Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation

PIERRE DE WIT and STEPHEN R. PALUMBI

Department of Biology, Hopkins Marine Station, Stanford University, 120 Ocean view Blvd., Pacific Grove, CA 93950, USA

Abstract

Global climate change is projected to accelerate during the next century, altering oceanic patterns in temperature, pH and oxygen concentrations. Documenting patterns of genetic adaptation to these variables in locations that currently experience geographic variation in them is an important tool in understanding the potential for natural selection to allow populations to adapt as climate change proceeds. We sequenced the mantle transcriptome of 39 red abalone (*Haliotis rufescens*) individuals from three regions (Monterey Bay, Sonoma, north of Cape Mendocino) distinct in temperature, aragonite saturation, exposure to hypoxia and disease pressure along the California coast. Among 1.17×10^6 Single Nucleotide Polymorphisms (SNPs) identified in this study (1.37% of the transcriptome), 21 579 could be genotyped for all individuals. A principal components analysis concluded that the vast majority of SNPs show no population structure from Monterey, California to the Oregon border, in corroboration with several previous studies. In contrast, an F_{ST} outlier analysis indicated 691 SNPs as exhibiting significantly higher than expected differentiation (experiment-wide $P < 0.05$). From these, it was possible to identify 163 genes through BLAST annotation, 34 of which contained more than one outlier SNP. A large number of these genes are involved in biomineralization, energy metabolism, heat-, disease- or hypoxia-tolerance. These genes are candidate loci for spatial adaptation to geographic variation that is likely to increase in the future.

Introduction

As Earth's climate continues to change, many species are likely to experience increased environmental stress (Reid *et al.* 2009). In particular, for marine species, sea level rise, higher temperatures and decreasing pH are affecting coastal and open ocean ecosystems throughout the world (Fabry *et al.* 2008; Pörtner & Peck 2010; Booth *et al.* 2011; Johnson *et al.* 2011; Philippart *et al.* 2011; Wassmann *et al.* 2011; Wernberg *et al.* 2011). For example, global sea surface temperature has risen by approximately 0.7 °C in the last century and is projected to increase 2–3 °C in the next 100 years, with the strongest impacts on more polar ecosystems (Solomon *et al.* 2007; Lough & Hobday 2011; Wassmann *et al.* 2011). The average pH of the ocean surface is expected to decrease from approximately 8.1 to nearly 7.7 by the year 2100 (Feely *et al.* 2009) because higher concentrations of atmospheric CO_2 continues to acidify surface waters (Sabine *et al.* 2004; Hauri *et al.* 2009; Hofmann *et al.* 2010). Changes in ocean circulation patterns may also be changing the distributions and severity of deep-sea, low oxygen zones (Stramma *et al.* 2010), which in combination with regional upwelling can contribute to hypoxic conditions in surface waters (Bograd *et al.* 2008).

The responses of marine organisms to these changing conditions are varied and complex. Transient high

Correspondence: Pierre De Wit, Fax: +(831) 375 0793;
E-mail: pdewit@stanford.edu

temperatures can cause coral bleaching (Anthony *et al.* 2011), death of intertidal organisms (Wethey *et al.* 2011) or biogeographic shifts in fish and other marine species (Jones *et al.* 2010; Pörtner & Peck 2010; Booth *et al.* 2011; Johnson *et al.* 2011; Somero 2012). Increasing temperature is also often associated with increased exposure to microbes and incidence of disease, such as the withering-foot syndrome in California abalone (Moore *et al.* 2002, 2011; Raimondi *et al.* 2002; Rogers-Bennett *et al.* 2010). Although not all organisms are negatively affected by decreasing pH (Kroeker *et al.* 2010), species of calcifying organisms have often been shown to experience strong fitness effects, because increased acidity (lower pH) makes it more energy-consuming to precipitate $CaCO_3$ into shell or skeleton matrices (Feely *et al.* 2004; Hofmann *et al.* 2010; Erez *et al.* 2011; Stumpp *et al.* 2011; Yu *et al.* 2011). Shelled molluscs seem particularly susceptible: a large body of evidence suggests a deleterious effect of decreasing pH (Miller *et al.* 2009a; Talmage & Gobler 2009; Beniash *et al.* 2010; Zippay & Hofmann 2010; Crim *et al.* 2011), possibly due to the fact that almost all mollusc shells to some extent consist of aragonite (Furuhashi *et al.* 2010), a relatively unstable form of calcium carbonate.

Although there has been much focus on possible future changes across the globe for pH, temperature, hypoxia and other environmental factors (Solomon *et al.* 2007; Bograd *et al.* 2008; Reid *et al.* 2009), it is difficult to predict how organisms will react. However, there are places, such as upwelling systems, that currently experience large environmental variations in all of the above-mentioned variables on small geographic or temporal scales (Aristegui *et al.* 2009; Belkin *et al.* 2009; Fransson *et al.* 2009; Hauri *et al.* 2009; Zhang *et al.* 2010). These systems might provide indicators for possible future genetic impacts of climate change.

The geographic variation in basic ocean environment may exert strong selection on populations to adapt to an environmental average or may allow local adaptation of populations to particular local conditions (Hereford 2009; North *et al.* 2011; Sanford & Kelly 2011). The effect of high dispersal was thought to be enough to keep a population genetically homogeneous along a range of environmental conditions (Díaz-Viloria *et al.* 2009; Hellberg 2009); however, population genetic theory shows that local adaptation depends on a balance of selection and gene flow (Sotka & Palumbi 2006). Recent studies of marine species at the whole-genome level increasingly demonstrate cases of particularly strong differentiation of some loci in the face of gene flow (Conover *et al.* 2006; Nielsen *et al.* 2009; Pespeni & Palumbi 2010). For example, Pespeni *et al.* (in prep) recently observed that there was enough standing genetic variation in natural populations of the purple

sea urchin such that experimental exposure to high acidity conditions led to marked gene frequency evolution. Such whole-genome approaches are well suited to understanding the relationship of coastal environmental mosaics to potential adaptation for climate change effects.

Along the upwelling-dominated California coast, strong variation in temperature (Checkley & Barth 2009), oxygen concentration (Bograd *et al.* 2008; Connolly *et al.* 2010) and pH (Feely *et al.* 2008; Hauri *et al.* 2009) suggests that there may be adaptive genes in species that live across this range associated with tolerance to these stressors (Somero 2010; Pespeni *et al.* 2011). Abalone (*Haliotis spp.*) are semi-sessile marine gastropods that have evolved and diversified along the Pacific coast of North America (Estes *et al.* 2005). All but one species is currently listed as under concern in California (CADFG 2011), and they have been shown to be negatively impacted both by increasing temperatures causing increased disease occurrence (Rogers-Bennett *et al.* 2010) and by decreasing pH causing abnormal larval development in an experimental setting (Crim *et al.* 2011). The red abalone (*Haliotis rufescens*), the only species for which there is still a limited recreational fishery, shows remarkably little population structure over large geographic distances, despite a low potential for gene flow (Burton & Tegner 2000; Gruenthal *et al.* 2007). In this study, we hypothesize that within an overall homogeneous genetic structure in *H. rufescens*, genes that are acted upon by local natural selection may diverge between areas of different environmental stressors along the California coast. As a result, we predict that genes related to shell biomineralization, resistance to hypoxia, temperature tolerance and resistance to pathogens will show the strongest patterns of local adaptation. To test this, we sequenced the mantle transcriptome of 39 *H. rufescens* individuals from three different regions in the northern part of the species' range between Monterey and the Oregon border (ca. 600 km) to compare Single Nucleotide Polymorphisms (SNPs) among locations. We scored all individuals at 21 579 loci and analysed them for geographic patterns and statistical outliers. We used these analyses to simultaneously estimate overall patterns of gene flow and to trace possible loci and gene functions under local environmental selection.

