Title

Comparative Metagenomics of Coral Reef Associated Marine Biodiversity Across a Pollution Gradient in Western Indonesia.

Author

Anggoro, Aji Wahyu

Publication Date

2022

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Comparative Metagenomics of Coral Reef Associated Marine Biodiversity Across a Pollution Gradient in Western Indonesia

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Biology

by

Aji Wahyu Anggoro

2022

ABSTRACT OF DISSERTATION

Comparative Metagenomics of Coral Reef Associated Marine Biodiversity Across a Pollution Gradient in Western Indonesia

by

Aji Wahyu Anggoro

Doctor of Philosophy in Biology

University of California, Los Angeles, 2022

Professor Paul Henry Barber, Chair

The increasing pace of global decline of marine ecosystems requires standardized monitoring methods that can more effectively capture changes in biodiversity composition, facilitating adaptive management efforts. Current monitoring methods focus on a small number of focal taxa, limiting our understanding of ecosystem change. Ideal methods would capture marine biodiversity from microbes to metazoans, ensuring that monitoring captures taxa most impacted by environmental degradation at sites of interest. Such data could help conservation practitioners manage ecosystems to preserve taxa that are most vulnerable to environmental perturbation, aiding management and conservation efforts focused on preserving local ecosystems and ecosystem function.

Environmental DNA (eDNA) and Autonomous Reef Monitoring Structures (ARMS) are increasingly used to enumerate local biodiversity as they are metabarcoding-based methods that allow for the sampling and reconstruction of entire communities across the tree of life from

microbial to metazoan communities. This study compares of eDNA and ARMS to understand changes in coral reef associated biodiversity in Pulau Seribu, an island chain in Western Indonesia where local reefs experience pronounced differences in pollution stress relative to their distance from Jakarta Bay. Chapter 1 compares the taxonomic range of ARMS and eDNA methods to capture local eukaryote marine biodiversity. Results demonstrate that these methods vary greatly in the total diversity and taxonomic composition of communities recovered, with minimal taxonomic overlap, indicating that ARMS and eDNA should be viewed as complementary rather than competing metabarcoding approaches for assessing marine biodiversity. Chapter 2 uses ARMS and eDNA to examine changes in marine communities and to identify indicator taxa across varied pollution levels in Kepulauan Seribu Jakarta. Results show a significant decrease in eDNA-ASV diversity and a slight increase in ARMS-ASV diversity with increasing Chlorophyll a concentration, a proxy for pollution stress, suggesting that eDNA may be a more effective tool for monitoring community change. Although taxa on ARMS appear to be impacted by pollution, both methods identified specific taxa that are indicators of pollution stress. Chapter 3 examines shifts in ARMS microbial communities relative to pollution stress across Kepulauan Seribu. Although microbial diversity was substantially different among the three ARMS fractions examined, there was no significant impact on total microbial community diversity, and a non-significant increase in bacterial diversity in sites with increased pollution stress. However, as with eukaryotes, there were taxa indicative of pollution stress. Combined, the results of this thesis highlight the power of ARMS and eDNA metabarcoding to capture changes in marine biodiversity in response to environmental degradation, providing important new tools for the study and management of marine biodiversity, both in Indonesia and globally.

The dissertation of Aji Wahyu Anggoro is approved.

Peggy Marie Fong

Thomas Bates Smith

Forest Rohwer

Paul Henry Barber, Committee Chair

University of California, Los Angeles,

2022

TABLE OF CONTENTS

LIST OF TABLES

TABLES

Chapter 1.

Chapter 2.

SUPPLEMENTAL TABLES

Chapter 1 Supplemental Tables.

Chapter 2 Supplemental Tables.

Chapter 3 Supplemental Tables.

LIST OF FIGURES

FIGURES

Chapter 1.

undertaken on individual fractions (106–500 μm, 500 μm -2mm, and sessile) and on all fractions combined across the seven sampling sites.

Figure 3-13. Plot of total OTUs as a function of Chlorophyll a concentration on including best fit line and R^2 value

SUPPLEMENTAL FIGURES

Chapter 3 Supplemental Figures

ACKNOWLEDGMENTS

I would like to acknowledge and give my warmest appreciation and gratitude to my supervisor Paul Barber who has made this difficult journey travelable. Pursuing a Ph.D. is definitely not for everybody, but his supervision, guidance, and advice have carried me through all the struggles and made this *a painless* journey. I enjoyed every moment of my Ph.D. life in his lab and will forever treasure his mentorship, thoughts, and advice for my future career. Although it has had its ups and downs, on the whole, my Ph.D. journey has been enjoyable moments, and I want to thank you and your family for making my life at UCLA and Los Angeles memorable.

I would also like to recognize the role and tremendous contributions of my committee in the success of my Ph.D. research and thesis. I thank Peggy Fong and Thomas Smith from UCLA and Forest Rohwer from San Diego State University for their impactful guidance during my research. Your feedback, suggestions, and thoughtful comments have greatly improved my research. I hope the results will catalyze important biodiversity research in Indonesia and offer important information that advances marine conservation in Indonesia. Similarly, I want to convey my appreciation to Christopher Meyer from the Smithsonian Institution. Thank you for your ideas, encouragement, and for spreading your enthusiasm for marine biodiversity research to young Indonesians. I will surely miss those early morning dives, late-night lab work, and diving in "*unforgettable"* Bidadari that we did together with Paul.

I also thank Tessa Villaseñor, Jocelyn Yamadera, Annelise Werhel, and the EEB community. Thank you for making things happen and for your patience and guidance in dealing

with campus life and never-ending paperwork. I also thank my cohort for the inspiring stories and all the great experiences along the way. I hope our path cross again.

My enormous appreciation and million thanks also for the past and current Barber Lab members; Samantha Cheng, Sara Simmons, Rita Rachmawati, Allison Fritts-Penniman, Abril Iñiguez, Zack Gold, Sam Degregori, Kelcie Chiquillo, Erick Zerecero, Onny Marwayana, Eric Caldera Candice Cross, Satoshi Tomano for the stories and the knowledge sharing. And a special thanks for Dita Cahyani for being a good listener and a place where could I dump all my academic problems. I also thank Dita for her integral part in completing lab work for all ARMS samples, a role with which I have a love-hate relationship. Your obsession with ensuring that every plan should be on track, and your appetite for things to be done orderly, is something that I'm unable to match.

I also want to thank Yayasan Biodiversitas Indonesia (BIONESIA); Andrianus Sembiring, Ni Putu Dian Pertiwi, Astria Yusmalinda, Yuliana Syamsuni, Eka Maya Kurniasih, Danie el Malik, for their constant support and for providing me a place where I can channel my passion for doing science in Indonesia. I hope the organization continues to expand and flourish in the future.

And a big thank you to Indonesian students and family in L.A.; thank you for the food, memorable road trips, and pictures. They have been instrumental in making L.A as my second home. I feel like L.A. is just another part of Indonesia when they are around.

I also thank my wife, Farah, for her love and sacrifice, leaving the job you love the most and supporting my crazy life decision. Thank you for your unconditional support and for agreeing that this path is good for us. Thanks also to Fariz and Zayna for making the world worth

living. Thank you to my mom, late father, brothers, and sister, who have kept me in their prayers and supported me all this time.

I also want to convey my gratitude to Lembaga Ilmu Pengetahuan Indonesia (LIPI): late Irma Arliza, Ismiliana Wirawati, and Dedy Kurnianto. This work was funded by the National Science Foundation with a Partnerships for International Research and Education (PIRE) grant (OISE-1243541), entitled "Assembly of Marine Biodiversity Along Geographic and Anthropogenic Stress Gradients". Thank you to the entire PIRE team: Forest Rohwer (San Diego State University), Paul Barber (University of California, Los Angeles), Jonathan Geller (Moss Landing Marine Laboratories and San Jose State University), Nancy Knowlton, Chris Meyer, and Allen Collins (The Smithsonian Institution, National Museum of Natural History), Russell Brainard, and Molly Timmers (NOAA Pacific Islands Fisheries Science Center), The National Evolutionary Synthesis Center (NESCent), I Gusti Ngurah Kade Mahardika (Udayana University), Ambariyanto (Diponegoro University), Hamid Toha (State University of Papua), Mark Erdmann (Conservation International), and Dr. Ir. Zainal Arifin (Indonesian Institute of Sciences or LIPI). And thank you to the Laboratory of Marine Molecular Genetics, Research Center for Oceanography, Indonesian Institute of Sciences, Jakarta, Indonesia, for providing us with the facility to perform research in Indonesia. I also thank Emma Ransome and Aaron Hartman for their important input for research design, writing, and analysis.

I am forever indebted to the Indonesia Endowment Fund for Education (LPDP) Scholarship. These life-changing experiences have exposed me to a world I had never seen before and gave me a view of the world and science that will forever be valuable for me as I continue with my life and career. I have learned so much from so many different and wonderful people along the way; I hope I can keep learning new things as I move forward with my life.

Finally, I thank the Center for Southeast Asia Study (CSEAS), UCLA, and EEB Department for providing a travel grant that funded my research in Indonesia.

BIOGRAPHICAL SKETCH

Previous Degrees Awarded

Bogor Agricultural University, Bogor, Indonesia – B.S. in Fisheries. Awarded 2006

University of The Ryukyus, Okinawa, Japan – M.S. in Biology. Awarded 2008

Awards and Grants

2021 - Biodiversity Information for Asia (BIFA): Advance Molecular Method for shark Identification: Moving forward toward Sustainable Shark Conservation (€19.900). Project Leader

2021 - Lessening the Gap: Nurturing Access for Young researchers in Eastern Indonesia to Analyze Molecular Data. Co-PI

2018 - Indonesian Studies Travel Grant from the UCLA Center for Southeast Asian Studies

2016 - Indonesian Studies Travel Grant from the UCLA Center for Southeast Asian Studies

2015 - Indonesian Endowment Fund for Education Presidential Scholarship program (Ph. D) (2015-2019)

Publications

1. Casey JM, Ransome E, Collins AG, Mahardini A, Kurniasih EM, Sembiring A, Schiettekatte NMD, Cahyani NKD, Anggoro AW, Moore M, Uehling A, Belcaid M, Barber PH, Geller JB, Meyer CP (2021) DNA metabarcoding marker choice skews perception of marine eukaryotic biodiversity. Environmental DNA 00:1–18

2. Fong CR, Chiquillo KL, Gaynus CJ, Grier SR, Hà BA, Ryznar ER, Smith LL, Sura SA, Zweng RC, Anggoro AW, Moore TN, Fong P (2021) Flip it and reverse it: Reasonable changes in designated controls can flip synergisms to antagonisms. Science of The Total Environment 772:145243

3. Sembiring, A., Pertiwi, N.P.D., Mahardini, A., Wulandari, R., Kurniasih, E.M., Kuncoro, A.W., Cahyani, N.K.D., Anggoro, A.W., Ulfa, M., Madduppa, H., Carpenter, K., Barber, P.H. and Mahardika, G.N. (2014). DNA Barcoding reveals targeted fisheries for endangered sharks in Indonesia. Fisheries Research. 164 (2015). 130-134

4. Barber, P. H., Ablan-lagman, M. C. A., Cahyani, D., Crandall, E. D., Ravago-gotanco, R., Juinio-meñez, M. A., Mahardika, I.G.N., Shanker, K., Starger, C.J., Toha, A.H.A., Anggoro, A.W. and Willette, D.A. (2014). Advancing biodiversity research in developing countries: the need for changing paradigms. Buletin of Marine Science 90(1). 2014 Rosenstiel School of Marine & Atmospheric Science of the University of Miami

Conference and Presentations

1. August 23rd-26th 2022, Carbon Accounting and Nature Climate Solution Training by The Nature Conservancy. Washington D.C. U.S.A
2. July 14th, 2019 - July 19th, 2019. Gordon Research Conference. HKUST, HK (Poster session). *Employing the power of metabarcoding and autonomous reef monitoring structures to assess Indonesian marine macro and micro biodiversity*

CHAPTER 1

Community Metabarcoding Results Vary Based on Method: Comparative study of eDNA and ARMS to detect Marine Biodiversity in Western Indonesia

Abstract

Environmental DNA (eDNA) and Autonomous Reef Monitoring Structure (ARMS) are widely used to assess marine biodiversity and have expanded our ability to document rare cryptic taxa. However, it is unclear whether these methods provide comparable results and can be used interchangeably or whether they capture different components of marine biodiversity. This study direct compares results of ARMS and eDNA metabarcoding from coral reefs of Pulau Seribu, West Java, Indonesia. Results from COI metabarcoding demonstrate a substantial variation in both total community diversity and taxonomic composition. ARMS captured nearly twice as many ASVs than eDNA (2907 vs 1538), with less than 3% of ASVs being recovered by both methods. Although ARMS were dominated by largely benthic taxa (Annelids, Arthropods, Cnidaria, Mollusca, Porifera, Rhodophyta), planktonic taxa (Ascomycota, Bacillariophyta and Chlorophyta) dominated eDNA; both methods had high numbers of unidentified taxa. Results indicate that ARMS and eDNA metabarcoding should be viewed as complementary, rather than competing, approaches to assessing marine biodiversity. The large number of unidentified taxa highlight the need for more research in high biodiversity marine ecosystems like Indonesia to build metabarcoding databases, and fully realize the power of metabarcoding approaches.

Introduction

Marine biodiversity plays a critical role in supporting the global economy (Bishop 1993, Jones-Walters & Mulder 2009), particularly the needs of local communities (Aylward & Barbier 1992), where they can provide as much as 80% of daily protein intake (Moberg & Folke 1999, Bell et al. 2009). A healthy marine ecosystem mitigates the severity of natural disasters (e.g., tsunamis and typhoons) by reducing coastal erosion (Hernández-Delgado 2015, Gracia et al. 2018) and allowing coastal communities to be more resilient (Hughes et al. 2003, Levin & Lubchenco 2008, Côté & Darling 2010). Marine biodiversity also plays an important role in regulation of global climate through carbon sequestration (Alongi 2014, Duarte & Krause-Jensen 2017). Finally, the combined ecological goods and services provided by marine biodiversity are valued at US $2.5 trillion annually (UNCTAD 2021), highlighting the enormous resources that the ecosystem provides and the needs of conserving biodiversity against future threats.

Despite its importance, marine biodiversity is in major decline worldwide (Sodhi et al. 2004, Wilson et al. 2010, 2015), with negative impacts for both humans and nature (Crain et al. 2008, Halpern et al. 2008). Marine ecosystems are threatened by local processes such as coastal development (Crain et al. 2009), pollution (Baum et al. 2015) and unsustainable fishing practices (Newton et al. 2007, Coll et al. 2008) as well as global processes such as ocean acidification and rising sea surface temperatures associated with increasing atmospheric CO_2 levels (Guinotte & Fabry 2008, Hofmann et al. 2010). To track the impacts of these threats, and provide data essential for biodiversity management, it is important to develop effective monitoring approaches.

Current approaches to monitoring marine biodiversity are often narrow in scope, focusing only on economically valuable or keystone species (Williams & Gaston 1994, Brooks et al.

2006, Bickford et al. 2007, Losey & Vaughan 2009). Results of such targeted monitoring efforts may be misleading as they ignore the majority of marine biodiversity and the role and magnitude of ecosystem functioning which remains unknown (Mora et al. 2011). Furthermore, the small number of monitored taxa may not be the most sensitive to anthropogenic stressors, potentially limiting an understanding of impacts and possible mitigation measures.

In response to the need for new approaches to monitoring of marine ecosystems, scientists are increasingly employing molecular methods such as metabarcoding to assess community diversity using environmental-derived samples (Deiner et al. 2017, Aylagas et al. 2018, Ríos-Castro et al. 2021). Metabarcoding is a method for elucidating the members of a biological community by combining DNA barcoding and high throughput DNA sequencing (Thomsen et al. 2012, Leray & Knowlton 2016b, DiBattista et al. 2020). Metabarcoding is a promising approach for monitoring marine ecosystems because of its ability to rapidly detect and enumerate local biodiversity, using specific primer sets to target large taxonomic groups such as teleost (Zhang et al. 2020, Madduppa et al. 2021), elasmobranchs (Bakker et al. 2017), metazoans (Leray & Knowlton 2015), or microbes (Ladin et al. 2021, Bairoliya et al. 2022). Due to the generality of these methods, which require only standard molecular lab facilities, as well as their sensitivity and cost-effectiveness, metabarcoding approaches are increasingly applied in various terrestrial (Schmidt et al. 2013, Evans et al. 2016, Valentin et al. 2020, Ladin et al. 2021) and aquatic environments (Leray et al. 2013, Carstensen et al. 2016, Leray & Knowlton 2016b, Wangensteen et al. 2018, Pearman et al. 2018, Madduppa et al. 2021), providing more detailed insights into local biodiversity compared to conventional methods.

Two metabarcoding approaches increasingly used to study biodiversity of marine ecosystems are Environmental DNA (eDNA) and Autonomous Reef Monitoring Structures

(ARMS). eDNA is based on the collection of cells and DNA that organisms shed into the environment. By metabarcoding this dissociated DNA, eDNA metabarcoding allows for non-destructive sampling and reconstruction of entire communities (D'Alessandro & Mariani 2021, Blackman et al. 2022). This method has a broad range of applications, including detection of rare/endangered species (Weltz et al. 2017) and invasive species (Ellis et al. 2022), assessing changes in community composition over time (Muha 2021, Reinholdt Jensen et al. 2021), and tracking ecologically important taxa (Djurhuus et al. 2020, Bonfil et al. 2021). This method is particularly useful in marine ecosystems where logistical, budgetary and safety concerns greatly limit biodiversity monitoring (Gold, Sprague, et al. 2021). In comparison to conventional biodiversity monitoring (e.g., visual census), eDNA captures much more biodiversity across a broader range of taxonomic groups (Zaiko et al. 2015, Stat et al. 2017, Holman et al. 2019, Gold, Sprague, et al. 2021). eDNA is particular useful when used in conjunction with conventional methods, by detecting taxa that are large and mobile, small and cryptic, or have activity patterns that make them easily missed by traditional methods of marine biodiversity monitoring (Closek et al. 2019, Steyaert et al. 2020, Gold, Sprague, et al. 2021, Marwayana et al. 2022, Klunder et al. 2022).

In contrast, ARMS metabarcoding is based on the collection and processing of entire marine communities that colonize an artificial structure comprised of 23 x 23cm PVC plates that are designed to approximate benthic marine habitats (Leray & Knowlton 2016b, Ransome et al. 2017, Pearman et al. 2018). Although less commonly used than eDNA, there are over 1600 ARMS deployed globally (Ransome et al. 2017) to understand distribution and dynamics of global marine biodiversity. Like eDNA, ARMS are designed to capture a broad cross section of marine biodiversity, particularly taxa that are overlooked by traditional survey methods (Leray &

Knowlton 2015, 2016a, Wangensteen et al. 2018), and can record changes in biodiversity composition across space and time. An important advantage of ARMS is that its standardized structure, surface area and processing protocols enable direct comparison among samples across a wide variety of taxonomic groups, habitats, geographic regions, and time.

Although ARMS and eDNA both employ DNA metabarcoding to survey marine biodiversity, particularly in support of marine ecosystem monitoring, each method has unique characteristics in terms of sampling and ability to characterize marine biodiversity. For example, eDNA samples freely associated cells and DNA isolated from water samples, whereas ARMS samples actual organisms that encrust and/or live associated with the surface of the ARMS unit. eDNA integrates DNA signals from vertebrates, invertebrates, and microbial communities from a localized but indeterminate area of marine habitat ranging from 10s of meters to kilometers (Port et al. 2016, Yamamoto et al. 2017, Jeunen et al. 2019), whereas ARMS specifically target marine invertebrates, algae and microbial communities that colonize settlement plates installed on a specific location on the sea floor. Conversely, because eDNA degrades over a span of hours to days (Collins et al. 2018, Jensen et al. 2022), it represents only a snapshot of biodiversity present during a relatively brief period of time, whereas ARMS were designed to capture and integrate marine biodiversity that accumulates over time, typically one to several years (Pearman et al. 2016, Wangensteen et al. 2018, Carvalho et al. 2019).

Although there haven't been published studies aiming to investigate succession pattern in ARMS, taxa accumulation across site or regions have been reported to differ substantially. Cahyani 2021 and author own observation for example, have reported a well-established ARMS plate consisting of hard coral, soft coral and other taxa (e.g., macro algae and Crustose coralline algae) growing in more pristine, high fish biomass and diversity, and in relatively undisturbed

area in ARMS deployed for three years in eastern Indonesia. In contrast, ARMS deployed in other areas with similar length of deployment, however with intensified anthropogenic pressure and poor fish biomass and diversity, have taxa coverage that were dominated by filamentous and macro algae with minimal hard coral growth. Composition and abundance of the colonizing assemblage on artificial plates have been reported to depend on local characteristic which includes proximity to source of disturbance, size of (e.g., reef size (Bohnsack et al. 1994), proximity of source populations (Burt et al. 2009), local hydrodynamics (Baynes & M. Szmant 1989) , and local taxa composition (Brown 2005). On this context ARMS methods can summarize local pollution context and can act as a proxy to ongoing and past disturbance within the proximity of ARMS deployment sites. And therefore, ARMS methods can be used to summarize local pollution context and can be used as a proxy to ongoing and past disturbance in the areas of deployment. Whereas eDNA tends to explain community dynamic within short period of time. Recent results showed however, albeit presumed to only represent snapshot of local diversity and coming from localized DNA, that eDNA might not be different across season in marine environment, and hence suggesting persistence of taxa composition across time (Salter 2018, Collins et al. 2018).

Although both eDNA and ARMS have been used widely and are often described as alternatives or complementary to traditional approaches for marine biodiversity monitoring, these two methods haven't been used in parallel. As such, it is unclear whether they capture local marine community diversity in relatively equivalent or dissimilar ways, and if the latter, how these methods differ in articulating local biodiversity and whether such differences matter for marine monitoring efforts. Answering these questions is important, particularly for high biodiversity regions like the Coral Triangle, where marine ecosystems are severely threatened

(Bruno & Selig 2007, Burke et al. 2011, Foale et al. 2013) but human and economic resources for monitoring of marine ecosystems are relatively limited (Barber et al. 2014).

The Coral Triangle is a region of Southeast Asia that is home to the world's largest and most diverse marine ecosystems (Allen & Adrim 2003, Allen & Erdmann 2009a). Despite the nutritional, economic, and cultural importance of marine ecosystems in this region (Hoegh-Guldberg et al. 2009, Foale et al. 2013), population growth and the need for economic development has resulted in pronounced environmental degradation, with 85 % of reef area being lost are threatened by local stressors (Burke et al. 2011). Although traditional marine biodiversity monitoring studies (e.g., underwater visual census) have produced important information, such as species conservation status and patterns of macrofauna distribution, supporting marine conservation in this biodiversity hot spot requires novel approaches to improve biodiversity monitoring and expand our knowledge of ecological processes within the valuable marine ecosystems of the Coral Triangle areas.

Cahyani (2021) used ARMS to demonstrate that broad-scale patterns of marine biodiversity observed in macrofauna like fish and corals are also observed in cryptofauna and microbes across the Indonesian Archipelago. Similarly, Marwayana et al. (2021) employed eDNA across the Indonesian Archipelago, demonstrating that this method captured a largely non-overlapping ichthyofauna, compared to traditional visual survey methods. Although both studies examined marine diversity across the Indonesian Archipelago, they each sampled different locations, and Marwayana et al. (2021) excluded invertebrates, preventing direct comparisons of eDNA and ARMS metabarcoding results.

Given the increasing use of eDNA and ARMS to study and monitor marine biodiversity, it is essential to understand how each of these methods perform, including their detection sensitivities

and taxonomic biases. This study examines the efficacy of eDNA and ARMS in marine ecosystem monitoring by comparing results from co-located samples obtained by each method from coral reef areas of the Seribu archipelago, Indonesia to better understand how these different metabarcoding approaches to environmental sampling might vary in their detection of marine biodiversity and to provide recommendations on how each method can best be employed in the monitoring of marine ecosystems

Material, Methods and Environmental Data

Study sites

Kepulauan Seribu is an archipelago comprised of 105 islands (11 inhabited) that spans 80km of the waters north and west of Jakarta, Indonesia. Due to its proximity to Jakarta and anthropogenic stressors related to terrestrial effluent, a number of studies have conducted biodiversity assessments using visual census and morphological identification approaches (Rachello-Dolmen & Cleary 2007, van der Meij et al. 2009, 2010, Madduppa et al. 2013, Cleary 2017), revealing 216 reef fish species and 61 genera of hard corals (Rachello-Dolmen & Cleary 2007, van der Meij et al. 2009, 2010, Madduppa et al. 2013, Cleary 2017). In this study, we use eDNA and ARMS metabarcoding to assess marine biodiversity across this archipelago and to compare how these methods vary in detecting local marine diversity.

ARMS Deployment, Collection, and Sampling

We deployed ARMS and collected eDNA samples from eight islands across the archipelago, with an average distance between sampling sites of 7-10 km. Our sampling sites encompass an area 10 km from Jakarta Bay to 60 km north direction away from bay (Table. 1).

In 2013, we deployed sets of three ARMS on SCUBA at a depth of 10m, 24 ARMS in a total for eight sites. Following standard protocols, each of these ARMS consisted of nine 23 cm x 23 cm stacked PVC plates separated by spacers and attached to a 35 cm x 45 cm base plate to provide cryptic habitats for reef organisms. After being deployed on the seafloor for three years, we recovered the ARMS units in 2016 using SCUBA. To recover ARMS units, we first enveloped the entire ARMS unit in a 40 μm nitex mesh-lined container to prevent motile organisms from escaping during recovery. After retrieval, we transported the ARMS unit to the Indonesian Institute of Science field lab in Pulau Pari using large plastic container filled with filtered, aerated sea water. At the lab, we carefully disassembled each ARMS unit individually and then separated and transferred each plate to individual trays filled with filtered seawater. To ensure that all motile organisms were detached from the ARMS plates, we vigorously shook each plate in the holding tank prior to transferring it to its individual tray. We then took high resolution photos of the top and bottom of each plate for future visual reference.

To obtain samples for metabarcoding, we processed each ARMS unit, following standardized disassembly and sampling protocols (Leray & Knowlton 2015, Ransome et al. 2017). To ensure that we collected all motile taxa and sediment, we filtered water from the container used for transporting ARMS through sets of sieves (2 mm, 500 μm and 106 μm) in two steps. In the first step, we stacked 2 mm sieve on top of the 500 μm sieve and allowed all the water and organisms from the transporting container to pass through both sieves and into a new bin. Next, we removed the 2 mm sieve and placed a 106 μm sieve under the 500 μm sieve filled with sediment. We then passed the water from the initial filtering step through the 500μm and 106 μm sieve. Through this sequential filtering, we obtained 3 fractions, >2 mm fraction that was saved for subsequent voucher-based DNA barcoding, as well as a 500 μm and 106 μm fraction,

hereafter referred to as the 500 and 100 μm fraction, respectively. Next, we concentrated these two fractions using a 40 μm nitex mesh stretched between fitted PVC pipes and then rinsed with 95% ethanol, and then transferred the resulting samples to falcon tubes preserved with 95% and stored at -20 °C. Lastly, to document non-motile taxa (hereafter referred to as the sessile fraction), we scraped all encrusting or sessile biota from ARMS plate into a tray then homogenized the tissues with a blender for 30 s at maximum speed. We then rinsed the homogenate with ethanol into a 45 μm Nitex mesh collection net. After allowing excess ethanol to drip away, we placed the sample in a 50 ml falcon tube filled with DMSO and stored in -20 °C.

eDNA Collection and Sampling

To create an eDNA dataset directly comparable to ARMS, we followed standard sampling protocols used in temperate ecosystems (Miya et al., 2015), collecting three one-liter replicate seawater samples on SCUBA at depths between 11-15m at each ARMS deployment site to maximize species diversity and to account for fine-scale heterogeneity in local eDNA signatures. To maximize comparability, we sampled eDNA within 1-2 days of ARMS retrieval.

To isolate eDNA from water samples, we filtered each replicate seawater sample through a 0.22-micron Sterivex™ filter (Millipore®, SIGMA MILLIPORE) following the methods of Miya et al. (2015) with one key modification; we collected individual water samples in sterile 1 liter Kangaroo™ Gravity Feeding Bags (similar to intravenous drip bags) that allow for gravity filtration through the Sterivex™ filters. In addition to the eDNA water samples, we also filtered one blank at each locality as a negative control. Filters were stored in a -20 freezer until eDNA was extracted.

DNA Extraction and Library Preparation

ARMS

To prepare the fractions for DNA extraction, we performed decantation on the 500 μm and 100 μm fractions to separate calcium carbonate and terrigenous sediment from organic matter. During the decantation, we suspended the entire fraction in sterile water in a 1 L Erlenmeyer flask, allowed the sediment to settle, and then decanted the suspended content through a geological sieve to recover the organic contents. We used a 106 μm sieve to decant the 500 μm fraction and a 45 μm sieve to decant the 100 μm fraction. We then used the decanted samples for further extraction and sequencing steps. Because of minimal inorganic content, the sessile fraction was not decanted; instead, the homogenized material was vigorously mixed before sub-sampling steps.

DNA extractions were performed at Laboratory of Marine Molecular Genetics, Research Center for Oceanography, Indonesian Institute of Sciences, Jakarta, Indonesia and sequencing were performed at Smithsonian Institution National Museum of Natural History in Washington DC, USA. We performed DNA extractions using 10 g of each ARMS sample digested overnight with the addition of 400 μg/ml Proteinase K in a shaking incubator at 56 ºC at 200 rpm. Following digestion, we extracted DNA using the Powermax Soil DNA Isolation Kit (MoBio, Carlsbad, USA) following the manufacturer's protocol. To remove potential PCR inhibiters from DNA extractions, we cleaned all DNA extracts using a PowerClean DNA Clean-Up Kit (MoBio, Carlsbad, USA) and then quantified DNA extractions using Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA) and a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA).

We amplified mitochondrial COI from ARMS samples using seven tailed primer pairs of m1COIintF and jgHCO2190 (Leray et al. 2013, Geller et al. 2013). We conducted PCR amplifications in 20 µl reactions volumes consisting of 1 µl of 10 µM forward and reverse primer, 1.4 µl of 0.2 mM dNTP, 2 µl Advantage 2 DNA Buffer (Takara Bio USA, Mountain View, CA, USA), 0.4 µl Advantage 2 Polymerase (Takara Bio USA, Mountain View, CA, USA), 13.2 µl distilled water, and 1 µl of 10 ng/l DNA template. We performed a two-step touchdown PCR; initial denaturation was at 95°C for 10 min, proceeded by the first step for sixteen cycles: 95°C for 10 s, 62°C (-1°C per cycle) for 30 s, and 72°C for 60 s. The second step was run for twenty cycles: 95°C for 10 s, 46°C for 30s, and 72°C for 7 min, followed by a final extension at 72°C for 7 min. We ran PCR reactions in triplicate and verified success on 1.2% agarose gels, then all successful reactions were pooled into a single product.

We quantified and pooled PCR products using a Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA) with a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). We pooled tailed primer pairs in equimolar concentrations, followed by bead cleaning at a concentration of 0.8x vol/vol with Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA). We then prepared sequencing libraries with a TruSeq DNA PCR-Free IT Library Prep Kit (Illumina, San Diego, CA, USA) following the manufacturer's protocols. Lastly, we quantified the indexed samples using a Qubit Flourometer, pooling 50 ng of each library, normalizing the sample to 4nM prior to sequencing on an Illumina MiSeq with a MiSeq using Reagent Kit v3, 600-cycle (Illumina, San Diego, CA, USA) and a 1% PhiX spike.

eDNA

We extracted eDNA samples using the modified DNeasy Blood & Tissue Kit (QIAGEN, Germany) of Spens et al. (2017) and then PCR amplified the extracted eDNA using the same primer sets as above, but using the Multiplex PCR Kit (QIAGEN, Germany). To account for potential PCR bias associated with low eDNA concentrations, we conduced PCR in triplicate (Andruszkiewicz et al., 2017; Miya et al., 2015; Taberlet et al., 2012). Each PCR reaction consisted of 12.5 µL Qiagen 2x Master Mix, 2.5 µL (2 mM) of the primer, 6.5µL nuclease free water, and 1 µL the DNA extract. Thermocycling parameters utilized a touchdown protocol, beginning with a 15-minute pre-denaturation step at a 95 °C, followed by a touchdown thermocycling profile consisting of 30 seconds denaturing at 94 °C, 30 seconds annealing at 69.5 °C, and 30 seconds extension at 72 °C, with the annealing temperature dropping by 1.5°C per cycle until50 °C. Following this initial touchdown phase, the main cycle consisted of 25 cycles of 94 °C for 30 seconds for denaturation, 50 °C for 30 seconds for annealing and 72 °C for 45 seconds for extension, concluding with a 10-minute final extension at 72 °C. Final PCR product were then verified on 2% agarose gels. We then pooled the triplicate PCR products, representing a single one-liter eDNA sample, into a single tube, and purified these pooled PCR products using Sera-Mag™ and Sera-Mag Speed Beads Magnetic Particles (SIGMA-ALDRICH®) following manufacturer's protocols. Next, we quantified the DNA concentration (ng/µL) of each pooled PCR sample using the Qubit™ 4 NGS Starter Kit (Thermo Fisher) following the manufacturer protocol and then adjusted concentrations of pooled PCRs to have equal concentrations across all samples. We then used the Nextera DNA Library Preparation Kit (Illumina®) to index each PCR amplified eDNA sample using a unique combination of Illumina Nextera i5 and i7 primers in a second PCR reaction, following the manufacturer protocol. The indexing PCR reaction consisted

of 12.5µL Kapa High Fidelity Master Mix, 0.625 µL of 1µM i5 Illumina Nextera indices, 0.625 µL of 1µM i7 Illumina Nextera indices, and 11.25 µL of PCR product for a total of 10ng of DNA.

To bioinformatically distinguish among samples, we added index barcodes to each sample utilizing an indexing PCR protocol that began with an initial denaturation of 95℃ for 5 minutes, followed by 8 cycles of: 98℃ denaturation for 30 seconds, 56℃ annealing for 30 seconds, and 72℃ extension for 3 minutes, ending with a 72℃ extension for 5 minutes. To ensure the indexing PCR was successful, we electrophoresed indexed PCR products at 120V for 45 minutes on a 2% agarose gel prepared with 6x SYBR™ Green. We then cleaned and quantified indexed PCR products, as above, to creating a final sequencing library that contained equal DNA concentrations (10 ng/µl) of each sample and sequenced the libraries at the UC Berkeley sequencing core on an Illumina MiSeq platform, utilizing 300 base pair paired end sequencing.

Sequence Processing and Taxonomic Assignments

ARMS

Because the resulting COI data contained mixed-oriented paired-end demultiplexed sequences (i.e., forward, and reverse-oriented DNA sequence reads within a single raw FASTQ file), we used several scripts and steps to preprocess our DNA barcodes to separate and group the sequences based on orientation. To split reads, we first demultiplexed all raw paired-end FASTQ reads using CUTADAPT 1.13 (Martin, 2011) by employing error rate of 0.15 (-e = 0.15), using --discard-untrimmed and --no-indels options. This step produced forward oriented and reverse oriented for each sample. Following this step we imported each sample from each orientation as

a qza file for further processing using QIIME22, ver. 2018.11 (Bolyen et al. 2019). Next, we used Divisive Amplicon Denoising Algorithm 2 (DADA2) (Callahan et al. 2016) implemented in QIIME22 to filter reads, remove poor quality base pairs from the 3' end of reads, remove sequencing error through denoising, combine forward and reverse reads, identify and remove chimeras, and cluster all reads into OTU's (amplicon sequence variants) to each orientation. We set --p-trunc-len for forward reads at position 240 and 210 for reverse reads, while for chimera removal methods we used 'consensus' method as implemented in QIIME22 chimera removal method. These steps generated representative sequences and an OTU table as qza files for each orientation. Next, we exported the qza representative sequences file into a FASTA file. Following this step, we reverse-complemented the reverse oriented FASTA file using fastx_reverse_complement command from FASTX Toolkit 0.0.14 (http://hannonlab.cshl/fastx_toolkit/) and rehashed it using FASTA-names-to-md5.py command from (https://gist.github.com/mkweskin). We then renamed reverse oriented ASV tables, using batch-find-replace.py from (https://gist.github.com/mkweskin). Following these steps, we imported the rehashed and renamed ASV and FASTA file into the QIIME22 qza format and merged with existing forward oriented file using QIIME2 feature-table merge-seqs and QIIME2 feature-table merge commands. Lastly, we applied LULU (Frøslev et al. 2017) to remove erroneous ASVs (minimum_match = 84 for COI, minimum_relative cooccurence = 0.90 for both markers) to the representative sequence and then clustered all FASTA sequences using vsearch (Rognes et al. 2016) at 97%

eDNA

We analyzed all eDNA sequences using the Anacapa pipeline (Curd et al., 2018). Briefly, Anacapa begins by creating a *de novo* sequence reference library by combining records from publicly accessible databases, such as European Molecular Biology Laboratory (EMBL) and National Center for Biotechnology Information (NCBI) using the CRUX package (Creating Reference libraries Using existing tools; https://github.com/limey-bean/Anacapa/tree/New-Master/Anacapa_db). Anacapa then de-multiplexes the amplicon reads based on the primer sequences and primers are trimmed from the reads. Next, the DADA2 algorithm (Callahan et al.2016) performs denoising and error correction on the raw sequence data, merges paired end reads, and assigns high quality reads to Amplicon Sequence Variants (ASVs) through ASV Parsing. Results were then assigned to taxa using both Bowtie 2 and the Bayesian Least Common Ancestor algorithm (BLCA; Gao et al.2017.). Following these steps, the resulting FASTA file, list of assigned ASV and number of reads per samples were then processed through LULU as detailed in ARMS method.

Merging eDNA and ARMS Data

To ensure consistency in taxonomic assignment and comparability across data sets, we merged FASTA files from both ARMS and eDNA prior to taxonomic assignment. We combined the FASTA files using cat command in UNIX, conducted alignment using MAfft software (MAFFT v7.481) and then clustered aligned sequences using *mothur* (at 97% similarity. We then conducted taxonomic assignment of the combined sequence files using BLASTN (Camacho et al. 2009) employing a 85% identity threshold. For assignment, we used custom database that combined NCBI database (downloaded on June 2020) and a local BIOCODE database (Meyer,

C. P. 2016) generated using ANACAPA toolkit (Curd et al. 2019). This resulted in an OTU table consisting of representative sequences and a taxonomic table that included assigned taxonomic names for each of these representative sequences. These files were then arranged as a single phyloseq object (*phyloseq* package version 1.26.1) (McMurdie & Holmes, 2013) to allow visualization and further analysis.

Data Analysis

We used *phyloseq* package (version 1.26.1) (McMurdie & Holmes, 2013) to wrangle data prior visualization steps, combining the taxonomic file, representative sequence and OTU table as a single *phyloseq* object/file. We also used the package to count alpha diversity (Observed OTUs, and Shannon), group samples based on fraction and sites and transformed from reads counts to relative read abundance (RRA). In order to visualize all plots we used *ggplot2* (version 3.3.2) (Wickham, 2016). To produce rarefaction curves and visualize sampling effort and OTU richness across sites and fractions we use *GGRARE* package (version 0.1.0.) (Kandlikar, 2020). To visualize taxonomic overlap among methods, we produced Venn diagrams using *VENN package* (version 1.10.)

To test for significant differences among alpha diversity in different sites and fractions, we first tested for normality in all alpha diversity values across sites ,using Shapiro test as employed in *stats package* (version 4.0.2) (R Core Team, 2020) and homogeneity of variances using using *bartlett.test* function from the same package. For data that have normal distribution and equal homogeneity we implemented one-way ANOVA and for non-normally distributed data we used Kruskal-Wallis test as implemented in *stats package*. To test if there was a significant difference between the number of OTUs generated by eDNA and ARMS, we used

individual non-summed replicate sample data to build a two-way ANOVA model with the formula number_of_OTUs~eDNA or ARMS*site (number of OTUs reflects number of OTUs at each sites, while eDNA/ARMS reflect the method used, and site reflect the deployment sites) implemented in R using the function *aov*. For the phylodiversity analysis we built a phylogenetic tree using Neighbor-Joining method as implemented in PHANGHORN 2.7.1 package. Following this step, we used *pml* and *optim.**pml*** to compute the likelihood of a phylogenetic tree and to optimize the different model parameters given a sequence alignment and a model and calculated phylogenetic diversity (PD) using the *estimate_pd* command from BTOOLS 0.0.1 (Battaglia, 2018). We then used Shapiro test again to check the normality in PD values and *bartlett test* for homogeneity detection. For data that have normal distribution and equal homogeneity we implemented one-way ANOVA and for non-normally distributed data we used Kruskal-Wallis test as previously stated

Results

General Reads and Sequence Results

ARMS

We recovered 18 of the 24 deployed ARMS; the remaining six either accidentally detached from the substrate or were stolen. We processed the 100um, 500um, and sessile fractions from the remaining ARMS, for a total of 54 samples for subsequent lab work and down streaming data analysis (a list of sample name is available in a Supplemental Table S1-1).

From the 54 samples, we generated a total of 2,049,305 sequencing reads after post quality filtering and the exclusion of chimeras, representing a total of 3385 OTUs with a mean read length of 312.3bp. Rarefaction curves plotting OTU discovery vs sequencing depth showed that most of the samples saturated for OTU discovery (Fig. 1-1). After creating the OTU tables and performing the final filtering step, we rarefied the dataset to even sequencing depths of 11,000 reads/sample to ensure results weren't impacted by uneven sequencing depth. Three samples (Pari-100 μm, Sepa-500 μm and Kotok-100 μm) were discarded as the sample had fewer than 11,000 reads individually. This process resulted in a final dataset of 561,000 high quality reads representing 2907 OTUs that was used for all subsequent analysis, including calculation of alpha and beta diversity and all statistical analysis.

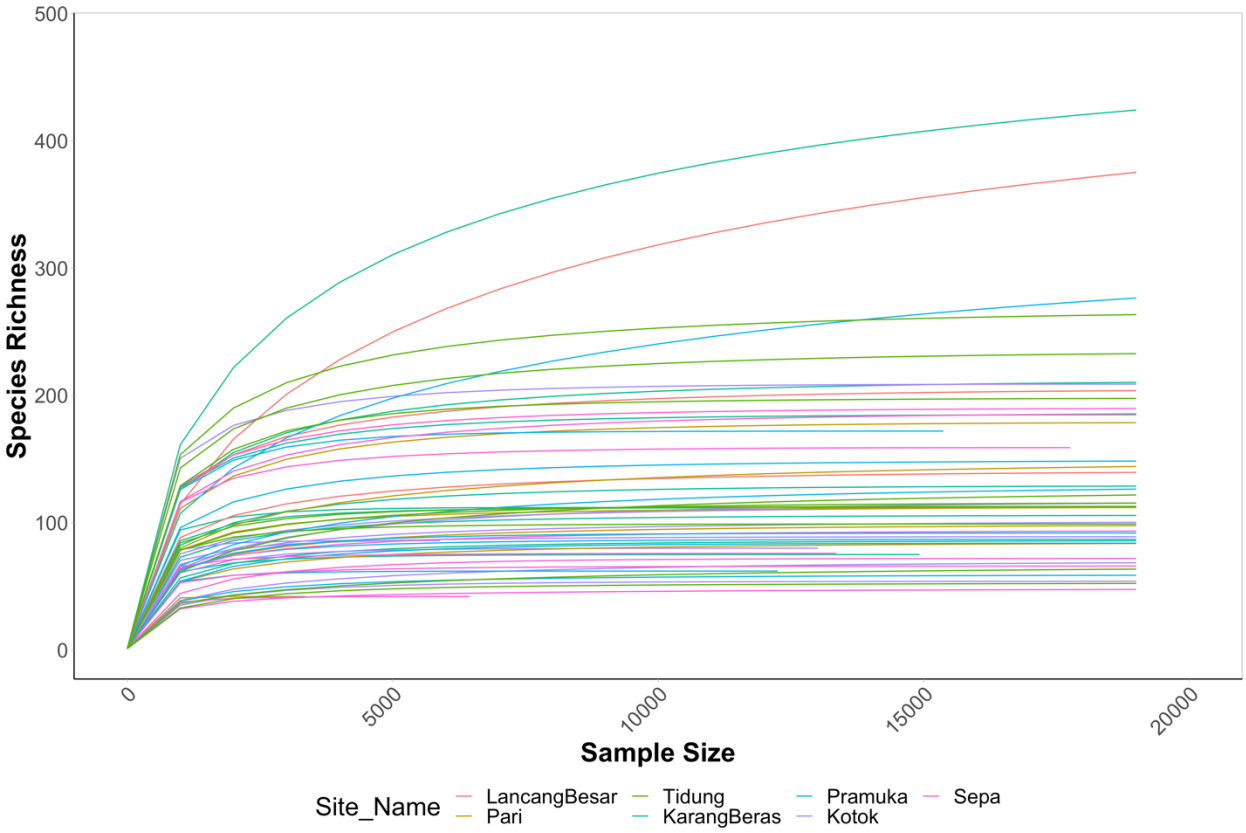


Figure 1-1. Rarefaction plots showing numbers of OTUs as a function of sequencing depth for each individual ARMS unit across sampling sites. Plot was truncated at 20.000 reads to show that each individual ARMS collected species at a saturation point.

eDNA

Following the removal of blanks and control samples, we collected 24 eDNA samples from across eight sites, from which we generated a total of 1,139,602 sequencing reads and 1,719 OTUs (a list of sample name is available in a Supplemental Table S1-2). Rarefaction curves show sequencing depth approached saturation for OTUs discovery for all of the samples (Fig. 1-2). To ensure comparisons of ARMS and eDNA were not impacted by sequencing depth,

we applied the same rarefaction threshold of 11,000 reads for each eDNA sample, resulting in a total of 231,000 reads and 1538 OTUs. We then use the resulting taxon table for all downstream biodiversity analysis and comparisons with ARMS data.

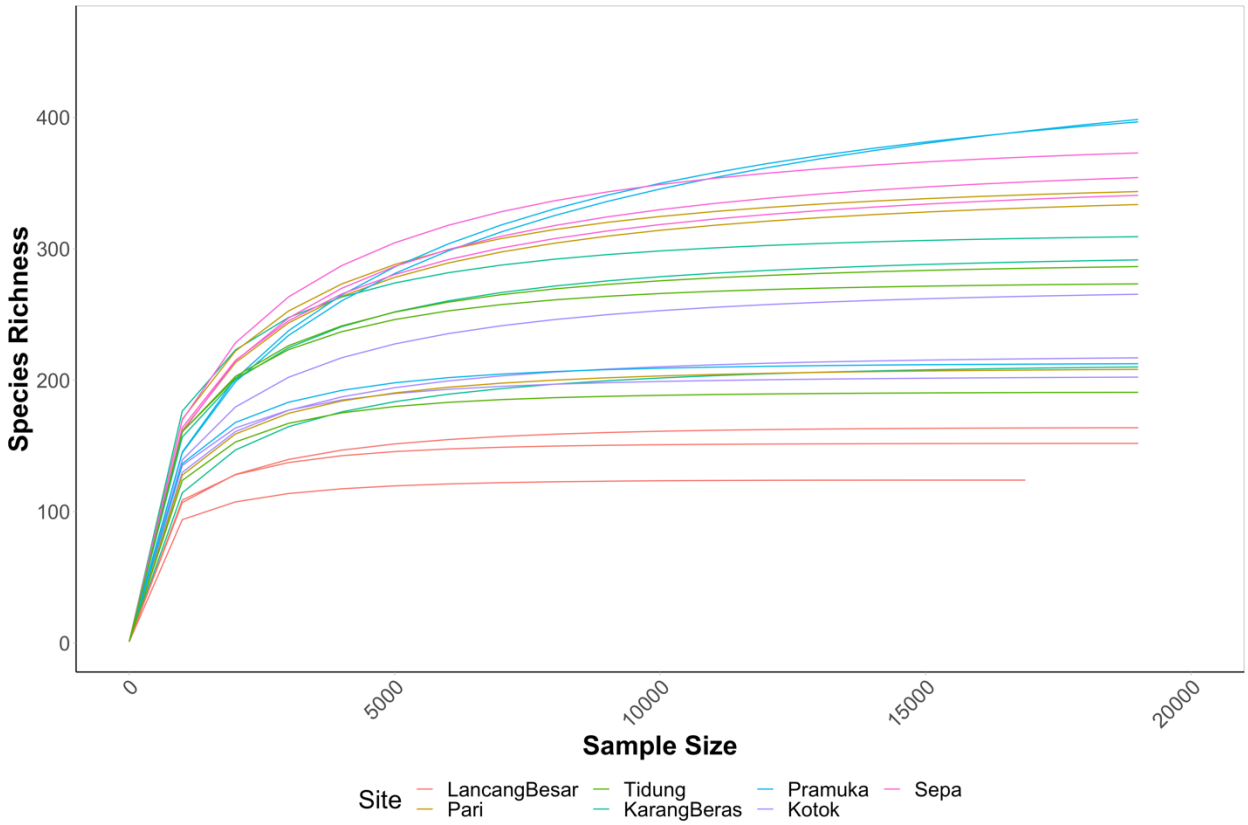


Figure 1-2. Rarefaction plots showing numbers of OTUs as a function of sequencing depth for each eDNA sample. Plot was truncated at 20.000 reads to show that each samples collected species at a saturation point.

Alpha Diversity

Across all sampling sites, ARMS had higher total OTU diversity (2907) than eDNA (1538) when combining data across the 100 µm, 500 µm and sessile fractions. Comparing individual ARMS fractions to eDNA revealed that the 100 µm fraction had similar total diversity

to eDNA (1550 vs 1538 OTUs, respectively), but that eDNA had more OTUs than the 500um (1052 OTUs) and Sessile (1159 OTUs) ARMS fraction.

Examining total OTU diversity on a site-by-site basis, ARMS recovered an average of 300.61 OTUs per site (Table 1; max= 531 at Lancang Besar, min=142 at Kotok). The 100 fractions recovered an average of 166.43 OTUs per site (max= 241 at Tidung, min=78 at Kotok), the 500 fraction recovered an average of 94.58 OTUs per site (max= 133 at Lancang Besar, min= 62 at Pramuka), and the sessile fraction an average of 111.28 OTUs per site (max= 347 at Karang Beras, min=45 at Sepa). In contrast, eDNA yielded an average of 255.14 OTUs per site (max= 322 at Sepa, min=151 at Lancang Besar). (Tab. 1-1.) (Fig. 1-3).

Table 1-1. Maximum and minimum number of OTUs across sites in both eDNA and ARMS. Kruskal-Wallis test demonstrated that only OTUs from eDNA samples were significantly different across sites.

Fractions	Site with highest OTUs (No. of OTUs)	Site with lowest OTUs (No. of OTUs)	Mean	Standard deviation	Kruskal-Wallis test
eDNA	Sepa (355)	Lancang Besar (123)	255.14	71.35	$p < 0.05$
Total ARMS	Lancang Besar (531)	Kotok (142)	300.61	97.26	$p > 0.05$
100	Tidung (241)	Kotok (78)	166.43	43.51	$p > 0.05$
500	Lancang Besar (133)	Pramuka (62)	94.58	18.58	$p > 0.05$
Sessile	Karang Beras (347)	Sepa (45)	111.27	83.97	$p > 0.05$

Wilcoxon signed-rank test showed that ARMS recovered significantly more OTUs than eDNA, when data is examined collectively, however when the data is split by individual fractions, we failed to reject the null hypothesis that both eDNA and ARMS have equal means of OTUs and the analysis suggest that eDNA has significantly more OTU than ARMS (paired Wilcoxon test, p

< 0.05). Further, a two-way ANOVA, testing the effects of sample types (eDNA or ARMS), site and their interaction on the number of OTUs, indicated a significant effect of the interaction between ARMS's location deployment – type of sampling methods used ($p < 0.001$) for 500 and Sessile fractions but not for 100 fraction and total ARMS (Table 1-2).

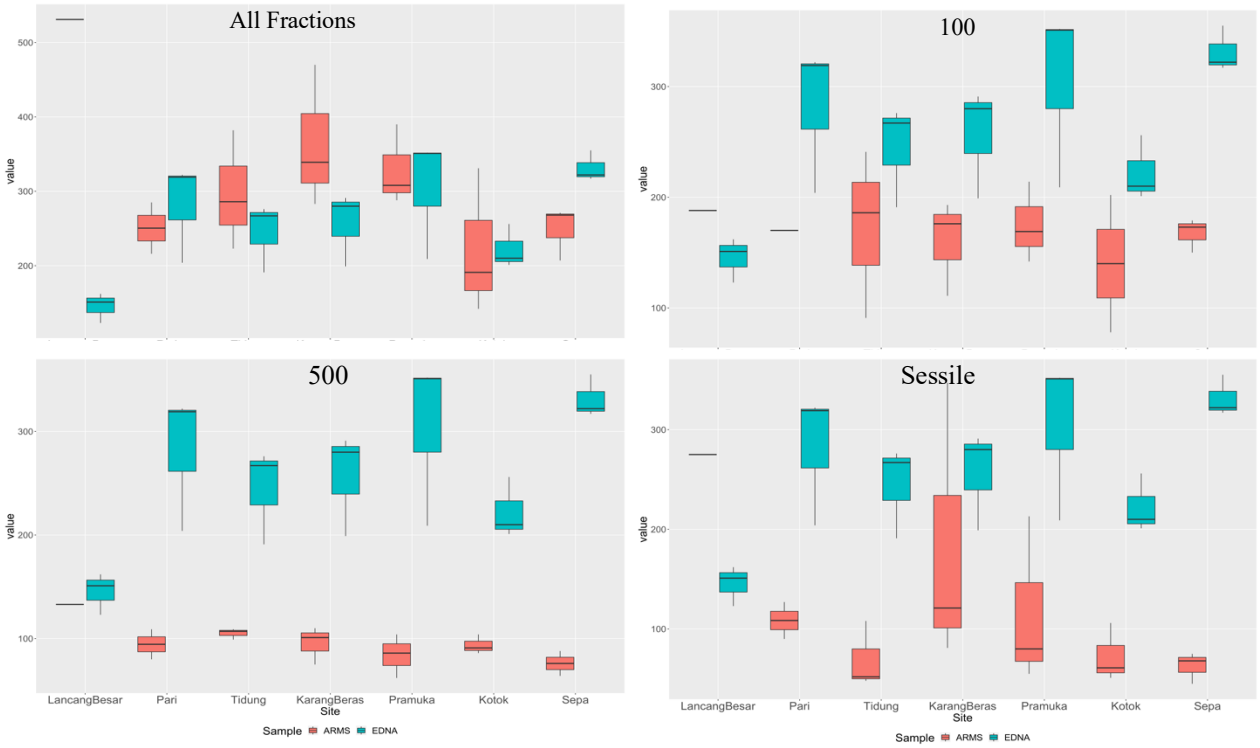


Figure 1-3. Boxplots of alpha diversity inferred by total number of OTUs generated using eDNA (Green) and ARMS (orange), showing that eDNA captures significantly more OTUs than ARMS (paired Wilcoxon test, $p < 0.05$).

