Marine Genomics 25 (2016) 57–68

Contents lists available at ScienceDirect

Marine Genomics

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

Fishing for divergence in a sea of connectivity: The utility of ddRADseq genotyping in a marine invertebrate, the black-lip pearl oyster *Pinctada margaritifera*☆,☆☆

Monal M. Lal *, Paul C. Southgate [1], Dean R. Jerry, Kyall R. Zenger

[a] Centre for Sustainable Tropical Fisheries and Aquaculture (CSTFA), James Cook University, Townsville Campus, Townsville, QLD 4811, Australia
[b] College of Marine and Environmental Sciences (CMES), James Cook University, Townsville Campus, Townsville, QLD 4811, Australia

ARTICLE INFO

Article history:
Received 30 September 2015
Received in revised form 27 October 2015
Accepted 27 October 2015
Available online 3 November 2015

Keywords:
Population genomics
Aquaculture
SNP
Fine-scale differentiation
Colour morphotype
Fiji Islands

ABSTRACT

Population genomic investigations on highly dispersive marine organisms typically require thousands of genome-wide SNP loci to resolve fine-scale population structure and detect signatures of selection. This information is important for species conservation efforts and stock management in both wild and captive populations, as well as genome mapping and genome wide association studies. Double digest Restriction site-Associated DNA Sequencing (ddRADseq) is a recent tool for delivering genome wide SNPs for non-model organisms. However, its application to marine invertebrate taxa has been limited, particularly given the complex and highly repetitive nature of many of these organisms' genomes. This study develops and evaluates an optimised ddRADseq technique together with associated analyses for generating genome-wide SNP data, and performs population genomic analyses to inform aquaculture and fishery management of a marine bivalve, the black-lip pearl oyster *Pinctada margaritifera*. A total of 5243 high-quality genome-wide SNP markers were detected, and used to assess population structure, genome diversity, detect F_{st} outliers and perform association testing in 156 individuals belonging to three wild and one hatchery produced populations from the Fiji Islands. Shallow but significant population structure was revealed among all wild populations (average pairwise $F_{st} = 0.046$) when visualised with DAPC and an individual network analysis (NetView P), with clear evidence of a genetic bottleneck in the hatchery population ($N_{eLD} = 6.1$), compared to wild populations ($N_{eLD} > 192.5$). F_{st} outlier detection revealed 42–62 highly differentiated SNPs ($p < 0.02$), while case–control association discovered up to 152 SNPs ($p < 0.001$). Both analyses were able to successfully differentiate individuals between the orange and black tissue colour morphotypes characteristic of this species. BLAST searches revealed that five of these SNPs were associated with a melanin biosynthesis pathway, demonstrating their biological relevance. This study has produced highly informative SNP and population genomic data in *P. margaritifera*, and using the same approach promises to be of substantial value to a range of other non-model, broadcast-spawning or marine invertebrate taxa.

1. Introduction

Broadcast spawning marine organisms that undergo prolonged planktonic larval development can achieve high rates of connectivity over large spatial scales. As a result, levels of population genetic differentiation in these taxa are often low, but can also be biologically relevant (Limborg et al., 2012). Furthermore, selective forces can still impact local populations even with high rates of gene flow (Nielsen et al., 2009). In order to detect fine-scale variability and the presence of local adaptation which may be overlooked using traditional

☆ **Competing interests, ethics and role of funding source**: The authors declare that they have no competing interests. The work described herein has been carried out (where appropriate) in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki) for animal experiments. The funding source (the Australian Centre for International Agricultural Research) had no direct involvement in the study design, collection, analysis and interpretation of the data, nor the decision to submit this article for publication.
☆☆ **Author contributions**. MML carried out all tissue collections, laboratory bench work, participated in the investigation design and conceptualisation, developed modifications to the genotyping protocol, performed all data analyses and drafted the manuscript. PCS developed the broad project concept, participated in the investigation design, provided supervisory support, all project funding, advice on pearl oyster biology and ecology, and edited the manuscript. DRJ participated in the investigation design, provided supervisory support and edited the manuscript. KRZ participated in the design and conceptualisation of the project, provided statistical advice and technical input on investigation design, developed modifications to the genotyping protocol, provided supervisory support and edited the manuscript. All authors read and approved the final manuscript.

