MOLECULAR APPROACHES IN CONSERVATION AND ECOLOGY RESEARCH IN MARINE SPONGES

SARAH MIRIAM GRIFFITHS

A thesis submitted in partial fulfilment of the requirements of the Manchester Metropolitan University for the degree of Doctor of Philosophy

School of Science and the Environment
Faculty of Science and Engineering
The Manchester Metropolitan University

2018

Table of Contents

General Abstract

Molecular approaches in conservation and ecology research in marine sponges

Marine ecosystems face a challenging future as a result of human activity, and as such, it is vital to understand the ecology of marine organisms in order to effectively conserve them. Increasingly, molecular ecology tools are used for this purpose. In this thesis, I developed molecular resources for the study of sponges, an ecologically important and diverse marine group. Firstly, I developed a bioinformatics pipeline for the *de novo* isolation of microsatellites - genetic markers used to study intraspecific genetic diversity. Using this method, I characterised microsatellites for two Tropical Western Atlantic sponge species, *Ircinia campana* and *Spheciospongia vesparium*. These species have suffered mass mortalities in Florida Bay (Florida Keys, USA), with devastating effects on the ecosystem, and are currently among the subjects of an ecological restoration program. With these microsatellites, I investigated the spatial genetic structure of populations of both species across the Greater Caribbean and within the Florida Keys. I found genetic isolation by distance patterns over the region, indicating distance-limited dispersal, and also identified oceanographic barriers to population connectivity. On smaller spatial scales, genetic structure was also present, but patterns of genetic patchiness suggested that stochastic processes were more influential in determining structure at this scale. Finally, I investigated the effect of the *Ircinia campana* genotype on the composition of its associated microbial community. I found that more genetically similar sponges hosted more similar microbiomes, supporting the theory of a close co-evolutionary relationship, and highlighting the importance of genetic diversity within species. With these insights on the distribution and effects of genetic diversity, I discuss the implications for sponge restoration in Florida Bay, and make recommendations for conservation practitioners to ensure resilience in sponge populations.

Acknowledgements

My thanks firstly go to my supervisor, Richard Preziosi, who has not only been a wise and supportive mentor, but has also offered me countless opportunities to grow as a researcher over the years. My colleagues both at the Faculty of Life Sciences at the University of Manchester and the School of Science and the Environment at Manchester Metropolitan University have been helpful, insightful and a pleasure to work with. I particularly thank members of the Preziosi and Rowntree lab groups, past and present, for the scientific insight and lab assistance that helped to progress my research, and for the beers that helped to maintain my sanity. I am especially grateful to Friederike Clever for her invaluable fieldwork and map-making assistance, Graeme Fox for his help in coding and bioinformatics, Eve Taylor-Cox for her hard work in the lab, Nate Truelove for his help especially at the beginning of my research career, and Jennifer Rowntree for advice in experimental design and analysis. Outside of Manchester, many members of the sponge research community have given me valuable feedback and discussions about my results, for which I am incredibly grateful.

Collecting samples from so many locations would not have been possible without assistance from a number of individuals and institutions. Many thanks to Mark Butler and Don Behringer for sharing their years of knowledge on the Florida Bay ecosystem and their tireless conservation efforts in restoring the Bay to its former glory. Thanks also to all the volunteers from the local community and various universities for doing the hard work of the sponge restoration. I am indebted to Thierry Pérez for his support in my PhD – taking part in the Pacotilles research cruise not only greatly enhanced my research, but was also the kind of experience I had dreamed of when I first wanted to become a biologist. I am extremely grateful to Belize Fisheries Department (especially the Caye Caulker office), FAMRACC, Chris Freeman, the summer 2014 Goshen crew from University of Florida and Old Dominion University, the crew of the Antea, the Pacotilles expedition team and Maycol Madrid. Thank

you to the many tourist diving operations that have supported me through discounted dives and local sponge knowledge in my quest for samples: thanks to Frenchies and Scuba Sensations in Caye Caulker, ReefCI in Tom Owens Caye, and La Buga in Bocas del Toro. I am eternally grateful to my dive instructor and friend Simon Exley for his patience, skill and generosity in training me – I am a far better diver for all of it. Thanks also to my original Manchester dive instructor Jay Washington for turning me into a real UK diver. Muchas gracias to Mariana Villalba de la Pena and Miguel Gómez for their help and patience in translating my permit applications to Spanish. Thanks to Paul Fletcher for his assistance with R.

My friends have been a wonderful support network throughout my PhD studies. I cannot name you all, but you know who you are – thanks for everything. Special thanks to Rachael Antwis, who has always been there with wise words in a crisis, and to Friederike Clever and Jess Lea – it's been a pleasure to go through the PhD journey together with you. Finally, and most of all, I thank my family for their endless support and belief. My parents Janet and John Griffiths have been there every step of the way in my education and it is hard to imagine that I could be in this position without their love and encouragement. Big thanks to Matthew Griffiths and Megan Finney, who proofread this entire thesis between them and have keenly listened to many sponge facts over the years. Andrew Fletcher has been a rock through my PhD – it truly would not be possible to list everything he has done to support me.

This thesis is dedicated to my grandpa Norman Rae, whose love of nature inspired my own.

Chapter 1: Thesis Introduction

Human development has changed the planet and its natural systems to the extent that a new geological epoch has been recognised the Anthropocene (Crutzen, 2006). This influence has come at great cost to ecosystems across the globe, and species are being lost at a magnitude and rate that suggests we are currently experiencing the sixth mass extinction event (Barnosky et al., 2011; Ceballos et al., 2015; Ceballos, Ehrlich & Dirzo, 2017). Biodiversity, defined as the variability among living organisms, including within species, among species, and of ecosystems (as described by the Convention on Biological Diversity, 1992), is declining rapidly, causing associated declines in ecosystem service provision. These losses in ecosystem services (defined as the benefits people obtain from ecosystems; Millennium Assessment 2005) are cause for serious concern for human health and wellbeing (Díaz et al., 2006; Cardinale et al., 2012). Conservation of all aspects of biodiversity is vital in halting the declines and restoring ecosystem services, but efforts are hampered by limits in our understanding. This ranges from an incomplete knowledge of Earth's biodiversity as it stands (with an estimated 86% of terrestrial species and 91% of marine species undescribed, Mora et al., 2011) to a lack of understanding of species ecology.

Molecular ecology tools in conservation

Advances in molecular biology have revolutionised the fields of conservation biology and ecology (Haig, 1998; Ekblom & Galindo, 2011). DNA sequencing technology developments have allowed biologists to expand research to

address previously inaccessible questions and gain important new insights to be used in conservation science and practice. Sanger sequencing, developed in the 1970s, was the first breakthrough in DNA sequencing, then just over 10 years ago, next-generation sequencing (NGS) technologies were developed and commercialised in response to the demands of the Human Genome Project. These high-throughput methods continue to develop in their output capacity whilst reducing in cost, making study in genetics possible for non-model organisms, and providing ecologists and conservationists with new tools and insights (DeSalle & Amato, 2004; Allendorf, Hohenlohe & Luikart, 2010).

Population genetics

The study of population dynamics has benefitted greatly from the development and increasing accessibility of molecular techniques. Population genetics details the microevolutionary processes, genetic diversity, and genetic structure of populations, and is critical for the understanding of both ecology and evolution of species. The fundamental Hardy-Weinberg principle was described in 1908, but population genetics as a research field gained traction in the 1930s, driven by major theoretical works by Ronald Fisher (1930), Sewell Wright (1931) and J.B.S Haldane (1932). During this period, known as the Modern Synthesis, important developments in the understanding of evolution occurred, as Mendelian genetics and Darwinian evolution could be definitively linked through processes occurring at the population level. Much of the mathematical theory developed during this time, including Wright's *F*-statistics, remain at the heart of modern population genetics. However, the arrival of DNA sequencing and the NGS revolution have unlocked a wealth of potential to study these processes in natural populations and apply genetics knowledge to conservation (Pearse & Crandall, 2004; Allendorf, 2017).

Population genetics is an important subdiscipline of conservation genetics that elucidates patterns and processes that are important to understand for successful conservation (DeSalle & Amato, 2004). In fragmented habitats (both due to natural patchiness such as in coral reef ecosystems and through destruction by anthropogenic means), populations can be connected to some

degree through dispersal or migration. Even for species in a seemingly continuous habitat, populations can be structured by historical and contemporary restrictions to gene flow (Reeb & Avise, 1990; Cabe et al., 2007; Rosel, Hansen & Hohn, 2009). Assessing the extent of these connectivity patterns and species genetic population structure can greatly aid conservation practice. Protected areas and restoration or translocation schemes can be designed to enhance resilience and maintain genetic diversity by preserving connectivity pathways between metapopulations (Palumbi, 2003; Almany et al., 2009). Similarly, knowledge of source-sink interactions between sites can allow the protection of important sites that produce emigrants for areas that experience low self-recruitment. The process of dispersal can be thought to weaken local adaptation, but conversely, the recruitment of emigrants to a population can also increase genetic diversity and aid its resilience to disturbances (Garant, Forde & Hendry, 2007; Frankham, 2015). Understanding natural population processes using genetics can help to reach a balanced equilibrium between these factors to maximise the success of applied conservation practices.

Reductions in the sizes of populations change their genetic properties. When populations undergo dramatic size reductions in a short time frame (due to, for example, sudden environmental changes or human activity), genetic diversity can also be rapidly lost, termed a bottleneck (Sbordoni et al., 1986; Bellinger et al., 2003; Bristol et al., 2013). Small populations with low genetic diversity are then vulnerable to future perturbations due to a lack of adaptive capacity. Small population sizes also mean that inbreeding levels can increase, causing recessive deleterious alleles to be exposed in homozygotes. Subsequent losses in fitness (inbreeding depression) then further damage the survival prospects of a population (Whitlock, 2000; Reed & Frankham, 2003; O'Grady et al., 2006; Charlesworth et al., 2009). In addition, in small populations the stochastic effects of genetic drift can influence allele frequencies more than the opposing force of selection, preventing local adaptation (Lande, 1976). Managing genetic diversity in small, vulnerable populations is therefore critical, and as such this

issue affects captive population management, reintroduction, restoration, in-situ species protection and protected area design.

Restoration genetics

The importance of genetic approaches in conservation is recognised by conservation biologists and practitioners, but the application of theory to practice can be limited (Shafer et al., 2015; Taylor, Dussex & van Heezik, 2017). Ecosystem restoration is an important conservation technique that can benefit greatly from genetic information. An umbrella term for a suite of potential actions and goals, 'restoration' generally involves an attempt to return an area that has been degraded by human action back to a previous state through human intervention, often with the hope of increasing biodiversity and enabling a return of ecosystem services (Corlett, 2016). Restoration ecology as a scientific research discipline emerged as a means to increase restoration success through the application of ecological theory to restoration practices and the scientific evaluation of their effectiveness (Bradshaw, 1993; Choi, 2007). As modern sequencing technology progresses and increases accessibility to genetic study for non-model organisms, the subfields of restoration genetics and restoration genomics have arisen and gained importance (Williams, Nevill & Krauss, 2014). This recognises the role that genetics and genetic diversity play in adaptation and resilience, and promotes the use of genetic tools in enabling predictions of demographic processes and selecting suitable source material for reintroductions (Baums, 2008; Mijangos et al., 2015).

Understanding patterns of population connectivity and levels of self-recruitment is important in the design of an efficient and robust restoration programme. Examining the extent of gene flow spatially and temporally can help to predict levels of migration and recruitment among restoration and natural sites, as well helping to identify barriers to dispersal and sites at risk of poor genetic health (Raeymaekers et al., 2008; van Strien et al., 2014; Coates et al., 2014; Bertrand et al., 2016). This can aid in identifying and designing appropriate restoration sites that not only can be sustained long-term through natural processes, but may also connect isolated populations to restore natural gene flow and maintain

genetic diversity in the species. Following restoration, molecular techniques can then be used to monitor genetic health of a population, such as identifying potential issues including inbreeding and founder effects (Cosentino et al., 2015), as well as examining how restoration influences demographic processes in general among local populations (Reynolds, Waycott & McGlathery, 2013).

Genetic tools can aid in the selection of restoration candidate organisms through the prediction of the scale of local adaptation, and thus the selection of individuals with favourable characters in the chosen restoration sites (Pickup et al., 2012; Jørgensen et al., 2016). However, local adaptation cannot be assumed without rigorous testing, as the degree of local adaptation can vary among populations (Gibson et al., 2016) and climate change and habitat fragmentation may render locally-sourced organisms maladapted (Gellie et al. 2016). It is therefore important to evaluate these factors to ensure the most suitable well-adapted source material is chosen for the current environmental characteristics at the restoration site.

Conversely, genetic diversity in a population can allow for multiple responses to stressors (Hanley et al., 2016; Rothäusler et al., 2016), and as such, selection of a genetically diverse founder population may be important for enhancing resilience of the restored area to perturbations and stressors (Ehlers, Worm & Reusch, 2008). Beyond the health and adaptive potential of the population, genetic diversity in foundation species can have community and ecosystem-level effects through the expression of a range of phenotypes, but the link between genotype and extended phenotype effects in the community is generally not as well recognised. Nonetheless, genetic diversity in foundation species can influence community composition and dynamics (Bangert et al., 2005; Lau et al., 2016), ecosystem processes (Rudman et al., 2015; Fitzpatrick et al., 2015; Salo & Gustafsson, 2016) and may ultimately enhance ecosystem service delivery (Hajjar, Jarvis & Gemmill-Herren, 2008; Bailey, 2011; Reynolds, McGlathery & Waycott, 2012). Accordingly, increasing genetic diversity whilst managing for potential outbreeding depression and adaptation should be a key consideration for restoration managers.

Study group: Marine sponges (Porifera)

In this thesis, I develop and use molecular resources to explore many of the concepts discussed above in marine sponges in the Greater Caribbean. Sponges are basal metazoans of the phylum Porifera that inhabit marine and freshwater ecosystems across the globe. Sponges have a simple body plan, and most of the 8, 873 species of sponges described (Van Soest et al., 2017) derive their nutrition from filter feeding (with the notable exception of the carnivorous sponges, and those which also derive nutrients from symbiotic microbes). This simple way of making a living has allowed sponges to inhabit aquatic ecosystems from the poles to the tropics, and the intertidal zones to the deep sea (Van Soest et al., 2012). It also makes them key in nutrient cycling and bentho-pelagic coupling (Maldonado et al., 2005; Jiménez & Ribes, 2007; de Goeij et al., 2013). Sponges are also important in the provision of habitat, both as microhabitats (Pearse, 1950; Wendt, Van Dolahh & O'Rourke, 1985; Fiore & Jutte, 2010) and as the main providers of habitat structure in some ecosystems, including sponge gardens or reefs (Whitney et al., 2005; Schönberg & Fromont, 2012; Maldonado et al., 2015). They are a prey item for many species, including invertebrates, fish and turtles (Dunlap & Pawlik, 1998; León & Bjorndal, 2002; Powell et al., 2015). They also have numerous effects on the substrate, including bioerosion (Rützler, 1975; Nava & Carballo, 2008) and rubble binding and stabilisation (Wulff, 1984; Rasser & Riegl, 2002). Sponges are also host to dense, diverse and very specific microbial communities living within their tissues (Taylor et al., 2007; Schmitt et al., 2012; Thomas et al., 2016) . The role of the microbiome in sponge biology, and the evolutionary and ecological relationship between sponge and microbiome is still being uncovered, but it is clear that it is complex, multifaceted and fundamentally interconnected (Freeman et al., 2013; Webster & Thomas, 2016; Moitinho-Silva et al., 2017).

As well as their role in the ecosystem functioning, sponges provide ecosystem services and associated economic benefits for humans. They (or their microbial symbionts) are the source of chemical compounds with biotechnological and

pharmaceutical application (Osinga et al., 2001; Wang, 2006; Laport, Santos & Muricy, 2009). It has been proposed that they may mitigate aquatic pathogen transmission though their filter feeding process (Longo et al., 2016; Burge et al., 2016). 'Bath' sponges have been harvested for centuries for their cleaning properties, and industry still exists today for wild harvesting or aquaculture of sponges (Pronzato & Manconi, 2008; Oronti et al., 2012). Due to the breadth of their ecological roles and their importance for human wellbeing, understanding sponge biology and ecology is an important avenue of research in order to help conserve and manage them.

Numerous threats face marine ecosystems, including overfishing (Jackson et al., 2001; Scheffer, Carpenter & Young, 2005), ocean warming and acidification (Hoegh-Guldberg et al., 2007; Cantin et al., 2010), destructive fishing techniques (Thrush & Dayton, 2002; Biju Kumar & Deepthi, 2006) and pollution (Derraik, 2002; Peterson et al., 2003; Naser, 2013). The decline in coral reef health and coral cover is well documented (Pandolfi et al., 2003; Bruno & Selig, 2007; De'ath et al., 2012), but less is known about the future of sponges in this changing world and their conservation status as a group globally (Bell et al., 2015). Some research suggests sponges may be 'winners' in some climate change scenarios as they are more resilient than corals to various stressors, and may be favoured in a regime shift, forming sponge reefs (Norström et al., 2009; Bell et al., 2013). However, mass mortalities in sponge communities have been recorded in the Caribbean (Galtsoff et al., 1939; Vicente, 1989; Wulff, 2006), Florida (Butler et al., 1995; Stevely et al., 2010), and the Mediterranean (Stachowitsch, 1984; Perez et al., 2000; Garrabou et al., 2009; Cebrian et al., 2011; Stabili et al., 2012), with variable causes including pathogens, ocean warming and cyanobacterial blooms. Sponge disease has been observed in numerous areas and species (reviewed by Webster, 2007), sometimes affecting commercial sponge fisheries (Smith, 1941; Pronzato, 1999), and has been suggested to be a threat to the health and ecology of reef systems globally (Webster, 2007). Additionally, climate change scenarios can cause shifts in the associated microbial community of sponges (Lesser et al., 2016), which when out of natural balance may cause disease (Rützler, 1988).

Despite their importance, much is still unknown about sponges. In the Caribbean, sponges are highly abundant and diverse, and comparatively well studied in terms of their biodiversity (Bell et al., 2015) (although new species from the area are regularly described; for example Zea and López-Victoria 2016; Zea and Pulido 2016) - however, knowledge of the ecology of individual species is lacking for most species. The two species I studied in this thesis are *Ircinia campana* (Lamarck, 1814) and *Spheciospongia vesparium* (Lamarck, 1815), which are common across the Greater Caribbean region with wide distributions in the Tropical Western Atlantic. These species were chosen due to their importance and vulnerability in hard bottom habitats in Florida Bay, Florida, USA.

Restoration of *Ircinia campana* and *Spheciospongia vesparium* in Florida Bay, USA

In Florida Bay's nearshore, shallow hard bottom habitats (characterised by a calcium carbonate-based substratum topped by a thin sediment layer), sponges, along with octocorals and macroalgae, form a dominant part of the benthos (Chiappone & Sullivan, 1994; Tellier & Bertelsen, 2008). Together with a network of mangrove and seagrass habitats, hard bottom is an important nursery area for reef species, including the commercially fished Caribbean spiny lobster, *Panulirus argus* (Bertelsen et al., 2009). However, since the late 1980s, Florida Bay has suffered a series of ecological disturbances with significant ecosystem-level consequences. This includes substantial die-offs of turtle grass (*Thalassia testudinum*) (Robblee et al., 1991; Hall et al., 2016), cyanobacterial blooms (Fourqurean & Robblee, 1999; Berry et al., 2015) and mass mortalities in sponge communities (Butler et al., 1995; Stevely et al., 2010). The direct cause of sponge mortality is not fully understood, but is correlated with occurrence of the blooms.

Sponge community declines in the area have been associated with lower local abundances and shifting habitat use in invertebrates (Butler et al., 1995; Herrnkind et al., 1997), and diminished soundscapes indicative of a loss of sponge epibionts and endosymbionts including snapping shrimp *Synalpheus*

(Butler, Stanley & Butler, 2016). Abundances of fish and invertebrates are likely to be further impacted by this change in soundscape, as acoustic cues are used in recruitment and larval settlement. Furthermore, the loss of these main suspension filter feeders is thought to have caused increased susceptibility of the area to further phytoplankton blooms (Peterson et al., 2006). For these reasons, a restoration program has been implemented for the most dominant sponges in Florida Bay, led by Mark Butler (Old Dominion University), and Donald Behringer (University of Florida). The program, which has involved volunteers from the local community, transplants whole sponges or sponge cuttings from areas in the Bay where sponge populations have been unaffected to sites where mass mortalities have occurred. Two species included in restoration efforts are *Spheciospongia vesparium* (loggerhead sponge; Clionaida, Clionaidae) and *Ircinia campana* (vase sponge; Dictyoceratida, Irciniidae) (Figure 1). These are large demosponges whose total biomass forms a significant component of the total biogenic structure of the area (Chiappone & Sullivan, 1994; Tellier & Bertelsen, 2008; Bertelsen et al., 2009). In two chapters of this thesis, I study the population genetics of these two species in Florida Bay as well as through locations in the Greater Caribbean. In the General Discussion, I discuss these results in the context of their restoration, and make practical recommendations for the restoration program based on my findings.

Figure 1: *Spheciospongia vesparium* (left) and *Ircinia campana* (right), two sponge species undergoing restoration in Florida Bay (USA) and the subject of part of this thesis (photographs: D. Behringer).

Thesis aims and chapters

The aims of this thesis are as follows:

1) The development of molecular tools and methods to aid ecological and conservation research in sponges:
    a. The creation of a microsatellite development bioinformatics pipeline
    b. The characterisation of microsatellite markers for ecologically important and vulnerable sponges, *Ircinia campana* and *Spheciospongia vesparium*

2) To study population genetics in *Ircinia campana* and *Spheciospongia vesparium* in the Greater Caribbean, and to make restoration recommendations based on the population genetic patterns.

3) To explore the effects of genetic variation in *Ircinia campana* hosts on their associated microbial communities

This thesis is composed of four data chapters:

In Chapter 2, I present a bioinformatics pipeline for the processing raw Illumina next-generation sequence data for microsatellite marker development. The contents of this chapter have been published in the journal *Conservation Genetics Resources*.

In Chapter 3, I develop microsatellite markers for *Spheciospongia vesparium* and use them to describe the population genetic structure of the species within Florida Bay and the northern Caribbean.

In Chapter 4, I develop microsatellite markers for *Ircinia campana* and use them to describe connectivity and genetic structure over a range of spatial scales in the Greater Caribbean and Florida Bay.

In Chapter 5, I go on to explore the ecological effects of genetic diversity in sponge populations. To do so, I investigate how microbiome variability in *Ircinia campana* is affected by host genotype.

REFERENCES

Allendorf FW. 2017. Genetics and the conservation of natural populations: allozymes to genomes. *Molecular Ecology* 26:420–430. DOI: 10.1111/mec.13948.

Allendorf FW., Hohenlohe P a., Luikart G. 2010. Genomics and the future of conservation genetics. *Nature Reviews Genetics* 11:697–709. DOI: 10.1038/nrg2844.

Almany GR., Connolly SR., Heath DD., Hogan JD., Jones GP., McCook LJ., Mills M., Pressey RL., Williamson DH. 2009. Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs. *Coral Reefs* 28:339–351. DOI: 10.1007/s00338-009-0484-x.

Bailey JK. 2011. From genes to ecosystems: A genetic basis to ecosystem services. *Population Ecology* 53:47–52. DOI: 10.1007/s10144-010-0251-4.

Bangert R., Turek R., Martinsen G., Wimp G., Bailey J., Whitham T. 2005. Benefits of conservation of plant genetic diversity to arthropod diversity. *Conservation Biology* 19:379–390. DOI:10.1111/j.1523-1739.2005.00450.x

Barnosky AD., Matzke N., Tomiya S., Wogan GOU., Swartz B., Quental TB., Marshall C., Mcguire JL., Lindsey EL., Maguire KC., Mersey B., Ferrer EA. 2011. Has the Earth's sixth mass extinction already arrived? *Nature* 471:51–57. DOI: 10.1038/nature09678.

Baums IB. 2008. A restoration genetics guide for coral reef conservation. *Molecular Ecology* 17:2796–2811. DOI: 10.1111/j.1365-294X.2008.03787.x.

Bell JJ., Davy SK., Jones T., Taylor MW., Webster NS. 2013. Could some coral reefs become sponge reefs as our climate changes? *Global Change Biology* 19:2613–24. DOI: 10.1111/gcb.12212.

Bell JJ., Mcgrath E., Biggerstaff A., Bates T., Cárdenas CA., Bennett H. 2015. Global conservation status of sponges. *Conservation Biology* 29:42–53. DOI: 10.1111/cobi.12447.

Bellinger MR., Johnson JA., Toepfer J., Dunn P. 2003. Loss of genetic variation in Greater Prairie Chickens following a population bottleneck in Wisconsin, U.S.A. *Conservation Biology* 17:717–724. DOI: 10.1046/j.1523-1739.2003.01581.x.

Berry DL., Goleski JA., Koch F., Wall CC., Peterson BJ., Anderson OR., Gobler CJ. 2015. Shifts in cyanobacterial strain dominance during the onset of harmful algal blooms in Florida Bay, USA. *Microbial Ecology* 70:361–371. DOI: 10.1007/s00248-014-0564-5.

Bertelsen RD., Butler MJ., Herrnkind WF., Hunt JH. 2009. Regional characterisation of hard-bottom nursery habitat for juvenile Caribbean spiny lobster (*Panulirus argus*) using rapid assessment techniques. *New Zealand Journal of Marine and Freshwater Research* 43:299–312. DOI: 10.1080/00288330909510002.

Bertrand KN., VanDeHey JA., Pilger TJ., Felts EA., Turner TF. 2016. Genetic structure of a disjunct peripheral population of mountain sucker *P. jordani* in the Black Hills, South Dakota, USA. *Conservation Genetics* 17:775–784. DOI: 10.1007/s10592-016-0820-y.

Biju Kumar A., Deepthi GR. 2006. Trawling and by-catch: Implications on marine ecosystem. *Current Science* 90:922–931.

Bradshaw AD. 1993. Restoration Ecology as a Science. *Restoration Ecology* 1:71–73. DOI: 10.1111/j.1526-100X.1993.tb00011.x.

Bristol RM., Tucker R., Dawson DA., Horsburgh G., Prys-Jones RP., Frantz AC., Krupa A., Shah NJ., Burke T., Groombridge JJ. 2013. Comparison of historical bottleneck effects and genetic consequences of re-introduction in a critically endangered island passerine. *Molecular Ecology* 22:4644–4662. DOI: 10.1111/mec.12429.

Bruno JF., Selig ER. 2007. Regional decline of coral cover in the Indo-Pacific: Timing, extent, and subregional comparisons. *PLoS ONE* 2:e711. DOI: 10.1371/journal.pone.0000711.

Burge CA., Closek CJ., Friedman CS., Groner ML., Jenkins CM., Shore-Maggio A., Welsh JE. 2016. The use of filter-feeders to manage disease in a changing world. *Integrative and Comparative Biology* 56:573–587. DOI: 10.1093/icb/icw048.

Butler M., Hunt J., Herrnkind W., Childress M., Bertelsen R., Sharp W., Matthews T., Field J., Marshall H. 1995. Cascading disturbances in Florida Bay, USA: cyanobacteria blooms, sponge mortality, and implications for juvenile spiny lobsters *Panulirus argus*. *Marine Ecology Progress Series* 129:119–125. DOI: 10.3354/meps129119.

Butler J., Stanley JA., Butler MJ. 2016. Underwater soundscapes in near-shore tropical habitats and the effects of environmental degradation and habitat restoration. *Journal of Experimental Marine Biology and Ecology* 479:89–96. DOI: 10.1016/j.jembe.2016.03.006.

Cabe PR., Page RB., Hanlon TJ., Aldrich ME., Connors L., Marsh DM. 2007. Fine-scale population differentiation and gene flow in a terrestrial salamander (*Plethodon cinereus*) living in continuous habitat. *Heredity* 98:53–60. DOI: 10.1038/sj.hdy.6800905.

Cantin NE., Cohen AL., Karnauskas KB., Tarrant AM., McCorkle DC. 2010. Ocean warming slows coral growth in the central Red Sea. *Science* 329:322–325. DOI: 10.1126/science.1190182.

Cardinale BJ., Duffy JE., Gonzalez A., Hooper DU., Perrings C., Venail P., Narwani A., Mace GM., Tilman D., A.Wardle D., Kinzig AP., Daily GC., Loreau M., Grace JB., Larigauderie A., Srivastava DS., Naeem S. 2012. Biodiversity loss and its impact on humanity. *Nature* 489:326–326. DOI: 10.1038/nature11373.

Ceballos G., Ehrlich PR., Barnosky AD., García A., Pringle RM., Palmer TM. 2015. Accelerated modern human–induced species losses: Entering the sixth mass extinction. *Science Advances* 1: e1400253. DOI: 10.1126/sciadv.1400253.

Ceballos G., Ehrlich PR., Dirzo R. 2017. Biological annihilation via the ongoing sixth mass extinction signaled by vertebrate population losses and declines. *Proceedings of the National Academy of Sciences*:201704949. DOI: 10.1073/pnas.1704949114.

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* 6:e20211. DOI: 10.1371/journal.pone.0020211.

Charlesworth D., Charlesworth D., Willis JH., Willis JH. 2009. The genetics of inbreeding

depression. *Nature Reviews Genetics* 10:783–96. DOI: 10.1038/nrg2664.

Chiappone M., Sullivan KM. 1994. Ecological structure and dynamics of nearshore hard-bottom communities in the Florida Keys. *Bulletin of Marine Science* 54:747–756.

Choi YD. 2007. Restoration Ecology to the Future: A Call for New Paradigm. *Restoration Ecology* 15:351–353. DOI: 10.1111/j.1526-100X.2007.00224.x.

Coates JH., Hovel KA., Butler JL., Bohonak AJ. 2014. Recruitment and recovery of pink abalone (*Haliotis corrugata*) in a historically overexploited kelp forest: Are local populations self-sustaining? *Journal of Experimental Marine Biology and Ecology* 460:184–192. DOI: 10.1016/j.jembe.2014.07.004.

Corlett RT. 2016. Restoration, reintroduction, and rewilding in a changing world. *Trends in Ecology & Evolution* 31:453–462. DOI: 10.1016/j.tree.2016.02.017.

Cosentino BJ., Schooley RL., Bestelmeyer BT., McCarthy AJ., Sierzega K. 2015. Rapid genetic restoration of a keystone species exhibiting delayed demographic response. *Molecular Ecology* 24:6120–6133. DOI: 10.1111/mec.13469.

Crutzen PJ. 2006. The anthropocene. In: *Earth System Science in the Anthropocene*. Berlin/Heidelberg: Springer-Verlag, 13–18. DOI: 10.1007/3-540-26590-2_3.

De'ath G., Fabricius KE., Sweatman H., Puotinen M. 2012. The 27-year decline of coral cover on the Great Barrier Reef and its causes. *Proceedings of the National Academy of Sciences* 109:17995–17999. DOI: 10.1073/pnas.1208909109.

Derraik JG. 2002. The pollution of the marine environment by plastic debris: a review. *Marine Pollution Bulletin* 44:842–852. DOI: 10.1016/S0025-326X(02)00220-5.

DeSalle R., Amato G. 2004. The expansion of conservation genetics. *Nature Reviews Genetics* 5:702–12. DOI: 10.1038/nrg1425.

Díaz S., Fargione J., Chapin FS., Tilman D. 2006. Biodiversity loss threatens human well-being. *PLoS Biology* 4:e277. DOI: 10.1371/journal.pbio.0040277.

Dunlap M., Pawlik JR. 1998. Spongivory by parrotfish in Florida mangrove and reef habitats. *Marine Ecology* 19:325–337. DOI: 10.1111/j.1439-0485.1998.tb00471.x.

Ehlers A., Worm B., Reusch TBH. 2008. Importance of genetic diversity in eelgrass *Zostera marina* for its resilience to global warming. *Marine Ecology Progress Series* 355:1–7. DOI: 10.3354/meps07369.

Ekblom R., Galindo J. 2011. Applications of next generation sequencing in molecular ecology of non-model organisms. *Heredity* 107:1–15. DOI: 10.1038/hdy.2010.152.

Fiore CL., Jutte PC. 2010. Characterization of macrofaunal assemblages associated with sponges and tunicates collected off the southeastern United States. *Invertebrate Biology* 129:105–120. DOI: 10.1111/j.1744-7410.2010.00184.x.

Fisher RA. 1930. *The Genetical Theory of Natural Selection.* Oxford: Clarendon Press. DOI:10.5962/bhl.title.27468

Fitzpatrick CR., Agrawal AA., Basiliko N., Hastings AP., Isaac ME., Preston M., Johnson MTJ. 2015. The importance of plant genotype and contemporary evolution for terrestrial ecosystem processes. *Ecology* 96:2632–2642. DOI: 10.1890/14-2333.1.

Fourqurean JW., Robblee MB. 1999. Florida Bay: A history of recent ecological changes. *Estuaries* 22:345–357. DOI: 10.2307/1353203.

Frankham R. 2015. Genetic rescue of small inbred populations: meta-analysis reveals large and consistent benefits of gene flow. *Molecular Ecology* 24:2610–2618. DOI: 10.1111/mec.13139.

Freeman C., Thacker R., Baker D., Fogel M. 2013. Quality or quantity: is nutrient transfer driven more by symbiont identity and productivity than by symbiont abundance? *The ISME Journal* 7:1116–1125. DOI: 10.1038/ismej.2013.7.

Galtsoff PS., Brown HH., Smith CL., Walton Smith FG. 1939. Sponge mortality in the Bahamas. *Nature* 143:807–808. DOI:10.1038/143807a0.

Garant D., Forde SE., Hendry AP. 2007. The multifarious effects of dispersal and gene flow on contemporary adaptation. *Functional Ecology* 21:434–443. DOI: 10.1111/j.1365-2435.2006.01228.x.

Garrabou J., Coma R., Bensoussan N., Bally M., Chevaldonné P., Cigliano M., Diaz D., Harmelin J-G., Gambi MC., Kersting DK., Ledoux J., Lejeusne C., Pérez T., Ribes M., Romano J-C., Serrano E., Linares C., Marschal C., Teixido N., Torrents O., Zabala M., Zuberer F., Cerrano C. 2009. Mass mortality in Northwestern Mediterranean rocky benthic communities : effects of the 2003 heat wave. *Global Change Biology* 15:1090–1103. DOI: 10.1111/j.1365-2486.2008.01823.x

Gibson AL., Espeland EK., Wagner V., Nelson CR. 2016. Can local adaptation research in plants inform selection of native plant materials? An analysis of experimental methodologies. *Evolutionary Applications* 9:1219–1228. DOI: 10.1111/eva.12379.

de Goeij JM., van Oevelen D., Vermeij MJA., Osinga R., Middelburg JJ., de Goeij AFPM., Admiraal W. 2013. Surviving in a marine desert: The sponge loop retains resources within coral reefs. *Science* 342:108–110. DOI: 10.1126/science.1241981.

Haig SM. 1998. Molecular contributions to conservation. *Ecology* 79:413–425. DOI: 10.1890/0012-9658(1998)079[0413:MCTC]2.0.CO;2.

Hajjar R., Jarvis DI., Gemmill-Herren B. 2008. The utility of crop genetic diversity in maintaining ecosystem services. *Agriculture, Ecosystems & Environment* 123:261–270. DOI: 10.1016/j.agee.2007.08.003.

Haldane JB. 1932. *The Causes of Evolution*. London: Longmans, Green & Co.

Hall MO., Furman BT., Merello M., Durako MJ. 2016. Recurrence of *Thalassia testudinum* seagrass die-off in Florida Bay, USA: Initial observations. *Marine Ecology Progress Series* 560:243–249. DOI: 10.3354/meps11923.

Hanley TC., Hughes AR., Williams B., Garland H., Kimbro DL. 2016. Effects of intraspecific diversity on survivorship, growth, and recruitment of the eastern oyster across sites. *Ecology* 97:1518–1529. DOI: 10.1890/15-1710.1.

Herrnkind WF., Butler IV MJ., Hunt JH., Childress M. 1997. Role of physical refugia: implications from a mass sponge die-off in a lobster nursery in Florida. *Marine and Freshwater Research* 48:759. DOI: 10.1071/MF97193.

Hoegh-Guldberg O., Mumby PJ., Hooten AJ., Steneck RS., Greenfield P., Gomez E., Harvell CD., Sale PF., Edwards AJ., Caldeira K., Knowlton N., Eakin CM., Iglesias-Prieto R., Muthiga N., Bradbury RH., Dubi A., Hatziolos ME. 2007. Coral reefs under rapid climate change and ocean acidification. *Science* 318:1737–42. DOI: 10.1126/science.1152509.

Jackson JBC., Kirby MX., Berger WH., Bjorndal KA., Botsford LW., Bourque BJ., Bradbury RH., Cooke R., Erlandson J., Estes JA., Hughes TP., Kidwell S., Lange CB., Lenihan HS., Pandolfi JM., Peterson CH., Steneck RS., Tegner MJ., Warner RR. 2001. Historical overfishing and the recent collapse of coastal ecosystems. *Science* 293:629–637. DOI: 10.1126/science.1059199.

Jiménez E., Ribes M. 2007. Sponges as a source of dissolved inorganic nitrogen: Nitrification mediated by temperate sponges. *Limnology and Oceanography* 52:948–958. DOI: 10.4319/lo.2007.52.3.0948.

Jørgensen MH., Elameen A., Hofman N., Klemsdal S., Malaval S., Fjellheim S. 2016. What's the meaning of local? Using molecular markers to define seed transfer zones for ecological restoration in Norway. *Evolutionary Applications* 9:673–684. DOI: 10.1111/eva.12378.

Lande R. 1976. Natural selection and random genetic drift in phenotypic evolution. *Evolution* 30:314. DOI: 10.2307/2407703.

Laport MS., Santos OCS., Muricy G. 2009. Marine sponges: potential sources of new antimicrobial drugs. *Current Pharmaceutical Biotechnology* 10:86–105. DOI: 10.2174/138920109787048625.

Lau MK., Keith AR., Borrett SR., Shuster SM., Whitham TG. 2016. Genotypic variation in foundation species generates network structure that may drive community dynamics and evolution. *Ecology*. DOI: 10.1890/15-0600.

León YM., Bjorndal KA. 2002. Selective feeding in the hawksbill turtle, an important predator in coral reef ecosystems. *Marine Ecology Progress Series* 245:249–258. DOI: 10.3354/meps245249.

Lesser MP., Fiore C., Slattery M., Zaneveld J. 2016. Climate change stressors destabilize the microbiome of the Caribbean barrel sponge, *Xestospongia muta. Journal of Experimental Marine Biology and Ecology* 475:11–18. DOI: 10.1016/j.jembe.2015.11.004.

Longo C., Cardone F., Corriero G., Licciano M., Pierri C., Stabili L. 2016. The co-occurrence of the demosponge Hymeniacidon perlevis and the edible mussel *Mytilus galloprovincialis* as a new tool for bacterial load mitigation in aquaculture. *Environmental Science and Pollution Research* 23:3736–3746. DOI: 10.1007/s11356-015-5587-z.

Maldonado M., Aguilar R., Bannister RJ., Bell JJ., Conway KW., Dayton PK., Díaz C., Gutt J., Kelly M., Kenchington ELR., Leys SP., Pomponi SA., Rapp HT., Rützler K., Tendal OS., Vacelet J., Young CM. 2015. Sponge grounds as key marine habitats: A synthetic review of types, structure, functional roles, and conservation concerns. In: *Marine Animal Forests*. Cham: Springer International Publishing, 1–39. DOI: 10.1007/978-3-319-17001-5_24-1.

Maldonado M., Carmona MC., Velásquez Z., Puig A., Cruzado A., López A., Young CM. 2005.

Siliceous sponges as a silicon sink: An overlooked aspect of benthopelagic coupling in the marine silicon cycle. *Limnology and Oceanography* 50:799–809. DOI: 10.4319/lo.2005.50.3.0799.

Mijangos JL., Pacioni C., Spencer PBS., Craig MD. 2015. Contribution of genetics to ecological restoration. *Molecular Ecology* 24:22–37. DOI: 10.1111/mec.12995.

Moitinho-Silva L., Díez-Vives C., Batani G., Esteves AI., Jahn MT., Thomas T. 2017. Integrated metabolism in sponge–microbe symbiosis revealed by genome-centered metatranscriptomics. *The ISME Journal* 11:1651–1666. DOI: 10.1038/ismej.2017.25.

Mora C., Tittensor DP., Adl S., Simpson AGB., Worm B. 2011. How many species are there on Earth and in the ocean? *PLoS Biology* 9:e1001127. DOI: 10.1371/journal.pbio.1001127.

Naser HA. 2013. Assessment and management of heavy metal pollution in the marine environment of the Arabian Gulf: A review. *Marine Pollution Bulletin* 72:6–13. DOI: 10.1016/j.marpolbul.2013.04.030.

Nava H., Carballo JL. 2008. Chemical and mechanical bioerosion of boring sponges from Mexican Pacific coral reefs. *Journal of Experimental Biology* 211: 2827-2831. DOI: 10.1242/jeb.019216.

Norström A., Nyström M., Lokrantz J., Folke C. 2009. Alternative states on coral reefs: beyond coral-macroalgal phase shifts. *Marine Ecology Progress Series* 376:295–306. DOI: 10.3354/meps07815.

O'Grady JJ., Brook BW., Reed DH., Ballou JD., Tonkyn DW., Frankham R. 2006. Realistic levels of inbreeding depression strongly affect extinction risk in wild populations. *Biological Conservation* 133:42–51. DOI: 10.1016/j.biocon.2006.05.016.

Oronti A., Danylchuk AJ., Elmore CE., Auriemma R., Pesle G. 2012. Assessing the feasibility of sponge aquaculture as a sustainable industry in The Bahamas. *Aquaculture International* 20:295–303. DOI: 10.1007/s10499-011-9457-5.

Osinga R., Armstrong E., Burgess JG., Hoffmann F., Reitner J., Schumann-Kindel G. 2001. Sponge – microbe associations and their importance for sponge bioprocess engineering. *Hydrobiologia* 461:55–62. DOI: doi.org/10.1023/A:1012717200362.

Palumbi S. 2003. Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications* 13:146–158. DOI: 10.1890/1051-0761(2003)013[0146:PGDCAT]2.0.CO;2.

Pandolfi JM., Bradbury RH., Sala E., Hughes TP., Bjorndal KA., Cooke RG., McArdle D., McClenachan L., Newman MJH., Paredes G., Warner RR., Jackson JBC. 2003. Global trajectories of the long-term decline of coral reef ecosystems. *Science* 301:955–958. DOI: 10.1126/science.1085706.

Pearse AS. 1950. Notes on the inhabitants of certain sponges at Bimini. *Ecology* 31:149–151. DOI: 10.2307/1931369.

Pearse DE., Crandall KA. 2004. Beyond FST : Analysis of population genetic data for conservation. *Conservation Genetics* 5:585–602. DOI: 10.1007/s10592-003-1863-4.

Perez T., Garrabou J., Sartoretto S., Harmelin JG., Francour P., Vacelet J. 2000. Massive

mortality of marine invertebrates: an unprecedented event in northwestern Mediterranean. *Comptes Rendus de l'Academie des Sciences. Serie III, Sciences de la Vie* 323:853–65. DOI: 10.1016/S0764-4469(00)01237-3.

Peterson B., Chester C., Jochem F., Fourqurean JW. 2006. Potential role of sponge communities in controlling phytoplankton blooms in Florida Bay. *Marine Ecology Progress Series* 328:93–103. DOI: doi:10.3354/meps328093.

Peterson CH., Rice SD., Short JW., Esler D., Bodkin JL., Ballachey BE., Irons DB. 2003. Long-term ecosystem response to the Exxon Valdez oil spill. *Science* 302: 2082-2086. DOI: 10.1126/science.1084282.

Pickup M., Field DL., Rowell DM., Young AG. 2012. Predicting local adaptation in fragmented plant populations: implications for restoration genetics. *Evolutionary Applications* 5:913–924. DOI: 10.1111/j.1752-4571.2012.00284.x.

Powell A., Jones T., Smith DJ., Jompa J., Bell JJ. 2015. Spongivory in the Wakatobi Marine National Park, Southeast Sulawesi, Indonesia 1. *Pacific Science* 69:487–508. DOI: 10.2984/69.4.5.

Pronzato R. 1999. Sponge-fishing, disease and farming in the Mediterranean Sea. *Aquatic Conservation: Marine and Freshwater Ecosystems* 9:485–493. DOI: 10.1002/(SICI)1099-0755(199909/10)9:5<485::AID-AQC362>3.0.CO;2-N.

Pronzato R., Manconi R. 2008. Mediterranean commercial sponges: over 5000 years of natural history and cultural heritage. *Marine Ecology* 29:146–166. DOI: 10.1111/j.1439-0485.2008.00235.x.

Raeymaekers JAM., Maes GE., Geldof S., Hontis I., Nackaerts K., Volckaert FAM. 2008. Modeling genetic connectivity in sticklebacks as a guideline for river restoration. *Evolutionary Applications* 1:475–488. DOI: 10.1111/j.1752-4571.2008.00019.x.

Rasser M., Riegl B. 2002. Holocene coral reef rubble and its binding agents. *Coral Reefs* 21:57–72. DOI: 10.1007/s00338-001-0206-5.

Reeb CA., Avise JC. 1990. A genetic discontinuity in a continuously distributed species: mitochondrial DNA in the American oyster, *Crassostrea virginica*. *Genetics* 124: 397-406.

Reed DH., Frankham R. 2003. Correlation between fitness and genetic diversity. *Conservation Biology* 17:230–237. DOI: 10.1046/j.1523-1739.2003.01236.x.

Reynolds L., McGlathery K., Waycott M. 2012. Genetic diversity enhances restoration success by augmenting ecosystem services. *PLoS ONE* 7:1–7. DOI: 10.1371/journal.pone.0038397.

Reynolds LK., Waycott M., McGlathery KJ. 2013. Restoration recovers population structure and landscape genetic connectivity in a dispersal-limited ecosystem. *Journal of Ecology* 101:1288–1297. DOI: 10.1111/1365-2745.12116.

Robblee M., Barber T., Carlson P., Durako M., Fourqurean J., Muehlstein L., Porter D., Yarbro L., Zieman R., Zieman J. 1991. Mass mortality of the tropical seagrass *Thalassia testudinum* in Florida Bay (USA). *Marine Ecology Progress Series* 71:297–299. DOI: 10.3354/meps071297.

Rosel PE., Hansen L., Hohn AA. 2009. Restricted dispersal in a continuously distributed marine species: common bottlenose dolphins *Tursiops truncatus* in coastal waters of the western North Atlantic. *Molecular Ecology* 18:5030–5045. DOI: 10.1111/j.1365-294X.2009.04413.x.

Rothäusler E., Sjöroos J., Heye K., Jormalainen V. 2016. Genetic variation in photosynthetic performance and tolerance to osmotic stress (desiccation, freezing, hyposalinity) in the rocky littoral foundation species *Fucus vesiculosus* (Fucales, Phaeophyceae). *Journal of Phycology* 52:877–887. DOI: 10.1111/jpy.12455.

Rudman SM., Rodriguez-Cabal MA., Stier A., Sato T., Heavyside J., El-Sabaawi RW., Crutsinger GM. 2015. Adaptive genetic variation mediates bottom-up and top-down control in an aquatic ecosystem. *Proceedings of the Royal Society B: Biological Sciences* 282:20151234. DOI: 10.1098/rspb.2015.1234.

Rützler K. 1975. The role of burrowing sponges in bioerosion. *Oecologia* 19:203–216. DOI: 10.1007/BF00345306.

Rützler K. 1988. Mangrove sponge disease induced by cyanobacterial symbionts: failure of a primitive immune system? *Diseases of Aquatic Organisms* 5:143–149. DOI: 10.3354/dao005143.

Salo T., Gustafsson C. 2016. The effect of genetic diversity on ecosystem functioning in vegetated coastal ecosystems. *Ecosystems* 19:1429–1444. DOI: 10.1007/s10021-016-0014-y.

Sbordoni V., De Matthaeis E., Cobolli Sbordoni M., La Rosa G., Mattoccia M. 1986. Bottleneck effects and the depression of genetic variability in hatchery stocks of *Penaeus japonicus* (Crustacea, Decapoda). *Aquaculture* 57:239–251. DOI: 10.1016/0044-8486(86)90202-4.

Scheffer M., Carpenter S., Young B. 2005. Cascading effects of overfishing marine systems. *Trends in Ecology & Evolution* 20:579–581. DOI: 10.1016/j.tree.2005.08.018.

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. *The ISME Journal* 6:564–76. DOI: 10.1038/ismej.2011.116.

Schönberg C., Fromont J. 2012. Sponge gardens of Ningaloo Reef (Carnarvon Shelf, Western Australia) are biodiversity hotspots. *Hydrobiologia* 687:143–161. DOI: 10.1007/s10750-011-0863-5.

Shafer ABA., Wolf JBW., Alves PC., Bergström L., Bruford MW., Brännström I., Colling G., Dalén L., De Meester L., Ekblom R., Fawcett KD., Fior S., Hajibabaei M., Hill JA., Hoezel AR., Höglund J., Jensen EL., Krause J., Kristensen TN., Krützen M., McKay JK., Norman AJ., Ogden R., Österling EM., Ouborg NJ., Piccolo J., Popović D., Primmer CR., Reed FA., Roumet M., Salmona J., Schenekar T., Schwartz MK., Segelbacher G., Senn H., Thaulow J., Valtonen M., Veale A., Vergeer P., Vijay N., Vilà C., Weissensteiner M., Wennerström L., Wheat CW., Zieliński P. 2015. Genomics and the challenging translation into conservation practice. *Trends in Ecology & Evolution* 30:78–87. DOI:

10.1016/j.tree.2014.11.009.

Smith FGW. 1941. Sponge disease in British Honduras, and its transmission by water currents. *Ecology* 22:415–421. DOI: 10.2307/1930719.

Stabili L., Cardone F., Alifano P., Tredici SM., Piraino S., Corriero G., Gaino E. 2012. Epidemic mortality of the sponge *Ircinia variabilis* (Schmidt, 1862) associated to proliferation of a Vibrio bacterium. *Microbial Ecology* 64:802–813. DOI: 10.1007/s00248-012-0068-0.

Stachowitsch M. 1984. Mass mortality in the Gulf of Tieste: The course of community destruction. *Marine Ecology* 5:243–264. DOI: 10.1111/j.1439-0485.1984.tb00124.x.

Stevely JM., Sweat DE., Bert TM., Sim-Smith C., Kelly M. 2010. Sponge Mortality at Marathon and Long Key, Florida : Patterns of Species Response and Population Recovery. *Proceedings of the 63rd Gulf and Caribbean Fisheries Institute* 63:384–400.

Taylor HR., Dussex N., van Heezik Y. 2017. Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners. *Global Ecology and Conservation* 10:231–242. DOI: 10.1016/j.gecco.2017.04.001.

Taylor MW., Hill RT., Piel J., Thacker RW., Hentschel U. 2007. Soaking it up: the complex lives of marine sponges and their microbial associates. *The ISME Journal* 1:187–90. DOI: 10.1038/ismej.2007.32.

Tellier M-AS., Bertelsen R. 2008. *Monitoring the flora and fauna of the nearshore hardbottom habitats of the Florida Keys*. F2196-05-08-F Final Report. Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute, Marathon, FL.

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., Victor Lopez J., Taylor MW., Thacker RW., Montoya JM., Hentschel U., Webster NS. 2016. Diversity, structure and convergent evolution of the global sponge microbiome. *Nature Communications* 7:11870. DOI: 10.1038/ncomms11870.

Thrush SF., Dayton PK. 2002. Disturbance to marine benthic habitats by trawling and dredging: Implications for marine biodiversity. *Annual Review of Ecology and Systematics* 33:449–473. DOI: 10.1146/annurev.ecolsys.33.010802.150515.

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* 7:e35105. DOI: 10.1371/journal.pone.0035105.

Van Soest, R.W.M; Boury-Esnault, N.; Hooper, J.N.A.; Rützler, K.; de Voogd, N.J.; Alvarez de Glasby, B.; Hajdu, E.; Pisera, A.B.; Manconi, R.; Schoenberg, C.; Klautau, M.; Picton, B.; Kelly, M.; Vacelet, J.; Dohrmann, M.; Díaz, M.-C.; Cárdenas, P.; Carballo, J. L.; Rios Lopez, P. (2017). World Porifera database. Accessed at http://www.marinespecies.org/porifera on 2017-09-23

van Strien MJ., Keller D., Holderegger R., Ghazoul J., Kienast F., Bolliger J. 2014. Landscape genetics as a tool for conservation planning: predicting the effects of landscape change on gene flow. *Ecological Applications* 24:327–339. DOI: 10.1890/13-0442.1.

Vicente VP. 1989. Regional commercial sponge extinctions in the West Indies: Are recent climatic changes responsible? *Marine Ecology* 10:179–191. DOI: 10.1111/j.1439-0485.1989.tb00073.x.

Wang G. 2006. Diversity and biotechnological potential of the sponge-associated microbial consortia. *Journal of Industrial Microbiology and Biotechnology* 33:545–51. DOI: 10.1007/s10295-006-0123-2.

Webster NS. 2007. Sponge disease: a global threat? *Environmental Microbiology* 9:1363–75. DOI: 10.1111/j.1462-2920.2007.01303.x.

Webster NS., Thomas T. 2016. The sponge hologenome. *mBio* 7:e00135-16. DOI: 10.1128/mBio.00135-16.

Wendt PH., Van Dolahh RF., O'Rourke CB. 1985. A comparative study of the invertebrate macrofauna associated with seven sponge and coral species collected from the south Atlantic Bight. *The Journal of the Elisha Mitchell Scientific Society* 101:187–203.

Whitlock MC. 2000. Fixation of new alleles and the extinction of small populations: drift load, beneficial alleles and sexual selection. *Evolution* 54:1855. DOI: 10.1554/0014-3820(2000)054[1855:FONAAT]2.0.CO;2.

Whitney F., Conway K., Thomson R., Barrie V., Krautter M., Mungov G. 2005. Oceanographic habitat of sponge reefs on the Western Canadian Continental Shelf. *Continental Shelf Research* 25:211–226. DOI: 10.1016/j.csr.2004.09.003.

Williams A V., Nevill PG., Krauss SL. 2014. Next generation restoration genetics: applications and opportunities. *Trends in Plant Science* 19:529–537. DOI: 10.1016/j.tplants.2014.03.011.

Wright S. 1931. Evolution in Mendelian populations. *Bulletin of Mathematical Biology* 52:241–295. DOI: 10.1007/BF02459575.

Wulff JL. 1984. Sponge-mediated coral reef growth and rejuvenation. *Coral Reefs* 3:157–163. DOI: 10.1007/BF00301960.

Wulff J. 2006. Rapid diversity and abundance decline in a Caribbean coral reef sponge community. *Biological Conservation* 127:167–176. DOI: 10.1016/j.biocon.2005.08.007.

Zea S., López-Victoria M. 2016. *Cliona acephala* (Porifera: Demospongiae: Clionaida), a new encrusting excavating reef sponge from the Colombian Caribbean belonging to the *Cliona viridis* species complex. *Zootaxa* 4178:583. DOI: 10.11646/zootaxa.4178.4.8.

Zea S., Pulido A. 2016. Taxonomy of the Caribbean sponge *Dragmacidon reticulatum* (Ridley & Dendy, 1886) (Porifera, Demospongiae, Axinellida), with the description of a new species. *Zootaxa* 4114:393. DOI: 10.11646/zootaxa.4114.4.2.

Chapter 2: A Galaxy-based bioinformatics pipeline for optimised, streamlined microsatellite development from Illumina next-generation sequencing data

Sarah M. Griffiths[1], Graeme Fox[1], Peter J. Briggs[2], Ian J. Donaldson[2], Simon Hood[3], Pen Richardson[3], George W. Leaver[3], Nathan K. Truelove[4] & Richard F. Preziosi[1]

[1] Faculty of Life Sciences, University of Manchester, Manchester, M13 9PT, UK.

[2] Bioinformatics Core Facility, Faculty of Life Sciences, University of Manchester, Manchester, M13 9PT, UK.

[3] Research Infrastructure Team, Research Computing, University of Manchester, Manchester, M13 9PL, UK.

[4] Smithsonian Marine Station, Fort Pierce, Florida 34949, USA.

Keywords: Microsatellite isolation; Pal_finder, PANDAseq, Trimmomatic; Pal_filter; Seq-SSR; SSRs.

Acknowledgments

This work was funded by the Faculty of Life Sciences, University of Manchester, and a studentship awarded to SMG from the Natural Environment Research Council (NERC). We are grateful to the Genomic Technologies Core Facility and the DNA Sequencing Facility at the University of Manchester for their services and advice in this project. Thanks to Ezemvelo KwaZulu Natal Wildlife and C. Weldon for the use of the *Amietia hymenopus* tissue sample (permit OP526/2014), and to A. Pretorius and R. Antwis for their work in the lab. Thanks also to J-D Hibbitt and Sea Life for providing the *Raja undulata* sample, and M. Butler IV and D. Behringer for samples used in the initial development of this tool.