Table 1-2. Model output for linear model with formula *number_of_OTUs~eDNAorARMS*Site*

100 Fraction Vs eDNA

	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
Sample	1	71456	71456	26.37	3.06e-05
Site	2	44509	114668	4.582	0.0352
Sample:Site	2	1322855	661428	26.431	4.01E-05
Residuals	12	300299	25025		

500 Fraction Vs eDNA

	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
Sample	1	393089	393089	98.29	3.93E-07
Site	2	89435	44717	11.18	0.001813
Sample:Site	2	118605	59303	14.83	0.000571
Residuals	12	47989	3999		

Sessile Fraction Vs eDNA

	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
Sample	1	4285616	4285616	171.254	1.83E-08
Site	2	229337	114668	4.582	0.0332
Sample:Site	2	1322855	661428	26.431	4.01E-05
Residuals	12	300299	25025		

Total ARMS Vs eDNA

	Df	Sum Sq	Mean Sq	F Value	Pr(>F)
Sample	1	20038	20038	5.179	0.031687
Site	6	37832	6305	1.630	0.180484
Sample:Site	6	128100	21350	5.519	0.000936
Residuals	25	96719	3869		

General Taxa Composition

The number of unassigned taxa varied both by site and by methods; however, on average 36.83 % of ARMS OTUs were unassigned (max= 54.44 % at Pari, min=23% at Tidung, sd = 10%). In contrast, unassigned OTUs were higher with eDNA; on average 56.32 % of OTUs could not be assigned to taxa (max= 73.78 % at Lancang Besar, min= 46.42 % at Sepa, sd = 9.37%). At thephylum level, the vast majority of ARMS reads across all sampling sites were dominated by Annelids, Arthropods, Cnidaria, Mollusca, Porifera, Rhodophyta. In contrast, Ascomycota, Bacillariophyta and Chlorophyta dominated eDNA reads. Overall, eDNA captured more taxonomic diversity associated with planktonic habitats while ARMS captured diversity largely associated with benthic habitats. For example, 3.4 % of eDNA reads were associated with planktonic diatoms in the order Bacillariophyta, while only 0.3 % reads from ARMS were from this same group. Similarly, red algae (Rhodophyta) were among the most common sequences recovered from ARMS, representing almost 13.92 % of total reads, whereas Rhodophyta was only 0.4 % to total reads in eDNA.

Examining taxonomic composition of individual ARMS fractions yielded similar results but varied by fraction. Taxonomic composition was relatively similar in the 100 and 500 fractions, with the sessile fraction having substantially higher reads of Rhodophyta and Porifera (± 16.79 % of Poriferan reads contributed to overall taxa composition in ARMS sample). Ascomycota, Bacilioryphyta and Chlorophyta have relatively zero to small reads contribution to overall taxa composition in 100µm and 500 µm ARMS fractions (only ± 0.007 and 0.004 % reads for Ascomycota and Chlorophyta respectively) however in contrast contributed 2.15 % (Ascomycota) and 25.75 % (Chlorophyta) being recovered in eDNA samples (Fig. 1-4).

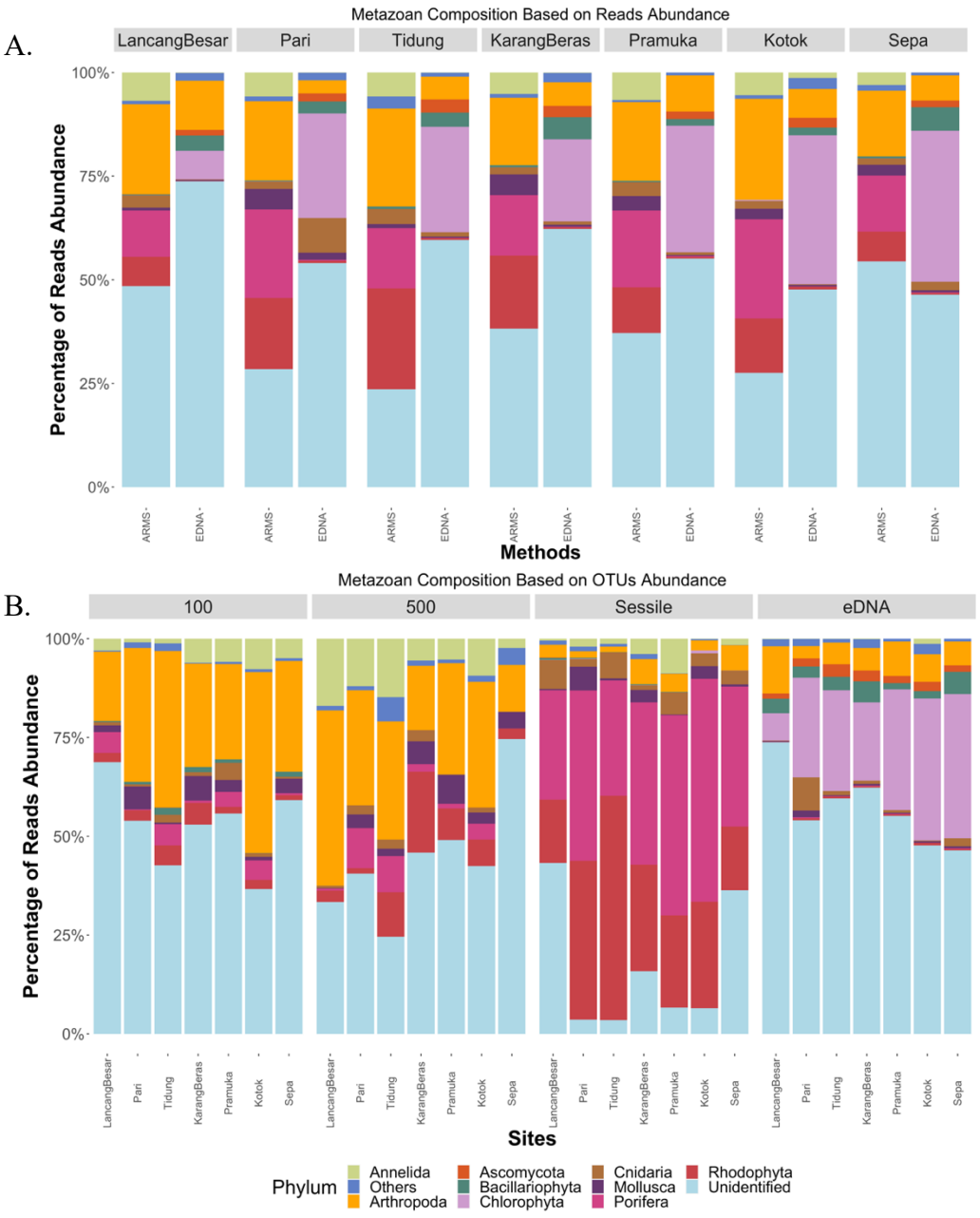


Figure 1-4. Contrasting differences in community composition between eDNA and ARMS method. Reads using ARMS were dominated by Porifera and Rhodophyta while in contrast Ascomycota and Bacillariophyta dominated reads collected using eDNA method. (A.) eDNA Vs ARMS total, (B.) eDNA Vs different fractions

Community Comparisons

Shared taxa

Despite sampling ARMS and eDNA from the same locations, comparison of total OTU diversity showed that only 132 (2.9%) of a combined 4313 OTUs were shared between the two datasets; 2775 OTUs (64.3%) were unique to ARMS and 1406 OTUs (32.6%) were unique to eDNA (Fig. 1-5A**.)**. These shared taxa represented 9.65 % of total sequencing reads, and were largely comprised of taxa from marine sponges (family Tedaniidae), polychaete worms (family Terebellidae) and "unassigned" taxa. The 2775 OTUs unique to ARMS represented 63.15 % of sequencing reads and were numerically dominated by red algae (family Peyssonneliaceae), marine sponges (family Microcionidae) and unassigned taxa. The 1406 OTUs unique to eDNA represented 27.19 % of sequencing reads and were numerically dominated by green algae (family Mamiellaceae), copepods (family Calanidae) and unassigned taxa.

Comparing eDNA to individual ARMS fractions showed that only 82 of 2615 (3.1%) taxa occurred in both the sessile fraction and eDNA (Fig. 1-5B.**)** and in general total OTU from each site is larger when eDNA method is used (Fig. 1-6). Sponges (family Pseudoceratinidae and Tedaniidae) and soft corals (family Nephtheidae) dominated the taxa shared among eDNA and the sessile ARMS fraction. Only 57 of 3031 (1.9%) of total OTUs were shared between eDNA and the 100 μm ARMS fraction shared (Fig. 1-5D**.)**. The most common shared taxa were diatoms (family Chaetocerotaceae), soft coral (family Nephtheidae) and polychaetes (framily Terebellidae). Lastly, only 52 taxa were shared between eDNA and the 500 μm ARMS fraction (Fig. 1-5C.). Shared taxa included diatoms (family Chaetocerotaceae), snails (family Haminoeidae) and polychaetes (family Terebellidae).

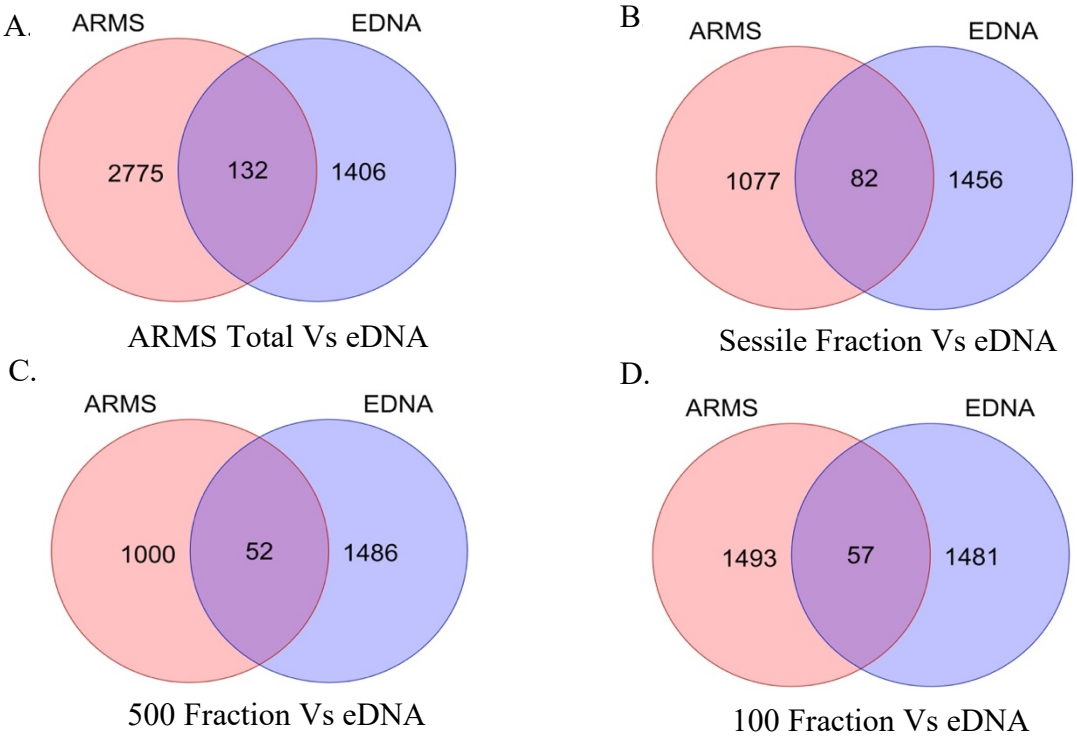


Figure 1-5. Venn diagram detailing number of shared and unique OTUs between fractions and eDNA. Samples from eDNA method have relatively more OTUs compared to ARMS especially when the samples are split by individual fractions. (A.) eDNA Vs ARMS total, (B.) eDNA Vs Sessile Fraction, (C.) eDNA Vs 500 Fraction, (D.) eDNA Vs 100 Fraction.

Phylogenetic Diversity Comparison Among Sites

Phylogenetic diversity (PD) varied across sampling sites. For eDNA, PD values ranged from a low of 34.54 in Lancang Besar to a high of 74.71 in Pramuka, with a mean of 55.42; however, the variation in PD among sites was marginally significant (Kruskal-Wallis test, $p = 0.06$ (Fig. 1-7.). For ARMS, PD values ranged from a low of 12.94 in Tidung to a high of 96.24 in Karang Beras, with a mean of 35.17; as with eDNA, no significant differences in PD were

observed across sample sites for the combined ARMS data or the 100 µm, 500 µm and sessile fractions (Kruskal-Wallis test, p = 0.2).

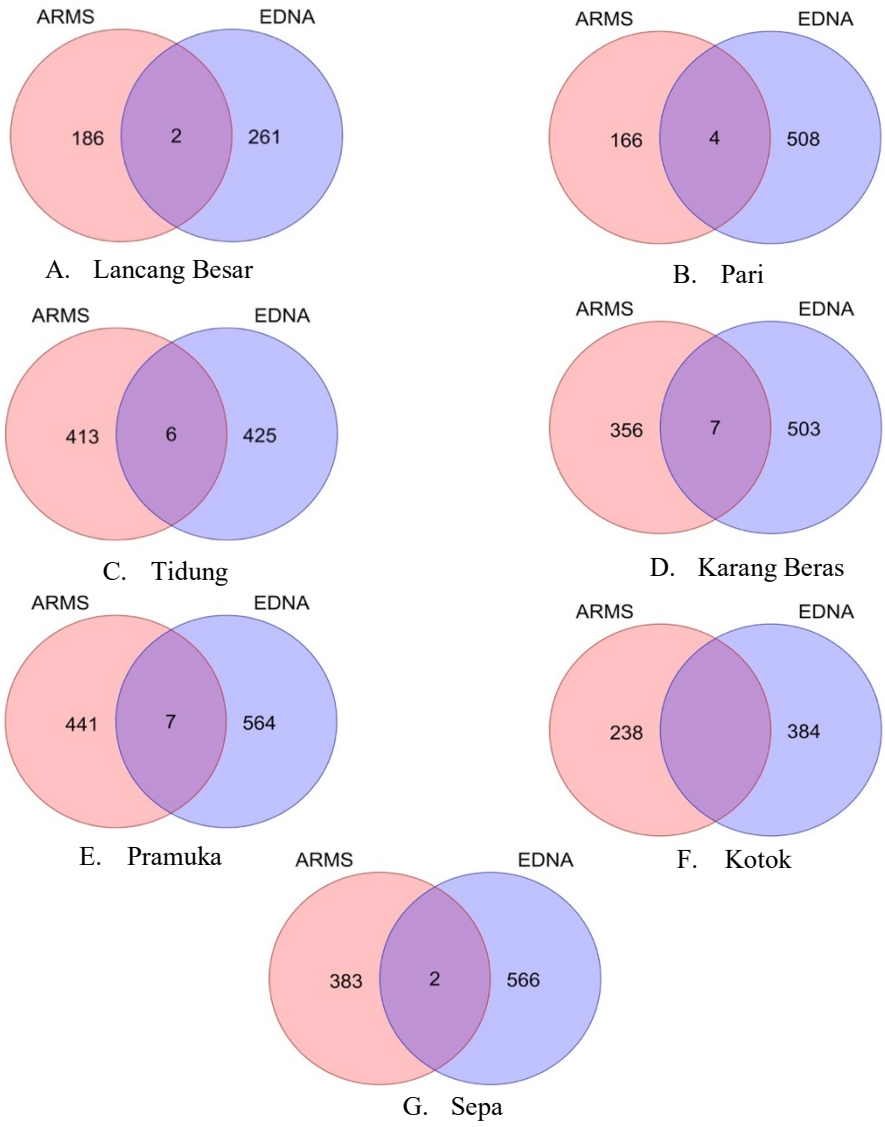


Figure 1-6. Venn diagram detailing number of shared and unique OTUs between ARMS-100 fraction and eDNA. Samples from eDNA across sites have relatively more OTUs compared to ARMS (A.) Lancang Besar, (B.) Pari, (C.) Tidung, (D.) Karang Beras, (E.) Pramuka, (F.) Kotok, (G.) Sepa

Although results showed no significant differences in PD across sites, PD was significantly higher in eDNA samples compared to ARMS when all fractions data were combined (Wilcoxon signed rank exact test, $p < 0.009$). This pattern was particularly pronounced when comparing eDNA to the 500 µm and sessile fractions (Wilcoxon signed rank exact test, $p = 0.0015$, green versus aqua color dots in Fig 6), but results were not significant when comparing eDNA and the 100 µm ARMS fraction (Wilcoxon signed rank exact test, $p > 0.05$, orange dots versus aqua color dots in Fig 6). Interestingly, the ARMS sample from Lancang Besar, the most polluted site, had significantly higher PD value compared to PD from eDNA sample from the same site (Fig. 1-6).

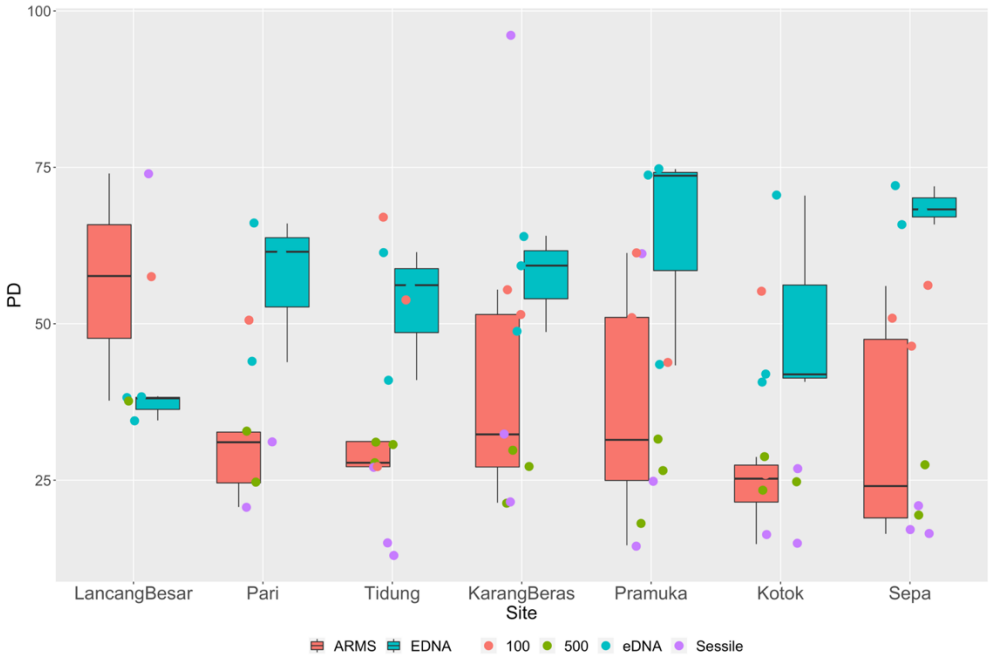


Figure 1-7. Differences in phylogenetic diversity between eDNA and ARMS method using collectively combined ARMS fraction data and eDNA. Taxa within a site collected with ARMS were generally more clustered compared to eDNA (i.e., lower value of phylogenetic diversity). Nonetheless, exception was occurred in Lancang Besar, site closest to Jakarta Bay with high pollution level.

Discussion

Metabarcoding approaches such as ARMS (Plaisance et al. 2011, Leray & Knowlton 2015, Pearman et al. 2019, Casey et al. 2021) and eDNA (Bohmann et al. 2014, Thomsen & Willerslev 2015, Madduppa et al. 2021, Ríos-Castro et al. 2021) are rapidly becoming important tools for ecosystem-based biodiversity monitoring in response to local and global threats to marine ecosystems (Plaisance et al. 2011, Bohmann et al. 2014, Aylagas et al. 2018), but are rarely used in concert. Direct comparison of ARMS and eDNA data demonstrates substantial variation in the total diversity and taxonomic composition of communities recovered from these two methods, with less than 3% of total taxonomic diversity being recovered by both methods. Variation in taxonomic diversity among metabarcoding primer sets is well-documented (Thomsen et al. 2012, Borrell et al. 2017, Aylagas et al. 2018, Valentin et al. 2020, Ellis et al. 2022), and ARMS were designed to capture benthic biodiversity (Leray & Knowlton 2015, Carvalho et al. 2019, Pearman et al. 2019, Casey et al. 2021), while marine eDNA studies disproportionately focus on fishes (e.g.,(Andruszkiewicz et al. 2017, Nakagawa et al. 2018, Jeunen et al. 2019, Closek et al. 2019, Madduppa et al. 2021, Gold, Sprague, et al. 2021, Marwayana et al. 2022). However, the minimal taxonomic overlap recovered by the COI metabarcoding primers from co-located sampling sites in this study is striking and suggests that ARMS and eDNA should be viewed as complementary rather than competing metabarcoding approaches to assessing marine biodiversity.

Despite both methods recovering a large number of OTUs, a large percentage of these were unidentified. Central to the effectiveness of metabarcoding approaches are reference databases (Mora et al. 2011, Hestetun et al. 2020, Bik 2021, De Santiago et al. 2022). Although shortcomings of reference databases are noted from eDNA studies on fishes in Indonesia

(Madduppa et al. 2021, Marques et al. 2021, Marwayana et al. 2022), the number of identified invertebrate metazoans identified in this study was substantially lower, with less than 80% of total OTUs identified even to phyla, compared to ~80% of OTUs being identified to species for fishes (Marwayana et al 2021). The inability to assign OTUs in this study, even to phyla, highlights the poor taxonomic representation of Indo-Pacific marine invertebrates in global references databases. This pattern is both a function of the exceptional diversity of the Coral Triangle (Allen 2008, Allen & Erdmann 2009b, Barber 2009, Froufe et al. 2016, Madduppa et al. 2021, Marwayana et al. 2022), particularly Indonesia, and biases in research effort that focus on less diverse temperate and tropical ecosystems (Donaldson et al. 2016, Bryan-Brown et al. 2017), highlighting a critical need for more biodiversity research in the global epicenter of marine biodiversity to support monitoring and conservation of marine biodiversity.

Total Diversity

Studies comparing metagenomic approaches to traditional measures of marine biodiversity are largely limited to comparisons of eDNA fish diversity to visual census techniques. These comparisons almost universally show that eDNA captures more local diversity, even when employing methods like site occupancy modelling that reduce total diversity by excluding taxa observed at low frequencies or in few replicates (Schmelzle & Kinziger 2016, Djurhuus et al. 2020, Gold, Sprague, et al. 2021). However, such comparisons of invertebrate metazoan communities are more challenging, because of challenges in sampling and identifying the diversity of marine invertebrates, particularly in biodiversity hotspots like the Coral Triangle.

Results of this study show that ARMS recovered more marine biodiversity as measured by total number of OTUs vs. eDNA across all sampling sites, with total OTU's from eDNA being similar to total OTU's from the 100 um ARMS fraction. However, in a site-by-site comparison, eDNA recovered significantly more OTUs at every site. Higher total diversity from ARMS while having less diversity on a site-by-site basis could be a function of the very different communities these two methods capture and the greater spatial heterogeneity in the benthic communities that encrust ARMS units. However, the ARMS processing protocol could also contribute to elevated numbers of OTUs as samples fractions are processed and sequenced individually (Ransome et al. 2017, Casey et al. 2021). As such, there is greater total sequencing depth for each ARMS unit, although rarefaction curves indicated that both ARMS and eDNA sequencing effort saturated. Whatever the cause, ARMS appear to capture more regional diversity than eDNA, viewed from the perspective of alpha diversity (e.g., OTUs).

Interestingly, in site-by-site comparisons, eDNA recovered significantly higher numbers of OTUs than ARMS. Because eDNA is mobile (Evans et al. 2017, Jeunen et al. 2019, Ely et al. 2021, Holman et al. 2022), it may be more effective in integrating biodiversity over a larger area. In contrast, ARMS samples represent the deposition and growth of encrusting taxa on a relatively small surface area over time, and the difference in spatial scale sampled by these methods may explain the higher diversity estimates from eDNA. Differences in diversity may also results from differences in DNA concentrations represented in these communities. Because ARMS samples are comprised of whole organisms, some taxa might have much higher DNA concentrations and dominate in the PCR and/or sequencing steps, reducing sensitivity to lower concentration DNA signatures. In contrast, genetic material in water samples are remnants of intracellular cell of the taxa from surrounding marine ecosystems and may be less prone to the

same magnitude of DNA concentration bias, potentially reducing competition in PCR and/or DNA sequencing phases, capturing more biodiversity.

Communities Vary by Method

Although ARMS recovered more total diversity across Kepulauan Seribu, and eDNA recovered more diversity at each sampling location, both methods recovered completely different taxa, indicating a high specificity of each method. As expected, ARMS were dominated by common benthic and cryptofauna such as annelids, arthropods, cnidaria, mollusca, porifera, as well as algae such as rhodophyta. In contrast, eDNA was more specific in capturing planktonic communities including marine fungi (Ascomycota), diatoms (Bacilioryphyta) and green algae (chlorophyta), which were likely phytoplankton. Although results showed significant variation among sites, as evidence by two-way ANOVA results, this taxonomic specificity by method was highly consistent across sites.

That COI metabarcoding of eDNA captured largely planktonic forms is unsurprising; previous marine eDNA studies indicate the predominance of planktonic organisms relative to benthic taxa (Pitz et al. 2020, Gallego et al. 2020). The taxa with highest read abundances, marine fungi (Ascomycota) and diatoms (Bacilioryphyta) are extremely common in planktonic communities and likely contribute represent a substantial amount of genetic material recovered from the water column. This is both because of their numerical abundance in the plankton where cell densities can achieve $3x10^8$ cells/L (Caputi et al. 2019), as well as because their DNA signal captured by eDNA sampling protocols would come largely from live organisms. In contrast, larger metazoans would be present in eDNA samples only as freely associated DNA so are more likely to have lower concentrations.

Similarly, the taxa recovered from ARMS and the high abundance of sponges (Porifera) and red algae (Rhodophyta) is unsurprising. Although we did not quantify the percent area covered, sponges and red algae were clearly the dominant encrusting taxa on ARMS plate surfaces, in some cases occupying easily > 50% ARMS plate surface area. Similar results have also been reported in an ARMS study located in Red Sea; Pearman et al. (2016) for example, found that Porifera, Bryozoa and Rhodophyta were the most abundant groups on the ARMS plates while Al-Rshaidat et al. (2016) suggested that Porifera drove clustering patterns in the PCoA results they produced. Other ARMS studies in Indonesia and French Polynesia show similar results (Ransome et al. 2017, Casey et al. 2021), suggesting that ARMS may favor colonization of specific taxa, regardless of the region where they are deployed. However, consistent detection of these three dominant taxa across multiple studies is likely a function of these taxa representing more biomass and higher DNA concentrations than other mobile taxa such as annelids, arthropods, and mollusks, leading to greater amplification and sequencing success. Combined, these results indicate that ARMS and eDNA methods capture different communities and that the choice of one method or another will depend greatly on the goals of a study, and that using both methods are needed to understand widest spectrum of diversity, especially in high biodiversity spots like Indonesia.

Community Similarities

Although taxonomic composition recovered by each method was highly specific, approximately 3% of the taxa were recovered in both ARMS and eDNA datasets, representing ~9.65 % of total sequence reads. The most common shared taxa included sponges (Tedaniidae), polychaetes (Terebellidae) and an unassigned group. Marine sponges are a versatile and

adaptable group that can survive a wide range of environmental conditions including highly polluted water (Done 1999, Bell & Barnes 2003, Venkateswara Rao et al. 2009). The proximity of Kepulauan Seribu to Jakarta results in the southern part of this archipelago having poor water quality (Uneputty & Evans 1997, Rees et al. 1999, Estradivari et al. 2009, van der Meij et al. 2010), and thus sponges are expected to be a prominent member of benthic communities. Moreover, the high volume of water filtering done by sponges may result in sloughing of cells or DNA into the water column, increasing the likelihood of detection through eDNA. Polychaetas are also an indicator of marine pollution (Hutchings 1998, Harlan K. Dean 2008). As such, recovery of sponges by both methods is unsurprising.

Although Polychaeta have not been studied in this region, polychaetes are also bioindicators of pollution in marine environments (REFS), so abundances are expected to be high. High abundance should increase signal in eDNA samples, increasing the likelihood of detecting polychaetes in both ARMS and eDNA samples. An alternative, but not mutually exclusive explanation is that high polychaetae sequence abundance in the eDNA dataset could represent the capture of larval forms during eDNA sampling. Although there are planktonic polychaetas, the family Terebellid is strictly benthic but has a meroplanktonic dispersive phase (Seitz & Schaffner 1995, Duchêne 2004, 2010). As with the fungi and diatoms, capture of the larval forms in eDNA could increase the probability of detecting Polychaetas, contributing to observation of shared taxa across both methods.

Phylogenetic Diversity

In addition to differences in total diversity and taxonomic composition, communities elucidated through ARMS and eDNA metabarcoding also varied greatly in taxonomic breadth.

Although ARMS recovered more total OTU diversity, phylogenetic diversity was significantly higher in eDNA samples in all but one site. Phylogenetic information derived from eDNA can inform the magnitude of local diversity and the levels of genetic clustering and dispersion in the sites of interests while also detecting species that are missed when visual census is implemented (Aylagas et al. 2018, Closek et al. 2019, Steyaert et al. 2020, DiBattista et al. 2020, Gold, Sprague, et al. 2021, Marwayana et al. 2022). In marine ecosystems, eDNA has captured taxa as diverse as plantae, chromists and animalia just from one single sample demonstrating its ability to detect a wide range of biodiversity, and the magnitude of diversity within eDNA samples seems to be only limited by local level of biodiversity, shedding and decay of genetic materials from each species and local transport processes (Weltz et al. 2017, Aylagas et al. 2018, Frühe et al. 2020).

In contrast, although ARMS capture a large number of species, suggested by the high number of OTUs in this study, taxa composition has less phylogenetic breadth because these structures represent a specialized niche that likely favors some taxa over others. ARMS were designed to provide habitat suitable for cryptofauna that live within the matrix of coral reefs (Plaisance et al. 2009, Leray & Knowlton 2015, Ransome et al. 2017, Wangensteen et al. 2018). As such, they will not capture the same range of biodiversity as eDNA that can theoretically capture all organisms within an ecosystem, although this method does have biases (Nichols et al. 2018, Ruppert et al. 2019, Mathieu et al. 2020). Moreover, because ARMS are a living community, the composition of that community is dependent on the biological interactions among taxa on each plate (Roman 2006, Vieira et al. 2018). As such, high abundance of some taxa might preclude settlement and growth of other taxa, potentially limiting total diversity on ARMS plates and promoting phylogenetic clustering that lowers overall phylogenetic diversity.

These differences likely explain why eDNA samples from all sites have, on average, twice the phylogenetic diversity of ARMS.

Conclusions

Given the diversity of challenges impacting marine biodiversity worldwide and the challenges associated with monitoring marine ecosystems (Brooks et al. 2006, Hoegh-Guldberg 2011, Taberlet et al. 2012, Bohmann et al. 2014, Hernández-Delgado 2015, Aylagas et al. 2018), metabarcoding approaches like ARMS and eDNA will continue to grow in popularity, particularly in global biodiversity hotspots like Indonesia where science capacity doesn't match local diversity (Barber et al. 2014). As the cost of sequencing continues to decrease and methods, bioinformatic pipelines, and reference databases improve, metabarcoding approaches such as ARMS and eDNA will likely become a staple of marine biodiversity researchers and resource managers, given the greater scope and resolution compared to visual methods (Bohmann et al. 2014, Yamamoto et al. 2017, Andújar et al. 2018, Ríos-Castro et al. 2021). However, the results of this study show that which methods used will depend on the goals of the study.

The advantage of eDNA is that it can capture the widest range of marine biodiversity and integrates signal over a larger area. Moreover, although we only employed COI to create directly comparable datasets, the use of different barcoding markers such as 12S, 18S (Casey et al. 2021), and 16S (Pearman et al. 2019) can allow researchers to focus on specific taxa such as fishes, metazoans, and microbes, respectively. In contrast, ARMS provide biodiversity information that are particularly aimed at understanding taxa composition aggregated in specific areas. Although costs and resources in implementing sampling efforts might be limiting factors, both methods should be considered if observation is aimed to get most comprehensive biodiversity information

possible. Further, as these two methods are increasingly utilized as a routine monitoring tool, it is critical to conduct more comparative studies to explore how these two methods vary in the detection of marine species.

Although our study demonstrates that both ARMS and eDNA metabarcoding are powerful tools to catalogue marine biodiversity in biodiversity hotspots like the Coral Triangle, maximizing the information obtained from these approaches requires more complete species databases. Although changes in marine communities can be documented focused only on OTUs, from a conservation perspective, it is important to identify OTUs to species. Incomplete databases lead to poor species assignment, impacting our understanding of local diversity (Deiner et al. 2017, Machida et al. 2017). Advancing metabarcoding as a tool for marine conservation will require focused efforts to develop well curated and accurate reference databases (Schenekar et al. 2020), particularly in mega diverse region like Indonesia and Coral Triangle (Veron et al. 2009, Foale et al. 2013, Bowen et al. 2013). Moreover, these databases should target regional fauna to produce the best species assignments (Gold, Curd, et al. 2021). Examples of such efforts include the Moorea Biocode Project (https://ocean.si.edu/ecosystems/coral-reefs/moorea-biocode-project) that catalogued all marine biodiversity on the island of Moorea, French Polynesia. The outcome of the project and the database it produces have provided a crucial biodiversity information for numerous research endeavors both locally and globally (Plaisance et al. 2009, Beaman & Cellinese 2012, Leray et al. 2013, 2015, Geller et al. 2013, Andersen et al. 2019), allowing local and international researchers to work together to monitor and advance biodiversity conservation in marine ecosystems across the globe.

Supplemental Tables and Figures

Supplemental Table S1-1. Deployment location and status following recovery for COI-ARMS samples.

Sample ID	Site Name	ARMS Name	Fraction_Size	Recovery Status
SBDI1ADMSO	Bidadari	SLCB1A	Sessile	Lost
SBDI1A500	Bidadari	SLCB1A	500	Lost
SBDI1A100	Bidadari	SLCB1A	100	Lost
SBDI1BDMSO	Bidadari	SLCB1B	Sessile	Lost
SBDI1B500	Bidadari	SLCB1B	500	Lost
SBDI1B100	Bidadari	SLCB1B	100	Lost
SBDI1CDMSO	Bidadari	SLCB1C	Sessile	Lost
SBDI1C500	Bidadari	SLCB1C	500	Lost
SBDI1C100	Bidadari	SLCB1C	100	Lost
SKBS1ADMSO	Karang Beras	SKBS1A	Sessile	Recovered
SKBS1BDMSO	Karang Beras	SKBS1B	Sessile	Recovered
SKBS1A100	Karang Beras	SKBS1A	100	Recovered
SKBS1B100	Karang Beras	SKBS1B	100	Recovered
SKBS1B500	Karang Beras	SKBS1B	500	Recovered
SKBS1C500	Karang Beras	SKBS1C	500	Recovered
SKBS1A500	Karang Beras	SKBS1A	500	Recovered
SKBS1C100	Karang Beras	SKBS1C	100	Recovered
SKBS1CDMSO	Karang Beras	SKBS1C	Sessile	Recovered
SKOT1A500	Kotok	SKOT1A	500	Recovered
SKOT1BDMSO	Kotok	SKOT1B	Sessile	Recovered
SKOT1ADMSO	Kotok	SKOT1A	Sessile	Recovered
SKOT1CDMSO	Kotok	SKOT1C	Sessile	Recovered
SKOT1A100	Kotok	SKOT1A	100	Recovered
SKOT1B500	Kotok	SKOT1B	500	Recovered
SKOT1C500	Kotok	SKOT1C	500	Recovered
SKOT1C100	Kotok	SKOT1C	100	Recovered
SKOT1B100	Kotok	SKOT1B	100	Recovered
SLCB1ADMSO	Lancang Besar	SLCB1A	Sessile	Recovered
SLCB1A500	Lancang Besar	SLCB1A	500	Recovered
SLCB1A100	Lancang Besar	SLCB1A	100	Recovered
SLCB1BDMSO	Lancang Besar	SLCB1B	Sessile	Lost
SLCB1B500	Lancang Besar	SLCB1B	500	Lost

Supplemental Table S1-1 *(continued)*

SampleID	Site Name	ARMS Name	Fraction Size	Recovery Status
SLCB1CDMSO	Lancang Besar	SLCB1C	Sessile	Lost
SLCB1C500	Lancang Besar	SLCB1C	500	Lost
SLCB1C100	Lancang Besar	SLCB1C	100	Lost
SPAR1CDMSO	Pari	SPAR1C	Sessile	Lost
SPAR1C500	Pari	SPAR1C	500	Lost
SPAR1C100	Pari	SPAR1C	100	Lost
SPAR1ADMSO	Pari	SPAR1A	Sessile	Recovered
SPAR1BDMSO	Pari	SPAR1B	Sessile	Recovered
SPAR1A100	Pari	SPAR1A	100	Recovered
SPAR1B100	Pari	SPAR1B	100	Recovered
SPAR1B500	Pari	SPAR1B	500	Recovered
SPAR1A500	Pari	SPAR1A	500	Recovered
STDN1C100	Pramuka	STDN1C	100	Recovered
SPRM1CDMSO	Pramuka	SPRM1C	Sessile	Recovered
SPRM1C500	Pramuka	SPRM1C	500	Recovered
SPRM1ADMSO	Pramuka	SPRM1A	Sessile	Recovered
SPRM1BDMSO	Pramuka	SPRM1B	Sessile	Recovered
SPRM1B100	Pramuka	SPRM1B	100	Recovered
SPRM1A500	Pramuka	SPRM1A	500	Recovered
SPRM1B500	Pramuka	SPRM1B	500	Recovered
SPRM1A100	Pramuka	SPRM1A	100	Recovered
SSEP1C100	Sepa	SSEP1C	100	Recovered
SSEP1CDMSO	Sepa	SSEP1C	Sessile	Recovered
SSEP1ADMSO	Sepa	SSEP1A	Sessile	Recovered
SSEP1BDMSO	Sepa	SSEP1B	Sessile	Recovered
SSEP1C500	Sepa	SSEP1C	500	Recovered
SSEP1A500	Sepa	SSEP1A	500	Recovered
SSEP1B100	Sepa	SSEP1B	100	Recovered
SSEP1A100	Sepa	SSEP1A	100	Recovered
SSEP1B500	Sepa	SSEP1B	500	Recovered
STDN1B100	Tidung	STDN1B	100	Recovered
STDN1C500	Tidung	STDN1C	500	Recovered
STDN1B500	Tidung	STDN1B	500	Recovered
STDN1CDMSO	Tidung	STDN1C	Sessile	Recovered
SPRM1C100	Tidung	SPRM1C	100	Recovered

STDN1A100	Tidung	STDN1A	100	Recovered
STDN1ADMSO	Tidung	STDN1A	Sessile	Recovered
STDN1BDMSO	Tidung	STDN1B	Sessile	Recovered
STDN1A500	Tidung	STDN1A	500	Recovered
SLCB1B100	Lancang Besar	SLCB1B	100	Lost

Supplemental Table S1-2. Sampling location for COI-eDNA samples.

Sample Name	Sampling Location
CO1_BDR.1.S1.L001	Bidadari
CO1_BDR.2.S2.L001	Bidadari
CO1_BDR.3.S3.L001	Bidadari
CO1_KBS.1.S4.L001	KarangBeras
CO1_KBS.2.S5.L001	KarangBeras
CO1_KBS.3.S6.L001	KarangBeras
CO1_KOT.1.S7.L001	Kotok
CO1_KOT.2.S8.L001	Kotok
CO1_KOT.3.S9.L001	Kotok
CO1_LCB.1.S10.L001	LancangBesar
CO1_LCB.2.S11.L001	LancangBesar
CO1_LCB.3.S12.L001	LancangBesar
CO1_PAR.1.S13.L001	Pari
CO1_PAR.2.S14.L001	Pari
CO1_PAR.3.S15.L001	Pari
CO1_PRM.1.S16.L001	Pramuka
CO1_PRM.2.S17.L001	Pramuka
CO1_PRM.3.S18.L001	Pramuka
CO1_SEP.1.S19.L001	Sepa
CO1_SEP.2.S20.L001	Sepa
CO1_SEP.3.S21.L001	Sepa
CO1_TDN.1.S22.L001	Tidung
CO1_TDN.2.S23.L001	Tidung
CO1_TDN.3.S24.L001	Tidung

References

Al-Rshaidat MMD, Snider A, Rosebraugh S, Devine AM, Devine TD, Plaisance L, Knowlton N, Leray M, Al-Rshaidat MMD, Snider A, Rosebraugh S, Devine AM, Devine TD, Plaisance L, Knowlton N, Leray M (2016) Deep COI sequencing of standardized benthic samples unveils overlooked diversity of Jordanian coral reefs in the northern Red Sea 1. Genome 59:724–737

Allen GR (2008) Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. Aquatic Conservation: Marine and Freshwater Ecosystems 18:541–556

Allen GR, Adrim M (2003) Review: Article Coral Reef Fishes of Indonesia. Zoological Studies 42:1–72

Allen GR, Erdmann M V. (2009a) Reef fishes of the Bird's Head Peninsula, West Papua, Indonesia. Check List 5:587–628

Allen GR, Erdmann M V (2009b) Reef fishes of the Bird ' s Head Peninsula , West Papua , Indonesia. Check List 5:587–628

Alongi DM (2014) Carbon sequestration in mangrove forests. http://dx.doi.org/104155/cmt1220 3:313–322

Andersen JC, Oboyski P, Davies N, Charlat S, Ewing C, Meyer C, Krehenwinkel H, Lim JY, Noriyuki S, Ramage T, Gillespie RG, Roderick GK (2019) Categorization of species as native or nonnative using DNA sequence signatures without a complete reference library. Ecological Applications 29:e01914

Andruszkiewicz EA, Sassoubre LM, Boehm AB (2017) Persistence of marine fish environmental DNA and the influence of sunlight. PLOS ONE 12:e0185043

Andújar C, Arribas P, Yu DW, Vogler AP, Emerson BC (2018) Why the COI barcode should be

the community DNA metabarcode for the metazoa. Molecular Ecology

Aylagas E, Borja Á, Muxika I, Rodríguez-Ezpeleta N (2018) Adapting metabarcoding-based benthic biomonitoring into routine marine ecological status assessment networks. Ecological Indicators 95:194–202

Aylward B, Barbier EB (1992) Valuing environmental functions in developing countries. Biodiversity and Conservation 1:34–50

Bairoliya S, Zhi Xiang JK, Cao B (2022) Extracellular DNA in Environmental Samples: Occurrence, Extraction, Quantification, and Impact on Microbial Biodiversity Assessment. Applied and Environmental Microbiology 88

Bakker J, Wangensteen OS, Chapman DD, Boussarie G, Buddo D, Guttridge TL, Hertler H, Mouillot D, Vigliola L, Mariani S (2017) Environmental DNA reveals tropical shark diversity in contrasting levels of anthropogenic impact. Scientific Reports 7:1–11

Barber PH (2009) The challenge of understanding the Coral Triangle biodiversity hotspot. Journal of Biogeography 36:1845–1846

Barber PH, Ablan-Lagman MCA, Ambariyanto A, Berlinck RGSG, Cahyani D, Crandall ED, Ravago-Gotanco R, Juinio-Meñez MA, Mahardika INGNIN, Shanker K, Starger CJ, Toha AHA, Anggoro AW, Willette DA (2014) Advancing biodiversity research in developing countries: the need for changing paradigms. Bulletin of Marine Science 90:187–210

Baum G, Januar HI, Ferse SSCA, Kunzmann A, Wilkinson C, Burke L, Reytar K, Spalding M, Perry A, Williams G, Gove J, Eynaud Y, Zgliczynski B, Sandin S, Graham N, Cinner J, Norström A, Nyström M, Sale P, Agardy T, Ainsworth C, Feist B, Bell J, Christie P, Blackburn S, Marques C, Pelling M, Blackburn S, Meij S Van der, Suharsono, Hoeksema B, Cleary D, Polónia A, Renema W, Hoeksema B, Wolstenholme J, Tuti Y, Fauzi A,

Buchary E, Rees J, Setiapermana D, Sharp V, Weeks J, Williams T, Williams T, Rees J, Setiapermana D, Cleary D, Suharsono, Hoeksema B, Cleary D, DeVantier L, Giyanto, Vail L, Manto P, Voogd N de, Madduppa H, Subhan B, Suparyani E, Siregar A, Arafat D, Tarigan S, Sluiter C, Umbgrove J, Bengen D, Knight M, Dutton I, Wolanski E, Moll H, Suharsono, Abrar M, Zamani N, Madduppa H, Ferse SSCA, Aktani U, Palm H, Tomascik T, Suharsono, Mah A, Ginsburg R, Kohler K, Gill S, Edinger E, Risk M, Froese R, Pauly D, Fabricius K, Cooper T, Humphrey C, Uthicke S, De'ath G, Davidson J, Storlazzi C, Field M, Bothner M, Clarke K, Gorley R, Clarke K, Ainsworth M, Clarke K, Green R, Bray J, Curtis J, Anderson M, Clarke K, Somerfield P, Gorley R, Veron J, Rogers C, Bak R, Povel G, Done T, Karlson R, Hurd L, Rachello-Dolmen P, Cleary D, Zaneveld J, Verstappen H, Verstappen H, Pelling M, Blackburn S, Pelling M, Blackburn S, Damar A, Colijnz F, Hesse K, Wardiatno Y, Bell P, Lapointe B, Elmetri I, Fabricius K, Dikou A, Woesik R Van, Huang D, Tun K, Chou L, Todd P, Koropitan A, Ikeda M, Damar A, Yamanaka Y, Hosono T, Su C, Delinom R, Umezawa Y, Toyota T, Kaneko S, Farhan A, Lim S, Rinawati, Koike T, Koike H, Kurumisawa R, Ito M, Sakurai S, Uneputty P, Evans S, Bruno J, Selig E, Selig E, Harvell C, Bruno J, Willis B, Page C, Casey K, Phinney J, Hoegh-Guldberg O, Kleypas J, Skirving W, Strong A, Murdoch T, Aronson R, Edmunds P, Berkelmans R, De'ath G, Kininmonth S, Skirving W, Bell P, Elmetri I, Lapointe B, Fox H, Caldwell R, Wenger A, McCormick M, Endo G, McLeod I, Kroon F, Jones G, Hughes T, Rodrigues M, Bellwood D, Ceccarelli D, Hoegh-Guldberg O, McCook L, Rasher D, Engel S, Bonito V, Fraser G, Montoya J, Hay M, Norström A, Nyström M, Lokrantz J, Folke C, Ban S, Graham N, Connolly S, Crain C, Kroeker K, Halpern B, Fabricius K, Dubinsky Z, Stambler N, Dutton I, Resosudarmo E (2015) Local and Regional Impacts of Pollution on Coral Reefs along the

Thousand Islands North of the Megacity Jakarta, Indonesia. (PA Todd, Ed.). PloS one 10:e0138271

Baynes T, M. Szmant A (1989) Effect of current on the sessile benthic community structure of an artificial reef. Bulletin of Marine Science 44:545–566

Beaman RS, Cellinese N (2012) Mass digitization of scientific collections: New opportunities to transform the use of biological specimens and underwrite biodiversity science. ZooKeys 209:7

Bell JJ, Barnes DKA (2003) Effect of Disturbance on Assemblages: An Example Using Porifera. Biological Bulletin

Bell JD, Kronen M, Vunisea A, Nash WJ, Keeble G, Demmke A, Pontifex S, Andrefouet S (2009) Planning the use of fish for food security in the Pacific. Marine Policy 33:64–76

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. Trends in Ecology & Evolution 22:148–155

Bik HM (2021) Just keep it simple? Benchmarking the accuracy of taxonomy assignment software in metabarcoding studies. Molecular Ecology Resources 21:2187–2189

Bishop RC (1993) Economic-Efficiency, Sustainability, and Biodiversity. Ambio 22:69–73

Blackman RC, Ho HC, Walser JC, Altermatt F (2022) Spatio-temporal patterns of multi-trophic biodiversity and food-web characteristics uncovered across a river catchment using environmental DNA. Communications Biology 2022 5:1 5:1–11

Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, Yu DW, Bruyn M de (2014) Environmental DNA for wildlife biology and biodiversity monitoring. Trends in Ecology & Evolution 29:358–367

Bohnsack JA, Harper DE, McClellan DB, Hulsbeck M (1994) Effects of reef size on colonization and assemblage structure of fishes at artificial reefs off southeastern Florida, U.S.A. Bulletin of Marine Science 55:796–823

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Silva R Da, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang K Bin, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik A V., Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, Hooft JJJ van der, Vargas F, Vázquez-Baeza Y, Vogtmann E, Hippel M von, Walters W, Wan Y, Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology

Bonfil R, Palacios-Barreto P, Vargas OUM, Ricaño-Soriano M, Díaz-Jaimes P (2021) Detection of critically endangered marine species with dwindling populations in the wild using eDNA gives hope for sawfishes. Marine Biology 168:1–12

Borrell YJ, Miralles L, Huu H Do, Mohammed-Geba K, Garcia-Vazquez E (2017) DNA in a bottle - Rapid metabarcoding survey for early alerts of invasive species in ports. PLoS ONE

Bowen BW, Rocha L a., Toonen RJ, Karl S a., Laboratory T (2013) The origins of tropical marine biodiversity. Trends in Ecology & Evolution 28:359–366

Brooks TM, Mittermeier RA, Fonseca GAB da, Gerlach J, Hoffmann M, Lamoreux JF, Mittermeier CG, Pilgrim JD, Rodrigues ASL (2006) Global biodiversity conservation priorities. Science (New York, NY) 313:58–61

Brown CJ (2005) Epifaunal colonization of the Loch Linnhe artificial reef: influence of substratum on epifaunal assemblage structure. Biofouling 21:73–85

Bruno JF, Selig ER (2007) Regional decline of coral cover in the Indo-Pacific: timing, extent, and subregional comparisons. PloS one 2:e711

Bryan-Brown DN, Brown CJ, Hughes JM, Connolly RM (2017) Patterns and trends in marine population connectivity research. Marine Ecology Progress Series 585:243–256

Burke L, Reytar K, Spalding M, Perry A (2011) Reefs at Risk Revisited.

Burt J, Bartholomew A, Bauman A, Saif A, Sale PF (2009) Coral recruitment and early benthic community development on several materials used in the construction of artificial reefs and breakwaters. Journal of Experimental Marine Biology and Ecology 373:72–78

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: Architecture and applications. BMC Bioinformatics 10:421

Caputi L, Carradec Q, Eveillard D, Kirilovsky A, Pelletier E, Pierella Karlusich JJ, Rocha Jimenez Vieira F, Villar E, Chaffron S, Malviya S, Scalco E, Acinas SG, Alberti A, Aury

JM, Benoiston AS, Bertrand A, Biard T, Bittner L, Boccara M, Brum JR, Brunet C, Busseni G, Carratalà A, Claustre H, Coelho LP, Colin S, D'Aniello S, Silva C Da, Core M Del, Doré H, Gasparini S, Kokoszka F, Jamet JL, Lejeusne C, Lepoivre C, Lescot M, Lima-Mendez G, Lombard F, Lukeš J, Maillet N, Madoui MA, Martinez E, Mazzocchi MG, Néou MB, Paz-Yepes J, Poulain J, Ramondenc S, Romagnan JB, Roux S, Salvagio Manta D, Sanges R, Speich S, Sprovieri M, Sunagawa S, Taillandier V, Tanaka A, Tirichine L, Trottier C, Uitz J, Veluchamy A, Veselá J, Vincent F, Yau S, Kandels-Lewis S, Searson S, Dimier C, Picheral M, Bork P, Boss E, Vargas C de, Follows MJ, Grimsley N, Guidi L, Hingamp P, Karsenti E, Sordino P, Stemmann L, Sullivan MB, Tagliabue A, Zingone A, Garczarek L, d'Ortenzio F, Testor P, Not F, d'Alcalà MR, Wincker P, Bowler C, Iudicone D (2019) Community-Level Responses to Iron Availability in Open Ocean Plankton Ecosystems. Global Biogeochemical Cycles 33:391–419

Carstensen J, Obst M, Høkedal J, Danovaro R, Carugati L, Berzano M, Cahill AE, Carvalho S, Chenuil A, Corinaldesi C, Cristina S, David R, Dell A, Dzhembekova N, Garcés E, Gasol JM, Goela P, Féral J-P, Ferrera I, Forster RM, Kurekin AA, Rastelli E, Marinova V, Miller PI, Moncheva S, Newton A, Pearman JK, Pitois SG, Reñé A, Rodríguez-Ezpeleta N, Saggiomo V, H Simis SG, Borja A, Kurekin AA, Cristina S, Ferrera I, Dell'Anno A, Reñé A, Forster RM, Cochrane SKJ, Saggiomo V, Féral J-P, Gasol JM, Martire M Lo, Greco S, Pitois SG, Stefanova K, Dzhembekova N, Berzano M, Goela P, Corinaldesi C, Rodríguez-Ezpeleta N, Newton A, Miller PI, Garcés E, Wilson C, Pearman JK, Carugati L, Marinova V, Danovaro R, Rastelli E, David R, Cahill AE, Moncheva S, Simis SGH, Mangoni O, Chenuil A (2016) Implementing and Innovating Marine Monitoring Approaches for Assessing Marine Environmental Status. Frontiers in Marine Science 3:1–25

Carvalho S, Aylagas E, Villalobos R, Kattan Y, Berumen M, Pearman JK (2019) Beyond the visual: using metabarcoding to characterize the hidden reef cryptobiome. Proceedings of the Royal Society B: Biological Sciences 286:20182697

Casey JM, Ransome E, Collins AG, Mahardini A, Kurniasih EM, Sembiring A, Schiettekatte NMD, Cahyani NKD, Anggoro AW, Moore M, Uehling A, Belcaid M, Barber PH, Geller JB, Meyer CP (2021) DNA metabarcoding marker choice skews perception of marine eukaryotic biodiversity. Environmental DNA 00:1–18

Cleary DFR (2017) Linking fish species traits to environmental conditions in the Jakarta Bay-Pulau Seribu coral reef system. Marine Pollution Bulletin 122:259–262

Closek CJ, Santora JA, Starks HA, Schroeder ID, Andruszkiewicz EA, Sakuma KM, Bograd SJ, Hazen EL, Field JC, Boehm AB (2019) Marine Vertebrate Biodiversity and Distribution Within the Central California Current Using Environmental DNA (eDNA) Metabarcoding and Ecosystem Surveys. Frontiers in Marine Science 6:732

Coll M, Libralato S, Tudela S, Palomera I, Pranovi F (2008) Ecosystem Overfishing in the Ocean. PLOS ONE 3:e3881

Collins RA, Wangensteen OS, O'Gorman EJ, Mariani S, Sims DW, Genner MJ (2018) Persistence of environmental DNA in marine systems. Communications Biology 2018 1:1 1:1–11

Côté IM, Darling ES (2010) Rethinking Ecosystem Resilience in the Face of Climate Change. PLOS Biology 8:e1000438

Crain CM, Halpern BS, Beck MW, Kappel C V. (2009) Understanding and Managing Human Threats to the Coastal Marine Environment. Annals of the New York Academy of Sciences 1162:39–62

Crain CM, Kroeker K, Halpern BS (2008) Interactive and cumulative effects of multiple human stressors in marine systems. Ecology Letters 11:1304–1315

Curd EE, Gold Z, Kandlikar GS, Gomer J, Ogden M, O'Connell T, Pipes L, Schweizer TM, Rabichow L, Lin M, Shi B, Barber PH, Kraft N, Wayne R, Meyer RS (2019) Anacapa Toolkit: An environmental DNA toolkit for processing multilocus metabarcode datasets. Methods in Ecology and Evolution

D'Alessandro S, Mariani S (2021) Sifting environmental DNA metabarcoding data sets for rapid reconstruction of marine food webs. Fish and Fisheries 22:822–833

Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, Creer S, Bista I, Lodge DM, Vere N de, Pfrender ME, Bernatchez L (2017) Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology

DiBattista JD, Reimer JD, Stat M, Masucci GD, Biondi P, Brauwer M De, Wilkinson SP, Chariton AA, Bunce M (2020) Environmental DNA can act as a biodiversity barometer of anthropogenic pressures in coastal ecosystems. Scientific Reports 10:1–15

Djurhuus A, Closek CJ, Kelly RP, Pitz KJ, Michisaki RP, Starks HA, Walz KR, Andruszkiewicz EA, Olesin E, Hubbard K, Montes E, Otis D, Muller-Karger FE, Chavez FP, Boehm AB, Breitbart M (2020) Environmental DNA reveals seasonal shifts and potential interactions in a marine community. Nature Communications 2020 11:1 11:1–9

Donaldson MR, Burnett NJ, Braun DC, Suski CD, Hinch SG, Cooke SJ, Kerr JT (2016) Taxonomic bias and international biodiversity conservation research. https://doi.org/101139/facets-2016-0011 1:105–113

Done TJ (1999) Coral Community Adaptability to Environmental Change at the Scales of

Regions , Reefs and Reef Zones ’. American Zoologist 39:66–79

Duarte CM, Krause-Jensen D (2017) Export from seagrass meadows contributes to marine carbon sequestration. Frontiers in Marine Science 4:13

Duchêne JC (2004) Early recognition of sediment during settlement of Eupolymnia nebulosa (Polychaeta: Terebellidae) larvae. Marine Biology 145:79–85

Duchêne JC (2010) Sediment recognition by post-larval stages of Eupolymnia nebulosa (Polychaeta, Terebellidae). Journal of Experimental Marine Biology and Ecology 386:69–76

Ellis MR, Clark ZSR, Treml EA, Brown MS, Matthews TG, Pocklington JB, Stafford-Bell RE, Bott NJ, Nai YH, Miller AD, Sherman CDH (2022) Detecting marine pests using environmental DNA and biophysical models. Science of The Total Environment 816:151666

Ely T, Barber PH, Man L, Gold Z (2021) Short-lived detection of an introduced vertebrate eDNA signal in a nearshore rocky reef environment. PLOS ONE 16:e0245314

Estradivari, Setyawan E, Yusri S (2009) Pengamatan Jangka Panjang Terumbu Karang Kepulauan Seribu (2003 - 2007). Jakarta

Evans DM, Kitson JJN, Lunt DH, Straw NA, Pocock MJO (2016) Merging DNA metabarcoding and ecological network analysis to understand and build resilient terrestrial ecosystems. Functional Ecology 30:1904–1916

Evans NT, Li Y, Renshaw MA, Olds BP, Deiner K, Turner CR, Jerde CL, Lodge DM, Lamberti GA, Pfrender ME (2017) Fish community assessment with eDNA metabarcoding: Effects of sampling design and bioinformatic filtering. Canadian Journal of Fisheries and Aquatic Sciences 74:1362–1374

Foale S, Adhuri D, Aliño P, Allison EH, Andrew N, Cohen P, Evans L, Fabinyi M, Fidelman P, Gregory C, Stacey N, Tanzer J, Weeratunge N (2013) Food security and the Coral Triangle Initiative. Marine Policy 38:174–183

Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, Hansen AJ (2017) Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nature Communications

Froufe E, Cabezas P, Alexandrino P, Pérez-Losada M (2016) Evolution and conservation of marine biodiversity in the Coral Triangle: insights from stomatopod Crustacea. :141–168

Frühe L, Cordier T, Dully V, Breiner | Hans-Werner, Lentendu | Guillaume, Pawlowski J, Martins C, Thomas |, Wilding A, Stoeck T (2020) S P E C I A L I S S U E : E N V I R O N M E N T A L D N A F O R B Supervised machine learning is superior to indicator value inference in monitoring the environmental impacts of salmon aquaculture using eDNA metabarcodes. Molecular Ecology 00:1–19

Gallego R, Jacobs-Palmer E, Cribari K, Kelly RP (2020) Environmental DNA metabarcoding reveals winners and losers of global change in coastal waters. Proceedings of the Royal Society B 287:20202424

Geller J, Meyer C, Parker M, Hawk H (2013) Redesign of PCR primers for mitochondrial cytochrome c oxidase subunit I for marine invertebrates and application in all-taxa biotic surveys. Molecular Ecology Resources 13:851–861

Gold Z, Curd EE, Goodwin KD, Choi ES, Frable BW, Thompson AR, Walker HJ, Burton RS, Kacev D, Martz LD, Barber PH (2021) Improving metabarcoding taxonomic assignment: A case study of fishes in a large marine ecosystem. Molecular Ecology Resources 21:2546–2564

Gold Z, Sprague J, Kushner DJ, Marin EZ, Barber PH (2021) eDNA metabarcoding as a biomonitoring tool for marine protected areas. PLOS ONE 16:e0238557

Gracia A, Rangel-Buitrago N, Oakley JA, Williams AT (2018) Use of ecosystems in coastal erosion management. Ocean & Coastal Management 156:277–289

Guinotte JM, Fabry VJ (2008) Ocean Acidification and Its Potential Effects on Marine Ecosystems. Annals of the New York Academy of Sciences 1134:320–342

Halpern BS, Walbridge S, Selkoe KA, Kappel C V, Micheli F, D'Agrosa C, Bruno JF, Casey KS, Ebert C, Fox HE, Fujita R, Heinemann D, Lenihan HS, Madin EMP, Perry MT, Selig ER, Spalding M, Steneck R, Watson R (2008) A global map of human impact on marine ecosystems. Science (New York, NY) 319:948–52

Harlan K. Dean (2008) The use of polychaetes (Annelida) as indicator species of marine pollution: a review. Revista de Biología Tropical 56

Hernández-Delgado EA (2015) The emerging threats of climate change on tropical coastal ecosystem services, public health, local economies and livelihood sustainability of small islands: Cumulative impacts and synergies. Marine Pollution Bulletin 101:5–28

Hestetun JT, Bye-Ingebrigtsen E, Nilsson RH, Glover AG, Johansen PO, Dahlgren TG (2020) Significant taxon sampling gaps in DNA databases limit the operational use of marine macrofauna metabarcoding. Marine Biodiversity 50:1–9

Hoegh-Guldberg O (2011) The impact of urbanisation on coral reef ecosystems

Hoegh-Guldberg O, Hoegh-Guldberg H, Veron JE., Green A, Gomez ED, Ambariyanto A, Hansen L (2009) THE CORAL TRIANGLE AND CLIMATE CHANGE: ECOSYSTEMS, PEOPLE AND SOCIETIES AT RISK.

