Aquaculture 437 (2015) 127–139

Contents lists available at ScienceDirect

Aquaculture

journal homepage: www.elsevier.com/locate/aqua-online

Inheritance of high-resolution melting profiles in assays targeting single nucleotide polymorphisms in protein-coding sequences of the Pacific oyster *Crassostrea gigas*: Implications for parentage assignment of experimental and commercial broodstocks

Xiujun Sun *, Grace Shin, Dennis Hedgecock

Department of Biological Sciences, University of Southern California, Los Angeles, CA 90089-0371, USA

ARTICLE INFO

Article history:
Received 11 August 2014
Received in revised form 4 November 2014
Accepted 10 November 2014
Available online 15 November 2014

Keywords:
Single nucleotide polymorphisms
High resolution melting
Non-canonical profiles
Mendelian inheritance
Segregation distortion
Parentage assignment

ABSTRACT

Single-nucleotide polymorphisms (SNPs) have become markers of choice in genetic studies and may be assayed by a variety of methods. For purposes of confirming pedigrees of experimental and commercial broodstocks of the Pacific oyster *Crassostrea gigas*, we developed PCR primers for exonic amplicons, ranging in size from 61 to 165 base pairs (average 95 bp) and high-resolution melting (HRM) assays for a set of 53 coding SNPs, which were previously identified from EST sequences and mapped in families. We define as canonical the three HRM profiles corresponding to the two homozygotes and the heterozygote expected at the target SNP (e.g. AA, GG, AG or AA, CC, AC), with variation at no other site in the PCR amplicon. We summarize data on Mendelian inheritance of HRM profiles in four, large, full-sib progenies from wild-caught parents and in 16 families with smaller numbers of progeny, a dataset of 24,065 HRM phenotypes. Not surprisingly, these family analyses confirm Mendelian inheritance of canonical HRM profiles, while showing evidence for non-amplifying null alleles and distortions of transmission ratios that can be ascribed to early viability selection. Unexpectedly, these family analyses reveal evidence for heritable, non-canonical HRM profiles, which we further describe. Non-canonical HRM phenotypes are found at 40 of the 53 markers but comprise only 5.37% of all phenotypes in our dataset (whereas technical artifacts account for only 0.4% of phenotypes and measured genotyping error from replicates is just 0.2%). Cloning and sequencing of 142, non-canonical HRM haplotypes, at 26 loci, reveal polymorphism at up to four sites besides the target nucleotide. Thus, HRM assays reveal extensive haplotype diversity within exons; interpreted correctly, this additional phenotypic variation increases the allelic diversity and power of parentage assignment for these Type-I markers.

1. Introduction

The Pacific oyster (*Crassostrea gigas*), which has been introduced from Asia to all continents but Antarctica (Mann, 1979), is one of the most commercially important species in aquaculture (FAO, 2013) and the subject of large-scale selection (Dégremont et al., 2010; Langdon et al., 2003) and crossbreeding programs (Hedgecock and Davis, 2007). Because of the evident potential for cross-contamination of research or commercial stocks of bivalve molluscs (Curole and Hedgecock, 2007; Foltz, 1986; Hedgecock and Davis, 2007; Hedgecock et al., 1995; Mallet et al., 1985; Zouros et al., 1994), molecular genetic markers must be used in breeding or experimental research programs in order to genotype broodstocks and their putative and potential parents and, thus, to confirm their pedigrees. In the past, parentage assignment of Pacific oysters was accomplished with allozymes (e.g. Hedgecock et al., 1995) and, more recently, with microsatellite DNA markers (e.g. Curole and Hedgecock, 2007; Hedgecock and Davis, 2007). Both types of markers have drawbacks for parentage assignment. The low number and low to moderate allelic diversity of allozyme markers in bivalves precludes rigorous parentage analysis. The number and allelic diversity of microsatellite DNA markers is high, but so too are genotyping errors, as determined by replicate typing of samples (2.5%; L. Plough, pers. comm.), and frequencies of null alleles (Li et al., 2003). With the advent of unparalleled genomic and biological resources for the highly polymorphic Pacific oyster (Hedgecock et al., 2005; Zhang et al., 2012), numerous single nucleotide polymorphisms (SNPs) are now available for this purpose with more than 100 SNPs having been tested in parentage assignment (Jin et al., 2014; Lapègue et al., 2014; Zhong et al., 2013).