Material and methods

Non-lethal mantle tissue samples of *Haliotis rufescens* were taken *in situ* with biopsy forceps from 39 individuals in the intertidal zone or underwater using SCUBA. Sampling sites were as follows: Hopkins Marine Station (H) (*n* = 13), 36° 37.244′N, 121° 54.118′W; Fort Ross (FR)

(n = 6), 38° 30.142′N, 123° 13.856′W; Salt Point (SP) (n = 1), 38° 33.729′N,123° 19.465′W; Gerstle Cove (GC) (n = 1), 38° 34.088′N,123° 20.195′W; Anchor Bay (AB) (n = 6), 38° 48.103′N, 123° 34.995′W; Trinidad Harbour (T) (n = 6), 41° 3.297′N, 124° 8.646′W; Patrick's Point (PP) (n = 4), 41° 7.968′N, 124° 9.799′W; Crescent City (CC) (n = 2): 41° 45.179′N, 124° 13.135′W (Fig. 1). These sampling locations correspond to three environmentally distinct regions along the California coast. Monterey Bay is characterized by mild upwelling during the summer months only, a fairly constant surface water temperature (12–14 °C year-round) (National Buoy Data Center, NOAA) and reports of occasional coastal hypoxia (Booth *et al.* 2010) as well as a long-term presence of the bacterium causing the lethal withering-foot syndrome during warm periods (Moore *et al.* 2002). Sonoma, with stronger upwelling during the summer (although primarily offshore), has temperatures between 9 and 12 °C (National Buoy Data Center, NOAA), a reported decrease during summer in aragonite saturation state (although still supersaturated in surface water) (Hauri *et al.* 2009) and only recent reports of the presence of the agent of the withering-foot syndrome (L. Rogers-Bennett, personal communication). Finally, the region located north of Cape Mendocino, up to the Oregon border, has year-round upwelling, which is very strong during the late summer months, a more
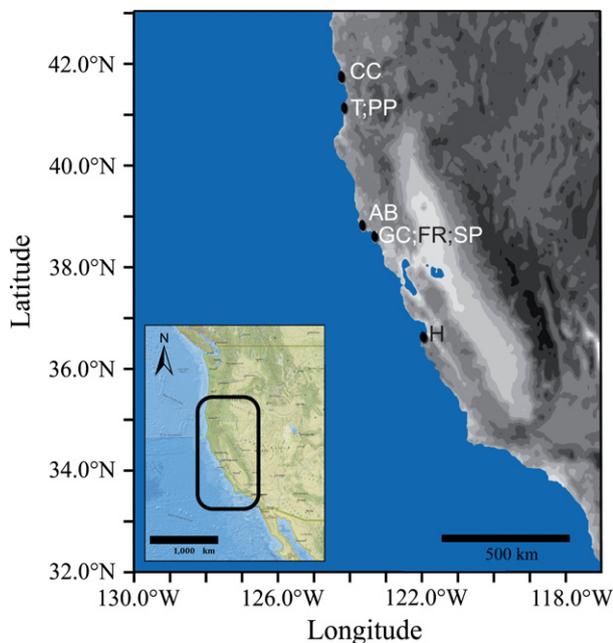

Fig. 1 Map of the species range of red abalone along the US West Coast, showing collecting locations. Location codes: CC, Crescent City; T, Trinidad Harbour; PP, Patrick's Point; AB, Anchor Bay; GC, Gerstle Cove; FR, Fort Ross; SP, Salt Point; H, Hopkins Marine Station.

variable temperature (between 8 and 15 °C) (National Buoy Data Center, NOAA), reported aragonite undersaturation in surface waters (Feely *et al.* 2008) but no presence of the agent of the withering-foot syndrome (Moore *et al.* 2002; L. Rogers-Bennett, personal communication).

Tissue samples were allowed to soak in RNA later (Qiagen, Valencia, CA, USA) at room temperature or at 4 °C for about 1 day, then stored at −80 °C. Total RNA was extracted from tissue using TriZol reagent (Invitrogen, Carlsbad, CA, USA). Tissue pieces were first cut into smaller pieces using a razorblade, after which they were flash frozen with liquid nitrogen and powderized with a micropestle before adding TriZol. Each sample was extracted in duplicate. A standard extraction protocol was used (Rio *et al.* 2010), at the end of which the RNA pellets were eluted in 25 μL of RNAse-free water; duplicates were combined for a total of 50 μL. Total RNA was quantified using QuBit 2.0 flourometer (Invitrogen) RNA assays.

For cDNA library construction, Illumina's TruSeq RNA sample prep kit (Illumina, San Diego, CA, USA) was used, following the standard protocol, using the full 50 μL of total RNA extract as input. No optional in-line controls were used. For multiplexing, we used six different indexing adapters. Complementary DNA was quantified using a QuBit 2.0 flourometer (Invitrogen), an Agilent Bioanalyser (Agilent, Santa Clara, CA, USA) as well as quantitative PCR, after which separate libraries were pooled into seven tubes [six with six samples each, and one with three samples (AB51, GC28 and H4)]. Cluster generation and parallel sequencing of single-end 50 base-reads was then performed in seven lanes in an Illumina HiSeq 2000 (Illumina) at the University of Utah's Bioinformatics Core Facility.

A detailed pipeline of all data processing steps can be found at http://sfg.stanford.edu (De Wit *et al.* 2012). Short reads were trimmed using a phred-scale quality score cut-off of 20; residual adapter sequences were clipped from the reads after which reads shorter than 20 bases were discarded. Quality score and nucleotide distributions as well as the proportions of duplicate reads were examined using the fastx toolkit (http://hannonlab.cshl.edu/fastx_toolkit/). Trimmed and clipped short reads from all individuals were imported to CLC genomics workbench (CLC bio, Cambridge, MA, USA) and combined to generate a *de novo* assembly with mismatch cost 1 and insertion and deletion cost 2 (limit 5), while voting for conflict resolution and ignoring all nonspecific matches and specifying a minimum contig length of 200 bases.

To annotate the *de novo* assembly, it was compared against the UniProt knowledgebase and GenBank's nr database using the BLASTx algorithm, and against Gen-

Bank's nt database and the *Lottia gigantea* genome project's (http://genome.jgi.doe.gov/Lotgi1/Lotgi1.home.html) 'all gene models' data set using the tBLASTx algorithm. In all cases, NCBI's default parameters were used. Hits with *e*-values $<10^{-5}$ were considered significant. BLASTing was conducted on Stanford University's Bio-X^2 cluster (http://biox2.stanford.edu/). Gene Ontology (GO) annotation was acquired from AGBase (http://www.agbase.msstate.edu/) using UniProt identifiers, when possible. To remove non-abalone contigs, ME-GAN 4 (http://ab.inf.uni-tuebingen.de/software/megan/) (Huson *et al.* 2011) was used with the results of the tBLASTx to nt, to create a list of contigs that were assigned to non-metazoa (minimum score 55, minimum support 2). These contigs were excluded from further analysis.

Samples were then individually mapped against the assembly using the Burrows–Wheeler Aligner (http://bio-bwa.sourceforge.net/) (Li & Durbin 2009), using default parameters (except $n = 0.005$ and $k = 5$). For SNP detection, the Genome Analysis Toolkit (McKenna *et al.* 2010; DePristo *et al.* 2011) was used, following the best practise protocol version 2 of the Broad Institute (http://www.broadinstitute.org/gsa/wiki/index.php/Best_Practice_Variant_Detection_with_the_GATK_v2); only deviating in the omission of the Base Quality Score Recalibration step, as this step requires known variant sites as input. An initial variant call set was obtained using a phred-scale quality threshold of 30 as well as applying hard filtering with all settings as recommended by the Broad Institute except cluster-window size = 10. This call set was used to train the Variant Quality Score Recalibration model and filter a call set with a variant quality score threshold of 4 (all other settings as recommended by the Broad Institute). The HARD_TO_VALIDATE filter was used to train the Gaussian mixture model but was ignored when applying the final filter, as recommended by the Broad Institute. The list of polymorphic sites was then filtered to extract only sites for which there were confident genotypes (phred-scale threshold 20) assigned to all 39 individuals.