Hofmann GE, Barry JP, Edmunds PJ, Gates RD, Hutchins DA, Klinger T, Sewell MA (2010)

The Effect of Ocean Acidification on Calcifying Organisms in Marine Ecosystems: An Organism-to-Ecosystem Perspective. http://dx.doi.org/101146/annurev.ecolsys110308120227 41:127–147

Holman LE, Bruyn M de, Creer S, Carvalho G, Robidart J, Rius M (2019) Detection of introduced and resident marine species using environmental DNA metabarcoding of sediment and water. Scientific Reports 2019 9:1 9:1–10

Holman LE, Chng Y, Rius M (2022) How does eDNA decay affect metabarcoding experiments? Environmental DNA 4:108–116

Hughes TP, Baird AH, Bellwood DR, Card M, Connolly SR, Folke C, Grosberg R, Hoegh-Guldberg O, Jackson JBC, Kleypas J, Lough JM, Marshall P, Nyström M, Palumbi SR, Pandolfi JM, Rosen B, Roughgarden J (2003) Climate change, human impacts, and the resilience of coral reefs. Science (New York, NY) 301:929–33

Hutchings P (1998) Biodiversity and functioning of polychaetes in benthic sediments. Biodiversity and Conservation

Jensen MR, Sigsgaard EE, Ávila M de P, Agersnap S, Brenner-Larsen W, Sengupta ME, Xing Y, Krag MA, Knudsen SW, Carl H, Møller PR, Thomsen PF (2022) Short-term temporal variation of coastal marine eDNA. Environmental DNA 00:1–16

Jeunen GJ, Knapp M, Spencer HG, Lamare MD, Taylor HR, Stat M, Bunce M, Gemmell NJ (2019) Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse marine habitats connected by water movement. Molecular Ecology Resources 19:426–438

Jones-Walters L, Mulder I (2009) Valuing nature: The economics of biodiversity. Journal for Nature Conservation 17:245–247

Kandlikar G (2020) ranacapa: Utility Functions and “shiny” App for Simple Environmental DNA Visualizations and Analyses.

Klunder L, Bleijswijk JDL van, Kleine Schaars L, Veer HW van der, Luttikhuizen PC, Bijleveld AI (2022) Quantification of marine benthic communities with metabarcoding. Molecular Ecology Resources 22:1043–1054

Ladin ZS, Ferrell B, Dums JT, Moore RM, Levia DF, Shriver WG, D’Amico V, Trammell TLE, Setubal JC, Wommack KE (2021) Assessing the efficacy of eDNA metabarcoding for measuring microbial biodiversity within forest ecosystems. Scientific Reports 2021 11:1 11:1–14

Leray M, Knowlton N (2015) DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proceedings of the National Academy of Sciences 112:201424997

Leray M, Knowlton N (2016a) Visualizing Patterns of Marine Eukaryotic Diversity from Metabarcoding Data Using QIIME. In: Humana Press, New York, NY, p 219–235

Leray M, Knowlton N (2016b) Censusing marine eukaryotic diversity in the twenty-first century. The Royal Society

Leray M, Meyer CP, Mills SC (2015) Metabarcoding dietary analysis of coral dwelling predatory fish demonstrates the minor contribution of coralmutualists to their highly partitioned, generalist diet. PeerJ 2015:e1047

Leray M, Yang JY, Meyer CP, Mills SC, Agudelo N, Ranwez V, Boehm JT, Machida RJ (2013) A new versatile primer set targeting a short fragment of the mitochondrial COI region for metabarcoding metazoan diversity: Application for characterizing coral reef fish gut contents. Frontiers in Zoology 10:1–14

Levin SA, Lubchenco J (2008) Resilience, Robustness, and Marine Ecosystem-based Management. BioScience 58:27–32

LOSEY JE, VAUGHAN M (2009) The Economic Value of Ecological Services Provided by Insects. http://dx.doi.org/101641/0006-3568(2006)56[311:TEVOES]20CO;2

Machida RJ, Leray M, Ho SL, Knowlton N (2017) Metazoan mitochondrial gene sequence reference datasets for taxonomic assignment of environmental samples. Scientific data 4

Madduppa HH, Beginner S, Suparyani E, Siregar AM, Arafat D, TARIGAN SA, ALIMUDDIN A, KHAIRUDI D, RAHMAWATI F, BRAHMANDITO A (2013) Dynamics of fish diversity across an environmental gradient in the Seribu Islands reefs off Jakarta. Biodiversitas Journal of Biological Diversity 14:17–24

Madduppa H, Cahyani NKD, Anggoro AW, Subhan B, Jefri E, Sani LMII, Arafat D, Akbar N, Bengen DG (2021) eDNA metabarcoding illuminates species diversity and composition of three phyla (chordata, mollusca and echinodermata) across Indonesian coral reefs. 30

Marques V, Milhau T, Albouy C, Dejean T, Manel S, Mouillot D, Juhel JB (2021) GAPeDNA: Assessing and mapping global species gaps in genetic databases for eDNA metabarcoding. Diversity and Distributions 27:1880–1892

Marwayana ON, Gold Z, Meyer CP, Barber PH (2022) Environmental DNA in a global biodiversity hotspot: Lessons from coral reef fish diversity across the Indonesian archipelago. Environmental DNA 4:222–238

Mathieu C, Hermans SM, Lear G, Buckley TR, Lee KC, Buckley HL (2020) A systematic review of sources of variability and uncertainty in eDNA data for environmental monitoring. Frontiers in Ecology and Evolution 8:1–14

McMurdie PJ, Holmes S (2013) Phyloseq: An R Package for Reproducible Interactive Analysis

and Graphics of Microbiome Census Data. PLoS ONE

Meij SET van der, Moolenbeek RG, Hoeksema BW (2009) Decline of the Jakarta Bay molluscan fauna linked to human impact. Marine pollution bulletin 59:101–7

Meij SET van der, Suharsono, Hoeksema BW (2010) Long-term changes in coral assemblages under natural and anthropogenic stress in Jakarta Bay (1920-2005). Marine Pollution Bulletin 60:1442–1454

Moberg F, Folke C (1999) Ecological goods and services of coral reef ecosystems. Ecological Economics 29:215–233

Mora C, Tittensor DP, Adl S, Simpson AGB, Worm B (2011) How many species are there on earth and in the ocean? (GM Mace, Ed.). PLoS Biology 9:e1001127

Muha TP (2021) Using eDNA Metabarcoding to Monitor Changes in Fish Community Composition After Barrier Removal. Frontiers in Ecology and Evolution 9:28

Nakagawa H, Yamamoto S, Sato Y, Sado T, Minamoto T, Miya M (2018) Comparing local- and regional-scale estimations of the diversity of stream fish using eDNA metabarcoding and conventional observation methods. Freshwater Biology 63:569–580

Newton K, Côté IM, Pilling GM, Jennings S, Dulvy NK (2007) Current and Future Sustainability of Island Coral Reef Fisheries. Current Biology 17:655–658

Nichols R V., Vollmers C, Newsom LA, Wang Y, Heintzman PD, Leighton M, Green RE, Shapiro B (2018) Minimizing polymerase biases in metabarcoding. Molecular Ecology Resources

Pearman JK, Anlauf H, Irigoien X, Carvalho S (2016) Please mind the gap - Visual census and cryptic biodiversity assessment at central Red Sea coral reefs. Marine Environmental Research 118:20–30

Pearman JK, Aylagas E, Voolstra CR, Anlauf H, Villalobos R, Carvalho S (2019) Disentangling the complex microbial community of coral reefs using standardized Autonomous Reef Monitoring Structures (ARMS). Molecular Ecology:mec.15167

Pearman JK, Leray M, Villalobos R, Machida RJ, Berumen ML, Knowlton N, Carvalho S (2018) Cross-shelf investigation of coral reef cryptic benthic organisms reveals diversity patterns of the hidden majority. Scientific Reports 8:8090

Pitz KJ, Guo J, Johnson SB, Campbell TL, Zhang H, Vrijenhoek RC, Chavez FP, Geller J (2020) Zooplankton biogeographic boundaries in the California Current System as determined from metabarcoding. PLOS ONE 15:e0235159

Plaisance L, Brainard R, Julian Caley M, Knowlton N (2011) Using DNA barcoding and standardized sampling to compare geographic and habitat differentiation of crustaceans: A Hawaiian Islands example. Diversity

Plaisance L, Knowlton N, Paulay G, Meyer C (2009) Reef-associated crustacean fauna: biodiversity estimates using semi-quantitative sampling and DNA barcoding. Coral Reefs 28:977–986

Port JA, O'Donnell JL, Romero-Maraccini OC, Leary PR, Litvin SY, Nickols KJ, Yamahara KM, Kelly RP (2016) Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. Molecular Ecology 25:527–541

R Core Team (2020) R: A language and environment for statistical computing. Retrieved from.

Rachello-Dolmen PG, Cleary DFR (2007) Relating coral species traits to environmental conditions in the Jakarta Bay/Pulau Seribu reef system, Indonesia. Estuarine, Coastal and Shelf Science

Ransome E, Geller JB, Timmers M, Leray M, Mahardini A, Sembiring A, Collins AG, Meyer

CP, Mahardini A, Sembiring A, Collins AG, Meyer CP (2017) The importance of standardization for biodiversity comparisons: A case study using autonomous reef monitoring structures (ARMS) and metabarcoding to measure cryptic diversity on Mo'orea coral reefs, French Polynesia (CA Chen, Ed.). PLoS ONE 12:e0175066

Rees JG, Setiapermana D, Sharp VA, Weeks JM, Williams TM (1999) Evaluation of the impacts of land-based contaminants on the benthic faunas of Jakarta Bay, Indonesia. Oceanologica Acta 22:627–640

Reinholdt Jensen M, Egelyng Sigsgaard E, Agersnap S, Jessen Rasmussen J, Baattrup-Pedersen A, Wiberg-Larsen P, Francis Thomsen P (2021) Seasonal turnover in community composition of stream-associated macroinvertebrates inferred from freshwater environmental DNA metabarcoding. Environmental DNA 3:861–876

Ríos-Castro R, Romero A, Aranguren R, Pallavicini A, Banchi E, Novoa B, Figueras A (2021) High-Throughput Sequencing of Environmental DNA as a Tool for Monitoring Eukaryotic Communities and Potential Pathogens in a Coastal Upwelling Ecosystem. Frontiers in Veterinary Science 8:1300

Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: A versatile open source tool for metagenomics. PeerJ

Roman J (2006) Diluting the founder effect: cryptic invasions expand a marine invader's range. Proceedings of the Royal Society B: Biological Sciences 273:2453–2459

Ruppert KM, Kline RJ, Rahman MS (2019) Past, present, and future perspectives of environmental DNA (eDNA) metabarcoding: A systematic review in methods, monitoring, and applications of global eDNA. Global Ecology and Conservation 17:e00547

Salter I (2018) Seasonal variability in the persistence of dissolved environmental DNA (eDNA)

in a marine system: The role of microbial nutrient limitation. PLOS ONE 13:e0192409

Santiago A De, Pereira TJ, Mincks SL, Bik HM (2022) Dataset complexity impacts both MOTU delimitation and biodiversity estimates in eukaryotic 18S rRNA metabarcoding studies. Environmental DNA 4:363–384

Schenekar T, Schletterer M, Lecaudey LA, Weiss SJ (2020) Reference databases, primer choice, and assay sensitivity for environmental metabarcoding: Lessons learnt from a re-evaluation of an eDNA fish assessment in the Volga headwaters. River Research and Applications 36:1004–1013

Schmelzle MC, Kinziger AP (2016) Using occupancy modelling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. Molecular Ecology Resources 16:895–908

Schmidt P-A, Bálint M, Greshake B, Bandow C, Römbke J, Schmitt I (2013) Illumina metabarcoding of a soil fungal community. Soil Biology and Biochemistry 65:128–132

Seitz RD, Schaffner LC (1995) Population ecology and secondary production of the polychaete Loimia medusa (Terebellidae). Marine Biology 1995 121:4 121:701–711

Sodhi NS, Koh LP, Brook BW, Ng PKL (2004) Southeast Asian biodiversity: an impending disaster. Trends in ecology & evolution 19:654–60

Stat M, Huggett MJ, Bernasconi R, DiBattista JD, Berry TE, Newman SJ, Harvey ES, Bunce M (2017) Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. Scientific Reports 7:12240

Steyaert M, Priestley V, Osborne O, Herraiz A, Arnold R, Savolainen V (2020) Advances in metabarcoding techniques bring us closer to reliable monitoring of the marine benthos. Journal of Applied Ecology 57:2234–2245

Taberlet P, Coissac E, Pompanon F, Brochmann C, Willerslev E (2012) Towards next-generation biodiversity assessment using DNA metabarcoding. Molecular Ecology 21:2045–50

Thomsen PF, Kielgast J, Iversen LL, Møller PR, Rasmussen M, Willerslev E (2012) Detection of a Diverse Marine Fish Fauna Using Environmental DNA from Seawater Samples (S Lin, Ed.). PLoS ONE 7:e41732

Thomsen PF, Willerslev E (2015) Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:4–18

UNCTAD (2021) Ocean-Based Economies : Trade Trends , Market Drivers and Market Access. :100

Uneputty PA, Evans SM (1997) Accumulation of beach litter on islands of the Pulau Seribu Archipelago, Indonesia. Marine Pollution Bulletin 34:652–655

Valentin RE, Fonseca DM, Gable S, Kyle KE, Hamilton GC, Nielsen AL, Lockwood JL (2020) Moving eDNA surveys onto land: Strategies for active eDNA aggregation to detect invasive forest insects. Molecular Ecology Resources 20:746–755

Venkateswara Rao J, Srikanth K, Pallela R, Gnaneshwar Rao T (2009) The use of marine sponge, Haliclona tenuiramosa as bioindicator to monitor heavy metal pollution in the coasts of Gulf of Mannar, India. Environmental monitoring and assessment 156:451–9

VERON JEN, DEVANTIER LM, TURAK E, GREEN AL, KININMONTH S, STAFFORD-SMITH M, PETERSON N (2009) Delineating the Coral Triangle. Galaxea, Journal of Coral Reef Studies 11:91–100

Vieira EA, Flores AAV V., Dias GM (2018) Persistence and space preemption explain species-specific founder effects on the organization of marine sessile communities. Ecology and Evolution 8:3430–3442

Wangensteen OS, Palacín C, Guardiola M, Turon X (2018) DNA metabarcoding of littoral hard-bottom communities: high diversity and database gaps revealed by two molecular markers. PeerJ 6:e4705

Weltz K, Lyle JM, Ovenden J, Morgan JAT, Moreno DA, Semmens JM (2017) Application of environmental DNA to detect an endangered marine skate species in the wild. PLOS ONE 12:e0178124

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

Williams PH, Gaston KJ (1994) Measuring more of biodiversity: Can higher-taxon richness predict wholesale species richness? Biological Conservation 67:211–217

Wilson MC, Chen X-Y, Corlett RT, Didham RK, Ding P, Holt RD, Holyoak M, Hu G, Hughes AC, Jiang L, Laurance WF, Liu J, Pimm SL, Robinson SK, Russo SE, Si X, Wilcove DS, Wu J, Yu M (2015) Habitat fragmentation and biodiversity conservation: key findings and future challenges. Landscape Ecology 31:219–227

Wilson SK, Fisher R, Pratchett MS, Graham NAJ, Dulvy NK, Turner RA, Cakacaka A, Polunin NVC (2010) Habitat degradation and fishing effects on the size structure of coral reef fish communities. Ecological Applications 20:442–451

Yamamoto S, Masuda R, Sato Y, Sado T, Araki H, Kondoh M, Minamoto T, Miya M (2017) Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. Scientific Reports 2017 7:1 7:1–12

Zaiko A, Samuiloviene A, Ardura A, Garcia-Vazquez E (2015) Metabarcoding approach for nonindigenous species surveillance in marine coastal waters. Marine Pollution Bulletin 100:53–59

Zhang S, Zhao J, Yao M (2020) A comprehensive and comparative evaluation of primers for

metabarcoding eDNA from fish. Methods in Ecology and Evolution 11:1609–1625

CHAPTER 2

Shifts in Eukaryotic Communities and Identification of Indicator Taxa Across a Marine Pollution Gradient

Abstract

Anthropogenic stressors are impacting marine ecosystems, but fully understanding these impacts requires methods of assessing changes in marine community diversity that is standardized and can detect changes across a broad range of taxonomic diversity. This study employs autonomous reef monitoring structures (ARMS) and environmental DNA (eDNA) with standardized metabarcoding approaches to examine changes in marine biodiversity across Pulau Seribu, an archipelago heavily impacted by pollution from Jakarta, Indonesia. Despite an eight fold difference in pollution stress as inferred from Chlorophyll-a concentrations, results from ARMS showed no significant impact on diversity, based on amplified sequence variants (ASVs) of cytochrome *c* oxidase subunit I (COI) and 18S rDNA (18S). However, eDNA data based on COI showed a significant decrease in ASV diversity on reefs with higher levels of pollution. Both methods captured indicator taxa that experienced significant increases or decreases associated with pollution stress. Chlorodendrophyceae (marine green algae) increased with pollution stress while Phaeophyceae (Brown algae) and Entoprocta decreased. Results indicate that marine communities captured by eDNA are more sensitive to pollution stress than benthic marine communities that colonize ARMS and could provide an important tool monitoring changes in marine ecosystems in our rapidly changing world.

Introduction

Marine ecosystems are vital to planetary health; they are home to almost one third of global biodiversity (Grosberg et al. 2012, Steele et al. 2019), produce nearly 70% of the global oxygen supply (Beaumont et al. 2007, Boeuf 2011, Gamfeldt et al. 2015) and contribute to global climate stability by sequestering nearly 50% of fixed global carbon (Beaumont et al. 2007, Howard et al. 2017, Armstrong McKay et al. 2021). Marine resources are similarly critical for human health and economic development, supporting food security (Duarte 2000, Worm et al. 2006, Foale et al. 2013), generating billions of dollars in economic activity globally (Moberg & Folke 1999, Jones-Walters & Mulder 2009), and are particularly important in countries with large coastal territories (Barber 2009, Foale et al. 2013, Jenkins & Van Houtan 2016).

Key to maintaining these important ecosystem goods and services is maintaining complexity and biodiversity in marine environments. More diverse marine ecosystems retain important functional diversity that promotes ecosystem resilience and recovery following disturbances (Costanza & Mageau 1999, Hughes et al. 2003, Anthony et al. 2015). The ecosystem also mitigate disturbance, with intact coral reef ecosystems serving as an important barrier to strong wave action resulting from typhoons and tsunamis (Narayan et al. 2016, Guannel et al. 2016). Marine biodiversity is also essential for economic and food security, particularly in developing maritime nations (Wilkinson 1996, Dixon 1998, Beaumont et al. 2008). For example, Indonesia derives 12.88 % of its national GDP from marine tourism and fisheries-related industries (setkab 2018) and 70% of its population are dependent on marine fisheries for their daily protein source (World Bank Group 2022). Given this importance, effective management of marine biodiversity is essential; foundational to advancing management efforts is effective biodiversity monitoring.

Marine resource managers employ a variety of monitoring methods to enumerate and quantify marine biodiversity in support of conservation of marine resources and habitats. Traditional methods include visual census, sediment grabs, plankton pumps, and suction samplers (reviewed in Costello et al. 2017). Although these methods provide important data to understand local ecosystem health, they typically require extensive taxonomic expertise, and are unreliable when performed by non-experts (Bernard et al. 2013, Mora 2015). Even when performed by experts these methods are prone to taxonomic biases and can be difficult to standardize (Bernard et al. 2013, Lindfield et al. 2014); they are also costly and time consuming. As such, many monitoring programs focus only on small areas, and/or with limited frequency, limiting their overall effectiveness (Sprague 2020) and our ability to monitor marine biodiversity on scales required to ensure sustained ecosystem function.

Another major shortcoming of current marine ecosystem monitoring is that, for efficiency, these methods frequently focus only on a narrow range of economically valuable taxa and important keystone species (Williams & Gaston 1994, Brooks et al. 2006, Bickford et al. 2007, Losey & Vaughan 2009). This focus on selective taxa is problematic in two key ways. First, resilience of marine ecosystems depends on ecosystem diversity (Hughes et al. 2003) and monitoring focal taxa does not permit assessment of overall ecosystem diversity and health. Second, the most commonly monitored taxa, particularly in coral reef ecosystems, are larger fishes and corals (Bouchet et al. 2002, Appeltans et al. 2012); these long-lived taxa may not be the most sensitive taxa to detect changes in ecosystem health. Combined, these monitoring approaches limit managers' understanding of overall ecosystem health and resilience and may limit their ability to respond proactively to threats before they lead to large-scale changes that could destabilize local marine communities.

An alternative approach to monitoring marine ecosystems is DNA metabarcoding. Metabarcoding generates 100s of thousands to millions of DNA barcode sequences (e.g. mitochondrial cytochrome c oxidase subunit I, 12S and 16S ribosomal DNA) from an unknown environmental sample and then identifies these barcodes to species by comparing them to a dataset of DNA barcode sequences with known taxonomy (e.g. BOLD (https://boldsystems.org/index.php) and NCBI (https://blast.ncbi.nlm.nih.gov/Blast.cgi)). Metabarcoding has been used across a broad range of questions ranging from understanding microbial composition in both terrestrial and marine ecosystem (Pearman et al. 2019, Joos et al. 2020), elucidating fish diets (Leray et al. 2015), testing for the presence of invasive species (Andersen et al. 2019) to understanding how biodiversity changes with changes in land use (Beng et al. 2016) or fishing pressure (Madduppa et al. 2021), and examining the diversity of known vs. cryptic species (Leray & Knowlton 2016). A key advantage of metabarcoding is that, with a relatively small amount of laboratory training, non-experts can conduct taxonomic identification across the tree of life at a scale impossible through traditional taxonomy.

Two common applications of metabarcoding in marine ecosystems are Autonomous Reef Monitoring Structures (ARMS) and environmental DNA (eDNA). ARMS are a highly standardized approach where a stack of 25x25 cm PCV plates is deployed on the benthos, and after a specified period of time (e.g., 3 years) the entire community is metabarcoded (Plaisance et al. 2011, Leray & Knowlton 2015, Ransome et al. 2017). In contrast, eDNA is based on the collection of freely associated DNA and/or cells that organisms shed into the environment, representing a snapshot of a local community at a given period of time (Thomsen & Willerslev 2015, Deiner et al. 2017, Stat et al. 2017, Marwayana et al. 2022). Although both of these methods capture unique communities (Chapter 1), they both capture a broad range of marine

biodiversity, making them ideal techniques for monitoring marine biodiversity, particularly in response to anthropogenic stressors.

Although a wide variety of anthropogenic stressors such as overfishing (Jackson et al. 2001, Coleman & Williams 2002, Coll et al. 2008), destructive fishing practices (Mcmanus et al. 1997, Pet-Soede et al. 1999, Cesar et al. 2003), climate change (Knowlton 2001, Jenkins 2003, Bellwood et al. 2004) and ocean acidification (Orr et al. 2005, Guinotte & Fabry 2008, Hofmann et al. 2010) impact marine communities, local point sources of pollution remain a significant concern. Research examining the impacts of pollution on marine biodiversity have largely focused on specific species of high economic value or keystone species that play important roles in supporting ecosystem function (Wilkinson 1996, Cleary et al. 2006, Cebrian et al. 2007, Cleary 2017). Although such studies provide important insights, economically important species are often long-lived, and may not be the most sensitive species to pollution. Similarly, some keystone species can adapt to changing conditions and are relatively unaffected by pollution stress (Pocklington & Wells 1992, Cleary et al. 2006, Cebrian et al. 2007, Kiruba-Sankar et al. 2016, Béguinot 2018). As such, monitoring efforts that are limited to economically valuable or keystone species may not provide an accurate gauge of the impact of pollution on local biodiversity.

Kepulauan Seribu is an archipelago north of Jakarta Indonesia. Comprised of 105 islands, 11 inhabited, this archipelago spans over 80 km, including areas heavily impacted by industrial activities and broad-scale riverine discharge into Jakarta Bay, highly populated islands that create localized pollution stress, and distant uninhabited islands with relatively low pollution stress. Previous studies show that pollution stress shapes the distribution of marine life across Kepulauan Seribu, with coral reefs closer to Jakarta Bay characterized by species that are more

tolerant to disturbance, while more sensitive species only survive in more distant islands with cleaner waters (van der Meij et al. 2009, 2010, Madduppa et al. 2013). Moreover, individual taxa respond differently, with abundance and diversity of fish being highest on the least polluted reefs, while diversity and abundance of urchins increases with pollution stress (Cleary et al. 2008, van der Meij et al. 2009, Madduppa et al. 2013). However, these studies, like most monitoring programs, focus only on a small amount of taxonomic diversity; it remains unclear whether these patterns are representative of eukaryotic diversity more broadly.

To improve the ability of resource managers to monitor marine ecosystems and promote the sustainability of these vital ecosystems, this study employs ARMS and eDNA metabarcoding across the anthropogenic stress gradients of Kepulauan Seribu. Given that ARMS and eDNA capture different, largely non-overlapping communities, we specifically test whether communities captured by these two methods provide comparable results in response to pollution stress. We then identify and compare taxa that are most sensitive to pollution stress and identify sites within the archipelago that are most unique in terms of biodiversity.

Material, Methods and Environmental Data

Deployment Design

In 2013, we deployed sets of three ARMS in eight islands across the archipelago with distance between sites of 7-10 km, for a total of 24 ARMS. The deployment sites were standardized to a depth of 10m and encompass sites with the most polluted level in Pulau Bidadari (10 km from Jakarta Bay) to gradually less polluted site in Pulau Sepa (60 km from Jakarta Bay) (Table 2-1.) (DeVantier L, Suharsono, Budiyanto A, Tuti Y, Imanto P 1998, van

der Meij et al. 2009, 2010, Polónia et al. 2014). ARMS remained *in situ* for three years, and were recovered in 2016

Table 2-1. ARMS deployment site and distance from the coast.

Site	Distance from Jakarta Bay
Pulau Bidadari	10.08
Pulau Lancang Besar	20.28
Pulau Pari	31.4
Pulau Tidung	33.81
Pulau Karang Beras	37.57
Pulau Pramuka	39.94
Pulau Kotok	46.45
Pulau Sepa	60.64

Environmental Data

Because direct measurement of water quality was not possible, we instead inferred pollution stress using Chlorophyll *a* concentration as a proxy. We obtained annual average Chlorophyll *a* concentrations and sea surface temperatures (SST) from MODIS A satellite imaging data archived on the NASA Ocean color website (https://ocean- color.gsfc.nasa.gov/). We downloaded data at a 4 km resolution from 2014-2016, corresponding to the dates of the ARMS deployment. We visualized this environmental data using Ocean Data View (version 5.0.0) based on the value associated with the nearest grid point to our sampling locations. Due to the spatial resolution of the data, some sites were assigned to the same grid point and thus had the same environmental data.

ARMS Collection and Sampling

To prevent motile organisms from escaping during underwater recovery, we covered ARMS units with 40um Nitex mesh-lined containers, brought them to the surface, and then transported them to an Indonesian Institute of Science field lab in Pulau Pari, using large plastic container filled with filtered, aerated sea water.

We processed ARMS following standardized disassembly and sampling protocol (Leray & Knowlton 2015, Ransome et al. 2017), using series of geologic sieves to separate all motile organisms into size fractions for downstream metabarcoding analysis: 1) a 2 mm to 500 μm sample or "500 μm fraction"; and 2) a 500 μm to 106 μm sample or "100 μm fraction". Each of these fraction samples were stored in 50mL falcon tubes, preserved with 95% ethanol, and stored at -20 °C until further processing. All organisms encrusting the ARMS plates were scraped off and homogenized in a blender and is subsequently referred to as the "sessile fraction". The processed homogenate was washed initially with 95% ethanol and preserved with DMSO in a 50mL stored in -20 °C.

Edna Collection and Sampling

We collected eDNA from each ARMS deployment site. Following standard sampling protocols (Miya et al., 2015), we collected three one-liter water samples on SCUBA at depths of 10-12m at each site to maximize species diversity and to account for fine-scale heterogeneity in local eDNA signatures. To isolate eDNA, we filtered individual seawater samples through a 0.22-micron Sterivex™ filter (Millipore®, SIGMA MILLIPORE) using the methods of Miya et al. (2015) with one key modification; we collected individual water samples in sterile 1 liter Kangaroo™ Gravity Feeding Bags (similar to intravenous drip bags) that allow for gravity

filtration through the Sterivex™ filters. In addition to the eDNA water samples, we also filtered one blank at each locality as a negative control. Filters were stored in a -20 freezer until eDNA was extracted.

ARMS Extraction and Library Preparation

To prepare the ARMS fractions for DNA extraction, we decanted both 500 μm and 100 μm fractions to separate calcium carbonate and terrigenous sediment. We performed decantation by resuspending samples with sterile water in 1 L Erlenmeyer flasks, allowing the sediment to settle, and then decanting the suspended sample contents through a geological sieve to recover the organic contents. We used a 106 μm sieve to decant the 500 μm fraction and a 45 μm sieve to decant the 100 μm fraction. We then used the decanted samples for further extraction and sequencing steps. For the sessile fraction, no decantation was necessary.

DNA extractions were performed at Laboratory of Marine Molecular Genetics, Research Center for Oceanography, Indonesian Institute of Sciences, Jakarta, Indonesia. We performed DNA extractions by digesting 10 g of each sample using 400 μg/ml Proteinase K and an overnight incubation at 56 ^{o}C at 200 rpm. We processed digested samples using Powermax Soil DNA Isolation Kit (MoBio, Carlsbad, USA), following manufacturer's protocol. To remove potential PCR inhibiters from DNA extractions, we cleaned all DNA extracts using a PowerClean DNA Clean-Up Kit (MoBio, Carlsbad, USA) following the manufacturer's protocols, and then quantified the final DNA extracts using Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA) and a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA).

We amplified mitochondrial Cytochrome Oxidase C, subunit 1 (COI) using seven tailed primer pairs of m1COIintF and jgHCO2190. We conducted PCR reactions in a total volume of 20 µl consisting of 1 µl of 10 µM forward and reverse primer, 1.4 µl dNTP, 2 µl Advantage 2 DNA Buffer (Takara Bio USA, Mountain View, CA, USA), 0.4 µl Advantage 2 Polymerase (Takara Bio USA, Mountain View, CA, USA), 13.2 µl of distilled water, and 10 ng DNA of extracted DNA. We performed PCR using a two-step touchdown profile; initial denaturation was at 95°C for 10 min, followed by sixteen cycles: 95°C for 10 s, 62°C (-1°C per cycle) for 30 s, and 72°C for 60 s, and then twenty cycles: 95°C for 10 s, 46°C for 30s, and 72°C for 7 min, followed by a final extension at 72°C for 7 min. We ran PCR reactions in triplicate and verified success on 1.2% agarose gels. All successful reactions were then pooled into a single product.

We quantified and pooled PCR products using a Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA) with a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA). We pooled tailed primer pairs in equimolar concentrations, then used bead cleaning at a concentration of 0.8x vol/vol with Agencourt AMPure XP beads (Beckman Coulter, Brea, CA, USA). Library preparation was performed with a TruSeq DNA PCR-Free IT Library Prep Kit (Illumina, San Diego, CA, USA) following the manufacturer's protocols. Prior to sequencing, we quantified the indexed samples using a Qubit Flourometer, pooling 50 ng of each library, then normalizing the sample to 4nM. Sequencing was conducted on an Illumina MiSeq with a MiSeq Reagent Kit v3, 600-cycle (Illumina, San Diego, CA, USA), using a 1% PhiX spike at the Smithsonian Institution National Museum of Natural History in Washington DC, USA

For 18S, we amplified and sequenced the V4 region using the V4_18SNext.For and V4_18SNext.Rev primers following the 2015 Ocean Sampling Day Protocol (Kopf et al. 2015).

Briefly, we ran each PCR reaction in a volume of 20 µl: 1.25 µl of 0.5 µM forward and reverse primer, 0.5 µl dNTP, 5 µl 5x High-Fidelity DNA Buffer (Thermo Fisher Scientific, Waltham, MA, USA), 0.5 µl of 1 U Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA), 16.0 µl distilled water, and 20 ng of extracted DNA. Thermocycling employed a two-step PCR protocol with an initial denaturation at 98°C for 30 s, then ten cycles of 98°C for 10 s, 44°C for 30 s, and 72°C for 15 s followed by fifteen cycles of 98°C for 10 s, 62°C for 30 s, and 72°C for 15 s, ending with a final extension at 72°C for 7 min. We verified amplification success on 1.2% agarose gels, and then bead cleaned PCR products with Agencourt AMPure XP Beads (Beckman Coulter, Brea, CA, USA) at a concentration of 1.2x vol/vol. We then quantified PCR concentrations using a Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA) to measure all PCR products with a Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA) to calculate the appropriate DNA concentration for the second round of PCR.

For library preparation, we used the dual index approach with the Nextera DNA Library Prep Kit (Illumina, San Diego, CA, USA) and the Nextera Index Kit (Illumina, San Diego, CA, USA). We ran each indexing PCR reaction in a volume of 50 µl: 5 µl of Index 1 (i7), 5 µl of Index 2 (i5), 1 µl dNTP, 10 µl 5x High-Fidelity DNA Buffer (Thermo Fisher Scientific, Waltham, MA, USA), 0.5 µl of 1 U Phusion High-Fidelity DNA Polymerase (Thermo Fisher Scientific, Waltham, MA, USA), a volume of DNA template required to yield 40 ng, and the remaining volume of distilled water. The PCR amplification included an initial denaturation at 98°C for 30 s, followed by five cycles as follows: 98°C for 10 s, 65°C for 30 s, and 72°C for 3 min. We used Agencourt AMPure XP Beads (Beckman Coulter, Brea, CA, USA) at a concentration of 0.6x vol/vol to clean the PCR products. Prior to sequencing, we quantified

indexed samples using a Qubit Fluorometer, then we pooled samples in equimolar ratios prior to normalizing libraries to 2nM and sequencing them on an Illumina MiSeq with a MiSeq Reagent Kit v2, 500-cycle (Illumina, San Diego, CA, USA).

eDNA Extraction and Library Preparation

We extracted eDNA samples and blanks using the DNeasy Blood & Tissue Kit (QIAGEN) following the modified extraction protocol of (Spens et al. 2017), adding 720 µl of ATL buffer and 80 µl of proteinase K directly into the filter cartridge. We amplified extracted eDNA using the Multiplex PCR Kit (QIAGEN, Germany), targeting only CO1, using the same primer set for ARMS. To minimize the impacts PCR bias, we performed PCRs in triplicate (Andruszkiewicz et al., 2017; Miya et al., 2015; Taberlet et al., 2012). Each PCR reaction consisted of 12.5 µL Qiagen 2x Master Mix, 2.5 µL (2 mM) of the primer, 6.5µL nuclease free water, and 1 µL the DNA extract. Thermocycling parameters utilized a touchdown protocol, beginning with a 15-minute pre-denaturation step at a 95 °C, followed by a touchdown thermocycling profile consisting of 30 seconds denaturing at 94 °C, 30 seconds annealing at 69.5 °C, and 30 seconds extension at 72 °C, with the annealing temperature dropping by 1.5°C per cycle until50 °C. Following this initial touchdown phase, the main cycle consisted of 25 cycles of 94 °C for 30 seconds for denaturation, 50 °C for 30 seconds for annealing and 72 °C for 45 seconds for extension, concluding with a 10-minute final extension at 72 °C. To visualize successful PCR reactions, we electrophoresed 5µL of all PCR products for 30 minutes at 150 volts on 2% agarose gels prepared with 6x SYBR™ Green (Invitrogen™, Thermo Fisher Scientific).

To prepare sequencing libraries, we pooled triplicate PCR products, representing each one-liter eDNA sample, into a single tube, and purified these pooled PCR products using Sera-Mag™ and Sera-Mag Speed Beads Magnetic Particles (SIGMA-ALDRICH®) following manufacturer's protocols. Next, we quantified the DNA concentration (ng/μL) of each pooled PCR sample using the Qubit™ 4 NGS Starter Kit (Thermo Fisher) following manufacturer protocols and adjusted concentrations of pooled PCRs to be equal. We then used the Nextera DNA Library Preparation Kit (illumine®) to index each PCR amplified eDNA sample using a unique combination of Illumina Nextera i5 and i7 primers in a second PCR reaction, following the manufacturer protocol. The indexing PCR reaction consisted of 12.5μL Kapa High Fidelity Master Mix, 0.625 μL of 1μM i5 Illumina Nextera indices, 0.625 μL of 1μM i7 Illumina Nextera indices, and 11.25 μL of PCR product for a total of 10ng of DNA. To bioinformatically distinguish among samples, we also added index barcodes to each sample utilizing an indexing PCR protocol that began with an initial denaturation of 95℃ for 5 minutes, followed by 8 cycles of: 98℃ denaturation for 30 seconds, 56℃ annealing for 30 seconds, and 72℃ extension for 3 minutes, ending with a 72℃ extension for 5 minutes. To ensure the indexing PCR was successful, we electrophoresed indexed PCR products at 120V for 45 minutes on a 2% agarose gel prepared with 6x SYBR™ Green. Indexed PCR products were then cleaned and quantified, as above, to creating a final sequencing library that contained equal DNA concentrations (10 ng/μl) of each sample. The final libraries were sequenced at the UC Berkeley sequencing core on an Illumina MiSeq platform utilizing 300 base pair paired end sequencing.

Sequence Processing and Taxonomic Assignments

ARMS

Because some of the COI sequence data was mixed-oriented (there are forward-oriented and reverse-oriented reads within one single raw FASTQ file) paired-end demultiplexed sequences, we included several scripts and steps to separate and group the sequences (i.e., forward and reverse oriented) for each sample. To split reads, we first demultiplexed all raw paired-end FASTQ reads using Cutadapt 1.13 (Martin, 2011) by employing error rate of 0.15 (-e = 0.15), --discard-untrimmed and --no-indels options. This step produced forward oriented and reverse oriented sequences for each sample. Following this step we imported each sample from each orientation as a qza file for further processing using QIIME2, ver. 2018.11 (Bolyen et al. 2019). Next, we used Divisive Amplicon Denoising Algorithm 2 (DADA2) (Callahan et al. 2016) implemented in QIIME2 to filter reads, remove poor quality base pairs from the 3' end of reads, remove sequencing error through denoising, combine forward and reverse reads, search for and removes chimeras, and cluster all reads into ASV's (amplicon sequence variants). We set --p-trunc-len for forward reads at position 240 and 210 for reverse reads, while for chimera removal methods we used ‘consensus’ method as implemented in QIIME2 chimera removal method. These steps generated representative sequences and an ASV table (a table that summarizes the distribution of ASV through all samples) as qza files for each orientation. Next, we exported the qza representative sequences file into a FASTA file. Following this step, we reverse-complemented the reverse oriented FASTA file using fastx_reverse_complement command from FASTX Toolkit 0.0.14 (http://hannonlab.cshl/fastx_toolkit/) and rehashed it using FASTA-names-to-md5.py command from (https://gist.github.com/mkweskin). We also renamed reverse oriented ASV table, using batch-find-replace.py from

(https://gist.github.com/mkweskin). Following these steps, we finally imported the renamed and rehashed ASV and FASTA file to QIIME2 qza format and merged these with existing forward oriented file using QIIME feature-table merge-seqs and QIIME feature-table merge commands.

For 18S sequences, we used Cutadapt 1.13 (Martin 2011) to remove primers by cutting the first 21 bases from each forward and reverse FASTQ file. We then merged both reads using merged paired-end reads using PEAR (Zhang et al. 2014). Next, we imported the merged reads as a single end FASTQ file to QIIME2 and performed denoising using DADA2 (this step filtered reads, removed poor quality base pairs from the 3' end of reads, removed sequencing error through denoising, searched for and removed chimeras, and clustered all reads into ASVs). The result was a table of representative sequences and ASVs.

Prior taxonomic assignment, we applied LULU (Frøslev et al. 2017) to remove erroneous ASVs (minimum_match = 84 for COI and 90 for 18S, minimum_relative_cooccurence = 0.90 for both markers) to the representative sequence. We then clustered the LULU outputs using vsearch (Rognes et al. 2016) (at 99% for 18S and 97% for COI) and assigned the sequences using BLASTN (Camacho et al. 2009) at 85% identity for both markers. For 18S taxonomic assignment we used SILVA SSU non-redundant database (132 release) (https://github.com/qiime2/q2-feature-classifier) for and for COI we used custom database that combined NCBI database (downloaded on June 2020) and a local BIOCODE database (Meyer, C. P. 2016) generated using ANACAPA toolkit (Curd et al. 2019) . This resulted in an ASV table, a table of representative sequences and a taxonomic table (assigned taxonomic name for each representative sequence).

eDNA

We analyzed all eDNA sequences using the Anacapa pipeline (Curd et al., 2018). Anacapa begins by creating a *de novo* sequence reference library by combining records from publicly accessible databases, such as European Molecular Biology Laboratory (EMBL) and National Center for Biotechnology Information (NCBI) using the CRUX package (Creating Reference libraries Using existing tools; https://github.com/limey-bean/Anacapa/tree/New-Master/Anacapa_db). Anacapa then de-multiplexes the amplicon reads based on the primer sequences, and then trims the primers from the reads. Next, the DADA2 algorithm (Callahan et al.2016) performs denoising and error correction on the raw sequence data, merges paired-end reads, and assigns high quality reads to Amplicon Sequence Variants (ASVs) through ASV Parsing. Finally, ASVs are assigned to taxa by Bowtie 2 and the Bayesian Least Common Ancestor algorithm using a 60% likelihood threshold (BLCA; Gao et al.2017.). Following this step, all generated FASTA file and ASV table will then also be feed into LULU using details explained in ARMS chapter.

Merging eDNA and ARMS Data of CO1 Marker

Next, FASTA files from ARMS and eDNA CO1 data were merged using the *cat* command in UNIX, aligned using MAfft software and then clustered using *mothur* at 97% similarity. We then conducted taxonomic assignment using BLASTN (Camacho et al. 2009) at 85% identity, using a custom database that combined NCBI database (downloaded on June 2020) and a local BIOCODE database (Meyer, C. P. 2016) generated using the ANACAPA toolkit (Curd et al. 2019). The resulting ASV table with representative sequences and a taxonomic

assignment was then arranged as a single phyloseq object (*phyloseq* package (version 1.26.1) (McMurdie & Holmes, 2013) to allow further analysis and visualization.

Data Analysis

We used the *PHYLOSEQ* package (version 1.26.1) (McMurdie & Holmes, 2013) to combine the representative sequences and ASV tables into a single *phyloseq* object/file, calculate alpha diversity (Observed ASVs, and Shannon), group samples based on fraction size and location, and transform sequence data from total reads counts to relative read abundance (RRA). We then visualized these data in *ggplot2* (version 3.3.2) (Wickham, 2016). To produce rarefaction curves and visualize sampling effort and ASV richness across sites and fractions, we use *ggrare* package (version 0.1.0.) (Kandlikar, 2020).

To test for significant differences among diversity metrics, we used one-way ANOVA (for non-normally distributed value) and Kruskal-Wallis test (for normally distributed value) as implemented in STATS package (version 4.0.2) (R Core Team, 2020). This step was done after checking for normality in all alpha diversity values across sites using *Saphiro.test* and homoscedasticity using *bartlett.test* function from the same package.

To test the effect of environmental variables, including distance from Jakarta Bay and deployment location against number of total OTUs from each site, we ran linear mixed model analysis using the LME4 package (version 1.1-23) (Bates et al. 2015). We used chlorophyll a, temperature, and distance as fixed effects, site name as a random effect and ASV abundances as the dependent variable. Following models were tested for the analysis

a. Total_OTUs ~ Distance + Temperature + (1|Site_Name)
b. Total_OTUs ~ Distance + Chlorophyll + (1|Site_Name)
c. Total_OTUs ~ Distance + (1|Site_Name)
d. Total_OTUs ~ Chlorophyll + (1|Site_Name)

e. Total_OTUs ~ 1 + (1|Site_Name)

Significance was calculated using the AFEX package (version 0.27-2) (Singmann et al. 2020) and to generate p-values for the mixed models. To compare the performance of different parameter combinations, we used AIC (Akike's information criterion) implemented in sjPLOT package (version 2.8.4) (Lüdecke 2020), and checked for collinearity between parameters using the CAR package (version 3.0.8) (Fox & Weisberg 2019). Lastly, we also performed linear regressions to test for correlations between taxa richness and chlorophyll a concentration. We then plotted data that were normally distributed and free of homoscedasticity.

To further compare community composition across sites, we conducted multivariate analyses (PERMANOVA) based on Jaccard distances in the *vegan* package (Ogle, 2017) in R (R development core team) and tested the statistical significance using 9999 permutations and a significance level of $\alpha = 0.05$. We then calculated the compositional dissimilarity using '*ADONIS*' command and the homogeneity of group dispersion using '*betadisper*' command in *vegan* package (Oksanen, 2017). We conducted Principles Coordinates Analyses (PCoA) using the *Ampvis2* package (Andersen et al., 2018) with the ordination function of *phyloseq* using Jaccard dissimilarity matrices, generating an ordination plot using *ggplot2* (Oksanen, 2017). Next, we tested whether individual sites were significantly different from each other, using *pairwise.ADONIS* from R pairwiseADONIS (version 4.0) package (Martinez Arbizu P, 2020). We then examined individual taxa that contributed sites differences using similarity percentage (SIMPER) based on the Jaccard dissimilarities, testing for statistical significance using the non-parametric Kruskal–Wallis rank-sum test based on simper.pretty and kruskal.pretty functions in R scripts simper_pretty.R and R_krusk.R. as explained in (Steinberger 2018). Only taxa that contributing more than 1% of the variance ($p<0.05$) are presented.

Results

Variability in Chlorophyll Content and Sea Surface Temperature

Annual averages of chlorophyll a concentration from 2014 – 2016 varied from 0.3 – 15 mg m^{-3} across the 8 ARMS deployment locations; the highest concentrations occurred near Pulau Bidadari (15.69 mg m^{-3}), the site closest to Jakarta Bay. Northern regions of Pulau Seribu generally had the lowest chlorophyll concentrations, although waters near Pulau Kotok also had relatively low chlorophyll concentration (0.6 – 1.5 mg m^{-3}) (Fig. 2-1). Variation in SST was smaller, ranging between 29°C and 31°C. Highest temperatures were observed closest to Jakarta Bay, near Bidadari and lowest temperatures were observed near the northern islands of the archipelago.

General Reads and Sequence Results

CO1-ARMS

We recovered only 18 of the 24 ARMS, resulting in 54 samples for subsequent lab work and downstream data analysis (Supplemental Table S2-1); the remainder were either lost or stolen. Following post quality filtering and the exclusion of chimeras, we generated 2,049,305 reads, yielding 3385 ASVs with a mean length of 312.3. Rarefaction curves indicate that ASV discovery largely saturated across all samples. The resulting ASV table was then rarefied to even sequencing depth to 11,000 reads per sample; 3 samples were discarded due to low read numbers (3351 reads from Pari 100 μm fraction, 6444 reads from Sepa 500 μm fraction and 5888 reads from Kotok 100 μm fraction). Following rarefaction (Fig. 2-2), a total of 561,000 reads representing 2907 ASVs remained, all of which were used for downstream analysis.

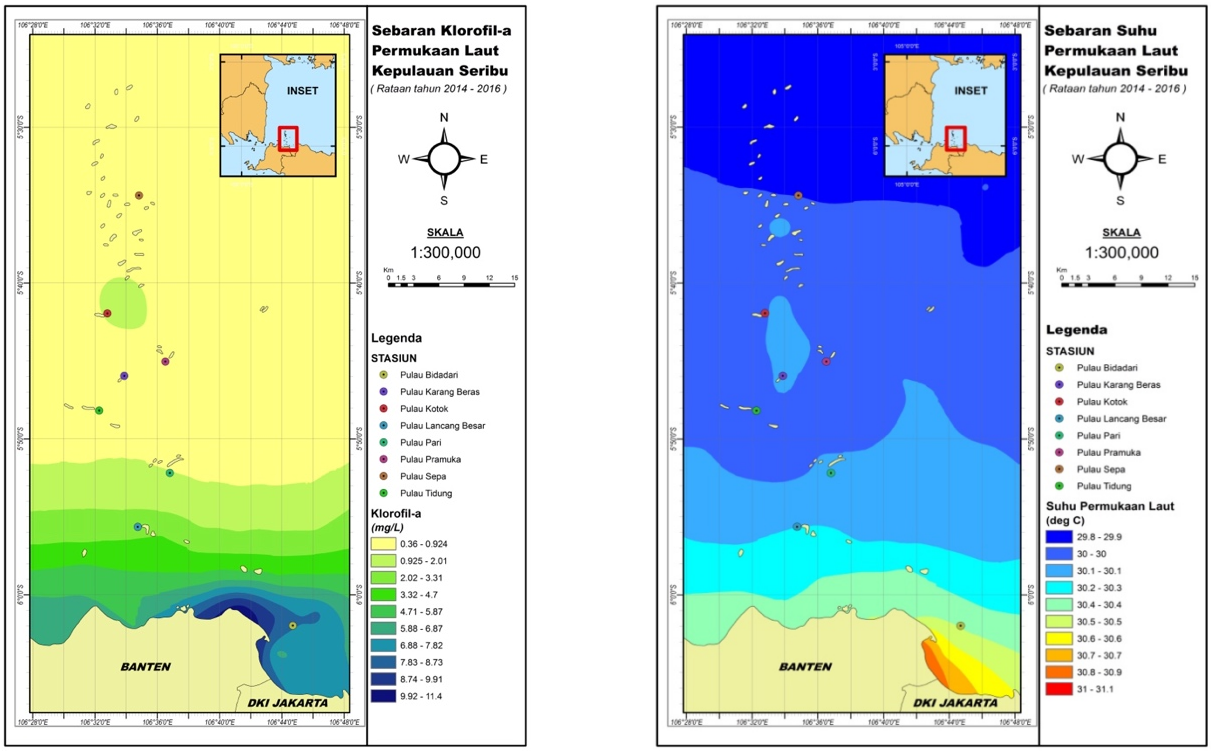


Figure 2-1. Annual average from three years Chlorophyll *a* concentration (a.) and Sea Surface Temperatures data (b.) (2014-2016) across Kepulauan Seribu from NASA's Oceancolor website (https://oceancolor.gsfc.nasa.gov/) derived from the MODIS A satellites at a 4 km resolution. Dots in the image indicate sites of deployment

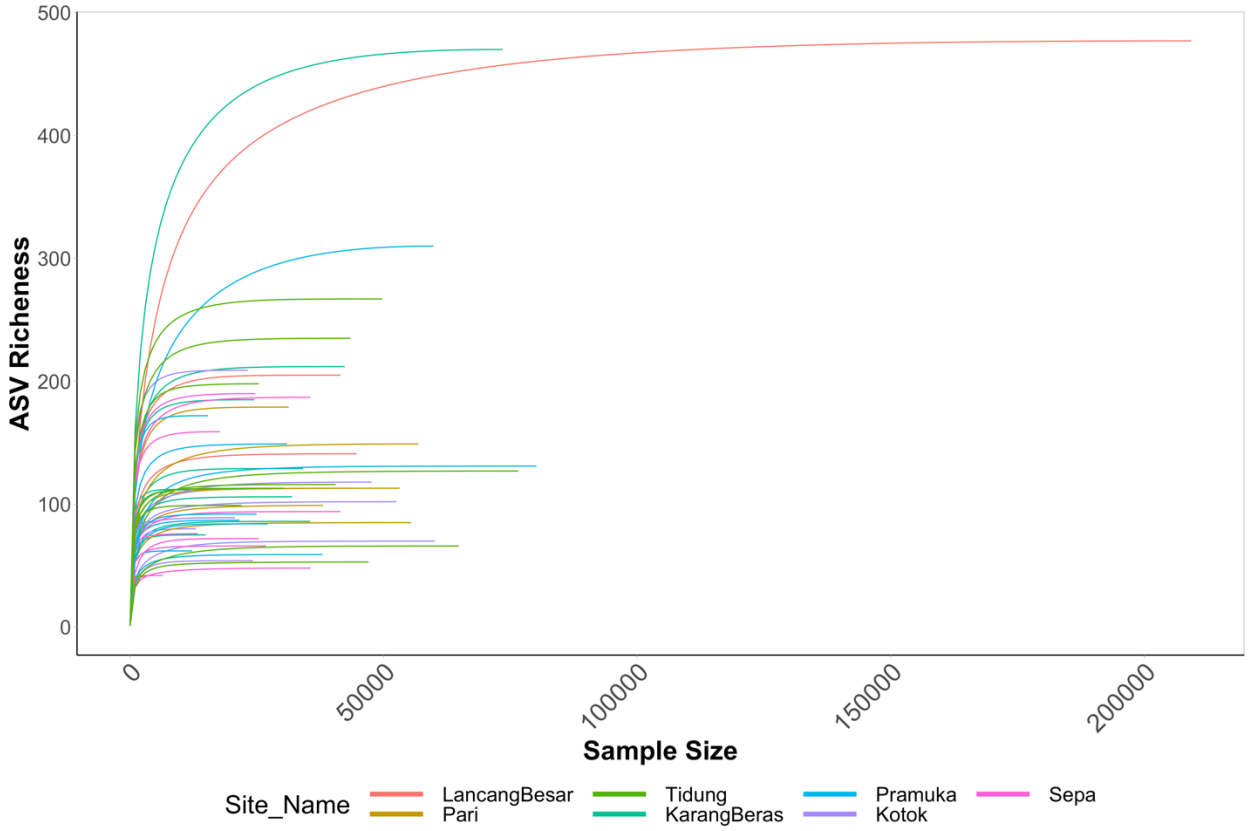


Figure 2-2. Rarefaction curves showing ASV richness vs. sequencing effort for COI-ARMS data.

