Please mind the gap – Visual census and cryptic biodiversity assessment at central red sea coral reefs

John K. Pearman, Holger Anlauf, Xabier Irigoien, Susana Carvalho

PII: S0141-1136(16)30061-7
DOI: 10.1016/j.marenvres.2016.04.011
Reference: MERE 4161

To appear in: Marine Environmental Research

Received Date:  3 March 2016
Revised Date:  20 April 2016
Accepted Date: 24 April 2016


This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.
Please mind the gap – visual census and cryptic biodiversity assessment at central Red Sea coral reefs

Authors

John K. Pearman*, Holger Anlauf, Xabier Irigoien, Susana Carvalho

Affiliation:

King Abdullah University of Science and Technology (KAUST), Red Sea Research Center (RSRC), Biological and Environmental Sciences and Engineering (BESE), Thuwal 23955-6900, Saudi Arabia.

*Corresponding author email: john.pearman@kaust.edu.sa

Abstract

Coral reefs are the most diverse assemblages in the ocean, however a large proportion of the diversity is cryptic and, therefore, undetected by standard visual census techniques. Cryptic and exposed communities differ considerably in species composition and ecological function. This study compares three different coral reef assessment protocols: i) visual benthic reef surveys: ii) visual census of Autonomous Reef Monitoring Structures (ARMS) plates; and iii) metabarcoding techniques of the ARMS (including sessile, 106 – 500 μm and 500 – 2000 μm size fractions) that target the cryptic and exposed communities of three reefs in the central Red Sea. Visual census showed a dominance of Cnidaria (Anthozoa) and Rhodophyta on the reef substrate, while Porifera, Bryozoa and Rhodophyta were the most abundant groups on the ARMS plates. Metabarcoding, targeting the 18S rRNA gene, significantly increased estimates of
the species diversity \( (p < 0.001) \); revealing that Annelida were generally the
dominant phyla (in terms of reads) of all fractions and reefs. Furthermore,
metabarcoding detected microbial eukaryotic groups such as Syndiniophyceae,
Mamiellophyceae and Bacillariophyceae as relevant components of the sessile
fraction. ANOSIM analysis showed that the three reef sites showed no differences
based on the visual census data. Metabarcoding showed a higher sensitivity for
identifying differences between reef communities at smaller geographic scales
than standard visual census techniques as significant differences in the
assemblages were observed amongst the reefs. Comparison of the techniques
showed no similar patterns for the visual techniques while the metabarcoding of
the ARMS showed similar patterns amongst fractions. Establishing ARMS as a
standard tool in reef monitoring will not only advance our understanding of local
processes and ecological community response to environmental changes, as
different faunal components will provide complementary information but also
improve the estimates of biodiversity in coral reef benthic communities. This
study lays the foundations for further studies looking at integrating traditional
reef survey methodologies with complementary approaches, such as
metabarcoding, which investigate other components of the reef community.

**Keywords:** Autonomous Reef Monitoring Structures; biodiversity surveys;
metabarcoding; coral reefs; Red Sea; Marine Ecology; benthos; monitoring

1. **Introduction**

It has been estimated that one billion people benefit directly from coral reefs
through a variety of services, including fisheries and tourism (Hughes et al., 2007). Furthermore, coral reefs provide substances for pharmaceutical application (Rocha et al., 2011) and protection of shores against the surge from storms and rising sea levels (Sheppard et al., 2005).

Coral reefs have been termed the “rainforests of the sea” (Reaka-Kudla 1997), as they are believed to host 25% to 33% of marine biodiversity (Plaisance et al. 2011). Conservative estimates suggest that ~176,000 species populate coral reefs globally (Glynn and Enochs 2010), but considering limitations in sampling reef associated habitats, up to one million species may reside in coral reefs (Reaka-Kudla 1997; Glynn and Enochs 2010). Recent reports show that the majority of the reef biodiversity is comprised of small invertebrates, which are cryptic (Dennis and Aldhous 2004). Cryptic habitat space, or small cavities within the coral limestone in which organisms can inhabit, has been estimated at 30 – 75% of the reef habitat (Ginsburg 1983; Scheffers et al., 2003). Exploration of these cryptic spaces has been achieved by using techniques that employ endoscopy of cavities in the limestone revealing, for example, abundant sponge communities in Red Sea coral reefs (Richter et al., 2001). Hence, a large cohort of biodiversity nested within coral reefs may remain unreported.

Coral reefs of the Red Sea have been previously described as biodiversity hotspots (Roberts et al., 2002), as well as refuge for potentially more temperature tolerant coral reef species, which might better endure the rise in sea surface temperature (Fine et al., 2013). Reef surveys following standard protocols assessing the exposed benthos will typically identify scleractinian corals as the dominant component of coral reefs, with them being vital elements
in the structuring of the habitat. Any deviation from the dominance of scleractinian corals in tropical coral reefs may indicate an ongoing process of shifting ecosystem stages (Bruno et al., 2009). In the Red Sea, there are around 260 species of scleractinian coral (Dubinsky and Stambler 2013), including 21 endemic hard coral species (DeVantier et al., 2000). A similar degree of biodiversity was observed for non-scleractinian cryptic taxa with 254 species being observed in reef cavities in the Gulf of Aqaba (Wunsch et al., 2000). Furthermore, reef survey protocols may target macroinvertebrates that play key roles in hard coral dominated reefs, (e.g. sea urchins, crown-of-thorns starfish) but cryptic species, which could account for a large proportion of biodiversity in coral reefs, are not usually assessed. Cryptic species may include sessile, encrusting, and mobile species, which find habitat and ecological niches in the cryptic spaces. Species assessment on hidden substrata may double species biodiversity and illustrates the limits of the visual census of the exposed benthos.

