

Genetic adaptation despite high gene flow in a range-expanding population

Andy Lee¹  | Benjamin N. Daniels²  | William Hemstrom¹  | Cataixa López³  |
Yuki Kagaya¹ | Daisuke Kihara^{1,4} | Jean M. Davidson² | Robert J. Toonen³  |
Crow White² | Mark R. Christie^{1,5} 

¹Department of Biological Sciences,
Purdue University, West Lafayette,
Indiana, USA

²Department of Biological Sciences,
California Polytechnic State University,
San Luis Obispo, California, USA

³Hawai'i Institute of Marine Biology,
University of Hawai'i at Mānoa, Kāne'ohe,
Hawaii, USA

⁴Department of Computer Science,
Purdue University, West Lafayette,
Indiana, USA

⁵Department of Forestry and Natural
Resources, Purdue University, West
Lafayette, Indiana, USA

Correspondence

Andy Lee and Mark R. Christie,
Department of Biological Sciences,
Purdue University, West Lafayette, IN
47907-2054, USA.
Email: lee3617@purdue.edu and
christ99@purdue.edu

Funding information

Division of Environmental Biology, Grant/
Award Number: 1856710

Handling Editor: Diana J. Rennison

Abstract

Signals of natural selection can be quickly eroded in high gene flow systems, curtailing efforts to understand how and when genetic adaptation occurs in the ocean. This long-standing, unresolved topic in ecology and evolution has renewed importance because changing environmental conditions are driving range expansions that may necessitate rapid evolutionary responses. One example occurs in Kellet's whelk (*Kelletia kelletii*), a common subtidal gastropod with an ~40- to 60-day pelagic larval duration that expanded their biogeographic range northwards in the 1970s by over 300 km. To test for genetic adaptation, we performed a series of experimental crosses with Kellet's whelk adults collected from their historical (HxH) and recently expanded range (ExE), and conducted RNA-Seq on offspring that we reared in a common garden environment. We identified 2770 differentially expressed genes (DEGs) between 54 offspring samples with either only historical range (HxH offspring) or expanded range (ExE offspring) ancestry. Using SNPs called directly from the DEGs, we assigned samples of known origin back to their range of origin with unprecedented accuracy for a marine species (92.6% and 94.5% for HxH and ExE offspring, respectively). The SNP with the highest predictive importance occurred on triosephosphate isomerase (*TPI*), an essential metabolic enzyme involved in cold stress response. *TPI* was significantly upregulated and contained a non-synonymous mutation in the expanded range. Our findings pave the way for accurately identifying patterns of dispersal, gene flow and population connectivity in the ocean by demonstrating that experimental transcriptomics can reveal mechanisms for how marine organisms respond to changing environmental conditions.

KEYWORDS

climate change, dispersal, experimental transcriptomics, gene flow, range expansion, rapid genetic adaptation

1 | INTRODUCTION

Understanding how and when genetic adaptation occurs within local populations has been a long-standing question in marine ecology (Hughes et al., 2003; Sanford & Kelly, 2011). Populations in marine environments are often connected by dispersal that occurs during a pelagic larval stage, a life-history trait shared among an estimated 80% of marine invertebrate species and over 95% of all marine fishes (Nelson et al., 2006; Thorson, 1950). Pelagic larvae can travel tens to hundreds of kilometres away from their spawning location (Christie et al., 2010; Cowen et al., 2007; D'Aloia et al., 2022), often facilitated by ocean currents (Gaylord & Gaines, 2000; Schraidt et al., 2023; White et al., 2010), which can create high gene flow and thus little to no population divergence (Liu et al., 2019; Maes et al., 2021; Quintero-Galvis et al., 2023; Singh et al., 2019). Given that signals of divergent or directional selection can be quickly eroded, or at the very least attenuated, in high gene-flow systems (Cornwell, 2020), it has long been speculated that genetic adaptation within local populations may be a rare phenomenon in marine species (Conover et al., 2006; Sanford & Kelly, 2011; Tigano & Friesen, 2016). This long-standing question has renewed importance in light of rapidly changing ocean environments (Lotterhos et al., 2021) and recent studies illustrating that many marine species are expanding their ranges polewards (Pinsky et al., 2013, 2019; Poloczanska et al., 2013).

