Environmental Pollution 203 (2015) 165–174

Contents lists available at ScienceDirect

Environmental Pollution

journal homepage: www.elsevier.com/locate/envpol

Metabarcoding of benthic eukaryote communities predicts the ecological condition of estuaries

Anthony A. Chariton [a, *], Sarah Stephenson [a], Matthew J. Morgan [b], Andrew D.L. Steven [c], Matthew J. Colloff [b], Leon N. Court [b], Christopher M. Hardy [b]

[a] CSIRO Oceans and Atmosphere, Locked Bag 2007, Kirrawee, NSW 2232, Australia
[b] CSIRO Land and Water, GPO Box 1700, Canberra, ACT 2601, Australia
[c] CSIRO Oceans and Atmosphere, GPO Box 2583, Brisbane, QLD 4001, Australia

ARTICLE INFO

Article history:
Received 7 December 2014
Received in revised form
25 March 2015
Accepted 29 March 2015
Available online 20 April 2015

Keywords:
Biomonitoring
Metabarcoding
Sediments
DNA
Eukaryotes
High-throughput sequencing
18S rRNA
Indicator taxa
Threshold analysis

ABSTRACT

DNA-derived measurements of biological composition have the potential to produce data covering all of life, and provide a tantalizing proposition for researchers and managers. We used metabarcoding to compare benthic eukaryote composition from five estuaries of varying condition. In contrast to traditional studies, we found biotic richness was greatest in the most disturbed estuary, with this being due to the large volume of extraneous material (i.e. run-off from aquaculture, agriculture and other catchment activities) being deposited in the system. In addition, we found strong correlations between composition and a number of environmental variables, including nutrients, pH and turbidity. A wide range of taxa responded to these environmental gradients, providing new insights into their sensitivities to natural and anthropogenic stressors. Metabarcoding has the capacity to bolster current monitoring techniques, enabling the decisions regarding ecological condition to be based on a more holistic view of biodiversity.

1. Introduction

The increasing human population and its activities are having pronounced deleterious effects on the ecological condition of the world's estuaries (Rabalais et al., 2009). These activities degrade the physical environment and modify the chemical composition of the water column and sediments and their associated biota (Davis and Koop, 2006). Ultimately, such activities are expressed as distinct changes in ecological composition and function (Dauer et al., 2000; Hooper et al., 2012). In Australia, more than 85% of the population of 22 million live within 50 km of the coast (ABS, 2003). With a population increase of 82% projected by 2056 (ABS, 2003), the pressures on estuarine environments in rapidly developing coastal regions such as southeast Queensland are likely to increase markedly.

In order to mitigate the pace of environmental degradation, fundamental information on the chemical, physical and ecological characteristics and components of estuaries is required. However, many of the variables which drive the ecology of estuaries are difficult to define and vary greatly across space and time (Morrisey et al., 1992; Wiens, 1989; Ysebaert and Herman, 2002). The most commonly monitored ecological component of estuaries is the macrobenthos, with many studies demonstrating its responsiveness to a range of natural and anthropogenic variables (Johnston and Roberts, 2009). This approach can lead to management decisions being made on the assumption that the macrobenthos accurately represents overall ecological condition (Chariton et al., 2010a), despite the knowledge that the meio- and microbiota are far more species-rich, have a greater diversity of life-histories and ecological niches, and are often more responsive to environmental change (Austen and Warwick, 1989; Kennedy and Jacoby, 1999). The inclusion of ecological data derived from these elements of the biota would provide a more representative, informative ecological picture.

In recent years, there have been considerable advances in applying DNA-based diversity methods using high-throughput

* Corresponding author.
E-mail address: Anthony.Chariton@csiro.au (A.A. Chariton).

http://dx.doi.org/10.1016/j.envpol.2015.03.047
0269-7491/

sequencing (Baird and Hajibabaei, 2012; Taberlet et al., 2012a), commonly referred to as metabarcoding (Taberlet et al., 2012b). Metabarcoding provides previously unattainable insights into communities and ecosystems, aiding our understanding of them. The approach has proven especially useful in deriving compositional data from samples containing organisms that are difficult to indentify because of small size, cryptic habits, their occurrence in the form of propagules (e.g. spores and zoocysts) or a lack of traditional identification keys (Medinger et al., 2010; Valentini et al., 2009).

In this study, we used high-throughput sequencing of the 18S rRNA gene to examine sub-macro benthic biotic composition of five estuaries in eastern Australia. These estuaries have been routinely sampled since 2000 as part of a larger monitoring program, but due to time and cost constraints the ecological health of these estuaries is determined using abiotic surrogates rather than ecological data. Our initial aim was to examine whether DNA-based eukaryotic composition could differentiate between estuaries. Secondly, we explored the relationships between the eukaryotic communities and environmental gradients observed among the estuaries. Finally, we examined whether metabarcoded eukaryotic data has the potential to produce relevant ecological information which can be used to further develop DNA-based approaches for the routine monitoring of estuarine sedimentary environments.