Author contributions

SMG, GF, NKT and RFP conceived and designed the pipeline; GF wrote the Pal_filter script; PJB and ID configured the Galaxy instance and wrappers; SH,

PR and GWL created the IT framework for website development; SMG wrote the manuscript; GF, PJB, ID, NKT and RFP edited drafts of the manuscript. The data used in the case studies were obtained from other projects carried out by SMG and GF.

Work on using Galaxy-based tools for microsatellite development began in SMG's masters project 'Isolating microsatellite markers in the marine sponge *Cinachyrella alloclada* for use in community and population genetics studies'. This chapter presents the entire pipeline that resulted from further developments made during this PhD.

A slightly modified version of this chapter is published in the journal Conservation Genetics Resources (Volume 8, Issue 4, pages 481-486) (Appendix 1).

ABSTRACT

Microsatellites are useful tools for ecologists and conservationist biologists, but are taxa-specific and traditionally expensive and time-consuming to develop. New methods using next-generation sequencing (NGS) have reduced these problems, but the plethora of software available for processing NGS data may cause confusion and difficulty for researchers new to the field of bioinformatics. We developed a bioinformatics pipeline for microsatellite development from Illumina paired-end sequences, which is packaged in the open-source bioinformatics tool Galaxy. This optimises and streamlines the design of a microsatellite panel and provides a user-friendly graphical user interface (GUI). The pipeline utilises existing programs along with our own novel program and wrappers to: quality-filter and trim reads (Trimmomatic); generate sequence quality reports (FastQC); identify potentially-amplifiable microsatellite loci (Pal_finder); design primers (Primer3); assemble pairs of reads to enhance marker amplification success rates (PANDAseq); and filter optimal loci (Pal_filter). The complete pipeline is freely available for use via a pre-configured Galaxy instance, accessible at https://palfinder.ls.manchester.ac.uk.

INTRODUCTION

Microsatellites are popular and effective genetic markers that are utilised in many conservation genetics studies and can inform natural resource management (for example, Maudetr et al., 2002; Jehle & Arntzen, 2002; Truelove et al., 2014). Their high rate of polymorphism, codominant mode of inheritance and their utility with even degraded DNA make microsatellites a go-to marker for many studies in ecology and conservation (Sunnucks, 2000; Selkoe & Toonen, 2006). However, these markers are taxa-specific, meaning primers must often be developed *de novo* for each new species or genus - historically an expensive and time-consuming process.

High-throughput next-generation sequencing (NGS) has decreased the cost-per-base of DNA sequencing significantly, while massively increasing the output (Wetterstrand 2012). Where random enrichment strategies were once used to target microsatellites, new methods to detect short sequence repeats (SSRs) directly from NGS datasets are being developed; the so-called Seq-SSR approach (Goldstein & Schlötterer, 1999; Castoe et al., 2012). It is now cost- and time-effective to perform shotgun genome sequencing, computationally identify SSRs in the raw sequencing reads and search their flanking regions for potential primer binding sites (Zalapa et al., 2012). Further cost reductions can be achieved by using Illumina paired-end sequencing, which involves sequencing from both ends of a read (Castoe et al., 2012). This gives greater read lengths than single-end sequencing (up to 2 x 300 base pairs [bp] with the Illumina MiSeq [Illumina 2016]) whilst at a cheaper cost per base than Roche 454 sequencing technology.

The reduced cost, increased number of loci, and more efficient development processes that NGS methods offer mean that microsatellite characterisation is now available to research groups that may have originally been too constrained by cost and time. However, effectively processing the huge amount of data resulting from an NGS run can be challenging for groups without bioinformatics support or previous experience with NGS data. The number of programs

available can be daunting, and many can be complicated and time-consuming for novices to master.

We have created a complete microsatellite development pipeline for raw Illumina paired-end data that incorporates existing computer programs and a novel filtering script described here (pal_filter). This pipeline has been developed within Galaxy, an open-source, web-based and user-friendly bioinformatics tool for handling large data sets, available on a free public server or to be downloaded as a local installation (Giardine et al., 2005; Goecks et al., 2010; Blankenberg et al., 2010). The use of Galaxy allows the programs within the pipeline be run in a single operational framework, streamlining the process, and providing a graphical user interface (GUI) to increase operational ease and accessibility. Galaxy is well supported, with video tutorials available to support first-time users in use and navigation (see http://galaxyproject.org). Our pipeline provides a complete workflow from receipt of raw sequencing files to production of a list of filtered, optimised microsatellite loci and primers with no further software required for preliminary or post processing.

MICROSATELLITE DEVELOPMENT PIPELINE PROCESSES

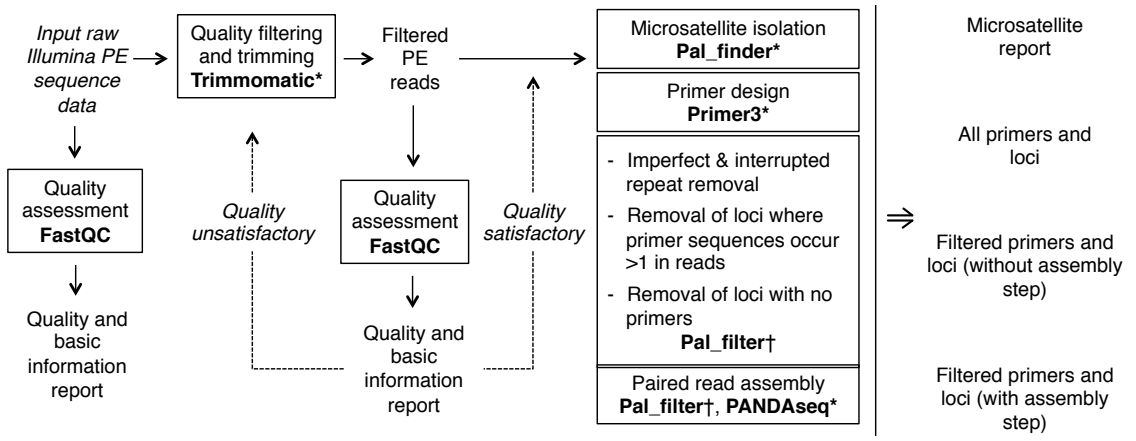


Figure 1: Pipeline processes (in boxes), the programs used (in bold), and pipeline output. *Novel wrapper enabling process step to be run in Galaxy; †Novel program developed by the authors.

Generating Illumina sequence data

This data-processing pipeline has been developed and optimised for Illumina paired-end sequence data. A single sample should be sequenced for each species intended for microsatellite development. Due to the large volume of data and potential microsatellite primers generated in a single sequencing run, more than one sample can be multiplexed in the same Illumina flow cell lane to allow microsatellite characterisation for multiple species for the same initial sequencing costs (Castoe et al., 2012; also see Table 1). The number of species that can be sequenced in one Illumina flow cell lane whilst still retaining an adequate number of suitable microsatellite primers depends on many factors, including the output capacity of the sequencer, microsatellite-richness of the genomes of the organisms and the types of microsatellite repeats the researchers are interested in (for example, dinucleotide repeats are more common in genomes than longer length repeats). We would advise potential users to consult a sequencing technician before making this choice.

A number of Illumina platforms are available, which offer users various read length, sequencing output and cost combinations (Illumina, 2016). Longer read

lengths are advantageous for microsatellite development purposes, as they allow more opportunity for suitable primer binding sites to be found in the microsatellite flanking regions. However, longer reads often suffer from reduced quality at their ends, and therefore they may have to be trimmed to ensure adequate quality (see 'Quality filtering of data', below). Additionally, longer read lengths allow for primers for larger PCR amplicons to be designed, which can be more prone to large allele dropout (Sefc, Payne & Sorenson, 2003). Currently, the MiSeq platform allows a maximum read length of 2 x 300 bp (Illumina 2016). However, Castoe et al., (2012) successfully used 2 x 116 bp read lengths generated by the GAIIx platform to develop microsatellite primers. As sequencing technology is constantly evolving, again we would recommend users to consult a sequencing technician to discuss the most appropriate platform and read length to use.

Quality filtering of data

Data resulting from automated sequencing processes inevitably contains error (especially at the end of reads), which can negatively affect downstream applications. In microsatellite development, miscalled bases in the microsatellite flanking regions could lead to ineffective primer design, non-binding or mis-priming with the target sequence during PCR, and subsequent amplification failure.

We have incorporated Trimmomatic v.0.32 (Bolger, Lohse & Usadel, 2014) into the pipeline to trim low-quality bases from reads and remove low-quality reads. Specially formulated for paired-end data, Trimmomatic discards both members of a pair if either one does not pass user-specified quality thresholds. This 'pair-awareness' results in two files in which the parity of the paired end reads is maintained, essential for the correct functioning of programs downstream. Users can also use Trimmomatic to remove adapter sequences from the reads that have been left over from the sequencing process.

Read quality and basic information report

FastQC v0.11.4 (Andrews 2014) is used to generate reports containing basic statistics on the reads and various quality assessments. Reports are generated both from the raw and quality-filtered data files, containing useful information such as Phred (quality) scores, GC content, sequence duplication levels, sequence length distribution, and amount and type of adapter content.

Microsatellite identification and primer design

The files containing surviving pairs from the Trimmomatic process are used for identification of microsatellites and PCR primer design. Sequences containing repeat motifs of up to 6 bp are identified using the program Pal_finder v.0.02.04 (Castoe et al., 2012). The program then examines the flanking regions for suitability as PCR priming sites (identifying 'PALs'; potentially amplifiable loci), and if suitable, uses Primer3 (Koressaar & Remm, 2007; Untergasser et al., 2012) to design primers according to parameters specified by the user (for example, melting temperature and primer length). Two tab delimited files are outputted (readable by Microsoft Excel); one comprising a list of the types of microsatellites found, and another giving a list of all the loci found including the motif, primer sequence, number of occurrences of the primer sequence in the total reads, and the sequence IDs of the forward and reverse reads.

Microsatellite loci filtering

We incorporated a series of optional filters into the pipeline (implemented via a novel Python script, which we have named Pal_filter) to select the optimal loci from the Pal_finder output text file of microsatellite loci and primers. This gives the user the option to filter out any or all of the following: 1) Loci for which primers could not be designed by Primer3; 2) Loci with imperfect or interrupted motifs (as these do not follow the stepwise mutation model, which many microsatellite population genetics analysis programs assume). If enabled, the loci are also ranked by size of motif (largest first); 3) Loci in which the primer sequences occur more than once in the total reads (to ensure a copy number of one and avoid genes with duplication in the genome). This generates an easy to navigate, tab delimited file and negates the need for manual sorting of

potentially thousands of results from the original Pal_finder output. The original file of all PALs and primers is still available (as are all outputs from the pipeline).

Improving PCR success: paired read assembly

Despite the many benefits of NGS workflows, pairs of primers must still be manually tested in the laboratory to ensure successful amplification. This can represent a considerable cost in both time and resources in the development of a panel of working microsatellite markers. We implemented an additional quality-filtering step with the specific aim of improving the rate of successful PCR and thus reducing these expenses. In brief, the paired-end read assembler PANDAseq (Masella et al., 2012) is used to provide confirmation that both primer sequences occur in the same region of DNA template and increase PCR success (Fox et al., unpublished). This additional quality check is implemented as part of the Pal_filter script. Selecting this option will generate another tab delimited file that again reduces the Pal_finder output to those loci in which the reads could be assembled, as well as incorporating any of the previous filters that have been applied (while still retaining all the other output files).

CASE STUDIES

Table 1 shows the number of microsatellites primers found and subsequent amplification success rates for a variety of configuration options in three species across different taxa (an amphibian, an elasmobranch and a mollusc). Total amplification success rates (S_T; percentage of primers tested that resulted in loci that were amplifiable and scorable by capillary electrophoresis) ranged from 53 to 73 per cent, providing proof of principle that the pipeline described here consistently results in successful microsatellite primer development. Table 1 also shows the percentage of primers tested that produced PCR products that could be visualised using agarose gel electrophoresis (S_G); it should be noted that this is consistently higher than the total amplification success rate. We have reported this to highlight that initial testing of primers on agarose gels may not reflect the actual number of usable loci that will be available when using capillary electrophoresis to measure allele sizes. This can be due to a number of reasons, including high levels of 'stutter' for a locus making the true allele difficult to distinguish, or non-specific binding resulting in multiple peaks on a sequencer trace.

The case studies also highlight the potential economy of this method. *Modiolus modiolus* was sequenced in an Illumina flowcell lane with seven other species for microsatellite development purposes, and 144 loci with primers were available after the most stringent filtering and assembly options were used. If the total amplification success rate for this species (53 per cent) is assumed to apply for all these loci, this would still mean that around 76 loci would be usable in a conservation genetics study. Currently, this far exceeds the number of microsatellites normally used for these purposes. This shows that pooling multiple samples in one lane of an Illumina flowcell can reduce the cost-per-species of microsatellite development considerably whilst still retaining an ample number of high-quality loci.

Table 1: Case studies of microsatellite development using the described pipeline.

Species	P	No. reads (2 x)		No. loci with primers *[total no. loci]*		S_T *[S_G]* (%)
		Raw	Filtered	Raw reads	Filtered reads	
Amietia hymenopus (Phofung river frog)	0.5	6,465,564	3,756,407	* 25,427 *[149,271]* † 1,345 ‡ 216	* 11,350 *[60,378]* † 1,097 ‡ 144	56 *[64]*
Raja undulata (Undulate ray)	0.5	11,019,590	10,174,420	* 267,431 *[130,894]* † 3,119 ‡ 428	* 107,470 *[31,876]* † 342 ‡ 148	73 *[80]*
Modiolus modiolus (Northern horsemussel)	0.125	4,647,211	4,455,417	* 64,489 *[44,408]* † 1,650 ‡ 225	* 39,232 *[16,814]* † 707 ‡ 144	53 *[74]*

All sequencing was paired end, carried out on the Illumina MiSeq, with sequence lengths of 2 x 250 bp. Trimmomatic settings (SLIDING WINDOW: WINDOW SIZE = 4 bp, QUALITY = 20; LEADING = 3; TRAILING = 3; MINLEN = 50) and primer design conditions (recommended settings for Qiagen Type-it® Microsatellite PCR kit) were constant across all tests. Minimum number of microsatellite repeats to be searched for was eight for all repeat types (2-6mer).

P, proportion of Illumina flow cell lane used; * without pal_filter or assembly; † with pal_filter (all filtering options selected), without assembly; ‡ with pal_filter (all filtering options selected) and assembly; S_T, total amplification success rate − percentage of loci tested that resulted in amplifiable loci that could be easily scored when fluorescently labeled and analysed using an automated capillary sequencer; S_G, amplification success rate using agarose gel electrophoresis − percentage of loci tested that resulted in clear bands when visualising PCR products of unlabeled primers on an agarose gel. Primers used in this test were developed from Trimmomatic-fitered reads, with all of the pal_filter and assembly options selected.

Filtering the reads using Trimmomatic removed between 4.1 and 41.9 per cent (*Raja undulata* and *Amietia hymenopus* respectively) of the raw reads. The settings used (see Table 1) ensured that the remaining reads had an average Phred score of 20 across every four bases, meaning a base call accuracy of 99 per cent. It is prudent to remove low quality reads and bases in order to reduce the likelihood of designing primers based on miscalled bases, as this may result in PCR amplification failure. This effect could be substantial when a high proportion of reads are low quality (as in *Amietia hymenopus*).

SUMMARY

This bioinformatics pipeline is a robust method for designing effective microsatellite primers, and its incorporation into Galaxy provides a user-friendly framework in which to operate the pipeline. In Chapter 4 and Chapter 5 of this thesis, I successfully used these methods to develop microsatellite markers for two sponge species, *Ircinia campana* and *Spheciospongia vesparium*. As microsatellite development becomes more accessible to researchers, it is important to consider both the positive and negative aspects of microsatellites as molecular markers before embarking on development projects. A number of articles discuss these potential issues (for example, Selkoe & Toonen, 2006; Väli et al., 2008; Guichoux et al., 2011; Putman & Carbone, 2014) and should be reviewed by any potential microsatellite users. Users of the pipeline described here are also encouraged to consult the articles cited for each of the programs utilised, as well as the user manual for the pipeline (see https://palfinder.ls.manchester.ac.uk/manual), which goes into detail on user-specified settings and use of the programs in Galaxy. We envision that this will be a useful tool for both academic and non-academic groups involved in conservation genetics research due to its comprehensiveness, effectiveness and ease of use.

Accessing the pipeline

There are three options available for potential users: 1) A public Galaxy instance (called Galaxy Palfinder Service) implementing the pipeline with complete functionality as described here is available online for research use at https://palfinder.ls.manchester.ac.uk. A manual including detailed instructions for use is available at https://palfinder.ls.manchester.ac.uk/manual; 2) Advanced users with access to their own local Galaxy server may download the Trimmomatic and Pal_finder (including Pal_filter) wrappers from https://toolshed.g2.bx.psu.edu/view/pjbriggs/, and the FastQC wrapper from https://toolshed.g2.bx.psu.edu/view/devteam/fastqc/; 3) Finally, all programs can be run outside the Galaxy environment at the command line (Unix) (for detailed instructions, see user manual).

REFERENCES

Andrews S. (2014). FastQC: a quality control tool for high throughput sequence data [Online]. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. [Accessed 7/4/16].

Blankenberg D., Von Kuster G., Coraor N., Ananda G., Lazarus R., Mangan M., Nekrutenko A., Taylor J. 2010. Galaxy: A web-based genome analysis tool for experimentalists. *Current Protocols in Molecular Biology* 89:19.10:19.10.1-19.10.21. DOI: 10.1002/0471142727.mb1910s89.

Bolger AM., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. DOI: 10.1093/bioinformatics/btu170.

Castoe T., Poole A., de Koning A., Jones K., Tomback D., Oyler-McCance S., Fike J., Lance S., Streicher J., Smith E., Pollock D. 2012. Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. *PLoS ONE* 7:e30953. DOI: 10.1371/journal.pone.0030953.

Giardine B., Riemer C., Hardison RC., Burhans R., Elnitski L., Shah P., Zhang Y., Blankenberg D., Albert I., Taylor J., Miller W., Kent WJ., Nekrutenko A. 2005. Galaxy: a platform for interactive large-scale genome analysis. *Genome Research* 15:1451–5. DOI: 10.1101/gr.4086505.

Goecks J., Nekrutenko A., Taylor J., The Galaxy Team. 2010. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology* 11:R86. DOI: 10.1186/gb-2010-11-8-r86.

Goldstein DB., Schlötterer C. (eds.) 1999. *Microsatellites — Evolution and Applications*. Oxford University Press, Oxford.

Guichoux E., Lagache L., Wagner S., Chaumeil P., Léger P., Lepais O., Lepoittevin C., Malausa T., Revardel E., Salin F., Petit RJ. 2011. Current trends in microsatellite genotyping. *Molecular Ecology Resources* 11:591–611. DOI: 10.1111/j.1755-0998.2011.03014.x.

Jehle R., Arntzen J. 2002. Review: microsatellite markers in amphibian conservation genetics. *Herpetological Journal* 12:1–9.

Koressaar T., Remm M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23:1289–91. DOI: 10.1093/bioinformatics/btm091.

Masella AP., Bartram AK., Truszkowski JM., Brown DG., Neufeld JD. 2012. PANDAseq: paired-end assembler for illumina sequences. *BMC Bioinformatics* 13:31. DOI: 10.1186/1471-2105-13-31.

Maudetr C., Miller C., Bassano B., Breitenmoser-Würsten C., Gauthier D., Obexer-Ruff G., Michallet J., Taberlet P., Luikart G. 2002. Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex [*Capra ibex*(*ibex*)]. *Molecular Ecology* 11:421–36. DOI: 10.1046/j.0962-1083.2001.01451.x

Putman AI., Carbone I. 2014. Challenges in analysis and interpretation of microsatellite data for population genetic studies. *Ecology and Evolution* 4:n/a-n/a. DOI: 10.1002/ece3.1305.

Sefc KM., Payne RB., Sorenson MD. 2003. Microsatellite amplification from museum feather samples: Effects of fragment size and template concentration on genotyping errors. *The Auk* 120:982–989. DOI: 10.2307/4090269.

Selkoe KA., Toonen RJ. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9:615–29. DOI: 10.1111/j.1461-0248.2006.00889.x.

Sunnucks P. 2000. Efficient genetic markers for population biology. *Trends in Ecology & Evolution* 15:199–203. DOI: 10.1016/S0169-5347(00)01825-5.

Truelove NK., Griffiths S., Ley-Cooper K., Azueta J., Majil I., Box SJ., Behringer DC., Butler MJ., Preziosi RF. 2014. Genetic evidence from the spiny lobster fishery supports international cooperation among Central American marine protected areas. *Conservation Genetics* 16:347–358. DOI: 10.1007/s10592-014-0662-4.

Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth BC., Remm M., Rozen SG. 2012. Primer3- new capabilities and interfaces. *Nucleic Acids Research* 40:e115. DOI: 10.1093/nar/gks596.

Väli Ü., Einarsson A., Waits L., Ellegren H. 2008. To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Molecular Ecology* 17:3808–3817. DOI: 10.1111/j.1365-294X.2008.03876.x.

Zalapa JE., Cuevas H., Zhu H., Steffan S., Senalik D., Zeldin E., McCown B., Harbut R., Simon P. 2012. Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. *American Journal of Botany* 99:193–208. DOI: 10.3732/ajb.1100394.

Appendix I: Publication in *Conservation Genetics Resources*

Conservation Genet Resour (2016) 8:481–486
DOI 10.1007/s12686-016-0570-7

METHODS AND RESOURCES ARTICLE

A Galaxy-based bioinformatics pipeline for optimised, streamlined microsatellite development from Illumina next-generation sequencing data

Sarah M. Griffiths[1] · Graeme Fox[1] · Peter J. Briggs[2] · Ian J. Donaldson[2] · Simon Hood[3] · Pen Richardson[3] · George W. Leaver[3] · Nathan K. Truelove[4] · Richard F. Preziosi[1]

Received: 19 February 2016 / Accepted: 1 July 2016 / Published online: 2 August 2016

Abstract Microsatellites are useful tools for ecologists and conservationist biologists, but are taxa-specific and traditionally expensive and time-consuming to develop. New methods using next-generation sequencing (NGS) have reduced these problems, but the plethora of software available for processing NGS data may cause confusion and difficulty for researchers new to the field of bioinformatics. We developed a bioinformatics pipeline for microsatellite development from Illumina paired-end sequences, which is packaged in the open-source bioinformatics tool Galaxy. This optimises and streamlines the design of a microsatellite panel and provides a user-friendly graphical user interface. The pipeline utilises existing programs along with our own novel program and wrappers to: quality-filter and trim reads (Trimmomatic); generate sequence quality reports (FastQC); identify potentially-amplifiable microsatellite loci (Pal_finder); design primers (Primer3); assemble pairs of reads to enhance marker amplification success rates (PANDAseq); and filter optimal loci (Pal_filter). The complete pipeline is freely available for use via a pre-configured Galaxy instance, accessible at https://palfinder.ls.manchester.ac.uk.

✉ Sarah M. Griffiths
griffiths.sarahm@gmail.com

[1] Faculty of Life Sciences, University of Manchester, Manchester M13 9PT, UK

[2] Bioinformatics Core Facility, Faculty of Life Sciences, University of Manchester, Manchester M13 9PT, UK

[3] Research Infrastructure Team, Research Computing, University of Manchester, Manchester M13 9PL, UK

[4] Smithsonian Marine Station, Fort Pierce, FL 34949, USA

Keywords Microsatellite isolation · Pal_finder · PANDAseq · Trimmomatic · Pal_filter · Seq-SSR · SSRs · Galaxy · Next-generation sequencing · Illumina

Introduction

Microsatellites are popular and effective genetic markers that are utilised in many conservation genetics studies and can inform natural resource management (for example, Maudetr et al. 2002; Jehle and Arntzen 2002; Truelove et al. 2014). Their high rate of polymorphism, codominant mode of inheritance and their utility with even degraded DNA make microsatellites a go-to marker for many studies in ecology and conservation (Sunnucks 2000; Selkoe and Toonen 2006). However, these markers are taxa-specific, meaning primers must often be developed de novo for each new species or genus—traditionally an expensive and time-consuming process.

High-throughput next-generation sequencing (NGS) has decreased the cost-per-base of DNA sequencing significantly, while massively increasing the output (Wetterstrand 2012). Where random enrichment strategies were once used to target microsatellites, new methods to detect short sequence repeats (SSRs) directly from NGS datasets are being developed; the so-called Seq-SSR approach (Goldstein and Schlotterer 1999; Castoe et al. 2012). It is now cost- and time-effective to perform shotgun genome sequencing, computationally identify SSRs in the raw sequencing reads and search their flanking regions for potential primer binding sites (Zalapa et al. 2012). Further cost reductions can be achieved by using Illumina paired-end sequencing, which involves sequencing from both ends of a read (Castoe et al. 2012). This gives greater read lengths than single-end sequencing (up to 2 × 300 base

pairs [bp] with the Illumina MiSeq [Illumina 2016]) whilst at a cheaper cost per base than Roche 454 sequencing technology.

The reduced cost, increased number of loci, and more efficient development processes that NGS methods offer mean that microsatellite characterisation is now available to research groups that may have originally been too constrained by cost and time. However, effectively processing the huge amount of data resulting from an NGS run can be challenging for groups without bioinformatics support or previous experience with NGS data. The number of programs available can be daunting, and many can be complicated and time-consuming for novices to master.

We have created a complete microsatellite development pipeline for raw Illumina paired-end data that incorporates existing computer programs and a novel filtering script described here (pal_filter). This pipeline has been developed within Galaxy, an open-source, web-based and user-friendly bioinformatics tool for handling large data sets, available on a free public server or to be downloaded as a local installation (Giardine et al. 2005; Blankenberg et al. 2010; Goecks et al. 2010). The use of Galaxy allows the programs within the pipeline be run in a single operational framework, streamlining the process, and providing a graphical user interface (GUI) to increase operational ease and accessibility. Galaxy is well supported, with video tutorials available to support first-time users in use and navigation (see http://galaxyproject.org). Our pipeline provides a complete workflow from receipt of raw sequencing files to production of a list of filtered, optimised microsatellite loci and primers with no further software required for preliminary or post processing (Fig. 1).

Microsatellite development pipeline processes

Generating Illumina sequence data

This data-processing pipeline has been developed and optimised for Illumina paired-end sequence data. A single sample should be sequenced for each species intended for microsatellite development. Due to the large volume of data and potential microsatellite primers generated in a single sequencing run, more than one sample can be multiplexed in the same Illumina flow cell lane to allow microsatellite characterisation for multiple species for the same initial sequencing costs (Castoe et al. 2012; also see Table 1). The number of species that can be sequenced in one Illumina flow cell lane whilst still retaining an adequate number of suitable microsatellite primers depends on many factors, including the output capacity of the sequencer, microsatellite-richness of the genomes of the organisms and the types of microsatellite repeats the researchers are interested in (for example, dinucleotide repeats are more common in genomes than longer length repeats). We would advise potential users to consult a sequencing technician before making this choice.

A number of Illumina platforms are available, which offer users various read length, sequencing output and cost combinations (Illumina 2016). Longer read lengths are advantageous for microsatellite development purposes, as they allow more opportunity for suitable primer binding sites to be found in the microsatellite flanking regions. However, longer reads often suffer from reduced quality at their ends, and therefore they may have to be trimmed to ensure adequate quality (see 'Quality filtering of data'

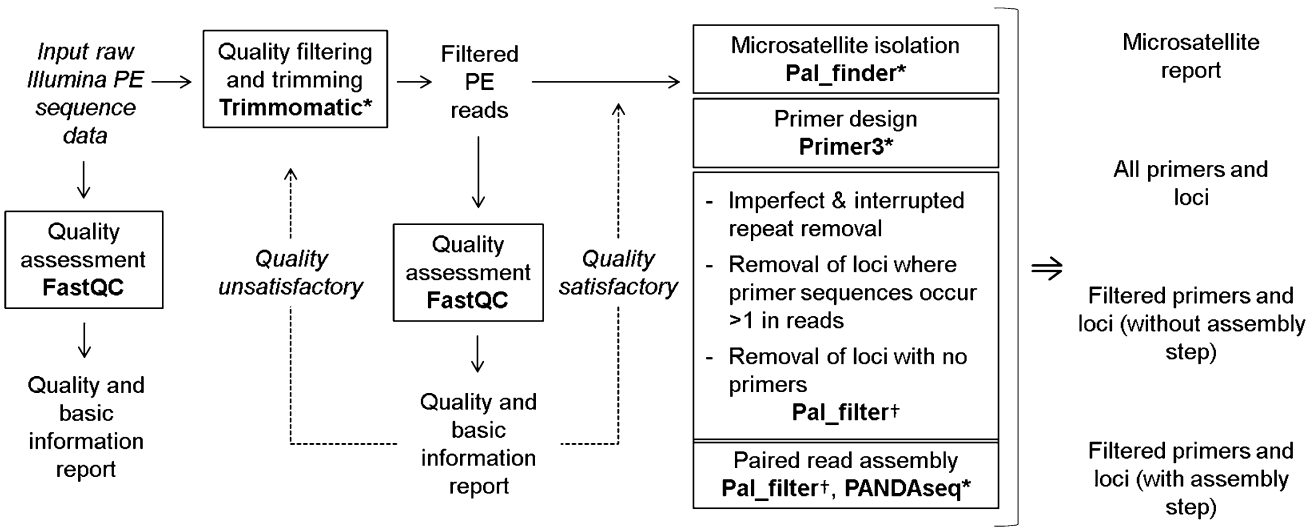


Fig. 1 Pipeline processes (*in boxes*), the programs used (*in bold*), and pipeline output. * novel wrapper enabling process step to be run in Galaxy; † novel program developed by the authors

Table 1 Case studies of microsatellite development using the described pipeline

Species	P	No. reads (2×)		No. loci with primers [total no. loci]		S_T [S_G] (%)
		Raw	Filtered	Raw reads	Filtered reads	
Amietia hymenopus (Phofung river frog)	0.5	6,465,564	3,756,407	25,427 [149,271]*	11,350 [60,378]*	56 [64]
				1,345†	1097†	
				216‡	144‡	
Raja undulata (Undulate ray)	0.5	11,019,590	10,174,420	267,431 [130,894]*	107,470 [31,876]*	73 [80]
				3119†	342†	
				428‡	148‡	
Modiolus modiolus (Northern horsemussel)	0.125	4,647,211	4,455,417	64,489 [44,408]*	39,232 [16,814]*	53 [74]
				1650†	707†	
				225‡	144‡	

All sequencing was paired end, carried out on the Illumina MiSeq, with sequence lengths of 2 × 250 bp. Trimmomatic settings (SLIDING WINDOW: WINDOW SIZE = 4 bp, QUALITY = 20; LEADING = 3; TRAILING = 3; MINLEN = 50) and primer design conditions (recommended settings for Qiagen Type-it® Microsatellite PCR kit) were constant across all tests. Minimum number of microsatellite repeats to be searched for was eight for all repeat types (2-6mer)

P, proportion of Illumina flow cell lane used; * without pal_filter or assembly; † with pal_filter (all filtering options selected), without assembly; ‡ with pal_filter (all filtering options selected) and assembly; S_T, total amplification success rate − percentage of loci tested that resulted in amplifiable loci that could be easily scored when fluorescently labeled and analysed using an automated capillary sequencer; S_G, amplification success rate using agarose gel electrophoresis − percentage of loci tested that resulted in clear bands when visualising PCR products of unlabeled primers on an agarose gel. Primers used in this test were developed from Trimmomatic-fitered reads, with all of the pal_filter and assembly options selected

section, below). Additionally, longer read lengths allow for primers for larger PCR amplicons to be designed, which can be more prone to large allele dropout (Sefc et al. 2003). Currently, the MiSeq platform allows a maximum read length of 2 × 300 bp (Illumina 2016). However, Castoe et al. (2012) successfully used 2 × 116 bp read lengths generated by the GAIIx platform to develop microsatellite primers. As sequencing technology is constantly evolving, again we would recommend users to consult a sequencing technician to discuss the most appropriate platform and read length to use.

Quality filtering of data

Data resulting from automated sequencing processes inevitably contains error (especially at the end of reads), which can negatively affect downstream applications. In microsatellite development, miscalled bases in the microsatellite flanking regions could lead to ineffective primer design, non-binding or mis-priming with the target sequence during PCR, and subsequent amplification failure.

We have incorporated Trimmomatic v.0.32 (Bolger et al. 2014) into the pipeline to trim low-quality bases from reads and remove low-quality reads. Specially formulated for paired-end data, Trimmomatic discards both members of a pair if either one does not pass user-specified quality thresholds. This 'pair-awareness' results in two files in which the parity of the paired end reads is maintained, essential for the correct functioning of programs downstream. Users can also use Trimmomatic to remove adapter sequences from the reads that have been left over from the sequencing process.

Read quality and basic information report

FastQC v0.11.4 (Andrews 2014) is used to generate reports containing basic statistics on the reads and various quality assessments. Reports are generated both from the raw and quality-filtered data files, containing useful information such as Phred (quality) scores, GC content, sequence duplication levels, sequence length distribution, and amount and type of adapter content.

Microsatellite identification and primer design

The files containing surviving pairs from the Trimmomatic process are used for identification of microsatellites and PCR primer design. Sequences containing repeat motifs of up to 6 bp are identified using the program Pal_finder v.0.02.04 (Castoe et al. 2012). The program then examines the flanking regions for suitability as PCR priming sites (identifying 'PALs'; potentially amplifiable loci), and if suitable, uses Primer3 (Koressaar and Remm 2007; Untergasser et al. 2012) to design primers according to parameters specified by the user (for example, melting temperature and primer length). Two tab delimited files are outputted (readable by Microsoft Excel); one comprising a list of the types of

microsatellites found, and another giving a list of all the loci found including the motif, primer sequence, number of occurrences of the primer sequence in the total reads, and the sequence IDs of the forward and reverse reads.

Microsatellite loci filtering

We incorporated a series of optional filters into the pipeline (implemented via a novel Python script, which we have named Pal_filter) to select the optimal loci from the Pal_finder output text file of microsatellite loci and primers. This gives the user the option to filter out any or all of the following: (1) Loci for which primers could not be designed by Primer3; (2) Loci with imperfect or interrupted motifs (as these do not follow the stepwise mutation model, which many microsatellite population genetics analysis programs assume). If enabled, the loci are also ranked by size of motif (largest first); (3) Loci in which the primer sequences occur more than once in the total reads (to ensure a copy number of one and avoid genes with duplication in the genome). This generates an easy to navigate, tab delimited file and negates the need for manual sorting of potentially thousands of results from the original Pal_finder output. The original file of all PALs and primers is still available (as are all outputs from the pipeline).

Improving PCR success: paired read assembly

Despite the many benefits of NGS workflows, pairs of primers must still be manually tested in the laboratory to ensure successful amplification. This can represent a considerable cost in both time and resources in the development of a panel of working microsatellite markers. We implemented an additional quality-filtering step with the specific aim of improving the rate of successful PCR and thus reducing these expenses. In brief, the paired-end read assembler PANDAseq (Masella et al. 2012) is used to provide confirmation that both primer sequences occur in the same region of DNA template and increase PCR success (Fox et al. unpublished). This additional quality check is implemented as part of the Pal_filter script. Selecting this option will generate another tab delimited file that again reduces the Pal_finder output to those loci in which the reads could be assembled, as well as incorporating any of the previous filters that have been applied (while still retaining all the other output files).

Case studies

Table 1 shows the number of microsatellites primers found and subsequent amplification success rates for a variety of configuration options in three species across different taxa (an amphibian, an elasmobranch and a mollusc). Total amplification success rates (S_T; percentage of primers tested that resulted in loci that were amplifiable and scorable by capillary electrophoresis) ranged from 53 to 73 %, providing proof of principle that the pipeline described here consistently results in successful microsatellite primer development. Table 1 also shows the percentage of primers tested that produced PCR products that could be visualised using agarose gel electrophoresis (S_G); it should be noted that this is consistently higher than the total amplification success rate. We have reported this to highlight that initial testing of primers on agarose gels may not reflect the actual number of usable loci that will be available when using capillary electrophoresis to measure allele sizes. This can be due to a number of reasons, including high levels of 'stutter' for a locus making the true allele difficult to distinguish, or non-specific binding resulting in multiple peaks on a sequencer trace.

The case studies also highlight the potential economy of this method. *Modiolus modiolus* was sequenced in an Illumina flowcell lane with seven other species for microsatellite development purposes, and 144 loci with primers were available after the most stringent filtering and assembly options were used. If the total amplification success rate for this species (53 %) is assumed to apply for all these loci, this would still mean that around 76 loci would be usable in a conservation genetics study. Currently, this far exceeds the number of microsatellites normally used for these purposes. This shows that pooling multiple samples in one lane of an Illumina flowcell can reduce the cost-per-species of microsatellite development considerably whilst still retaining an ample amount of high-quality loci.

Filtering the reads using Trimmomatic removed between 4.1 and 41.9 % (*Raja undulata* and *Amietia hymenopus* respectively) of the raw reads. The settings used (see Table 1) ensured that the remaining reads had an average Phred score of 20 across every four bases, meaning a base call accuracy of 99 %. It is prudent to remove low quality reads and bases in order to reduce the likelihood of designing primers based on miscalled bases, as this may result in PCR amplification failure. This effect could be substantial when a high proportion of reads are low quality (as in *Amietia hymenopus*).

Summary

This bioinformatics pipeline is a robust method for designing effective microsatellite primers, and its incorporation into Galaxy provides a user-friendly framework in which to operate the pipeline. Our lab group has successfully used this method to develop microsatellite markers in

a number of species, including vertebrates (Bertolotti et al. 2015), invertebrates and plants (data unpublished, also see Table 1).

As microsatellite development becomes more accessible to researchers, it is important to consider both the positive and negative aspects of microsatellites as molecular markers before embarking on development projects. A number of articles discuss these potential issues (for example, Selkoe and Toonen 2006; Väli et al. 2008; Guichoux et al. 2011; Putman and Carbone 2014) and should be reviewed by any potential microsatellite users. Users of the pipeline described here are also encouraged to consult the articles cited for each of the programs utilised, as well as the user manual for the pipeline (see https://palfinder.ls.manchester.ac.uk/manual), which goes into detail on user-specified settings and use of the programs in Galaxy. We envision that this will be a useful tool for both academic and non-academic groups involved in conservation genetics research due to its comprehensiveness, effectiveness and ease of use.

Accessing the pipeline

There are three options available for potential users: (1) A public Galaxy instance (called Galaxy Palfinder Service) implementing the pipeline with complete functionality as described here is available online for research use at https://palfinder.ls.manchester.ac.uk. A manual including detailed instructions for use is available at https://palfinder.ls.manchester.ac.uk/manual; (2) Advanced users with access to their own local Galaxy server may download the Trimmomatic and Pal_finder (including Pal_filter) wrappers from https://toolshed.g2.bx.psu.edu/view/pjbriggs/, and the FastQC wrapper from https://toolshed.g2.bx.psu.edu/view/devteam/fastqc/; (3) Finally, all programs can be run outside the Galaxy environment at the command line (Unix) (for detailed instructions, see user manual).

Acknowledgments This work was funded by the Faculty of Life Sciences, University of Manchester, and a studentship awarded to SMG from the Natural Environment Research Council (NERC). We are grateful to the Genomic Technologies Core Facility and the DNA Sequencing Facility at the University of Manchester for their services and advice in this project. Thanks to Ezemvelo KwaZulu Natal Wildlife and C. Weldon for the use of the *A. hymenopus* tissue sample (permit OP526/2014). Thanks also to J-D Hibbitt and Sea Life for providing the *R. undulata* sample, and M. Butler IV and D. Behringer for samples used in the initial development of this tool.

References

Andrews S (2014) FastQC: a quality control tool for high throughput sequence data [Online]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 7 April 2016

Bertolotti AC, Griffiths SM, Truelove NK, Box SJ, Preziosi RF, Salinas de Leon P (2015) Isolation and characterization of 10 polymorphic microsatellite loci for the endangered Galapagos-endemic whitespotted sandbass (*Paralabrax albomaculatus*). PeerJ 3:e1253. doi:10.7717/peerj.1253

Blankenberg D, Von Kuster G, Coraor N, Ananda G, Lazarus R, Mangan M, Nekrtenko A, Taylor J (2010) Galaxy: a web-based genome analysis tool for experimentalists. Curr Protoc Mol Biol. doi:10.1002/0471142727.mb1910s89

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30:2114–2120. doi:10.1093/bioinformatics/btu170

Castoe TA, Poole AW, de Koning J, Jones KL, Tomback DF, Oyler-McCance SJ, Fike JA, Lance SL, Streicher JW, Smith EN, Pollock dd (2012) Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. PLoS ONE 7:e3095. doi:10.1371/journal.pone.0030953

Giardine B, Riemer C, Hardison RC, Burhans R, Elnitski L, Shah P, Zhang Y, Blankenberg D, Albert I, Taylor J, Miller W, Kent WJ, Nekrutenko A (2005) Galaxy: a platform for interactive large-scale genome analysis. Genome Res 15:1451–1455. doi:10.1101/gr.4086505

Goecks J, Nekrutenko A, Taylor J, The Galaxy Team (2010) Galaxy: a comprehensive approach for supporting accessible, reproducible and transparent computational research in the life sciences. Genome Biol 11(8):R26. doi:10.1186/gb-2010-11-8-r86

Goldstein DB, Schlotterer C (1999) Microsatellites: evolution and applications. Oxford University Press, Oxford

Guichoux E, Lagache L, Wagner S, Chaumeil P, Léger P, Lepais O, Lepoittevin C, Malausa T, Revardel E, Salin F, Petit RJ (2011) Current trends in microsatellite genotyping. Mol Ecol Resour 11:591–611. doi:10.1111/j.1755-0998.2011.03014.x

Illumina (2016) Illumina next-generation sequencing platforms. http://www.illumina.com/systems/sequencing-platform-comparison.html. Accessed 7 April 16

Jehle R, Arntzen J (2002) Review: microsatellite markers in amphibian conservation genetics. Herpetol J 12:1–9

Koressaar T, Remm M (2007) Enhancements and modifications of primer design program Primer3. Bioinformatics 23:1289–1291. doi:10.1093/bioinformatics/btm091

Masella AP, Bartram AK, Truszkowski JM, Brown DG, Neufeld JD (2012) PANDAseq: paired-end assembler for Illumina sequences. BMC Bioinf 13:31. doi:10.1186/1471-2105-13-31

Maudetr C, Miller C, Bassano B, Breitenmoser-Würsten C, Gauthier D, Obexer-Ruff G, Michallet J, Taberlet P, Luikart G (2002) Microsatellite DNA and recent statistical methods in wildlife conservation management: applications in Alpine ibex [*Capra ibex*(*ibex*)]. Mol Ecol 11:421–436. doi:10.1046/j.0962-1083.2001.01451.x

Putman AI, Carbone I (2014) Challenges in analysis and interpretation of microsatellite data for population genetic studies. Ecol Evol. doi:10.1002/ece3.1305

Sefc KM, Payne RB, Sorenson MD (2003) Microsatellite amplification from museum feather samples: effects of fragment size and template concentration on genotyping errors. Auk 120:982–989. doi:10.2307/4090269

Selkoe KA, Toonen RJ (2006) Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. Ecol Lett 9:615–629. doi:10.1111/j.1461-0248.2006.00889.x

Sunnucks P (2000) Efficient genetic markers for population biology. Trends Ecol Evol 15:199–203. doi:10.1016/S0169-5347(00)01825-5

Truelove NK, Griffiths S, Ley-Cooper K, Azueta J, Majil I, Box SJ, Behringer DC, Butler MJ IV, Preziosi RF (2014) Genetic evidence from the spiny lobster fishery supports international cooperation among Central American marine protected areas. Conserv Genet 16:347–358. doi:10.1007/s10592-014-0662-4

Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3—new capabilities and interfaces. Nucleic Acids Res 40:e115. doi:10.1093/nar/gks596

Väli Ü, Einarsson A, Waits L, Ellegren H (2008) To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? Mol Ecol 17:3808–3817. doi:10.1111/j.1365-294X.2008.03876.x

Wetterstrand KA (2012) DNA sequencing costs: data from the NHGRI large scale genome sequencing program. http://www.genome.gov/SequencingCosts/. Accessed 13 Feb 2016

Zalapa JE, Cuevas H, Zhu H, Steffan S, Senalik D, Zeldin E, McCown B, Harbut R, Simon P (2012) Using next-generation sequencing approaches to isolate simple sequence repeat (SSR) loci in the plant sciences. Am J Bot 99:193–208. doi:10.3732/ajb.1100394

Chapter 3: Genetic population structure and genetic diversity of the vulnerable loggerhead sponge *Spheciospongia vesparium* in the Florida Keys and the northern Greater Caribbean

Sarah M. Griffiths[1], Evelyn Taylor-Cox[1], Donald C. Behringer[2,3], Mark J. Butler IV[4], Thierry Pérez[5] & Richard F. Preziosi[1]

[1] Division of Biology and Conservation Ecology, Manchester Metropolitan University, Chester Street, Manchester, UK.

[2] School of Forest Resources and Conservation, Program in Fisheries and Aquatic Sciences, University of Florida, Gainesville, Florida, USA.

[3] Emerging Pathogens Institute, University of Florida, Gainesville, Florida, USA.

[4] Department of Biological Sciences, Old Dominion University, Norfolk, Virginia, USA.

[5] Institut Méditerranéen de Biodiversité et d'Ecologie Marine et Continentale, UMR CNRS 7263, IRD 237, Aix Marseille Université, Marseille, France.

Keywords: microsatellite; Florida Bay; Florida Keys; connectivity; Porifera; marine; Clionaida; Clionaidae.

Acknowledgements

This work was funded by a PhD studentship awarded to SMG by the Natural Environment Research Council (NERC) and by grants awarded to MJB and DCB from The Nature Conservancy/ National Oceanic and Atmospheric Administration - Community Restoration Partnership (GMT-ODU SPONGE-091512) and to MJB from the Everglades National Park. Fieldwork in Barbuda was funded by ship time awarded to TP and Cécile Fauvelot for the PACOTILLES campaigns by Flotte Océanographique Française (http://dx.doi.org/10.17600/15005300; http://dx.doi.org/10.17600/15005200). Thanks to Belize Fisheries Department (permit 0.00023-13), Florida Fish and Wildlife Commission (permit SAL-13-0582A-SR), the US Department of the Interior National Park Service (permit EVER-2014-SCI-0050), the government of Antigua and Barbuda, and the Department of Marine Resources of the

Bahamas (permit MAMR/FIS/1, MAMR/FIS/17) for permission to collect and export samples.

Many thanks for assistance in sampling to FAMRACC (Forest and Marine Reserves Association of Caye Caulker) and F. Clever in Caye Caulker, L. Gittens in Abaco, the Bahamas, Flotte Oceanographique Francias and the PACOTILLES 2015 campaign team in Barbuda, and finally the Old Dominion University and University of Florida 2014 field teams and C. Freeman in the Florida Keys. Thanks to students at the University of Manchester for assistance in processing samples for molecular work, to F. Clever for the Florida Keys ArcGIS map, to N. Truelove for analysis advice, to P. Fullwood, F. Combe and G. Fox at the University of Manchester DNA Sequencing Facility, and to P. Briggs and I. Donaldson at the University of Manchester Bioinformatics Core Facility.

Author contributions

SMG, RFP, DCB and MJB designed the study, SMG, DCB, MJB, and TP collected samples, SMG and ETC carried out lab work, SMG analysed the data and wrote the chapter, RFP, DCB, MJB and TP secured funding for the work.

ABSTRACT

Sponges are dominant, habitat-providing benthic invertebrates in Florida Bay's hard bottom habitats (Florida Keys, USA). As such, mass mortalities in these sponge communities have seen numerous negative consequences befall the ecosystem. Understanding genetic diversity and the genetic population structure of sponges in Florida Bay, and the scale at which their populations are structured will help in conservation efforts to predict resilience and recovery patterns. Here, we developed and used twelve microsatellite markers to study population genetics in the largest Florida Bay sponge, *Spheciospongia vesparium*, among four locations in the northern Greater Caribbean area: Belize, the Florida Keys, the Bahamas and Barbuda, and within fourteen sites across the Florida Keys. We found that the four main locations were all genetically different, suggesting limited connectivity over large spatial scales. However we also identified the Florida Current as an oceanographic barrier to connectivity between the Bahamas and Florida. Within Florida, we found a weak isolation by distance pattern with spatial genetic patchiness. This shows that as well as distance-limited dispersal, other factors influence small-scale spatial structure in *S. vesparium*, potentially including demographic stochasticity coupled with temporal variability in water circulation. Genetic diversity levels were broadly similar across all sites, but signs of inbreeding and bottleneck signatures were apparent in Florida. Together, these results can be used to assess the genetic health of vulnerable *S. vesparium* populations and predict their resilience and recovery in the face of mortalities.

INTRODUCTION

Marine sponges are a fundamental part of the Greater Caribbean seascape (Diaz & Rutzler, 2001). This sessile, benthic, invertebrate taxon is key in nutrient cycling dynamics (de Goeij et al., 2013; Fiore, Freeman & Kujawinski, 2017), and it has numerous interactions with other organisms as both a habitat and food source (Pearse, 1950; Dunlap & Pawlik, 1998). In hard bottom habitats of Florida Bay, a subtropical lagoon located between the Florida Keys and the Everglades, sponge communities have an additional role as the dominant, structure-forming component of the benthos (Chiappone & Sullivan, 1994; Tellier & Bertelsen, 2008; Bertelsen et al., 2009). Hard bottom areas are interspersed with seagrass beds to form a productive nursery habitat, where commercially valuable juvenile spiny lobsters (*Panulirus argus*) are among the many species that shelter here before migrating out to coral reefs or other habitats in adulthood (Forcucci, Butler & Hunt, 1994; Poulakis & Seitz, 2004). Unfortunately, the Florida Bay ecosystem has suffered declining health since the 1980s with a series of recurrent ecological disturbances. This began with seagrass die-offs (Robblee et al., 1991), followed by blooms of cyanobacteria, which then caused mass mortalities in sponge communities (Butler et al., 1995; Paul et al., 2005; Berry et al., 2015). It is clear that these sponge mortalities have had many negative ecosystem-level consequences, notably habitat loss (Butler et al., 1995; Herrnkind et al., 1997), increased risk of phytoplankton blooms (Peterson et al., 2006) and loss of acoustic recruitment cues for larvae (Butler, Stanley & Butler, 2016).

The cyanobacterial blooms are transient and occur in localised areas, and thus there are still surviving sponge communities in the Bay. Despite this, repopulation of barren areas by new sponge recruits has been lacking. This raises a number of interesting questions regarding the dispersal capacity of Florida Bay's sponges and the degree of connectivity among geographically proximate sites. Sponges have small lecithotrophic larvae (Maldonado, 2006), and thus have limited energy supplies to sustain themselves during the pelagic phase. As such, larval duration can be a matter of hours to days in most sponge

species (e.g. Maldonado and Young 1999; Whalan et al., 2008; Bergquist and Sinclair 2010), and so dispersal capacity is expected to be restricted. Indeed, population structure in Mediterranean encrusting sponge *Scopalina lophyropoda* was found to occur among individuals collected within 100 meters of each other (Blanquer, Uriz & Caujapé-Castells, 2009), and populations of reef-dwelling sponges have been found to be structured in the Florida Keys (DeBiasse, Richards & Shivji, 2010; Chaves-Fonnegra et al., 2015; Richards et al., 2016). However, the population genetics of dominant sponge species in Florida Bay have not yet been investigated.

Using genetic markers to investigate the distribution of genetic diversity can allow us to infer gene flow patterns and effective dispersal among sites. Connectivity (the movement of larvae or gametes among sites) bolsters resilience and persistence, whereas isolated sites that rely heavily on self-recruitment can be vulnerable to population declines (Cowen & Sponaugle, 2009; Saenz-Agudelo et al., 2011; van der Meer et al., 2013). Connectivity also helps genetic diversity to be maintained, which is an important component of resilience through the maintenance of evolutionary potential (Ehlers, Worm & Reusch, 2008; Evans, Vergés & Poore, 2017). In populations that have experienced rapid declines in abundance, genetic variation can be lost due to bottleneck effects, leading to a potential loss in adaptive capacity as well as increased inbreeding (Reed & Frankham, 2003; Bristol et al., 2013). *Spongia lamella*, *S. officinalis* and *Ircinia fasciculata* are three sponge species that have suffered notable population declines, and while genetic diversity is low and bottleneck signatures are present in *S. lamella* (Pérez-Portela, Noyer & Becerro, 2015), *S. officinalis* and *I. fasciculata* still have high levels of genetic diversity (Dailianis et al., 2011; Riesgo et al., 2016). It is therefore important to examine genetic diversity in potentially bottlenecked populations on a case-by-case basis in order to implement species-specific conservation action plans.

To this end, this study aimed to investigate patterns of population structure and to measure genetic variation in the sponge with the largest biomass in Florida Bay (Tellier & Bertelsen, 2008), the loggerhead sponge *Spheciospongia*

vesparium (Lamarck, 1815; Clionaida; Clionaidae). *Spheciospongia vesparium* provides a microhabitat (Pearse, 1950; Duffy, 1996) and nutrient cycling services (Fiore, Freeman & Kujawinski, 2017), as well as providing structure in Florida Bay. In addition to sampling multiple sites in Florida, we also sampled three other locations in the Greater Caribbean area in order to compare genetic diversity levels and to observe population structure at larger spatial scales.

METHODS

Sample collection and preservation

Sampling was conducted within four areas of the northern Greater Caribbean – Belize, the Florida Keys, the Bahamas and Barbuda (Figure 1, Table 1). Within the Florida Keys, sampling was conducted in multiple sites from the Upper, Middle and Lower Keys to study connectivity on a smaller spatial scale (5 - 160 km). Most sampling sites were located on the Bay side (Gulf side) of the Keys, but in the Middle Keys, sampling was also conducted at two sites on the Atlantic side to examine potential population structure between opposing sides of the islands. In each of the other areas (Belize, the Bahamas and Barbuda), samples were collected at a single site. In all cases, samples were collected whilst snorkelling in shallow lagoons (< 2 m depth); small pieces of sponge tissue were cut (avoiding new recruits) and immediately transferred to 95% ethanol upon surfacing. Ethanol was replaced after 24 hours and samples were stored at the coldest temperature available (ranging from -20°C to 20°C), before shipment to Manchester where they were stored at -80°C until further processing commenced.

DNA extraction

Before processing for DNA extraction, samples were checked under a dissecting microscope and any visible endosymbiotic invertebrates were removed. Total DNA was extracted from sponge tissue using the DNeasy® Blood and Tissue Kit (Qiagen) (spin column method) according to the manufacturers protocol. DNA was checked for quality and quantity using a NanoDrop™ 3000 spectrophotometer (Thermo Fisher Scientific) and agarose gel electrophoresis, before storage at -20°C.

Microsatellite development and genotyping

For this study, twelve tri- and tetra-nucleotide microsatellite loci were newly characterised. A single sample collected from Long Key (Florida, USA) was processed for Illumina sequencing. The sample was studied under a stereomicroscope to remove any contaminating invertebrates and the tissue

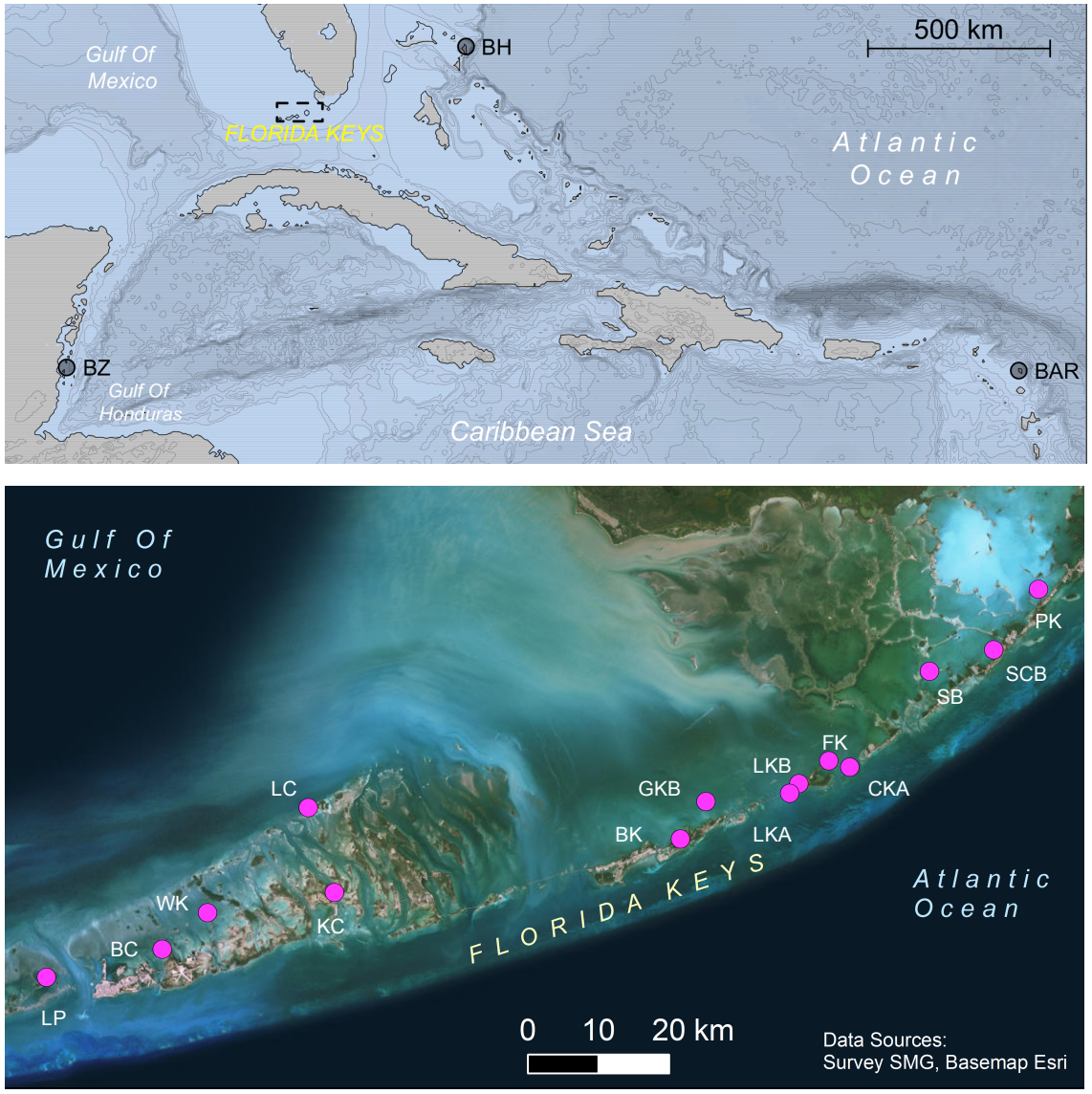


Figure 1: Maps showing *Spheciospongia vesparium* sampling locations. Top map shows the four main sampling locations (data source: ETOPO1, NOAA), bottom map shows the sites sampled within the Florida Keys (see Table 1 for full site names).

