Ecological Indicators 85 (2018) 440–450

Contents lists available at ScienceDirect

Ecological Indicators

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

Research paper

Can high-throughput sequencing detect macroinvertebrate diversity for routine monitoring of an urban river?

Melissa E. Carew*, Claudette R. Kellar, Vincent J. Pettigrove, Ary A. Hoffmann

Centre for Aquatic Pollution Identification and Management (CAPIM), School of BioSciences, Bio21 Institute, 30 Flemington Rd, The University of Melbourne, 3010, Victoria, Australia

ARTICLE INFO

Keywords:
Macroinvertebrates
DNA barcoding
High-throughput DNA sequencing
Biological monitoring
Metabarcoding
Chironomidae

ABSTRACT

Macroinvertebrate families identified through morphological examination have traditionally been used in routine assessment of freshwater ecosystems. However, high throughput DNA sequencing (HTS) promises to improve routine assessment by providing rapid and cost-effective identification of macroinvertebrate species. In freshwater ecosystems in urbanised areas where family diversity is often low, new insights into ecosystem condition and impacting factors are likely through species-level assessments. Here we compare morphological identification to HTS based identification of macroinvertebrate families by considering 12 sites in an urban river system. Some taxa detected morphologically were not detected by HTS and vice versa. However, this had only a small impact on computed family-level metrics of ecological condition. We detected multiple species using HTS in the Chironomidae, Coenagrionidae, Hydrobiidae, Leptoceridae, Ceratopogonidae, Corixidae, Veliidae, Oligochaeta and Acarina. The highest species diversity was found in the Chironomidae, and for many of these species we had prior knowledge of their likely pollution sensitivity. In the Chironomidae, we showed that species level data was congruent with expectations based on measured levels of pollutants at sites and other family level metrics. Importantly, we also identified many species in the same family that differed in their distribution and likely pollution sensitivity in this urban river system. Therefore, HTS provided similar levels of information to traditional methods at the family level, but also generated new information for more accurate condition monitoring at the species level.

1. Introduction

Freshwater biological monitoring has traditionally been based on mostly family level identification of macroinvertebrates using morphological characteristics (Barbour et al., 1999; Rosenberg and Resh, 1993). Currently, well-established protocols exist in many countries to sample macroinvertebrate family biodiversity and compare it with environmental objectives set in policy (Barbour et al., 1992; Bonada et al., 2006; King and Richardson, 2003; Metcalfe, 1989; Metzeling et al., 2006). However, DNA technologies that allow rapid species-level identification can change how macroinvertebrate biodiversity is used to assess freshwater ecosystems (Baird and Hajibabaei, 2012; Carew et al., 2013; Hajibabaei et al., 2011; Jackson et al., 2014; Pfrender et al., 2010). DNA based approaches, like DNA barcoding, can recognise more species compared to morphological identification (Hebert et al., 2003), as morphologically cryptic, immature taxa or those that represent little known groups can easily be distinguished in samples (Pilgrim et al., 2011; Shackleton and Rees, 2015; Sweeney et al., 2011).

By identifying macroinvertebrate species, species-specific responses to pollutants and environmental characteristics can be used as bioindicators (Baird and Hajibabaei, 2012). This is especially useful when species within families vary in environmental responses (Bailey et al., 2001; Carew et al., 2011; Lenat and Resh, 2001; Resh and Unzicker, 1975). For example, Chironomidae species can vary greatly in their response to different environmental stressors, and show specific responses to certain pollutants (Carew et al., 2007; Marshall et al., 2010; Pettigrove and Hoffmann, 2005b; Sharley et al., 2008). By considering these groups only at higher levels of taxonomic resolution, valuable environmental signal is lost. There is greater sensitivity in distinguishing between impacting factors where species level identification using DNA barcoding are performed (Jackson et al., 2014; Sweeney et al., 2011). Species level identifications could provide new sensitive and diagnostic monitoring tools for environmental condition in freshwater systems.