The evaluation of current biodiversity baselines is important to better understand species richness in the marine realm, the economic value of biodiversity and alterations of ecosystems services due to changes in community structures in the Anthropocene. Like other marine habitats, coral reefs are jeopardized by multiple factors on a local, regional and global scale that will alter biodiversity and ecosystem services during the 21st century (Hughes et al., 2007; Sandin et al., 2008; Kennedy et al., 2013; Salvat et al., 2015). Indeed, different biological components may respond differently to environmental pressures, and thus neglecting a relevant component may lead to less efficient conservation initiatives. Even within the cryptic component different colonisation and
succession patterns may be observed between sessile and vagile epibenthic assemblages (Moura et al., 2008).

Recently, the ARMS (Autonomous Reef Monitoring Structure) sampling tool has been introduced to provide estimates of the cryptic reef biodiversity (Leray and Knowlton 2015). ARMS units provide a standardized method that has been utilized at different locations around the globe since 2006 to conduct visual census of species abundance and molecular estimates of biodiversity (for more information, see http://www.pifsc.noaa.gov/cred/survey_methods/arms/overview.php). The ARMS collection protocol includes the application of targeted metabarcoding sequencing of the community to estimate the expected biodiversity richness of both sessile and mobile (sessile encrusting fauna; meiofauna, 106-500 μm; macrofauna, 500-2000 μm; large macrofauna, >2000 μm). The application of a standardized protocol that relies on amplicon sequencing provides an approach that is independent of the observer and worldwide datasets targeting the same gene regions can easily be combined and compared. Here, we present results from the standard visual benthic census of three Red Sea nearshore reefs (Jeddah, Kingdom of Saudi Arabia) and cryptic diversity assessments using ARMS. Results from three assessment methods that include 1) standard reef survey, 2) photo survey of ARMS plates and 3) metabarcoding of ARMS units, were compared and analysed for similar patterns. We hypothesize that the different methods, which investigate different faunal components, will provide complementary information on the biodiversity pattern, thus increasing the information available to detect responses to pressures affecting coral reef
2. **Material and Methods**

2.1 Study area

Three nearshore reefs were randomly selected along the shoreline off Jeddah in order to represent a gradient of human pressure. One reef (JH) was located off of Jeddah harbour, the main harbour on the Saudi Arabian Red Sea coast, serving a city of over 3 million inhabitants. The other two reefs were located approximately 20 km (JS1 – in the vicinity of the most recent sewage treatment plant) and 40 km (JS2) south of JH (Table 1 and Figure 1). Field permits for sampling were granted by the Saudi Arabian coastguard.
Figure 1: Map showing the position of the three reefs sampled in the Central Red Sea. See Table 1 for details of the sampling positions. Made in ArcGIS.
Table 1: Descriptions of the reefs (Co-ordinates, depth of sampling, temperature, salinity, nutrient and chlorophyll \( a \) concentrations) where reef surveys were undertaken and ARMS units placed. NA = not available.

<table>
<thead>
<tr>
<th>Reef</th>
<th>Reef name</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Depth (m)</th>
<th>Temperature (°C)</th>
<th>Salinity</th>
<th>( \text{PO}_4 ) (uM)</th>
<th>( \text{SiO}_4 ) (uM)</th>
<th>( \text{NO}_3 ) (uM)</th>
<th>( \text{NO}_2 ) (uM)</th>
<th>Chl a (ug/L)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JH</td>
<td>Janib Sa'ara</td>
<td>39 06.661</td>
<td>21 27.253</td>
<td>10</td>
<td>26.4</td>
<td>38.6</td>
<td>0.148</td>
<td>1.519</td>
<td>1.045</td>
<td>0.122</td>
<td>0.604</td>
</tr>
<tr>
<td>JS1</td>
<td>NA</td>
<td>39 07.237</td>
<td>21 13.508</td>
<td>10</td>
<td>26.3</td>
<td>38.7</td>
<td>0.062</td>
<td>1.608</td>
<td>3.597</td>
<td>0.107</td>
<td>0.729</td>
</tr>
<tr>
<td>JS2</td>
<td>Qaham</td>
<td>39 12.063</td>
<td>21 04.921</td>
<td>10</td>
<td>25.8</td>
<td>38.9</td>
<td>0.082</td>
<td>1.102</td>
<td>0.568</td>
<td>0.125</td>
<td>0.211</td>
</tr>
</tbody>
</table>
2.2 Standard reef benthic diversity surveys

Strategy

Reef surveys were carried out in March 2014, just before the retrieval of ARMS units. In each reef, three replicate transects of 20 m each (5 m gap between transects) were assessed at 10 m depth. Along each transect, a photo (1 m²) was taken every two meters. Quantification and identification of benthic categories were conducted using the Coral Point Count with Excel extensions (Kohler and Gill 2006). Forty-eight points were randomly distributed within a total of 12 cells on each substrate image and the features underlying the points user-identified. In general, given the revisions on coral’s taxonomy being undertaken worldwide, coral identification was limited to family or genus level, following Veron (2000).

Overall, benthic community composition was analysed in terms: 1) of number of morphological Operational Taxonomic Units (morphOTU; organisms with similar morphological characteristics being classified as a taxon); 2) cover of corals (both soft and scleractinian corals), sponges, hydroids and other invertebrates (e.g. bivalves, echinoderms), algae (subdivided into the following functional groups: macroalgae, turf, coralline algae); 3) cover of abiotic substrates (sand, rock, rubble); and 4) the occurrence of coral diseases and mortality.