For data analysis purposes, the individuals were grouped into three populations: Monterey ($n = 13$), Sonoma (including FR, GC, SP and AB) ($n = 14$) and north of Cape Mendocino (including T, PP and CC) ($n = 12$). A principal components analysis (PCA) was performed using the EIGENSOFT software package (http://genepath.med.harvard.edu/~reich/Software.htm) (Patterson *et al.* 2006). The results of the PCA led us to remove two deviating individuals from Patrick's Point (PP1 and PP3) and then repeat the SNP detection pipeline as described above with the remaining 37 individuals, after which another PCA was performed.

Finally, the data set was further trimmed by removing variant sites for which all individuals had been genotyped as heterozygotes (likely to be due to either contamination or paralogous genes, $n = 70$) and by removing all sites in which the minor allele frequency was <2.5%. Mean global and pairwise F_{ST} were calculated using GenePop (http://kimura.univ-montp2.fr/~rousset/Genepop.htm) (Rousset 2008). To further assess the species status of individuals PP1 and PP3, genomic DNA was back-extracted from Trizol, and the VERL gene, which has been shown to be useful for abalone species identification (Gruenthal & Burton 2005), was amplified using the protocol and primers described in Swanson *et al.* (2001).

To scan the transcriptome for F_{ST} outlier SNPs, the Lositan selection workbench (http://popgen.eu/soft/lositan/) (Antao *et al.* 2008) was used, using a 'neutral' mean F_{ST}, running 500 000 simulations and a 95% confidence interval. From the resulting list of outlier SNPs, genes for which there were more than one outlier SNP as well as at least one significant BLAST hit were extracted. Pairwise F_{ST} between the three populations were calculated for outlier SNPs in these contigs using GenePop and plotted in a 3D scatterplot in R.

Results

Transcriptomic coverage

The number of reads per individual after quality trimming and adapter clipping ranged from 22 885 713 to 25 237 104 in the lane with only three individuals, and from 9 721 592 to 21 646 749 in the remaining lanes. 355 678 562 reads were used in the *de novo* assembly, creating 162 928 contigs of average length 522 bases (N50 = 653) for a total length of 85 063 104 bases (supplementary material 1). BLAST annotations could be assigned to 48 004 out of 162 928 contigs (*e*-value threshold 10^{-5}) (supplementary material 3). BLASTx against NCBI's nr database resulted in 33 049 hits, while tBLASTx against NCBI's nt database returned 42 603 hits. BLASTx against the UniProt knowledgebase returned 28 819 hits. tBLASTx against the *Lottia gigantea* genome resulted in 34 994 hits. The MEGAN algorithm was able to identify 546 contigs in the reference assembly that were non-metazoan (supplementary material 4).

This study provides the first transcriptomic resource for abalone. The large number of contigs (162 928) in the *de novo* assembly is disconcerting (although a known issue when dealing with short reads). After removing the 546 non-metazoan contigs, this is still many more than the number of genes described for *Lottia gigantea* (23 851) (Lotgi v1.0), or ca. 28 000 in the sea urchin *Strongylocentrotus purpuratus* (Sodergren *et al.*

2006; Pespeni *et al.* 2011). However, 114 924 contigs remain un-annotated. It is likely that a large fraction of these are contaminant sequences from bacteria, fungi or unicellular eukaryotes present inside or on the surface of the mantle tissue. In addition, the short average lengths of the contigs indicate that in most cases, several contigs represent different regions of the same genes. Even with a slight overlap, the *de novo* assembly algorithm might not combine adjacent contigs from the same exon in many cases (Martin & Wang 2011). The results from the BLASTx against the nr database (which contains only protein-coding sequences) support this in that 33% of the annotated contigs (10 876 out of 33 049) have the same top hit as one or more additional contig (s) in the assembly. By counting the number of unique hits from the BLASTx against nr, the number of genes in the assembly was approximated to ca. 22 000. Further work is needed to more completely assemble, annotate and remove contaminant sequences from this abalone transcriptome, but the stringent filters that we have used here for annotation give us confidence that we have identified abalone SNPs correctly, at the cost of some data loss.

We also designed primers to amplify outlier-containing fragments of four genes from genomic DNA, using the *Lottia* genome to predict intron-exon boundaries (Pif (F-AGCCTGACTTGTACCTCCC; R-ACAAGAAAGAG-ATGACCAGAACG), Chitin synthase 2 (F-ACACTGA-AACATGAATTGGAACC; R-ATTGTGTAGTTTGTGCG CCA), Sodium glucose cotransporter (F-GAACATCTC CAACGCCAAGG; R-GCCAGCATGAGACCTCGGA) and ATP synthase β chain (F-AGCAGCACCACCAA-ACAGAT; R-CAACAACCCAGACTTCCTTTG)). All primer pairs successfully amplified fragments in the predicted size range in all study individuals; Sanger sequencing of the amplicons confirmed all outlier SNPs in Pif and Chitin synthase 2, and some of the SNPs in the remaining two genes (length-polymorphic introns prohibited validation of all SNPs in Sodium glucose cotransporter and in ATP synthase β chain) (P. De Wit, unpublished data).

Genotyping and population structure

An average of 62.86% (SD 4.41%) of the reads aligned to only one position in the assembly (mapping quality >20). The initial SNP detection resulted in 1 178 549 high-quality variant sites, of which 21 285 had confident genotype calls for all 39 individuals. We removed 70 of these due to failure to conform to Hardy–Weinberg expectations.

A Principal Components Analysis (Fig. 2A) suggests a lack of strong population structure, but finds two deviating individuals (PP1, PP3) on the only significant axis (PC 1, $P = 3 \times 10^{-21}$). Subsequent analysis of SNP weights indicates that the differences observed derive from subtle patterns at many base pair sites and that these two individuals are heterozygous at 498 of the 500 most highly weighted SNPs on PC axis one. The VERL sequences of the two individuals were compared to the data set published on GenBank by Swanson *et al.* (2001) for *Haliotis rufescens*; both were found to be identical to sequence AF250893. Still, because of the possibility that these two individuals represent hybrids (first generation or later backcross, see Discussion), this data set was discarded and the SNP detection pipeline was repeated, resulting in 1 166 275 variant sites ($Q > 20$) (supplementary material 2), of which 21 579 have confident genotype calls ($Q > 20$) for all 37 individuals. A PCA of this data set (Fig. 2B) continues to suggest that there is no strong genetic differentiation between geographic locations. In addition, mean F_{ST} indicates low genetic differentiation among regions for all SNPs (Global $F_{ST} = 0.0003$). The distribution of minor allele frequencies in the data set is given in Fig. 3.

Only 1.85% of the detected variant sites have confident genotypes for all 37 individuals. Due to randomness in sample preparation and the multiplexing of six individuals in each HiSeq lane, it is likely that only the most highly expressed genes were sequenced to sufficient depth to be confidently genotyped for all individuals. The GATK uses a combination of sequencing depth and read qualities at a given site to calculate Bayesian posterior probabilities for the different genotypes and log odds ratios between them, emitting a genotype only if the log odds ratio exceeds a threshold ($P > 0.99$ in our case). Deeper re-sequencing would undoubtedly increase genotype data significantly and provide information on less common transcripts. For this initial study, however, genotype information at 21 579 SNP sites allows us to explore patterns of differentiation across genes likely to be important in organism function.

