Genomic evidence for ecological divergence against a background of population homogeneity in the marine snail *Chlorostoma funebralis*

LANI U. GLEASON*† and RONALD S. BURTON*

*Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, CA 92093-0202, USA, †Department of Biology, Loyola Marymount University, Los Angeles, CA 90045, USA

Abstract

The balance between natural selection, gene flow and genetic drift is difficult to resolve in marine invertebrates with extensive dispersal and fluctuating population sizes. The intertidal snail *Chlorostoma funebralis* has planktonic larvae and previous work using mtDNA polymorphism reported no genetic population structure. Nevertheless, recent studies have documented differences in thermal tolerance and transcriptomic responses to heat stress between northern and southern California, USA, populations. To gain insight into the dynamics influencing adaptive divergence, we used double-digest restriction site-associated DNA (ddRAD) sequencing to identify 1861 genomewide, quality-filtered single-nucleotide polymorphism (SNP) loci for *C. funebralis* collected from three northern and three southern California sites (15 individuals per population). Considering all SNPs, there was no evidence for genetic differentiation among populations or regions (average F_{ST} = 0.0042). However, outlier tests revealed 34 loci putatively under divergent selection between northern and southern populations, and structure and SNP tree analyses based on these outliers show clear genetic differentiation between geographic regions. Three of these outliers are known or hypothesized to be involved in stress granule formation, a response to environmental stress such as heat. Combined with previous work that found thermally tolerant southern populations show high baseline expression of stress response genes, these results further suggest that thermal stress is a strong selective pressure across *C. funebralis* populations. Overall, this study increases our understanding of the factors constraining local adaptation in marine organisms, while suggesting that ecologically driven, strong differentiation can occur at relevant loci in a species with planktonic larvae.

Keywords: adaptation, ecological divergence, mollusc, RAD sequencing, rocky intertidal

Received 25 November 2015; revision received 29 April 2016; accepted 3 May 2016

Introduction

Many marine organisms are distributed widely across heterogeneous landscapes, and across these ranges,

Correspondence: Lani U. Gleason, Fax: 310 338 5317;
E-mail: lani.gleason@lmu.edu

The drop-a-drifter tool used in this publication can be accessed at http://west.rssoffice.com/ca_roms?drifter=active. It is developed and maintained by Dr. Yi Chao and his team at Remote Sensing Solutions, Inc. and Joint Institute for Regional Earth System Science and Engineering at UCLA.

natural selection can promote genetic differentiation and local adaptation. However, evolution of ecological divergence may be impeded if high rates of migration homogenize the gene pool among populations (Mayr 1963; Lewontin 1974; Slatkin 1985; Lenormand 2002). In many species with planktonic larvae and possible long-distance pelagic dispersal, the potential for local adaptation is unclear because the balance between selection for local adaptation and the rate of interpopulation gene flow is largely unknown. Nevertheless, numerous studies suggest that marine populations are not as

connected as might be presumed (Burton 1983; Kyle & Boulding 2000; Levin 2006; Marshall *et al.* 2010), and adaptive differentiation has been observed in various marine invertebrates with planktonic larvae such as mussels, oysters, barnacles, sea urchins, and abalone (Koehn *et al.* 1980; Karl & Avise 1992; Schmidt & Rand 1999, 2001; Pespeni *et al.* 2012; De Wit & Palumbi 2013).

The advent and increasingly widespread use of high-throughput next-generation sequencing, such as double-digest restriction site-associated DNA (ddRAD) sequencing (Peterson *et al.* 2012), facilitates the identification of such local adaptation through genomewide scans (Davey *et al.* 2011). For instance, RAD sequencing has been used to investigate local adaptation in marine gastropods such as the green abalone *Haliotis fulgens* (Gruenthal *et al.* 2014), the North Atlantic rocky intertidal snail *Nucella lapillus* (Chu *et al.* 2014) and the marine periwinkle *Littorina saxatilis* (Westram *et al.* 2014; Ravinet *et al.* 2016). Using such methods to investigate populations undergoing early stages of ecological divergence is particularly useful because they provide good opportunities to detect genomic loci under divergent selection (see Russello *et al.* 2012; Lemay & Russello 2015), with such loci expected to stand out with high F_{ST} estimates against a background of low genetic divergence (Pérez-Figueroa *et al.* 2010; Vilas *et al.* 2012). However, it is important to note that these genome scan approaches also have limitations; for instance, the high F_{ST} outlier loci identified can sometimes be due to chance or demographic processes rather than local adaptation (Bierne *et al.* 2011). Nevertheless, such scans can serve as a useful first pass (especially for nonmodel organisms) to identify candidate loci for ecological divergence that can then be subjected to further functional tests. Overall, these types of genomewide studies that examine both background levels of genetic divergence as well as locus-specific measures of population differentiation, especially in the sea, are valuable because the factors that promote or constrain local adaptation in marine organisms are still relatively unexplored.

Chlorostoma funebralis, a mid-intertidal snail, is an excellent system to lend insight into the balance between local adaptation and gene flow because it is influenced by several factors that alternately encourage and deter local adaptation. *C. funebralis* has a wide latitudinal range: it is found in rocky habitat along the Pacific coast of North America from Vancouver Island, British Columbia to Baja California, Mexico (Abbott & Haderlie 1980; Sagarin & Gaines 2002). Across this range, individuals encounter highly heterogeneous environments, suggesting that natural selection might favour local adaptation; in fact, previous phenotypic and genetic work has shown that northern and

southern California populations are locally adapted to heat stress (Gleason & Burton 2013, 2015). Although several previous studies have indicated that intertidal thermal stress along the west coast of North America follows a 'mosaic' pattern rather than correlating with latitude (Helmuth *et al.* 2002, 2006), in situ temperature data indicate that chronic thermal stress (calculated as the average of all daily high temperatures) in *C. funebralis* does significantly correlate with latitude at the six geographic sampling sites used in this study (L. U. Gleason & R. S. Burton, in review). Therefore, latitudinal variation and associated differences in thermal stress between the northern and southern California populations used in this study could potentially be driving local adaptation in this species. However, *C. funebralis* also has planktonic larvae (Moran 1997) and no apparent genetic structure at cytochrome oxidase subunit I (COI) in individuals sampled from Oregon down to Santa Barbara (Kelly & Palumbi 2010; Kelly *et al.* 2010), which suggests a high level of gene flow that could preclude local adaptation.

This study's objective was to lend insight into the balance between local adaptation and gene flow in geographically distinct *C. funebralis* populations. We sampled 15 *C. funebralis* individuals from each of three northern and three southern California populations and then performed double-digest restriction site-associated (ddRAD) DNA sequencing to examine genomewide population structure and to test for genetic signatures of divergent selection.

Methods

Collection and animal maintenance

Small to medium-sized *C. funebralis* adults (15–20 mm in shell diameter) were collected in the winter of 2014 from three northern California sites: Slide Ranch (SR), Marin Co. (37°52′N, 122°35′W); Pescadero (PES; 37°15′N, 122°24′W) and Pigeon Point (PP; 37°11′N, 122°23′W), San Mateo Co., and from three southern California sites: Aliso Beach (AB), Orange Co. (33°30′N, 117°45′W), and La Jolla (LJ; 32°52′N, 117°15′W) and Bird Rock (BR; 32°48′N, 117°15′W), San Diego Co. (Fig. 1). Snails were collected across the length of all available intertidal habitat at each site, and individuals were obtained from all available locations (under rocks, in rock crevices, from both shaded and sun-exposed rock faces, etc.) at each site to avoid potential confounding effects of microhabitat (Byers 1983; Prada *et al.* 2014). Snails were transported to Scripps Institution of Oceanography (SIO) within 24 h of collection. Once at SIO, snails were maintained in flow-through running seawater aquaria and regularly

Fig 1 *Chlorostoma funebralis* collecting sites along the California coastline.

Slide Ranch
Pescadero
Pigeon Point

N

California

150 km

Aliso Beach
La Jolla
Bird Rock

Pacific Ocean

150 km

fed freshly collected kelp, *Macrocytis pyrifera*, until sacrificed.

DNA extraction, ddRAD library preparation and sequencing

Live animals (15 animals from each of the six populations) were placed in 95% ethanol for 24 h, and then tweezers and scissors were used to take a tissue sample of each animal's foot. DNA was extracted using the Qiagen DNeasy kit (Qiagen, Germantown, MD, USA). ddRAD libraries were constructed following a protocol adapted from Peterson *et al.* (2012). Briefly, DNA was digested with the restriction enzymes *MseI* and *XhoI*, and six ddRAD libraries were constructed (one library for each geographic site). Each library was created by uniquely barcoding each of the 15 individuals from the respective site and then pooling these 15 individually barcoded samples. The barcodes used were seven nucleotides in length, and each differed by at least two bases. The libraries were run on agarose gels, and the 300–400 bp range was manually excised, purified and enriched with 12 amplification cycles in six individual PCRs containing 4 μL of Phusion High-Fidelity HF

Buffer (NEB), 2 μL of each amplification primer (10 μM), 1 μL of dNTP mix (10 mM), 1 μL Phusion polymerase and 10 μL of library (~30 ng/μL; total reaction volume 20 μL). The six libraries (along with two libraries for another project) were sequenced across two lanes of an Illumina HiSeq2000 (paired end, 2×100 bp) at the University of California Irvine Genomics High-Throughput Facility.