Table 1: Sampling locations for *Spheciospongia vesparium*.

Location	Site	ID	Latitude, Longitude	Date	N
Florida Keys (Upper Keys), USA	Pigeon Key	PK	25.0594, -80.4979	7/14	18
	Snake Creek Basin	SCB	24.9831, -80.5602	7/14	17
	Steamboat Channel	SB	24.9559, -80.6492	7/14	19
Florida Keys (Middle Keys, Bay side) USA	Fiesta Key	FK	24.8430, -80.7890	7/14	18
	Long Key (Bay side)	LKB	24.8143, -80.8307	7/14	18
	Grassy Key Bank	GKB	24.7917, -80.9598	7/14	17
	Bamboo Key	BK	24.7442, -80.9950	7/14	19
Florida Keys (Middle Keys, Atlantic side), USA	Craig Key (Atlantic side)	CKA	24.8350, -80.7599	6/16	10
	Long Key (Atlantic side)	LKA	24.8021, -80.8435	6/16	17
Florida Keys (Lower Keys), USA	Kemp Channel	KC	24.6768, -81.4757	7/14	20
	Little Crane Key	LC	24.7840, -81.5120	7/14	20
	Waltz Key	WK	24.6510, -81.6521	7/14	17
	Boca Chica Channel	BC	24.6049, -81.7150	7/14	19
	Lakes Passage	LP	24.5694, -81.8757	7/14	32
The Bahamas	Mermaids Reef, Abaco	BH	26.5537, -77.0527	7/15	12
Antigua and Barbuda	Barbuda lagoon	BAR	17.6547, -61.8527	5/15	20
Belize	Caye Caulker	BZ	17.7422, -88.0354	5/13	10

N: Number of samples successfully genotyped and used in analysis.

was processed to separate eukaryote and prokaryote cells (as far as possible) following the methods of Freeman and Thacker (2011). DNA was extracted from the resulting pellet of 'eukaryote' cells using the Qiagen DNeasy® Blood and Tissue Kit, and DNA was subsequently concentrated by vacuum centrifugation. Paired-end library construction was performed using 50 ng of DNA and the Nextera® DNA Sample Preparation Kit, before paired-end sequencing (2 x 250 bp) in half a flow cell lane on the Illumina MiSeq platform (i.e. the lane was shared with one other sample unrelated to this study). 2 x 3,051,330 reads were produced by the sequencing run, which were processed using the Galaxy bioinformatics pipeline of Griffiths et al. (2016) (Chapter 2), using the settings recommended by the authors. Following the quality-filtering step, 2 x 2,960,928 reads remained. Thirty-six loci were tested; twelve could be successfully amplified and scored, and were thus subsequently used in this study. Loci were excluded from further analysis due to inconsistent amplification, complicated stutter patterns or non-specific amplification.

To facilitate faster and more economic genotyping of samples, ten of the twelve primer pairs were combined in two multiplex (5-plex) PCRs designed with Multiplex Manager (Holleley & Geerts, 2009), using the fluorophores 6FAM and HEX (see Table 2). A three-primer universal tail system, as described in Blacket et al. (2012) and Culley et al. (2013), was used to fluorescently label PCR products using the following tail and dye combinations: Tail 1: 6FAM-GCCTCCCTCGCGCCA; Tail 2: HEX- GCCTTGCCAGCCCGC (Blacket et al. 2012). Two primer pairs, Vesp36 and Vesp9, were run in singleplex PCRs with the 'Tail 1' combination due to problems encountered in multiplexing these loci. PCRs were carried out using the Type-it® Microsatellite PCR Kit (Qiagen) in 5 µl reaction volumes, according to the manufacturers instructions. Cycling conditions used were: 95°C for 5 minutes, 28 x (95°C for 30 seconds, 60°C for 90 seconds, 72°C for 30 seconds), 60°C for 30 minutes. Any loci that failed to amplify were repeated in singleplex PCR with lowered (50-59°C) annealing temperatures.

PCR products were sized using capillary electrophoresis on a 3730 DNA Analyzer (Thermo Fisher Scientific) either at the University of Manchester DNA Sequencing Facility (UK) with GeneScan™ 500, 600 or 1200 LIZ® size standard (Thermo Fisher Scientific), or at the Smithsonian Institute in Washington D.C. (USA) using a homemade ROX-based size standard. On all plates, both positive and negative controls were included. Alleles were scored using Genemapper® v3.7 software (Thermo Fisher Scientific), and allele sizes were adjusted according to the positive controls to account for differences in allele length based on the machine or size standard used. Alleles were then binned using the R package MsatAllele v1.02 (Alberto, 2009). PGDSpider v2.1.0.3 (Lischer & Excoffier, 2012) was used to convert data files to the formats required by subsequently used programs if they did not have their own internal conversion functions.

Quality control and summary statistics

Samples in which over half of the loci did not amplify or produced ambiguous plots were excluded from further analysis. Probability of linkage disequilibrium

between pairs of loci was calculated using Genepop on the Web v4.2 (Raymond and Rousset 1995; Rousset 2008) with significance levels adjusted for multiple tests using the 'B-Y' false discovery rate (FDR) correction (Benjamini & Yekutieli, 2001; Narum, 2006).

Null allele frequency at each locus was estimated using the EM algorithm (Dempster, Laird & Rubin, 1977) in FreeNA (Chapuis and Estoup 2007). Due to the presence of null alleles (see Results), and the propensity of null alleles to cause overestimation of F_{ST} values and levels of population differentiation (Chapuis & Estoup, 2007), a post hoc analysis was conducted as follows to test the extent of any bias: Average null allele frequency was calculated for all loci, and global F_{ST} was calculated with and without ENA correction for null alleles (as implemented in FreeNA; Chapuis and Estoup 2007); the locus with the highest null allele frequency was removed and the uncorrected and corrected global F_{ST} values were recalculated. This was repeated systematically until just one locus remained. This allowed us to observe the cumulative effects of each locus and their null allele loads on F_{ST} by comparing the corrected and uncorrected values (see Results).

Observed heterozygosity (H_O) and expected heterozygosity (H_E) were calculated in GenoDive v2.032b (Meirmans & Van Tiendener, 2004); probability of departures from Hardy-Weinberg Equilibrium (HWE) were tested in the same program using the AMOVA (least squares) method and 50,000 permutations (significance corrected for multiple tests using the B-Y FDR method). INEst v2.1 (Chybicki & Burczyk, 2009) was used to estimate inbreeding coefficients (Avg Fi) corrected for the presence of null alleles. The program includes three possible parameters that can affect inbreeding coefficient estimation: null alleles ('n'), inbreeding ('f') and genotyping failure ('b'). The individual inbreeding model (IMM) was run for all combinations of these parameters, and the Deviance Information Criteria (DIC) were calculated for each run to determine the best model fit for the data. The program was run with 500,000 Monte Carlo Markov Chain (MCMC) cycles with 50,000 burnin cycles. Average allelic richness and private allele richness rarefied to the lowest sample size (maximum g = 10)

were calculated in ADZE v1.0 (Szpiech, Jakobsson & Rosenberg, 2008). These analyses were repeated with all the Florida Keys sites grouped as one population, as well as testing each separately.

Population genetic structure

Population differentiation was estimated by calculating pairwise F_{ST} (Wright, 1943, 1949) and *D* (Jost 2008) between sites in GenoDive. The significance of the pairwise F_{ST} values was assessed using 50,000 permutations (also carried out in GenoDive), and B-Y FDR correction applied to the p values. To test the presence of isolation by distance (IBD) signals in the data, a Mantel test was performed to detect association between matrices of linearised pairwise genetic distances (F_{ST}/[1-F_{ST}]) and the logarithm of geographic distances. Least-cost oceanographic distances between sites were calculated using the 'marmap' package (Pante & Simon-Bouhet, 2013) in R v3.3.3 (R Core Team 2017), and the Mantel test was carried out in the 'ade4' package (Dray, Dufour & Chessel, 2007) in R, with 9999 permutations to calculate significance. This was carried out over all sites and within Florida.

Two different approaches were used to infer the number of population clusters ('*K*') in the data. Firstly we used the Bayesian individual-based assignment model implemented in the 'Geneland' package (Guillot, Mortier & Estoup, 2005; Guillot, Santos & Estoup, 2008) in R. This uses spatial information (i.e., geographic coordinates for sampling locations) and genetic data to infer *K* and calculate the probability of individual assignment. The spatial and null allele models were used, and the model was run with 1,000,000 MCMC iterations, 100 thinning and 1000 burnin, and uncertainty on coordinates set to 0.0005. The maximum number of nuclei was set to 909, and the maximum rate of the Poisson process to 303. This was repeated for 10 independent runs with *K* ranging from one to ten. The model was run twice to test both the correlated and uncorrelated allele frequencies models. Due to the assumptions of this model, only the six loci that did not deviate from HWE in over half of the sites were used for this analysis.

I also used FLOCK v3.1 (Duchesne & Turgeon, 2012) to infer membership to population clusters, which uses an alternative method for estimating *K* and partitioning samples into *K* clusters based on iterated reallocation rather than the MCMC sampling-based methods of Geneland. This method uses no *a priori* information such as sampling locations, and does not assume populations are in HWE. The program was run with 20 iterations per value of *K* and 50 runs, and plateau analysis as described by Duchesne and Turgeon (2012) was used to infer the most likely value of *K*. Both Geneland and FLOCK were carried out using all sites, and then on any clusters identified by the programs in a hierarchical clustering approach.

Multivariate analysis was also used to investigate the relationship among sites, due to its power and flexibility in its lack of HWE assumptions. A Principle Coordinates Analysis (PCoA) was conducted in GenAlEx v6.502 (Peakall & Smouse, 2012) using pairwise F_{ST} as calculated previously. Analyses were conducted first on all sites separately, and then on only the Florida sites.

Analyses of Molecular Variance (AMOVAs) were conducted in GenoDive using the Infinite Allele Model with 50,000 permutations to examine the partitioning of genetic diversity within and among sites and to test their significance. As AMOVA allows hierarchical structuring, the sites were further grouped by location (i.e., Florida, the Bahamas, Barbuda and Belize). A second AMOVA was also conducted on the Florida sites only, which were further grouped into Upper Keys, Middle Keys (Bay side), Middle Keys (Atlantic side) and Lower Keys as described in Table 1.

GENECLASS2 v2 (Piry et al., 2004) was used to detect the number of first generation migrants among the sampling locations, and their putative population origins. The Bayesian criteria of Rannala and Mountain (1997) was used for likelihood estimation, and the Monte Carlo method of Paetkau et al. (2004) was used for probability computation, with the L_{home} criterion, as source populations for all individuals were unlikely to have been sampled. Significance

was tested at p = 0.01 and p = 0.05, and simulations were run with 10,000 individuals.

INest v2.1 (Chybicki & Burczyk, 2009) was used to find evidence of genetic signatures of recent population bottleneck events. The program implements two tests – the first identifies heterozygosity excesses in respect to allelic richness (Cornuet & Luikart, 1996), and the second identifies M-ratio (mean ratio of allelic richness to allelic size) deficiencies (Garza & Williamson, 2001). Both phenomena have been observed when populations experience rapid reductions in population sizes. The two-phase mutation model was used, as the loci did not always mutate in a stepwise fashion, and significance was tested with a Wilcoxon signed-rank test with 1000 permutations.

RESULTS

Quality control and summary statistics

In total, 326 individuals were collected from seventeen sites (Table 1) and genotyped at the twelve microsatellite loci (Table 2). Twenty-two samples were removed from the final dataset due to amplification failure in over 50% of the loci, leaving 304 individuals. Two individuals from the Lakes Passage (LP) had identical genotypes, one of which was removed from the dataset for analysis, yielding 303 individuals. Following B-Y FDR correction, no significant linkage disequilibrium was found between pairs of loci. Null allele frequency was high in some markers (Table 2, Table 3). However, the post hoc analysis conducted on the loci showed that the null allele-corrected global F_{ST} value was only marginally higher (0.002) than the uncorrected value when all loci were included in the analysis (Table 3). Furthermore, the difference between the uncorrected and corrected F_{ST} did not significantly increase as more loci were added (linear regression, r^2 = -0.03608, p = 0.4504), and therefore all loci were retained for the population genetics analysis. The number of alleles observed per locus ranged from four (Vesp23) to 27 (Vesp30). Significant departures from HWE were found in a number of loci and populations (Table 4), and all populations showed lower than expected levels of heterozygosity (Table 5).

The DIC analysis in INEst determined either the 'nfb' (null allele, inbreeding and genotyping failure) or 'nb' (null allele and genotyping failure) models to be the best fit for the sites in this study (Table 5). This indicates that null alleles and genotyping failure would affect inbreeding coefficient estimations in all the sites, but in ten of the sites, inbreeding was also an influential component of the model. The null allele- corrected inbreeding coefficients were positive in all locations, ranging from 0.036 (the Bahamas) to 0.343 (CKA). However the posterior 95% probability intervals included zeros in all sites when tested separately, and therefore F_{IS} cannot be considered to be significantly above zero. When the Florida sites were grouped together, the posterior 95% probability interval was above zero, which may be indicative of significant inbreeding across the area (but see Discussion for further examination

Table 2: Characterisation of 12 polymorphic microsatellite loci and two multiplexes for *Spheciospongia vesparium.*

Locus name	Motif	Primer sequences (5'→ 3')	MP	L	Na	Size range (bp)	N	GenBank accession no.
Vesp15	TTC	F: AGAAGGGTTTAAAAGAAGCAGCAGAAGGG R: TATTGTGAGATATCACTTCCACGACCAGC	A	1	17	223-300	0.012	KX758634
Vesp23	TTC	F: CTAGAAGATCAACTCCTTGACCTTGGGC R: TGAGGATGATTTCGATGAAGTACCG	A	2	4	202-238	0.201	KX758641
Vesp35	AGG	F: ACCCCAGTCCGAGTACATCATCAGG R: ATGATTCCCGAACAGAAGTGAGTGC	A	2	14	447-468	0.040	KX758643
Vesp3	AAC	F: TATTATGCTGCAGTGTATTCAGCATCTCC R: CTCTTCCCTTTGGCTCACAGTATCCC	A	1	10	380-411	0.027	KX758633
Vesp27	ACC	F: TTCTTACACAATCTACCAATCCTTGCAGC R: CACACTGTGATCTATTTTAATGTCCCTCC	A	2	25	291-391	0.252	KX758642
Vesp1	ATAC	F: TGGTTCATAATTGTAGCAACTAATCCCGC R: AAGTATGCGTTTGAGCAAGTCTGAAAAGG	B	2	14	174-255	0.232	KX758638
Vesp30	ACGC	F: GGATCATCAAGATGTTTCTCAAGGTCAGC R: TTTGGTCCTGTTACACACAAATTGTAGCC	B	2	27	278-404	0.188	KX758636
Vesp17	AGTG	F: CTAACTTTAGAATGCACTGCAGCAGAAGG R: ATAGTGAGCCTACTACACTGCTGACCTGC	B	1	17	391-445	0.027	KX758635
Vesp19	TTG	F: CTTAGGGTGCGTCTGTTACCCATTACG R: CCATACGCTTAGCGAAACTTCATTCTACG	B	1	10	330-354	0.122	KX758639
Vesp22	ATAC	F: CTAGTATGTGTGATCCTGATATTGTACTGC R: GTTATTGCTATGTTATTACCCTGAGGTGG	B	1	19	228-286	0.273	KX758640
Vesp36	ATG	F: GGCCACGGACACTAACAGAAAATGG R: TGGAGTTACGAAAGAATCTCACTTTGTTGG	-	1	6	110-131	0.127	KX758644
Vesp9	TCC	F: ACCATCACTTCCTCCACCTCCC R: TCAGTCAAA GCAAAACCTAGACTGAGGG	-	1	12	273-324	0.192	KX758637

NA: number of alleles per locus; MP: multiplex; L: tail sequence/ florescent label combination (1: 6FAM-GCCTCCCTCGCGCCA; 2: HEX-GCCTTGCCAGCCCGC); bp: base pairs; N: frequency of null alleles.

Table 3: Cumulative average null allele frequency (from lowest to highest individual locus null allele frequency), with global F_{ST} corrected with ENA (with 1000 bootstrap replicates) and uncorrected for null alleles.

No. loci →	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
Global F_{ST}	0.092	0.163	0.124	0.142	0.122	0.105	0.102	0.105	0.101	0.095	0.087	0.087
Global F_{ST} (corrected)	0.091	0.159	0.121	0.140	0.121	0.105	0.099	0.101	0.097	0.093	0.086	0.085
Mean null allele frequency	0.012	0.019	0.022	0.026	0.045	0.059	0.077	0.092	0.104	0.117	0.129	0.141

Table 4: Probability of departure from Hardy-Weinberg Equilibrium for each population and locus.

	Vesp15	**Vesp23**	**Vesp35**	**Vesp3**	**Vesp27**	**Vesp1**	**Vesp30**	**Vesp17**	**Vesp19**	**Vesp22**	**Vesp36**	**Vesp9**
PK	0.579	**0.005**	0.554	-	**0.007**	**0.005**	**0.004**	0.507	0.012	0.941	0.290	**0.002**
SCB	0.433	0.090	**0.003**	1.000	0.008	0.490	**0.001**	0.680	0.452	**0.000**	**0.000**	**0.000**
SB	0.526	0.038	0.216	0.964	**0.000**	**0.000**	0.129	0.192	**0.002**	**0.000**	0.554	**0.000**
FK	0.727	**0.000**	0.868	0.037	**0.000**	**0.000**	0.559	0.523	0.008	**0.000**	0.185	**0.003**
LKB	0.115	**0.000**	0.015	0.767	**0.000**	**0.000**	**0.000**	0.554	0.093	**0.000**	**0.000**	**0.000**
GKB	0.623	**0.003**	0.669	0.178	**0.000**	**0.000**	0.016	0.276	0.013	**0.000**	0.488	0.843
BK	0.078	**0.001**	0.132	1.000	**0.000**	**0.003**	0.092	0.582	**0.000**	**0.000**	0.441	**0.000**
CKA	0.200	**0.001**	1.000	0.695	**0.000**	**0.000**	0.008	0.011	0.029	**0.000**	0.060	0.008
LKA	0.312	**0.002**	0.796	0.599	**0.000**	**0.000**	0.018	0.399	0.044	**0.000**	**0.001**	**0.002**
KC	0.266	0.010	0.226	0.219	**0.000**	**0.000**	0.026	0.498	0.135	**0.000**	0.370	**0.000**
LC	0.351	0.016	0.104	0.324	**0.000**	**0.000**	**0.002**	0.336	0.439	**0.000**	**0.007**	0.103
WK	0.499	**0.000**	0.562	0.098	0.019	**0.000**	**0.001**	0.054	**0.000**	0.016	0.026	**0.000**
BC	0.579	0.056	0.140	0.960	**0.000**	0.011	0.034	0.221	0.008	**0.000**	0.569	**0.004**
LP	0.171	**0.003**	0.444	0.335	**0.000**	**0.000**	**0.000**	0.534	0.253	**0.000**	**0.000**	**0.005**
BH	0.600	**0.002**	0.403	0.642	**0.004**	0.111	**0.000**	0.358	0.264	**0.000**	0.104	0.597
BAR	0.444	0.021	0.233	0.422	0.008	**0.000**	**0.000**	0.381	0.024	**0.001**	**0.001**	0.127
BZ	0.256	0.134	0.680	0.501	**0.000**	0.395	**0.000**	0.474	0.237	**0.001**	0.158	0.234

Significant p values after B-Y FDR highlighted in bol

Table 5: Genetic diversity and summary statistics for *Spheciospongia vesparium* sites.

Site	AR *(SE)*	PR *(SE)*	H_O	H_E	Avg Fi	95% HPDI
PK	3.408 *(0.421)*	0.054 *(0.029)*	0.397	0.569	0.0995	0 – 0.2662
SCB	3.698 *(0.378)*	0.115 *(0.073)*	0.368	0.595	**0.1875**	0 – 0.3516
SB	3.857 *(0.379)*	0.030 *(0.011)*	0.427	0.655	**0.1089**	0 – 0.2566
FK	3.699 *(0.411)*	0.158 *(0.070)*	0.386	0.611	0.0499	0 – 0.1574
LKB	4.399 *(0.360)*	0.211 *(0.063)*	0.339	0.711	**0.1698**	0 – 0.3972
GKB	3.638 *(0.267)*	0.065 *(0.033)*	0.394	0.636	0.0546	0 – 0.1623
BK	3.429 *(0.315)*	0.022 *(0.013)*	0.383	0.612	**0.1608**	0 – 0.3016
CKA	3.844 *(0.349)*	0.112 *(0.061)*	0.251	0.650	**0.3433**	0 – 0.6069
LKA	3.742 *(0.211)*	0.077 *(0.034)*	0.365	0.638	**0.1509**	0 – 0.2773
KC	3.862 *(0.297)*	0.134 *(0.086)*	0.379	0.659	0.0785	0 – 0.2180
LC	3.823 *(0.304)*	0.020 *(0.008)*	0.418	0.645	0.0403	0 – 0.1213
WK	3.819 *(0.256)*	0.139 *(0.050)*	0.356	0.662	**0.2187**	0 – 0.4180
BC	3.883 *(0.376)*	0.100 *(0.045)*	0.417	0.644	**0.1138**	0 – 0.2655
LP	3.806 (0.315)	0.122 (0.063)	0.412	0.635	**0.0622**	0 – 0.1760
BH	4.375 *(0.340)*	0.819 *(0.253)*	0.504	0.735	0.0361	0 – 0.1191
BAR	3.940 *(0.331)*	0.956 *(0.351)*	0.404	0.653	**0.0816**	0 – 0.1700
BZ	4.390 *(0.414)*	0.929 *(0.310)*	0.481	0.713	0.0367	0 – 0.1277
FL	4.080 *(0.332)*	1.095 *(0.133)*	0.383	0.655	**0.0547**	0.0025 – 0.1002
BH	4.375 *(0.340)*	1.538 *(0.324)*	0.504	0.735	0.0361	0 – 0.1191
BAR	3.940 *(0.331)*	1.344 *(0.392)*	0.404	0.653	**0.0816**	0 – 0.1700
BZ	4.390 *(0.414)*	1.568 *(0.386)*	0.481	0.713	0.0367	0 – 0.1277

AR *(SE)*: Average rarefied allelic richness (Standard error); PR *(SE)*: Average rarefied private allelic richness (Standard error); H_O: Observed heterozygosity; H_E: Expected heterozygosity; Avg Fi: Null allele corrected inbreeding coefficient (values in bold denote sites where the 'nfb' [null alleles, inbreeding and genotyping error] model has the lowest DIC, values not in bold denote where the 'nb' [null alleles and genotyping] model has the lowest DIC; 95% HDPI: Posterior 95% probability intervals. Analyses repeated for all Florida sites grouped together ('FL').

of this result). Private allelic richness and allelic richness were largely consistent among sites (Table 5).

Population structure

Global F_{ST} was 0.085 (corrected for null alleles), indicating moderate population structure, and pairwise F_{ST} ranged from -0.019 (no differentiation) between the Atlantic side sites of Craig's Key (CKA) and Long Key (LKA), to 0.273 (great differentiation) between Pigeon Key, Florida (PK) and Barbuda (BAR) (Table 6). Among the four main geographical locations (Florida, the Bahamas, Barbuda, Belize), F_{ST} values were large and significant, showing strong differentiation. Among sites within Florida, differentiation was lower (≤ 0.116), but differentiation was present between many pairs of sites, and many pairwise comparisons were significant. In general, more differentiation could be observed between the Upper and Lower Keys sites than among comparisons involving the Middle Keys sites, but patchiness can be observed throughout. Patterns of *D* were similar, and ranged from -0.035 (between the Atlantic side sites, as previously) to 0.668 (between the Bahamas [BH] and Waltz Key [WK]) (Table 6). Isolation by distance was very strong and significant over all sites ($r = 0.792$, $p < 0.001$), and was also significant within Florida, although the effect was much weaker ($r = 0.229$, $p = 0.031$) (Figure 2). When the analysis was replicated with Jost's *D*, results were very similar (all sites: $r = 0.717$, $p = 0.009$; Florida: $r = 0.225$, $p = 0.033$). Without the six loci that deviated from HWE in more than half the populations, IBD over all sites was very similar to estimates with the full set of loci ($r = 0.767$, $p = 0.001$) but was not significant when tested in Florida only ($r = 0.138$, $p = 0.132$).

For both the uncorrelated and correlated allele frequency models in Geneland, *K*=4 was found for each independent run. This showed each sampling location as a separate population cluster, with probability of assignment to the cluster as follows (uncorrelated model): Florida: 0.492; the Bahamas: 0.482; Barbuda: 0.481; Belize: 0.481. In contrast, FLOCK showed strong evidence for *K*=2. Samples were broadly partitioned into a Florida cluster, and then a cluster with Belize, Barbuda and the Bahamas. Two individuals from Florida (one from

Craig's Key Atlantic side [CKA] and one from the Lakes Passage [LP]) were put into the Belize, Barbuda and the Bahamas cluster, otherwise clustering was concurrent with sampling locations. When FLOCK was carried out on the Barbuda, Belize and the Bahamas samples separately, the samples were partitioned into K = 3 concurrent with sampling locations. When running Geneland with just the Florida samples, the model was unable to converge, indicating that either K=1 or strong isolation by distance in the data, and no plateau was obtained in FLOCK, indicating K=1.

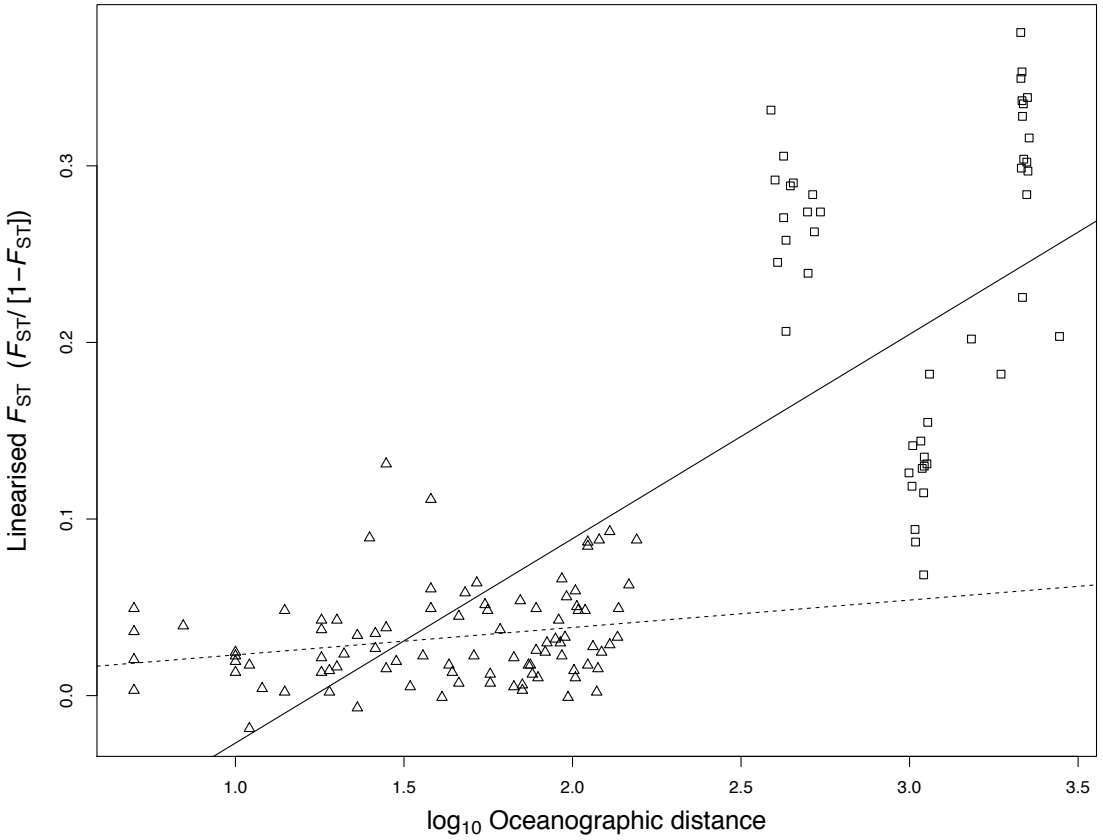


Figure 2: Genetic isolation by distance in *Spheciospongia vesparium* using pairwise calculations of linearised F_{ST} (F_{ST}/[1-F_{ST}]) and the logarithm of oceanographic distance. Triangles represent pairwise comparisons among only the Florida sites; squares represent pairwise comparisons including Belize, the Bahamas and Barbuda. Dashed line shows a regression for the Florida sites only, solid line shows a regression for all of the sites together.

In the PCoA carried out on all sites (Figure 3), the first coordinate accounted for 45.86% of the observed variation, and the second coordinate explained 17.28% of the variation. The first axis separated Florida from the Bahamas, Belize and

Barbuda, and the second separated the Upper Keys from the Lower Keys and Atlantic side Middle Keys; the Bay side Middle Keys were distributed among both. In the Florida-only PCoA (Figure 3), the first coordinate explained 48.24% of the variation, the second 17.72% of the variation. In this analysis, samples were distributed in a loose isolation by distance fashion, but notably the Atlantic side Middle Keys sites were clustered with Waltz Key (WK), and separated from the Bay side Middle Keys sites. When the analysis was replicated with Jost's *D* instead of F_{ST}, the patterns observed were very similar (data not shown).

The AMOVA showed that 14% of variance observed was found among the four locations and only 1.4% found among sites within locations (Table 7). As only the Florida region contained multiple sites, this shows that variation within Florida was small in comparison to variation overall. A further test carried out only within Florida showed that 1.4% of variance were found among sites within the Upper, Middle, Atlantic side and Lower Keys groups and 0.9% among the groups. In both tests, more variation was found among individuals within sites (25.5% and 30.7% respectively).

Migrants analysis

Three putative first generation migrants were detected under the significance threshold of p = 0.01. All potential migrants were found in the Florida Keys, and their origin was within the Florida Keys. Two migrants were found in Long Key Bay side (LKB), with origins in Waltz Key (WK; $p = 0.0007$; distance 89 km) and Fiesta Key (FK, p = 0.0019, distance 43 km), and the final migrant was found in Little Crane Key (LC) with origins inferred in Kemp Channel (KC; $p < 0.0001$, distance 12 km).

Bottleneck analysis

Deficiencies in M-ratios were found at four sites, giving evidence of recent bottleneck events. The sites where this was detected were Boca Chica Channel (BC; p = 0.0385), Little Crane Key (LC; p = 0.000) and both the Atlantic side sites (Craig's Key and Long Key [CKA, LKA; p = 0.000 for both]). However,

Table 6: Pairwise F_{ST} (below diagonal, shaded grey) and Jost's *D* (above diagonal, not shaded) values between population pairs.

	PK	SCB	SB	LKB	FK	GKB	BK	LKA	CKA	KC	LC	WK	BC	LP	BH	BAR	BZ
PK	-	0.027	0.028	0.031	0.089	0.074	0.055	0.091	0.166	0.141	0.133	0.147	0.078	0.138	0.613	0.593	0.311
SCB	0.019	-	0.029	0.010	0.058	0.071	0.035	0.077	0.205	0.098	0.081	0.149	0.048	0.100	0.564	0.577	0.280
SB	0.016	0.017	-	-0.016	0.023	0.042	-0.002	0.029	0.076	0.081	0.055	0.094	0.003	0.059	0.557	0.560	0.276
LKB	0.017	0.005	-0.007	-	0.006	0.004	0.027	0.041	0.082	0.012	0.004	0.071	-0.002	0.052	0.544	0.476	0.172
FK	**0.057**	**0.037**	0.013	0.003	-	0.003	0.054	0.038	0.055	0.031	0.020	0.118	0.016	0.025	0.620	0.606	0.257
GKB	**0.046**	**0.043**	0.022	0.002	0.002	-	0.082	0.024	0.039	0.024	0.013	0.033	0.044	0.025	0.626	0.607	0.259
BK	**0.036**	0.022	-0.001	0.014	**0.033**	**0.047**	-	0.081	0.145	0.113	0.087	0.093	0.018	0.094	0.592	0.523	0.273
LKA	**0.055**	**0.047**	0.015	0.020	0.022	0.013	**0.046**	-	-0.035	0.010	0.038	0.056	0.041	0.030	0.564	0.597	0.234
CKA	**0.100**	**0.116**	0.041	0.038	0.035	0.023	**0.082**	-0.019	-	0.045	0.085	0.062	0.083	0.037	0.616	0.620	0.270
KC	**0.080**	**0.056**	**0.041**	0.006	0.017	0.012	**0.060**	0.005	0.025	-	0.008	0.074	0.049	0.024	0.547	0.537	0.198
LC	**0.078**	**0.048**	**0.029**	0.003	0.012	0.007	**0.049**	0.021	**0.047**	0.004	-	0.080	0.035	0.012	0.600	0.553	0.176
WK	**0.085**	**0.081**	**0.046**	0.031	**0.062**	0.017	**0.051**	0.029	0.032	**0.036**	**0.041**	-	0.047	0.064	0.668	0.648	0.310
BC	**0.047**	0.028	0.002	-0.001	0.010	0.024	0.010	0.022	**0.046**	0.026	0.019	0.024	-	0.038	0.580	0.544	0.246
LP	**0.081**	**0.059**	**0.032**	**0.027**	0.015	0.014	**0.053**	0.017	0.024	0.013	0.007	**0.034**	0.021	-	0.575	0.565	0.245
BH	**0.249**	**0.226**	**0.197**	**0.171**	**0.234**	**0.224**	**0.225**	**0.205**	**0.213**	**0.193**	**0.215**	**0.221**	**0.208**	**0.215**	-	0.407	0.527
BAR	**0.273**	**0.259**	**0.230**	**0.184**	**0.261**	**0.251**	**0.233**	**0.247**	**0.252**	**0.221**	**0.232**	**0.253**	**0.229**	**0.240**	**0.154**	-	0.425
BZ	**0.154**	**0.134**	**0.116**	**0.064**	**0.119**	**0.114**	**0.126**	**0.103**	**0.115**	**0.086**	**0.080**	**0.124**	**0.106**	**0.112**	**0.168**	**0.169**	-

Significant (following FDR correction) pairwise comparisons in bold ($p < 0.009$)

none of the sites showed significant heterozygote excesses in comparison to allelic richness.

Table 7: Analysis of Molecular Variance (AMOVA) results for *Spheciospongia vesparium* performed in two site groupings: 1) All sites included, split into four main locations (Florida, the Bahamas, Barbuda and Belize); 2) Only Florida sites, split into Upper Keys (U), Middle Keys Bay side (M), Middle Keys Atlantic side (A), and Lower Keys (L).

	Source of variation	SSD	Variance components	% Variance	*F*- statistics
1)	Within individuals	844.500	2.787	59.2	0.408
	Among individuals, within sites	1482.551	1.198	25.5	0.301
	Among sites, within locations	98.128	0.064	1.4	0.016
	Among locations	121.265	0.659	14.0	0.140
2)	Within individuals	687.500	2.634	67.0	0.330
	Among individuals, within sites	1242.097	1.208	30.7	0.314
	Among sites, within U/M/A/ L	76.832	0.056	1.4	0.014
	Among U/M/A/L	34.970	0.036	0.9	0.009

SSD: Sum of squared differences; F-statistics: Fixation indices.

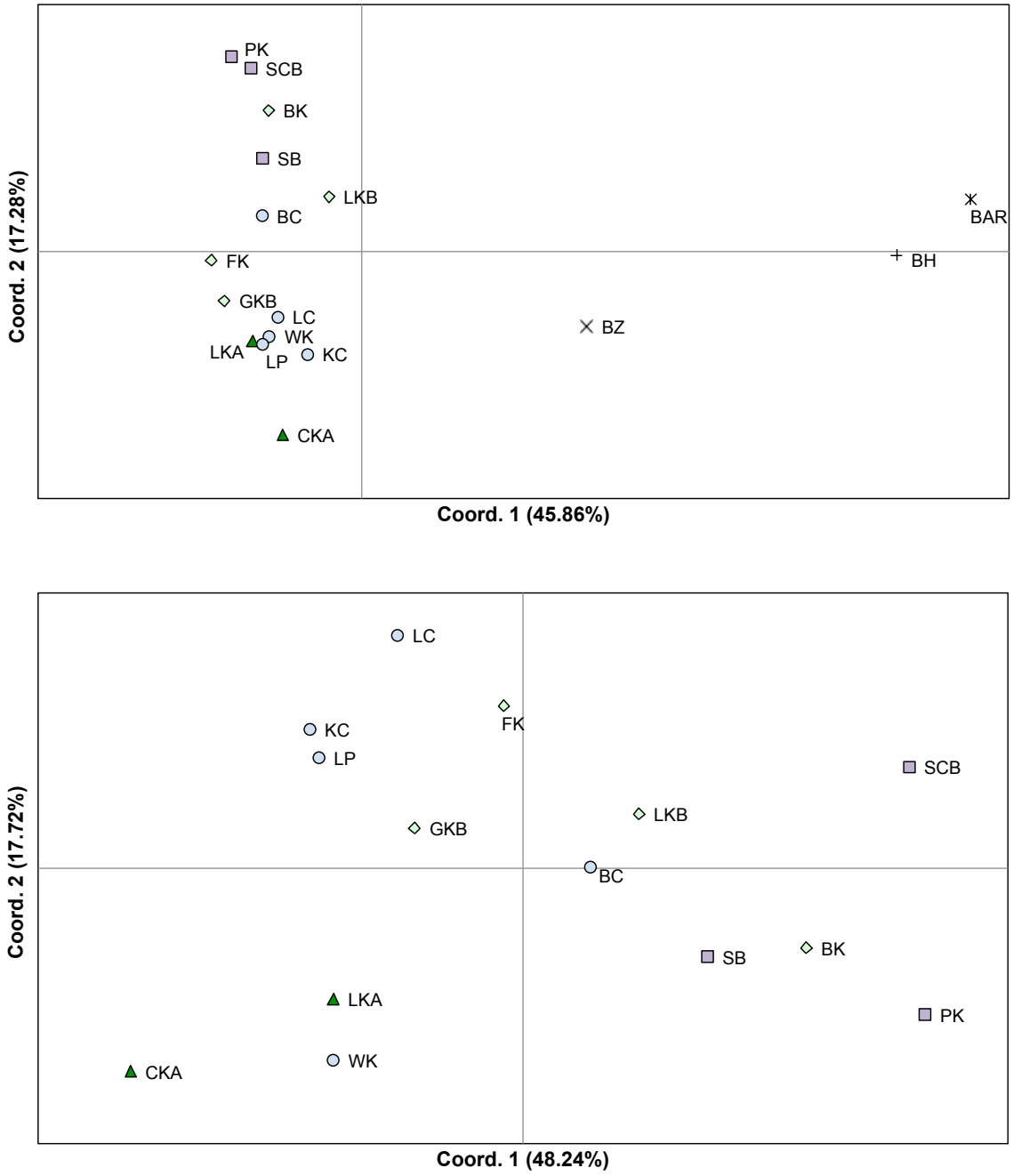


Figure 3: Principle Coordinates Analysis using null allele-corrected pairwise F_{ST} values for *Spheciospongia vesparium* at all sampling sites (top) and only Florida Keys sites (bottom). Florida Keys sites are grouped by symbol into Upper Keys (purple squares), Middle Keys Bay side (light green diamonds), Middle Keys Atlantic side (dark green triangles) and Lower Keys (blue circles).

DISCUSSION

Marker characteristics and summary statistics

A total of twelve polymorphic microsatellite markers for *Spheciospongia vesparium* were developed in this study; these are the only microsatellites developed for this species, and will be useful tools for future molecular studies. Ten of the markers were combined into two multiplex PCRs to reduce the number of PCR reactions and capillary sequencer runs necessary, and in addition a three-primer approach was used to avoid purchase of an expensive fluorescently-labelled primer for every locus. These two optimisations increase the cost effectiveness of using these markers, increasing their accessibility for other research groups and projects in the future. The number of alleles per locus ranged from four to 27, which was sufficiently polymorphic to detect population structure at small spatial scales.

Moderate levels of amplification failure and null alleles were present at a number of loci. Null alleles are caused by mutations in primer binding regions that prevent primers from binding and subsequent amplification of the locus in PCR, causing either non-amplification in both alleles (resulting in missing data), or non-amplification of only one allele (resulting in false homozygotes). High null allele frequencies have been found in other studies on sponges using microsatellites (Dailianis et al., 2011; Guardiola, Frotscher & Uriz, 2012, 2016; Chaves-Fonnegra et al., 2015; Pérez-Portela, Noyer & Becerro, 2015; Richards et al., 2016), suggesting that the problem may be common in the phylum. It is known for some groups (for example, molluscs and insects) to have higher occurrences of null alleles (Chapuis & Estoup, 2007), and it may be the case that Porifera fall into this bracket (although high null allele frequencies have not been reported in some sponge studies, e.g. Bell et al., [2014]). Spermcasters may be particularly susceptible to null alleles due to higher amounts of sperm production, as this means there are more cell cycles and therefore more opportunities for mutation (Addison & Hart, 2005).

High null allele frequencies can cause heterozygote deficiencies, and can bias population genetic studies by exaggerating differences among populations. Here, I ameliorated these effects using various approaches. I conducted a post-hoc analysis to examine the extent of any bias in F_{ST} caused by the null alleles, and found it to be minimal. Nonetheless, the null allele model was used in Geneland to account for null allele bias in population clustering analysis. I also corrected F_{IS} estimates for null alleles in INest (which were indeed found to affect F_{IS} estimates in all sites). Departure from HWE was found at a number of loci in all populations, which could be caused by inbreeding (discussed below in detail) or the null alleles present, but could also be due to other biological factors. This includes Wahlund effects (population structure within a site or group sampled), which have been proposed by other authors to be the result of temporal variation caused by reproductive asynchronicity or by recruitment of different genetic cohorts (Duran et al., 2004; Chaves-Fonnegra et al., 2015; Riesgo et al., 2016).

Greater Caribbean population structure

The four areas sampled (Florida, the Bahamas, Barbuda and Belize) all showed strong genetic differentiation from each other in genetic distance measures (F_{ST} and D) and clustering analyses. Pairwise F_{ST} values between the areas were all significant ($p < 0.05$, corrected for multiple comparisons) and ranged between 0.103 and 0.273. Analysis in Geneland grouped the four locations into four separate clusters. Iterative reallocation program FLOCK only detected two clusters initially - a Florida cluster and a cluster containing samples from the remaining sites (plus two individuals from Florida). However when the analysis was performed on only the Barbuda, Belize and the Bahamas samples, the three sites formed separate clusters. This reflects the results from the PCoA, as the first axis split the Florida sites and the remaining sites. The first-generation migrants analysis conducted did not find any evidence of putative migrants among any of the four locations. The AMOVA found that 14% of genetic variation in the dataset was found among the four locations.

These results are congruent with those of other sponge species. Richards et al. (2016) sampled the sponge *Xestospongia muta* in locations in similar geographic areas to those sampled in this study (Florida Keys, Utila [Honduras], Crooked Island [the Bahamas] and St Croix [US Virgin Islands]), and found similar results, including Geneland analyses splitting the four locations into four separate clusters. Other sponge studies both within the Caribbean (López-Legentil & Pawlik, 2009; Chaves-Fonnegra et al., 2015; de Bakker et al., 2016) and in other regions (Duran et al., 2004; Xavier et al., 2010; Pérez-Portela, Noyer & Becerro, 2015; Riesgo et al., 2016; Brown, Davis & Leys, 2017) have shown strong population structure at large spatial scales. This is likely to be due to poor dispersal abilities in the phylum. Dispersal in marine species is affected by a number of factors and the complex interactions between them (Cowen, Paris & Srinivasan, 2006; Cowen & Sponaugle, 2009). Life history characteristics such as pelagic larval duration, larval behaviour and reproductive strategies in adults are highly influential on dispersal (Butler MJ et al., 2011; Selkoe & Toonen, 2011; Kough & Paris, 2015; Coelho & Lasker, 2016). The lecithotrophic larvae of sponges are only in the water column for a matter of hours to days before they settle (Maldonado, 2006), and therefore their dispersal is restricted by time, and so generally, high levels of philopatry are expected (although higher dispersal may be found in some species).

In this study, isolation by distance (IBD) was observed among the sampling sites in the Greater Caribbean ($r = 0.792$, $p < 0.001$). This may suggest that dispersal is distance-limited at the larger spatial scale in *S. vesparium*, reflecting short larval durations as is found throughout the phylum. However, to understand fully the extent of IBD forces in shaping population structure in this species, further study with intermediate populations would be required, as the sites sampled here were great distances apart. Oceanographic features were found to influence the population structure of *S. vesparium* in the region. The Florida sites were more similar to Caye Caulker in Belize (BZ) than the more geographically proximate Abaco in the Bahamas (BH). Similar patterns have been found in previous studies in sponges (López-Legentil & Pawlik, 2009; Richards et al., 2016) and corals (Brazeau, Sammarco & Gleason, 2005;

Baums et al., 2010), and predicted in reef fish through biophysical modelling (Cowen, Paris & Srinivasan, 2006). This is due to regional current patterns: The strong Florida Current running through the Straits of Florida separating the Bahamas and Florida is likely to act as a barrier to dispersal, while the Caribbean Current and Loop Current could transport larvae from Belize towards Florida. From the northern Mesoamerican Barrier Reef System (MBRS), of which the Belizean site is part, the strong effects of currents mean that larvae are estimated to be transported to Florida in as little as seven to ten days (Muhling et al., 2013). Although this could be longer than the time from release of gametes to larval settlement in *S. vesparium*, sites between the MBRS and Florida could act as intermediate stepping stone sites in this dispersal pathway. These patterns have not been found to be universal in sponges, however, as Chaves-Fonnegra et al. (2015) did find connectivity between Florida and the Bahamas in *Cliona delitrix*, suggesting that this could occur through deep-water currents between the locations, therefore indicating that this barrier to dispersal is not absolute. However, sampling for this study was not carried out in the same areas as in the *C. delitrix* study: we sampled mostly Bay side locations in Florida (which were very shallow), and Chaves-Fonnegra et al. (2015) sampled Atlantic side sites, and so these spatial variants could explain the different patterns observed between the studies.

Private alleles were found in all sites, and this can also be used as evidence for isolation and limited gene flow. Private allelic richness tended towards being slightly lower in Florida than in the other sites (when considered as one site), although standard errors did slightly overlap. Lower private allelic richness in Florida was found by Chaves-Fonnegra et al., (2015) in their study of *C. delitrix*, which was suggested to be because Florida is at the edge of the range for the species, and therefore more recent expansion has meant the population here has not had as much time to accrue unique mutations in the microsatellite genes. This may also mean that Florida is a sink for gene flow from other areas due to prevailing currents.

Florida population structure

Although clustering analyses (Geneland, Flock) grouped all the Florida sites into a single genetic group, there was some genetic structure detected among the sites, as can be observed in pairwise genetic distance calculations, in the PCoA conducted on the Florida sites only, and in the Mantel test for IBD. The presence of structure in Floridian sponge populations has also been found in other studies using microsatellites (DeBiasse, Richards & Shivji, 2010; Chaves-Fonnegra et al., 2015; Richards et al., 2016), and has been found using genomics in the coral *Acropora cervicornis* (Drury et al., 2017), which also has a short (3-5 day) larval duration. Limited dispersal capacity is likely to contribute to this structuring, and indeed, a species with a longer larval duration (*Diadema antillarum;* 36 days, [Eckert 1998]) has been observed to show no population structure in Florida (Chandler et al., 2017). However, IBD has not been detected in the above species, indicating that distance-limited dispersal is not the primary driver of their population structures.

In *S. vesparium*, we also found that IBD was not a prevailing force driving population structure in Florida; IBD was significant, but accounted for only a small proportion of the population structure observed ($r = 0.229$, $p = 0.03$). In the PCoA, a loose IBD pattern can be seen in the Bay side sites, but the Atlantic side sites (CKA, LKA) cluster together away from the Bay side sites together with Waltz Key (WK). This indicates that there is less connectivity between geographically proximate Bay side and Atlantic side sites than among more distant sites on the same side of the islands. This may be attributable to water movement patterns; inflow and outflow through the interisland channels between the Keys are variable both spatially and temporally (Smith, 1994; Yeung et al., 2001; Lee & Smith, 2002), and therefore connectivity may be limited.

The clustering of Waltz Key (WK) with the Atlantic side sites (CKA, LKA) is hard to explain; the site also shows small (F_{ST} = 0.034 - 0.047) but significant differentiation from geographically proximate Lower Keys sites (Kemp Channel [KC], Little Crane Key [LC], Lakes Passage [LP]). Waltz Key is located in an

area of emergent carbonate mud banks, which could form a retentive environment that does not lend itself well to larval dispersal. This could explain the genetic differences between it and other Bay side sites, but its connection to Middle Keys Atlantic side sites is perplexing. Unexpected patterns of fine-scale genetic structure that do not obviously coincide with structural or oceanographic features were observed in other instances in pairwise F_{ST} and differentiation tests in *S. vesparium*, and have also been found in other sponge species in Florida (DeBiasse, Richards & Shivji, 2010; Chaves-Fonnegra et al., 2015). Genetic patchiness is suggested to be caused by 'sweepstakes reproductive success' (Hedgecock, 1982, 1994; Hedgecock & Pudovkin, 2011) – the random survival of certain larval cohorts due to oceanographic conditions, found in species with high fecundity and high larval mortality. Furthermore, in variable current regimes (as can be found in the Florida Keys [Lee et al., 1992]), temporally variable dispersal pathways could arise. Both of these situations could lead to spatially heterogeneous genetic structure through the mechanism of drift, however differential selection may also be the cause of such structure. Cyanobacteria blooms exert a strong selection pressure, and it is feasible that sponges in areas that have been affected by blooms are genetically different to those that have not been exposed to blooms.

Although there was structure present, non-significant pairwise comparisons among sites, clustering analyses and first generation migrant analysis all show that connectivity is present throughout the Bay. Sites such as Long Key Bay side (LKB) in the Middle Keys and Boca Chica Channel (BC) in the Lower Keys were apparently well connected to sites over the Keys range. A putative migrant from Waltz Key (WK) was found in Long Key Bay side (LKB), sites which have an oceanographic distance of ~89 km separating them. This is not unusual; Chaves-Fonnegra et al. (2015) found dispersal up to ~315 km in *C. delitrix* in Florida. The complex currents found across the Florida Keys are likely to aid in connectivity among disparate sites. Although the main current dominating the area is the Florida Current, the continuation of the Loop Current that flows easterly then northerly through the Straits of Florida and around the Florida Peninsula, there are many other ocean processes in the area that could affect

larval dispersal patterns. Westerly running counter currents arise as a result of downwelling winds and offshore eddies and gyres (Lee & Williams, 1999; Yeung et al., 2001), and eddies themselves also forge connectivity in the area (Sponaugle et al., 2005).

Genetic diversity, inbreeding and bottlenecks

Inbreeding was potentially present in all populations, with positive average F_{IS} values in all populations when corrected for null alleles. Values ranged from 0.036 in the Bahamas (BH) to 0.343 in Craig's Key Atlantic side (CKA). However, the posterior 95% probability intervals from the model ranged from zero in all sites but in Florida as a whole. This indicates that there is insufficient evidence to show that inbreeding is present in the Bahamas, Barbuda and Belize sites, or to specify particular sites within Florida that it occurs. The posterior 95% probability intervals were fairly large, potentially due to small sample sizes, and so inbreeding cannot be excluded as a cause of heterozygosity deficiencies at these sites. However, the data here can only statistically support that over all sites in Florida, on average, individuals are more related than would be expected under a model of random mating, even taking into account null alleles and genotyping failure. Grouping the Floridian sites in a single analysis may have caused Wahlund effects to arise, which could affect F_{IS} estimates. However, Wahlund effects are unlikely to alter F_{IS} estimates in comparison to null alleles or inbreeding when genetic diversity does not greatly differ among mixed groups (Zhivotovsky, 2015). Nevertheless, understanding of inbreeding in this species across all locations would benefit greatly from increased sampling, and the use of genetic markers with low null allele prevalence.

Inbreeding can have negative consequences for fitness, and so the results here could be of concern for the populations sampled. In sponges, high inbreeding values have been found in many species from different regions and taxonomic groups (Guardiola, Frotscher & Uriz, 2012; Bell et al., 2014; Chaves-Fonnegra et al., 2015; Pérez-Portela, Noyer & Becerro, 2015; Giles et al., 2015; Riesgo et al., 2016; Padua, Cunha & Klautau, 2017). This indicates that there are

characteristics common in the Porifera that cause widespread inbreeding; this is likely to be limited dispersal and limited connectivity among locations, meaning high levels of philopatry. In addition, sponges can be hermaphroditic, and therefore could self-fertilise. However, Blanquer and Uriz (2010) did not find inbreeding in Mediterranean demosponge *Scopalina lophyropoda*, despite this species exhibiting philopatry in larval dispersal, which the authors suggest could be due to sperm dispersal or strong selection pressures against inbreeding. In *S. vesparium*, limited dispersal and reduced population sizes due to cyanobacterial blooms could have lead to higher inbreeding levels in Florida. In the Clionaidae family, both gonochorism and hermaphroditism are found, and therefore inbreeding rates could also be partially the result of self-fertilization in this species.

We only found one pair of identical multilocus genotypes in our dataset, which were found within the same site, the Lakes Passage (LP). These samples may have been clones, but there is a small possibility that the identical genotypes occurred by chance. The low prevalence of replicate multilocus genotypes could indicate that asexual reproduction is low in *S. vesparium* in the sites sampled, but alternatively could be due to insufficient sampling to detect clones. In other sponge studies using microsatellites, replicate genotypes have been found at varying rates: Bell et al. (2014) found none in *Xestospongia sp*., whereas Chaves-Fonnegra et al. (2015) found 60 out of 495 samples in *C. delitrix* (12% incidence).

Genetic diversity levels (allelic richness) were very similar across locations when the data were rarefied for sample sizes, although some Floridian sites did show slightly lower levels than the other Caribbean sites. Genetic diversity can be lost in populations that have suffered from rapid declines due to bottleneck effects. In this study, I found signatures of bottlenecks in four Florida sites (both Middle Keys Atlantic sites [CKA, LKA], Boca Chica Channel [BC], and Little Crane Key [LC] in the lower Keys. Dailianis et al. (2011) also investigated for genetic signals of population bottlenecks in *Spongia officinalis*, a species that has been exploited for centuries for the bath sponge market, as well as having

suffered mass mortalities. Unexpectedly, genetic diversity levels were high, and population bottleneck signatures were not detected. The authors attributed this to 3 possibilities – adequate population connectivity, potential regeneration of partially harvested sponges and robustness of populations able to recolonize affected areas. In contrast, congener *S. lamella*, which has also experienced disease and exploitation, was found to have low genetic diversity in Western Mediterranean and Portuguese populations (Pérez-Portela, Noyer & Becerro, 2015), and did show bottleneck signatures. In *Ircinia fasciculata* populations (also in the Mediterranean), disease has caused mass mortalities, and evidence of bottlenecks have been found at many sites, but at one site where mortalities have been known to have occurred, no such signature was detected. This was suggested by the authors to potentially be because of high levels of connectivity with other sites, which could protect the population against the effects of genetic drift (Riesgo et al., 2016).