2.3 Assessment of cryptic diversity: ARMS

Deployment and sampling of ARMS units

Three replicate units of ARMS were deployed by scuba divers at three reef sites off Jeddah (Figure 1) in April-May 2013 and were left undisturbed for one year.
The ARMS were deployed on a flat rocky area at a depth of approximately 10 m. In order to characterize each reef area, water samples were collected and analysed for nutrients and chlorophyll $a$ (see Pearman et al., 2016), for details on the analytical procedures). When the nine ARMS units were retrieved, they were covered by a 106 μm net (to avoid the loss of mobile fauna), placed in individual boxes filled with filtered (106 μm) seawater, and transported to the laboratory for processing and analysis.

The ARMS samples were processed as described by (Leray and Knowlton 2015). Briefly, the ARMS were dismantled in the transport box with each plate (9 plates in all per ARMS) being gently brushed to dislodge motile organisms from the plates into the filtered seawater. Individual plates were placed in labelled trays (top side up) and filled with filtered (0.2 μm) seawater. Pictures of the top and bottom sides of each plate were taken (Figure 2), as well as close ups of each quarter of the plates. After the plates had been photographed, they were scraped clean to remove the sessile community. The sessile community from all plates (of a single ARMS unit) was combined and washed through a 106 μm net. The sessile fraction was then blended, washed with deionized water and ethanol before sub samples (10 g) were preserved at – 20 °C in 96% ethanol.
Figure 2: An uncolonised ARMS unit (A) positioned on the reef. Example pictures from the top of two plates (B and C) and the bottom of the same plates (D and E respectively).

For the analysis of the mobile fractions, the seawater from each transport box was filtered through a selection of mesh sizes (2000 μm, 500 μm and 106 μm).
The material retained in each sieve is stored separately. The greater than 2000 μm size fraction, which is analysed by a combination of traditional taxonomy and barcoding, was not considered in this study. The smaller size fractions (106 μm – 500 μm and 500 μm – 2000 μm) were washed in deionized water until no organic particles were observed before two replicate subsamples of 10 g from each size fraction were stored in 50 mL centrifuge tubes. The centrifuge tubes were filled with 96% ethanol and stored at −20 °C until further processing.

**Photograph analysis of plates**

A subset of plates (plates 1, 4 and 8; both top and bottom surfaces) was analysed using the software photoquad (http://www.mar.aegean.gr/sonarlab/photoquad/index.php). A total of 64 evenly stratified random points were spawned over each plate for visual census following a point count approach. When the selected points fell within an area, which was uncolonisable (e.g. where the cross between plates was situated or where screws were present), the point was assigned to a category named “uncolonisable”. This was taken into account in the calculation of the areas colonized by the different taxa. Biotic elements were then identified to the lowest possible taxonomic level using appropriate identification keys and similar to the visual analysis of the reef survey classified as morphOTUs.

**DNA extraction**

The 10 g subsamples of the sessile community and the two size fractions (106 μm – 500 μm and 500 μm – 2000 μm) were extracted using the Powermax Soil DNA isolation kit (MoBio) following the kits instructions with the exception of step three. Instead of the bead beating step, 0.4 mg ml⁻¹ of Proteinase K was
added to the Powermax Bead solution and incubated in a shaking incubator overnight at 56 °C, as detailed in (Leray and Knowlton 2015). The DNA was purified using the Powerclean DNA Clean-up kit (MoBio) following manufacturer’s instructions and quantified with a Qubit Fluorometer (Invitrogen) prior to PCR amplification and normalised to and approximately 5 ng DNA. Amplicons were generated, in triplicate 50 μl reaction volumes for each sample, using general eukaryotic primers (TAReuk454FWD1: CASCYCGGTAATTCC and TAReukREV3: ACTTTCGTTCTTGATYRA; Stoeck et al 2010). Each reaction volume contained 2.5 U Taq polymerase, 1X Taq reaction Buffer, 200 μM dNTPs, 1.5mM MgSO₄, 0.05 mg Bovine Serum Albumin (BSA) and 0.2 μM of each primer. Amplification involved an initial denaturation step at 95 °C for 5 min followed by 10 cycles consisting of 94 °C for 30 s, 57 °C for 45 s and 72 °C for 1 min followed by 25 cycles of 94 °C for 30 s, 49 °C for 45 s and 72 °C for 1 min followed by a final extension at 72 °C for 5 min. The triplicate samples were pooled and cleaned up using SequelPrep Normalization (ThermoFisher) plates (which normalise concentrations to ~ 1-2 ng μl⁻¹ before continuing with PCR 2 of the Illumina 16S metagenomic sequencing library preparation protocol to index the samples. Indexed samples were pooled (96 samples per pool) and sent for sequencing at the King Abdullah University of Science and Technology core laboratory on a MiSeq platform. Raw reads were submitted to the NCBI SRA achieve and can be found under the project accession number: SRP068409.

2.4 Data analysis

Bioinformatics
Automatic demultiplexing of samples occurred on the MiSeq sequencing machine and forward and reverse reads were joined using the join_seqs.py script in QIIME (Caporaso et al., 2010). Subsequently, the joined reads were quality filtered (PHRED quality = 25) and reads were truncated at the reverse primer having previously removed the forward primer. A two-step clustering process was undertaken to generate genomic Operational Taxonomic Units (gOTUs). First, the CD-HIT algorithm (Li and Godzik 2006) was implemented in QIIME using the trie function and representative sequences were generated for each gOTU (longest sequence was selected). These representative sequences were then used as the basis for a second round of clustering using USEARCH (version 5.2.236) (Edgar 2010) at 97% similarity with a minimum cluster size of two (singletons were removed). Reference based chimera detection was undertaken using UCHIME (Edgar et al., 2011) against the PR2 database (Guillou et al., 2013) and representative sequences from the non chimeric gOTUs were taxonomically classified using QIIME against the PR2 database.

