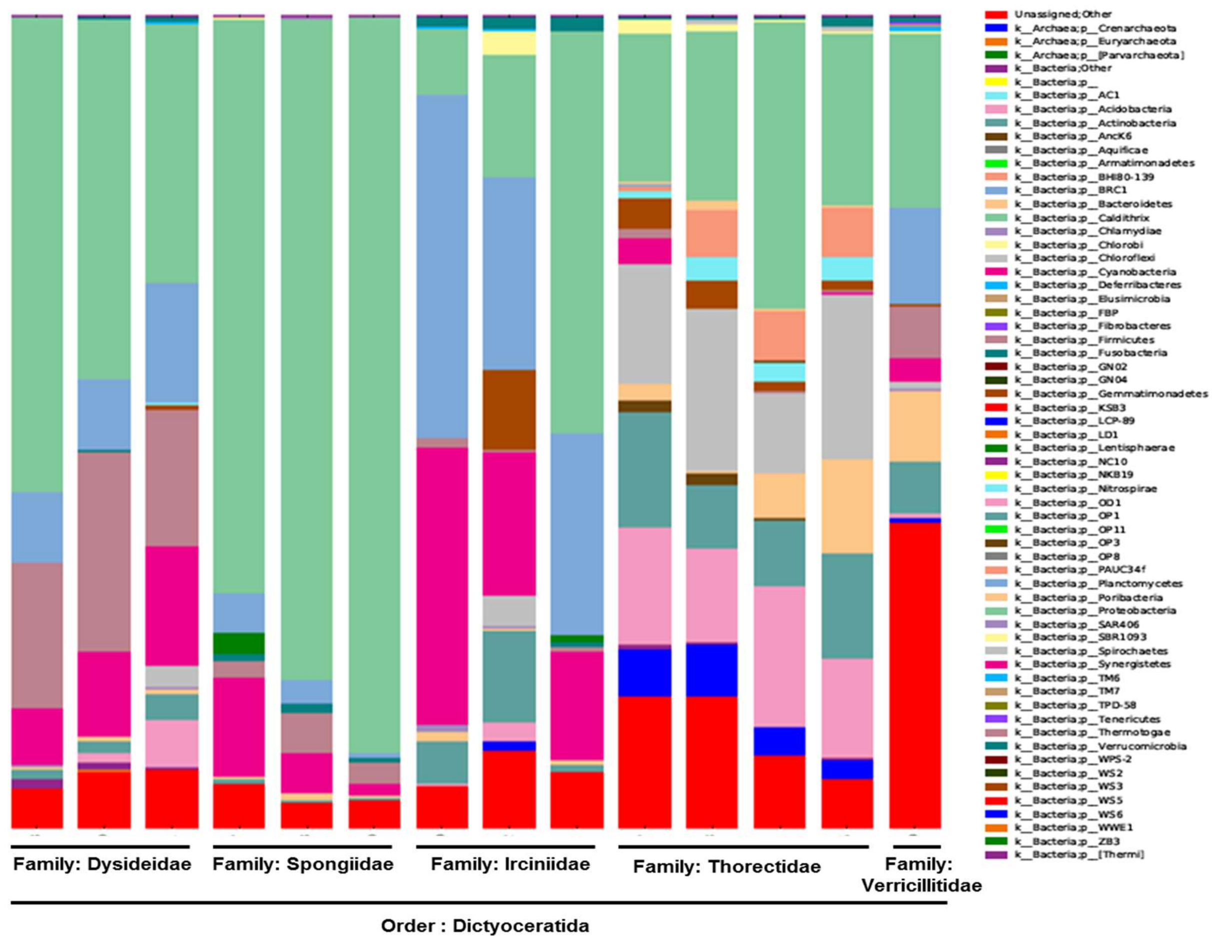


Figure 5 - 2 Microbial community composition and structure of 14 sponge species belonging to five taxonomic families within one order based on the phylum-level microbial OTUs

5.3.3 OTU richness of the sponge microbiome at the order level

The rarefied tables are the basis for calculating alpha diversity metrics, which describe the richness and/or evenness of taxa in a single sample. The rarefaction curves of the 16S rRNA gene indicate the diversity of microbial community within a given sample. In the Figure 5-3, the four different sponge orders were analysed by various metrics (PD whole tree, chao1, observed species, and Shannon). The metrics supported by different principles compare the microbial communities belonging to different sponge orders on different aspects. PD whole tree is a divergence-based qualitative metric and considers each microbial OTUs in the sample are phylogenetically distinct. Chao1 and observed species are all species-based qualitative metrics. Shannon is a species-based quantitative metric to take the relative abundance of each district microbial OTU into account when computing the rarefaction curves. The four types of the

metrics all concluded the consensus that the sponge order Verongida had higher diversity of the associated microbial community than Poecilosclerida, Suberitida, and Dendroceratida. The flat curve in the rarefaction using Shannon metric could be found in all the four sponge orders, which implied the sequencing depth was valid to reveal the entire microbial community.

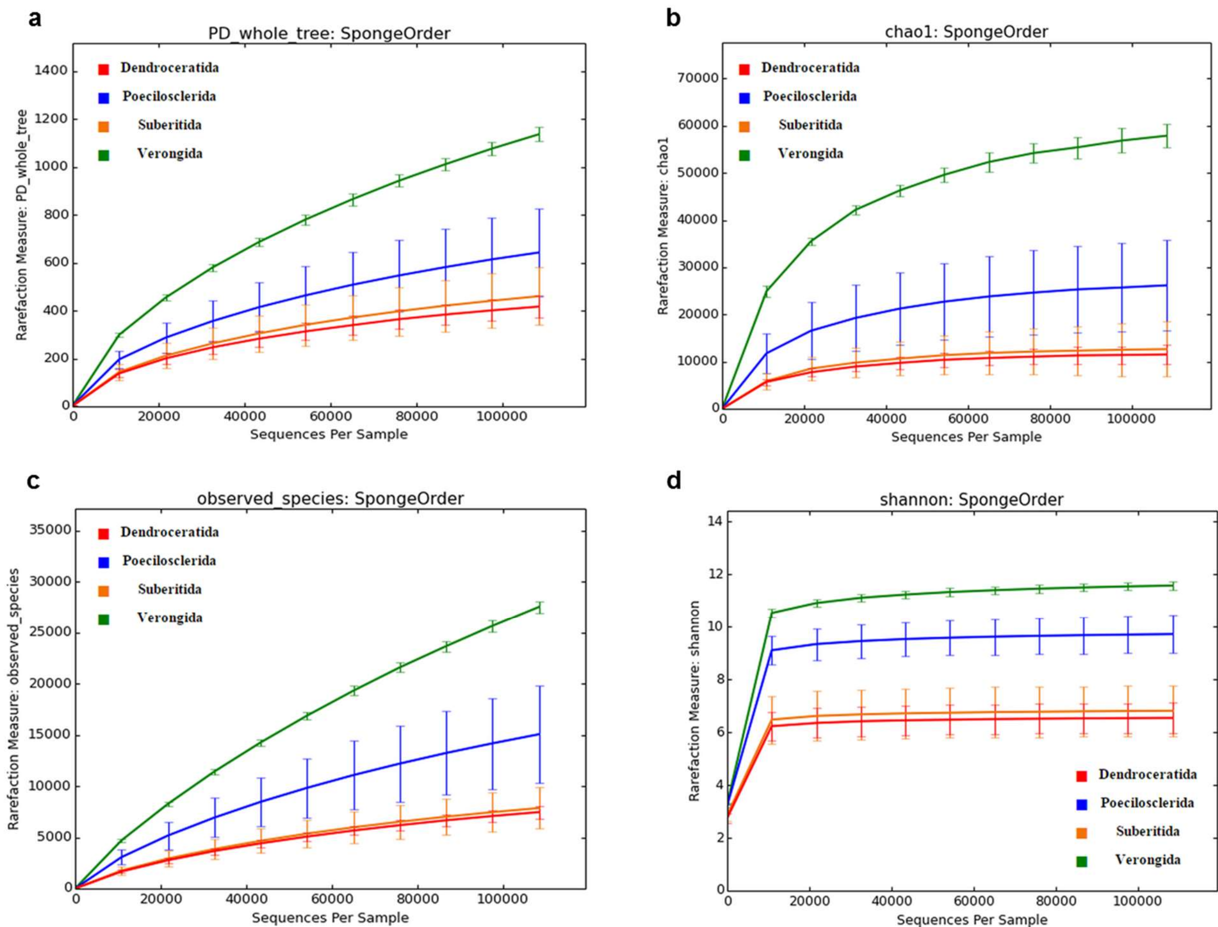


Figure 5 - 3 Rarefaction curves of 16S rRNA gene diversity showing the microbial communities of the sponge species in four taxonomic orders. Rarefaction using a. PD whole tree metric; b. Chao1 metric; c. Observed species metric; d. Shannon metric.

5.3.4 OTU richness of the sponge microbiome at the family level

The rarefaction curves of the 16S rRNA gene of five different sponge families within the same order are shown in Figure 5-4. The three types of the metric (PD whole tree, chao1, and observed species) provided one conclusion that the sponge family Irciniidae had the highest community diversity followed by the family Dysideidae, then Verricillitidae, Thorectidae, and

Spongiidae. However, based on the Shannon metric, we obtained a different rank by swapping family Dysideidae and Verricillitidae, which was impacted probably by the uneven contribution of the sponge species in the two families. In family Dysideidae, three species in different genera were analysed. In family Verricillitidae, only one specie was included in the comparison. The variation of the structure of the microbial communities of different species in the family Dysideidae compensates the significant value of the variation compared to the family Verricillitidae with only one genus. Again, the flat curve in the rarefaction using Shannon metric were found in all the five sponge families reached the proper sequencing depth for a valid analysis of the microbial community in this study.

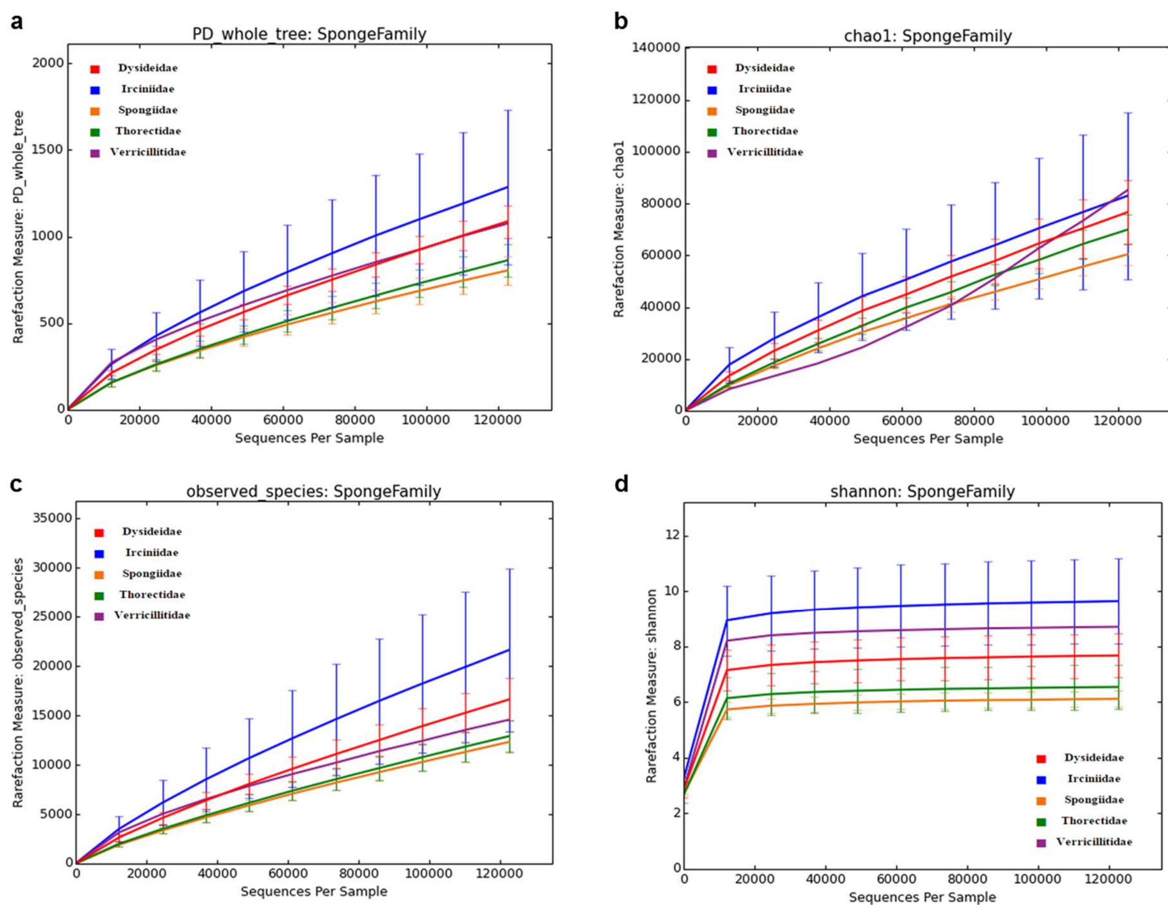


Figure 5 - 4 Rarefaction curves of 16S rRNA gene diversity showing the microbial communities of the sponge species in five taxonomic families within one order. Rarefaction using a. PD whole tree metric; b. Chao1 metric; c.Observed species metric; d. Shannon metric.

5.3.5 Comparison of microbial community at the sponge order level

Beta diversity represents the explicit comparison of microbial (or other) communities based on their composition. Beta diversity metrics thus assess the differences between microbial communities. The fundamental output of these comparisons is a square, hollow matrix where a “distance” or dissimilarity is calculated between every pair of community samples, reflecting the dissimilarity between those samples. The data in this distance matrix can be visualised with analyses such as Principal Coordinates Analysis (PCoA). To directly measure the robustness of individual Unweighted Pair Group Method with Arithmetic mean (UPGMA) clusters and clusters in PCoA plots, jackknifing approach, such jackknifed beta diversity (repeatedly resampling a subset of the available data from each sample) can be applied.

The beta diversity and jackknifed beta diversity analyses of the microbial community of 19 sponge species belonging to four taxonomic orders are shown in Figure 5-5. The Principal Coordinates Analysis (PCoA) plots demonstrate the similarity of microbial communities of the sponge species using both unweighted and weighted UniFrac metrics. UniFrac is a distance metric used for comparing biological communities. It differs from dissimilarity measures such as Bray-Curtis dissimilarity in that it incorporates information on the relative relatedness of community members by incorporating phylogenetic distances between observed organisms in the computation. Both weighted (quantitative) and unweighted (qualitative) variants of UniFrac are widely used in microbial ecology, where the former accounts for abundance of observed organisms, while the latter only considers their presence or absence. On that basis, it was found that using the weighted UniFrac metric performed better cluster of the sponge species within the same order than the unweighted UniFrac metric, regardless of the beta diversity or jackknifed beta diversity. It implies that the microbial community has the similar structure than the composition of the microbes within a sponge order.

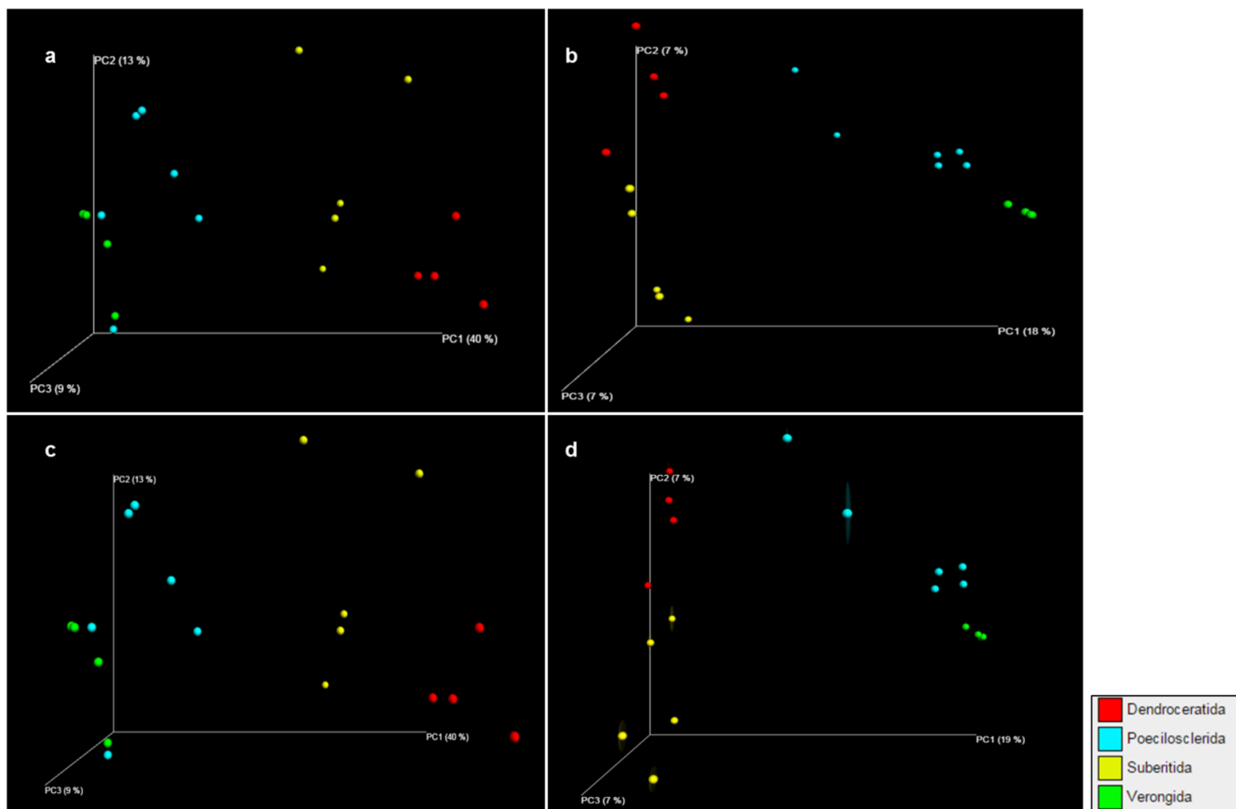


Figure 5 - 5 Principal Coordinates Analysis (PCoA) plots showing the similarity of microbial communities of the 19 sponge species in four taxonomic orders. a. Beta diversity using unweighted UniFrac metric; b. Beta diversity using weighted UniFrac metric; c. Jackknifed beta diversity using unweighted UniFrac metric; d. Jackknifed beta diversity using weighted UniFrac metric.

5.3.6 Comparison of microbial community at the sponge family level

The beta diversity and jackknifed beta diversity analyses of the microbial community of 14 sponge species belonging to five taxonomic families in the same order are shown in Figure 5-6. The Principal Coordinates Analysis (PCoA) plots demonstrate the similarity of microbial communities of the sponge species using both unweighted and weighted UniFrac metrics. Similarly, based on the weighted UniFrac analysis, the communities associated with the sponge species within the same family clustered together and showed the significant differences between each other. Only one exception was the microbial community of sponge species *Vaceletia* sp. (family Thorectidae), which clustered with the communities for sponges in the family Irciniidae. At sponge family level, the microbial community was of structure-specificity rather than diversity-specificity.

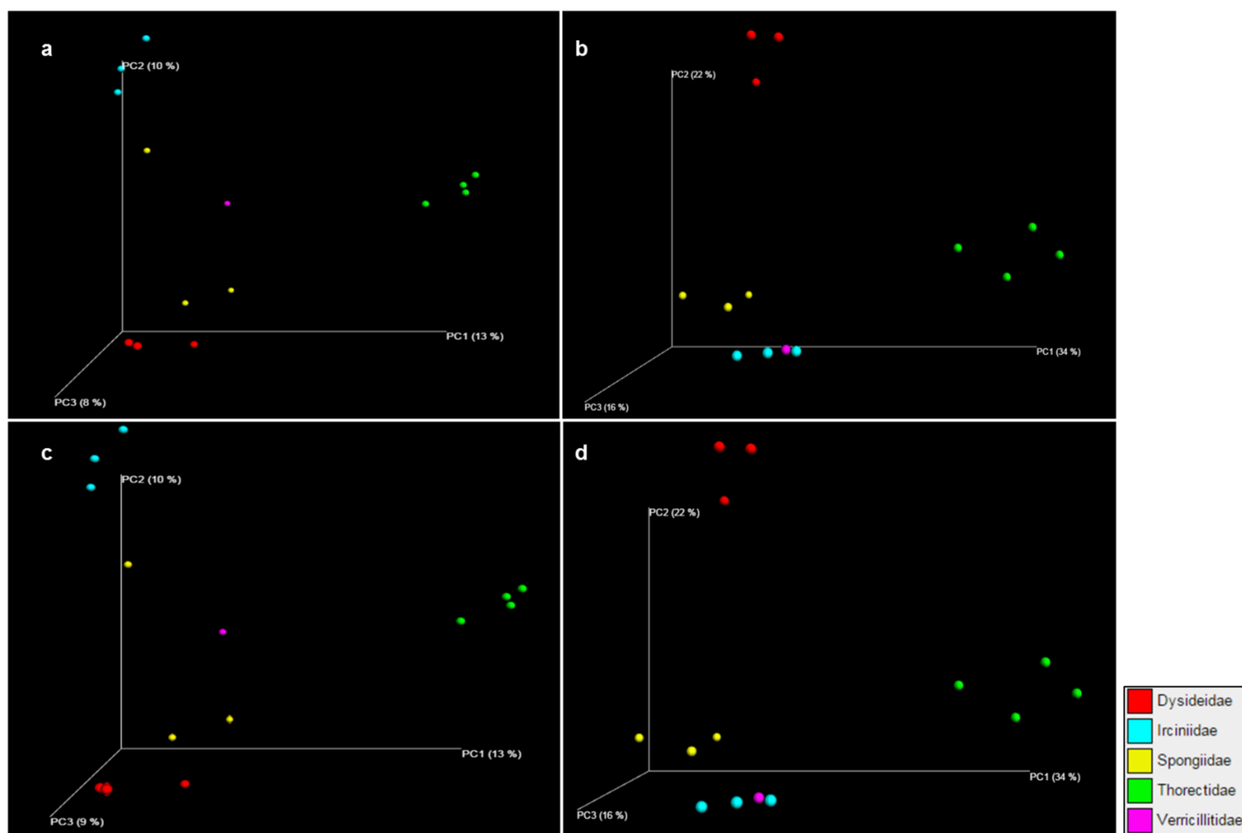


Figure 5 - 6 Principal Coordinates Analysis (PCoA) plots showing the similarity of microbial communities of the 14 sponge species in five taxonomic families within one order. a. Beta diversity using unweighted UniFrac metric; b. Beta diversity using weighted UniFrac metric; c. Jackknifed beta diversity using unweighted UniFrac metric; d. Jackknifed beta diversity using weighted UniFrac metric.

5.3.7 Dominant microbial OTUs

The microbial taxa belonging to the studied sponge samples can be presented together with the sponge samples in the PCoA plot. The coordinates of a given taxon are plotted as a weighted average of the coordinates of all samples, where the weights are the relative abundances of the taxon in the samples. The size of the sphere representing a taxon is proportional to the mean relative abundance of the taxon across all samples. The five most abundant class-level taxa among the 19 sponge species in four orders are presented in a biplot display (Figure 5-7). The class-level OTU of 'Unassigned' contributed the highest abundance among the microbial communities belonging to four sponge orders, followed by the four class-level OTUs of Alpha-proteobacteria, Gamma-proteobacteria, Synechococcophycideae, and Thaumarchaeota.

For the microbial community comparison of the five families in the same order, the PCoA plot using weighted top five class-level OTUs are shown in Figure 5-8. The most abundant was Gamma-proteobacteria, followed by the four class-level OTUs of Alpha-proteobacteria, Unassigned, planctomycetia, and Synechococcophycideae.

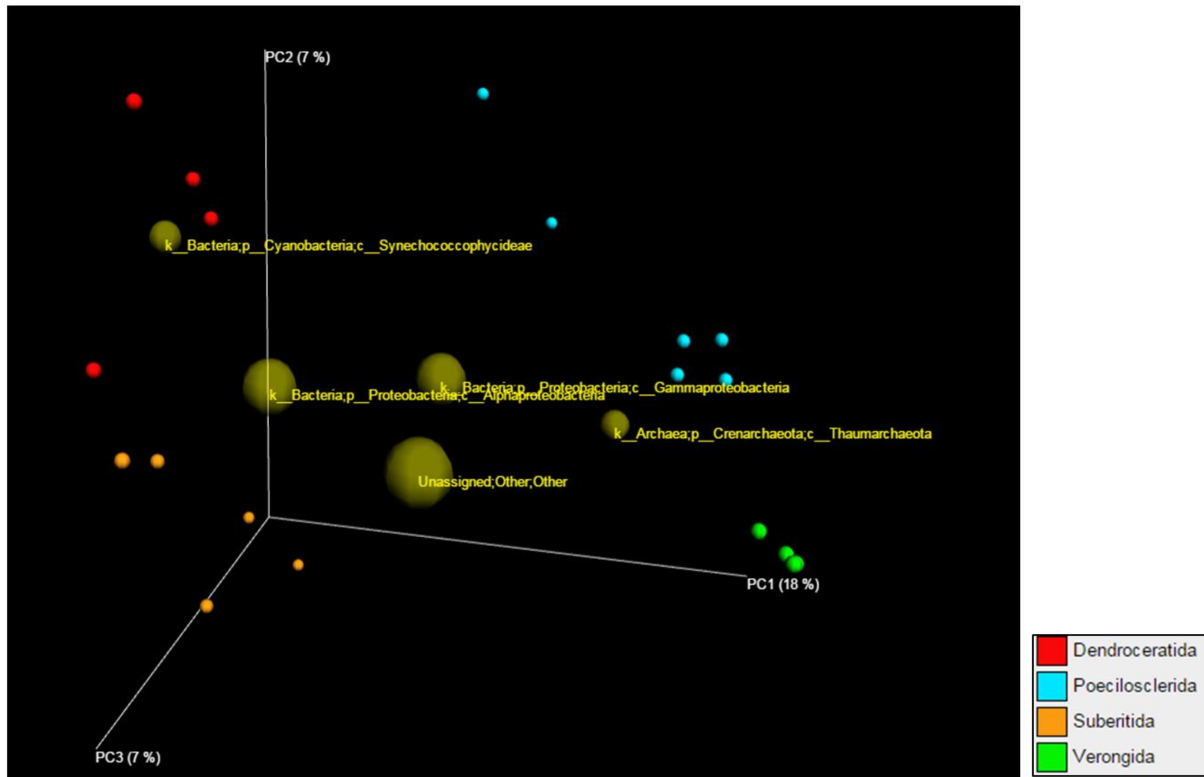


Figure 5 - 7 PCoA plot showing the five most abundant class-level microbial OTUs among the 19 sponge species in four taxonomic orders. The coordinates of a given microbial taxon are plotted as a weighted average of the coordinates of all sponge samples, where the weights are the relative abundances of the taxon in the samples. The size of the sphere representing a taxon is proportional to the mean relative abundance of the taxon across all samples.

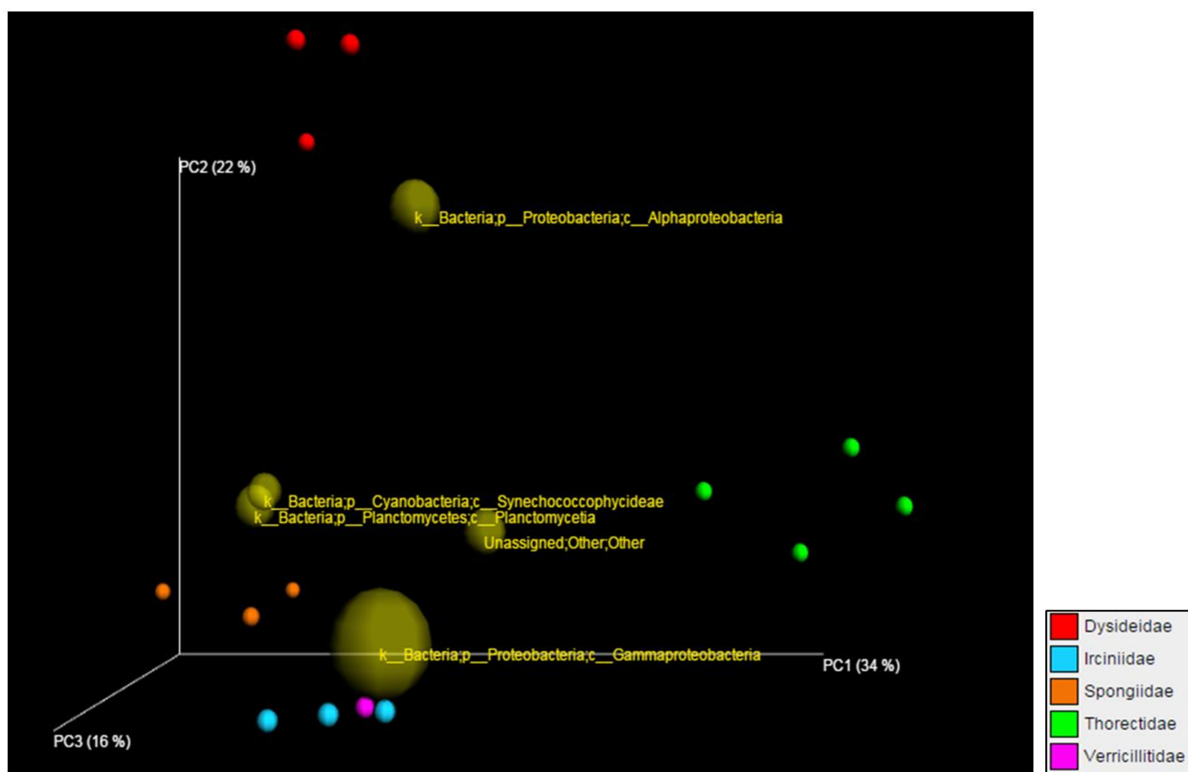


Figure 5 - 8 PCoA plot showing the five most abundant class-level microbial OTUs among the 14 sponge species in five taxonomic families within one order. The coordinates of a given microbial taxon are plotted as a weighted average of the coordinates of all samples, where the weights are the relative abundances of the taxon in the sponge samples. The size of the sphere representing a taxon is proportional to the mean relative abundance of the taxon across all samples.

5.3.8 Specific microbial OTUs

The dominant microbial OTUs of different taxonomic levels for four sponge orders and five families in the fifth order are documented in Tables 5-3 and 5-4. The dominant OTUs are defined as the OTUs with more than 1% of the sequences among the whole dataset. The dominant microbial OTUs for a given sponge order are the OTUs that are dominant for every sponge species in the order and are shared between all the sponge species. Based on these definitions, the dominant microbial OTUs belonging to each of the four sponge orders were compared and analysed to find out the unique and shared OTUs. As shown in the Table 5-3, apart from one sponge order (Suberitida) which had no unique dominant OTUs, the three other sponge orders all had one or two phylum-level unique OTUs. Among the 19 sponge species belonging to four orders, there were two shared phylum-level dominant OTUs of Proteobacteria and Unassigned. For the five sponge families in the order Dictyoceratida (Table 5-4), two of

them (Dysideidae and Spongiidae) had no phylum-level unique dominant OTUs. However, family Spongiidae had the unique microbial OTUs at the class level and lower taxonomic levels, and family Dysideidae had the unique ones at the family and genus levels. In contrast, the family Irciniidae only presented unique dominant OTUs at phylum level. In terms of the other two families Thorectidae and Verricillitidae, they had unique microbial OTUs at all the five taxonomic levels. Overall, the 14 sponge species in the five families within the same order shared two dominant phylum-level OTUs (Proteobacteria and Unassigned).

For the qualitative analysis, the common microbial OTUs among all the sponge species in the same order were used as the representatives to find the shared OTUs between the four orders (Table 5-5). The unique microbial OTUs for each sponge order are defined as the OTUs that only exist in one sponge order (any species) but do not show up in the three other sponge orders (Table 5-5). Order Dendroceratida had the largest number of the unique phylum-level microbial OTUs (10), and order Poecilosclerida had the second largest number of nine OTUs. The orders Verongiida and Suberitida had one and two unique phylum-level microbial OTUs, respectively. All of the 19 sponge species shared 16 phylum-level OTUs. For the five sponge families in the order Dictyoceratida (Table 5-6), family Thorectidae, Dysideidae, and Verricillitidae had four, two, and one unique phylum-level microbial OTUs, respectively. The other two families Spongiidae and Irciniidae had no unique OTUs at the phylum level. The 14 sponge species shared 16 phylum-level OTUs.

Table 5 - 3 Comparison for shared and unique dominant OTUs (>1%) between different sponge orders

	Dendroceratida (4 species in 3 families)	Poecilosclerida (6 species in 6 families)	Suberitida (5 species in 3 families)	Verongiida (4 species in 2 families)	Shared (19 species)
Phylum-level OTUs	4 [a] (1) [b] Planctomycetes	7 (0)	4 (1) Bacteroidetes	9 (2) Actinobacteria PAUC34f	2 [c] Proteobacteria Unassigned
Class-level OTUs	6 (3) Chloroplast Synechococcophycideae Planctomycetia	7 (0)	4(1) Flavobacteriia	14 (7) Solibacteres Acidimicrobiia Anaerolineae SAR202 TK17 PAUC34f;c__ Deltaproteobacteria	3 Alphaproteobacteria Gammaproteobacteria Unassigned
Order-level OTUs	5 (4) Chloroplast;Other Synechococcales Pirellulales Rhizobiales	8 (0)	2 (1) Flavobacteriales	17 (9) iii1-15 Solibacterales Acidimicrobiales Caldilineales SAR202;o__ PAUC34f;c__;o__ Rhodospirillales Entotheonellales Syntrophobacterales	1 Unassigned

Family-level OTUs	4 (3) Chloroplast;Other;Other Synechococcaceae Pirellulaceae	7 (0)	2 (1) Flavobacteriaceae	18 (9) iii1-15;f__ Acidimicrobiales;f__ wb1_P06 Caldilineaceae SAR202;o__;f__ PAUC34f;c__;o__;f__ Rhodospirillaceae Syntrophobacteraceae Ectothiorhodospiraceae	1 Unassigned
Genus-level OTUs	4 (3) Chloroplast;Other;Other;Other *Synechococcus* Pirellulaceae;g__	5 (0)	1 (0)	18 (12) iii1-15;f__;g__ PAUC26f;g__ Acidimicrobiales;f__;g__ wb1_P06;g__ Caldilineaceae;g__ SAR202;o__;f__;g__ PAUC34f;c__;o__;f__;g__ Rhodospirillaceae;g__ Syntrophobacteraceae;g__ Chromatiales;f__;g__ Ectothiorhodospiraceae;g__ HTCC2089;g__	1 Unassigned

[a] The total number of phylum/ class/ order/ family/ genus level OTUs existing in all the species in the order with > 1% sequence abundance; [b] The unique phylum/ class/ order/ family/ genus level OTUs only existing in all of the species in one order with > 1% sequence abundance; [c] The shared phylum/ class/ order/ family/ genus level OTUs among all of the species in four orders with > 1% sequence abundance.

Table 5 - 4 Comparison for shared and unique dominant OTUs (>1%) between different sponge families in one order

	Dysideidae (3 species in 3 genera)	Spongiidae (3 species in 3 genera)	Irciniidae (3 species in 3 genera)	Thorectidae (4 species in 4 genera)	Verricillitidae (1 species in 1 genus)	Shared (14 species)
Phylum-level OTUs	6 [a] (0) [b]	4 (0)	6 (1) Verrucomicrobia	7 (4) Crenarchaeota Acidobacteria Chloroflexi Gemmatimonadetes	7 (1) Bacteroidetes	2 [c] Proteobacteria Unassigned
Class-level OTUs	6 (0)	5 (1) Epsilonproteobacteria	4(0)	12 (7) Thaumarchaeota Acidobacteria-6 Solibacteres Anaerolineae SAR202 TK17 Gemm-2	13 (6) Actinobacteria Coriobacteriia Bacteroidia Chloroplast Bacilli Betaproteobacteria	2 Gammaproteobacteria Unassigned
Order-level OTUs	6 (0)	7 (2) Campylobacterales Oceanospirillales	4 (0)	9 (7) Cenarchaeales Solibacterales Caldilineales SAR202;o__ Gemm-2;o__ Rhodospirillales Chromatiales	15 (8) Coriobacteriales Bacteroidales Chloroplast;o__ Lactobacillales Planctomycetales Rhodobacterales EC94 Legionellales	1 Unassigned

Family-level OTUs	7 (3) Actinomycetales;f__ Clostridiaceae Phyllobacteriaceae	4 (1) Campylobacteraceae	3 (0)	11 (9) Cenarchaeaceae BPC015;f__ PAUC26f TK06 wb1_P06 Caldilineaceae SAR202;o__;f__ Gemm-2;o__;f__ f__Rhodospirillaceae	15 (11) Actinomycetaceae Coriobacteriaceae S24-7 Chloroplast;o__;f__ Lactobacillaceae Streptococcaceae Clostridiales;f__ Planctomycetaceae Rhodobacteraceae EC94;f__ Coxiellaceae	1 Unassigned
Genus-level OTUs	7 (3) Actinomycetales;f__;g__ *Clostridium* Phyllobacteriaceae;g__	7 (3) *Arcobacter* *Photobacterium* Vibrionaceae;Other	3 (0)	11 (9) *Nitrosopumilus* BPC015;f__;g__ PAUC26f;g__ TK06;g__ wb1_P06;g__ Caldilineaceae;g__ SAR202;o__;f__;g__ Gemm-2;o__;f__;g__ Rhodospirillaceae;g__	16 (11) *Actinomyces* *Atopobium* S24-7;g__ Chloroplast;o__;f__;g__ Lactobacillus Streptococcus Clostridiales;f__;g__ Planctomyces Rhodobacteraceae;g__ EC94;f__;g__ Coxiellaceae;g__	1 Unassigned

[a] The total number of phylum/ class/ order/ family/ genus level OTUs existing in all the species in the family with > 1% sequence abundance; [b] The unique phylum/ class/ order/ family/ genus level OTUs only existing in all of the species in one family with > 1% sequence abundance; [c] The shared phylum/ class/ order/ family/ genus level OTUs among all of the species in five families with > 1% sequence abundance.

Table 5 - 5 Comparison for shared and unique OTUs between different sponge orders

	Dendroceratida (4 species in 3 families)	Poecilosclerida (6 species in 6 families)	Suberitida (5 species in 3 families)	Verongiida (4 species in 2 families)	Shared (19 species)
Phylum-level OTUs	27 [a] /1 [b]	19/2	26/10	27/9	16 [c]
	Deferribacteres	Nanoarchaeota	Parvarchaeota	AD3	Crenarchaeota
		Aquificae	Elusimicrobia	FBP	Bacteria;Other
			Fibrobacteres	GN04	Acidobacteria
			Fusobacteria	LCP-89	Actinobacteria
			GN02	MVS-104	Bacteroidetes
			Kazan-3B-28	NC10	Chloroflexi
			KSB3	WS1	Cyanobacteria
			Lentisphaerae	WS3	Firmicutes
			OD1	WS4	Gemmatimonadetes
			WS5		Nitrospirae
					Planctomycetes
					Proteobacteria
					SBR1093
					Verrucomicrobia
					Unassigned
Class-level OTUs	60/15	37/9	52/38	63/32	30
	AT-s2-57	Archaeoglobi	Thermoplasmata	Methanomicrobia	Unassigned
	SJA-176	Halobacteria	Parvarchaea	AD3;c__	Thaumarchaeota
	C0119	Methanopyri	Chloracidobacteria	Acidobacteria;Other	Bacteria;Other;Other
	Ktedonobacteria	Thermococci	KIST-JJY010	Acidobacteria;c__	Acidobacteria-6
	4C0d-2	Nanoarchaeoti	Rubrobacteria	Acidobacteria-5	Solibacteres
	Oscillatoriophycideae	Aquificae	BRC1;c__	RB25	Sva0725
	Deferribacteres	Bacteroidetes;c__	NPL-UPA2	TM1	Acidimicrobiia
	VHS-B5-50	Chlorobia	SM1A07	Actinobacteria;Other	Actinobacteria
	Brachyspirae	OP11;c__	Sphingobacteriia	Actinobacteria;c__	Bacteroidia
	F38		OPB56	Chloroflexi;Other	Cytophagia
	TM7;c__		Elusimicrobia	Dehalococcoidetes	Flavobacteriia
	TM7-1		Endomicrobia	S085	Rhodothermi
	TM7-3		Fibrobacteria	SHA-26	Anaerolineae
	Tenericutes;c__		TG3	TK10	SAR202

	Spartobacteria		Fusobacteriia GN02;c__ BD1-5 KSB3;c__ Kazan-3B-28;c__ Lentisphaeria NKB19;Other SHAB590 TSBW08 Mb-NB09 SM2F11 OP3;c__ OP8_1 OP8_2 028H05-P-BN-P5 OM190 Brocadiae A712011 Leptospirae TM6;c__ SC3 Pedosphaerae SHA-109 WS5;c__	Thermobacula FBP;c__ GN04;c__ GN15 Gemmatimonadetes;Other Gemm-1 Gemm-3 LCP-89;c__ SAW1_B44 MVS-104;c__ NC10;c__12-24 SAW1_B6 OPB46 Proteobacteria;c__ WS1;c__ Kazan-3B-09 PRR-12 WS4;c__	TK17 Chloroplast Synechococcophycideae Bacilli Clostridia Gemm-2 Gemm-4 Nitrospira Planctomycetia Alphaproteobacteria Betaproteobacteria Deltaproteobacteria Gammaproteobacteria EC214 Spirochaetes Opitutae
Order-level OTUs	96/14	65/13	105/78	118/19	51
Family-level OTUs	165/32	87/24	176/123	168/29	65

[a] The total number of phylum/ class/ order/ family/ genus level OTUs existing in all the species in the order; [b] The unique phylum/ class/ order/ family/ genus level OTUs that exist in any species in one order but are absent in other orders; [c] The shared phylum/ class/ order/ family/ genus level OTUs among all of the species in four orders.