* Corresponding author.
E-mail addresses: monal.lal@my.jcu.edu.au (M.M. Lal), psouthgate@usc.edu.au (P.C. Southgate), dean.jerry@jcu.edu.au (D.R. Jerry), kyall.zenger@jcu.edu.au (K.R. Zenger).
[1] Current address: Faculty of Science, Health, Education and Engineering, University of the Sunshine Coast, Maroochydore, QLD 4558, Australia.

http://dx.doi.org/10.1016/j.margen.2015.10.010
1874-7787/

molecular markers, new genome-wide genetic resources incorporating thousands of Single Nucleotide Polymorphism (SNP) loci are typically required (Krück et al., 2013; Limborg et al., 2012; Nielsen et al., 2009; Palumbi, 2003; Pujolar et al., 2014). The detection of fine-scale structure and signatures of selection are important for delineating conservation and management units (CUs and MUs; Funk et al., 2012) for both wild and captive species, as populations may otherwise appear homogenous in their organisation and distribution.

To reveal these underlying differences, marker sets with sufficient sensitivity and resolving power are required to accurately inform conservation and fishery management efforts, but are also necessary in aquaculture applications. Given the availability of such marker sets, genome mapping and association studies are now real possibilities for a number of established species in aquaculture, including quantitative trait locus (QTL) identification and marker assisted selection (MAS; Allendorf et al., 2010; Angeloni et al., 2012; McAndrew and Napier, 2011; Yue, 2013). Genome-wide SNPs are powerful and highly versatile markers capable of addressing this problem, and are gaining broader use in population genomic and phylogenetic investigations, as well as genome structure evaluation and genome wide association studies (Davey et al., 2011; Helyar et al., 2011; Kai et al., 2014; Pool et al., 2010; Rasic et al., 2014).

The advent of Restriction-site Associated DNA (RADseq) genotyping methods combined with next generation sequencing (NGS) technologies, has enabled the delivery of genome-wide SNPs for both model and non-model organisms, at considerably lower costs and less hands-on time investment than traditional techniques (Davey and Blaxter, 2011; Davey et al., 2011; Nielsen et al., 2011). A recent Restriction-site Associated DNA-based genotyping method is double digest RADseq (ddRADseq), which permits precise selection of the proportion of the genome required for sequencing and maximises the number of sequence reads incorporated in the analysis, without the need for a reference genome (Hohenlohe et al., 2011; Kai et al., 2014; Peterson et al., 2012). This is particularly important for non-model marine organisms which include many species used for aquaculture (McAndrew and Napier, 2011), where genomic resources are often lacking, or absent altogether.

Despite its popularity, ddRADseq along with other RADseq methodologies have seen limited application to marine invertebrate taxa due to challenges such as a lack of a priori knowledge of genome size, variability and frequency of restriction enzyme cut sites (Toonen et al., 2013). This is especially true for marine bivalves, as their genomes are characterised by a high density of repetitive non-coding regions and increased levels of polymorphism (Suárez-Ulloa et al., 2013). To date, draft genomes for only two bivalve species are available; the Pacific oyster *Crassostrea gigas* and the Akoya pearl oyster *Pinctada fucata* (Takeuchi et al., 2012; Zhang et al., 2012).

Given the increased occurrence of genome-wide polymorphism rates characteristic of marine bivalves (e.g. up to 1 SNP per 40 base pairs reported for *C. gigas*; Harrang et al., 2013; Hedgecock et al., 2005; Zhang et al., 2012), variability at restriction enzyme cut sites may erode the genotyping power of RADseq approaches by altering fragment distributions and causing locus allele drop outs (Davey et al., 2011; Huang and Knowles, 2014). For genomic investigations involving bivalves and other highly polymorphic taxa, determination of the extent to which genotyping efficiency is affected and thorough analysis to maintain data integrity are required.