To enable comparison between samples in the metabarcoding dataset, reads were rarefied to an even depth (69,600 reads per sample) multiple times (n=100). Samples not meeting this threshold were removed from analysis. Diversity statistics for all datasets were calculated in R (R development core team 2015) using the package vegan (Oksanen et al., 2015). One-way Analysis of Variance (ANOVA) for factor “Reef” (orthogonal, three levels) was used to investigate patterns of variation in the univariate measures calculated in relation to the reefs surveyed. For the metabarcoding dataset, two-way ANOVA for factors “Fraction” (orthogonal, three levels, sessile, 106 – 500 μm and 500 – 2000...
μm) and “Reef” (orthogonal, three levels) was also investigated. Community composition was visualised using ggplot (Wickham and Chang 2015) in R by clustering OTUs based on taxonomic classifications at the phyla level.

To visualize patterns of change in the composition and structure of reef macrobenthic assemblages, a Principal Coordinate Analysis (PCO) was undertaken using PRIMER v6 (Clarke and Gorley 2006). Taxa and benthic categories showing Pearson’s correlations higher than $r = 0.6$ and $r = 0.7$, respectively, with the biological model were added to the plots. A one-way (reef surveys, photo analysis of the plates datasets) or two-way (metabarcoding dataset) ANOSIM (Analysis of similarities) (Clarke and Ainsworth 1993) for factor “Reef” or factors “Reef” and “Fraction”, respectively, was performed, in order to test the null hypothesis of no differences on the multivariate structure and composition of reef macrobenthic assemblages among reefs and faunal size fractions. For the reef survey dataset and for the analysis of the sessile component based on the photo analysis, all multivariate methods were based on Jaccard (presence/absence) and Bray-Curtis (percentage occurrence) similarities calculated from square-root transformed data. All multivariate methods for the metabarcoding dataset were based in the same indices after log $(x+1)$ transformation of the reads abundance. Comparative (Mantel-type) tests were undertaken on the similarity matrices to compare if different methods showed similar patterns using the RELATE (Spearman’s rank method) function in PRIMER.

### 3. Results
For the metabarcoding a total of 9,609,425 raw reads (from an extended dataset) were obtained from the sequencing and after quality filtering 9,281,518 reads remained. Clustering resulted in 2420 gOTUs at the 97% similarity level. However 557 gOTUs were found to be chimeric and removed from the analysis and subsequent to rarefaction at 69,000 reads, a total of 1700 gOTUs with an average amplicon length of 395 bp were observed (for the dataset described in this manuscript).

Diversity

Comparison between methods showed that the metabarcoding resulted in a significantly higher number of OTUs than the other two methods (F=350.3, p < 0.001). For both the reef surveys and photo analysis of ARMS plates there was no significant difference in the number of observed morphOTUs among reefs (F = 2.844, p = 0.135; F = 0.037, p = 0.964, respectively). There was, however, a significant difference in the number of gOTUs observed in the metabarcoding dataset among reefs (F = 10.709; p < 0.001) with JH having the highest number of OTUs (Table 2) and JS1 being significantly lower in terms of observed gOTUs compared with JH and JS2. There was also a significant difference in the number of gOTUs observed between the size fractions (F = 9.491; p = 0.002) with the 106-500 μm fraction showing the highest number of gOTUs and being significantly different from the sessile and 500 - 2000 μm size fractions.
Table 2: Number of gOTUs and chao1 estimation for each reef and size fraction of each reef for the metabarcoding. Rarefaction was at 69,600 reads per sample, n = 100.

<table>
<thead>
<tr>
<th>Reef</th>
<th>Photo</th>
<th>JH</th>
<th>JS1</th>
<th>JS2</th>
<th>JH</th>
<th>JS1</th>
<th>JS2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Total</td>
<td>Sessile</td>
<td>106 - 500 μm</td>
<td>500 - 2000 μm</td>
<td>Total</td>
<td>Sessile</td>
</tr>
<tr>
<td>OTUs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>JH</td>
<td>JS1</td>
<td>JS2</td>
<td>JH</td>
<td>JS1</td>
<td>JS2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OTUs (rarefied)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1388</td>
<td>993</td>
<td>1006</td>
<td>916</td>
<td>1122</td>
<td>737</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chao1</td>
<td>(rarefied)</td>
<td>85.69</td>
<td>40.53</td>
<td>116.09</td>
<td>36.33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>45.69</td>
<td>(3.64)</td>
<td>40.53</td>
<td>85.69</td>
<td>13.34</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chao1 (rarefied)</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1744.88</td>
<td>1263.18</td>
<td>1144.99</td>
<td>1105.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1291.66</td>
<td>1085.16</td>
<td>967.93</td>
<td>956.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1266.89</td>
<td>1085.16</td>
<td>856.33</td>
<td>956.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1233.55</td>
<td>1085.16</td>
<td>856.33</td>
<td>956.76</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unidentified OTUs</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td>--</td>
<td></td>
<td></td>
</tr>
<tr>
<td>213</td>
<td>164</td>
<td>83</td>
<td>64</td>
<td>77</td>
<td>21</td>
<td>53</td>
<td>47</td>
</tr>
</tbody>
</table>


A substantial proportion of the gOTUs were shared between all reefs in all size fractions of the metabarcoding (Table 3). For the reef survey, plates’ photo analysis and the 500 – 2000 μm size fraction of the metabarcoding dataset, more gOTUs were shared between JH and JS1 than there were shared between either of those reefs and JS2. For the other two fractions of the metabarcoding (sessile and 106 – 500 μm), the reefs JH and JS2 presented higher numbers of shared gOTUs. In terms of gOTUs the majority were shared between all size fractions (961 gOTUs relating to ~56%) with the highest number of unique gOTUs belonging to the sessile fraction (228 gOTUs which is ~13% of the total number of gOTUs) (Figure 3). In total there were ~ 19% (320 gOTUs) which were only found in the mobile size fractions.