Although genetic diversity levels are not substantially lower than in other populations, there is evidence that mortalities have caused bottlenecks and led to lower genetic diversity in Florida *S. vesparium* populations. This is concerning, as genetic diversity is the foundation of adaptive variation in a species, and as such the ability of the species to evolve to new environmental changes may be threatened. At the moment however, there is a case for cautious optimism as there are still similar levels of genetic diversity as in unaffected populations. Similar scenarios have been observed in other taxa that have suffered mass mortalities in the Florida reef tract; in coral *Acropora cervicornis* and sea urchin *Diadema antillarum*, genetic diversity is not lower that in other Caribbean sites tested (Chandler et al., 2017; Drury et al., 2017).

Conclusion

Spheciospongia vesparium populations are highly structured at the large spatial scale (388 km to 2785 km), and are influenced by both low dispersal abilities and oceanographic barriers to dispersal. At the smaller spatial scale (5 km to 155 km), while connectivity is present, some population structure is found in weak isolation by distance patterns but with some seemingly random spatial

heterogeneity. This is likely influenced by a combination of demographic stochasticity, spatial and temporal fine scale variation in currents, and bathymetric features. Despite mortalities and bottleneck signatures, genetic diversity in Florida populations is comparable to the other locations sampled. Due to this and the evidence of connectivity over the Keys, populations should show some resilience to further stressors. However, current stunted recovery of sponge communities in the Bay indicates that factors other than demographics are important for ensuring their conservation. In addition, the fine-scale structuring that can occur and its apparent unpredictability should be taken into account in management decisions.

REFERENCES

Addison JA., Hart MW. 2005. Spawning, copulation and inbreeding coefficients in marine invertebrates. *Biology Letters* 1:450–453. DOI: 10.1098/rsbl.2005.0353.

Alberto F. 2009. MsatAllele_1.0: An R package to visualize the binning of microsatellite alleles. *Journal of Heredity* 100:394–7. DOI: 10.1093/jhered/esn110.

de Bakker DM., Meesters EHWG., van Bleijswijk JDL., Luttikhuizen PC., Breeuwer HJAJ., Becking LE. 2016. Population genetic structure, abundance, and health status of two dominant benthic species in the Saba Bank National Park, Caribbean Netherlands: *Montastraea cavernosa* and *Xestospongia muta*. *PLoS ONE* 11:e0155969. DOI: 10.1371/journal.pone.0155969.

Baums IB., Johnson ME., Devlin-Durante MK., Miller MW. 2010. Host population genetic structure and zooxanthellae diversity of two reef-building coral species along the Florida Reef Tract and wider Caribbean. *Coral Reefs* 29:835–842. DOI: 10.1007/s00338-010-0645-y.

Bell JJ., Smith D., Hannan D., Haris A., Jompa J., Thomas L. 2014. Resilience to disturbance despite limited dispersal and self-recruitment in tropical barrel sponges: implications for conservation and management. *PLoS ONE* 9:e91635. DOI: 10.1371/journal.pone.0091635.

Benjamini Y., Yekutieli D. 2001. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* 29:1165–1188. DOI: 10.2307/2674075.

Bergquist PR., Sinclair ME. 2010. The morphology and behaviour of larvae of some intertidal sponges. *New Zealand Journal of Marine and Freshwater Research* 2:426–437. DOI: 10.1080/00288330.1968.9515247.

Berry DL., Goleski JA., Koch F., Wall CC., Peterson BJ., Anderson OR., Gobler CJ. 2015. Shifts in cyanobacterial strain dominance during the onset of harmful algal blooms in Florida Bay, USA. *Microbial Ecology* 70:361–371. DOI: 10.1007/s00248-014-0564-5.

Bertelsen RD., Butler MJ., Herrnkind WF., Hunt JH. 2009. Regional characterisation of hardbottom nursery habitat for juvenile Caribbean spiny lobster (*Panulirus argus*) using rapid assessment techniques. *New Zealand Journal of Marine and Freshwater Research* 43:299–312. DOI: 10.1080/00288330909510002.

Blacket M., Robin C., Good R., Lee S., Miller A. 2012. Universal primers for fluorescent labelling of PCR fragments- an efficient and cost-effective approach to genotyping by fluorescence. *Molecular Ecology Resources* 12:456–63. DOI: 10.1111/j.1755-0998.2011.03104.x.

Blanquer A., Uriz MJ. 2010. Population genetics at three spatial scales of a rare sponge living in fragmented habitats. *BMC Evolutionary Biology* 10:1–9. DOI: 10.1186/1471-2148-10-13.

Blanquer A., Uriz M., Caujapé-Castells J. 2009. Small-scale spatial genetic structure in Scopalina lophyropoda, an encrusting sponge with philopatric larval dispersal and frequent fission and fusion events. *Marine Ecology Progress Series* 380:95–102. DOI:

10.3354/meps07931.

Brazeau DA., Sammarco PW., Gleason DF. 2005. A multi-locus genetic assignment technique to assess sources of *Agaricia agaricites* larvae on coral reefs. *Marine Biology* 147:1141–1148. DOI: 10.1007/s00227-005-0022-5.

Bristol RM., Tucker R., Dawson DA., Horsburgh G., Prys-Jones RP., Frantz AC., Krupa A., Shah NJ., Burke T., Groombridge JJ. 2013. Comparison of historical bottleneck effects and genetic consequences of re-introduction in a critically endangered island passerine. *Molecular Ecology* 22:4644–4662. DOI: 10.1111/mec.12429.

Brown RR., Davis CS., Leys SP. 2017. Clones or clans: the genetic structure of a deep-sea sponge, *Aphrocallistes vastus*, in unique sponge reefs of British Columbia, Canada. *Molecular Ecology* 26:1045–1059. DOI: 10.1111/mec.13982.

Butler M., Hunt J., Herrnkind W., Childress M., Bertelsen R., Sharp W., Matthews T., Field J., Marshall H. 1995. Cascading disturbances in Florida Bay, USA: cyanobacteria blooms, sponge mortality, and implications for juvenile spiny lobsters *Panulirus argus*. *Marine Ecology Progress Series* 129:119–125. DOI: 10.3354/meps129119.

Butler MJ I., Paris C., Goldstein J., Matsuda H., Cowen R. 2011. Behavior constrains the dispersal of long-lived spiny lobster larvae. *Marine Ecology Progress Series* 422:223–237. DOI: 10.3354/meps08878.

Butler J., Stanley JA., Butler MJ. 2016. Underwater soundscapes in near-shore tropical habitats and the effects of environmental degradation and habitat restoration. *Journal of Experimental Marine Biology and Ecology* 479:89–96. DOI: 10.1016/j.jembe.2016.03.006.

Chandler LM., Walters LJ., Sharp WC., Hoffman EA. 2017. Genetic structure of natural and broodstock populations of the long-spined sea urchin, *Diadema antillarum*, throughout the Florida Keys. *Bulletin of Marine Science* 93:1–9. DOI: 10.5343/bms.216.1101.

Chapuis M-P., Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24:621–31. DOI: 10.1093/molbev/msl191.

Chaves-Fonnegra A., Feldheim KA., Secord J., Lopez J V. 2015. Population structure and dispersal of the coral-excavating sponge *Cliona delitrix*. *Molecular Ecology* 24:1447–66. DOI: 10.1111/mec.13134.

Chiappone M., Sullivan KM. 1994. Ecological structure and dynamics of nearshore hard-bottom communities in the Florida Keys. *Bulletin of Marine Science* 54:747–756.

Chybicki IJ., Burczyk J. 2009. Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* 100:106–13. DOI: 10.1093/jhered/esn088.

Coelho MAG., Lasker HR. 2016. Larval behavior and settlement dynamics of a ubiquitous Caribbean octocoral and its implications for dispersal. *Marine Ecology Progress Series* 561:109–121. DOI: 10.3354/meps11941.

Cornuet JM., Luikart G. 1996. Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. *Genetics* 144:2001–2014.

Cowen RK., Paris CB., Srinivasan A. 2006. Scaling of connectivity in marine populations. *Science* 311:522–7. DOI: 10.1126/science.1122039.

Cowen R., Sponaugle S. 2009. Larval dispersal and marine population connectivity. *Annual Review of Marine Science* 1:443–466. DOI: 10.1146/annurev.marine.010908.163757

Culley TM., Stamper TI., Stokes RL., Brzyski JR., Hardiman N a., Klooster MR., Merritt BJ. 2013. An efficient technique for primer development and application that integrates fluorescent labeling and multiplex PCR. *Applications in Plant Sciences* 1:1300027. DOI: 10.3732/apps.1300027.

Dailianis T., Tsigenopoulos CS., Dounas C., Voultsiadou E. 2011. Genetic diversity of the imperilled bath sponge *Spongia officinalis* Linnaeus, 1759 across the Mediterranean Sea: patterns of population differentiation and implications for taxonomy and conservation. *Molecular Ecology* 20:3757–72. DOI: 10.1111/j.1365-294X.2011.05222.x.

DeBiasse MB., Richards VP., Shivji MS. 2010. Genetic assessment of connectivity in the common reef sponge, *Callyspongia vaginalis* (Demospongiae: Haplosclerida) reveals high population structure along the Florida reef tract. *Coral Reefs* 29:47–55. DOI: 10.1007/s00338-009-0554-0.

Dempster AP., Laird NM., Rubin DB. 1977. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* 39:1–38.

Diaz M., Rutzler K. 2001. Sponges: an essential component of Caribbean coral reefs. *Bulletin of Marine Science* 69:535–546.

Dray S., Dufour AB., Chessel D. 2007. The ade4 Package - II: Two-table and K-table methods. *R News* 7:47–52.

Drury C., Schopmeyer S., Goergen E., Bartels E., Nedimyer K., Johnson M., Maxwell K., Galvan V., Manfrino C., Lirman D. 2017. Genomic patterns in *Acropora cervicornis* show extensive population structure and variable genetic diversity. *Ecology and Evolution*. DOI: 10.1002/ece3.3184.

Duchesne P., Turgeon J. 2012. FLOCK provides reliable solutions to the “number of populations” problem. *Journal of Heredity* 103:734–743. DOI: 10.1093/jhered/ess038.

Duffy JE. 1996. Eusociality in a coral-reef shrimp. *Nature* 381:512–514. DOI: 10.1038/381512a0.

Dunlap M., Pawlik JR. 1998. Spongivory by parrotfish in Florida mangrove and reef habitats. *Marine Ecology* 19:325–337. DOI: 10.1111/j.1439-0485.1998.tb00471.x.

Duran S., Pascual M., Estoup A., Turon X. 2004. Strong population structure in the marine sponge *Crambe crambe* (Poecilosclerida) as revealed by microsatellite markers. *Molecular Ecology* 13:511–522. DOI: 10.1046/j.1365-294X.2004.02080.x.

Eckert G. 1998. Larval development, growth and morphology of the sea urchin *Diadema antillarum*. *Bulletin of Marine Science* 63:443–451.

Ehlers A., Worm B., Reusch TBH. 2008. Importance of genetic diversity in eelgrass *Zostera marina* for its resilience to global warming. *Marine Ecology Progress Series* 355:1–7. DOI: 10.3354/meps07369.

Evans SM., Vergés A., Poore AGB. 2017. Genotypic diversity and short-term response to shading stress in a threatened seagrass: Does low diversity mean low resilience?

Frontiers in Plant Science 8:1417. DOI: 10.3389/fpls.2017.01417.

Fiore CL., Freeman CJ., Kujawinski EB. 2017. Sponge exhalent seawater contains a unique chemical profile of dissolved organic matter. *PeerJ* 5:e2870. DOI: 10.7717/peerj.2870.

Forcucci D., Butler MJ., Hunt JH. 1994. Population dynamics of juvenile Caribbean spiny lobster, *Panulirus argus*, in Florida Bay, Florida. *Bulletin of Marine Science* 54:805–818.

Freeman CJ., Thacker RW. 2011. Complex interactions between marine sponges and their symbiotic microbial communities. *Limnology and Oceanography* 56:1577–1586. DOI: 10.4319/lo.2011.56.5.1577.

Garza JC., Williamson EG. 2001. Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* 10:305–18. DOI: 10.1046/j.1365-294X.2001.01190.x

Giles EC., Saenz-Agudelo P., Hussey NE., Ravasi T., Berumen ML. 2015. Exploring seascape genetics and kinship in the reef sponge *Stylissa carteri* in the Red Sea. *Ecology and Evolution* 5:2487–502. DOI: 10.1002/ece3.1511.

de Goeij JM., van Oevelen D., Vermeij MJA., Osinga R., Middelburg JJ., de Goeij AFPM., Admiraal W. 2013. Surviving in a marine desert: The sponge loop retains resources within coral reefs. *Science* 342:108–110. DOI: 10.1126/science.1241981.

Griffiths SM., Fox G., Briggs PJ., Donaldson IJ., Hood S., Richardson P., Leaver GW., Truelove NK., Preziosi RF. 2016. A Galaxy-based bioinformatics pipeline for optimised, streamlined microsatellite development from Illumina next-generation sequencing data. *Conservation Genetics Resources* 8:481–486. DOI: 10.1007/s12686-016-0570-7.

Guardiola M., Frotscher J., Uriz J. 2012. Genetic structure and differentiation at a short-time scale of the introduced calcarean sponge *Paraleucilla magna* to the western Mediterranean. *Hydrobiologia* 687:71–84. DOI: 10.1007/s10750-011-0948-1.

Guardiola M., Frotscher J., Uriz M-J. 2016. High genetic diversity, phenotypic plasticity, and invasive potential of a recently introduced calcareous sponge, fast spreading across the Atlanto-Mediterranean basin. *Marine Biology* 163:123. DOI: 10.1007/s00227-016-2862-6.

Guillot Gi., Mortier F., Estoup A. 2005. Geneland: a computer package for landscape genetics. *Molecular Ecology Notes* 5:712–715. DOI: 10.1111/j.1471-8286.2005.01031.x.

Guillot G., Santos F., Estoup A. 2008. Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics* 24:1406–1407. DOI: 10.1093/bioinformatics/btn136.

Hedgecock D. 1982. Genetic consequences of larval retention: theoretical and methodological aspects. In: *Estuarine Comparisons*. Elsevier, 553–568. DOI: 10.1016/B978-0-12-404070-0.50039-9.

Hedgecock D. 1994. Does variance in reproductive success limit effective population sizes of marine organisms? In Beaumont A (ed.) *Genetics and Evolution of Aquatic Organisms* (122–134). Chapman & Hall, London.

Hedgecock D., Pudovkin AI. 2011. Sweepstakes reproductive success in highly fecund marine fish and shellfish: A review and commentary. *Bulletin of Marine Science* 87:971–1002. DOI: 10.5343/bms.2010.1051.

Herrnkind WF., Butler IV MJ., Hunt JH., Childress M. 1997. Role of physical refugia: implications from a mass sponge die-off in a lobster nursery in Florida. *Marine and Freshwater Research* 48:759. DOI: 10.1071/MF97193.

Holleley CE., Geerts PG. 2009. Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. *BioTechniques* 46:511–7. DOI: 10.2144/000113156.

Jost L. 2008. GST and its relatives do not measure differentiation. *Molecular Ecology* 17:4015–4026. DOI: 10.1111/j.1365-294X.2008.03887.x.

Kough AS., Paris CB. 2015. The influence of spawning periodicity on population connectivity. *Coral Reefs* 34:753–757. DOI: 10.1007/s00338-015-1311-1.

Lee TN., Rooth C., Williams E., McGowan M., Szmant AF., Clarke ME. 1992. Influence of Florida Current, gyres and wind-driven circulation on transport of larvae and recruitment in the Florida Keys coral reefs. *Continental Shelf Research* 12:971–1002. DOI: 10.1016/0278-4343(92)90055-O.

Lee TN., Smith N. 2002. Volume transport variability through the Florida Keys tidal channels. *Continental Shelf Research* 22:1361–1377. DOI: 10.1016/S0278-4343(02)00003-1.

Lee TN., Williams E. 1999. Mean distribution and seasonal variability of coastal currents and temperature in the Florida Keys with implications for larval recruitment. *Bulletin of Marine Science* 64:35–56.

Lischer HEL., Excoffier L. 2012. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics* 28:298–299. DOI: 10.1093/bioinformatics/btr642.

López-Legentil S., Pawlik JR. 2009. Genetic structure of the Caribbean giant barrel sponge *Xestospongia muta* using the I3-M11 partition of COI. *Coral Reefs* 28:157–165. DOI: 10.1007/s00338-008-0430-3.

Maldonado M. 2006. The ecology of the sponge larva. *Canadian Journal of Zoology* 84:175–194. DOI: 10.1139/z05-177.

Maldonado M., Young CM. 1999. Effects of the duration of larval life on postlarval stages of the demosponge *Sigmadocia caerulea*. *Journal of Experimental Marine Biology and Ecology* 232:9–21. DOI: 10.1016/S0022-0981(98)00076-8.

van der Meer MH., Horne JB., Gardner MG., Hobbs J-PA., Pratchett M., van Herwerden L. 2013. Limited contemporary gene flow and high self-replenishment drives peripheral isolation in an endemic coral reef fish. *Ecology and Evolution* 3:n/a-n/a. DOI: 10.1002/ece3.584.

Meirmans PG., Van Tiendener PH. 2004. Genotype and Genodive: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4:792–794. DOI: 10.1111/j.1471-8286.2004.00770.x.

Muhling BA., Smith RH., Vásquez-Yeomans L., Lamkin JT., Johns EM., Carrillo L., Sosa-Cordero E., Malca E. 2013. Larval fish assemblages and mesoscale oceanographic structure along the Mesoamerican Barrier Reef System. *Fisheries Oceanography* 22:409–428. DOI: 10.1111/fog.12031.

Narum SR. 2006. Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics* 7:783–787. DOI: 10.1007/s10592-005-9056-y.

Padua A., Cunha HA., Klautau M. 2017. Gene flow and differentiation in a native calcareous sponge (Porifera) with unknown dispersal phase. *Marine Biodiversity*:1–11. DOI: 10.1007/s12526-017-0742-z.

Paetkau D., Slade R., Burden M., Estoup A. 2004. Genetic assignment methods for the direct, real-time estimation of migration rate: a simulation-based exploration of accuracy and power. *Molecular Ecology* 13:55–65. DOI: org/10.1046/j.1365-294X.2004.02008.x

Pante E., Simon-Bouhet B. 2013. marmap: A package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE* 8:e73051. DOI: 10.1371/journal.pone.0073051.

Paul VJ., Thacker RW., Banks K., Golubic S. 2005. Benthic cyanobacterial bloom impacts the reefs of South Florida. *Coral Reefs* 24:693–697. DOI: 10.1007/s00338-005-0061-x.

Peakall R., Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research- an update. *Bioinformatics* 28:2537–2539. DOI: 10.1093/bioinformatics/bts460.

Pearse AS. 1950. Notes on the inhabitants of certain sponges at Bimini. *Ecology* 31:149–151.

Pérez-Portela R., Noyer C., Becerro MA. 2015. Genetic structure and diversity of the endangered bath sponge *Spongia lamella*. *Aquatic Conservation: Marine and Freshwater Ecosystems* 25:365–379. DOI: 10.1002/aqc.2423.

Peterson B., Chester C., Jochem F., Fourqurean JW. 2006. Potential role of sponge communities in controlling phytoplankton blooms in Florida Bay. *Marine Ecology Progress Series* 328:93–103. DOI:10.3354/meps328093.

Piry S., Alapetite A., Cornuet J-M., Paetkau D., Baudouin L., Estoup A. 2004. GENECLASS2: A software for genetic assignment and first-generation migrant detection. *Journal of Heredity* 95:536–539. DOI: 10.1093/JHERED/ESH074.

Poulakis GR., Seitz JC. 2004. Recent occurrence of the smalltooth sawfish, *Pristis pectinata* (Elasmobranchiomorphi: Pristidae), in Florida Bay and the Florida Keys, with comments on sawfish ecology. *Florida Scientist* 67:27–35. DOI: 10.2307/24321195.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Rannala B., Mountain JL. 1997. Detecting immigration by using multilocus genotypes. *Proceedings of the National Academy of Sciences* 94:9197–9201. DOI: 10.1073/pnas.94.17.9197.

Raymond M., Rousset F. 1995. GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86:248–249. DOI: 10.1093/oxfordjournals.jhered.a111573.

Reed DH., Frankham R. 2003. Correlation between fitness and genetic diversity. *Conservation Biology* 17:230–237. DOI: 10.1046/j.1523-1739.2003.01236.x.

Richards VP., Bernard AM., Feldheim KA., Shivji MS. 2016. Patterns of population structure and

dispersal in the long-lived “redwood” of the coral reef, the giant barrel sponge (*Xestospongia muta*). *Coral Reefs* 35:1097–1107. DOI: 10.1007/s00338-016-1435-y.

Riesgo A., Pérez-Portela R., Pita L., Blasco G., Erwin PM., López-Legentil S. 2016. Population structure and connectivity in the Mediterranean sponge *Ircinia fasciculata* are affected by mass mortalities and hybridization. *Heredity* 117:427–439. DOI: 10.1038/hdy.2016.41.

Robblee M., Barber T., Carlson P., Durako M., Fourqurean J., Muehlstein L., Porter D., Yarbro L., Zieman R., Zieman J. 1991. Mass mortality of the tropical seagrass *Thalassia testudinum* in Florida Bay (USA). *Marine Ecology Progress Series* 71:297–299. DOI: 10.3354/meps071297.

Saenz-Agudelo P., Jones GP., Thorrold SR., Planes S. 2011. Connectivity dominates larval replenishment in a coastal reef fish metapopulation. *Proceedings of the Royal Society B: Biological Sciences* 278:2954–61. DOI: 10.1098/rspb.2010.2780.

Selkoe K., Toonen R. 2011. Marine connectivity: a new look at pelagic larval duration and genetic metrics of dispersal. *Marine Ecology Progress Series* 436:291–305. DOI: 10.3354/meps09238.

Smith NP. 1994. Long-term Gulf to Atlantic transport through tidal channels in the Florida Keys. *Bulletin of Marine Science* 54:602–609.

Sponaugle S., Lee T., Kourafalou V., Pinkard D. 2005. Florida Current frontal eddies and the settlement of coral reef fishes. *Limnology and Oceanography* 50:1033–1048. DOI: 10.4319/lo.2005.50.4.1033.

Szpiech ZA., Jakobsson M., Rosenberg NA. 2008. ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* 24:2498–2504. DOI: 10.1093/bioinformatics/btn478.

Tellier M-AS., Bertelsen R. 2008. *Monitoring the flora and fauna of the nearshore hardbottom habitats of the Florida Keys*. F2196-05-08-F Final Report. Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute, Marathon, FL.

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* 27:783–786. DOI: 10.1007/s00338-008-0400-9.

Wright S. 1943. Isolation by Distance. *Genetics* 28:114–38.

Wright S. 1949. The genetical structure of populations. *Annals of Eugenics* 15:323–354. DOI: 10.1111/j.1469-1809.1949.tb02451.x.

Xavier JR., Soest RWM Van., Breeuwer JAJ., Menken SBJ. 2010. Phylogeography, genetic diversity and structure of the poecilosclerid sponge *Phorbas fictitius* at oceanic islands. *Contributions to Zoology* 79:119–129.

Yeung C., Jones DL., Criales MM., Jackson TL., Richards WJ. 2001. Influence of coastal eddies and counter-currents on the influx of spiny lobster, *Panulirus argus*, postlarvae into Florida Bay. *Marine And Freshwater Research* 52:1217–1232. DOI: 10.1071/MF01110.

Zhivotovsky LA. 2015. Relationships between Wright’s FST and FIS statistics in a context of Wahlund effect. *Journal of Heredity* 106:306–309. DOI: 10.1093/jhered/esv019.

Chapter 4: Genetic population structure of the vase sponge *Ircinia campana* in the Caribbean is shaped by isolation by distance effects and oceanographic features

Sarah M. Griffiths[1], Mark J. Butler IV[2], Donald C. Behringer[3,4], Thierry Pérez[5] & Richard F. Preziosi[1]

[1] Division of Biology and Conservation Ecology, Manchester Metropolitan University, Chester Street, Manchester, UK.

[2] Department of Biological Sciences, Old Dominion University, Norfolk, Virginia, USA.

[3] School of Forest Resources and Conservation, Program in Fisheries and Aquatic Sciences, University of Florida, Gainesville, Florida, USA.

[4] Emerging Pathogens Institute, University of Florida, Gainesville, Florida, USA.

[5] Institut Méditerranéen de Biodiversité et d'Ecologie Marine et Continentale, UMR CNRS 7263, IRD 237, Aix Marseille Université, Marseille, France.

Keywords: Porifera; Dictyoceratida; Irciniidae; marine; Lesser Antilles; Florida Bay; Florida Keys; South Atlantic Bight; Mesoamerican Barrier Reef System; MBRS; connectivity; microsatellite.

Acknowledgments

We thank the following for granting permission to collect samples: Panamá Ministerio de Ambiente (SC/A-36-16), Belize Fisheries Department (000023-13 and 000010-14), US Department of the Interior National Park Service (Everglades) (EVER-2014-SCI-0050), Florida Fish and Wildlife Commission (SAL-13-0582A-SR), NOAA (GRNMS-2010-001) and all the agencies involved in granting permission to the PACOTILLES campaign (FAUVELOT Cécile (2015) PACOTILLES-1 cruise, RV Antea, http://dx.doi.org/10.17600/15005200; PEREZ Thierry (2015) PACOTILLES-2 cruise, RV Antea, http://dx.doi.org/10.17600/15005300).

Thanks to the following groups and people for sampling assistance: the PACOTILLES campaign team and Antea crew, M. Madrid, R. Grant, F. Clever,

C. Freeman, T. Allen, P. Alford, ReefCI, Frenchies (Caye Caulker), Scuba Sensations (Caye Caulker), and the summer 2014 field work team from University of Florida and Old Dominion University. Thanks to F. Clever for the ArcGIS map of the Florida Keys. I am grateful to G. Fox, P. Fullwood and F. Combe of the DNA Sequencing Facility, and the staff at the Genomic Technologies Core Facility and Bioinformatics Core Facility (all University of Manchester). This was funded by a Natural Environment Research Council (NERC) PhD studentship awarded to SMG, ship time from Flotte Océanographique Française awarded to TP and C. Fauvelot, a grant from the Nature Conservancy and NOAA - Community Restoration Partnership to MJB and DCB, and a grant from the Everglades National Park to MJB.

Author contributions

MJB, DCB, TP and RFP obtained funding for this study; SMG, MJB, DCB and TP collected samples; SMG carried out laboratory work; SMG analysed the data; SMG wrote the chapter.

ABSTRACT

Understanding species' genetic population structure can help us to infer their dispersal patterns and predict population resilience. Spatial population structure is affected by a combination of life history, environment, and the interplay between them, and as such can be difficult to predict. This can be especially true of benthic marine invertebrates that are widely distributed, but occur in naturally fragmented habitats, and whose dispersal occurs in an open environment. This lack of understanding impedes our ability to effectively manage and conserve these species. Here, we describe the population structure of *Ircinia campana*, a sponge that has suffered mass mortalities in the Florida Keys, at various spatial scales. We developed a new set of ten microsatellite loci, and genotyped 440 individuals collected from nineteen sites across the Greater Caribbean, including the South Atlantic Bight, Florida Keys, Mesoamerican Barrier Reef System, southern Caribbean Sea and Lesser Antilles. We found strong genetic structure among sites through the Greater Caribbean, with genetic differences detected even among some sites a few kilometres distance from each other. Isolation by distance effects indicate that structure at larger spatial scales is largely driven by limited dispersal, however, oceanographic features also appeared to correlate with patterns of connectivity and isolation in some areas. Genetic diversity levels were broadly similar across sites, but evidence of genetic bottlenecks was observed at many of the sites. This research aids understanding of Caribbean sponge population genetics and will help to inform management of the species in vulnerable populations.

INTRODUCTION

Understanding how genetic diversity is distributed among populations can reveal fundamental details about species' ecology and evolution. We can use this population genetics data to infer the dispersal capacity of a species, the levels of connectivity (the movement of larvae or gametes among sites) and gene flow among subpopulations, and to better understand species distributions. These factors can have important implications for management and conservation, including for planning protected areas and species restoration (Palumbi, 2003; Baums, 2008; Botsford et al., 2009). For marine species, fully open, panmictic populations were historically assumed to be universal due to the lack of obvious physical or environmental barriers to dispersal in the marine environment, coupled with the presence of a larval stage in the life cycle of many organisms. However, many marine species have been found to have differentiated populations (for example, Knutsen et al., 2003; Duran et al., 2004; Hoffman et al., 2011; Morrison et al., 2011; Vignaud et al., 2014), and it is now clear that there is a dynamic interplay between life history, oceanographic features, and demographic stochasticity that can give rise to complex patterns of genetic structuring.

Dispersal is important in dictating the genetic structure of populations (Bohonak, 1999), and dispersal in turn is shaped by the life history of the organism; species with a larval phase have higher dispersal distances than those without, and those with dispersing larvae have higher dispersal ability than those with non-dispersing larvae (Chust et al., 2016). The duration of the free-drifting larval phase (pelagic larval duration; PLD) also affects dispersal distance, with species that have larger PLDs generally experiencing longer dispersal distances and more highly connected populations (Cowen, Paris & Srinivasan, 2006; Shanks, 2009). Larval swimming behaviour can also impact dispersal distance, with vertical migration able to restrict it (Butler et al., 2011) or enhance it (Paris, Chérubin & Cowen, 2007). These characteristics at the larval life stage can be key in determining connectivity for species with a sessile, benthic adult life.

Ocean circulation is also a powerful determinant of the direction and magnitude of larval dispersal (White et al., 2010), interplaying with life history characters to influence connectivity (Cowen, Paris & Srinivasan, 2006; Kough & Paris, 2015; Pascual et al., 2017). Water movement can create physical connectivity between locations, and therefore networks of dispersal pathways (Christie et al., 2010; Rossi et al., 2014), but can also form barriers to dispersal, create retentive environments, and isolate populations (Baums, Miller & Hellberg, 2005; Palero et al., 2011; Pascual et al., 2017). Biophysical models that take into account ocean circulation patterns and life history parameters can be used to great effect in predicting connectivity and genetic population structure (Paris, Chérubin & Cowen, 2007; Foster et al., 2012; Treml et al., 2012). However, long-term temporal variation in ocean circulation, as well as demographic stochasticity, can make connectivity patterns more chaotic and difficult to predict (Siegel et al., 2008; Hepburn et al., 2008; Hogan, Thiessen & Heath, 2010; Castorani et al., 2017). Furthermore, if detailed life history data are unavailable for the study species, or insufficient parameters are taken into account, predictions can lose their precision (Galarza et al., 2009; Herrera et al., 2016). As such, there is still a great need for empirical evidence of genetic population structure from natural populations of little-studied taxa. In this study, we explore population structure in the marine sponge *Ircinia campana* (Lamarck, 1814) at a range of spatial scales in the Greater Caribbean region.

The Caribbean offers an interesting oceanographic area in which to study population structure, due to its complex and temporally variable current regime (Alvera-Azcárate, Barth & Weisberg, 2009) and heterogeneous, fragmented habitats. The main pattern of water movement is a westerly flow through the Caribbean Sea; water enters from the Atlantic in the east through channels between the Lesser Antilles islands. This forms the Caribbean Current, which travels westward and northward through the Yucatan Channel to the Gulf of Mexico, where it becomes the Loop Current. Upon exiting the Gulf of Mexico, the Loop Current is named the Florida Current as it exits northward through the Straights of Florida. Temporary eddies, permanent and semi-permanent gyres, counter currents and temporal variation produce a complex system of water

movement that could affect the dispersal of larvae and population structure of marine species in space and time.

In Caribbean marine ecosystems, sponges have high abundance, species richness and functional diversity (Diaz & Rutzler, 2001; Bell, 2008). Their biomass and species diversity in coral reef systems is high in the Caribbean compared to other regions (Wilkinson & Cheshire, 1990), and recently their role in the cycling of dissolved organic matter was revealed to be of fundamental importance to sustaining the ecosystem food web in Caribbean coral reefs (de Goeij et al., 2013). In other habitats, such as seagrass, hard bottom and mangrove, they also have a number of important ecological roles (e.g. Butler et al., 1995; Peterson et al., 2006; Archer, Stoner & Layman, 2015). Similarly to many marine animals, sponge communities have suffered mass mortalities in the greater Caribbean region (Butler et al., 1995; Wulff, 2006; Stevely et al., 2010). Due to their ecological importance and vulnerability to die-offs, understanding the spatial ecology of this group may be important for creating ecosystem-based management plans for the region. Furthermore, sponges display a range of reproductive strategies, including viviparity, oviviparity and asexuality (both gemmule production and fragmentation) (Maldonado & Riesgo, 2008), with multiple reproductive modes potentially present in one species (e.g. Zilberberg, Solé-Cava & Klautau, 2006) making them interesting subjects for population genetics studies.

There have been few studies so far on Caribbean sponge population genetics considering their importance and abundance in the region, with only three species studied to date (*Xestospongia muta*: López-Legentil & Pawlik, 2009; de Bakker et al., 2016; Richards et al., 2016; *Callyspongia vaginalis*: DeBiasse, Richards & Shivji, 2010; *Cliona delitrix*: Chaves-Fonnegra et al., 2015) out of at least 676 species found in the region (Van Soest et al., 2012). In this chapter, I investigate genetic diversity and genetic population structure in *Ircinia campana* at a range of spatial scales in the Greater Caribbean region. *Ircinia campana*, also known as the vase sponge, is a common species in coral reef and lagoon habitats in the Tropical Western Atlantic region (van Soest 2010). It is a

keratose sponge in the largest class of sponge, Demospongiae, in the order Dictyoceratida and the family Irciniidae. It is conspicuous in Caribbean ecosystems due to its large size and vase-shaped morphology, which provides habitat structure. The species has, however, suffered mass mortalities in the Florida Keys (Butler et al., 1995; Stevely et al., 2010), where in hard bottom habitats it is among the most dominant benthic invertebrates. These demographic changes make *I. campana* an important candidate for population genetic study in order to understand levels of connectivity among populations, their potential resilience, and the factors shaping its dispersal.

A recent study characterised mitochondrial haplotypes in *I. campana* from five locations across the Caribbean and South Atlantic Bight: Bocas del Toro, Panama; Carrie Bow Caye, Belize; Key Largo, Florida Keys, Florida USA; Gray's Reef, Georgia, USA; and Wilmington, North Carolina, USA (Marino et al., 2017). Two cytochrome oxidase I (*COI*) genes were sequenced in fifteen individuals (the Folmer fragment [Folmer et al., 1994] and *I3-M11* partition [Erpenbeck, Hooper & Wörheide, 2006]), and although the objective of the study was not the assessment of population structure, the results do indicate that these genes are insufficiently variable to study population structure in this species. Only two haplotypes were found across the range; the first was found in North Carolina and Georgia (USA), the second in Panama and Belize, and a mix of haplotypes was found in Florida. In this study, we develop microsatellite markers for the species and we study of fine-scale population structure and genetic diversity throughout the Greater Caribbean.

METHODS

Sampling

All sampling was carried out with the permission of relevant national authorities (permit numbers listed in the Acknowledgements), and was non-lethal to the sponge, with as minimal amounts of damage incurred as possible. Sampling was conducted at nineteen sites in the Greater Caribbean (Table 1, Figure 1), with between ten and 41 individuals sampled per site. Samples were collected while scuba diving or snorkelling, from depths of one to 25 m. Small pieces of sponge tissue were cut from the edge of the sponge, stored in plastic bags with seawater and preserved in 95% ethanol (or the highest concentration available) as soon as possible after surfacing. Ethanol was replaced after 24 hours to avoid its dilution by the seawater held in the sponge tissue.

Microsatellite development and multiplex characterisation

A single *Ircinia campana* individual was used for microsatellite development. To avoid contamination from microbial DNA and associated invertebrates, tissue was dissected under a stereomicroscope, and prokaryote and eukaryote cells were separated (as far as possible) using the centrifugation method of Freeman & Thacker (2011). DNA was extracted from the 'eukaryotic' cell pellet with the DNeasy® Blood and Tissue Kit (Qiagen), and 50ng of DNA was used for library construction with the Nextera® DNA Sample Preparation Kit. Paired-end sequencing on the Illumina MiSeq (2 x 250 base pairs) resulted in 2 x 1,939,933 reads. The bioinformatics pipeline developed by Griffiths et al. (2016), and detailed in Chapter 2 of this thesis, was used to quality filter reads (Trimmomatic; Bolger, Lohse & Usadel, 2014), locate microsatellite repeat sequences (pal_finder; Castoe, Poole & Koning, 2012), design primers (Primer3; Untergasser et al., 2012) and select optimal loci. Optimal loci were considered as such for containing perfectly repeating motifs and having primer sequences that were only found once in the entire set of reads. Following quality filtering using the settings described within the cited article and Chapter 2, 2 x 1,175,084 reads remained, and thirty-six tri- and tetra-nucleotide loci were selected for testing.

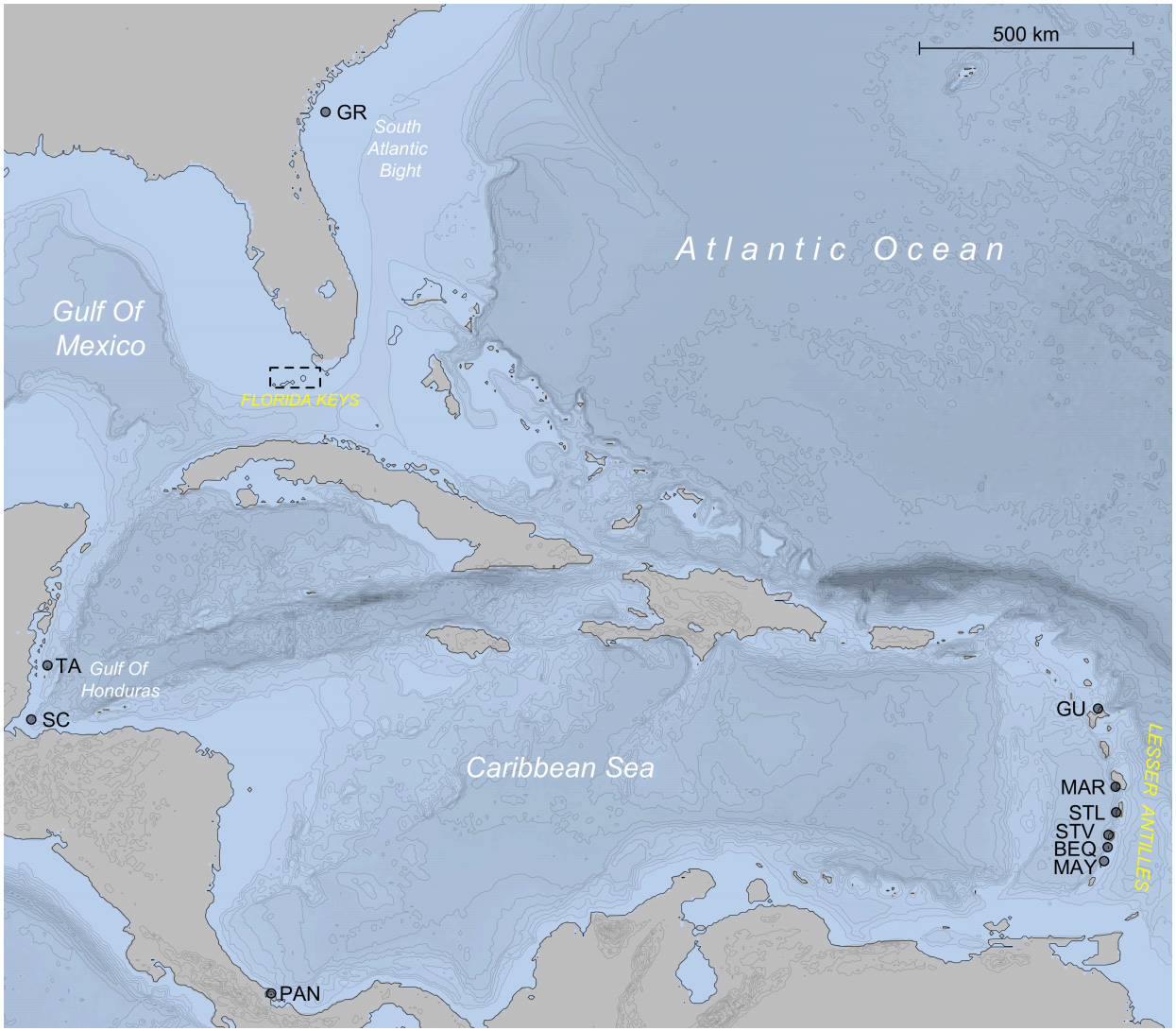


Figure 1A: *Ircinia campana* sampling sites. Dashed-line box outlines Florida Keys; see Figure 1B for detailed map of sampling sites in this area. Data source: ETOPO1, NOAA.

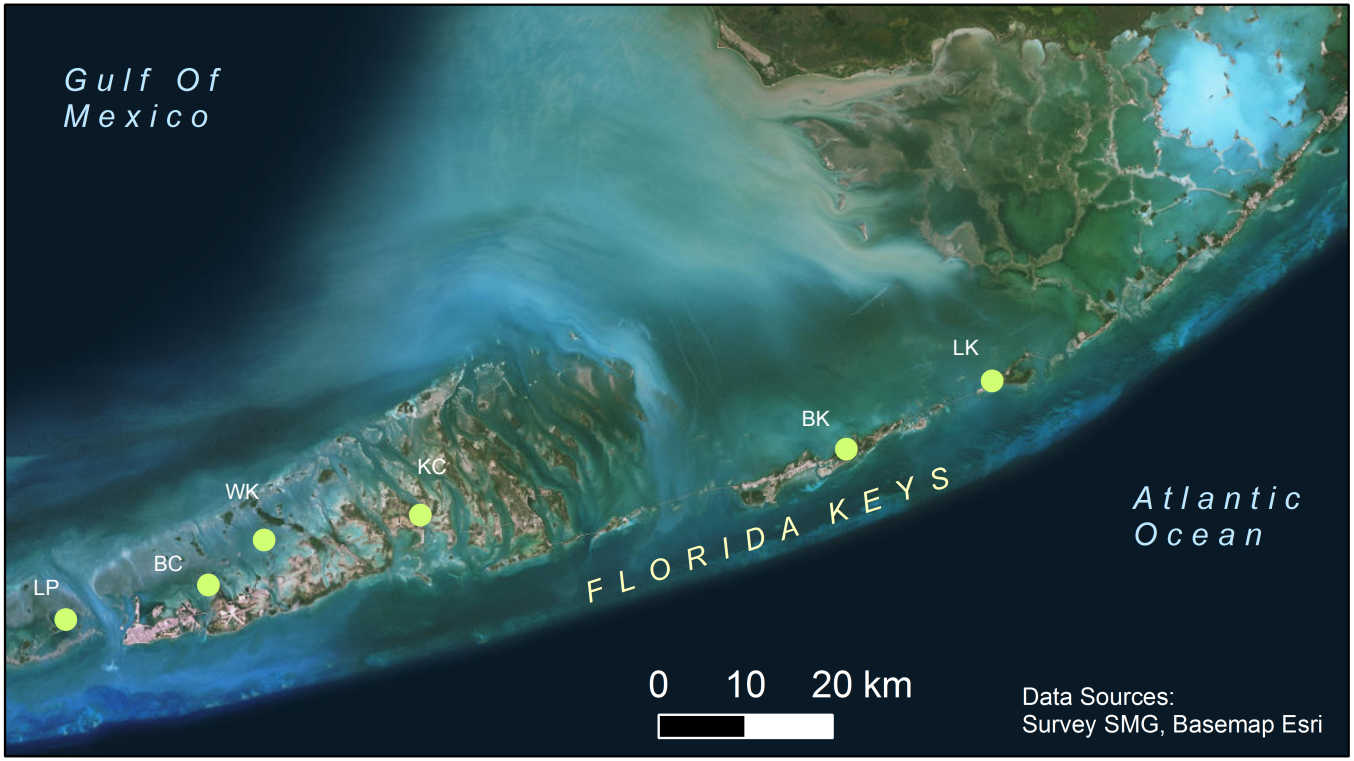


Figure 1B: *Ircinia campana* sampling sites in the Florida Keys.

Table 1: *Ircinia campana* sampling information.

Country	Site	Site code	n	Latitude, Longitude	Date (M/YY)
Georgia, USA	Gray's Reef National Marine Sanctuary	GR	10	31.40480, -80.86677	6/13
Florida, USA	Long Key	LK	20	24.81437, -80.83073	7/14
	Bamboo Key	BK	16	24.74429, -80.99504	7/14
	Kemp Channel	KC	20	24.67687, -81.47577	7/14
	Waltz Key	WK	17	24.65108, -81.65213	7/14
	Boca Chica Channel	BC	18	24.60495, -81.71508	7/14
	Lakes Passage	LP	20	24.56948, -81.87572	7/14
Belize	Turneffe Atoll	TA	35	17.54436, -87.82664	4/13
	Tom Owen's Caye, Sapodilla Cayes	SC	39	16.18898, -88.23277	4/13
Panama	Bocas Del Toro 1	PAN1	15	9.32833, -82.22668	7/16
	Bocas Del Toro 2	PAN2	16	9.30604, -82.23275	7/16
	Bocas Del Toro 3	PAN3	11	9.2413, -82.1737	7/16
St. Vincent & The Grenadines	Mayreau	MAY	20	12.64218, -61.37975	5/15
	Bequia	BEQ	27	12.99128, -61.29043	5/15
	St. Vincent	STV	40	13.18303, -61.26945	5/15
St. Lucia	Anse Chastenet	STL	41	13.86413, -61.07882	5/15
Martinique	Grande Anse d'Arlet	MAR1	17	14.5059, -61.0932	5/15
	Diamond Rock	MAR2	26	14.4426, -61.04013	5/15
Guadeloupe	Grande-Terre north	GU	32	16.46233, -61.53062	5/15

n: Number of samples.

PCR amplifications were carried out with the Type-it® Microsatellite Kit (Qiagen) in 5µl reaction volumes with the following thermal cycling conditions: 95°C initial denaturation for 5 minutes, 28 cycles of 95°C for 30 seconds, 60°C for 90 seconds and 72°C for 30 seconds, and a final extension at 60°C for 30 minutes. Optimisation was carried out through alteration of cycle number and annealing temperature when PCR products for certain loci showed poor amplification on agarose gels.

After this process, eighteen loci produced clear bands on agarose gels. To further examine their suitability and size ranges, PCRs were repeated as above, but with a universal primer tail fluorophore labelling system using the M13(-21) tail (TGTAAAACGACGGCCAGT) and the 6-FAM dye to fluorescently label PCR products (Culley et al., 2013). These were then analysed on a DNA Analyzer 3730 (Thermo Fisher Scientific) at the University of Manchester DNA

Sequencing Facility with the GeneScan™ LIZ® 1200 (Thermo Fisher Scientific) size standard. Sequencer traces were then viewed using Genemapper v3.7 (Thermo Fisher Scientific), and eleven loci were deemed suitable for use in the study, as they exhibited easy-to-score peaks, consistent amplification and minimal stutter. Two multiplexes incorporating the eleven loci were designed using Multiplex Manager (Holleley & Geerts, 2009), with the following universal tails to fluorescently label PCR products: 6FAM- TGTAAAACGACGGCCAGT (M13-21; Culley et al., 2013), HEX- CGGAGAGCCGAGAGGTG (Tail D; Blacket et al., 2012) and PET-CACTGCTTAGAGCGATGC (M13 modified B; Culley et al., 2013) (Table 2).

Sample genotyping

All samples were dissected under a stereomicroscope to remove visible endo- and epibionts from the sponge tissue, and DNA was extracted using DNeasy® Blood and Tissue Kit (Qiagen). Multiplex PCRs were carried out with the Type-it® Microsatellite PCR Kit (Qiagen) in 5µl reaction volumes and with primer concentrations altered to those specified in Culley et al., (2013) (i.e. 2µM of the forward primer, 2µM of the reverse primer, and 0.5µM of the third universal tail primer), but otherwise according to the manufacturers protocol. All plates contained positive and negative controls. Thermal cycling conditions used were the same as the singleplex conditions listed above, except with the annealing temperature increased to 63°C in multiplex B. PCR products were sized using the DNA Analyzer 3730 at the DNA Sequencing Facility at the University of Manchester, using the GeneScan™ LIZ® 1200 size standard. Alleles were scored with Genemapper v3.7, sizes corrected according to positive controls (if necessary), and binned in MsatAllele v1.02 (Alberto, 2009) in R 3.3.3 (R Core Team, 2017).

Quality control and summary statistics

Linkage disequilibrium between loci was examined using Genepop on the Web (Raymond & Rousset, 1995; Rousset, 2008), and significance levels were corrected using Benjamini and Yekutieli's (2001) false discovery rate (FDR) correction (B-Y correction; Narum, 2006). Null allele frequencies were

calculated in FreeNA (Chapuis & Estoup, 2007). A post hoc analysis was then conducted to determine the influence of null alleles on population differentiation estimation. This analysis was carried out by calculating global F_{ST} with and without ENA correction for null alleles using FreeNA, then removing the locus with the highest null allele frequency and recalculating both values. This was repeated sequentially until only one locus remained.

Probability of deviation from Hardy-Weinberg Equilibrium (HWE), observed (H_O) and expected (H_E) heterozygosity were calculated in GenoDive v2.0b23 (Meirmans & Van Tiendener, 2004). Inbreeding coefficients were calculated corrected for null alleles in INEST v2.1 (Chybicki & Burczyk, 2009) using the Bayesian Interacting Multiple Model (IMM) approach, with 500,000 Markov chain Monte Carlo (MCMC) cycles and 50,000 burnin cycles. The model was run using all combinations of parameters for possible null allele causes ('n': null alleles; 'b': genotyping error; 'f': inbreeding), and Bayesian deviance information criterion (DIC) was used to infer which parameters contributed more to the observed data.

Population genetics analysis

The traditional measure of subpopulation differentiation F_{ST} can be found to underestimate differentiation when variation in markers is high (Charlesworth, 1998; Jost, 2008), therefore we also used Jost's *D* to estimate population differentiation (Jost, 2008), as high numbers of alleles were present in a number of loci (Table 2). Population pairwise *D* and F_{ST} were calculated in GenoDive using the eight loci with the lowest null allele frequencies, and significance of the pairwise F_{ST} values was tested using 50,000 permutations (with B-Y correction applied for multiple tests). F_{ST} was also calculated between population pairs for all ten loci with correction for null alleles in FreeNA using ENA correction (Chapuis & Estoup, 2007), in order to compare these with the non-corrected values calculated from the eight loci.

Isolation by distance (IBD) patterns in the data were identified by testing the correlation between pairwise matrices of linearised genetic distances (F_{ST}/1-F_{ST}

[null allele corrected] and *D*/1-*D*) and the logarithm of oceanographic distances with Mantel tests in the 'ade4' package in R, using 9999 permutations. Pairwise oceanographic distance between sites was calculated using the 'marmap' package in R (Pante & Simon-Bouhet, 2013).

Various methods were used to investigate the presence of population clusters. Discriminant Analysis of Principle Components (DAPC) (Jombart, Devillard & Balloux, 2010) was carried out using the 'adegenet' package (v2.0.1) in R (Jombart, 2008). This method uses principle components analysis (PCA) to transform the data, and then retained principle components are used in discriminant analysis (DA). This is effective in minimising within-group genetic variation and maximising between-group variation, and does not make assumptions regarding HWE in populations. The number of principle components retained varied among analyses; if too many are retained, resulting membership probabilities can be unstable, and therefore the maximum number without compromising stability were retained (displayed graphically in each DAPC plot). The analysis was repeated methodically in a hierarchical clustering approach *sensu* Vaha et al. (2007). Principle Coordinates Analysis was carried out in GenAlEx 6.503 (Peakall & Smouse, 2012) using population pairwise F_{ST} (null-allele corrected 10 loci) and *D* (8 loci) as calculated above.

A spatially-explicit Bayesian approach was used in Geneland v4.0.6 (Guillot, Mortier & Estoup, 2005; Guillot, Santos & Estoup, 2008) in R to identify the number of population clusters *K* in the dataset. This program functions in a similar way to STRUCTURE (Pritchard, Stephens & Donnelly, 2000), but can also use geographical coordinates of sampling sites to aid clustering. Geneland was run with 1,000,000 iterations, 100 thinning and 500 burnin using the uncorrelated allele frequencies, spatial and null allele models, and repeated for ten independent runs of K from one to eight. The maximum number of nuclei was set to 1320, the maximum rate of the Poisson process to 440, and the spatial uncertainty on coordinates to 0.0005.

Genetic diversity and bottleneck signatures

We tested for genetic signatures of population bottlenecks at the sampling sites using the methods implemented in INEST v2.1. This uses two techniques – testing for heterozygosity excess in comparison to number of alleles per locus (as in popular software Bottleneck; Piry, Luikart & Cornuet, 1999) and testing for M-ratio deficiency (Garza & Williamson, 2001). Heterozygosity excesses in comparison to number of alleles occurs because when bottlenecks occur, the number of alleles reduces faster than heterozygosity (which can remain fairly stable). M-ratio is the ratio of allele number to allele size range; in the case of population size reductions, the number of alleles is expected to reduce, but the size range is expected to remain. The two-phase mutation model was run with 100,000 coalescent simulations, and significance tested using the Wilcoxon signed-rank test calculated based on 1,000,000 permutations.

Genetic diversity of sites was assessed using ADZE (Szpiech, Jakobsson & Rosenberg, 2008), which uses a rarefaction method to calculate allelic richness and private allelic richness corrected for sampling size. Gray's Reef was excluded from this analysis due to its small sample size, and because Icam3 did not amplify in any of the samples from this site.

RESULTS

Marker development and summary statistics

From the eleven loci developed (Table 2), one locus was excluded from analysis due to significant linkage disequilibrium. Following its exclusion, the remaining loci were in linkage equilibrium after B-Y FDR correction. One pair of identical multilocus genotypes was identified (both individuals originating from Turneffe Atoll [TA]); one of the individuals was removed from the dataset for further analyses. The number of alleles per locus across all populations ranged from four (Icam32) to ninety-seven (Icam23) (Table 2). Average null allele frequencies were high at many loci, overall ranging from <0.001 (Icam32) to 0.323 (Icam34) (Table 2). The post hoc analysis showed that the two loci with the highest null allele frequencies, Icam34 and Icam10 skewed null-uncorrected global F_{ST} estimates over 0.01 compared to the null-corrected F_{ST} (Table 3). As a result of this, in analyses when correction for null alleles could be used (Geneland analysis, F_{ST} and F_{IS} calculations), all ten loci were used, but in all other analyses Icam34 and Icam10 were excluded.

The average number of alleles per site ranged from 3.78 (GR; Gray's Reef) to 13.1 (STV; St. Vincent) (Table 4), however, due to variation in sample sizes among sites this does not reflect differences among sites well. Rarefied allelic richness per site (not including Gray's Reef due to non-amplification of Icam3 in all samples) ranged from 3.850 (TA; Turneffe Atoll) to 5.500 (MAY; Mayreau), and rarefied private allelic richness ranged from 0.532 (TA; Turneffe Atoll) to 1.419 (SC; Sapodilla Cayes) (Table 4, Figure 2). However, standard error of the mean was large and thus many of the error bars overlapped.

Observed heterozygosity ranged from 0.305 (KC; Kemp Channel) to 0.565 (PANC) within sites, and expected heterozygosity ranged from 0.528 (GR; Gray's Reef) to 0.761 (MAY; Mayreau) (Table 4). Significant departures from HWE were observed in all loci but Icam32, at between two and seventeen sites (Table 5). Five loci departed from HWE in over half of the sites. Null allele corrected average F_{IS} values were all positive, ranging from 0.0342 in St.

Vincent (STV) to 0.0351 in Bamboo Key, Florida (BK) (Table 4), but as the posterior 95% probability intervals included zero for all sites, the F_{IS} estimates are not significantly different from zero. Model comparison using DIC values revealed that null alleles were important in affecting F_{IS} estimates in all of the populations, along with genotyping error and/or inbreeding in some of the populations (Table 6). The best-suited model was identified as that with the lowest DIC value.

Table 2: Characterisation of 11 *Ircinia campana* microsatellite loci and two multiplexes.