The black-lip pearl oyster *Pinctada margaritifera* is the basis of a valuable aquaculture industry in the Fiji Islands that is almost exclusively dependent on oysters collected from the wild. However, there are currently no comprehensive fishery management guidelines nor genomic resources available for this species, as a country-wide stock assessment has never been undertaken. In this study, we examine the suitability of ddRADseq genotyping for detecting genome-wide SNPs in *P. margaritifera*, and assess the utility of the marker set for a range of population genomic and association analyses. In addition, we provide an optimised ddRADseq methodology for this species, including analysis recommendations for high-quality SNP data recovery in this taxon.

To demonstrate the power and versatility of the genome-wide SNPs discovered, population structure and variability were investigated in four closely-related Fijian populations. Furthermore, genome-wide association testing and F_{st} outlier detection were used to distinguish between individual oysters belonging to the two different tissue colour morphotypes characteristic of this species within separate populations. Finally, we assess the suitability of the genome-wide SNPs discovered for their application to further population genomic investigations involving this highly important bivalve mollusc, and illustrate the applicability of ddRADseq for similar analyses in other highly dispersive marine taxa.

2. Methods and materials

2.1. Specimen collection, tissue sampling and DNA extraction

Adult *P. margaritifera* (n = 156) sized between 7 and 15 cm in dorso-ventral measurement (DVM) were collected from four Fijian sites representing one hatchery-produced population (n = 25), and three populations of wild oysters (n = 50, 32 and 49 at Namarai, Raviravi and Savusavu respectively). All wild oysters were sampled either directly from natural coral reef habitats or spat collectors (Southgate, 2008; see Fig. 1 for a site map). Proximal mantle and adductor muscle tissues (3 and 6 cm respectively) were removed and transferred to tubes containing 20% salt saturated dimethyl sulfoxide (DMSO-salt) preservative (Dawson et al., 1998). Genomic DNA was extracted using a modified cetyl trimethyl ammonium bromide (CTAB, Amresco, cat. #0833-500G) chloroform/isoamyl alcohol protocol with a warm (30 °C) isopropanol precipitation (Adamkewicz and Harasewych, 1996). To clean up all DNA extractions, a Sephadex G50 (GE, 2007) spin column protocol was used prior to quantification with a Nanodrop 1000 Spectrophotometer (Thermo Scientific).

2.2. ddRADseq library preparation, in silico size selection and sequencing

Double digest restriction site-associated (ddRAD) libraries were prepared following Peterson et al. (2012), with the following modifications. Briefly, after being restriction digested for 18 h using 10 U each of MspI and EcoRI-HF endonucleases (New England Biolabs, cat. #R0106S and #R3101S) and ~1.5 μg gDNA at 37 °C, all samples were cleaned up using Sera-Mag SpeedBeads Carboxylate-modified Microparticles (Thermofisher Scientific cat. # 4515-2105-050250) and quantified with an AccuBlue High Sensitivity dsDNA kit (Biotium, cat. #31006). Custom barcoded adapters (P1 and P2; refer to Peterson et al., 2012 for sequences) were then ligated to 500 ng of digested DNA fragments using a T4 DNA ligase (New England Biolabs, cat. #M0202S).

To determine the optimum fragment size selection window, in silico restriction digest simulations were performed on the Pacific oyster *C. gigas* (NCBI GenBank ID 10758) and Akoya pearl oyster *P. fucata* (http://marinegenomics.oist.jp/pinctada_fucata) genomes respectively. Simulations were carried out in the *R* package *SimRAD* (Lepais and Weir, 2014) at varying size selection windows, and the results extrapolated to the estimated size of the *P. margaritifera* genome (0.824 Gb). Results indicated 16,141 and 16,938 fragments to be expected at a target window of 490 ± 40 bp based on the *C. gigas* and *P. fucata* genomes respectively. Selection of this fragment size window was based on sufficient DNA recovery during the size selection step. These results agree well with simulations by Peterson et al. (2012) using the same enzyme combination (~17,000 fragments at >7× coverage).

Samples were pooled into sets of 48 according to their P1 adapter barcodes, quantified using a Nanodrop 1000 Spectrophotometer (Thermo Scientific) and size selected at 490 ± 40 bp using a Pippin Prep automated size selector (PIP0001) and 2% agarose gel Pippin