18S-ARMS

Sequencing of 18S produced 1,048,552 reads and 5149 ASVs with a mean length of 303.2 bases per read and 19,784 reads per sample. Rarefaction curves indicate that sequencing effort largely saturated except for some samples from Pari, Pramuka and Sepa (Fig. 2-3). The resulting ASV table was then rarefied to 19,784 reads per sample. One sample was discarded due to potential mislabeling and therefore only total 53 samples used for downstream analysis (Supplemental Table S2-2).

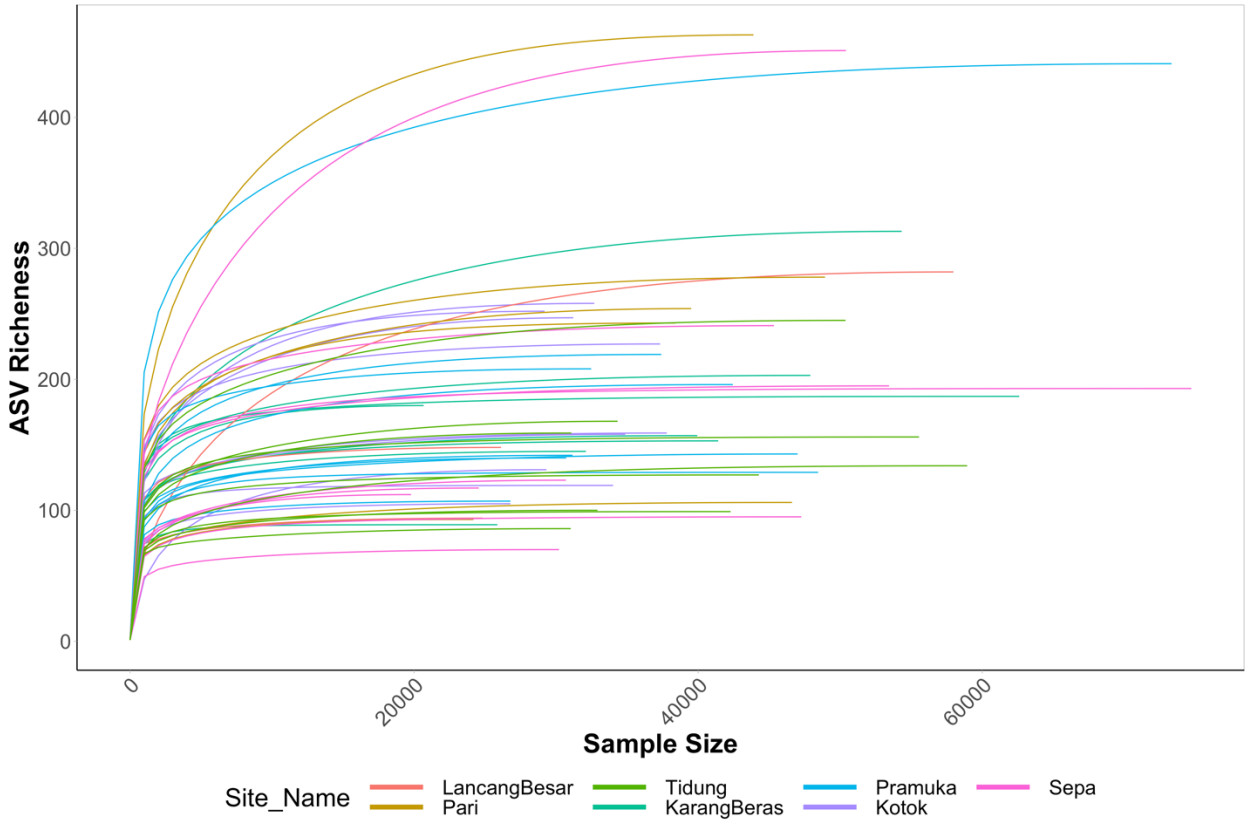


Figure 2-3. Rarefaction curves showing ASV richness vs. sequencing effort for 18S-ARMS data.

CO1-eDNA

Following the removal of blanks and control samples, we collected 24 samples from eight sites (Supplemental Table S2-3). Following post quality filtering and the exclusion of chimeras, we generated 1,260,826 reads and 1837 ASVs. Rarefaction curves show sequencing depth was sufficient to saturate for ASV discovery (Fig. 2-4). Following rarefaction to 11,000 reads per sample, the final data set included 264,000 reads and 1652 ASVs for downstream analysis.

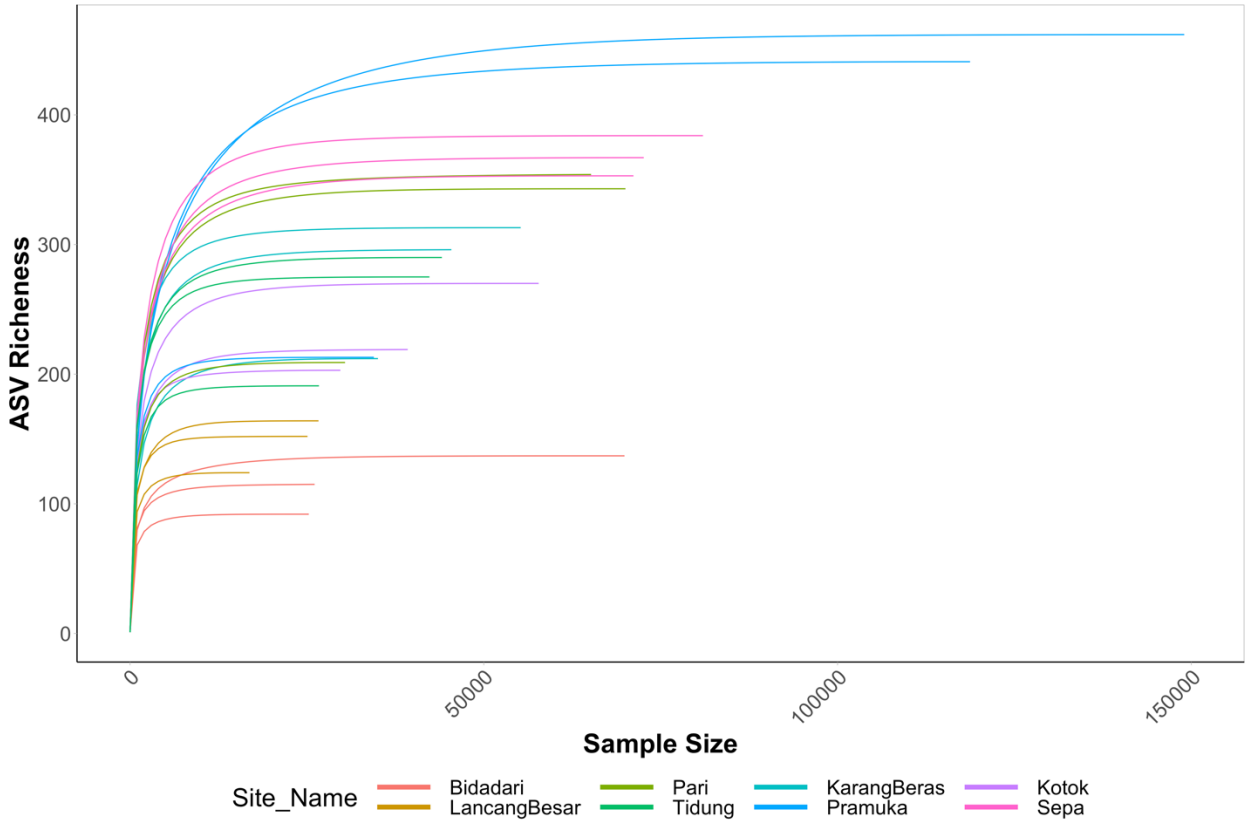


Figure 2-4. Rarefaction curves showing ASV richness vs. sequencing effort for COI-eDNA data.

Number of ASVs

Total ASVs Diversity across Pollution Gradient

For COI ARMS data, total ASV diversity ranged from a high of 531 ASVs at Lancang Besar and a low of 147 ASVs at Kotok, with an average of 319.9 ASVs per locality across all sites (Table 2-2). For 18S ARMS data, total ASV diversity ranged from a high of 710 ASVs at Pari and a low of 247 ASVs at Karang Beras, with an average of 465.9ASVs per locality across all sites (Table 2-2). For COI eDNA data, total ASV diversity ranged from a high of 355 ASVs at Sepa and a low of 92 ASVs at Bidadari, with an average of 236.6 ASVs per locality across all sites.

Following tests of normality and homoscedasticity, ANOVA only found significant differences in ASV diversity in the eDNA-COI dataset; no significant differences were observed in either ARMS data set (Table 2-3).

Table 2-2. Maximum and minimum number ASVs across sites in both eDNA and ARMS. No ARMS can be recovered Bidadari and we lost two ARMS in Lancang Besar, therefore no data were available for the sites

Locality	Chlorophyll (mg/l)	Temp. (°C)	COI-eDNA ASV			18S-ARMS ASV			COI-ARMS ASV		
			Mean	Max	Min	Mean	Max	Min	Mean	Max	Min
Bidadari	8.53	30.40	110	126	92						
Lancang Besar	3.33	30.15	145	162	123	478	-	-	531	-	-
Pari	0.82	29.98	281	322	204	609	710	509	250	285	216
Tidung	0.43	29.97	244	276	191	361	366	358	297	382	223
Karang Beras	0.44	29.95	256	291	199	405	526	247	364	470	283
Pramuka	0.66	29.96	304	352	209	489	600	433	328	390	288
Kotok	0.45	29.91	222	256	201	459	603	327	221	331	142
Sepa	0.47	29.85	331	355	317	460	695	328	248	271	207
Mean	1.89	30.02	236.6	267.5	192	465.9	583.3	367	319.9	354.8	265.5

Linear Regression between Chlorophyll a Concentration and Total Taxonomic Richness

Linear regression showed that Chlorophyll a concentration, a proxy for pollution stress, was not a significant predictor for ARMS COI (R^2= 0.15, *p-value* = 0.05798) or ARMS 18S (R^2= 0.013, *p-value* = 0.283) diversity (Fig. 2-5A). However, Chlorophyll a concentration was a

significant predictor for total ASVs for COI-eDNA data (R^2= 0.67, *p-value* = 0.000595) (Figure 2-5B).

Table 2-3. Normality and Homoscedasticity test based on marker and sampling methods. All markers in all sampling method fulfill the homoscedasticity requirement hence can undergo linear regression analysis

	Normality (Shapiro Test)	**Homoscedasticity (Bartlett Test)**	**One-Way Anova Test**
COI eDNA ASV	0.3967	0.3748	0.000375
18S ARMS ASV	0.3316	0.05574	0.601
COI ARMS ASV	0.3967	0.8241	0.0559

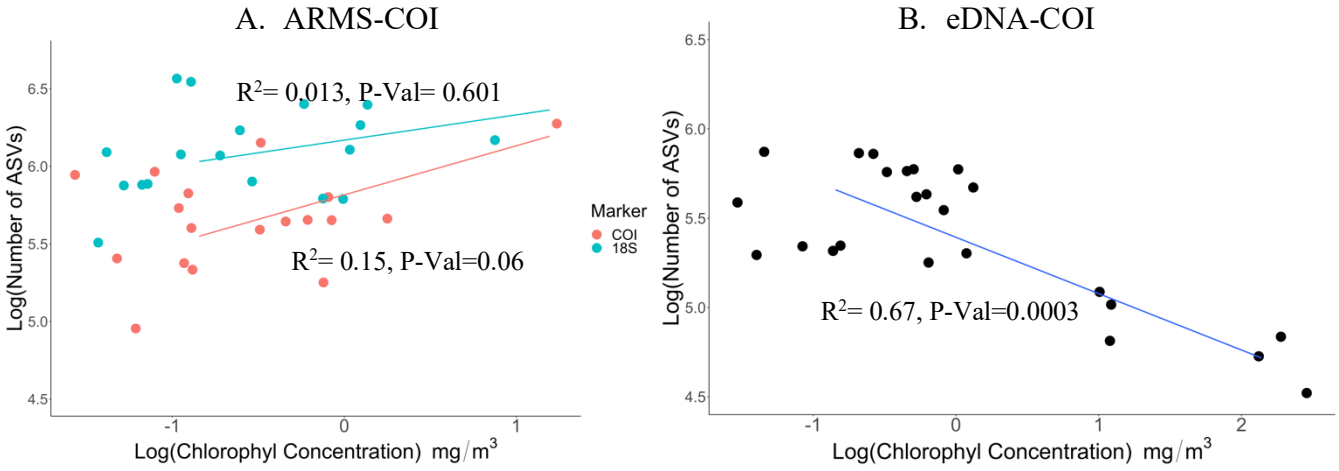


Figure 2-5. Plot of total ASVs as a function of Chlorophyll a concentration on A) ARMS-COI ARMS-18S B) COI-eDNA including best fit line and R^2 values. Only plot from COI-eDNA data is significant.

In addition to linear regression of ASV diversity and Chlorophyll a concentration, we also ran linear mixed models that incorporated both distance and temperature. However, upon testing we found collinearity between distance and temperature (model a); as such, we discarded the model and used only model b that included distance and chlorophyll. We found that Distance + Chlorophyll (model b.) was the best model with the lowest AIC value (195.179). This model found a significant main effect of Chlorophyll a (beta = 78.67, t = 2.65, p = .04) but not with the distance (beta = -1.917, t = -0.891, p = 0.4) indicating that ASVs abundance on ARMS as inferred by COI changed in response to chlorophyll concentrations, a proxy for pollution. For the 18S data, the same model also has the best AIC value (603.90), but unlike COI, there was no significant effects of Chlorophyll (beta = -6.41, t = -0.32, p = .75) or Distance (beta = -0.22, t = -1.88, p = .86) on ASV diversity across sites.

For eDNA data, models a and b had a rank deficiency, suggesting insufficient variation or not enough sampling points to estimate the desired model. Therefore, we only tested the remaining three models. Unlike ARMS, AIC scores for eDNA model c were the lowest (166.0733) compared to model d (175.11) and e (231.37), but there were no significant effects of distance to number of ASVs across sites (beta = 186.00, t = 1.75, p = .1).

Taxonomic Richness of Phyla across Pollution Gradient

Linear Regression between Taxa Richness and Chlorophyl A. Concentration

For common phyla that passed tests of normality (Saphiro-Wilk test) and homogeneity (Breusch-Pagan test), we conducted linear regression of ASV richness and chlorophyl concentration. Results showed that Chlorophyll a concentration is a not significant predictor of ASV diversity for most of the phyla detected with COI, for either the ARMS or eDNA data sets (except for

Ascomycota in COI-eDNA). Although with no significant correlation (Table 2-4.), exceptions were a positive relationship between Chlorophyll a concentrations and Platyhelminthes diversity and in both COI-ARMS and COI-eDNA datasets. Results also show similar pattern for Annelid diversity from the COI-ARMS data and Chaetognatha from COI-eDNA data (Fig. 2-6A and B).

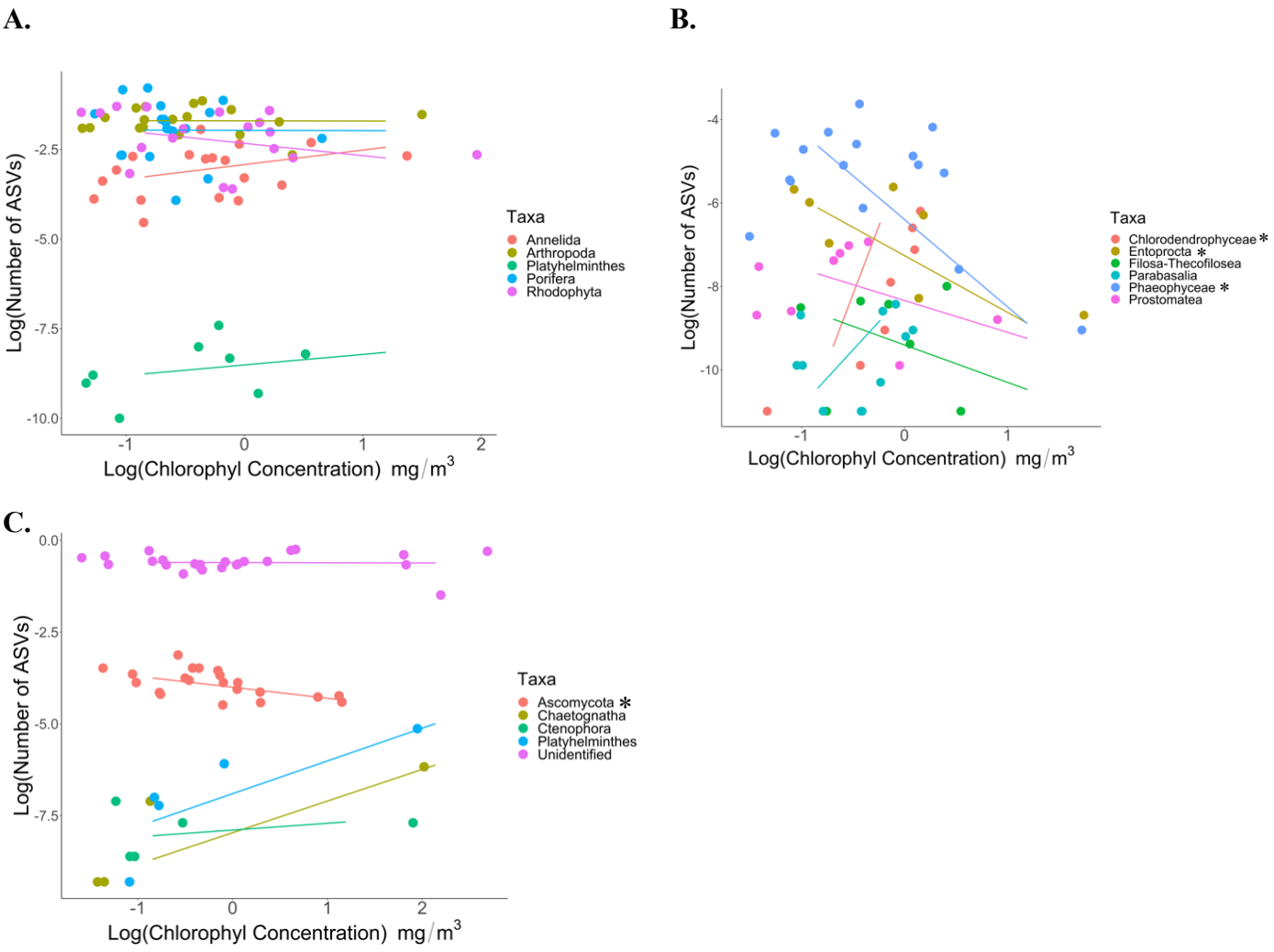


Figure 2-6. Plot of total ASVs per taxa as a function of Chlorophyll a concentration on A). ARMS-COI B). ARMS-18S and C). COI-eDNA. Only data that pass normality and heteroscedasticity test are presented. Taxa that are present in less than ten data point were also not presented. Asterixis indicate taxa with significant p-values.

Table 2-4. P-value calculated from linear regression modelling between ASV richness per taxa versus Chlorophyll a concentration. Taxa with significant p-value were highlighted in bold

A. COI-ARMS

Variable	r.squared	adj.r. squared	sigma	statistic	p. value	df	logLik	AIC	BIC	deviance	df. residual	nobs
Annelida	0.073	0.015	0.461	1.254	0.279	1	-10.542	27.083	29.754	3.4	16	18
Arthropoda	0.0000433	-0.062	0.479	0.001	0.979	1	-11.22	28.441	31.112	3.666	16	18
Platyhel-minthes	0.059	-0.098	0.704	0.373	0.564	1	-7.392	20.785	21.023	2.973	6	8
Porifera	0.0000182	-0.062	0.479	0	0.987	1	-11.221	28.441	31.113	3.667	16	18
Rhodophyta	0.045	-0.015	0.468	0.754	0.398	1	-10.806	27.612	30.283	3.501	16	18

B. 18S-ARMS

Variable	r.squared	adj.r. squared	sigma	statistic	p. value	df	logLik	AIC	BIC	deviance	df. residual	nobs
Chloroden drophyceae	0.594	0.513	0.148	7.313	**0.043**	1	4.614	-3.228	-3.391	0.11	5	7
Entoprocta	0.576	0.491	0.499	6.799	**0.048**	1	-3.886	13.771	13.609	1.244	5	7
Filosa-Thecofilosea	0.222	0.092	0.607	1.709	0.239	1	-6.205	18.411	18.649	2.21	6	8
Parabasalia	0.353	0.281	0.188	4.908	0.054	1	3.906	-1.811	-0.618	0.317	9	11
Phaeophyceae	0.525	0.491	0.345	15.463	**0.002**	1	-4.587	15.173	17.491	1.662	14	16
Prostomatea	0.232	0.122	0.606	2.109	0.19	1	-7.136	20.272	20.864	2.573	7	9

C. COI-eDNA

Variable	r.squared	adj.r. squared	sigma	statistic	p. value	df	logLik	AIC	BIC	deviance	df. residual	nobs
Ascomycota	0.29	0.253	0.576	7.761	**0.012**	1	-17.146	40.292	43.426	6.294	19	21
Chaetognatha	0.606	0.409	1.103	3.078	0.221	1	-4.68	15.36	13.519	2.431	2	4
Ctenophora	0.058	-0.255	0.965	0.186	0.695	1	-5.638	17.276	16.104	2.792	3	5
Platyhel-minthes	0.511	0.347	1.007	3.13	0.175	1	-5.852	17.704	16.532	3.041	3	5
Unidentified	0	-0.045	1.063	0.01	0.919	1	-34.478	74.957	78.491	24.863	22	24

In contrast, three phyla showed strong relationships between Chlorophyll a concentration and ASV diversity as inferred from 18S ARMS data. Diversity of Chlorodendrophyceae (marine green algae) and increased with higher Chlorophyll a concentration (R^2=0.594, p = 0.043). Parabasalia (protist) also increased and had a strong R^2=0.353 but this value was just short of significant, p = 0.54. In contrast, Phaeophyceae (Brown algae) and Entoprocta richness all significantly decreased with increasing Chlorophyll concentration ($R^{2.}$ = 0.535, p = 0.002; $R^{2.}$ = 0.576, p = 0.048, respectively, Fig. 2-6C). Although Protosmatea (ciliates) and Filosathecofilosea (protist) also decreased, their R^2values were not significant.

Taxa Contributing to Dissimilarities Among Sites

Simper analysis *of* COI-ARMS data showed 15 unique taxa (e.g., Scaphandridae, Coelosphaeridae, Microcionidae and Rhodomelaceae etc.) that significantly contributed to the dissimilarities among sites (Supplemental Table S2-4); For example, ASV2475 (Demospongidae) has a significantly higher relative abundance in Kotok (~28%) where Chlorophyll a concentrations are lower and low abundance in Lancang Besar and Pari (~ 0.1%) (Fig. 2-7) where Chlorophyll a concentrations are higher.

For 18S-ARMS data, *Simper* analysis identified, 47 ASVs that contributed to dissimilarities across sites, although none were statistically significant following Kruskal-Wallis rank-sum test. Taxa that exhibited large changes in relative abundance include ASV16946 (Florideophyceae), which was approximately 20% of sequence reads from Lancang Besar where Chlorophyll A concentrations are relatively high, and lower abundance in other sites (~ 0.1%) where Chlorophyll concentrations are lower. Similarly, the crustacean ASV14015 had higher abundances in Lancang Besar (~ 3%) where Chlorophyll a concentrations are relatively high but lower relative abundance in Kotok, Pramuka and Tidung (~ 0.1%) (Fig. 2-7) where Chlorophyll concentrations are lower. A full list of taxa identified by SIMPER is available in supplemental Table S2-5.

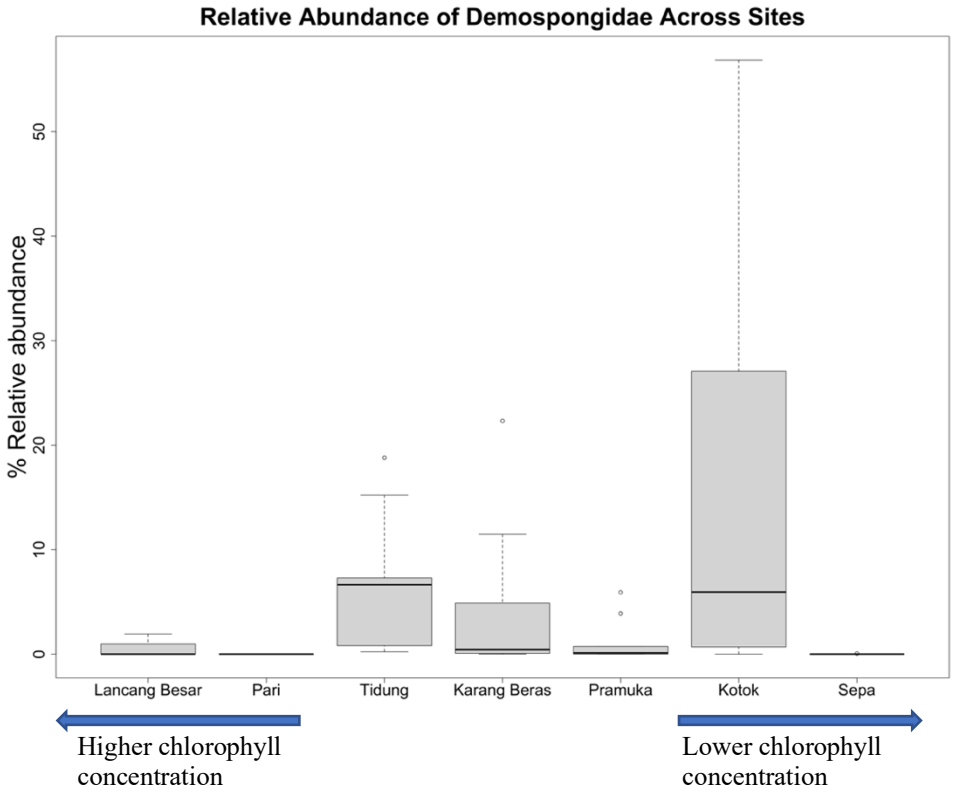


Figure 2-7. A box plot describing relative abundance of Demospongiae across sites. Kotok has significantly higher abundance compared to the remaining site

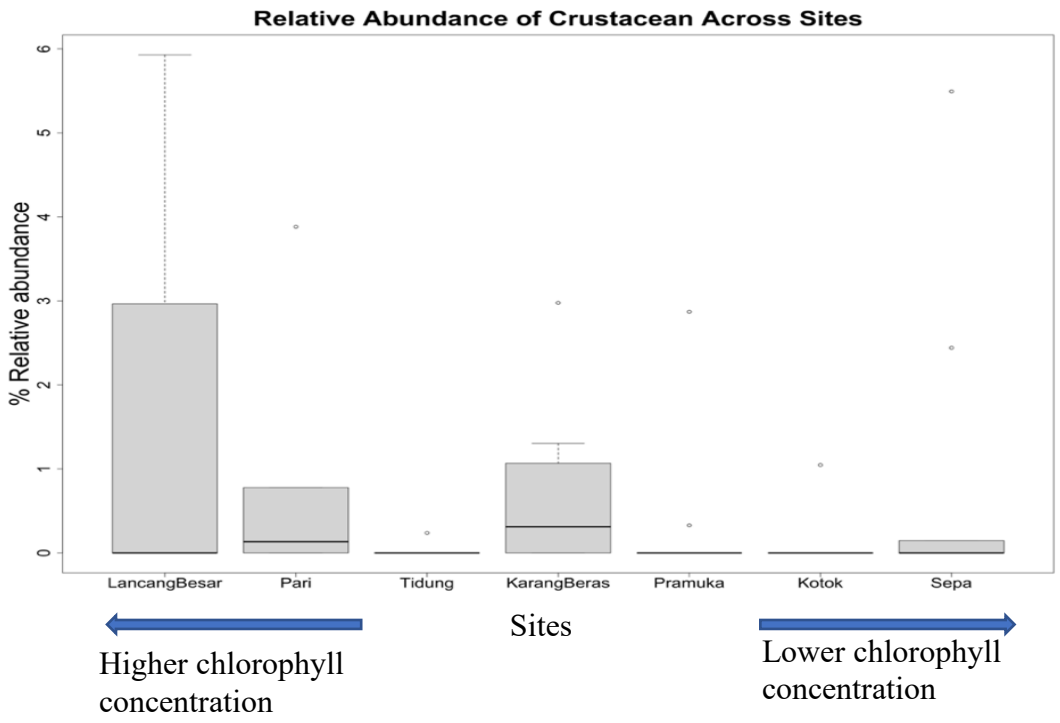


Figure 2-8. A box plot describing relative abundance of Crustaceans across sites. Lancang Besar has significantly higher abundance compared to the remaining sites

eDNA data from COI showed 30 taxa (e.g., Micromonas, Ilyonectria, Undinula and parvocalanus etc.) that significantly contributed to the dissimilarities among sites. For example, ASV1329 (Micromonas) has a significantly higher relative abundance in all sites (~ 13 – 35 %) where Chlorophyll a concentrations are lower than in Bidadari and Lancang Besar (~ 0.1 – 5 %) (Fig. 2-9) where Chlorophyll a concentrations are higher. A full list of taxa identified by SIMPER is available in supplemental Table S2-6.

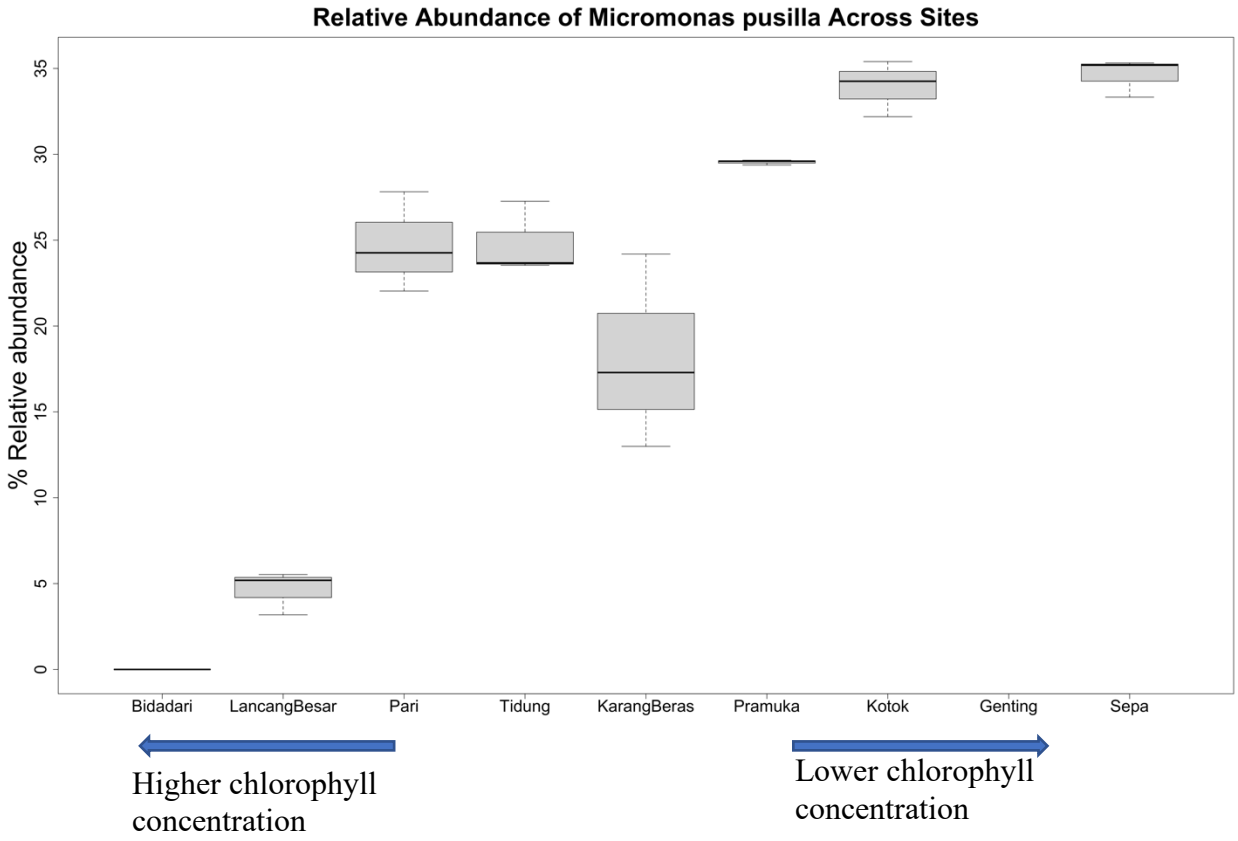


Figure 2-9. A box plot describing relative abundance of Micromonas across sites. Kotok and Sepa has significantly higher abundance compared to the remaining sites.

Beta Diversity Across Deployment Sites

Analysis of COI-ARMS data with both *ADONIS* and *PERMANOVA* of Jaccard distances showed significant differences in species composition and heterogeneity based on location ($p < 0.05$ and *ADONIS* has $R^2 = 0.44$). In total, deployment sites contributed 44 % of the variance in species composition. Differentiation among sites is particularly noticeable among ARMS deployed furthest from pollution sources relatively to those deployed closer to polluted sites (Fig. 2-10). Although pairwise *ADONIS* of ARMS-COI data shows that many of the sites are different in their species composition, all are non-significant ($p > 0.05$) (Table 5).

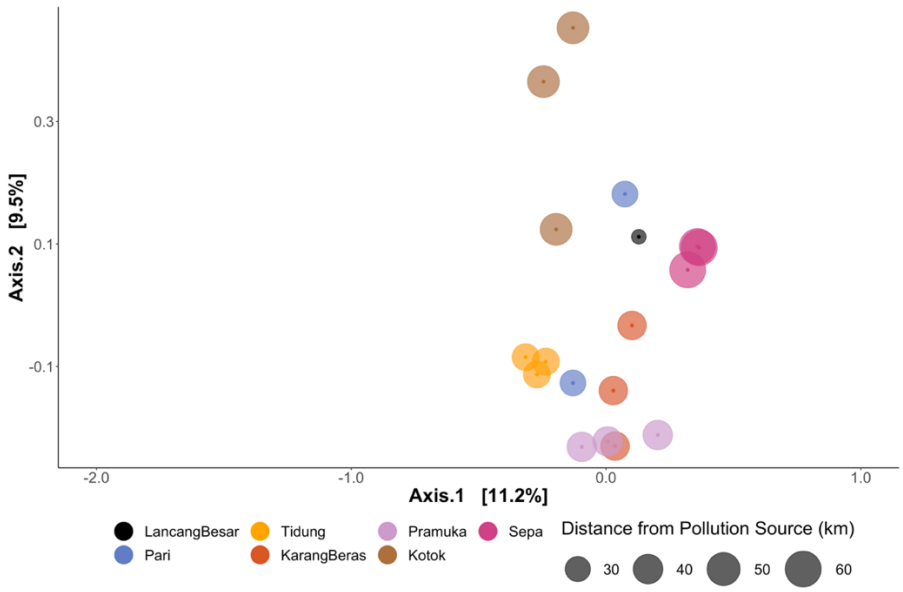


Figure 2-10. Principal Coordinates Analysis (PCoA) analysis illustrating dissimilarities in community composition using COI in 18 ARMS. Pollution levels are commensurate to distance (km) from Capital city, Jakarta which are reflected by the point size. Lancang Besar is the most polluted site while Kotok is least polluted.

Similarly, analysis of 18S on ARMS data with *ADONIS* and *PERMANOVA* tests using Jaccard distances showed significant differences based on location (Fig. 2-11. 18S total, $p < 0.05$ and ADONIS has R^2 of 0.43), indicating significant differences in overall species composition and species heterogeneity among sites. Deployment sites contributed ~43 % of the variation in species composition among sites. The signal of separation between sites is particularly noticeable among the sites closest and furthest from pollution sources as inferred by Chlorophyll a concentration (Fig. 10). However, site's position within PCoA plot doesn't represent clear pollution gradient (i.e., one ARMS from Tidung was within the proximity of ARMS from Lancang Besar)

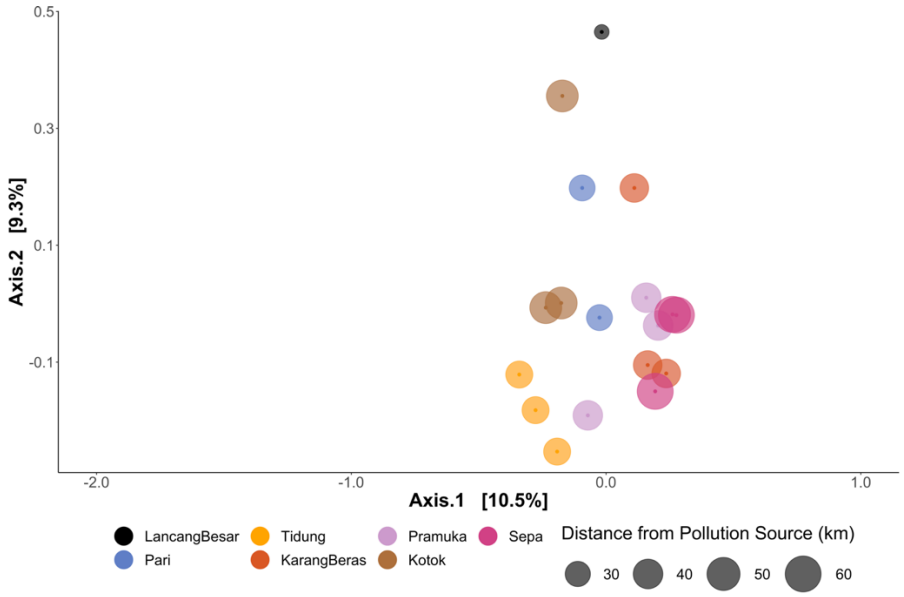


Figure 2-11. Principal Coordinates Analysis (PCoA) analysis illustrating dissimilarities in community composition from 18 ARMS deployed across Kepulauan Seribu. Analysis was undertaken using Bray-Curtis similarities on the full dataset across sampling locations. Pollution levels are commensurate to distance (km) from Capital city, Jakarta which are reflected by the point size. Lancang Besar is the most polluted site while Kotok is least polluted.

Pairwise *ADONIS* showed that Kotok, the second furthest site from pollution source in Jakarta, was significantly different from Karang Beras, Pari, Pramuka, Tidung and Sepa (Table 6, row 2-6). In contrast, Lancang Besar, the most polluted site was significantly different only from Tidung (Table 6, row 11). Similarly, Sepa, the furthest site from pollution source in Jakarta, was significantly different than Tidung and Pari (Table 6, row 17 and 21).

Table 2-5. P-values from pairwise ADONIS test on the Jaccard diversity metric using COI-ARMS data. P-values within a given pair were all non-significant.

	pairs	Df	SumsOfSqs	F.Model	R2	p.value
1	KarangBeras vs Kotok	1	0.6666135	1.942163	0.3268445	0.1000000
2	KarangBeras vs LancangBesar	1	0.4300184	1.206651	0.3762963	0.2500000
3	KarangBeras vs Pari	1	0.5479526	1.532687	0.3381410	0.1000000
4	KarangBeras vs Pramuka	1	0.4462042	1.212110	0.2325565	0.1000000
5	KarangBeras vs Sepa	1	0.5169565	1.457516	0.2670658	0.1000000
6	KarangBeras vs Tidung	1	0.5717560	1.596103	0.2852169	0.1000000
7	Kotok vs LancangBesar	1	0.4917004	1.489589	0.4268666	0.2500000
8	Kotok vs Pari	1	0.5555519	1.634026	0.3526148	0.1000000
9	Kotok vs Pramuka	1	0.5932843	1.671315	0.2946962	0.1000000
10	Kotok vs Sepa	1	0.6719883	1.967512	0.3297039	0.1000000
11	Kotok vs Tidung	1	0.5678711	1.645627	0.2914871	0.1000000
12	LancangBesar vs Pari	1	0.4376128	1.216314	0.5488004	0.3333333
13	LancangBesar vs Pramuka	1	0.4042167	1.064093	0.3472783	0.5000000
14	LancangBesar vs Sepa	1	0.4711356	1.334688	0.4002438	0.2500000
15	LancangBesar vs Tidung	1	0.4581399	1.272377	0.3888235	0.2500000
16	Pari vs Pramuka	1	0.4856152	1.301306	0.3025375	0.1000000
17	Pari vs Sepa	1	0.5661325	1.593584	0.3469152	0.1000000
18	Pari vs Tidung	1	0.5230099	1.452915	0.3262840	0.1000000
19	Pramuka vs Sepa	1	0.5169339	1.410725	0.2607275	0.1000000
20	Pramuka vs Tidung	1	0.4323134	1.168516	0.2260834	0.2000000
21	Sepa vs Tidung	1	0.6297657	1.766377	0.3063235	0.1000000

As with ARMS data, *ADONIS* analysis of COI-eDNA data showed significant variation in ASV composition among sites ($p < 0.05$, $R^2 = 0.62$); however, *PERMANOVA* results were non-

significant ($p > 0.05$). Deployment sites contributed 62 % to the differences in species composition. Lancang Besar and Bidadari, the most polluted sites, clustered away from the remaining sites, while Karang Beras, also clustered separately (Fig. 2-12).

Table 2-6. P-values from pairwise ADONIS test on the Jaccard diversity metric using 18S-ARMS data. Significant P-values were coded with asterisk in the table.

	pairs	Df	SumsOfSqs	F.Model	R2	p.value	p.adjusted	sig
1	Kotok vs LancangBesar	1	0.5442194	1.206098	0.10762869	0.027	0.120	
2	Kotok vs KarangBeras	1	0.6542262	1.490983	0.09041202	0.001	0.021	*
3	Kotok vs Pari	1	0.5684724	1.270055	0.08900142	0.001	0.021	*
4	Kotok vs Pramuka	1	0.5937026	1.321783	0.07630756	0.001	0.021	*
5	Kotok vs Sepa	1	0.6564622	1.460434	0.08364250	0.001	0.021	*
6	Kotok vs Tidung	1	0.5751869	1.300288	0.07515990	0.004	0.044	*
7	LancangBesar vs KarangBeras	1	0.5663919	1.298217	0.12606233	0.005	0.045	*
8	LancangBesar vs Pari	1	0.5194303	1.149368	0.14103767	0.024	0.120	
9	LancangBesar vs Pramuka	1	0.5277710	1.164692	0.10431923	0.027	0.120	
10	LancangBesar vs Sepa	1	0.5571027	1.227991	0.10936873	0.032	0.120	
11	LancangBesar vs Tidung	1	0.5804615	1.312552	0.11602614	0.003	0.036	*
12	KarangBeras vs Pari	1	0.5518569	1.265465	0.09539546	0.002	0.028	*
13	KarangBeras vs Pramuka	1	0.5195890	1.180703	0.07296980	0.017	0.119	
14	KarangBeras vs Sepa	1	0.5529236	1.255447	0.07723241	0.004	0.044	*
15	KarangBeras vs Tidung	1	0.7097069	1.639808	0.09854727	0.001	0.021	*
16	Pari vs Pramuka	1	0.5248954	1.168843	0.08249391	0.018	0.119	
17	Pari vs Sepa	1	0.5522212	1.228582	0.08634604	0.005	0.045	*
18	Pari vs Tidung	1	0.5805478	1.317374	0.09201228	0.002	0.028	*
19	Pramuka vs Sepa	1	0.5424239	1.203522	0.06995788	0.024	0.120	
20	Pramuka vs Tidung	1	0.5975269	1.347138	0.07765767	0.001	0.021	*
21	Sepa vs Tidung	1	0.6788003	1.529234	0.08723905	0.001	0.021	*

Pairwise ADONIS using eDNA-COI showed that many of the sites differed in their species composition, however none of the pairwise were significant (Table 2-6). For example, Karang Beras and Lancang Besar are different in their species composition with Kotok, Pari, Pramuka and Sepa.

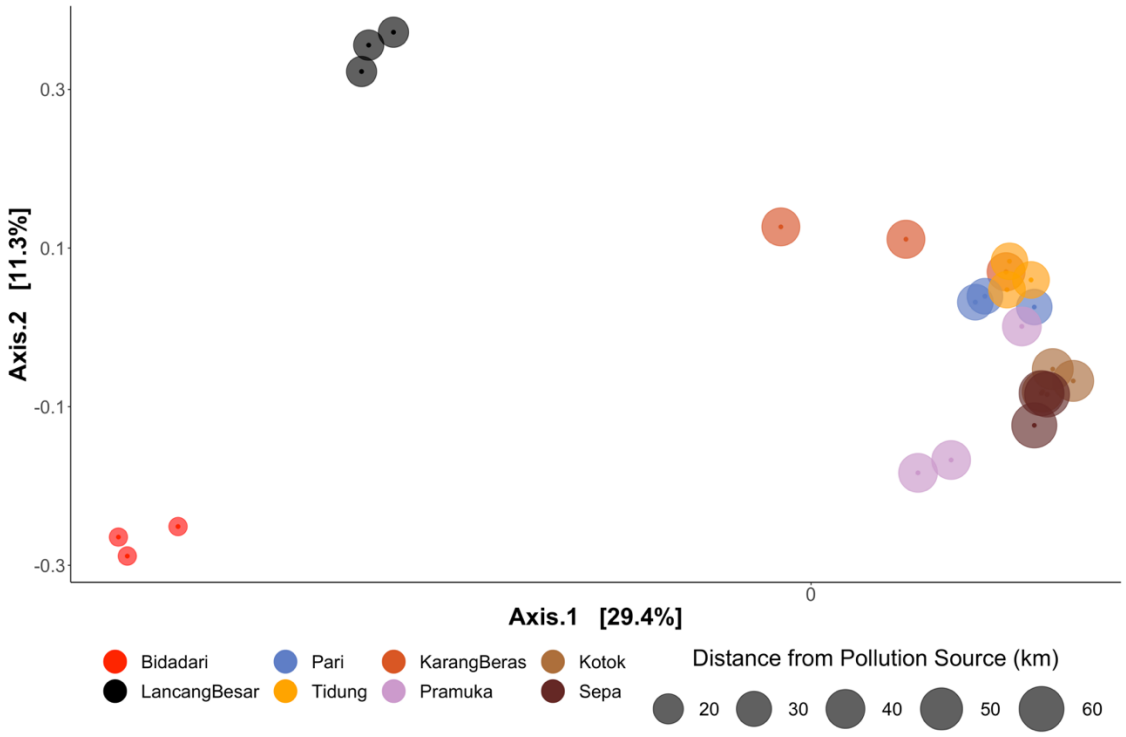


Figure 2-12. Principal Coordinates Analysis (PCoA) analysis illustrating dissimilarities in community composition using COI-eDNA in 18 ARMS. Pollution levels are commensurate to distance (km) from Capital city, Jakarta which are reflected by the point size. Lancang Besar is the most polluted site while Kotok is least polluted.

Table 2-7. P-values from pairwise ADONIS test on the Jaccard diversity metric using COI-eDNA data. No significant P-value were found among the pairs

	pairs	Df	SumsOfSqs	F.Model	R2	p.value
1	KarangBeras vs Kotok	1	0.3149801	1.498928	0.2725854	0.1
2	KarangBeras vs LancangBesar	1	0.5215688	2.332571	0.3683450	0.1
3	KarangBeras vs Pari	1	0.3231376	1.423164	0.2624232	0.1
4	KarangBeras vs Pramuka	1	0.3333854	1.402492	0.2596010	0.1
5	KarangBeras vs Sepa	1	0.2942595	1.395160	0.2585947	0.1
6	KarangBeras vs Tidung	1	0.3017553	1.386498	0.2574024	0.1
7	Kotok vs LancangBesar	1	0.6571705	3.569450	0.4715600	0.1
8	Kotok vs Pari	1	0.3501551	1.866867	0.3182051	0.1
9	Kotok vs Pramuka	1	0.2399068	1.210327	0.2322939	0.2
10	Kotok vs Sepa	1	0.2838961	1.656127	0.2928022	0.1
11	Kotok vs Tidung	1	0.2766598	1.552997	0.2796683	0.1
12	LancangBesar vs Pari	1	0.6290469	3.129142	0.4389227	0.1
13	LancangBesar vs Pramuka	1	0.6560719	3.099327	0.4365663	0.1
14	LancangBesar vs Sepa	1	0.6674710	3.610150	0.4743862	0.1
15	LancangBesar vs Tidung	1	0.6143330	3.206142	0.4449180	0.1
16	Pari vs Pramuka	1	0.3287603	1.528156	0.2764314	0.1
17	Pari vs Sepa	1	0.3439345	1.826130	0.3134380	0.1
18	Pari vs Tidung	1	0.2414276	1.237681	0.2363032	0.1
19	Pramuka vs Sepa	1	0.2711270	1.362488	0.2540775	0.1
20	Pramuka vs Tidung	1	0.2937009	1.427686	0.2630377	0.1
21	Sepa vs Tidung	1	0.3205920	1.791784	0.3093665	0.1

Discussion

ARMS and eDNA metabarcoding on the coral reefs of Pulau Seribu captured a wide variety of marine biodiversity, and this community diversity changed as a function of pollution stress. Specifically, ASV diversity inferred from eDNA demonstrated a significant, sharp decrease with increasing Chlorophyll a concentration, a pattern previously reported in multiple taxa including coral (Smith et al. 2008, Estradivari et al. 2009, van der Meij et al. 2010, Baum et al. 2015), reef fish (Manikandan et al. 2014, Brown et al. 2017, Ling et al. 2018) and many

benthic macro invertebrates (Rees et al. 1999, Cebrian et al. 2007, van der Meij et al. 2009, Johnston & Roberts 2009). In contrast, ASV diversity from ARMS showed a slight, but not significant, increase. Given that ARMS and eDNA capture largely non-overlapping taxa (Chapter 1), these results suggest that the taxa that colonize ARMS may not be as sensitive to pollution stress as those recovered by eDNA. As such, eDNA may be a more effective tool for monitoring community change.

Despite the differences in ARMS and eDNA to detect overall community change, both methods captured specific taxa that increased or decreased in relation to pollution stress. Previous studies indicate taxonomic differences in sensitivity to pollution stress (Johnston & Roberts 2009, Pelletier et al. 2010, Ivanina & Sokolova 2015, Gissi et al. 2016). As such, monitoring for these specific taxa—either using ARMS or eDNA—could provide insights into changing environmental conditions. Given the limited overlap between methods, application of both methods would maximize pollution monitoring sensitivity, providing conservation managers the greatest ability to detect environmental changed through regular monitoring activities.

Sensitivity of Monitoring Methods to detect Species impacted by Pollution gradient

Previous studies examining the impact of pollution on marine communities have relied primarily on visual census and species-specific observation data (Sale & Douglas 1981, Nagelkerken et al. 2000, Willis 2001, Cleary et al. 2005, Campbell et al. 2011). Results from these studies suggested that taxa have different sensitivities depending on the severity of the pollution and the structural and habitat complexity of where the taxa are settled. For example, the majority of branching coral in Kaneohe Bay exhibited reduced growth rates and smaller colony

size in response to pollution stress, but massive and non-branching corals did not (Pastorok & Bilyard 1985). Similarly, benthic macroinvertebrate diversity and richness can decline in response to pollution (Widbornl & Elmgren 1988, Pinedo et al. 2015, Ellis et al. 2017, Piroddi et al. 2021), some taxa can adapt to tolerate elevated pollution levels (reviewed by Sanford & Kelly 2011).

Many studies have examined the impact of pollution stress on reef ecosystems of Kepulauan Seribu, (Rees et al. 1999, Rachello-Dolmen & Cleary 2007, van der Meij et al. 2009, Hadi 2011, Polónia et al. 2014, Baum et al. 2015, Cleary 2017). These studies focus on taxa with significant ecological roles or that are important for economic and tourism activities (e.g., coral, reef fish, many taxa from Echinodermata, although one examines responses of microbial communities (Polónia et al. 2014). In general, results indicate shifts in community composition relative to pollution stress as measured by distance from Jakarta. For example, Maduppa et al. 2013, showed that the diversity, richness and abundance of reef fish declined substantially on reefs closer to pollution sources. Polluted areas also had more herbivorous fishes, potentially a response to increased algae coverage on eutrophied reefs. Similarly, pollution decreased richness and percent coral cover on reefs closer to pollution source, with massive coral species (*Oulastrea crispate* and *Favia maxima*) dominating sites with higher sedimentation and nutrient enrichment (Cleary et al. 2005, 2008). Studies that examine a larger diversity of taxa, including mollusc, sponge, echinoderms, coral, fish, large benthic foraminifera report lower diversity on reefs closer to Jakarta Bay (Cleary et al. 2016). The broad similarity of metabarcoding results to the above visual monitoring studies suggests that metabarcoding can be a useful tool for monitoring changes in marine communities in response to anthropogenic stressors.

ARMS vs eDNA for Marine Ecosystem Monitoring

Because both ARMS and eDNA employ community based metabarcoding and our samples were co-located, we expected these two methods to provide similar results. However, community diversity on ARMS did not significantly change with respect to chlorophyll-a concentrations, although some taxonomic groups had significant increases or decreases. In contrast, eDNA captured a sharp decrease in total community diversity with increasing Chlorophyll-a concentrations.