Raw sequences filtering

Raw sequences were filtered using the *process_radtags* pipeline in STACKS version 1.21 (Catchen *et al.* 2011, 2013). Low-quality reads with a Phred quality score <10 as well as any reads with an uncalled base were removed, and reads were trimmed to 90 base pairs in length.

Assembly

There is no reference genome for *C. funebralis*, so reads were assembled de novo using the *denovo_map.pl* pipeline in Stacks. The number of raw reads required to form a RAD-tag (stack depth parameter, m) was set to

5, and the number of allowed nucleotide mismatches between two RAD-tags (mismatch parameter, M) was set to 2.

After assembly and genotyping, the data were filtered further to ensure maximal quality. As indicated in the Results, generally similar genomic patterns were observed with more stringent filtering (i.e. requiring that loci be present in a higher proportion of individuals per region in order to be analysed). Previous studies have documented differences in thermal tolerance and transcriptomic responses to heat stress between northern and southern California populations (Gleason & Burton 2013, 2015). For this study, we also focus on regional differences between northern and southern California to gain further insight into the genomewide patterns that could be contributing to these already established differences. SR, PES and PP were analysed as sites within the northern region, and AB, LJ and BR were analysed as sites within the southern region. Samples were pooled by geographic region ($n = 45$ samples per region), and using the POPULATIONS module in STACKS, we kept only those loci that were genotyped in more than 50% of individuals from each region, had an overall minor allele frequency above 0.1 and had a minimum coverage of $10\times$ per allele for each individual. To meet the assumptions for subsequent population genetics analysis and to prevent the analysis of physically linked loci, if a RAD-tag had more than one polymorphism, only one was retained. For the outlier analysis, to ensure that no potential outlier SNPs were missed, analyses were rerun with different SNP subsets from the same RAD-tags (Puebla et al. 2014). Lastly, to exclude artefacts such as duplicated loci and loci with null alleles (Lynch & Milligan 1994), which are likely to be more frequent in ddRAD studies compared to the original RAD-seq method (DaCosta & Sorenson 2014), we used GENEPOP version 4.2 (Raymond & Rousset 1995; Rousset 2008) to remove all loci that significantly deviated from Hardy–Weinberg equilibrium (HWE; Guo & Thompson 1992). This analysis tested for HWE separately in each of the two geographic regions (north and south). Any locus that significantly deviated from HWE in both regions was removed from the data set.

Population genetics statistics and isolation by distance

Pairwise F_{ST} for each locus (both between north/south geographic regions and also between each possible pair of six populations) was calculated in the POPULATIONS module of STACKS (Weir & Cockerham 1984). Geographic (shortest straight line) distances between each site were calculated using Google Earth. These measures of genetic and geographic distance between each of the six sites were used to test for isolation by distance as implemented in IBDWS version 3.23 (Bohonak 2002). Following Slatkin's recommendation (Slatkin 1993), the log of both genetic and geographic distance was used as input for the Mantel test for matrix correlation between genetic and geographic distance. Reduced major axis regression using 1000 randomizations to calculate the intercept and slope of the regression line of genetic distance vs. geographic distance was also performed.

F_{ST} outlier analysis

Using 1825 SNPs (overall MAF >0.1) and the 90 individuals, LOSITAN (Antao et al. 2008) was run using parameter settings of 50 000 simulations, confidence interval of 0.995, false discovery rate set to 0.1, subsample size of 50, simulated F_{ST} of 0.0073 and an attempted F_{ST} of 0.014. We considered loci candidates for positive selection above a probability level of 0.995. We tested for linkage disequilibrium (LD) in each pair of loci that were identified as candidates for positive selection using GENEPOP v.4.2 (Raymond & Rousset 1995).

To make sure the outlier loci we identified were not false positives, the SNP genotypes of all individuals were shuffled and randomly assigned to each of the 90 C. funebralis individuals, thereby creating two different groups of individuals that were determined by chance alone. Next we ran LOSITAN on these randomized groups, using the same parameters that were run with the actual data. We repeated the random genotype assignment and subsequent outlier detection analyses three times.

To assess whether a greater proportion of variation occurred among regions in the outlier loci data set compared to a neutral data set with outliers removed, an analysis of molecular variance (AMOVA) was performed on both the outlier and neutral (nonoutlier) loci. ARLEQUIN v.3.5.2.2 (Excoffier & Lischer 2010) was used to perform this analysis, which was based on 1000 permutations.

Annotating loci

Each RAD-tag that was identified as a potential LOSITAN outlier was subjected to both a BLASTX (default parameters) and a BLASTN (wordsize = 16; mismatch scores = 2, −3; maximum e-value = 10^{-15}; Altschul et al. 1997) search of all sequences in the NCBI database. Because each RAD-tag was relatively short (90 bp), we used stringent criteria to prevent inaccurate annotations. We only retained annotations with a unique Blast hit or with a top hit with an e-value an order of magnitude lower than the next closest Blast hit. Furthermore,

all RAD-tags were also mapped back to a *de novo* transcriptome derived from RNA-seq data from four populations of *C. funebralis* (Gleason & Burton 2015; L. U. Gleason & R. S. Burton, unpublished) using CLC Genomics Workbench 5.1 (read mapping parameters: minimum fraction length of read overlap = 0.8, minimum sequence similarity = 0.95). Any contig in the assembly that a RAD-tag uniquely mapped to was also subjected to a BLASTX and BLASTN search (default parameters for both), and this annotation was retained for the RAD-tag as well.

Structure analysis

We used STRUCTURE version 2.3 (Pritchard *et al.* 2000) to test for the presence of genetic population structure. Following 100 000 burn-in steps, the entire data set was analysed using 100 000 Markov chain steps. For this analysis, the admixture and LOCPRIOR models were used. Providing the two geographic regions as location data was deemed appropriate because the LOCPRIOR model was specifically designed for data sets with weak signatures of selection (see Results below), and perhaps more importantly, these models do not spuriously detect false population structure (Pritchard 2010 Structure manual). It is worth noting that while all analyses reported in this manuscript are based on regional geographic comparisons of northern and southern California, preliminary STRUCTURE runs using all identified SNPs (see Results below) and providing each of the six individual geographic sites as location data showed that $K = 1$ was most likely (based on highest mean ln $\mathrm{PR}(X|K)$, data not shown). This result indicates that the three sites within each geographic region were genetically similar enough to be pooled. We used five replicates for each value of K, and we tested K values (number of clusters) ranging from 1 to 6. Two different methods were used to infer the most likely number of clusters present in the data set. First, the number of clusters was determined by selecting the K value with the highest ln $\mathrm{Pr}(X|K)$ or the one after which the trend plateaus and that also provides consistent groupings across repeated runs (Pritchard *et al.* 2000). We also ran STRUCTURE HARVESTER (Earl & vonHoldt 2012), which utilizes the Delta K method (Evanno *et al.* 2005) to determine the most likely number of clusters, and CLUMP v.1.1.2 (Jakobsson & Rosenberg 2007) to average each individual's membership coefficient across the K value replicates. It is important to note that the ad hoc ΔK statistic does not apply when $K = 1$.

Genetic structure was also analysed with statistically supported outlier loci (see F_{ST} Outlier Analysis above) to investigate whether a signal of differentiation between the two geographic regions at several (potentially ecologically relevant) loci was evident once the effects of genomewide neutral background similarity were removed.

SNP trees

Maximum-likelihood trees were constructed from SNP data that was exported using STACKS, filtered to remove any invariant sites and then concatenated. Because all sites are variable (a requirement of the tree-building program) an ascertainment bias exists; this bias can be problematic because mean rates of evolution will be overestimated, resulting in overestimated branch lengths and possible errors in topology inference (Lewis 2001). To account for this issue, we applied an ascertainment bias correction to the likelihood calculations (Lewis 2001; Puebla *et al.* 2014). However, these trees should still be interpreted with these caveats in mind.

For the clustering analysis, all RAD-tags with coverage $\geq10\times$ in ≥23 individuals per geographic region were retained, keeping a single SNP per RAD-tag. The data were analysed with JMODELTEST version 2.1.7 (Guindon & Gascuel 2003; Darriba *et al.* 2012) and the Akaike information criterion, Bayesian information criterion and decision theory all indicated that a GTR+G model of nucleotide substitution was appropriate. RAXML version 8.1.17 (Stamatakis 2014) was used for the SNP tree construction, using the GTR+G model with ascertainment bias correction as recommended by the author for SNP data requiring a model of rate heterogeneity (RAXML v.8.1.17 manual, February, 2015; Puebla *et al.* 2014). The rapid bootstrap procedure was implemented (Stamatakis *et al.* 2008) with 100 replicates per run. Analyses were run with the entire SNP data set and then also repeated with outlier SNPs considered for the STRUCTURE clustering analyses.