While substantial effort and resources are needed to individually DNA barcode macroinvertebrate species, high-throughput DNA sequencing (HTS) transforms DNA barcoding into a cost-effective tool for routine identification (Baird and Hajibabaei, 2012; Carew et al., 2013;

* Corresponding author.
E-mail address: mecarew@unimelb.edu.au (M.E. Carew).

http://dx.doi.org/10.1016/j.ecolind.2017.11.002
Received 28 July 2016; Received in revised form 26 October 2017; Accepted 1 November 2017

Stein et al., 2014). Bulk DNA extraction and simultaneous processing of multiple samples enables rapid and accurate detection of species. However, successfully identifying species detected with HTS relies on DNA barcoding libraries with comprehensive coverage of accurately identified species (Pfrender et al., 2010; Shackleton and Rees, 2015; Webb et al., 2012). When comprehensive DNA barcoding libraries are available, species detection using HTS and subsequent identification is relatively straightforward (Carew et al., 2013; Hajibabaei et al., 2011, 2012). However, it is more difficult when DNA barcoding libraries are incomplete, as PCR artefacts can be confused with new species if there are no DNA barcodes from voucher specimens available (Creer et al., 2010; Tedersoo et al., 2010). DNA barcoding libraries need to be developed alongside HTS to turn this approach into a practical tool (Ekrem et al., 2007; Vivien et al., 2015; Zimmermann et al., 2014).

A challenge in using HTS in routine assessment is potential bias in detecting taxa. In multi-family HTS analysis of invertebrate samples or 'environmental barcoding', around 90% of species in samples are detected (Gibson et al., 2014; Hajibabaei et al., 2011, 2012). However, some taxonomic groups might be overrepresented in the fraction of species not detected. Moreover, additional taxa can also be detected through HTS through amplification of egg masses, tissue fragments, gut content or environmental DNA present in invertebrate samples (Gibson et al., 2014; Hajibabaei et al., 2011). While taxa originating from these sources could be included in environmental assessments as they often reflect resident invertebrates, others might need to be excluded, such as tissues transported aerially into a site. These potential biases need to be examined by comparing the results of an HTS-based approach with those obtained using morphological approaches.

In this study, we test how HTS performs in routine assessment of sites along an urban river system. These systems are likely to benefit from species-level identification, as macroinvertebrate communities in urban environments can lack high family diversity and be dominated by a few speciose groups with diverse environmental responses, such as the Chironomidae and Oligochaeta (Beavan et al., 2001; Carew et al., 2007; Rosenberg, 1992; Vivien et al., 2015). Furthermore, factors responsible for deterioration of aquatic and riparian habitats in urban river systems can be difficult to identify because of multiple impacting factors, such as invasion by exotic species, changes to hydrology, and increased water temperatures, turbidity and contamination (Ellis and Hvitved-Jacobsen, 1996; Marsalek, 1998; Paul and Meyer, 2001; Pettigrove and Hoffmann, 2005b). If species-specific stressor responses could be included in assessments, they may assist in identifying the dominant factors responsible for environmental degradation.

We test HTS in routine assessments of environmental condition by first applying three sets of DNA barcoding primers to bulk-extracted macroinvertebrate samples, and examine their ability to detect species within morphologically identified families. We compare family diversity metrics calculated from morphologically- and HTS- derived data sets, and examine how HTS-derived data compares to existing protocols and objectives for routine assessment. We then explore species diversity within families and whether species level identification based on DNA barcodes provides additional information. Finally, we draw on a database of Chironomidae species distributions along environmental gradients to determine if species composition is diagnostic of pollution levels at sites.