Although ours is the first study to directly compare ARMS and eDNA, the different results from ARMS and eDNA communities align with previous studies. For example, Pearman et al (2019) show that microbial communities on ARMS shifted in response to pollution stress (Pearman et al. 2019), but other ARMS studies examining macrobiota report equivocal results (David et al. 2019). In contrast, eDNA studies more consistently detect the impacts of anthropogenic stress on coastal ecosystems (Bakker et al. 2017, Polanco et al. 2022), although in some cases in unexpected ways. For example, DiBattista et al. (2020) showed higher marine biodiversity in regions of Japan characterized by medium to high anthropogenic pollution stress, suggesting that "intermediate disturbance theory", could elevate total biodiversity (Reynolds et al. 1993, Townsend et al. 1997).

Given that samples were co-located and chlorophyll a concentration differed by an order of magnitude among sites, it is unclear why ARMS and eDNA data showed such different patterns. One potential explanation is that chlorophyll-a concentration, as a proxy for pollution stress, is among many factors that can impact coral reefs communities (Dubinksy & Stambler 1996, Bachoon et al. 2010, Jessen et al. 2014, Baum et al. 2015, Guo et al. 2019). Sediments, and toxic substances are also important pollutants that drive changes in marine community

composition (Pastorok & Bilyard 1985). Pollution from Jakarta and surrounding communities almost certainly contains all three components, yet sediment and toxins cannot be inferred from Chlorophyll-a concentrations. Given significant interactions between anthropogenic stressors (Fong et al. 2018), by focusing only on Chlorophyll-a, our study likely provides an incomplete picture of how pollution stress in Jakarta Bay is impacting these marine ecosystems.

Additionally, ARMS and eDNA capture very different communities with limited overlap (chapter 1). Moreover, these communities are very different in their ecologies. Communities from eDNA analysis captured a community comprised largely of small, planktonic taxa. In contrast, ARMS communities include a number of large, encrusting organisms, particularly within the sessile fraction. As such, accumulation of biodiversity on ARMS can be influenced by priority effects (Benedetti-Cecchi 2000, Adam et al. 2022), whereby an established resident influences/inhibits the colonization of newer individuals entering the plate, influencing community composition. For example, many ARMS plates include marine sponges (genus Cinachyrella) that can inhibit other taxa (Singh & Thakur 2017), potentially limiting settlement and accumulation of other benthic taxa. Moreover, because ARMS are deployed for three years the community is likely influenced by many factors beyond pollution stress, including dynamics of top-down consumers, climate variability (Rasher et al., 2020), disturbance history (Reed et al., 2011), and community diversity (Nell et al., 2018).

Lastly, the inability of ARMS to detect significant differences in community composition may result from sampling design. Pearman et al (2019) deployed 3 x 3 ARMS in each site to get. This increased sampling could contribute to greater ability to detect differences. Perhaps more importantly, although we deployed three ARMS per sites across 8 sites spanning varied pollution stress levels, we lost five ARMS, all from the two most polluted sites. Loss of ARMS from sites

likely to be most impacted by heavy pollution stress likely reduced our power to capture significant biodiversity changes, although what changes were observed appeared to be increases in diversity, rather than decreases, as observed in eDNA.

Beta Diversity Across Pollution Gradient

Although the impacts of pollution stress on alpha diversity differed between ARMS and eDNA, beta diversity analysis confirms that sites are strongly structured based on chlorophyll a concentrations. Both ARMS and eDNA data sets showed significant differences in beta diversity among reefs with different levels of pollution stress. Species composition and species heterogeneity among sites were different, with the greatest difference observed among sites closest and furthest from mainland pollution sources. However, these patterns were most pronounced for the eDNA-COI dataset.

Given the strong differences in alpha diversity in the eDNA dataset, differences in beta diversity are not unexpected. One explanation for the observation of significant differences in beta diversity in ARMS despite no differences in alpha diversity is that there is species replacement across our sites based on pollution stress, as suggested in other studies (Baselga 2010, Carvalho et al. 2012, Legendre 2014). The rationale behind this hypothesis is that interactions between pollution levels and the physical complexity of the reefs produces a wide variety of microhabitats. If these microhabitats are colonized by different groups of organisms, differences in taxa composition could result, increasing beta diversity. Further, the variability in niches could also promote specific co-occurrences, particularly from species that have similar ecological functions, hence creating site specific taxa composition. Hoeksema et al. (2019) report that variation in structural complexity impacts biodiversity distribution, suggesting niche

variation across Kepulauan Seribu. In addition, other factors such as wave energy, substrate stability, and sediment could interact to shape microhabitats (Trapon et al. 2013, Hamilton et al. 2017, Waltham & Sheaves 2018), resulting in greater beta diversity.

Sensitivity of Marine Phyla to Pollution Gradient

From a monitoring perspective, an important advantage of AMRS and eDNA metabarcoding is the ability to expand taxonomic coverage (Thomsen et al. 2012, Thomsen & Willerslev 2015, Valentini et al. 2016, Bakker et al. 2017, Valentin et al. 2020, Gold et al. 2021). By expanding taxonomic coverage, it is possible to identify species that are sensitive to pollution stress. Our results showed that many taxa groups increased or decreased as a function of pollution stress. Although coral reef monitoring typically focuses on fish and corals (REFS), our results indicate that smaller taxa are also good indicators of environmental stressors. Previous studies indicate *Platyhelminthes* are indicators for degraded marine environments (Sluys 1999, Bayoumy et al. 2008, Gilbert & Avenant-Oldewage 2021), as are *Annelids* due to their plasticity in using an array of organic material to survive (Reish 1970, Dauer & Simon 1976, Pocklington & Wells 1992, Giangrande et al. 2005). Although our data showed both *Platyhelminthes* and *Annelids* increasing with Chlorophyll-a concentrations, these increases were not significant. However, many other taxa (e.g., Chlorodendrophycea, Entroprocta, Phaeophyceae) did exhibit significant changes with pollution stress, suggesting that they could be excellent taxa to use for monitoring changes in local ecosystems.

Primer and Database Contribution to Gradient in Biodiversity

Metabarcoding markers have tradeoffs between taxonomic resolution and amplification efficiency (Drummond et al. 2015, Piñol et al. 2019, Macheriotou et al. 2019, Cordier et al. 2019, Bik 2021). As such, we used the slower evolving 18S rRNA marker to provide a broader overview of the eukaryotic domain and the faster evolving COI gent to provide higher taxonomic resolution across metazoans (Baird & Hajibabaei 2012, Deagle et al. 2014, Clarke et al. 2017, Andújar et al. 2018). Comparison of these two markers in the ARMS dataset showed similar patterns with respect to biodiversity changes across pollution levels. Despite similar patterns, these markers detected different suites of taxa responding to pollution stress.

Different primer sets bias metabarcoding results (Casey et al. 2021), likely due to differences in amplification efficiency. Although previous studies suggest that 18S is better than COI for capture diversity gradients (Tytgat et al. 2019) or these markers have equivalent performance (Pearman et al. 2018), our results indicated that COI on eDNA had the greatest power to detect community changes associated with pollution stress. However, this result is likely more a function of the communities captured by eDNA than the performance of COI, given the inability of the COI-ARMS dataset to detect significant changes in diversity with pollution stress.

Rather than indicating one marker outperformed another, our study highlights the value of both markers, depending on the question being asked. COI was a powerful marker for detecting shifts in alpha diversity related to pollution stress in the eDNA dataset. 18S performed best in identifying specific taxa in the ARMS data set responding to changes in pollution stress. [It is possible that 18S could have performed similarly well in the eDNA dataset, but our study did not collect 18S eDNA data, as COI (for metazoans) and 12S (for fishes) are more commonly

used]. As such, these markers should be viewed as complementary, each providing data and insights that the other cannot. Given that doubling markers doubles the costs, studies using single markers should choose carefully, contingent upon the study system, question, and the type of samples being analyzed.

Conservation Implications

Although Kepulauan Seribu is not the center of marine biodiversity in Indonesia, the archipelago has relatively rich coral and fish diversity, with 60 coral genera, 174 reef fish species, and 216 macro-benthic species (Estradivari et al. 2009, Madduppa et al. 2013), contributing to ~ 20% percent of total Indonesian marine species richness. Results from our documenting over 5000 COI ASVs, and nearly 9000 18S ASVs for 18S marker suggests that biodiversity of Kepulauan Seribu is an order of magnitude higher.

With escalating local and global anthropogenic stressors, it is essential to better understand how marine communities in this archipelago are affected by anthropogenic disturbances, primarily from the capital city Jakarta. Application of standardized metabarcoding monitoring protocols could greatly increase the number of taxa observed and identify taxa most sensitive to environmental changes for local resource managers to monitor. Expanding taxonomic coverage of monitoring efforts beyond fishes and coral could identify indicator species that can be used for early detection of degrading environmental conditions. With marine habitats under severe pressure, it is vitally important for conservation programs to detect changes as early as possible, allowing for early interventions to promote the sustainability of these important marine ecosystems, and the local populations who are dependent on them.

Supplemental Tables and Figures

Supplemental Table S2-1. Deployment location and status following recovery for COI-ARMS samples.

Sample ID	Site Name	ARMS Name	Fraction Size	Recovery Status	Remarks
SBDI1ADMSO	Bidadari	SLCB1A	Sessile	Lost	
SBDI1A500	Bidadari	SLCB1A	500	Lost	
SBDI1A100	Bidadari	SLCB1A	100	Lost	
SBDI1BDMSO	Bidadari	SLCB1B	Sessile	Lost	
SBDI1B500	Bidadari	SLCB1B	500	Lost	
SBDI1B100	Bidadari	SLCB1B	100	Lost	
SBDI1CDMSO	Bidadari	SLCB1C	Sessile	Lost	
SBDI1C500	Bidadari	SLCB1C	500	Lost	
SBDI1C100	Bidadari	SLCB1C	100	Lost	
SKBS1ADMSO	Karang Beras	SKBS1A	Sessile	Recovered	
SKBS1BDMSO	Karang Beras	SKBS1B	Sessile	Recovered	
SKBS1A100	Karang Beras	SKBS1A	100	Recovered	
SKBS1B100	Karang Beras	SKBS1B	100	Recovered	
SKBS1B500	Karang Beras	SKBS1B	500	Recovered	
SKBS1C500	Karang Beras	SKBS1C	500	Recovered	
SKBS1A500	Karang Beras	SKBS1A	500	Recovered	
SKBS1C100	Karang Beras	SKBS1C	100	Recovered	
SKBS1CDMSO	Karang Beras	SKBS1C	Sessile	Recovered	
SKOT1A500	Kotok	SKOT1A	500	Recovered	
SKOT1BDMSO	Kotok	SKOT1B	Sessile	Recovered	
SKOT1ADMSO	Kotok	SKOT1A	Sessile	Recovered	
SKOT1CDMSO	Kotok	SKOT1C	Sessile	Recovered	
SKOT1A100	Kotok	SKOT1A	100	Recovered	
SKOT1B500	Kotok	SKOT1B	500	Recovered	
SKOT1C500	Kotok	SKOT1C	500	Recovered	
SKOT1C100	Kotok	SKOT1C	100	Recovered	
SKOT1B100	Kotok	SKOT1B	100	Recovered	Low Reads
SLCB1ADMSO	Lancang Besar	SLCB1A	Sessile	Recovered	
SLCB1A500	Lancang Besar	SLCB1A	500	Recovered	
SLCB1A100	Lancang Besar	SLCB1A	100	Recovered	
SLCB1BDMSO	Lancang Besar	SLCB1B	Sessile	Lost	
SLCB1B500	Lancang Besar	SLCB1B	500	Lost	

SLCB1CDMSO	Lancang Besar	SLCB1C	Sessile	Lost	
SLCB1C500	Lancang Besar	SLCB1C	500	Lost	
SLCB1C100	Lancang Besar	SLCB1C	100	Lost	
SPAR1CDMSO	Pari	SPAR1C	Sessile	Lost	
SPAR1C500	Pari	SPAR1C	500	Lost	
SPAR1C100	Pari	SPAR1C	100	Lost	
SPAR1ADMSO	Pari	SPAR1A	Sessile	Recovered	
SPAR1BDMSO	Pari	SPAR1B	Sessile	Recovered	
SPAR1A100	Pari	SPAR1A	100	Recovered	
SPAR1B100	Pari	SPAR1B	100	Recovered	Low Reads
SPAR1B500	Pari	SPAR1B	500	Recovered	
SPAR1A500	Pari	SPAR1A	500	Recovered	
STDN1C100	Pramuka	STDN1C	100	Recovered	
SPRM1CDMSO	Pramuka	SPRM1C	Sessile	Recovered	
SPRM1C500	Pramuka	SPRM1C	500	Recovered	
SPRM1ADMSO	Pramuka	SPRM1A	Sessile	Recovered	
SPRM1BDMSO	Pramuka	SPRM1B	Sessile	Recovered	
SPRM1B100	Pramuka	SPRM1B	100	Recovered	
SPRM1A500	Pramuka	SPRM1A	500	Recovered	
SPRM1B500	Pramuka	SPRM1B	500	Recovered	
SPRM1A100	Pramuka	SPRM1A	100	Recovered	
SSEP1C100	Sepa	SSEP1C	100	Recovered	
SSEP1CDMSO	Sepa	SSEP1C	Sessile	Recovered	
SSEP1ADMSO	Sepa	SSEP1A	Sessile	Recovered	
SSEP1BDMSO	Sepa	SSEP1B	Sessile	Recovered	
SSEP1C500	Sepa	SSEP1C	500	Recovered	
SSEP1A500	Sepa	SSEP1A	500	Recovered	Low Reads
SSEP1B100	Sepa	SSEP1B	100	Recovered	
SSEP1A100	Sepa	SSEP1A	100	Recovered	
SSEP1B500	Sepa	SSEP1B	500	Recovered	
STDN1B100	Tidung	STDN1B	100	Recovered	
STDN1C500	Tidung	STDN1C	500	Recovered	
STDN1B500	Tidung	STDN1B	500	Recovered	
STDN1CDMSO	Tidung	STDN1C	Sessile	Recovered	
SPRM1C100	Tidung	SPRM1C	100	Recovered	
STDN1A100	Tidung	STDN1A	100	Recovered	
STDN1ADMSO	Tidung	STDN1A	Sessile	Recovered	
STDN1BDMSO	Tidung	STDN1B	Sessile	Recovered	
STDN1A500	Tidung	STDN1A	500	Recovered	

SLCB1B100	Lancang Besar	SLCB1B	100	Lost	

Supplemental Table S2-2. Deployment location and samples used for 18S-ARMS analysis.

SampleID	Site_Name	ARMS	Fraction_Size
SKOT1BDMSO	Kotok	1B	Sessile
SLCB1ADMSO	LancangBesar	1A	Sessile
SKBS1A100	KarangBeras	1A	100
SKBS1A500	KarangBeras	1A	500
SKBS1B100	KarangBeras	1B	100
SKBS1B500	KarangBeras	1B	500
SKBS1BDMSO	KarangBeras	1B	Sessile
SKBS1C100	KarangBeras	1C	100
SKBS1C500	KarangBeras	1C	500
SKBS1CDMSO	KarangBeras	1C	Sessile
SKOT1A100	Kotok	1A	100
SKOT1A500	Kotok	1A	500
SKOT1ADMSO	Kotok	1A	Sessile
SKOT1B100	Kotok	1B	100
SKOT1B500	Kotok	1B	500
SKOT1C100	Kotok	1C	100
SKOT1C500	Kotok	1C	500
SKOT1CDMSO	Kotok	1C	Sessile
SLCB1A100	LancangBesar	1A	100
SLCB1A500	LancangBesar	1A	500
SPAR1A100	Pari	1A	100
SPAR1A500	Pari	1A	500
SPAR1ADMSO	Pari	1A	Sessile
SPAR1B100	Pari	1B	100
SPAR1B500	Pari	1B	500
SPAR1BDMSO	Pari	1B	Sessile
SPRM1A100	Pramuka	1A	100
SPRM1A500	Pramuka	1A	500
SPRM1ADMSO_202	Pramuka	1A	Sessile
SPRM1B100	Pramuka	1B	100

SPRM1B500	Pramuka	1B	500
SPRM1BDMSO	Pramuka	1B	Sessile
SPRM1C100	Pramuka	1C	100
SPRM1C500	Pramuka	1C	500
SPRM1CDMSO	Pramuka	1C	Sessile
SSEP1A100	Sepa	1A	100
SSEP1A500	Sepa	1A	500
SSEP1ADMSO	Sepa	1A	Sessile
SSEP1B100	Sepa	1B	100
SSEP1B500	Sepa	1B	500
SSEP1BDMSO	Sepa	1B	Sessile
SSEP1C100	Sepa	1C	100
SSEP1C500	Sepa	1C	500
SSEP1CDMSO	Sepa	1C	Sessile
STDN1A100	Tidung	1A	100
STDN1A500	Tidung	1A	500
STDN1ADMSO	Tidung	1A	Sessile
STDN1B100	Tidung	1B	100
STDN1B500	Tidung	1B	500
STDN1BDMSO	Tidung	1B	Sessile
STDN1C100	Tidung	1C	100
STDN1C500	Tidung	1C	500
STDN1CDMSO	Tidung	1C	Sessile

Supplemental Table S2-3. Deployment location and samples used for down streaming analysis in CO1-eDNA samples.

Sample Name	Sampling Location
CO1_BDR.1.S1.L001	Bidadari
CO1_BDR.2.S2.L001	Bidadari
CO1_BDR.3.S3.L001	Bidadari
CO1_KBS.1.S4.L001	KarangBeras
CO1_KBS.2.S5.L001	KarangBeras
CO1_KBS.3.S6.L001	KarangBeras
CO1_KOT.1.S7.L001	Kotok
CO1_KOT.2.S8.L001	Kotok

CO1_KOT.3.S9.L001	Kotok
CO1_LCB.1.S10.L001	LancangBesar
CO1_LCB.2.S11.L001	LancangBesar
CO1_LCB.3.S12.L001	LancangBesar
CO1_PAR.1.S13.L001	Pari
CO1_PAR.2.S14.L001	Pari
CO1_PAR.3.S15.L001	Pari
CO1_PRM.1.S16.L001	Pramuka
CO1_PRM.2.S17.L001	Pramuka
CO1_PRM.3.S18.L001	Pramuka
CO1_SEP.1.S19.L001	Sepa
CO1_SEP.2.S20.L001	Sepa
CO1_SEP.3.S21.L001	Sepa
CO1_TDN.1.S22.L001	Tidung
CO1_TDN.2.S23.L001	Tidung
CO1_TDN.3.S24.L001	Tidung

Supplemental Table S2-4. Results of SIMPER analysis for studies using CO1-ARMS data

No	OTU	Family	Genus	Species
1	OTU4574	Peyssonneliaceae	Ramicrusta	Ramicrusta appressa
2	OTU2475	Microcionidae	Clathria	Clathria abietina
3	OTU365	Unidentified	Unidentified	Unidentified
4	OTU2364	Unidentified	Unidentified	Unidentified
5	OTU677	Hesionidae	Gyptis	Gyptis sp. 1 AN-2008
6	OTU2998	Rhodomelaceae	Palisada	Palisada sp.
7	OTU1186	Terebellidae	Polycirrus	Unidentified
8	OTU5851	Nephtheidae	Stereonephthya	Stereonephthya sp. A CSM-2013
9	OTU2458	Unidentified	Unidentified	Unidentified
10	OTU5249	Peyssonneliaceae	Ramicrusta	Ramicrusta bonairenesis
11	OTU1238	Coelosphaeridae	Lissodendoryx	Lissodendoryx flabellata
12	OTU5320	Stenothoidae	Unidentified	Stenothoidae sp. 31
13	OTU5975	Unidentified	Unidentified	Unidentified
14	OTU1163	Scaphandridae	Scaphander	Scaphander lignarius
15	OTU2801	Unidentified	Unidentified	Unidentified

Supplemental Table S2-5. Results of SIMPER analysis for studies using 18S-ARMS data.

No	OTU	Family	Genus	Species
1	ASV10986	Maxillopoda	Maxillopoda_X	Maxillopoda_X_sp.
2	ASV25194	Malacostraca	Ogyrides	Ogyrides_sp.
3	ASV1093	Gigartinales_X	Ramicrusta	Ramicrusta_sp.
4	ASV1162	Unidentified	Unidentified	Unidentified
5	ASV10738	Demospongiae	Clathria	Clathria_venosa
6	ASV16946	Corallinales_X	Corallinales_XX	Corallinales_XX_sp.
7	ASV11867	Demospongiae	Tedania	Tedania_strongylostyla
8	ASV18755	Demospongiae	Antho	Antho_prima
9	ASV18399	Ostracoda	Xestoleberis	Xestoleberis_hanaii
10	ASV12978	Annelida_XX	Terebellidae	Lysilla_sp.
11	ASV19220	Polyplacophora_X	Plaxiphora	Plaxiphora_albida
12	ASV17383	Ascidiacea	Botryllus	Botryllus_schlosseri
13	ASV19260	Echinodermata_XX	Echinodermata_XXX	Echinodermata_XXX_sp.
14	ASV9361	Annelida_XX	Prosphaerosyllis	Prosphaerosyllis_longipapillata
15	ASV10706	Malacostraca	Eriphia	Eriphia_scabricula
16	ASV2321	Limoida	Limaria	Limaria_hians
17	ASV1954	Anthozoa	Porites	Porites_cylindrica
18	ASV14015	Maxillopoda	Lucicutia	Lucicutia_ovaliformis
19	ASV10746	Annelida_XX	Cirriformia	Cirriformia_tentaculata
20	ASV13390	Annelida_XX	Exogone	Exogone_fustifera
21	ASV17928	Insecta	Strepsylla	Strepsylla_villai
22	ASV13531	Malacostraca	Coralliocaris	Coralliocaris_superba
23	ASV13254	Mytiloida	Musculus	Musculus_lateralis
24	ASV12962	Demospongiae	Demospongiae_X	Demospongiae_X_sp.
25	ASV11401	Anthozoa	Pocillopora	Pocillopora_meandrina
26	ASV13186	Annelida_XX	Syllis	Syllis_ferrani
27	ASV21998	Malacostraca	Pagurus	Pagurus_longicarpus
28	ASV9884	Annelida_XX	Armandia	Armandia_sp.
29	ASV9486	Polyplacophora_X	Acanthopleura	Acanthopleura_granulata
30	ASV16970	Anthozoa	Pavona	Pavona_varians
31	ASV2195	Annelida_XX	Notomastus	Notomastus_latericeus
32	ASV10944	Annelida_XX	Cirriformia	Cirriformia_tentaculata
33	ASV22637	Ascidiacea	Polycarpa	Polycarpa_mytiligera

34	ASV14326	Annelida_XX	Thelepus	Thelepus_crispus
35	ASV12686	Gigartinales_X	Ramicrusta	Ramicrusta_sp.
36	ASV15699	Suessiaceae	Symbiodinium	Symbiodinium_sp._Clade_D
37	ASV11565	Rhodomelaceae	Acanthophora	Acanthophora_spicifera
38	ASV1361	Ostracoda	Vargula	Vargula_hilgendorfii
39	ASV21643	Echinodermata_XX	Ophioderma	Ophioderma_cenereum
40	ASV8373	Pterioida-Pinnidae	Pinna	Pinna_nobilis
41	ASV20715	Chromadorea_X	Tarvaia	Tarvaia_sp.
42	ASV6969	Maxillopoda	Typhlamphiascus	Typhlamphiascus_typhlops
43	ASV16520	Pirsonia_Clade_XX	Pirsonia	Pirsonia_guinardiae
44	ASV10183	Demospongiae	Neofibularia	Neofibularia_hartmani
45	ASV11412	Annelida_XX	Lysidice	Lysidice_sp.
46	ASV12282	Ascidiacea	Perophora	Perophora_japonica
47	ASV15434	Annelida_XX	Exogone	Exogone_heterosetosa

Supplemental Table S2-6. Results of SIMPER analysis for studies using CO1-eDNA data.

No	OTU	Family	Genus	Species
1	Otu1329	Mamiellaceae	Micromonas	Micromonas pusilla
2	Otu0407	Unidentified	Unidentified	Unidentified
3	Otu1405	Unidentified	Unidentified	Unidentified
4	Otu0249	Unidentified	Unidentified	Unidentified
5	Otu5264	Unidentified	Unidentified	Unidentified
6	Otu5088	Nectriaceae	Ilyonectria	Ilyonectria destructans
7	Otu0375	Calanidae	Undinula	Undinula vulgaris
8	Otu5248	Unidentified	Unidentified	Unidentified
9	Otu0822	Unidentified	Unidentified	Unidentified
10	Otu5223	Unidentified	Unidentified	Unidentified
11	Otu5273	Unidentified	Cymbomonas	Cymbomonas tetramitiformis
12	Otu0100	Chaetocerotaceae	Chaetoceros	Chaetoceros socialis
13	Otu5207	Paracalanidae	Parvocalanus	Parvocalanus crassirostris
14	Otu1910	Unidentified	Unidentified	Unidentified
15	Otu0209	Chloropicaceae	Chloropicon	Chloropicon primus
16	Otu5093	Unidentified	Unidentified	Unidentified
17	Otu0404	Unidentified	Unidentified	Unidentified
18	Otu1433	Unidentified	Unidentified	Unidentified
19	Otu5190	Unidentified	Unidentified	Unidentified

20	Otu0743	Unidentified	Unidentified	Unidentified
21	Otu5157	Aequoreidae	Aequorea	Aequorea sp. USHIKKK
22	Otu0471	Unidentified	Haptoglossa	Haptoglossa sp. DM2 4/12
23	Otu0656	Halictidae	Lasioglossum	Lasioglossum (Dialictus) sp. 1 PLG-2016
24	Otu0616	Unidentified	Unidentified	Amphipoda sp. LPdivOTU79
25	Otu5169	Unidentified	Unidentified	Unidentified
26	Otu1240	Dinobryaceae	Dinobryon	Dinobryon sp.
27	Otu0566	Acartiidae	Acartia	Acartia spinicauda
28	Otu0962	Unidentified	Unidentified	Unidentified
29	Otu0453	Campanulariidae	Clytia	Clytia folleata
30	Otu1794	Zerconidae	Unidentified	Zerconidae sp. BIOUG26189-D02

References

Adam TC, Holbrook SJ, Burkepile DE, Speare KE, Brooks AJ, Ladd MC, Shantz AA, Vega Thurber R, Schmitt RJ (2022) Priority effects in coral–macroalgae interactions can drive alternate community paths in the absence of top-down control. Ecology:e3831

Andersen JC, Oboyski P, Davies N, Charlat S, Ewing C, Meyer C, Krehenwinkel H, Lim JY, Noriyuki S, Ramage T, Gillespie RG, Roderick GK (2019) Categorization of species as native or nonnative using DNA sequence signatures without a complete reference library. Ecological Applications 29:e01914

Andújar C, Arribas P, Yu DW, Vogler AP, Emerson BC (2018) Why the COI barcode should be the community DNA metabarcode for the metazoa. Molecular Ecology

Anthony KRN, Marshall PA, Abdulla A, Beeden R, Bergh C, Black R, Eakin CM, Game ET, Gooch M, Graham NAJ, Green A, Heron SF, Hooidonk R van, Knowland C, Mangubhai S, Marshall N, Maynard JA, Mcginnity P, Mcleod E, Mumby PJ, Nystr??m M, Obura D, Oliver J, Possingham HP, Pressey RL, Rowlands GP, Tamelander J, Wachenfeld D, Wear S (2015) Operationalizing resilience for adaptive coral reef management under global environmental change. Global Change Biology 21:48–61

Appeltans W, Ahyong ST, Anderson G, Angel M V., Artois T, Bailly N, Bamber R, Barber A, Bartsch I, Berta A, Błażewicz-Paszkowycz M, Bock P, Boxshall G, Boyko CB, Brandão SN, Bray RA, Bruce NL, Cairns SD, Chan TY, Cheng L, Collins AG, Cribb T, Curini-Galletti M, Dahdouh-Guebas F, Davie PJF, Dawson MN, Clerck O De, Decock W, Grave S De, Voogd NJ de, Domning DP, Emig CC, Erséus C, Eschmeyer W, Fauchald K, Fautin DG, Feist SW, Fransen CHJM, Furuya H, Garcia-Alvarez O, Gerken S, Gibson D, Gittenberger A, Gofas S, Gómez-Daglio L, Gordon DP, Guiry MD, Hernandez F,

Hoeksema BW, Hopcroft RR, Jaume D, Kirk P, Koedam N, Koenemann S, Kolb JB, Kristensen RM, Kroh A, Lambert G, Lazarus DB, Lemaitre R, Longshaw M, Lowry J, Macpherson E, Madin LP, Mah C, Mapstone G, McLaughlin PA, Mees J, Meland K, Messing CG, Mills CE, Molodtsova TN, Mooi R, Neuhaus B, Ng PKL, Nielsen C, Norenburg J, Opresko DM, Osawa M, Paulay G, Perrin W, Pilger JF, Poore GCB, Pugh P, Read GB, Reimer JD, Rius M, Rocha RM, Saiz-Salinas JI, Scarabino V, Schierwater B, Schmidt-Rhaesa A, Schnabel KE, Schotte M, Schuchert P, Schwabe E, Segers H, Self-Sullivan C, Shenkar N, Siegel V, Sterrer W, Stöhr S, Swalla B, Tasker ML, Thuesen E V., Timm T, Todaro MA, Turon X, Tyler S, Uetz P, Land J van der, Vanhoorne B, Ofwegen LP Van, Soest RWM Van, Vanaverbeke J, Walker-Smith G, Walter TC, Warren A, Williams GC, Wilson SP, Costello MJ, Błazewicz-Paszkowycz M, Bock P, Boxshall G, Boyko CB, Brandão SN, Bray RA, Bruce NL, Cairns SD, Chan TY, Cheng L, Collins AG, Cribb T, Curini-Galletti M, Dahdouh-Guebas F, Davie PJF, Dawson MN, Clerck O De, Decock W, Grave S De, Voogd NJ de, Domning DP, Emig CC, Erséus C, Eschmeyer W, Fauchald K, Fautin DG, Feist SW, Fransen CHJM, Furuya H, Garcia-Alvarez O, Gerken S, Gibson D, Gittenberger A, Gofas S, Gómez-Daglio L, Gordon DP, Guiry MD, Hernandez F,
Hoeksema BW, Hopcroft RR, Jaume D, Kirk P, Koedam N, Koenemann S, Kolb JB, Kristensen RM, Kroh A, Lambert G, Lazarus DB, Lemaitre R, Longshaw M, Lowry J, Macpherson E, Madin LP, Mah C, Mapstone G, McLaughlin PA, Mees J, Meland K, Messing CG, Mills CE, Molodtsova TN, Mooi R, Neuhaus B, Ng PKL, Nielsen C, Norenburg J, Opresko DM, Osawa M, Paulay G, Perrin W, Pilger JF, Poore GCB, Pugh P, Read GB, Reimer JD, Rius M, Rocha RM, Saiz-Salinas JI, Scarabino V, Schierwater B, Schmidt-Rhaesa A, Schnabel KE, Schotte M, Schuchert P, Schwabe E, Segers H, Self-

Sullivan C, Shenkar N, Siegel V, Sterrer W, Stöhr S, Swalla B, Tasker ML, Thuesen E V., Timm T, Todaro MA, Turon X, Tyler S, Uetz P, Land J van der, Vanhoorne B, Ofwegen LP Van, Soest RWM Van, Vanaverbeke J, Walker-Smith G, Walter TC, Warren A, Williams GC, Wilson SP, Costello MJ (2012) The Magnitude of Global Marine Species Diversity. Current Biology 22:2189–2202

Armstrong McKay DI, Cornell SE, Richardson K, Rockström J (2021) Resolving ecological feedbacks on the ocean carbon sink in Earth system models. Earth System Dynamics 12:797–818

Bachoon DS, Markand S, Otero E, Perry G, Ramsubaugh A (2010) Assessment of non-point sources of fecal pollution in coastal waters of Puerto Rico and Trinidad. Marine Pollution Bulletin

Baird DJ, Hajibabaei M (2012) Biomonitoring 2.0: A new paradigm in ecosystem assessment made possible by next-generation DNA sequencing. Molecular Ecology 21:2039–2044

Bakker J, Wangensteen OS, Chapman DD, Boussarie G, Buddo D, Guttridge TL, Hertler H, Mouillot D, Vigliola L, Mariani S (2017) Environmental DNA reveals tropical shark diversity in contrasting levels of anthropogenic impact. Scientific Reports 7:1–11

Barber PH (2009) The challenge of understanding the Coral Triangle biodiversity hotspot. Journal of Biogeography 36:1845–1846

Baselga A (2010) Partitioning the turnover and nestedness components of beta diversity. Global Ecology and Biogeography 19:134–143

Bates D, Mächler M, Bolker B, Walker S (2015) Fitting Linear Mixed-Effects Models Using {lme4}. Journal of Statistical Software 67:1–48

Baum G, Januar HI, Ferse SCA, Kunzmann A (2015) Local and Regional Impacts of Pollution

on Coral Reefs along the Thousand Islands North of the Megacity Jakarta, Indonesia. PloS one 10:e0138271

Bayoumy EM, Osman HAM, El-Bana LF, Hassanain MA (2008) Monogenean parasites as bioindicatores for heavy metals status in some egyptian Red Sea fishes. Global Veterinaria 2:117–122

Beaumont NJ, Austen MC, Atkins JP, Burdon D, Degraer S, Dentinho TP, Derous S, Holm P, Horton T, Ierland E van, Marboe AH, Starkey DJ, Townsend M, Zarzycki T (2007) Identification, definition and quantification of goods and services provided by marine biodiversity: Implications for the ecosystem approach. Marine Pollution Bulletin 54:253–265

Beaumont NJ, Austen MC, Mangi SC, Townsend M (2008) Economic valuation for the conservation of marine biodiversity. Marine Pollution Bulletin 56:386–396

Béguinot J (2018) Analysing the Role of Environmental Stresses on Species Richness and the Process of Hierarchical Structuring of Species Abundances in Marine Gastropods communities at Suva (Fiji Islands). International Journal of Environment and Climate Change

Bellwood DR, Hughes TP, Folke C, Nyström M (2004) Confronting the coral reef crisis. Nature 429:827–833

Benedetti-Cecchi L (2000) Priority effects, taxonomic resolution, and the prediction of variable patterns of colonisation of algae in littoral rock pools. Oecologia 2000 123:2 123:265–274

Beng KC, Tomlinson KW, Shen XH, Surget-Groba Y, Hughes AC, Corlett RT, Slik JWF (2016) The utility of DNA metabarcoding for studying the response of arthropod diversity and composition to land-use change in the tropics. Scientific Reports

Bernard ATF, Götz A, Kerwath SE, Wilke CG (2013) Observer bias and detection probability in underwater visual census of fish assemblages measured with independent double-observers. Journal of Experimental Marine Biology and Ecology 443:75–84

Bickford D, Lohman DJ, Sodhi NS, Ng PKL, Meier R, Winker K, Ingram KK, Das I (2007) Cryptic species as a window on diversity and conservation. Trends in Ecology & Evolution 22:148–155

Bik HM (2021) Just keep it simple? Benchmarking the accuracy of taxonomy assignment software in metabarcoding studies. Molecular Ecology Resources 21:2187–2189

Boeuf G (2011) Marine biodiversity characteristics. Comptes Rendus Biologies 334:435–440

Bolyen E, Rideout JR, Dillon MR, Bokulich NA, Abnet CC, Al-Ghalith GA, Alexander H, Alm EJ, Arumugam M, Asnicar F, Bai Y, Bisanz JE, Bittinger K, Brejnrod A, Brislawn CJ, Brown CT, Callahan BJ, Caraballo-Rodríguez AM, Chase J, Cope EK, Silva R Da, Diener C, Dorrestein PC, Douglas GM, Durall DM, Duvallet C, Edwardson CF, Ernst M, Estaki M, Fouquier J, Gauglitz JM, Gibbons SM, Gibson DL, Gonzalez A, Gorlick K, Guo J, Hillmann B, Holmes S, Holste H, Huttenhower C, Huttley GA, Janssen S, Jarmusch AK, Jiang L, Kaehler BD, Kang K Bin, Keefe CR, Keim P, Kelley ST, Knights D, Koester I, Kosciolek T, Kreps J, Langille MGI, Lee J, Ley R, Liu YX, Loftfield E, Lozupone C, Maher M, Marotz C, Martin BD, McDonald D, McIver LJ, Melnik A V., Metcalf JL, Morgan SC, Morton JT, Naimey AT, Navas-Molina JA, Nothias LF, Orchanian SB, Pearson T, Peoples SL, Petras D, Preuss ML, Pruesse E, Rasmussen LB, Rivers A, Robeson MS, Rosenthal P, Segata N, Shaffer M, Shiffer A, Sinha R, Song SJ, Spear JR, Swafford AD, Thompson LR, Torres PJ, Trinh P, Tripathi A, Turnbaugh PJ, Ul-Hasan S, Hooft JJJ van der, Vargas F, Vázquez-Baeza Y, Vogtmann E, Hippel M von, Walters W, Wan Y,

Wang M, Warren J, Weber KC, Williamson CHD, Willis AD, Xu ZZ, Zaneveld JR, Zhang Y, Zhu Q, Knight R, Caporaso JG (2019) Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. Nature Biotechnology

Bouchet P, Lozouet P, Maestrati P, Heros V (2002) Assessing the magnitude of species richness in tropical marine environments: Exceptionally high numbers of molluscs at a New Caledonia site. Biological Journal of the Linnean Society 75:421–436

Brown CJ, Jupiter SD, Lin HY, Albert S, Klein C, Maina JM, Tulloch VJD, Wenger AS, Mumby PJ (2017) Habitat change mediates the response of coral reef fish populations to terrestrial run-off. Marine Ecology Progress Series 576:55–68

Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJA, Holmes SP (2016) DADA2: High-resolution sample inference from Illumina amplicon data. Nature Methods

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: Architecture and applications. BMC Bioinformatics 10:421

Campbell SJ, Kartawijaya T, Sabarini EK (2011) Connectivity in reef fish assemblages between seagrass and coral reef habitats. Aquatic Biology 13:65–77

Carvalho JC, Cardoso P, Gomes P (2012) Determining the relative roles of species replacement and species richness differences in generating beta-diversity patterns. Global Ecology and Biogeography 21:760–771

Cebrian E, Uriz MJ, Turon X (2007) Sponges as biomonitors of heavy metals in spatial and temporal surveys in northwestern Mediterranean: Multispecies comparison. Environmental Toxicology and Chemistry

Cesar H, Burke L, Pet-Soede L (2003) The economics of worldwide coral reef degradation.

Clarke LJ, Beard JM, Swadling KM, Deagle BE (2017) Effect of marker choice and thermal

cycling protocol on zooplankton DNA metabarcoding studies. Ecology and Evolution 7:873–883

Cleary DFR (2017) Linking fish species traits to environmental conditions in the Jakarta Bay-Pulau Seribu coral reef system. Marine Pollution Bulletin 122:259–262

Cleary DFR, Becking LE, Voogd NJ de, Renema W, Beer M de, Soest RWM van, Hoeksema BW (2005) Variation in the diversity and composition of benthic taxa as a function of distance offshore, depth and exposure in the Spermonde Archipelago, Indonesia. Estuarine, Coastal and Shelf Science 65:557–570

Cleary DFR, Polónia ARM, Renema W, Hoeksema BW, Rachello-Dolmen PG, Moolenbeek RG, Budiyanto A, Yahmantoro, Tuti Y, Giyanto, Draisma SGA, Prud'homme van Reine WF, Hariyanto R, Gittenberger A, Rikoh MS, Voogd NJ de (2016) Variation in the composition of corals, fishes, sponges, echinoderms, ascidians, molluscs, foraminifera and macroalgae across a pronounced in-to-offshore environmental gradient in the Jakarta Bay–Thousand Islands coral reef complex. Marine Pollution Bulletin

Cleary DFR, Suharsono, Hoeksema BW (2006) Coral diversity across a disturbance gradient in the Pulau Seribu reef complex off Jakarta, Indonesia. Biodiversity and Conservation 15:3653–3674

Cleary DFR, Vantier L De, Giyanto, Vail L, Manto P, Voogd NJ de, Rachello-Dolmen PG, Tuti Y, Budiyanto A, Wolstenholme J, Hoeksema BW, Suharsono (2008) Relating variation in species composition to environmental variables: a multi-taxon study in an Indonesian coral reef complex. Aquatic Sciences 70:419–431

Coleman FC, Williams SL (2002) Overexploiting marine ecosystem engineers: potential consequences for biodiversity. Trends in Ecology & Evolution 17:40–44

Coll M, Libralato S, Tudela S, Palomera I, Pranovi F (2008) Ecosystem Overfishing in the Ocean. PLOS ONE 3:e3881

Cordier T, Frontalini F, Cermakova K, Apothéloz-Perret-Gentil L, Treglia M, Scantamburlo E, Bonamin V, Pawlowski J (2019) Multi-marker eDNA metabarcoding survey to assess the environmental impact of three offshore gas platforms in the North Adriatic Sea (Italy). Marine Environmental Research 146:24–34

Costanza R, Mageau M (1999) What is a healthy ecosystem? Aquatic Ecology 33:105–115

Costello MJ, Basher Z, McLeod L, Asaad I, Claus S, Vandepitte L, Yasuhara M, Gislason H, Edwards M, Appeltans W, Enevoldsen H, Edgar GJ, Miloslavich P, Monte S De, Pinto IS, Obura D, Bates AE (2017) Methods for the Study of Marine Biodiversity. In: The GEO Handbook on Biodiversity Observation Networks. Springer International Publishing, p 129–163

Curd EE, Gold Z, Kandlikar GS, Gomer J, Ogden M, O'Connell T, Pipes L, Schweizer TM, Rabichow L, Lin M, Shi B, Barber PH, Kraft N, Wayne R, Meyer RS (2019) Anacapa Toolkit: An environmental DNA toolkit for processing multilocus metabarcode datasets. Methods in Ecology and Evolution

Dauer DM, Simon JL (1976) Habitat expansion among polychaetous annelids repopulating a defaunated marine habitat. Marine Biology 1976 37:2 37:169–177

David R, Uyarra MC, Carvalho S, Anlauf H, Borja A, Cahill AE, Carugati L, Danovaro R, Jode A De, Feral J-P, Guillemain D, Martire M Lo, D'Avray LTDV, Pearman JK, Chenuil A (2019) Lessons from photo analyses of Autonomous Reef Monitoring Structures as tools to detect (bio-)geographical, spatial, and environmental effects. Marine Pollution Bulletin 141:420–429

Deagle BE, Jarman SN, Coissac E, Pompanon F, Taberlet P (2014) DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. Biology Letters 10

Deiner K, Bik HM, Mächler E, Seymour M, Lacoursière-Roussel A, Altermatt F, Creer S, Bista I, Lodge DM, Vere N de, Pfrender ME, Bernatchez L (2017) Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology

DeVantier L, Suharsono, Budiyanto A, Tuti Y, Imanto P LR (1998) Status of coral communities of Pulau Seribu (Indonesia). In: S S (ed) Proceedings of coral reef evaluation workshop, Pulau Seribu, Jakarta, Indonesia, 1995. UNESCO, Jakarta, Jakarta, p 1–24

DiBattista JD, Reimer JD, Stat M, Masucci GD, Biondi P, Brauwer M De, Wilkinson SP, Chariton AA, Bunce M (2020) Environmental DNA can act as a biodiversity barometer of anthropogenic pressures in coastal ecosystems. Scientific Reports 10:1–15

Dixon J a (1998) Economic valuation of coral reefs. Coral Reefs: Challenges and Opportunities for Sustainable Management:157–179

Drummond AJ, Newcomb RD, Buckley TR, Xie D, Dopheide A, Potter BCM, Heled J, Ross HA, Tooman L, Grosser S, Park D, Demetras NJ, Stevens MI, Russell JC, Anderson SH, Carter A, Nelson N (2015) Evaluating a multigene environmental DNA approach for biodiversity assessment. GigaScience

Duarte CM (2000) Marine biodiversity and ecosystem services: an elusive link. Journal of Experimental Marine Biology and Ecology 250:117–131

Dubinksy Z, Stambler N (1996) Marine pollution and coral reefs. Global Change Biology 2:511–526

Ellis JI, Clark D, Atalah J, Jiang W, Taiapa C, Patterson M, Sinner J, Hewitt J (2017) Multiple

stressor effects on marine infauna: responses of estuarine taxa and functional traits to sedimentation, nutrient and metal loading. Scientific Reports 2017 7:1 7:1–16

Estradivari, Setyawan E, Yusri S (2009) Pengamatan Jangka Panjang Terumbu Karang Kepulauan Seribu (2003 - 2007). Jakarta

Foale S, Adhuri D, Aliño P, Allison EH, Andrew N, Cohen P, Evans L, Fabinyi M, Fidelman P, Gregory C, Stacey N, Tanzer J, Weeratunge N (2013) Food security and the Coral Triangle Initiative. Marine Policy 38:174–183

Fong CR, Bittick SJ, Fong P (2018) Simultaneous synergist, antagonistic and additive interactions between multiple local stressors all degrade algal turf communities on coral reefs. Journal of Ecology 106:1390–1400

Fox J, Weisberg S (2019) An {R} Companion to Applied Regression, Third. Sage, Thousand Oaks {CA}

Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, Hansen AJ (2017) Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nature Communications

Gamfeldt L, Lefcheck JS, Byrnes JEK, Cardinale BJ, Duffy JE, Griffin JN (2015) Marine biodiversity and ecosystem functioning: what's known and what's next? Oikos 124:252–265

Giangrande A, Licciano M, Musco L (2005) Polychaetes as environmental indicators revisited. Marine Pollution Bulletin 50:1153–1162

Gilbert BM, Avenant-Oldewage A (2021) Monogeneans as bioindicators: A meta-analysis of effect size of contaminant exposure toward Monogenea (Platyhelminthes). Ecological Indicators 130:108062

Gissi F, Stauber JL, Binet MT, Golding LA, Adams MS, Schlekat CE, Garman ER, Jolley DF (2016) A review of nickel toxicity to marine and estuarine tropical biota with particular reference to the South East Asian and Melanesian region. Environmental Pollution 218:1308–1323

Gold Z, Sprague J, Kushner DJ, Marin EZ, Barber PH (2021) eDNA metabarcoding as a biomonitoring tool for marine protected areas. PLOS ONE 16:e0238557

Grosberg RK, Vermeij GJ, Wainwright PC (2012) Biodiversity in water and on land. Current Biology 22:R900–R903

Guannel G, Arkema K, Ruggiero P, Verutes G (2016) The Power of Three: Coral Reefs, Seagrasses and Mangroves Protect Coastal Regions and Increase Their Resilience. PLOS ONE 11:e0158094

Guinotte JM, Fabry VJ (2008) Ocean Acidification and Its Potential Effects on Marine Ecosystems. Annals of the New York Academy of Sciences 1134:320–342

Guo J, Yu K, Wang Y, Zhang R, Huang X, Qin Z (2019) Potential impacts of anthropogenic nutrient enrichment on coral reefs in the South China Sea: Evidence from nutrient and chlorophyll a levels in seawater. Environmental Science: Processes and Impacts

Hadi TA (2011) Keragaman Jenis Spons Pada Ekosistem Terumbu Karang Di Gugus Pulau Pari, Kepulauan Seribu. Oseanografi dan Limnologi di Indonesia 37(3): 383:383–396

Hamilton RJ, Almany GR, Brown CJ, Pita J, Peterson NA, Howard Choat J (2017) Logging degrades nursery habitat for an iconic coral reef fish. Biological Conservation 210:273–280

Hoeksema B, Giyanto, Suharsono (2019) The Role of Maximum Shelf Depth versus Distance from Shore in Explaining a Diversity Gradient of Mushroom Corals (Fungiidae) off Jakarta. Diversity 11:46

Hofmann GE, Barry JP, Edmunds PJ, Gates RD, Hutchins DA, Klinger T, Sewell MA (2010) The Effect of Ocean Acidification on Calcifying Organisms in Marine Ecosystems: An Organism-to-Ecosystem Perspective. http://dx.doi.org/101146/annurev.ecolsys110308120227 41:127–147

Howard J, Sutton-Grier A, Herr D, Kleypas J, Landis E, Mcleod E, Pidgeon E, Simpson S (2017) Clarifying the role of coastal and marine systems in climate mitigation. Frontiers in Ecology and the Environment 15:42–50

Hughes TP, Baird AH, Bellwood DR, Card M, Connolly SR, Folke C, Grosberg R, Hoegh-Guldberg O, Jackson JBC, Kleypas J, Lough JM, Marshall P, Nyström M, Palumbi SR, Pandolfi JM, Rosen B, Roughgarden J (2003) Climate change, human impacts, and the resilience of coral reefs. Science (New York, NY) 301:929–33

Ivanina A V., Sokolova IM (2015) Interactive effects of metal pollution and ocean acidification on physiology of marine organisms. Current Zoology 61:653–668

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

Jenkins M (2003) Prospects for biodiversity. Science (New York, NY) 302:1175–7

Jenkins CN, Houtan KS Van (2016) Global and regional priorities for marine biodiversity protection. Biological Conservation 204:333–339

Jessen C, Wild C, Voolstra CR, Wild C (2014) In-Situ Effects of Simulated Overfishing and Eutrophication on Settlement of Benthic Coral Reef Invertebrates in the Central Red Sea. Marine Ecology Progress Series 2:e339

Johnston EL, Roberts DA (2009) Contaminants reduce the richness and evenness of marine communities: a review and meta-analysis. Environmental pollution (Barking, Essex : 1987) 157:1745–1752

Jones-Walters L, Mulder I (2009) Valuing nature: The economics of biodiversity. Journal for Nature Conservation 17:245–247

Joos L, Beirinckx S, Haegeman A, Debode J, Vandecasteele B, Baeyen S, Goormachtig S, Clement L, Tender C De (2020) Daring to be differential: metabarcoding analysis of soil and plant-related microbial communities using amplicon sequence variants and operational taxonomical units. BMC Genomics 21:1–17

Kandlikar G (2020) ranacapa: Utility Functions and "shiny" App for Simple Environmental DNA Visualizations and Analyses.

Kiruba-Sankar R, Chadha N, Dam-Roy S, Sawant P, Saharan N, Krishnan P (2016) Marine sponges as biological indicator of oligotrophic Andaman waters. Indian Journal of Geo-Marine Sciences (IJMS) 45:338–341

Knowlton N (2001) The future of coral reefs. Proceedings of the National Academy of Sciences of the United States of America 98:5419–25

Kopf A, Bicak M, Kottmann R, Schnetzer J, Kostadinov I, Lehmann K, Fernandez-Guerra A, Jeanthon C, Rahav E, Ullrich M, Wichels A, Gerdts G, Polymenakou P, Kotoulas G, Siam R, Abdallah RZ, Sonnenschein EC, Cariou T, O'Gara F, Jackson S, Orlic S, Steinke M, Busch J, Duarte B, Caçador I, Canning-Clode J, Bobrova O, Marteinsson V, Reynisson E, Loureiro CM, Luna GM, Quero GM, Löscher CR, Kremp A, DeLorenzo ME, Øvreås L, Tolman J, LaRoche J, Penna A, Frischer M, Davis T, Katherine B, Meyer CP, Ramos S, Magalhães C, Jude-Lemeilleur F, Aguirre-Macedo ML, Wang S, Poulton N, Jones S, Collin

R, Fuhrman JA, Conan P, Alonso C, Stambler N, Goodwin K, Yakimov MM, Baltar F, Bodrossy L, Kamp J Van De, Frampton DMF, Ostrowski M, Ruth P Van, Malthouse P, Claus S, Deneudt K, Mortelmans J, Pitois S, Wallom D, Salter I, Costa R, Schroeder DC, Kandil MM, Amaral V, Biancalana F, Santana R, Pedrotti ML, Yoshida T, Ogata H, Ingleton T, Munnik K, Rodriguez-Ezpeleta N, Berteaux-Lecellier V, Wecker P, Cancio I, Vaulot D, Bienhold C, Ghazal H, Chaouni B, Essayeh S, Ettamimi S, Zaid EH, Boukhatem N, Bouali A, Chahboune R, Barrijal S, Timinouni M, Otmani F El, Bennani M, Mea M, Todorova N, Karamfilov V, Hoopen P Ten, Cochrane G, L'Haridon S, Can Bizsel K, Vezzi A, Lauro FM, Martin P, Jensen RM, Hinks J, Gebbels S, Rosselli R, Pascale F De, Schiavon R, Santos A Dos, Villar E, Pesant S, Cataletto B, Malfatti F, Edirisinghe R, Herrera Silveira JA, Barbier M, Turk V, Tinta T, Fuller WJ, Salihoglu I, Serakinci N, Ergoren MC, Bresnan E, Iriberri J, Nyhus PAF, Bente E, Karlsen HE, Golyshin PN, Gasol JM, Moncheva S, Dzhembekova N, Johnson Z, Sinigalliano CD, Gidley ML, Zingone A, Danovaro R, Tsiamis G, Clark MS, Costa AC, Bour M El, Martins AM, Eric Collins R, Ducluzeau AL, Martinez J, Costello MJ, Amaral-Zettler LA, Gilbert JA, Davies N, Field D, Glöckner FO (2015) The ocean sampling day consortium. GigaScience

Legendre P (2014) Interpreting the replacement and richness difference components of beta diversity. Global Ecology and Biogeography 23:1324–1334

Leray M, Knowlton N (2015) DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proceedings of the National Academy of Sciences 112:201424997

Leray M, Knowlton N (2016) Censusing marine eukaryotic diversity in the twenty-first century. The Royal Society

Leray M, Meyer CP, Mills SC (2015) Metabarcoding dietary analysis of coral dwelling predatory fish demonstrates the minor contribution of coralmutualists to their highly partitioned, generalist diet. PeerJ 2015

Lindfield SJ, Harvey ES, McIlwain JL, Halford AR (2014) Silent fish surveys: bubble-free diving highlights inaccuracies associated with SCUBA-based surveys in heavily fished areas. Methods in Ecology and Evolution 5:1061–1069

Ling SD, Davey A, Reeves SE, Gaylard S, Davies PL, Stuart-Smith RD, Edgar GJ (2018) Pollution signature for temperate reef biodiversity is short and simple. Marine Pollution Bulletin 130:159–169

LOSEY JE, VAUGHAN M (2009) The Economic Value of Ecological Services Provided by Insects. http://dx.doi.org/101641/0006-3568(2006)56[311:TEVOES]20CO;2

Lüdecke D (2020) sjPlot: Data Visualization for Statistics in Social Science.