Mitochondrial sequencing

ddRAD data (see Results below) and preliminary SNPs based on transcriptomes of southern LJ and northern PP individuals (L. U. Gleason & R. S. Burton, unpublished) indicated there were differences in the mitochondrial genome between northern and southern California populations of *C. funebralis.* To further investigate these mitochondrial differences, we extracted DNA from 92 individuals across the six populations collected from the field in the winter of 2014 (specifics listed below) using a Qiagen DNeasy kit (Qiagen) and designed primers for portions of the cytochrome b (forward primer: 5′AAAGTAAAAACAGCTAATACCACTC 3′, reverse primer: 5′ TTCCTGCTAATCCTTTAGT TACTCC 3′) and NADH dehydrogenase subunit

4 (forward primer: 5′ AGGCAAAGAAGCTGTAATAGT GT 3′, reverse primer: 5′ GAGTGGTATTAGCTGTTTT-TACTT 3′) genes based on previous transcriptome data (Gleason & Burton 2015). These two regions of the mitochondrial genome were amplified via PCR. The reactions for cytochrome b had a final volume of 25 μL and contained 1 μL of the DNA template, 10 μM of forward primer, 10 μM of reverse primer and 1X GoTaq Green Master Mix DNA polymerase (Promega Corporation, Madison, WI, USA). Reactions for NADH dehydrogenase subunit 4 were the same, except 5 μM of each primer was used. The thermal cycler profile for cytochrome b reactions was as follows: 95 °C for 2 min, 35 cycles of 95 °C for 30 s, 54 °C for 1 min and 72 °C for 2 min and a final extension of 72 °C for 7 min. PCR conditions for NADH dehydrogenase subunit 4 were identical, except an annealing temperature of 56 °C was used. For cytochrome b, a 572-bp fragment was amplified for 92 individuals total, 45 northern and 47 southern individuals (17 SR, 15 PES, 15 PP, 20 AB, 8 LJ and 17 BR). For NADH dehydrogenase subunit 4, a 614-bp fragment was amplified for 91 individuals total, 49 northern and 42 southern individuals (15 SR, 17 PES, 17 PP, 17 AB, 8 LJ and 17 BR).

After PCR, all products were run on 1.5% agarose gels and visualized with Gel Red Nucleic Acid Gel Stain (Biotium Inc., Hayward, CA, USA) to verify successful amplification. All successfully amplified PCR products were then cleaned with a QIAquick PCR purification kit (Qiagen), and purified PCR products were sequenced by Retrogen Inc. (San Diego, CA, USA). Sequences were then cleaned and trimmed with SEQUENCHER v. 5.3 (Gene Codes Corporation, Ann Arbor, MI, USA) and aligned in GENEIOUS v. 9 (Biomatters Limited, Auckland, New Zealand). As in the ddRAD data analysis, we focused on regional differences between northern and southern California; thus, sequences from SR, PP and PES were grouped together and treated as northern populations, and sequences from AB, LJ, and BR were grouped together and treated as southern populations. AMOVAs were run in PopART (http://popart.otago.ac.nz) to identify variation between the two geographic regions, and SNPs between northern and southern California were identified by eye from the Geneious alignments for each gene and tested for significance using Fisher's exact tests and 2 × 2 contingency tables.

Dispersal simulation

To investigate whether movement of larval migrants may be working against local adaptation in northern and southern California *C. funebralis* populations, we performed a dispersal simulation. *C. funebralis* are broadcast spawners, meaning individuals release gametes into the water column (Moran 1997). San Diego *C. funebralis* are known to spawn at least twice a year, once in the winter in January–February and once in the summer in July–August (Cooper 2010). (To date, no one has investigated the frequency of spawning events in northern California populations.) In addition, the estimated planktonic duration of larvae ranges from ~5 to 13 days at 13–15 °C (Moran 1997). Based on all of this information, a surface transport model available at http://west.rssoffice.com/CA/drifter.jsp was used to create particle drifter predictions as a proxy to estimate the path of *C. funebralis* planktonic larvae released from each of the six collecting sites used in this study. This approximation relies on the assumption that larvae act as passive particles at the ocean surface. In support of this assumption, Moran (1997) has shown that *C. funebralis* larvae swim to the water surface, at least in the laboratory. The surface transport models were set for a duration of 13 days (the maximum larval duration estimated by Moran (1997)), with 15 total particles released at the water surface at each of the six sites on 30 July and 17 August 2014, to estimate summer dispersal patterns, and on 1 January and 13 February 2015, to estimate winter dispersal patterns.

Results

Sequencing produced ~3.8 million reads per individual. Assembly and quality filtering (keeping only those SNP loci that were genotyped in more than 50% of individuals from each region, had an overall minor allele frequency above 0.1 and had a minimum coverage of 10× per allele for each individual) resulted in the identification of 1861 RAD-tags across all 90 individuals in both northern and southern California that contained a total of 2371 SNPs. Based on the fact that each RAD-tag is 90-bp long, these 1861 RAD-tags represent ~0.01% of the 1.5 pg genome for *C. funebralis* (Hinegardner 1974). Table 1 presents summary genomic data broken down by geographic region (the number of total nucleotide sites examined across the genome, number and proportion of polymorphic SNP loci identified, mean number of individuals sampled per SNP locus, nucleotide diversity (π) and expected heterozygosity for northern and southern California separately). The two regions had very similar parameters, with marginally higher number of total nucleotide sites and number of polymorphic SNP loci for southern California.

Population genetics statistics

Average F_{ST} between the two regions (considering all SNPs) was estimated to 0.0042. An estimate of 0.0021

Table 1 Number of total nucleotide sites examined, number and proportion of polymorphic SNP loci, mean number of individuals sampled per SNP locus, nucleotide diversity (π) and expected heterozygosity of the two geographic regions considered in this study

	Northern California ($n = 45$)	Southern California ($n = 45$)
Number of nucleotide sites	1 119 516	1 119 545
Number of polymorphic SNP loci	1906	1937
Proportion of polymorphic SNP loci (%)	0.1703	0.173
Mean n per SNP locus	32.9	31.8
Nucleotide diversity (π)	0.0006	0.0006
Expected heterozygosity	0.0006	0.0006

was obtained when considering only one SNP per locus, and an estimate of 0.0027 was obtained when applying more stringent filtering (loci present in 75% of the individuals per region instead of 50%). The distribution of F_{ST} estimates was characterized by a sharp mode close to 0 and a long tail extending to a value of 0.59 (Fig. 2).

Isolation by distance

For this and all subsequent population genetics analyses, we further filtered the full SNP data set by retaining only a single SNP per RAD-tag (the first one), resulting in 1861 SNPs. (410 tags had more than one SNP). From this, 228 SNPs were removed that significantly deviated from HWE in both northern and southern populations, resulting in 1633 high-quality SNPs. Using these 1633 SNPs, a Mantel test for a significant relationship between the genetic and geographic distances between sampling sites found no evidence for matrix correlation, with a Z score of -75.516, an r value of 0.367 and a one-sided P-value of 0.203 (Fig. S1, Supporting information). Reduced major axis regression estimated an R^2 of 0.135 for the regression line of genetic vs. geographic distance.

F_{ST} outlier analysis

The same 1633 SNP loci used in the isolation-by-distance analysis above were also used for the outlier analysis, but as multiple SNPs within the same RAD-tag were examined (see Methods above), a total of 1825 SNPs were analysed. (SNPs in the additional data set for outlier analysis were also tested for HWE, and 0 was removed). A total of 34 outliers were identified, representing 0.019% of the SNPs analysed (Fig. S2, Supporting information). Six pairs of loci showed evidence for linkage disequilibrium, and thus, one of each pair (the locus with the lower probability of being an outlier, as calculated by LOSITAN) was removed from all subsequent analyses. Notably, one of the remaining 28 loci, 23143, contained two outlier SNPs at base pair positions 14 and 78 and thus is a
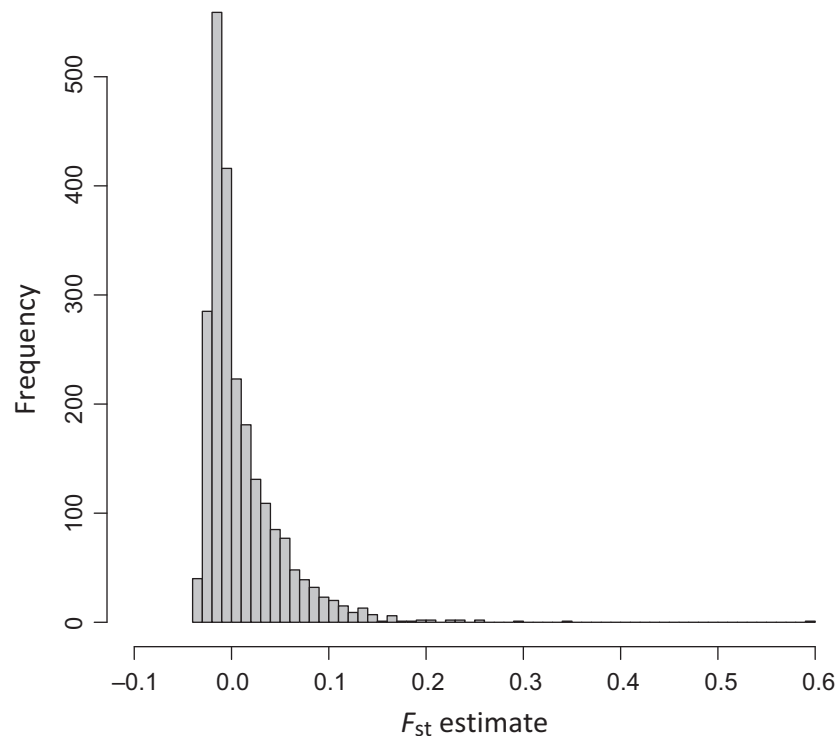

Fig 2 Frequency distribution of F_{ST} values obtained across all loci ($n = 2371$; mean $F_{ST} = 0.0042$).

particularly good candidate for being a target of divergence caused by natural selection (De Wit & Palumbi 2013). However, this locus was not annotated, and thus, its identity and function remain unknown. AMOVA results indicate that a relatively large proportion of variation (~19%) occurs between the geographic regions of northern and southern California for the outlier loci, compared to only 0.9% for the nonoutlier loci (Table 2).