Locus name	Motif	Primer sequences (5'→ 3')	MP	L	Na	Size range (bp)	N	GenBank accession no.
Icam10	ATC	F: TATGCCGATACCCAATGACATCACC R: GCTGTGTGGATACAGTAAATGTCCAACG	A	3	19	331-414	0.250	MF987878
Icam3	ATAC	F: ACAAGTGCAGCATGGAGAATGTGC R: CCTGTGTGTATCCATCACAAGTGTCC	A	1	32	503-616	0.097	MF987882
Icam4	AATC	F: ACAGCATGGCAGTGTTTTCTGATCG R: ACATATCGACAGGACAAGCTGATGG	A	1	19	137-275	0.042	MF987879
Icam18	ATAC	F: TCTTGGCAGCCTTAGATTGAACAGC R: TGCAGTGGCTTCTATGACTTTAAACAAAGC	A	2	96	346-617	0.069	MF987883
Icam31	ATAC	F: TGTTATAAACTGCGGCTATGGATGTACG R: GTCATGCATCCAGAATGACCACTCC	A	1	21	343-394	0.059	MF987880
Icam32	TTC	F: GCATTACAAATAGGTTGGCCTTTGTGG R: GCAAGAAAGCAAATGTTAGAGCGAACC	A	3	4	277-282	< 0.001	MF987887
Icam23	ATAC	F: TGCTGGACAAGAGAGGTTTCACTGC R: TTGAACTCAGGCCTCCTGACATACG	B	1	97	396-825	0.236	MF987884
Icam24	ATAC	F: CATTGGTTAACAAACCTATAGCAACCC R: ACTGCCTGTACAAACATTAACATGC	B	2	83	428-830	0.239	MF987885
Icam26	ATAC	F: GTCTTGTGTGGACCTTCAGATCACC R: GGGATGATGTGATAAAGCATTTCC	B	1	21	350-409	0.200	MF987886
Icam34	ATAC	F: TGCACATAACCCCTATTTTCTCATGCC R: GCTATTTTCGTGCCATGATTTCAGC	B	3	58	201-459	0.323	MF987881

MP: multiplex; L: tail sequence/ florescent label combination (1: 6FAM- TGTAAAACGACGGCCAGT; 2: HEX- CGGAGAGCCGAGAGGTG; 3: PET-CACTGCTTAGAGCGATGC); Na: number of alleles per locus; bp: base pairs; N: frequency of null alleles.

Table 3: Cumulative average null allele frequency and global F_{ST} calculated with and without ENA correction for null alleles.

Number of loci →	1	2	3	4	5	6	7	8	9	10
Average null allele frequency	0.000	0.042	0.059	0.069	0.097	0.200	0.236	0.239	0.250	0.323
Global F_{ST} without ENA	0.035	0.126	0.162	0.127	0.132	0.152	0.141	0.128	0.134	0.141
Global F_{ST} with ENA	0.035	0.140	0.162	0.127	0.130	0.143	0.133	0.120	0.123	0.126
Difference in F_{ST} with and without ENA	0.000	-0.014	0.000	0.000	0.002	0.009	0.008	0.008	0.011	0.015

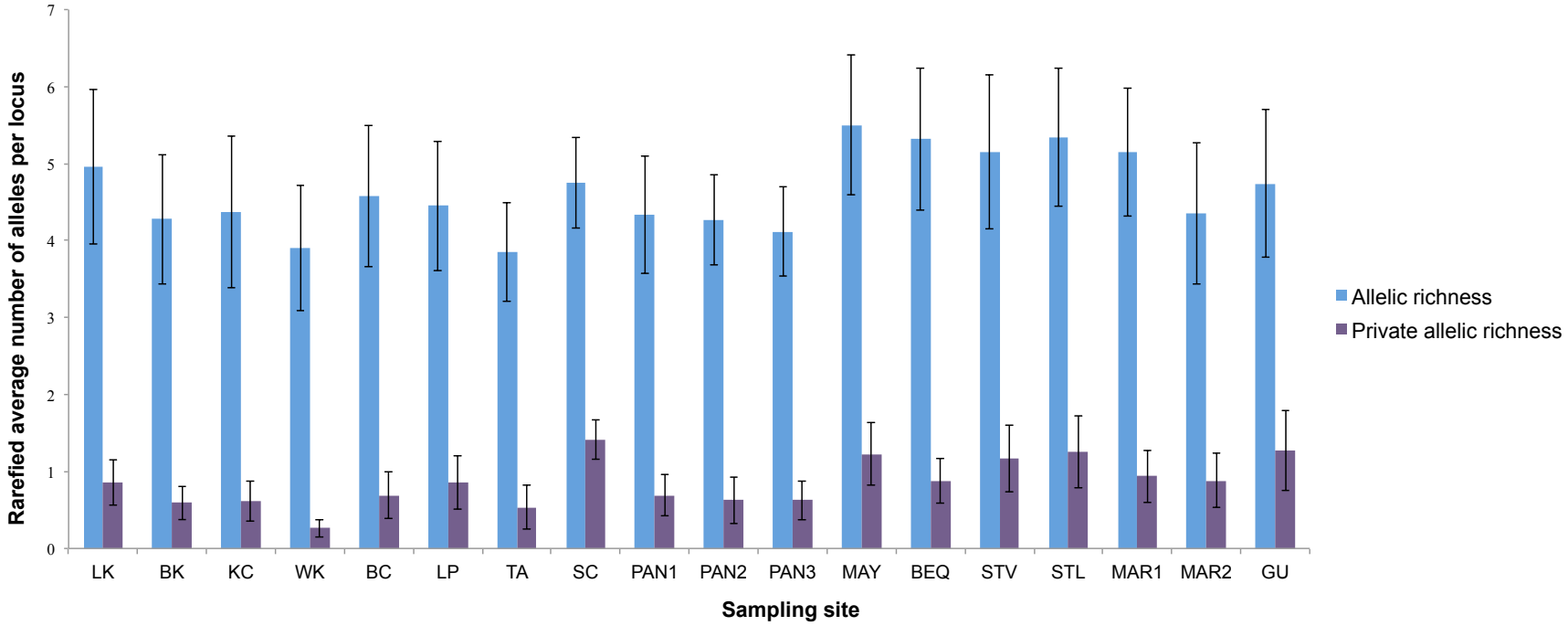


Figure 2: Average allelic richness and private allelic richness per site rarefied for sample size. Error bars are +/- 1 SE.

Table 4: Summary statistics by site for *Ircinia campana* microsatellite analysis.

Site	$H_{O\ (8)}$	$H_{E\ (8)}$	AvF_{IS}	Model	95% HPDI	AR	RAR *(SE)*	RPAR *(SE)*
GR	0.325	0.528	0.0371	nb	0 – 0.1306	3.778	-	-
LK	0.414	0.665	0.1486	nb	0 – 0.3394	8.600	4.957 *(1.002)*	0.851 *(0.294)*
BK	0.365	0.641	0.3531	nb	0 – 0.5757	6.000	4.275 *(0.839)*	0.590 *(0.215)*
KC	0.305	0.578	0.0992	n	0 – 0.2905	7.300	4.370 *(0.983)*	0.616 *(0.263)*
WK	0.317	0.557	0.2259	nf	0 – 0.4156	5.300	3.902 *(0.813)*	0.264 *(0.115)*
BC	0.377	0.650	0.1483	nb	0 – 0.3710	6.700	4.576 *(0.923)*	0.693 *(0.308)*
LP	0.365	0.644	0.092	nb	0 – 0.2719	6.900	4.449 *(0.839)*	0.854 *(0.346)*
TA	0.397	0.603	0.0471	nf	0 – 0.1166	6.500	3.850 *(0.641)*	0.532 *(0.288)*
SC	0.512	0.725	0.0888	nfb	0 – 0.1534	9.600	4.747 *(0.587)*	1.419 *(0.260)*
PAN1	0.509	0.634	0.077	nb	0 – 0.2285	5.900	4.329 *(0.761)*	0.690 *(0.273)*
PAN2	0.565	0.695	0.0326	nb	0 – 0.1106	5.600	4.271 *(0.587)*	0.624 *(0.300)*
PAN3	0.466	0.661	0.0573	nb	0 – 0.1672	4.800	4.113 *(0.580)*	0.628 *(0.255)*
MAY	0.535	0.761	0.1588	nfb	0 – 0.3043	9.900	5.500 *(0.912)*	1.229 *(0.414)*
BEQ	0.496	0.692	0.0927	nb	0 – 0.2372	11.800	5.313 *(0.918)*	0.878 *(0.289)*
STV	0.457	0.654	0.0342	nb	0 – 0.1043	13.100	5.151 *(0.999)*	1.163 *(0.435)*
STL	0.473	0.740	0.0752	nb	0 – 0.1948	13.000	5.344 *(0.898)*	1.255 *(0.470)*
MAR1	0.495	0.687	0.1022	n	0 – 0.2507	8.700	5.152 *(0.917)*	0.938 *(0.349)*
MAR2	0.502	0.627	0.0483	n	0 – 0.1354	7.800	4.352 *(0.835)*	0.883 *(0.341)*
GU	0.368	0.642	0.0857	n	0 – 0.2424	10.100	4.739 *(0.955)*	1.269 *(0.519)*

$H_{O\ (8)}$: Observed heterozygosity (over 8 loci); $H_{E\ (8)}$: Expected heterozygosity (over 8 loci); AvF_{IS}: Average inbreeding coefficient corrected for null alleles; Model: Model with lowest Deviance Information Criteria for estimating the inbreeding coefficient in the presence of combinations of null alleles (n), inbreeding (f) and genotyping error (b); 95% HPDI: 95% highest posterior density interval; AR: Average allelic richness; RAR *(SE)*: Average rarefied allelic richness (Standard error of the mean); PRAR *(SE)*: average rarefied private allelic richness (Standard error of the mean).

Table 5: Probability of deviation from Hardy-Weinberg Equilibrium for each population and locus.

	lcam23	lcam24	lcam26	lcam34	lcam10	lcam18	lcam3	lcam31	lcam32	lcam4
GR	0.659	0.334	**0.006**	0.066	**0.001**	0.286	---	0.158	0.843	---
LK	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.310	0.174	0.849	1.000
BK	**0.000**	0.028	**0.004**	**0.000**	**0.000**	**0.000**	0.636	**0.000**	0.438	---
KC	**0.000**	**0.000**	**0.000**	0.085	**0.000**	0.300	**0.005**	**0.002**	---	---
WK	**0.001**	**0.000**	**0.000**	0.078	**0.000**	**0.003**	0.364	0.014	0.908	---
BC	**0.005**	**0.000**	**0.000**	**0.001**	**0.000**	**0.005**	0.014	0.386	0.916	---
LP	**0.000**	**0.000**	**0.000**	**0.002**	**0.002**	0.042	0.035	**0.001**	0.925	---
TA	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.036	0.155	0.100	0.638	0.060
SC	0.059	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.053	0.592	0.727	0.058
PAN1	**0.000**	**0.000**	0.166	**0.002**	0.109	0.381	0.432	0.543	0.537	0.889
PAN2	**0.000**	0.333	0.033	**0.001**	0.009	0.070	0.188	0.149	0.180	0.693
PAN3	**0.000**	**0.002**	0.111	**0.001**	---	0.580	0.397	0.338	0.856	0.857
MAY	**0.000**	**0.000**	**0.004**	**0.000**	**0.000**	0.015	**0.001**	0.520	0.207	**0.003**
BEQ	**0.000**	**0.000**	**0.007**	**0.000**	**0.000**	**0.005**	**0.000**	0.521	0.530	0.239
STV	**0.000**	**0.000**	**0.000**	**0.000**	**0.002**	0.645	**0.000**	0.152	0.255	0.082
STL	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.003**	0.013	0.009	0.613	**0.000**
MAR1	**0.000**	**0.000**	0.086	**0.000**	**0.004**	0.011	0.181	0.493	0.882	0.980
MAR2	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.322	0.631	0.819	0.909
GU	**0.000**	**0.000**	**0.000**	**0.000**	**0.000**	0.290	**0.000**	0.563	0.985	0.046

Significant p values after correction for multiple tests highlighted in bold ($p < 0.0086$).

Table 6: Deviance information criterion (DIC) using various combinations of parameters in inbreeding coefficient (F_{IS}) estimation models in *Ircinia campana*.

Site	Model					
	nfb	**nf**	**nb**	**bf**	**n**	**b**
GR	291.989	293.466	**290.59**	-	292.542	-
LK	1067.314	1090.907	**1065.222**	1097.398	1081.915	1227.870
BK	701.601	711.539	**701.122**	707.175	711.035	807.151
KC	926.179	925.259	926.460	946.704	**925.147**	1056.005
WK	672.205	**672.110**	676.515	684.412	677.906	738.229
BC	857.693	862.754	**857.453**	867.695	862.041	957.969
LP	972.074	971.871	**971.097**	986.096	971.113	1056.805
TA	1485.159	**1483.566**	1488.071	1551.031	1488.037	1683.175
SC	**2188.793**	2196.549	2202.342	2239.740	2214.965	2384.137
PAN1	487.430	496.328	**486.003**	497.951	495.728	515.232
PAN2	703.262	713.151	**701.930**	721.475	711.647	731.595
PAN3	806.103	813.974	**805.861**	814.440	813.670	829.725
MAY	**1246.373**	1247.470	1249.000	1256.874	1248.051	1338.043
BEQ	1720.730	1730.438	**1718.836**	1739.943	1728.436	1852.124
STV	2389.176	2428.908	**2387.637**	2448.622	2427.530	2635.384
STL	2596.634	2605.839	**2594.866**	2644.272	2603.508	2894.867
MAR1	1020.657	1019.150	1020.065	1044.973	**1019.061**	1120.244
MAR2	1406.457	1406.817	1405.281	1438.552	**1405.035**	1504.624
GU	1677.984	1674.430	1676.310	1705.252	**1673.548**	1900.984

n: null alleles; f: inbreeding; b: genotyping error. Lowest DIC values are highlighted in bold; these correspond to the best-fitting model for the data.

Population structure

Strong population structure was present over the region. F_{ST} values calculated from the eight loci set ranged from 0.008 (between Long Key [LK] and Boca Chica Channel [BC]) and 0.265 (between Gray's Reef [GR] and Panama 1 [PAN1]) (Table 7). F_{ST} calculated from the eight loci set and F_{ST} calculated from ten loci set with ENA correction for null alleles were broadly similar, in both the F_{ST} values themselves and the pattern among sites (Table 7, Table 8). *D*, which is able to range to higher values than F_{ST}, ranged from 0.009 (between Long Key [LK] and Boca Chica Channel [BC]) to 0.590 (Gray's Reef [GR] and Sapodilla Cayes [SC]). Pairwise F_{ST} values (calculated from the 8 loci set) were significant in 162 out of the 171 comparisons; seven non-significant pairwise comparisons were between Floridian sites, one was between Gray's Reef (Georgia, USA) and Bamboo Key (Florida USA), and one was between Panamanian sites (PAN2 and PAN3) (Table 7).

Isolation by distance was strong and significant over all sites (3 – 2975 km; F_{ST}: r = 0.599, $p < 0.001$; *D*: r = 0.442, p < 0.001) (Figure 2A), and within the Lesser Antilles subset (15-443 km; F_{ST} r = 0.793, p = 0.001; *D*: r = 0.688, p = < 0.001) (Figure 2B) but not within the Florida subset (10 – 115 km; *D*: r = 0.025, p = 0.476; F_{ST}: r = 0.036, p = 0.483) (Figure 2C). However, Figure 3A shows that there is a large amount of variation in genetic distance found at the longer oceanographic distances (i.e. $\log_{10}$ 3-3.5 km).

Geneland analysis identified five population clusters in all independent runs (Figure 3). However, two different solutions were reached among different runs. The six runs with the highest log posterior probabilities converged upon the same solution (proportion of membership to the cluster in parentheses): Cluster 1: Gray's Reef and Florida (0.3188); Cluster 2: Turneffe Atoll (Belize) and Panama (0.406); Cluster 3: Sapodilla Cayes (0.345); Cluster 4: St. Vincent and the Grenadines and St. Lucia (0.332); Cluster 5: Martinique and Guadeloupe (0.380). The remaining Geneland runs converged on an alternative clustering solution, but had the lowest log posterior probability scores out of all runs, and are therefore less well supported. The cluster membership reached in these

solutions was as follows: Cluster 1: Gray's Reef and Florida (0.470); Cluster 2: Turneffe Atoll (Belize), Panama and Guadeloupe (0.416); Cluster 3: Sapodilla Cayes (0.432); Cluster 4: St. Vincent and the Grenadines (0.450); Cluster 5: Martinique and St. Lucia (0.430).

The DAPC conducted on all sites shows the Sapodilla Cayes (SC) separated from the rest of the sites (Figure 4A). The DAPC carried out on all sites but SC then broadly split the sites into three clusters as follows: 1) All USA sites; 2) St. Vincent, the Grenadines (Mayreau and Bequia), St. Lucia and both Martinique sites (STV, MAY, BEQ, STL, MAR1 and MAR2); 3) All Panama sites, Turneffe Atoll and Guadeloupe (PAN1, PAN2, PAN3, TA, GU) (Figure 4B). DAPC analyses conducted on these clusters revealed further substructure (Figures 4C, 4D, 4E). Notably, in the USA cluster analysis, the Lakes Passage (LP) separated from the other sites (Figure 4C). Also of interest was that St. Vincent, Mayreau and Bequia remained as a tight cluster (STV, MAY, BEQ), while the Martinique sites were separated from each other (MAR1, MAR2) (Figure 4D).

The PCoA showed many similarities with the patterns revealed by the DAPC (Figure 5A). For example, there was also clustering of the USA sites, which were separated from all sites but Martinique2 (MAR2) by the first axis. St. Vincent and the Grenadines also formed a cluster, both in the global analysis (Figure 5A) and in the Lesser Antilles only analysis (Figure 5B). Sapodilla Cayes (SC) and Turneffe Atoll (TA) were also very distant from each other (Figure 5A).

Genetic bottlenecks

A significant heterozygosity excess in comparison to allelic richness was found at Bamboo Key (BK; p = 0.020), and M-ratio deficiencies were found at Long Key (LK; p = 0.014), Bamboo Key (BK; p = 0.010), Boca Chica Channel (BC; 0.014), Kemp Channel (p = 0.020), Lakes Passage (LP; p = 0.004), Turneffe Atoll (TA; p = 0.007), Sapodilla Cayes (SC; p = 0.000), Panama 1 (PAN1; p = 0.037), Panama 2 (PAN2; p = 0.042), Panama 3 (PAN3; p = 0.007),

Guadeloupe (GU; p = 0.019), Martinique 1 (MAR1; p = 0.042), Bequia (BEQ; p = 0.032) and Mayreau (MAY; p = 0.0417).

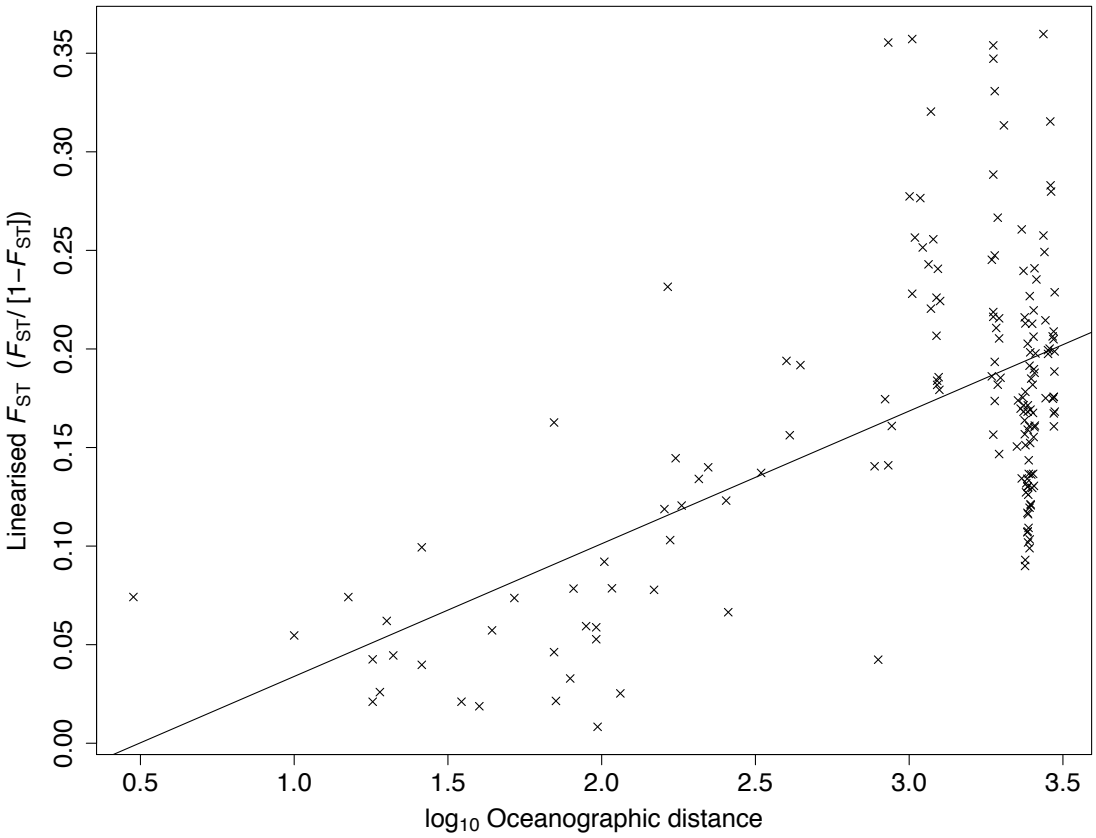
Table 7 – Pairwise F_{ST} (below diagonal, shaded; calculated using 10 loci and ENA correction for null alleles) and Jost's *D* (above diagonal, non shaded; calculated using 8 loci).

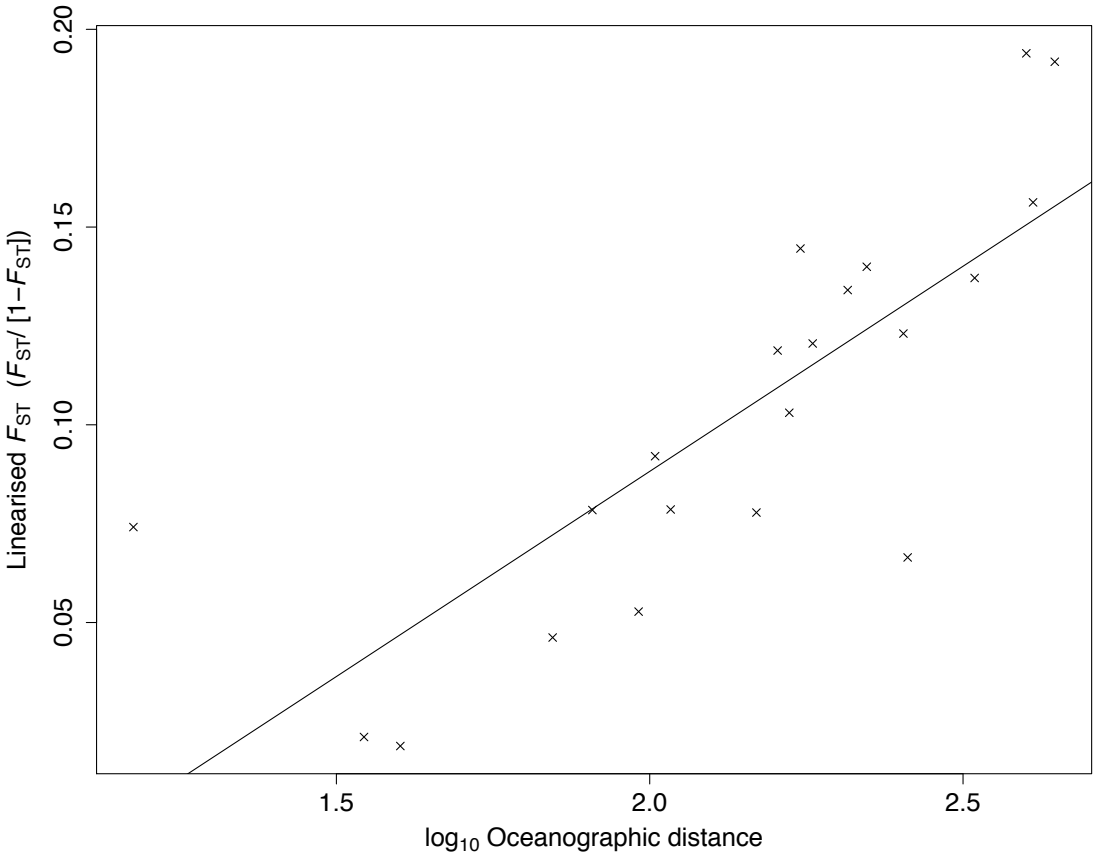
	GR	LK	BK	KC	WK	BC	LP	TA	SC	PAN1	PAN2	PAN3	MAY	BEQ	STV	STL	MAR1	MAR2	GU
GR	-	0.202	0.085	0.201	0.332	0.197	0.236	0.432	0.590	0.427	0.377	0.276	0.507	0.437	0.470	0.390	0.277	0.304	0.303
LK	0.123	-	0.056	0.036	0.066	0.009	0.032	0.322	0.468	0.237	0.185	0.186	0.324	0.206	0.258	0.179	0.124	0.105	0.189
BK	0.041^{X}	0.025^{X}	-	0.114	0.167	0.065	0.110	0.375	0.499	0.329	0.262	0.229	0.390	0.290	0.323	0.288	0.225	0.170	0.227
KC	0.149	0.021^{X}	0.069	-	0.012	0.075	0.080	0.378	0.519	0.328	0.314	0.276	0.403	0.281	0.295	0.266	0.231	0.150	0.261
WK	0.262	0.056	0.140	0.041^{X}	-	0.100	0.112	0.376	0.540	0.343	0.349	0.309	0.385	0.250	0.274	0.330	0.295	0.209	0.307
BC	0.124	0.008^{X}	0.032^{X}	0.038	0.052	-	0.030	0.282	0.444	0.254	0.223	0.180	0.316	0.212	0.239	0.243	0.148	0.124	0.166
LP	0.139	0.025^{X}	0.056	0.054	0.090	0.021^{X}	-	0.369	0.492	0.298	0.272	0.251	0.397	0.269	0.299	0.250	0.181	0.159	0.243
TA	0.258	0.201	0.217	0.204	0.263	0.186	0.217	-	0.456	0.294	0.341	0.275	0.337	0.231	0.261	0.322	0.273	0.287	0.247
SC	0.239	0.183	0.194	0.204	0.243	0.181	0.195	0.188	-	0.402	0.424	0.414	0.555	0.496	0.552	0.495	0.405	0.484	0.398
PAN1	0.265	0.177	0.210	0.198	0.261	0.179	0.197	0.171	0.155	-	0.091	0.085	0.465	0.332	0.360	0.256	0.149	0.212	0.250
PAN2	0.205	0.128	0.154	0.162	0.224	0.135	0.157	0.184	0.154	0.069	-	0.064	0.399	0.293	0.356	0.218	0.155	0.182	0.208
PAN3	0.199	0.156	0.170	0.174	0.249	0.148	0.178	0.157	0.152	0.043	0.058^{X}	-	0.317	0.225	0.257	0.245	0.195	0.145	0.193
MAY	0.219	0.132	0.156	0.171	0.194	0.138	0.165	0.170	0.159	0.193	0.144	0.146	-	0.060	0.104	0.212	0.382	0.296	0.381
BEQ	0.221	0.107	0.145	0.144	0.159	0.115	0.139	0.149	0.166	0.178	0.141	0.136	0.018	-	0.032	0.192	0.269	0.194	0.248
STV	0.240	0.139	0.165	0.154	0.180	0.134	0.158	0.149	0.186	0.176	0.151	0.131	0.044	0.021	-	0.234	0.311	0.219	0.315
STL	0.165	0.097	0.126	0.120	0.175	0.115	0.120	0.143	0.144	0.118	0.097	0.105	0.072	0.073	0.084	-	0.117	0.160	0.251
MAR1	0.149	0.085	0.116	0.120	0.185	0.094	0.108	0.149	0.138	0.104	0.092	0.112	0.123	0.108	0.126	0.050	-	0.125	0.062
MAR2	0.177	0.082	0.113	0.098	0.161	0.090	0.108	0.171	0.173	0.144	0.115	0.115	0.118	0.093	0.106	0.073	0.177	-	0.168
GU	0.190	0.131	0.148	0.145	0.207	0.118	0.149	0.166	0.167	0.169	0.137	0.146	0.161	0.135	0.162	0.121	0.062	0.110	-

^{X}: Non-significant pairwise comparison following correction for multiple comparisons.

Table 8 – Pairwise F_{ST} with ENA correction for null alleles between *Ircinia campana* sites (carried out on 10 loci).

	GR	LK	BK	KC	WK	BC	LP	TA	SC	PAN1	PAN2	PAN3	MAY	BEQ	STV	STL	MAR1	MAR2	GU
GR	-																		
LK	0.111	-																	
BK	0.053	0.023	-																
KC	0.135	0.017	0.055	-															
WK	0.214	0.049	0.112	0.030	-														
BC	0.104	0.006	0.028	0.033	0.055	-													
LP	0.119	0.020	0.045	0.039	0.072	0.016	-												
TA	0.226	0.166	0.171	0.172	0.219	0.158	0.176	-											
SC	0.222	0.167	0.167	0.182	0.215	0.166	0.173	0.154	-										
PAN1	0.233	0.160	0.176	0.175	0.225	0.165	0.168	0.139	0.142	-									
PAN2	0.165	0.107	0.118	0.135	0.186	0.113	0.127	0.150	0.138	0.058	-								
PAN3	0.190	0.161	0.167	0.173	0.235	0.158	0.168	0.139	0.150	0.046	0.054	-							
MAY	0.204	0.132	0.143	0.162	0.183	0.136	0.154	0.152	0.146	0.186	0.135	0.161	-						
BEQ	0.198	0.096	0.119	0.122	0.140	0.103	0.118	0.121	0.153	0.163	0.125	0.143	0.024	-					
STV	0.218	0.128	0.143	0.138	0.162	0.127	0.139	0.122	0.172	0.163	0.141	0.139	0.048	0.016	-				
STL	0.157	0.090	0.109	0.110	0.160	0.107	0.110	0.122	0.126	0.121	0.091	0.120	0.060	0.067	0.079	-			
MAR1	0.127	0.065	0.086	0.095	0.152	0.071	0.085	0.113	0.128	0.090	0.074	0.112	0.119	0.094	0.113	0.051	-		
MAR2	0.151	0.068	0.089	0.079	0.135	0.075	0.090	0.145	0.169	0.128	0.099	0.116	0.120	0.089	0.102	0.075	0.060	-	
GU	0.167	0.110	0.124	0.122	0.180	0.099	0.124	0.140	0.149	0.160	0.119	0.151	0.153	0.123	0.151	0.107	0.057	0.098	-


Linearised F_{ST} (F_{ST}/ [1−F_{ST}])
0.00
0.05
0.10
0.15
0.20
0.25
0.30
0.35
0.5
1.0
1.5
2.0
2.5
3.0
3.5
$\log_{10}$ Oceanographic distance


Linearised F_{ST} (F_{ST}/ [1−F_{ST}])
0.05
0.10
0.15
0.20
1.5
2.0
2.5
$\log_{10}$ Oceanographic distance

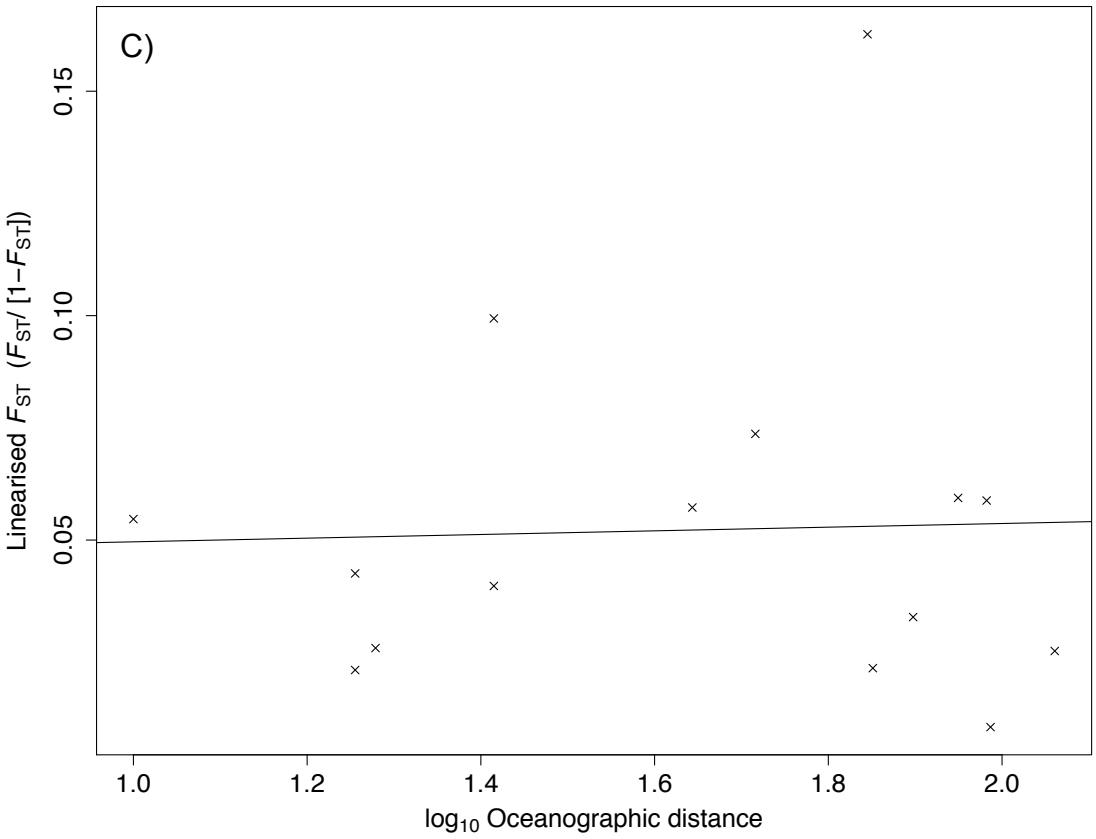


Figure 3: Genetic isolation by distance for *Ircinia campana;* A) over all sites, B) in the Lesser Antilles, and C) in Florida, USA.

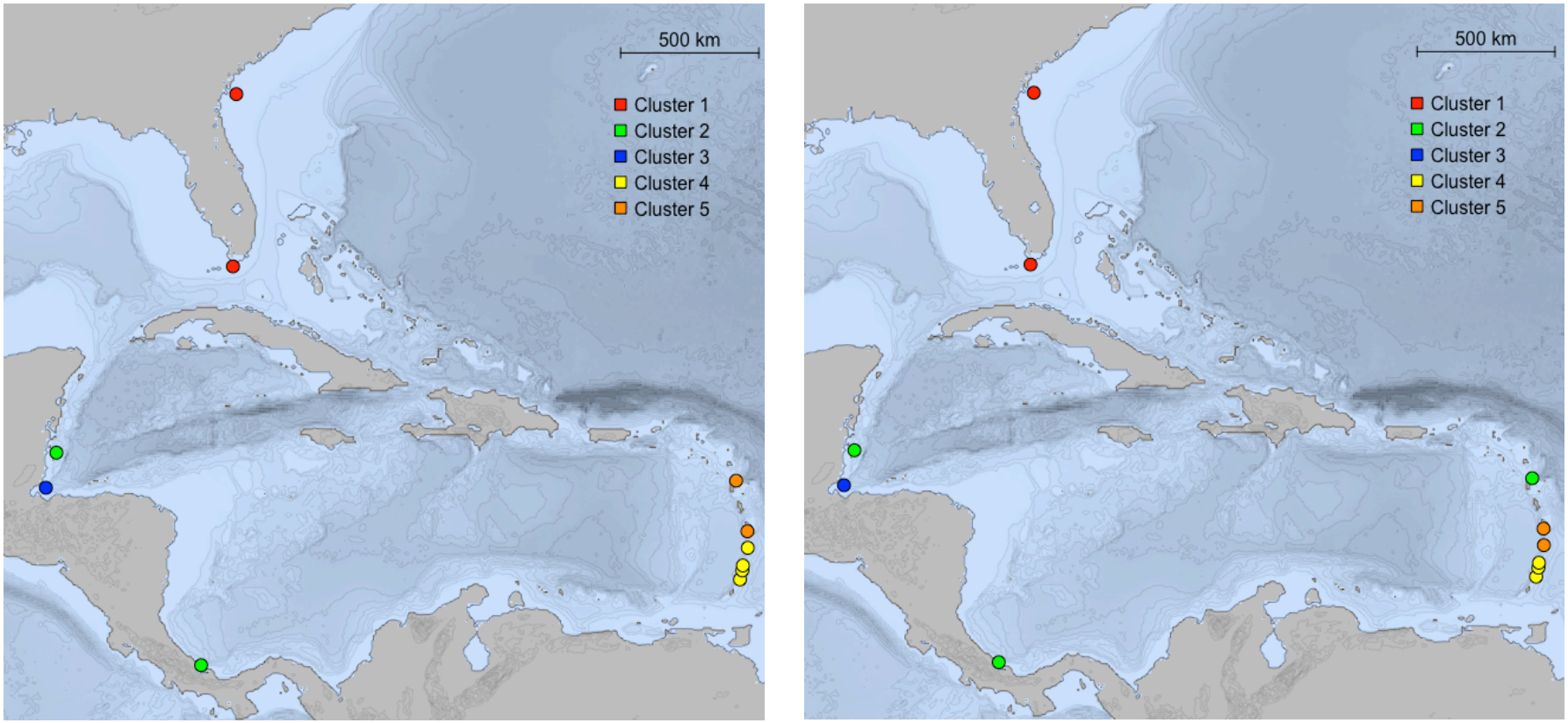
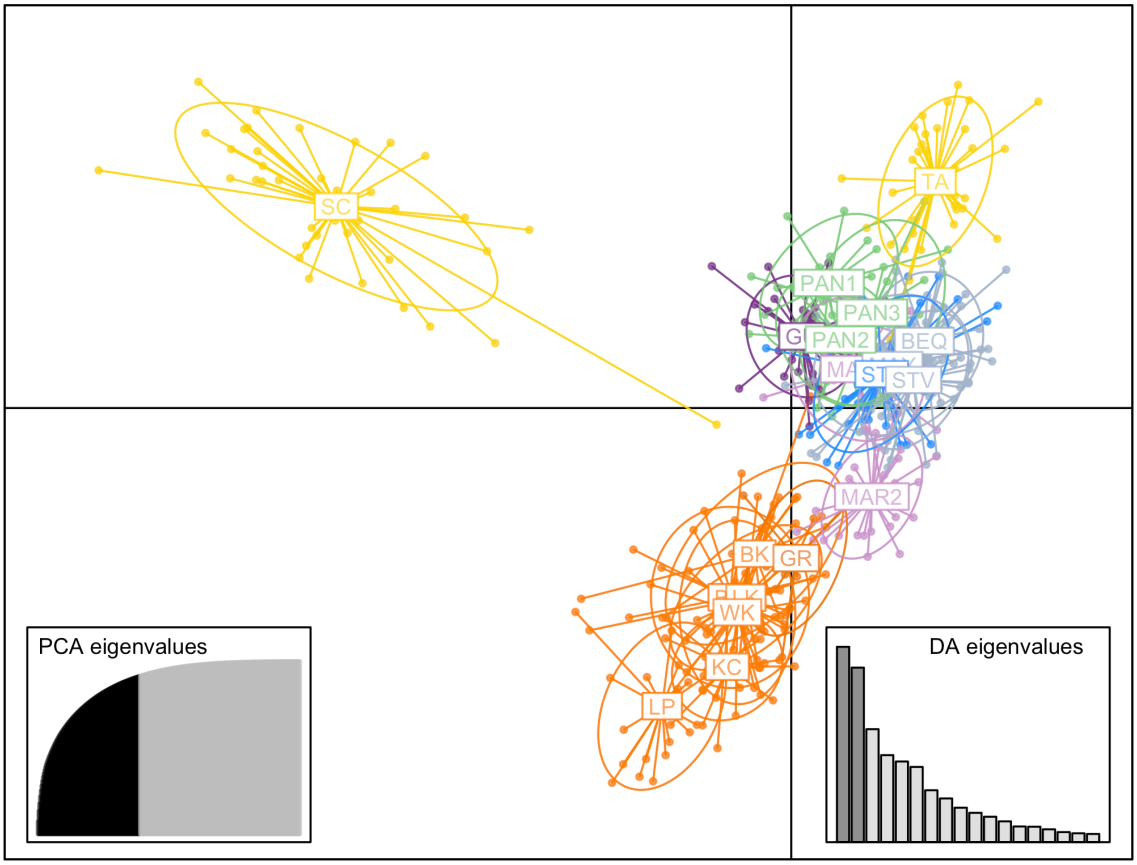
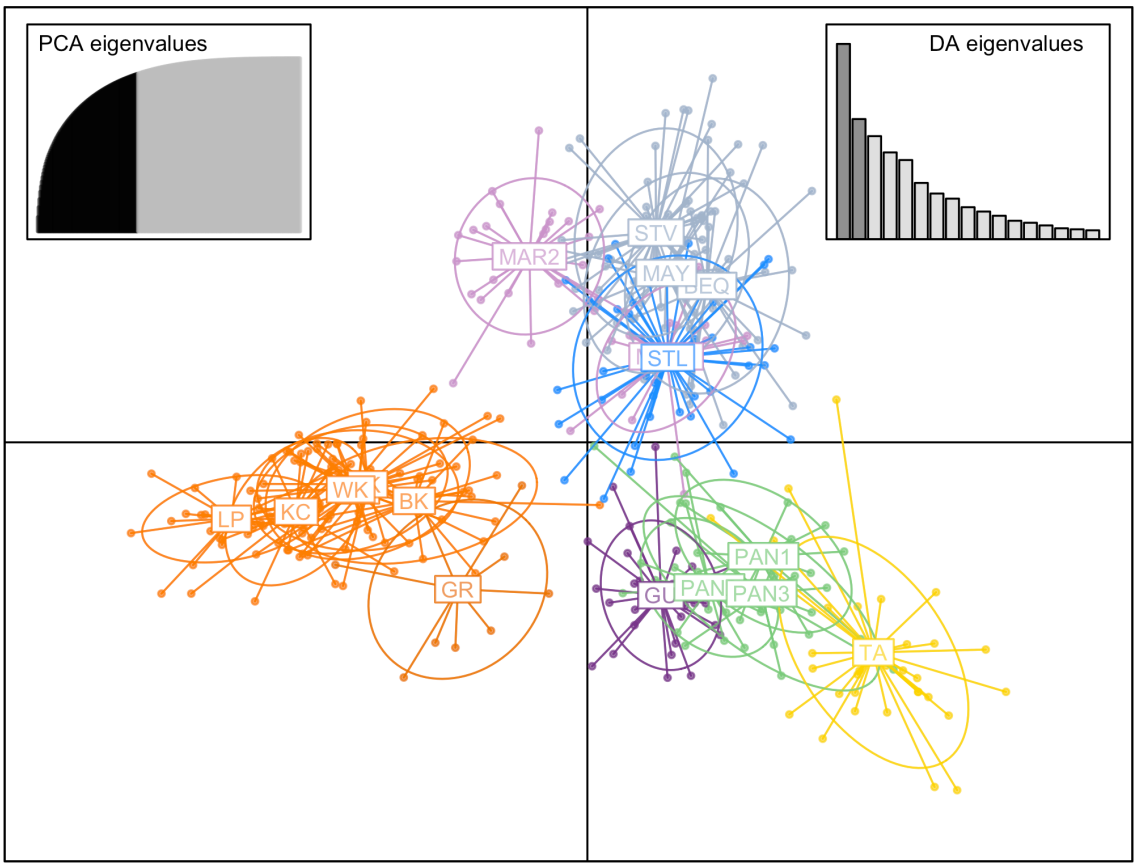


Figure 4: Clustering solutions reached by Bayesian spatially-explicit genetic clustering program Geneland for *K* = 5. Left hand pane shows the solution reached with the highest log posterior probabilities (i.e. a better fit to the data); right hand pane shows the alternative clustering solution reached with lower log posterior probabilities. Multiple sites within Panama (PAN1, PAN2 and PAN3) and Florida (LK, BK, KC, WK, BC and LP) represented by a single circle.

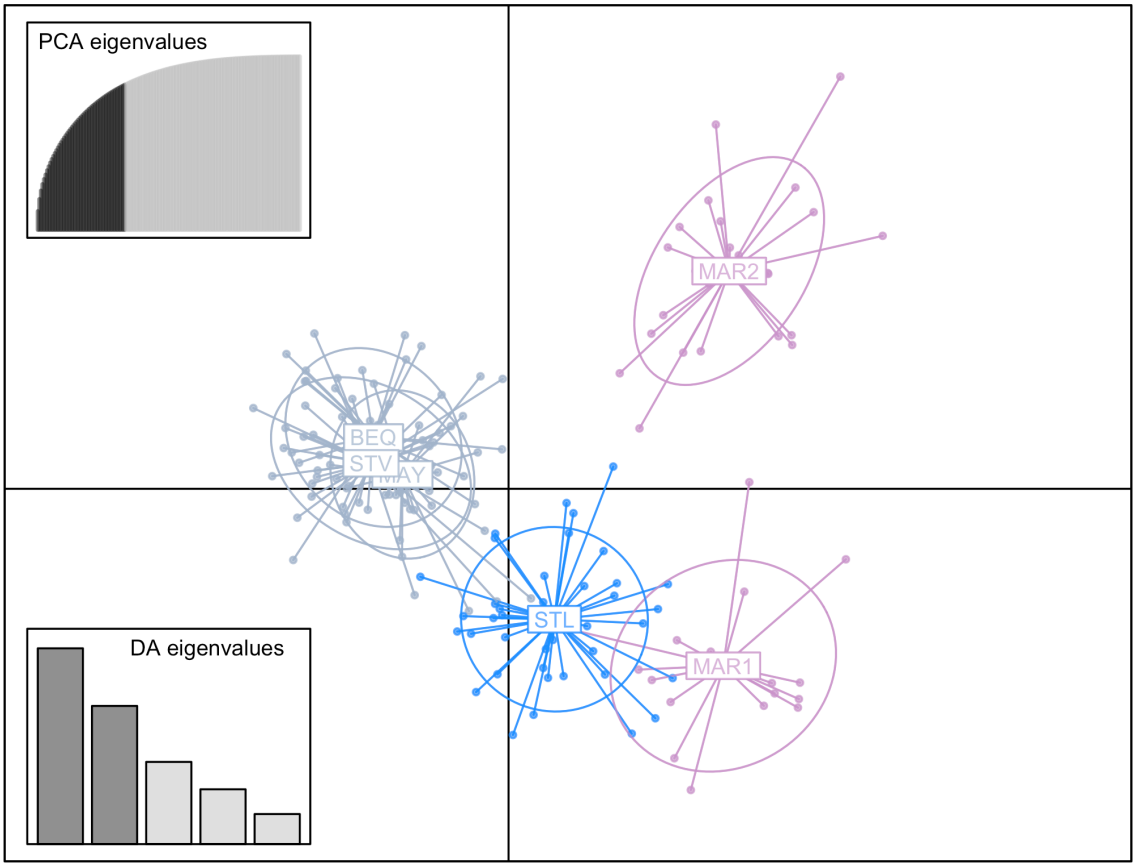
A)


SC
TA
PAN1
PAN3
G
PAN2
BEQ
STV
MAR2
BK
GR
WK
KC
LP
PCA eigenvalues
DA eigenvalues

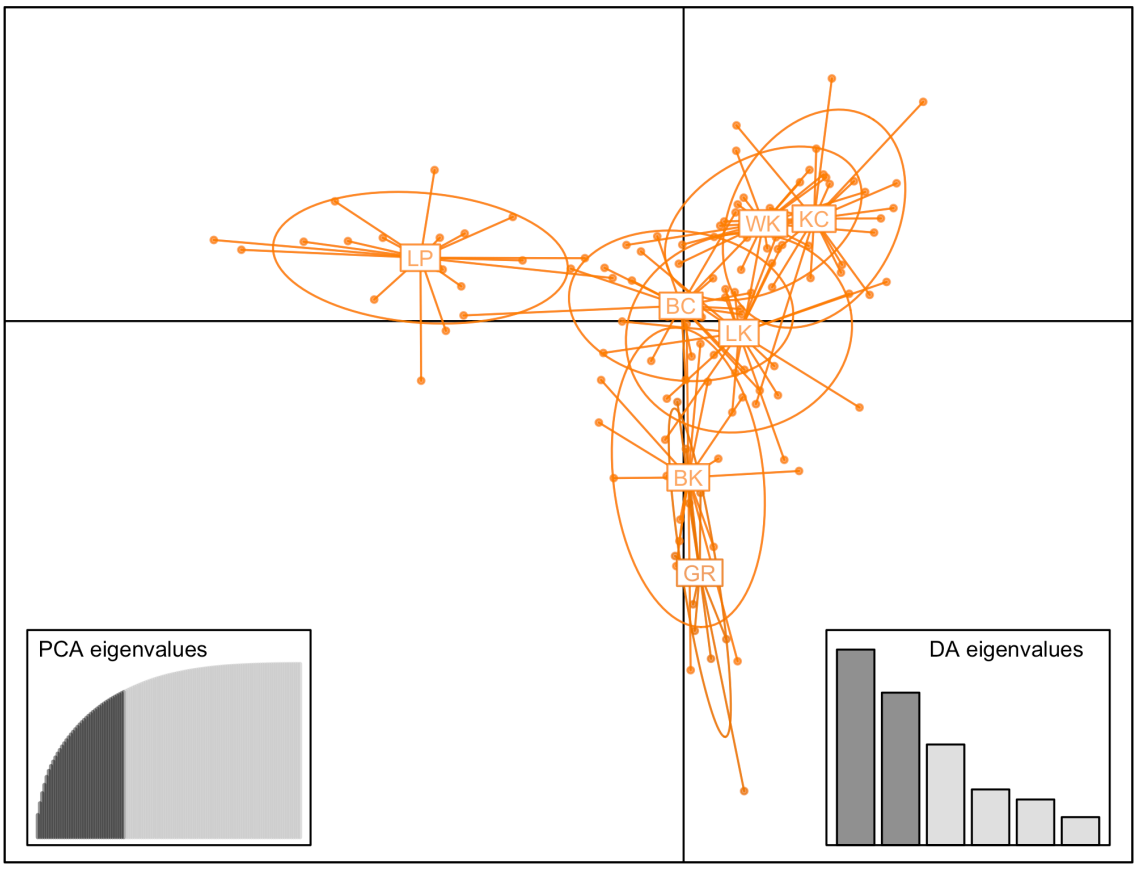
B)


PCA eigenvalues
DA eigenvalues
MAR2
STV
MAY
EQ
STL
WK
BK
LP
KC
GR
PAN1
PAN3
TA

C)


PCA eigenvalues
MAR2
BEQ
STV
STL
MAR1
DA eigenvalues

D)


LP
WK
KC
BC
LK
BK
GR
PCA eigenvalues
DA eigenvalues

E)

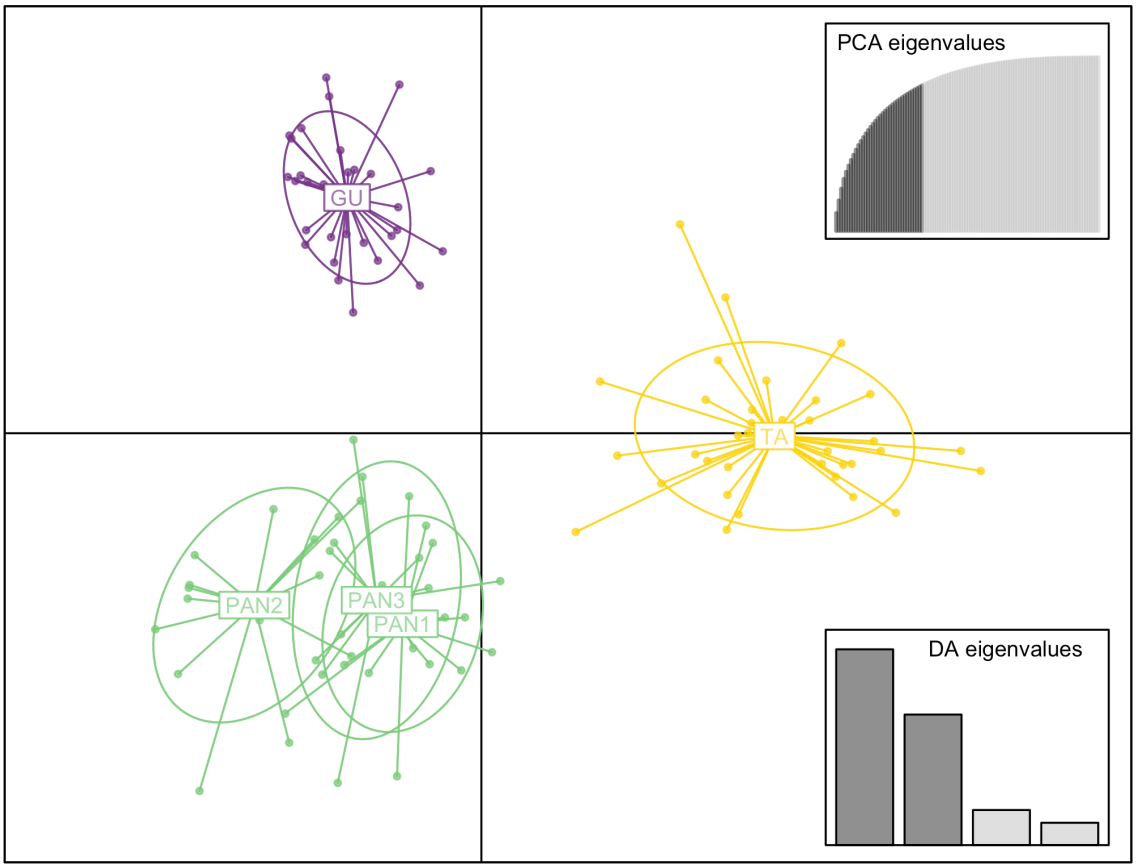


Figure 5: Discriminant analysis of principle components (DAPC) conducted on A) All sites; B) All sites but Sapodilla Cayes; C) St. Vincent and the Grenadines, St. Lucia and Martinique cluster; D) USA cluster; E) Panama, Turneffe Atoll and Guadeloupe cluster. Individuals are represented by dots, and sites are represented by inertia ellipses. The proportion of variance captured by the principle components and the discriminant analysis eigenvalues are displayed graphically (inset). Countries are grouped by colour: USA: orange; Belize: yellow; Panama: green; St. Vincent and the Grenadines: pale blue; St. Lucia: royal blue; Martinique: lilac; Guadeloupe; deep purple.

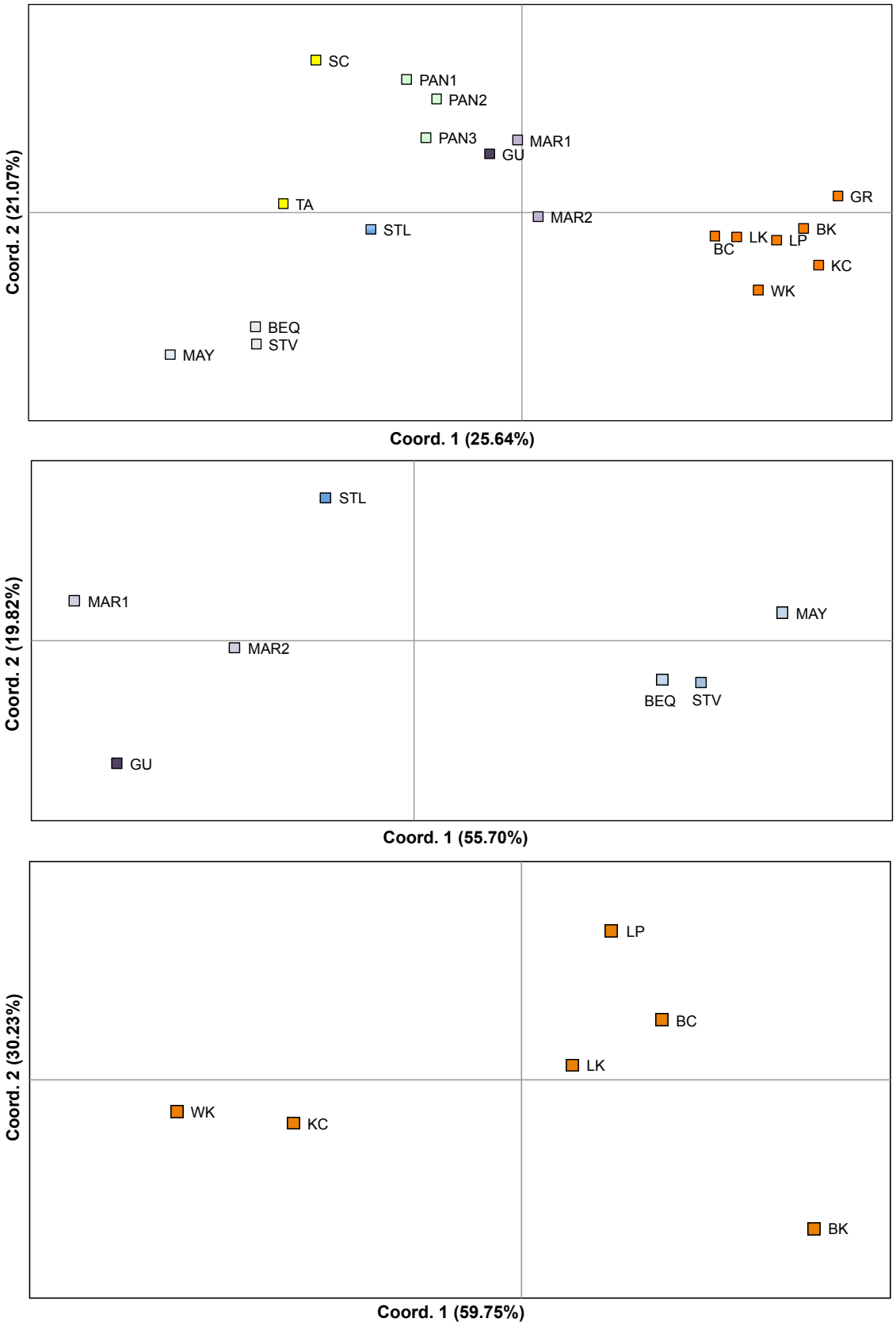


Figure 6: Principle coordinates analysis (PCoA) using site pairwise *D* calculations. Top pane: all sites; middle pane: Lesser Antilles sites; bottom pane: Florida sites. Axis labels show the percentage variance explained by each principle coordinate. Countries are grouped by colour: USA: orange; Belize: yellow; Panama: green; St. Vincent and the Grenadines: pale blue; St. Lucia: royal blue; Martinique: lilac; Guadeloupe; deep purple.