Macheriotou L, Guilini K, Bezerra TN, Tytgat B, Nguyen DT, Phuong Nguyen TX, Noppe F, Armenteros M, Boufahja F, Rigaux A, Vanreusel A, Derycke S (2019) Metabarcoding free-living marine nematodes using curated 18S and CO1 reference sequence databases for species-level taxonomic assignments. Ecology and Evolution 9:1211–1226

Madduppa HH, Beginner S, Suparyani E, Siregar AM, Arafat D, TARIGAN SA, ALIMUDDIN A, KHAIRUDI D, RAHMAWATI F, BRAHMANDITO A (2013) Dynamics of fish diversity across an environmental gradient in the Seribu Islands reefs off Jakarta. Biodiversitas Journal of Biological Diversity 14:17–24

Madduppa H, Cahyani NKD, Anggoro AW, Subhan B, Jefri E, Sani LMII, Arafat D, Akbar N, Bengen DG (2021) eDNA metabarcoding illuminates species diversity and composition of three phyla (chordata, mollusca and echinodermata) across Indonesian coral reefs. 30

Manikandan B, Ravindran J, Shrinivaasu S, Marimuthu N, Paramasivam K (2014) Community structure and coral status across reef fishing intensity gradients in Palk Bay reef, southeast coast of India. Environmental Monitoring and Assessment 186:5989–6002

Martin M (2011) Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet.journal

Marwayana ON, Gold Z, Meyer CP, Barber PH (2022) Environmental DNA in a global biodiversity hotspot: Lessons from coral reef fish diversity across the Indonesian archipelago. Environmental DNA 4:222–238

Mcmanus JW, Reyes RB, Nañola CL (1997) Effects of some destructive fishing methods on coral cover and potential rates of recovery. Environmental Management 21:69–78

McMurdie PJ, Holmes S (2013) Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE

Meij SET van der, Moolenbeek RG, Hoeksema BW (2009) Decline of the Jakarta Bay molluscan fauna linked to human impact. Marine pollution bulletin 59:101–7

Meij SET van der, Suharsono, Hoeksema BW (2010) Long-term changes in coral assemblages under natural and anthropogenic stress in Jakarta Bay (1920-2005). Marine Pollution Bulletin 60:1442–1454

Moberg F, Folke C (1999) Ecological goods and services of coral reef ecosystems. Ecological Economics 29:215–233

Mora C (2015) Ecology of Fishes on Coral Reefs. Cambridge University Press

Nagelkerken I, Velde G van der, Gorissen MW, Meijer GJ, Van't Hof T, Hartog C den (2000) Importance of Mangroves, Seagrass Beds and the Shallow Coral Reef as a Nursery for Important Coral Reef Fishes, Using a Visual Census Technique. Estuarine, Coastal and

Shelf Science 51:31–44

Narayan S, Beck MW, Reguero BG, Losada IJ, Wesenbeeck B Van, Pontee N, Sanchirico JN, Ingram JC, Lange GM, Burks-Copes KA (2016) The Effectiveness, Costs and Coastal Protection Benefits of Natural and Nature-Based Defences. PLOS ONE 11:e0154735

Orr JC, Fabry VJ, Aumont O, Bopp L, Doney SC, Feely RA, Gnanadesikan A, Gruber N, Ishida A, Joos F, Key RM, Lindsay K, Maier-Reimer E, Matear R, Monfray P, Mouchet A, Najjar RG, Plattner G-K, Rodgers KB, Sabine CL, Sarmiento JL, Schlitzer R, Slater RD, Totterdell IJ, Weirig M-F, Yamanaka Y, Yool A (2005) Anthropogenic ocean acidification over the twenty-first century and its impact on calcifying organisms. Nature 437:681–6

Pastorok R, Bilyard G (1985) Effects of sewage pollution on coral-reef communities. Marine Ecology Progress Series 21:175–189

Pearman JK, Aylagas E, Voolstra CR, Anlauf H, Villalobos R, Carvalho S (2019) Disentangling the complex microbial community of coral reefs using standardized Autonomous Reef Monitoring Structures (ARMS). Molecular Ecology:mec.15167

Pearman JK, Leray M, Villalobos R, Machida RJ, Berumen ML, Knowlton N, Carvalho S (2018) Cross-shelf investigation of coral reef cryptic benthic organisms reveals diversity patterns of the hidden majority. Scientific Reports 8:8090

Pelletier MC, Gold AJ, Heltshe JF, Buffum HW (2010) A method to identify estuarine macroinvertebrate pollution indicator species in the Virginian Biogeographic Province. Ecological Indicators 10:1037–1048

Pet-Soede L, Cesar HSJ, Pet JS (1999) An economic analysis of blast fishing on Indonesian coral reefs. Environmental Conservation 26:83–93

Pinedo S, Jordana E, Ballesteros E (2015) A critical analysis on the response of

macroinvertebrate communities along disturbance gradients: description of MEDOCC (MEDiterranean OCCidental) index. Marine Ecology 36:141–154

Piñol J, Senar MA, Symondson WOC (2019) The choice of universal primers and the characteristics of the species mixture determine when <scp>DNA</scp> metabarcoding can be quantitative. Molecular Ecology 28:407–419

Piroddi C, Akoglu E, Andonegi E, Bentley JW, Celić I, Coll M, Dimarchopoulou D, Friedland R, Mutsert K de, Girardin R, Garcia-Gorriz E, Grizzetti B, Hernvann PY, Heymans JJ, Müller-Karulis B, Libralato S, Lynam CP, Macias D, Miladinova S, Moullec F, Palialexis A, Parn O, Serpetti N, Solidoro C, Steenbeek J, Stips A, Tomczak MT, Travers-Trolet M, Tsikliras AC (2021) Effects of Nutrient Management Scenarios on Marine Food Webs: A Pan-European Assessment in Support of the Marine Strategy Framework Directive. Frontiers in Marine Science 8:179

Plaisance L, Brainard R, Julian Caley M, Knowlton N (2011) Using DNA barcoding and standardized sampling to compare geographic and habitat differentiation of crustaceans: A Hawaiian Islands example. Diversity

Pocklington P, Wells PG (1992) Polychaetes Key taxa for marine environmental quality monitoring. Marine Pollution Bulletin 24:593–598

Polanco A, Waldock | Conor, Keggin | Thomas, Marques V, Rozanski R, Valentini A, Dejean T, Vermeij M, Albouy C, Loïc Pellissier |, Nature B (2022) Ecological indices from environmental DNA to contrast coastal reefs under different anthropogenic pressures. Ecology and Evolution 12:e9212

Polónia ARM, Cleary DFR, Duarte LN, Voogd NJ de, Gomes NCM (2014) Composition of Archaea in Seawater, Sediment, and Sponges in the Kepulauan Seribu Reef System,

Indonesia. Microbial Ecology 67:553–567

R Core Team (2020) R: A language and environment for statistical computing. Retrieved from.

Rachello-Dolmen PG, Cleary DFR (2007) Relating coral species traits to environmental conditions in the Jakarta Bay/Pulau Seribu reef system, Indonesia. Estuarine, Coastal and Shelf Science

Ransome E, Geller JB, Timmers M, Leray M, Mahardini A, Sembiring A, Collins AG, Meyer CP, Mahardini A, Sembiring A, Collins AG, Meyer CP (2017) The importance of standardization for biodiversity comparisons: A case study using autonomous reef monitoring structures (ARMS) and metabarcoding to measure cryptic diversity on Mo'orea coral reefs, French Polynesia (CA Chen, Ed.). PLoS ONE 12:e0175066

Rees JG, Setiapermana D, Sharp VA, Weeks JM, Williams TM (1999) Evaluation of the impacts of land-based contaminants on the benthic faunas of Jakarta Bay, Indonesia. Oceanologica Acta 22:627–640

Reish DJ (1970) The effects of varying concentrations of nutrients, chlorinity, and dissolved oxygen on polychaetous annelids. Water Research 4:721–735

Reynolds CS, Padisák J, Sommer U (1993) Intermediate disturbance in the ecology of phytoplankton and the maintenance of species diversity: a synthesis. Hydrobiologia 249:183–188

Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: A versatile open source tool for metagenomics. PeerJ

Sale PF, Douglas WA (1981) Precision and accuracy of visual census technique for fish assemblages on coral patch reefs. Environmental Biology of Fishes 6:333–339

Sanford E, Kelly MW (2011) Local adaptation in marine invertebrates. Annual Review of

Marine Science 3:509–535

setkab (2018) Sekretariat Kabinet Republik Indonesia | GDP of Indonesia□s Fisheries Sector Rises Significantly: Minister of Marine and Fisheries.

Singh A, Thakur NL (2017) Influence of spatial competitor on asexual reproduction of the marine sponge Cinachyrella cf. cavernosa (Porifera, Demospongiae). Hydrobiologia 2017 809:1 809:247–263

Singmann H, Bolker B, Westfall J, Aust F, Ben-Shachar MS (2020) afex: Analysis of Factorial Experiments.

Sluys R (1999) Global diversity of land planarians (Platyhelminthes, Tricladida, Terricola): a new indicator-taxon in biodiversity and conservation studies. Biodiversity & Conservation 1999 8:12 8:1663–1681

Smith TB, Nemeth RS, Blondeau J, Calnan JM, Kadison E, Herzlieb S (2008) Assessing coral reef health across onshore to offshore stress gradients in the US Virgin Islands. Marine Pollution Bulletin 56:1983–1991

Spens J, Evans AR, Halfmaerten D, Knudsen SW, Sengupta ME, Mak SST, Sigsgaard EE, Hellström M (2017) Comparison of capture and storage methods for aqueous macrobial eDNA using an optimized extraction protocol: advantage of enclosed filter. Methods in Ecology and Evolution 8:635–645

Sprague J (2020) Personal Observation Joshua Sprague. Personal Observation Joshua Sprague

Stat M, Huggett MJ, Bernasconi R, Dibattista JD, Berry TE, Newman SJ, Harvey ES, Bunce M (2017) Ecosystem biomonitoring with eDNA: metabarcoding across the tree of life in a tropical marine environment. Scientific Reports 2017 7:1 7:1–11

Steele JH, Brink KH, Scott BE (2019) Comparison of marine and terrestrial ecosystems:

suggestions of an evolutionary perspective influenced by environmental variation. ICES Journal of Marine Science 76:50–59

Steinberger AJA (2018) Asteinberger9/seq_scripts: Release v1.

Thomsen PF, Kielgast J, Iversen LL, Møller PR, Rasmussen M, Willerslev E (2012) Detection of a Diverse Marine Fish Fauna Using Environmental DNA from Seawater Samples (S Lin, Ed.). PLoS ONE 7:e41732

Thomsen PF, Willerslev E (2015) Environmental DNA – An emerging tool in conservation for monitoring past and present biodiversity. Biological Conservation 183:4–18

Townsend CR, Scarsbrook MR, Dolédec S (1997) The intermediate disturbance hypothesis, refugia, and biodiversity in streams. Limnology and Oceanography

Trapon ML, Pratchett MS, Hoey AS, Baird AH (2013) Influence of fish grazing and sedimentation on the early post-settlement survival of the tabular coral Acropora cytherea. Coral Reefs 32:1051–1059

Tytgat B, Nguyen DT, Nguyen TXP, Pham TM, Long PK, Vanreusel A, Derycke S (2019) Monitoring of marine nematode communities through 18S rRNA metabarcoding as a sensitive alternative to morphology. Ecological Indicators 107:105554

Valentin RE, Fonseca DM, Gable S, Kyle KE, Hamilton GC, Nielsen AL, Lockwood JL (2020) Moving eDNA surveys onto land: Strategies for active eDNA aggregation to detect invasive forest insects. Molecular Ecology Resources 20:746–755

Valentini A, Taberlet P, Miaud C, Civade R, Herder J, Thomsen PF, Bellemain E, Besnard A, Coissac E, Boyer F, Gaboriaud C, Jean P, Poulet N, Roset N, Copp GH, Geniez P, Pont D, Argillier C, Baudoin JM, Peroux T, Crivelli AJ, Olivier A, Acqueberge M, Brun M Le, Møller PR, Willerslev E, Dejean T (2016) Next-generation monitoring of aquatic

biodiversity using environmental DNA metabarcoding. Molecular Ecology

Waltham NJ, Sheaves M (2018) Eco-engineering rock pools to a seawall in a tropical estuary: Microhabitat features and fine sediment accumulation. Ecological Engineering 120:631–636

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

Widbornl B, Elmgren R (1988) Response of benthic meiofauna to nutrient enrichment of experimental marine ecosystems. Marine Ecology Progress Series 42:257–288

Wilkinson CR (1996) Global change and coral reefs: impacts on reefs, economies and human cultures. Global Change Biology 2:547–558

Williams PH, Gaston KJ (1994) Measuring more of biodiversity: Can higher-taxon richness predict wholesale species richness? Biological Conservation 67:211–217

Willis TJ (2001) Visual census methods underestimate density and diversity of cryptic reef fishes. Journal of Fish Biology 59:1408–1411

World Bank Group (2022) INDONESIA'S SUSTAINABLE OCEANS PROGRAM.

Worm B, Barbier EB, Beaumont N, Duffy JE, Folke C, Halpern BS, Jackson JBC, Lotze HK, Micheli F, Palumbi SR, Sala E, Selkoe KA, Stachowicz JJ, Watson R (2006) Impacts of biodiversity loss on ocean ecosystem services. Science (New York, NY) 314:787–90

Zhang J, Kobert K, Flouri T, Stamatakis A (2014) PEAR: a fast and accurate illumina Paired-End reAd mergeR Genome analysis PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics

CHAPTER 3

Marine Microbial Communities Across a Pollution Gradient in Kepulauan, Seribu, Jakarta Indonesia

Abstract

Autonomous Reef Monitoring Structure (ARMS) are widely used for assessing cryptic marine biodiversity on coral reefs, but few ARMS studies examine microbial diversity. In this study, we use ARMS and DNA metabarcoding to examine shifts in microbial communities across Pulau Seribu, an island chain north of Jakarta, Indonesia that is heavily impacted by pollution stress. Results from 16S rRNA metabarcoding indicate substantial differentiation among microbial communities associated with the 100 μm, 500 μm, and sessile ARMS fractions. Although microbial diversity on ARMS did not varying significantly with pollution intensity as measured by annual chlorophyll-a concentrations, results show a clear separation of bacterial community composition between sites with higher and lower pollution stress. Abundance of Alpha and Gammaproteobacteria varied considerably across sites associated with pollution stress. Sulfate-reducing bacteria Desulfovibrionaceae and Clostridiaceae were taxa most sensitive to pollution stress, potentially serving as indicators for higher pollution levels, while alphaproteobacteria were indicative of less polluted environments. As environmental pressures on global coral reef ecosystems intensify, monitoring these ecosystems becomes increasingly important. Results of this study suggests that microbial diversity associated with ARMS can provide important insights into environmental differences, offering an additional tool to monitor micro to macrobial marine biodiversity in our rapidly changing world.

I. Introduction

Coral reefs are among the most biologically and economically valuable ecosystems on Earth (Pendleton 1995, Cesar et al. 2003, Spalding et al. 2017). Despite occupying highly oligotrophic waters and being only a small part of the world's oceans, coral reefs support a disproportionately high amount of marine biodiversity (Allen & Adrim 2003, Allen 2008a, Allen & Erdmann 2009). In turn, this biodiversity provides critical ecosystem goods and services, such as coastal protection and fisheries, as well as tourism and recreational activities with a combined global value estimated at almost $ 125-140 trillion per year globally (OECD 2019)

Unfortunately, coral reef ecosystems are undergoing rapid degradation, resulting in significant global loss of reef ecosystems. Anthropogenic stressors such as habitat destruction (van der Meij et al. 2009, Baum et al. 2015), overfishing (Jackson et al. 2001, Newton et al. 2007, Warren & Steenbergen 2021), increased sedimentation (Dubinksy & Stambler 1996, Fabricius et al. 2005), and eutrophication (Jessen et al. 2013, Vega Thurber et al. 2014), among others, are compounded by the increasingly severe effects of climate change (e.g., sea surface warming and ocean acidification) (Pandolfi et al. 2003, Hughes et al. 2003). Combined, these stressors threaten the ability of coral reefs to provide the essential ecological and economic services on which our civilization relies (Knowlton 2001).

The vast majority of research focused on human impacts on coral reefs and reef resilience examines macrofauna such corals (Done 1992, DeVantier L, Suharsono, Budiyanto A, Tuti Y, Imanto P 1998, White et al. 2000, Bouchet et al. 2002, Bruno & Selig 2007, Allen & Erdmann 2009) and fishes (Green & Bellwood, Allen 2008b, Hubert et al. 2011, Selig et al. 2014). However, there is a growing body of literature focused on the importance of microorganisms (e.g. bacteria, archaea, and viruses) on coral reefs and their essential role in promoting the health

of coral reef ecosystems (Webster & Reusch 2017, Glasl et al. 2018). Microbes perform a multitude of ecological and physiological roles, often through symbioses, that benefit their hosts ,among others, by facilitating nitrogen fixation in oligotrophic waters (Lesser et al. 2004, Lema et al. 2012), aiding the cycling process of sulfur compounds (González et al. 2003) and providing protection against pathogens (Rohwer et al. 2002, Shnit-Orland & Kushmaro 2009). Microbes can also play a crucial role in supporting the acclimatization and maintenance of homeostasis under changing environmental conditions (Mortzfeld et al. 2016), a function of growing significance given the increasing environmental threats to the coral reefs ecosystem.

Although coral reef management typically focuses on corals or fishes (Carpenter et al. 2008, Pen et al. 2009, Foale et al. 2013, Selig et al. 2014, Jenkins & Van Houtan 2016), microbes can play an important role in coral reef sustainability by facilitating rapid responses to changes in environmental conditions (Bruno et al. 2003, Vega Thurber et al. 2014, Webster et al. 2016). Such responses might include alterations in microbial community composition, taxonomic diversity, changes in relative abundance of taxa, and proliferation of pathogenic and opportunistic bacterial groups (Rohwer et al. 2001, Horner-Devine et al. 2003, Mao-Jones et al. 2010, Barott & Rohwer 2012, Walsh et al. 2015, Hernandez-Agreda et al. 2018). Environmental stressors driving these changes are numerous (reviewed in Nogales et al. 2011, Logue et al. 2015); however, a combination of sedimentation from coastal development, nutrient enrichment from domestic waste, and hydrocarbon discharge from industrial activities are some of the primary sources of pollution on coral reef ecosystems (Cho & Kim 2000, Abed et al. 2002, Jessen et al. 2013, Ziegler et al. 2016).

Field studies on the response of microbial communities to changing environments indicate changes that are either irregular (i.e. haphazard in pattern despite the known source of

pollutants; (Rohwer et al. 2001, Lozupone & Knight 2007, Quero et al. 2015, Hernandez-Agreda et al. 2018)) or that produce gradients in community composition, where microbial communities closer to pollution sources have marked changes in taxonomic compositions, relative to populations further from the pollution source (Fuhrman et al. 2008, Brakstad & Lødeng 2013, Yu et al. 2018). Two primary factors typically influence these patterns; 1) the type of pollution impacting local ecosystems (i.e., whether it is a point source like sewage treatment disposal/outflows or river discharges and runoff from land), and 2) the strength of local oceanographic/hydrological processes influencing the dispersion of the pollutants (Fuhrman et al. 2006, Bachoon et al. 2010, Sinigalliano et al. 2010, Lindström & Langenheder 2012). The type of pollution can influence the bacteria proliferating during and after pollution exposure. For example, non-point or diffuse sources of pollution tend to favor specific fecal indicator bacteria such as *Escherichia coli*, enterococci, and *Clostridium sp.* (Abdelzaher et al. 2010). In contrast, point source pollution tends to heighten specific taxa that are related to the kind of pollution being discharged (e.g., hydrocarbon, phosphorus, and nitrogen) (Bachoon et al. 2010, Mukherjee et al. 2016). Oceanic factors, such as the pattern and strength of ocean currents and/or wind direction, can influence the magnitude of pollution in local areas. In some cases, pollutants can be pushed away from the pollution source, lessening the impact of pollution proximal to the source, while elevating pollution in areas more distant from the source (Zhang et al. 2009).

There is growing interest in characterizing changes in bacterial community composition with respect to pollution stress, including the identification of indicator taxa most sensitive to pollution (Zhang et al., 2014, Quero et al. 2015, Chen et al. 2019). Although molecular techniques have greatly advanced our understanding of how microbial communities respond to pollution (Hugerth & Andersson 2017, Mukherjee et al. 2018), there is a wide range of variation

how these microbial communities are sampled. Sampling approaches vary from collecting water (Shibata et al. 2004, Enns et al. 2012, Ponsero et al. 2021) and sea floor sediment (Gillan et al. 2005, Enns et al. 2012, Yu et al. 2018, Chen et al. 2019) to targeted sampling of microbe-bearing host taxa such as corals and sponges, among others (Harder et al. 2003, Selvin et al. 2009, Barott & Rohwer 2012). In the case of targeted sampling, microbial communities are typically sampled from a particular host body part, and then compared across areas with varying pollution levels (Glasl et al. 2016). Although such studies can document changes in microbial diversity across varied environmental conditions (Blackall et al. 2015, Hester et al. 2016), the non-standardized sampling approach limits our ability to compare results broadly across studies, as bacterial communities can vary greatly across sampled habitats or parts a hosts body (Enomoto et al. 2012, Chiarello et al. 2015, Kramar et al. 2019). Moreover, the physiological and biological characteristics of hosts (e.g., size, age, etc.) in these studies are often unknown, even though hosts of different sizes and age can host different bacterial communities (Krediet et al. 2013, Ainsworth et al. 2015, Weiler et al. 2018). Furthermore, variation in sample processing and preservation can also introduce variation in the results (Bahl et al. 2012, Araújo-Pérez et al. 2012). Advancing our understanding of how pollution stress impacts microbial communities on coral reefs requires sampling methods that are standardized, allowing for direct comparisons of results across studies.

Comprised of 10 25x25cm PVC plates, Autonomous Reef Monitoring Structures (ARMS; Fig. 3-1) are a highly standardized way of sampling and comparing biodiversity in marine ecosystems using DNA metabarcoding (Leray & Knowlton 2016, Ransome et al. 2017, Pearman et al. 2018). ARMS are deployed, processed, and document marine biodiversity using highly standardized protocols, making them ideal for comparative studies (Leray & Knowlton

2015). However, despite being widely used to examine marine metazoans (Leray & Knowlton 2016, Ransome et al. 2017, Casey et al. 2021), only one study has used this method to investigate bacterial diversity (Pearman et al. 2019), documenting how microbial communities changed along an environmental gradient in the Red Sea.

Kepulauan Seribu is an archipelago of 110 islands north of Indonesia's capital, Jakarta (Fig. 3-2), where discharges of pollutants from Jakarta (Willoughby 1986, Rees et al. 1999, Cleary et al. 2006, Rachello-Dolmen & Cleary 2007) combine with domestic waste from local island populations (Uneputty & Evans 1997, Hutomo & Moosa 2005) to create pronounced differences in local pollution stress (Williams et al. 2000). Previous studies show that increased pollution stress in Kepulauan Seribu drives changes in coral and fish communities and lowers alpha diversity (Rees et al. 1999, Cleary et al. 2006, van der Meij et al. 2009, Cleary 2017). However, using ARMS and DNA metabarcoding, Anggoro et al. (chapter 2) showed no changes in alpha diversity of metazoan communities across a 10-fold difference in pollution stress but found significant decrease in beta diversity and significant changes in particular indicator taxa with increasing pollution stress.

In this study, we examine bacterial diversity on ARMS to understand the impact of pollution stress on reef-associated microbial communities. Specifically, we examine bacterial diversity to test how microbial communities shift in response to pollution stress, and whether there are microbial taxa that are indicative of pollution stress that could be used to aid the monitoring of marine ecosystems.

Material and Methods

Study site

Industrial activities and riverine discharge into Jakarta Bay combined with point source pollution from the 11 inhabited islands of Pulau Seribu are major contributors to pollution across the archipelago (Willoughby 1986, Tomascik et al. 1994, Uneputty & Evans 1997). These local stressors are then acted upon by seasonal changes in wind and current directions (Lubis & Yosi 2012), shaping the distribution of marine life in Kepulauan Seribu. For example, coral reefs closer to Jakarta Bay include species that are more tolerant to disturbance, while more sensitive species are only found on the islands furthest from pollution sources (Rees et al. 1999, Cleary et al. 2006, van der Meij et al. 2009). Community composition patterns also shift as a function of pollution levels, with higher fish abundance and richness in less polluted regions, while sea urchin abundance and richness peaks in more polluted regions (Cleary et al. 2008, van der Meij et al. 2009, Madduppa et al. 2013).

ARMS Deployment, Collection, and Sampling

To capture differences in local pollution stress levels, in 2013 we deployed ARMS units in eight islands spanning Pulau Seribu with an distance range between sites of 7-10 km. The deployments encompass highly polluted sites close to Jakarta (e.g., Pulau Bidadari 10 km from Jakarta Bay) to less polluted sites that are more distant (e.g., Pulau Sepa 60 km from Jakarta Bay) (Table. 3-1). Following standard deployment protocols, we deployed three ARMS per location at a depth of 10m, for a total of 24 ARMS. After three years on the sea floor, ARMS were recovered in summer 2016.

Figure. 3-1. ARMS prior to underwater deployment (a.) and fully covered surface following three years deployment in Raja Ampat, Eastern Indonesia (b.)

We recovered ARMS using standard protocols. First, we enveloped ARMS in mesh-lined crates to prevent motile organisms from escaping during recovery. We then brought the ARMS to the surface and transported them in large plastic containers filled with filtered aerated sea water to the Indonesian Institute of Science field lab in Pulau Pari. Next, we disassembled the ARMS,

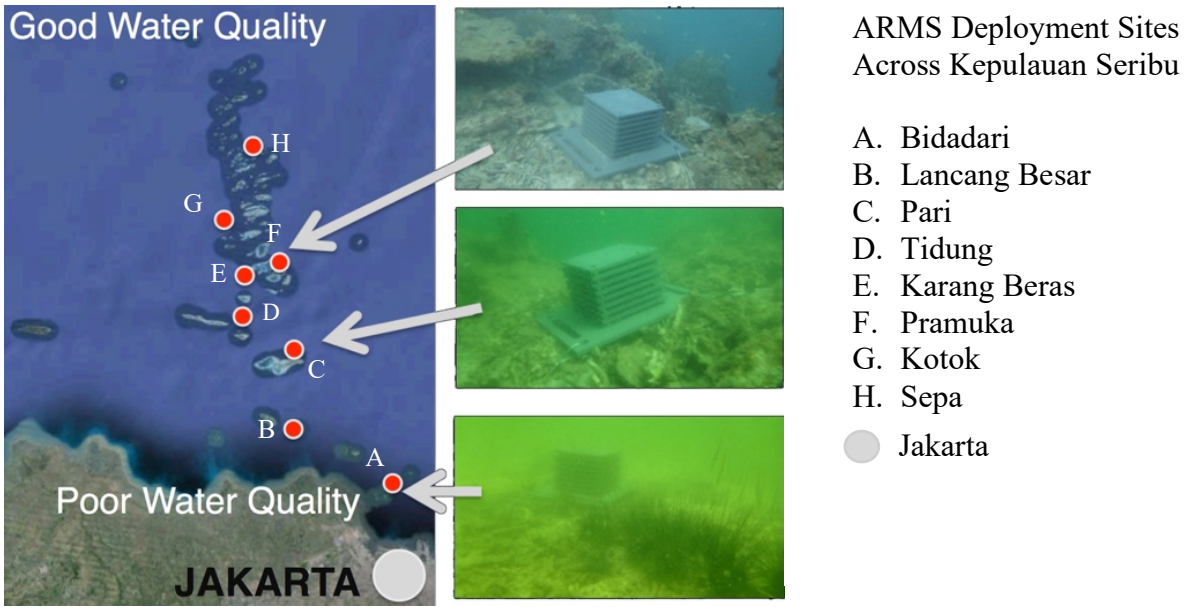


Figure 3-2. Location of ARMS deployment in Pulau Seribu, with representative water conditions

carefully separating each plate and then transferring each plate to individual containers filled with filtered seawater. We then processed the ARMS following a standard protocol (Leray & Knowlton 2015, Ransome et al. 2017), shaking plates vigorously to ensure that all motile organisms detached from the plates. Next, to ensure that higher biomass animals didn't swamp out sequencing signals from smaller organisms, we size fractionated ARMS samples by passing all mobile taxa and associated sediment through a set of geological sieves (2 mm, 500 µm and 106 µm) in two steps. In the first step, we stacked 2 mm sieve on top of the 500 µm sieve and allowed all the water and organisms from the transporting container to pass through both sieves and into a new bin. This step isolated motile taxa >2 mm in size for voucher-based DNA barcoding. We then passed the water and associated sediment through the 500 µm sieve stacked above the 106 µm sieve. The result was two size fractioned samples: a 2mm to 500 µm sample

(hereafter referred to as the 500 μm fraction) and a 500 - 106 μm sample (hereafter referred to as the 100 μm fraction). For preservation and storage, we concentrated each of these fractions using a 40 μm nitex mesh and then rinsed the sample with 95% ethanol before preserving each fraction in separate falcon tubes with 95% ethanol that we stored at -20 °C until further processing. Lastly, we scraped all encrusting sessile biota from ARMS plates into a tray (hereafter referred to as the sessile fraction), and then homogenized the sample with a blender for 30 s at maximum speed. We then rinsed the homogenate with 95% ethanol in a 40 μm Nitex mesh and placed the sample in a 50 ml falcon tube filled with DMSO and stored the sample at -20 °C.

DNA extraction and library preparation

To prepare the fractions for DNA extraction, we performed a decantation step on both 500 μm and 100 μm fractions to separate organic contents from inorganic calcium carbonate and terrigenous sediment. Decantation consisted of adding individual samples to a 1 L Erlenmeyer flask with sterile water, agitating the slurry to suspend the entire fraction, allowing the inorganic sediment to settle, and then decanting the suspended organic content into a geological sieve to recover the organic contents. We used a 106 μm sieve to decant the 500 μm fraction and a 45 μm sieve to decant the 100 μm fraction. No decantation was necessary for the sessile fraction.

We performed DNA extractions using 1 g of the decanted 500 μm and 106 μm fractions, and 10g homogenized sessile sample. To break down organic tissue, we incubated samples with 80 μl of 400 μg/ml Proteinase K overnight in a shaker-incubator at 56 ºC and 200 rpm, and then extracted the DNA using the Powermax Soil DNA Isolation Kit (MoBio, Carlsbad, USA) following the manufacturer protocols. To remove potential PCR inhibitors, we further cleaned all DNA extracts using a Power Clean DNA Clean-Up Kit (MoBio, Carlsbad, USA) following the

manufacturer protocols, eluting the DNA into nuclease free water instead of buffer. We then quantified each DNA extraction using a Qubit Fluorometer (Invitrogen, Carlsbad, CA, USA) and the Qubit dsDNA HS Assay Kit (Thermo Fisher Scientific, Waltham, MA, USA).

To assess microbial diversity, we followed Earth Microbiome Project protocols to amplify the V4 region of the 16SrRNA gene region using primers 515f and 806r (Caporaso et al. 2012) (Walters et al. 2015). Library preparation followed a single indexing approach where barcodes incorporated into the forward primer of the 515f-806r primers pair to facilitate multiplexing of up to 96 samples per run. PCR were performed in triplicate, using 5 ng of DNA from each sample. The following PCR conditions were used: initial denaturing at 94 °C for 3 min, 35 cycles each at 94°C for 45 s, 50°C for 60 s, and 72 °C for 90 s, followed by a final extension step at 72 °C for 10 min. Triplicate PCR products were visualized individually on 1.2% agarose gel to confirm correct product size and then pooled and sequenced on MiSeq Illumina using V2 300-cycles kit with 20% PhiX DNA added to improve data quality. Sequencing of final libraries was performed at Smithsonian Institution National Museum of Natural History, Washington DC.

Operational Taxonomic Unit (OTU) and Taxonomic Assignments

We performed quality control on all raw paired-end FASTQ reads using the Quantitative Insights Into Microbial Ecology 2 program (QIIME2, ver. 2018.11 Caporaso et al., 2010) and then demultiplexed sequences using QIIME's native plugin. Next, we used Divisive Amplicon Denoising Algorithm 2 (DADA2) (Callahan et al., 2016) to quality filter, trim, denoise, and merge pairs the data. The completion of these steps resulted a feature table that included representative sequences in FASTA format and the abundance of each of these sequences in the

dataset. We then merged sequences into operational taxonomic units (OTUs), clustering representative sequences at 97% sequence similarity and performing *de novo* taxonomic assignment using QIIME *vsearch* cluster-features-de-novo command from QIIME2 and SILVA SSU non-redundant database (132 release). The result was a new feature/OTU table, taxonomy table and representative sequences, that were then used for subsequent analysis.

Prior to taxonomic assignment, we used LULU (Frøslev et al. 2017) to remove erroneous OTUs (minimum_match = 84, minimum_relative_cooccurence = 0.90 for both markers). We then clustered the LULU outputs using *vsearch* (Rognes et al. 2016) using a 97% clustering level, and assigned the sequences using BLASTN (Camacho et al. 2009) at 85% identity using the 132 release version of the SILVA SSU non-redundant database (https://www.arb-silva.de/download/archive/qiime/), resulting in an OTU table, a table of representative sequences and a taxonomic table (assigned taxonomic name for each representative sequence).

Data Analysis

To evaluate whether sequencing depth was sufficient to capture local microbial diversity, we generated rarefaction curves, plotting OTU richness vs. sequencing depth, using *ggrare* R package (version 0.1.0.) (Kandlikar, 2020). We then rarefied samples of each fraction to even sequencing depths to ensure that diversity metrics were not impacted by variation in sequencing effort. Next, we used the *phyloseq* package (version 1.26.1) (McMurdie & Holmes, 2013) to combine the representative sequences and OTU tables into a single *phyloseq* object/file, calculate alpha diversity (Observed, Inverse Simpson, Shannon, Simpson), group samples based on fraction size and location, and transform sequence data from total reads counts to relative read abundance (RRA). We then visualized these data in *ggplot2* (version 3.3.2) (Wickham, 2016).

To test for significant differences among biodiversity metrics, we first tested the data for normality using the *Shapiro.test* and for homoscedasticity using the *bartlett.test* function in Stats package (version 4.0.2) (R Core Team, 2020). For normally distributed data, we tested for differences in diversity statistics using a one-way ANOVA and for data that were not normally distributed, we used a Kruskal-Wallis test as implemented in Stats package (version 4.0.2) (R Core Team, 2020). We then plotted data that were normally distributed and free of homoscedasticity.

To examine similarities and differences of samples based on fraction size, we constructed a venn diagram using *Venn Diagram* R package (Chen & Boutros, 2011) in R. To examine similarities and differences based on fraction size and location, we performed hierarchical clustering using the count zero multiplicative *(CZM)* method as implemented in *cmultRepl* command in zComposition package (1.3.4), and then visualized the results using *hclust* from *stats* package (4.0.2). To visualize microbial community composition by fraction, we constructed a box plot using *amp_boxplot* command from Ampvis2 R package (2.6.4), with only the eight most abundance taxa visualized, and to visually discern variation of microbial community composition across sites and fractions we constructed bar plots using *ggplot2* (version 3.3.2) (Wickham, 2016), visualizing the 16 most abundant taxa.

To further compare differences in community composition among sites and fractions, we conducted multivariate analyses (PERMANOVA) based on Jaccard distances in the *vegan* package (Ogle, 2017) in R (R development core team), using 9999 permutations to test for statistical significance. We then calculated the compositional dissimilarity using '*ADONIS*' command and the homogeneity of group dispersion using '*betadisper*' command in *vegan* package (Oksanen, 2017). To visualize potential differences, we conducted Principles

Coordinates Analyses (PCoA) using the *Ampvis2* package (Andersen et al., 2018) with the ordination function of *phyloseq,* using Jaccard dissimilarity matrices, and then generated an ordination plot using *ggplot2* (Oksanen, 2017). To compute compositional turnover and the number of OTUs shared by multiple ARMS deployment sites we used *Zeta.decline.mc* function as implemented in zetadiv *R package (1.2.0).*

To identify bacterial taxa indicative of local pollution stress, we conducted two analyses. First, we performed linear discriminant analysis (LDA) and effect size (LEfSe) analyses (Segata et al., 2011), employing a threshold of >2.0 for the logarithmic LDA score in order to take into account discriminant features, using the online Huttenhower Galaxy server (hutternhower.sph.harvard.edu/galaxy). Second, we performed a random forest analysis on the 161 OTUs that have more than 0.05% relative abundance in the dataset, using the R *randomForest* (version 4.6-14) and *rfPermute* package (R 3.5.1) using ntree = 10,000. To identify indicator species from each site we also performed indicator species analysis using R package *indicspeceis* (version 1.7.6) using *multipatt* and *nperm* function with 1000 random permutations to identify taxa most strongly associated with each site. The method calculates the IndVal index between the OTUs and each site group and then looks for the group corresponding to the highest association value (Dufrˆene and Legendre 1997). The maximum value of 1 indicates that the species have high associations with their site groups. In this analysis we considered each ARMS deployment site as an individual group. Random forest analyses did not employ perform cross-validation (splitting the data into training and test sets) as this approach is not recommended for relatively small sample sizes as is common in microbial study (Wainberg et al. 2016). Instead, we examined out-of-bag error (OOB), the calculated error from each model/tree produced during the calculation. OOB error can be calculated by inspecting the

performance of each tree on data they have not been trained on, providing a form of internal cross-validation (Ramosaj & Pauly 2019).

To test for associations between pollution and microbial communities, we plotted microbial diversity against Chlorophyll *a* concentrations derived from remote sensing data. We used Chlorophyll *a* as a proxy for pollution stress because it was not practical to monitor water quality at each site over the three-year deployment period, and productivity can be used as an indicator of integrated aspects of water quality (Boyer et al. 2009). To obtain Chlorophyll a concentrations, we downloaded annual average Chlorophyll *a* concentrations and (MODIS A) at a 4 km resolution across three years (2014-2016) from the NASA Ocean color website (https://ocean- color.gsfc.nasa.gov/). These data were cleaned to remove unnecessary commas and other punctuation using Microsoft excel (Version 16.3) and imported to Ocean Data View (version 5.0.0) for visualization. We then determined Chlorophyll *a* concentrations for each ARMS deployment site was based on satellite data values from the nearest grid point. Due to the spatial resolution of the data, some sites were assigned to the same grid point and thus had the same environmental data. We then performed linear regressions to test for correlations between taxa richness and chlorophyll a concentration using *lm* function from stats r package (4.0.2) and visualized it using r package ggplot2 (3.3.6)

Results

Sequencing and Rarefaction

We recovered only 18 of the 24 deployed ARMS; the remaining 6 could not be located.

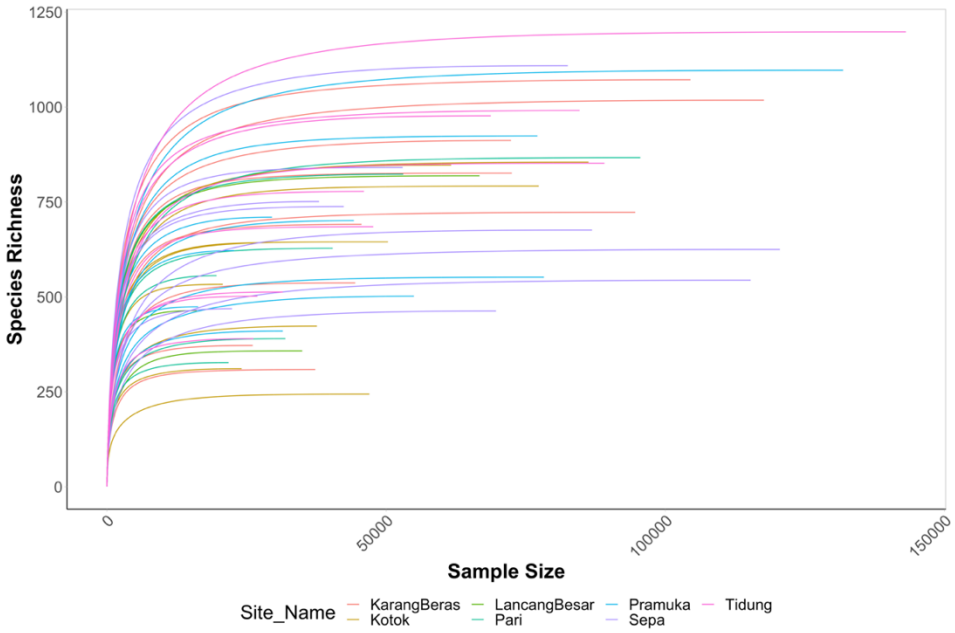


Figure 3-3. Alpha diversity rarefaction plot generated with ggrare (Kandlikar et al., 2018) in R environment. Number of amplified sequence variants (OTUs) (left axis) plotted against sequencing depth (bottom axis) each individual ARMS unit.

From the recovered ARMS, we collected 100um, 500um and sessile fractions, resulting in 54 total samples (Supplemental Table S3-1). After post quality filtering and the exclusion of chimeras, these samples generated 4,074,730 sequence reads with a mean length of 250.92 base pairs. After removing all eukaryotes, archaea, mitochondria, and chloroplast sequences, 3,108,439 sequences remained, representing 8,470 Operational Taxonomic Units (OTUs). Rarefaction plots indicated that this sequencing depth was sufficient to saturate OTU discovery

for nearly all samples (Fig. 3-3). After rarefying our data to a uniform 16,069 sequences per sample, a total of 7,190 OTUs remained.

Composition of Microbial Communities by Fraction

Total microbial communities from ARMS were primarily dominated by Proteobacteria followed by Bacteroidetes and Chloroflexi, with mean relative abundances of 48.82 %, 10.44 % and 8.24 % respectively (Fig. 3-4A). However, when parsed at *Class* level, differences among the fractions emerged. Excluding the dominant proteobacteria classes, the 100um fraction was dominated by Bacteroidia, Oxyproteobacteria and Plancomycetecia with relative read abundances of 16.29 %, 6.59 % and 3.73 %, respectively (Fig. 3-4B), in addition to unknown or uncultured microbes. The 500 μm fractions had the same dominant bacterial classes with relative abundance of 11.64 %, 9.24 % and 3.47 % (Fig. 3-4C), respectively, in addition to unidentified taxa. In contrast, excluding proteobacteria, the sessile fraction was dominated by Dehalococcoidia, Acidimicrobia, Anaerolineae, and Oxyproteobacteria with relative read abundances of 13.89%, 6.96 %, 4.5 %, and 3.55%, respectively (Fig. 3-4D). Unlike the 100 μm and 500 μm fractions, the sessile fraction did not have a substantial number of unidentified taxa.

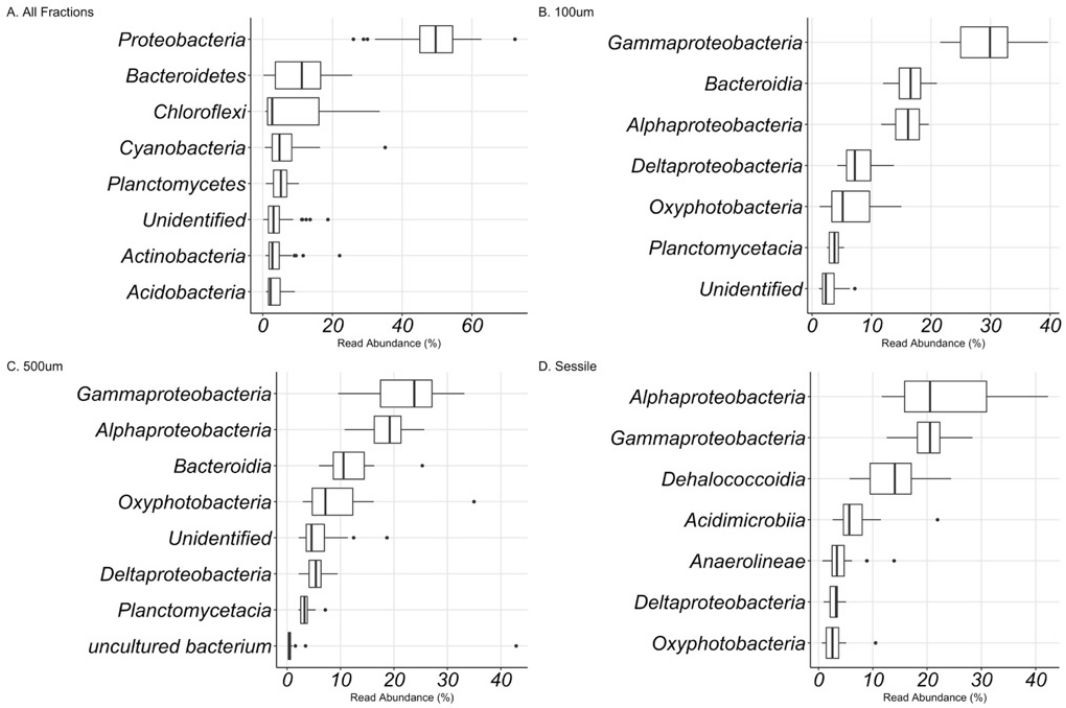


Figure 3-4. A box plot showing taxonomic composition microbial communities at Phylum level for A) and at Class level for B, C and D) for each of the three size fractions. Plot is showing taxa relative abundance of the sample across eight sites in the archipelago. The box plot constructed based on phyla contribute more than 2% of the relative abundance of each sample.

Comparison of Microbial Communities Across Size Fractions

Comparison of microbial community composition across the three fractions demonstrates some overlap, but also substantial differences in OTU composition. Venn diagrams (Fig. 3-5) show that of a total of 7,190 OTUs, only 1,190 (16.6%) were present in all three fractions. The most commonly shared taxa among the three fractions were OTUs in the phylum *Proteobacteria, Chloroflexi* and *Cyanobacteria.* Of 4,456 OTUs from the 100um fraction, 1,814 (40.7%) were

unique to this fraction. Similarly, 1,814 of 4,380 OTUs from the 500um fraction (40.0%) were unique to this fraction and 799 of 2,502 OTUs from the sessile fraction (31.9%) were unique to this fraction.

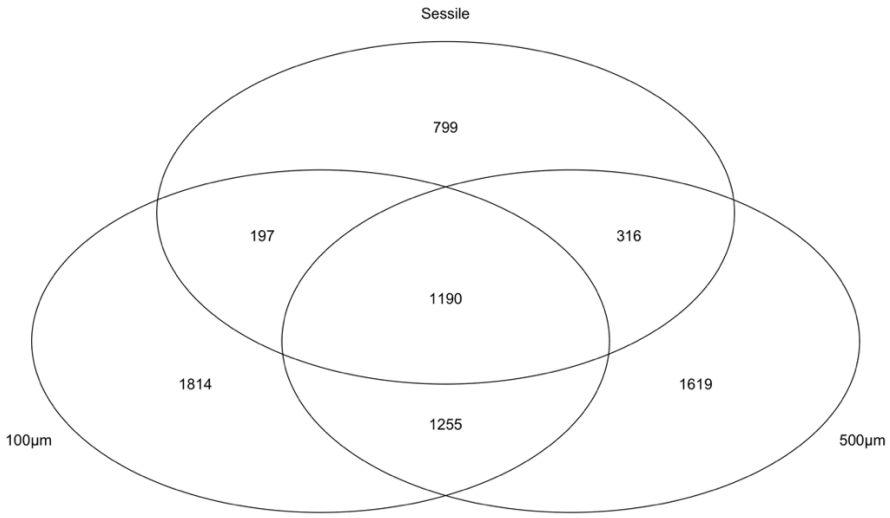


Figure 3-5. Number and distribution of microbial OTUs revealed from16S rRNA metabarcoding of 100 µm, 500 µm, and sessile size fractions. Plot represents data rarefied to an even depth of 16.069 reads per ARMS unit.

Of the 6,391 combined OTUs from the 100 µm and 500 µm fractions, 2,455 (38.3%) were found in both fractions. In contrast, of 5,571 combined OTUs from the 100 µm and sessile fraction, only 1,387 (24.9%) were observed in both fractions, and of 5,376 OTUs combined from the 500um and sessile fraction, only 1,506 (28.0%) were shared among these fractions.

Hierarchical clustering showed that the microbial communities from ARMS samples grouped in two major clades, one comprised largely of all sessile samples and a second comprised of 100 µm and 500 µm fractions, (Fig. 3-6). Although some communities clustered by location (e.g., Tidung and Kotok), others did not, with no clear geographic pattern.

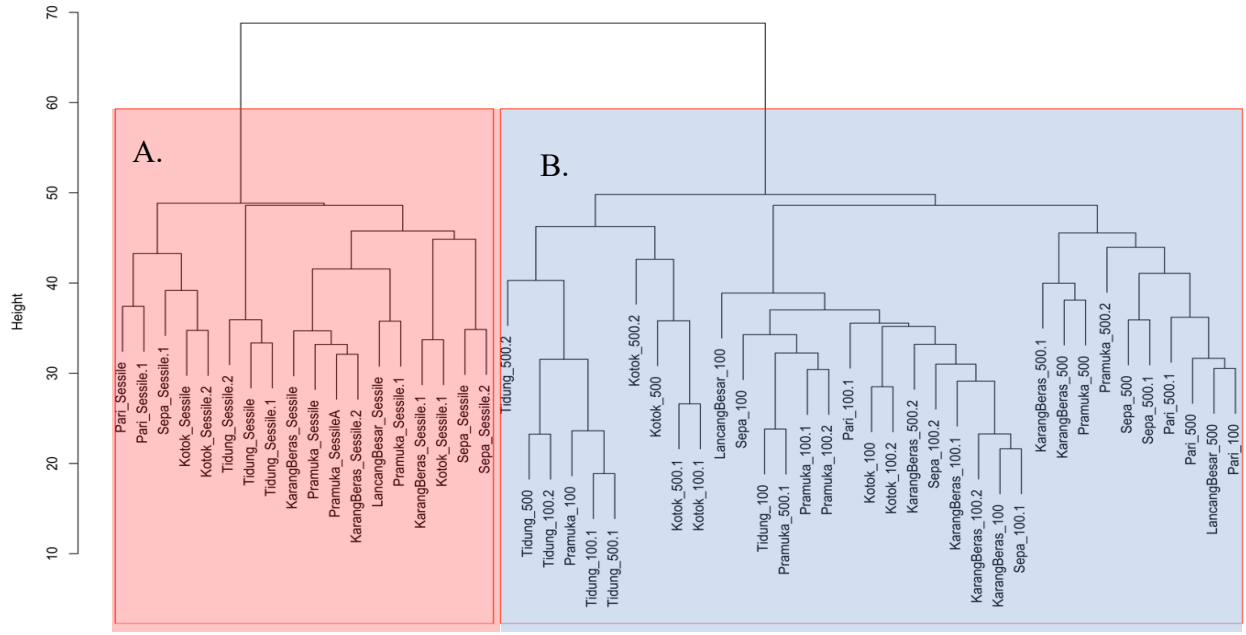


Figure 3-6. Individual sample clustering explaining sample grouping based on fraction sizes. Samples were grouped based on fraction sizes; a sessile fraction created separated clade (A.) while 500 and 100 fractions were lumped as a single clade (B.) Plot was developed using count zero multiplicative (CZM) method implemented in *Zcomposition* package (Version 1.3.4) in R environment

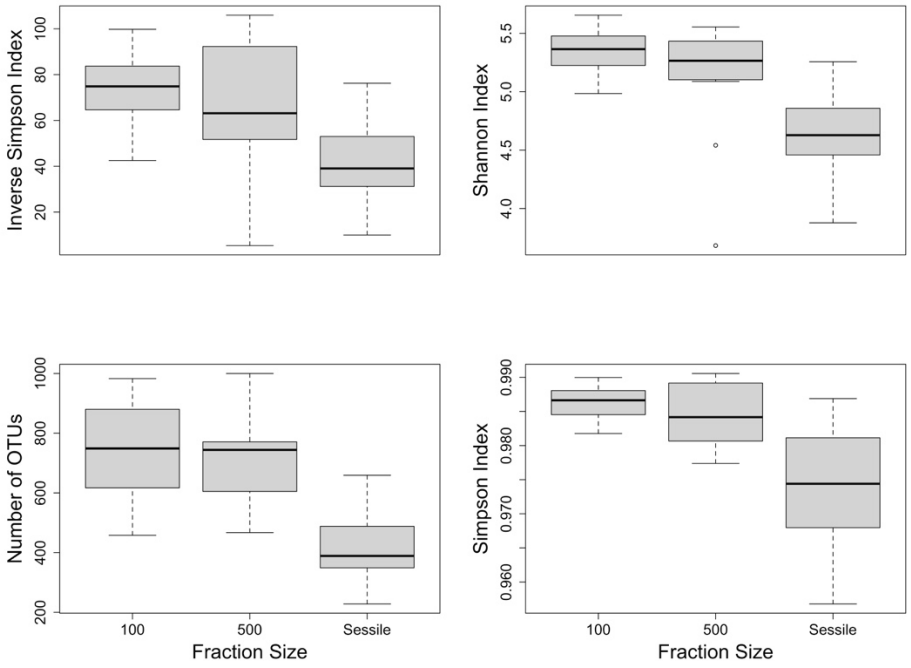


Figure 3-7. Boxplots showing the microbial diversity indices (Inverse Simpson, Shannon Observed and Simpson) across Sites based on fraction size.

Microbial Diversity Across Size Fractions

Normality tests of Observed, Shannon, Simpson, Inverse Simpson, measures of alpha diversity showed that only Observed OTUs conformed to assumptions of a normal distribution (W = 0.97, p-value = 0.3118; Bartlet K-squared = 2.35, df = 2, p-value = 0.3096). ANOVA results from Observed OTU diversity indicate significant differences among microbial OTU diversity among fractions (F-value = 33.49, df = 2, p-value < 0.05) (Fig. 3-7). Similarly, results of Kruskal-Wallis tests show that the Shannon diversity index (Chi square = 26.98, p = < 0.0005, df = 2), Simpson diversity index (Chi square = 19.86, p = < 0.0005, df = 2), and InvSimpson diversity index (Chi square = 19.86, p = < 0.0005, df = 2) are also significantly different.

Figure 3-8. Non-metric multidimensional scaling (NMDS) analysis illustrating dissimilarities in bacterial community composition in different fraction sizes deployed across the archipelago using Jaccard index similarities.

In addition to difference in alpha diversity, non-metric multi-dimensional scaling indicated that beta diversity was significantly different among all three fractions (Fig. 3-8) ($p < 0.05$ and *ADONIS* has $R^2 = 0.07$). Similar to the alpha diversity metrics, beta diversity of the sessile fraction was the most dissimilar to the 100 μm and 500 μm fractions.

Analysis of zeta diversity showed a steep decline in the first orders, especially from (ζ_1) to (ζ_2) across all fraction sizes, with the steepest declines in the 100 µm and 500 µm fractions (1 to 0.24) (Fig. 3-9B); the slowest declines were in the combined fractions (1 to – 0.28). Following zeta order 2 (ζ_2), declines are still observed but are not as steep. For example, a decline of zeta diversity in all fractions from zeta order 5 (ζ_5) to zeta order 6 (ζ_6) is from (0.08– 0.07) and in 100 fraction the value is from (0.05– 0.04). Across all fractions, no zeta diversity values reached zero.

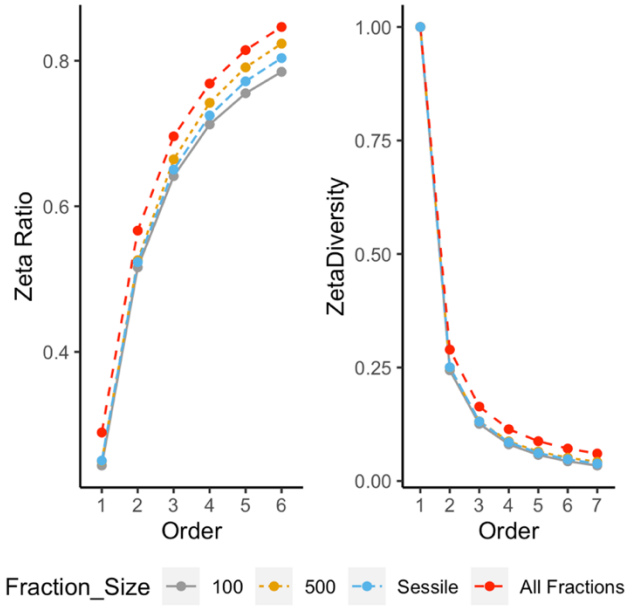


Figure 3-9. Normalized zeta diversity decline for all fractions showing how the number of shared OTUs decreases with the zeta order (a) and The species retention rate using the zeta ratio, which shows the degree to which common OTUs are more likely to be retained in additional cases or samples than rare ones with an increase in zeta order across sites in all fraction and combined fraction (b).