Three negative control LOSITAN runs (based on randomizing the data) identified a total of nine, seven and nine outliers (0.005, 0.004 and 0.005% of all SNPs examined) compared to 34 (0.019%) outliers identified in the actual data set (exact binomial test comparing 9 vs. 34 outliers, P-value <0.0001). These results suggest that the genomic patterns reported in this study are not expected by chance alone.

Within the outlier data set, seven RAD-tags matched sequences from the NCBI nr database (Table 3). Three outlier loci, cytoplasmic dynein 2 heavy chain, MAP/microtubule affinity-regulating kinase 1 (MARK1) and poly [ADP-ribose] polymerase 14 (PARP14), were annotated to genes known or hypothesized to be involved in cytoplasmic stress granule formation, which is triggered by environmental stress such as heat. In addition, one locus is a small conductance calcium-activated potassium channel and another, dipeptidyl-peptidase 6 (DPP6), has been reported to regulate potassium channel activity (Soh & Goldstein 2008); high differentiation at these two outliers could be a result of thermally driven divergent selection between northern and southern California, as potassium channel activation has been suggested to be an early event in initiation of the heat-shock response (Saad & Hahn 1992; Pell et al. 1997), and northern and southern California populations have unique profiles of heat-shock protein gene expression following thermal stress (Gleason & Burton 2015). Another outlier matched a mitochondrial gene, NADH dehydrogenase subunit 1.

It is worth noting that underlying population structure is a major limitation of F_{ST} outlier approaches (Excoffier & Lischer 2010); however, given that we found only slight population structure with the outlier loci (see Structure analysis and SNP trees below), this is unlikely to be a confounding factor in this instance. In other words, minimal population structure between northern and southern California provides a less noisy neutral background upon which adaptive markers can be more reliably distinguished as outliers (Pérez-Figueroa et al. 2010; Hess et al. 2013).

Lastly, while all results and analyses discussed in this manuscript compare the geographic regions of northern and southern California, preliminary LOSITAN runs did identify outlier loci when each of the six geographic sampling locations was considered its own site (data not shown). However, because the goal of this study was to gain further insight into the previously documented regional differences in thermal tolerance (Gleason & Burton 2013) and transcriptomic response to heat stress (Gleason & Burton 2015) between northern and southern California C. funebralis populations, the data from these preliminary LOSITAN runs are not presented here.

Structure analysis

Results of the STRUCTURE analysis are shown in Fig. 3a–f and further detailed in Table S1 (Supporting information). Using the filtered data set of 1633 SNPs (considering a single SNP per RAD-tag), no consistent evidence
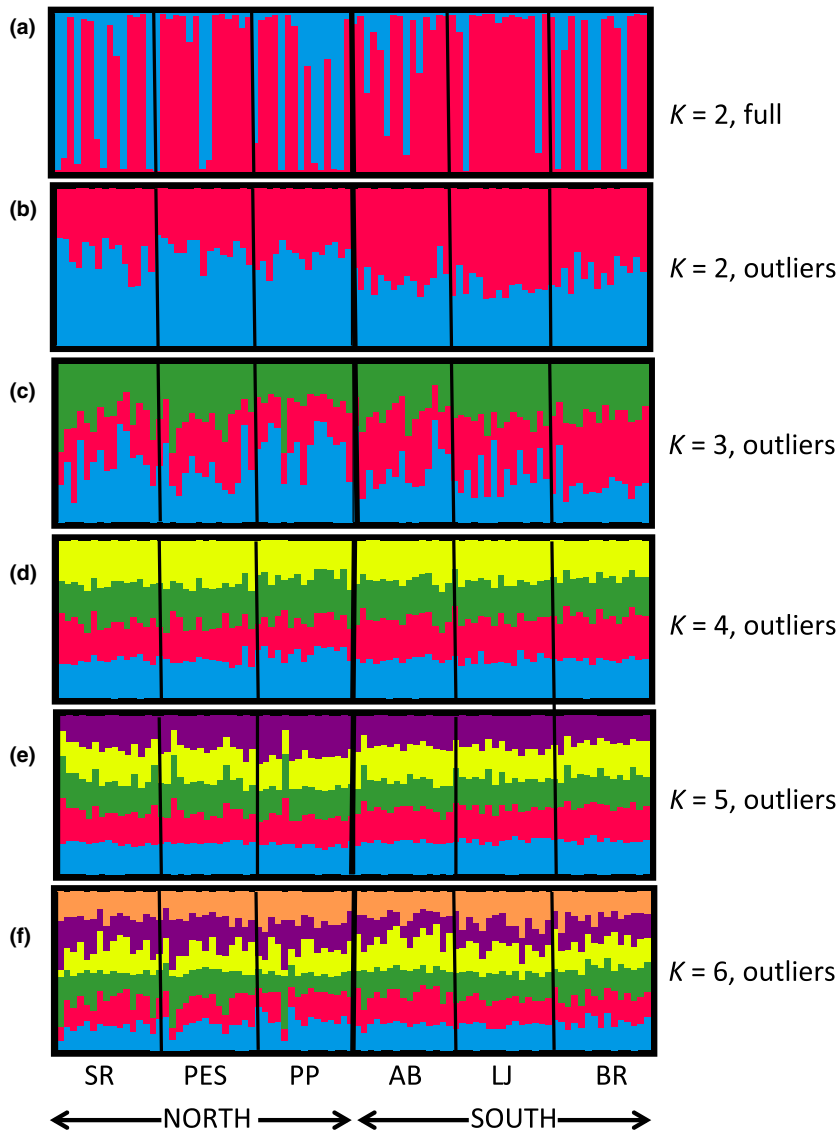
Table 2 AMOVA results showing the percentage of variation for nonoutlier and outlier loci among groups (northern or southern California), among populations within groups, and within populations

	Percentage of variation		
Data set	Among groups (northern vs. southern California)	Among populations within groups	Within populations
Nonoutlier loci	0.88	−3.47	102.59
Outlier loci	18.91	3.4	77.68

Table 3 Summary of annotated (out of 28 total) outlier loci identified using LOSITAN (Antao et al. 2008) at false discovery rate (FDR) = 0.1

Locus	Heterozygosity	F_{ST}	Abbreviated description	Annotation NCBI accession no.
50064_14	0.673	0.18	Dipeptidyl-peptidase 6 (DPP6)	NG_033878
52740_81	0.477	0.166	NADH dehydrogenase subunit 1 (ND1)	JF909879
87446_46	0.293	0.163	Cytoplasmic dynein 2 heavy chain 1	XP_012944204
1540_30	0.248	0.161	Small conductance calcium-activated potassium channel	LOC105344664
51870_70	0.294	0.139	MAP/microtubule affinity-regulating kinase 1 (MARK1)	XM_007435954
78140_52	0.297	0.131	RAB3 GTPase-activating protein subunit 2 (noncatalytic) (RAB3GAP2)	XM_006140235
38893_23	0.129	0.115	Poly[ADP-ribose] polymerase 14-like	XP_003390636

Fig 3 STRUCTURE clustering results for $K = 2$ for (a) the entire data set (1 SNP per locus, 1633 SNPs) and (b) through (f) for $K = 2$–6 for the outlier SNPs (1 SNP per locus, 28 SNPs).

for clustering was found. For the five replicate runs, ln Pr(X|K) was marginally higher for $K = 2$, although $K = 1$ and $K = 3$ were only slightly less likely (Table S1, Supporting information). However, the outlier SNPs showed evidence of regional clustering; 80% of northern individuals were assigned to one cluster, and 82% of southern individuals were assigned to the other cluster when $K = 2$ (Fig. 3b). The highest mean ln PR(X|K) corresponded to $K = 1$, but $K = 2$ was only slightly less likely. The ΔK statistic presented a sharp peak at $K = 5$ (Table S1, Supporting information), but the corresponding STRUCTURE plot showed no discernible population structure (Fig. 3e). The five replicate runs resulted in the same groupings. The discrepancy between the two different methods used to infer the most likely number of clusters present in the outlier data set indicates that although there is genetic differentiation between

northern and southern California at a small number of potentially adaptive loci, this regional structure is relatively subtle.