Despite theoretical predictions, there is a growing body of evidence that genetic adaptation within local populations can occur in marine species characterized by high rates of gene flow. Environmental conditions such as salinity, temperature gradients and dissolved oxygen levels can impose strong selection in pan-mictic species (Gagnaire et al., 2012, 2015; Takeuchi et al., 2020; Vandamme et al., 2014). Furthermore, signals of a response to this strong selection have been identified at putatively adaptive loci (Benestan et al., 2016; Conover et al., 2006; Wilder et al., 2020) even when moderate rates of gene flow can erase signals of population differentiation at neutral loci (Cornwell, 2020; Whitlock & McCauley, 1999). Several marine taxa show identifiable patterns of local adaptation and rapid evolution in response to increased temperatures (DuBois et al., 2022; O'Donnell et al., 2018; Sanford & Kelly, 2011), presumably driven by a response to selection at loci underlying key quantitative traits (Mackay et al., 2009).

Traditional genome-wide scans to identify loci responding to selection may not always be sufficient to identify genetic adaptation in species with high gene flow, particularly if selection acts on the pre-existing standing genetic variation. While local adaptation on chromosomal rearrangements and large inversions (Barth et al., 2017; Meyer et al., 2024; Schaal et al., 2022; Wilder et al., 2020) or hitchhiking caused by hard selective sweeps (Bierne, 2010; Haenel et al., 2019) can generate clear peaks of differentiation and signals of adaptation in high gene flow systems, selection on standing genetic variation will produce much smaller peaks of differentiation that can be difficult to identify (Hermisson & Pennings, 2005). Nevertheless, selection at putatively adaptive

loci can manifest as downstream changes in mRNA expression profiles or sequences, such that RNA sequencing (RNA-Seq) offers an opportunity to identify adaptive genetic variation (McGuigan et al., 2011; Paaby & Rockman, 2014; Thorstensen et al., 2021; Yin et al., 2021). Furthermore, coupling RNA-Seq with a common garden breeding design that minimizes the effects of environmental conditions on gene expression can improve our ability to identify genes affected by selection on regulatory elements (Christie et al., 2016; Gil & Ulitsky, 2020; Harder et al., 2020; Signor & Nuzhdin, 2018). We applied this approach to Kellet's whelks (*Kelletia kelletii*), a marine species with an extensive larval dispersal period that is exhibiting a northward range expansion, potentially in response to rapidly changing climatic conditions (Herrlinger, 1981; Zacherl et al., 2003).

Kellet's whelk is a kelp forest gastropod subtidally distributed throughout the eastern Pacific Ocean. This species underwent a recent range expansion from its historical range in Southern California, United States, and Baja California, Mexico, into an ecologically distinct, much colder bioregion north of a natural biogeographic break at Point Conception (Figure 1) (Herrlinger, 1981; Zacherl et al., 2003). Despite being colder, the expanded range has comparatively little seasonal thermal variation relative to the historical range which may result in less thermal stress for whelk populations in the expanded range (Figure 1c). Adult whelks were first identified in the expanded range in 1980 (Herrlinger, 1981) and likely settled in the new range in the 1970s (White et al., 2022). Similar to many marine species, Kellet's whelk shows very little population structure across its biogeographic range in California, United States, and Baja California, Mexico (global $F_{ST}=0.0009$; Selkoe et al., 2010), due to a highly dispersive, 40–60 day pelagic larval stage (Romero et al., 2012). Kellet's whelk can be bred in captivity (Vendetti, 2020), facilitating experimental transcriptomics (Figure 1).

Here, we used a common garden experimental design to allow for the comparison of individuals from different ancestries while holding environmental factors constant. Adult wild-origin whelks were collected from the northern (expanded, E) and the southern (historical, H) regions of the species' range and housed in a common garden where the temperature of seawater was identical for all individuals (Figure 1). We performed a series of crosses between adults from the historical range (HxH) and the expanded range (ExE) (Figure 1c) and performed RNA-Seq on their offspring. Because adult whelks were maintained in aquaria with identical conditions, differences in gene expression in their offspring are attributable to genetic effects, heritable epigenetic effects or maternal/paternal effects (Christie et al., 2016), which can uncover loci responding to the targets of directional selection. We found several such loci which underlie genes attributed to important functional differences, and we illustrate that these genes reveal population structuring not identifiable with microsatellite loci and can support high-accuracy population assignment. By coupling transcriptomics and common garden experiments, we present a powerful approach to resolve population structure, identify genetic adaptation and assess population connectivity in high gene flow systems.