2. Materials and methods

2.1. Study location

The Maribyrnong River is located in the Greater Melbourne Area, Victoria, Australia (Fig. 1). It is ~180 km long flowing south from the Great Dividing Range to the Yarra Estuary. The upper reaches are in rural areas, while areas surrounding the lower reaches are mostly urbanised with some industrialised areas. While the majority of riparian vegetation along the upper and middle reaches has been preserved, many tributaries are affected by urbanisation and development that has reduced riparian vegetation and water quality.

This study investigates 12 sites along the freshwater section of the lower Maribyrnong River system consisting of eight sites on the Maribyrnong River, three sites on Arundel Creek, and a single site on Taylors Creek, both tributaries (Table 1, Fig. 1). Arundel Creek receives discharge from Melbourne International Airport and sites were located upstream and downstream of the discharge outfall. The Taylors Creek site receives urban runoff from recently developed residential areas.

2.2. Macroinvertebrates surveys

Aquatic macroinvertebrates were collected from edge habitats according to the Rapid Bioassessment Protocols developed by EPA Victoria as part of their Guidelines for Environmental Management (available at: http://www.epa.vic.gov.au/~/media/Publications/604%201.pdf). Samples were enumerated and morphologically identified to family level, except that Acarina and Oligochaeta were identified to order, and Chironomidae to sub-family using a binocular dissecting microscope (Leica Microsystem and Instruments, Wetzlar, Germany). Whole specimens were sorted to order and the relevant family-level taxonomic keys where followed for each group (see Hawking, 2000)

Four standard biological metrics were calculated and were compared to the EPA biological objectives for Victoria in the State Environment Protection Policy (available at: http://www.epa.vic.gov.au/~/media/Publications/905.pdf). These measures included Number of Families, Key Families, the Stream Invertebrate Grade Number Average Level or 'SIGNAL' biotic index (Chessman, 1995) and AUSRIVAS predictive model scores and bands (Reynoldson et al., 1997). Sites MAR1 and MAR2 (Fig. 1) were compared to biological objectives for region B4, while the remaining sites were compared to biological objectives for 'urban streams and rivers in region B4′ (available at: http://www.epa.vic.gov.au/~/media/Publications/793%202.pdf). Key families were not compared to biological objectives as data from two season is required.

2.3. High throughput sequencing

For each sample, a leg or small amount of tissue from individuals was dissected and placed into a 1.5 ml microcentrifuge tube for bulk extraction. A Qiagen DNeasy® Tissue Kit (Qiagen, Hilden, Germany) was used to extract total genomic DNA from tubes of tissues following the manufacturer's protocol. A biological replicate of one sample (MAR8 CW) was used to examine the robustness of species identification. We also included a chironomid sample (BR08) as a control because all taxa in this sample had previously been identified individually to species level and were bulk sequenced using 454 pyrosequencing (see Carew et al., 2013). The purpose of this control was to examine data quality and test a bioinformatic pipeline for data analyses.

A two-step PCR process was used to obtain amplicons for Illumina MiSeq sequencing. The first PCR involved amplifying the three amplicons within the Cytochrome Oxidase I (COI) DNA barcode region (Folmer et al., 1994) using three overlapping primer sets. We used the COI DNA barcoding region as this is the standard region for performing species identification of animals (Hebert et al., 2003). Two sets A/E and B/F were taken from Hajibabaei et al. (2012) while a third set used B with a primer developed for this study- COIBrev (5′-GATCARACAA-AYARWGGYATWCGRTC-3′). The primer set LCOI1490 (Folmer et al., 1994)/MLepR2 (Hebert et al., 2013) was used if any of the first three sets of primers failed to amplify. Primers were selected based on their ability to amplify a broad range of taxa. Template specific primers had Illumina adaptors incorporated onto the 5′ end for the attachment of Nextera-XT illumina indexes (Illumina Corporation, San Diego, CA, USA) in the second round PCR.

First round PCR reactions contained 2 µl of DNA template, 16.4 µl molecular biology grade water, 2.5 µl PCR buffer, 1 µl $MgCl_2$ (50 mM),