SNP trees

As with the STRUCTURE analysis, for this analysis, we further filtered the full SNP data set by retaining only a single SNP per RAD-tag and by removing loci that significantly deviated from HWE in both northern and southern populations, resulting in 1633 high-quality SNPs. As observed with the STRUCTURE analysis, no phylogeographic signal was identified when considering the whole data set, with individuals from different regions mixed throughout the tree (Fig. 4a). The outlier SNPs, however, produced a tree in which individuals generally grouped by geographic region. For instance,
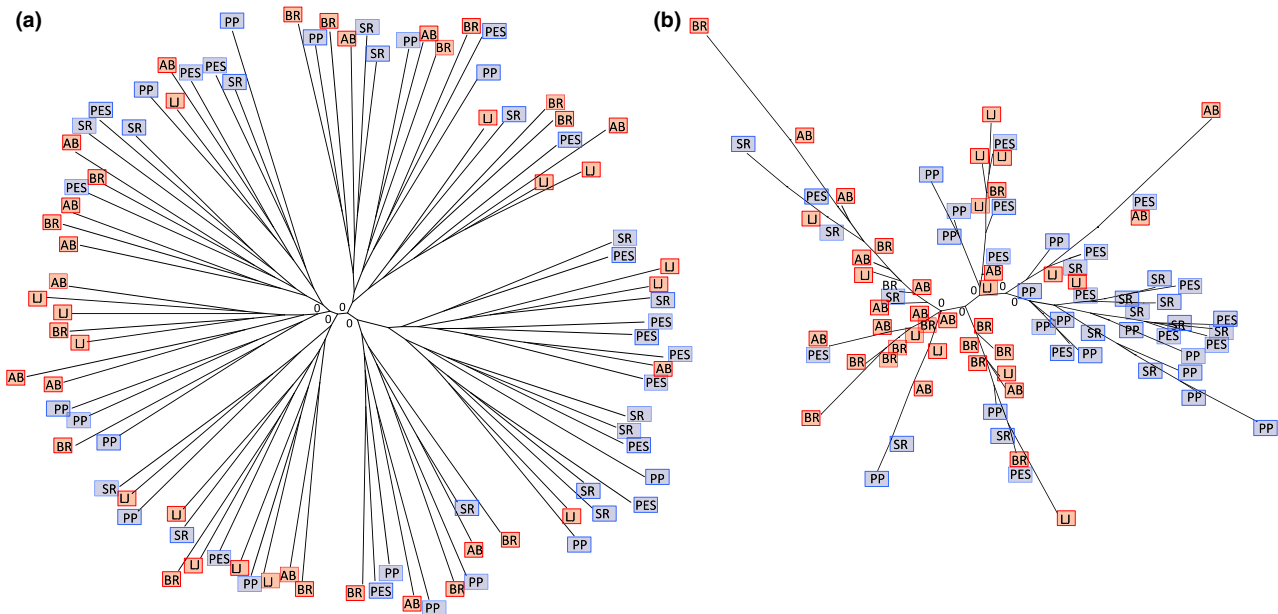
Fig 4 Maximum-likelihood SNP tree based on (a) all the data and (b) the outlier SNPs. Northern California individuals are highlighted with a small blue box, and southern individuals are highlighted with a small red box. Bootstrap values within groups not shown.

there was one large clade composed exclusively of 23 northern California individuals, and another clade composed of 20 southern California individuals (and 6 interspersed northern individuals) (Fig. 4b).

Mitochondrial sequencing

While no significant differences were found between geographic regions for cytochrome b (Tables S2 and S3, Supporting information), one statistically significant SNP was found in NADH dehydrogenase subunit 4 (ND4, Table S2, Supporting information), and the AMOVA between the two regions for ND4 was also significant (Table S3, Supporting information). Specifically, at base pair position 366 in ND4, 43 northern individuals were found to have the major allele (A), and six to have the minor allele (G); 42 and zero southern individuals were found to have the major and minor alleles, respectively ($P = 0.0288$, Fisher's exact test). In combination with the LOSITAN analysis detailed above that identified one outlier locus in NADH dehydrogenase subunit 1, these sequencing results further indicate that regional differentiation between northern and southern California is occurring in mitochondrial genes (in contrast to the previous COI results of Kelly & Palumbi 2010).

Dispersal simulation

Overall, the predicted dispersal trajectory for *C. funebralis* planktonic larvae suggest larvae only disperse between populations within a geographic region, at least on the specific dates in the two years investigated. Larval connectivity was observed among the three northern populations SR, PES and PP (Fig. S3a, Supporting information) and among the three southern populations AB, LJ and BR (Fig. S3b,c, Supporting information). However, this analysis revealed no evidence for direct gene flow between the three northern and the three southern populations. The farthest south any larvae from northern California reached was Big Sur in San Luis Obispo County in central California, and the farthest north any larvae from southern California reached was Ventura County in southern California.

Discussion

We utilized ddRAD to examine genomewide population structure and to gain insight into the dynamics influencing ecological divergence in northern and southern California *C. funebralis* populations. We identified 28 F_{ST} outlier loci as candidates for positive selection; these loci show evidence for regional differentiation between northern and southern California. In contrast, analysis based on all SNP loci showed no evidence for genetic structuring. Only seven of the outlier loci could be annotated (most ddRAD-seq SNPs are expected to be in noncoding regions of the genome); three of these seven are involved in cytoplasmic stress granule formation, which is a known response to environmental stress. Given previously documented

differences in in situ thermal stress exposure (L. U. Gleason & R. S. Burton, in review), thermal tolerance (Gleason & Burton 2013) and transcriptomic response to heat stress (Gleason & Burton 2015) in northern and southern California *C. funebralis* populations, we hypothesize that variation in thermal stress along the California coast could be the primary selective pressure driving differentiation between northern and southern populations.

Causes of background genetic homogeneity: dispersal of migrants?

We found no evidence for population structure when F_{ST}'s are averaged across all SNPs, but we did find evidence for population differentiation based on the outlier SNPs. Thus, it seems that the phylogeographic signal from the most highly differentiated SNPs is swamped out by the background genetic homogeneity at the majority of the SNPs. This background genetic similarity can be due to gene flow (dispersal of migrants), lack of drift in large population sizes or both. The particle drifter prediction analyses suggest larvae are dispersing within but not across regions, thus opening up the potential for divergence between northern and southern California *C. funebralis* populations, as observed in this study.

The findings of our drifter simulations agree with previous drifter studies conducted in southern California. For instance, Tegner and Butler released drifter tubes from various sites along the California coast to investigate dispersal of abalone larvae, and they found that 80–90% of the tubes released in Orange County and San Diego were recovered locally (Tegner & Butler 1985). In addition, Levin released drifter tubes from Mission Bay in San Diego to estimate larval dispersal potential for the polychaete *Pseudopolydora paucibranchiata*; the farthest north any tube travelled was to Palos Verdes, 125 km upcoast, during a winter storm in February (Levin 1983). In this drifter simulation study, we similarly observed that particles released from LJ only travel north to AB in February.

It is worth noting that the time period we used for the drifter simulations (13 days) is the longest available estimate of pelagic larval duration, but settlement can occur in as little as 5 days (Moran 1997). It has also been suggested that high temperatures, which are common in the southern portion of *C. funebralis*' range, can further reduce larval duration (Hahn 1989). Thus, these larval dispersal simulations, which already found no evidence for larval migration between northern and southern California, are potentially overestimating dispersal. Overall, *C. funebralis*' relatively short pelagic larval duration could impede long-distance dispersal and

instead promote larval settlement close to natal grounds, similar to what has been observed in abalone larvae (Prince *et al.* 1987).

This dispersal simulation examined populations in the two general regions of northern and southern California; however, there are *C. funebralis* populations in between the two regions, in central California near San Simeon for instance. Thus, hypothetically, stepping stone migration could be occurring over a few generations (Kimura & Weiss 1964), with no direct exchange of migrants occurring between northern and southern California, but instead with genes being exchanged between the two regions through intermediate sites in central California. However, the data from this study do not support this stepping stone hypothesis. Stepping stone migration leads to isolation by distance (IBD; Kimura & Weiss 1964), but the Mantel test examining whether a significant relationship exists between the genetic and geographic distances between northern and southern California populations did not provide any evidence for such patterns of IBD (Fig. S1, Supporting information).

Lastly, the observed lack of dispersal in the larval migration simulation suggests the genomic patterns of background neutral similarity between northern and southern California *C. funebralis* populations could be due to large population sizes, as previously hypothesized by Kelly & Palumbi (2010). Populations of *C. funebralis* are extremely abundant in low intertidal habitats; historical estimates of population density have reported up to 500 individuals/m^2 in central California sites such as Pacific Grove (Frank 1975). Alternatively, populations might not be at drift–migration equilibrium (Slatkin 1993), with recent population shifts due to glacial cycles obscuring genetic differentiation and causing patterns that mimic high contemporary larval exchange (Kelly & Palumbi 2010). Ultimately, more work is needed to definitively determine the cause of the background genetic homogeneity between northern and southern California populations, but the results of this drifter simulation suggest dispersal is not a large factor.

Targets of divergent selection: stress granules

Given that this study only surveyed approximately 0.01% of the *C. funebralis* genome, the fact that we identified loci potentially related to thermal stress response suggests this is a strong selective pressure in this intertidal species. Three of the seven annotated F_{ST} outliers putatively under selection are known or hypothesized to be involved in stress granule (SG) formation. These dynamic cytosolic aggregations comprised of RNAs and RNA-binding proteins form in response to cellular or environmental stress; under stressful conditions, the

translation of control or 'housekeeping' mRNAs are inhibited, and these untranslated mRNAs enter SGs (Anderson & Kedersha 2002; Kedersha & Anderson 2002). The granules function to sequester and preserve untranslated control mRNAs as part of a mechanism that adapts patterns of local RNA translation to facilitate the stress response, for instance by ensuring the translation of mRNAs such as chaperones and enzymes immediately needed for the stress response. In addition to being critical for mRNA regulation during stress, SGs also prevent apoptosis by sequestering pro-apoptosis factors and silencing their activities (Kim *et al.* 2005; Arimoto *et al.* 2008; Buchan & Parker 2009); growing evidence suggests that SG assembly and apoptotic cell death are mutually exclusive (Shih & Lee 2014). A recent study has reported that SGs also harbour antioxidant activity and may thus act as rapidly inducible antioxidant machinery that protects cells from reactive oxygen species-induced apoptosis (Takahashi *et al.* 2013). Lastly, SG formation also mediates recovery following exposure to environmental stress; once the stress passes, the granules dissociate and the preserved mRNAs that had been stored in the SG resume translation (Kedersha *et al.* 1999, 2002), as evidenced by the highly active protein synthesis in close vicinity to heat-shock granules during recovery (Nover *et al.* 1989). Stress granules are hypothesized to play a role in anoxia response in the marine snail *Littorina littorea* (Larade & Storey 2009), and heat shock has also been shown to cause the appearance of stress granules in trypanosomes, tomato cell cultures, plants, *Caenorhabditis elegans* nematodes, *Drosophila* fruit flies, *Saccharomyces cervisiae* yeast and other various mammalian cells (Nover *et al.* 1989; Kedersha & Anderson 2007; Kobayashi *et al.* 2007; Kramer *et al.* 2008; Farny *et al.* 2009; Grousl *et al.* 2009; Morton & Lamitina 2013).