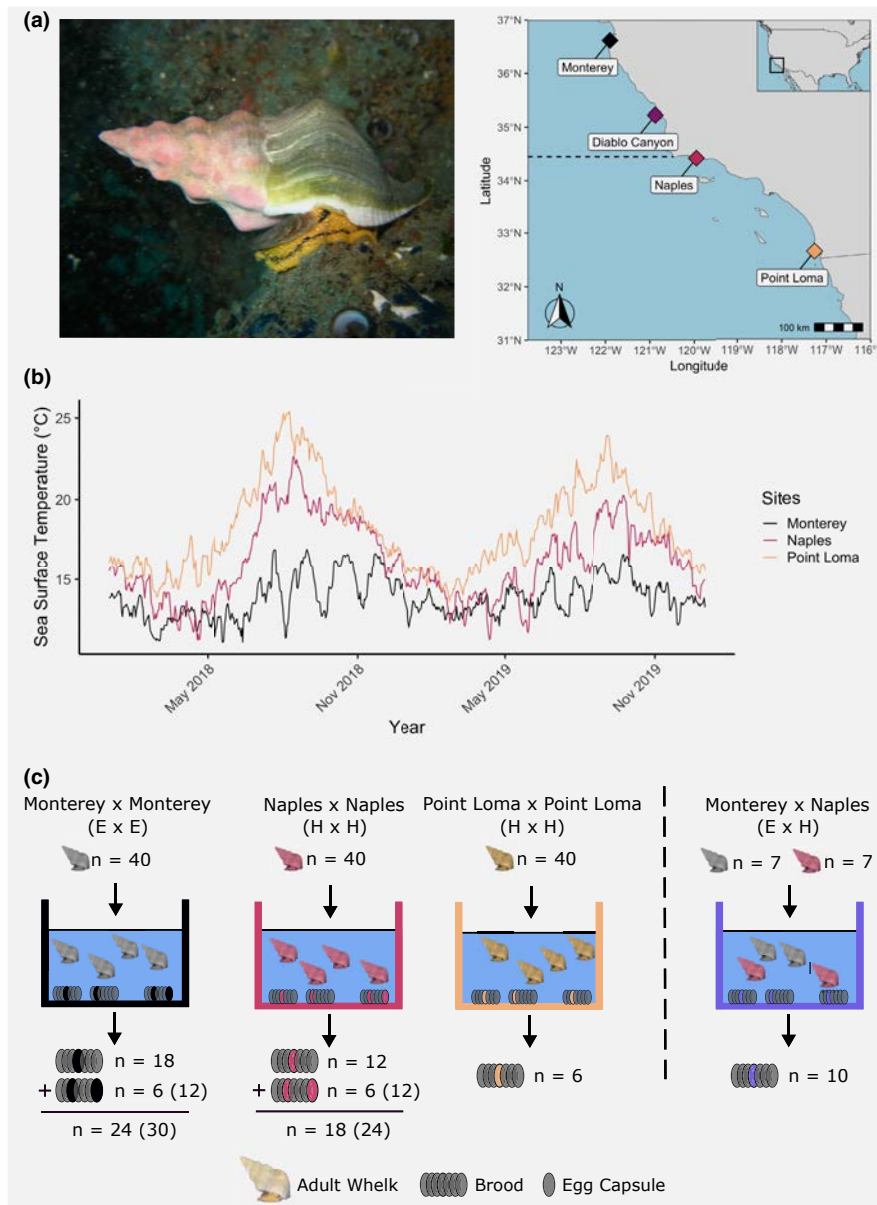


FIGURE 1 (a) Adult *Kelletia kelletii* on rocky substrate (left). Collection sites from Kellet's whelks' historical range south of Pt. Conception (dashed line) and expanded range north of Pt. Conception (right). (b) Sea surface temperature (SST) across the 2 years during which adult whelks were collected from each sampling location. Populations south of Pt. Conception experience warmer water temperatures. (c) Illustration of the experimental design. Adult whelks were collected from the wild and maintained in aquaria until crosses were performed between historic (HxH) and expanded (ExE) range adults. The number of adult whelks allowed to freely cross within each tank is indicated with an n above each tank (e.g. $n=40$ for Monterey). The numbers depicted below each tank represents the number of egg capsules sequenced from each tank. For most broods, we collected one egg capsule for RNA sequencing (illustrated as a single coloured capsule within a brood; e.g. $n=18$ for Monterey); for six additional broods from Monterey x Monterey and Naples x Naples crosses, we collected two egg capsules per brood (illustrated as two coloured capsules within a brood). Here, the first number represents the number of brood collected and the second number within parentheses indicate egg capsules sequenced (e.g. $n=6(12)$ for Monterey). For Monterey and Naples, the total number of broods is presented first, followed by the total number of egg capsules sequenced in parentheses (e.g. $n=24(30)$ for Monterey).

2 | MATERIALS AND METHODS

2.1 | Field collections

Using SCUBA, we collected adult Kellet's whelks by hand from sub-tidal (approximately 15m depth) locations across California

(Figure 1a, CDFW SCP #8018 to C.W.). Adult whelks from three locations: Monterey (MON, north of Point Conception; 36.6181670N, 121.89W), Naples (NAP, just south of Point Conception; 34.4219670N, 119.952283W) and Point Loma (POL, Southern California; 32.665333N, 117.261517W) were used in our main experimental crosses (described below). Additionally, adult whelks

from Diablo Canyon (DIC in the expanded range; 35.22445N, 120.877483W) were collected and used to validate the results of the main experimental crosses.

2.2 | Main experimental crosses

We acclimated 40 adults each from MON, NAP and POL in common garden aquaria using the flow-through filtered seawater system at the California Polytechnic State University Research Pier in Avila, California (35.169817N, 120.740838W). Additionally, an experimental reciprocal cross between MON and NAP was created with seven individuals from each site in an attempt to test for maternal/paternal effects. In total, these four experimental crosses used 134 adult whelks (Table S2). The wild-collected adults