Figure 3: Number of unique and shared gOTUs amongst the various size fractions of the metabarcoding. Samples were rarefied at 69,600 reads per sample (n=100).
Table 3: Number of OTUs shared between reefs for the different methods and size fractions and also the number of shared gOTUs among different size fractions (sessile, 106 - 500 μm and 500 - 2000 μm) in the metabarcoding (samples rarefied at 69,600 reads, n=100).

<table>
<thead>
<tr>
<th>Locations</th>
<th>Locations comparisons</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>JH</td>
</tr>
<tr>
<td>Reef</td>
<td>41</td>
</tr>
<tr>
<td>Photos</td>
<td>70</td>
</tr>
<tr>
<td>Metabarcoding - Combined Fractions</td>
<td>1388</td>
</tr>
<tr>
<td>Metabarcoding - Sessile</td>
<td>993</td>
</tr>
<tr>
<td>Metabarcoding - 106-500 μm</td>
<td>1006</td>
</tr>
<tr>
<td>Metabarcoding - 500 - 2000 μm</td>
<td>916</td>
</tr>
</tbody>
</table>
328 Composition of benthic assemblages

329 Overall, the standard reef surveys identified morphOTUs belonging to a total of 6
330 phyla with Cnidaria accounting for the majority of morphOTUs (83% of total
331 morphOTUs) (Table S1). The standard photo-transects reef surveys resulted in a
332 total of 53 morphOTUs comprising soft corals (Xeniidae - genera Lobophyton and
333 Sinularia), turf algae, and the scleractinian corals Porites (massive form),
334 Pocillopora, Goniastrea, Pavona, Acropora (branching form), and Echinopora
335 (Figure 4a). With the exception of Lobophytum (which was absent from JS1) all
336 the dominant taxa were present across the reefs surveyed.
Figure 4: Composition of the assemblages at the phyla level for a) standard visual reef survey; b) photo analysis of ARMS plates; c) metabarcoding analysis of the sessile fraction; d) metabarcoding analysis of the 106 – 500 μm fraction; and e) metabarcoding analysis of the 500 – 2000 μm fraction. For a) and b) percentages do not add up to 100% due to the occurrence of non-biological substrates. “Other” comprises those OTUs which could not be reliably assigned to a single taxa including turf algae in a) and b) and those reads that were not assigned to the phyla level in the metabarcoding.

Photo analysis of the ARMS plates resulted in a slightly higher number of phyla being identified (9 phyla) with Porifera and Rhodophyta contributing with the highest number of morphOTUs (28 and 18 morphOTUs respectively), followed
A photo analysis of ARMS - plates 1, 4 and 8 (both surfaces) identified a total of 109 taxa with the assemblages being dominated by Porifera, Bryozoa, algae (including crustose coralline algae – CrCA - and turf algae), and Polychaeta (Figure 4b). However, in several replicate units, a high proportion of the total coverage could not be assigned to any taxa (either because they could not be assigned to a single taxa or the PVC was uncolonised).

Table S1: Numbers of OTUs belonging to various phyla for all methods and size fractions for individual reefs and as a total of all reefs for each method. The category “Other” accounts for those OTUs that could not be attributed to a single phylum.

The metabarcoding dataset resulted in the highest number of identified phyla in all size fractions (36, 39 and 38 for the sessile, 106 – 500 μm and 500 – 2000 μm fractions, respectively). In each of the size fractions, large numbers of Rhodophyta, Annelida and Playhelminthes were observed (Table S1). Phyla such as Stramenopiles, Dinophyta, Maxillopoda, which were missing from the other two methods, had relatively high numbers of observed gOTUs. A similar number of Cnidaria gOTUs were observed in the metabarcoding analysis and the reef surveys although the metabarcoding also included Hydrozoa and Scyphozoa as well as Anthozoa. In the metabarcoding dataset, in terms of read assignments, Annelida (mainly Polychaeta) were substantial contributors to the assemblage in all size fractions (Figure 4c-e). In the sessile and 500 – 2000 μm fractions, Porifera and Mollusca were the other main contributors. The only exception was observed for the sessile fraction in two replicates of sessile fraction of JS1 which had a high proportion of reads attributed to the unicellular microbial groups
Syndiniophyceae, Mamiellophyceae, Alveolata group RAD - A and Bacillariophyceae. In the 106 – 500 μm fraction, Platyhelminthes were observed across all reefs in substantial proportions, with a decrease in the contribution of Mollusca and Porifera compared with the other size fractions. In JH there was also a large contribution from Chordata.

Structure of benthic assemblages and relationship with environmental variables

PCO analysis revealed different patterns for the three analysed datasets (reef surveys, photo analyses of the plates, and the metabarcoding) (Figure 5). In both the reef survey and the photo analysis of the ARMS plates, no significant differences were observed among the reefs (Table S2) for either morphOTU occurrence (Jaccard) or abundance (Bray-Curtis). ANOSIM analysis of the metabarcoding dataset showed that all the reefs were significantly different based both on the presence/absence of gOTUs (Jaccard similarity) (Figure 5e and Table S2) and the abundance of reads (Bray-Curtis) (Figure 5f and Table S2). For the size fractions, the sessile component was different from the other size fractions (Table S2), with the 106-500 μm and 500 - 2000 μm size fractions showing no significant difference from each other.
Figure 5: PCO plots for a) reef survey using Jaccard similarity; b) reef survey using Bray–Curtis similarity; c) photo analysis of plates using Jaccard similarity; d) photo analysis of plates based on Bray–Curtis similarity; e) metabarcoding analysis using Jaccard similarity, and f) metabarcoding analysis based on Bray–Curtis similarity.