While both oxygen levels (Bograd *et al.* 2008) and temperature (Helmuth *et al.* 2006) vary along the California coast, based on previous work that has shown *C. funebralis* populations differ in their thermal tolerance and in their transcriptome-wide response to heat stress, we hypothesize that the formation of stress granules is under divergent selection in northern and southern California populations in response to the differences in environmental thermal stress they encounter (Fig. S4, Supporting information; L. U. Gleason & R. S. Burton, in review; Gleason & Burton 2015). One of the identified F_{ST} outliers, microtubule affinity-regulating kinase, regulates microtubule dynamics. Microtubules play a role in stress granule formation; once stress granules begin to form as small foci, they are transported along microtubules by molecular motors enabling coalescence of stress granules into larger foci (Bartoli *et al.* 2011). Moreover, disruption of microtubules blocks the appearance of stress granules (Ivanov *et al.* 2003). Another outlier locus, cytoplasmic dynein 2 heavy chain, is also involved in motor activity. Heavy chain dyneins have been localized to stress granules, and stress granule formation is severely impaired by inhibiting dynein function (Kwon *et al.* 2007; Loschi *et al.* 2009. Lastly, poly [ADP-ribose] polymerase 14, the third annotated outlier, is known to cooperate with RNA-binding poly [ADP-ribose] polymerases following activation of the stress response to form stress granules (Vyas & Chang 2014). The strong genetic differentiation between northern and southern *C. funebralis* populations in these outlier genes associated with SGs suggests that these granules could play an important role in this species' thermal stress response and that the higher exposure to thermal stress in southern populations could be imposing differential selective pressure on the formation and function of SGs in geographically separated *C. funebralis* populations. Important next steps would include localizing SG markers such as PAB1 (Swisher & Parker 2010; Takahara & Maeda 2012) in northern and southern California populations before, during and following heat stress and furthermore disrupting microtubule and dynein function (hence blocking SG formation) in southern populations to see whether this reduces their thermal tolerance metrics such as survival and recovery to match those of northern populations (Gleason & Burton 2013).

Mitochondrial differentiation and thermal adaptation

In contrast to previous work that found no evidence for genetic structure in populations of *C. funebralis* sampled from Oregon to Santa Barbara, California, based on the mitochondrial gene COI (Kelly & Palumbi 2010), we found one SNP in the mitochondrial gene NADH dehydrogenase subunit 4 (ND4) with significantly different allele frequencies in northern vs. southern California *C. funebralis* individuals. In addition, outlier detection analysis also identified one region in NADH dehydrogenase subunit 1 (ND1) that is potentially under divergent selection. Based on the predicted open reading frame for this sequence (OrfPredictor; Min *et al.* 2005), this outlier SNP is expected to cause an amino acid change from alanine to glycine. As NADH dehydrogenase is the first enzyme in the electron transport chain (Complex 1), it is directly associated with energy metabolism (Porcelli *et al.* 2015). In nonmodel systems, including other marine molluscs such as the bivalve *Macoma balthica* (Pante *et al.* 2012) and the red abalone *Haliotis rufescens* (De Wit & Palumbi 2013), it has been found that energy metabolism genes are most frequently under selection due to differences in thermal environment (Porcelli *et al.* 2015). This class of genes has also been

observed to play a role in thermal adaptation in the coral *Acropora palmate* (Polato *et al.* 2011), and Lemay *et al.* (2013) discovered two distinct haplotypes within NADH dehydrogenase subunit 5 associated with elevation and climate differences in the American pika, *Ochotona princeps*. Similarly, in this study we have identified nucleotide differences in ND1 and ND4 between geographically separated populations of *C. funebralis* that differ in their thermal response. At least in mammals, ND4 is considered to be one of the actual protein pumps in Complex 1. Thus, variation in these subunits may affect the efficiency of the proton-pumping process, for example through chemical changes that may alter proton translocation (Da Fonseca *et al.* 2008). In addition, ND1 is known to have a critical role in the assembly of Complex 1 (Cardol *et al.* 2002). Ultimately, further work is needed to predict the likely effects of the identified SNPs in NADH dehydrogenase on the structure and function of this mitochondrial enzyme in *C. funebralis* populations and to determine whether significant genetic differences in NADH dehydrogenase between northern and southern California could be playing a role in the differential thermal adaptation of *C. funebralis*.

Patterns of genetic divergence in other marine species with planktonic larvae

While ddRAD sequencing provided no genomewide evidence for population structure in *C. funebralis*, outlier analysis identified candidates under positive selection that do show regional differentiation between northern and southern California geographic sites separated by ~850 km; several of these outlier loci are involved in responses to environmental stress such as heat. Whether this pattern of high background genetic similarity with few locally adapted outliers is also observed in other marine species with similar potential for dispersal (based on planktonic larval duration, or PLD) seems to depend on the geographic range of sampling. For instance, different species of abalone with a PLD of 4–15 days (Leighton 1972, 1974, 2000) have been examined across various geographic scales, and the results of these studies differed in whether or not they identified outlier loci. Green abalone sampled across ~200 km within southern California showed genomewide evidence for panmixia throughout this range, but unlike this study, outlier tests revealed no significant loci under selection (Gruenthal *et al.* 2014). In contrast, De Wit & Palumbi (2013) found that the majority of SNPs across the transcriptome show no population structure in red abalone individuals sampled from Monterey to Oregon (~1000 km range, similar to this study on *C. funebralis*), and these authors also identified outliers

involved in biomineralization, hypoxia and disease stress, and metabolism. Lastly, in the sea anemone *Nematostella vectensis* (PLD ~7 days; Reitzel *et al.* 2007) sampled from Nova Scotia to South Carolina (~2500 km range), strong population structure was evident when examining the whole genome (Reitzel *et al.* 2013). Although this latter finding differs from the results of the current study, this may be due at least in part to the fact that individuals were sampled across a geographic range roughly three times larger than the distance between the northern and southern California sampling sites used for *C. funebralis.* Nevertheless, despite these different genomewide results, Reitzel *et al.* did find a similar number of markers (37) potentially under selection. Moreover, as in this study, one of these outliers is involved in response to heat stress; this intergenic candidate SNP was located nearest to heat-shock factor 1, the main transcription factor that regulates the downstream expression of heat-shock response genes (Reitzel *et al.* 2013). In sum, previous studies on species with similar PLD compared to *C. funebralis* conducted at comparable geographic scales of sampling have obtained similar patterns of genomewide and outlier-specific divergence, while studies with different sampling scales have given conflicting results. Ultimately, more data regarding genomewide differentiation in other systems are necessary before we can conclude whether the patterns observed in *C. funebralis* are typical for marine species with planktonic larvae.

Conclusions

This study demonstrates that several (potentially) ecologically relevant loci in northern and southern *C. funebralis* populations show strong genetic differentiation and therefore evidence for local adaptation, despite the presence of genomewide neutral background similarity. In agreement with previous work that indicated these geographically separated *C. funebralis* populations have different thermal tolerances and distinct transcriptomic responses to heat, the functions of the F_{ST} outlier loci potentially under selection indicate that variation in thermal stress between northern and southern California could be driving the genetic differences observed between *C. funebralis* populations. Overall, our results provide further insight into the factors that encourage and deter local adaptation in a marine species with planktonic larvae.

Acknowledgements

References

Abbott DP, Haderlie EC (1980) Prosobranchia: marine snails. In: *Intertidal Invertebrates of California* (eds Morris RH, Abbott DP, Haderlie EC), pp. 230–307. Stanford University Press, Stanford, California.

Altschul SF, Madden TL, Schaffer AA *et al.* (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research*, **25**, 3389–3402.

Anderson P, Kedersha N (2002) Stressful initiations. *Journal of Cell Science*, **115**, 3227–3234.

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.

Arimoto K, Fukuda H, Imajoh-Ohmi S, Saito H, Takekawa M (2008) Formation of stress granules inhibits apoptosis by suppressing stress-responsive MAPK pathways. *Nature Cell Biology*, **10**, 1324–1332.

Bartoli KM, Bishop DL, Saunders WS (2011) The role of molecular microtubule motors and the microtubule cytoskeleton in stress granule dynamics. *International Journal of Cell Biology*, **2011**, 939848.

Bierne N, Welch J, Loire E, Bonhomme F, David P (2011) The coupling hypothesis: why genome scans may fail to map local adaptation genes. *Molecular Ecology*, **20**, 2044–2072.