Table 5 - 6 Comparison for shared and unique OTUs between different sponge families in one order

	Dysideidae (3 species in 3 genera)	Spongiidae (3 species in 3 genera)	Irciniidae (3 species in 3 genera)	Thorectidae (4 species in 4 genera)	Verricillitidae (1 species in 1 genus)	Shared (14 species)
Phylum-level OTUs	33 [a]/2 [b] LD1 TPD-58	25/0	30/0	21/4 AC1 AD3 FBP WS6	43/1 WWE1	19 [c] Crenarchaeota Euryarchaeota Bacteria;Other Acidobacteria Actinobacteria AncK6 Bacteroidetes Chloroflexi Cyanobacteria Firmicutes Gemmatimonadetes Nitrospirae Planctomycetes Proteobacteria SBR1093 Spirochaetes TM7 Verrucomicrobia Unassigned
Class-level OTUs	69/6 MCG iii1-8 GN02;c__ LD1;c__ Brevinematae TPD-58;c__	63/6 Archaeoglobi BPC102 BHI80-139;c__ At12OctB3 OP8_2 ZB3;c__	72/10 OS-K SJA-176 SJA-28 Gloeobacterophycideae Elusimicrobia;c__ ODP123 TM6;Other TM6;c__ RF3 SC72	46/14 SHA-114 AD3;c__ EC1113 Actinobacteria;Other MB-A2-108 MD2896-B26 Chloroflexi;c__ TK10 FBP;c__ GN04;c__	113/10 Rubrobacteria Thermomicrobia TG3 OPB54 Mb-NB09 SM2F11 OP11-4 A712011 GN05 Cloacamonae	37 Unassigned;Other;Other Thaumarchaeota Bacteria;Other;Other Acidobacteria-6 Solibacteres Sva0725 Chloracidobacteria Acidimicrobiia Actinobacteria AncK6;c__

				Gemm-3		Bacteroidia
				NC10;c__		Cytophagia
				OP3;Other		Flavobacteriia
				SAW1_B6		Rhodothermi
						Saprospirae
						Anaerolineae
						SAR202
						TK17
						Chloroplast
						Oscillatoriophycideae
						Synechococcophycideae
						Bacilli
						Clostridia
						Gemm-2
						Gemm-4
						Nitrospira
						Planctomycetia
						Proteobacteria;Other
						Alphaproteobacteria
						Betaproteobacteria
						Deltaproteobacteria
						Epsilonproteobacteria
						Gammaproteobacteria
						EC214
						Spirochaetes
						Opitutae
						Verrucomicrobiae
Order-level OTUs	127/14	112/12	135/13	71/28	214/17	57
Family-level OTUs	215/26	193/25	229/36	95/37	365/25	71

[a] The total number of phylum/ class/ order/ family/ genus level OTUs existing in all the species in the family; [b] The unique phylum/ class/ order/ family/ genus level OTUs that exist in any species in one family but are absent in other families; [c] The shared phylum/ class/ order/ family/ genus level OTUs among all of the species in five families.

5.3.9 Performing procrustes analysis

The procrustes analysis evaluates whether the beta diversity is reproducible, regardless of which region specific primer set is used to compare the samples. The unweighted and weighted UniFrac PCoA matrices were applied to demonstrate the evaluation of the 19 sponge species belonging to the four taxonomic orders (Fig. 5-9, 5-10). The sponge species in the order Verongida showed highly consistent results using the three primer sets. In contrast, the sponge species in the orders Dendroceratida, Poecilosclerida, and Suberitida showed significant divergences using the three primer sets (Fig. 5-9). Considering the relative abundance of each microbial OTUs within the host sponges (weighted UniFrac metric), both order Verongida and Poecilosclerida showed consistency between the three primer sets used for profiling the sponge-associated microbial communities (Fig. 5-10). The other two orders (Dendroceratida and Suberitida) showed significant divergences when using different primer sets to reveal the same sponge microbial community.

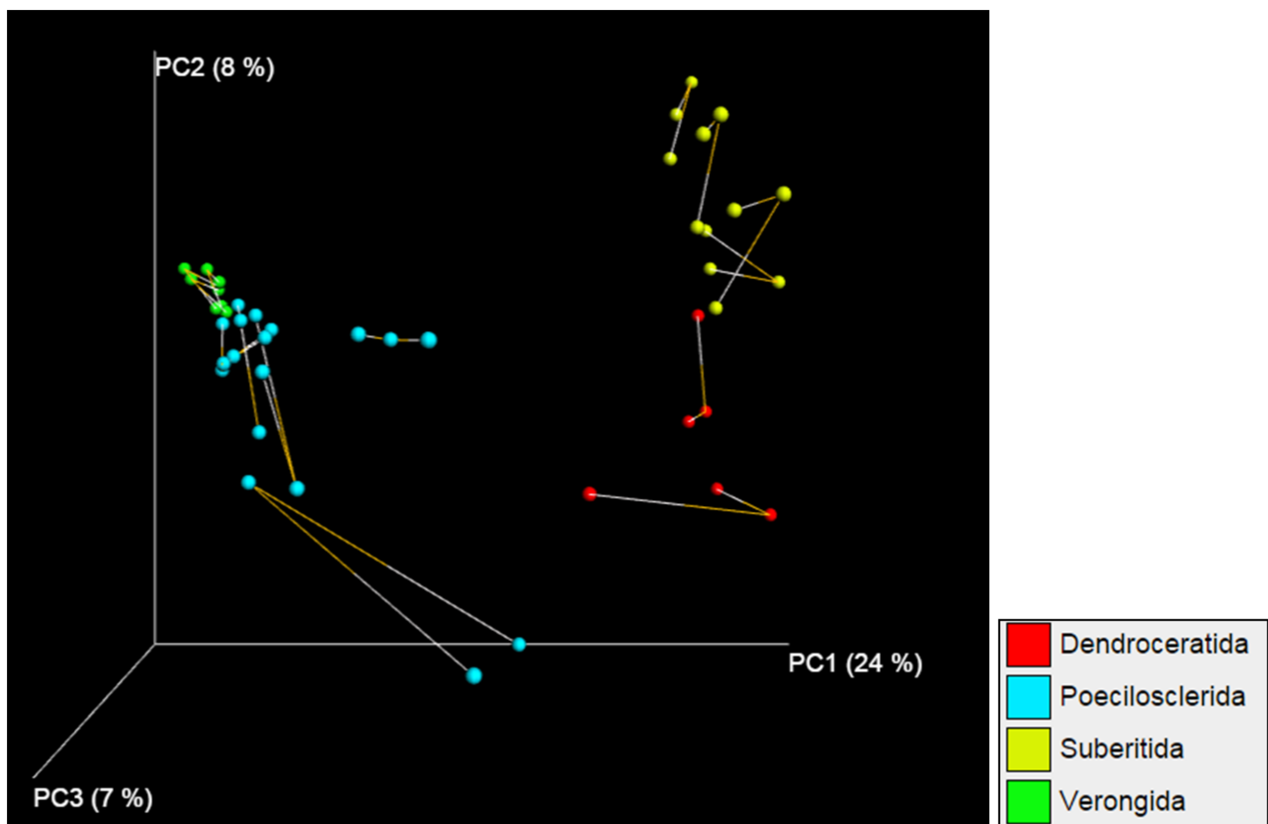


Figure 5 - 9 Procrustes analysis of the 19 sponge species belonging to four orders using the unweighted UniFrac metric generated by three 16S rRNA gene region-specific primer sets

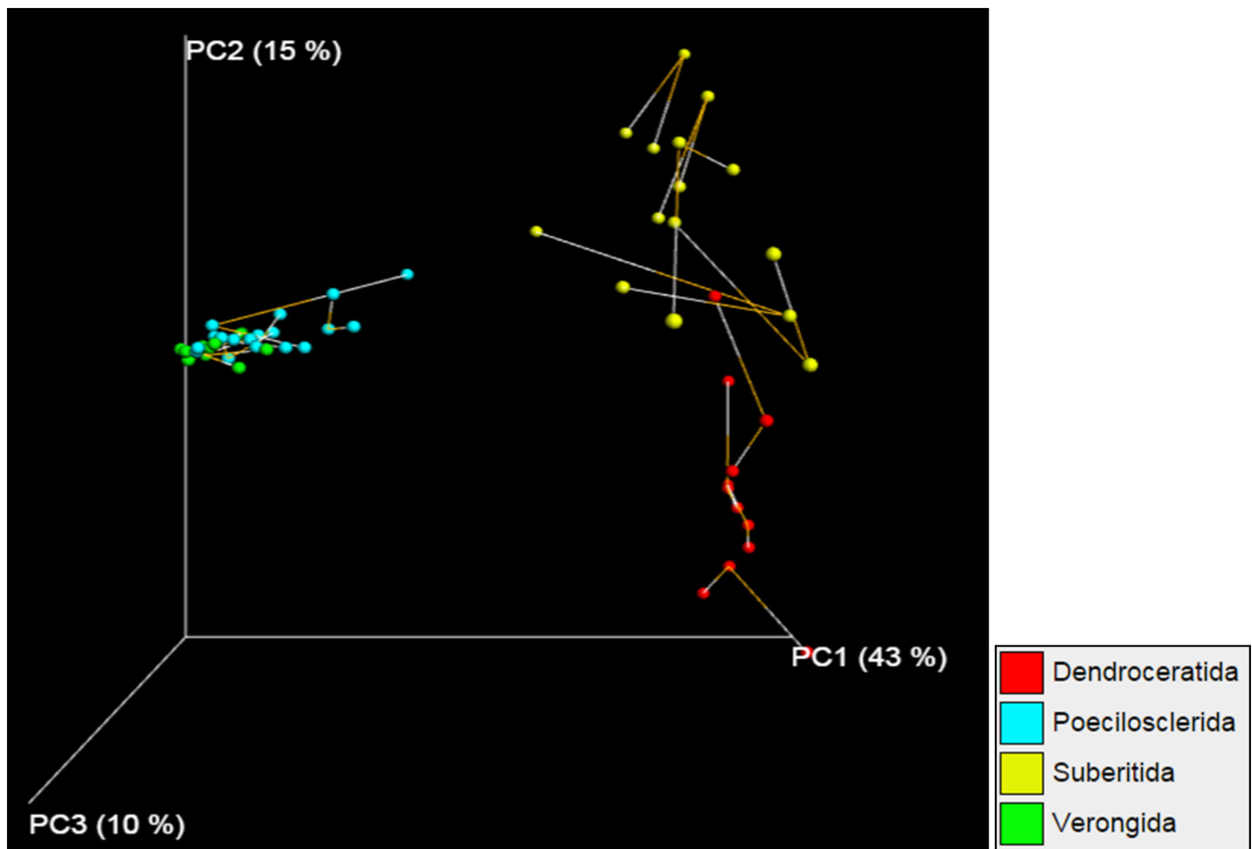


Figure 5 - 10 Procrustes analysis of the 19 sponge species belonging to four orders using the weighted UniFrac metric generated by three 16S rRNA gene region-specific primer sets

Similarly, the five sponge families within one order, including 14 sponge species, were evaluated by the procrustes analysis (Fig. 5-11, 5-12). Based on the comparison using unweighted UniFrac PCoA metric, it was found that the sponge species in the families Spongiidae and Thorectidae showed better consistency than the other three families Dysideidae, Irciniidae, and Verricillitidae (Fig. 5-11). For the comparison using weighted UniFrac PCoA metric, all the 14 sponge species showed high divergence when applying different primer sets to reveal the associated microbial community (Fig. 5-12).

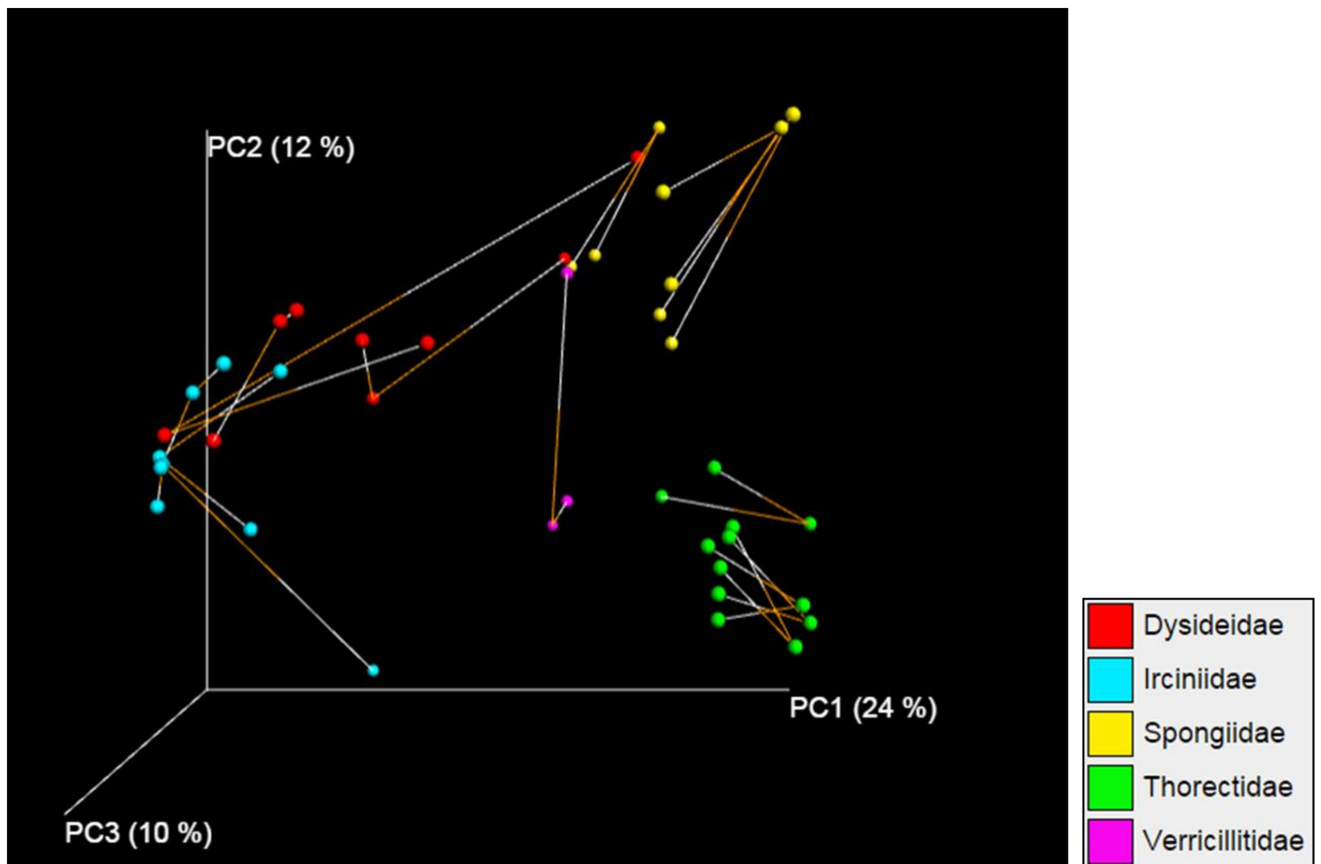


Figure 5 - 11 Procrustes analysis of the 14 sponge species belonging to five families in one order using the unweighted UniFrac metric generated by three 16S rRNA gene region-specific primer sets

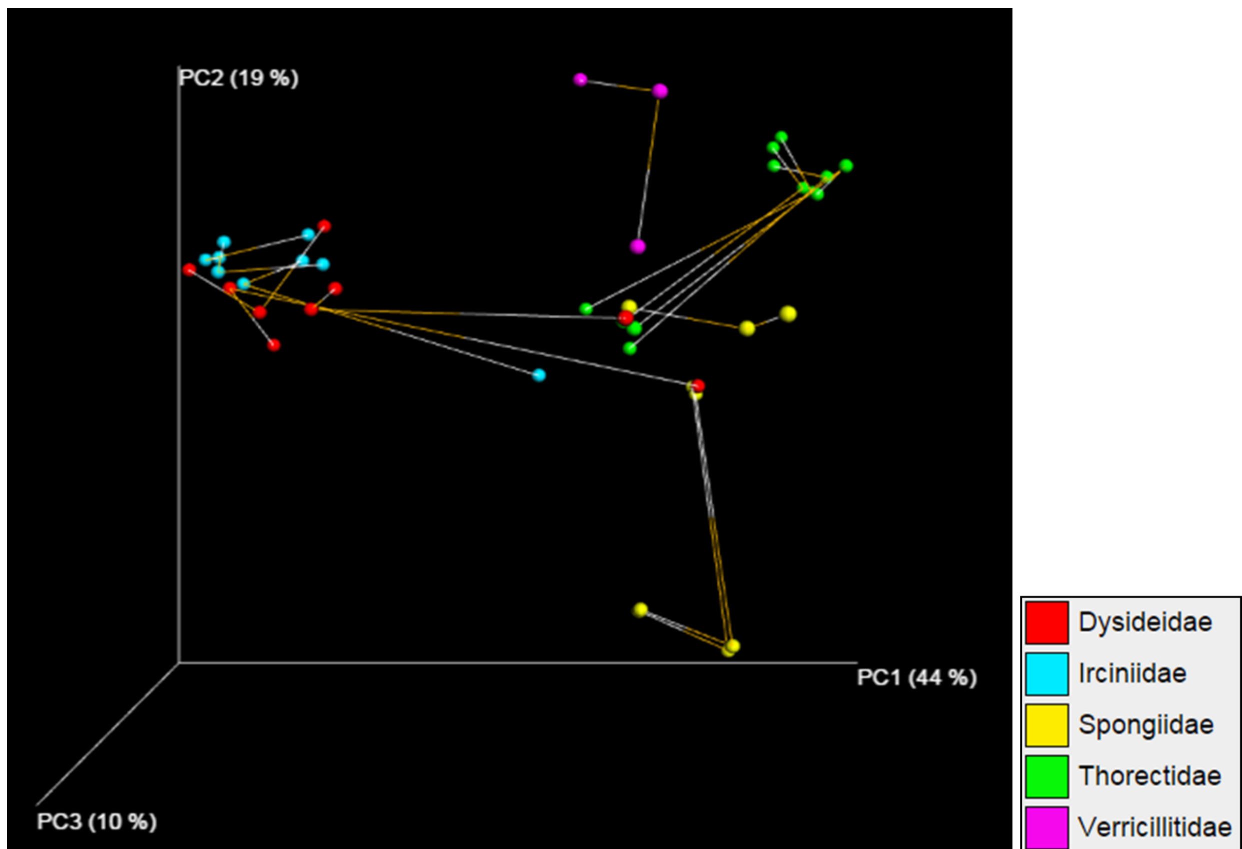


Figure 5 - 12 Procrustes analysis of the 14 sponge species belonging to five families in one order using the weighted UniFrac metric generated by three 16S rRNA gene region-specific primer sets

5.3.10 Phylogenetic relationship of studied sponges

The phylogenetic tree, based on the 28S rRNA gene of the sponges in the study, demonstrated that the evolutionary relationship between the 19 species belonging to four taxonomic orders (Fig. 5-13). The sponge species belonging to the orders Poecilosclerida and Verongida clustered together in the same clade (Fig. 5-13 purple colour marked). The orders Dendroceratida and Suberitida showed closer relationship and clustered in one clade (Fig. 5-13 red colour marked). Similarly, the phylogenetic relationship between another 14 sponge species belonging to five families in the fifth order Dictyoceratida was shown in the Figure 5-14. The species belonging to four families (Irciniidae, Spongiidae, Thorectidae, and Verricillitidae) clustered as one clade (Fig. 5-14 purple colour marked). The species in the family Dysideidae clustered separately as a single clade (Fig. 5-14 red colour marked).

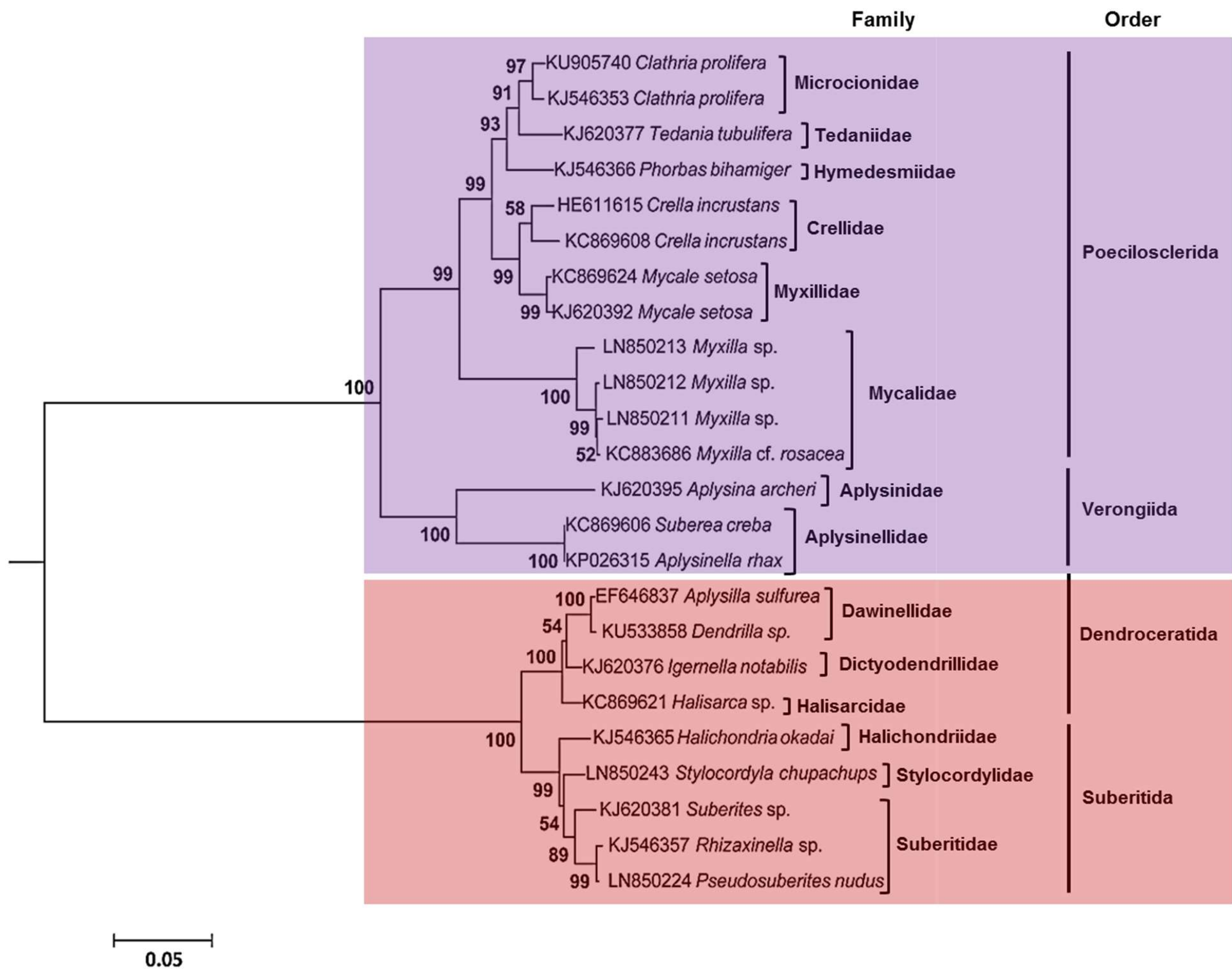


Figure 5 - 13 Phylogenetic relationship of the studied sponge species belonging to four taxonomic orders based on the 28S rRNA gene using Neighbor Joining method

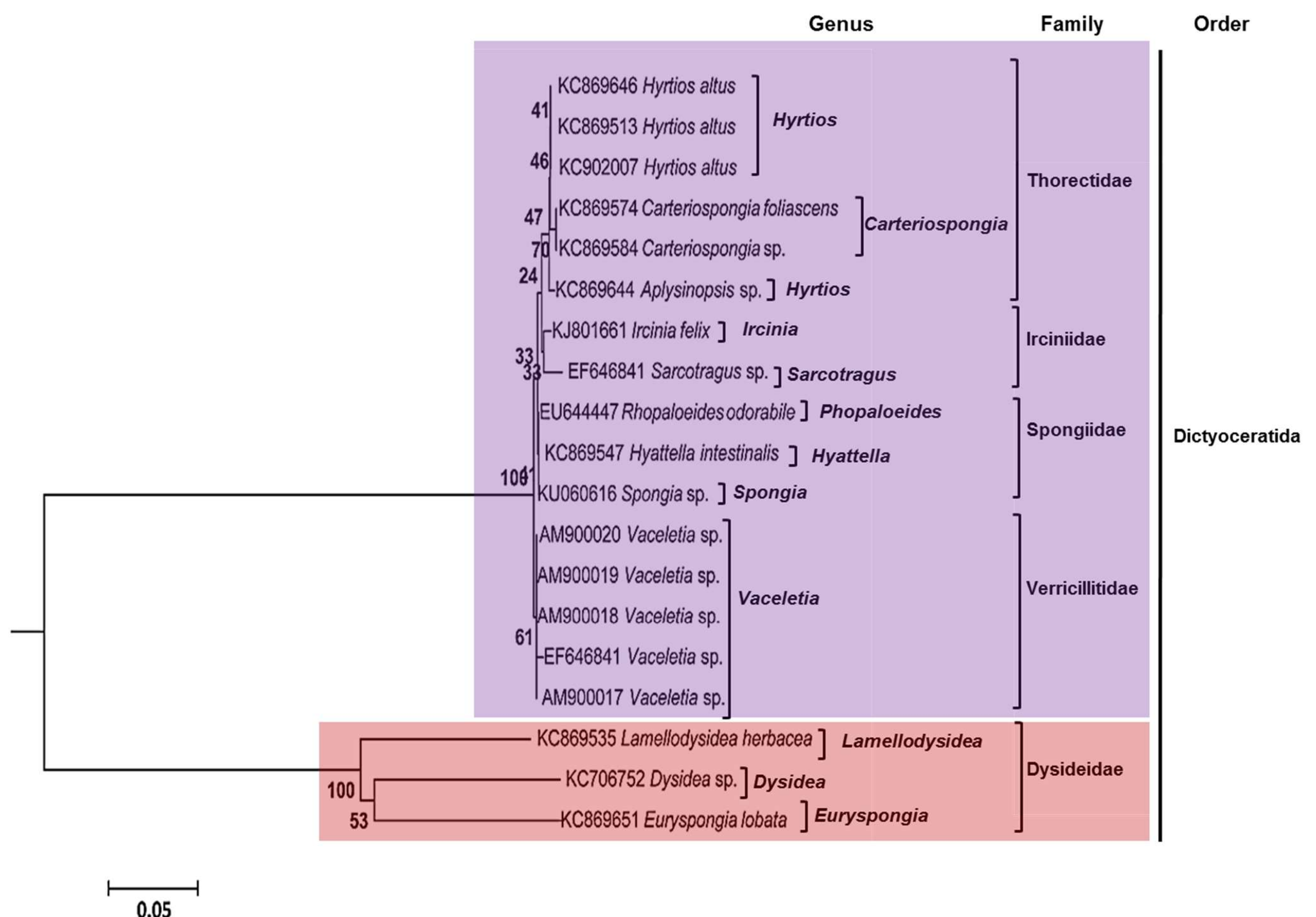


Figure 5 - 14 Phylogenetic relationship of the studied sponge species belonging to five taxonomic families in one order based on the 28S rRNA gene using Neighbor Joining method

5.4 Discussion

5.4.1 A combination of multiple region-specific primer sets required for comprehensive sponge microbiome analysis

The different primer sets targeting different regions of 16S rRNA gene could provide vastly different microbial community profiles when using the amplicon sequencing approach. This has caused the difficulty for comparison between different studies, as they utilised the single-primer based sequencing data and different primer sets only provided partial microbial profile of a given sample. This study demonstrated the significant divergence between the microbial profiles revealed by primers for regions V1V3, V4V5, and V5V8 of 16S rRNA gene for the same sponge sample (Fig. 5-9, 5-10, 5-11, and 5-12). Both unweighted and the weighted analyses illustrated that using different primer sets on the same sponge community DNA provided

different outputs on the microbial diversity and the relative abundance of each microbial taxa within the same sample.

Additionally, using the combined datasets generated from three specific primers for regions V1V3, V4V5, and V5V8, sufficient sequencing depth and consistent diversity trend were achieved when comparing the communities of the four sponge orders by alpha diversity analysis (Fig. 5-3). The order Verongiida (four species in two families) showed the highest microbial community diversity, followed by the order Poecilosclerida (six species in six families), then the order Suberitida (five species in three families) and Dendroceratida (four species in three families). Based on the rarefaction using the Shannon metric, the resulting flat trend confirmed the qualified sequencing depth for all the samples. In Figure 5-4, the sponge-associated microbial communities were compared at the sponge family level. The combined datasets indicated that the communities of the family Irciniidae (three species in three genera) showed higher diversity than the sponges in the other four families studied. Among them, the families Dysideidae (three species in three genera) and Verricillitidae (one species in one genus) had a higher diversity than the families Thorectidae (four species in four genera) and Spongiidae (three species in three genera). However, the families Dysideidae and Verricillitidae showed an inconsistent trend when using different metrics. For example, using PD whole tree (qualitative divergence-based measure), observed species (qualitative species-based measure), and chao1 (qualitative species-based measure), the family Dysideidae had higher microbial community diversity than the family Verricillitidae. In contrast, the associated microbial community of the sponge family Verricillitidae had a higher diversity when analysed by Shannon metric (quantitative species-based measure). But the sequencing depth could be confirmed by the output of the alpha diversity –Shannon metric.

Importantly, the strategy applying a combination of multiple primer sets performed significantly better than the single-primer based method (Table 5-2). In total, seventy-one microbial OTUs, including 32 affiliated (known) OTUs and 39 unaffiliated OTUs, were revealed from the 33 sponge species. In contrast, the study of Thomas et al. (2016) in Sponge Microbiome Project reported only 25 affiliated (known) OTUs and 16 unaffiliated OTUs revealed from 81 sponge species of 804 samples collected globally. Considering the same OTU picking method of both two studies, our proposed multiple primer sets based amplicon sequencing approach is extremely efficient compared to single primer set based approach, particularly for untapped microbial taxa discovery (39 unaffiliated OTUs vs 16 unaffiliated OTUs).

5.4.2 Structure-specificity rather than diversity-specificity

Based on the phylum-level taxa summary chart, the microbial communities of the sponge

species within the same taxonomic order (Fig. 5-1) showed similar and consistent structure, which was different with other sponge orders. Similarly, for the sponge species belonging to five different families within one order, it was also found a consistent microbial taxa structure within the same family and distinct structures between the families (Fig. 5-2).

If applying beta diversity analysis to compare the microbial communities, it was found that the microbial profiles of the sponge species in the same taxonomic order clustered better using a weighted metric than using an unweighted metric (Fig. 5-5). It suggested that the relative abundance of each taxon within a microbial community played an important role in influencing the grouping of the communities belonging to different sponge species. It also implied that the microbial communities derived from the sponge species within the same taxonomic order shared a similar taxa structure to allow them to cluster together and to be distinguished from the others; though various extents of the divergences still existed among the different families within the same order. In terms of the associated microbial community analysis at the sponge family level, Figure 5-6 demonstrated that the weighted metric again showed a better clustering for the five families within the order Dictyoceratida than the unweighted metric. Notably, the families of Irciniidae and Verricillitidae plotted in the same area, which implied the microbial communities in these two families could share a similar structure. However, as only one species was analysed to represent the family Verricillitidae, more data is required for a more reliable conclusion for this family.

Considering the phylogenetic relationship between the sponge species, we found the four orders were grouped into two clades: one clade included orders Poecilosclerida and Verongiida; the other clade included orders Suberitida and Dendroceratida (Fig. 5-13), which matched the weighted microbial community comparison in the Figure 5-5. Again, the weighted microbial community comparison between the sponge species in five families in one order (Fig. 5-6) correlated with the phylogenetic relationship in the Figure 5-14. The families Irciniidae, Verricillitidae, Spongiidae, and Thorectidae plotted much closer than the family Dysideidae. The phylogenetic tree also confirmed that those four families were in the same clade and the family Dysideidae was in a different clade.

5.4.3 Order-specificity and family-specificity of sponge microbial community

It is interesting to find that the microbial taxa derived from the sponge species within the same taxonomic order shared different numbers of taxa. When considering the dominant OTUs only (sequence abundance >1%), there were four, seven, four, and nine phylum-level OTUs commonly existing in the sponge orders Dendroceratida, Poecilosclerida, Suberitida, and Verongiida, respectively (Table 5-3). These microbial OTUs could be the representative

microbial taxa for the sponge order, which can be called sponge order-specific microbial OTUs. Comparing these specific OTUs representing the four sponge orders, two OTUs shared between the four orders, one unique OTU (Planctomycetes) for order Dendroceratida, one (Bacteroidetes) for order Suberitida, and two (Actinobacteria, PAUC34f) for order Verongiida. However, there was no unique phylum-level OTUs for the order Poecilosclerida. In terms of the class-level, order-level, family-level, and genus-level microbial OTUs, we could obtain various numbers of the sponge order-specific and the unique taxa. For the five families in the same order, their specific microbial OTUs could be also obtained (Table 5-4). There were six, four, six, seven, and seven common phylum-level microbial OTUs for the sponge families Dysideidae, Spongiidae, Irciniidae, Thorectidae, and Verricillitidae, respectively. These family-specific microbial OTUs could be the representative taxa for the particular sponge family. The understanding of these order-specific and family-specific microbial taxa could help establish a rationale guided isolation of unique microbial resources.

5.4.4 Unique microbial taxa as signature for a sponge order or family

If considering the entire microbial profile without the relative abundance of each OTU, the number of the common shared OTUs within a sponge order significantly increased (Table 5-5), though some of them were not the dominant ones (relative abundance >1%). Importantly, the unique microbial OTUs that only exist in one order (any sponge species) but are absent from any other sponge orders provided great information to find the signature of a certain sponge order. These microbial OTUs are not only specific to a certain sponge order but also are unique for the order. For the associated microbial community at the sponge family level, the unique microbial OTUs for each sponge family showed lower number than the unique microbial OTUs at sponge order level (Table 5-5, 5-6) as the different families within the same order have closer relationship than the different orders. These unique OTUs are the signature of the sponge order or family and offer efficient guidance to target on these microbial groups producing valuable secondary metabolites. Importantly, most of the signature microbial taxa are unaffiliated OTUs, which could be the untapped microbial resources.

5.5 Conclusion

In conclusion, the proposed amplicon sequencing approach, using a combination of multiple region-specific primer sets, had a powerful and unprecedented capacity to reveal the comprehensive microbial community of a given sponge sample on Illumina MiSeq platform. Seventy-one phylum-level microbial OTUs were revealed, including 32 affiliated (known) OTUs and 39 unaffiliated OTUs (unassigned and candidate OTUs), surpassing any of the existing studies. Based on the integrated sequencing data, sponge microbial community showed the

order-level and family-level specificity, which showed structure-specificity rather than the diversity-specificity. Different sponge orders and families have signature microbial OTUs, which are unique for the particular order/ family. Importantly, most unique microbial OTUs are unaffiliated taxa, which could be a promising source of untapped microorganisms.

References

Angermeier, H, Glöckner, V, Pawlik, JR, Lindquist, NL & Hentschel1, U 2012, 'Sponge white patch disease affecting the Caribbean sponge *Amphimedon compressa*', *Diseases of Aquatic Organisms*, vol. 99, pp. 95-102.

Beckers, B, Op De Beeck, M, Thijs, S, Truyens, S, Weyens, N, Boerjan, W & Vangronsveld, J 2016, 'Performance of 16s rdna primer pairs in the study of rhizosphere and endosphere bacterial microbiomes in metabarcoding studies', *Frontiers in Microbiology*, vol. 7, no. 605, pp. 1-15.

Benson, DA, Karsch-Mizrachi, I, Lipman, DJ, Ostell, J & Sayers, EW 2009, 'GenBank', *Nucleic Acids Research*, vol. 37, no. Database issue, pp. D26-31.

Burgsdorf, I, Erwin, PM, López-Legentil, S, Cerrano, C, Haber, M, Frenk, S & Steindler, L 2014, 'Biogeography rather than association with cyanobacteria structures symbiotic microbial communities in the marine sponge *Petrosia ficiformis*', *Frontiers in Microbiology*, vol. 10, no. 529, pp. 1-11.

Caporaso, JG, Kuczynski, J, Stombaugh, J, Bittinger, k, Bushman, FD, Costello, EK, Fierer, N, Peña, AG, J.K., G, Gordon, JI, Huttley, GA, Kelley, ST, Knights, D, Koenig, JE, Ley, RE, Lozupone, CA, McDonald, D, Muegge, BD, Pirrung, M, Reeder, J, Sevinsky, JR, Turnbaugh, PJ, Walters, WA, Widmann, J, Yatsunenko, T, Zaneveld, J & Knight, R 2010, 'QIIME allows analysis of high-throughput community sequencing data', *Nature Methods*, vol. 7, no. 5, pp. 335-6.

Caporaso, JG, Lauber, CL, Walters, WA, Berg-Lyons, D, Lozupone, CA, Turnbaugh, PJ, Fierer, N & Knight, R 2011, 'Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample', *Proceedings of the National Academy of Sciences*, vol. 108, pp. 4516-22.

Cebrian, E, Uriz, MJ, Garrabou, J & Ballesteros, E 2011, 'Sponge mass mortalities in a warming mediterranean sea: Are cyanobacteria-harboring species worse off?', *PLoS ONE*, vol. 6, no. 6, p. e20211.

Croué, J, West, NJ, Escande, M-L, Intertaglia, L, Lebaron, P & Suzuki, MT 2013, 'A single betaproteobacterium dominates the microbial community of the crambescidine-containing sponge Crambe crambe', *Scientific Reports*, vol. 3, p. 2583.

Cuvelier, ML, Blake, E, Mulheron, R, McCarthy, PJ, Blackwelder, P, Vega Thurber, RL & Lopez, JV 2014, 'Two distinct microbial communities revealed in the sponge Cinachyrella', *Frontiers in Microbiology*, vol. 5, no. 581, pp. 1-14.

De Voogd, NJ, Cleary, DFR, Polónia, ARM & Gomes, NCM 2015, 'Bacterial community composition and predicted functional ecology of sponges, sediment and seawater from the thousand islands reef complex, West Java, Indonesia', *FEMS Microbiology Ecology*, vol. 91, no. 4.

DeSantis, TZ, Hugenholtz, P, Larsen, N, Rojas, M, Brodie, EL, K. Keller, Huber, T, Dalevi, D, Hu, P & G.L., A 2006, 'Greengenes, a chimera-checked 16s rrna gene database and workbench compatible with ARB', *Applied and Environmental Microbiology*, vol. 72, no. 7, pp. 5069-72.

Easson, CG & Thacker, RW 2014, 'Phylogenetic signal in the community structure of host-specific microbiomes of tropical marine sponges', *Frontiers in Microbiology*, vol. 5, no. 532, pp. 1-11.

Edgar, RC 2010, 'Search and clustering orders of magnitude faster than BLAST', *Bioinformatics*, vol. 26, no. 19, pp. 2460-1.

Engelbrektson, A, Kunin, V, Wrighton, KC, Zvenigorodsky, N, Chen, F, Ochman, H & Hugenholtz, P 2010, 'Experimental factors affecting PCR-based estimates of microbial species richness and evenness', *ISME Journal*, vol. 4, no. 5, pp. 642-7.

Erwin, PM, Pita, L, López-Legentil, S & Turona, X 2012, 'Stability of Sponge-Associated Bacteria over Large Seasonal Shifts in Temperature and Irradiance', *Applied and Environmental Microbiology*, vol. 78, no. 20, pp. 7358-68.

Galtier, N, Gouy, M & Gautier, C 1996, 'SEA VIEW and PHYLO_ WIN: two graphic tools for sequence alignment and molecular phylogeny', *Computer Applications in the Bioscience*, vol. 12, no. 6, pp. 543-8.

Ghyselinck, J, Pfeiffer, S, Heylen, K, Sessitsch, A & De Vos, P 2013, 'The effect of primer choice and short read sequences on the outcome of 16s rRNA gene based diversity studies', *PLoS ONE*, vol. 8, no. 8, p. e71360.

Gilbert, JA, Jansson, JK & Knight, R 2014, 'The Earth Microbiome project: successes and aspirations', *BMC Biology*, vol. 12, no. 69, pp. 1-4.

Gower, JC 1975, 'Generalized procrustes analysis', *Psychometrika*, vol. 40, pp. 33-51.

Hall, TA 1999, 'BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT', *Nucleic Acids Symposium Series*, vol. 41, pp. 95-8.

Hentschel, U, Hopke, J, Horn, M, Friedrich, AB, Wagner, M, Hacker, J & Moore, BS 2002, 'Molecular evidence for a uniform microbial community in sponges from different oceans', *Applied and Environmental Microbiology*, vol. 68, no. 9, pp. 4431-40.