Table S2: Statistical results from the ANOSIM analysis of the methods using Jaccard and Bray-Curtis distance matrices. Results in bold are significant.

Comparative (Mantel-type) tests were undertaken on the similarity matrices resulting from each dataset based on Jaccard and Bray-Curtis indices to analyse the consistency of the biodiversity patterns. The reef survey had no similarity to
the datasets resulting from the ARMS units. The photo analysis of the plates identified a similar distribution to the 106 – 500 μm size fraction of the metabarcoding component (Table 4) whilst all size fractions of the metabarcoding showed similar distributions.

Table 4: Results from the comparative tests (comparative mantel tests using RELATE (Spearman’s rank method) in PRIMER) between reefs using both Jaccard and Bray-Curtis distance matrices. Comparisons in bold are significant

<table>
<thead>
<tr>
<th></th>
<th>Jaccard</th>
<th>Bray-Curtis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>statistics</td>
<td>P</td>
</tr>
<tr>
<td>Coral Reef surveys</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Photos</td>
<td>0.022</td>
<td>0.429</td>
</tr>
<tr>
<td>Combined ARMS</td>
<td>0.162</td>
<td>0.24</td>
</tr>
<tr>
<td>Sessile</td>
<td>0.18</td>
<td>0.24</td>
</tr>
<tr>
<td>106-500</td>
<td>0.311</td>
<td>0.141</td>
</tr>
<tr>
<td>500-2000</td>
<td>-0.156</td>
<td>0.735</td>
</tr>
<tr>
<td>Photo_plates</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Combined ARMS</td>
<td>0.152</td>
<td>0.196</td>
</tr>
<tr>
<td>Sessile</td>
<td>-0.031</td>
<td>0.549</td>
</tr>
<tr>
<td>106-500</td>
<td><strong>0.374</strong></td>
<td><strong>0.013</strong></td>
</tr>
<tr>
<td>500-2000</td>
<td>0.217</td>
<td>0.111</td>
</tr>
<tr>
<td>Sessile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>106-500</td>
<td><strong>0.462</strong></td>
<td><strong>0.044</strong></td>
</tr>
<tr>
<td>500-2000</td>
<td><strong>0.486</strong></td>
<td><strong>0.014</strong></td>
</tr>
</tbody>
</table>

**Discussion**

In the current scenario of environmental change due to natural and anthropogenic pressures, it is vital to have a sound understanding of biodiversity patterns across multiple spatial and temporal scales. With the current rates of biodiversity loss, it is becoming increasingly necessary to improve survey protocols for biodiversity assessments. Commonly, the number of species
present in a certain area or habitat is used to investigate changes in the biodiversity patterns. However, many studies examining diversity patterns in coral reefs, have been undertaken using visual census of the exposed area alone, thus ignoring cryptic benthic fauna that due to their small size and/or location are not identified during reef surveys. In fact, numerous species, especially small invertebrates, have yet to be described, and may be being pushed to extinction before their roles in ecosystem functioning can be understood (Costello and Wilson 2011). There are large discrepancies between morphological and genetic surveys of eukaryotic diversity, with morphological studies biased towards taxa containing larger organisms (Pawlowski et al., 2012). Indeed, the microbial component of coral reefs is generally neglected despite influencing biogeochemical and ecological processes within the reef system (Ainsworth et al., 2010).

The current study highlighted the limitation of visual census protocols in coral reef biodiversity assessments that focus only on the exposed benthos. Although the addition of the photo analysis of the ARMS plates complemented that of the reef survey by revealing some of the visually cryptic species, it was observed that the metabarcoding of the sessile fraction of the plates provided a substantially higher number of OTUs compared with the photo analysis of the same plates. A large proportion of this discrepancy was due to the identification of microbial phyla by the metabarcoding survey that could not be identified visually. Also, OTUs from mobile taxa (e.g. Platyhelminthes, Malacostraca and Maxillopoda) contributed a high number of OTUs to the sessile fraction (although a low proportion of reads). The presence of mobile fauna in the sessile fraction could be a result of organisms not being detached during processing (including broken
fragments of individuals) or the production of extracellular DNA by these organisms that was detected by the sensitivity of the metabarcoding approach. Either way, the visual analysis would be unlikely to detect the presence of these motile groups, due to their small sizes. However, some of the discrepancies observed in the present study between the photo analysis of the plates and the metabarcoding of the sessile fraction may be partially related to the use of a subset of plates in the photo analysis. The metabarcoding approach took into account the scrapings from all the plates instead of just the subset, but the current study did not ascertain the scale of this bias. Another potential source of discrepancy are the limitations associated with the identification of taxa based on the pictures. Identification keys for the Red Sea are often inadequate, and taxonomic expertise is limited, thus, sometimes it is hard to distinguish between some groups like Porifera, Asciidiacea and Bryozoa. For the photo analysis of the plates, it is, however, likely that several experts will be required for the morphological classification of taxa to low taxonomic levels (e.g. genus or species). This would require a re-analysis of the photos, when taxonomies are updated, which is in itself time consuming and expensive. Similar issues are observed with the reef survey exacerbated by the lower quality images resulting from the 1 m² frames used during the survey. The metabarcoding approach is less dependent on the observer, making comparisons between datasets generated from different research groups more accurate. There are limitations to the classification of molecular data. Current reference databases are incomplete (Carugati et al., 2015) and for those taxa, which are represented there is low confidence in the reliability of taxonomic assignments past the class or family level (Cowart et al., 2015; Bik et al., 2012). However taxonomic assignments can
quickly be repeated against updated reference databases.