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**, L12607.

Bohonak AJ (2002) IBD (isolation by distance): a program for analyses of isolation by distance. *Journal of Heredity*, **93**, 153–154.

Buchan JR, Parker R (2009) Eukaryotic stress granules: the ins and outs of translation. *Molecular Cell*, **36**, 932–941.

Burton RS (1983) Protein polymorphisms and genetic differentiation of marine invertebrate populations. *Marine Biology Letters*, **4**, 193–206.

Byers B (1983) Enzyme polymorphism associated with habitat choice in the intertidal snail *Tegula funebralis*. *Behavior Genetics*, **13**, 65–75.

Cardol P, Matagne RF, Remacle C (2002) Impact of mutations affecting ND mitochondria-encoded subunits on the activity and assembly of complex 1 in *Chlamydomonas*: implication for the structural organization of the enzyme. *Journal of Molecular Biology*, **319**, 1211–1221.

Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping loci de novo from short-read sequences. *G3-Genes Genomes Genetics*, **1**, 171–182.

Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Molecular Ecology*, **22**, 3124–3140.

Chu ND, Kaluziak ST, Trussell GC, Vollmer SV (2014) Phylogenomic analyses reveal latitudinal population structure and polymorphisms in heat stress genes in the North Atlantic snail *Nucella lapillus*. *Molecular Ecology*, **23**, 1863–1873.

Cooper EE (2010) *Population biology and reproductive ecology of Chlorostoma (Tegula) funebralis, an intertidal gastropod*. Doctoral Dissertation, University of Oregon, Eugene, Oregon.

Da Fonseca RR, Johnson WE, O'Brien SJ, Ramos MJ, Antunes A (2008) The adaptive evolution of the mammalian mitochondrial genome. *BMC Genomics*, **9**, 119.

DaCosta JM, Sorenson MD (2014) Amplification biases and consistent recovery of loci in a double-digest RAD-seq protocol. *PLoS ONE*, **9**, e106713.

Darriba D, Taboada GL, Doallo R, Posada D (2012) jModelTest 2: more models, new heuristics and parallel computing. *Nature Methods*, **9**, 772.

Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nature Reviews Genetics*, **12**, 499–510.

De Wit P, Palumbi SR (2013) Transcriptome-wide polymorphisms of red abalone (*Haliotis rufescens*) reveal patterns of gene flow and local adaptation. *Molecular Ecology*, **22**, 2884–2897.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genetics Resources*, **4**, 359–361.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology*, **14**, 2611–2620.

Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular Ecology Resources*, **10**, 564–567.

Farny NG, Kedersha NL, Silver PA (2009) Metazoan stress granule assembly is mediated by P-eIF2 alpha-dependent and -independent mechanisms. *RNA-A Publication of the RNA Society*, **15**, 1814–1821.

Frank PW (1975) Latitudinal variation in life-history features of black turban snail *Tegula funebralis* (Prosobranchia: Trochidae). *Marine Biology*, **31**, 181–192.

Gleason LU, Burton RS (2013) Phenotypic evidence for local adaptation to heat stress in the marine snail *Chlorostoma* (formerly *Tegula*) *funebralis*. *Journal of Experimental Marine Biology and Ecology*, **448**, 360–366.

Gleason LU, Burton RS (2015) RNA-seq reveals regional differences in transcriptome response to heat stress in the marine snail *Chlorostoma funebralis*. *Molecular Ecology*, **24**, 610–627.

Grousl T, Ivanov P, Frydlova I *et al.* (2009) Robust heat shock induces eIF2 alpha-phosphorylation-independent assembly of stress granules containing eIF3 and 40S ribosomal subunits in budding yeast, *Saccharomyces cerevisiae*. *Journal of Cell Science*, **122**, 2078–2088.

Gruenthal KM, Witting DA, Ford T *et al.* (2014) Development and application of genomic tools to the restoration of green abalone in southern California. *Conservation Genetics*, **15**, 109–121.

Guindon S, Gascuel O (2003) A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. *Systematic Biology*, **52**, 696–704.

Guo SW, Thompson EA (1992) Performing the exact test of Hardy-Weinberg proportion for multiple alleles. *Biometrics*, **48**, 361–372.

Hahn KO (1989) *Handbook of Culture of Abalone and other Marine Gastropods*. CRC Press, Florida, USA.

Helmuth B, Harley CDG, Halpin PM, O'Donnell MJ, Hofmann GE, Blanchette CA (2002) Climate change and latitudinal patterns of intertidal thermal stress. *Science*, **298**, 1015–1017.

Helmuth B, Broitman BR, Blanchette CA *et al.* (2006) Mosaic patterns of thermal stress in the rocky intertidal zone: implications for climate change. *Ecological Monographs*, **76**, 461–479.

Hess JE, Campbell NR, Close DA, Docker MF, Narum SR (2013) Population genomics of Pacific lamprey: adaptive variation in a highly dispersive species. *Molecular Ecology*, **22**, 2898–2916.

Hinegardner R (1974) Cellular DNA content of the Mollusca. *Comparative Biochemistry and Physiology*, **47A**, 447–460.

Ivanov PA, Chudinova EM, Nadezhdina ES (2003) Disruption of microtubules inhibits cytoplasmic ribonucleoprotein stress granule formation. *Experimental Cell Research*, **290**, 227–233.

Jakobsson M, Rosenberg NA (2007) CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, **23**, 1801–1806.

Karl SA, Avise JC (1992) Balancing selection at allozyme loci in oysters: implications from nuclear RFLPs. *Science*, **256**, 100–102.

Kedersha N, Anderson P (2002) Stress granules: sites of mRNA triage that regulate mRNA stability and translatability. *Biochemical Society Transactions*, **30**, 963–969.

Kedersha N, Anderson P (2007) Mammalian stress granules and processing bodies. *Methods in Enzymology*, **431**, 61–81.

Kedersha NL, Gupta M, Li W, Miller I, Anderson P (1999) RNA-binding proteins TIA-1 and TIAR link the phosphorylation of eIF-2 alpha to the assembly of mammalian stress granules. *Journal of Cell Biology*, **147**, 1431–1441.

Kedersha N, Chen S, Gilks N *et al.* (2002) Evidence that ternary complex (eIF2-GTP-tRNA(i)(Met))-deficient preinitiation complexes are core constituents of mammalian stress granules. *Molecular Biology of the Cell*, **13**, 195–210.

Kelly RP, Palumbi SR (2010) Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PLoS ONE*, **5**, e8594.

Kelly RP, Oliver TA, Sivasundar A, Palumbi SR (2010) A method for detecting population genetic structure in diverse, high gene-flow species. *Journal of Heredity*, **101**, 423–436.

Kim WJ, Back SH, Kim V, Ryu I, Jang SK (2005) Sequestration of TRAF2 into stress granules interrupts tumor necrosis factor signaling under stress conditions. *Molecular and Cellular Biology*, **25**, 2450–2462.

Kimura M, Weiss GH (1964) The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, **49**, 561–576.

Kobayashi K, Otegui MS, Krishnakumar S, Mindrinos M, Zambryski P (2007) Increased size exclusion limit 2 encodes a putative DEVH box RNA helicase involved in plasmodesmata function during Arabidopsis embryogenesis. *Plant Cell*, **19**, 1885–1897.

Koehn RK, Newell RIE, Immermann FW (1980) Maintenance of an aminopeptidase allele frequency cline by natural selection. *Proceedings of the National Academy of Sciences of the United States of America*, **77**, 5385–5389.

Kramer S, Queiroz R, Ellis L *et al.* (2008) Heat shock causes a decrease in polysomes and the appearance of stress granules in trypanosomes independently of eIF2 alpha phosphorylation at Thr169. *Journal of Cell Science*, **121**, 3002–3014.

Kwon S, Zhang Y, Matthias P (2007) The deacetylase HDAC6 is a novel critical component of stress granules involved in the stress response. *Genes and Development*, **21**, 3381–3394.

Kyle CJ, Boulding EG (2000) Comparative population genetic structure of marine gastropods (*Littorina* spp.) with and without pelagic larval dispersal. *Marine Biology*, **137**, 835–845.

Larade K, Storey KB (2009) Living without oxygen: anoxia-responsive gene expression and regulation. *Current Genomics*, **10**, 76–85.

Leighton DL (1972) Laboratory observations on the early growth of the abalone *Haliotis sorenseni*, and the effect of temperature on larval development and settling success. *US National Marine Fisheries Services Fisheries Bulletin*, **70**, 373–381.

Leighton DL (1974) The influence of temperature on larval and juvenile growth in three species of southern California abalones. *US National Marine Fisheries Services Fisheries Bulletin*, **72**, 1137–1145.

Leighton DL (2000) *The Biology and Culture of the California Abalones*. Dorrance Publishing Co. Inc, Philadelphia.

Lemay MA, Russello MA (2015) Genetic evidence for ecological divergence in kokanee salmon. *Molecular Ecology*, **24**, 798–811.

Lemay MA, Henry P, Lamb CT, Robson KM, Russello MA (2013) Novel genomic resources for a climate change sensitive mammal: characterization of the American pika transcriptome. *BMC Genomics*, **14**, 311.

Lenormand T (2002) Gene flow and the limits to natural selection. *Trends in Ecology and Evolution*, **17**, 183–189.

Levin LA (1983) Drift tube studies of bay-ocean water exchange and implications for larval dispersal. *Estuaries*, **6**, 364–371.