Hentschel, U, Piel, J, Degnan, SM & Taylor, MW 2012, 'Genomic insights into the marine sponge microbiome', *Nature Reviews Microbiology*, vol. 10, pp. 641-54.

Kim, M, Morrison, M & Yu, Z 2011, 'Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes', *Journal of microbiological methods*, vol. 84, pp. 81-7.

Lee, OO, Wang, Y, Yang, J, Lafi, FF, Al-Suwailem, A & Qian, PY 2011, 'Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea', *ISME Journal*, vol. 5, no. 4, pp. 650-64.

Luter, HM, Widder, S, Botté, ES, Wahab, MA, Whalan, S, Moitinho-Silva, L, Thomas, T & Webster, NS 2015, 'Biogeographic variation in the microbiome of the ecologically important

sponge, *Carteriospongia foliascens*', *PeerJ*, vol. 3, p. e1435.

Nelson, MC, Morrison, HG, Benjamino, J, Grim, SL & Graf, J 2014, 'Analysis, Optimization and Verification of Illumina-Generated 16S rRNA Gene Amplicon Surveys', *PLoS ONE*, vol. 9, no. 4, p. e94249.

Reveillaud, J, Maignien, L, Eren, AM, Huber, JA, Apprill, A, Sogin, ML & Vanreusel, A 2014, 'Host-specificity among abundant and rare taxa in the sponge microbiome', *ISME Journal*, vol. 8, pp. 1198-209.

Rideout, JR, He, Y, Navas-Molina, JA, Walters, WA, Ursell, LK, Gibbons, SM, Chase, J, McDonald, D, Gonzalez, A, Robbins-Pianka, A, Clemente, JC, Gilbert, JA, Huse, SM, Zhou, H-W, Knight, R & Caporaso, JG 2014, 'Subsampled open-reference clustering creates consistent, comprehensive OTU definitions and scales to billions of sequences', *PeerJ*, vol. 2, p. e545.

Rodríguez-Marconi, S, De la Iglesia, R, Díez, B, Fonseca, CA, Hajdu, E & Trefault, N 2015, 'Characterization of bacterial, archaeal and eukaryote symbionts from antarctic sponges reveals a high diversity at a three-domain level and a particular signature for this ecosystem', *PLoS ONE*, vol. 10, no. 9, p. e0138837.

Schloss, PD 2010, 'The Effects of Alignment Quality, Distance Calculation Method, Sequence Filtering, and Region on the Analysis of 16S rRNA Gene-Based Studies', *PLoS Computational Biology*, vol. 6, no. 7, p. e1000844.

Schmitt, S, Tsai, P, Bell, J, Fromont, J, Ilan, M, Lindquist, N, Perez, T, Rodrigo, A, Schupp, PJ, Vacelet, J, Webster, N, Hentschel, U & Taylor, MW 2012, 'Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges', *ISME Journal*, vol. 6, no. 3, pp. 564-76.

Selvin, J, Priya, SS, Kiran, GS, Thangavelu, T & SapnaBai, N 2009, 'Sponge-associated marine bacteria as indicators of heavy metal pollution', *Microbiological Research*, vol. 164, pp. 352-63.

Simister, RL, Schmitt, S & Taylor, MW 2011, 'Evaluating methods for the preservation and extraction of DNA and RNA for analysis of microbial communities in marine sponges', *Journal of Experimental Marine Biology and Ecology*, vol. 397, no. 1, pp. 38-43.

Taylor, MW, Radax, R, Steger, D & Wagner, M 2007, 'Sponge-associated microorganisms: evolution, ecology, and biotechnological potential', *Microbiology and Molecular Biology Reviews*, vol. 71, pp. 295-347.

Taylor, MW, Schupp, PJ, Dahllöf, I, Kjelleberg, S & Steinberg, PD 2004, 'Host specificity in marine sponge-associated bacteria, and potential implications for marine microbial diversity', *Environmental Microbiology*, vol. 6, no. 2, pp. 121-30.

Taylor, MW, Tsai, P, Simister, RL, Deines, P, Botte, E, Ericson, G, Schmitt, S & Webster, NS 2013, ''Sponge-specific' bacteria are widespread (but rare) in diverse marine environments', *ISME Journal*, vol. 7, pp. 438-43.

Thomas, T, Moitinho-Silva, L, Lurgi, M, Björk, JR, Easson, C, Astudillo-García, C, Olson, JB, Erwin, PM, López-Legentil, S, Luter, H, Chaves-Fonnegra, A, Costa, R, Schupp, PJ, Steindler, L, Erpenbeck, D, Gilbert, J, Knight, R, Ackermann, G, Lopez, JV, Taylor, MW, Thacker, RW, Montoya, JM, Hentschel, U & Webster, NS 2016, 'Diversity, structure and convergent evolution of the global sponge microbiome', *Nature Communications*, vol. 7, p. 11870.

Tremblay, J, Singh, K, Fern, A, Kirton, ES, He, S, Woyke, T, Lee, J, Chen, F, Dangl, JL & Tringe, SG 2015, 'Primer and platform effects on 16S rRNA tag sequencing', *Frontiers in Microbiology*, vol. 6.

Van Soest, RWM, Boury-Esnault, N, Hooper, JNA, Rützler, K, De Voogd, NJ, Alvarez de Glasby, B, Hajdu, E, Pisera, AB, Manconi, R, Schoenberg, C, Klautau, M, Picton, B, Kelly, M, Vacelet, J, Dohrmann, M, Díaz, M-C, Cárdenas, P & Carballo, JL 2016, *World Porifera database*, http://www.marinespecies.org/porifera, 2016-11-03.

Webster, NS, Webb, RI, Ridd, MJ, Hill, RT & Negri, AP 2001, 'The effects of copper on the microbial community of a coral reef sponge', *Environmental Microbiology*, vol. 3, no. 1, pp. 19-31.

Webster, NS, Xavier, JR, Freckelton, M, Motti, CA & Cobb, R 2008, 'Shifts in microbial and chemical patterns within the marine sponge *Aplysina aerophoba* during a disease outbreak', *Environmental Microbiology*, vol. 10, no. 12, pp. 3366–76.

Weigela, BL & Erwin, PM 2016, 'Intraspecific variation in microbial symbiont communities of the sun sponge, *Hymeniacidon heliophila*, from intertidal and subtidal habitats', *Applied and Enrivonmental Micorbiology*, vol. 82, pp. 650–8.

White, JR, Patel, J, Ottesen, A, Arce, G, Blackwelder, P & Lopez, JV 2012, 'Pyrosequencing of Bacterial Symbionts within *Axinella corrugate* Sponges: Diversity and Seasonal Variability', *PLoS ONE*, vol. 7, no. 6, p. e38204.

CHAPTER 6 OVERALL CONCLUSION AND FUTURE DIRECTIONS

In large part, this thesis is about finding rational solutions to the various limitations in the current study of sponge microbiome, such as sponge identification and the associated microbial diversity; as well as about validating and developing new methodology to ensure its relevance and effectiveness for a comprehensive sponge microbiome survey.

6.1 Summary of the project key findings

The key findings have been obtained by answering the following hypotheses:

I. A multilocus-based molecular identification protocol can be effective and reliable for sponge (phylum Porifera) classification.

A new Sponge Identification Protocol (SIP) has been developed based primarily on a molecular identification protocol using multiple loci - COI mtDNA, 28S rRNA gene, and ITS region. The reliability of the SIP was validated by the phylogenetic analysis. The SIP can be an important tool to guide the re-examination of the morphological classification leading to a system that is more effective and reliable for sponge (Porifera) identification. For example, the SIP may point to sponges where alternative morphological characters may be required.

II. The extraction method selected is effective for the recovery of microbial community DNA as demonstrated by actinobacteria (spores and mycelia) within the sponge microbial community.

Actinobacterial spores and mycelium from different genera were added to sponge samples to demonstrate that the methods used for DNA isolation are optimal for the extraction of DNA from these target organisms. This basic step is often missing from other studies that explore microbial diversity. In addition, it was found that per unit weight, actinobacterial spores have a higher amount of DNA than mycelia. This has a potential impact on the calculation of the relative abundance of actinobacterial OTUs. Therefore, cautions should be taken regarding the DNA extraction bias toward sponge microbiome analysis.

III. The influence of the inhibitor(s) if existing in DNA preparation to PCR amplification can be relieved by optimising the conditions for microbial 16S rRNA gene PCR.

Inhibitors of the PCR reaction for the 16S rRNA gene were identified during the amplification of the DNA templates extracted from the artificially mixed sponge-actinobacteria

samples. The effect of these inhibitors was relieved by dilution of the DNA templates. This provided the microbial 16S rRNA gene amplicons for the analysis of the microbial community.

IV. A more comprehensive and reliable sponge microbiome can be revealed through the use of an optimum number of primer-sets targeting different hypervariable regions of the 16S rRNA gene.

The approach developed by applying a combination of multiple primer sets, targeting regions V1V3, V4V5 and V5V8, has revealed a more comprehensive and reliable sponge microbiome than the use of any single primer set. As a result, the OTU coverage generated from the combined datasets of three sequencing markers improved significantly. For example, the primer set specific to region V5V8 uncovered a larger percentage of the unique OTUs. Therefore, the multiple primers targeting different 16S rRNA gene regions are required to reveal the complete sponge microbiome. The baseline study developed an effective approach to improve the understanding of the sponge-associated microbial community. This new finding has profound implications not only for the study of the sponge microbiome but also for most other microbial ecosystems.

V. The structure and the composition of the microbial community of the sponges collected within the same location and the same season are highly specific to the host phylogenetic status (at order and family levels).

The structure and the composition of the sponge microbial community have been proven highly specific to the host sponge phylogeny status. Based on the integrated data using the multi-primer based amplicon sequencing approach, the findings demonstrated that the specificity was not only at the sponge species level but also at the higher taxonomic ranks of the order and family levels. The concept of sponge species-specific microbial OTUs has been enlarged to family-specific and order-specific microbial OTUs, which refer to the dominant OTUs (relative abundance >1%) commonly existing in all the studies sponge species within a sponge order or family. Moreover, the specificity performed on the structure of the microbial community (relative abundance of each taxon within the sample) rather than the diversity of the microbial community (the composition of the taxa within the sample). At the same time, each sponge order or family had the unique associated microbial OTUs that only showed in the studied sponge species belonging to one order or family. They could be regarded as the signature of the particular sponge order or family to offer a rationale guided discovery for the pharmaceutically valuable microorganisms. Importantly, the large number of the unaffiliated microbial OTUs (unassigned and candidate OTUs) were uncovered using this proposed approach, which provides promising untapped microbial resources.

Overall, this PhD project established a pipeline for a more complete characterisation of the sponge microbiome, which includes reliable identification of sponge samples, efficient extraction of community DNA, PCR optimisation, evaluation of region-specific primer sets for 16S rRNA gene based amplicon sequencing, and bioinformatics analysis, for a rationale guided discovery of untapped marine sponge associated microbial resources.

6.2 Limitations and challenges of this study

Considering the project hypotheses and the experimental design, some limitations still exist and can be overcome in the future. In sponge taxonomy study, a sample of 37 sponge species (three individual of each) used to evaluate the Sponge Identification Protocol (SIP) developed in this study was considered by journal reviewers as the minimum requirement for a technically sound methodology establishment. In sponge microbiome analysis, the last hypervariable region V9 of the 16S rRNA gene should be evaluated if significant more microbial OTUs can be revealed. Additionally, a detailed assessment on the different primer sets targeting the same region of 16S rRNA gene (V1-V9) is also necessary to further optimise the multiple primer sets based amplicon sequencing approach proposed in this study.

The data and the analyses answered the hypotheses of the project, though some challenges could not be overcome when conducting the experiments and the analyses. One of such limitations is the sampling times and the availability of the sponges in South Australia for such a big study. For example, for microbiome comparison at the sponge family level, one of the five families had only one species included. If the higher diversity of the representative sponge species could have been collected and analysed, the microbial community comparison at different taxonomic levels (e.g. family and order levels) would be more comprehensive.

6.3 Future research directions

Based on the methodology validation and development in this project, the new paradigm of efficient research tools established allow us to design experiments to answer many different scientific questions. The distinct advantages of this approach can be applied not only in the studies of marine sponge microbiome, but also in any other environmental microbial ecosystem, including the human microbiome analysis. Given the unprecedented capacity in revealing a comprehensive microbiome, the approach developed in this study will revolutionise our scientific understanding of many microbiomes toward practical applications in studies of environment, ecology, biology and human health.

Therefore, many further research directions can be proposed specific to sponge

microbiome. Some of the immediate research questions can be addressed as below.

Firstly, to better understand the concept of 'sponge host specific microbiome', this project could further optimise the methodology and data analysis. For example,

a. To optimise the combination of different region-specific primer sets covering the full length of the 16S rRNA gene (V1-V9) to see if further coverage of new microbial communities could be achieved for a more comprehensive sponge microbiome;

b. To conduct more in-depth keep finish the and comprehensive analysis of the massive volume of sequencing data by thoroughly comparing and discussing the diversity of the unaffiliated microbial OTUs (different taxonomic levels) revealed from different sponge species by different various primer sets in this study to extend our findings in Chapter 5;

c. To collect more sponge species with high diversity for a more comprehensive sponge microbiome survey to generalise our understanding of sponge microbiome;

d. Guided by the findings in this study, the isolation approach can be rationally designed to obtain novel culturable stains, especially those unaffiliated microbes, with bioactivities and further to produce novel natural products for pharmaceutical application.

In addition, understanding how different environmental factors influence the sponge-associated microbial communities is important and valuable for the understanding of sponge ecology. This could be another potential topic for the future research. For example,

a. To compare the microbial community of the same sponge species living at different depths. The surrounding seawater or sediment samples will be the control to analyse the sponge microbial community shift influenced by the environmental factor(s);

b. To compare the microbial community of the same sponge species/ genus living at distinct geographic locations. The microbial community of the seawater or sediment samples will be used as controls to understand the geographic environmental impact on sponge microbiome;

c. To compare the microbiome of the healthy and disease-affected sponge species to identify the pathogen strains threatening sponges.

References

Anderson, SA, Northcote, PT & Page, MJ 2010, 'Spatialand temporal variabilityof the bacterial community in different chemotypes of the New Zealand marine sponge *Mycale hentscheli*', *FEMS Microbiology Ecology*, vol. 72, pp. 328-42.

Bauvais, C, Zirah, S, Piette, L, Chaspoul, F, Domart-Coulon, I, Chapon, V, Gallice, P, Rebuffat, S, Pérez, T & Bourguet-Kondracki, M-L 2015, 'Sponging up metals: Bacteria associated with the marine sponge *Spongia officinalis*', *Marine Environmental Research*, vol. 104, pp. 20-30.

Burgsdorf, I, Erwin, PM, López-Legentil, S, Cerrano, C, Haber, M, Frenk, S & Steindler, L 2014, 'Biogeography rather than association with cyanobacteria structures symbiotic microbial communities in the marine sponge *Petrosia ficiformis*', *Frontiers in Microbiology*, vol. 10, no. 529, pp. 1-11.

Cárdenas, CA, Bell, JJ, Davy, SK, Hoggard, M & Taylor, MW 2014, 'Influence of environmental variation on symbiotic bacterial communities of two temperate sponges', *FEMS Microbiology Ecology*, vol. 88, no. 3, pp. 516-27.

Cebrian, E, Uriz, MJ, Garrabou, J & Ballesteros, E 2011, 'Sponge mass mortalities in a warming mediterranean sea: Are cyanobacteria-harboring species worse off?', *PLoS ONE*, vol. 6, no. 6, p. e20211.

Cebrian, E, Uriz, MJ & Turon, X 2007, 'Sponges as biomonitors of heavy metals inspatial and temporal surveys in Northwestern Mediterranean: Multispecies comparison', *Environmental Toxicology and Chemistry*, vol. 26, pp. 2430-9.

Cleary, DFR, Becking, LE, De Voogd, NJ, Pires, ACC, Polónia, ARM, Egas, C & Gomes, NCM 2013, 'Habitat- and host-related variation in sponge bacterial symbiont communities in Indonesian waters', *FEMS Microbiology Ecology*, vol. 85, no. 3, pp. 465-82.

de Mestre, C, Maher, W, Roberts, D, Broad, A, Krikowa, F & Davis, AR 2012, ' Sponges as sentinels: Patterns of spatial and intra-individual variation in tracemetal concentration', *Marine Pollution Bulletin,* vol. 64, pp. 80-9.

Erwin, PM, Coma, R, López-Sendino, P, Serrano, E & Ribes, M 2015, 'Stable symbionts across the HMA-LMA dichotomy: low seasonal and interannual variation in sponge-associated bacteria from taxonomically diverse hosts', *FEMS Microbial Ecology*, vol. 91, pp. 1-11.

Erwin, PM, Pita, L, López-Legentil, S & Turona, X 2012, 'Stability of Sponge-Associated Bacteria over Large Seasonal Shifts in Temperature and Irradiance', *Applied and Environmental Microbiology*, vol. 78, no. 20, pp. 7358-68.

Fan, L, Liu, M, Simister, R, Webster, NS & Thomas, T 2013, 'Marine microbial symbiosis heats up: the phylogenetic and functional response of a sponge holobiont to thermal stress', *ISME Journal*, vol. 7, no. 5, pp. 991-1002.

Friedrich, AB, Fischer, I, Proksch, P, Hacker, J & Hentschel, U 2001, 'Temporal variation of the microbial community associated with the mediterranean sponge *Aplysina aerophoba*', *FEMS Microbiology Ecology*, vol. 38, pp. 105-13.

Gochfeld, DJ, Easson, CG, Freeman, CJ, Thacker, RW & Olson, JB 2012, 'Disease and nutrient enrichment as potential stressors on the Caribbean sponge *Aplysina cauliformis* and its bacterial symbionts', *Marine Ecology Progress Series*, vol. 456, pp. 101–11.

Hansen, IV, Weeks, JM & Depledge, MH 1995, 'Accumulation of copper, zinc,cadmium and chromium by the marine sponge *Halichondria panacea* Pallas and the implications for biomonitoring', *Marine Pollution Bulletin,* vol. 31, pp. 133-8.

Hentschel, U, Hopke, J, Horn, M, Friedrich, AB, Wagner, M, Hacker, J & Moore, BS 2002, 'Molecular evidence for a uniform microbial community in sponges from different oceans', *Applied and Environmental Microbiology*, vol. 68, no. 9, pp. 4431-40.

Kaluzhnaya, OV & Itskovich, VB 2015, 'Bleaching of baikalian sponge affects the taxonomic composition of symbiotic microorganisms', *Russian Journal of Genetics*, vol. 51, no. 11, pp. 1153-7.

Lee, YK, Jung, HJ & Lee, HK 2006, 'Marine bacteria associated with the Korean brown alga, *Undaria pinnatifida*', *Journal of Microbiology*, vol. 44, pp. 694-8.

Lemoine, N, Buell, N, Hill, A & Hill, M 2007, 'Assessing the utility of sponge microbial symbiont communities as models to study global climate change: a case study with *Halichondria bowerbanki*', *Porifera research: Biodiversity, innovation and sustainability*, pp. 419-26.

López-Legentil, S, Erwin, PM, Pawlik, JR & Song, B 2010, 'Effects of sponge bleaching on ammonia-oxidizing *Archaea*: Distribution and relative expression of ammonia monooxygenase genes associated with the barrel sponge *Xestospongia muta*', *Microbial Ecology*, vol. 60, p. 561.

López-Legentil, S, Song, B, Mcmurray, SE & Pawlik, JR 2008, 'Bleaching and stress in coral reef ecosystems: *hsp70* expression by the giant barrel sponge *Xestospongia muta*', *Molecular Ecology*, vol. 17, no. 7, pp. 1840–9.

Luter, HM, Whalan, S & Webster, NS 2010, 'Prevalence of tissue necrosis and brown spot lesions in a common marine sponge', *Marine and Freshwater Research*, vol. 61, pp. 484-9.

Luter, HM, Widder, S, Botté, ES, Wahab, MA, Whalan, S, Moitinho-Silva, L, Thomas, T & Webster, NS 2015, 'Biogeographic variation in the microbiome of the ecologically important sponge, *Carteriospongia foliascens*', *PeerJ*, vol. 3, p. e1435.

Mangano, S, Michaud, L, Caruso, C & Giudice, AL 2014, 'Metal and antibiotic resistance in psychrotrophic bacteria associated with the Antarctic sponge *Hemigellius pilosus* (Kirkpatrick, 1907)', *Polar Biology*, vol. 37, pp. 227-35.

Mohamed, NM, Enticknap, JJ, Lohr, JE, McIntosh, SM & Hill, RT 2008, 'Changes in bacterial communities of the marine sponge *Mycale laxissima* on transfer into aquaculture', *Applied and Environmental Microbiology*, vol. 74, pp. 1209-22.

Mohamed, NM, Rao, V, Hamann, MT, Kelly, M & Hill, RT 2008, 'Monitoring bacterial diversity of the marine sponge *Ircinia strobilina* upon transfer into aquaculture', *Applied and Environmental Microbiology*, vol. 74, pp. 4133-43.

Montalvo, NF & Hill, RT 2011, 'Sponge-Associated Bacteria Are Strictly Maintained in Two Closely Related but Geographically Distant Sponge Hosts', *Applied and Enrivonmental Micorbiology*, vol. 77, no. 20, pp. 7207-16.

Morrow, KM, Bourne, DG, Humphrey, C, Botté, ES, Laffy, P, Zaneveld, J, Uthicke, S, Fabricius, KE & Webster, NS 2015, 'Natural volcanic CO2 seeps reveal future trajectories for host–microbial associations in corals and sponges', *ISME Journal*, vol. 9, pp. 894–908.

Noyer, C, Casamayor, EO & Becerro, MA 2014, 'Environmental heterogeneity and microbial inheritance influence sponge-associated bacterial composition of *Spongia lamella*', *Microbial Ecology*, vol. 68, pp. 611-20.

Olson, JB & Gao, X 2013, 'Characterizing the bacterial associates of three Caribbean sponges along a gradient from shallow to mesophotic depths', *FEMS Microbiology Ecology*, vol. 85, pp. 74–84.

Olson, JB, Thacker, RW & Gochfeld, DJ 2014, 'Molecular community profiling reveals impacts of time, space, and disease status on the bacterial community associated with the Caribbean sponge Aplysina cauliformis', *FEMS Microbiology Ecology*, vol. 87, no. 1, pp. 268-79.

Pantile, R & Webster, N 2011, 'Strict thermal threshold identified by quantitative PCR in the sponge *Rhopaloeides odorabile*', *Marine Ecology Progress Series,* vol. 431, pp. 97-105.

Patel, B, Balani, MC & Patel, S 1985, 'Sponge 'sentinel' of heavy metals', *Science of the Total Environment*, vol. 41, pp. 143-52.

Pérez, T, Longet, D, Schembri, T, Rebouillon, P & Vacelet, J 2005, 'Effect of 12 years'operation of a sewage treatment plant on trace metal occurrence within a Mediterranean commercial sponge (*Spongia officianalis*, Demospongiae)', *Marine Pollution Bulletin,* vol. 50, pp. 301-9.

Pita, L, Erwin, PM, Turon, X & López-Legentil, S 2013, 'Till death do us part: Stable sponge-bacteria associations under thermal and food shortage stresses', *PLoS ONE*, vol. 8, no. 11, p. e80307.

Sacristán-Soriano, O, Banaigs, B & Becerro, MA 2016, 'Can light intensity cause shifts in natural product and bacterial profiles of the sponge *Aplysina aerophoba*?', *Marine Ecology*, vol. 37, pp. 88-105.

Santos-Gandelman, JF, Cruz, K, Crane, S, Muricy, G, Giambiagi-deMarval, M, Barkay, T & Laport, MS 2014, 'Potential application in mercury bioremediation of a marine sponge-isolated *Bacilluscereus* strain pj1', *Current Microbiology*, vol. 69, no. 3, pp. 374-80.

Schmitt, S, Hentschel, U & Taylor, M 2012, 'Deep sequencing reveals diversity and community structure of complex microbiota in five Mediterranean sponges', in Maldonado M, Turon X, Becerro M & Jesús Uriz M (eds), vol. 219, pp. 341-51, *Ancient Animals, New Challenges*, Springer Netherlands.

Schmitt, S, Tsai, P, Bell, J, Fromont, J, Ilan, M, Lindquist, N, Perez, T, Rodrigo, A, Schupp, PJ, Vacelet, J, Webster, N, Hentschel, U & Taylor, MW 2012, 'Assessing the complex sponge microbiota: core, variable and species-specific bacterial communities in marine sponges', *ISME Journal*, vol. 6, no. 3, pp. 564-76.

Simister, R, Taylor, MW, Tsai, P, Fan, L, Bruxner, TJ, Crowe, ML & Webster, N 2012, 'Thermal stress responses in the bacterial biosphere of the Great Barrier Reef sponge, *Rhopaloeides odorabile*', *Environmental Microbiology*, vol. 14, no. 12, pp. 3232–46.

Simister, R, Taylor, MW, Tsai, P & Webster, N 2012, 'Sponge-microbe associations survive high nutrients and temperatures', *PLoS ONE*, vol. 7, no. 12, p. e52220.

Taylor, MW, Radax, R, Steger, D & Wagner, M 2007, 'Sponge-associated microorganisms: evolution, ecology, and biotechnological potential', *Microbiology and Molecular Biology Reviews*, vol. 71, pp. 295-347.

Taylor, MW, Schupp, PJ, de Nys, R, Kjelleberg, S & Steinberg, PD 2005, 'Biogeography of bacteria associated with the marine sponge *Cymbastela concentrica*', *Environmental Microbiology*, vol. 7, pp. 419-33.

Tian, R-M, Lee, OO, Wang, Y, Cai, L, Bougouffa, S, Chiu, JMY, Wu, RSS & Qian, P-Y 2015, 'Effect of polybrominated diphenyl ether (PBDE) treatment on the composition and function of the bacterial community in the sponge *Haliclona cymaeformis*', *Frontiers in Microbiology*, vol. 5, pp. 1-8.

Tian, R-M, Wang, Y, Bougouffa, S, Gao, Z-M, Cai, L, Zhang, W-P, Bajic, V & Qian, P-Y 2014, 'Effect of copper treatment on the composition and function of the bacterial community in the sponge *Haliclona cymaeformis*', *mBio*, vol. 5, no. 6, pp. e01980-14.

Turque, AS, Batista, D, Silveira, CB, Cardoso, AM, Vieira, RP, Moraes, FC, Clementino, MM, Albano, RM, Paranhos, R, Martins, OB & Muricy, G 2010, 'Environmental shaping of sponge associated archaeal communities', *PLoS ONE*, vol. 5, no. 12, p. e15774.

Wanick, RC, de Sousa Barbosa, H, Frazão, LR, Santelli, RE, Arruda, MAZ & Coutinho, CC 2013, 'Evaluation of differential protein expression in *Haliclona aquarius* and sponge-associated microorganisms under cadmium stress', *Analytical and Bioanalytical Chemistry*, vol. 405, pp. 7661-70.

Webster, NS, Botté, ES, Soo, RM & Whalan, S 2011, 'The larval sponge holobiont exhibits high thermal tolerance', *Environmental Microbiology Report*, vol. 3, no. 6, pp. 756-62.

Webster, NS, Cobb, RE & Negri, AP 2008, 'Temperature thresholds for bacterial symbiosis with a sponge', *ISME Journal*, vol. 2, pp. 830-42.

Webster, NS, Cobb, RE, Soo, R, Anthony, SL, Battershill, CN, Whalan, S & Evans-Illidge, E 2011, 'Bacterial community dynamics in the marine sponge *Rhopaloeides odorabile* under *in situ* and *ex situ* cultivation', *Marine Biotechnology*, vol. 13, pp. 296-304.

Webster, NS & Hill, RT 2001, 'The culturable microbial community of the Great Barrier Reef sponge *Rhopaloeides odorabile* is dominated by an α-Proteobacterium', *Marine Biology*, vol. 138, no. 4, pp. 843-51.

Webster, NS, Webb, RI, Ridd, MJ, Hill, RT & Negri, AP 2001, 'The effects of copper on the microbial community of a coral reef sponge', *Environmental Microbiology*, vol. 3, no. 1, pp. 19-31.

Webster, NS, Xavier, JR, Freckelton, M, Motti, CA & Cobb, R 2008, 'Shifts in microbial and chemical patterns within the marine sponge *Aplysina aerophoba* during a disease outbreak', *Environmental Microbiology*, vol. 10, no. 12, pp. 3366–76.

Weigela, BL & Erwin, PM 2016, 'Intraspecific variation in microbial symbiont communities of the sun sponge, *Hymeniacidon heliophila*, from intertidal and subtidal habitats', *Applied and Enrivonmental Micorbiology*, vol. 82, pp. 650–8.

Whalan, S, Ettinger-Epstein, P & de Nys, R 2008, 'The effect of temperature on larval pre-

settlement duration and metamorphosis for the sponge, *Rhopaloeides odorabile*', *Coral Reefs*, vol. 27, pp. 783–6.

White, JR, Patel, J, Ottesen, A, Arce, G, Blackwelder, P & Lopez, JV 2012, 'Pyrosequencing of Bacterial Symbionts within *Axinella corrugata* Sponges: Diversity and Seasonal Variability', *PLoS ONE*, vol. 7, no. 6, p. e38204.

Wichels, A, Würtz, S, Döpke, H, Schütt, C & Gerdts, G 2006, 'Bacterial diversity in the breadcrumb sponge *Halichondria panicea*', *FEMS Microbial Ecology*, vol. 56, pp. 102-18.

APPENDICES

Appendix Table 2 - 1 Morphological description of thirty-seven sponges

Museum Voucher	Morphological classification	Brief morphological description	Category
SAMA S1962	*Ecionemia* sp.	Massive; russet brown live, cream in ethanol; firm; oscules on ridge. Anatriaenes, oxeas, spiny microrhabds, oxyasters.	Cat I
SAMA S1964	*Tethya* cf. *bergquistae*	Small spherical (3.5cm diam.); orange externally; brown internally; thick cortex (5 mm).	Cat I
SAMA S1966	*Mycale* (*Arenochalina*) sp.	Red fibres. Mucoid. Very soft. Falls apart out of water. Large open oscules. Mycalostyles, no microscleres.	Cat I
SAMA S1972	*Echinodictyum mesenterinum*	Stalked; bilaminate; sculptured ventral surface; reddish brown. Oxeas (many), acanthostyles (many) some with swollen tip.	Cat I
SAMA S1973	*Aplysina lendenfeldi*	Erect; digitate/lobate; black (oxidised) oscule on top of lobe; slimy to touch; fibrous. Yellow internally oxidating to dark blue.	Cat I
SAMA S1974	*Ircinia* sp.	Massive; black surface; creamy yellow inside; firm-compressible. Irciniid filaments.	Cat I
SAMA S1981	*Suberites* sp.	Digitate; orange; very firm, barely compressible; oscules 0.5 - 1 mm diam. obscured by emerging hydroid.	Cat I
SAMA S1987	*Mycale* (*Zygomycale*) sp.	Small soft lobes; heavily fouled; oscules inconspicuous. Styles/ mycalostyles (465 µm), Anisochelae (45 µm). Palmate isochelae (15 µm)	Cat I
SAMA S1996	*Ircinia* sp.	Massive, with ventral crater; beige; conulose; tough/firm.	Cat I
SAMA S1960	*Cliona* sp.	Massive; bright orange; orange internally; firm; oscules 3-4 mm diam. No pore sieves seen. Spicules are tylostyles (tyle on top or just below top of spicule) and spirasters. This is put in Clionaidae as it is not encrusting.	Cat II
SAMA S1971	*Callyspongia* (*Callyspongia*) sp.	Digitate; beige with pink tinge; soft; 2-3 mm oscules along branches. Very firm, springy. Very little tissue between fibres. Strong multispicular fibre network, different size meshes on surface, and not so many spicules. Oxeas very small (60 µm).	Cat II
SAMA S1975	*Crella* sp. 1	Springy small; beige with pink tinge; beige in ethanol; oscules in recesses 2 mm diam; Spicules are tornotes in bunches, acanthostyles, arcuate chelae. Ectosomal tangential layer of acanthostyles.	Cat II
SAMA S1977	*Crella* sp. 1	Small sponge lacunate structure; beige with pink tinge; beige in ethanol; oscules in recesses 2 mm diam. Acanthostyles (60 µm), oxeas - tornotes (165 µm) no styles. Arcuate chelae.	Cat II
SAMA S1978	*Chondropsis* sp.	Lobate to amorphous; beige; skin-like ectosome, does not feel arenaceous; firm not soft. Surface armoured with sand and foreign spicules, the only whole spicules are thin strongyles.	Cat II
SAMA S1989	*Aplysilla rosea*	Pink, fleshy, conulose, encrusting.	Cat II
SAMA S1995	*Echinodictyum mesenterinum*	Bilamellate fan; brown; rugose; firm, oxeas (many) and acanthostyles (few) with swelling at base end.	Cat II
SAMA S1963	Ancorinid sp.	Spherical; dark red with crustose red algae on surface; thick cortex; beige internally; firm; oscules 5 mm diam. Oxeas, euasters, microrhabds (no triaenes).	Cat III
SAMA S1965	*Caulospongia* sp.	Stalked with pagoda-like branches; light brown; easily compressible. Spicules are tylostyles with flattened tyles.	Cat III

SAMA S1982	*Chondropsis* sp.	Lobate; beige; compressible. Oscule at top of lobe. Foreign spicules and strongyles.	Cat III
SAMA S1983	Geodiid sp.	Spherical; black; hard; thick cortex (4 mm), brown internally. Triaenes, oxeas, oxyasters, microrhabds, surface sterrasters.	Cat III
SAMA S1984	*Chondropsis* sp.	Lobate; beige; quite thick surface armour; oscules not seen. Spicules are strongyles, sigmas and chelae. Very soft.	Cat III
SAMA S1991	*Chondropsis* sp.	Lobate; brown; firm-compressible; oscules 6 mm diam. on top of lobes. Whole spicules are strongyles (not numerous) and sigmas.	Cat III
SAMA S1992	*Tedania* cf. *anhelans*	Red; firm-compressible; skin-like ectosome. Microspined tylotes, styles, onychaetes (with no tyle).	Cat III
SAMA S1994	*Chondropsis* sp.	Lobate; orange; firm-compressible; oscules 3 mm diam. at tip of lobes.	Cat III
SAMA S1961	*Spheciospongia* sp.	Massive; pale mottled; very firm; thick skin, ectosomal layer; oscules on ridge 2-3 mm diam. Spicules are styles and diplasters (3 categories).	Cat IV
SAMA S1968	Astrophorin sp.	Massive, pale. Encrusted by soft orange sponge. Thick cortex. Oscules on peaks 2 mm diam. Large oxeas, sanidasters, no oxyasters.	Cat IV
SAMA S1980	*Haliclona* sp.	Creeping digitate; soft; oscules 1mm diam. along branches. Regular ladder-like network - joining fibres one spicule long. Very thin oxeas (90-120 µm).	Cat IV
SAMA S1985	*Pseudoceratina* sp.	Massive; black (oxidised); firm/hard; many small blunt conules close together. Fibres heavily cored with sand. Very slightly lighter inside then externally. Sand in fibres. Although the sponge tissue looks solid, under the microscope it is lacunose in areas. Fibre walls are not thick.	Cat V
SAMA S1988	*Pseudoceratina* sp.	Massive; firm/hard; black (oxidised); blunt conules. Pigment granules.	Cat V
SAMA S1979	*Acanthodendrilla* sp.	Encrusting; pink; soft, but not slippery; conulose. Fibres are reticulate (but not neatly) Many whole poecilosclerid spicules (chelae, styles, acanthostyles) in tissue. Surface is armoured and some of the primary fibres completed cored with sand.	Cat V
SAMA S1967	*Callyspongia bilamellata)*	Bilaminate; stalked; russet brown; sculptured underside; sticky to touch; Pale brown inside. Spicules are oxeas.	Cat V
SAMA S1969	Spongiid sp.	Erect; bright orange. Firm-compressible. Fibres obvious through ectosome. Clear fibres. Dictyoceratida, Spongiidae. No irciniid fibres.	Cat V
SAMA S1970	Thorectid sp.	Fibrous. Slightly mucoid. Pale. Striated surface. Oscules recessed. Laminated reticulate uncored fibres. Fibres darker than choanosome.	Cat V
SAMA S1967	*Chondrosia* sp.	Spherical; brown externally (ectosomal skin), beige internally. Compressible. Spicules seem to be foreign, mixed eg. oxeas, styles, anisooxeas, acanthostyles and many broken spicules. Fibrous choanosome.	Cat V

Appendix Table 2 - 2 DNA loci sequencing results with Bit Score, sequence similarity, coverage percentage and the morphological classifications of thirty-seven sponges. No. 3-1 ~ 3-5 are the sponges with three valid loci sequence information; No. 2-1 ~ 2-17are the sponges with two loci sequence information; No. 1-1 ~ 1-12 are the sponges with only one locus information; No. 0-1 ~ 0-3 are the sponges without inferences according to the Sponge Identification Protocol (SIP) in this study; v indicates the one inferring the final identity; Underline labels the excluded data with substandard coverage region (<50%).

No.	**Museum Voucher**	**COI Identities**				**28S Identities**				**ITS Identities**				**Morphological Classification (Order-Family-Genus/Species)**
		PCR	Seq	Accession No.	Result (Order-Family-Genus/Species)	PCR	Seq	Accession No.	Result (Order-Family-Genus/Species)	PCR	Seq	Accession No.	Result (Order-Family-Genus/Species)	
3-1	SAMA S1981	v	v	KJ546357	Suberitida; Suberitidae; *Rhizaxinella* sp. (608, 98%, 96%)	v	v	KJ620381	Suberitida; Suberitidae; *Suberites aurantiacus* (975 v, 98%, 97%)	v	v	KJ782592	Suberitida; Halichondriidae; *Hymeniacidon heliophila* (274, 84%, 62%)	Suberitida; Suberitidae; *Suberites* sp.
3-2	SAMA S1960	v	v	KJ620399	Clionaida; Clionaidae; *Clionaopsis platei* (605, 97%, 96%)	v	v	KJ620386	Clionaida; Spirastrellidae; *Spirastrella hartmani* (1012 v, 99%, 98%)	v	v	KJ782595	Clionaida; Spirastrellidae; *Spirastrella hartmani* (480, 88%, 99%)	Clionaida; Clionaidae; *Cliona* sp.
3-3	SAMA S1965	v	v	KJ620404	Suberitida; Suberitidae; '*Protosuberites*' sp. (598, 97%, 97%)	v	v	KJ620391	Suberitida; Halichondriidae; *Hymeniacidon heliophila* (1004 v, 99%, 96%)	v	v	KJ782600	Suberitida; Halichondriidae; *Hymeniacidon heliophila* (388, 84%, 98%)	Suberitida; Suberitidae; *Caulospongia* sp.
3-4	SAMA S1971	v	v	KJ620407	Haplosclerida; Callyspongiidae; *Callyspongia siphonella* (581, 98%, 91%)	v	v	KJ620394	Haplosclerida; Chalinidae; *Cladocroce* sp. (1085 v, 99%, 98%)	v	v	KJ782602	Haplosclerida; Chalinidae; *Haliclona* sp. (183, 80%, 55%)	Haplosclerida; Callyspongiidae; *Callyspongia* sp.
3-5	SAMA S1973	v	v	KJ620409	Verongiida; Aplysinidae; *Aplysina lacunose* (611, 98%, 96%)	v	v	KJ620395	Verongiida; Aplysinidae; *Aplysina archeri* (1005 v, 98%, 97%)	v	v	KJ782604	Verongiida; Aplysinidae; *Verongula gigantean* (335, 90%, 81%)	Verongiida; Aplysinidae; *Aplysina lendenfeldi*

2-1	SAMA S1989	v	v	KJ546351	Dendroceratida; Dictyodendrillidae; *Igernella notabilis* (629, 98%, 98%)	v	v	KJ620376	Dendroceratida; Dictyodendrillidae; *Igernella notabilis* (977 v, 98%, 98%)	v	×		CANCEL (The other two are all valid and matched at the species level.)	Dendroceratida; Darwinellidae; *Aplysilla rosea*
2-2	SAMA S1991	v	v	KJ546352	Poecilosclerida; Desmacididae; *Desmapsamma anchorata* (612, 98%, 97%)	v	v	KJ620377	Poecilosclerida; Tedaniidae; *Tedania tubulifera* (992 v, 99%, 97%)	v	v	-	Polymastiida; Polymastiidae; *Polymastia pachymastia* (34, 98%, 3%)	Poecilosclerida; Chondropsidae; *Chondropsis* sp.
2-3	SAMA S1993	v	v	KJ546354	Poecilosclerida; Desmacididae; *Desmapsamma anchorata* (614 v, 98%, 97%)	×	×		Try the second pair of primers (Positive)	v	v		Not sponge	Poecilosclerida; Tedaniidae; *Tedania* cf. *anhelans*
2-4	SAMA S1994	v	v	KJ546355	Poecilosclerida; Tedaniidae; *Tedania ignis* (604, 98%, 95%)	v	v	KJ620378	Poecilosclerida; Tedaniidae; *Tedania tubulifera* (1000 v, 99%, 96%)	×	×		CANCEL (The other two are all valid and matched at the genus level.)	Poecilosclerida; Chondropsidae; *Chondropsis* sp.
2-5	SAMA S1995	v	v	KJ546356	Axinellida; Raspailiidae; *Echinodictyum cancellatum* (470, 93%, 86%)	v	v	KJ620379	Axinellida; Raspailiidae; *Raspailia vestigifera* (822 v, 96%, 82%)	v	×		CANCEL (The other two are all valid and matched at the family level.)	Axinellida; Raspailiidae; *Echinodictyum mesenterium*
2-6	SAMA S1996	×	×		CANCEL (The other two are all valid and matched at the genus level.)	v	v	KJ620380	Dictyoceratida; Irciniidae; *Ircinia strobilina* (845 v, 94%, 94%)	v	v	KJ801654	Dictyoceratida; Irciniidae; *Ircinia felix* f. *felix* (595, 90%, 96%)	Dictyoceratida; Irciniidae; *Ircinia* sp.
2-7	SAMA S1982	v	v	KJ546358	Poecilosclerida; Desmacididae; *Desmapsamma anchorata* (615, 99%, 93%)	v	v	KJ620382	Poecilosclerida; Tedaniidae; *Tedania tubulifera* (979 v, 98%, 98%)	v	v	-	Poecilosclerida; Tedaniidae; *Tedania ignis* (115, 90%, 20%)	Poecilosclerida; Chondropsidae; *Chondropsis* sp.
2-8	SAMA S1983	v	v	KJ546359	Tetractinellida; Ancorinidae; *Ancorina* sp. (600, 98%, 94%)	v	v	KJ620383	Tetractinellida; Ancorinidae; *Tethyopsis mortenseni* (967 v, 98%, 97%)	v	v	-	Tetractinellida; Geodiidae; *Pachymatisma johnstonia* (153, 96%, 17%)	Tetractinellida; Geodiidae; Geodiid sp.
2-9	SAMA S1984	v	v	KJ546360	Poecilosclerida; Desmacididae; *Desmapsamma anchorata* (612, 98%, 96%)	v	v	KJ620384	Poecilosclerida; Tedaniidae; *Tedania tubulifera* (870 v, 95%, 92%)	v	×		CANCEL (The other two are all valid and matched at the Order level.)	Poecilosclerida; Chondropsidae; *Chondropsis* sp.
2-10	SAMA S1987	v	v	KJ546362	Poecilosclerida; Podospongiidae; *Diacarnus spinipoculum* (173, 75%, 97%)	v	v	KJ620385	Poecilosclerida; Mycalidae; *Mycale setosa* (305 v, 84%, 55%)	×	×		CANCEL (The other two are all valid and matched at the Order level.)	Poeciliosclerida; Mycalidae; *Mycale* (*Zygomycale*) sp.