In contrast to the previous ARMS study undertaken by Leray and Knowlton (2015), the current study investigated the eukaryotic fraction using primers targeted to the 18S rDNA. There are limitations to using a multicopy gene such as the 18S rDNA (e.g. differences between copies within species, (Galluzzi et al., 2009) or different copy numbers between species (Prokopowich et al., 2003)). However, it has the advantage over the mt COI gene in that it can currently provide a more comprehensive survey of protists, which are an important component of coral reef environments, due to its previous usage in microbial molecular studies (Prokopowich et al., 2003). Eukaryotic microbial species (e.g. Bacillariophyta, Dinophyceae and Syndiniophyceae) comprised a large proportion of the sessile community in two of the replicates of JS1 ARMS units, which, as expected, had been completely neglected in the photo plate analysis. Beyond the symbiosis of zooxanthellae (genus *Symbiodinium*) with corals, the contribution of other microbial eukaryotes (as well as bacteria and Archaea) has scarcely been studied for coral reefs (Knowlton and Rohwer 2003). Nevertheless, they are likely to play vital roles in biogeochemical and ecological processes within a coral reef such as food webs, organism life cycles, and chemical and nutrient cycling (Ainsworth et al., 2010). Indeed, chemical cues from microbial biofilms (prokaryotic communities) have been shown to influence the settlement of larvae of many keystone species such as corals and sea urchins (Webster et al., 2004; Huggett et al., 2008). Also, some endolithic algae have been suggested to play a crucial role in coral survival during bleaching events allowing corals to survive until a normal complement of zooxanthellae is restored (Fine and Loya 2002).
Another relevant advantage of using ARMS units is the ability of analysing reef biodiversity of mobile specimens with different size-ranges. The 106 - 500 μm and 500 - 2000 μm size fractions are disregarded in traditional reef surveys as they are hard to distinguish during visual reef surveys but are vital components of coral reef food webs (e.g. (Glynn 2004)). Besides, sessile and vagile epibenthic assemblages colonizing artificial units in rocky areas have been found to present different colonization and succession patterns (Moura et al., 2008). Despite the use of different primers, we found, similar to (Leray and Knowlton 2015), that the motile fractions (106 – 500 μm and 500 – 2000 μm) clustered together more closely compared with the sessile fraction. Also the 106 – 500 μm size fraction had a higher diversity compared with the 500 – 2000 μm fraction. This could be due to the presence of juvenile stages of the 500 – 2000 μm fraction being present in the smaller size fraction (106 – 500 μm) thus increasing the diversity. However, caution has to be taken in the analysis of these motile size fractions because ~ 56% of gOTUs were shared amongst all three size fractions with both mobile fractions containing sessile gOTUs (e.g. Porifera), which are likely to be byproducts of the processing. Fragments from the sessile community could have been detached as the ARMS units were disassembled and were subsequently retained on the sieves whilst, conversely, extracellular DNA from mobile organisms may have been retained on the plates.

The community composition of ARMS units could be affected by a variety of either biological (e.g. grazing (Fitzhardinge and Bailey-Brock 1989), chemical cues from biofilms (Morse and Morse 1996), seasonality of species (Cummings 1994)) or physical conditions (e.g. disturbances (Cummings 1994), sedimentation (Rodgers 1990), salinity (Hédouin et al., 2015), nutrients
(D’Angelo and Wiedenmann 2014) or turbidity (Bell 1992) during the colonisation of the plates. In the current study, the ARMS units were in situ for one year, which could mean the community inhabiting the ARMS unit was still immature compared with the surrounding reef. For the ARMS units to reach maturity, an age-class structure and 3D structure similar to the surrounding natural community would have to be created (Gittings et al., 1993). Carter & Prekel (2008) showed that diversity progressively increased on an artificial reef between 9 and 36 months without reaching equilibrium. Stable communities may take many years to reach with previous studies comparing benthic assemblages colonizing artificial reef units and natural surrounding reefs, finding that even for the same organism size-ranges and after 16 years of colonization, assemblages were still different (Carvalho et al., 2013). Compared with the reef survey data (mature community) there were substantially fewer cnidarian OTUs and lower percentage abundance identified on the ARMS units when analysed using photo analysis. Although the metabarcoding revealed similar numbers of cnidarian OTUs compared to the reef survey, a large proportion of the OTUs were not attributed to Anthozoa and relative abundance was low. This may partly be due to a lack of time for corals to establish themselves on the substrate, as coral larvae often remain inconspicuous upon settlement (Edmunds 2000). Carter & Prekel (2008) attributed this fact to the reason why they observed differences in their study before and after 12 months post-deployment. However, the fact that the metabarcoding also revealed a smaller number of Anthozoa gOTUs suggests that coral larvae may not be substantial components of the cryptic fauna (Wunsch et al., 2000). Also ARMS plates may not be suitable for coral larval recruitment due to, amongst other
reasons, incompatible light regimes (Maida et al., 1994), substrate characterisation (including chemical cues from biofilms (Morse and Morse 1996)) or competition from other biotic assemblages (e.g. sponges (González-Rivero et al., 2011), macroalgae (McCook et al., 2014), and bryozoan (Fine and Loya 2003)). Although relative abundance was low in the metabarcoding results, both the metabarcoding and reef surveys showed Pocillopora, Porites and Xenia to be the dominant genera suggesting that as with previous studies the surrounding biota is important for recruitment (Bohnsack 1991; Gittings et al., 1993; Spieler et al., 2001).