Microbial Community Across Sites Based on Proximities to Source of Pollution

Total microbial diversity and community composition varied across sites. Kotok, the second site furthest from Jakarta, had the highest number of OTUs (1789) and Sepa had the lowest (781); the remaining sites (Pari, Karang Beras, Pramuka) have OTUs between 800 and 1720

Table 3-1. Summary of number of OTUs across sites using 16S rRNA

Site Name	Mean	Max	Min
Lancang Besar	1569	1569	1569
Pari	1533.5	1580	1487
Tidung	1052.6	1246	901
Karang Beras	1186.3	1389	868
Pramuka	1454.3	1715	1311
Kotok	1617.6	1789	1511
Sepa	1109.6	1553	781

Measures of alpha diversity varied across sites (Fig. 3-10). Examining total observed OTUs alpha diversity peaked in Kotok, the second site furthest from Jakarta, and was lowest in Sepa and Karang Beras (Fig. 3-10 A). Similar patterns were also observed in other indices where diversity also peaked at Kotok (Fig. 3-10 B, C, and D). Although Lancang Besar has only one ARMS unit recovered, it had the most diversity based on Simpson's, inverse Simpson's and Shannon indices and total observed OTUs. However, despite these absolute differences, none were significantly different (Shannon index; F value = 1.39 and P value > 0.05 and Bartlett's K-squared = 5.242, df = 4, P value > 0.05: Observed OTUs; F value = 2.32 and P value > 0.05, Bartlett's K-squared = 2.02, df = 4, P value > 0.05: Simpson index; Kruskal-Wallis chi-squared = 9.32, df = 6, P value > 0.05: InvSimpson index; Kruskal-Wallis chi-squared = 9.32, df = 6, P value > 0.05)

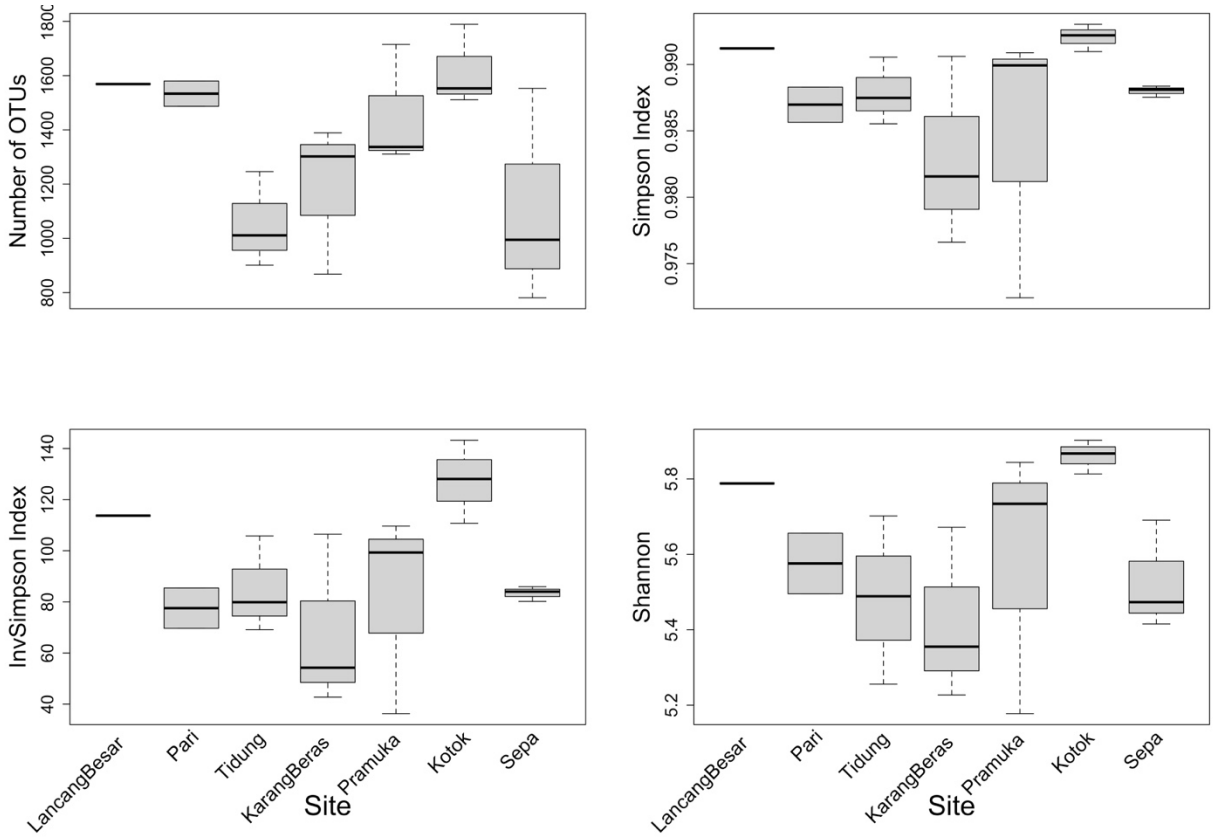


Figure 3-10. Total Operational Taxonomic Units (OTUs) diversity captured in across sites measured with different diversity indices. Highest mean observed number of OTUs was found in Tidung, however using three other indices we found that Karang Beras has relatively higher diversity compared to other sites

Plots of microbial community composition varied across both sites and fractions (Fig. 3-11). Although composition was generally similar across sites in the 100 μm and 500 μm, both Oxyphotobacteria and Deltaproteobacteria had elevated representation in sites with intermediate (e.g., Karang Beras) and short distances to pollution source (e.g., Lancang Besar) (Fig. 3-11 A). Changes in composition, however, were much more pronounced in the sessile fraction. Alphaproteobacteria, Gammaproteobacteria and Dehalococcoidia were generally lowest in

Lancang Besar and higher in sites further from Jakarta. In contrast Acidiomicrobiia and Anaerolineae were highest in Lancang Besar, and lower in sites further from Jakarta. Taxa categorized into 'other', were also generally higher in sites that were closer and further from Jakarta (Fig. 3-11 A), as were *unidentified* taxa (e.g., Pari).

Sepa, the site furthest from Jakarta, has the highest number of unique OTUs (759) consisting mainly of Patescibacteria, Cyanobacteria, Chloroflexi and Protebacteria. The lowest number of unique OTUs (183) were in Lancang Besar, the site closest to Jakarta, and was dominated by Cyanobacteria and Proteobacteria, whereas sites located in the center of Kepulauan Seribu (Tidung and Pramuka) were primarily dominated by Proteobacteria from class Gammaproteobacteria and Dehalococcoidia. Further, as many as 435 OTUs were shared across sites. This shared OTUs were dominated by Alphaproteobacteria, Gammaproteobacteria and Dehalococcoidia mostly from species *Photobacterium rosenbergii*.

Although there was variation in taxonomic composition among sites based on relative read abundance (Fig. 3-11 A), examining microbial communities based on OTU relative abundance were more subtle (Fig. 3-11 B.). In general, there was a decrease in OTU abundance of Bacteroidia and Unidentified taxa in the sessile fraction and an increase in Alphaproteobacteria and Dehalococcoidia.

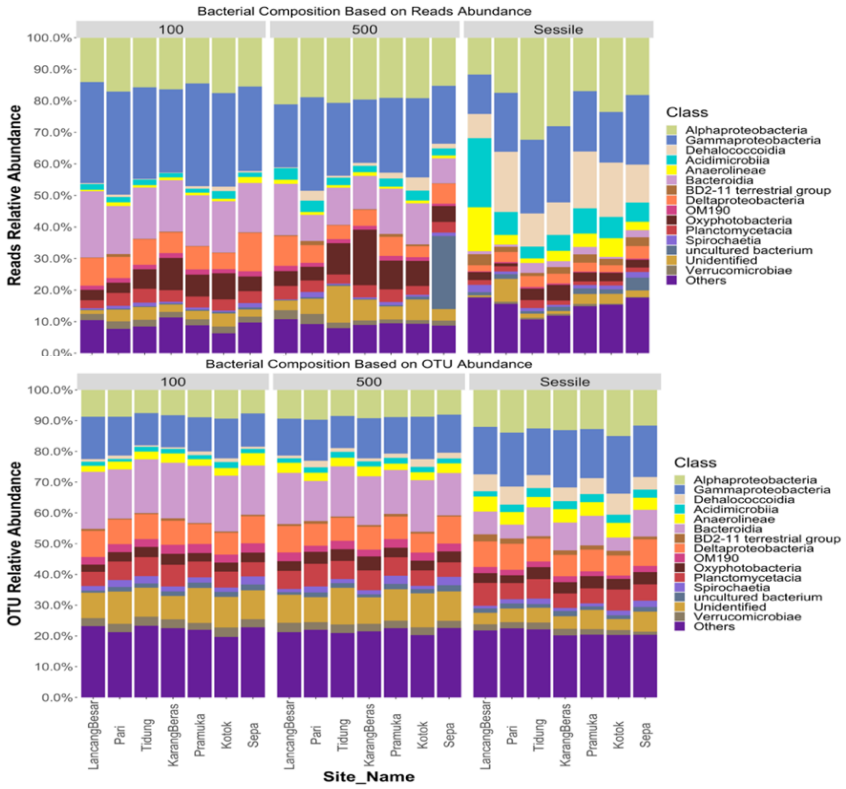


Figure 3-11. Taxonomic composition microbial communities at Class level based on read abundance A) and based on OTU abundance B) for each of the three size fractions at Class level. Bar plot is showing abundance across eight sites in the archipelago. The bar plot constructed based on phyla contribute more than 2% of the relative abundance of each sample.

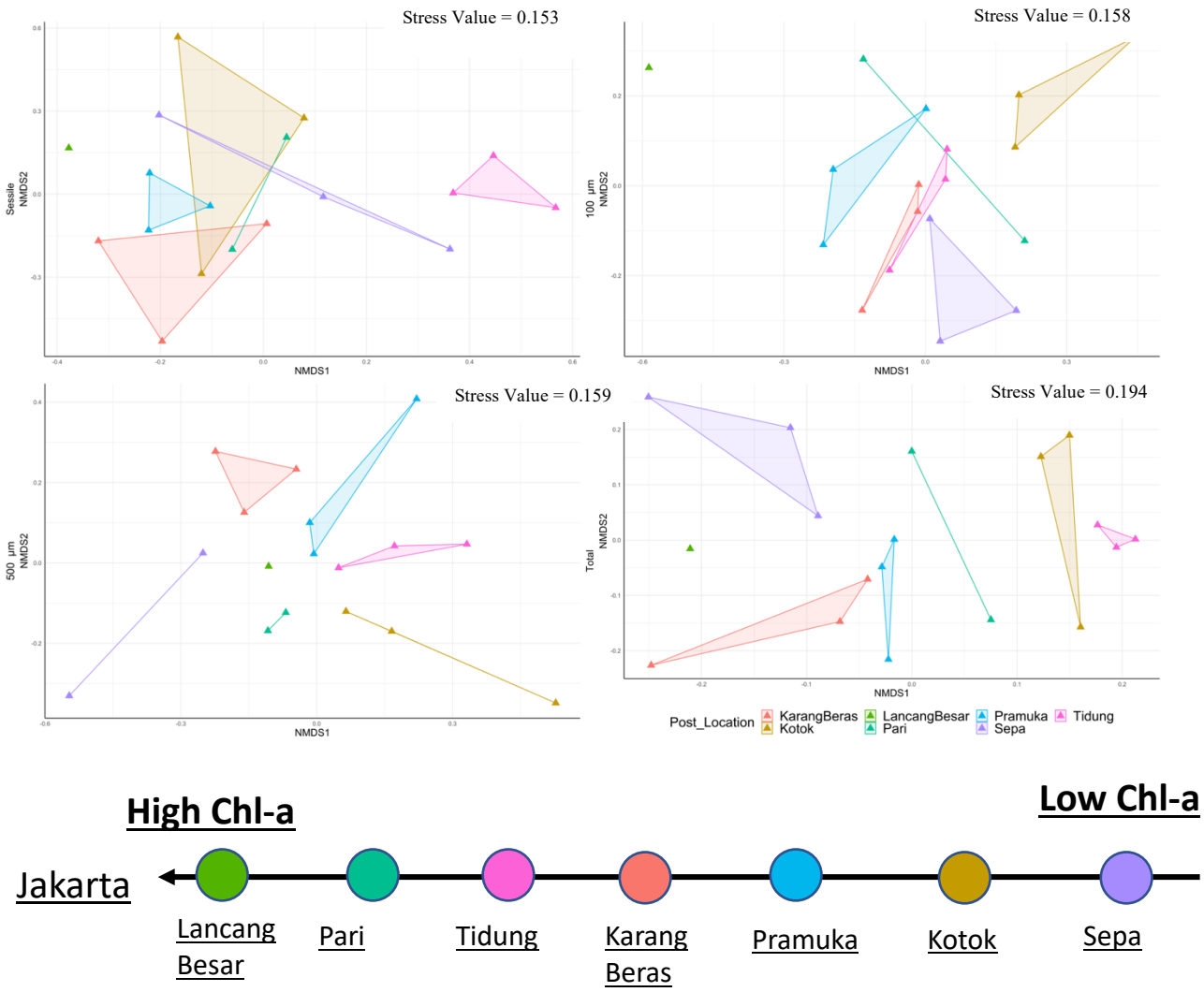


Figure 3-12. Nonmetric multidimensional analysis illustrating dissimilarities in microbial community composition across sites. Analyses using Jaccard similarity were undertaken on individual fractions (106–500 µm, 500 µm -2mm, and sessile) and on all fractions combined across the seven sampling sites.

Results of non-metric multidimensional scaling, merging read abundance from each fraction into a single data point, showed significant differences in the microbial composition

across sites. Betadisper showed confirmed homogeneity in group dispersions ($F= 0.36$, $p > 0.05$), with Adonis returning significant differences among sites based on total microbial diversity (PERMANOVA, $R^2 = 0.14$, $p < 0.005$; Fig. 3-12). Similar results were obtained examining all fractions combined or independently. Both *adonis* and *betadisper* revealed significant differences among locations for all three fractions, with $p < 0.005$ and $R^2 \sim 0.43$ (Adonis) and F value ranges 5 – 10, p value < 0.005 (betadisper) respectively (Table 3-2). Separation among sites located closest and furthest to pollution is apparent in all fractions.

Table 3-2. Beta diversity summary (PERMANOVA) of microbial diversity across sites based on fraction size using 16S rRNA

100 µm Fraction				**500 µm Fraction**			
Adonis		**Betadisper**		**Adonis**		**Betadisper**	
R2	***p-val***	**F**	***p-val***	**R2**	***p-val.***	**F**	***p-val.***
0.43228	**1.00E-03**	6.038	**0.005**	0.428	**1.00E-04**	10.749	**0.001**

Sessile			
Adonis		**Betadisper**	
R2	***p-val.***	**F**	***p-val.***
0.430	**0.0016**	5.6087	**0.004**

Linear Regression between Chlorophyll a Concentration and Total Taxonomic Richness

Three-year averages of Chlorophyll a ranged between 0.50 – 8.59 mg m^{-3}, with the highest concentration observed around Pulau Bidadari (8.59 mg m^{-3}), the site closest to Jakarta. The lowest chlorophyll concentrations were observed in Pramuka and two other sites furthest from Jakarta (Table 3-3).

Linear regression showed that although there was a small increase in total OTU diversity on ARMS with increasing Chlorophyll a concentration (R^2= 0.12), this relationship was not significant (p = 0.1507) (Fig. 3-13). Expanding this approach to examine the relationship between Chlorophyll a concentration and relative abundance of specific microbial taxa that passed tests of normality (Saphiro-Wilk test) and homogeneity (Bartlett test) showed that Chlorophyll a concentration is a not significant predictor of OTU diversity for most phyla detected. The one exception was a significant positive relationship between Chlorophyll a concentration and *Margulisbacteri*a (Cyanobacteria group) (Table 3-4).

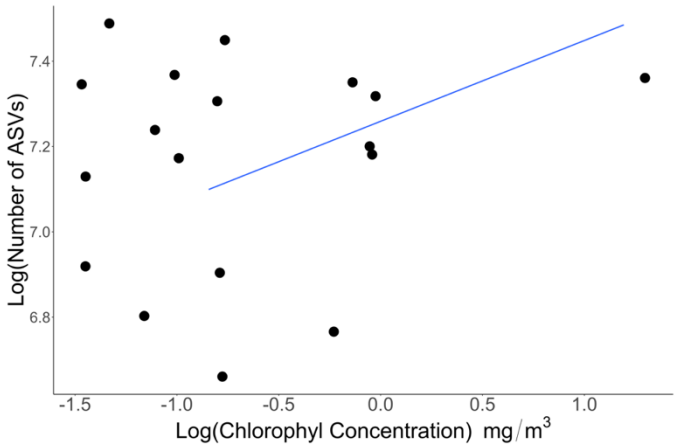


Figure 3-13. Plot of total OTUs as a function of Chlorophyll a concentration on including best fit line and R^2 value

Table 3-3. Annual average Chlorophyll Chlorophyll *a* concentrations and (MODIS A) at a 4 km resolution across three years (2014-2016)

Site Name	**2014**	**2015**	**2016**	Three years Average
Bidadari	9.38	7.57	8.66	8.54
Lancang Besar	3.20	2.01	4.80	3.34
Pari	0.64	0.82	0.94	0.80
Tidung	0.44	0.67	0.47	0.53
Karang Beras	0.65	0.65	0.68	0.66
Pramuka	0.50	0.48	0.54	0.50
Kotok	0.52	0.56	0.47	0.52
Sepa	0.52	0.56	0.47	0.52

Table 3-4. P-value calculated from linear regression modelling between OTU richness per taxa versus Chlorophyll a concentration. Taxa with significant p-value were highlighted in bold

No.	Variable	r. Squ.	adj.r. squ	Sigma	Stat.	p. Val.	df	logLik	AIC	BIC	Dev.	df. Res.	nobs
1	Acetothermia	0.16	0.08	0.54	1.91	0.20	1	-8.58	23.15	24.61	2.93	10	12
2	Acidobacteria	0.01	-0.05	0.48	0.24	0.63	1	-11.09	28.18	30.85	3.61	16	18
3	Bacteroidetes	0.04	-0.02	0.47	0.74	0.40	1	-10.81	27.63	30.30	3.50	16	18
4	BRC1	0.01	-0.06	0.48	0.10	0.75	1	-11.16	28.33	31.00	3.64	16	18
5	Chlamydiae	0.01	-0.06	0.48	0.09	0.77	1	-11.17	28.34	31.01	3.65	16	18
6	Chloroflexi	0.04	-0.02	0.47	0.62	0.44	1	-10.88	27.76	30.43	3.53	16	18
7	Deferribacteres	0.27	0.09	0.19	1.49	0.29	1	2.69	0.62	-0.01	0.14	4	6
8	Dependentiae	0.00	-0.06	0.48	0.08	0.79	1	-11.18	28.36	31.03	3.65	16	18
9	Gemmatimonadetes	0.01	-0.05	0.48	0.18	0.68	1	-11.12	28.24	30.92	3.63	16	18
10	Kiritimatiellaeota	0.02	-0.04	0.47	0.35	0.56	1	-11.03	28.05	30.72	3.59	16	18
11	Latescibacteria	0.03	-0.03	0.47	0.43	0.52	1	-10.98	27.97	30.64	3.57	16	18
12	Lentisphaerae	0.03	-0.03	0.47	0.49	0.49	1	-10.95	27.89	30.57	3.56	16	18
13	**Margulisbacteria**	**0.39**	**0.33**	0.42	**7.02**	**0.02**	**1**	**-6.20**	**18.41**	**20.10**	**1.98**	**11**	**13**
14	Marinimicrobia (SAR406 clade)	0.01	-0.19	0.77	0.03	0.88	1	-6.89	19.77	19.61	2.93	5	7
15	Modulibacteria	0.29	0.20	0.19	3.21	0.11	1	3.46	-0.92	-0.01	0.29	8	10
16	Nitrospinae	0.10	0.04	0.45	1.75	0.20	1	-10.29	26.57	29.24	3.30	16	18
17	Nitrospirae	0.00	-0.06	0.48	0.00	0.97	1	-11.22	28.44	31.11	3.67	16	18
18	PAUC34f	0.02	-0.04	0.47	0.31	0.59	1	-11.05	28.10	30.77	3.60	16	18
19	Planctomycetes	0.01	-0.06	0.48	0.10	0.76	1	-11.16	28.33	31.00	3.64	16	18
20	Proteobacteria	0.01	-0.05	0.48	0.24	0.63	1	-11.09	28.17	30.84	3.61	16	18
21	Schekmanbacteria	0.21	0.08	0.15	1.60	0.25	1	5.19	-4.39	-4.15	0.13	6	8
22	Spirochaetes	0.08	0.02	0.46	1.34	0.26	1	-10.50	26.99	29.66	3.38	16	18
23	Verrucomicrobia	0.04	-0.02	0.47	0.64	0.43	1	-10.87	27.73	30.40	3.52	16	18
24	WPS-2	0.01	-0.09	0.57	0.10	0.76	1	-9.15	24.30	25.75	3.23	10	12
25	WS2	0.03	-0.07	0.60	0.32	0.59	1	-8.92	23.85	25.04	3.26	9	11

Bacteria Taxa indicative of Pollution impact on ARMS

Lineal discriminant analysis (LDA) and the effect size (LEfSe) revealed bacterial bioindicators significantly associated with deployment sites. In total, 50 taxa with LDA score higher than 2.0 were identified (Supplemental Figure S3-1). The most polluted sites, Lancang Besar and Pari had 17 indicator taxa each, greatly exceeding the five from Tidung, two for Pramuka, six for Kotok, one for Karang Beras and two for Sepa.

Results from random forest analyses shows an OOB error of 24.53% with the lowest class of error observed in Karang Beras and Sepa (Supplemental Table S3-2). Analyses showed that different taxa were most important in differentiating among sites. In the site closest to Jakarta, Lancang Besar, Desulfovibrionaceae has the highest importance value (0.011) followed by bacteria from family A4B (0.0078), unidentified uncultured bacterium (0.0061) and Clostridiaceae (0.0056). In the sites intermediate distance from Jakarta (Karang Beras and Pramuka), Unassigned taxa from OTU 22408 (0.0041) and Phormidesmiaceae (0.0039) are two bacteria family with highest value of importance respectively. Finally, sites furthest from Jakarta (Kotok and Sepa) Kordiimonadaceae from OTU11384 (0.0053) and Unidentified taxa from class Acidobacteria OTU1237 (0.0025) are taxa with the highest importance values respectively (Supplemental Figure S3-2).

Indicator species analysis showed that each site has varying number of indicator taxa. The most polluted site, Lancang Besar, had the highest number of indicator taxa (95 total), while Pramuka had the lowest (15 taxa). Taxa with the highest association values in Lancang Besar were Oxyphotobacteria and Gammaproteobacteria with association value of 0.816 respectively (p-value < 0.05). Meanwhile at Pramuka the highest association value were unidentified taxa and Gammaproteobacteria with association value of 0.745 and 0.715 respectively (p-value < 0.05)

Intermediate sites, Pari, Tidung and Karang beras have 68, 32 and 45 taxa indicative of the site respectively, with Spirochaetia and Gammaproteobacteria having the highest association value with in Tidung (0.852 and 0.828 respectively and p-value < 0.05). In contrast, the site furthest from Jakarta Bay, Sepa, has 47 indicative taxa with the highest association value of 0.866 and 0.766 (p-value < 0.05) owned by Verrucomicrobiae and and Kotok has 28 indicator taxa with the highest association value of 0.882 owned by Alphaproteobacteria. Full list of indicator taxa and their associated sites are presented in supplemental table S3-3

Discussion

Studies examining the impacts of pollution on marine ecosystems largely focus on macrofauna, particularly corals and fishes (Waldichuk 1974, Pastorok & Bilyard 1985, Dubinksy & Stambler 1996, Ward-Paige et al. 2005, Cebrian et al. 2007, van der Meij et al. 2009, Rochman et al. 2014, Butnariu 2022), but studies increasingly show that microbes play an essential role in the health of coral reef ecosystems (Mao-Jones et al. 2010, Hester et al. 2016, Webster & Reusch 2017, Glasl et al. 2018). Despite substantial variation in local pollution stress across Pulau Seribu, there was no significant impact on total microbial community diversity from ARMS. Like metazoan diversity from ARMS across these pollution gradients (Chapter 2), there was a small, but non-significant increase in bacterial diversity in sites with increased pollution stress. This result stands in contrast to previous studies showing increased (Montenegro et al. 2020) or decreased (Ding et al. 2017, Xiong et al. 2018) richness and diversity of marine microbial communities in more polluted sites, and to previous metazoan studies from Pulau Seribu reporting decreased fish abundance and richness (Madduppa 2013), and increased sea urchin abundance and richness increased (Cleary et al. 2008) with increased pollution levels. Combined these results highlight taxonomic specific responses to pollution.

Although there were no clear shifts in overall diversity with pollution stress, there were significant differences in microbial community composition among the three fractions, with surprisingly limited taxonomic overlap, suggesting that the microbial communities are strongly associated with the metazoan communities recovered in each fraction. Previous studies report differences in microbial composition as results of habitat partitioning (Sawall et al. 2012, Tout et al. 2014, Glasl et al. 2019, Frade et al. 2020) . For example, Tout et al. (2014) demonstrated clear

shifts in microbial communities between the water column and sandy substrates on coral reefs. There were also pronounced differences in beta diversity among sampling sites, a result observed in plant communities in response to pollution stress (Trubina & Vorobeichik 2012, Montenegro et al. 2020). In addition, specific microbial taxa responded differently to pollution stress, indicating that although total microbial community diversity may not shift in response to pollution stress, communities change in important ways in response to the pollution across Pulau Seribu.

Distinct but Overlapping Microbial Communities

Despite the growing application of ARMS to studies of marine biodiversity, the majority of these studies focus on eukaryotic, rather than microbial communities (Plaisance et al. 2011, Leray & Knowlton 2015, Al-Rshaidat et al. 2016, Pearman et al. 2016, Ransome et al. 2017). Of the few studies that examine microbial communities, they either only examine the sessile ARMS fraction (Pearman et al, 2019) or don't directly compare diversity in the three fractions (Ip et al. 2022). Detailed examination of microbial communities on each of the three ARMS fractions revealed that the metazoans captured in each of the size fractions host unique microbial communities, with surprisingly little overlap.

Across all biodiversity indices, results showed that the 100 μm and 500 μm fractions had substantially higher community diversity than the sessile fraction; although the 100 μm fraction had higher diversity than the 500 μm fraction, these differences were not significant. Elevated microbial diversity in the 100 μm fraction could result from carry over of small fragments of organisms captured by the 500 μm sieve, resulting in the accumulation of bacterial taxa. Although both the Venn diagram and clustering analyses show the greatest similarities in the 100

μm and 500 μm microbial communities, the 100 μm fraction also had the highest number of unique microbial ASVs suggesting that elevated diversity in this fraction is likely not the result of sample carryover. Given that the 100 μm fraction traps sediment, it is possible that the elevated microbial diversity is a function of microbes associated with silt and sediment trapped in 100 μm fraction. However, silt and sediment are removed via the decantation steps. Rather, elevated microbial diversity in the 100 μm fraction is likely a function of higher metazoan diversity in the 100 μm fraction, as shown in chapter 2 and Cahyani (2021).

The low diversity of bacterial communities in the sessile fraction is likely a function of lower diversity of metazoans in this fraction. Our sample design captures bacteria associated with metazoans; as metazoan diversity decreases, so too, should bacterial diversity. The sessile metazoan community on ARMS plates can be dominated by particular taxa, and these taxa can prevent other taxa from settling and flourishing (Kuffner et al. 2006). For example, although marine sponges harbor a large number of bacteria (Taylor et al. 2004, Schmitt et al. 2012), they also have chemical defenses that can prevent other bacteria and marine pathogens from thriving (Helber et al. 2018), potentially reducing overall bacterial diversity within the sessile community encrusting ARMS plates. Although we cannot test this hypothesis because we didn't sample sponges separately, photographs of the ARMS plates confirm high sponge coverage on individual ARMS plates, suggesting that sponges could reduce total diversity through inhibition of other taxa.

In addition to sponges, micro and macroalgae can also reduce bacterial community diversity in the sessile fraction. Many ARMS units had higher sediment loads and more microalgae coverage in Pulau Seribu than ARMS from other regions of Indonesia (pers, obs). As with sponges, marine algae can exert an antifouling effect that prevents settlement and growth of

other marine species (Paul & Fenical 1986), potentially reducing the diversity associated microbial taxa.

Community Diversity Across Sites

Comparison of total bacterial community diversity to pollution levels, as inferred by Chlorophyll-a concentration, showed a positive relationship between pollution and bacterial diversity, although this relationship was not significant. Previous studies examining bacterial diversity and pollution stress show contrasting results. Montenegro et al. (2020) found that the mucus coating of fish skin had higher richness and diversity of microbes in more contaminated areas and Cho & Kim (2000) showed that bacterial communities in the livestock wastewater and the contaminated aquifer were much more diverse than those in the uncontaminated aquifer. However, studies of marine (Ding et al 2017), terrestrial (Xiong et al 2018), and estuarine (Jose et al. 2011) sediments report decreased bacterial diversity with increased contamination. Similarly, Cao et al. 2012 report that the alpha diversity of pathogenic microbes is decreased in healthy coral ecosystems and higher at more polluted sites.

Similarly, although PCA plots indicate clear and significant differences in microbial communities across Pulau Seribu, there is no clear pattern related to pollution stress or proximity to Jakarta Bay. Microbial communities from Lancang Besar and Pari, the two sites closest to Jakarta cluster together suggesting that bacterial taxa composition within ARMS are affected by pollution from Jakarta Bay. However, Sepa and Kotok do not cluster together despite these sites being the most distant from Jakarta. This lack of a clear pattern could result if microbial communities respond in inconsistent ways to pollution stress, or if pollution stress must exceed a threshold before microbial communities change(Nogales et al. 2007, Nogales, Lanfranconi, Piña,

et al. 2011) . However, the most likely explanation for this clustering pattern is that while sites closest to Jakarta Bay are dominated by pollution originating from Jakarta and surrounding communities, sites more distant from Jakarta Bay are influenced more by local anthropogenic activities. For example, during ARMS deployment, domestic tourist visits to the archipelago, particularly Tidung, increased by 200% (Annual report 2017 on Pulau Tidung); with nonexistent local wastewater treatment, all this domestic waste is dumped into the ocean. Because our Chlorophyll a data is derived from 4km grids, such local pollution sources could shape local marine biodiversity without being captured in remote sensing data, contributing to our anomalous results.

The lack of clear patterns in microbial diversity relative to distance from Jakarta Bay is surprising given previous studies examining higher taxa (e.g., coral, fish, sponge) Kepulauan Seribu. For example, Madduppa et al. (2013) shows clear shifts in fish abundance and diversity across the archipelago, with populations grouped into three distinct clusters representing sites closest, intermediate, and furthest from pollution originating from Jakarta. Similarly, coral community composition also changes with pollution levels across the archipelago (Rachello-Dolmen & Cleary 2007), with large/massive stress-tolerant species dominating polluted areas and smaller/branching species in areas with less pollution.

Multiple factors may contribute to our inability to detect a significant relationship, positive or negative, between pollution stress and microbial diversity in our study. First, ARMS processing protocols result in the recovery of microbial communities associated with eukaryotes. As such, the overall communities may be more reflective of metazoan diversity, and how metazoans respond to pollution. In contrast, sediments include freely-associated bacterial communities (decomposers, chemophiles, etc), and sediments can accumulate pollution

contaminants (Olsen et al. 1982, Weng et al. 2008, Sharifuzzaman et al. 2015), supporting elevated bacterial diversity and abundance (Wang et al. 2012, McDevitt-Irwin et al. 2017, Catania et al. 2018). Moreover, many of the microbes in contaminated sediment are taxa that degrade pollution contaminants under aerobic, microaerophilic, and anaerobic conditions (Xue et al. 2015), which likely would not be present on, or inside of, living organisms. Lastly, and perhaps most importantly, the loss of ARMS from our most polluted sites likely reduced our ability to detect differences, as the remaining sites were so similar in Chlorophyll-a concentrations. Inclusion of replicate samples from Bididari and Lancang Besar were important in elucidating a pattern of reduced diversity with increased pollution in eDNA analyses of metazoans (Chapter 2); their absence in the microbial datasets likely reduces our power to detect patterns. Additionally, pollution originating from Jakarta Bay might impact reefs closest to this pollution source, but that distant reefs are more impacted by local pollution stress or that pollution impacts may be more consistent on reefs close to Jakarta, and that seasonal ocean currents in this region (Tomascik et al. 1995) may result in more variable pollution impacts on more distant reefs.

Indicators of Pollution Stress

Despite broad similarities in bacterial taxa composition across sites, both composition and reads abundance varied greatly across sites, particularly in the sessile fraction. For example, both Alpha- and Gamma-proteobacteria abundance decreased markedly at Lancang Besar and Pari, the two sites closest to Jakarta and with highest Chlorophyll-a levels and increased with increasing distance from Jakarta. Both Anaerolineae and Acidimircrobiia had higher abundances in Lancang Besar than in sites with lower pollution stress. These results agree with previous

studies showing higher alphaproteobacteria (e.g., SAR11 or SAR86) abundance in sites with higher abundance of gammaproteobacteria (e.g., Oceanospirillales, Alteromonadales, Enterobacteriales, and Vibrionales) in more polluted ecosystems (Zhang et al. 2007, 2009, Nogales et al. 2007, Nogales, Lanfranconi, Pi??a-Villalonga, et al. 2011). Similarly, Acidimircrobiia, which peaked in abundance in Lancang Besar and Pari, is associated with chemical pollution (Huang & Jaffé 2019). In particular, Acidimicrobium bacterium A6, metabolizes chemicals such as perfluorooctanoic acid and perfluorooctane sulfonate, chemicals abundant in Jakarta Bay as a result of increasing chemical and boating activities (Hosono et al. 2011, Irnidayanti 2015, Riani et al. 2018).

Many other bacterial taxa found in Lancang Besar (e.g., Rhodobacterales, Microtrichales, Caldilineales, and Cellvibrionales) are taxa commonly found in areas with high pollution levels. For example, *Candidatus microthrix* (family Microthricaceae) is known to inhabit sludge water (Blackall et al. 199,6), and the Cellvibrionales group is known to utilize xylene (Iwaki et al. 2018), which originates from various industrial activities and gasoline residue from boating activities (Fay et al. 2007). Sludge water, including xylene, is commonly found in Jakarta Bay, as a result of effluent from industrial activities around Jakarta and from shipping originating from the port of Tanjung Priok (Gilbert & James 1994, Undap et al. 2013).

Further evidence of pollution impacts in Lancang Besar and Pari come from linear discriminant analysis (LDA) and the effect size (LEfSe) that reveal a 3- 8-fold increase in bacterial bioindicators associated with pollution stress. Microbes in the family Burkholdeceae had the largest response in the most polluted site (Lancang Besar), including species such as *Burkholderia pseudomallei* and *B. mallei*, which are animals and human pathogens (Ribot & Ulrich 2006, Kaewpan et al. 2022), *B. caryophylli, B. gladioli*, and B. *solanacearum*, which are

plant pathogens (Maeda et al. 2006, Elshafie & Camele 2021) and *B. pickettii*, an opportunistic bacterium (Ryan & Adley 2014). Although not common in marine environments, Burkholdeceae are known to utilize ammonium and nitrite via ammonia oxidation and nitrification in various aquatic environments (Salcher et al., 2019). The presence of nitrites in surface waters in conjunction with high ammonia levels can be indicative of sewage effluent, which is commonly discharged into Jakarta Bay (Uneputty & Evans 1997, Willoughby et al. 1997, Undap et al. 2013, Simanjuntak et al. 2018). Lactobacilliceae was also a significant indicator in Lancang Besar, a group of fermenting bacteria associated with nutrient rich environments. In contrast, although Pari had many more indicator taxa than other sites with less pollution stress, Burkholdeceae and Lactobacilliceae were not among them, suggesting these are indicative of particularly high pollution stress.

Random forest analyses is commonly used in microbial studies to detect associations between bacterial taxa and environmental conditions (Delgado-Baquerizo et al. 2016, Moitinho-Silva et al. 2017, Corrigan et al. 2018). Random forest analyses indicate that Clostridiales and Desulfovibrionaceae are strongly associated with polluted sites like Lancang Besar, and Kordiimonadales and Synechococcales are taxa associated with less polluted sites like Kotok. Clostridiales is a common fecal pollution indicator in coastal environments (Halliday et al. 2014), and Desulfovibrionaceae is a sulfate-reducing bacteria linked to pollution that is used in bioremediation (Zhang et al. 2008, Kondo et al. 2012, Zouch et al. 2017) and abundant in site closest to the pollution level. In contrast, Kordiimonadales is from a class of alphaproteobacteria that is associated with oligotrophic water (Nogales et al. 2007, Aguiló-Ferretjans et al. 2008).

Conclusions

Although ARMS were designed to study invertebrate cryptofauna in marine ecosystems, our results show their utility in the study of microbial diversity. ARMS provide the ability to sample both marine invertebrates and associated microbial communities in a highly standardized manner. The results of this study, however, indicate that even the microbial communities recovered from ARMS are highly dependent the fractions analyzed, as well as the local characteristics where ARMS are deployed. Although microbial communities from ARMS deployed closest to Jakarta Bay showed strong indications of pollution stress, there were no clear gradients with increasing distance from this pollution source. Other factors such as local pollution stress, current patterns and strength, water circulation, bottom slope, rugosity, etc. may also be impacting local communities. Studies using ARMS need to consider these details and, as much as possible, ensure comparable deployment locations to reduce variation and get comparable results.

To date, conservation planning has primarily focused on macro-organisms, particularly keystone species and those of high economic value (Mcmanus et al. 1997, Hughes et al. 2003, Hoegh-Guldberg et al. 2007). Although changes in coral cover and fish biomass over time are important metrics of ecosystem health and resilience, these taxa are long-lived and may not respond quickly to environmental conditions. Results of this study shows how studies of microbial communities using a standardize platform like ARMS can provide insights into pollution stress on coral reef ecosystem, potentially allowing for more rapid detection of changes in environmental than larger macrofauna. Given that ARMS data is being collected on a global scale, data from microbial communities could provide important information on environmental conditions to include in integrative conservation management decisions.

Supplemental Tables and Figures

Supplemental Table S3-1. Location and number of metabarcoding samples used on this study.

SampleID	Site Name	ARMS Name	Fraction_Size
SKBS1A100	Karang Beras	SKBS1A	100
SKBS1A500	Karang Beras	SKBS1A	500
SKBS1ADMSO	Karang Beras	SKBS1A	Sessile
SKBS1B100	Karang Beras	SKBS1B	100
SKBS1B500	Karang Beras	SKBS1B	500
SKBS1BDMSO	Karang Beras	SKBS1B	Sessile
SKBS1C100	Karang Beras	SKBS1C	100
SKBS1C500	Karang Beras	SKBS1C	500
SKBS1CDMSO	Karang Beras	SKBS1C	Sessile
SKOT1A100	Kotok	SKOT1A	100
SKOT1A500	Kotok	SKOT1A	500
SKOT1ADMSO	Kotok	SKOT1A	Sessile
SKOT1B100	Kotok	SKOT1B	100
SKOT1B500	Kotok	SKOT1B	500
SKOT1BDMSO	Kotok	SKOT1B	Sessile
SKOT1C100	Kotok	SKOT1C	100
SKOT1C500	Kotok	SKOT1C	500
SKOT1CDMSO	Kotok	SKOT1C	Sessile
SLCB1A100	Lancang Besar	SLCB1A	100
SLCB1A500	Lancang Besar	SLCB1A	500
SLCB1ADMSO	Lancang Besar	SLCB1A	Sessile
SPAR1A100	Pari	SPAR1A	100
SPAR1A500	Pari	SPAR1A	500
SPAR1ADMSO	Pari	SPAR1A	Sessile
SPAR1B100	Pari	SPAR1B	100
SPAR1B500	Pari	SPAR1B	500
SPAR1BDMSO	Pari	SPAR1B	Sessile
SPRM1A100	Pramuka	SPRM1A	100
SPRM1A500	Pramuka	SPRM1A	500
SPRM1ADMSOA	Pramuka	SPRM1A	Sessile
SPRM1B100	Pramuka	SPRM1B	100

SPRM1B500	Pramuka	SPRM1B	500
SPRM1BDMSO	Pramuka	SPRM1B	Sessile
SPRM1C100	Pramuka	SPRM1C	100
SPRM1C500	Pramuka	SPRM1C	500
SPRM1CDMSO	Pramuka	SPRM1C	Sessile
SSEP1A100	Sepa	SSEP1A	100
SSEP1A500	Sepa	SSEP1A	500
SSEP1ADMSO	Sepa	SSEP1A	Sessile
SSEP1B100	Sepa	SSEP1B	100
SSEP1B500	Sepa	SSEP1B	500
SSEP1BDMSO	Sepa	SSEP1B	Sessile
SSEP1C100	Sepa	SSEP1C	100
SSEP1CDMSO	Sepa	SSEP1C	Sessile
STDN1A100	Tidung	STDN1A	100
STDN1A500	Tidung	STDN1A	500
STDN1ADMSO	Tidung	STDN1A	Sessile
STDN1B100	Tidung	STDN1B	100
STDN1B500	Tidung	STDN1B	500
STDN1BDMSO	Tidung	STDN1B	Sessile
STDN1C100	Tidung	SPRM1C	100
STDN1C500	Tidung	STDN1C	500
STDN1CDMSO	Tidung	STDN1C	Sessile

Supplemental Table S3-2. Results from random forest analysis and Out of the Bag (OOB) error form seven sites observed in this study

Type of random forest: classification
Number of trees: 10000
No. of variables tried at each split: 84
OOB estimate of error rate: 24.5%

	Karang Beras	Kotok	Lancang Besar	Pari	Pramuka	Sepa	Tidung	Class of Error
Karang Beras	8	0	1	0	0	0	0	0.111
Kotok	0	7	2	0	0	0	0	0.222
Lancang Besar	0	1	2	0	0	0	0	0.333
Pari	0	0	2	3	1	0	0	0.5
Pramuka	0	0	2	0	6	0	1	0.333
Sepa	0	1	0	0	0	7	0	0.125
Tidung	0	1	0	0	1	0	7	0.222

Supplemental Table S3-3 Summary of significant bacterial taxa associated with sites based on indicspecies analysis. The association value indicates the strength of the association for the respective OTU with the tested sample group

Group Karang Beras 45

	stat	p.value
OTU22408_Unidentified	0.903	0.003
OTU45735_Bacteroidia	0.873	0.001
OTU538_Gammaproteobacteria	0.821	0.001
OTU35950_Bacteroidia	0.816	0.003
OTU44015_Oxyphotobacteria	0.816	0.002
OTU49651_Gammaproteobacteria	0.776	0.011
OTU2487_Oxyphotobacteria	0.747	0.003
OTU5146_Bacteroidia	0.745	0.011
OTU19601_Alphaproteobacteria	0.745	0.005
OTU27943_Deltaproteobacteria	0.745	0.001
OTU18017_Bacteroidia	0.728	0.006
OTU28541_Planctomycetacia	0.697	0.006
OTU12740_Alphaproteobacteria	0.691	0.011
OTU22704_Bacteroidia	0.685	0.018
OTU10292_Unidentified	0.671	0.012
OTU5409_Unidentified	0.667	0.011
OTU31474_BD2-11 terrestrial	0.655	0.015
OTU33037_Bacteroidia	0.652	0.017
OTU6495_Gammaproteobacteria	0.637	0.021
OTU25708_WCHB1-81	0.633	0.025
OTU14359_Anaerolineae	0.626	0.022
OTU20997_Bacteroidia	0.584	0.038
OTU15879_Mollicutes	0.577	0.029
OTU17300_Unidentified	0.577	0.022
OTU19196_Deltaproteobacteria	0.577	0.032
OTU16260_Campylobacteria	0.577	0.028
OTU35016_Spirochaetia	0.577	0.032
OTU44607_Fibrobacteria	0.577	0.022
OTU47902_Unidentified	0.577	0.018
OTU28667_Anaerolineae	0.577	0.04
OTU41838_Unidentified	0.577	0.026
OTU4851_Clostridia	0.577	0.023

OTU21195_Unidentified	0.577	0.032
OTU40841_Pla4 lineage	0.577	0.043
OTU10450_Babeliae	0.577	0.013
OTU22219_Omnitrophia	0.577	0.025
OTU31470_Subgroup 5	0.577	0.03
OTU43878_OM190	0.577	0.023
OTU28940_Phycisphaerae	0.577	0.032
OTU35408_Unidentified	0.577	0.013
OTU17754_Gammaproteobacteria	0.577	0.03
OTU17201_Gammaproteobacteria	0.57	0.04
OTU10224_Bacteroidia	0.551	0.044
OTU21208_Unidentified	0.55	0.041
OTU45465_Unidentified	0.55	0.036

Group Kotok #sps. 28

OTU11384_Alphaproteobacteria	0.882	0.001
OTU18361_Unidentified	0.869	0.004
OTU1237_Unidentified	0.816	0.001
OTU5767_Gammaproteobacteria	0.736	0.004
OTU15288_Unidentified	0.728	0.014
OTU2189_Unidentified	0.714	0.009
OTU21556_Bacteroidia	0.667	0.012
OTU26162_Oxyphotobacteria	0.661	0.012
OTU7022_Unidentified	0.657	0.05
OTU38044_Deltaproteobacteria	0.656	0.008
OTU28138_Unidentified	0.641	0.018
OTU13185_Deltaproteobacteria	0.623	0.025
OTU12536_Planctomycetacia	0.616	0.022
OTU18030_Verrucomicrobiae	0.606	0.038
OTU9325_Bacteroidia	0.577	0.02
OTU15976_Alphaproteobacteria	0.577	0.033
OTU33889_Mollicutes	0.577	0.021
OTU28195_Bacteroidia	0.577	0.023
OTU32973_Mollicutes	0.577	0.035
OTU12899_Bacteroidia	0.577	0.024
OTU17667_Unidentified	0.577	0.029
OTU18376_Deltaproteobacteria	0.577	0.03
OTU10643_Gammaproteobacteria	0.576	0.035
OTU18437_Gammaproteobacteria	0.568	0.03

OTU20848_Gammaproteobacteria	0.565	0.042
OTU15897_Unidentified	0.56	0.039
OTU35065_Unidentified	0.549	0.029
OTU15208_uncultured organism	0.542	0.041

Group LancangBesar #sps. 95

	stat p	.value
OTU17491_Gammaproteobacteria	0.816	0.002
OTU15169_Oxyphotobacteria	0.816	0.002
OTU33000_Phycisphaerae	0.816	0.002
OTU45655_Bacteroidia	0.816	0.002
OTU6736_Unidentified	0.816	0.002
OTU44196_Bacteroidia	0.816	0.002
OTU11431_Bacteroidia	0.816	0.002
OTU52948_Clostridia	0.816	0.002
OTU26616_Spirochaetia	0.816	0.002
OTU15762_Deltaproteobacteria	0.816	0.002
OTU53003_Unidentified	0.816	0.002
OTU47417_Unidentified	0.816	0.002
OTU5369_Kiritimatiellae	0.816	0.002
OTU27849_Campylobacteria	0.816	0.002
OTU50539_Bacteroidia	0.816	0.002
OTU26106_Clostridia	0.816	0.002
OTU20120_Clostridia	0.816	0.002
OTU35104_Anaerolineae	0.816	0.002
OTU52412_Bacteroidia	0.816	0.002
OTU23892_Unidentified	0.816	0.002
OTU38208_Clostridia	0.816	0.001
OTU12072_Deltaproteobacteria	0.814	0.002
OTU21955_Bacteroidia	0.813	0.002
OTU12862_OM190	0.813	0.004
OTU22402_Alphaproteobacteria	0.811	0.002
OTU47316_Bacteroidia	0.81	0.002
OTU11164_Fusobacteriia	0.804	0.002
OTU12803_Gammaproteobacteria	0.768	0.002
OTU34776_Gammaproteobacteria	0.765	0.004
OTU35946_Phycisphaerae	0.761	0.002
OTU32109_Deltaproteobacteria	0.761	0.002
OTU45710_Bacteroidia	0.756	0.005

OTU24237_Lentisphaeria	0.749	0.009
OTU25948_Kiritimatiellae	0.747	0.003
OTU13371_Anaerolineae	0.745	0.004
OTU32951_Anaerolineae	0.741	0.013
OTU42535_Unidentified	0.733	0.015
OTU44318_PAUC43f marine		0.7
OTU35621_Alphaproteobacteria	0.723	0.009
OTU46250_Alphaproteobacteria	0.72	0.011
OTU13373_Babeliae	0.707	0.002
OTU30239_Bacteroidia	0.695	0.004
OTU18487_Clostridia	0.694	0.004
OTU40007_Anaerolineae	0.694	0.005
OTU11756_Verrucomicrobiae	0.686	0.005
OTU12336_Bacteroidia	0.686	0.008
OTU47456_Verrucomicrobiae	0.68	0.012
OTU5041_Entotheonellia	0.667	0.011
OTU16817_Gammaproteobacteria	0.662	0.013
OTU2679_Bacteroidia	0.659	0.013
OTU51294_Verrucomicrobiae	0.657	0.007
OTU27447_Gammaproteobacteria	0.655	0.006
OTU11514_Verrucomicrobiae	0.645	0.01
OTU23160_vadinHA49	0.641	0.015
OTU18495_Bacteroidia	0.632	0.016
OTU25540_Planctomycetacia	0.631	0.031
OTU32622_Bacteroidia	0.627	0.021
OTU4669_Anaerolineae	0.626	0.019
OTU20981_Phycisphaerae	0.577	0.05
OTU5985_Planctomycetacia	0.577	0.05
OTU15102_Bacteroidia	0.577	0.05
OTU12083_Deltaproteobacteria	0.577	0.05
OTU19887_Bacteroidia	0.577	0.05
OTU47468_Babeliae	0.577	0.05
OTU11074_Deltaproteobacteria	0.577	0.05
OTU12931_Alphaproteobacteria	0.577	0.05
OTU21686_Parcubacteria	0.577	0.05
OTU23256_Deltaproteobacteria	0.577	0.05
OTU32856_Chlamydiae	0.577	0.05
OTU19738_uncultured bacteriu		0.5
OTU20591_Leptospirae	0.577	0.05
OTU2931_OM190	0.577	0.05

OTU4160_Deltaproteobacteria	0.577	0.05
OTU47339_Deltaproteobacteria	0.577	0.05
OTU10538_Unidentified	0.577	0.05
OTU13700_Unidentified	0.577	0.05
OTU18700_ABY1	0.577	0.05
OTU31015_Unidentified	0.577	0.05
OTU35582_Unidentified	0.577	0.05
OTU45993_Lentisphaeria	0.577	0.05
OTU15264_Deltaproteobacteria	0.577	0.05
OTU20658_Unidentified	0.577	0.05
OTU40693_Gammaproteobacteria	0.577	0.05
OTU4498_uncultured bacterium	0.577	0.05
OTU2020_Alphaproteobacteria	0.574	0.05
OTU39928_Gracilibacteria	0.566	0.043
OTU13410_Chlamydiae	0.559	0.05
OTU46558_Phycisphaerae	0.555	0.042
OTU12411_Anaerolineae	0.554	0.047
OTU33373_Deltaproteobacteria	0.548	0.046
OTU33388_Anaerolineae	0.535	0.048
OTU19566_Spirochaetia	0.535	0.048
OTU18711_Lineage IIb	0.535	0.05
OTU6316_Mollicutes	0.53	0.05
OTU33980_Deltaproteobacteria	0.514	0.049

Group Pari #sps. 68

	stat p	.value
OTU22542_Gammaproteobacteria	0.855	0.001
OTU48956_Gammaproteobacteria	0.816	0.001
OTU27087_Gammaproteobacteria	0.816	0.001
OTU44626_Gammaproteobacteria	0.812	0.001
OTU31170_Deltaproteobacteria	0.798	0.001
OTU26071_Gammaproteobacteria	0.77	0.003
OTU5777_BD2-11 terrestrial g	0	0.707
OTU35726_Alphaproteobacteria	0.707	0.004
OTU33063_Dehalococcoidia	0.707	0.005
OTU13090_TK17	0.707	0.005
OTU30311_Verrucomicrobiae	0.707	0.005
OTU8492_Deltaproteobacteria	0.707	0.005
OTU44779_Deltaproteobacteria	0.707	0.005
OTU16238_Verrucomicrobiae	0.706	0.006

OTU1632_Bacteroidia	0.697	0.034
OTU38393_Phycisphaerae	0.691	0.004
OTU24459_Unidentified	0.688	0.01
OTU17574_Bacteroidia	0.673	0.007
OTU30781_Gammaproteobacteria	0.657	0.01
OTU46027_Clostridia	0.656	0.016
OTU29018_Gammaproteobacteria	0.653	0.029
OTU45327_Phycisphaerae	0.65	0.017
OTU41381_Alphaproteobacteria	0.647	0.016
OTU21452_Bacteroidia	0.623	0.014
OTU2277_Unidentified	0.577	0.011
OTU14168_Bacteroidia	0.577	0.011
OTU545_Alphaproteobacteria	0.577	0.012
OTU10208_Deltaproteobacteria	0.577	0.011
OTU26345_Unidentified	0.577	0.011
OTU13549_Unidentified	0.577	0.011
OTU17040_Unidentified	0.577	0.011
OTU20010_Campylobacteria	0.577	0.011
OTU15370_Spirochaetia	0.577	0.012
OTU22318_Thermoleophilia	0.577	0.012
OTU44972_Unidentified	0.577	0.011
OTU11406_Anaerolineae	0.577	0.012
OTU26504_Anaerolineae	0.577	0.012
OTU23216_Mollicutes	0.577	0.012
OTU4517_Anaerolineae	0.577	0.011
OTU9917_Deltaproteobacteria	0.577	0.011
OTU28857_Unidentified	0.577	0.012
OTU32932_Deltaproteobacteria	0.577	0.012
OTU37938_Unidentified	0.577	0.012
OTU46404_Unidentified	0.577	0.012
OTU47638_Leptospirae	0.577	0.013
OTU10884_Unidentified	0.577	0.011
OTU23167_Bacteroidia	0.577	0.011
OTU270_uncultured bacterium	0.577	0.011
OTU3567_Unidentified	0.577	0.02
OTU35571_Unidentified	0.577	0.012
OTU9145_LD1-PA32	0.577	0.012
OTU11774_Melainabacteria	0.577	0.012
OTU23729_uncultured organism	0.577	0.021
OTU24799_Alphaproteobacteria	0.577	0.011

OTU11829_Deltaproteobacteria	0.577	0.012
OTU14073_Deltaproteobacteria	0.577	0.012
OTU22239_Deltaproteobacteria	0.577	0.011
OTU48369_Spirochaetia	0.568	0.034
OTU2600_Fibrobacteria	0.567	0.041
OTU33158_Unidentified	0.567	0.032
OTU3776_Unidentified	0.564	0.034
OTU38900_Clostridia	0.557	0.033
OTU1205_Unidentified	0.533	0.03
OTU33045_Anaerolineae	0.524	0.026
OTU43057_vadinHA49	0.522	0.043
OTU28095_vadinHA49	0.521	0.049
OTU39781_Unidentified	0.52	0.046
OTU23987_Mollicutes	0.518	0.049

Group Pramuka #sps. 15

	stat p	.value
OTU9991_Unidentified	0.745	0.01
OTU34901_Gammaproteobacteria	0.715	0.021
OTU12132_Unidentified	0.667	0.016
OTU37259_Unidentified	0.667	0.016
OTU44401_Chitinivibrionia	0.667	0.013
OTU46838_Unidentified	0.636	0.015
OTU5689_Phycisphaerae	0.635	0.02
OTU31105_Deltaproteobacteria	0.577	0.024
OTU4043_Planctomycetacia	0.577	0.02
OTU10429_vadinHA49	0.577	0.02
OTU10263_Unidentified	0.577	0.018
OTU45164_OM190	0.577	0.027
OTU10424_Unidentified	0.576	0.047
OTU12727_Spirochaetia	0.56	0.036
OTU7190_Unidentified	0.555	0.049

Group Sepa #sps. 47

	stat	p.value
OTU20047_uncultured bacterium	0.86	0.0
OTU49036_Unidentified	0.796	0.002
OTU1097_Verrucomicrobiae	0.766	0.003
OTU42950_Spirochaetia	0.755	0.003
OTU1318_Oxyphotobacteria	0.707	0.008