2-11	SAMA S1976	v	v	KJ546365	Suberitida; Halichondriidae; *Halichondria okadai* (576, 96%, 95%)	v	×		CANCEL (The other two are all valid and matched at the species level.)	v	v	KJ801656	Suberitida; Halichondriidae; *Halichondria okadai* (617 v, 94%, 99%)	Chondrillida Chondrillidae *Chondrosia* sp.
2-12	SAMA S1961	v	v	KJ620400	Poecilosclerida; Podospongiidae; *Diacarnus spinipoculum* (657, 99%, 97%)	v	v	KJ620387	Poecilosclerida; Podospongiidae*;* *Diacarnus spinipoculum* (953 v, 99%, 92%)	v	v	-	Poecilosclerida; Tedaniidae; *Tedania ignis* (153, 92%, 21%)	Clionaida; Spirastrellidae; *Spheciospongia* sp.
2-13	SAMA S1962	v	v	KJ620401	Tetractinellida; Ancorinidae; *Ecionemia robusta* (1010 v, 99%, 99%)	v	v	KJ620388	Tetractinellida; Ancorinidae; *Stelletta clavosa* (642, 99%, 95%)	v	v	-	Tetractinellida; Geodiidae; *Pachymatisma johnstonia* (154, 97%, 18%)	Tetractinellida; Ancorinidae; *Ecionemia* sp.
2-14	SAMA S1963	v	v	KJ620402	Tetractinellida; Ancorinidae; *Ecionemia* sp. (641, 99%, 94%)	v	v	KJ620389	Tetractinellida; Ancorinidae; *Stelletta clavosa* (1023 v, 99%, 97%)	v	v	-	Tetractinellida; Tetillidae; *Cinachyrella apion* (169, 98%, 18%)	Tetractinellida; Ancorinidae; Ancorinid sp.
2-15	SAMA S1964	v	v	KJ620403	Tethyida; Tethyidae; *Tethya californiana* (576, 96%, 96%)	v	v	KJ620390	Tethyida; Tethyidae; *Tethya* sp. (973 v, 98%, 98%)	v	v	-	Tethyida; Tethyidae; *Tethya* sp. (163, 100%, 38%)	Tethyida; Tethyidae; *Tethya* cf. *bergquistae*
2-16	SAMA S1966	v	v	KJ620405	Poecilosclerida; Mycalidae; *Mycale mirabilis* (633, 99%, 94%)	v	v	KJ620392	Poecilosclerida; Mycalidae; *Mycale setosa* (1008 v, 99%, 85%)	v	v	-	Agelasida; Agelasidae; *Agelas schmidti* (132, 93%, 167/912=18%)	Poecilosclerida; Mycalidae; *Mycale* sp.
2-17	SAMA S1968	v	v	KJ620406	Poecilosclerida; Microcionina; *Clathria rugosa* (510 v, 93%, 94%)	v	×		CANCEL (The other two are all available.)	v	v	KJ801658	Tetractinellida; Ancorinidae; *Stryphnus mucronatus* (380, 90%, 65%)	Tetractinellida; Astrophorin sp.
1-1	SAMA S1985	v	v	KJ546361	Verongiida; Pseudoceratinidae; *Pseudoceratina* sp. (634 v, 99%, 95%)	v	×		Try the second pair of primers (Negative)	v	×		Try other primers (Not done in this study.)	Verongiida; Pseudoceratinidae; *Pseudoceratina* sp.
1-2	SAMA S1988	v	v	KJ546363	Verongiida; Pseudoceratinidae; *Pseudoceratina* sp. (619 v, 99%, 93%)	×	×		Try the second pair of primers (Negative)	v	v	-	R: not belong to Porifera	Verongiida; Pseudoceratinidae; *Pseudoceratina* sp.
1-3	SAMA S1975	v	v	KJ546364	Poecilosclerida; Hymedesmiidae; *Phorbas bihamiger* (469 v, 91%, 95%)	×	×		Try the second pair of primers (Negative)	×	×		Try other primers (Not done in this study.)	Poecilosclerida; Crellidae; *Crella* sp.

1-4	SAMA S1977	v	v	KJ546366	Poecilosclerida; Hymedesmiidae; *Phorbas bihamiger* (453 v, 91%, 90%)	×	×		Try the second pair of primers (Negative)	×	×		Try other primers (Not done in this study.)	Poecilosclerida; Crellidae; *Crella* sp.
1-5	SAMA S1978	v	v	KJ546367	Poecilosclerida; Desmacididae; *Desmapsamma anchorata* (581 v, 97%, 93%)	×	×		Try the second pair of primers (Negative)	×	×		Try other primers (Not done in this study.)	Poecilosclerida; Chondropsidae; *Chondropsis* sp.
1-6	SAMA S1979	v	v	KJ546368	Dendroceratida; Dictyodendrillidae; *Acanthodendrilla australis* (526 v, 97%, 84%)	×	×		Try the second pair of primers (Negative)	v	v	-	R: not belong to Porifera	Dendroceratida; Dictyodendrillidae; *Acanthodendrilla* sp.
1-7	SAMA S1980	v	v	KJ620398	Suberitida; Suberitidae; '*Protosuberites*' sp. (404 v, 87%, 96%)	×	×		Try the second pair of primers (Negative)	v	×		Try other primers (Not done in this study.)	Haplosclerida; Chalinidae; *Haliclona* sp.
1-8	SAMA S1967	v	v	-	F: not belong to Porifera	v	v	KJ620393	Haplosclerida; Petrosiidae; *Petrosia lignosa* (826 v, 93%, 96%)	v	v	-	Haplosclerida; Niphatidae; *Amphimedon chloros* (130, 93%, 19%)	Haplosclerida; Callyspongiidae; *Callyspongia bilamellata*
1-9	SAMA S1969	×	×		Try nested-PCR (Negative)	×	×		Try the second pair of primers (Positive)	v	v	KJ801659	Dictyoceratida; Irciniidae; *Ircinia felix* f. *felix* (358 v, 87%, 70%)	Dictyoceratida; Spongiid sp.
1-10	SAMA S1970	×	×		Try nested-PCR (Negative)	×	×		Try the second pair of primers (Negative)	v	v	KJ801660	Dictyoceratida; Irciniidae; *Ircinia felix* f. *felix* (349 v, 81%, 99%)	Dictyoceratida; Thorectid sp.
1-11	SAMA S1972	v	v	KJ620408	Axinellida; Raspailiidae; *Echinodictyum cancellatum* (467 v, 93%, 85%)	v	v	-	R: not belong to Porifera	v	v	-	Axinellida; Axinellidae; *Acanthella pulcherrima* (250, 95%, 36%)	Axinellida; Raspailiidae; *Echinodictyum mesenterinum*
1-12	SAMA S1974	×	×		Try nested-PCR (Negative)	×	×		Try the second pair of primers (positive)	v	v	KJ801661	Dictyoceratida; Irciniidae; *Ircinia felix* f. *felix* (389 v, 83%, 100%)	Dictyoceratida; Irciniidae; *Ircinia* sp.
0-1	SAMA S1990	×	×		Try nested-PCR (Negative)	×	×		Try the second pair of primers (Positive)	v	v	-	Dendroceratida; Dictyodendrillidae; *Igernella notabilis* (163, 99%, 19%)	Dendroceratida; Darwinellidae ; *Aplysilla rosea*

0-2	SAMA S1992	v	v	-	Poecilosclerida; Microcionidae; *Microciona prolifera* (124, 97%, 80%) F: only 170 bp in total	×	×		CANCEL (The other two are all available.)	v	v	-	Poecilosclerida; Microcionidae; *Microciona prolifera* (109, 91%, 155/853=18%)	Poecilosclerida; Poecilosclerid sp.
0-3	SAMA S1986	×	×		CANCEL (The other two are all available.)	v	v	-	Haplosclerida; unclassified Haplosclerida; *Haplosclerine* sp. (705, 88%, 93%) R: not belong to Porifera	v	v	-	Haplosclerida; Niphatidae; *Amphimedon queenslandica* (119, 91% , 24%) F: not belong to Porifera	Haplosclerida; Chalinidae; *Chalinula* sp.

Appendix Table 2 - 3 Sponges with any two matched identities among the three DNA loci and the morphological classification at different taxonomic levels

Order level match

No.	Sponge Code	Loci providing the identities [a]	Identity Information (Order)
1	SAMA S1991	COI, 28S, M	Poecilosclerida
2	SAMA S1993	COI, M	
3	SAMA S1982	COI, 28S, M	
4	SAMA S1984	COI, 28S, M	
5	SAMA S1975	COI, M	
6	SAMA S1977	COI, M	
7	SAMA S1978	COI, M	
8	SAMA S1968	ITS, M	Tetractinellida
9	SAMA S1970	ITS, M	Dictyoceratida
10	SAMA S1969	ITS, M	
11	SAMA S1967	28S, M	Haplosclerida

Family level match

No.	Sponge Code	Loci providing the identities	Identity Information (Family-Order)	
1	SAMA S1983	COI, 28S	Ancorinidae	Tetractinellida
2	SAMA S1963	COI, M		

Genus level match

No.	Sponge Code	Loci providing the identities	Identity Information (Genus-Family-Order)		
1	SAMA S1994	COI, 28S	*Tedania*	Tedaniidae	Poecilosclerida
2	SAMA S1987	28S, M	*Mycale*	Mycalidae	
3	SAMA S1966	COI, 28S, M			
4	SAMA S1995	COI, M	*Echinodictyum*	Raspailiidae	Axinellida
5	SAMA S1972	COI, M			
6	SAMA S1981	28S, M	*Suberite*	Suberitidae	Suberitida
7	SAMA S1964	COI, 28S, M	*Tethya*	Tethyidae	Tethyida

8	SAMA S1971	COI, M	*Callyspongia*	Callyspongiidae	Haplosclerida
9	SAMA S1996	28S, ITS, M	*Ircinia*	Irciniidae	Dictyoceratida
10	SAMA S1974	ITS, M			
11	SAMA S1962	COI, M	*Ecionemia*	Ancorinidae	Tetractinellida
12	SAMA S1973	COI, 28S, M	*Aplysina*	Aplysinidae	Verongiida
13	SAMA S1985	COI, M	*Pseudoceratina*	Pseudoceratinidae	
14	SAMA S1988	COI, M			
15	SAMA S1979	COI, M	*Acanthodendrilla*	Dictyodendrillidae	Dendroceratida

Species level match

No.	Sponge Code	Loci providing the identities	Identity Information (Species-Family-Order)		
1	SAMA S1965	28S, ITS	*Hymeniacidon heliophila*	Hymeniacidonidae	Suberitida
2	SAMA S1976	COI, ITS	*Halichondria okadai*	Halichondriidae	
3	SAMA S1960	28S, ITS	*Spirastrella hartmani*	Spirastrellidae	Clionaida
4	SAMA S1961	COI, 28S	*Diacarnus spinipoculum*	Podospongiidae	Poecilosclerida
5	SAMA S1989	COI, 28S	*Igernella notabilis*	Dictyodendrillidae	Dendroceratida

[a] The COI, 28S, ITS and M in the item of Loci providing the identities represent the COI mtDNA, 28S rRNA gene, nuclear ITS region and the morphological classification.

Appendix Table 4 - 1 The genera revealed by different primer sets for four sponge species

Aplysina archeri				*Halichondria okadai*			
V1V3	**V4**	**V4V5**	**V5V8**	**V1V3**	**V4**	**V4V5**	**V5V8**
Acidipila	*Aciditerrimonas*	*Aciditerrimonas*	*Acetomicrobium*	*Anabaena*	*Aciditerrimonas*	*Aciditerrimonas*	*Aciditerrimonas*
Acidobacterium	*Alkalispirillum*	*Alkalispirillum*	*Acetonema*	*Aquibacter*	*Algibacter*	*Algibacter*	*Actibacter*
Aquihabitans	*Aquihabitans*	*Aquibacter*	*Acidimicrobium*	*Aquihabitans*	*Alkalispirillum*	*Alkalispirillum*	*Afifella*
Arthrobacter	*Bartonella*	*Aquihabitans*	*Aciditerrimonas*	*Arthrobacter*	*Aquihabitans*	*Aquibacter*	*Algibacter*
Brevibacterium	*Blastopirellula*	*Blastopirellula*	*Alcanivorax*	*Azospirillum*	*Blastopirellula*	*Aquihabitans*	*Alkalilimnicola*
Caenispirillum	*Candidatus Methylomirabilis*	*Conexibacter*	*Anaerolinea*	*Bauldia*	*Conexibacter*	*Bartonella*	*Anaerolinea*
Candidatus Methylomirabilis	*Candidatus Puniceispirillum*	*Cyanothece*	*Blastopirellula*	*Brumimicrobium*	*Cyanothece*	*Blastopirellula*	*Blastopirellula*
Candidatus Puniceispirillum	*Candidatus Solibacter*	*Dehalobacter*	*Bryobacter*	*Caenispirillum*	*Dehalobacter*	*Candidatus Methylomirabilis*	*Brevundimonas*
Candidatus Solibacter	*Clostridium*	*Desulfomonile*	*Calderihabitans*	*Candidatus Methylomirabilis*	*Ectothiorhodospira*	*Candidatus Puniceispirillum*	*Candidatus Portiera*
Clostridium	*Conexibacter*	*Ectothiorhodospira*	*Caldilinea*	*Candidatus Puniceispirillum*	*Ekhidna*	*Candidatus Solibacter*	*Chlamydia*
Coraliomargarita	*Cyanothece*	*Endozoicomonas*	*Cobetia*	*Candidatus Solibacter*	*Endozoicomonas*	*Clostridium*	*Chroococcales*
Cryobacterium	*Dehalobacter*	*Ferrimicrobium*	*Conexibacter*	*Coxiella*	*Ensifer*	*Conexibacter*	*Cohaesibacter*
Cyanobium	*Desulfomonile*	*Formivibrio*	*Dehalococcoides*	*Cryobacterium*	*Ferrimicrobium*	*Cyanothece*	*Conexibacter*
Deferribacter	*Ectothiorhodospira*	*Granulosicoccus*	*Desulfacinum*	*Cyanobium*	*Geodermatophilus*	*Dehalobacter*	*Cystobacter*
Desulfomonile	*Endozoicomonas*	*Halioglobus*	*Desulfothermus*	*Desulfomonile*	*Granulosicoccus*	*Desulfomonile*	*Dehalococcoides*
Desulfovibrio	*Ferrimicrobium*	*Halochromatium*	*Dissulfuribacter*	*Dethiosulfovibrio*	*Halioglobus*	*Ectothiorhodospira*	*Desulfacinum*
Devosia	*Halochromatium*	*Henriciella*	*Ectothiorhodospira*	*Devosia*	*Halochromatium*	*Ekhidna*	*Dissulfuribacter*
Formivibrio	*Henriciella*	*Iamia*	*Eggerthella*	*Granulicella*	*Henriciella*	*Endozoicomonas*	*Ekhidna*
Frigoribacterium	*Hoeflea*	*Jannaschia*	*Endozoicomonas*	*Hippea*	*Hoeflea*	*Ensifer*	*Endozoicomonas*
Frondihabitans	*Iamia*	*Laceyella*	*Flavobacterium*	*Hydrogenivirga*	*Jannaschia*	*Ferrimicrobium*	*Filomicrobium*
Granulicella	*Jannaschia*	*Lacibacterium*	*Gaiella*	*Iamia*	*Kordia*	*Formivibrio*	*Fischerella*
Halorhodospira	*Laceyella*	*Lactobacillus*	*Gemmatimonas*	*Laceyella*	*Lactobacillus*	*Geodermatophilus*	*Flavobacterium*
Hippea	*Lactobacillus*	*Litorilinea*	*Geobacter*	*Lactobacillus*	*Litorilinea*	*Granulosicoccus*	*Gaetbulibacter*
Hydrogenivirga	*Litorilinea*	*Nitrosopumilus*	*Geodermatophilaceae*	*Magnetovibrio*	*Nitrospira*	*Halioglobus*	*Gaiella*
Hydrogenobacter	*Nitrosopumilus*	*Nitrospira*	*Geothermobacter*	*Moraxella*	*Oceanibaculum*	*Halochromatium*	*Gangjinia*
Hydrogenobaculum	*Nitrospira*	*Octadecabacter*	*Geothermobacterium*	*Niastella*	*Octadecabacter*	*Henriciella*	*Geobacter*
Iamia	*Oceanibaculum*	*Persephonella*	*Hahella*	*Nisaea*	*Olleya*	*Hoeflea*	*Geothermobacterium*
Lacibacterium	*Octadecabacter*	*Phaeocystidibacter*	*Halochromatium*	*Pelagicola*	*Persephonella*	*Iamia*	*Gilvibacter*
Lactobacillus	*Olleya*	*Planifilum*	*Hippea*	*Persephonella*	*Phaeocystidibacter*	*Jannaschia*	*Halorhodospira*
Limimonas	*Persephonella*	*Prochlorococcus*	*Holophaga*	*Pseudothermotoga*	*Planifilum*	*Kordia*	*Hoeflea*
Magnetovibrio	*Prochlorococcus*	*Rhodopirellula*	*Iamia*	*Psychroserpens*	*Prochlorococcus*	*Laceyella*	*Iamia*
Maricaulis	*Pseudomonas*	*Rhodothermus*	*Kordiimonas*	*Rhodopirellula*	*Pseudofulvibacter*	*Litorilinea*	*Kordiimonas*
Niastella	*Rhodopirellula*	*Rhodovulum*	*Levilinea*	*Rhodothermus*	*Pseudomonas*	*Nitrospira*	*Kribbella*
Nitratiruptor	*Rhodothermus*	*Rudaea*	*Longilinea*	*Roseospirillum*	*Rhodopirellula*	*Oceanibaculum*	*Lacibacterium*
Olsenella	*Rhodovulum*	*Sneathiella*	*Magnetospira*	*Sphaerobacter*	*Rhodothermus*	*Octadecabacter*	*Lactobacillus*
Pelagicola	*Roseibacillus*	*Solirubrobacter*	*Marinobacter*	*Sphingopyxis*	*Roseibium*	*Olleya*	*Megasphaera*
Planifilum	*Rudaea*	*Sphaerobacter*	*Megasphaera*	*Spirochaeta*	*Rudaea*	*Phaeocystidibacter*	*Meridianimaribacter*
Pseudothermotoga	*Sneathiella*	*Spirochaeta*	*Methylogaea*	*Streptomyces*	*Sagittula*	*Phyllobacterium*	*Methylogaea*
Rhodopirellula	*Solirubrobacter*	*Stanieria*	*Nitrosopumilus*	*Synechococcus*	*Solirubrobacter*	*Planifilum*	*Microbulbifer*
Rhodothermus	*Sphaerobacter*	*Tepidibacter*	*Nitrospina*	*Syntrophorhabdus*	*Spirochaeta*	*Prochlorococcus*	*Neorickettsia*
Roseiflexus	*Spirochaeta*	*Thermaerobacter*	*Ornatilinea*	*Tahibacter*	*Stanieria*	*Pseudofulvibacter*	*Nitrosopumilus*
Roseomonas	*Tepidibacter*	*Thermobaculum*	*Patulibacter*	*Terracoccus*	*Tahibacter*	*Pseudomonas*	*Nitrospina*

Roseospirillum	Thermaerobacter	Thermodesulforhabdus	Pectinatus	Thermoflexus	Thermaerobacter	Rhodopirellula	Oceanibacterium
Sphaerobacter	Thermobaculum	Thermovenabulum	Pelobacter	Thioalbus	Thermovenabulum	Roseibacillus	Ochrobactrum
Spirochaeta	Thermodesulforhabdus	Thioalbus	Planomicrobium	Thiobios	Thioalbus	Roseibium	Ornatilinea
Streptomyces	Thermovenabulum	Thioalkalivibrio	Propionibacterium	Thiococcus	Thioalkalivibrio	Rudaea	Owenweeksia
Synechococcus	Thioalbus	Thioprofundum	Propionicimonas	Thiohalocapsa		Sagittula	Pectinatus
Syntrophorhabdus	Thioalkalivibrio		Pseudomonas	Trichormus		Sneathiella	Pelobacter
Thermodesulfobium	Thioprofundum		Rhodobacter	Vibrio		Solirubrobacter	Phyllobacterium
Thermodesulfovibrio			Rhodovulum			Spirochaeta	Planktotalea
Thermoflexus			Rubrivirga			Stanieria	Propionibacterium
Thiobios			Skermanella			Tepidibacter	Pseudovibrio
Thiococcus			Solirubrobacter			Thermaerobacter	Rhodobacter
Thiohalocapsa			Spiribacter			Thermobaculum	Rhodopirellula
			Spirochaeta			Thermodesulforhabdus	Rhodothermus
			Steroidobacter			Thermovenabulum	Rhodovulum
			Synechococcus			Thioalkalivibrio	Roseibacterium
			Syntrophomonas			Thioprofundum	Ruegeria
			Syntrophus				Sneathiella
			Thalassobaculum				Solirubrobacter
			Thermasporomyces				Sphaerobacter
			Thermoanaerobaculum				Spiribacter
			Thermodesulfobium				Spirochaeta
			Thermoleophilum				Synechococcus
			Thermolithobacter				Terasakiella
			Thioalbus				Thermoleophilum
			Thioalkalivibrio				Thioalkalivibrio
			Thiohalocapsa				Thiohalophilus
			Thiohalophilus				Wenyingzhuangia
			Thioprofundum				

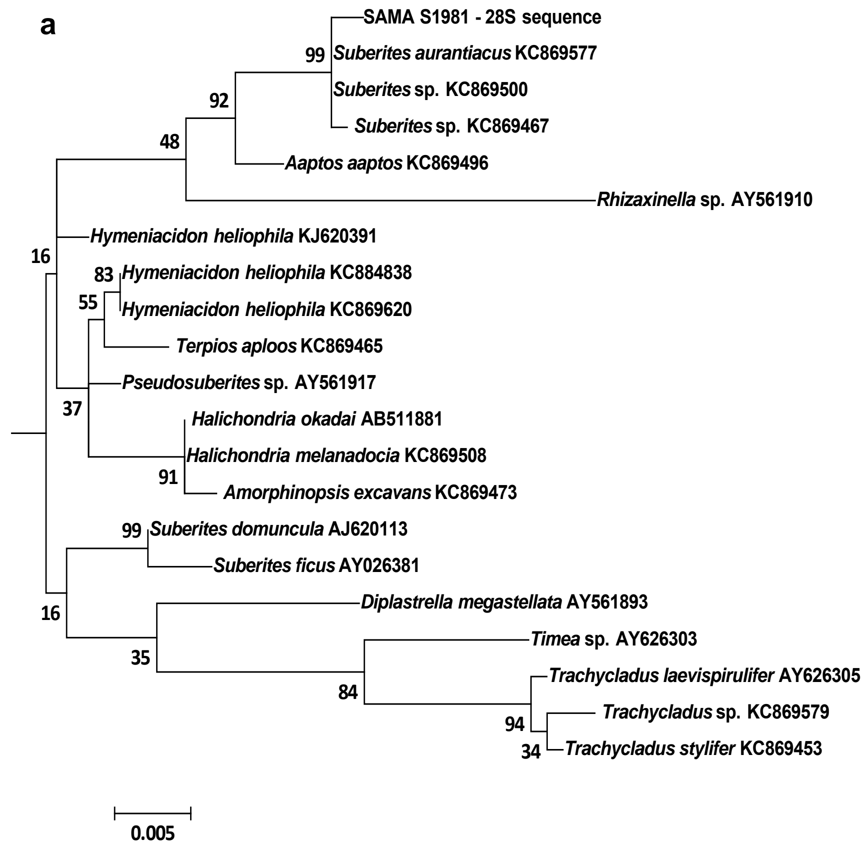
Igernella notabilis				Tedania tubulifera			
V1V3	**V4**	**V4V5**	**V5V8**	**V1V3**	**V4**	**V4V5**	**V5V8**
Acidipila	Aciditerrimonas	Aciditerrimonas	Acetomicrobium	Acidithiomicrobium	Aciditerrimonas	Aciditerrimonas	Acetomicrobium
Acidobacterium	Algibacter	Algibacter	Acetonema	Acidobacterium	Alkalispirillum	Alkalispirillum	Acetonema
Alkalispirillum	Alkalispirillum	Aquibacter	Aciditerrimonas	Anabaena	Aquihabitans	Aquibacter	Aciditerrimonas
Aquihabitans	Aquibacter	Aquihabitans	Algibacter	Aquihabitans	Bartonella	Aquihabitans	Aliihoeflea
Arthrobacter	Aquihabitans	Bartonella	Aliihoeflea	Arthrobacter	Blastopirellula	Bartonella	Alphaproteobacteria
Brevibacterium	Blastopirellula	Blastopirellula	Anaerolinea	Brevibacterium	Clostridium	Blastopirellula	Anaerolinea
Brumimicrobium	Candidatus Methylomirabilis	Candidatus Solibacter	Aquamicrobium	Caenispirillum	Conexibacter	Candidatus Methylomirabilis	Blastopirellula
Caenispirillum	Candidatus Puniceispirillum	Clostridium	Blastopirellula	Candidatus Methylomirabilis	Dehalobacter	Candidatus Puniceispirillum	Bryobacter
Candidatus Methylomirabilis	Candidatus Solibacter	Conexibacter	Bryobacter	Candidatus Pelagibacter	Desulfomonile	Candidatus Solibacter	Calderihabitans
Candidatus Pelagibacter	Clostridium	Dehalobacter	Calderihabitans	Candidatus Puniceispirillum	Ectothiorhodospira	Clostridium	Caldilinea
Candidatus Puniceispirillum	Conexibacter	Desulfomonile	Caldilinea	Candidatus Solibacter	Ekhidna	Conexibacter	Chroococcales
Coraliomargarita	Cyanothece	Ectothiorhodospira	Candidatus Portiera	Chloroflexus	Endozoicomonas	Dehalobacter	Cobetia
Cryobacterium	Dehalobacter	Ekhidna	Chroococcales	Cryobacterium	Ferrimicrobium	Desulfomonile	Conexibacter
Cyanobium	Desulfomonile	Endozoicomonas	Cobetia	Cyanobium	Formivibrio	Ectothiorhodospira	Dehalococcoides
Deferribacter	Ectothiorhodospira	Ferrimicrobium	Conexibacter	Deferribacter	Geodermatophilus	Ekhidna	Desulfacinum
Desulfomonile	Endozoicomonas	Formivibrio	Dehalococcoides	Desulfomonile	Granulosicoccus	Endozoicomonas	Ectothiorhodospira
Desulfovibrio	Ferrimicrobium	Geodermatophilus	Desulfacinum	Desulfovibrio	Halioglobus	Ferrimicrobium	Eggerthella
Dethiosulfovibrio	Granulosicoccus	Granulosicoccus	Desulfothermus	Dethiosulfovibrio	Henriciella	Formivibrio	Filomicrobium

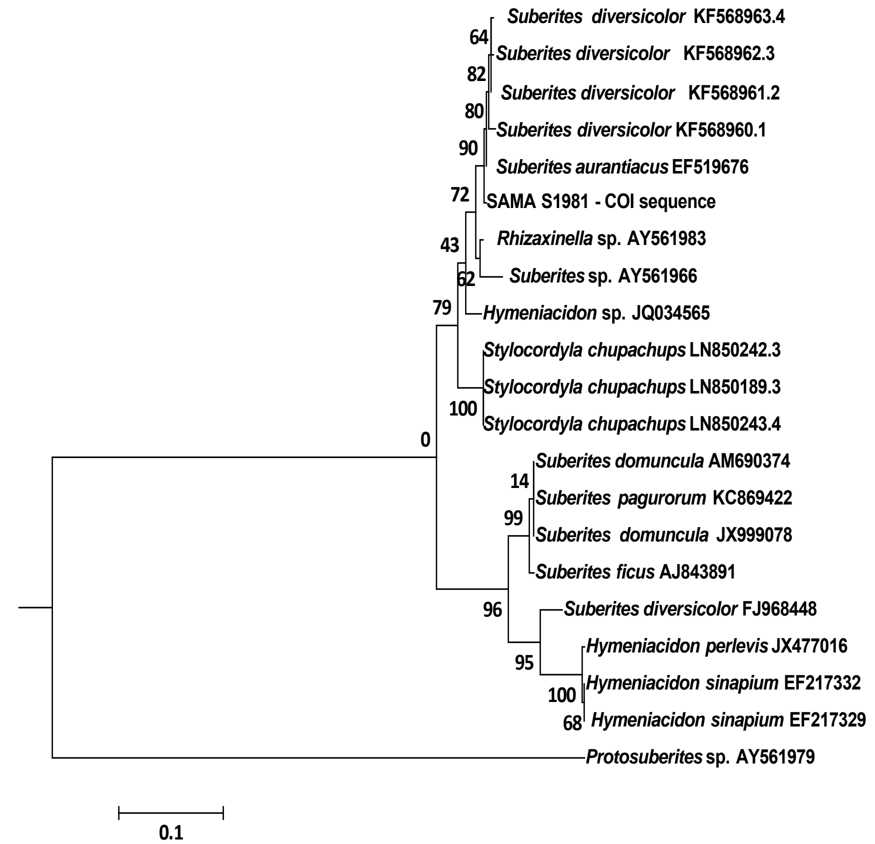
Devosia	Halioglobus	Halioglobus	Desulfovibrio	Devosia	Hoeflea	Geodermatophilus	Fischerella
Formivibrio	Halochromatium	Halochromatium	Dissulfuribacter	Frigoribacterium	Iamia	Granulosicoccus	Gaiella
Frigoribacterium	Henriciella	Henriciella	Ectothiorhodospira	Halochromatium	Jannaschia	Halioglobus	Gemmatimonas
Frondihabitans	Hoeflea	Hoeflea	Eggerthella	Halorhodospira	Kordia	Halochromatium	Geobacter
Geodermatophilus	Iamia	Iamia	Endozoicomonas	Hippea	Laceyella	Henriciella	Geodermatophilaceae
Granulicella	Jannaschia	Jannaschia	Filomicrobium	Hydrogenivirga	Lactobacillus	Hoeflea	Geothermobacter
Halorhodospira	Kordia	Kordia	Fischerella	Iamia	Litorilinea	Iamia	Geothermobacterium
Hippea	Laceyella	Laceyella	Flavobacterium	Lacibacterium	Nitrospira	Jannaschia	Hahella
Hydrogenivirga	Lacibacterium	Lacibacterium	Gaetbulibacter	Lactobacillus	Octadecabacter	Laceyella	Holophaga
Hydrogenobacter	Lactobacillus	Lactobacillus	Gaiella	Magnetovibrio	Persephonella	Lacibacterium	Iamia
Hydrogenobaculum	Litorilinea	Litorilinea	Gangjinia	Maritalea	Planifilum	Lactobacillus	Kordiimonas
Iamia	Nitrosopumilus	Nitrosopumilus	Gemmatimonas	Olsenella	Prochlorococcus	Litorilinea	Kribbella
Lactobacillus	Nitrospira	Nitrospira	Geobacter	Pelagicola	Rhodopirellula	Nitrosopumilus	Legionella
Limimonas	Octadecabacter	Octadecabacter	Geodermatophilaceae	Persephonella	Rhodothermus	Nitrospira	Levilinea
Magnetovibrio	Olleya	Olleya	Geothermobacter	Propionibacterium	Rhodovulum	Octadecabacter	Longilinea
Maricaulis	Persephonella	Persephonella	Geothermobacterium	Pseudothermotoga	Roseibium	Phaeocystidibacter	Marinobacter
Maritalea	Phaeocystidibacter	Phaeocystidibacter	Gilvibacter	Psychroserpens	Rudaea	Phyllobacterium	Megasphaera
mitochondria-1	Phyllobacterium	Phyllobacterium	Hahella	Rhodopirellula	Sagittula	Planifilum	Methylogaea
Niastella	Planifilum	Planifilum	Halochromatium	Rhodothermus	Sneathiella	Prochlorococcus	Microbulbifer
Nitratiruptor	Prochlorococcus	Prochlorococcus	Hippea	Roseiflexus	Solirubrobacter	Pseudofulvibacter	Neorickettsia
Olsenella	Pseudofulvibacter	Pseudofulvibacter	Holophaga	Roseospirillum	Sphaerobacter	Rhodopirellula	Nitrosopumilus
Pelagicola	Pseudomonas	Pseudomonas	Iamia	Sphaerobacter	Spirochaeta	Rhodothermus	Nitrospira
Pseudothermotoga	Rhodopirellula	Rhodopirellula	Joostella	Spirochaeta	Stanieria	Roseibacillus	Ochrobactrum
Psychroserpens	Rhodothermus	Rhodothermus	Kordiimonas	Streptomyces	Tepidibacter	Roseibium	Ornatilinea
Rhodopirellula	Rhodovulum	Rhodovulum	Kribbella	Syntrophorhabdus	Thermaerobacter	Rudaea	Patulibacter
Rhodothermus	Roseibacillus	Roseibium	Legionella	Terracoccus	Thermobaculum	Sagittula	Pectinatus
Roseibacillus	Roseibium	Rudaea	Levilinea	Thermobaculum	Thermodesulforhabdus	Sneathiella	Pelobacter
Roseiflexus	Rudaea	Sagittula	Longilinea	Thermodesulfovibrio	Thermovenabulum	Solirubrobacter	Propionicimonas
Roseomonas	Sagittula	Sneathiella	Magnetospira	Thiobios	Thioalbus	Spirochaeta	Pseudomonas
Roseospirillum	Sneathiella	Solirubrobacter	Marinobacter	Thiococcus	Thioalkalivibrio	Stanieria	Rhodobacter
Ruegeria	Solirubrobacter	Sphaerobacter	Megasphaera	Trichormus	Thioprofundum	Tepidibacter	Rhodopirellula
Sphaerobacter	Sphaerobacter	Spirochaeta	Meridianimaribacter	Vibrio		Thermaerobacter	Rhodovulum
Spirochaeta	Spirochaeta	Stanieria	Mesorhizobium			Thermodesulforhabdus	Rubrivirga
Streptomyces	Stanieria	Tepidibacter	Methylogaea			Thermovenabulum	Ruegeria
Synechococcus	Tepidibacter	Thermaerobacter	Microbulbifer			Thioalbus	Skermanella
Syntrophorhabdus	Thermaerobacter	Thermobaculum	Neorickettsia			Thioalkalivibrio	Sneathiella
Terracoccus	Thermobaculum	Thermodesulforhabdus	Nitrosopumilus			Thioprofundum	Solirubrobacter
Thermodesulfobium	Thermodesulforhabdus	Thermovenabulum	Nitrospina				Spiribacter
Thermodesulfovibrio	Thermovenabulum	Thioalbus	Ochrobactrum				Spirochaeta
Thermoflexus	Thioalbus	Thioalkalivibrio	Ornatilinea				Steroidobacter
Thiobios	Thioalkalivibrio	Thioprofundum	Owenweeksia				Synechococcus
Thiococcus	Thioprofundum		Patulibacter				Syntrophomonas
Thiohalocapsa			Pectinatus				Syntrophus
Trichormus			Pelobacter				Thalassobaculum
			Planctomyces				Thermoanaerobaculum
			Planktotalea				Thermoleophilum
			Planomicrobium				Thermolithobacter
			Propionibacterium				Thioalbus

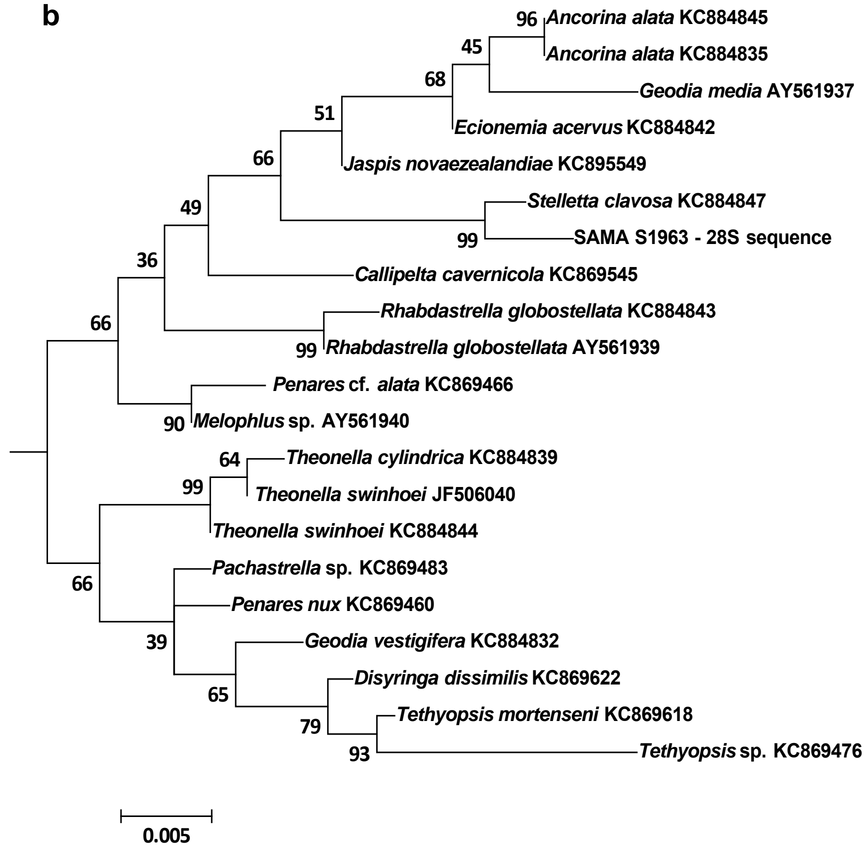
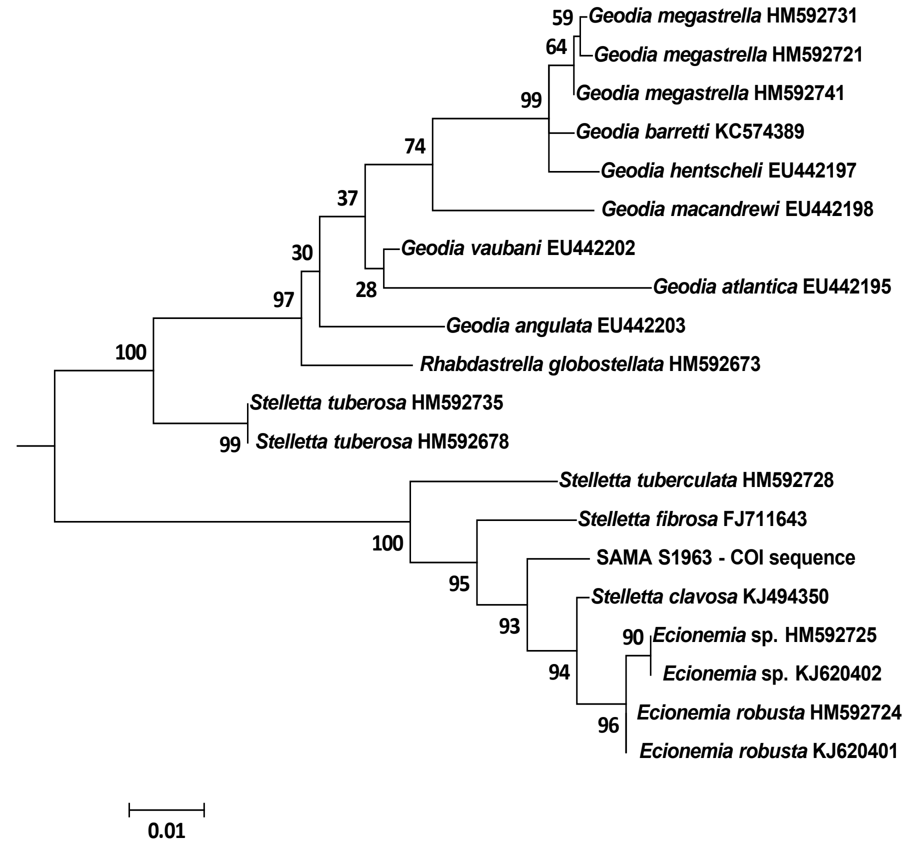
	Propionicimonas	*Thioalkalivibrio*
	Pseudahrensia	*Thiohalocapsa*
	Rhizobium	*Thiohalophilus*
	Rhodobacter	
	Rhodopirellula	
	Rhodovulum	
	Rubrivirga	
	Ruegeria	
	Skermanella	
	Solirubrobacter	
	Spiribacter	
	Spirochaeta	
	Steroidobacter	
	Synechococcus	
	Syntrophomonas	
	Syntrophus	
	Thalassobaculum	
	Thalassobius	
	Thermasporomyces	
	Thermoanaerobaculum	
	Thermodesulfobium	
	Thermoleophilum	
	Thermolithobacter	
	Thioalkalivibrio	
	Thiohalocapsa	
	Thiohalophilus	
	Wenyingzhuangia	