were acclimated for 7–10 months, fed identical diets of mixed seafood (scallops, shrimp, blue mussels and squid) and then crossed within 92L common aquaria (Figure 1c). Because all offspring were reared in a common garden and thus environmental conditions were identical, differences in their gene expression can largely be attributed to genetic and/or trans-generational epigenetic differences between populations (Christie et al., 2016; Fitz-James & Cavalli, 2022; Roberge et al., 2008). A small subset of these differentially expressed genes may also be driven by maternal or paternal effects (Mousseau & Fox, 1998; Wolf & Wade, 2009).

Kellett's whelks have discrete sexes, but males and females have similar morphology (Gosnell et al. 2024; although females tend to be larger) and cannot be sexed conclusively without sacrificing and dissection (Rosenthal, 1970). As a result, we allowed whelks to freely cross within each tank. Kellett's whelk reaches reproductive maturity (~60mm shell length; Rosenthal, 1970) after 5–6 years (White et al., 2022). Female Kellett's whelks have been observed to mate with multiple males, sometimes in aggregations. Prior observations revealed that females release eggs during or immediately after copulation (Rosenthal, 1970). F1 offspring in the form of egg capsules deposited by females on aquaria walls were collected in spring and summer of 2020 from the common garden aquaria when embryos inside the capsules reached the veliger stage (19–44 days after the capsule was laid). Whelk egg capsules are laid in broods of 3–82 capsules (Vendetti, 2020); we primarily collected only one capsule per brood to maximize the number of parental crosses represented in the study. Because not much is known about Kellett's whelk's reproductive biology, we also collected two capsules from the same brood for each of six broods from both the Monterey (MONxMON) and Naples (NAPxNAP) crosses to examine genetic relationships between capsules from the same broods (Figure 1c, Supporting Information Method and Result, Figure S1). All egg capsules were immediately flash frozen in liquid nitrogen and stored at –80°C until mRNA extraction. We performed RNA-Seq on 70 egg capsules (pooled samples of 500–1000 veliger larvae) produced from our common garden experiment: 30 from MONxMON crosses, 24 from NAPxNAP crosses, 6 from POLxPOL crosses and 10 from reciprocal crosses

between MON and NAP (Figure 1c; Table S2). We collected six capsules from the POLxPOL crosses because only six broods of egg capsules were laid. We aligned an average of 21,948,826 reads per sample back to the Kellett's whelk reference transcriptome (Daniels, Andrasz, et al., 2023) (Table S3), performed tests of differential expression and called genetic variants (SNPs; see details below).