Outlier and gene association analysis

From the 21 579 SNP sites for which there was genotype information for all 37 individuals, 70 were removed due to all individuals being genotyped as heterozygotes and 8 292 were removed due to a minor allele frequency <2.5%, leaving 13 217 variant sites from 3 342 contigs for the F_{ST} outlier analysis. Lositan indicated 691 SNPs as significantly more divergent between geographic locations than likely from chance alone (supplementary material 5). The SNPs were distributed across 486 contigs, from which 163 genes were identifiable by BLAST search. Of these 163, 34 genes contain more than one outlier SNP and thus are particularly good candidates for being targets of differentiation caused by natural selection (Table 1). This list includes genes basic

(A)

(B)
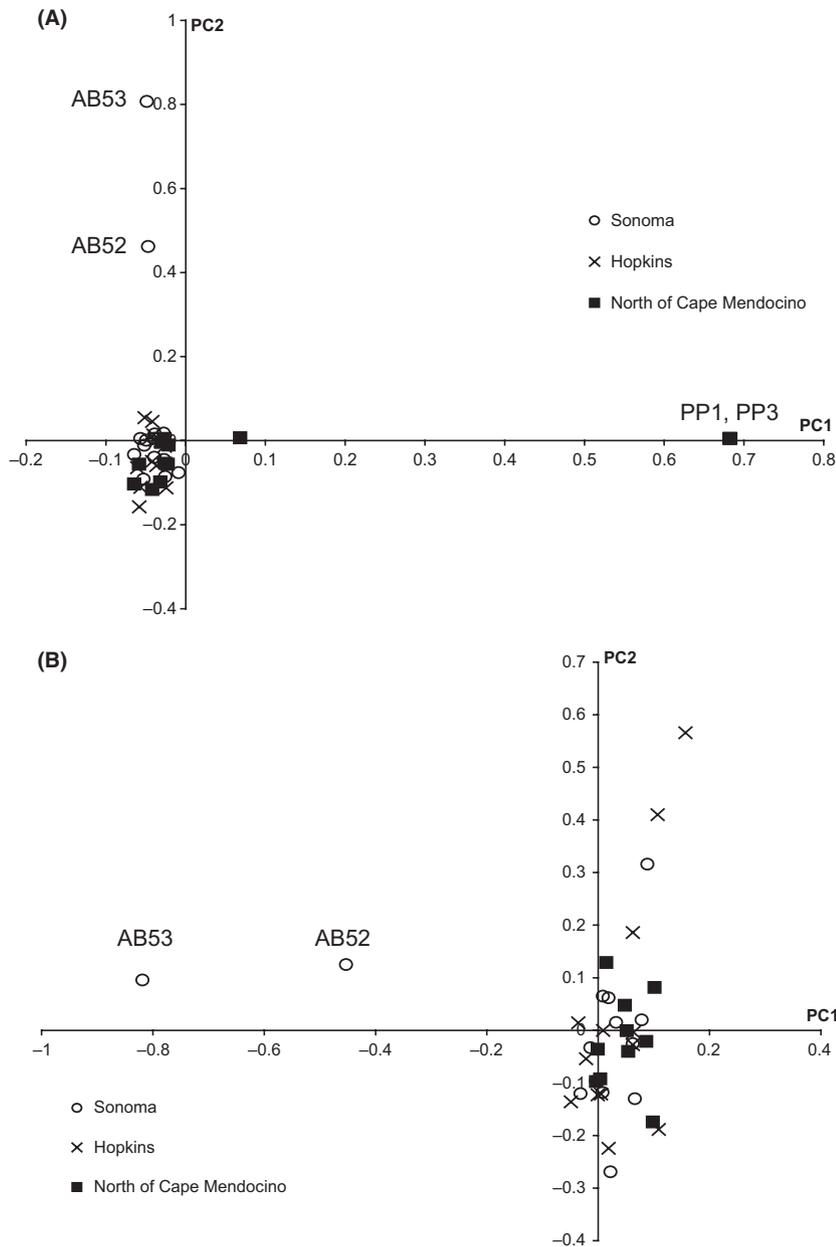
Fig. 2 Principal components plot of PC 1 vs. PC 2. (A) Using 21 222 Single Nucleotide Polymorphisms (SNPs) with high-quality genotypes for all 39 individuals. Axis 1 is significant ($P = 3 \times 10^{-21}$). (B) Using 21 509 SNPs with high-quality genotypes for all 37 individuals after re-running the SNP detection without the two deviant individuals PP1 and PP3. None of the axes are significant. In both (A) and (B), SNPs genotyped as heterozygotes in all individuals have been omitted.

to shell formation and Calcium regulation, genes involved in hypoxia and disease resistance, as well as genes involved in energy metabolism. We compared pairwise F_{ST} among the three populations for each SNP within the 34 genes. For 18 genes where all SNPs in a gene showed a clear consistent pattern, we list which of our three main populations was deviant (Table 1).

Discussion

Our data broadly agree with prior results that show very low genetic differentiation of red abalone along

the California coast (Burton & Tegner 2000; Gruenthal *et al.* 2007), despite their predicted low dispersal ability (Prince *et al.* 1987). Nevertheless, we also detected outlier SNPs with higher than expected differentiation in 486 contigs, suggesting that overlain on a pattern of high gene flow is the signature of local adaptation in some genes.

Possible hybrids

Two individuals from Patrick's Point (PP1 & 3) are clearly deviant in the principal component analysis
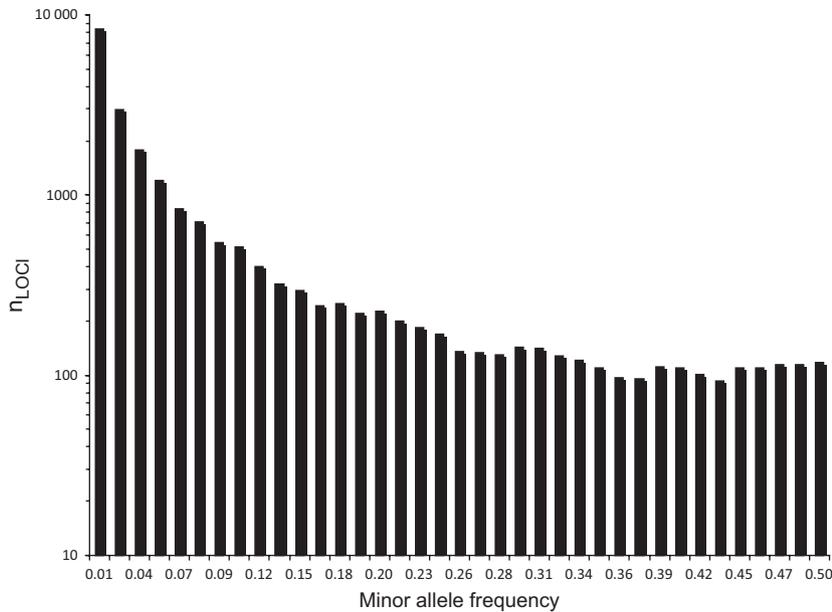
(Fig. 2A). The strongest component of this difference was derived from more than 500 loci that loaded heavily onto PCA axis 1 (Fig. 2A). Individuals PP1 and PP3 were uniquely heterozygous at all but two of these SNPs, suggesting that they might be hybrids between red and another co-occurring abalone species, such as *Haliotis walallensis* or *Haliotis kamtschatkana*. Abalone hybridization is presumed rare because of the species specific interaction of gamete recognition proteins (Swanson & Vacquier 1995; Clark *et al.* 2009), yet natural hybrids have been reported between *Haliotis rufescens* and all other North American species of abalone at an average rate of 0.2% (de la Cruz & Gallardo-Escarate 2010). This rate was observed to be higher in areas of ecological disturbance or at the extreme range limits of one species. The collection point for PP1 and PP3 is north of Cape Mendocino, near the northern limit of *H. rufescens*. VERL sequences of the two individuals were unambiguous and identical to published sequences for *H. rufescens* (Swanson *et al.* 2001). The VERL primers used (Swanson *et al.* 2001) have previously been used to successfully amplify VERL fragments of all North American abalone species, as the polymorphism rate in VERL is very low between species (Gruenthal & Burton 2005). Still, this result cannot rule out that the individuals are hybrids or backcrosses. Confirmation of the hybrid nature of these individuals must await genomic data from the other potential parent species.

Population differentiation

The principal component analyses (Fig. 2) and the mean F_{ST} show that there is no genetic differentiation at the vast majority of SNPs between geographic locations. This is consistent with previous findings (Burton & Tegner 2000; Gruenthal *et al.* 2007) and indicates that red abalone larvae have dispersed among populations along the California coast, despite their short pelagic life stage. Overall F_{ST} is 0.0003, suggesting high rates of historic gene flow at most SNPs.