Appendix Table 4 - 2 The shared genera between the four sponge species revealed by different primer sets

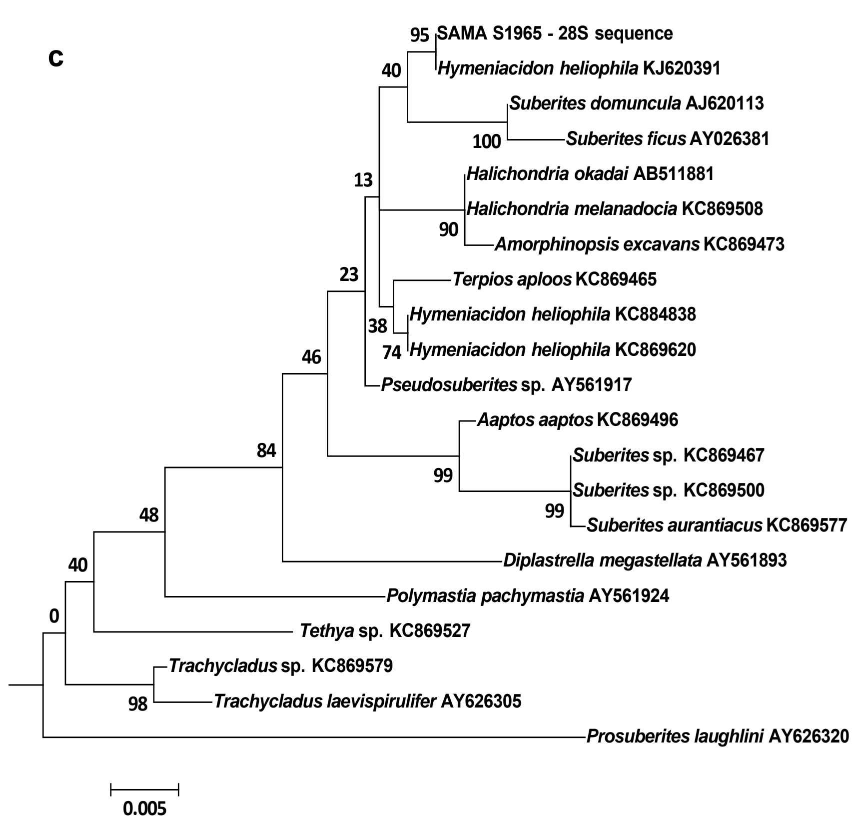
Region	V1V3	V4	V4V5	V5V8
Shared genera	Aquihabitans	*Aciditerrimonas*	*Aciditerrimonas*	*Aciditerrimonas*
	Arthrobacter	*Alkalispirillum*	*Aquibacter*	*Anaerolinea*
	Caenispirillum	*Aquihabitans*	*Aquihabitans*	*Blastopirellula*
	Candidatus Methylomirabilis	*Blastopirellula*	*Blastopirellula*	*Conexibacter*
	Candidatus Puniceispirillum	*Conexibacter*	*Conexibacter*	*Dehalococcoides*
	Cryobacterium	*Dehalobacter*	*Dehalobacter*	*Desulfacinum*
	Cyanobium	*Ectothiorhodospira*	*Desulfomonile*	*Gaiella*
	Desulfomonile	*Endozoicomonas*	*Ectothiorhodospira*	*Geobacter*
	Devosia	*Ferrimicrobium*	*Endozoicomonas*	*Geothermobacterium*
	Hippea	*Henriciella*	*Ferrimicrobium*	*Kordiimonas*
	Hydrogenivirga	*Hoeflea*	*Formivibrio*	*Megasphaera*
	Iamia	*Jannaschia*	*Granulosicoccus*	*Methylogaea*
	Lactobacillus	*Lactobacillus*	*Halioglobus*	*Nitrosopumilus*
	Magnetovibrio	*Litorilinea*	*Halochromatium*	*Nitrospira*
	Pelagicola	*Nitrospira*	*Henriciella*	*Ornatilinea*
	Pseudothermotoga	*Octadecabacter*	*Iamia*	*Pectinatus*
	Rhodopirellula	*Persephonella*	*Jannaschia*	*Pelobacter*
	Rhodothermus	*Prochlorococcus*	*Laceyella*	*Rhodobacter*
	Roseospirillum	*Rhodopirellula*	*Litorilinea*	*Rhodovulum*
	Sphaerobacter	*Rhodothermus*	*Nitrospira*	*Solirubrobacter*
	Spirochaeta	*Rudaea*	*Octadecabacter*	*Spiribacter*
	Streptomyces	*Solirubrobacter*	*Phaeocystidibacter*	*Spirochaeta*
	Syntrophorhabdus	*Spirochaeta*	*Planifilum*	*Synechococcus*
	Thiobios	*Thermaerobacter*	*Prochlorococcus*	*Thermoleophilum*
	Thiococcus	*Thermovenabulum*	*Rhodopirellula*	*Thioalkalivibrio*
		Thioalbus	*Rudaea*	*Thiohalophilus*
		Thioalkalivibrio	*Sneathiella*	*Iamia*
			Solirubrobacter	
			Spirochaeta	
			Stanieria	
			Tepidibacter	
			Thermaerobacter	
			Thermodesulforhabdus	
			Thermovenabulum	
			Thioalkalivibrio	
			Thioprofundum	
Sum	25	27	36	27

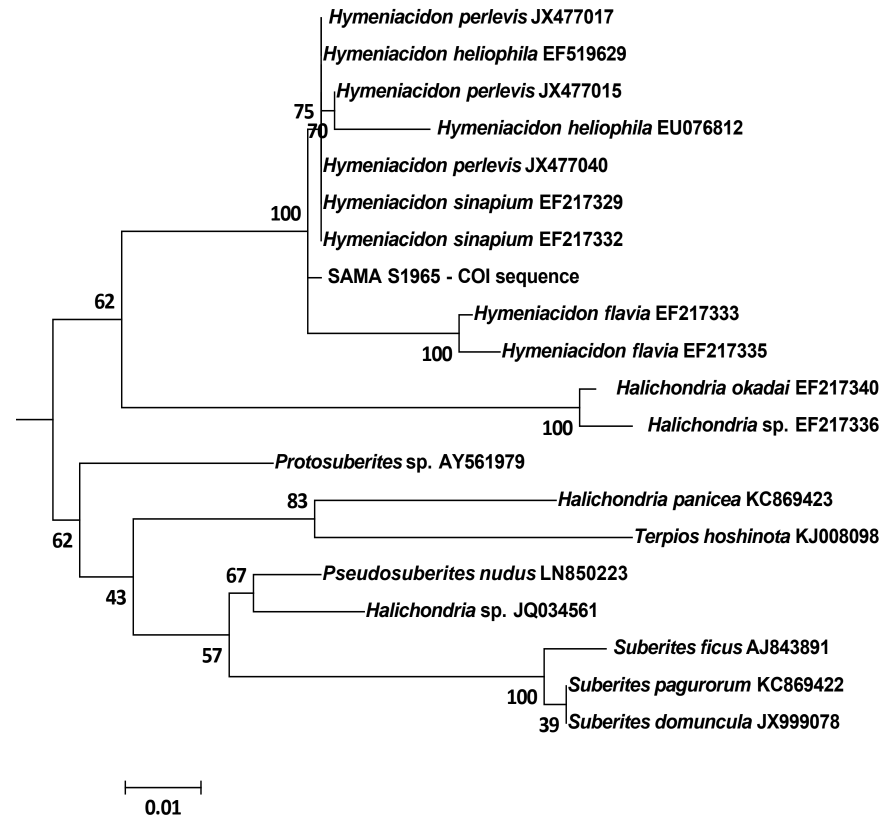

a
SAMA S1981 - 28S sequence
Suberites aurantiacus KC869577
Suberites sp. KC869500
Suberites sp. KC869467
Aaptos aaptos KC869496
Rhizaxinella sp. AY561910
Hymeniacidon heliophila KJ620391
Hymeniacidon heliophila KC884838
Hymeniacidon heliophila KC869620
Terpios aploos KC869465
Pseudosuberites sp. AY561917
Halichondria okadai AB511881
Halichondria melanadocia KC869508
Amorphinopsis excavans KC869473
Suberites domuncula AJ620113
Suberites ficus AY026381
Diplastrella megastellata AY561893
Timea sp. AY626303
Trachycladus laevispirulifer AY626305
Trachycladus sp. KC869579
Trachycladus stylifer KC869453
0.005

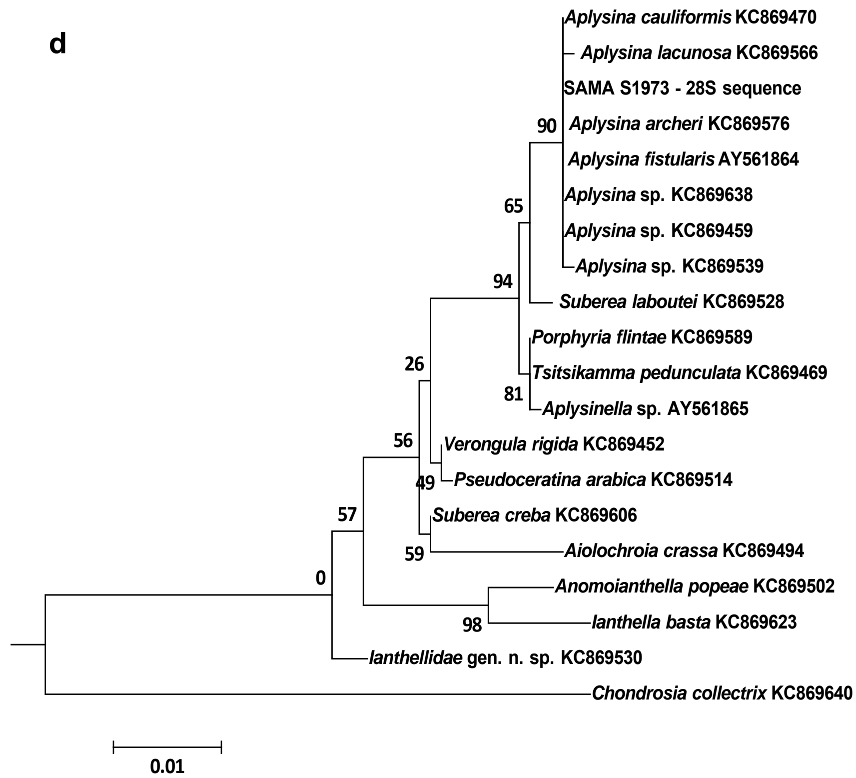

Suberites diversicolor KF568963.4
Suberites diversicolor KF568962.3
Suberites diversicolor KF568961.2
Suberites diversicolor KF568960.1
Suberites aurantiacus EF519676
SAMA S1981 - COI sequence
Rhizaxinella sp. AY561983
Suberites sp. AY561966
Hymeniacidon sp. JQ034565
Stylocordyla chupachups LN850242.3
Stylocordyla chupachups LN850189.3
Stylocordyla chupachups LN850243.4
Suberites domuncula AM690374
Suberites pagurorum KC869422
Suberites domuncula JX999078
Suberites ficus AJ843891
Suberites diversicolor FJ968448
Hymeniacidon perlevis JX477016
Hymeniacidon sinapium EF217332
Hymeniacidon sinapium EF217329
Protosuberites sp. AY561979
0.1

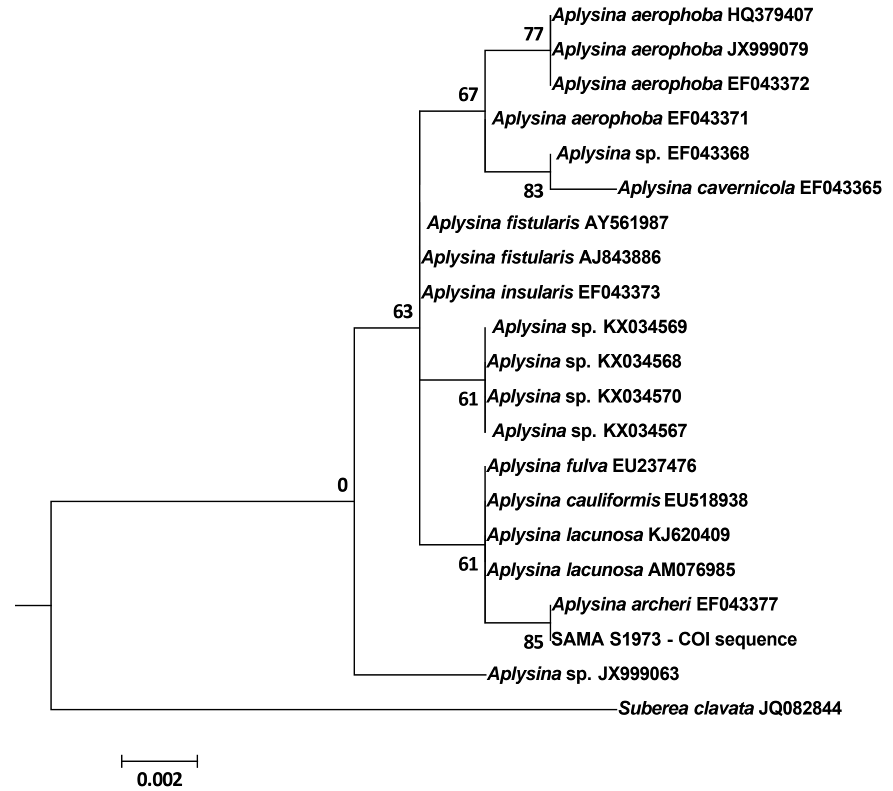

b
96 Ancorina alata KC884845
Ancorina alata KC884835
45
68
Geodia media AY561937
51
Ecionemia acervus KC884842
66
Jaspis novaezealandiae KC895549
49
Stelletta clavosa KC884847
99
SAMA S1963 - 28S sequence
36
Callipelta cavernicola KC869545
66
Rhabdastrella globostellata KC884843
99 Rhabdastrella globostellata AY561939
Penares cf. alata KC869466
90 Melophlus sp. AY561940
64 Theonella cylindrica KC884839
99 Theonella swinhoei JF506040
Theonella swinhoei KC884844
Pachastrella sp. KC869483
66
Penares nux KC869460
39
Geodia vestigifera KC884832
65
Disyringa dissimilis KC869622
79
Tethyopsis mortenseni KC869618
93
Tethyopsis sp. KC869476
0.005

59 Geodia megastrella HM592731
64 Geodia megastrella HM592721
99 Geodia megastrella HM592741
Geodia barretti KC574389
74
Geodia hentscheli EU442197
Geodia macandrewi EU442198
37
Geodia vaubani EU442202
30
28
Geodia atlantica EU442195
97
Geodia angulata EU442203
100
Rhabdastrella globostellata HM592673
Stelletta tuberosa HM592735
99 Stelletta tuberosa HM592678
Stelletta tuberculata HM592728
100
Stelletta fibrosa FJ711643
95
SAMA S1963 - COI sequence
93
Stelletta clavosa KJ494350
90 Ecionemia sp. HM592725
94
Ecionemia sp. KJ620402
96
Ecionemia robusta HM592724
Ecionemia robusta KJ620401
0.01

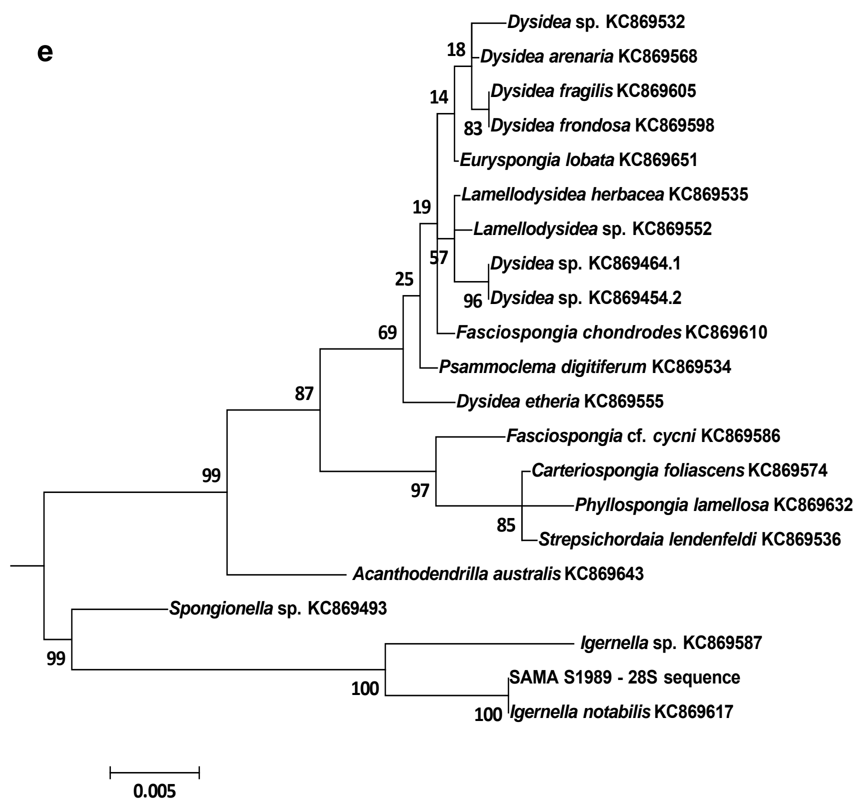
c

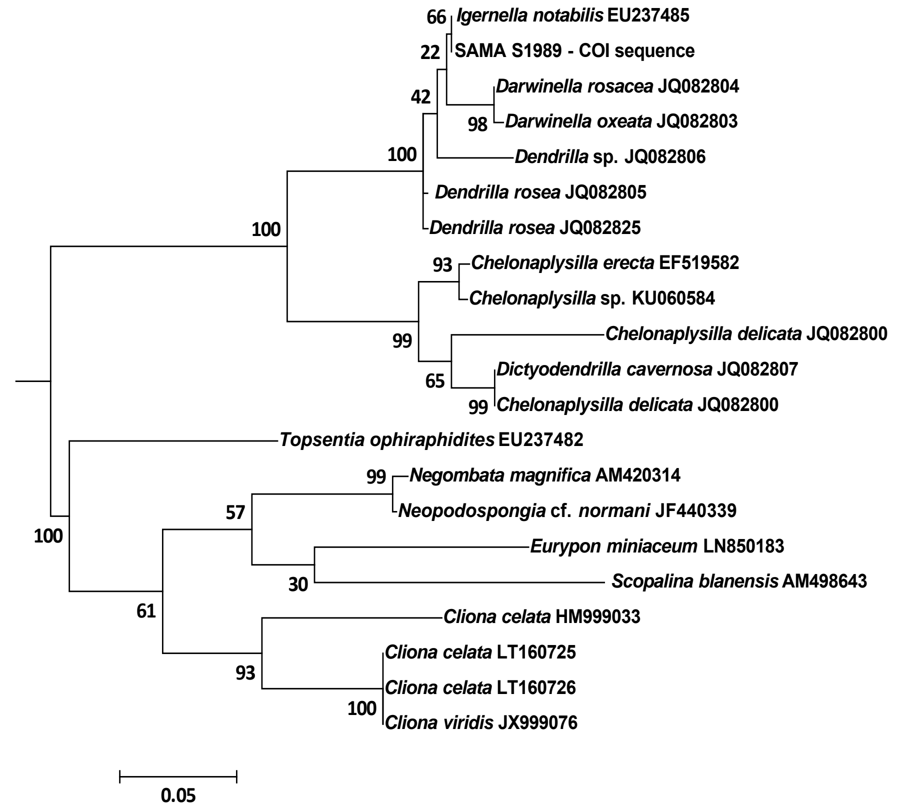


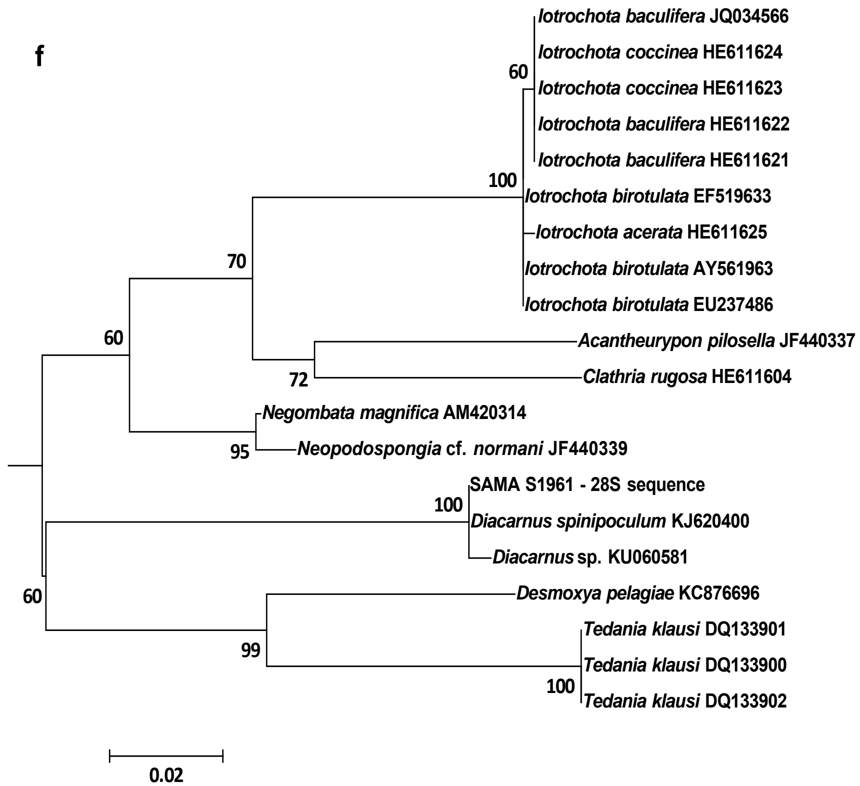


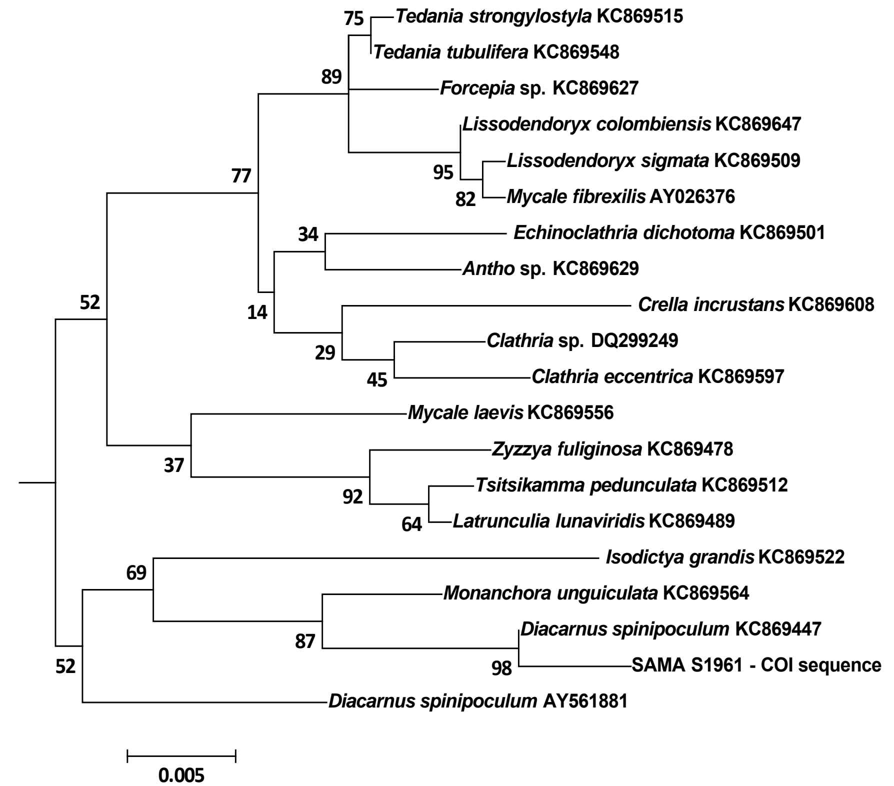

d
Aplysina cauliformis KC869470
Aplysina lacunosa KC869566
SAMA S1973 - 28S sequence
Aplysina archeri KC869576
Aplysina fistularis AY561864
Aplysina sp. KC869638
Aplysina sp. KC869459
Aplysina sp. KC869539
Suberea laboutei KC869528
Porphyria flintae KC869589
Tsitsikamma pedunculata KC869469
Aplysinella sp. AY561865
Verongula rigida KC869452
Pseudoceratina arabica KC869514
Suberea creba KC869606
Aiolochroia crassa KC869494
Anomoianthella popeae KC869502
Ianthella basta KC869623
Ianthellidae gen. n. sp. KC869530
Chondrosia collectrix KC869640
90
65
94
26
81
56
49
57
59
0
98
0.01

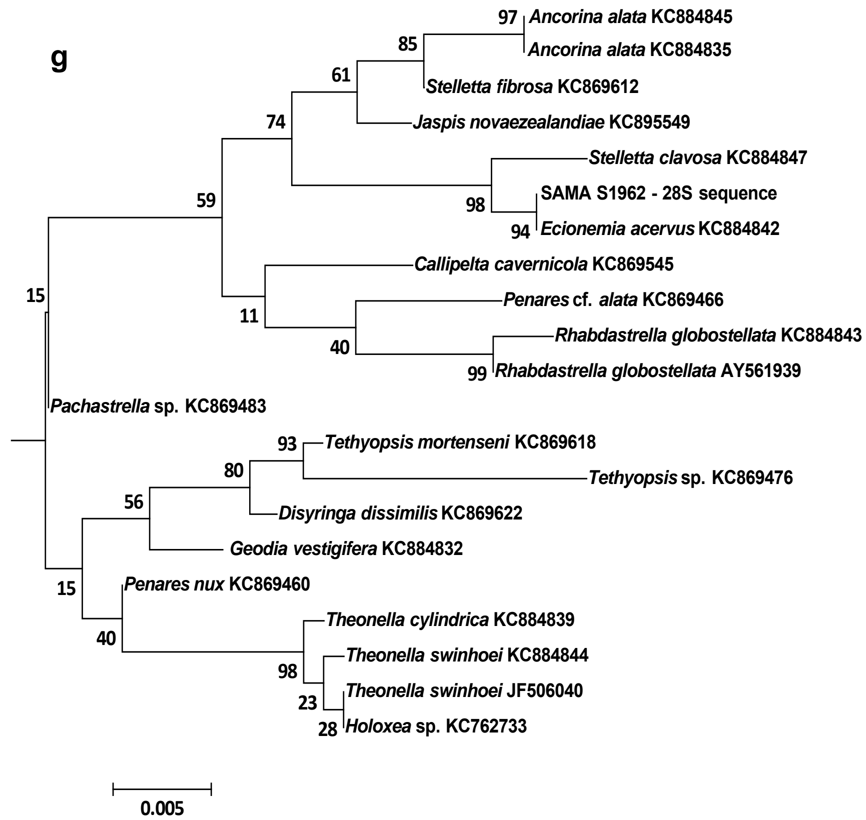
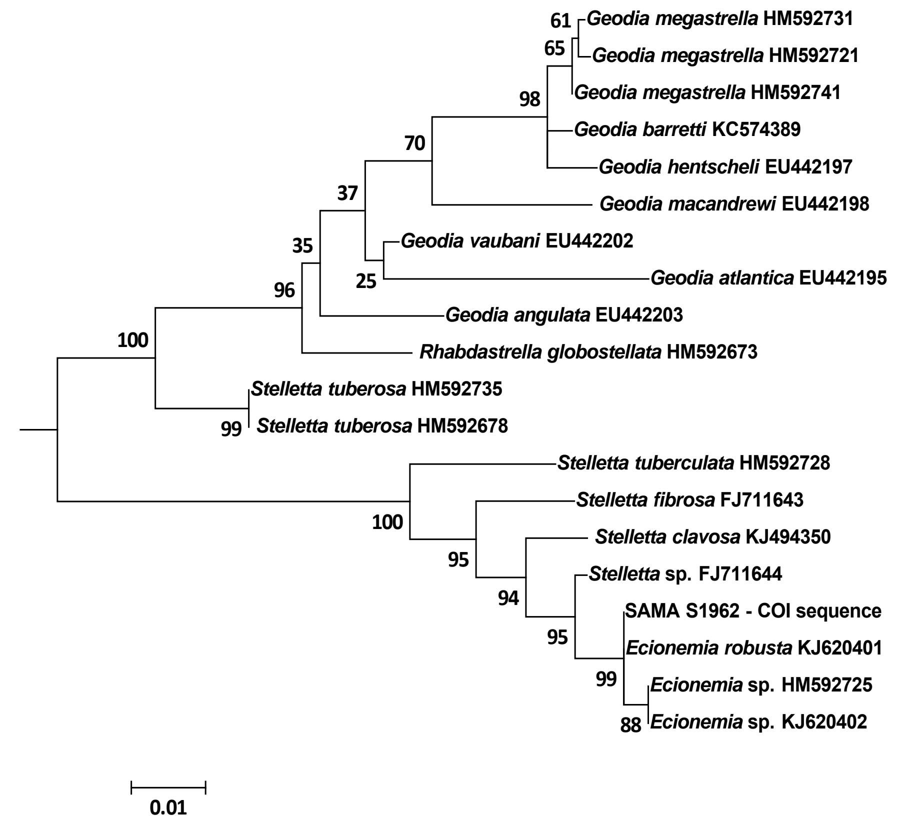

Aplysina aerophoba HQ379407
Aplysina aerophoba JX999079
Aplysina aerophoba EF043372
Aplysina aerophoba EF043371
Aplysina sp. EF043368
Aplysina cavernicola EF043365
Aplysina fistularis AY561987
Aplysina fistularis AJ843886
Aplysina insularis EF043373
Aplysina sp. KX034569
Aplysina sp. KX034568
Aplysina sp. KX034570
Aplysina sp. KX034567
Aplysina fulva EU237476
Aplysina cauliformis EU518938
Aplysina lacunosa KJ620409
Aplysina lacunosa AM076985
Aplysina archeri EF043377
SAMA S1973 - COI sequence
Aplysina sp. JX999063
Suberea clavata JQ082844
77
67
83
63
61
0
61
85
0.002

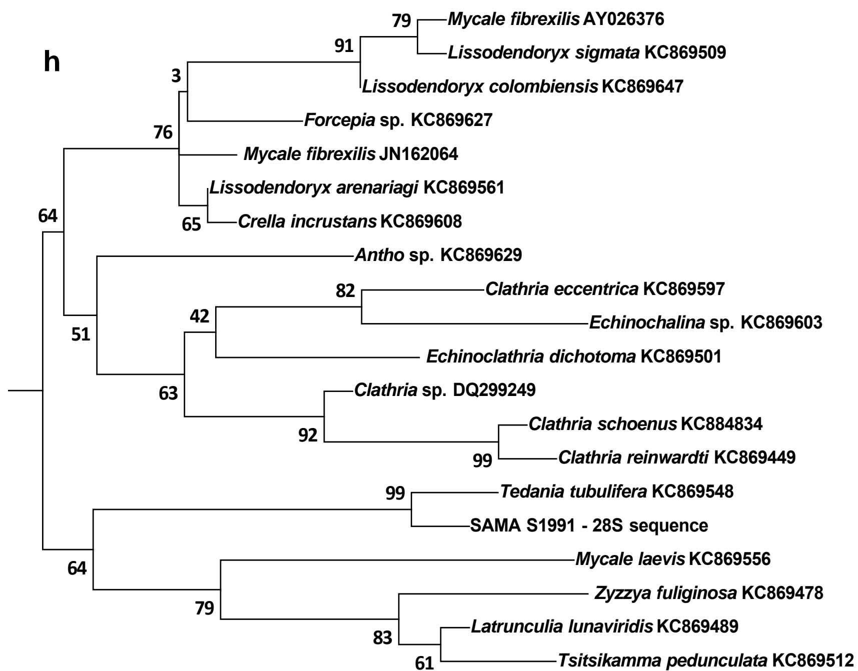

e
Dysidea sp. KC869532
Dysidea arenaria KC869568
Dysidea fragilis KC869605
Dysidea frondosa KC869598
Euryspongia lobata KC869651
Lamellodysidea herbacea KC869535
Lamellodysidea sp. KC869552
Dysidea sp. KC869464.1
Dysidea sp. KC869454.2
Fasciospongia chondrodes KC869610
Psammoclema digitiferum KC869534
Dysidea etheria KC869555
Fasciospongia cf. cycni KC869586
Carteriospongia foliascens KC869574
Phyllospongia lamellosa KC869632
Strepsichordaia lendenfeldi KC869536
Acanthodendrilla australis KC869643
Spongionella sp. KC869493
Igernella sp. KC869587
SAMA S1989 - 28S sequence
Igernella notabilis KC869617
18
14
83
19
57
96
25
69
87
97
85
99
99
100
100
0.005

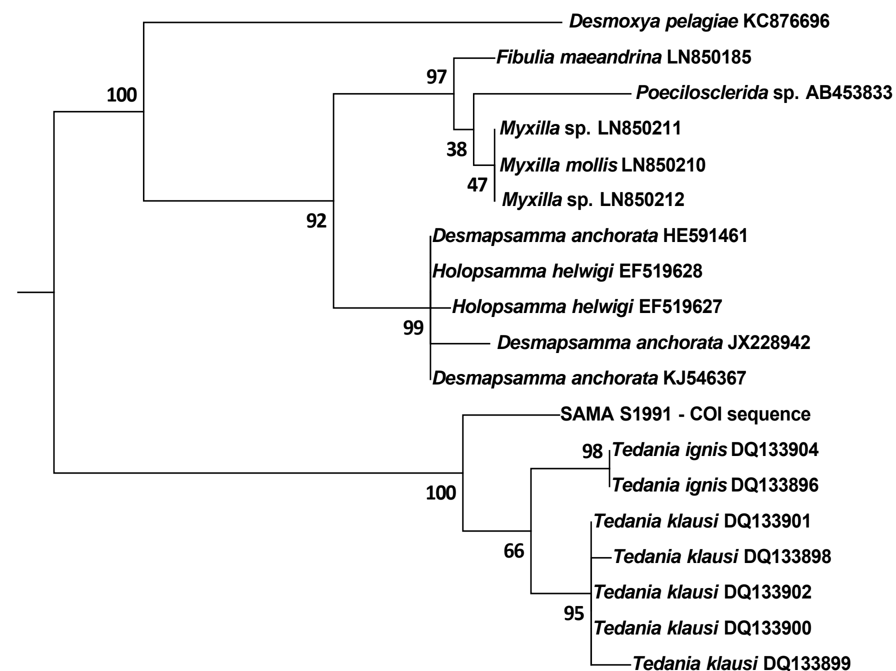

Igernella notabilis EU237485
SAMA S1989 - COI sequence
Darwinella rosacea JQ082804
Darwinella oxeata JQ082803
Dendrilla sp. JQ082806
Dendrilla rosea JQ082805
Dendrilla rosea JQ082825
Chelonaplysilla erecta EF519582
Chelonaplysilla sp. KU060584
Chelonaplysilla delicata JQ082800
Dictyodendrilla cavernosa JQ082807
Chelonaplysilla delicata JQ082800
Topsentia ophiraphidites EU237482
Negombata magnifica AM420314
Neopodospongia cf. normani JF440339
Eurypon miniaceum LN850183
Scopalina blanensis AM498643
Cliona celata HM999033
Cliona celata LT160725
Cliona celata LT160726
Cliona viridis JX999076
66
22
42
98
100
100
93
99
65
99
99
57
100
30
61
93
100
0.05

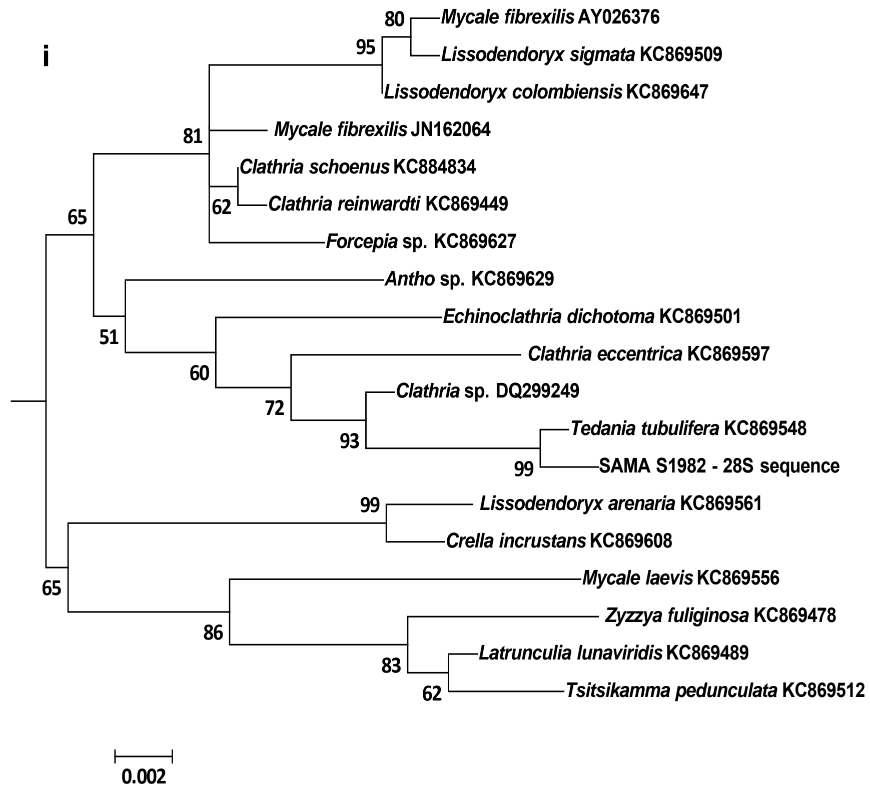
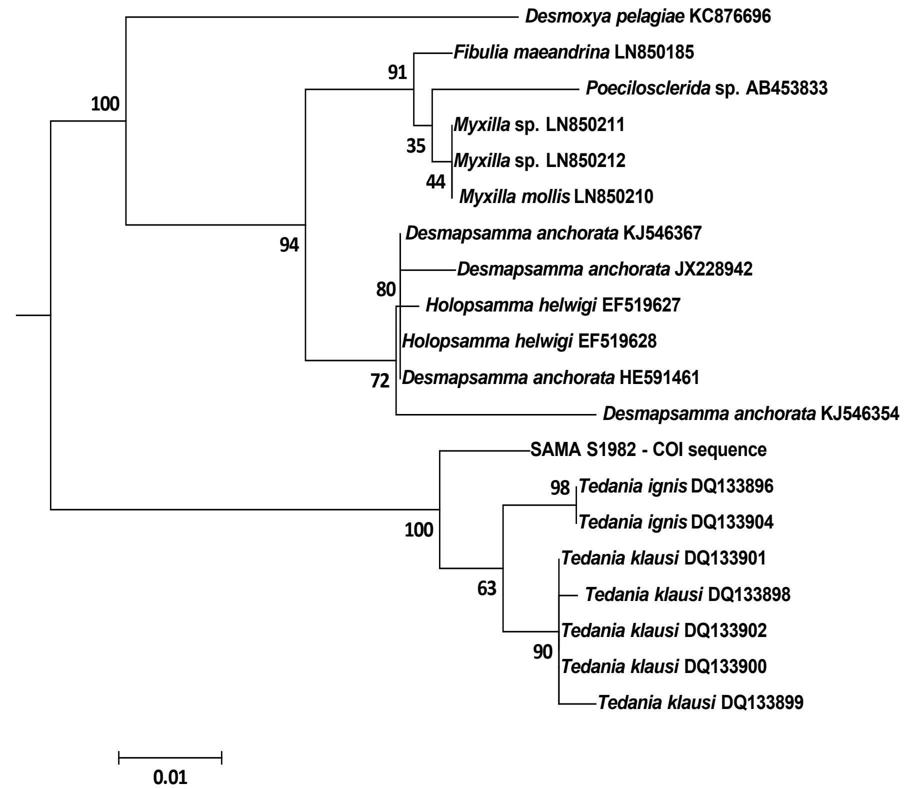

f
Iotrochota baculifera JQ034566
Iotrochota coccinea HE611624
Iotrochota coccinea HE611623
Iotrochota baculifera HE611622
Iotrochota baculifera HE611621
Iotrochota birotulata EF519633
Iotrochota acerata HE611625
Iotrochota birotulata AY561963
Iotrochota birotulata EU237486
Acantheurypon pilosella JF440337
Clathria rugosa HE611604
Negombata magnifica AM420314
Neopodospongia cf. normani JF440339
SAMA S1961 - 28S sequence
Diacarnus spinipoculum KJ620400
Diacarnus sp. KU060581
Desmoxya pelagiae KC876696
Tedania klausi DQ133901
Tedania klausi DQ133900
Tedania klausi DQ133902
0.02