Using comparative tests, we showed that the ecological patterns were in general not consistent amongst the different sampling methods. While similar patterns were obtained from the different size fractions of the metabarcoding, only the patterns detected for the 106 - 500 μm size fraction was similar to the photo analysis of the plates. However this finding is not entirely surprising considering that the methods vary in scale, size-range of organisms and level of taxonomic identification. While it would be thought that the photo analysis of the plates and the metabarcoding of the sessile fraction (derived from the scraping of the plates) may give similar patterns, there are still substantial differences in the analytical procedures (e.g. use of a subset of plates in the photo analysis). For example, metabarcoding allows for the identification of microbial organisms that will be ignored using other approaches. The lack of taxonomists and updated taxonomic keys, as well as the limitations associated with identifying benthic groups from photographs, likely contributed to the discrepancies between survey methods. Often, morphological studies and monitoring programs focus on the abundance of each species, however this is a drawback with molecular
techniques where abundance cannot be determined, especially for multicellular organisms. Primer mismatches and variations in gene copy numbers between species (Prokopowich et al., 2003) can bias estimations but the number of reads appears to be more correlated with morphological biomass (Lindeque et al., 2013) or dry weight (Hirai et al., 2015) than abundance. This currently limits the use of sequencing efforts for applications such as monitoring.

In summary, a combination of ARMS and reef surveys, can contribute to a more comprehensive knowledge of the diversity of reef systems as different biological components of the reef may provide different patterns. While reef surveys can only assess the large establish flora and fauna of the system, they can be undertaken with a higher frequency than ARMS analysis, as these units need to be in situ for a year or longer. On the other hand, the advantage of the ARMS units is that they are uniform, in place for the same time and thus comparisons, across a large temporal and spatial scale, can more easily be made. Also, the analysis of ARMS units can allow the identification of not only a larger range of phyla but also cryptic benthic fauna and protistan biofilms that are important components of the reef community which would be difficult to sample in the mature reef without using destructive methods. Also, the current study indicates that the metabarcoding approach was able to discriminate differences in the biodiversity of the reefs whereas the visual methods were unable to. This could suggest that finer scale differences can be observed when using molecular methods possibly due to the increased breadth of taxonomy analysed. As monitoring moves from the use of indicator species to a more comprehensive analysis of alpha and beta diversity including multivariate statistical methods
(Hewitt et al., 2005), the combination of reef surveys and metabarcoding of ARMS could give a more accurate state of the health of the reefs at different spatial and temporal scales. Furthermore, future molecular studies using ARMS could target more than one region (e.g. a combination of 18S rRNA, mt COI and internal transcript spacer (ITS)) to have a more comprehensive understanding of the diversity, as this would mitigate some of the issues associated with primer biases (for example the uneven amplification of taxa (Cowart et al., 2015)). Also, molecular methods could not only investigate diversity but also the functional ability of the community using shotgun metagenomic techniques, as well as successional patterns during the establishment of communities on the ARMS units.

**Acknowledgement**

The authors would like to thank the skippers and crew from CMOR, KAUST who enabled the sampling to be performed. The authors are also indebted to the help given during sampling and laboratory work from Eva Aylagas Martinez and Miguel Viegas as well as Saskia Kürten for undertaking the nutrient analysis. We would also like to thank Ute Langner for the production of the map. Furthermore the authors would like to acknowledge the assistance provide in the improvement of this manuscript from the two reviewers. The research reported in this publication was supported by funding from King Abdullah University of Science and Technology (KAUST). This study was undertaken within the framework of the project DEVOTES (EU - ENV.2012.6.2-3 - Grant agreement 308392; http://www.devotes-project.eu).
References


Carvalho S, Moura A, Curdia J, Cancela da Fonseca L, Santos MN. How


Galluzzi L, Bertozzini E, Penna A, Perini F, Garcés E, Magnani M. Analysis of rRNA
gene content in the Mediterranean dinoflagellate *Alexandrium catenella* and
*Alexandrium taylori*: implications for the quantitative real-time PCR-based

Ginsburg RN. Geological and biological roles of cavities in coral reefs. Barnes DJ,

Gittings SR, Bright TJ, Hagman DK, Walton Smith FG. The M/V Wellwood and
other large vessel groundings: coral reef damage and recovery. Case Histories for
the Colloquium and Forum on Global Aspects of Coral Reefs: Health, Hazard and
History. 1993; 174-180

Glynn PW. High Complexity Food Webs in Low-diversity Eastern Pacific Reef-

Glynn PW, Enochs IC. Invertebrates and Their Roles in Coral Reef Ecosystems. In
Coral Reefs: An Ecosystem in Transition (pp. 273–325). Dordrecht: Springer
Netherlands. 2010. doi: 10.1007/978-94-007-0114-4_18

González-Rivero M, Yakob L, Mumby PJ. The role of sponge competition on coral
reef alternative steady states. Ecological Modelling. 2011; 222(11): 1847-53. doi:

Vargas C, Decelle J, Del Campo J, Dolan JR, Dunthorn M, Edvardsen B, Holzmann
catalog of unicellular eukaryote Small Sub-Unit rRNA sequences with curated taxonomy. Nucleic Acids Research. 2013; 41; D597-D604


R development core team, 2015


Wickham H, Chang W. ggplot2. R package version 2.0.0. 2015


• Wide range of diversity observed with molecular techniques
• Molecular and visual techniques did not show similar patterns
• Autonomous Reef Monitoring Structures (ARMS) complement current survey techniques