2.3 | Samples for validation

2.3.1 | Additional experimental crosses

In order to validate the SNPs called from the main experimental crosses, we collected and sequenced egg capsules from a subsequent cross experiment conducted in 2021. In addition to MON, NAP and POL populations from the main cross, this experiment included adults collected from Diablo Canyon (DIC) (Figure 1a). Because these egg capsules were raised in different tanks and thermal conditions, this set of crosses eliminated any possible tank bias in the SNP data sets called on our main experimental crosses. In brief, these experiments involved a second set of HxH and ExE crosses using the same adults from common garden crosses described above (with the new addition of Diablo Canyon adults), but where the crosses were performed and the offspring were collected 1 year later (2021), then raised in different tanks under various thermal regimes (Figure S3, Table S2, Supporting Information Methods).

2.3.2 | Wild-caught adults

Because the above additional experimental crosses were conducted on the same putative set of adults (though the actual crosses were likely different), we cannot rule out that the population structure patterns we observe are confounded by non-independent family structure. To address this bias, we validated the predictive power of the SNP loci called from the main crosses by assigning a completely independent sample of wild-caught adult whelks collected in the field. These wild whelks were collected from MON ($n=70$) and NAP ($n=99$) in 2015 and 2016. DNA was extracted from whelk tissue using a Salting-out protocol (Daniels, Nurge, et al., 2023) and sequenced with Genotyping-in-Thousands by sequencing (GTSeq) panel of 155 SNPs (Campbell et al., 2015) identified from highly divergent SNPs found on differentially expressed genes (Supporting Information Methods).

2.4 | RNA-Seq and sequence processing

For each egg capsule, we separated 500–1000 veliger larvae (F1 offspring) from the maternal egg capsule tissue for use in subsequent RNA extraction and sequencing. Thus, each 'sample' in our

analyses represents a pool of larvae from a single egg capsule. Details on RNA extractions, sequencing and data quality control of the resulting pools of individuals are available as [Supporting Information Methods](#).

2.5 | Differential gene expression

RNA reads were aligned to a de novo reference transcriptome assembled with samples from MON, NAP and POL (Daniels, Andrasz, et al., 2023). We aligned raw reads using bowtie2/2.4.2 with the very sensitive flag and default parameters (Langmead & Salzberg, 2012). We then sorted the resulting alignments using samtools/1.9 (Li et al., 2009), merged the assemblies using stringtie2/2.1.1 (Kovaka et al., 2019) and created a count matrix using the featureCounts tool of subread/2.0.1 (Liao et al., 2014), which exudes chimeric fragments (-C flag) and requires that both reads in a pair to be successfully mapped (-B flag). Additionally, featureCounts does not count reads with multiple alignments or read pairs that overlap multiple features. An average of 21,948,826 reads per sample were aligned back to the Kellet's whelk reference transcriptome (Table S3). We used the R package DESeq2/1.34 (Love et al., 2014) to identify differentially expressed genes (DEGs) between the offspring of our Monterey and Naples crosses using a minimum significance threshold of 0.05 after false discovery rate correction via the Benjamini–Hochberg method (Benjamini & Hochberg, 1995). DESeq2 performs independent filtering based on the mean of normalized counts for each gene and internally corrects for library size (Love et al., 2014).

2.6 | SNP variant discovery

SNP variants on reads aligned to the de novo transcriptome (Daniels, Andrasz, et al., 2023) were identified using ANGSD (Korneliussen et al., 2014), a maximum likelihood approach to identify major and minor alleles (Skotte et al., 2012), the GATK genotyping algorithm (Van der Auwera & O'Connor, 2020) and a uniform prior. Bases were only retained if they had a Phred-scaled quality score of at least 20. Genotypes with a posterior probability of less than 95% and/or a depth less than 20 in an individual or less than 100 across all individuals were removed, and only loci genotyped in every individual were retained. Lastly, we removed any loci with a minor allele frequency of less than 0.05 in every population using the R package snpR/1.2.79 (Hemstrom & Jones, 2022). Because selection at putatively adaptive loci may manifest as downstream changes in expression, we also identified SNP variants that occurred on differentially expressed contigs between MON and NAP. Although genetic changes to expression are due to changes at nearby cis or distant trans-regulatory elements (Gil & Ulitsky, 2020; Signor & Nuzhdin, 2018), we identified SNPs within the DEGs to search for functional (structural) differences in the genes themselves. We called 122,599 SNPs throughout the transcriptome, of which 4168 were located on DEG contigs (see [Supporting Information Methods](#) and Table S1 for filtering details).