2. Methods

2.1. Study region

In February 2010, we sampled five estuaries (Noosa, Maroochydore, Pine, Logan and Currumbin) in south-east Queensland, Australia (Fig. 1). All estuaries are monitored monthly by the Queensland Government as part of the Ecosystem Health Monitoring Program (EHMP) (http://www.health-e-waterways.org). The EHMP includes an Ecological Health Index based on algal productivity derived from measurement of chlorophyll *a*, concentrations of dissolved oxygen, major nutrients and turbidity. The conformation of these variables with national guideline values, together with estimates of seagrass and riparian vegetation cover is used to develop an annual report card for each estuary (Table S1 in Supplementary material).

The five estuaries were located no more than 190 km apart (Fig. 1) and represent a range of ecological conditions (Table 1). There were large differences in morphology between the five estuaries, with the Currumbin Creek considerably smaller than the others (Table 1). Within each estuary, five sites were sampled, and in all but two cases (one in each the Noosa and Currumbin) the sites have been routinely monitored under the EHMP.

2.2. Collection and analysis of environmental variables

Sediment collection was confined to non-sandy substrates. Five sediment samples were collected at each site from ca. 2 m below low water using a Van Veen grab. Sub-samples were taken from the surficial layer (1.5–2 cm) of each sample for DNA, grain size and total organic carbon analysis. All samples for DNA analysis were transferred into clean 50 mL Greiner tubes and placed on ice immediately, then frozen within 6 h of collection and thawed only just prior to DNA extraction. All materials used for the collection and storage of DNA samples were pre-rinsed for at least 24 h in 5% sodium hypochlorite, and rinsed thoroughly five times with Milli-Q water (Millipore, Academic Water Systems, Australia). To minimise cross contamination, sediments were only sub-sampled from the centre of each grab sample.

The physico-chemical properties of the water column were measured at each sampling site approximately 0.5 m above the sediment surface using a calibrated YSI 6920 multi-sonde. At all EHMP sites (excluding one site in both the Noosa and Pine), water samples were collected for nutrient and chlorophyll *a* analyses. Water samples for nutrient analysis were filtered upon collection, with the filtrate stored in clean foil-wrapped containers stored on ice. Total phosphorus, filterable reactive phosphorus, total nitrogen, organic nitrogen, inorganic nitrates, ammonia and chlorophyll *a* analyses was performed using standard methods (Clesceri et al., 1998). Total organic carbon (TOC) and grain size analysis for the following grain size classes: <63 μm (fines), 0.63 μm–1 mm (sand), >1 mm (coarse) were performed as previously described (Chariton et al., 2010c).

2.3. DNA extraction, amplification and sequencing

DNA was extracted from 1.5 g of sediment and purified using UltraCleanSoil DNA extraction kits (MO BIO, Carlsbad, CA) following the manufacturer's protocols. In addition to the sediment samples, three internal reference samples containing sixteen clones from a range of eukaryotic taxa were also processed, as previously described (Morgan et al., 2013). Polymerase chain reaction (PCR) amplification of a 200–500-bp fragment of the 18S rRNA gene was carried out with the 'universal' primers All18SF-TGGTGCATGGCCGTTCTTAGT and All18SR-CATCTAAGGGCATCACAGACC (Hardy et al., 2010), and sample preparation was conducted as previously described (Baldwin et al., 2013). Sequencing was performed by the Australian Genome Research Facility (St Lucia, Queensland) using a single plate of Roche 454 GS FLX Titanium. Demultiplexing and the removal of potential PCR artefacts, sequencing errors and chimeric sequences were performed using the Amplicon Pyrosequence Denoising Program (APDP) (Morgan et al., 2013). Taxon identification of each unique sequence, herein referred to as a Molecular Operational Taxonomic Unit (MOTU) was inferred using the RDP classifier with the SILVA 18S rRNA database (release 113) (www.arb-silva.de).

2.4. Statistical analysis

As there is a weak statistical relationship between the number of sequence reads and organism biomass or abundance (Egge et al., 2013), all MOTU data were converted to presence or absence prior to computation (Chariton et al., 2014). Ordination of MOTU data was performed by non-metric multidimensional scaling (nMDS) using the Jaccard similarity coefficient in the Primer 6 + statistical package (Plymouth Marine Laboratory, UK). Statistical differences between estuaries were tested by a two-factor permutational multivariate analysis of variance (PERMANOVA), with 'sites' nested within 'estuary'. Differences between treatments were identified by pairwise *a posteriori* tests based on 9999 random permutations. The proportions of explained variation at spatial scales of estuary, site and residual were calculated using the procedure described by Quinn and Keough (2002). Differences in richness of total MOTU and dominant taxonomic groups were examined using a two-factor nested ANOVA. Residuals were assessed for skewness, kurtosis, and omnibus normality using D'Agostino's tests (D'Agostino et al., 1990) with homogeneity of variances examined using a modified Levene equal variance test (Levene, 1960). When assumptions of homogeneity were violated, appropriate transformations were performed (Sokal and Rolf, 1995). In cases in which the data remained heteroscedastic, the level of statistical significance was set at $P < 0.01$. All ANOVAs were performed in NCSS v8 (NCSS, Kaysville, UT). MOTUs indicative of each estuary and combinations of estuaries were identified using the R package *Indispecies*, with Indictor Values (IV) reflecting both the conditional probability of the MOTU as an indicator of a particular estuary and the probability of finding