DISCUSSION

Marker evaluation and summary statistics

In this study, I characterised ten microsatellite markers and two PCR multiplexes for *Ircinia campana* for the purpose of studying population structure and genetic diversity over the Greater Caribbean. These markers have sufficient variability and sensitivity to describe population structure compared to mitochondrial *COI* markers previously used in this species (Marino et al., 2017), and consequently can be used in the future for other intraspecific molecular ecology research in *I. campana*. However, high frequencies of null alleles were found at many of the loci, and deviations from Hardy-Weinberg Equilibrium (HWE) were found widely across loci and sites. Both of these are common in sponge studies (Dailianis et al., 2011; Chaves-Fonnegra et al., 2015; Pérez-Portela, Noyer & Becerro, 2015; Guardiola, Frotscher & Uriz, 2016; Richards et al., 2016), potentially reflecting shared life history characters. However, these characteristics are not universally found in the phylum (Bell et al., 2014); interestingly a study with congener *I. fasciculata* showed low levels of null alleles (Riesgo et al., 2016).

Null alleles are those that fail to amplify in PCR due to the presence of mutations in the flanking regions of the microsatellite where PCR primers bind. In diploid species, both or one of the alleles at a locus fail to amplify, causing either missing data or false homozygotes. This is problematic for population genetics studies - such deficiencies in heterozygotes cause departures from HWE and inflate F_{ST} and F_{IS} values, thus obscuring important population genetic inferences. Model comparison using DIC values in the INEST analysis of inbreeding coefficients indicated that models with null alleles as a component best fitted the data, showing that null alleles did skew F_{IS} estimates in this study. Therefore, we can also conclude that loci that departed from HWE did so at least partly because of null alleles. Loci also showed high polymorphism in many cases, which may have led to deviations from HWE. This is because significance testing for HWE is very sensitive to individuals who are

homozygous for rare alleles (Morin et al., 2009), and rare alleles are more likely when there are large numbers of alleles per locus.

Although the technical issues discussed above may have led to departures from HWE, there may have been additional biological causes to this phenomenon. F_{IS} estimates were still positive when corrected for null alleles (albeit not significantly so), indicating that inbreeding could have contributed to HWE departures. Inbreeding in *I. campana* could be the result of low dispersal in planktonic sperm and low dispersal of larvae, leading to philopatry and higher incidences of non-random mating. Asexual reproduction was apparently low, with only one pair of identical mutlilocus genotypes found throughout the dataset (this was also true of congener *I. fasciculata*, which showed no identical genotypes in a dataset of 194 individuals; Riesgo et al., 2016). Asexual reproduction serves to maintain heterozygosity in populations (Stoeckel et al., 2014), therefore its low prevalence may contribute to the heterozygosity deficiencies observed in the populations. Wahlund effects - genetic structure within a set of samples - can also cause departure from HWE. This can include temporal structure within a single site (Tesson et al., 2014; Truelove et al., 2015). In the marine environment, the random survival of larval cohorts in temporally variable ocean current patterns can lead to patterns of gene flow varying through time (Hedgecock & Pudovkin, 2011). In sponges, it is difficult to age individuals accurately and so temporal genetic population structure is hard to test, however, long term monitoring of sites and sampling of new recruits could offer a solution.

High F_{IS} values and resultant HWE departures are a common trait among marine invertebrates with free-spawned planktonic sperm (Addison & Hart, 2005), a category in which *I. campana* belongs. Addison and Hart (2005) proposed that this could link to higher levels of null alleles in this group, due to higher numbers of cell cycles for sperm production causing increased mutation rates. This could also help to explain why the loci in this study were very polymorphic, if the mechanisms governing mutation rate in the microsatellite gene are the same as those in the binding regions. Interestingly, Romiguier et

al. (2014) found life history was the key explainer of genetic diversity across the Metazoa, with broadcast spawners with small larvae (r-strategists) consistently the most genetically diverse animals across the genome, compared to animals that invest resources in few slow-growing offspring (K-strategists). However they attribute this to high effective population sizes, rather than high mutation rate, as theory predicts that in fact K-strategists would have increased mutation rates. Bivalves are well known for their high null allele frequencies (Foltz, 1986; Hedgecock et al., 2004), and also are the most genetically diverse group in the Romiguier et al. (2014) analysis. In this study, such links are purely speculative, however, it would be interesting for future research to delve further into the links between life history, reproductive strategy, genetic diversity, null alleles, high F_{IS} and departures from HWE. In particular for this study, it would be valuable to know the amount of sperm produced by *I. campana*, as this can vary among broadcast spawners and may not necessarily be high.

Regardless of the causes for HWE departures, there can be some analytical constraints as a result of them. Many standard population genetics methods (for example, STRUCTURE [Pritchard, Stephens & Donnelly, 2000]) assume loci and populations are in HWE, restricting the analyses or loci that can be used by researchers. Furthermore, when loci depart from HWE, estimates of genetic distance such as F_{ST} no longer show differences between populations due to gene flow or connectivity, as differences may be due to selection, non-random mating or any of the other assumptions of HWE; rather they show absolute differentiation. This is a weakness of the study that must be considered when interpreting the genetic distance measures and analyses that use them (significance testing for population differentiation, principle coordinates analysis). However, multivariate methods such as discriminant analysis of principle components (DAPC) that do not hold such assumptions can be used as an alternative (Jombart, 2008).

As mentioned above, the numbers of alleles per locus observed were unusually high in a number of loci (up to 97 alleles in Icam23). Richards et al., (2016) also found a high allelic richness in the microsatellite loci of another sponge species,

Xestospongia muta, with up to 70 alleles per locus, suggesting that this could be more common amongst sponges than other taxa. Polymorphism is clearly an asset of microsatellites for their use in population genetic studies, as it allows population to be differentiated on ecological timescales. However, very high levels of polymorphism can be problematic concerning the use of F_{ST} and similar subpopulation differentiation measures, as they are strongly influenced by within-population diversity. High levels of polymorphism in multi-allelic markers such as microsatellites can therefore limit the maximum levels of differentiation that F_{ST} can show, and large differentiation between subpopulations can be masked by low F_{ST} values (Charlesworth, 1998; Meirmans & Hedrick, 2011). Because of this, I also used *D*, which is based on effective number of alleles rather than heterozygosity, and unlike heterozygosity, scales linearly with increased diversity, and can range up to 1 in cases of total differentiation (Jost, 2008).

Population structure and connectivity

The results of this study show the presence of strong population structure in *Ircinia campana* in the Greater Caribbean region. Isolation by distance (IBD) was significant on the regional scale over all sampling sites (3 – 2975 km), and within the Lesser Antilles sites on a medium spatial scale (15 – 443 km) but not on the small spatial scale within the Florida sites (10 – 115 km). In the Lesser Antilles (medium scale), the effects were strongest (r = 0.69 – 0.79, compared to 0.44 – 0.60 over all sites; range accounts for differences using *D* and F_{ST} in analyses). In the marine environment, there are problems with using crude oceanographic distance measures, as they do not take into account the directions of the currents that act as dispersal pathways among locations. However, without complex modelling and long-term current data, such measures cannot be accurately taken, and therefore oceanographic distance is the best measure that can be used to get an indication of the effects of distance in shaping population genetic structure in this study. The results here do show a distance-decay relationship at pairwise distances over 115 km, indicating dispersal ability is limited in the species at these distances.

The short larval durations found in sponges are likely to drive the IBD relationship, and indeed IBD has been found in many other studies of sponges on larger spatial scales (Duran et al., 2004; Blanquer & Uriz, 2010; Guardiola, Frotscher & Uriz, 2012; Chaves-Fonnegra et al., 2015; Riesgo et al., 2016). Sponge larvae are lecithotrophic, meaning that they do not feed while in the plankton, and so rely on their existing energy supplies until settlement and metamorphosis. This constrains their ability to disperse, as they have a limited time before energy supplies are depleted. The larvae and reproduction of *I. campana* itself has not yet been studied, but we may be able to make inferences from information available for members of the genus and family that have been studied. Larvae in the Dictyoceratida, including the tufted parenchmyella of the Irciniidae, are better swimmers than other sponge larvae (Ereskovsky & Tokina, 2004; Mariani et al., 2006), which could further their dispersal ability. In addition, another member of the Irciniidae, *Ircinia oros*, was found to have relatively large lipid stores (Ereskovsky & Tokina, 2004), which could enable longer dispersal distances. Nevertheless, despite these features, sponge larvae are still very short lived. Furthermore, sponges in the Irciniidae family are viviparous, brooding embryos until their release as larvae. This reduces their potential dispersal capability compared to a non-brooding species in which dispersal also occurs for the ova and developing embryo. However, it is unknown whether *I. campana* exhibits other forms of reproduction, such as asexually by fragmentation (e.g. Wulff, 1991). In addition, even small fragments of sponge broken away by physical or biological damage can contain embryos, which can be retained and nourished to full development even in small fragments (Maldonado & Uriz, 1999). Another dispersal opportunity could be thorough detachment of whole sponges from the substrate (which I have observed while collecting samples), which could then be potentially transported in ocean currents.

In this study, it was not possible to use the popular Bayesian clustering and individual assignment program STRUCTURE to infer population structure patterns. Despite testing long Monte Carlo Markov Chain runs (up to 1 million iterations), convergence was not reached, with repeat runs at the same value of

K producing different clustering solutions. The inability to reach robust and replicable solutions was also encountered when TESS3 v1.1.0 (Caye et al., 2016) and FLOCK v3.1 (Duchesne & Turgeon, 2012) were used, programs which have similar objectives to STRUCTURE but slightly different methodologies. The spatially explicit Bayesian approach of Geneland performed best, but still did not reach the same solution in all replicate runs. Isolation by distance models of population structure do not fit well with the model assumptions of STRUCTURE and its family of related programs, as there are not discreet population units in which to partition samples. Furthermore, assigning individuals of unknown ancestry to an unknown number of genetic clusters is computationally and statistically more complicated than testing differences between pre-defined 'populations', such as in analyses based on F_{ST} measurements. The loss of power caused by the use of only loci that do not depart from HWE is likely to cause further difficulties. Perhaps due to its inclusion of spatial coordinates, Geneland was able to give some more consistent clustering results. However, the fact that two solutions were reached, in addition to the failure of the other programs to converge, indicates that there is an absence of discreet population clusters in the dataset. This strongly supports the isolation by distance 'stepping stone' model as the most likely population structure for *I. campana*, with admixture occurring among proximate locations. However, despite the lack of defined population units, there were some clear population structure patterns in the dataset where other influences appear to have been present along with distance-limited dispersal effects.

The Sapodilla Cayes (SC), in the south of Belize's Mesoamerican Barrier Reef System (MBRS), consistently showed great genetic differentiation from other sites. It formed its own cluster in the DAPC and Geneland analyses, and high F_{ST} and *D* values were found when comparing the site with all others, even Turneffe Atoll (TA), which was the most geographically proximate site sampled, in the northern MBRS. It also contained higher amounts of private alleles than many sites, including Turneffe Atoll. Two main circulation regimes are present in the MBRS: in the north, the northward flowing Yucatan Current, and in the south, weak southward flowing coastal currents and the anti-clockwise

Honduras Gyre (Ezer et al., 2005; Carrillo et al., 2015). A third, temporally variable regime is the Cayman current, which forms a barrier between the north and south MBRS. In the south MBRS, these features are suspected to promote a retentive environment for larvae, restricting dispersal and promoting high levels of self-recruitment. In addition, the temporally variable river discharge into the Gulf of Honduras (Soto et al., 2009) may act as a barrier to connectivity. There is evidence of distinction between the north and south MBRS in larval assemblages (Muhling et al., 2013), in genetic and modelling data for coral *Montastrea annularis* (Foster et al., 2012), and genetic evidence that there are more *Panulirus argus* (spiny lobster) migrants in the north and more self-recruitment in the south (Truelove et al., 2014). Specifically for the Sapodilla Cayes, additional retention could be caused by local water movement patterns caused by the distinctive hooked 'J' shape of the barrier reef in that area. Although the outer morphology of the sponges did not show any visible differences in this area, it may be the case that the isolation this area experiences could cause speciation, considering the great amount of differentiation shown. Future studies using other genes such as mitochondrial COI, may be useful in determining if cryptic speciation has occurred in this area in *Ircinia campana* (e.g. Pöppe et al., 2010), as well as studies comparing skeletal fibres and architecture.

Connectivity between the MBRS and Florida was apparently poor. Drifter tracks have shown there to be high potential for connectivity between the north MBRS and the Florida Keys, with larvae potentially being transported in 7-10 days (Muhling et al., 2013), but sponge larvae generally have PLDs lower than this. In addition, as Irciniidae are viviparous, dispersal from the maternal site does not begin until the larva is developed, unlike oviparous sponges, whose eggs also disperse.

High differentiation was found between Panama (PAN1, PAN2 and PAN3) and the southern Lesser Antilles islands (MAY, BEQ, STV). One may expect connectivity due to the westward Caribbean current flowing through the Lesser Antilles towards the north coast of Panama. However, there are a number of

potential barriers to dispersal. There is evidence of a counter current running along the Venuzuelan coast from the Panama-Colombia gyre (Andrade, Barton & Mooers, 2003) which could serve to prevent connection between the sites, as a front could form between the westward and eastward flowing water. Another dispersal barrier could be the plume of the Río Magdalena, which discharges into the Caribbean Sea at Cartagena, Colombia. This is the largest river discharging into the Caribbean Sea, annually discharging a mean of 228 km^3 of water (Restrepo et al., 2006). Foster et al. (2012) also found a similar genetic pattern in the coral *Montastraea annularis* and suggested that this could be a barrier. Instead, Panama was found to show a higher degree of connectivity to Turneffe Atoll in Belize (TA) and to Guadeloupe (GU) (PCoA, Geneland and DAPC analyses). The connection to both areas is surprising. The anti-clockwise Colombia-Panama gyre may be expected to promote larval retention, and additionally, one may expect the archipelago of Bocas Del Toro itself to promote retention due to the complexity of the islands. However, this was apparently not the case, with Geneland (clustering solution 2), the DAPC and PCoA all supporting a degree of genetic similarity between Panama, Guadeloupe and Turneffe. This could be the result of few long-distance dispersal events of sponge fragments or larvae, however, intermediate populations could act as stepping stones in a connection pathway.

No study has yet described population structure in sponges within the Lesser Antilles specifically. We found strong IBD patterns, which as discussed above, is likely to be caused in a large part by limitations in the dispersal ability of sponge larvae. However, water flows into the Caribbean Sea from the east through gaps between islands, forming channels that could also act to weaken dispersal. The strongest flow found in these channels are through the Grenada, St. Vincent and St. Lucia passages (Johns et al., 2002); the absence of strong channels separating St. Vincent and the Grenadines coupled with the potential dispersal-weakening St. Vincent passage separating St. Lucia and St. Vincent may explain the clustering of St. Vincent and the Grenadines observed in the PCoA. Another interesting feature of the Lesser Antilles data was that we found strong differentiation between the two Martinique sites. As there are no

apparent physical barriers to connectivity between these locations, the cause of this is unknown; however, there is limited knowledge on local current patterns around Martinique.

Within Florida, IBD effects were not present. Similarly, studies on other sponge species have found absences of IBD in Florida (DeBiasse, Richards & Shivji, 2010; Chaves-Fonnegra et al., 2015; Richards et al., 2016). Like *I. campana*, these species do show population structure, indicating that connectivity patterns are driven by other factors than distance-limited dispersal. In this study, the DAPC carried out on the USA sites showed the Lakes Passage (LP) separated from the cluster of other sites. A deep channel to the west of Key West with currents running through may limit dispersal between this site and the other sites in Florida (M. Butler pers. comm.). Other patterns in the data were not clearly related to oceanographic patterns. Long Key (LK) shows low (and non-significant) differentiation from all Florida sites except Waltz Key (WK), whilst Waltz Key shows higher (and statistically significant) differentiation from all sites but Bamboo Key. This genetic patchiness may be the result of more stochastic processes, such as rare long-distance dispersal events, or sweepstakes reproductive success. Grays's Reef (GR) appeared to be well connected to Florida, which is likely to be driven by the strong northerly-flowing Florida Current aiding dispersal.

Genetic diversity and genetic bottlenecks

Evidence for sites having experienced genetic bottlenecks were found in fourteen of the nineteen sites. However, thirteen of the sites only exhibited M-ratio deficiencies, while one site (Bamboo Key) showed both heterozygosity excess and M-ratio deficiency. This may be due to M-ratio tests showing higher sensitivity, but alternatively could be due to a higher propensity for false positives. Riesgo et al., (2016) used microsatellites to detect heterozygosity excesses in *Ircinia fasciculata*, and found that although many sites exhibited signatures of bottlenecks, a site where mass mortalities have been recorded did not have any significant heterozygosity excesses. However, Hoban et al., (2013) found that in species with high variability in individual reproductive

success, false positives can be produced in both types of bottleneck test. As spawning marine invertebrates, sponges are likely to fall into this category, with 'sweepstake' reproductive success caused by random pairings of larval cohorts with favourable oceanographic conditions (Hedgecock, 1982; Hedgecock & Pudovkin, 2011). In addition, both methods are affected by the mutation model selected, and false positives can occur with incorrect selection of the mutation model for the loci used, particularly the proportion of multi-step mutations (Peery et al., 2012). Mass mortality and disease-like conditions are common in *Ircinia* species (Maldonado, Sánchez-Tocino & Navarro, 2010; Cebrian et al., 2011; Stabili et al., 2012; Rivetti et al., 2014). Indeed, it is known that *I. campana* suffered mass mortalities in Florida Keys (Butler et al., 1995; Stevely et al., 2010), and a disease-like condition was observed in a few individuals with necrotic tissue in Bequia (personal observation). Therefore it is may be possible that the signs of bottlenecks observed here are indeed true signatures.

Allelic richness tests were largely uninformative in comparing differences among sites due to low sample sizes and high variance around the mean. Sites in Florida, Panama and Belize did generally show lower levels of allelic richness than sites in the Lesser Antilles, however large standard errors preclude definitive comparisons and statements regarding the effects of mortalities or potential bottlenecks on genetic diversity. As the preservation of genetic diversity is important for conservation, accurate identification of bottlenecks and monitoring genetic diversity are important in species management. As such, populations of *Ircinia campana* should continue to be monitored in the future using genetic techniques. This would enable the comparison of genetic diversity at different points in time, enabling less uncertain assessments of genetic health of the population than those that can be gained from sampling at a single point in time.

Conclusions

In this chapter, I demonstrated that the sponge *Ircinia campana* shows strong population structure through the Greater Caribbean region. At larger spatial scales, this is affected by dispersal limitation and ocean currents, with

oceanographic barriers identified that correlate with modelling and empirical studies. This finding is in common with other sponge studies, including congener *Ircinia fasciculata* in the Mediterranean (Riesgo et al., 2016). At the smaller spatial scale, population structure was also present but was not affected by distance. Many sites showed genetic bottleneck signatures, indicating that genetic diversity losses may have occurred in the recent past.

REFERENCES

Addison JA., Hart MW. 2005. Spawning, copulation and inbreeding coefficients in marine invertebrates. *Biology Letters* 1:450–453. DOI: 10.1098/rsbl.2005.0353.

Alberto F. 2009. MsatAllele_1.0: An R package to visualize the binning of microsatellite alleles. *Journal of Heredity* 100:394–7. DOI: 10.1093/jhered/esn110.

Alvera-Azcárate A., Barth A., Weisberg RH. 2009. The surface circulation of the Caribbean Sea and the Gulf of Mexico as inferred from satellite altimetry. *Journal of Physical Oceanography* 39:640–657. DOI: 10.1175/2008JPO3765.1.

Andrade CA., Barton ED., Mooers CNK. 2003. Evidence for an eastward flow along the Central and South American Caribbean Coast. *Journal of Geophysical Research* 108:3185. DOI: 10.1029/2002JC001549.

Archer SK., Stoner EW., Layman CA. 2015. A complex interaction between a sponge (*Halichondria melanadocia*) and a seagrass (*Thalassia testudinum*) in a subtropical coastal ecosystem. *Journal of Experimental Marine Biology and Ecology* 465:33–40. DOI: 10.1016/j.jembe.2015.01.003.

de Bakker DM., Meesters EHWG., van Bleijswijk JDL., Luttikhuizen PC., Breeuwer HJAJ., Becking LE. 2016. Population genetic structure, abundance, and health status of two dominant benthic species in the Saba Bank National Park, Caribbean Netherlands: *Montastraea cavernosa* and *Xestospongia muta*. *PLoS ONE* 11:e0155969. DOI: 10.1371/journal.pone.0155969.

Baums IB. 2008. A restoration genetics guide for coral reef conservation. *Molecular Ecology* 17:2796–2811. DOI: 10.1111/j.1365-294X.2008.03787.x.

Baums IB., Miller MW., Hellberg ME. 2005. Regionally isolated populations of an imperiled Caribbean coral, *Acropora palmata*. *Molecular Ecology* 14:1377–1390. DOI: 10.1111/j.1365-294X.2005.02489.x.

Bell JJ. 2008. The functional roles of marine sponges. *Estuarine, Coastal and Shelf Science* 79:341–353. DOI: 10.1016/j.ecss.2008.05.002.

Bell JJ., Smith D., Hannan D., Haris A., Jompa J., Thomas L. 2014. Resilience to disturbance despite limited dispersal and self-recruitment in tropical barrel sponges: implications for conservation and management. *PLoS ONE* 9:e91635. DOI: 10.1371/journal.pone.0091635.

Benjamini Y., Yekutieli D. 2001. The control of the false discovery rate in multiple testing under dependency. *The Annals of Statistics* 29:1165–1188. DOI: 10.2307/2674075.

Blacket M., Robin C., Good R., Lee S., Miller A. 2012. Universal primers for fluorescent labelling of PCR fragments- an efficient and cost-effective approach to genotyping by fluorescence. *Molecular Ecology Resources* 12:456–63. DOI: 10.1111/j.1755-0998.2011.03104.x.

Blanquer A., Uriz MJ. 2010. Population genetics at three spatial scales of a rare sponge living in fragmented habitats. *BMC Evolutionary Biology* 10:1–9. DOI: 10.1186/1471-2148-10-13.

Bohonak AJ. 1999. Dispersal, gene flow, and population structure. *The Quarterly Review of Biology* 74:21–45. DOI: 10.2307/2664376.

Bolger AM., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. DOI: 10.1093/bioinformatics/btu170.

Botsford LW., White JW., Coffroth M-A., Paris CB., Planes S., Shearer TL., Thorrold SR., Jones GP. 2009. Connectivity and resilience of coral reef metapopulations in marine protected areas: matching empirical efforts to predictive needs. *Coral Reefs* 28:327–337. DOI: 10.1007/s00338-009-0466-z.

Butler M., Hunt J., Herrnkind W., Childress M., Bertelsen R., Sharp W., Matthews T., Field J., Marshall H. 1995. Cascading disturbances in Florida Bay, USA: cyanobacteria blooms, sponge mortality, and implications for juvenile spiny lobsters *Panulirus argus*. *Marine Ecology Progress Series* 129:119–125. DOI: 10.3354/meps129119.

Butler MJ I., Paris C., Goldstein J., Matsuda H., Cowen R. 2011. Behavior constrains the dispersal of long-lived spiny lobster larvae. *Marine Ecology Progress Series* 422:223–237. DOI: 10.3354/meps08878.

Carrillo L., Johns EM., Smith RH., Lamkin JT., Largier JL. 2015. Pathways and hydrography in the Mesoamerican Barrier Reef System Part 1: Circulation. *Continental Shelf Research* 109:164–176. DOI: 10.1016/j.csr.2015.09.014.

Castoe T., Poole A., de Koning A., Jones K., Tomback D., Oyler-McCance S., Fike J., Lance S., Streicher J., Smith E., Pollock D. 2012. Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. *PLoS ONE* 7:e30953. DOI: 10.1371/journal.pone.0030953.

Castorani MCN., Reed DC., Raimondi PT., Alberto F., Bell TW., Cavanaugh KC., Siegel DA., Simons RD. 2017. Fluctuations in population fecundity drive variation in demographic connectivity and metapopulation dynamics. *Proceedings of the Royal Society B: Biological Sciences* 284: 20162086. DOI: 10.1098/rspb.2016.2086.

Caye K., Deist TM., Martins H., Michel O., François O. 2016. TESS3: fast inference of spatial population structure and genome scans for selection. *Molecular Ecology Resources* 16:540–548. DOI: 10.1111/1755-0998.12471.

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* 6:e20211. DOI: 10.1371/journal.pone.0020211.

Chapuis M-P., Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24:621–31. DOI: 10.1093/molbev/msl191.

Charlesworth B. 1998. Measures of divergence between populations and the effect of forces that reduce variability. *Molecular Biology and Evolution* 15:538–543. DOI: 10.1093/oxfordjournals.molbev.a025953.

Chaves-Fonnegra A., Feldheim KA., Secord J., Lopez J V. 2015. Population structure and dispersal of the coral-excavating sponge *Cliona delitrix*. *Molecular Ecology* 24:1447–66.

DOI: 10.1111/mec.13134.

Christie MR., Tissot BN., Albins MA., Beets JP., Jia Y., Ortiz DM., Thompson SE., Hixon MA. 2010. Larval connectivity in an effective network of marine protected areas. *PLoS ONE* 5:e15715. DOI: 10.1371/journal.pone.0015715.

Chust G., Villarino E., Chenuil A., Irigoien X., Bizsel N., Bode A., Broms C., Claus S., Fernández de Puelles ML., Fonda-Umani S., Hoarau G., Mazzocchi MG., Mozetič P., Vandepitte L., Veríssimo H., Zervoudaki S., Borja A. 2016. Dispersal similarly shapes both population genetics and community patterns in the marine realm. *Scientific Reports* 6:28730. DOI: 10.1038/srep28730.

Chybicki IJ., Burczyk J. 2009. Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* 100:106–13. DOI: 10.1093/jhered/esn088.

Cowen RK., Paris CB., Srinivasan A. 2006. Scaling of connectivity in marine populations. *Science* 311:522–7. DOI: 10.1126/science.1122039.

Culley TM., Stamper TI., Stokes RL., Brzyski JR., Hardiman N a., Klooster MR., Merritt BJ. 2013. An efficient technique for primer development and application that integrates fluorescent labeling and multiplex PCR. *Applications in Plant Sciences* 1:1300027. DOI: 10.3732/apps.1300027.

Dailianis T., Tsigenopoulos CS., Dounas C., Voultsiadou E. 2011. Genetic diversity of the imperilled bath sponge *Spongia officinalis* Linnaeus, 1759 across the Mediterranean Sea: patterns of population differentiation and implications for taxonomy and conservation. *Molecular Ecology* 20:3757–72. DOI: 10.1111/j.1365-294X.2011.05222.x.

DeBiasse MB., Richards VP., Shivji MS. 2010. Genetic assessment of connectivity in the common reef sponge, *Callyspongia vaginalis* (Demospongiae: Haplosclerida) reveals high population structure along the Florida reef tract. *Coral Reefs* 29:47–55. DOI: 10.1007/s00338-009-0554-0.

Diaz M., Rutzler K. 2001. Sponges: an essential component of Caribbean coral reefs. *Bulletin of Marine Science* 69:535–546.

Duchesne P., Turgeon J. 2012. FLOCK provides reliable solutions to the “number of populations” problem. *Journal of Heredity* 103:734–743. DOI: 10.1093/jhered/ess038.

Dunlap M., Pawlik JR. 1998. Spongivory by parrotfish in Florida mangrove and reef habitats. *Marine Ecology* 19:325–337. DOI: 10.1111/j.1439-0485.1998.tb00471.x.

Duran S., Pascual M., Estoup A., Turon X. 2004. Strong population structure in the marine sponge Crambe crambe (Poecilosclerida) as revealed by microsatellite markers. *Molecular Ecology* 13:511–522. DOI: 10.1046/j.1365-294X.2004.02080.x.

Ereskovsky AV., Tokina DB. 2004. Morphology and fine structure of the swimming larvae of *Ircinia oros* (Porifera, Demospongiae, Dictyoceratida). *Invertebrate Reproduction & Development* 45:137–150. DOI: 10.1080/07924259.2004.9652583.

Erpenbeck D., Hooper JNA., Wörheide G. 2006. CO1 phylogenies in diploblasts and the “Barcoding of Life” - are we sequencing a suboptimal partition? *Molecular Ecology Notes* 6:550–553. DOI: 10.1111/j.1471-8286.2005.01259.x.

Ezer T., Thattai D V., Kjerfve B., Heyman WD. 2005. On the variability of the flow along the Meso-American Barrier Reef system: A numerical model study of the influence of the Caribbean current and eddies. In: *Ocean Dynamics*. Springer-Verlag, 458–475. DOI: 10.1007/s10236-005-0033-2.

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* 3:294–9.

Foltz DW. 1986. Null alleles as a possible cause of heterozygote deficiencies in the oyster *Crassostrea virginica* and other bivalves. *Evolution* 40:869–870. DOI: 10.1111/j.1558-5646.1986.tb00549.x.

Foster NiL., Paris CB., Kool JT., Baums IB., Stevens JR., Sanchez JA., Bastidas C., Agudelo C., Bush P., Day O., Ferrari R., Gonzalez P., Gore S., Guppy R., McCartney MA., McCoy C., Mendes J., Srinivasan A., Steiner S., Vermeij MJA., Weil E., Mumby PJ. 2012. Connectivity of Caribbean coral populations: complementary insights from empirical and modelled gene flow. *Molecular Ecology* 21:1143–1157. DOI: 10.1111/j.1365-294X.2012.05455.x.

Freeman CJ., Thacker RW. 2011. Complex interactions between marine sponges and their symbiotic microbial communities. *Limnology and Oceanography* 56:1577–1586. DOI: 10.4319/lo.2011.56.5.1577.

Galarza JA., Carreras-Carbonell J., Macpherson E., Pascual M., Roques S., Turner GF., Rico C. 2009. The influence of oceanographic fronts and early-life-history traits on connectivity among littoral fish species. *Proceedings of the National Academy of Sciences* 106:1473–8. DOI: 10.1073/pnas.0806804106.

Garza JC., Williamson EG. 2001. Detection of reduction in population size using data from microsatellite loci. *Molecular Ecology* 10:305–18. DOI: doi.org/10.1046/j.1365-294X.2001.01190.x.

de Goeij JM., van Oevelen D., Vermeij MJA., Osinga R., Middelburg JJ., de Goeij AFPM., Admiraal W. 2013. Surviving in a marine desert: The sponge loop retains resources within coral reefs. *Science* 342:108–110. DOI: 10.1126/science.1241981.

Griffiths SM., Fox G., Briggs PJ., Donaldson IJ., Hood S., Richardson P., Leaver GW., Truelove NK., Preziosi RF. 2016. A Galaxy-based bioinformatics pipeline for optimised, streamlined microsatellite development from Illumina next-generation sequencing data. *Conservation Genetics Resources* 8:481–486. DOI: 10.1007/s12686-016-0570-7.

Guardiola M., Frotscher J., Uriz J. 2012. Genetic structure and differentiation at a short-time scale of the introduced calcarean sponge *Paraleucilla magna* to the western Mediterranean. *Hydrobiologia* 687:71–84. DOI: 10.1007/s10750-011-0948-1.

Guardiola M., Frotscher J., Uriz M-J. 2016. High genetic diversity, phenotypic plasticity, and invasive potential of a recently introduced calcareous sponge, fast spreading across the Atlanto-Mediterranean basin. *Marine Biology* 163:123. DOI: 10.1007/s00227-016-2862-6.

Guillot Gi., Mortier F., Estoup A. 2005. Geneland: a computer package for landscape genetics.

Molecular Ecology Notes 5:712–715. DOI: 10.1111/j.1471-8286.2005.01031.x.

Guillot G., Santos F., Estoup A. 2008. Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics* 24:1406–1407. DOI: 10.1093/bioinformatics/btn136.

Hedgecock D. 1982. Genetic consequences of larval retention: theoretical and methodological aspects. In: *Estuarine Comparisons*. Elsevier, 553–568. DOI: 10.1016/B978-0-12-404070-0.50039-9.

Hedgecock D., Li G., Hubert S., Bucklin K., Ribes V. 2004. Widespread null alleles and poor cross-species amplification of microsatellite DNA loci cloned from the Pacific oyster, *Crassostrea gigas*. *Journal of Shellfish Research* 23:379–386. DOI: http://dx.doi.org/10.1890/09-1249.1.

Hedgecock D., Pudovkin AI. 2011. Sweepstakes reproductive success in highly fecund marine fish and shellfish: A review and commentary. *Bulletin of Marine Science* 87:971–1002. DOI: 10.5343/bms.2010.1051.

Hepburn RI., Sale PF., Dixon B., Heath DD. 2008. Genetic structure of juvenile cohorts of bicolor damselfish (*Stegastes partitus*) along the Mesoamerican barrier reef: chaos through time. *Coral Reefs* 28:277–288. DOI: 10.1007/s00338-008-0423-2.

Herrera M., Nanninga GB., Planes S., Jones GP., Thorrold SR., Saenz-Agudelo P., Almany GR., Berumen ML. 2016. Seascape and life-history traits do not predict self-recruitment in a coral reef fish. *Biology Letters* 12:422–428. DOI: 10.1098/rsbl.2016.0309.

Hoban SM., Mezzavilla M., Gaggiotti OE., Benazzo A., van Oosterhout C., Bertorelle G. 2013. High variance in reproductive success generates a false signature of a genetic bottleneck in populations of constant size: a simulation study. *BMC Bioinformatics* 14:309. DOI: 10.1186/1471-2105-14-309.

Hoffman JI., Peck LS., Linse K., Clarke A. 2011. Strong Population Genetic Structure in a Broadcast-Spawning Antarctic Marine Invertebrate. *Journal of Heredity* 102:55–66. DOI: 10.1093/jhered/esq094.

Hogan JD., Thiessen RJ., Heath DD. 2010. Variability in connectivity indicated by chaotic genetic patchiness within and among populations of a marine fish. *Marine Ecology Progress Series* 417:263–275. DOI: 10.2307/24875317.

Holleley CE., Geerts PG. 2009. Multiplex Manager 1.0: a cross-platform computer program that plans and optimizes multiplex PCR. *BioTechniques* 46:511–7. DOI: 10.2144/000113156.

Johns WE., Townsend TL., Fratantoni DM., Wilson WD. 2002. On the Atlantic inflow to the Caribbean Sea. *Deep-Sea Research Part I: Oceanographic Research Papers* 49:211–243. DOI: 10.1016/S0967-0637(01)00041-3.

Jombart T. 2008. Adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24:1403–1405. DOI: 10.1093/bioinformatics/btn129.

Jombart T., Devillard S., Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics* 11:94. DOI: 10.1186/1471-2156-11-94.

Jost L. 2008. GST and its relatives do not measure differentiation. *Molecular Ecology* 17:4015–4026. DOI: 10.1111/j.1365-294X.2008.03887.x.

Knutsen H., Jorde PE., Andre C., Stenseth NC. 2003. Fine-scaled geographical population structuring in a highly mobile marine species: the Atlantic cod. *Molecular Ecology* 12:385–394. DOI: 10.1046/j.1365-294X.2003.01750.x.

Kough AS., Paris CB. 2015. The influence of spawning periodicity on population connectivity. *Coral Reefs* 34:753–757. DOI: 10.1007/s00338-015-1311-1.

López-Legentil S., Pawlik JR. 2009. Genetic structure of the Caribbean giant barrel sponge *Xestospongia muta* using the I3-M11 partition of COI. *Coral Reefs* 28:157–165. DOI: 10.1007/s00338-008-0430-3.

Maldonado M., Riesgo A. 2008. Reproduction in Porifera: a synoptic overview. *Treballs de la SCB* 59:29–49. DOI: 10.2436/20.1501.02.56.

Maldonado M., Sánchez-Tocino L., Navarro C. 2010. Recurrent disease outbreaks in corneous demosponges of the genus *Ircinia*: Epidemic incidence and defense mechanisms. *Marine Biology* 157:1577–1590. DOI: 10.1007/s00227-010-1431-7.

Maldonado M., Uriz M. 1999. Sexual propagation by sponge fragments. *Nature* 398:1999. DOI: 10.1038/19007.

Mariani S., Uriz MJ., Turon X., Alcoverro T. 2006. Dispersal strategies in sponge larvae: Integrating the life history of larvae and the hydrologic component. *Oecologia* 149:174–184. DOI: 10.1007/s00442-006-0429-9.

Marino C., Pawlik J., López-Legentil S., Erwin P. 2017. Latitudinal variation in the microbiome of the sponge Ircinia campana correlates with host haplotype but not anti-predatory chemical defense. *Marine Ecology Progress Series* 565:53–66. DOI: 10.3354/meps12015.

Meirmans PG., Hedrick PW. 2011. Assessing population structure: FST and related measures. *Molecular Ecology Resources* 11:5–18. DOI: 10.1111/j.1755-0998.2010.02927.x.

Meirmans PG., Van Tiendener PH. 2004. Genotype and Genodive: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4:792–794. DOI: 10.1111/j.1471-8286.2004.00770.x.

Morin PA., Leduc RG., Archer FI., Martien KK., Huebinger R., Bickham JW., Taylor BL. 2009. Significant deviations from Hardy-Weinberg equilibrium caused by low levels of microsatellite genotyping errors. *Molecular Ecology Resources* 9:498–504. DOI: 10.1111/j.1755-0998.2008.02502.x.

Morrison CL., Ross SW., Nizinski MS., Brooke S., Järnegren J., Waller RG., Johnson RL., King TL. 2011. Genetic discontinuity among regional populations of *Lophelia pertusa* in the North Atlantic Ocean. *Conservation Genetics* 12:713–729. DOI: 10.1007/s10592-010-0178-5.

Muhling BA., Smith RH., Vásquez-Yeomans L., Lamkin JT., Johns EM., Carrillo L., Sosa-Cordero E., Malca E. 2013. Larval fish assemblages and mesoscale oceanographic structure along the Mesoamerican Barrier Reef System. *Fisheries Oceanography* 22:409–428. DOI: 10.1111/fog.12031.

Narum SR. 2006. Beyond Bonferroni: Less conservative analyses for conservation genetics. *Conservation Genetics* 7:783–787. DOI: 10.1007/s10592-005-9056-y.

Palero F., Abelló P., Macpherson E., Beaumont M., Pascual M. 2011. Effect of oceanographic barriers and overfishing on the population genetic structure of the European spiny lobster (*Palinurus elephas*). *Biological Journal of the Linnean Society* 104:407–418. DOI: 10.1111/j.1095-8312.2011.01728.x.

Palumbi S. 2003. Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications* 13:146–158. DOI: 10.1890/1051-0761(2003)013[0146:PGDCAT]2.0.CO;2.

Pante E., Simon-Bouhet B. 2013. marmap: A package for importing, plotting and analyzing bathymetric and topographic data in R. *PLoS ONE* 8:e73051. DOI: 10.1371/journal.pone.0073051.

Paris C., Chérubin L., Cowen R. 2007. Surfing, spinning, or diving from reef to reef: effects on population connectivity. *Marine Ecology Progress Series* 347:285–300. DOI: 10.3354/meps06985.

Pascual M., Rives B., Schunter C., Macpherson E., Airoldi L., Serrão E. 2017. Impact of life history traits on gene flow: A multispecies systematic review across oceanographic barriers in the Mediterranean Sea. *PLoS ONE* 12:e0176419. DOI: 10.1371/journal.pone.0176419.

Peakall R., Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research- an update. *Bioinformatics* 28:2537–2539. DOI: 10.1093/bioinformatics/bts460.

Peery ZM., Kirby R., Reid BN., Stoelting R., Doucet-Bëer E., Robinson S., Vásquez-Carrillo C., Pauli JN., Palsboll PJ. 2012. Reliability of genetic bottleneck tests for detecting recent population declines. *Molecular Ecology* 21:3403–3418. DOI: 10.1111/j.1365-294X.2012.05635.x.

Pérez-Portela R., Noyer C., Becerro MA. 2015. Genetic structure and diversity of the endangered bath sponge *Spongia lamella*. *Aquatic Conservation: Marine and Freshwater Ecosystems* 25:365–379. DOI: 10.1002/aqc.2423.

Peterson B., Chester C., Jochem F., Fourqurean JW. 2006. Potential role of sponge communities in controlling phytoplankton blooms in Florida Bay. *Marine Ecology Progress Series* 328:93–103. DOI: 10.3354/meps328093.

Piry S., Luikart G., Cornuet J-M. 1999. Computer note. BOTTLENECK: a computer program for detecting recent reductions in the effective size using allele frequency data. *Journal of Heredity* 90:502–503. DOI: 10.1093/jhered/90.4.502.

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* 5:e9950. DOI: 10.1371/journal.pone.0009950.

Pritchard JK., Stephens M., Donnelly P. 2000. Inference of Population Structure Using

Multilocus Genotype Data. *Genetics* 155:945–959.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Raymond M., Rousset F. 1995. GENEPOP (Version 1.2): Population genetics software for exact tests and ecumenicism. *Journal of Heredity* 86:248–249. DOI: 10.1093/oxfordjournals.jhered.a111573.

Restrepo JD., Kjerfve B., Hermelin M., Restrepo JC. 2006. Factors controlling sediment yield in a major South American drainage basin: the Magdalena River, Colombia. *Journal of Hydrology* 316:213–232. DOI: 10.1016/j.jhydrol.2005.05.002.

Richards VP., Bernard AM., Feldheim KA., Shivji MS. 2016. Patterns of population structure and dispersal in the long-lived "redwood" of the coral reef, the giant barrel sponge (*Xestospongia muta*). *Coral Reefs* 35:1097–1107. DOI: 10.1007/s00338-016-1435-y.

Riesgo A., Pérez-Portela R., Pita L., Blasco G., Erwin PM., López-Legentil S. 2016. Population structure and connectivity in the Mediterranean sponge *Ircinia fasciculata* are affected by mass mortalities and hybridization. *Heredity* 117:427–439. DOI: 10.1038/hdy.2016.41.

Rivetti I., Fraschetti S., Lionello P., Zambianchi E., Boero F. 2014. Global warming and mass mortalities of benthic invertebrates in the Mediterranean Sea. *PLoS ONE* 9:e115655. DOI: 10.1371/journal.pone.0115655.

Romiguier J., Gayral P., Ballenghien M., Bernard A., Cahais V., Chenuil A., Chiari Y., Dernat R., Duret L., Faivre N., Loire E., Lourenco JM., Nabholz B., Roux C., Tsagkogeorga G., Weber AA-T, Weinert LA., Belkhir K., Bierne N., Glémin S., Galtier N. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* 515:261–263. DOI: 10.1038/NATURE13685.

Rossi V., Ser-Giacomi E., López C., Hernández-García E. 2014. Hydrodynamic provinces and oceanic connectivity from a transport network help designing marine reserves. *Geophysical Research Letters* 41:2883–2891. DOI: 10.1002/2014GL059540.

Rousset F. 2008. genepop'007: a complete re-implementation of the genepop software for Windows and Linux. *Molecular Ecology Resources* 8:103–6. DOI: 10.1111/j.1471-8286.2007.01931.x.

Shanks AL. 2009. Pelagic larval duration and dispersal distance revisited. *The Biological Bulletin* 216:373–85. DOI: 10.1086/BBLv216n3p373.

Siegel DA., Mitarai S., Costello CJ., Gaines SD., Kendall BE., Warner RR., Winters KB. 2008. The stochastic nature of larval connectivity among nearshore marine populations. *Proceedings of the National Academy of Sciences* 105:8974–9. DOI: 10.1073/pnas.0802544105.

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* 7:e35105. DOI: 10.1371/journal.pone.0035105.

Soto I., Andréfouët S., Hu C., Muller-Karger FE., Wall CC., Sheng J., Hatcher BG. 2009. Physical connectivity in the Mesoamerican Barrier Reef System inferred from 9 years of

ocean color observations. *Coral Reefs* 28:415–425. DOI: 10.1007/s00338-009-0465-0.

Stabili L., Cardone F., Alifano P., Tredici SM., Piraino S., Corriero G., Gaino E. 2012. Epidemic mortality of the sponge *Ircinia variabilis* (Schmidt, 1862) associated to proliferation of a Vibrio bacterium. *Microbial Ecology* 64:802–813. DOI: 10.1007/s00248-012-0068-0.

Stevely JM., Sweat DE., Bert TM., Sim-Smith C., Kelly M. 2010. Sponge Mortality at Marathon and Long Key , Florida : Patterns of Species Response and Population Recovery. *Proceedings of the 63rd Gulf and Caribbean Fisheries Institute* 63:384–400.

Stoeckel S., Masson J-P., Jr JL., Jr MB., Garcia M. 2014. The exact distributions of FIS under partial asexuality in small finite populations with mutation. *PLoS ONE* 9:e85228. DOI: 10.1371/journal.pone.0085228.

Szpiech ZA., Jakobsson M., Rosenberg NA. 2008. ADZE: a rarefaction approach for counting alleles private to combinations of populations. *Bioinformatics* 24:2498–2504. DOI: 10.1093/bioinformatics/btn478.

Tesson SVM., Montresor M., Procaccini G., Kooistra WHCF. 2014. Temporal changes in population structure of a marine planktonic diatom. *PLoS ONE* 9:e114984. DOI: 10.1371/journal.pone.0114984.

Treml EA., Roberts JJ., Chao Y., Halpin PN., Possingham HP., Riginos C. 2012. Reproductive output and duration of the pelagic larval stage determine seascape-wide connectivity of marine populations. *Integrative and Comparative Biology* 52:525–537. DOI: 10.1093/icb/ics101.

Truelove NK., Griffiths S., Ley-Cooper K., Azueta J., Majil I., Box SJ., Behringer DC., Butler MJ., Preziosi RF. 2014. Genetic evidence from the spiny lobster fishery supports international cooperation among Central American marine protected areas. *Conservation Genetics* 16:347–358. DOI: 10.1007/s10592-014-0662-4.

Truelove NK., Ley-Cooper K., Segura-García I., Briones-Fourzán P., Lozano-Álvarez E., Phillips BF., Box SJ., Preziosi RF. 2015. Genetic analysis reveals temporal population structure in Caribbean spiny lobster (*Panulirus argus*) within marine protected areas in Mexico. *Fisheries Research* 172:44–49. DOI: 10.1016/j.fishres.2015.05.029.

Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth BC., Remm M., Rozen SG. 2012. Primer3- new capabilities and interfaces. *Nucleic Acids Research* 40:e115. DOI: 10.1093/nar/gks596.

Vaha J-P., Erkinaro J., Niemela E., Primmer CR. 2007. Life-history and habitat features influence the within-river genetic structure of Atlantic salmon. *Molecular Ecology* 16:2638–2654. DOI: 10.1111/j.1365-294X.2007.03329.

van Soest, R. (2010). *Ircinia campana* (Lamarck, 1814). In: Van Soest, R.W.M; Boury-Esnault, N.; Hooper, J.N.A.; Rützler, K.; de Voogd, N.J.; Alvarez de Glasby, B.; Hajdu, E.; Pisera, A.B.; Manconi, R.; Schoenberg, C.; Klautau, M.; Picton, B.; Kelly, M.; Vacelet, J.; Dohrmann, M.; Díaz, M.-C.; Cárdenas, P.; Carballo, J. L. (2017). World Porifera Database. Accessed at [http://www.marinespecies.org/porifera/porifera.php?p=taxdetails&id=165009] on 2017-08-29.

Vignaud TM., Maynard JA., Leblois R., Meekan MG., Vázquez-Juárez R., Ramírez-Macías D., Pierce SJ., Rowat D., Berumen ML., Beeravolu C., Baksay S., Planes S. 2014. Genetic structure of populations of whale sharks among ocean basins and evidence for their historic rise and recent decline. *Molecular Ecology* 23:2590–601. DOI: 10.1111/mec.12754.

White C., Selkoe KA., Watson J., Siegel DA., Zacherl DC., Toonen RJ. 2010. Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences* 277:1685–1694. DOI: 10.1098/rspb.2009.2214.

Wilkinson CR., Cheshire AC. 1990. Comparisons of sponge populations across the Barrier Reefs of Australia and Belize: evidence for higher productivity in the Caribbean. *Marine Ecology Progress Series* 67:285–294. DOI: 10.3354/meps067285.

Wulff JL. 1991. Asexual fragmentation, genotype success, and population dynamics of erect branching sponges. *Journal of Experimental Marine Biology and Ecology* 149:227–247. DOI: 10.1016/0022-0981(91)90047-Z.

Wulff J. 2006. Rapid diversity and abundance decline in a Caribbean coral reef sponge community. *Biological Conservation* 127:167–176. DOI: 10.1016/j.biocon.2005.08.007.

Zilberberg C., Solé-Cava AM., Klautau M. 2006. The extent of asexual reproduction in sponges of the genus *Chondrilla* (Demospongiae: Chondrosida) from the Caribbean and the Brazilian coasts. *Journal of Experimental Marine Biology and Ecology* 336:211–220. DOI: 10.1016/j.jembe.2006.05.010.

Chapter 5: Host genotype influences microbiome composition in the sponge *Ircinia campana*

Sarah M. Griffiths[1], Luca Lenzi[2], Anita Lucaci[2], Donald C. Behringer[3, 4], Mark J. Butler IV[5] and Richard F. Preziosi[1]

[1] Division of Biology and Conservation Ecology, Manchester Metropolitan University, Chester Street, Manchester, UK.
[2] Centre for Genomic Research, Institute of Integrative Biology, University of Liverpool, Liverpool, UK.
[3] School of Forest Resources and Conservation, Program in Fisheries and Aquatic Sciences, University of Florida, Gainesville, Florida, USA.
[4] Emerging Pathogens Institute, University of Florida, Gainesville, Florida, USA.
[5] Department of Biological Sciences, Old Dominion University, Norfolk, Virginia, USA.

Keywords: Associated microbial communities; Bacteria; microsatellite; Porifera; genetic diversity; Chloroflexi; HMA.

Acknowledgements

This study was funded by a NERC-NBAF pilot project grant awarded to SMG and RFP, and a NERC PhD studentship awarded to SMG. Many thanks to J. Rowntree and G. Fox for their helpful advice in the design of this study, and to P. Fullwood at the University of Manchester DNA Sequencing Facility for the use of lab equipment. Thanks to students from University of Florida and Old Dominion University for their assistance in collecting samples. Thanks to the staff at CGR Liverpool for their advice and assistance.

Author contributions

SG and RFP conceived and designed the study, SG, RFP, MB and DCB secured funding, SG, MB and DCB collected the samples, SG and AL carried out the lab work, LL and SG did the analysis, SG wrote the manuscript.

ABSTRACT

Marine sponges are host to incredibly large and diverse communities of microorganisms. The composition of the microbiome differs among species, indicating a close co-evolutionary relationship and a key role in host identity in shaping its microbial community. However, the factors governing intraspecific microbiome variability are underexplored, and may shed light on the evolutionary and ecological relationship between host and microbiome. Here, we show that variability in microbiome composition correlates with host genotypic variation in the marine sponge *Ircinia campana* in two locations in the Florida Keys, USA. We used multilocus microsatellite genotyping to characterise the host, V4 16S rRNA Illumina amplicon sequencing to characterise their microbial communities, and Mantel tests to examine the correlation between them. More genetically similar sponges hosted significantly more similar microbial communities in both locations sampled. These results show that microbiome specificity extends beyond the host species level to the level of genotype. This may be due to stable vertical transmission of the microbial community from parent to offspring, making microbiomes more similar by descent. Alternatively, sponge genotypic variation may reflect variation in functional traits that impose differential selective pressures on horizontally acquired environmental microbes. This study provides further evidence for the strong co-evolutionary relationship between the sponge and its microbiome, and the importance of intraspecific variability in mediating eco-evolutionary dynamics of host-associated microbiomes.

INTRODUCTION

Research into the microbial communities associated with plants and animals (microbiomes) has attracted much interest in recent years, bringing about increased recognition of their roles in host biology and ecology (Gilbert, Jansson & Knight, 2014; Vandenkoornhuyse et al., 2015). This includes influence in nutrition (Stevens & Hume, 1998; Flint et al., 2012), disease susceptibility (Honda & Littman, 2012), behaviour (Ezenwa et al., 2012; Archie & Tung, 2015) and physiology (Blaser et al., 2013) amongst others, making microbiomes of significant interest in wildlife conservation and resource management (Redford et al., 2012; Bahrndorff et al., 2016; Busby et al., 2017), as well as ecology and evolution research. In turn, microbiome composition is affected by numerous features in both the host and environment, and determining what these are is essential to furthering our understanding of these relationships (Antwis et al., 2017).

The microbial communities associated with marine sponges (phylum Porifera) have received considerable attention due to their incredible size and diversity (Lee et al., 2011; Schmitt et al., 2012; Thomas et al., 2016). Microbes can form up to 35% of the biomass of sponge tissue in high microbial abundance (HMA) sponges (Vacelet & Donadey, 1977), and their diversity encompasses 32 bacterial phyla and candidate phyla that are found regularly, and an additional 20 found rarely (Thomas et al., 2016). In contrast, low microbial abundance (LMA) sponges generally have a lower diversity of microbes, with a different taxonomic composition than HMA sponges (Giles et al., 2013; Blanquer, Uriz & Galand, 2013; Erwin et al., 2015). Microbes have been found to confer various benefits to the host, including roles in nutrition and waste removal (Thomas, Rusch & DeMaere, 2010; Freeman & Thacker, 2011; Freeman et al., 2013), and production of compounds that protect their host against predation (Garate, Blanquer & Uriz, 2015) and surface fouling (On, Lau & Qian, 2006). The diversity of microbes found in sponges and the variability among species and environments has driven proposals for sponges to be used as natural models to study animal-microbial symbioses (Taylor et al., 2007; Pita, Fraune &

Hentschel, 2016). However, disentangling the relationship between sponges and members of the microbiome is challenging due to the incredible diversity and abundance of microbes in a single host.

As most sponges are filter feeders, they are exposed to an immense variety of ambient microbes, but their own communities remain distinct from seawater (Hentschel et al., 2002; Taylor et al., 2004; Schmitt et al., 2012). Environmental characters including depth (Olson & Gao, 2013; Morrow, Fiore & Lesser, 2016), geographical location (Fiore, Jarett & Lesser, 2013) and habitat (Cleary et al., 2013; Weigel & Erwin, 2015) have been shown to correlate with sponge microbiome variability in some studies, although these factors have not been found to be universal (for example, Taylor et al., 2004; Giles et al., 2013; Pita, López-Legentil & Erwin, 2013; Pita et al., 2013). It is, however, clear that host-associated factors have the most influence in determining microbiome composition, as different sponge species have been repeatedly found to host different microbial communities (Webster et al., 2010; Schmitt et al., 2012; Gloeckner et al., 2012; Blanquer, Uriz & Galand, 2013; Pita et al., 2013; Thomas et al., 2016).

In HMA sponges, relatively few microbial OTUs (operational taxonomic units) are shared among sponge species, and most are specialist to only one or a few sponge species (Schmitt et al., 2012). There are also few opportunist OTUs indicating a close and stable relationship between a sponge and its microbiome (Thomas et al., 2016). In fact, the importance of microbial symbionts and their close relationship with the host has prompted characterisation of the sponge and microbiome together as a holobiont, and their combination of genomes as a hologenome (Webster & Thomas, 2016). This relationship may be driven by vertical transmission of microbial associates (i.e. parent to offspring transmission), by horizontal transfer of microbes from the seawater coupled with high selection pressure from the host, or by a combination of both (Thacker & Freeman, 2012).