A comparison of heterozygosity and F_{ST} may help explain the conundrum between short larval duration and low differentiation. We find nearly 1.2 million high quality, variant sites in the abalone transcriptome, resulting in a polymorphism rate of about 1.37%. This is much higher than the variation described in the human genome (0.1%) (Gibbs *et al.* 2003), yet similarly high numbers have been reported from sea urchins (4%) (Britten *et al.* 1978) and oysters (1.67%) (Curole & Hedgecock 2005; Sauvage *et al.* 2007). The genome-wide polymorphism rate is expected to be even higher than the rate in protein-coding expressed DNA reported here. High polymorphism in abalone suggests, as in urchins and oysters, the build-up of genetic diversity in a consistently large population with a high mutation rate. Kang *et al.* (2010) also found high SNP heterozygosity in *Haliotis iris* in New Zealand (3.2%, $n = 2$) and multiple microsatellite studies also demonstrate high allele numbers and high heterozygosity in many species of abalone (Huang *et al.* 2000; Maynard *et al.* 2004; Tang *et al.* 2004; Gutierrez-Gonzalez *et al.* 2007; Gruenthal & Burton 2008; Shu *et al.* 2008; Díaz-Viloria *et al.* 2009; Miller *et al.* 2009b).

High polymorphism rate (1.37%) suggests a large population size. If the average mutation rate per base is on the order of 10^{-8}–10^{-9} per generation (or about 0.2% to 0.02% mutation per million years with a 5-year

Table 1 Genes potentially under differential selective pressure along the California coast in *Haliotis rufescens*

Top BLAST hit		Contig(s) containing outlier SNPs	Number of outlier SNPs	Divergent population
Gene	Organism			
78 kDa glucose-regulated protein	*Crassostrea gigas*	Contig147173, contig148731	2	North
Actin-related protein 2/3	*Saccostrea kegaki*	Contig83864	2	
Alpha-amidating enzyme precursor 2 (LPAM)	*Lymnaea stagnalis*	Contig89165	3	
Ankyrin-1	*Pediculus humanus corporis*	Contig24507	2	North
Arginine kinase	*Haliotis madaka*	Contig90217	3	Monterey
ATP synthase B chain	*Ixodes scapularis*	Contig104796	2	North
Calmodulin 2	*Haliotis diversicolor*	Contig46443	2	Hopkins
Capping protein (actin filament) muscle Z-line, beta (Capzb)	*Mus musculus*	Contig83915	2	
Catalase	*Haliotis discus discus*	Contig88626	2	
Crumbs complex protein (Smp_130280)	*Schistosoma mansoni*	Contig14237	2	
Cytosolic malate dehydrogenase	*Mytilus californianus*	Contig26566	3	
Dermatopontin	*Haliotis discus discus*	Contig115444	3	
EGF-like domain containing protein	*Haliotis diversicolor*	Contig83774	4	
Galectin 4-like protein transcript	*Haliotis discus hannai*	Contig83577	3	North
Glucose-6-phosphate 1-dehydrogenase (g6PD)	*Crassostrea gigas*	Contig25921	3	North
Glutamine synthetase	*Crassostrea gigas*	Contig90598	2	
Glycine transporter 2	*Danio rerio*	Contig102343	2	Sonoma
Hepatopancreas 5-aminolevulinate synthase (ALS)	*Limulus polyphemus*	Contig88857	3	North
Hillarin	*Hirudo medicinalis*	Contig88612	2	
Malate dehydrogenase precursor	*Haliotis discus discus*	Contig83945	2	Sonoma
Molluscan defence molecule precursor (MDM)	*Lymnaea stagnalis*	Contig84742	2	
Nervous system adducin (ADD)	*Aplysia californica*	Contig24566	4	Sonoma
Pif177 Aragonite-binding protein	*Pinctada fucata*	Contig82770	3	North
Plancitoxin	*Haliotis cracherodii*	Contig88923	2	
Poly-U binding splicing factor a (puf60a)	*Danio rerio*	Contig24491	3	Monterey
Proline rich 15-like (PRR15L) ATPase	*Homo sapiens*	Contig147114	2	North
Proteasome alpha type 2	*Haliotis discus discus*	Contig27448	3	North
Protein disulfide isomerase 1	*Ixodes scapularis*	Contig16119	2	
RAS-related protein	*Ixodes scapularis*	Contig13864	3	
Sarco/endoplasmic reticulum calcium ATPase isoform C (SERCA)	*Pinctada fucata*	Contig137187, contig137221	2	Sonoma
Sarcophaga 26,29 kDa proteinase	*Periplaneta americana*	Contig83712	4	Sonoma
Sodium/glucose cotransporter	*Crassostrea gigas*	Contig73178, contig102168	8	Monterey
Spectrin alpha chain	*Ixodes scapularis*	Contig24945	2	
TNFAIP3 interacting protein 1	*Danio rerio*	Contig88834	3	Monterey

Divergent population is noted only when pairwise $F_{ST} > 0.1$ in comparisons with both other populations in some or all of the outlier SNPs within a gene, and no conflicting pattern is seen.

Hypoxia and disease stress

Several genes with functions related to environmental stress or disease are indicated as under differential selective pressure along the California coast. Arginine kinase has been shown to be upregulated in shrimp exposed to low oxygen concentrations and is thought to facilitate oxygen re-uptake after a hypoxic event (Abe *et al.* 2007). In our assembly, 5 contigs match Arginine kinase from *Haliotis madaka*, out of which one (contig90360) contains two linked outlier SNPs. Arginine kinase is different in Monterey compared to Sonoma and northern California (Table 1). The canyons of Monterey Bay bring deep water from the Oxygen Minimum Zone close to shore, and hypoxic water (<ca. 60 μmol/kg) has been regularly observed in shallow water (Bograd *et al.* 2008; Booth *et al.* 2010). Incidence of hypoxia at the other locations is currently unknown.

The 78-kDa glucose-regulated protein (GRP78 (different in northern California)) is a heat-shock protein, induced with heat treatment in oysters (Yokoyama *et al.* 2006) and implicated as having a function salvaging incorrectly folded proteins. Our data set has two contigs of this gene with three outlier SNPs.

Disease-related genes include the sodium/glucose co-transporter protein, which has been shown to be important in resistance to bacteria during warm conditions in the oyster *Crassostrea gigas* (Huvet *et al.* 2004) and is different in Monterey compared to Sonoma and northern California (contigs73178 and 102168, eight outliers). The Rickettsiales-like bacterium that causes withering-foot syndrome in California abalone had until recently only been documented south of the mouth of San Francisco Bay (Moore *et al.* 2002, L. Rogers-Bennett, personal communication).

Metabolism

An ATP synthase, a 5-aminolevulinate synthase (critical for erythrocyte formation in vertebrates (Ferreira & Gong 1995)) and a Proline-rich ATPase are all divergent in the northern population compared to the other two. Within contig 104796 (one of the ATP synthases mentioned above), 2 SNPs are outliers. Five of 18 total SNPs in this contig code for amino acid substitutions, out of which one (position 194) is an outlier showing a strong latitudinal clinal pattern. In experimental ocean acidification settings, sea urchin larvae have been found to up-regulate ATP synthase (Stumpp *et al.* 2011) and mantle tissue ATP levels in oysters have been seen to decrease significantly (Lannig *et al.* 2010), indicating an increased energy demand under high CO_2 stress.

Diversity, dispersal and selection in red abalone

Simultaneous analysis of transcriptome-wide heterozygosity, population differentiation and outlier SNPs suggests a view of abalone population structure that reconciles different views of this species. Low planktonic duration suggests that short-distance dispersal is common in abalone, and other species in this genus show genetic differentiation across short spatial scales (Huang *et al.* 2000; Gruenthal & Burton 2008). However, high heterozygosity suggests very long-term, high population size in red abalone, similar to that of other benthic invertebrates such as sea urchins and oysters (Palumbi 1994). Under such conditions, a small amount of gene flow or small selective difference can be more powerful in generating differentiation than genetic drift. One prediction of this model of population genetics is dominance of genetic patterns by a large number of neutral alleles with no differentiation, overlain on a set of loci with greater than expected genetic variation due to selection. We observe this pattern in red abalone: 21 000 SNPs with no significant difference among populations and about 700 (ca. 3.3%) with greater than expected gene frequency shifts.