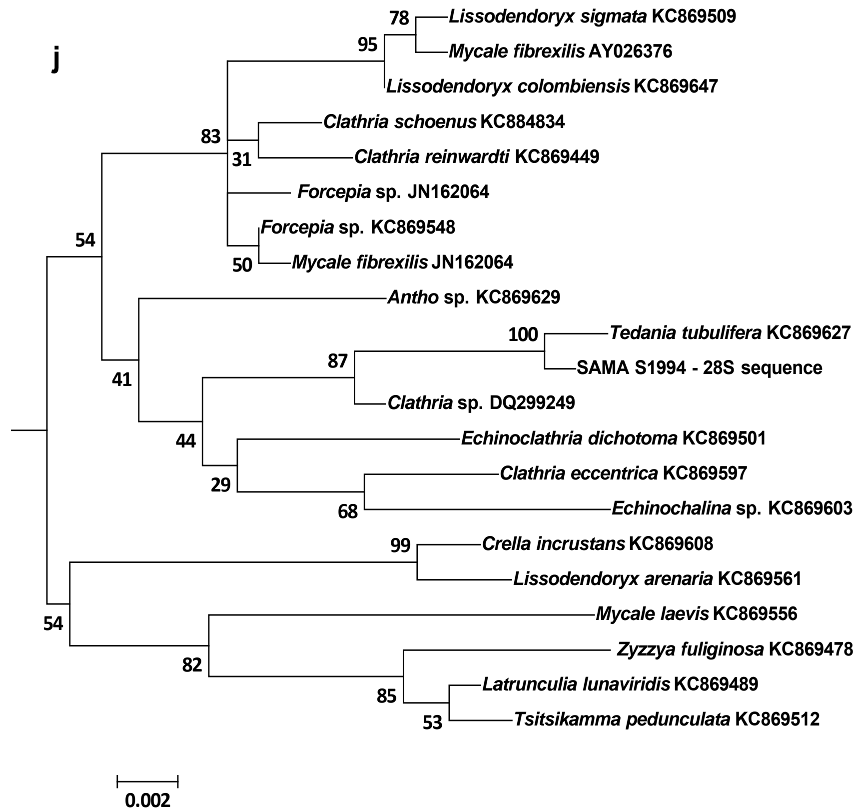

Tedania strongylostyla KC869515
Tedania tubulifera KC869548
Forcepia sp. KC869627
Lissodendoryx colombiensis KC869647
Lissodendoryx sigmata KC869509
Mycale fibrexilis AY026376
Echinoclathria dichotoma KC869501
Antho sp. KC869629
Crella incrustans KC869608
Clathria sp. DQ299249
Clathria eccentrica KC869597
Mycale laevis KC869556
Zyzzya fuliginosa KC869478
Tsitsikamma pedunculata KC869512
Latrunculia lunaviridis KC869489
Isodictya grandis KC869522
Monanchora unguiculata KC869564
Diacarnus spinipoculum KC869447
SAMA S1961 - COI sequence
Diacarnus spinipoculum AY561881
0.005

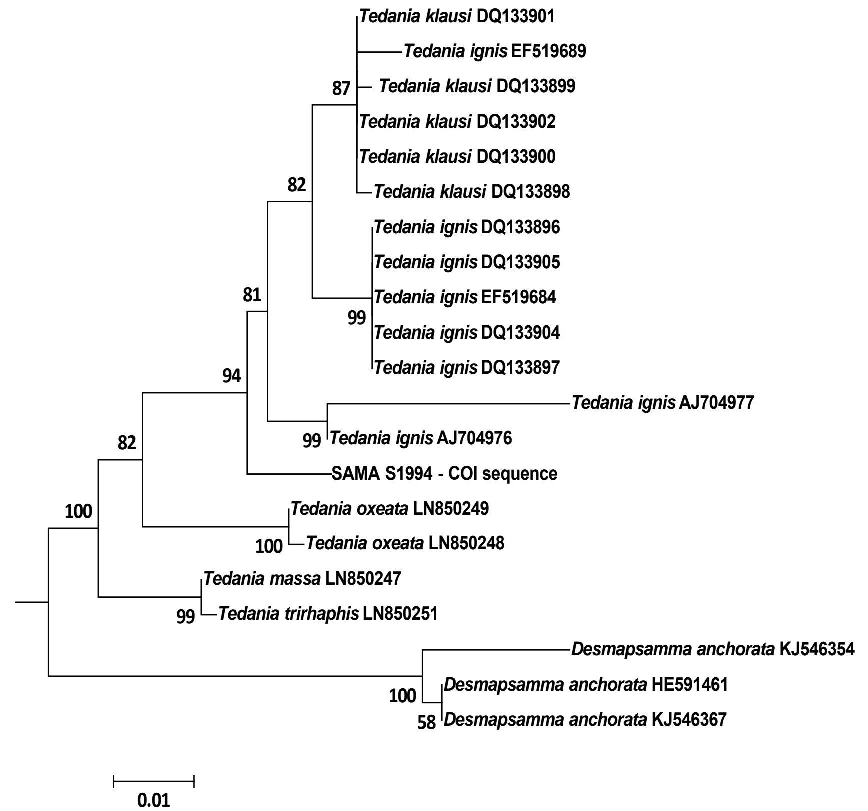

g
Ancorina alata KC884845
Ancorina alata KC884835
Stelletta fibrosa KC869612
Jaspis novaezealandiae KC895549
Stelletta clavosa KC884847
SAMA S1962 - 28S sequence
Ecionemia acervus KC884842
Callipelta cavernicola KC869545
Penares cf. alata KC869466
Rhabdastrella globostellata KC884843
Rhabdastrella globostellata AY561939
Pachastrella sp. KC869483
Tethyopsis mortenseni KC869618
Tethyopsis sp. KC869476
Disyringa dissimilis KC869622
Geodia vestigifera KC884832
Penares nux KC869460
Theonella cylindrica KC884839
Theonella swinhoei KC884844
Theonella swinhoei JF506040
Holoxea sp. KC762733
0.005

Geodia megastrella HM592731
Geodia megastrella HM592721
Geodia megastrella HM592741
Geodia barretti KC574389
Geodia hentscheli EU442197
Geodia macandrewi EU442198
Geodia vaubani EU442202
Geodia atlantica EU442195
Geodia angulata EU442203
Rhabdastrella globostellata HM592673
Stelletta tuberosa HM592735
Stelletta tuberosa HM592678
Stelletta tuberculata HM592728
Stelletta fibrosa FJ711643
Stelletta clavosa KJ494350
Stelletta sp. FJ711644
SAMA S1962 - COI sequence
Ecionemia robusta KJ620401
Ecionemia sp. HM592725
Ecionemia sp. KJ620402
0.01

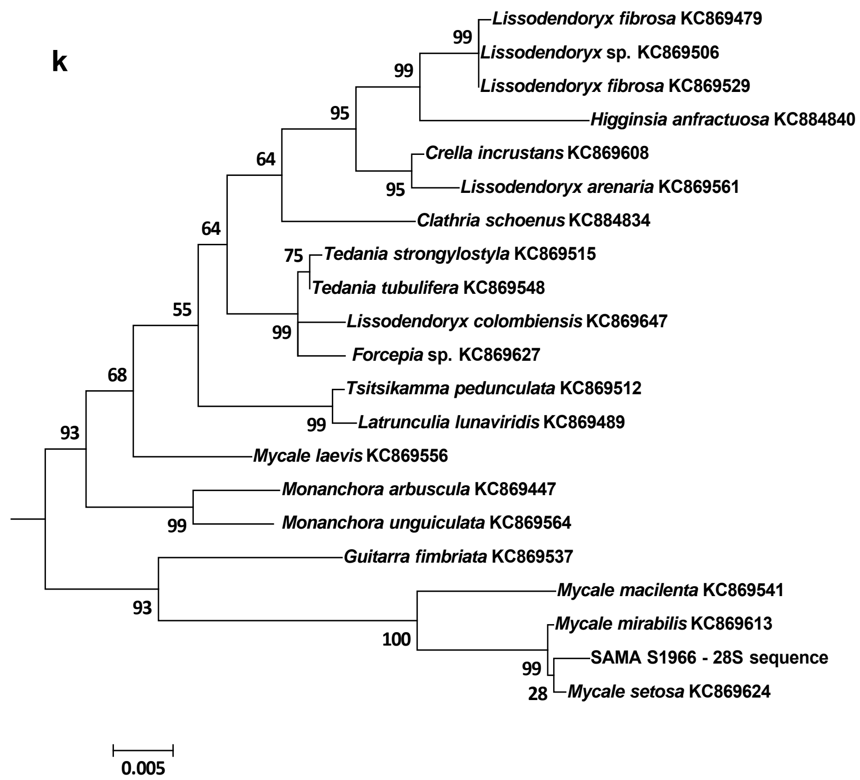
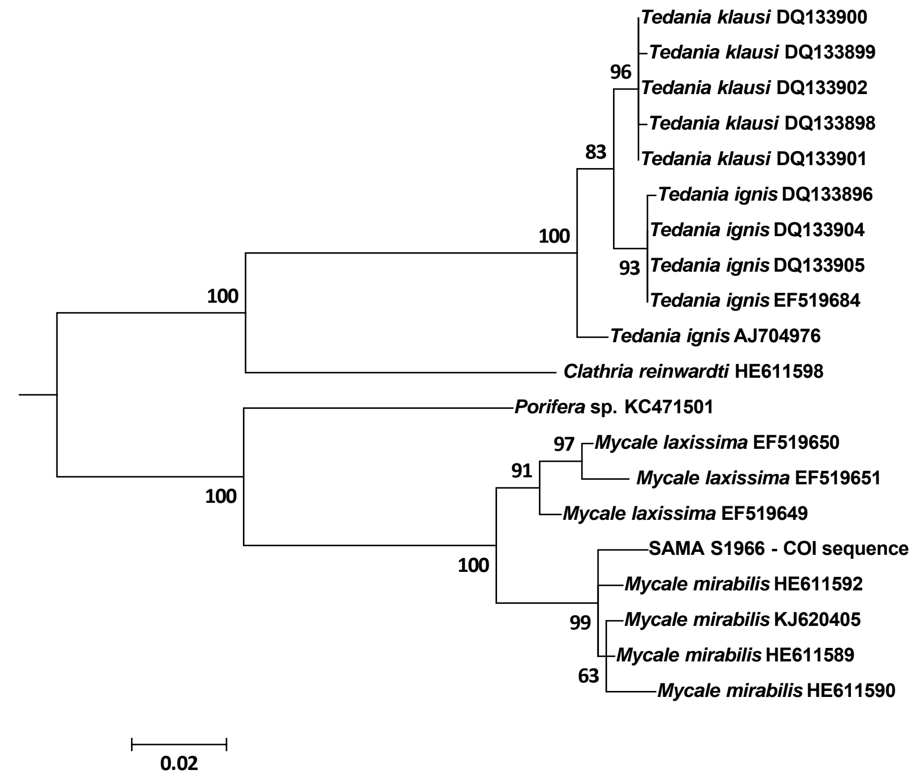

h
Mycale fibrexilis AY026376
Lissodendoryx sigmata KC869509
Lissodendoryx colombiensis KC869647
Forcepia sp. KC869627
Mycale fibrexilis JN162064
Lissodendoryx arenariagi KC869561
Crella incrustans KC869608
Antho sp. KC869629
Clathria eccentrica KC869597
Echinochalina sp. KC869603
Echinoclathria dichotoma KC869501
Clathria sp. DQ299249
Clathria schoenus KC884834
Clathria reinwardti KC869449
Tedania tubulifera KC869548
SAMA S1991 - 28S sequence
Mycale laevis KC869556
Zyzzya fuliginosa KC869478
Latrunculia lunaviridis KC869489
Tsitsikamma pedunculata KC869512
0.002


Desmoxya pelagiae KC876696
Fibulia maeandrina LN850185
Poecilosclerida sp. AB453833
Myxilla sp. LN850211
Myxilla mollis LN850210
Myxilla sp. LN850212
Desmapsamma anchorata HE591461
Holopsamma helwigi EF519628
Holopsamma helwigi EF519627
Desmapsamma anchorata JX228942
Desmapsamma anchorata KJ546367
SAMA S1991 - COI sequence
Tedania ignis DQ133904
Tedania ignis DQ133896
Tedania klausi DQ133901
Tedania klausi DQ133898
Tedania klausi DQ133902
Tedania klausi DQ133900
Tedania klausi DQ133899
0.01

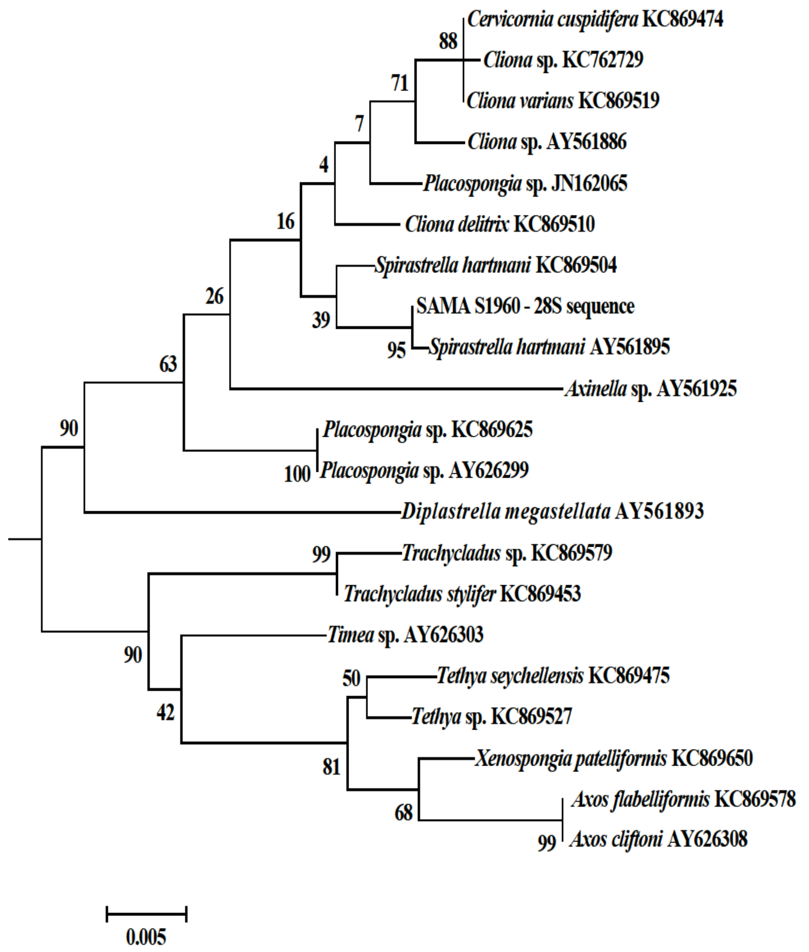

i
Mycale fibrexilis AY026376
Lissodendoryx sigmata KC869509
Lissodendoryx colombiensis KC869647
Mycale fibrexilis JN162064
Clathria schoenus KC884834
Clathria reinwardti KC869449
Forcepia sp. KC869627
Antho sp. KC869629
Echinoclathria dichotoma KC869501
Clathria eccentrica KC869597
Clathria sp. DQ299249
Tedania tubulifera KC869548
SAMA S1982 - 28S sequence
Lissodendoryx arenaria KC869561
Crella incrustans KC869608
Mycale laevis KC869556
Zyzzya fuliginosa KC869478
Latrunculia lunaviridis KC869489
Tsitsikamma pedunculata KC869512
80
95
81
62
65
51
60
72
93
99
99
65
86
83
62
0.002

Desmoxya pelagiae KC876696
Fibulia maeandrina LN850185
Poecilosclerida sp. AB453833
Myxilla sp. LN850211
Myxilla sp. LN850212
Myxilla mollis LN850210
Desmapsamma anchorata KJ546367
Desmapsamma anchorata JX228942
Holopsamma helwigi EF519627
Holopsamma helwigi EF519628
Desmapsamma anchorata HE591461
Desmapsamma anchorata KJ546354
SAMA S1982 - COI sequence
Tedania ignis DQ133896
Tedania ignis DQ133904
Tedania klausi DQ133901
Tedania klausi DQ133898
Tedania klausi DQ133902
Tedania klausi DQ133900
Tedania klausi DQ133899
100
91
35
44
94
80
72
98
100
63
90
0.01

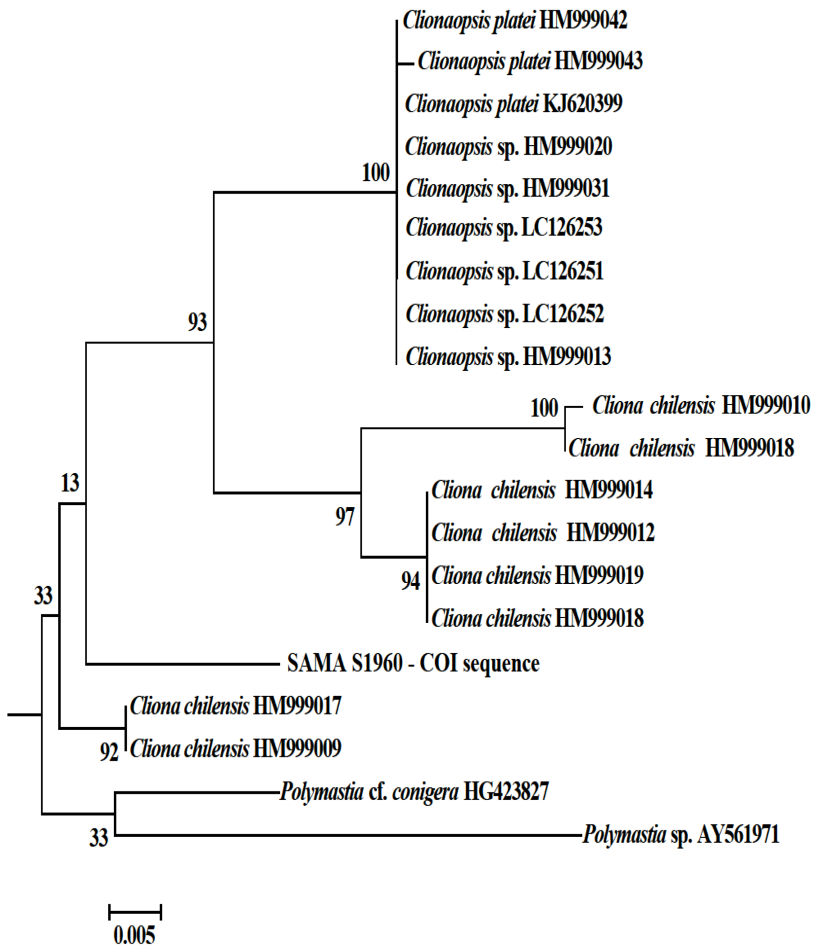

j
Lissodendoryx sigmata KC869509
Mycale fibrexilis AY026376
Lissodendoryx colombiensis KC869647
Clathria schoenus KC884834
Clathria reinwardti KC869449
Forcepia sp. JN162064
Forcepia sp. KC869548
Mycale fibrexilis JN162064
Antho sp. KC869629
Tedania tubulifera KC869627
SAMA S1994 - 28S sequence
Clathria sp. DQ299249
Echinoclathria dichotoma KC869501
Clathria eccentrica KC869597
Echinochalina sp. KC869603
Crella incrustans KC869608
Lissodendoryx arenaria KC869561
Mycale laevis KC869556
Zyzzya fuliginosa KC869478
Latrunculia lunaviridis KC869489
Tsitsikamma pedunculata KC869512
0.002


Tedania klausi DQ133901
Tedania ignis EF519689
Tedania klausi DQ133899
Tedania klausi DQ133902
Tedania klausi DQ133900
Tedania klausi DQ133898
Tedania ignis DQ133896
Tedania ignis DQ133905
Tedania ignis EF519684
Tedania ignis DQ133904
Tedania ignis DQ133897
Tedania ignis AJ704977
Tedania ignis AJ704976
SAMA S1994 - COI sequence
Tedania oxeata LN850249
Tedania oxeata LN850248
Tedania massa LN850247
Tedania trirhaphis LN850251
Desmapsamma anchorata KJ546354
Desmapsamma anchorata HE591461
Desmapsamma anchorata KJ546367
0.01

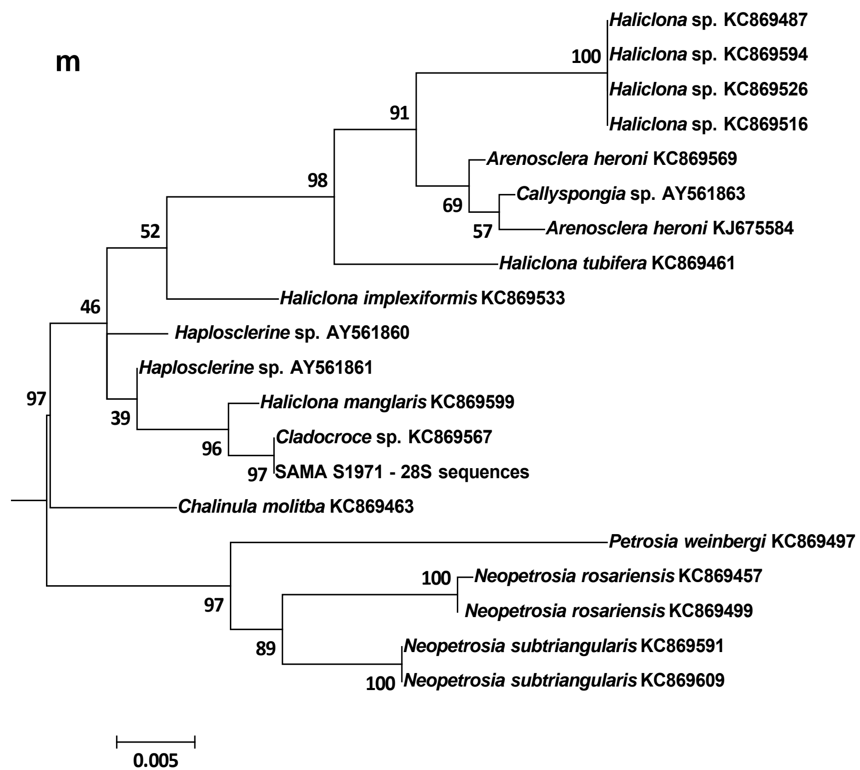

k
Lissodendoryx fibrosa KC869479
Lissodendoryx sp. KC869506
Lissodendoryx fibrosa KC869529
Higginsia anfractuosa KC884840
Crella incrustans KC869608
Lissodendoryx arenaria KC869561
Clathria schoenus KC884834
Tedania strongylostyla KC869515
Tedania tubulifera KC869548
Lissodendoryx colombiensis KC869647
Forcepia sp. KC869627
Tsitsikamma pedunculata KC869512
Latrunculia lunaviridis KC869489
Mycale laevis KC869556
Monanchora arbuscula KC869447
Monanchora unguiculata KC869564
Guitarra fimbriata KC869537
Mycale macilenta KC869541
Mycale mirabilis KC869613
SAMA S1966 - 28S sequence
Mycale setosa KC869624
0.005

Tedania klausi DQ133900
Tedania klausi DQ133899
Tedania klausi DQ133902
Tedania klausi DQ133898
Tedania klausi DQ133901
Tedania ignis DQ133896
Tedania ignis DQ133904
Tedania ignis DQ133905
Tedania ignis EF519684
Tedania ignis AJ704976
Clathria reinwardti HE611598
Porifera sp. KC471501
Mycale laxissima EF519650
Mycale laxissima EF519651
Mycale laxissima EF519649
SAMA S1966 - COI sequence
Mycale mirabilis HE611592
Mycale mirabilis KJ620405
Mycale mirabilis HE611589
Mycale mirabilis HE611590
0.02

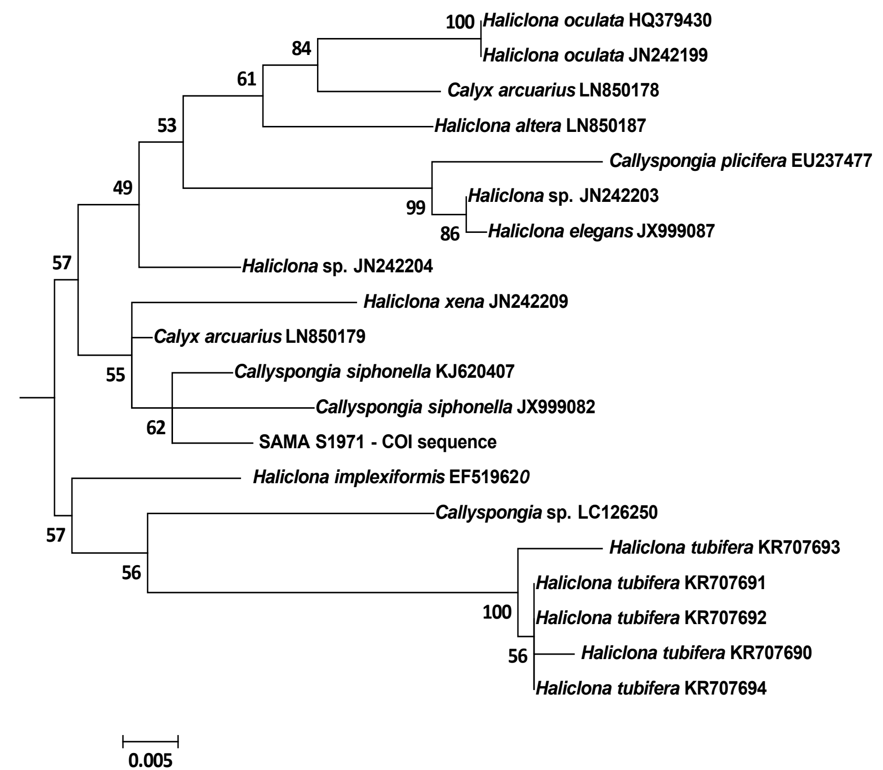
I

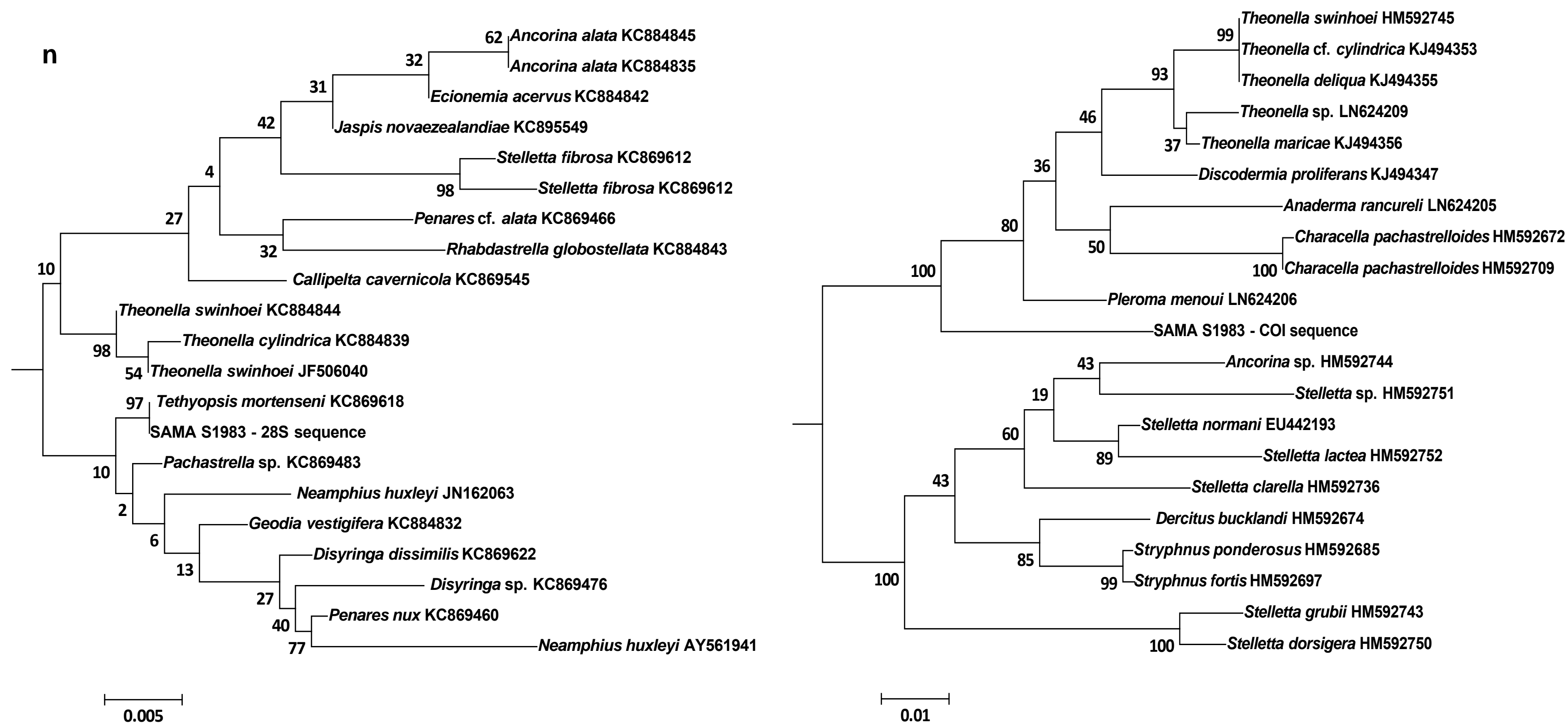




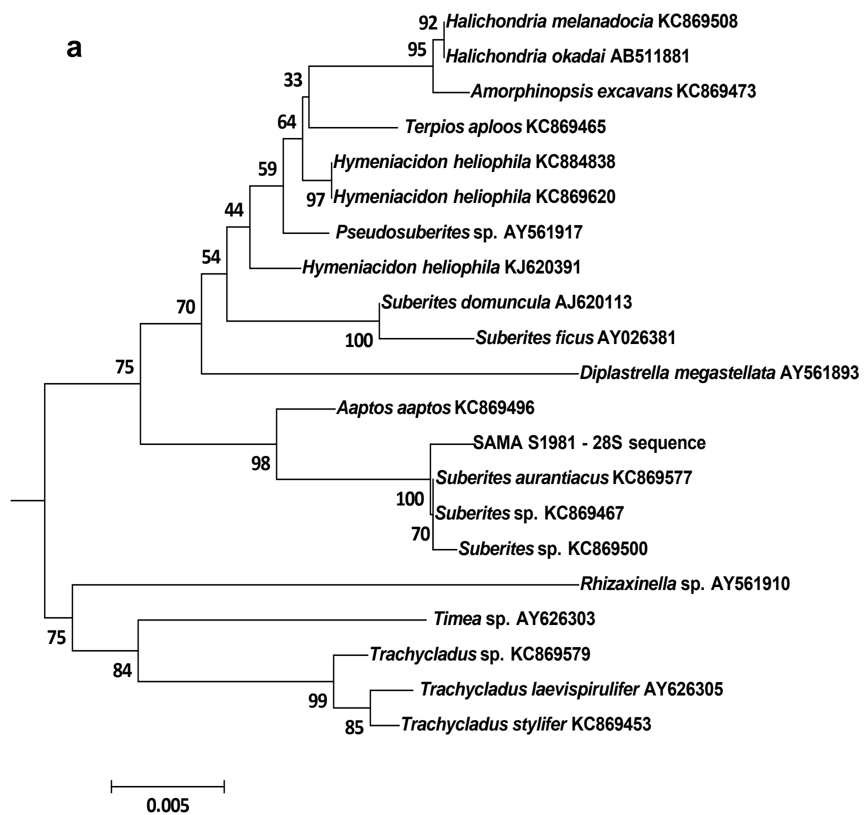
m

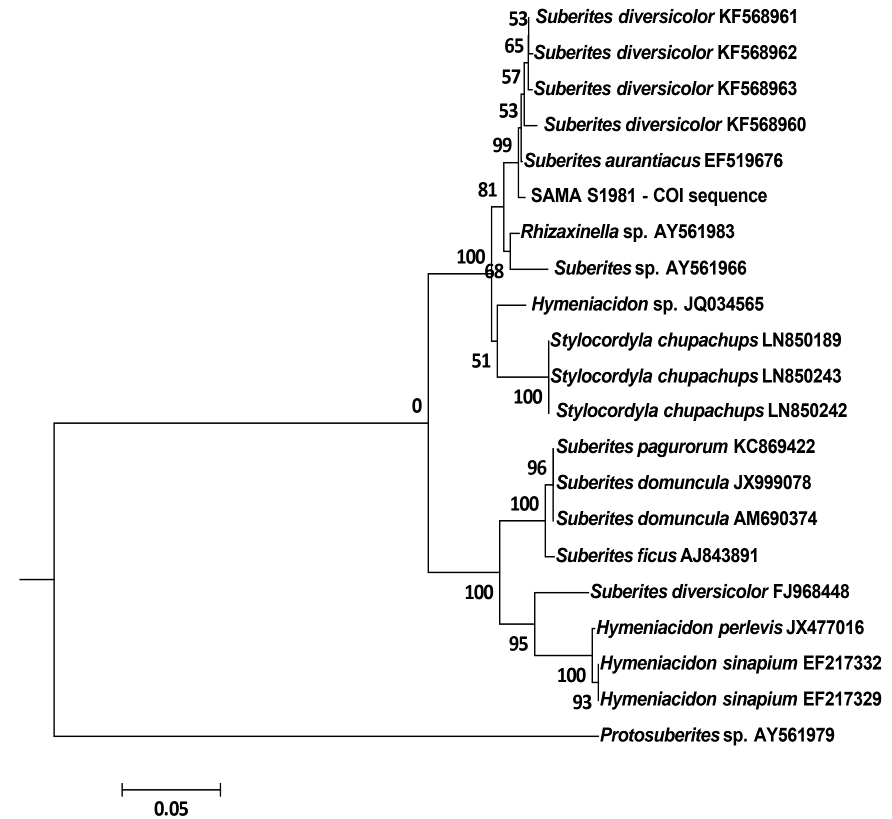


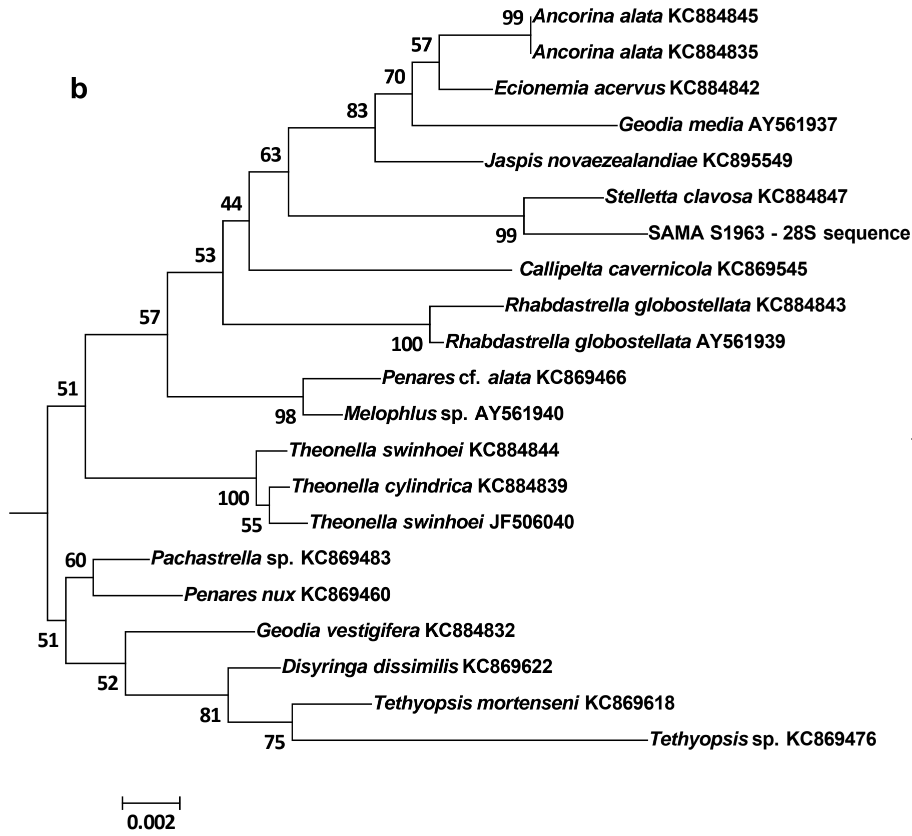




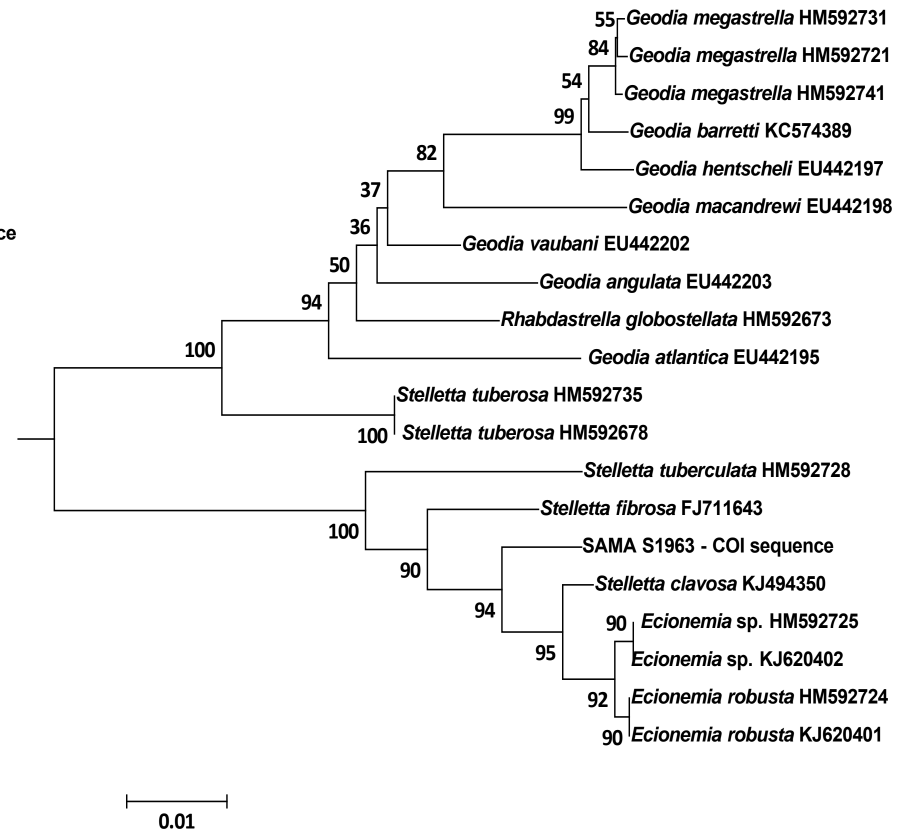
Appendix Figure 2 - 1 Phylogenetic relationship of fourteen sponges using Maximum Likelihood method based on 28S rRNA gene and COI mtDNA. a. Phylogenetic relationship of sponge SAMA S1981. b. Phylogenetic relationship of sponge SAMA S1963. c. Phylogenetic relationship of sponge SAMA S1965. d. Phylogenetic relationship of sponge SAMA S1973. e. Phylogenetic relationship of sponge SAMA S1989. f. Phylogenetic relationship of sponge SAMA S1961. g. Phylogenetic relationship of sponge SAMA S1962. h. Phylogenetic relationship of sponge SAMA S1991. i. Phylogenetic relationship of sponge SAMA S1982. j. Phylogenetic relationship of sponge SAMA S1994. k. Phylogenetic relationship of sponge SAMA S1966. l. Phylogenetic relationship of sponge SAMA S1960. m. Phylogenetic relationship of sponge SAMA S1971. n. Phylogenetic relationship of sponge SAMA S1983.


a
Halichondria melanadocia KC869508
Halichondria okadai AB511881
Amorphinopsis excavans KC869473
Terpios aploos KC869465
Hymeniacidon heliophila KC884838
Hymeniacidon heliophila KC869620
Pseudosuberites sp. AY561917
Hymeniacidon heliophila KJ620391
Suberites domuncula AJ620113
Suberites ficus AY026381
Diplastrella megastellata AY561893
Aaptos aaptos KC869496
SAMA S1981 - 28S sequence
Suberites aurantiacus KC869577
Suberites sp. KC869467
Suberites sp. KC869500
Rhizaxinella sp. AY561910
Timea sp. AY626303
Trachycladus sp. KC869579
Trachycladus laevispirulifer AY626305
Trachycladus stylifer KC869453
0.005