Over the last decade, SNPs have become genetic markers of choice, because they are abundant in genomes and are amenable to high-throughput, automated genotyping assays (Garvin et al., 2010; Helyar

* Corresponding author at: Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071, China. Tel.: +86 532 85811982.
E-mail address: xjsun@ysfri.ac.cn (X. Sun).

http://dx.doi.org/10.1016/j.aquaculture.2014.11.009
0044-8486/

et al., 2011; Morin et al., 2004). Genotyping of SNP variation has become routine in studies as diverse as QTL mapping (Edwards, 2013), population genetics (Helyar et al., 2011), and genome-wide association (Hunter et al., 2007). Recent improvements in the speed, cost and accuracy of next generation sequencing (NGS) and advances in bioinformatic tools are revolutionizing opportunities for high-throughput SNP discovery and genotyping at the genome-wide level (Baird et al., 2008; Hohenlohe et al., 2010; Romay et al., 2013; Wang et al., 2012). Nevertheless, low-throughput methods for genotyping a limited set of known SNPs often suffice for certain applications, such as parentage analysis, which we routinely perform to confirm the pedigrees of experimental broodstocks of the Pacific oyster *C. gigas* (Curole and Hedgecock, 2007; Hedgecock and Davis, 2007).

High-resolution melting (HRM) analysis is a relatively efficient, accurate, and inexpensive method for SNP genotyping (Montgomery et al., 2007; Smith et al., 2010; Wu et al., 2008). HRM simply measures decay in fluorescence of a saturated, intercalating, DNA dye, as temperature is incrementally increased to melt double-stranded DNA (usually, a PCR product containing a target SNP). Population variation in the target DNA-sequence produces variation among individuals in the melting curve or profile of a given PCR amplicon (Distefano et al., 2012; Lévesque et al., 2011; Liew et al., 2004). In species with low levels of polymorphism, such as humans, mammals and fruit flies, HRM polymorphisms will usually correspond to variation at a single, target SNP; however, in highly polymorphic species, such as the Pacific oyster (Zhang et al., 2012), HRM polymorphism may reflect multiple SNPs or even small insertions or deletions (INDELs) in the target sequence. To some extent, non-target variation can be avoided in the initial stages of marker development (Zhong et al., 2013), but this may not always be effective, as we shall show. Still, HRM provides a single-tube genotyping assay, with no need for separating PCR products in a gel or other matrix to detect sequence variation (Montgomery et al., 2007).

In this study, we develop HRM assays for a set of 53 coding SNPs, which were previously identified from EST sequences and placed on the Pacific oyster linkage map (Hedgecock et al., 2011; D. Hedgecock, personal communication). We first define the canonical HRM profiles corresponding to the target SNPs assayed and summarize data on their expected Mendelian inheritance in random-bred and inbred full-sib progenies. Canonical HRM profiles are those having a one-to-one correspondence with the genotyping calls for the two homozygotes and the heterozygote expected at the target SNP (e.g. A/A, G/G, A/G or A/A, C/C, A/C), with variation at no other site in the PCR amplicon. However, we also find numerous non-canonical (NC) HRM phenotypes, which are caused by polymorphism at sites other than the target nucleotide. Non-canonical HRM profiles are distinct from the three canonical melting profiles at any one locus and usually comprise two or more peaks, suggesting that they represent heterozygotes for canonical and non-canonical haplotypes.

These non-canonical HRM phenotypes comprise a small but non-trivial proportion of phenotypes and show Mendelian inheritance. When correctly interpreted, non-canonical phenotypes increase allelic diversity at HRM markers and, thus, increase power in parentage assignment.