After filtering (Hemstrom et al., 2024), 94,654 SNP loci were identified throughout the transcriptome (hereafter 'all-SNPs') with 3118 loci located on DEG contigs (hereafter 'DEG-SNPs').

2.7 | Population structure

To investigate population structure, we conducted principal components analysis (PCA) on offsprings of the main experimental crosses via snpR/1.2.79 (Hemstrom & Jones, 2022) using the smartPCA method (Patterson et al., 2006; Price et al., 2006). For comparison, we also used STRUCTURE software (Pritchard et al., 2000) to group individuals into k clusters for each k value from 1 to 4 using 20,000 burn-in and 100,000 estimation iterations. To determine the degree of variance across STRUCTURE runs, we ran the programme 10 times for each value of k . We used the Δk method to determine the 'optimal' number of k values (Evanno et al., 2005), although, given that this method does not reliably account for higher order population structuring (Janes et al., 2017), we also visualized our results across all k values by condensing results across runs using the 'greedy' option in CLUMPP (Jakobsson & Rosenberg, 2007). We used both snpR (Hemstrom & Jones, 2022) and pophelper (Francis, 2017) to run, organize and plot these results. Lastly, we constructed neighbour-joining trees for the SNP data sets using fasttreeR/1.2.0 (Gkanogiannis, 2023). The package fasttreeR/1.2.0 first calculates a zero-diagonal symmetric distance matrix between the samples of a VCF file and then performs hierarchical clustering without a resampling algorithm (Gkanogiannis, 2023). These trees were created to illustrate genetic differentiation rather than actual phylogenetic relationships. We visualized the results using ggtree/3.6.2 (Yu et al., 2017). For each of the above, we used three data sets: (1) all-SNPs loci, (2) DEG-SNPs loci and (3) a nine microsatellite data set developed by White and Toonen (2008) and analysed previously by Selkoe et al. (2010) and White et al. (2010). Since analysis of microsatellite loci is not currently supported by snpR or fasttreeR, we used adegenet/2.1.10 (Jombart, 2008) for the PCA and ape/5.7 (Paradis et al., 2004) to construct neighbour-joining trees for the microsatellite data. Lastly, we filtered our SNP data sets down to one SNP per assembled transcript/contig to remove any potential bias due to non-independence among loci caused by linkage, but saw no substantial changes (Figure S2).

2.8 | Population assignment, validation and cross-evaluation

We conducted population assignments using the R package Rubias/0.3.3, which implements genetic stock identification (GSI) using a conditional Bayesian model (Moran & Anderson, 2019). Using both reference data sets (all-SNPs and DEG-SNPs), we assigned 'unknown individuals,' which really represented individuals of known ancestry from a subsequent crossing experiment (see *Additional Experimental Crosses* above). In addition to testing for

population assignment accuracy back to collection sites (MON, NAP and POL), we assigned these unknown individuals back to their respective ranges (Historical or Expanded). For comparison, we ran a random forest (RF) model using 1,000,000 trees and *mtry* equal to the number of SNP loci (the number of variables to randomly sample as candidates at each split) using the R package Ranger (Wright & Ziegler, 2017). The RF model was trained using both all-SNPs and DEG-SNPs as the reference data set to make predictions about samples from the additional experimental crosses.