Loci indicated as experiencing differential selective pressures along the California coast are enriched for genes that affect the calcareous shell formation, across a mosaic of ocean chemistry that is known to affect calcification (Kroeker *et al.* 2010). As in purple sea urchins, this may allow local selection for alleles that operate better in high CO_2, low pH conditions, given a high enough pre-existing genetic variability. Stressful environmental conditions are expected to increase in severity in future oceans due to build-up of anthropogenic CO_2 (Sabine *et al.* 2004; Feely *et al.* 2009), so the existence of standing genetic variation for adaptation in a future acidic ocean is an important finding. The candidate genes uncovered by our analysis do not represent a comprehensive list—our analysis will have missed selection on larval genes, reproductive genes or genes with low transcript abundance in mantle tissue. The patterns we report here suggest further work on the small-scale geographic variation of variants sites in these genes, functional studies of the relationship between alleles and fitness, and direct experiments on the impact of high CO_2 on gene frequencies.

Conclusion

By using RNA-Seq to conduct a transcriptome-wide scan of the red abalone, *Haliotis rufescens*, we can confirm low levels of genetic differentiation between geographic locations, along with a substantial number of

loci under differential selective pressure along the California coast. The combination of high heterozygosity and low differentiation of abalone despite low dispersal potential points to a widely distributed large population linked by small amounts of gene flow. Under these circumstances, the impact of selection on gene frequencies is predicted to be particularly visible. Loci that appear to be under selection are involved in biomineralization, resistance to hypoxia, response to heat/pathogens and energy metabolism. Upwelling of low pH water to the surface in northern California (Feely *et al.* 2008) may be responsible for some of the differences in biomineralization genes. Exposure to hypoxia (Bograd *et al.* 2008; Booth *et al.* 2010) and pathogens (Bergen & Raimondi 2001) in Monterey Bay might also be the cause of differential selective pressures.

Acknowledgements

References

Abe H, Hirai S, Okada S (2007) Metabolic responses and arginine kinase expression under hypoxic stress of the kuruma prawn *Marsupenaeus japonicus*. *Comparative Biochemistry and Physiology. Part A, Molecular & integrative physiology*, **146**, 40–46.

Antao T, Lopes A, Lopes RJ, Beja-Pereira A, Luikart G (2008) LOSITAN: a workbench to detect molecular adaptation based on a F_{ST}-outlier method. *BMC Bioinformatics*, **9**, 323.

Anthony KRN, Maynard JA, Diaz-Pulido G *et al.* (2011) Ocean acidification and warming will lower coral reef resilience. *Global Change Biology*, **17**, 1798–1808.

Aristegui J, Barton ED, Alvarez-Salgado XA *et al.* (2009) Sub-regional ecosystem variability in the Canary Current upwelling. *Progress in Oceanography*, **83**, 33–48.

Belkin IM, Cornillon PC, Sherman K (2009) Fronts in large marine ecosystems. *Progress in Oceanography*, **81**, 223–236.

Beniash E, Ivanina A, Lieb NS, Kurochkin I, Sokolova IM (2010) Elevated level of carbon dioxide affects metabolism and shell formation in oysters *Crassostrea virginica*. *Marine Ecology Progress Series*, **419**, 95–108.

Bergen L, Raimondi PT (2001) PISCO update: withering syndrome in black abalone. In: *MBNMS Ecosystem Observations* (ed. Carless J), pp. 9–10. NOAA, Monterey, CA.

Bograd SJ, Castro CG, Di Lorenzo E *et al.* (2008) Oxygen declines and the shoaling of the hypoxic boundary in the California Current. *Geophysical Research Letters*, **35**, 6.

Booth JAT, McPhee-Shaw E, Phillips R, Chua P, Gilly W (2010) Naturally occurring coastal hypoxia in Monterey Bay. Poster presented at MBNMS Sanctuary Currents Symposium 2010. Monterey, CA.

Booth DJ, Bond N, Macreadie P (2011) Detecting range shifts among Australian fishes in response to climate change. *Marine and Freshwater Research*, **62**, 1027–1042.

Britten RJ, Cetta A, Davidson EH (1978) The single-copy DNA sequence polymorphism of the sea urchin *Strongylocentrotus purpuratus*. *Cell*, **15**, 1175–1186.

Burton RS, Tegner MJ (2000) Enhancement of red abalone *Haliotis rufescens* stocks at San Miguel Island: reassessing a success story. *Marine Ecology Progress Series*, **202**, 303–308.

CADFG (2011) Special Animals (898 taxa). California Department of Fish and Game. Available online at http://www.dfg.ca.gov/biogeodata/cnddb/pdfs/SPAnimals.pdf

Checkley DM Jr, Barth JA (2009) Patterns and processes in the California current system. *Progress in Oceanography*, **83**, 49–64.

Clark NL, Gasper J, Sekino M, Springer SA, Aquadro CF, Swanson WJ (2009) Coevolution of interacting fertilization proteins. *PLoS Genetics*, **5**, e1000570.

Connolly TP, Hickey BM, Geier SL, Cochlan WP (2010) Processes influencing seasonal hypoxia in the northern California Current System. *Journal of Geophysical Research*, **115**, C03021.

Conover DO, Clarke LM, Munch SB, Wagner GN (2006) Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *Journal of Fish Biology*, **69**, 21–47.

Crim RN, Sunday JM, Harley CDG (2011) Elevated seawater CO_2 concentrations impair larval development and reduce larval survival in endangered northern abalone (*Haliotis kamtschatkana*). *Journal of Experimental Marine Biology and Ecology*, **400**, 272–277.

Curole JP, Hedgecock D (2005) Estimation of preferential pairing rates in second-generation autotetraploid Pacific oysters (Crassostrea gigas). *Genetics*, **171**, 855–859.

de la Cruz FL, Gallardo-Escarate C (2010) Intraspecies and interspecies hybrids in *Haliotis*: natural and experimental evidence and its impact on abalone aquaculture. *Reviews in Aquaculture*, **3**, 74–99.

De Wit P, Pespeni MH, Ladner JT *et al.* (2012) The simple fool's guide to population genomics via RNA-Seq: gene expression and SNP data analysis in the age of high-throughput sequencing. *Molecular Ecology Resources*, doi: 10.1111/1755-0998.12003.

De Pristo M, Banks E, Poplin R *et al.* (2011) A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nature Genetics*, **43**, 491–498.

Díaz-Viloria N, Cruz P, Guzmán del Próo SA, Perez-Enriquez R (2009) Genetic connectivity among pink abalone *Haliotis corrugata* populations. *Journal of Shellfish Research*, **28**, 599–608.

Erez J, Reynaud S, Silverman J, Schneider K, Allemand D (2011) Coral Calcification Under Ocean Acidification and Global Change In: *Coral Reefs: An Ecosystem in Transition* (ed. Dubinsky ZSN), pp. 151–176.

Estes JA, Lindberg DR, Wray C (2005) Evolution of large body size in abalones (*Haliotis*): patterns and implications. *Paleobiology*, **31**, 591–606.

Fabry VJ, Siebel BA, Feely RA, Orr JC (2008) Impacts of ocean acidification on marine fauna and ecosystem processes. *ICES Journal of Marine Science*, **65**, 414–432.

Fan W, Li C, Li S, Feng Q, Xie L, Zhang R (2007) Cloning, characterization, and expression patterns of three sarco/endoplasmic reticulum Ca^{2+}-ATPase isoforms from pearl oyster (*Pinctada fucata*). *Acta Biochimica et Biophysica Sinica*, **39**, 722–730.