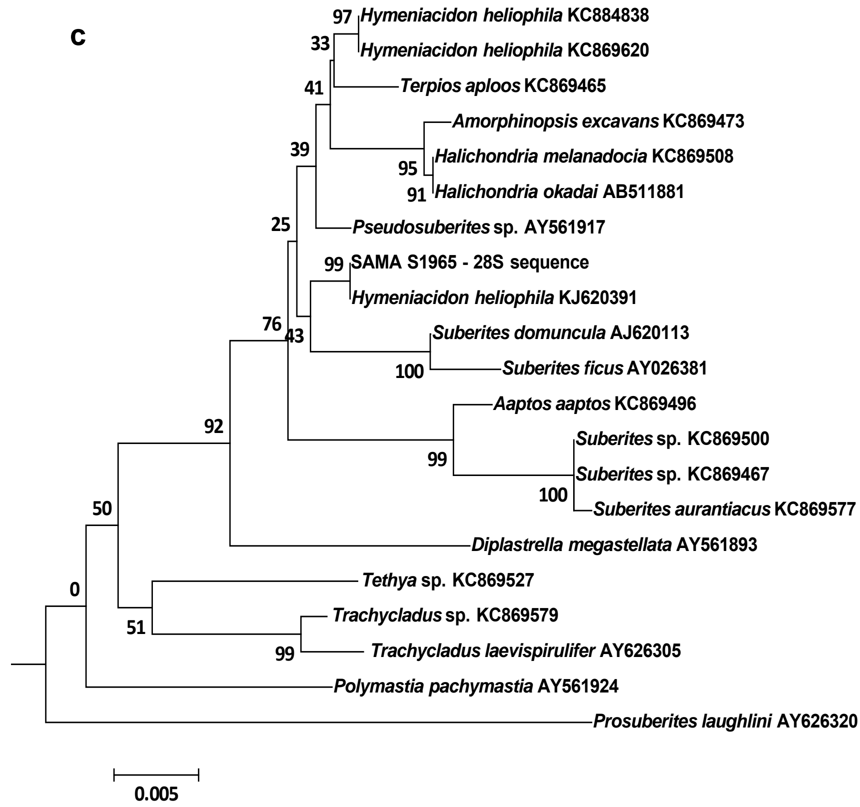

Suberites diversicolor KF568961
Suberites diversicolor KF568962
Suberites diversicolor KF568963
Suberites diversicolor KF568960
Suberites aurantiacus EF519676
SAMA S1981 - COI sequence
Rhizaxinella sp. AY561983
Suberites sp. AY561966
Hymeniacidon sp. JQ034565
Stylocordyla chupachups LN850189
Stylocordyla chupachups LN850243
Stylocordyla chupachups LN850242
Suberites pagurorum KC869422
Suberites domuncula JX999078
Suberites domuncula AM690374
Suberites ficus AJ843891
Suberites diversicolor FJ968448
Hymeniacidon perlevis JX477016
Hymeniacidon sinapium EF217332
Hymeniacidon sinapium EF217329
Protosuberites sp. AY561979
0.05

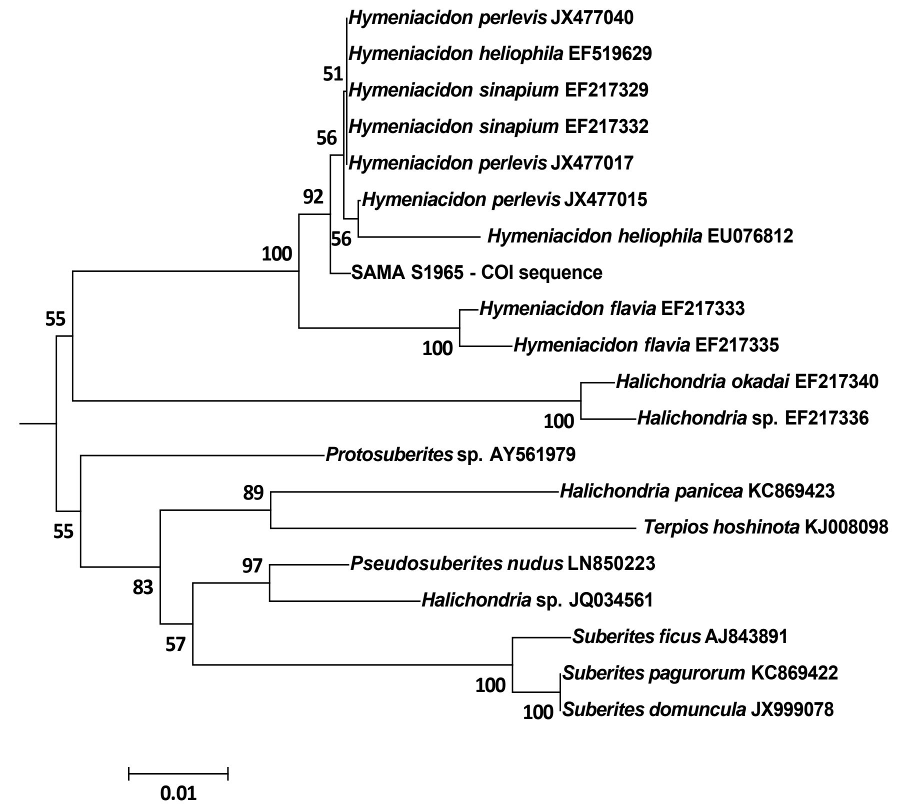


b
99 Ancorina alata KC884845
Ancorina alata KC884835
57
70
Ecionemia acervus KC884842
83
Geodia media AY561937
63
Jaspis novaezealandiae KC895549
Stelletta clavosa KC884847
44
99
SAMA S1963 - 28S sequence
53
Callipelta cavernicola KC869545
Rhabdastrella globostellata KC884843
57
100
Rhabdastrella globostellata AY561939
Penares cf. alata KC869466
51
98
Melophlus sp. AY561940
Theonella swinhoei KC884844
100
Theonella cylindrica KC884839
55
Theonella swinhoei JF506040
60
Pachastrella sp. KC869483
Penares nux KC869460
51
Geodia vestigifera KC884832
52
Disyringa dissimilis KC869622
81
Tethyopsis mortenseni KC869618
75
Tethyopsis sp. KC869476
0.002

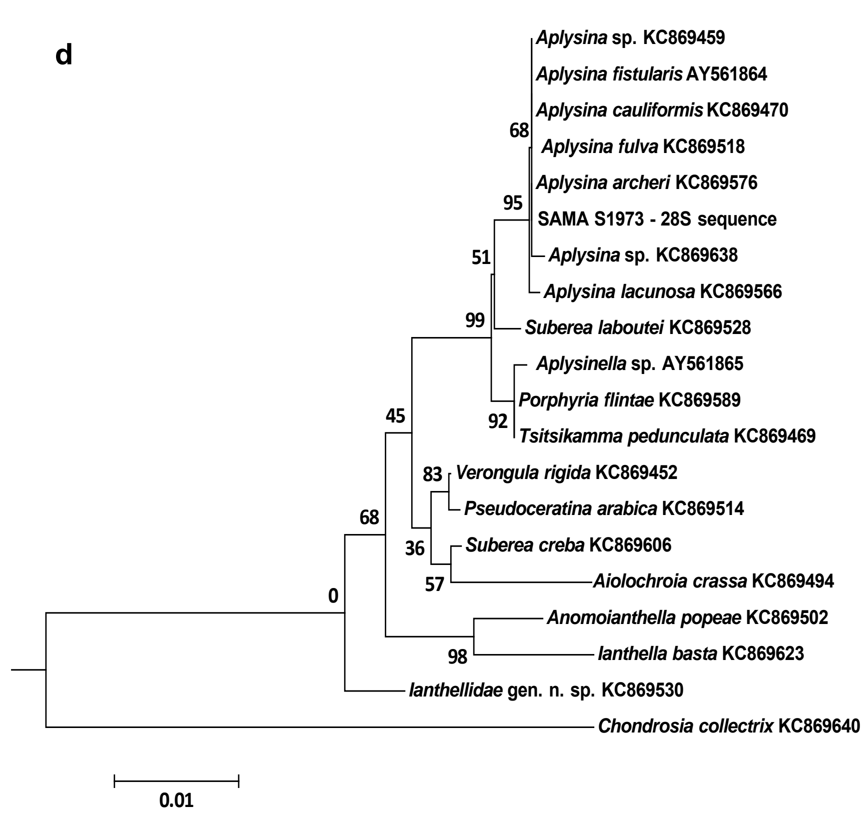


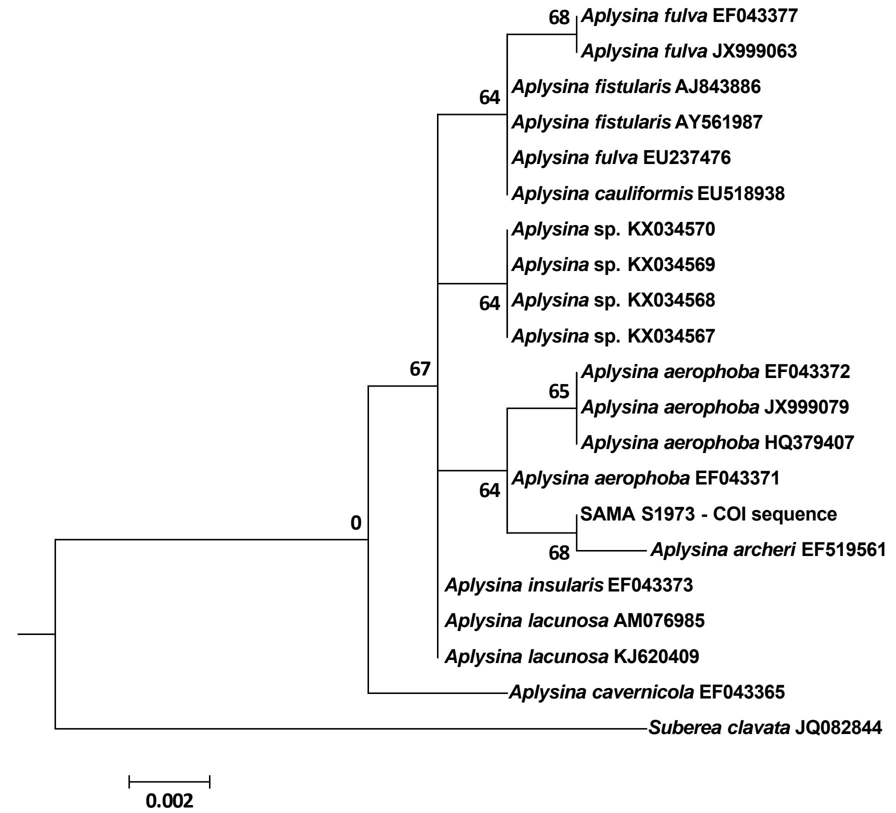
55 Geodia megastrella HM592731
84 Geodia megastrella HM592721
54
Geodia megastrella HM592741
99
Geodia barretti KC574389
82
Geodia hentscheli EU442197
37
Geodia macandrewi EU442198
36
Geodia vaubani EU442202
50
Geodia angulata EU442203
94
Rhabdastrella globostellata HM592673
100
Geodia atlantica EU442195
Stelletta tuberosa HM592735
100
Stelletta tuberosa HM592678
Stelletta tuberculata HM592728
Stelletta fibrosa FJ711643
100
SAMA S1963 - COI sequence
90
Stelletta clavosa KJ494350
94
90 Ecionemia sp. HM592725
95
Ecionemia sp. KJ620402
92 Ecionemia robusta HM592724
90 Ecionemia robusta KJ620401
0.01


c
97 Hymeniacidon heliophila KC884838
33 Hymeniacidon heliophila KC869620
41 Terpios aploos KC869465
Amorphinopsis excavans KC869473
39
95 Halichondria melanadocia KC869508
91 Halichondria okadai AB511881
25 Pseudosuberites sp. AY561917
99 SAMA S1965 - 28S sequence
Hymeniacidon heliophila KJ620391
76
43 Suberites domuncula AJ620113
100 Suberites ficus AY026381
92 Aaptos aaptos KC869496
Suberites sp. KC869500
99 Suberites sp. KC869467
50
100 Suberites aurantiacus KC869577
Diplastrella megastellata AY561893
0 Tethya sp. KC869527
Trachycladus sp. KC869579
51
99 Trachycladus laevispirulifer AY626305
Polymastia pachymastia AY561924
Prosuberites laughlini AY626320
0.005

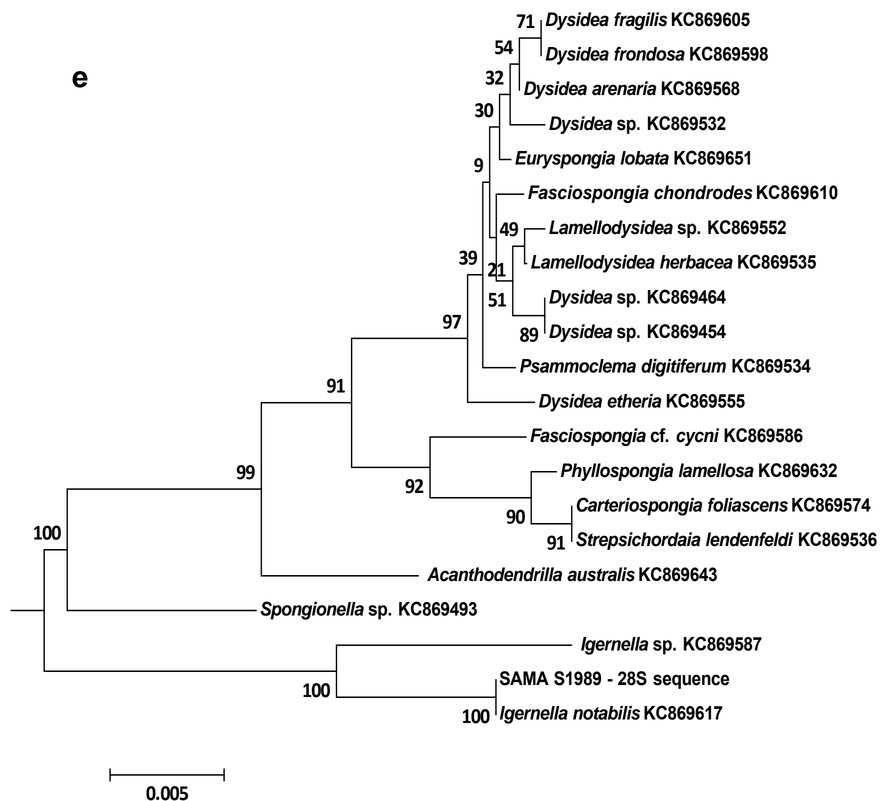

Hymeniacidon perlevis JX477040
Hymeniacidon heliophila EF519629
51 Hymeniacidon sinapium EF217329
Hymeniacidon sinapium EF217332
56 Hymeniacidon perlevis JX477017
92 Hymeniacidon perlevis JX477015
Hymeniacidon heliophila EU076812
56
100 SAMA S1965 - COI sequence
55 Hymeniacidon flavia EF217333
100 Hymeniacidon flavia EF217335
Halichondria okadai EF217340
100 Halichondria sp. EF217336
Protosuberites sp. AY561979
55
89 Halichondria panicea KC869423
Terpios hoshinota KJ008098
97 Pseudosuberites nudus LN850223
83 Halichondria sp. JQ034561
57 Suberites ficus AJ843891
Suberites pagurorum KC869422
100
100 Suberites domuncula JX999078
0.01

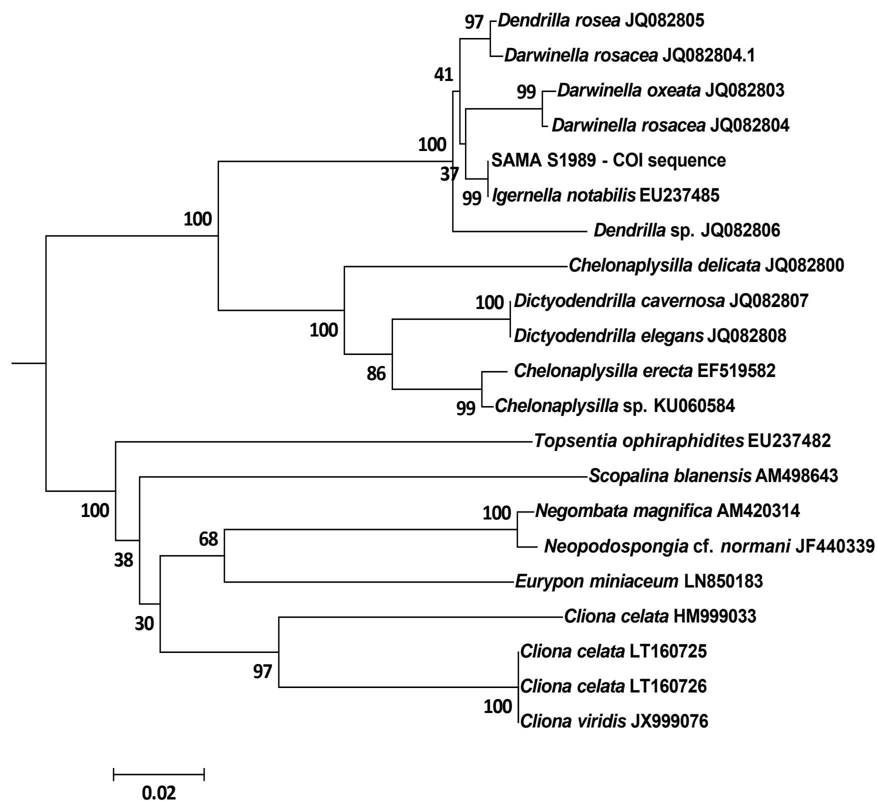
d

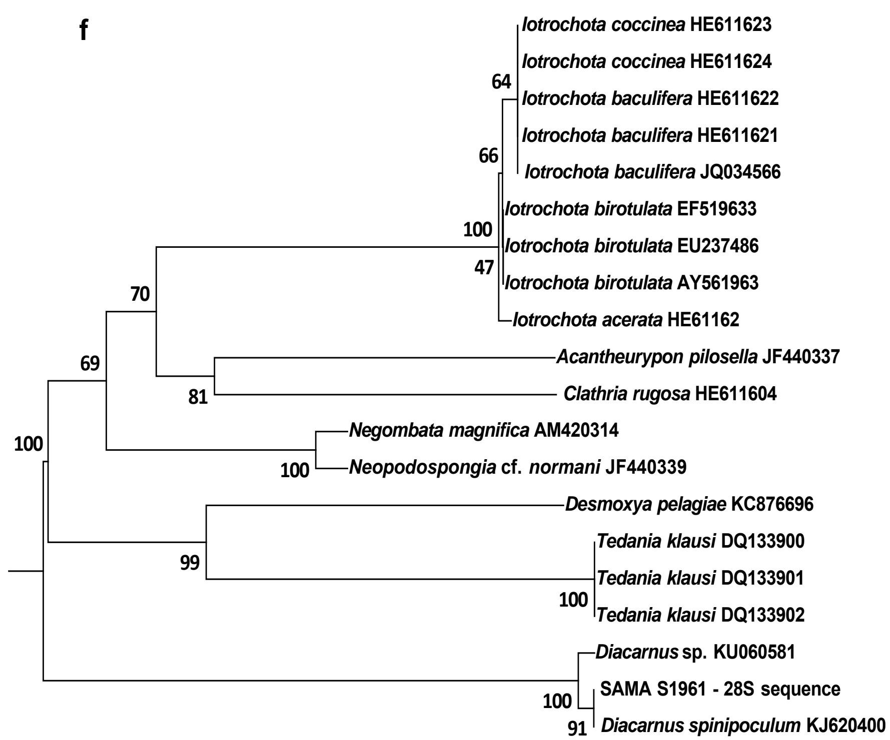
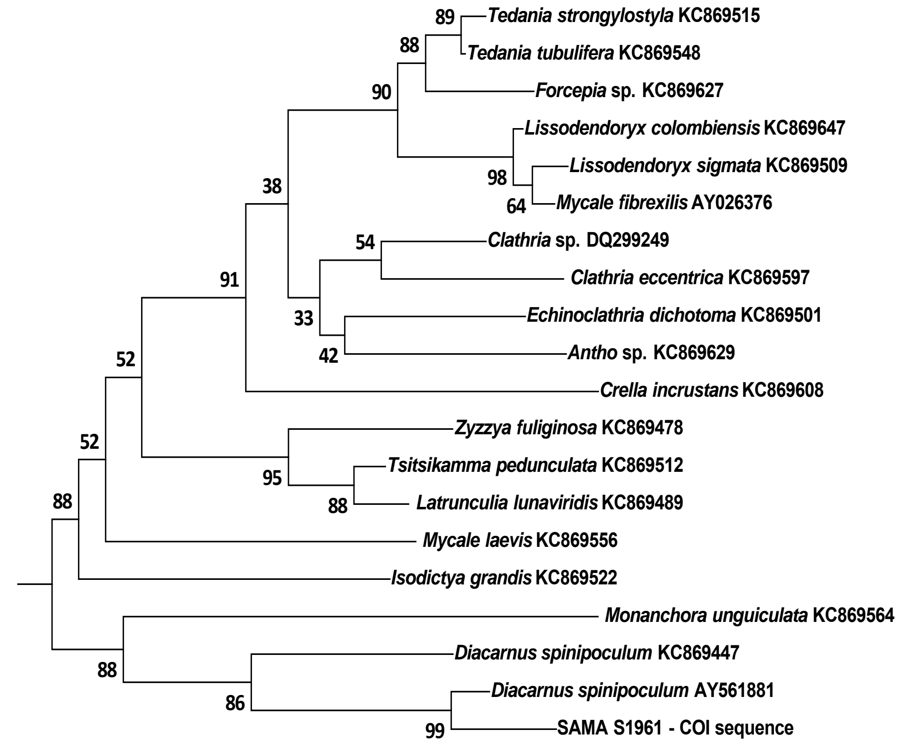


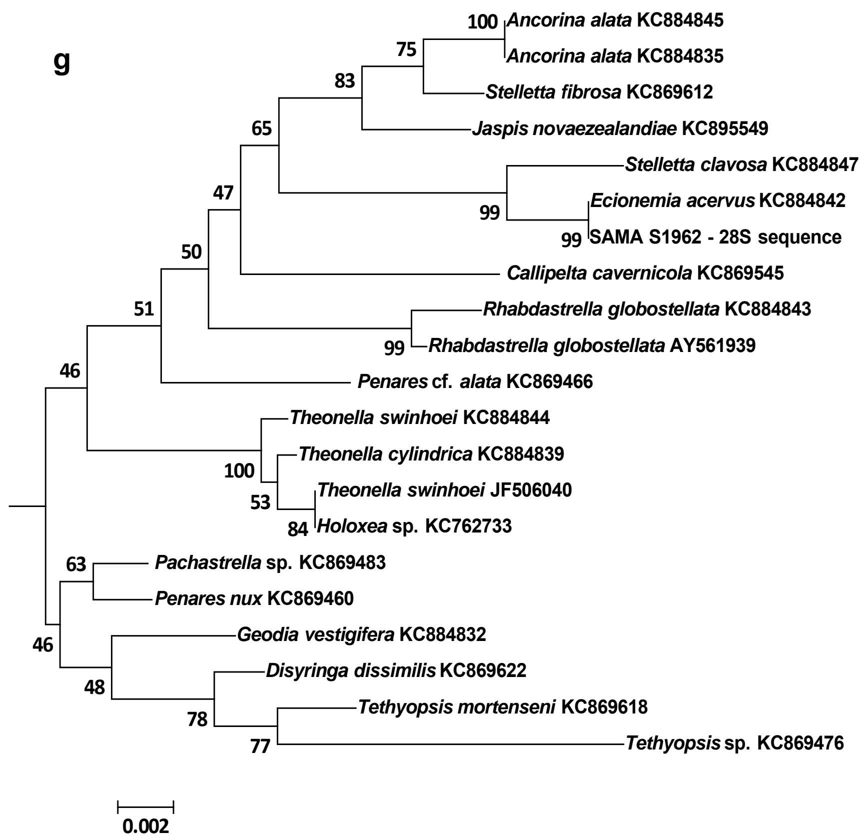
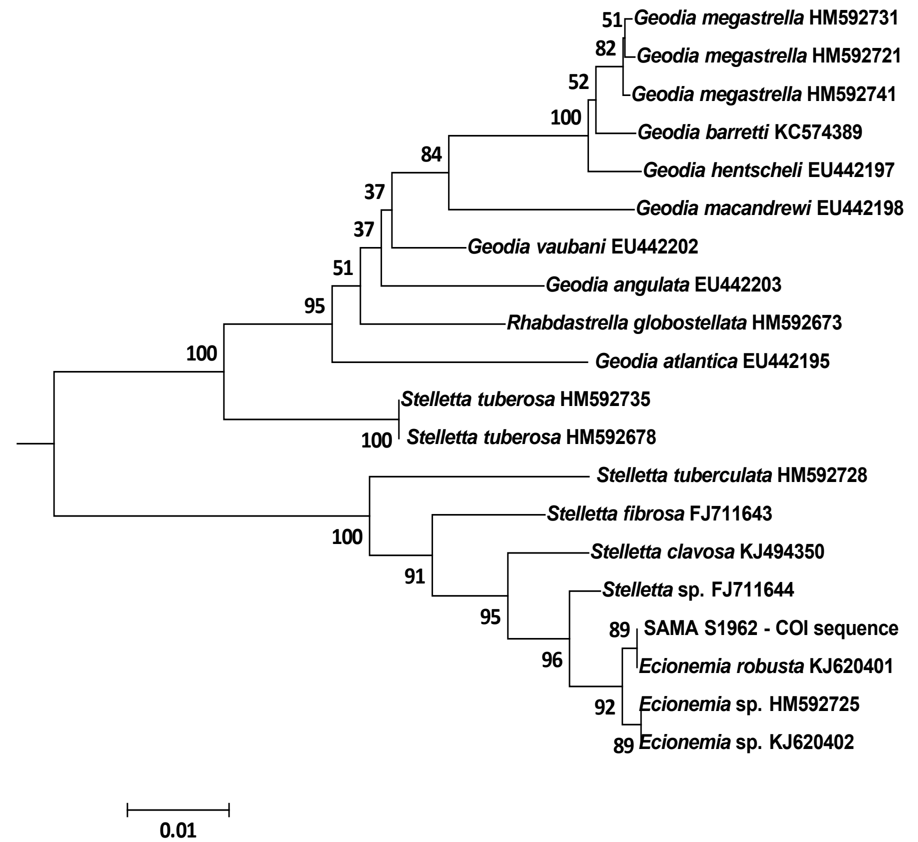


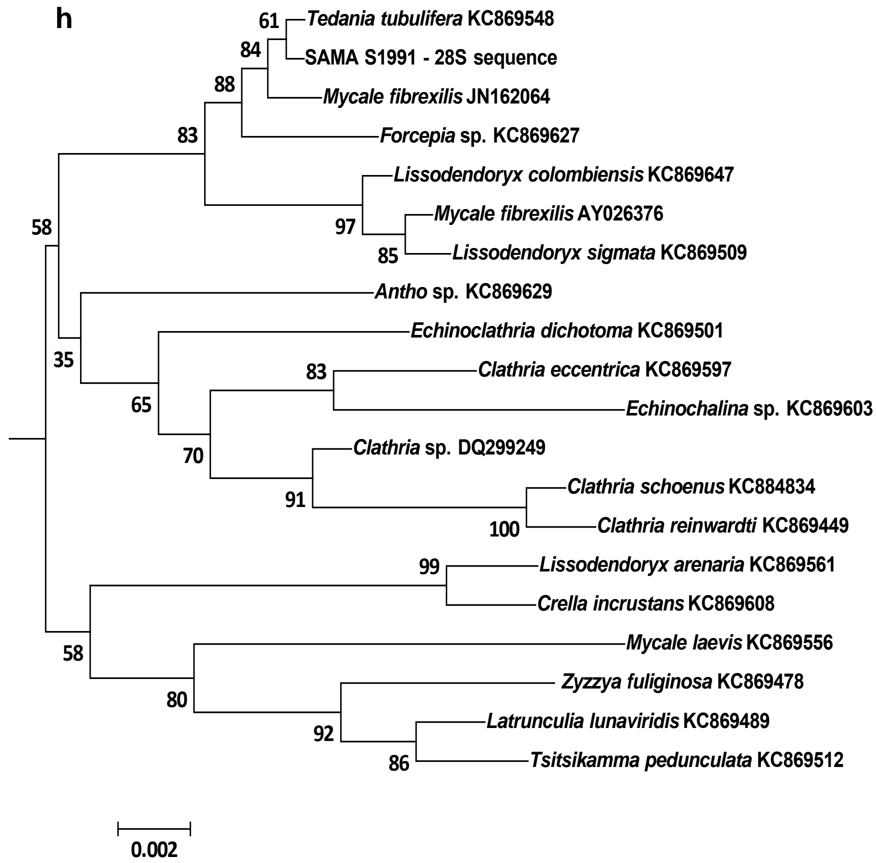
e

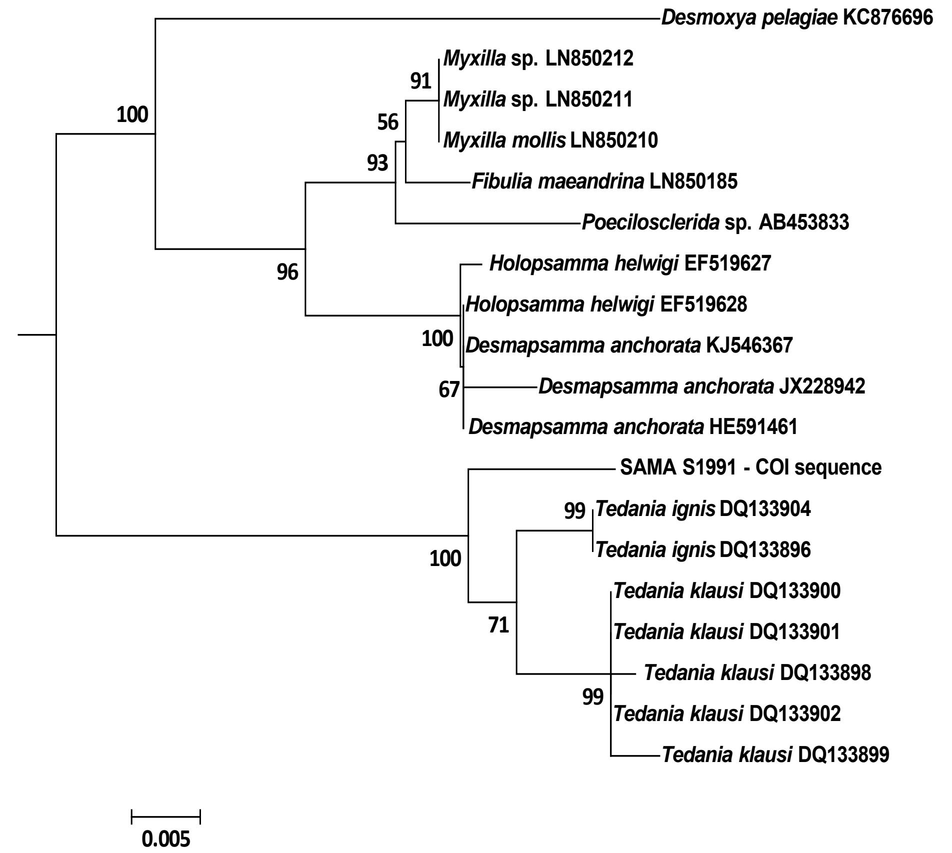


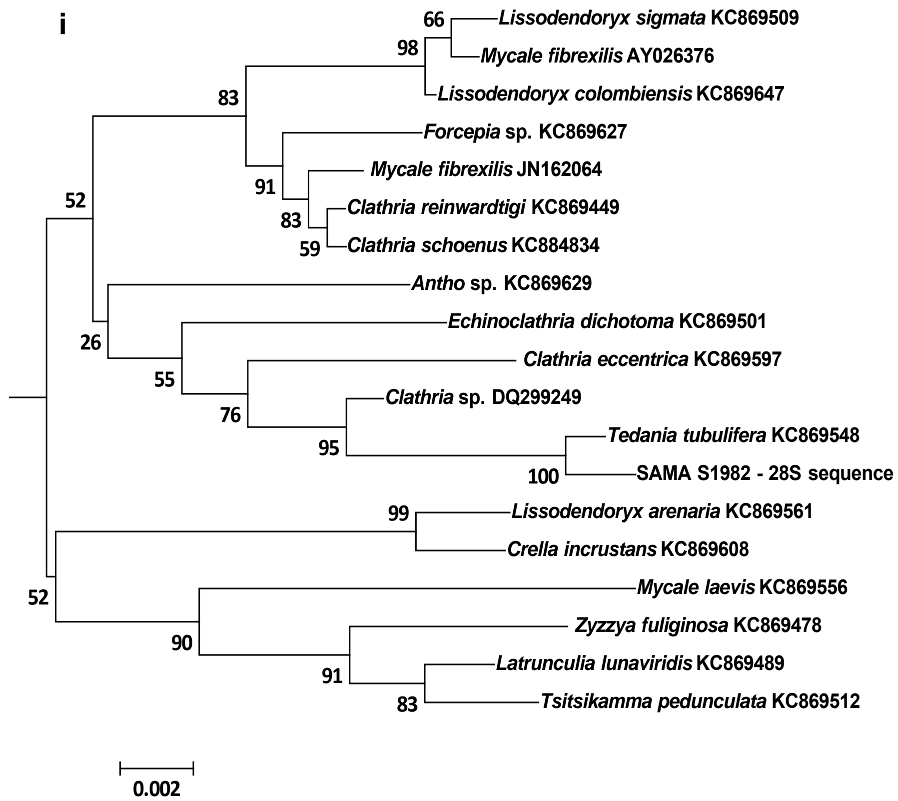
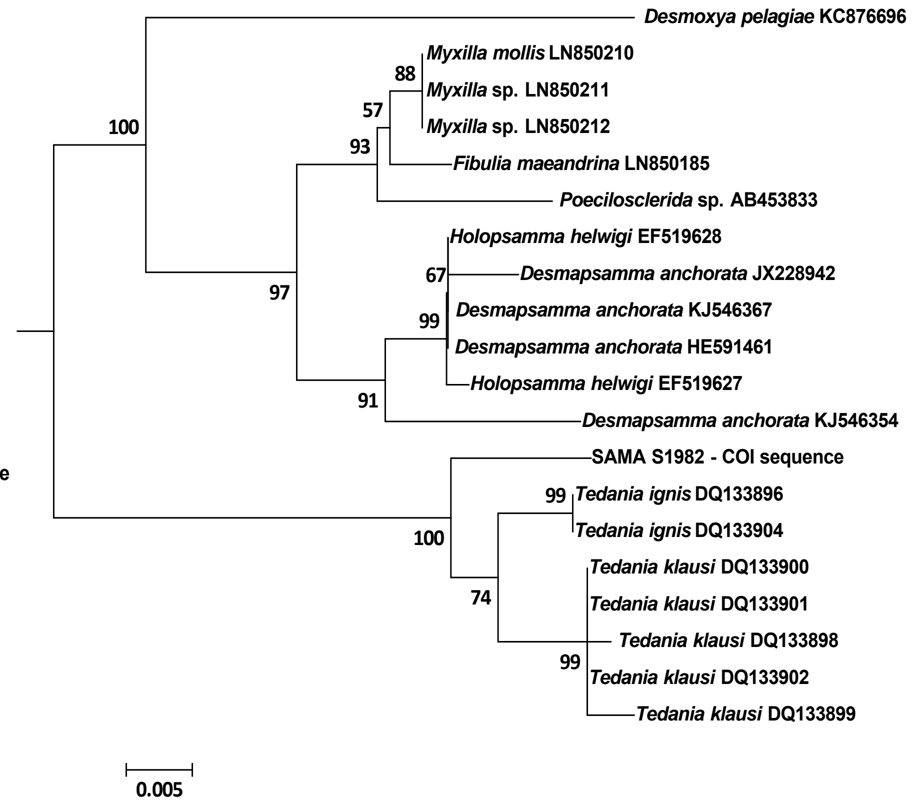


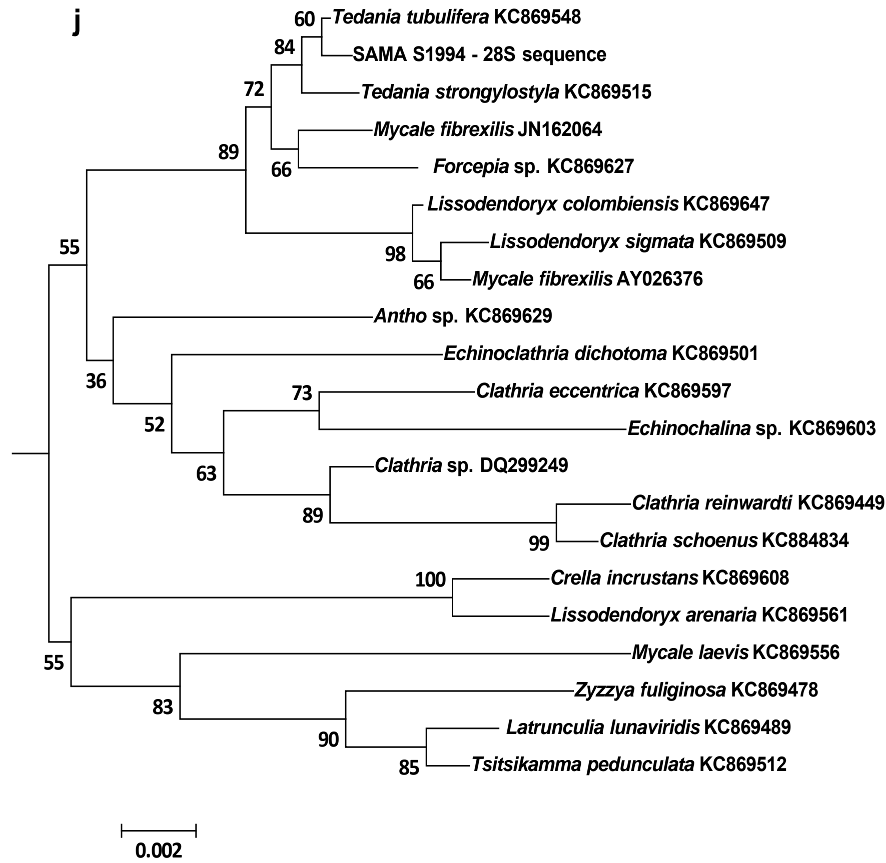

f
Iotrochota coccinea HE611623
Iotrochota coccinea HE611624
Iotrochota baculifera HE611622
Iotrochota baculifera HE611621
Iotrochota baculifera JQ034566
Iotrochota birotulata EF519633
Iotrochota birotulata EU237486
Iotrochota birotulata AY561963
Iotrochota acerata HE61162
Acantheurypon pilosella JF440337
Clathria rugosa HE611604
Negombata magnifica AM420314
Neopodospongia cf. normani JF440339
Desmoxya pelagiae KC876696
Tedania klausi DQ133900
Tedania klausi DQ133901
Tedania klausi DQ133902
Diacarnus sp. KU060581
SAMA S1961 - 28S sequence
Diacarnus spinipoculum KJ620400
64
66
100
47
70
69
81
100
100
99
100
100
91
0.01

Tedania strongylostyla KC869515
Tedania tubulifera KC869548
Forcepia sp. KC869627
Lissodendoryx colombiensis KC869647
Lissodendoryx sigmata KC869509
Mycale fibrexilis AY026376
Clathria sp. DQ299249
Clathria eccentrica KC869597
Echinoclathria dichotoma KC869501
Antho sp. KC869629
Crella incrustans KC869608
Zyzzya fuliginosa KC869478
Tsitsikamma pedunculata KC869512
Latrunculia lunaviridis KC869489
Mycale laevis KC869556
Isodictya grandis KC869522
Monanchora unguiculata KC869564
Diacarnus spinipoculum KC869447
Diacarnus spinipoculum AY561881
SAMA S1961 - COI sequence
89
88
90
98
64
38
54
91
33
42
52
52
95
88
88
88
86
99
0.005