We implemented three approaches to validate the results of the population assignment tests: (1) an independent GT-Seq panel with wild-caught samples, (2) cross-validation and (3) an independent reciprocal cross between MON and NAP. First, we validated the predictive power of the SNP loci called from DEGs by genotyping 169 wild-caught adult whelks with GTSeq (see *Samples for Validation* above) back to MON and NAP. Here, we tested the efficacy of these SNPs found on DEGs in assigning a completely independent set of wild-caught individuals back to their site of origin using the leave-one-out approach in both Rubias and Ranger (SI). Second, to rule out the possibility that prediction success was due to tank effects, we also cross-validated our population assignment results by instead training our models using genotype calls of individuals from the additional experimental crosses as reference and treated samples from NAP, MON and POL (main experimental crosses) as 'unknowns'. By using the offspring from the *additional experimental crosses* as reference, we avoid overfitting the model in our predictions (i.e. avoid covariance across samples due to environmental effects rather than population differences). Lastly, to further validate the assignment power of our reference SNPs, we also assigned samples from our reciprocal cross (Monterey×Naples in Figure 1c, Supporting Information Methods).

2.9 | Candidate loci driving population differences and adaptation

To investigate the roles of DEGs in describing population structure, we compared the $\log_2$ fold change of each DEG to the random forest importance value of each DEG-SNP. To identify genes associated with the historic and expanded ranges, we ran an additional RF model with samples from MON (ExE offspring) and NAP (HxH offspring) using 1,000,000 trees and a *mtry* equal to the number of SNPs, as before, in the R package Ranger (Wright & Ziegler, 2017) and calculated the importance of individual SNPs to model predictions ('importance values') using the permutation method (Altmann et al., 2010). We then compared importance values for each SNP to their pairwise F_{ST} values (MON vs. NAP) to determine if SNPs with high explanatory values had congruently high F_{ST} values.

We used BLASTx (Altschul et al., 1990; Camacho et al., 2009) to identify the gene containing the SNP with the highest importance from the RF model (*triosephosphate isomerase*, *TPI*). We used ORFfinder (Rombel et al., 2002) to identify the open reading frames (ORFs) within the assembled transcript. To investigate the effects of

missense mutations on the protein's function, we used an approach that combines structure prediction and mutation mapping. First, the wild-type protein sequence was used as input for AlphaFold v2.3.2 (Jumper et al., 2021) using default settings for monomer and all templates. The predicted wild-type structure was used as input for the Missense3D web server (Ittisoponpisan et al., 2019). Since the *TPI* homologues in the PDB are registered as homodimers, we superimposed the predicted structure onto a *TPI* homologue (PDB ID: 2J27) and mapped the mutation positions.

To further understand the functions of differentially expressed genes, we conducted Gene Ontology (GO) enrichment and weighted gene co-expression network analysis (WGCNA). To do so, we first annotated the entire transcriptome and DEG contigs using eggNOG mapper (Cantalapiedra et al., 2021). We then conducted GO enrichment analysis for the DEGs using the full transcriptome as a reference in the R package topGO/2.50 (Alexa & Rahnenfuhrer, 2023) using the 'weight01' elimination algorithm to limit the proportion of false positives (Alexa et al., 2006). WGCNA analysis between Naples and Monterey samples was conducted using the WGCNA package (1.72–1) (Langfelder & Horvath, 2008) with a minimum module size of 30 genes and merged correlated modules ($r^2 > 0.9$). Using TopGo, we identified overrepresented GO terms within each significant module using the 'weight01' algorithm. Lastly, we constructed a network of the top significantly overrepresented GO terms in each significant WGCNA module ($p < 0.001$) using *metacoder* 0.3.6 (Foster et al., 2017), with internal nodes pruned from each network for the ease of visualization.

3 | RESULTS

3.1 | Differential expression

Differential expression analyses identified 2770 significant differentially expressed genes (DEGs with FDR-corrected p -values < 0.05) between MON (ExE offspring) and NAP (HxH offspring). Of these DEGs, 1773 were upregulated (64%) in the expanded range offspring (MON) and 997 (36%) were downregulated. The mean $\log_2$ fold change (absolute value) for all DEGs was 0.68, where 351 and 58 DEGs had a $\log_2$ fold change (absolute value) of greater than 1 and 2, respectively.