Until recently, the amount of intraspecific variation in sponge microbiomes was largely unknown due to the small sample sizes traditionally utilized in such studies. However, recently Thomas et al. (2016) characterised the microbiomes of up to 109 individuals of the same species, and showed differing amounts of intraspecific microbiome variation among host species. Elucidating the factors that may drive intraspecific variation is therefore an interesting question worthy of further study. In general, properties of the host can affect the composition of associated communities, and in turn these properties (i.e. phenotypes) can be affected by their genotype, which varies within species as well as among species. Genetic identity of a host can therefore be influential in structuring the communities associated with it (Zytynska et al., 2011; Whitham et al., 2012; Crutsinger et al., 2013; Bálint et al., 2013). However, it is not known if these effects are present in sponge-associated microbiomes.

In this study, we aim to explore this question in the HMA demosponge *Ircinia campana* (Lamarck, 1814; Dictyoceratida, Irciniidae). The microbiome of *I. campana* has been the focus of one previous study, in which the authors assessed its associated microbial community over a range of latitudes and between host haplotypes (Marino et al., 2017). The authors found that microbiome composition varied along a latitudinal gradient, which correlated with the distribution of mitochondrial cytochrome oxidase I (*COI*) gene haplotypes. However, mitochondrial DNA evolves slowly in sponges and as such, intraspecific variability is generally low at this gene and the study identified only two haplotypes. Haplotypes were correlated with location (latitude), and therefore it was not possible to separate the effects of geographical location and host genetics on the microbiome. To study the relationship between host genetic identity and the composition of its microbiome, there is a need for more variable genetic markers to be used to characterise the host sponge. Microsatellites (tandem repeats of up to six nucleotides) are one such marker that can show high levels of polymorphism in sponges (Uriz & Turon, 2012), and may be more suitable for an in-depth analysis of the effects of host genetics on microbiome composition.

Another study by Noyer & Becerro (2012) investigated the relationship between genetic and bacterial diversity (and additionally, chemical diversity) in *Spongia lamella* in the Mediterranean and Atlantic Iberian coast. Although again, latitude was found to be an important factor, there was no significant relationship found between host genetics and bacterial communities. Microsatellites were used to characterise the hosts in this study, but microbial communities were characterised using denaturing gradient gel electrophoresis (DGGE), which gives lower taxonomic resolution than next-generation sequencing (NGS) methods. Therefore it may be the case that the methods used gave insufficient statistical power to adequately describe bacterial diversity and correlate it with host genetics.

In this chapter, I investigate this question using molecular techniques that allow a higher resolution in characterising both host genetics and microbiome composition than those employed in previous studies. I use highly polymorphic microsatellites to genotype *Ircinia campana* host sponges, and 16S Illumina amplicon sequencing to characterise their associated microbial communities. I then explore the relationship between genetic variability in *I. campana* and variability in its microbial community.

METHODS

Sample collection

Ircinia campana individuals were sampled by snorkelling at two shallow (< 2 m) nearshore hard bottom habitats in Florida Bay, Florida Keys (FL, USA): Long Key in the Middle Keys (24.81437, -80.8307) and Kemp Channel in the Lower Keys (24.6768, -81.4757). Samples were taken in a single collection instance at each site to eliminate temporal variability. A piece of tissue from each individual was cut and immediately preserved in absolute ethanol upon surfacing, and twenty individuals were sampled per site. The ethanol was replaced firstly to act as a rinse, removing loosely attached seawater bacteria, and secondly to prevent dilution of the ethanol in order to aid preservation of DNA. Samples were shipped to Manchester, UK, and stored at -80°C until processing. Prior to DNA extraction, the tissue was dissected under a stereomicroscope to avoid contamination with commensal organisms such as polychaetes. Total DNA was then extracted using aseptic technique with the DNeasy® Blood and Tissue Kit (Qiagen) and was normalised to 1ng/µl.

Host genotyping and summary statistics

Sponges were genotyped at nine polymorphic microsatellite markers developed in Chapter 4 (Icam23, Icam24, Icam26, Icam10, Icam18, Icam3, Icam31, Icam32 and Icam4), using the methodology described within. One locus from the set used in Chapter 4, Icam34, showed poor amplification with high amounts of missing data, so was excluded from the analysis here. Null allele frequencies were estimated using the ENA method in FreeNA (Chapuis & Estoup, 2007). Observed and expected heterozygosity was calculated in GenoDive v2.0 b27 (Meirmans & Van Tiendener, 2004). F_{ST} between the two sites was calculated in GenoDive, and its significance calculated based on 50,000 permutations. A Principles Coordinates Analysis (PCoA) was carried out using GenAlEx v6.503 (Peakall & Smouse, 2006, 2012) using pairwise Euclidean genetic distances between individuals. The R package ‘ggplot2’ (Wickham 2009) was then used to plot the Eigenvalues for the first and second coordinates.

Microbiome characterisation

PCR, library preparation and sequencing

PCR, sequencing and subsequent bioinformatics analyses were carried out at the Centre for Genomics Research, University of Liverpool, UK. Amplification of the V4 region of the *16S* rRNA gene was carried out in a two-stage nested PCR, using the primers described by Caporaso et al. (2011) (ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNNGTGCCAGCMGCCGCGGTAA; GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCTGGACTACHVGGGTWTCTAAT) in 5µl reaction volumes, with the following thermal conditions: 15 x 95°C for 20 seconds, 65°C for 15 seconds and 70°C for 30 seconds; 1 x 72°C for 5 minutes. PCR products were then purified using AMPure SPRI beads (Beckman Coulter), before entering into a second stage of PCR performed using the same conditions as above for 20 cycles, to incorporate Illumina sequencing adapter sequences containing indexes (i5 and i7) for sample identification. Following PCR, the samples were again purified, and successfully generated amplicon libraries identified by Qubit quantification and the Fragment Analyzer. Using this data, the final libraries were pooled in equimolar amounts and size selected with a Pippin Prep (Sage Science) using a size range of 300-600 base pairs (bp). Quantity and quality of each pool was assessed using a Bioanalyzer (Agilent Genomics) and qPCR with the Illumina® Library Quantification Kit (Kapa Biosystems) on a LightCycler® (Roche), following the manufacturer's instructions. Paired-end 250 bp sequencing was then carried out using the Illumina MiSeq, with fragmented PhiX bacteriophage genome added to increase sequence complexity.

Quality filtering

CASAVA v1.8.2 (Illumina) was used to base call and de-multiplex indexed reads, and Cutadapt v1.2.1 (Martin, 2011) was used to remove Illumina adapter sequences. Low quality bases were trimmed from the reads using Sickle v1.200 (Joshi & Fass, 2011) (minimum window quality score 20), and following trimming, reads under 10 bp in length were removed. Reads were subject to a second trimming step with Cutadapt to remove primer sequence from the PCR process. Sequencing errors were corrected in forward and reverse reads using

the error-correct module in SPAdes v3.1.0 (Bankevich et al., 2012). Read pairs were aligned to produce a single sequence for each pair of reads using USEARCH8 (Edgar, 2010) 'fast-mergepairs' command, and a size selection between 200bp and 600bp was applied to each merged sequence set. To remove any residual contaminating PhiX sequences, BLASTN (Altschul et al., 1990) was used to search for PhiX sequences (GenBank GI:9626372) in each sample; matching sequences (E-value $<10^{-5}$) were then filtered out of the dataset. Sequences containing Ns were discarded to remove low-quality reads.

Metagenomic analysis

Sequences were clustered into operational taxonomic units (OTUs) with 99% sequence similarity. Two different clustering algorithms were used for OTU picking; the first implemented in VSEARCH 1.1.3 (Edgar, 2010) using the function '-cluster-smalmem' with 99% identity threshold, and the second in Swarm (Mahé et al., 2014). Clusters containing fewer than two sequences were removed to reduce error, and the results from both clustering steps were merged to create a non-redundant sequence set. Chimera detection was carried out in VSEARCH using both the reference-based and *de novo* methods. The 'usearch_global' function in VSEARCH was used to define the abundance of each OTU. OTUs were taxonomically classified in QIIME 1.9.0 (Caporaso et al., 2010) using pick_rep_set.py to select the most representative sequence in the OTU, and assign_taxonomy.py to match sequences to those in the SILVA 119 database (Quast et al., 2013). OTU tables were produced with counts of all OTUs for all samples.

As coverage was uneven amongst samples, rarefaction was carried out in order to standardize sequencing depth. Repeated subsampling (33 repetitions) was carried out on the OTU count table at sampling depths from 2000 to 350000 (multi_rarefaction.py), following which Chao 1 alpha diversity measures were calculated and plotted as rarefaction curves (alpha_diversity.py, collate_alpha.py, make rarefaction_plots.py). From this, 173,000 sequences was chosen as the best trade-off between capturing OTU richness and retaining samples for further analysis. The QIIME script single_rarefaction.py was used

for repeated subsampling (without replacement) at a depth of 173,000 sequences; samples not reaching this threshold were removed from subsequent analysis (nine samples were removed, 25 were retained). Non-metric multidimensional scaling (NMDS) was performed on the samples that remained following rarefaction using the QUIIME script nmds.py. NMDS1 and NMDS2 were plotted using ggplot2 in R (R Core Team, 2017).

Statistical analysis

Microbiome and host genetic data from thirteen samples from Long Key and twelve samples from Kemp Channel that remained after microbiome rarefaction were used in the following analyses. Mantel tests were carried out using 9999 permutations in the package 'ade4' in R between matrices of microbial community dissimilarity and host genetic distance. This follows the approach of Zytynska et al. (2011), and is appropriate when multilocus genotypes are utilized, and genetic variation in hosts is continuous rather than discreet groups of genotypes. The two sampling sites were analysed separately to avoid environmental or geographic variability caused by differences between the locations (for example, in seawater bacterial communities). Pairwise microbial community distance between individuals was measured using both Bray-Curtis dissimilarity and the Jaccard index. Bray Curtis dissimilarity takes into account abundance, whereas the Jaccard index uses only presence/absence data. Pairwise Euclidean genetic distances between individuals were calculated from the multilocus genotypes using GenoDive v2.0b27. In distance-based calculations, null alleles and missing data can both bias results, overestimating differences between samples. In order to minimize the risks of false positive results, GenoDive was used to randomly fill in missing data in all loci based on overall allele frequencies. A further more conservative data file was created with those loci with high (>0.16) null allele frequencies removed (Icam24, Icam26 and Icam10 in both sites, Icam3 in Kemp Channel samples). Following the removal of these loci, missing data was only present in Icam23 (0.077) and Icam3 (0.462) in Long Key; Icam3 was also removed at this site due to the high amount of missing data. Missing data in the five loci dataset was filled in randomly based on the allele frequencies over all sites as previously.

RESULTS

Host sponge genotyping

No identical multilocus genotypes were found in the dataset, indicating no clones were present amongst the sampled individuals. All loci were polymorphic, ranging up to 16 alleles per locus (Icam18, Kemp Channel), but two of the loci were monomorphic in Kemp Channel (Icam32 and Icam4) (Table 1). Null allele frequencies and the proportion of missing alleles (genotyping failures) were high for many loci, and heterozygosity deficiencies were also observed in many cases (Table 1).

Pairwise genetic differentiation between Long Key and Kemp Channel was low and not significant, albeit marginally (F_{ST} = 0.033, p = 0.054). The first and second principle coordinates of the PCoA explained only 19.99% of the total variation among the samples, and the plot shows that the individuals are not separated by site (Figure 1). The sites can therefore be considered to be well-mixed genetically, with only a minimal amount of differentiation found between the sponges at each site.

Microbial community analysis

PCR amplification and Illumina sequencing was successfully carried out on 34 samples (17 each from Long Key and Kemp Channel). Between 97.10 and 99.61% of reads were successfully assembled per sample, yielding between 32,161 and 527,441 assembled sequences between 200 and 600 bp in length per sample. One sample, KC20, yielded a far lower number of assembled sequences than the remaining samples (32,161); the next lowest number obtained was 183,418 (LK01). No chimeras were detected among the sequences. Between 80.77 and 87.57% of the filtered sequence set could be aligned to any taxa, and 31,567 OTUs were found among all samples, ranging from 4,165 to 14,503 OTUs in a single sample. Rarefaction curves showed sequencing depth was sufficient to capture the majority of 99% OTU diversity in many of the samples, with a number of curves approaching asymptotes (Figure 2). However, continued sampling would have revealed further diversity.

Among all OTUs, a total of 22 bacterial and one archaeal phyla were present. By far the most abundant phylum was Chloroflexi (62.6% of the OTUs), followed by Proteobacteria (17.5%), Acidobacteria (6.4%), PAUC34f (4.3%), SBR1093 (3.8%), Gemmatimonadetes (1.6%) and Actinobacteria (1.5%) (all bacterial phyla). The remaining phyla formed less than 1% of the OTUs present. Within the Chloroflexi, the class Anaerolineae was the most dominant, forming large proportions of the microbiomes of all samples (Figure 3). Following rarefaction, twenty-five samples remained (thirteen samples from Long Key, twelve samples from Kemp Channel). 889 core OTUs (OTUs found in all samples) were present in the rarefied dataset out of a total of 30,715 OTUs found across all samples. Sequences belonging to the core OTUs accounted for between 20.85 and 43.38% of the total rarefied sequence sets for each sample. The taxonomic composition of the core OTU set was 402 Proteobacteria OTUs, 284 Chloroflexi, 132 Acidobacteria, 35 PAUC34f, 20 SBR1093, 9 Actinobacteria, 4 Bacteroidetes and 3 unclassified Bacteria (Appendix I). Between 85 and 290 OTUs were only found in a single sample in the rarefied dataset. The NMDS ordination of Bray-Curtis dissimilarities shows some separation between the Long Key (LK) and Kemp Channel (KC) samples, and between samples within the Long Key site (Figure 4).

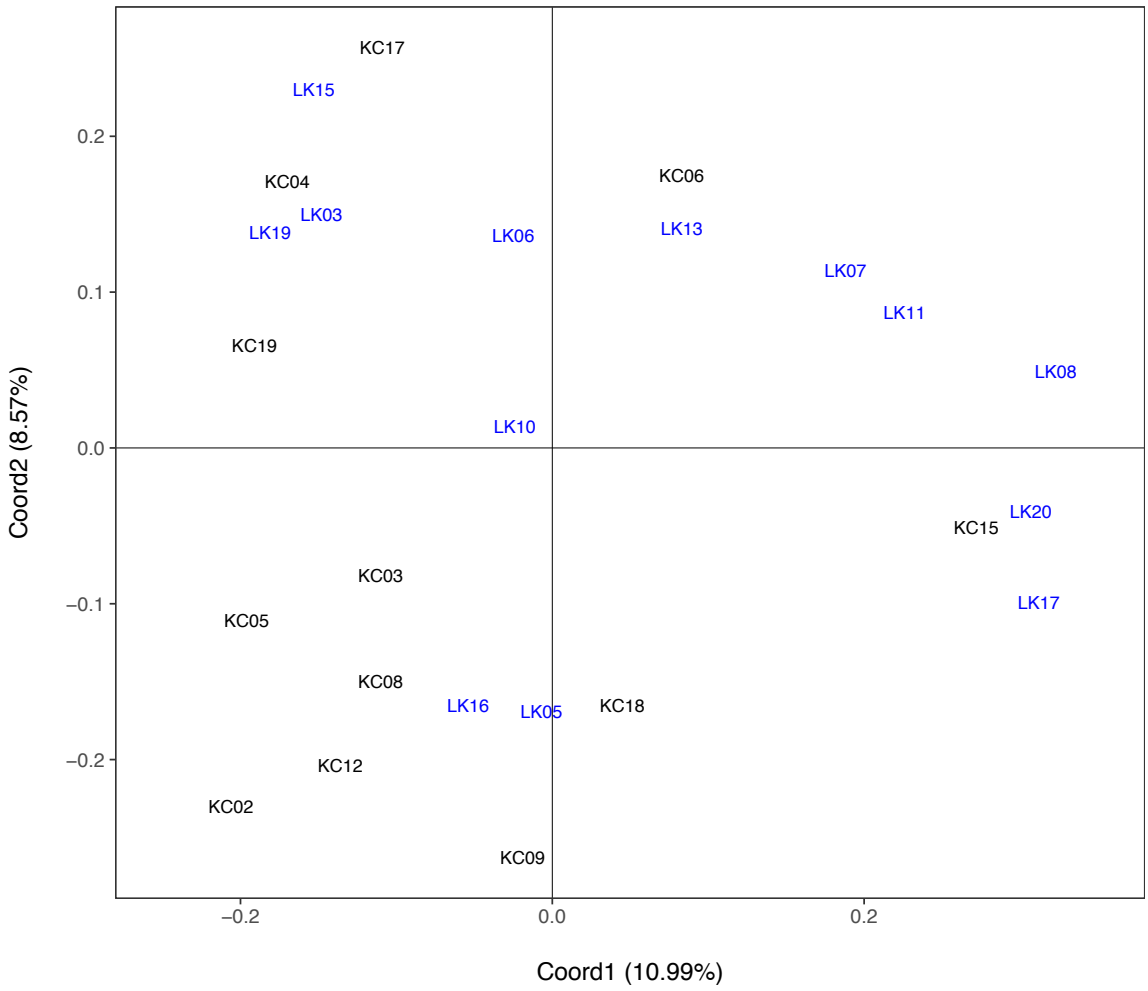


Figure 1: Principle coordinates analysis (PCoA) of pairwise Euclidean genetic distances among *Ircinia campana* individuals from Long Key (LK; blue) and Kemp Channel (KC; black). Only individuals included in the analyses after the rarefaction step are displayed.

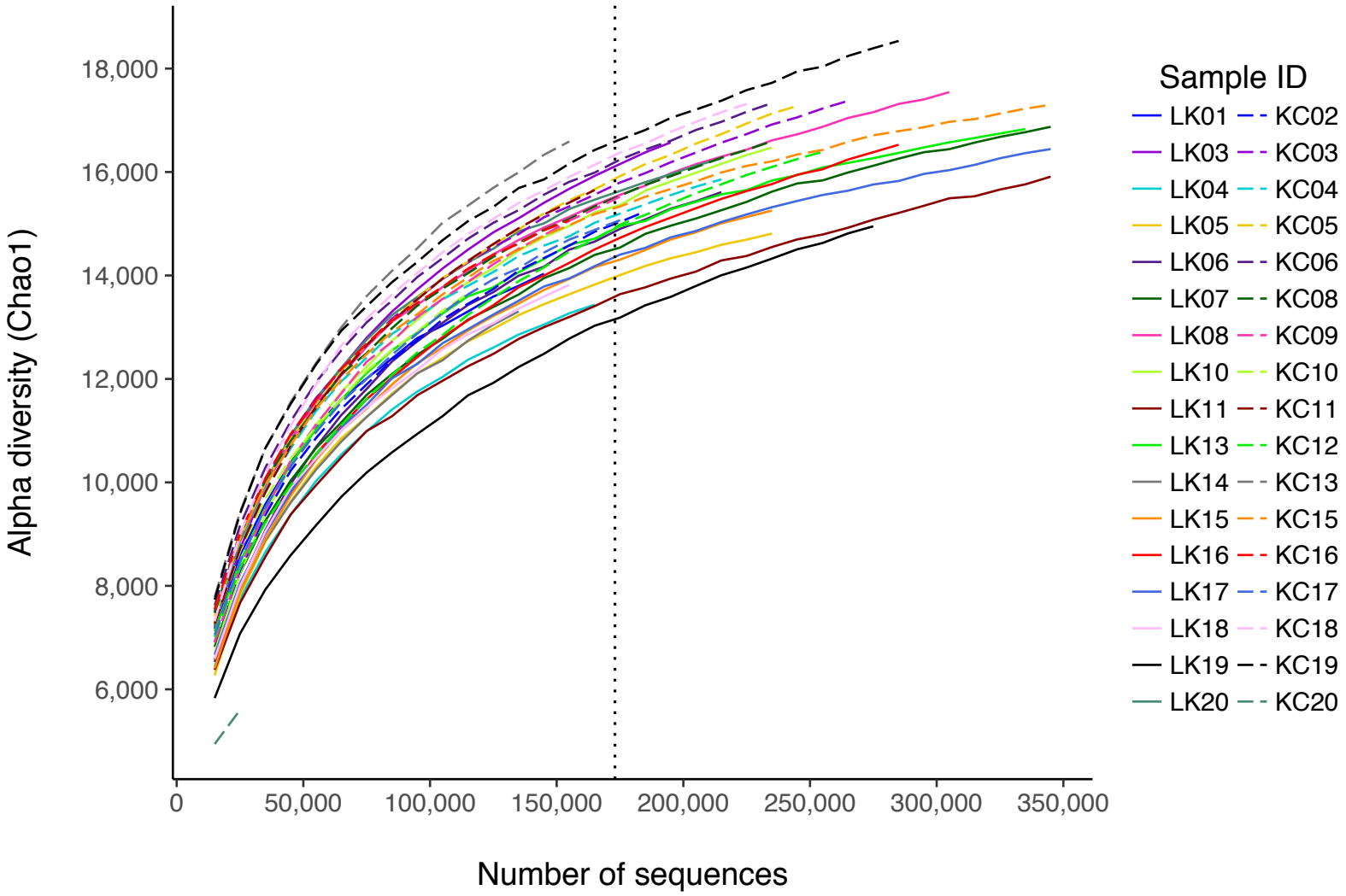


Figure 2: Rarefaction curves showing Chao1 alpha diversity for microbial 99% OTUs in *Ircinia campana* samples at increasing sequencing depths. Vertical dotted line shows 173,000 sequences, the sampling depth chosen for rarefacation of the data.

Table 1: Summary statistics for *Ircinia campana* microsatellites at Long Key (LK) and Kemp Channel (KC).

	Icam23	Icam24	Icam26	Icam10	Icam18	Icam3	Icam31	Icam32	Icam4	Overall
Long Key (LK)										
Number of alleles	15	7	7	5	15	5	3	2	2	6.778
Proportion of missing alleles	0.077	0.385	0.077	0.154	0.000	0.462	0.000	0.000	0.000	0.065
Null allele frequency	0.016	0.389	0.344	0.412	0.154	<0.000	0.061	<0.001	<0.001	0.183
Observed heterozygosity	0.833	0.125	0.167	0.000	0.615	0.857	0.538	0.231	0.077	0.383
Expected heterozygosity	0.955	0.946	0.830	0.764	0.958	0.726	0.583	0.212	0.077	0.672
Kemp Channel (KC)										
Number of alleles	6	10	5	5	16	3	5	1	1	5.778
Proportion of missing alleles	0.000	0.250	0.000	0.083	0.000	0.250	0.000	0.000	0.000	0.128
Null allele frequency	0.149	0.283	0.321	0.255	<0.001	0.298	0.157	0.001	0.001	0.229
Observed heterozygosity	0.250	0.333	0.167	0.273	0.917	0.000	0.417	0.000	0.000	0.262
Expected heterozygosity	0.564	0.944	0.750	0.736	0.966	0.417	0.697	0.000	0.000	0.564

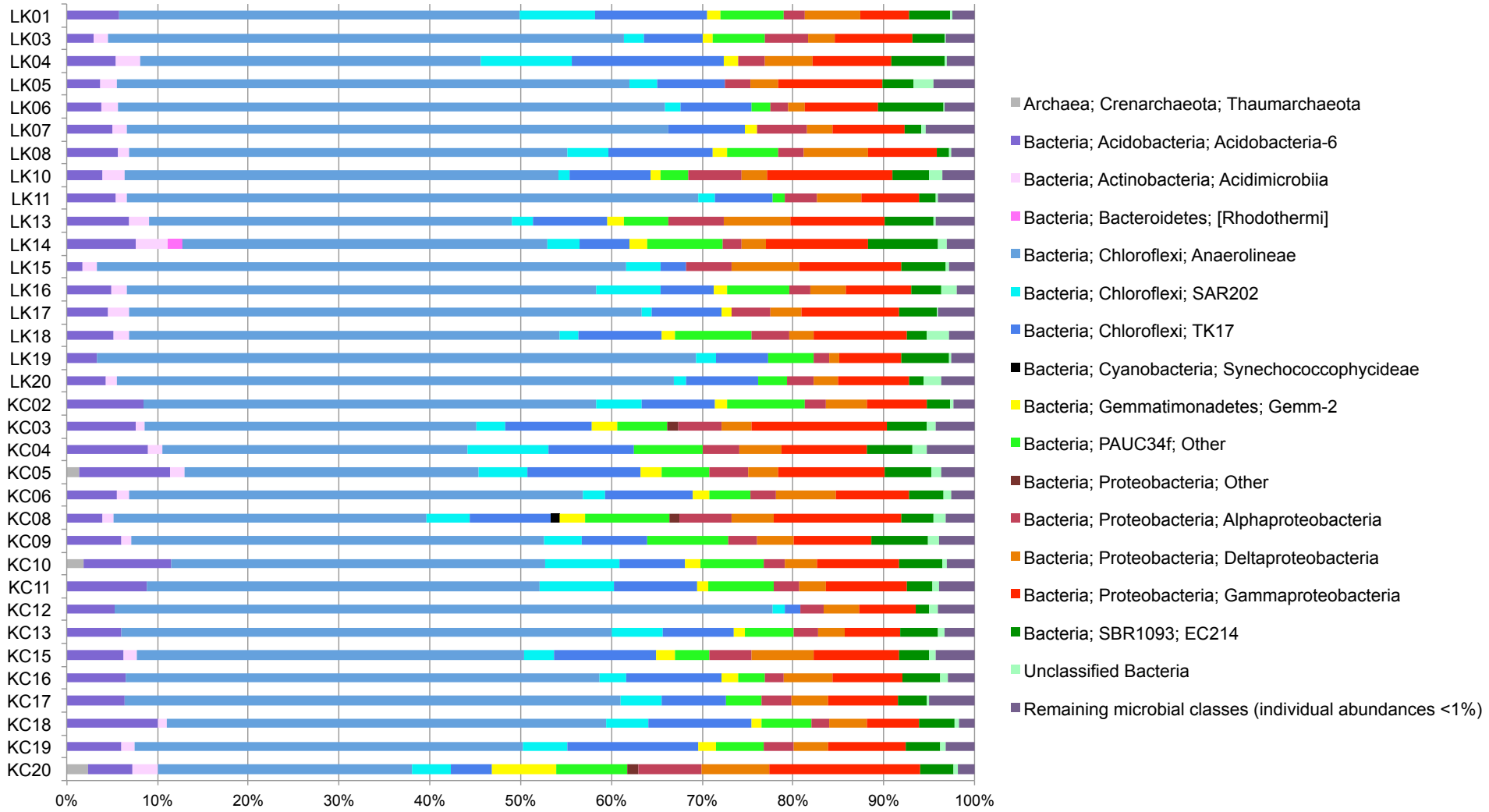


Figure 3: Stacked bar charts showing the relative abundance of microbial classes (>1% abundance) among *Ircinia campana* samples. Square brackets indicate proposed taxa. Classes that formed less than 1% relative abundance in an individual were grouped together under 'Remaining microbial classes'.

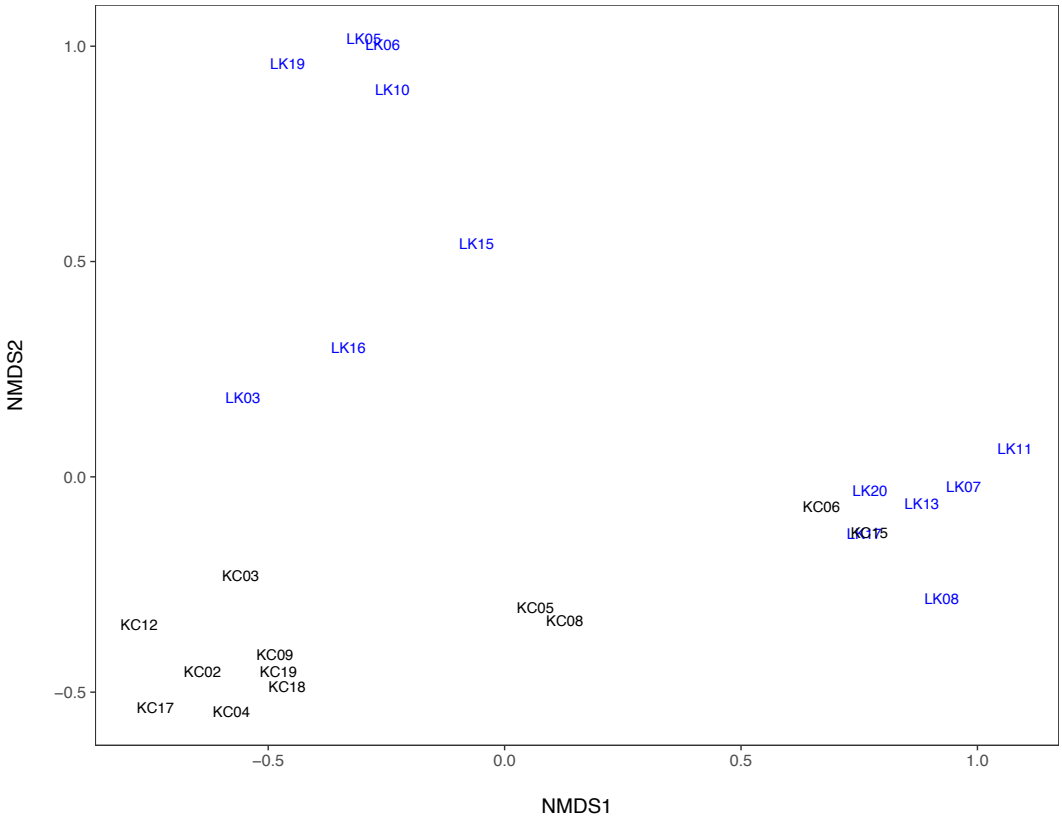


Figure 4: Non-metric multi-dimensional scaling (NMDS) ordination plot of microbial communities associated with *Ircinia campana* from Long Key (LK; blue) and Kemp Channel (KC; black).

Relationship between host genotype and microbiome composition

Positive, significant relationships between microbial community Bray-Curtis dissimilarity and Euclidean genetic distance (calculated from the nine loci) among host sponges were found at both Long Key and Kemp Channel (Mantel tests: Long Key r = 0.408, p = 0.003; KC r = 0.361, p = 0.042) (Figure 5). This was also true when the Jaccard index was used (Long Key r = 0.407, p = 0.003, Kemp Channel r = 0.509, p = 0.009). When the more conservative five loci datasets were used to calculate genetic distances, significant positive correlations were found in Long Key using both Bray Curtis dissimilarity and the Jaccard index (Bray Curtis: r = 0.429, p = 0.003; Jaccard Index: r = 0.430, p = 0.002), but only when using the Jaccard index in Kemp Channel (Bray Curtis: r = 0.268, p = 0.091; Jaccard index: r = 0.417, p = 0.045).

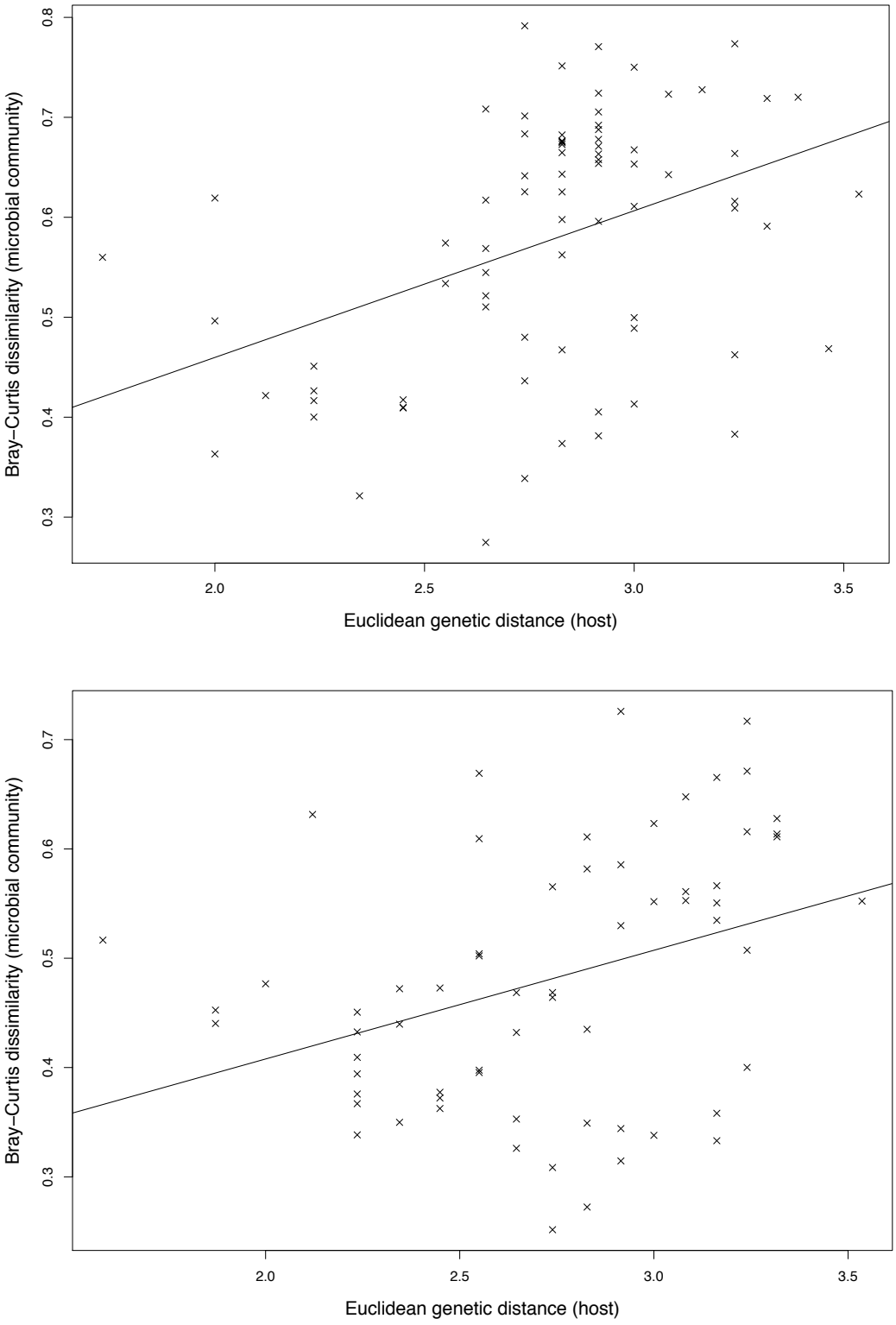


Figure 5: Scatter plots with regression lines showing microbial community dissimilarity and host genetic distance in *Ircinia campana* at Long Key (top panel) and Kemp Channel (bottom panel), using Euclidean distance between multilocus microsatellite genotypes (9 loci) and Bray-Curtis community dissimilarity between microbial communities.

DISCUSSION

Microbial community composition in *Ircinia campana*

We found high OTU richness in *I. campana* microbial communities, as has been found in other *Ircinia* species previously (Thomas et al., 2016). The microbial community composition found in this study showed similarities to the results of the *I. campana* microbiome study by Marino et al., (2017) and included the presence of bacterial phyla characteristic of HMA sponges (Chloroflexi, Acidobacteria, PAUC34f) (Schmitt et al., 2011; Bayer, Kamke & Hentschel, 2014; Moitinho-Silva et al., 2017). Marino et al. (2017) found higher taxonomic diversity in their study (33 bacterial and archaeal phyla, as opposed to 22 phyla characterised here). This may be because the authors sampled from a wider range of locations than explored here, or could be due to methodological differences between the studies; for example, different primer sets were used in each study, and methodology-based bias in metagenomics studies is well documented (Brooks et al., 2015; D'Amore et al., 2016).

Chloroflexi dominated the *I. campana* microbiome in this study, which also contrasts with the results of Marino et al. (2017), who found Proteobacteria the dominant microbial phylum for *I. campana* (the second most abundant phylum in this study). Chloroflexi are diverse and abundant in HMA sponges, which may reflect an ecological importance of the phylum (Schmitt et al., 2011), however this is as yet unknown. The relatively low proportion of cyanobacteria found in this study was surprising, given that samples were collected in very shallow water (<2 meters) with plenty of light penetration, where phototrophic microbes may be predicted to make up a more sizable amount of the community. Furthermore, cyanobacteria made up a larger proportion of the *I. campana* microbial community in the Marino et al. (2017) study. However, cyanobacteria are usually concentrated in the outer surface of the sponge, which was avoided in this study to reduce epibiotic contamination and uneven sampling of ectodermal communtites.

Influence of host genotype on microbiome composition

More genetically similar sponges hosted more similar microbiomes in *I. campana* in both the Long Key and Kemp Channel sites sampled in this study. This result adds to the evidence showing a strong influence of the host in determining microbiome composition, a determinant shown many times previously in sponges to the level of the species (Taylor et al., 2004; Schmitt et al., 2012; Giles et al., 2013; Thomas et al., 2016), including among species of the genus *Ircinia* (Erwin et al., 2012; Pita et al., 2013), but never before to the level of the genotype. This effect was apparently stronger in Long Key, where the r test statistic values were almost always higher, and p values significant in all combinations tested. This could be an artefact of the reduced sample size in Kemp Channel giving diminished statistical power (one less individual was used in the Kemp Channel analysis, leading to 12 fewer pairwise comparisons in the Mantel test). However, it is possible that other factors may have more influence in determining microbiome composition in sponges at Kemp Channel than those at Long Key.

The correlations observed in this study were moderate in strength, with the Mantel test r statistic in Long Key ranging between 0.407 and 0.430 and in Kemp channel ranging between 0.361 to 0.507, depending on the microsatellite loci set and community dissimilarity index used. Nonetheless, there are clearly additional factors driving microbial community composition in *I. campana* within both locations. Phenotypic plasticity in the host (which in sponges can be considerable; Hill & Hill, 2002; Morley et al., 2016; Guardiola, Frotscher & Uriz, 2016) could provide a non-genotypic source of host variability that could influence microbiome composition. Environmental factors may also be important, and in a complex biological system, many factors are likely to be interacting with each other to shape sponge-associated microbial communities. However, in studies on other *Ircinia* species in the Mediterranean (Pita et al., 2013) and Caribbean (Pita, López-Legentil & Erwin, 2013), location was not found to significantly affect microbiome composition within a species, with the exception of *I. variabilis* in the Mediterranean. Marino et al., (2017) showed a latitudinal gradient in microbiome composition in *I. campana*, indicating a

potential influence of environment - however as latitude was also correlated with host mitochondrial haplotype, the cause of the microbial variation observed in the study cannot be isolated. This study was designed to near-eliminate environmental (i.e. seawater bacterial community) variability by analysing sampling sites separately, and carrying out sampling in a single collection instance. Nonetheless, there is likely to be interacting effects of the genotype and the environment in shaping microbial communities. The NMDS plot shows some separation between the microbial communities present in the two sites and within Long Key, however the PCoA and F_{ST} results for the host genetics show that the sites are not differentiated in their population genetic structure. Although the purpose of this study was not to compare the microbial communities found between the sites, this does indicate that site-specific characters had more influence than host genetic identity in structuring microbial communities, but that individual host genotype was still influential within sites. Whether this is environmentally based or due to spatially variable aspects of the host biology is unknown and requires further study in order to elucidate.

The relationship between host genetics and microbiome composition may be driven by vertical transmission of microbial communities, in which more genetically similar sponges host more similar microbiomes by descent. Evidence for vertical transmission of the microbiome has been observed in sympatric congener *Ircinia felix* (Schmitt et al., 2007), as well as other species (Ereskovsky, Gonobobleva & Vishnyakov, 2005; Sharp et al., 2007; Lee et al., 2009; Sipkema et al., 2015), and is thought to be a significant driver of the high host-species fidelity of microbiomes in sponges. Many evolutionary advantages can be gained from the inheritance of parental microbiomes, as favourable symbionts that are important for sponge health and physiology are already present in growing larvae. Yet there are also disadvantages to strict vertical transmission – symbionts may not be optimal for the selection pressures exerted by new environments that larvae disperse to, and over generations microbial genome sizes reduce and functional genes are lost, thus narrowing their functional diversity (Gil et al., 2002; Moran, McCutcheon & Nakabachi,

2008; Bright & Bulgheresi, 2010). Horizontal transmission of microbial communities therefore offers its own set of advantages to sponge hosts.

It is thought that a model of 'leaky vertical transmission' - a mix of both vertical (transferred from the maternal sponge to the oocyte or larva) and horizontal (acquired from the environment) transmission - could be the mode of symbiont acquisition for many sponge species (Thacker & Freeman, 2012; Hentschel et al., 2012). Bacteria that were first thought to be 'sponge specific' bacteria are indeed found at low abundances in seawater, indicating that symbionts can be acquired from the environment (Taylor et al., 2013), and similar types of microbes can be transmitted both vertically and horizontally (Sipkema et al., 2015). Mixed transmission of microbes was found in LMA sponge *Amphimedon queenslandica*; vertically transmitted communities altered dramatically during larval settlement and development, and although the microbiome largely reverts back in the adult sponge, horizontally transmitted OTUs from the settlement stage were found to persist (Fieth et al., 2016).

For horizontally transmitted microbes, host genotype may influence which symbionts are acquired; intraspecific genetic variation in the host species may cause variation in functional traits that could influence symbiont acquisition through environmental selection. In this way, the genotype could be underlying a selective pressure that a host exerts on seawater bacteria exposed to it during the filtering process. Microsatellite genes themselves are theoretically selectively neutral, however can be linked to, or even found within, protein-encoding genes under selection, including those involved in immune response (Li et al., 2004; Santucci et al., 2007; Jensen et al., 2008; Gemayel et al., 2010; Tollenaere et al., 2012). Therefore variation in multilocus microsatellite genotypes could correspond to variation in functional genes. Such selective pressures imposed by genetic variation could include the secondary metabolites produced by the sponge, which are highly diverse (Genta-Jouve & Thomas, 2012), include antimicrobial compounds (McCaffrey & Endean, 1985; Kelman et al., 2001) and can be intraspecifically variable (Noyer, Thomas & Becerro, 2011; Puyana et al., 2015). Noyer & Becerro (2012) proposed a

framework in which genetic, chemical and bacterial diversity could be linked and influence each; increased genetic diversity in sponge populations may promote chemical variation and therefore habitat heterogeneity for microbial communities. Noyer & Becerro (2012) did not find a significant relationship among these three levels in their study of *Spongia lamella*, and chemical profile was not assessed in this study; further research into genetically-underlined heterogeneous traits that could affect microbial niches would be interesting to explore further.

Microbiome variation by host genetic variation may also arise as a result of variable responses in the host immune system. Although they do not have an acquired immune system, the innate immune system in sponges is more complex than one might imagine when considering their evolutionary basal status in the Metazoa and their relative morphological simplicity (Müller & Müller, 2003). However, to maintain specific, distinct extracellular microbial communities in the mesohyl tissue where phagocytosis of food bacteria takes place, and in the face of constant exposure to seawater bacteria, it is necessary to have a sophisticated innate immune system capable of discriminating 'food' and 'symbiont' bacteria, which has indeed been found (Wilkinson, Garrone & Vacelet, 1984; Wehrl, Steinert & Hentschel, 2007). This system includes receptor proteins at the interface between the organism and the environment that can recognise and differentiate bacteria (Wiens et al., 2005, 2007). In addition, antimicrobial compounds can be produced by the sponge in response to potential invaders (Schröder et al., 2003; Thakur, Anil & Müller, 2004).

Genetic variation has been shown to influence immune response in other organisms. Lazzaro, Sceurman & Clark (2004) found that nucleotide polymorphism in sixteen innate immunity-related genes produced widely different responses to a pathogenic bacteria in *Drosophila melanogaster*. In the coral *Acropora millepora*, variable patterns of gene expression were found among different genotypes in response to potentially-pathogenic *Vibrio spp.* (Wright et al., 2017). In one case, such effects have been found to influence the microbiome - Bolnick et al., (2014) found that polymorphism in Major

Histocompatibility Complex (MHC) class IIb genes influenced gut microbiome composition and diversity in the threespined stickleback *Gasterosteus aculeatus*. There is no evidence yet to show that immune response can be intraspecifically variable in sponges. However, there is evidence of polymorphism of the *Amphimedon queenslandica AqNLR* (nucleotide-binding domain and Leucine-rich repeat containing) genes, which are pattern recognition receptors involved in detecting and binding a range of microbial ligands (Degnan, 2015). Therefore, this pathway may be worthy of further study as a possible mechanism for intraspecific genotype-driven microbiome variation.

Further to these potential mechanisms, the sponge itself cannot be considered in isolation; selection on seawater bacteria will be performed by the holobiont. The timing and order in which microbes enter the communities will have secondary effects on determining succession and ultimately community composition (historical contingency; Costello et al., 2012), with competitive interactions occurring among community members (Esteves, Cullen & Thomas, 2017). Because of this, influence of the host genotype on even a relatively small proportion of the microbiome could increase its reach in shaping community composition.

Although complex to disentangle, it is recognized that intraspecific variation has community-level impacts (Bolnick et al., 2011), and links between the host genetics and the composition of associated communities have been found in many systems (Zytynska et al., 2011; Whitham et al., 2012; Crutsinger et al., 2013). This includes host-microbiome systems: In plants, genotypic variation in pea aphid *Acyrthosiphon pisum* affects the size of heritable bacterial communities (Chong & Moran, 2016), genotype of balsam poplar trees *Populus balsamifera* affect the fungal microbiome of the leaf (Bálint et al., 2013), and the leaf and root microbiomes of perennial wild mustard *Boechera stricta* are affected by genotype (Wagner et al., 2016). In animals, host genetics influence the gut microbiome composition in humans (Zoetendal et al., 2001; Goodrich et al., 2014), chickens (Zhao et al., 2013) mice (Benson et al., 2010) and

threespine stickleback fish (Bolnick et al., 2014), and the genotype of tadpole *Amietia hymenopus* is correlated with the microbiome associated with the mouthparts (Griffiths et al., in press).

This work is important in highlighting the importance of host intraspecific variation in determining microbiome composition in *Ircinia campana*. However it also brings with it a number of questions to be answered in order to better understand this relationship and its ecological and evolutionary implications. It is important to understand if vertical transmission is driving the association, or phenotypic variation creates differential selection pressures and niches for colonizing microbes. It would also be enlightening to investigate if the microbiome community differences seen here correspond to any functional diversity, and have subsequent effects on host biology and ecology. Recently, as techniques such as whole genome sequencing and transcriptomics have become more accessible, studies have begun to identify functional roles of members of sponge microbiomes, including metabolic interactions and pathways (Moitinho-Silva et al., 2017). The temporal stability of the relationship, and its interaction with environmental characters are also as yet unknown.

With microbial imbalances triggered by ocean warming implicated in disease (and subsequent mass mortalities) in *Ircinia fasciculata* in the Mediterranean (Blanquer et al., 2016), understanding the relationship between *I. campana* and its microbiome becomes ever more crucial, especially given their dominant role in the ecosystem in Florida Bay (Chiappone & Sullivan, 1994; Tellier & Bertelsen, 2008). This is especially compelling given the instability in the Florida Bay ecosystem (Robblee et al., 1991; Butler et al., 1995; Fourqurean & Robblee, 1999; Kearney et al., 2015; Hall et al., 2016), and the mass mortalities already suffered by sponges in the area (Butler et al., 1995; Stevely et al., 2010). Additionally, the microbiota in coral reef organisms (including sponges) has been proposed as a potentially important mechanism for acclimation and resilience to climate change scenarios (Webster & Reusch, 2017). With this in mind, understanding individual-level drivers of microbiome variation may assist in species management and conservation in the face of future stressors.

The correlation between host genetics and microbiome composition suggests that more genetically variable populations of host sponges may support higher levels of microbial community variability. This could have implications in the light of sponge mass mortalities: losses of genetic diversity through bottleneck effects may be associated with reductions in microbiome variability. As microbes are suspected to be the producers of some 'sponge'-derived compounds (Thakur & Anil, 2000; Thomas, Kavlekar & LokaBharathi, 2010), this could have negative consequences for the future discovery of pharmaceutically important compounds. In fact, sponges of the Ircinidae family have been proposed as good models for microbiology and biochemistry research for drug discovery due to their diverse and stable microbiomes, the interesting compounds produced by the holobiont, and their suitability for captivity (Hardoim & Costa, 2014). Reductions in microbial diversity in sponge populations may also have effects on ecosystem processes if functional redundancy is not retained among remaining members of microbial communities.

Conclusions

Host genetic identity has an important role in structuring *Ircinia campana*-associated microbial communities within locations, with more genetically similar individuals hosting more similar microbiomes. These results demonstrate the close co-evolutionary relationship between hosts and their microbiomes in sponges, and highlight the importance of intraspecific variability in a host influencing the communities associated with it.

REFERENCES

Altschul SF., Gish W., Miller W., Myers EW., Lipman DJ. 1990. Basic local alignment search tool. *Journal of Molecular Biology* 215:403–410. DOI: 10.1016/S0022-2836(05)80360-2.

Antwis RE., Griffiths SM., Harrison XA., Aranega-Bou P., Arce A., Bettridge AS., Brailsford FL., de Menezes A., Devaynes A., Forbes KM., Fry EL., Goodhead I., Haskell E., Heys C., James C., Johnston SR., Lewis GR., Lewis Z., Macey MC., McCarthy A., McDonald JE., Mejia-Florez NL., O'Brien D., Orland C., Pautasso M., Reid WDK., Robinson HA., Wilson K., Sutherland WJ. 2017. Fifty important research questions in microbial ecology. *FEMS Microbiology Ecology* 93. DOI: 10.1093/femsec/fix044.

Archie EA., Tung J. 2015. Social behavior and the microbiome. *Current Opinion in Behavioral Sciences* 6:28–34. DOI: 10.1016/j.cobeha.2015.07.008.

Bahrndorff S., Alemu T., Alemneh T., Lund Nielsen J. 2016. The microbiome of animals: Implications for conservation biology. *International Journal of Genomics* 2016:1–7. DOI: 10.1155/2016/5304028.

Bálint M., Tiffin P., Hallström B., O'Hara RB., Olson MS., Fankhauser JD., Piepenbring M., Schmitt I. 2013. Host genotype shapes the foliar fungal microbiome of balsam poplar (*Populus balsamifera*). *PLoS ONE* 8:e53987. DOI: 10.1371/journal.pone.0053987.

Bankevich A., Nurk S., Antipov D., Gurevich AA., Dvorkin M., Kulikov AS., Lesin VM., Nikolenko SI., Pham S., Prjibelski AD., Pyshkin A V., Sirotkin A V., Vyahhi N., Tesler G., Alekseyev MA., Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *Journal of Computational Biology* 19:455–77. DOI: 10.1089/cmb.2012.0021.

Bayer K., Kamke J., Hentschel U. 2014. Quantification of bacterial and archaeal symbionts in high and low microbial abundance sponges using real-time PCR. *FEMS Microbiology Ecology* 89:679–690. DOI: 10.1111/1574-6941.12369.

Benson AK., Kelly SA., Legge R., Ma F., Low SJ., Kim J., Zhang M., Oh PL., Nehrenberg D., Hua K., Kachman SD., Moriyama EN., Walter J., Peterson DA., Pomp D. 2010. Individuality in gut microbiota composition is a complex polygenic trait shaped by multiple environmental and host genetic factors. *Proceedings of the National Academy of Sciences* 107:18933–8. DOI: 10.1073/pnas.1007028107.

Blanquer A., Uriz MJ., Cebrian E., Galand PE. 2016. Snapshot of a bacterial microbiome shift during the early symptoms of a massive sponge die-off in the Western Mediterranean. *Frontiers in Microbiology* 7:752. DOI: 10.3389/fmicb.2016.00752.

Blanquer A., Uriz MJ., Galand PE. 2013. Removing environmental sources of variation to gain insight on symbionts vs. transient microbes in high and low microbial abundance sponges. *Environmental Microbiology* 15:3008–3019. DOI: 10.1111/1462-2920.12261.

Blaser M., Bork P., Fraser C., Knight R., Wang J. 2013. The microbiome explored: recent insights and future challenges. *Nature Reviews Microbiology* 11:213–217. DOI:

10.1038/nrmicro2973.

Bolnick DI., Amarasekare P., Araújo MS., Bürger R., Levine JM., Novak M., Rudolf VHW., Schreiber SJ., Urban MC., Vasseur DA. 2011. Why intraspecific trait variation matters in community ecology. *Trends in Ecology & Evolution* 26:183–92. DOI: 10.1016/j.tree.2011.01.009.

Bolnick DI., Snowberg LK., Caporaso JG., Lauber C., Knight R., Stutz WE. 2014. Major Histocompatibility Complex class IIb polymorphism influences gut microbiota composition and diversity. *Molecular Ecology* 23:4831–4845. DOI: 10.1111/mec.12846.

Bright M., Bulgheresi S. 2010. A complex journey: transmission of microbial symbionts. *Nature Reviews Microbiology* 8:218–30. DOI: 10.1038/nrmicro2262.

Brooks JP., Edwards DJ., Harwich MD., Rivera MC., Fettweis JM., Serrano MG., Reris RA., Sheth NU., Huang B., Girerd P., Strauss JF., Jefferson KK., Buck GA. 2015. The truth about metagenomics: quantifying and counteracting bias in 16S rRNA studies. *BMC Microbiology* 15:66. DOI: 10.1186/s12866-015-0351-6.

Busby PE., Soman C., Wagner MR., Friesen ML., Kremer J., Bennett A., Morsy M., Eisen JA., Leach JE., Dangl JL. 2017. Research priorities for harnessing plant microbiomes in sustainable agriculture. *PLoS Biology* 15:e2001793. DOI: 10.1371/journal.pbio.2001793.

Butler M., Hunt J., Herrnkind W., Childress M., Bertelsen R., Sharp W., Matthews T., Field J., Marshall H. 1995. Cascading disturbances in Florida Bay, USA: cyanobacteria blooms, sponge mortality, and implications for juvenile spiny lobsters *Panulirus argus*. *Marine Ecology Progress Series* 129:119–125. DOI: 10.3354/meps129119.

Caporaso JG., Kuczynski J., Stombaugh J., Bittinger K., Bushman FD., Costello EK., Fierer N., Peña AG., Goodrich JK., 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* 7:335–336. DOI: 10.1038/nmeth.f.303.

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* 108:4516–22. DOI: 10.1073/pnas.1000080107.

Chapuis M-P., Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24:621–31. DOI: 10.1093/molbev/msl191.

Chiappone M., Sullivan KM. 1994. Ecological structure and dynamics of nearshore hard-bottom communities in the Florida Keys. *Bulletin of Marine Science* 54:747–756.

Chong RA., Moran NA. 2016. Intraspecific genetic variation in hosts affects regulation of obligate heritable symbionts. *Proceedings of the National Academy of Sciences* 113:13114–13119. DOI: 10.1073/pnas.1610749113.

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* 85:465–82. DOI: 10.1111/1574-6941.12135.

Costello EK., Stagaman K., Dethlefsen L., Bohannan BJM., Relman DA. 2012. The application of ecological theory toward an understanding of the human microbiome. *Science* 336:1255–1262. DOI: 10.1126/science.1224203.

Crutsinger GM., Collins MD., Fordyce JA., Gompert Z., Nice CC., Sander NJ. 2013. Plant genotypic diversity predicts community structure and governs an ecosystem process. *Science* 313:966–968. DOI: 10.1126/science.1128326.

D'Amore R., Ijaz UZ., Schirmer M., Kenny JG., Gregory R., Darby AC., Shakya M., Podar M., Quince C., Hall N. 2016. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. *BMC Genomics* 17:55. DOI: 10.1186/s12864-015-2194-9.

Degnan SM. 2015. The surprisingly complex immune gene repertoire of a simple sponge, exemplified by the NLR genes: A capacity for specificity? *Developmental & Comparative Immunology* 48:269–274. DOI: 10.1016/j.dci.2014.07.012.

Edgar RC. 2010. Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26:2460–2461. DOI: 10.1093/bioinformatics/btq461.

Ereskovsky A V., Gonobobleva E., Vishnyakov A. 2005. Morphological evidence for vertical transmission of symbiotic bacteria in the viviparous sponge *Halisarca dujardini* Johnston (Porifera, Demospongiae, Halisarcida). *Marine Biology* 146:869–875. DOI: 10.1007/s00227-004-1489-1.

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 Microbiology Ecology* 91:fiv115. DOI: 10.1093/femsec/fiv115.

Erwin PM., López-Legentil S., González-Pech R., Turon X. 2012. A specific mix of generalists: bacterial symbionts in Mediterranean *Ircinia spp*. *FEMS Microbiology Ecology* 79:619–637. DOI: 10.1111/j.1574-6941.2011.01243.x.

Esteves AIS., Cullen A., Thomas T. 2017. Competitive interactions between sponge-associated bacteria. *FEMS Microbiology Ecology* 321:fix008. DOI: 10.1093/femsec/fix008.

Ezenwa VO., Gerardo NM., Inouye DW., Medina M., Xavier JB. 2012. Animal behavior and the microbiome. *Science* 338:198–199. DOI: 10.1126/science.1227412.

Fieth RA., Gauthier M-EA., Bayes J., Green KM., Degnan SM. 2016. Ontogenetic Changes in the Bacterial Symbiont Community of the Tropical Demosponge *Amphimedon queenslandica*: Metamorphosis Is a New Beginning. *Frontiers in Marine Science* 3:1–20. DOI: 10.3389/fmars.2016.00228.