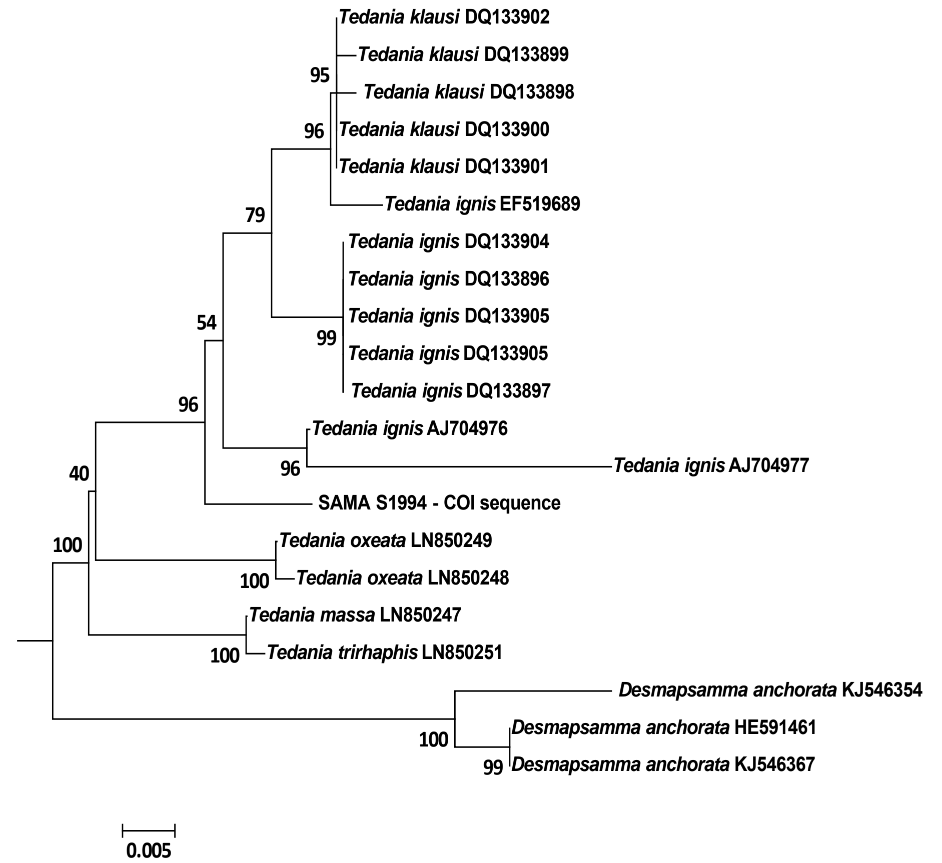

g
100 Ancorina alata KC884845
Ancorina alata KC884835
75
83
Stelletta fibrosa KC869612
Jaspis novaezealandiae KC895549
65
Stelletta clavosa KC884847
47
Ecionemia acervus KC884842
99
99 SAMA S1962 - 28S sequence
50
Callipelta cavernicola KC869545
51
Rhabdastrella globostellata KC884843
99 Rhabdastrella globostellata AY561939
46
Penares cf. alata KC869466
Theonella swinhoei KC884844
Theonella cylindrica KC884839
100
Theonella swinhoei JF506040
53
84 Holoxea sp. KC762733
63
Pachastrella sp. KC869483
Penares nux KC869460
46
Geodia vestigifera KC884832
48
Disyringa dissimilis KC869622
78
Tethyopsis mortenseni KC869618
77
Tethyopsis sp. KC869476
0.002

51 Geodia megastrella HM592731
82 Geodia megastrella HM592721
52
Geodia megastrella HM592741
100
Geodia barretti KC574389
84
Geodia hentscheli EU442197
37
Geodia macandrewi EU442198
37
Geodia vaubani EU442202
51
Geodia angulata EU442203
95
Rhabdastrella globostellata HM592673
100
Geodia atlantica EU442195
Stelletta tuberosa HM592735
100 Stelletta tuberosa HM592678
Stelletta tuberculata HM592728
Stelletta fibrosa FJ711643
100
Stelletta clavosa KJ494350
91
Stelletta sp. FJ711644
95
89 SAMA S1962 - COI sequence
96
Ecionemia robusta KJ620401
92 Ecionemia sp. HM592725
89 Ecionemia sp. KJ620402
0.01

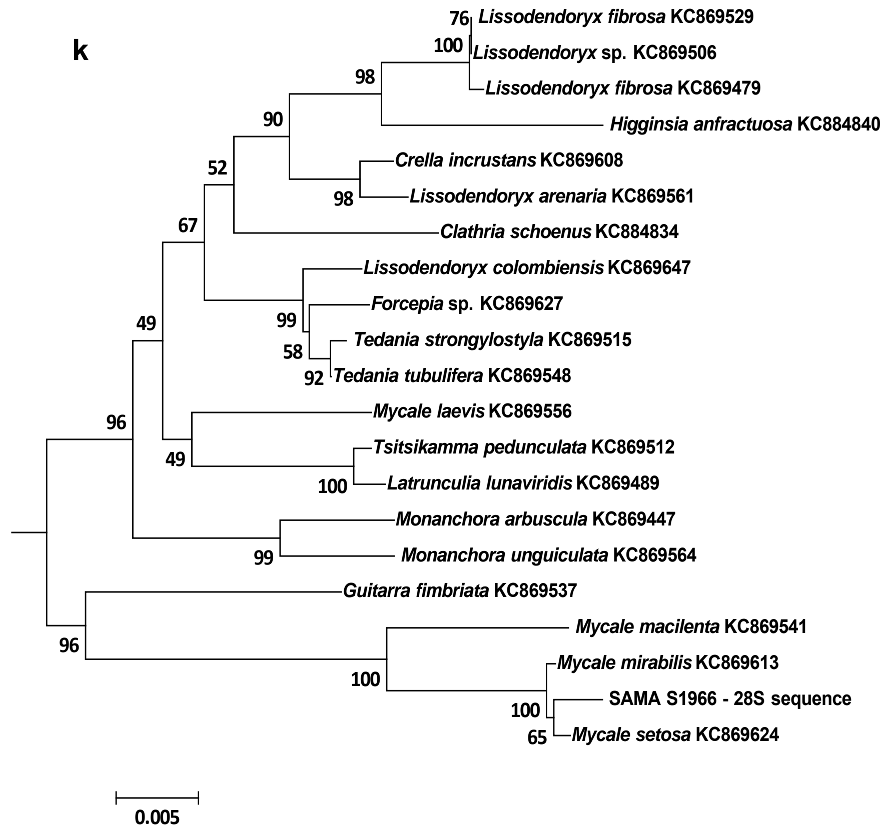

h
Tedania tubulifera KC869548
SAMA S1991 - 28S sequence
Mycale fibrexilis JN162064
Forcepia sp. KC869627
Lissodendoryx colombiensis KC869647
Mycale fibrexilis AY026376
Lissodendoryx sigmata KC869509
Antho sp. KC869629
Echinoclathria dichotoma KC869501
Clathria eccentrica KC869597
Echinochalina sp. KC869603
Clathria sp. DQ299249
Clathria schoenus KC884834
Clathria reinwardti KC869449
Lissodendoryx arenaria KC869561
Crella incrustans KC869608
Mycale laevis KC869556
Zyzzya fuliginosa KC869478
Latrunculia lunaviridis KC869489
Tsitsikamma pedunculata KC869512
0.002

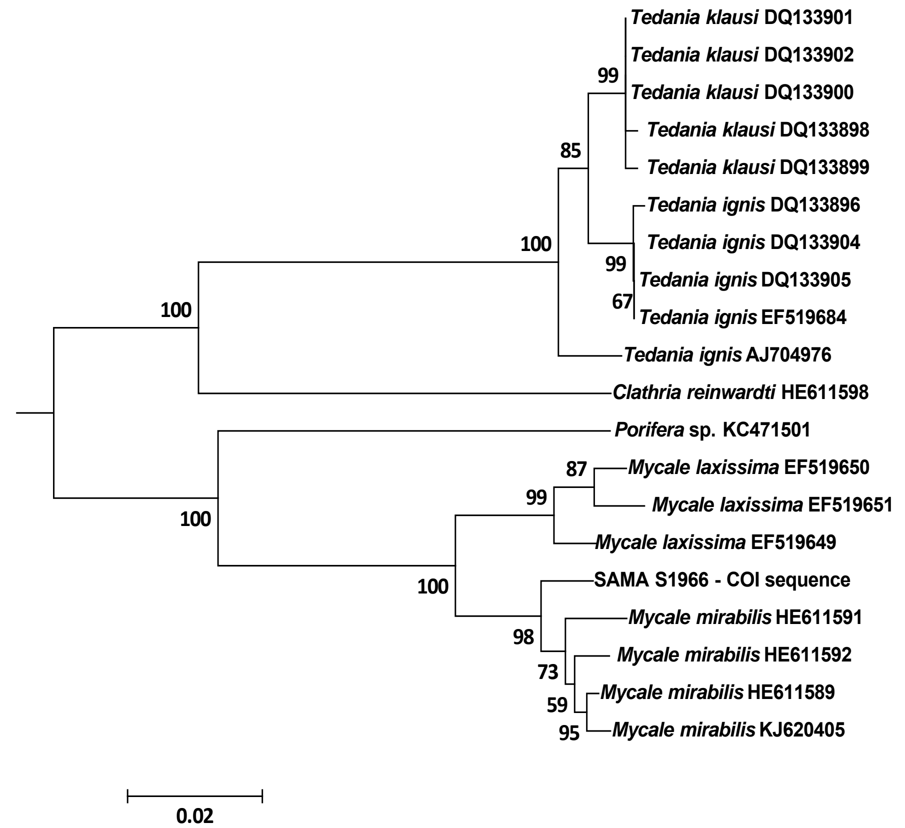

Desmoxya pelagiae KC876696
Myxilla sp. LN850212
Myxilla sp. LN850211
Myxilla mollis LN850210
Fibulia maeandrina LN850185
Poecilosclerida sp. AB453833
Holopsamma helwigi EF519627
Holopsamma helwigi EF519628
Desmapsamma anchorata KJ546367
Desmapsamma anchorata JX228942
Desmapsamma anchorata HE591461
SAMA S1991 - COI sequence
Tedania ignis DQ133904
Tedania ignis DQ133896
Tedania klausi DQ133900
Tedania klausi DQ133901
Tedania klausi DQ133898
Tedania klausi DQ133902
Tedania klausi DQ133899
0.005

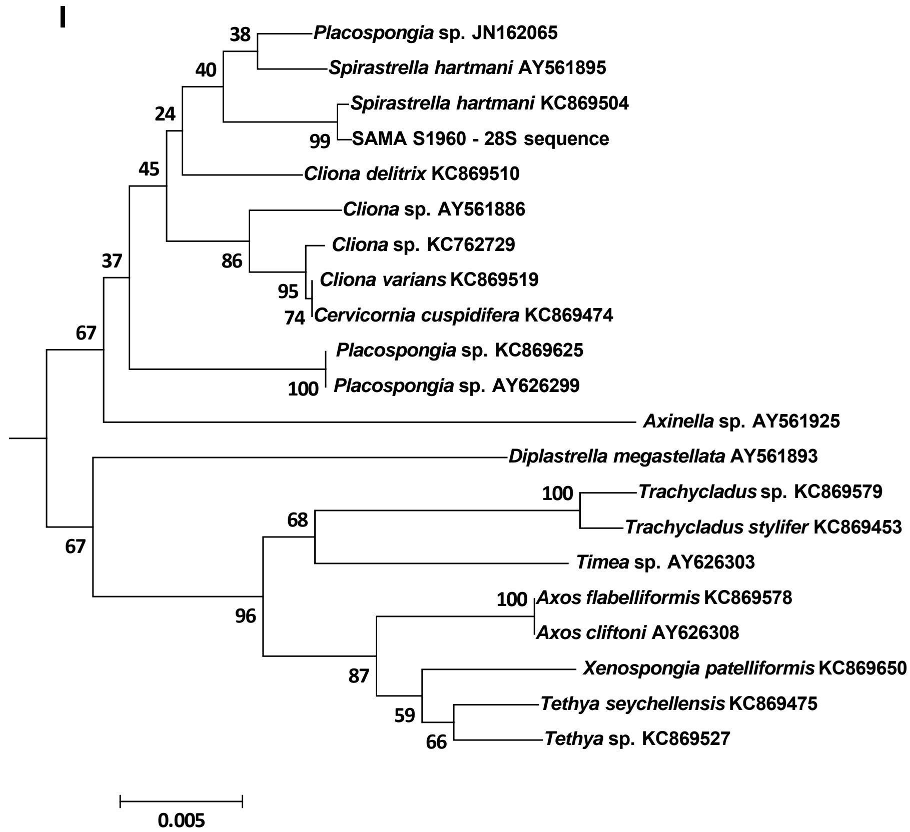

i
Lissodendoryx sigmata KC869509
Mycale fibrexilis AY026376
Lissodendoryx colombiensis KC869647
Forcepia sp. KC869627
Mycale fibrexilis JN162064
Clathria reinwardtigi KC869449
Clathria schoenus KC884834
Antho sp. KC869629
Echinoclathria dichotoma KC869501
Clathria eccentrica KC869597
Clathria sp. DQ299249
Tedania tubulifera KC869548
SAMA S1982 - 28S sequence
Lissodendoryx arenaria KC869561
Crella incrustans KC869608
Mycale laevis KC869556
Zyzzya fuliginosa KC869478
Latrunculia lunaviridis KC869489
Tsitsikamma pedunculata KC869512
66
98
83
91
83
59
52
26
55
76
95
100
99
52
90
91
83
0.002

Desmoxya pelagiae KC876696
Myxilla mollis LN850210
Myxilla sp. LN850211
Myxilla sp. LN850212
Fibulia maeandrina LN850185
Poecilosclerida sp. AB453833
Holopsamma helwigi EF519628
Desmapsamma anchorata JX228942
Desmapsamma anchorata KJ546367
Desmapsamma anchorata HE591461
Holopsamma helwigi EF519627
Desmapsamma anchorata KJ546354
SAMA S1982 - COI sequence
Tedania ignis DQ133896
Tedania ignis DQ133904
Tedania klausi DQ133900
Tedania klausi DQ133901
Tedania klausi DQ133898
Tedania klausi DQ133902
Tedania klausi DQ133899
100
88
57
93
67
97
99
91
99
100
74
99
0.005

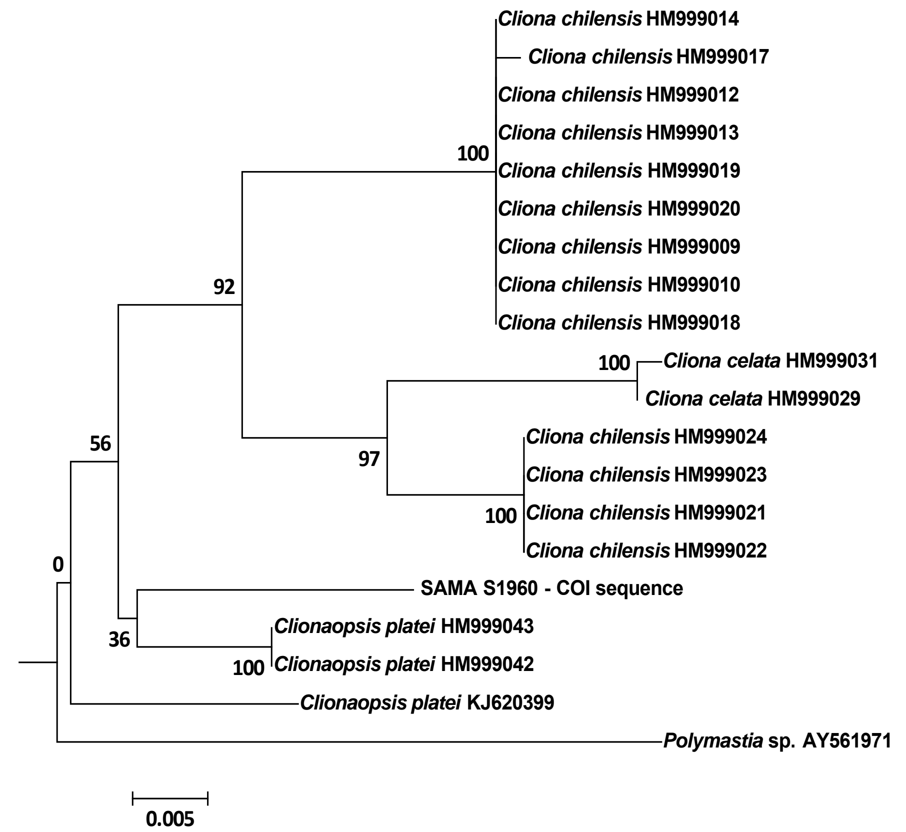

j
Tedania tubulifera KC869548
SAMA S1994 - 28S sequence
Tedania strongylostyla KC869515
Mycale fibrexilis JN162064
Forcepia sp. KC869627
Lissodendoryx colombiensis KC869647
Lissodendoryx sigmata KC869509
Mycale fibrexilis AY026376
Antho sp. KC869629
Echinoclathria dichotoma KC869501
Clathria eccentrica KC869597
Echinochalina sp. KC869603
Clathria sp. DQ299249
Clathria reinwardti KC869449
Clathria schoenus KC884834
Crella incrustans KC869608
Lissodendoryx arenaria KC869561
Mycale laevis KC869556
Zyzzya fuliginosa KC869478
Latrunculia lunaviridis KC869489
Tsitsikamma pedunculata KC869512
0.002

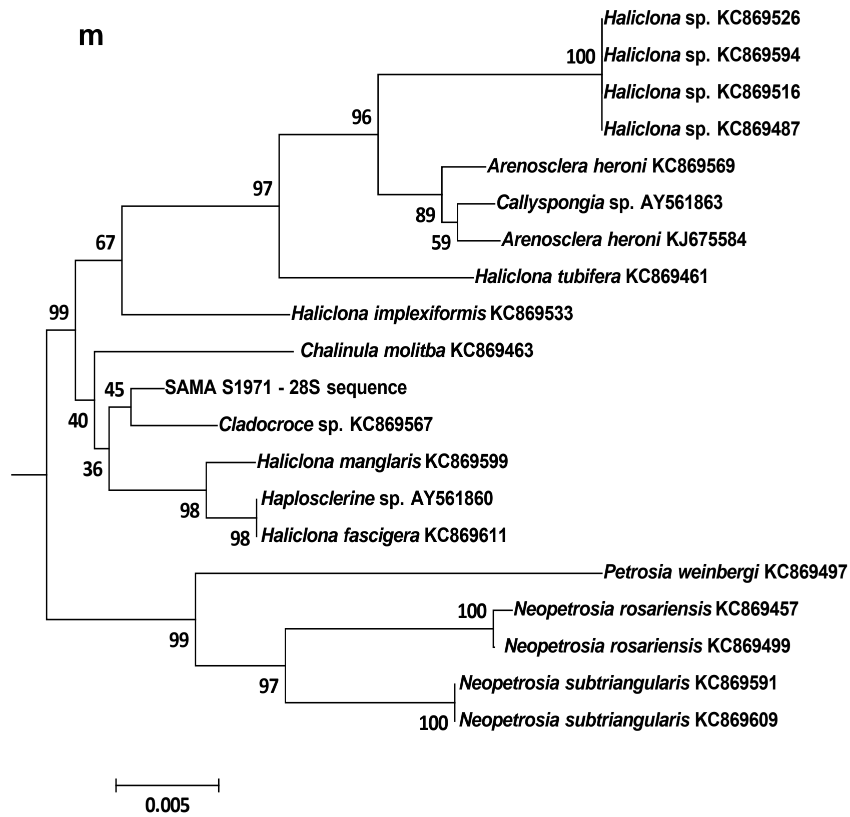

Tedania klausi DQ133902
Tedania klausi DQ133899
Tedania klausi DQ133898
Tedania klausi DQ133900
Tedania klausi DQ133901
Tedania ignis EF519689
Tedania ignis DQ133904
Tedania ignis DQ133896
Tedania ignis DQ133905
Tedania ignis DQ133905
Tedania ignis DQ133897
Tedania ignis AJ704976
Tedania ignis AJ704977
SAMA S1994 - COI sequence
Tedania oxeata LN850249
Tedania oxeata LN850248
Tedania massa LN850247
Tedania trirhaphis LN850251
Desmapsamma anchorata KJ546354
Desmapsamma anchorata HE591461
Desmapsamma anchorata KJ546367
0.005


k
Lissodendoryx fibrosa KC869529
Lissodendoryx sp. KC869506
Lissodendoryx fibrosa KC869479
Higginsia anfractuosa KC884840
Crella incrustans KC869608
Lissodendoryx arenaria KC869561
Clathria schoenus KC884834
Lissodendoryx colombiensis KC869647
Forcepia sp. KC869627
Tedania strongylostyla KC869515
Tedania tubulifera KC869548
Mycale laevis KC869556
Tsitsikamma pedunculata KC869512
Latrunculia lunaviridis KC869489
Monanchora arbuscula KC869447
Monanchora unguiculata KC869564
Guitarra fimbriata KC869537
Mycale macilenta KC869541
Mycale mirabilis KC869613
SAMA S1966 - 28S sequence
Mycale setosa KC869624
76
100
98
90
98
52
67
49
99
58
92
96
49
100
99
96
100
100
65
0.005

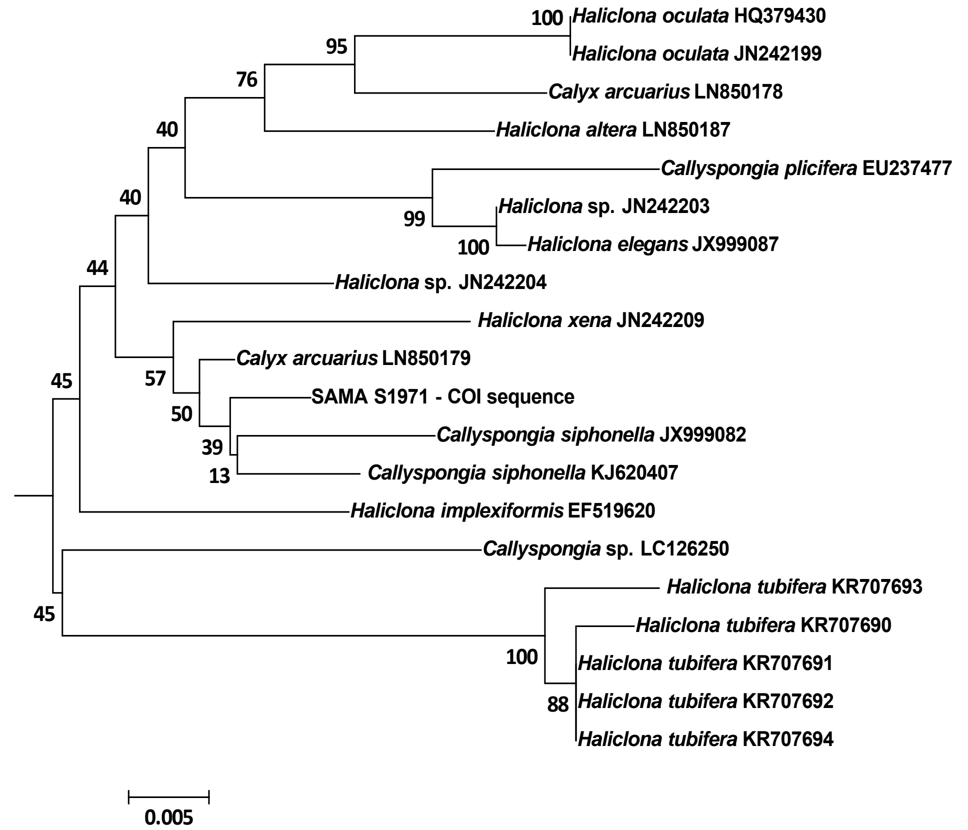

Tedania klausi DQ133901
Tedania klausi DQ133902
Tedania klausi DQ133900
Tedania klausi DQ133898
Tedania klausi DQ133899
Tedania ignis DQ133896
Tedania ignis DQ133904
Tedania ignis DQ133905
Tedania ignis EF519684
Tedania ignis AJ704976
Clathria reinwardti HE611598
Porifera sp. KC471501
Mycale laxissima EF519650
Mycale laxissima EF519651
Mycale laxissima EF519649
SAMA S1966 - COI sequence
Mycale mirabilis HE611591
Mycale mirabilis HE611592
Mycale mirabilis HE611589
Mycale mirabilis KJ620405
99
85
100
99
67
100
87
99
100
100
98
73
59
95
0.02


I
Placospongia sp. JN162065
Spirastrella hartmani AY561895
Spirastrella hartmani KC869504
SAMA S1960 - 28S sequence
Cliona delitrix KC869510
Cliona sp. AY561886
Cliona sp. KC762729
Cliona varians KC869519
Cervicornia cuspidifera KC869474
Placospongia sp. KC869625
Placospongia sp. AY626299
Axinella sp. AY561925
Diplastrella megastellata AY561893
Trachycladus sp. KC869579
Trachycladus stylifer KC869453
Timea sp. AY626303
Axos flabelliformis KC869578
Axos cliftoni AY626308
Xenospongia patelliformis KC869650
Tethya seychellensis KC869475
Tethya sp. KC869527
0.005

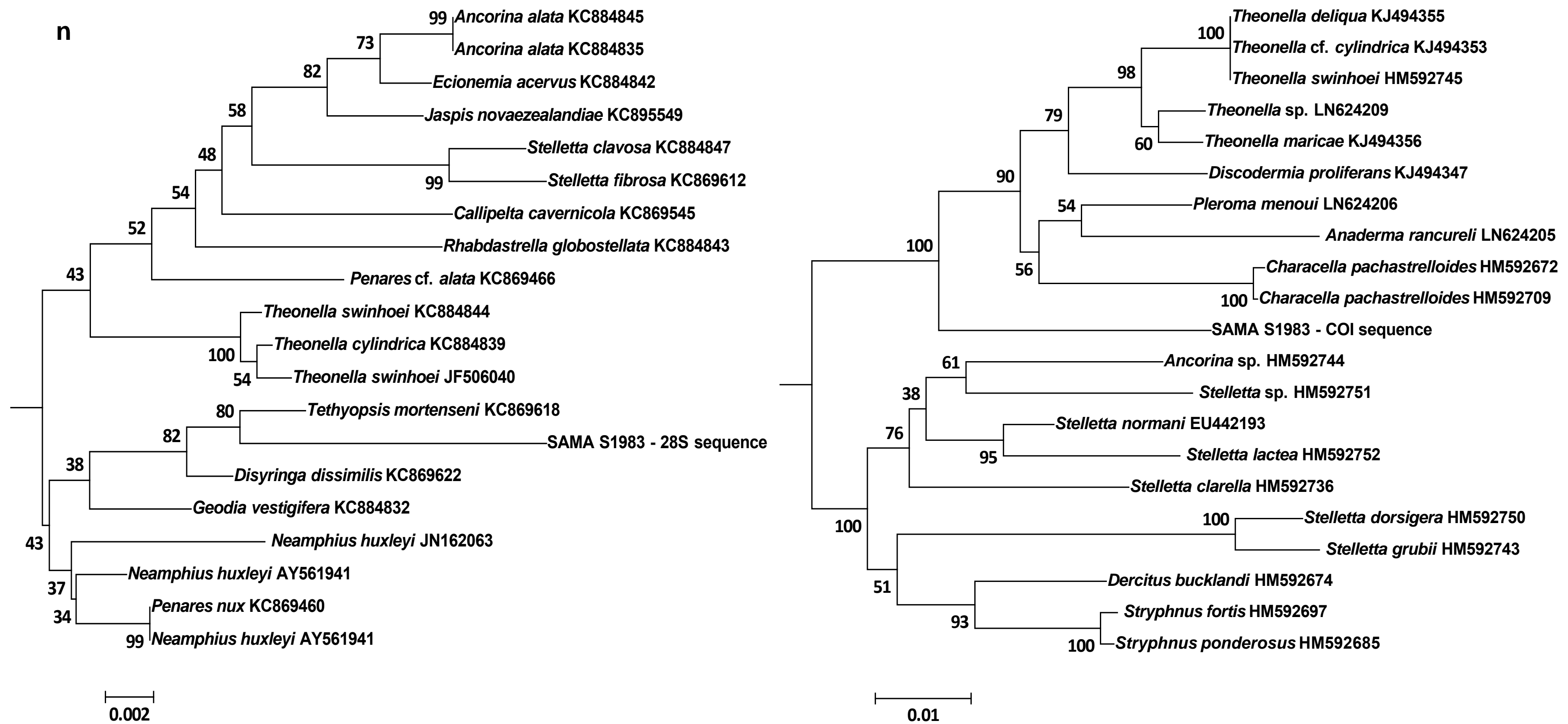

Cliona chilensis HM999014
Cliona chilensis HM999017
Cliona chilensis HM999012
Cliona chilensis HM999013
Cliona chilensis HM999019
Cliona chilensis HM999020
Cliona chilensis HM999009
Cliona chilensis HM999010
Cliona chilensis HM999018
Cliona celata HM999031
Cliona celata HM999029
Cliona chilensis HM999024
Cliona chilensis HM999023
Cliona chilensis HM999021
Cliona chilensis HM999022
SAMA S1960 - COI sequence
Clionaopsis platei HM999043
Clionaopsis platei HM999042
Clionaopsis platei KJ620399
Polymastia sp. AY561971
0.005



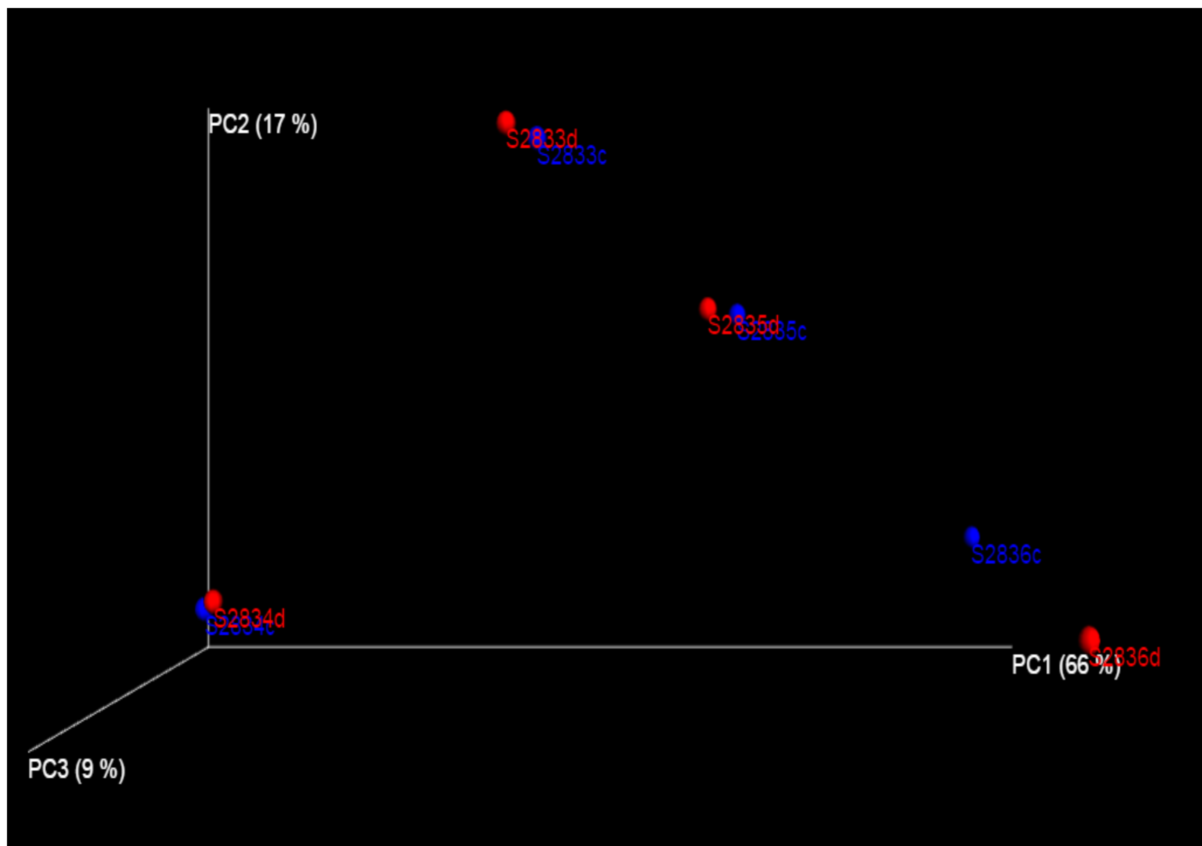
m
Haliclona sp. KC869526
Haliclona sp. KC869594
Haliclona sp. KC869516
Haliclona sp. KC869487
Arenosclera heroni KC869569
Callyspongia sp. AY561863
Arenosclera heroni KJ675584
Haliclona tubifera KC869461
Haliclona implexiformis KC869533
Chalinula molitba KC869463
SAMA S1971 - 28S sequence
Cladocroce sp. KC869567
Haliclona manglaris KC869599
Haplosclerine sp. AY561860
Haliclona fascigera KC869611
Petrosia weinbergi KC869497
Neopetrosia rosariensis KC869457
Neopetrosia rosariensis KC869499
Neopetrosia subtriangularis KC869591
Neopetrosia subtriangularis KC869609
0.005



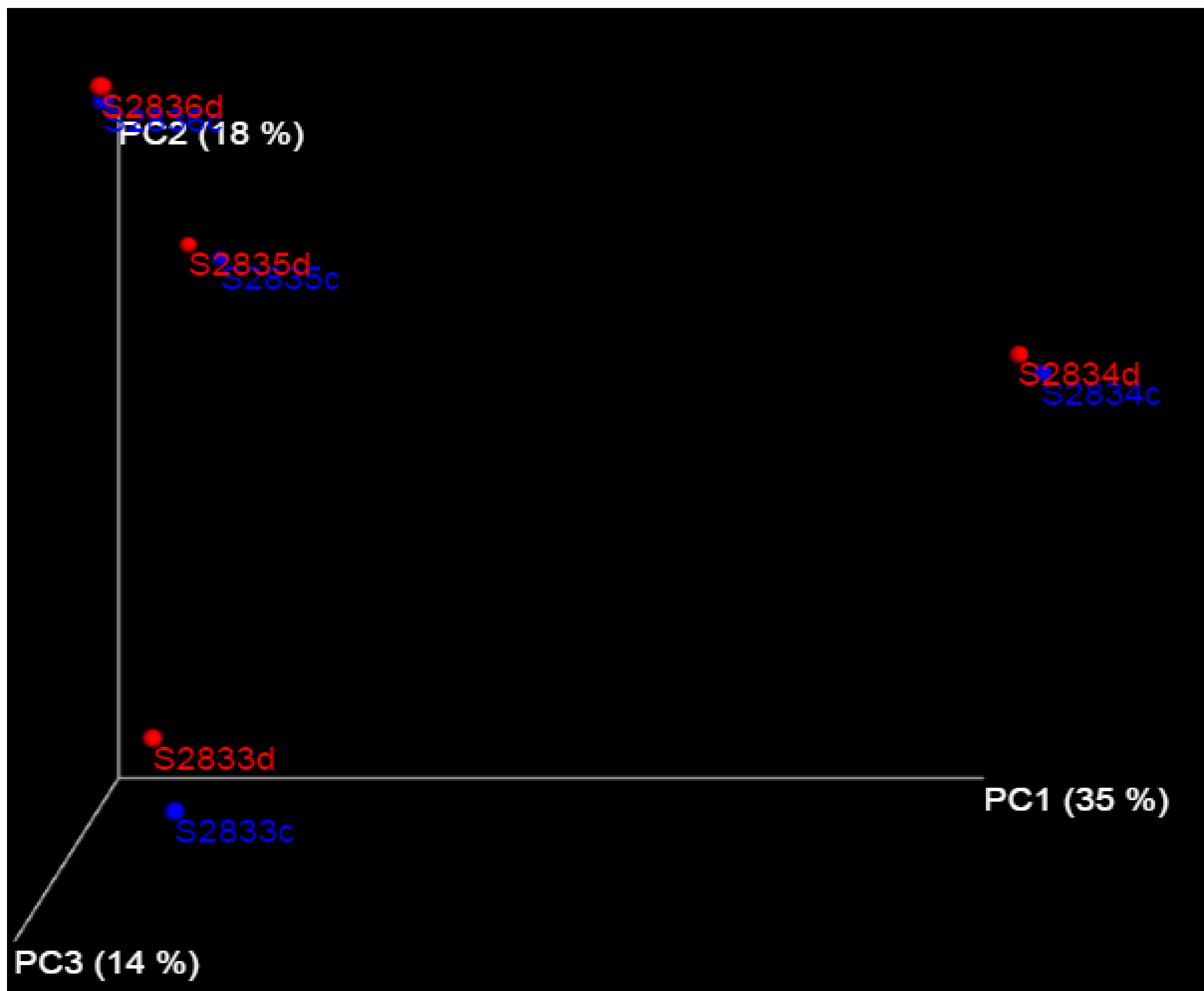
Haliclona oculata HQ379430
Haliclona oculata JN242199
Calyx arcuarius LN850178
Haliclona altera LN850187
Callyspongia plicifera EU237477
Haliclona sp. JN242203
Haliclona elegans JX999087
Haliclona sp. JN242204
Haliclona xena JN242209
Calyx arcuarius LN850179
SAMA S1971 - COI sequence
Callyspongia siphonella JX999082
Callyspongia siphonella KJ620407
Haliclona implexiformis EF519620
Callyspongia sp. LC126250
Haliclona tubifera KR707693
Haliclona tubifera KR707690
Haliclona tubifera KR707691
Haliclona tubifera KR707692
Haliclona tubifera KR707694
0.005



Appendix Figure 2 - 2 Phylogenetic relationship of fourteen sponges using Neighbor Joining method based on 28S rRNA gene and COI mtDNA. a. Phylogenetic relationship of sponge SAMA S1981. b. Phylogenetic relationship of sponge SAMA S1963. c. Phylogenetic relationship of sponge SAMA S1965. d. Phylogenetic relationship of sponge SAMA S1973. e. Phylogenetic relationship of sponge SAMA S1989. f. Phylogenetic relationship of sponge SAMA S1961. g. Phylogenetic relationship of sponge SAMA S1962. h. Phylogenetic relationship of sponge SAMA S1991. i.Phylogenetic relationship of sponge SAMA S1982. j. Phylogenetic relationship of sponge SAMA S1994. k. Phylogenetic relationship of sponge SAMA S1966. l. Phylogenetic relationship of sponge SAMA S1960. m. Phylogenetic relationship of sponge SAMA S1971. n. Phylogenetic relationship of sponge SAMA S1983.



Appendix Figure 5 - 1 Principal Coordinates Analysis (PCoA) plots showing the similarity of microbial communities of four sponge species in four taxonomic orders revealed by primer set for 16S rRNA gene region V1V3. S2833: *Aplysina archeri*, S2834: *Halichondria okadai*, S2835: *Igernella notabilis*, S2836: Tedania tubulifera; a and b distinguish the two duplicates.



Appendix Figure 5 - 2 Principal Coordinates Analysis (PCoA) plots showing the similarity of microbial communities of four sponge species in four taxonomic orders revealed by primer set for 16S rRNA gene region V5V8. S2833: *Aplysina archeri*, S2834: *Halichondria okadai*, S2835: *Igernella notabilis*, S2836: Tedania tubulifera; a and b distinguish the two duplicates.

Appl Microbiol Biotechnol
DOI 10.1007/s00253-015-6875-8

METHODS AND PROTOCOLS

Sponge-associated actinobacterial diversity: validation of the methods of actinobacterial DNA extraction and optimization of 16S rRNA gene amplification

Qi Yang[1,2] · Christopher M. M. Franco[1,2] · Wei Zhang[1,2,3]

Received: 18 May 2015 / Revised: 16 July 2015 / Accepted: 21 July 2015
© Springer-Verlag Berlin Heidelberg 2015

Abstract Experiments were designed to validate the two common DNA extraction protocols (CTAB-based method and DNeasy Blood & Tissue Kit) used to effectively recover actinobacterial DNA from sponge samples in order to study the sponge-associated actinobacterial diversity. This was done by artificially spiking sponge samples with actinobacteria (spores, mycelia and a combination of the two). Our results demonstrated that both DNA extraction methods were effective in obtaining DNA from the sponge samples as well as the sponge samples spiked with different amounts of actinobacteria. However, it was noted that in the presence of the sponge, the bacterial 16S rRNA gene could not be amplified unless the combined DNA template was diluted. To test the hypothesis that the extracted sponge DNA contained inhibitors, dilutions of the DNA extracts were tested for six sponge species representing five orders. The results suggested that the inhibitors were co-extracted with the sponge DNA, and a high dilution of this DNA was required for the successful PCR amplification for most of the samples. The optimized PCR conditions, including primer selection, PCR reaction system and program optimization, further improved the PCR performance. However, no single PCR condition was found to be suitable for the diverse sponge samples using various primer sets. These results highlight for the first time that the DNA extraction methods used are effective in obtaining actinobacterial DNA and that the presence of inhibitors in the sponge DNA requires high dilution coupled with fine tuning of the PCR conditions to achieve success in the study of sponge-associated actinobacterial diversity.

Keywords Sponge (Porifera) · DNA extraction efficiency · Inhibitor · PCR optimization · Validation · Actinobacteria

✉ Christopher M. M. Franco
chris.franco@flinders.edu.au

✉ Wei Zhang
wei.zhang@flinders.edu.au

[1] Centre for Marine Bioproducts Development, Adelaide, SA 5042, Australia

[2] Department of Medical Biotechnology, School of Medicine, Flinders University, Adelaide, SA 5042, Australia

[3] Centre for Marine Drugs, Renji Hospital, Shanghai Jiaotong University, Shanghai 200240, China

Introduction

Marine sponges (phylum Porifera) are considered to be the oldest multicellular animals with a history of more than 600 million years. They have attracted substantial research interests because of their ecological importance and their production of a wide range of bioactive compounds for pharmacological use (Ando et al. 2010; Blunt et al. 2010; Leal et al. 2012; Sirirak et al. 2013; Vogel 2008; Waters et al. 2010). One striking characteristic of sponges is their association with a remarkable array of microorganisms, such as archaea (Preston et al. 1996; Radax et al. 2012; Turque et al. 2010), bacteria (Hentschel et al. 2001; Hentschel et al. 2012; Montalvo and Hill 2011; Radwan et al. 2010; Richardson et al. 2012; Schmitt et al. 2011; Webster and Hill 2001) including actinobacteria (Abdelmohsen et al. 2014a) and cyanobacteria (Alex et al. 2012; Thacker and Starnes 2003), unicellular algae (Annenkova et al. 2011; He et al. 2014; Hentschel et al. 2012; Wecker et al. 2015) and fungi (Gopi et al. 2012; Maldonado et al. 2005). These microorganisms are reported to comprise between 35 and 40 % of the total tissue volume in some sponge species (Hentschel et al. 2012; Taylor et al. 2007; Vacelet and Donadey 1977) and exceed a density

Published online: 07 August 2015