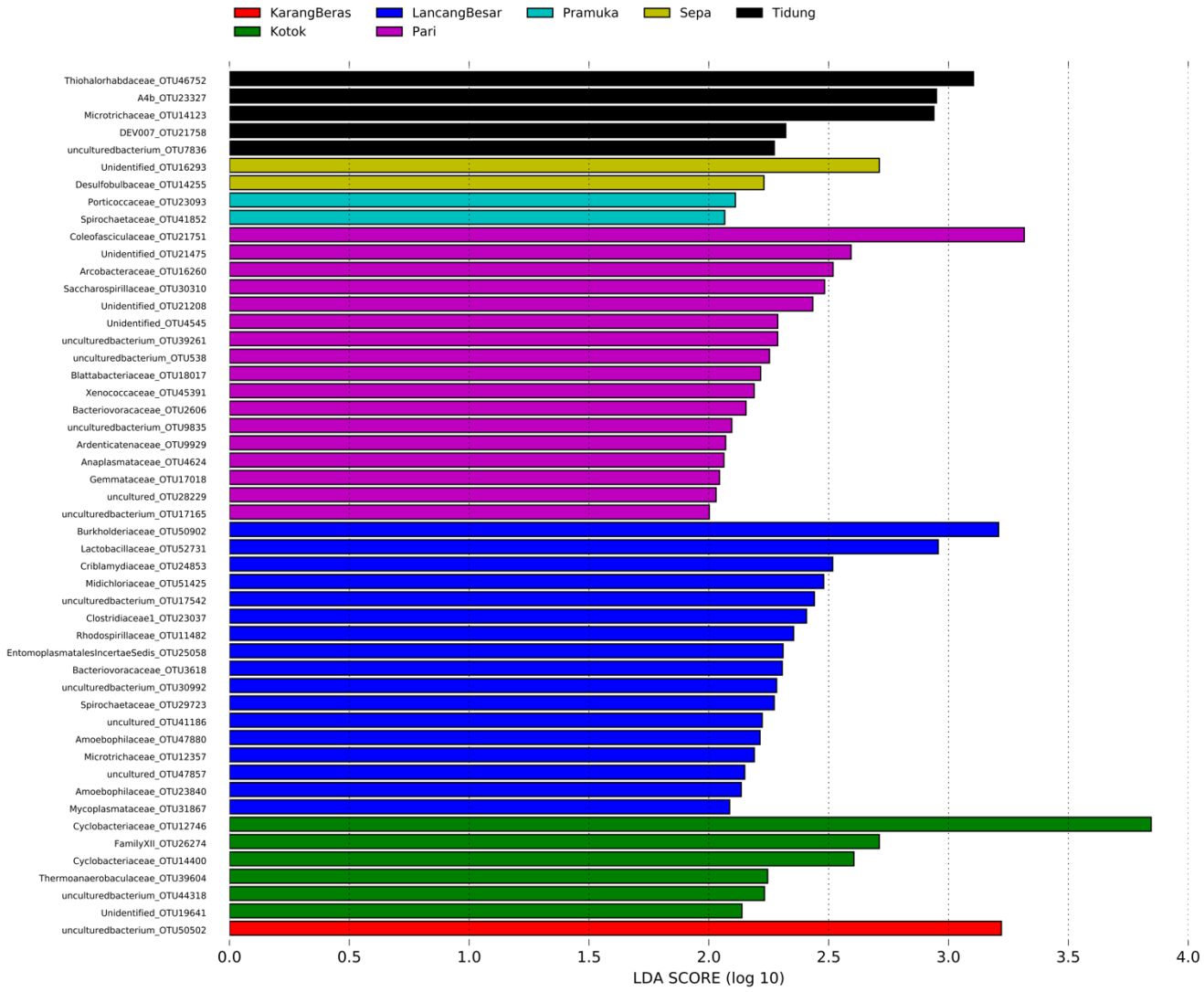
OTU42633_Deltaproteobacteria	0.707	0.008
OTU43401_Deltaproteobacteria	0.707	0.005
OTU44736_Bacteroidia	0.707	0.007
OTU1053_Phycisphaerae	0.707	0.007
OTU23079_Unidentified	0.707	0.004
OTU30611_Deltaproteobacteria	0.693	0.002
OTU23485_Subgroup 17	0.681	0.018
OTU10714_Oxyphotobacteria	0.68	0.016
OTU25220_Unidentified	0.671	0.03
OTU3681_Bacteroidia	0.669	0.013
OTU16970_Oxyphotobacteria	0.631	0.015
OTU34997_TK17	0.612	0.006
OTU24110_Deltaproteobacteria	0.612	0.008
OTU10383_Gammaproteobacteria	0.612	0.011
OTU18732_Bacteroidia	0.612	0.003
OTU23380_Bacteroidia	0.612	0.01
OTU36365_Spirochaetia	0.612	0.015
OTU14790_Deltaproteobacteria	0.612	0.006
OTU16967_Deltaproteobacteria	0.612	0.008
OTU47823_Anaerolineae	0.612	0.016
OTU42363_Ignavibacteria	0.612	0.01
OTU15219_uncultured bacterium	0.61	2 0.0
OTU52474_Anaerolineae	0.612	0.007
OTU51215_OM190	0.612	0.011
OTU21793_Unidentified	0.612	0.011
OTU36322_Anaerolineae	0.598	0.018
OTU51989_Gammaproteobacteria	0.595	0.05
OTU24056_Phycisphaerae	0.593	0.016
OTU18934_Deltaproteobacteria	0.591	0.019
OTU14876_Microgenomatia	0.588	0.035
OTU24727_Alphaproteobacteria	0.587	0.019
OTU44081_Anaerolineae	0.582	0.017
OTU1909_Clostridia	0.577	0.022
OTU3159_Phycisphaerae	0.577	0.021
OTU26910_Fibrobacteria	0.569	0.03
OTU41413_Actinobacteria	0.568	0.032
OTU15740_Gammaproteobacteria	0.557	0.017
OTU41852_Spirochaetia	0.548	0.033
OTU17114_Unidentified	0.547	0.047
OTU48170_BD7-11	0.538	0.04

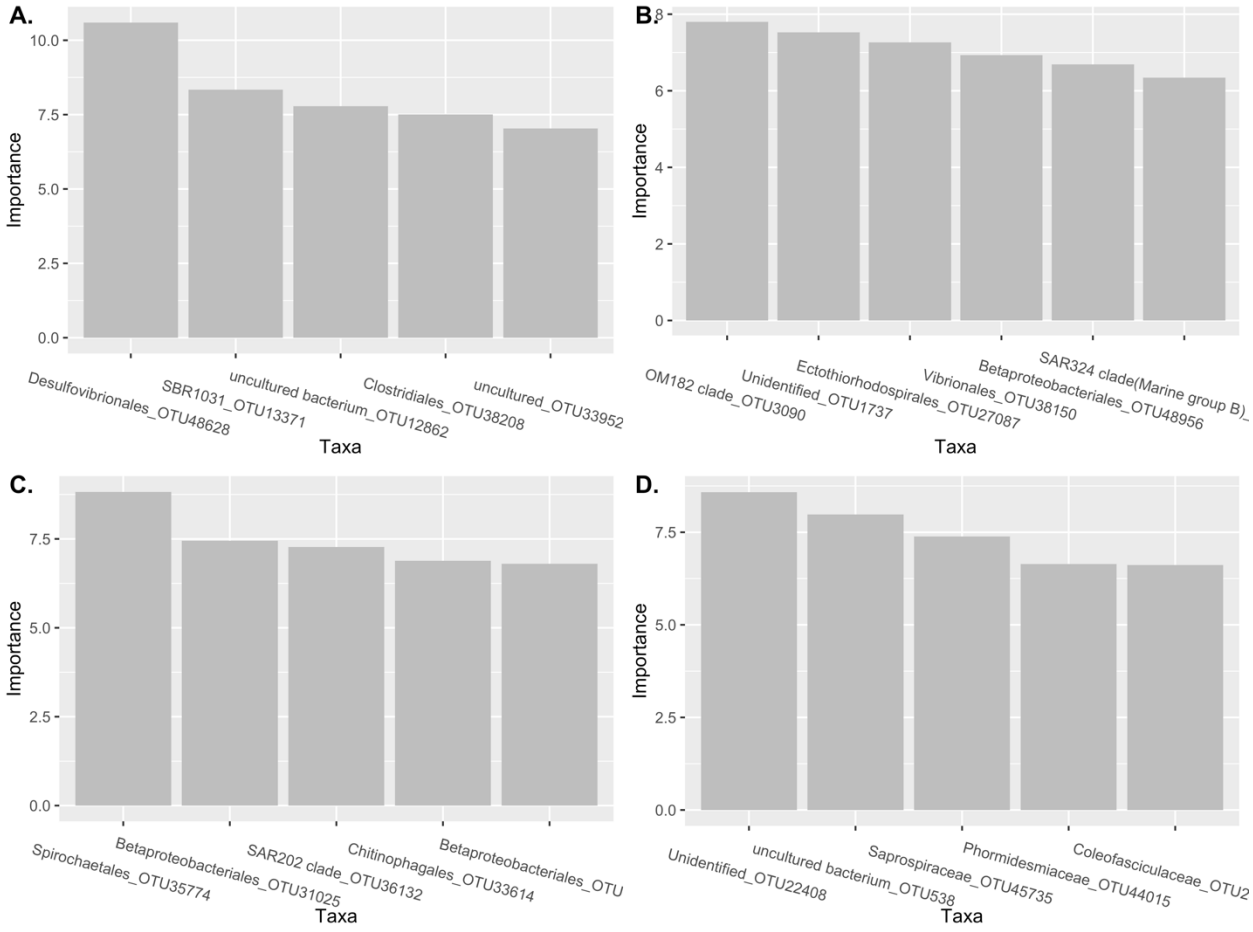
OTU6987_OM190	0.536	0.033
OTU12526_Thermoanaerobaculia	0.529	0.045

Group Tidung #sps. 32

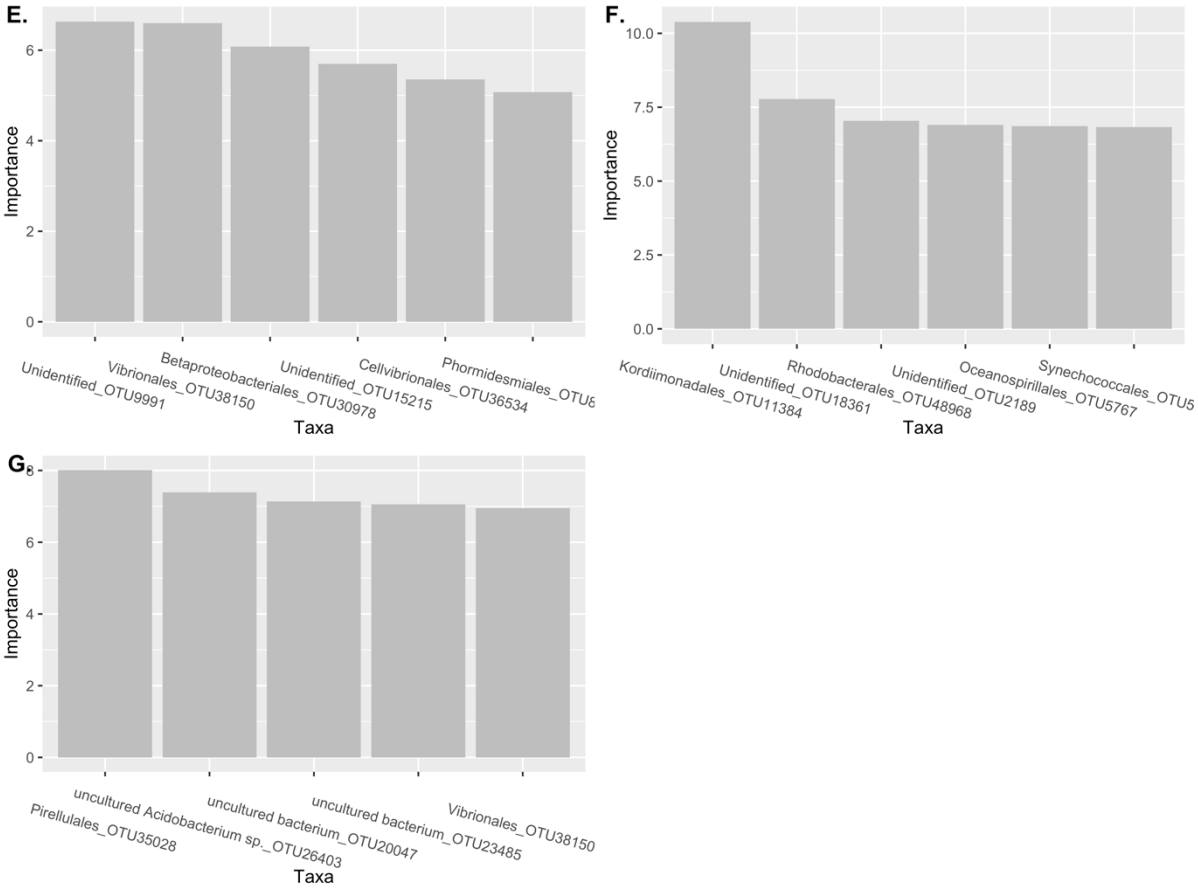
	stat p	.value
OTU35774_Spirochaetia	0.852	0.001
OTU31025_Gammaproteobacteria	0.828	0.001
OTU33614_Bacteroidia	0.808	0.003
OTU37853_Gammaproteobacteria	0.731	0.031
OTU16435_Unidentified	0.729	0.018
OTU51004_Acidimicrobiia	0.726	0.014
OTU41080_Unidentified	0.705	0.009
OTU26555_Deltaproteobacteria	0.68	0.005
OTU26671_Gammaproteobacteria	0.667	0.021
OTU25900_Alphaproteobacteria	0.667	0.01
OTU20402_Deltaproteobacteria	0.667	0.011
OTU12257_Phycisphaerae	0.667	0.015
OTU27747_Alphaproteobacteria	0.663	0.029
OTU39731_Deltaproteobacteria	0.65	0.011
OTU34959_Subgroup 22	0.645	0.034
OTU44849_OM190	0.638	0.02
OTU19517_Verrucomicrobiae	0.626	0.028
OTU46070_Gammaproteobacteria	0.603	0.033
OTU43860_Bacteroidia	0.577	0.029
OTU18717_Bacteroidia	0.577	0.026
OTU2364_Phycisphaerae	0.577	0.033
OTU43318_Spirochaetia	0.575	0.041
OTU23553_Unidentified	0.567	0.023
OTU2092_Verrucomicrobiae	0.564	0.037
OTU30005_Blastocatellia		0.56
OTU44576_Bacteroidia	0.559	0.047
OTU19688_Deltaproteobacteria	0.558	0.033
OTU16873_Deltaproteobacteria	0.556	0.037
OTU17962_Elusimicrobia	0.555	0.049
OTU17863_Phycisphaerae	0.554	0.036
OTU51686_Gammaproteobacteria	0.553	0.025
OTU47995_Unidentified	0.545	0.049

Supplementary Figure S3-1. Histogram of the linear discriminant analysis (LDA) scores for differentially abundant bacterial taxa across sites. Both Pari and Lancang Besar has the highest number of bacterial taxa associated with the sites





Supplementary Figure S3-2. Random forest predictions for taxa that were associated with deployments site. Only maximum six taxa that have maximum importance value were presented in the plot. (A. Lancang Besar, B. Pari, C. Tidung, D. Karang Beras, E. Pramuka, F. Kotok, G. Sepa)



Supplementary Figure S3-2. (continued)

References

Abdelzaher AM, Wright ME, Ortega C, Solo-Gabriele HM, Miller G, Elmir S, Newman X, Shih P, Bonilla JA, Bonilla TD, Palmer CJ, Scott T, Lukasik J, Harwood VJ, McQuaig S, Sinigalliano C, Gidley M, Plano LRW, Zhu X, Wang JD, Fleming LE (2010) Presence of pathogens and indicator microbes at a non-point source subtropical recreational marine beach. Applied and environmental microbiology 76:724–32

Abed RMM, Safi NMD, Köster J, Beer D De, El-Nahhal Y, Rullkötter J, Garcia-Pichel F (2002) Microbial diversity of a heavily polluted microbial mat and its community changes following degradation of petroleum compounds. Applied and Environmental Microbiology

Aguiló-Ferretjans MM, Bosch R, Martín-Cardona C, Lalucat J, Nogales B (2008) Phylogenetic analysis of the composition of bacterial communities in human-exploited coastal environments from Mallorca Island (Spain). Systematic and Applied Microbiology

Ainsworth TD, Krause L, Bridge T, Torda G, Raina JB, Zakrzewski M, Gates RD, Padilla-Gamiño JL, Spalding HL, Smith C, Woolsey ES, Bourne DG, Bongaerts P, Hoegh-Guldberg O, Leggat W (2015) The coral core microbiome identifies rare bacterial taxa as ubiquitous endosymbionts. ISME Journal

Al-Rshaidat MMD, Snider A, Rosebraugh S, Devine AM, Devine TD, Plaisance L, Knowlton N, Leray M, Al-Rshaidat MMD, Snider A, Rosebraugh S, Devine AM, Devine TD, Plaisance L, Knowlton N, Leray M (2016) Deep COI sequencing of standardized benthic samples unveils overlooked diversity of Jordanian coral reefs in the northern Red Sea 1. Genome 59:724–737

Allen GR (2008a) Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. Aquatic Conservation: Marine and Freshwater Ecosystems 18:541–556

Allen GR (2008b) Conservation hotspots of biodiversity and endemism for Indo-Pacific coral reef fishes. Aquatic Conservation: Marine and Freshwater Ecosystems 18:541–556

Allen GR, Adrim M (2003) Review: Article Coral Reef Fishes of Indonesia. Zoological Studies 42:1–72

Allen GR, Erdmann M V. (2009) Reef fishes of the Bird's Head Peninsula, West Papua, Indonesia. Check List 5:587–628

Araújo-Pérez F, McCoy AN, Okechukwu C, Carroll IM, Smith KM, Jeremiah K, Sandler RS, Asher GN, Keku TO (2012) Differences in microbial signatures between rectal mucosal biopsies and rectal swabs. Gut Microbes 3:530–535

Bachoon DS, Markand S, Otero E, Perry G, Ramsubaugh A (2010) Assessment of non-point sources of fecal pollution in coastal waters of Puerto Rico and Trinidad. Marine Pollution Bulletin

Bahl MI, Bergström A, Licht TR (2012) Freezing fecal samples prior to DNA extraction affects the Firmicutes to Bacteroidetes ratio determined by downstream quantitative PCR analysis. FEMS Microbiology Letters 329:193–197

Barott KL, Rohwer FL (2012) Unseen players shape benthic competition on coral reefs. Elsevier Current Trends

Baum G, Januar HI, Ferse SCA, Kunzmann A (2015) Local and Regional Impacts of Pollution on Coral Reefs along the Thousand Islands North of the Megacity Jakarta, Indonesia. PloS one 10:e0138271

Blackall LL, Stratton H, Bradford D, Dot T Del, Sjörup C, Seviour EM, Seviour RJ (1996) "Candidatus microthrix parvicella," a filamentous bacterium from activated sludge sewage treatment plants. International Journal of Systematic Bacteriology

Blackall LL, Wilson B, Oppen MJH Van (2015) Coral-the world's most diverse symbiotic ecosystem. Molecular Ecology 24:5330–5347

Bouchet P, Lozouet P, Maestrati P, Heros V (2002) Assessing the magnitude of species richness in tropical marine environments: Exceptionally high numbers of molluscs at a New Caledonia site. Biological Journal of the Linnean Society 75:421–436

Boyer JN, Kelble CR, Ortner PB, Rudnick DT (2009) Phytoplankton bloom status: Chlorophyll a biomass as an indicator of water quality condition in the southern estuaries of Florida, USA. Ecological Indicators 9:S56–S67

Brakstad OG, Lødeng AGG Microbial Diversity during Biodegradation of Crude Oil in Seawater from the North Sea. Microbial Ecology 49:94–103

Bruno JF, Petes LE, Harvell CD, Hettinger A (2003) Nutrient enrichment can increase the severity of coral diseases. Ecology Letters 6:1056–1061

Bruno JF, Selig ER (2007) Regional decline of coral cover in the Indo-Pacific: timing, extent, and subregional comparisons. PloS one 2:e711

Butnariu M (2022) Aquatic pollution and marine ecosystems. Bacterial Fish Diseases:1–29

Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, Bealer K, Madden TL (2009) BLAST+: Architecture and applications. BMC Bioinformatics 10:421

Cao H, Hong Y, Li M, Gu JD (2012) Community shift of ammonia-oxidizing bacteria along an anthropogenic pollution gradient from the Pearl River Delta to the South China Sea. Applied Microbiology and Biotechnology 94:247–259

Caporaso JG, Lauber CL, Walters WA, Berg-Lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, Gormley N, Gilbert JA, Smith G, Knight R (2012) Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. The

ISME journal 6:1621–4

Carpenter KE, Abrar M, Aeby G, Aronson RB, Banks S, Bruckner A, Chiriboga A, Cortés J, Delbeek JC, DeVantier L, Edgar GJ, Edwards AJ, Fenner D, Guzmán HM, Hoeksema BW, Hodgson G, Johan O, Licuanan WY, Livingstone SR, Lovell ER, Moore JA, Obura DO, Ochavillo D, Polidoro BA, Precht WF, Quibilan MC, Reboton C, Richards ZT, Rogers AD, Sanciangco J, Sheppard A, Sheppard C, Smith J, Stuart S, Turak E, Veron JEN, Wallace C, Weil E, Wood E, Cortes J, Delbeek JC, DeVantier L, Edgar GJ, Edwards AJ, Fenner D, Guzman HM, Hoeksema BW, Hodgson G, Johan O, Licuanan WY, Livingstone SR, Lovell ER, Moore JA, Obura DO, Ochavillo D, Polidoro BA, Precht WF, Quibilan MC, Reboton C, Richards ZT, Rogers AD, Sanciangco J, Sheppard A, Sheppard C, Smith J, Stuart S, Turak E, Veron JEN, Wallace C, Weil E, Wood E (2008) One-third of reef-building corals face elevated extinction risk from climate change and local impacts. Science (New York, NY) 321:560–563

Casey JM, Ransome E, Collins AG, Mahardini A, Kurniasih EM, Sembiring A, Schiettekatte NMD, Cahyani NKD, Wahyu Anggoro A, Moore M, Uehling A, Belcaid M, Barber PH, Geller JB, Meyer CP (2021) DNA metabarcoding marker choice skews perception of marine eukaryotic biodiversity. Environmental DNA 3:1229–1246

Catania V, Cappello S, Giorgi V Di, Santisi S, Maria R Di, Mazzola A, Vizzini S, Quatrini P (2018) Microbial communities of polluted sub-surface marine sediments. Marine Pollution Bulletin 131:396–406

Cebrian E, Uriz MJ, Turon X (2007) Sponges as biomonitors of heavy metals in spatial and temporal surveys in northwestern Mediterranean: Multispecies comparison. Environmental Toxicology and Chemistry

Cesar H, Burke L, Pet-soede L (2003) The Economics of Worldwide Coral Reef Degradation.

Chen J, McIlroy SE, Archana A, Baker DM, Panagiotou G (2019) A pollution gradient contributes to the taxonomic, functional, and resistome diversity of microbial communities in marine sediments. Microbiome 7:104

Chiarello M, Villéger S, Bouvier C, Bettarel Y, Bouvier T (2015) High diversity of skin-associated bacterial communities of marine fishes is promoted by their high variability among body parts, individuals and species. FEMS Microbiology Ecology 91

Cho JC, Kim SJ (2000) Increase in bacterial community diversity in subsurface aquifers receiving livestock wastewater input. Applied and Environmental Microbiology

Cleary DFR (2017) Linking fish species traits to environmental conditions in the Jakarta Bay-Pulau Seribu coral reef system. Marine Pollution Bulletin 122:259–262

Cleary DFR, Suharsono, Hoeksema BW (2006) Coral diversity across a disturbance gradient in the Pulau Seribu reef complex off Jakarta, Indonesia. Biodiversity and Conservation 15:3653–3674

Cleary DFR, Vantier L De, Giyanto, Vail L, Manto P, Voogd NJ de, Rachello-Dolmen PG, Tuti Y, Budiyanto A, Wolstenholme J, Hoeksema BW, Suharsono (2008) Relating variation in species composition to environmental variables: a multi-taxon study in an Indonesian coral reef complex. Aquatic Sciences 70:419–431

Corrigan A, Russell N, Welge M, Auvil L, Bushell C, White BA, Murphy RA (2018) The use of random forests modelling to detect yeast-mannan sensitive bacterial changes in the broiler cecum. Scientific Reports 8:13270

Delgado-Baquerizo M, Grinyer J, Reich PB, Singh BK (2016) Relative importance of soil properties and microbial community for soil functionality: insights from a microbial swap

experiment (E Allen, Ed.). Functional Ecology 30:1862–1873

DeVantier L, Suharsono, Budiyanto A, Tuti Y, Imanto P LR (1998) Status of coral communities of Pulau Seribu (Indonesia). In: S S (ed) Proceedings of coral reef evaluation workshop, Pulau Seribu, Jakarta, Indonesia, 1995. UNESCO, Jakarta, Jakarta, p 1–24

Ding Q, Huang X, Hu H, Hong M, Zhang D, Wang K (2017) Impact of pyrene and cadmium co-contamination on prokaryotic community in coastal sediment microcosms. Chemosphere 188:320–328

Done TJ (1992) Phase shifts in coral reef communities and their ecological significance. Hydrobiologia 247:121–132

Dubinksy Z, Stambler N (1996) Marine pollution and coral reefs. Global Change Biology 2:511–526

Elshafie HS, Camele I (2021) An Overview of Metabolic Activity, Beneficial and Pathogenic Aspects of Burkholderia Spp. Metabolites 2021, Vol 11, Page 321 11:321

Enns AA, Vogel LJ, Abdelzaher AM, Solo-Gabriele HM, Plano LRW, Gidley ML, Phillips MC, Klaus JS, Piggot AM, Feng Z, Reniers AJHM, Haus BK, Elmir SM, Zhang Y, Jimenez NH, Abdel-Mottaleb N, Schoor ME, Brown A, Khan SQ, Dameron AS, Salazar NC, Fleming LE (2012) Spatial and temporal variation in indicator microbe sampling is influential in beach management decisions. Water Research 46:2237–2246

Enomoto M, Nakagawa S, Sawabe T (2012) Microbial Communities Associated with Holothurians: Presence of Unique Bacteria in the Coelomic Fluid. Microbes and Environments 27:300–305

Fabricius K, De'ath G, McCook L, Turak E, Williams DMB (2005) Changes in algal, coral and fish assemblages along water quality gradients on the inshore Great Barrier Reef. Marine

Pollution Bulletin 51:384–398

Fay M, Risher JF WJ (2007) Toxicological Profile for Xylene. Atlanta

Foale S, Adhuri D, Aliño P, Allison EH, Andrew N, Cohen P, Evans L, Fabinyi M, Fidelman P, Gregory C, Stacey N, Tanzer J, Weeratunge N (2013) Food security and the Coral Triangle Initiative. Marine Policy 38:174–183

Frade PR, Glasl B, Matthews SA, Mellin C, Serrão EA, Wolfe K, Mumby PJ, Webster NS, Bourne DG (2020) Spatial patterns of microbial communities across surface waters of the Great Barrier Reef. Communications Biology 2020 3:1 3:1–14

Frøslev TG, Kjøller R, Bruun HH, Ejrnæs R, Brunbjerg AK, Pietroni C, Hansen AJ (2017) Algorithm for post-clustering curation of DNA amplicon data yields reliable biodiversity estimates. Nature Communications

Fuhrman JA, Hewson I, Schwalbach MS, Steele JA, Brown M V., Naeem S (2006) Annually reoccurring bacterial communities are predictable from ocean conditions. Proceedings of the National Academy of Sciences

Fuhrman JA, Steele JA, Hewson I, Schwalbach MS, Brown M V., Green JL, Brown JH (2008) A latitudinal diversity gradient in planktonic marine bacteria. Proceedings of the National Academy of Sciences

Gilbert A, James D (1994) Water Pollution in Jakarta Bay. In: Springer, Dordrecht, p 111–142

Gillan DC, Danis B, Pernet P, Joly G, Dubois P (2005) Structure of sediment-associated microbial communities along a heavy-metal contamination gradient in the marine environment. Applied and environmental microbiology 71:679–90

Glasl B, Bourne DG, Frade PR, Thomas T, Schaffelke B, Webster NS (2019) Microbial indicators of environmental perturbations in coral reef ecosystems. Microbiome 7:1–13

Glasl B, Bourne DG, Frade PR, Webster NS (2018) Establishing microbial baselines to identify indicators of coral reef health. Microbiology Australia

Glasl B, Herndl GJ, Frade PR (2016) The microbiome of coral surface mucus has a key role in mediating holobiont health and survival upon disturbance. Isme J:1–13

González JM, Covert JS, Whitman WB, Henriksen JR, Mayer F, Scharf B, Schmitt R, Buchan A, Fuhrman JA, Kiene RP, Moran MA (2003) Silicibacter pomeroyi sp. nov. and Roseovarius nubinhibens sp. nov., dimethylsulfoniopropionate-demethylating bacteria from marine environments. International Journal of Systematic and Evolutionary Microbiology 53:1261–1269

Green AL, Bellwood DR Monitoring Functional Groups of Herbivorous Reef Fishes as Indicators of Coral Reef Resilience A practical guide for coral reef managers in the Asia Pacifi c Region.

Halliday E, McLellan SL, Amaral-Zettler LA, Sogin ML, Gast RJ (2014) Comparison of Bacterial Communities in Sands and Water at Beaches with Bacterial Water Quality Violations (Z Zhou, Ed.). PLoS ONE 9:e90815

Harder T, Lau SC., Dobretsov S, Fang TK, Qian P-Y (2003) A distinctive epibiotic bacterial community on the soft coral Dendronephthya sp. and antibacterial activity of coral tissue extracts suggest a chemical mechanism against bacterial epibiosis. FEMS Microbiology Ecology 43:337–347

Helber SB, Hoeijmakers DJJ, Muhando CA, Rohde S, Schupp PJ (2018) Sponge chemical defenses are a possible mechanism for increasing sponge abundance on reefs in Zanzibar. PloS one 13:e0197617

Hernandez-Agreda A, Leggat W, Bongaerts P, Herrera C, Ainsworth TD (2018) Rethinking the

Coral Microbiome: Simplicity Exists within a Diverse Microbial Biosphere. mBio 9:e00812-18

Hester ER, Barott KL, Nulton J, Vermeij MJ, Rohwer FL (2016) Stable and sporadic symbiotic communities of coral and algal holobionts. The ISME Journal 10:1157–1169

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 (New York, NY) 318:1737–42

Horner-Devine MC, Leibold MA, Smith VH, Bohannan BJM, Claire Horner-Devine M, Leibold MA, Smith VH, Bohannan BJM (2003) Bacterial diversity patterns along a gradient of primary productivity. Ecology Letters 6:613–622

Hosono T, Su CC, Delinom R, Umezawa Y, Toyota T, Kaneko S, Taniguchi M (2011) Decline in heavy metal contamination in marine sediments in Jakarta Bay, Indonesia due to increasing environmental regulations. Estuarine, Coastal and Shelf Science

Huang S, Jaffé PR (2019) Defluorination of Perfluorooctanoic Acid (PFOA) and Perfluorooctane Sulfonate (PFOS) by Acidimicrobium sp. Strain A6. Environmental Science and Technology

Hubert N, Paradis E, Bruggemann H, Planes S (2011) Community assembly and diversification in Indo-Pacific coral reef fishes. Ecology and evolution 1:229–77

Hugerth LW, Andersson AF (2017) Analysing microbial community composition through amplicon sequencing: From sampling to hypothesis testing. Frontiers in Microbiology

Hughes TP, Baird AH, Bellwood DR, Card M, Connolly SR, Folke C, Grosberg R, Hoegh-Guldberg O, Jackson JBC, Kleypas J, Lough JM, Marshall P, Nyström M, Palumbi SR,

Pandolfi JM, Rosen B, Roughgarden J (2003) Climate change, human impacts, and the resilience of coral reefs. Science (New York, NY) 301:929–33

Hutomo M, Moosa MK (2005) Indonesian marine and coastal biodiversity : Present status. Indian Journal of Marine Sciences 34:88–97

Ip YCA, Chang JJM, Oh RM, Quek ZBR, Chan YKS, Bauman AG, Huang D (2022) Seq' and ARMS shall find: DNA (meta)barcoding of Autonomous Reef Monitoring Structures across the tree of life uncovers hidden cryptobiome of tropical urban coral reefs. Molecular Ecology

Irnidayanti Y (2015) Toxicity and traces of Hg, Pb and Cd in the hepatopancreas, gills and muscles of Perna Viridis from Jakarta bay, Indonesia. Pakistan Journal of Biological Sciences

Iwaki H, Yamamoto T, Hasegawa Y (2018) Isolation of marine xylene-utilizing bacteria and characterization of Halioxenophilus aromaticivorans gen. nov., sp. nov. and its xylene degradation gene cluster. FEMS Microbiology Letters 365

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

Jenkins CN, Houtan KS Van (2016) Global and regional priorities for marine biodiversity protection. Biological Conservation 204:333–339

Jessen C, Villa Lizcano JF, Bayer T, Roder C, Aranda M, Wild C, Voolstra CR (2013) In-situ Effects of Eutrophication and Overfishing on Physiology and Bacterial Diversity of the Red Sea Coral Acropora hemprichii (JA Gilbert, Ed.). PLoS ONE 8:e62091

Jose J, Giridhar R, Anas A, Loka Bharathi PA, Nair S (2011) Heavy metal pollution exerts reduction/adaptation in the diversity and enzyme expression profile of heterotrophic bacteria in Cochin estuary, India. Environmental Pollution 159:2775–2780

Kaewpan A, Duangurai T, Rungruengkitkun A, Muangkaew W, Kanjanapruthipong T, Jitprasutwit N, Ampawong S, Sukphopetch P, Chantratita N, Pumirat P (2022) Burkholderia pseudomallei pathogenesis in human skin fibroblasts: A Bsa type III secretion system is involved in the invasion, multinucleated giant cell formation, and cellular damage. PLOS ONE 17:e0261961

Kandlikar G (2020) ranacapa: Utility Functions and "shiny" App for Simple Environmental DNA Visualizations and Analyses.

Knowlton N (2001) The future of coral reefs. Proceedings of the National Academy of Sciences of the United States of America 98:5419–25

Kondo R, Shigematsu K, Kawahara N, Okamura T, Yoon YH, Sakami T, Yokoyama H, Koizumi Y (2012) Abundance of sulphate-reducing bacteria in fish farm sediments along the coast of Japan and South Korea. Fisheries Science

Kramar MK, Tinta T, Lučić D, Malej A, Turk V (2019) Bacteria associated with moon jellyfish during bloom and post-bloom periods in the Gulf of Trieste (northern Adriatic). PLOS ONE 14:e0198056

Krediet CJ, Ritchie KB, Paul VJ, Teplitski M (2013) Coral-associated micro-organisms and their roles in promoting coral health and thwarting diseases. Proceedings Biological sciences / The Royal Society 280:20122328

Kuffner IBIB, Walters LJ, Becerro MA, Paul VJ, Ritson-Williams R, Beach KS (2006) Inhibition of coral recruitment by macroalgae and cyanobacteria. Marine Ecology Progress

Series 323:107–117

Lema KA, Willis BL, Bourneb DG (2012) Corals form characteristic associations with symbiotic nitrogen-fixing bacteria. Applied and Environmental Microbiology 78:3136–3144

Leray M, Knowlton N (2015) DNA barcoding and metabarcoding of standardized samples reveal patterns of marine benthic diversity. Proceedings of the National Academy of Sciences 112:201424997

Leray M, Knowlton N (2016) Visualizing Patterns of Marine Eukaryotic Diversity from Metabarcoding Data Using QIIME. In: Humana Press, New York, NY, p 219–235

Lesser MP, Mazel CH, Gorbunov MY, Falkowski PG (2004) Discovery of symbiotic nitrogen-fixing cyanobacteria in corals. Science (New York, NY) 305:997–1000

Lindström ES, Langenheder S (2012) Local and regional factors influencing bacterial community assembly. Environmental Microbiology Reports

Logue JB, Findlay SEG, Comte J (2015) Editorial: Microbial Responses to Environmental Changes. Frontiers in microbiology 6:1364

Lozupone CA, Knight R (2007) Global patterns in bacterial diversity. Proceedings of the National Academy of Sciences 104:11436–11440

Lubis A, Yosi M (2012) CONDITIONS OF MARITIME METEOROLOGY AND OCEANOGRAPHY IN THE VICINITY OF KOTOK ISLAND, SERIBU ISLANDS: APRIL 2011. Jurnal Ilmu dan Teknologi Kelautan Tropis 4:24

Madduppa HH, Beginner S, Suparyani E, Siregar AM, Arafat D, TARIGAN SA, ALIMUDDIN A, KHAIRUDI D, RAHMAWATI F, BRAHMANDITO A (2013) Dynamics of fish diversity across an environmental gradient in the Seribu Islands reefs off Jakarta. Biodiversitas Journal of Biological Diversity 14:17–24

Maeda Y, Shinohara H, Kiba A, Ohnishi K, Furuya N, Kawamura Y, Ezaki T, Vandamme P, Tsushima S, Hikichi Y (2006) Phylogenetic study and multiplex PCR-based detection of Burkholderia plantarii, Burkholderia glumae and Burkholderia gladioli using gyrB and rpoD sequences. International Journal of Systematic and Evolutionary Microbiology 56:1031–1038

Mao-Jones J, Ritchie KB, Jones LE, Ellner SP (2010) How microbial community composition regulates coral disease development. PLoS Biology 8

McDevitt-Irwin JM, Baum JK, Garren M, Vega Thurber RL (2017) Responses of Coral-Associated Bacterial Communities to Local and Global Stressors. Frontiers in Marine Science 4:262

Mcmanus JW, Reyes RB, Nañola CL (1997) Effects of some destructive fishing methods on coral cover and potential rates of recovery. Environmental Management 21:69–78

McMurdie PJ, Holmes S (2013) Phyloseq: An R Package for Reproducible Interactive Analysis and Graphics of Microbiome Census Data. PLoS ONE

Meij SET van der, Moolenbeek RG, Hoeksema BW (2009) Decline of the Jakarta Bay molluscan fauna linked to human impact. Marine pollution bulletin 59:101–7

Moitinho-Silva L, Steinert G, Nielsen S, Hardoim CCP, Wu Y-C, McCormack GP, López-Legentil S, Marchant R, Webster N, Thomas T, Hentschel U (2017) Predicting the HMA-LMA Status in Marine Sponges by Machine Learning. Frontiers in Microbiology 8:752

Montenegro D, Astudillo-García C, Hickey T, Lear G (2020) A non-invasive method to monitor marine pollution from bacterial DNA present in fish skin mucus. Environmental Pollution 263:114438

Mortzfeld BM, Urbanski S, Reitzel AM, Künzel S, Technau U, Fraune S (2016) Response of

bacterial colonization in Nematostella vectensis to development, environment and biogeography. Environmental Microbiology 18:1764–1781

Mukherjee N, Bartelli D, Patra C, Chauhan B V., Dowd SE, Banerjee P (2016) Microbial diversity of source and point-of-use water in rural Haiti - A pyrosequencing-based metagenomic survey. PLoS ONE

Mukherjee C, Beall CJ, Griffen AL, Leys EJ (2018) High-resolution ISR amplicon sequencing reveals personalized oral microbiome. Microbiome

Newton K, Côté IM, Pilling GM, Jennings S, Dulvy NK (2007) Current and Future Sustainability of Island Coral Reef Fisheries. Current Biology 17:655–658

Nogales B, Aguiló-Ferretjans MM, Martín-Cardona C, Lalucat J, Bosch R (2007) Bacterial diversity, composition and dynamics in and around recreational coastal areas. Environmental Microbiology 9:1913–1929

Nogales B, Lanfranconi MP, Pi??a-Villalonga JM, Bosch R (2011) Anthropogenic perturbations in marine microbial communities. FEMS Microbiology Reviews

Nogales B, Lanfranconi MP, Piña JM, Piña-Villalonga P, Bosch R, Pi??a-Villalonga JM, Bosch R, Piña-Villalonga JM, Bosch R (2011) Anthropogenic perturbations in marine microbial communities. FEMS Microbiology Reviews 35:275–298

OECD (2019) Biodiversity : Finance and the Economic and Business Case for Action Annexes to the Report - Executive Summary and Synthesis. :5–6

Olsen CR, Cutshall NH, Larsen IL (1982) Pollutant—particle associations and dynamics in coastal marine environments: a review. Marine Chemistry 11:501–533

Pandolfi JM, Bradbury RH, Sala E, Hughes TP, Bjorndal K a, 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 (New York, NY) 301:955–958

Pastorok R, Bilyard G (1985) Effects of sewage pollution on coral-reef communities. Marine Ecology Progress Series 21:175–189

Paul VJ, Fenical W (1986) Chemical defense in tropical green algae, order Caulerpales.

Pearman JK, Anlauf H, Irigoien X, Carvalho S (2016) Please mind the gap - Visual census and cryptic biodiversity assessment at central Red Sea coral reefs. Marine Environmental Research 118:20–30

Pearman JK, Aylagas E, Voolstra CR, Anlauf H, Villalobos R, Carvalho S (2019) Disentangling the complex microbial community of coral reefs using standardized Autonomous Reef Monitoring Structures (ARMS). Molecular Ecology:mec.15167

Pearman JK, Leray M, Villalobos R, Machida RJ, Berumen ML, Knowlton N, Carvalho S (2018) Cross-shelf investigation of coral reef cryptic benthic organisms reveals diversity patterns of the hidden majority. Scientific Reports 8:8090

Pen O, Mcclanahan TR, Graham NAJ, Wilson SK, Letourneur Y, Fisher R (2009) Effects of fisheries closure size , age , and history of compliance on coral reef fish communities in the western Indian Ocean. Marine Ecology Progress Series 396:99–109

Pendleton LH (1995) Valuing coral reef protection. Ocean & Coastal Management 26:119–131

Plaisance L, Brainard R, Julian Caley M, Knowlton N (2011) Using DNA barcoding and standardized sampling to compare geographic and habitat differentiation of crustaceans: A Hawaiian Islands example. Diversity

Ponsero AJ, Bomhoff M, Blumberg K, Youens-Clark K, Herz NM, Wood-Charlson EM, Delong EF, Hurwitz BL (2021) Planet Microbe: a platform for marine microbiology to discover and

analyze interconnected ‘omics and environmental data. Nucleic Acids Research 49:D792–D802

Quero GM, Cassin D, Botter M, Perini L, Luna GM (2015) Patterns of benthic bacterial diversity in coastal areas contaminated by heavy metals, polycyclic aromatic hydrocarbons (PAHs) and polychlorinated biphenyls (PCBs). Frontiers in Microbiology

R Core Team (2020) R: A language and environment for statistical computing. Retrieved from.

Rachello-Dolmen PG, Cleary DFR (2007) Relating coral species traits to environmental conditions in the Jakarta Bay/Pulau Seribu reef system, Indonesia. Estuarine, Coastal and Shelf Science

Ramosaj B, Pauly M (2019) Consistent estimation of residual variance with random forest Out-Of-Bag errors. Statistics & Probability Letters 151:49–57

Ransome E, Geller JB, Timmers M, Leray M, Mahardini A, Sembiring A, Collins AG, Meyer CP, Mahardini A, Sembiring A, Collins AG, Meyer CP (2017) The importance of standardization for biodiversity comparisons: A case study using autonomous reef monitoring structures (ARMS) and metabarcoding to measure cryptic diversity on Mo’orea coral reefs, French Polynesia (CA Chen, Ed.). PLoS ONE 12:e0175066

Rees JG, Setiapermana D, Sharp VA, Weeks JM, Williams TM (1999) Evaluation of the impacts of land-based contaminants on the benthic faunas of Jakarta Bay, Indonesia. Oceanologica Acta 22:627–640

Riani E, Cordova MR, Arifin Z (2018) Heavy metal pollution and its relation to the malformation of green mussels cultured in Muara Kamal waters, Jakarta Bay, Indonesia. Marine Pollution Bulletin 133:664–670

Ribot WJ, Ulrich RL (2006) The Animal Pathogen-Like Type III Secretion System Is Required

for the Intracellular Survival of Burkholderia mallei within J774.2 Macrophages. Infection and Immunity 74:4349

Rochman CM, Kurobe T, Flores I, Teh SJ (2014) Early warning signs of endocrine disruption in adult fish from the ingestion of polyethylene with and without sorbed chemical pollutants from the marine environment. Science of The Total Environment 493:656–661

Rognes T, Flouri T, Nichols B, Quince C, Mahé F (2016) VSEARCH: A versatile open source tool for metagenomics. PeerJ

Rohwer F, Breitbart M, Jara J, Azam F, Knowlton N (2001) Diversity of bacteria associated with the Caribbean coral Montastraea franksi. Coral Reefs 20:85–91

Rohwer F, Seguritan V, Azam F, Knowlton N (2002) Diversity and distribution of coral-associated bacteria. Marine Ecology Progress Series 243:1–10

Ryan MP, Adley CC (2014) Ralstonia spp.: Emerging global opportunistic pathogens. European Journal of Clinical Microbiology and Infectious Diseases 33:291–304

Sawall Y, Richter C, Ramette A (2012) Effects of Eutrophication, Seasonality and Macrofouling on the Diversity of Bacterial Biofilms in Equatorial Coral Reefs. PLOS ONE 7:e39951

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 6:564–576

Selig ER, Turner WR, Troëng S, Wallace BP, Halpern BS, Kaschner K, Lascelles BG, Carpenter KE, Mittermeier RA (2014) Global priorities for marine biodiversity conservation. PloS one 9:e82898

Selvin J, Shanmugha Priya S, Seghal Kiran G, Thangavelu T, Sapna Bai N (2009) Sponge-

associated marine bacteria as indicators of heavy metal pollution. Microbiological research 164:352–63

Sharifuzzaman SM, Rahman H, Ashekuzzaman SM, Islam MM, Chowdhury SR, Hossain MS (2015) Heavy metals accumulation in coastal sediments. Environmental Remediation Technologies for Metal-Contaminated Soils:21–42

Shibata T, Solo-Gabriele HM, Fleming LE, Elmir S (2004) Monitoring marine recreational water quality using multiple microbial indicators in an urban tropical environment. Water Research 38:3119–3131

Shnit-Orland M, Kushmaro A (2009) Coral mucus-associated bacteria: A possible first line of defense. FEMS Microbiology Ecology 67:371–380

Simanjuntak SW, Supriharyono, Haeruddin (2018) Analysis of Suitability and Carrying Capacity of Tourism in Tidung Island, Kepulauan Seribu of Indonesia. Russian Journal of Agricultural and Socio-Economic Sciences 78:151–159

Sinigalliano CD, Fleisher JM, Gidley ML, Solo-Gabriele HM, Shibata T, Plano LRW, Elmir SM, Wanless D, Bartkowiak J, Boiteau R, Withum K, Abdelzaher AM, He G, Ortega C, Zhu X, Wright ME, Kish J, Hollenbeck J, Scott T, Backer LC, Fleming LE (2010) Traditional and molecular analyses for fecal indicator bacteria in non-point source subtropical recreational marine waters. Water Research

Spalding M, Burke L, Wood SA, Ashpole J, Hutchison J, zu Ermgassen P (2017) Mapping the global value and distribution of coral reef tourism. Marine Policy 82:104–113

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

Tomascik T, Suharsono, Mah AJ (1994) Case histories: a historical perspective of the natural and anthropogenic impacts in the Indonesian Archipelago with a focus on the Kepulauan Seribu, Java Sea. Proceedings of the colloquium on global aspects of coral reefs, Miami, 1993 8:683–684

Tomascik T, Suharsono, Mah AJ (1995) Case histories: a historical perspective of the natural and anthropogenic impacts in the Indonesian Archipelago with a focus on the Kepulauan Seribu, Java Sea. Oceanographic Literature Review 8:683–684

Tout J, Jeffries TC, Webster NS, Stocker R, Ralph PJ, Seymour JR (2014) Variability in microbial community composition and function between different niches within a coral reef. Microbial ecology 67:540–552

Trubina MR, Vorobeichik EL (2012) Severe industrial pollution increases the β-diversity of plant communities. Doklady Biological Sciences 2012 442:1 442:17–19

Undap SL, Nirmala K, Miki S, Inoue S, Qiu X, Honda M, Shimasaki Y, Oshima Y (2013) High tributyltin contamination in sediments from ports in Indonesia and Northern Kyushu, Japan. Journal of the Faculty of Agriculture, Kyushu University

Uneputty PA, Evans SM (1997) Accumulation of beach litter on islands of the Pulau Seribu Archipelago, Indonesia. Marine Pollution Bulletin 34:652–655

Vega Thurber RL, Burkepile DE, Fuchs C, Shantz AA, McMinds R, Zaneveld JR (2014) Chronic nutrient enrichment increases prevalence and severity of coral disease and bleaching. Global Change Biology 20:544–554

Wainberg M, Alipanahi B, Frey BJ (2016) Are Random Forests Truly the Best Classifiers? Journal of Machine Learning Research 17:1–5

Waldichuk M (1974) Coastal marine pollution and fish. Ocean Management 2:1–60

Walsh EA, Kirkpatrick JB, Rutherford SD, Smith DC, Sogin M, D 'hondt S, D'Hondt S (2015) Bacterial diversity and community composition from seasurface to subseafloor. The ISME Journal 10:979–989

Walters W, Hyde ER, Berg-lyons D, Ackermann G, Humphrey G, Parada A, Gilbert J a, Jansson JK (2015) Improved Bacterial 16S rRNA Gene (V4 and V4-5) and Fungal Internal Transcribed Spacer Marker Gene Primers for Microbial Community Surveys. mSystems 1:1–15

Wang Y, Sheng H-F, He Y, Wu J-Y, Jiang Y-X, Tam NF-Y, Zhou H-W (2012) Comparison of the Levels of Bacterial Diversity in Freshwater, Intertidal Wetland, and Marine Sediments by Using Millions of Illumina Tags. Applied and Environmental Microbiology 78:8264–8271

Ward-Paige CA, Risk MJ, Sherwood OA, Jaap WC (2005) Clionid sponge surveys on the Florida Reef Tract suggest land-based nutrient inputs. Marine Pollution Bulletin

Warren C, Steenbergen DJ (2021) Fisheries decline, local livelihoods and conflicted governance: An Indonesian case. Ocean & Coastal Management 202:105498

Webster NS, Negri AP, Botté ES, Laffy PW, Flores F, Noonan S, Schmidt C, Uthicke S, Orr JC, Hoegh-Guldberg O, Anthony KRN, Kline DI, Diaz-Pulido G, Dove S, Hoegh-Guldberg O, Schneider K, Erez J, Jokiel P, Wernberg T, Smale DA, Thomsen MS, Lesser MP, Mazel CH, Gorbunov MY, Falkowski PG, Kimes NE, Nostrand JD Van, Weil E, Zhou J, Morris PJ, Wegley L, Edwards RA, Rodriguez-Brito B, Liu H, Rohwer F, Raina J-B, Dinsdale E, Willis BL, Bourne DG, Raina J-B, Tapiolas D, Willis BL, Bourne DG, Castillo I, Lodeiros C, Nunez M, Campos I, Ritchie KB, Sweet MJ, Croquer A, Bythell JC, Negri AP, Webster NS, Hill RT, Heyward AJ, Tebben J, Johnson CR, Sutton DC, Joint I n., Doney SC, Karl

DM, Bourne DG, Iida Y, Uthicke S, Smith-Keune C, Littman R, Willis BL, Bourne DG, Thurber RLVV, Webster NS, Soo R, Cobb R, Negri AP, Fan L, Liu M, Simister R, Webster NS, Thomas T, Hutchins DA, Beman JM, Shi D, Xu Y, Hopkinson BM, Morel FMM, Webster NS, Uthicke S, Botté ES, Flores F, Negri AP, Webster NS, Meron D, Morrow KM, Thurber RV, Thurber RLV, Webster NS, Cobb RE, Negri AP, Bayer T, Bourne DG, Schmidt C, Heinz P, Kucera M, Uthicke S, Schmidt C, Kucera M, Uthicke S, Dam JW van, Negri AP, Mueller JF, Altenburger R, Uthicke S, Uthicke S, Vogel N, Doyle J, Schmidt C, Humphrey C, Momigliano P, Uthicke S, Langer MR, Silk MT, Lipps JH, Littman RA, Willis BL, Pfeffer C, Bourne DG, Chen C-P, Tseng C-H, Chen CA, Tang S-L, Littman RA, Bourne DG, Willis BL, Santos HF, Noonan S, Fabricius K, Lomas MW, Arakaki Y, Uehara T, Fagoonee I, Uthicke S, Soars N, Foo S, Byrne M, Schloss PD, Gevers D, Westcott SL, Pruesse E, Edgar RC, Haas BJ, Clemente JC, Quince C, Knight R, Altschul SF, DeSantis TZ, Legendre P, Anderson MJ (2016) Host-associated coral reef microbes respond to the cumulative pressures of ocean warming and ocean acidification. Scientific Reports 6:19324

Webster NS, Reusch TBH (2017) Microbial contributions to the persistence of coral reefs. ISME Journal 11:2167–2174

Weiler BA, Verhoeven JTP, Dufour SC (2018) Bacterial Communities in Tissues and Surficial Mucus of the Cold-Water Coral Paragorgia arborea. Frontiers in Marine Science

Weng HX, Zhu YM, Qin YC, Chen JY, Chen XH (2008) Accumulation discrepancy of heavy metal and organic pollutants in three near-shore depositional environments, southeastern China. Journal of Asian Earth Sciences 31:522–532

White AT, Vogt HP, Arin T (2000) Philippine Coral Reefs Under Threat: The Economic Losses Caused by Reef Destruction. Marine Pollution Bulletin 40:598–605

Wickham H (2016) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York

Williams TM, Rees JG, Setiapermana D (2000) Metals and Trace Organic Compounds in Sediments and Waters of Jakarta Bay and the Pulau Seribu Complex, Indonesia. Marine Pollution Bulletin 40:277–285

Willoughby NG (1986) Man-made litter on the shores of the Thousand Island archipelago, Java. Marine Pollution Bulletin 17:224–228

Willoughby NG, Sangkoyo H, Lakaseru BO (1997) Beach litter: an increasing and changing problem for Indonesia. Marine Pollution Bulletin 34:469–478

Xiong T, Yuan X, Wang H, Leng L, Li H, Wu Z, Jiang L, Xu R, Zeng G (2018) Implication of graphene oxide in Cd-contaminated soil: A case study of bacterial communities. Journal of Environmental Management 205:99–106

Xue J, Yu Y, Bai Y, Wang L, Wu Y (2015) Marine Oil-Degrading Microorganisms and Biodegradation Process of Petroleum Hydrocarbon in Marine Environments: A Review. Current Microbiology

Yu S-X, Pang Y-L, Wang Y-C, Li J-L, Qin S (2018) Distribution of bacterial communities along the spatial and environmental gradients from Bohai Sea to northern Yellow Sea. PeerJ

Zhang D, Chen H, Wang K, Ye X, Zhu J, Hu C, Xiong J (2014) Biogeography of the Sediment Bacterial Community Responds to a Nitrogen Pollution Gradient in the East China Sea. Applied and Environmental Microbiology

Zhang R, Lau SCK, Ki J-S, Thiyagarajan V, Qian P-Y (2009) Response of bacterioplankton community structures to hydrological conditions and anthropogenic pollution in contrasting subtropical environments. FEMS Microbiology Ecology 69:449–460

Zhang R, Liu B, Lau SCK, Ki JS, Qian PY (2007) Particle-attached and free-living bacterial

communities in a contrasting marine environment: Victoria Harbor, Hong Kong. FEMS Microbiology Ecology

Zhang W, Song L sheng, Ki JS, Lau CK, Li XD, Qian PY (2008) Microbial diversity in polluted harbor sediments II: Sulfate-reducing bacterial community assessment using terminal restriction fragment length polymorphism and clone library of dsrAB gene. Estuarine, Coastal and Shelf Science

Ziegler M, Roik A, Porter A, Zubier K, Mudarris MS, Ormond R, Voolstra CR (2016) Coral microbial community dynamics in response to anthropogenic impacts near a major city in the central Red Sea. Marine Pollution Bulletin 105:629–640

Zouch H, Karray F, Armougom F, Chifflet S, Hirschler-Réa A, Kharrat H, Kamoun L, Hania W Ben, Ollivier B, Sayadi S, Quéméneur M (2017) Microbial diversity in sulfate-reducing marine sediment enrichment cultures associated with anaerobic biotransformation of coastal stockpiled phosphogypsum (Sfax, Tunisia). Frontiers in Microbiology