Feely RA, Sabine CL, Lee K *et al.* (2004) Impact of anthropogenic CO_2 on the $CaCO_3$ system in the oceans. *Science*, **305**, 362–366.

Feely RA, Sabine CL, Hernandez-Ayon M, Ianson D, Hales B (2008) Evidence for upwelling of corrosive "acidified" water onto the continental shelf. *Science*, **320**, 1490–1492.

Feely RA, Doney SC, Cooley SR (2009) Ocean acidification: present conditions and future changes in a high-CO_2 world. *Oceanography*, **22**, 36–47.

Ferreira GC, Gong J (1995) 5-aminolevulinate synthase and the first step of heme-biosynthesis. *Journal of Bioenergetics and Biomembranes*, **27**, 151–159.

Fransson A, Chierici M, Nojiri Y (2009) New insights into the spatial variability of the surface water carbon dioxide in varying sea ice conditions in the Arctic Ocean. *Continental Shelf Research*, **29**, 1317–1328.

Furuhashi T, Miksik I, Smrz M *et al.* (2010) Comparison of aragonitic molluscan shell proteins. *Comparative Biochemistry and Physiology. Part B, Biochemistry & Molecular Biology*, **155**, 195–200.

Gibbs RA, Belmont JW, Hardenbol P *et al.* (2003) The International HapMap Project. *Nature*, **426**, 789–796.

Gruenthal KM, Burton RS (2005) Genetic diversity and species identification in the endangered white abalone (*Haliotis sorenseni*). *Conservation Genetics*, **6**, 929–939.

Gruenthal KM, Burton RS (2008) Genetic structure of natural populations of the California black abalone (*Haliotis cracherodii* Leach, 1814), a candidate for endangered species status. *Journal of Experimental Marine Biology and Ecology*, **355**, 47–58.

Gruenthal KM, Acheson LK, Burton RS (2007) Genetic structure of natural populations of California red abalone (*Haliotis rufescens*) using multiple genetic markers. *Marine Biology*, **152**, 1237–1248.

Gutierrez-Gonzalez JL, Cruz P, del Rio-Portilla MA, Perez-Enriquez R (2007) Genetic structure of green abalone *Haliotis fulgens* population off Baja California, Mexico. *Journal of Shellfish Research*, **26**, 839–846.

Hauri C, Gruber N, Plattner G-K *et al.* (2009) Ocean acidification in the California current system. *Oceanography*, **22**, 60–71.

Hellberg ME (2009) Gene flow and isolation among populations of marine animals. *Annual Review of Ecology, Evolution, and Systematics*, **40**, 291–310.

Hereford J (2009) A quantitative survey of local adaptation and fitness trade-offs. *The American Naturalist*, **173**, 579–588.

Hofmann GE, Barry JP, Edmunds PJ *et al.* (2010) The effect of ocean acidification on calcifying organisms in marine ecosystems: an organism-to-ecosystem perspective. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 127–147.

Huang BX, Peakall R, Hanna PJ (2000) Analysis of genetic structure of blacklip abalone (*Haliotis rubra*) populations using RAPD, minisatellite and microsatellite markers. *Marine Biology*, **136**, 207–216.

Huson DH, Mitra S, Ruscheweyh H-J, Weber N, Schuster SC (2011) Integrative analysis of environmental sequences using MEGAN4. *Genome Research*, **21**, 1552–1560.

Huvet A, Herpin A, Dégremont L, Labreuche Y, Samain J-F, Cunningham C (2004) The identification of genes from the oyster *Crassostrea gigas* that are differentially expressed in progeny exhibiting opposed susceptibility to summer mortality. *Gene*, **343**, 211–220.

Johnson CR, Banks SC, Barrett NS *et al.* (2011) Climate change cascades: shifts in oceanography, species' ranges and subtidal marine community dynamics in eastern Tasmania. *Journal of Experimental Marine Biology and Ecology*, **400**, 17–32.

Jones SJ, Lima FP, Wethey DS (2010) Rising environmental temperatures and biogeography: poleward range contraction of the blue mussel, Mytilus edulis L., in the western Atlantic. *Journal of Biogeography*, **37**, 2243–2259.

Kang J-H, Appleyard SA, Elliott NG *et al.* (2010) Development of genetic markers in abalone through construction of a SNP database. *Animal Genetics*, **42**, 309–315.

Kroeker KJ, Kordas RL, Crim RN, Singh GG (2010) Meta-analysis reveals negative yet variable effects of ocean acidification on marine organisms. *Ecology Letters*, **13**, 1419–1434.

Kröger N (2009) The molecular basis of nacre formation. *Science*, **325**, 1351–1352.

Lannig G, Eilers S, Portner HO, Sokolova IM, Bock C (2010) Impact of ocean acidification on energy metabolism of oyster, *Crassostrea gigas*—changes in metabolic pathways and thermal response. *Marine Drugs*, **8**, 2318–2339.

Li H, Durbin R (2009) Fast and accurate short read alignment with Burrows-Wheeler Transform. *Bioinformatics*, **25**, 1754–1760.

Lough JM, Hobday AJ (2011) Observed climate change in Australian marine and freshwater environments. *Marine and Freshwater Research*, **62**, 984–999.

Martin JA, Wang Z (2011) Next-generation transcriptome assembly. *Nature Reviews Genetics*, **12**, 671–682.

Maynard BT, Hanna PJ, Benzie JAH (2004) Microsatellite DNA analysis of southeast Australian Haliotis laevigata (Donovan) populations—implications for Ranching in Port Phillip Bay. *Journal of Shellfish Research*, **23**, 1195–1200.

McKenna A, Hanna M, Banks E *et al.* (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Research*, **20**, 1297–1303.

Miller AW, Reynolds AC, Sobrino C, Riedel GF (2009a) Shellfish face uncertain future in high CO_2 world: influence of acidification on oyster larvae calcification and growth in estuaries. *PLoS ONE*, **4**, e5661.

Miller KJ, Maynard BT, Mundy CN (2009b) Genetic diversity and gene flow in collapsed and healthy abalone fisheries. *Molecular Ecology*, **18**, 200–211.

Moore JD, Finley CA, Robbins TT, Friedman CS (2002) Withering syndrome and restoration of southern California abalone populations. *CalCOFI Report*, **43**, 112–117.

Moore JD, Marshman BC, Chun CSY (2011) Health and survival of red abalone *Haliotis rufescens* from San Miguel Island, California, USA, in a laboratory simulation of La Nina and El Nino conditions. *Journal of Aquatic Animal Health*, **23**, 78–84.

Narum SR, Hess JE (2011) Comparison of F_{ST} outlier tests for SNP loci under selection. *Molecular Ecology Resources*, **11**, 184–194.

Nielsen EE, Hemmer-Hansen J, Poulsen NA *et al.* (2009) Genomic signatures of local directional selection in a high gene

flow marine organism; the Atlantic cod (*Gadus morhua*). *BMC Evolutionary Biology*, **9**, 276.

North A, Pennanen J, Ovaskainen O, Laine A-L (2011) Local adaptation in a changing world: the roles of gene-flow, mutation, and sexual reproduction. *Evolution*, **65**, 79–89.

Palumbi SR (1994) Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics*, **25**, 547–572.

Patterson N, Price AL, Reich D (2006) Population structure and eigenanalysis. *PLoS Genetics*, **2**, e190.

Pespeni MH, Palumbi SR (2010) The purple sea urchin genome suggests local adaptation along a latitudinal gradient despite high gene flow. *Integrative and Comparative Biology*, **50**, E136.

Pespeni MH, Garfield DA, Manier MK, Palumbi SR (2011) Genome-wide polymorphisms show unexpected targets of natural selection. *Proceedings of the Royal Society of London. Series B*, **279**, 1412–1420.

Philippart CJM, Anadon R, Danovaro R et al. (2011) Impacts of climate change on European marine ecosystems: observations, expectations and indicators. *Journal of Experimental Marine Biology and Ecology*, **400**, 52–69.