Levin LA (2006) Recent progress in understanding larval dispersal: new directions and digressions. *Integrative and Comparative Biology*, **46**, 282–297.

Lewis PO (2001) A likelihood approach to estimating phylogeny from discrete morphological character data. *Systematic Biology*, **50**, 913–925.

Lewontin RC (1974) Analysis of variance and analysis of causes. *American Journal of Human Genetics*, **26**, 400–411.

Loschi M, Leishman CC, Berardone N, Boccaccio GL (2009) Dynein and kinesin regulate stress-granule and P-body dynamics. *Journal of Cell Science*, **122**, 3973–3982.

Lynch M, Milligan BG (1994) Analysis of population genetic-structure with RAPD markers. *Molecular Ecology*, **3**, 91–99.

Marshall DJ, Monro K, Bode M, Keough MJ, Swearer S (2010) Phenotype-environment mismatches reduce connectivity in the sea. *Ecology Letters*, **13**, 128–140.

Mayr E (1963) *Animal Species and Evolution*. 797 pp. Harvard University Press, Cambridge, Massachusetts.

Min XJ, Butler G, Storms R, Tsang A (2005) OrfPredictor: predicting protein-coding regions in EST-derived sequences. *Nucleic Acids Research*, **1**, W677–W680.

Moran AL (1997) Spawning and larval development of the black turban snail *Tegula funebralis* (Prosobranchia: Trochidae). *Marine Biology*, **128**, 107–114.

Morton EA, Lamitina T (2013) *Caenorhabditis elegans* HSF-1 is an essential nuclear protein that forms stress granule-like structures following heat shock. *Aging Cell*, **12**, 112–120.

Nover L, Scharf KD, Neumann D (1989) Cytoplasmic heat-shock granules are formed from precursor particles and are associated with a specific set of messenger-RNAs. *Molecular and Cellular Biology*, **9**, 1298–1308.

Pante E, Rohfritsch A, Becquet V, Belkhir K, Bierne N, Garcia P (2012) SNP detection from *de novo* transcriptome sequencing in the bivalve *Macoma balthica*: marker development for evolutionary studies. *PLoS ONE*, **7**, e52302.

Pell TJ, Yellon DM, Goodwin RW, Baxter GF (1997) Myocardial ischemic tolerance following heat stress is abolished by ATP-sensitive potassium channel blockade. *Cardiovascular Drugs and Therapy*, **11**, 679–686.

Pérez-Figueroa A, Garcia-Pereira MJ, Saura M, Rolan-Alvarez E, Caballero A (2010) Comparing three different methods to detect selective loci using dominant markers. *Journal of Evolutionary Biology*, **23**, 2267–2276.

Pespeni MH, Garfield DA, Manier MK, Palumbi SR (2012) Genome-wide polymorphisms show unexpected targets of natural selection. *Proceedings of the Royal Society B-Biological Sciences*, **279**, 1412–1420.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE (2012) Double digest RADseq: an inexpensive method for *de novo* SNP discovery and genotyping in model and non-model species. *PLoS ONE*, **7**, e37135.

Polato NR, Vera JC, Baums IB (2011) Gene discovery in the threatened elkhorn coral: 454 sequencing of the *Acropora palmate* transcriptome. *PLoS ONE*, **6**, e28634.

Porcelli D, Butlin RK, Gaston KJ, Joly D, Snook RR (2015) The environmental genomics of metazoan thermal adaptation. *Heredity*, **114**, 502–514.

Prada C, McIlroy SE, Beltrán DM *et al.* (2014) Cryptic diversity hides host and habitat specialization in a gorgonian-algal symbiosis. *Molecular Ecology*, **23**, 3330–3340.

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.

Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics*, **155**, 945–959.

Puebla O, Bermingham E, McMillan WO (2014) Genomic atolls of differentiation in coral reef fishes (*Hypoplectrus* spp., Serranidae). *Molecular Ecology*, **23**, 5291–5303.

Ravinet M, Westram A, Johannesson K, Butlin R, André C, Panova M (2016) Shared and nonshared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. *Molecular Ecology*, **25**, 287–305.

Raymond M, Rousset F (1995) An exact test for population differentiation. *Evolution*, **49**, 1280–1283.

Reitzel AM, Burton PM, Krone C, Finnerty JR (2007) Comparison of developmental trajectories in the starlet sea anemone *Nematostella vectensis*: embryogenesis, regeneration, and two forms of asexual fission. *Invertebrate Biology*, **126**, 99–112.

Reitzel AM, Herrera S, Layden MJ, Martindale MQ, Shank TM (2013) Going where traditional markers have not gone before: utility of and promise for RAD sequencing in marine invertebrate phylogeography and population genomics. *Molecular Ecology*, **22**, 2953–2970.

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

Russello MA, Kirk SL, Frazer KK, Askey PJ (2012) Detection of outlier loci and their utility for fisheries management. *Evolutionary Applications*, **5**, 39–52.

Saad AH, Hahn GM (1992) Activation of potassium channels: relationship to the heat-shock response. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 9396–9399.

Sagarin RD, Gaines SD (2002) Geographical abundance distributions of coastal invertebrates: using one-dimensional ranges to test biogeographic hypotheses. *Journal of Biogeography*, **29**, 985–997.

Schmidt PS, Rand DM (1999) Intertidal microhabitat and selection at Mpi: interlocus contrasts in the northern acorn barnacle, *Semibalanus balanoides*. *Evolution*, **53**, 135–146.

Schmidt PS, Rand DM (2001) Adaptive maintenance of genetic polymorphism in an intertidal barnacle: habitat- and life-stage-specific survivorship of Mpi genotypes. *Evolution*, **55**, 1336–1344.

Shih JW, Lee YHW (2014) Human DExD/H RNA helicases: emerging roles in stress survival regulation. *Clinica Chimica Acta*, **436**, 45–58.

Slatkin M (1985) Gene flow in natural populations. *Annual Review of Ecology and Systematics*, **16**, 393–430.

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

Soh H, Goldstein SAN (2008) I-SA channel complexes include four subunits each of DPP6 and Kv4.2. *Journal of Biological Chemistry*, **283**, 15072–15077.

Stamatakis A (2014) RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics*, **30**, 1312–1313.

Stamatakis A, Hoover P, Rougemont J (2008) A rapid bootstrap algorithm for the RAxML web servers. *Systematic Biology*, **57**, 758–771.

Swisher KD, Parker R (2010) Localization to, and effects of Pbp1, Pbp4, Lsm12, Dhh1, and Pab1on stress granules in *Saccharomyces cerevisiae*. *PLoS ONE*, **5**, e10006.

Takahara T, Maeda T (2012) Transient sequestration of TORC1 into stress granules during heat stress. *Molecular Cell*, **47**, 242–252.

Takahashi M, Higuchi M, Matsuki H *et al.* (2013) Stress granules inhibit apoptosis by reducing reactive oxygen species production. *Molecular and Cellular Biology*, **33**, 815–829.

Tegner MJ, Butler RA (1985) Drift-tube study of the dispersal potential of green abalone (*Haliotis fulgens*) larvae in the Southern California Bight: implications for recovery of depleted populations. *Marine Ecology Progress Series*, **26**, 73–84.

Vilas A, Pérez-Figueroa A, Caballero A (2012) A simulation study on the performance of differentiation-based methods to detect selected loci using linked neutral markers. *Journal of Evolutionary Biology*, **25**, 1364–1376.

Vyas S, Chang P (2014) New PARP targets for cancer therapy. *Nature Reviews Cancer*, **14**, 502–509.

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population structure. *Evolution*, **38**, 1358–1370.

Westram AM, Galindo J, Rosenblad MA, Grahame JW, Panova M, Butlin RK (2014) Do the same genes underlie parallel phenotypic divergence in different *Littorina saxatilis* populations? *Molecular Ecology*, **23**, 4603–4616.

L.U.G. and R.S.B. designed the research, L.U.G. performed the experiments and analysed the data, and L.U.G. and R.S.B. wrote the manuscript.

Data accessibility

Multilocus SNP genotypes for each individual at all 2371 loci (in GENEPOP format) have been archived in the Dryad digital data repository (doi: 10.5061/dryad.g9k62). Raw Illumina sequence reads are available in the NCBI Sequence Read Archive (BioProject ID: 286949).

Supporting information

Additional supporting information may be found in the online version of this article.

Fig S1 Isolation by distance results, plotting the log of geographic distance by the log of genetic distance.

Fig S2 Results from the LOSITAN F_{ST} outlier analysis for 1825 SNPs.

Fig S3 Drop-a-Drifter simulated larval trajectories released from (a) Slide Ranch on July 30, 2014, (b) La Jolla on August 17, 2014, and (c) La Jolla on February 13, 2015.

Fig S4 'Acute' (99th percentile of temperatures) and 'chronic' high-temperature (peak average daily maximum) exposures calculated from in situ temperatures recorded at each of the six geographic sites using biomimetic loggers (robosnails; L. U. Gleason & R. S. Burton, in review).

Table S1 Results of the STRUCTURE analyses. Inferred number of clusters based on the Evanno method are highlighted in bold.

Table S2 Biallelic SNPs identified in cytochrome b and NADH dehydrogenase subunit 4.

Table S3 Summary of AMOVA results examining differences between northern and southern California for cytochrome b and NADH dehydrogenase subunit 4.