2. Materials and methods

2.1. HRM analyses of target SNPs

Target SNPs were selected from among nearly 1550 SNPs first identified in EST sequences (Hedgecock et al., 2011; D. Hedgecock, pers. comm.). These SNPs were localized within specific exons of the oyster genome (Zhang et al., 2012), typed in parents and offspring of several families, by means of Sequenome or Illumina Golden Gate methods, and placed on previously published linkage maps of microsatellite DNA markers. All but two target SNPs and their corresponding HRM markers are given a binomial consisting of the genome scaffold number followed by the nucleotide position of the SNP, which in turn is flanked by the substitution identified in the EST sequences (e.g. 1583-T699272C). One of the original cDNA clones, 12-431, coding for one of 88 HSP70 proteins, was not wholly present in any scaffold of the final genome assembly. Another cDNA clone is tentatively assigned to scaffold 728, although BLAST alignment shows a match in scaffold C21892 as well.

Candidate SNPs for the development of HRM assays were selected from among these previously characterized SNPs so as to maximize discrimination of alternative homozygotes as well as heterozygotes (i.e., class 1, G/A SNPs, or class 2, C/A SNPs; Liew et al., 2004), to avoid proximity to other SNPs identified in the EST sequences, and to have an even distribution over the linkage map. Altogether, 192 target SNPs were identified under these criteria. We then designed primers for PCR amplification from genomic DNA, using Primer3 (http://gmdd.shgmo.org/primer3/?seqid=47). Primer pairs were chosen to satisfy the usual criteria for PCR and to amplify a product greater than 60 base pairs (bp) but less than 200 bp. All PCR products were initially genotyped by HRM of the parents and six progeny of two, random-bred, full-sib families (G_0 families 12 and 45), which had previously been typed by Illumina Golden Gate assay, to confirm robust PCR amplification and polymorphism for the target SNPs. At this stage, primer pairs were rejected, if they produced HRM melting profiles indicative of multiple SNPs in the same amplicon.

Genomic DNA was extracted from the muscle or mantle samples of oysters using the QIAGEN DNeasy kit (Qiagen Inc., Valencia, CA) according to the manufacturer's standard protocol. DNA concentration of each sample was determined by Quant-iT PicoGreen® dsDNA Reagent and Kits (Invitrogen) and normalized to 5 ng/μl. The HRM assay was performed in 96-well or 384-well plates in a final volume of 10 μl containing 5 ng of genomic DNA, 0.175 μmol of each primer, 5 μl HRM Master Mix (Roche Diagnostics), and 3 or 3.5 mmol $MgCl_2$. The DNA amplifications and subsequent melting curve analyses were carried out using a LightCycler® 480 System (Roche Diagnostics), under the following conditions: 10 m at 95 °C followed by 40–45 cycles of amplification, each including 10 s denaturation at 95 °C, 10 s annealing at 57–62 °C (depending on the primer pairs) and 5 s extension at 72 °C.

After PCR amplification, the products were immediately heated to 95 °C for 1 m and then cooled to 40 °C, followed by a continuous temperature increase of 0.05 °C/s to a target temperature of 95 °C. Fluorescence was monitored during the entire heating process, and genotypes were determined by examining the normalized melting curve and the melting-peaks display of the GeneScanning Analysis module of the LightCycler® 480 Software, which plots the negative first derivative of sample fluorescence against temperature. We refer generically to the melting curve and melting-peak plot as the HRM profile.

2.2. Families and parentage assignments

A set of 50 random-bred families (G_0) were established by pair crosses of wild-caught Pacific oysters (*C. gigas*) from Dabob Bay, Washington, in 2006, at the Taylor Shellfish Farms hatchery. Four of these G_0 families, 12, 20, 24, and 45 were sampled in 2009 and used to confirm Mendelian inheritance in this study ($n = 126$, 81, 94, and 123 respectively). In 2008, first-generation inbred families (G_1) were created by a single brother–sister mating within many of these G_0 families. In 2012 and 2013, when the G_1 oysters had matured and families were being screened for use in experimental diallel crosses, we assigned potential G_1 broodstock to their G_0 parents to confirm pedigree. For this study, we used parents and progeny from 15, G_1 families (7, 8, 15, 16, 19, 20, 21, 24, 27, 30, 32, 33, 34, 39, and 45) and one additional unrelated family 05 x 8.039; maximum samples sizes per family ranged from six to 20. Potential inbred broodstocks were genotyped and compared to the genotypes of their putative parents, as well as 50 pairs of wild-caught G_0-family parents, using CERVUS 3.0 (Kalinowski et al., 2007). All broodstock populations were combined to calculate allele-