Pinsky ML, Montes HR Jr, Palumbi SR (2010) Using isolation by distance and effective density to estimate dispersal scales in anemonefish. *Evolution*, **64**, 2688–2700.

Pörtner HO, Peck MA (2010) Climate change effects on fishes and fisheries: towards a cause-and-effect understanding. *Journal of Fish Biology*, **77**, 1745–1779.

Prince JD, Sellers TL, Ford WB, Talbot SR (1987) Experimental evidence for limited dispersal of haliotid larvae (genus *Haliotis*; Mollusca: Gastropoda). *Journal of Experimental Marine Biology and Ecology*, **106**, 243–263.

Raimondi PT, Wilson CM, Ambrose RF, Engle JM, Minchinton TE (2002) Continued declines of black abalone along the coast of California: are mass mortalities related to El Niño events? *Marine Ecology Progress Series*, **242**, 143–152.

Reid PC, Fischer AC, Lewis-Brown E et al. (2009) Impacts of the oceans on climate change. In: *Advances in Marine Biology*, Vol 56 (ed. Sims DW), pp. 1–150. Elsevier, Cambridge, MA.

Rio DC, Ares M Jr, Hannon GJ, Nilsen TW (2010) Purification of RNA Using TRIzol (TRI Reagent). *Cold Spring Harbor Protocols*, **6**, doi: 10.1101/pdb.prot5439.

Rogers-Bennett L, Dondanville RF, Moore JD, Vilchis LI (2010) Response of red abalone reproduction to warm water, starvation, and disease stressors: implications of ocean warming. *Journal of Shellfish Research*, **29**, 599–611.

Rousset F (1999) Genetic differentiation within and between two habitats. *Genetics*, **151**, 397–407.

Rousset F (2008) Genepop'007: a complete reimplementation of the Genepop software for Windows and Linux. *Molecular Ecology Resources*, **8**, 103–106.

Sabine CL, Feely RA, Gruber N et al. (2004) The oceanic sink for anthropogenic CO_2. *Science*, **305**, 367–371.

Sanford E, Kelly MW (2011) Local adaptation in marine invertebrates. *Annual Review of Marine Science*, **3**, 509–535.

Sauvage C, Bierne N, Lapegue S, Boudry P (2007) Single Nucleotide polymorphisms and their relationship to codon usage bias in the Pacific oyster Crassostrea gigas. *Gene*, **406**, 13–22.

Shu J, Li Q, Yu R-H, Tian C-Y (2008) Microsatellites analysis on genetic variation between wild and cultured populations of Pacific abalone (*Haliotis discus hannai*). *Zhongguo Haiyang Daxue Xuebao*, **38**, 52–58.

Slatkin M (1993) Isolation by distance in equilibrium and non-equilibrium populations. *Evolution*, **47**, 264–279.

Sodergren E, Weinstock GM, Davidson EH et al. (2006) Research article—the genome of the sea urchin Strongylocentrotus purpuratus. *Science*, **314**, 941–952.

Solomon S, Qin D, Manning M et al. (2007) Climate Change 2007: The Physical Science Basis. Cambridge University Press, Cambridge.

Somero GN (2010) The physiology of climate change: how potentials for acclimatization and genetic adaptation will determine 'winners' and 'losers'. *Journal of Experimental Biology*, **213**, 912–920.

Somero GN (2012) The physiology of global change: linking patterns to mechanisms. In: *Annual Review of Marine Science*, Vol 4 (ed. Carlson CAGSJ), pp. 39–61. Annual Reviews, Palo Alto, CA.

Sotka EE, Palumbi SR (2006) The use of genetic clines to estimate dispersal distances of marine larvae. *Ecology*, **87**, 1094–1103.

Stramma L, Schmidtko S, Levin LA, Johnson GC (2010) Ocean oxygen minima expansions and their biological impacts. *Deep-Sea Research I*, **57**, 587–595.

Stumpp M, Dupont S, Thorndyke MC, Melzner F (2011) CO(2) induced seawater acidification impacts sea urchin larval development II: gene expression patterns in pluteus larvae. *Comparative Biochemistry and Physiology. Part A, Molecular & Integrative Physiology*, **160**, 320–330.

Suzuki M, Saruwatari K, Kogure T et al. (2009) An acidic matrix protein, Pif, is a key macromolecule for nacre formation. *Science*, **325**, 1388–1390.

Swanson WJ, Vacquier VD (1995) Extraordinary divergence and positive Darwinian selection in a fusagenic protein coating the acrosomal process of abalone spermatozoa. *Proceedings of the National Academy of Sciences*, **92**, 4957–4961.

Swanson WJ, Aquadro CF, Vacquier VD (2001) Polymorphism in abalone fertilization genes is consistent with the neutral evolution of the egg's receptor for lysin (VERL) and positive selection of sperm lysin. *Molecular Biology and Evolution*, **18**, 376–383.

Talmage SC, Gobler CJ (2009) The effects of elevated carbon dioxide concentrations on the metamorphosis, size, and survival of larval hard clams (*Mercenaria mercenaria*), bay scallops (*Argopecten irradians*), and eastern oysters (*Crassostrea virginica*). *Limnology and Oceanography*, **54**, 2072–2080.

Tang S, Tassanakajon A, Klinbunga S, Jarayabhand P, Menasveta P (2004) Population structure of tropical abalone (*Haliotis asinina*) in coastal waters of Thailand determined using microsatellite markers. *Marine Biotechnology*, **6**, 604–611.

Wang X, Fan W, Xie L, Zhang R (2008) Molecular cloning and distribution of a plasma membrane Calcium ATPase homolog from the Pearl Oyster *Pinctada fucata*. *Tsinghua Science and Technology*, **13**, 439–446.

Waples RS (1998) Separating the wheat from the chaff: patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, **89**, 438–450.

Wassmann P, Duarte CM, Agusti S, Sejr MK (2011) Footprints of climate change in the Arctic marine ecosystem. *Global Change Biology*, **17**, 1235–1249.

Wernberg T, Russell BD, Moore PJ et al. (2011) Impacts of climate change in a global hotspot for temperate marine

biodiversity and ocean warming. *Journal of Experimental Marine Biology and Ecology*, **400**, 7–16.

Wethey DS, Woodin SA, Hilbish TJ, Jones SJ, Lima FP, Brannock PM (2011) Response of intertidal populations to climate: effects of extreme events versus long term change. *Journal of Experimental Marine Biology and Ecology*, **400**, 132–144.

Yokoyama Y, Hashimoto H, Kubota S *et al.* (2006) cDNA cloning of Japanese oyster stress protein homologous to the mammalian 78-kDa glucose regulated protein and its induction by heatshock. *Fisheries Science*, **72**, 402–409.

Yu PC, Matson PG, Martz TR, Hofmann GE (2011) The ocean acidification seascape and its relationship to the performance of calcifying marine invertebrates: laboratory experiments on the development of urchin larvae framed by environmentally-relevant pCO(2)/pH. *Journal of Experimental Marine Biology and Ecology*, **400**, 288–295.

Zhang J, Gilbert D, Gooday AJ *et al.* (2010) Natural and human-induced hypoxia and consequences for coastal areas: synthesis and future development. *Biogeosciences*, **7**, 1443–1467.

Zippay ML, Hofmann GE (2010) Effect of pH on gene expression and thermal tolerance of early life history stages of red abalone (*Haliotis rufescens*). *Journal of Shellfish Research*, **29**, 429–439.

Zoccola D, Tambutte E, Kulhanek E *et al.* (2004) Molecular cloning and localization of a PMCA P-type calcium ATPase from the coral *Stylophora pistillata*. *Biochimica Et Biophysica Acta-Biomembranes*, **1663**, 117–126.

Data accessibility

Supplementary materials, including the *de novo* transcriptome assembly, a list of all detected variant sites plus genotypes for all individuals at those sites and information on all top hits from BLAST searches to NCBI's nr, nt, the Uniprot knowledgebase, and the *Lottia gigantea* genome, as well as GO annotation are available for download at the Dryad Digital depository (doi:10.5061/dryad.85p80).