Fiore CL., Jarett JK., Lesser MP. 2013. Symbiotic prokaryotic communities from different populations of the giant barrel sponge, *Xestospongia muta*. *MicrobiologyOpen* 2:938–52. DOI: 10.1002/mbo3.135.

Flint HJ., Scott KP., Louis P., Duncan SH. 2012. The role of the gut microbiota in nutrition and

health. *Nature Reviews Gastroenterology & Hepatology* 9:577–589. DOI: 10.1038/nrgastro.2012.156.

Fourqurean JW., Robblee MB. 1999. Florida Bay: A history of recent ecological changes. *Estuaries* 22:345–357. DOI: 10.2307/1353203.

Freeman CJ., Thacker RW. 2011. Complex interactions between marine sponges and their symbiotic microbial communities. *Limnology and Oceanography* 56:1577–1586. DOI: 10.4319/lo.2011.56.5.1577.

Freeman C., Thacker R., Baker D., Fogel M. 2013. Quality or quantity: is nutrient transfer driven more by symbiont identity and productivity than by symbiont abundance? *The ISME Journal* 7:1116–1125. DOI: 10.1038/ismej.2013.7.

Garate L., Blanquer A., Uriz M. 2015. Calcareous spherules produced by intracellular symbiotic bacteria protect the sponge *Hemimycale columella* from predation better than secondary metabolites. *Marine Ecology Progress Series* 523:81–92. DOI: 10.3354/meps11196.

Gemayel R., Vinces MD., Legendre M., Verstrepen KJ. 2010. Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annual Review of Genetics* 44:445–477. DOI: 10.1146/annurev-genet-072610-155046.

Genta-Jouve G., Thomas OP. 2012. *Sponge Chemical Diversity. From Biosynthetic Pathways to Ecological Roles*. DOI: 10.1016/B978-0-12-394283-8.00004-7.

Gil R., Sabater-Muñoz B., Latorre A., Silva FJ., Moya AS. 2002. Extreme genome reduction in *Buchnera spp*.: Toward the minimal genome needed for symbiotic life. *Proceedings of the National Academy of Sciences* 99:4454–4458. DOI: 10.1073/pnas.062067299.

Gilbert JA., Jansson JK., Knight R. 2014. The Earth Microbiome project: successes and aspirations. *BMC Biology* 12:69. DOI: 10.1186/s12915-014-0069-1.

Giles EC., Kamke J., Moitinho-Silva L., Taylor MW., Hentschel U., Ravasi T., Schmitt S. 2013. Bacterial community profiles in low microbial abundance sponges. *FEMS Microbiology Ecology* 83:232–41. DOI: 10.1111/j.1574-6941.2012.01467.x.

Gloeckner V., Hentschel U., Ereskovsky A V., Schmitt S. 2012. Unique and species-specific microbial communities in Oscarella lobularis and other Mediterranean *Oscarella* species (Porifera: Homoscleromorpha). *Marine Biology* 160:781–791. DOI: 10.1007/s00227-012-2133-0.

Goodrich JK., Waters JL., Poole AC., Sutter JL., Koren O., Blekhman R., Beaumont M., Van Treuren W., Knight R., Bell JT., Spector TD., Clark AG., Ley RE. 2014. Human genetics shape the gut microbiome. *Cell* 159:789–99. DOI: 10.1016/j.cell.2014.09.053.

Griffiths SM., Harrison XA., Weldon C., Wood MD., Pretorius A., Hopkins K., Fox G., Preziosi RF., Antwis RE. In press. Genetic variability and ontogeny predict microbiome structure in a disease- challenged montane amphibian. *The ISME Journal*.

Guardiola M., Frotscher J., Uriz M-J. 2016. High genetic diversity, phenotypic plasticity, and invasive potential of a recently introduced calcareous sponge, fast spreading across the Atlanto-Mediterranean basin. *Marine Biology* 163:123. DOI: 10.1007/s00227-016-2862-6.

Hall MO., Furman BT., Merello M., Durako MJ. 2016. Recurrence of *Thalassia testudinum*

seagrass die-off in Florida Bay, USA: Initial observations. *Marine Ecology Progress Series* 560:243–249. DOI: 10.3354/meps11923.

Hardoim C., Costa R. 2014. Microbial communities and bioactive compounds in marine sponges of the family Ircíniidae—a review. *Marine Drugs* 12:5089–5122. DOI: 10.3390/md12105089.

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* 68:4431–4440. DOI: 10.1128/AEM.68.9.4431-4440.2002.

Hentschel U., Piel J., Degnan SM., Taylor MW. 2012. Genomic insights into the marine sponge microbiome. *Nature Reviews Microbiology* 10:641–654. DOI: 10.1038/nrmicro2839.

Hill MS., Hill AL. 2002. Morphological plasticity in the tropical sponge *Anthosigmella varians*: responses to predators and wave energy. *The Biological Bulletin* 202:86–95. DOI: 10.2307/1543225.

Honda K., Littman DR. 2012. The microbiome in infectious disease and inflammation. *Annual Review of Immunology* 30:759–795. DOI: 10.1146/annurev-immunol-020711-074937.

Jensen LF., Hansen MM., Mensberg K-LD., Loeschcke V. 2008. Spatially and temporally fluctuating selection at non-MHC immune genes: evidence from TAP polymorphism in populations of brown trout (*Salmo trutta*, L.). *Heredity* 100:79–91. DOI: 10.1038/sj.hdy.6801067.

Joshi NA., Fass JN. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.200) [Software]. Available at https://github.com/najoshi/sickle.

Kearney KA., Butler M., Glazer R., Kelble CR., Serafy JE., Stabenau E. 2015. Quantifying Florida Bay habitat suitability for fishes and invertebrates under climate change scenarios. *Environmental Management* 55:836–856. DOI: 10.1007/s00267-014-0336-5.

Kelman D., Kashman Y., Rosenberg E., Ilan M., Ifrach I., Loya Y. 2001. Antimicrobial activity of the reef sponge *Amphimedon viridis* from the Red Sea: evidence for selective toxicity. *Aquatic Microbial Ecology* 24:9–16. DOI: 10.3354/ame024009.

Lazzaro BP., Sceurman BK., Clark AG. 2004. Genetic basis of natural variation in *D. melanogaster* antibacterial immunity. *Science* 303:1873–1876. DOI: 10.1126/science.1092447.

Lee OO., Chui PY., Wong YH., Pawlik JR., Qian P-Y. 2009. Evidence for vertical transmission of bacterial symbionts from adult to embryo in the Caribbean sponge *Svenzea zeai. Applied and Environmental Microbiology* 75:6147–56. DOI: 10.1128/AEM.00023-09.

Lee OO., Wang Y., Yang J., Lafi FF., Al-Suwailem A., Qian P-Y. 2011. Pyrosequencing reveals highly diverse and species-specific microbial communities in sponges from the Red Sea. *The ISME Journal* 5:650–64. DOI: 10.1038/ismej.2010.165.

Li YC., Korol AB., Fahima T., Nevo E. 2004. Microsatellites within genes: Structure, function, and evolution. *Molecular Biology and Evolution* 21:991–1007. DOI: 10.1093/molbev/msh073.

Mahé F., Rognes T., Quince C., de Vargas C., Dunthorn M. 2014. Swarm: robust and fast clustering method for amplicon-based studies. *PeerJ* 2:e593. DOI: 10.7717/peerj.593.

Marino C., Pawlik J., López-Legentil S., Erwin P. 2017. Latitudinal variation in the microbiome of the sponge *Ircinia campana* correlates with host haplotype but not anti-predatory chemical defense. *Marine Ecology Progress Series* 565:53–66. DOI: 10.3354/meps12015.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet* 17:10. DOI: 10.14806/ej.17.1.200.

McCaffrey EJ., Endean R. 1985. Antimicrobial activity of tropical and subtropical sponges. *Marine Biology* 89:1–8. DOI: 10.1007/BF00392871.

Meirmans PG., Van Tiendener PH. 2004. Genotype and Genodive: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* 4:792–794. DOI: 10.1111/j.1471-8286.2004.00770.x.

Moitinho-Silva L., Díez-Vives C., Batani G., Esteves AI., Jahn MT., Thomas T. 2017. Integrated metabolism in sponge–microbe symbiosis revealed by genome-centered metatranscriptomics. *The ISME Journal* 11:1651–1666. DOI: 10.1038/ismej.2017.25.

Moran NA., McCutcheon JP., Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. *Annual Review of Genetics* 42:165–190. DOI: 10.1146/annurev.genet.41.110306.130119.

Morley SA., Berman J., Barnes DKA., de Juan Carbonell C., Downey R V., Peck LS. 2016. Extreme phenotypic plasticity in metabolic physiology of Antarctic demosponges. *Frontiers in Ecology and Evolution* 3:157. DOI: 10.3389/fevo.2015.00157.

Morrow KM., Fiore CL., Lesser MP. 2016. Environmental drivers of microbial community shifts in the giant barrel sponge, *Xestospongia muta*, over a shallow to mesophotic depth gradient. *Environmental Microbiology* 18:2025–2038. DOI: 10.1111/1462-2920.13226.

Müller WEG., Müller IM. 2003. Origin of the metazoan immune system: identification of the molecules and their functions in sponges. *Integrative and Comparative Biology* 43:281–92. DOI: 10.1093/icb/43.2.281.

Noyer C., Becerro M. 2012. Relationship between genetic, chemical, and bacterial diversity in the Atlanto-Mediterranean bath sponge *Spongia lamella*. *Hydrobiologia* 687:85–99. DOI: 10.1007/s10750-011-0884-0.

Noyer C., Thomas OP., Becerro MA. 2011. Patterns of chemical diversity in the Mediterranean sponge *Spongia lamella. PLoS ONE* 6:e20844. DOI: 10.1371/journal.pone.0020844.

Olson JB., Gao X. 2013. Characterizing the bacterial associates of three Caribbean sponges along a gradient from shallow to mesophotic depths. *FEMS Microbiology Ecology* 85:74–84. DOI: 10.1111/1574-6941.12099.

On OL., Lau SCK., Qian PY. 2006. Defense against epibiosis in the sponge *Mycale adhaerens*: Modulating the bacterial community associated with its surface. *Aquatic Microbial Ecology* 43:55–65. DOI: 10.3354/ame043055.

Peakall R., Smouse PE. 2006. genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes* 6:288–295. DOI: 10.1111/j.1471-

8286.2005.01155.x.

Peakall R., Smouse PE. 2012. GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research- an update. *Bioinformatics* 28:2537–2539. DOI: 10.1093/bioinformatics/bts460.

Pita L., Fraune S., Hentschel U. 2016. Emerging sponge models of animal-microbe symbioses. *Frontiers in Microbiology* 7:2102. DOI: 10.3389/fmicb.2016.02102.

Pita L., López-Legentil S., Erwin PM. 2013. Biogeography and host fidelity of bacterial communities in *Ircinia spp*. from the Bahamas. *Microbial Ecology* 66:437–47. DOI: 10.1007/s00248-013-0215-2.

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* 86:268–76. DOI: 10.1111/1574-6941.12159.

Puyana M., Pawlik J., Blum J., Fenical W., Puyana M., Pawlik J., Blum J., Fenical W. 2015. Metabolite variability in Caribbean sponges of the genus *Aplysina*. *Revista Brasileira de Farmacognosia* 25:592–599. DOI: 10.1016/j.bjp.2015.08.002.

Quast C., Pruesse E., Yilmaz P., Gerken J., Schweer T., Yarza P., Peplies J., Glöckner FO. 2013. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41:D590-6. DOI: 10.1093/nar/gks1219.

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at https://www.R-project.org/.

Redford KH., Segre JA., Salafsky N., Martinez del Rio C., McAloose D. 2012. Conservation and the microbiome. *Conservation Biology* 26:195–7. DOI: 10.1111/j.1523-1739.2012.01829.x.

Robblee M., Barber T., Carlson P., Durako M., Fourqurean J., Muehlstein L., Porter D., Yarbro L., Zieman R., Zieman J. 1991. Mass mortality of the tropical seagrass *Thalassia testudinum* in Florida Bay (USA). *Marine Ecology Progress Series* 71:297–299. DOI: 10.3354/meps071297.

Santucci F., Ibrahim KM., Bruzzone A., Hewit GM. 2007. Selection on MHC-linked microsatellite loci in sheep populations. *Heredity* 99:340–348. DOI: 10.1038/sj.hdy.6801006.

Schmitt S., Deines P., Behnam F., Wagner M., Taylor MW. 2011. Chloroflexi bacteria are more diverse, abundant, and similar in high than in low microbial abundance sponges. *FEMS Microbiology Ecology* 78:497–510. DOI: 10.1111/j.1574-6941.2011.01179.x.

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. *The ISME Journal* 6:564–76. DOI: 10.1038/ismej.2011.116.

Schmitt S., Weisz JB., Lindquist N., Hentschel U. 2007. Vertical transmission of a phylogenetically complex microbial consortium in the viviparous sponge *Ircinia felix*. *Applied and Environmental Microbiology* 73:2067–78. DOI: 10.1128/AEM.01944-06.

Schröder HC., Ushijima H., Krasko A., Gamulin V., Thakur NL., Diehl-Seifert B., Müller IM., Müller WEG. 2003. Emergence and disappearance of an immune molecule, an

antimicrobial lectin, in basal metazoa: A tachylectin-related protein in the sponge *Suberites domuncula*. *Journal of Biological Chemistry* 278:32810–32817. DOI: 10.1074/jbc.M304116200.

Sharp KH., Eam B., Faulkner DJ., Haygood MG. 2007. Vertical transmission of diverse microbes in the tropical sponge *Corticium* sp. *Applied and Environmental Microbiology* 73:622–9. DOI: 10.1128/AEM.01493-06.

Sipkema D., de Caralt S., Morillo JA., Al-Soud WA., Sørensen SJ., Smidt H., Uriz MJ. 2015. Similar sponge-associated bacteria can be acquired via both vertical and horizontal transmission. *Environmental Microbiology* 17:3807–3821. DOI: 10.1111/1462-2920.12827.

Stevely JM., Sweat DE., Bert TM., Sim-Smith C., Kelly M. 2010. Sponge Mortality at Marathon and Long Key , Florida : Patterns of Species Response and Population Recovery. *Proceedings of the 63rd Gulf and Caribbean Fisheries Institute* 63:384–400.

Stevens CE., Hume ID. 1998. Contributions of microbes in vertebrate gastrointestinal tract to production and conservation of nutrients. *Physiological Reviews* 78:393–427. DOI: 10.1152/physrev.1998.78.2.393.

Taylor MW., Hill RT., Piel J., Thacker RW., Hentschel U. 2007. Soaking it up: the complex lives of marine sponges and their microbial associates. *The ISME Journal* 1:187–90. DOI: 10.1038/ismej.2007.32.

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* 6:121–130. DOI: 10.1046/j.1462-2920.2003.00545.x.

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. *The ISME Journal* 7:438–43. DOI: 10.1038/ismej.2012.111.

Tellier M-AS., Bertelsen R. 2008. *Monitoring the flora and fauna of the nearshore hardbottom habitats of the Florida Keys*. F2196-05-08-F Final Report. Florida Fish and Wildlife Conservation Commission, Fish and Wildlife Research Institute, Marathon, FL.

Thacker RW., Freeman CJ. 2012. Sponge-microbe symbioses: recent advances and new directions. *Advances in Marine Biology* 62:57–111. DOI: 10.1016/B978-0-12-394283-8.00002-3.

Thakur NL., Anil A. 2000. Antibacterial activity of the sponge *Ircinia ramosa*: importance of its surface-associated bacteria. *Journal of Chemical Ecology* 26:57–71. DOI:10.1023/A:1005485310488.

Thakur N., Anil A., Müller W. 2004. Culturable epibacteria of the marine sponge *Ircinia fusca*: temporal variations and their possible role in the epibacterial defense of the host. *Aquatic Microbial Ecology* 37:295–304. DOI: 10.3354/ame037295.

Thomas TRA., Kavlekar DP., LokaBharathi PA. 2010. Marine drugs from sponge-microbe association - a review. *Marine Drugs* 8:1417–68. DOI: 10.3390/md8041417.

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., Victor Lopez J., Taylor MW., Thacker RW., Montoya JM., Hentschel U., Webster NS. 2016. Diversity, structure and convergent evolution of the global sponge microbiome. *Nature Communications* 7:11870. DOI: 10.1038/ncomms11870.

Thomas T., Rusch D., DeMaere M., Yung PY., Lewis M., Halpern A., Heidelberg KB., Egan S., Steinberg PD., Kjelleberg S. 2010. Functional genomic signatures of sponge bacteria reveal unique and shared features of symbiosis. *The ISME Journal* 4:1577–1567. DOI: 10.1038/ismej.2010.74.

Tollenaere C., Ivanova S., Duplantier J-M., Loiseau A., Rahalison L., Rahelinirina S., Brouat C. 2012. Contrasted patterns of selection on MHC-linked microsatellites in natural populations of the Malagasy plague reservoir. *PLoS ONE* 7:e32814. DOI: 10.1371/journal.pone.0032814.

Uriz MJ., Turon X. 2012. Sponge ecology in the molecular era. *Advances in Marine Biology* 61:345–410. DOI: 10.1016/B978-0-12-387787-1.00006-4.

Vacelet J., Donadey C. 1977. Electron microscope study of the association between some sponges and bacteria. *Journal of Experimental Marine Biology and Ecology* 30:301–314. DOI: 10.1016/0022-0981(77)90038-7.

Vandenkoornhuyse P., Quaiser A., Duhamel M., Le Van A., Dufresne A. 2015. The importance of the microbiome of the plant holobiont. *New Phytologist* 206:1196–1206. DOI: 10.1111/nph.13312.

Wagner MR., Lundberg DS., del Rio TG., Tringe SG., Dangl JL., Mitchell-Olds T. 2016. Host genotype and age shape the leaf and root microbiomes of a wild perennial plant. *Nature Communications* 7:12151. DOI: 10.1038/ncomms12151.

Webster NS., Reusch TBH. 2017. Microbial contributions to the persistence of coral reefs. *The ISME Journal*. DOI: 10.1038/ismej.2017.66.

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* 12:2070–82. DOI: 10.1111/j.1462-2920.2009.02065.x.

Webster NS., Thomas T. 2016. The sponge hologenome. *mBio* 7:e00135-16. DOI: 10.1128/mBio.00135-16.

Wehrl M., Steinert M., Hentschel U. 2007. Bacterial uptake by the marine sponge *Aplysina aerophoba*. *Microbial Ecology* 53:355–365. DOI: 10.1007/s00248-006-9090-4.

Weigel BL., Erwin PM. 2015. Intraspecific variation in microbial symbiont communities of the sun sponge, *Hymeniacidon heliophila*, from intertidal and subtidal habitats. *Applied and Environmental Microbiology* 82:650–8. DOI: 10.1128/AEM.02980-15.

Whitham TG., Gehring C a., Lamit LJ., Wojtowicz T., Evans LM., Keith AR., Smith DS. 2012. Community specificity: life and afterlife effects of genes. *Trends in Plant Science* 17:271–81. DOI: 10.1016/j.tplants.2012.01.005.

Wickham H. (2009) ggplot2: Elegant graphics for data analysis. Springer-Verlag New York.

Wiens M., Korzhev M., Krasko A., Thakur NL., Perović-Ottstadt S., Breter HJ., Ushijima H., Diehl-Seifert B., Müller IM., Müller WEG. 2005. Innate immune defense of the sponge *Suberites domuncula* against bacteria involves a MyD88-dependent signaling pathway. Induction of a perforin-like molecule. *The Journal of Biological Chemistry* 280:27949–59. DOI: 10.1074/jbc.M504049200.

Wiens M., Korzhev M., Perović-Ottstadt S., Luthringer B., Brandt D., Klein S., Müller WEG. 2007. Toll-like receptors are part of the innate immune defense system of sponges (Demospongiae: Porifera). *Molecular Biology and Evolution* 24:792–804. DOI: 10.1093/molbev/msl208.

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 B: Biological Sciences* 220:519–528. DOI: 10.1098/rspb.1984.0018.

Wright RM., Kenkel CD., Dunn CE., Shilling EN., Bay LK., Matz M V. 2017. Intraspecific differences in molecular stress responses and coral pathobiome contribute to mortality under bacterial challenge in *Acropora millepora*. *Scientific Reports* 7:2609. DOI: 10.1038/s41598-017-02685-1.

Zhao L., Wang G., Siegel P., He C., Wang H., Zhao W., Zhai Z., Tian F., Zhao J., Zhang H., Sun Z., Chen W., Zhang Y., Meng H. 2013. Quantitative genetic background of the host influences gut microbiomes in chickens. *Scientific Reports* 3:1163. DOI: 10.1038/srep01163.

Zoetendal EG., Akkermans ADL., Akkermans-van Vliet WM., de Visser JAGM., de Vos WM. 2001. The host genotype affects the bacterial community in the human gastronintestinal tract. *Microbial Ecology in Health and Disease* 13:129–134. DOI: 10.1080/089106001750462669.

Zytynska SE., Fay MF., Penney D., Preziosi RF. 2011. Genetic variation in a tropical tree species influences the associated epiphytic plant and invertebrate communities in a complex forest ecosystem. *Philosophical Transactions of the Royal Society B: Biological sciences* 366:1329–36. DOI: 10.1098/rstb.2010.0183.

Appendix I: Core operational taxonomic units (OTUs) for *Ircinia campana*

k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__BPC015; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__BPC015; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__BPC015; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__BPC015; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__BPC015; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__BPC015; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__BPC015; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__BPC015; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__BPC015; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__BPC015; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__

k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Acidobacteria-6; o__iii1-15; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Solibacteres; o__Solibacterales; f__PAUC26f; g__; s__
k__Bacteria; p__Acidobacteria; c__Solibacteres; o__Solibacterales; f__PAUC26f; g__; s__
k__Bacteria; p__Acidobacteria; c__Solibacteres; o__Solibacterales; f__PAUC26f; g__; s__
k__Bacteria; p__Acidobacteria; c__Solibacteres; o__Solibacterales; f__PAUC26f; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Acidobacteria; c__Sva0725; o__Sva0725; f__; g__; s__
k__Bacteria; p__Actinobacteria; c__Acidimicrobiia; o__Acidimicrobiales;
k__Bacteria; p__Actinobacteria; c__Acidimicrobiia; o__Acidimicrobiales;
k__Bacteria; p__Actinobacteria; c__Acidimicrobiia; o__Acidimicrobiales; f__; g__; s__
k__Bacteria; p__Actinobacteria; c__Acidimicrobiia; o__Acidimicrobiales; f__; g__; s__
k__Bacteria; p__Actinobacteria; c__Acidimicrobiia; o__Acidimicrobiales; f__TK06; g__; s__
k__Bacteria; p__Actinobacteria; c__Acidimicrobiia; o__Acidimicrobiales; f__TK06; g__; s__
k__Bacteria; p__Actinobacteria; c__Acidimicrobiia; o__Acidimicrobiales; f__TK06; g__; s__
k__Bacteria; p__Actinobacteria; c__Acidimicrobiia; o__Acidimicrobiales; f__TK06; g__; s__
k__Bacteria; p__Actinobacteria; c__Acidimicrobiia; o__Acidimicrobiales; f__TK06; g__; s__
k__Bacteria; p__Bacteroidetes; c__[Rhodothermi]; o__[Rhodothermales]; f__Rhodothermaceae; g__; s__
k__Bacteria; p__Bacteroidetes; c__[Rhodothermi]; o__[Rhodothermales]; f__Rhodothermaceae; g__; s__
k__Bacteria; p__Bacteroidetes; c__[Rhodothermi]; o__[Rhodothermales]; f__Rhodothermaceae; g__; s__
k__Bacteria; p__Bacteroidetes; c__[Rhodothermi]; o__[Rhodothermales]; f__Rhodothermaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__

k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae;
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae;
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae;
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__

k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__Caldilineales; f__Caldilineaceae; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Anaerolineae; o__SBR1031; f__A4b; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__Ktedonobacteria; o__TK10; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__

k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__mle1-48; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__

k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__Chloroflexi; c__TK17; o__TK18; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__PAUC34f; c__; o__; f__; g__; s__
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;

k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae; g__;Rhodovulum;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodobacterales; f__Rhodobacteraceae; g__;Rhodovulum; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;

k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae;
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Alphaproteobacteria; o__Rhodospirillales; f__Rhodospirillaceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__

k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__[Entotheonellales]; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Deltaproteobacteria; o__Syntrophobacterales; f__Syntrophobacteraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales;

k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Chromatiales; f__Ectothiorhodospiraceae; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089;
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__HTCC2188; f__HTCC2089; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__

```
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__Proteobacteria; c__Gammaproteobacteria; o__Thiohalorhabdales; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria; p__SBR1093; c__EC214; o__; f__; g__; s__
k__Bacteria;
k__Bacteria;
k__Bacteria;
```

Chapter 6: General Discussion

Ecosystems worldwide are suffering the impacts of an expanding and technologically advancing human population, including loss of habitat, declining biodiversity, reduced ecological functionality and loss of ecosystem services (Vitousek et al., 1997; Dirzo & Raven, 2003). Understanding species ecology is essential for conservation of populations and ecosystems, and molecular ecology and conservation genetics methods can be extremely useful in pursuit of this goal. In light of this, the research presented in this thesis aimed to develop molecular tools for an important marine group - sponges - and to then use them to study various aspects of ecology and biodiversity. Specifically, I aimed to develop microsatellite markers and investigate the population genetic structure of two ecologically important and vulnerable sponge species in the Greater Caribbean (*Ircinia campana* and *Spheciospongia vesparium*), and investigate links between the genotype of *I. campana* and the composition of its microbiome.

The first aspect of this thesis involved the development of molecular resources to aid the ecological study of sponges. In Chapter 2, I developed a bioinformatics pipeline for microsatellite isolation and PCR primer design. Microsatellites are species- or genus-specific genetic markers that are popular with ecologists and conservationists for studying important aspects of species' biology, including genetic diversity, population genetic structure, migration rates and relatedness (Selkoe & Toonen, 2006). I described a comprehensive pipeline for processing raw Illumina paired-end sequence data to finish with a list of microsatellite loci primers tailored to the user, set within the 'Galaxy' web-based bioinformatics environment (Giardine et al., 2005; Goecks et al., 2010;

Blankenberg et al., 2010). The pipeline brought together existing programs (FastQC [Andrews 2014], Trimmomatic [Bolger, Lohse & Usadel, 2014], Pal_finder [Castoe, Poole & Koning, 2012], Primer3 [Koressaar & Remm, 2007; Untergasser et al., 2012]) in addition to a new program (Pal_filter, [Griffiths et al., 2016]) to simplify their use within a single online tool. This was used in Chapter 3 and Chapter 4 to develop ten and twelve markers respectively for the sponges *I. campana* and *S. vesparium*; these are the first microsatellites to be isolated for these species. The open-access, easy-to-operate, web-based resource created as part of this work increases the accessibility of bioinformatics tools to ecologists and conservation biologists who may have limited bioinformatics expertise to hand.

In the second aspect of this thesis, I used the microsatellite markers that I developed to investigate the ecology of *I. campana* and *S. vesparium*. In Chapter 3 and Chapter 4, I investigated the population structure of both species in the Greater Caribbean region. Due to differences in habitat use and local distribution, the species were only sampled sympatrically in six Florida Bay sites. Both species showed strong population structure, consistent with other sponge species in the Greater Caribbean and in other regions (global F_{ST} values were 0.085 for *S. vesparium* and 0.126 for *I. campana*). Genetic isolation by distance was observed in both species, indicating the presence of distance-limited dispersal. I also observed the effects of oceanographic features acting as barriers to dispersal in both species, such as the Gulf of Honduras gyre and the Florida Current, as well as more local-scale barriers. In some instances, patterns were observed that could not be ascribed to ocean currents or distance. For example, in *I. campana*, the two sites sampled in Martinique showed genetic differentiation levels higher than would be expected considering the short (15 km) oceanographic distance between them. In both species, I also found physical structures (reef and landmass) acting as barriers to gene flow beyond their effects on increasing oceanographic distances, suggesting they affect water movement and subsequent larval dispersal patterns. There was a general pattern of lower F_{ST} and D values for *S. vesparium* compared to *I. campana* in sites where they were both sampled within Florida. This indicates

that there is less population differentiation in *S. vesparium*, and higher levels of gene flow, and may be because the former is thought to be oviparous and the latter viviparous (based on the life history of other species in each family; Maldonado & Riesgo, 2008); the dispersal of the egg before it becomes a larva may therefore provide additional connectivity in the species. These results give insight into the population dynamics of these species, and can be used to guide restoration efforts of the species in Florida Bay (see below).

In Chapter 5, I showed that genetic variation in *I. campana* hosts and the variation in the composition of their microbial communities are linked. More genetically similar sponges (measured by multilocus microsatellite Euclidian distance) hosted microbiomes that were more similar in terms of their composition (measured by Bray-Curtis dissimilarity and the Jaccard index). This relationship was robust; two Florida Bay sites were sampled separately and the same relationship was found in both (Mantel tests: Bray-Curtis: Long Key $r = 0.408$, $p = 0.003$; KC $r = 0.361$, $p = 0.042$; Jaccard: Long Key $r = 0.407$, $p = 0.003$, Kemp Channel $r = 0.509$, $p = 0.009$). The cause of this relationship is unclear, and could either be due to vertical transmission of microbial communities, (which has been found in other sponges including the congener *Ircinia felix* [Schmitt et al., 2007]), or could be due to linkage between microsatellite genes and genes for functional traits. In this manner, microsatellite variability could correlate to variability in phenotypic traits, including innate immune system features, or the secondary metabolite profile, which could affect the microbial communities able to live in the sponge. This study helps to shed light on the sponge-microbiome relationship, and indicates the important of intraspecific genetic diversity in influencing associated communities.

Evaluation of methods

As highlighted in the thesis introduction, Chapter 1, there is much value in using molecular techniques in ecological research. The work presented here utilised various technologies.

In this thesis, I use microsatellite markers in the study of population genetic structure and genetic diversity. For these studies, microsatellites were chosen for a number of reasons: Firstly, they have relatively simple sample preservation requirements, which was essential for the fieldwork necessary to carry out this research. Samples were often collected in remote areas without sufficient scientific infrastructure to allow more sophisticated preservation methods to be used (for example, liquid nitrogen or freezing at -80°C). Secondly, polymorphism levels in microsatellites are generally high, unlike at some mitochondrial genes, for example, which show unusually slow mutation rates in sponges. High polymorphism was essential for the purposes of these studies – to study population structure on an ecologically relevant timescale and small spatial scales (Chapters 3 and 4), and to characterise groups of individuals with sufficient diversity to study the effects of genetic variation (Chapter 5).

There is a common misconception that some molecular markers are inherently 'better' than others (Karl et al., 2012), and new techniques outperform older established techniques – however, this depends on the question that a researcher is trying to address. Although genomic techniques (single nucleotide polymorphisms [SNPs], restriction site-associated DNA sequencing [RADSeq]) are increasing in popularity, there is still an important place for microsatellites in conservation genetics and population genetics research. In sponges, the sizable microbial communities in their tissue can be problematic for using non-species-specific markers, as variation detected may be from microbial symbionts instead of the sponge host. Additionally, genomic techniques are still relatively expensive, and the bioinformatics expertise and computational power required may be prohibitive to some research groups.

Nevertheless, there were some limitations encountered in the use of microsatellites in this study. In Chapter 4, problems were encountered with linkage disequilibrium in the microsatellite markers developed for *Ircinia campana*, meaning that one locus had to be dropped from analysis. Null alleles were also a common occurrence in many loci for both *I. campana* and *S. vesparium*, a phenomenon that appears to be common in sponges (Dailianis et

al., 2011; Chaves-Fonnegra et al., 2015; Pérez-Portela, Noyer & Becerro, 2015; Guardiola, Frotscher & Uriz, 2016; Richards et al., 2016). Null alleles can be problematic, as they falsely amplify population differences, however methods exist to correct null allele presence, implemented in programs including FreeNA (Chapuis & Estoup, 2007), Inest (Chybicki & Burczyk, 2009) and Geneland (Guillot, Santos & Estoup, 2008). Deviations from Hardy-Weinberg equilibrium (HWE) were also a common theme. These can be caused by null alleles, or by numerous biological effects that are likely to be found in natural populations (migration, drift, mutation and natural selection). Many researchers choose to discard loci out of HWE, however, this can be an overly conservative and unnecessary waste of loci, and is also unnecessary with the use of methods that do not have HWE assumptions, such as multivariate methods (Jombart, Pontier & Dufour, 2009).

Illumina sequencing, currently one of the more popular NGS platforms, was used in this thesis to develop microsatellite markers for two sponge species, and to characterise microbial communities associated with *I. campana*. For microsatellite marker isolation, NGS methods offer many advantages over traditional methods, including a cheaper cost-per-species when multiple species' DNA are run in the same Illumina lane, and thousands more markers from which to test and choose. For characterising microbial communities, NGS methods have been revolutionary. The majority of microbes, including in sponges, are not possible to cultivate and therefore culture-independent methods are essential for describing microbial diversity. However, any method that involves a PCR amplification step will have its own bias, and the sequencer used also affects the composition and diversity captured (D'Amore et al., 2016). The V4 region of the 16S gene was amplified in the study in Chapter 6, however, there is evidence that in sponges this does not capture as much diversity as a combined, multi-primer approach (Zhang et al., unpublished). Nonetheless, the aim of the study in Chapter 5 was to explore the relationship between host genetic identity and associated microbial community composition within a species, and as each sample was analysed consistently, any platform- or PCR-based bias will not affect the conclusions of this study.

Restoration implications for *Ircinia campana* and *Spheciospongia vesparium*

The population genetics approach taken in Chapter 3 and Chapter 4 gives valuable insight to guide sponge restoration in Florida Bay. The restoration of sponges in Florida Bay has so far shown good success: The transplants and cuttings have good survival and growth (pers. comm. M. Butler and D. Behringer), and there is evidence of the return of some of their functional roles in the ecosystem (Butler, Stanley & Butler, 2016). Despite the simplicity of the restoration techniques utilised, they involve dedication of significant resources – boats, fuel, manpower and time. In addition, the cyanobacterial blooms that cause sponge mortalities in Florida Bay are not a single event for which we need to correct without further intervention – the blooms have returned many times since their original occurrence in 1991 (Butler et al., 1995; Paul et al., 2005; Berry et al., 2015). For these reasons, we must use the tools at our disposal to ensure that restoration is conducted with the highest levels of efficacy possible, and with the maximum chance of success. The information that has been gained through the genetics research presented in this thesis will allow practitioners valuable insight into optimising their restoration strategy.

Sites sampled in Florida Bay (along with Gray's Reef, Georgia for *I. campana*) were genetically distinct from other populations in the Greater Caribbean. In both species, the Bayesian Geneland clustering assignment program grouped the USA sites into separate population clusters in both species. Furthermore the first axes of the Principle Coordinates Analyses in both species split the USA sites from the rest of the sites sampled (the only exception being one Martinique site for *I. campana*). These results emphasise and support the need for ongoing sponge restoration in Florida Bay: Firstly, they hold unique genetic diversity, which is an important component of biodiversity often overlooked in conservation planning and implementation (Laikre et al., 2010; Laikre, 2010), but identified as a conservation priority; conserving genetic diversity is one of the Convention on Biological Diversity's Aichi Biodiversity Targets (Target 13). A species' genetic diversity can encode variation in responses to stressors, forming the basis of adaptive capacity (along with epigenetic processes and

plasticity). Secondly, the indicated lack of connectivity between Floridian sites and other sites in the Caribbean increases the vulnerability of Florida Bay to local extinctions. If sponge larvae cannot be recruited from other areas, as a result of distance or oceanographic barriers, Florida is in danger of losing its sponge population if mortalities continue to recur in the area. Local extinctions could have catastrophic effects on the Bay, as shown in the ecosystem changes in sites where sponge populations dramatically declined as a result of the blooms (Butler et al., 1995; Herrnkind et al., 1997; Peterson et al., 2006). The Floridian sites for *I. campana* did show connectivity with Gray's Reef in Georgia, however, the upstream position of Florida in comparison to Georgia in the Florida Current indicates that there is unlikely to be gene flow in the Georgia-Florida direction, so this cannot be relied on as source of new recruits.

Genetic diversity levels were generally not observed to be lower in Florida than in other populations, but sites sampled there did tend towards higher inbreeding coefficients and at some sites, evidence of bottlenecks. Although higher inbreeding coefficients are a common occurrence in sponge populations (Duran et al., 2004; Dailianis et al., 2011; Bell et al., 2014; Pérez-Portela, Noyer & Becerro, 2015; Giles et al., 2015; Guardiola, Frotscher & Uriz, 2016; Riesgo et al., 2016), and inbreeding could not be statistically confirmed as the source of positive F_{IS} values in *I. campana*, this is concerning given that inbreeding depression is arguably the most dangerous genetic consequence of population sizes reducing. This occurs when deleterious recessive alleles are unable to be purged from the inbreeding population. Genetic diversity in populations is important for adaptability to future stressors, such as ocean warming, or extreme cold water events (such an event caused sponge mortalities in Florida in 2010; [Colella et al., 2012]) as it may encode resistance. However, diversity at genetic markers such as microsatellites does not necessarily reflect genome-wide levels of diversity (Väli et al., 2008), or diversity at functionally-relevant loci. Therefore although future studies may benefit from the explicit linking of diversity at functional genes to diversity in an adaptive trait through genomics approaches, the strength of information currently available suggests that for effective restoration, practitioners should plan for increased genetic diversity

(Harrisson et al., 2014). The results of this study suggest that genetic diversity should be maintained when selecting cuttings to ensure that sites are not repopulated by clones, causing genetic homogeneity. The results of Chapter 6 also support maintaining high genetic diversity levels in the restored sponge populations. In this chapter, I found that in *I. campana*, more genetically similar host sponges have more similar associated microbial communities. If we extrapolate this finding, more genetically diverse sponge populations may support a higher diversity in their associated microbial communities, at least in terms of beta diversity.

The dispersal barriers identified in Chapters 3 and 4 indicate how best to focus restoration efforts. By understanding the factors that can limit natural dispersal, we can make predictions about which mortality-affected sites in the Keys are less likely to experience natural recruitment, and thus are in more need of restoration attention. Isolation by distance (IBD) was found in *S. vesparium* within Florida and at the Greater Caribbean level, whereas for *I. campana*, it was found at the Greater Caribbean level and within the Lesser Antilles but not within Florida. The lack of IBD in Florida in *I. campana* could however be because there was less geographic range across the sampling sites (this species could not be found in the Upper Keys sites sampled). An IBD pattern indicates distance-limited dispersal, which is unsurprising given the short larval duration found in sponges (Maldonado & Young, 1999; Bergquist & Sinclair, 2010; Wang et al., 2012). Therefore sites experiencing mortalities that are not in close proximity to healthy sites would be important candidates for restoration. It is also important to note that locations on the Atlantic side of the Keys in close proximity to mortality-affected Bay side sites may not be well connected, as observed in the case of *S. vesparium* in Chapter 3.

I also found evidence of oceanographic barriers to connectivity, both within Florida and at other locations in the Greater Caribbean. In *I. campana*, the Marquesas Keys were separated from the other Florida sites in the Discriminant Analysis of Principle Components (DACP) of the Florida Keys sites. This pattern of genetic structure may be explained by the deep channel and currents

between Key West and the Lakes Passage (personal communication M. Butler), which could be forming a barrier to dispersal of larvae to and from the Lakes Passage. I also found an indication of a combined effect of oceanographic and structural features in the DACP of all *I. campana* sites, with the Sapodilla Cayes forming a separate cluster to all other sites. This site is subject to the anti-clockwise gyre of the Gulf of Honduras, as well as an unusual hook-shaped reef structure (Figure 1). Together, I hypothesise that these features promote very high levels of larval retention, as well as preventing recruitment of larvae from other sites. In *S. vesparium*, the landmasses of the Keys themselves also appeared to alter dispersal, disrupting isolation by distance patterns so that the Atlantic side sites in the Middle Keys were less related to Bay side Middle Keys sites than Bay side Lower Keys sites, despite shorter oceanographic distances. In light of these barriers to connectivity, local hydrology and bathymetric features of Florida Bay should be particularly well studied to find areas characterised by potential barriers, and these should receive restoration priority.

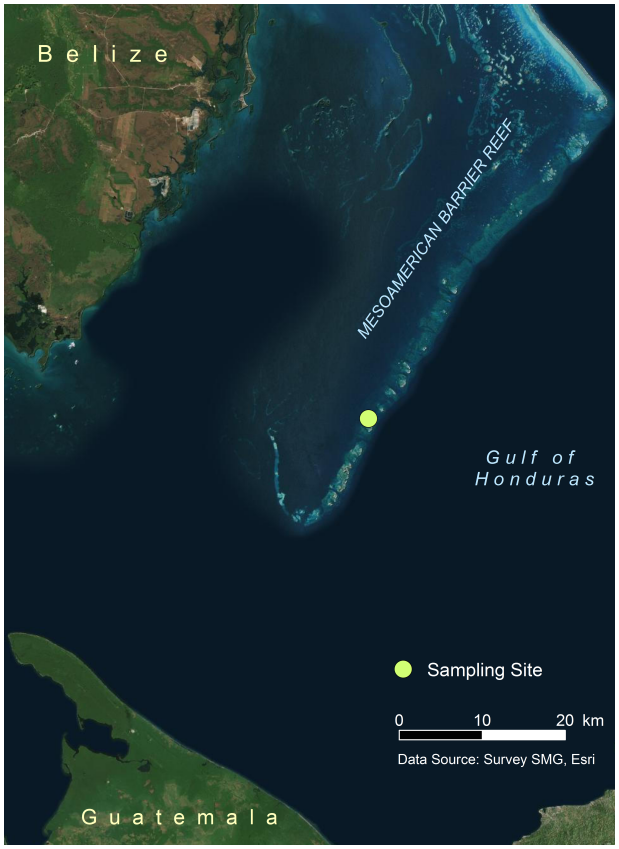


Figure 1: Map showing the location and reef shape of the Sapodilla Cayes in Belize. Credit: F. Clever.

Thesis conclusions and future directions

Although the utility of genetic approaches in conservation is well accepted, in practice they are often not used due to barriers in funding and access to expertise (Taylor, Dussex & van Heezik, 2017). The gap between research and practice is not unique to conservation genetics, with a 'research-implementation gap' widely acknowledged in conservation biology as a whole (Salafsky et al., 2002; Knight et al., 2008; Sunderland et al., 2009; Gossa, Fisher & Milner-Gulland, 2015). In this thesis, I used genetic techniques to research sponge ecology and to advance conservation of this important benthic group. I created a microsatellite development online tool to aid characterisation of these useful and popular markers for researchers with little bioinformatics experience, and used the method to develop markers for two Tropical Western Atlantic sponge species: *Ircinia campana* and *Spheciospongia vesparium*. With these markers, I investigated the population genetic structure and diversity of these species in the Greater Caribbean, and identified likely oceanographic, structural and life history barriers to dispersal. I then applied these population genetics results to make recommendations on restoration of these species in Florida Bay, where mass mortalities in sponge communities have occurred with great cost to the ecosystem. Finally, I investigated how genotype in one of these species, *Ircinia campana*, affects their associated microbial communities, highlighting an ecological effect of genetic diversity that is often overlooked.

Many avenues of research could be taken to expand upon these findings. It is important to combine conservation genetics work with an understanding of species demography and life history (Lande, 1988), and therefore research on the life histories of *S. vesparium* and *I. campana*, such as their pelagic larval durations, larval behaviours and physiologies, and settlement cues, could provide valuable conservation insight into their population dynamics. This would add further ecological information that could be used to aid restoration, including in approaches such as biophysical modelling of connectivity. It would also be valuable to examine the population structure of these species over time. Temporal variability in recruitment and ocean currents could affect population

structure; assessing the stability of connectivity patterns would be advantageous for planning restoration for population resilience in the future. Genomic or transcriptomic approaches could also be utilised to consider mechanisms of resilience and adaptation to the mortality-inducing cyanobacterial blooms. It is also important to carry out further research into the cyanobacterial blooms and their causes, the mechanism of bloom-associated sponge death, and how they can be prevented.

To advance our understanding of the interaction between host genetics and associated microbial communities in sponges, future work could involve a functional approach to determine if any of the differences in microbiome composition correlated with genetic variation is functional in nature. Finding the cause of the genotype-microbiome relationship (vertical transmission versus phenotype-based selection) would be essential in further understanding the nature of the relationship between the sponge host and its microbiome.

REFERENCES

Andrews S (2014) FastQC: a quality control tool for high throughput sequence data [Online]. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/. Accessed 7 April 2016.

Bell JJ., Smith D., Hannan D., Haris A., Jompa J., Thomas L. 2014. Resilience to disturbance despite limited dispersal and self-recruitment in tropical barrel sponges: implications for conservation and management. *PLoS ONE* 9:e91635. DOI: 10.1371/journal.pone.0091635.

Bergquist PR., Sinclair ME. 2010. The morphology and behaviour of larvae of some intertidal sponges. *New Zealand Journal of Marine and Freshwater Research* 2:426–437. DOI: 10.1080/00288330.1968.9515247.

Berry DL., Goleski JA., Koch F., Wall CC., Peterson BJ., Anderson OR., Gobler CJ. 2015. Shifts in cyanobacterial strain dominance during the onset of harmful algal blooms in Florida Bay, USA. *Microbial Ecology* 70:361–371. DOI: 10.1007/s00248-014-0564-5.

Blankenberg D., Von Kuster G., Coraor N., Ananda G., Lazarus R., Mangan M., Nekrutenko A., Taylor J. 2010. Galaxy: A web-based genome analysis tool for experimentalists. *Current Protocols in Molecular Biology* 89:19.10:19.10.1-19.10.21. DOI: 10.1002/0471142727.mb1910s89.

Bolger AM., Lohse M., Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. DOI: 10.1093/bioinformatics/btu170.

Butler M., Hunt J., Herrnkind W., Childress M., Bertelsen R., Sharp W., Matthews T., Field J., Marshall H. 1995. Cascading disturbances in Florida Bay, USA: cyanobacteria blooms, sponge mortality, and implications for juvenile spiny lobsters *Panulirus argus*. *Marine Ecology Progress Series* 129:119–125. DOI: 10.3354/meps129119.

Butler J., Stanley JA., Butler MJ. 2016. Underwater soundscapes in near-shore tropical habitats and the effects of environmental degradation and habitat restoration. *Journal of Experimental Marine Biology and Ecology* 479:89–96. DOI: 10.1016/j.jembe.2016.03.006.

Castoe T., Poole A., de Koning A., Jones K., Tomback D., Oyler-McCance S., Fike J., Lance S., Streicher J., Smith E., Pollock D. 2012. Rapid microsatellite identification from Illumina paired-end genomic sequencing in two birds and a snake. *PLoS ONE* 7:e30953. DOI: 10.1371/journal.pone.0030953.

Chapuis M-P., Estoup A. 2007. Microsatellite null alleles and estimation of population differentiation. *Molecular Biology and Evolution* 24:621–31. DOI: 10.1093/molbev/msl191.

Chaves-Fonnegra A., Feldheim KA., Secord J., Lopez J V. 2015. Population structure and dispersal of the coral-excavating sponge *Cliona delitrix*. *Molecular Ecology* 24:1447–66. DOI: 10.1111/mec.13134.

Chybicki IJ., Burczyk J. 2009. Simultaneous estimation of null alleles and inbreeding coefficients. *Journal of Heredity* 100:106–113. DOI: 10.1093/jhered/esn088.

Colella MA., Ruzicka RR., Kidney JA., Morrison JM., Brinkhuis VB. 2012. Cold-water event of

January 2010 results in catastrophic benthic mortality on patch reefs in the Florida Keys. *Coral Reefs* 31:621–632. DOI: 10.1007/s00338-012-0880-5.

D'Amore R., Ijaz UZ., Schirmer M., Kenny JG., Gregory R., Darby AC., Shakya M., Podar M., Quince C., Hall N. 2016. A comprehensive benchmarking study of protocols and sequencing platforms for 16S rRNA community profiling. *BMC Genomics* 17:55. DOI: 10.1186/s12864-015-2194-9.

Dailianis T., Tsigenopoulos CS., Dounas C., Voultsiadou E. 2011. Genetic diversity of the imperilled bath sponge *Spongia officinalis* Linnaeus, 1759 across the Mediterranean Sea: patterns of population differentiation and implications for taxonomy and conservation. *Molecular Ecology* 20:3757–72. DOI: 10.1111/j.1365-294X.2011.05222.x.

Dirzo R., Raven PH. 2003. Global state of biodiversity loss. *Annual Review of Environment and Resources* 28:137–167. DOI: 10.1146/annurev.energy.28.050302.105532.

Duran S., Pascual M., Estoup A., Turon X. 2004. Strong population structure in the marine sponge *Crambe crambe* (Poecilosclerida) as revealed by microsatellite markers. *Molecular Ecology* 13:511–522. DOI: 10.1046/j.1365-294X.2004.02080.x.

Giardine B., Riemer C., Hardison RC., Burhans R., Elnitski L., Shah P., Zhang Y., Blankenberg D., Albert I., Taylor J., Miller W., Kent WJ., Nekrutenko A. 2005. Galaxy: a platform for interactive large-scale genome analysis. *Genome Research* 15:1451–5. DOI: 10.1101/gr.4086505.

Giles EC., Saenz-Agudelo P., Hussey NE., Ravasi T., Berumen ML. 2015. Exploring seascape genetics and kinship in the reef sponge *Stylissa carteri* in the Red Sea. *Ecology and Evolution* 5:2487–502. DOI: 10.1002/ece3.1511.

Goecks J., Nekrutenko A., Taylor J., The Galaxy Team. 2010. Galaxy: a comprehensive approach for supporting accessible, reproducible, and transparent computational research in the life sciences. *Genome Biology* 11:R86. DOI: 10.1186/gb-2010-11-8-r86.

Gossa C., Fisher M., Milner-Gulland EJ. 2015. The research–implementation gap: how practitioners and researchers from developing countries perceive the role of peer-reviewed literature in conservation science. *Oryx* 49:80–87. DOI: 10.1017/S0030605313001634.

Griffiths SM., Fox G., Briggs PJ., Donaldson IJ., Hood S., Richardson P., Leaver GW., Truelove NK., Preziosi RF. 2016. A Galaxy-based bioinformatics pipeline for optimised, streamlined microsatellite development from Illumina next-generation sequencing data. *Conservation Genetics Resources* 8:481–486. DOI: 10.1007/s12686-016-0570-7.

Guardiola M., Frotscher J., Uriz M-J. 2016. High genetic diversity, phenotypic plasticity, and invasive potential of a recently introduced calcareous sponge, fast spreading across the Atlanto-Mediterranean basin. *Marine Biology* 163:123. DOI: 10.1007/s00227-016-2862-6.

Guillot G., Santos F., Estoup A. 2008. Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. *Bioinformatics* 24:1406–1407. DOI: 10.1093/bioinformatics/btn136.

Harrisson KA., Pavlova A., Telonis-Scott M., Sunnucks P. 2014. Using genomics to characterize evolutionary potential for conservation of wild populations. *Evolutionary Applications*

7:1008–1025. DOI: 10.1111/eva.12149.

Herrnkind WF., Butler IV MJ., Hunt JH., Childress M. 1997. Role of physical refugia: implications from a mass sponge die-off in a lobster nursery in Florida. *Marine and Freshwater Research* 48:759. DOI: 10.1071/MF97193.

Jombart T., Pontier D., Dufour A-B. 2009. Genetic markers in the playground of multivariate analysis. *Heredity* 102:330–341. DOI: 10.1038/hdy.2008.130.

Karl SA., Toonen RJ., Grant WS., Bowen BW. 2012. Common misconceptions in molecular ecology: Echoes of the modern synthesis. *Molecular Ecology* 21:4171–4189. DOI: 10.1111/j.1365-294X.2012.05576.x.

Knight AT., Cowling RM., Rouget M., Balmford A., Lombard AT., Campbell BM. 2008. Knowing but not doing: Selecting priority conservation areas and the research-implementation gap. *Conservation Biology* 22:610–617. DOI: 10.1111/j.1523-1739.2008.00914.x.

Koressaar T., Remm M. 2007. Enhancements and modifications of primer design program Primer3. *Bioinformatics* 23:1289–91. DOI: 10.1093/bioinformatics/btm091.

Laikre L. 2010. Genetic diversity is overlooked in international conservation policy implementation. *Conservation Genetics* 11:349–354. DOI: 10.1007/s10592-009-0037-4.

Laikre L., Allendorf FW., Aroner LC., Baker CS., Gregovich DP., Hansen MM., Jackson JA., Kendall KC., McKelvey K., Neel MC., Olivieri I., Ryman N., Schwartz MK., Bull RS., Stetz JB., Tallmon DA., Taylor BL., Vojta CD., Waller DM., Waples RS. 2010. Neglect of genetic diversity in implementation of the Convention of Biological Diversity. *Conservation Biology* 24:86–88. DOI: 10.2307/40419633.

Lande R. 1988. Genetics and demography in biological conservation. *Science* 241:1455–1460. DOI: 10.1126/science.3420403.

Maldonado M., Riesgo A. 2008. Reproduction in Porifera: a synoptic overview. *Treballs de la SCB* 59:29–49. DOI: 10.2436/20.1501.02.56.

Maldonado M., Young CM. 1999. Effects of the duration of larval life on postlarval stages of the demosponge *Sigmadocia caerulea*. *Journal of Experimental Marine Biology and Ecology* 232:9–21. DOI: 10.1016/S0022-0981(98)00076-8.

Paul VJ., Thacker RW., Banks K., Golubic S. 2005. Benthic cyanobacterial bloom impacts the reefs of South Florida. *Coral Reefs* 24:693–697. DOI: 10.1007/s00338-005-0061-x.

Pérez-Portela R., Noyer C., Becerro MA. 2015. Genetic structure and diversity of the endangered bath sponge *Spongia lamella*. *Aquatic Conservation: Marine and Freshwater Ecosystems* 25:365–379. DOI: 10.1002/aqc.2423.

Peterson B., Chester C., Jochem F., Fourqurean JW. 2006. Potential role of sponge communities in controlling phytoplankton blooms in Florida Bay. *Marine Ecology Progress Series* 328:93–103. DOI:10.3354/meps328093.

Richards VP., Bernard AM., Feldheim KA., Shivji MS. 2016. Patterns of population structure and dispersal in the long-lived “redwood” of the coral reef, the giant barrel sponge (*Xestospongia muta*). *Coral Reefs* 35:1097–1107. DOI: 10.1007/s00338-016-1435-y.

Riesgo A., Pérez-Portela R., Pita L., Blasco G., Erwin PM., López-Legentil S. 2016. Population

structure and connectivity in the Mediterranean sponge *Ircinia fasciculata* are affected by mass mortalities and hybridization. *Heredity* 117:427–439. DOI: 10.1038/hdy.2016.41.

Salafsky N., Margoluis R., Redford KH., Robinson JG. 2002. Improving the practice of conservation: A conceptual framework and research agenda for conservation science. *Conservation Biology* 16:1469–1479. DOI: 10.1046/j.1523-1739.2002.01232.x.

Schmitt S., Weisz JB., Lindquist N., Hentschel U. 2007. Vertical transmission of a phylogenetically complex microbial consortium in the viviparous sponge *Ircinia felix*. *Applied and Environmental Microbiology* 73:2067–78. DOI: 10.1128/AEM.01944-06.

Selkoe KA., Toonen RJ. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters* 9:615–29. DOI: 10.1111/j.1461-0248.2006.00889.x.

Sunderland T., Sunderland-Groves J., Shanley P., Campbell B. 2009. Bridging the gap: How can information access and exchange between conservation biologists and field practitioners be improved for better conservation outcomes? *Biotropica* 41:549–554. DOI: 10.1111/j.1744-7429.2009.00557.x.

Taylor HR., Dussex N., van Heezik Y. 2017. Bridging the conservation genetics gap by identifying barriers to implementation for conservation practitioners. *Global Ecology and Conservation* 10:231–242. DOI: 10.1016/j.gecco.2017.04.001.

Untergasser A., Cutcutache I., Koressaar T., Ye J., Faircloth BC., Remm M., Rozen SG. 2012. Primer3- new capabilities and interfaces. *Nucleic Acids Research* 40:e115. DOI: 10.1093/nar/gks596.

Väli Ü., Einarsson A., Waits L., Ellegren H. 2008. To what extent do microsatellite markers reflect genome-wide genetic diversity in natural populations? *Molecular Ecology* 17:3808–3817. DOI: 10.1111/j.1365-294X.2008.03876.x.

Vitousek PM., Mooney HA., Lubchenco J., Melillo JM. 1997. Human domination of Earth's ecosystems. *Science* 277:494–499. DOI: 10.1126/science.277.5325.494.

Wang JT., Hirose E., Hsu CM., Chen YY., Meng PJ., Chen CA. 2012. A coral-killing sponge, *Terpios hoshinota*, releases larvae harboring cyanobacterial symbionts: An implication of dispersal. *Zoological Studies* 51:314–320.

Zhang W., Yang Q., Franco CMM. (unpublished) Multi-primers targeting 16S rDNA variable regions essential for sponge (Porifera) microbiome study.