3.2 | Population structure in a high gene flow system

Principal component analysis (PCA) showed no population structure among these three populations when using nine microsatellite loci (Figure 2a), replicating the results of previous population genetic studies on this species (Selkoe et al., 2010; White et al., 2010). In contrast, PCA using the all-SNPs data set identified discrete population structure among the three adult collection sites: MON versus NAP ($F_{ST} = 0.02$), MON versus POL ($F_{ST} = 0.04$) and NAP versus POL

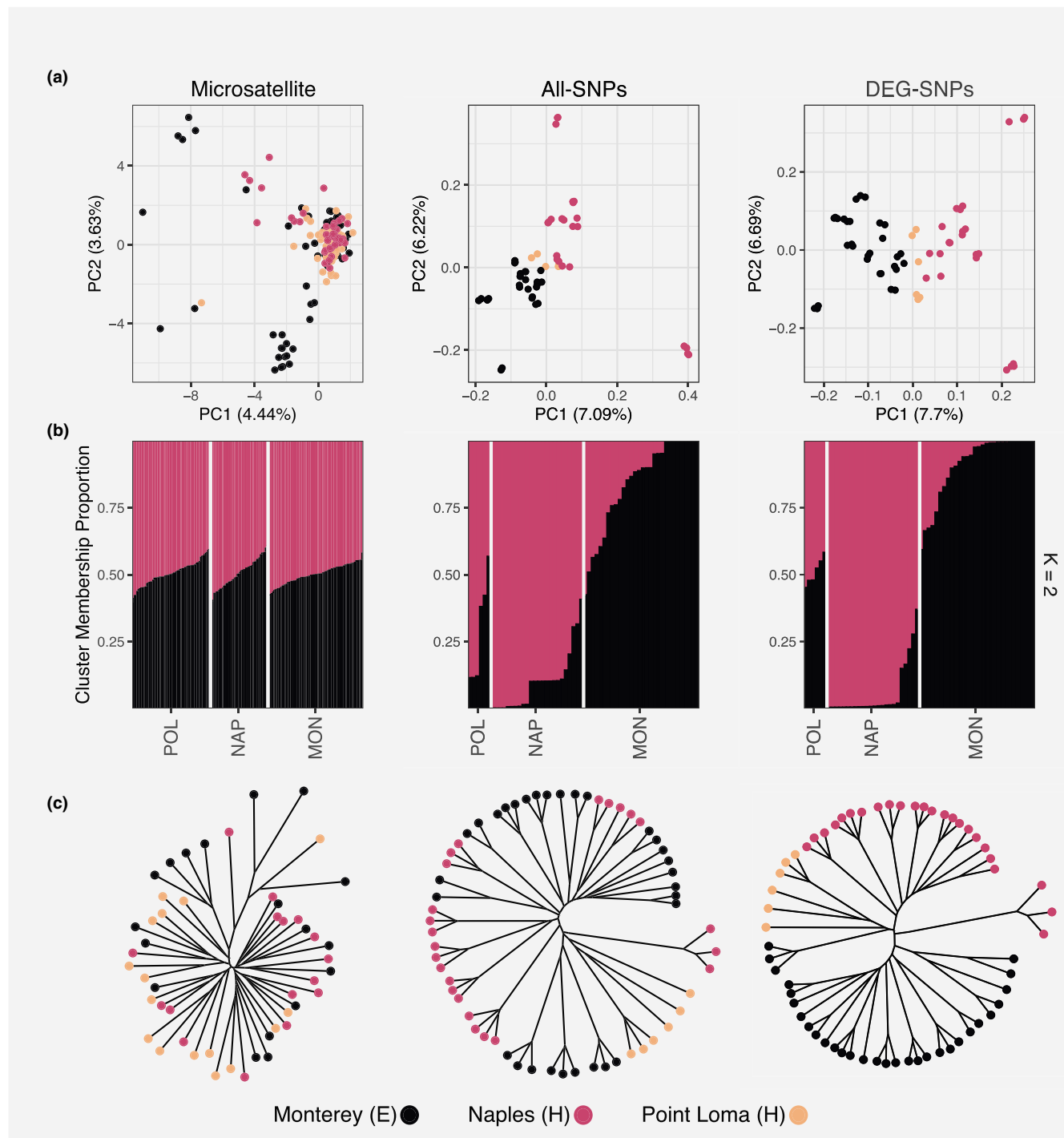


FIGURE 2 (a) Principal component analysis, (b) STRUCTURE analysis and (c) Neighbour-joining analysis conducted using (left to right) microsatellite data, all SNPs identified across the entire transcriptome ('All-SNPs'; $n = 94,654$), and SNPs identified on contigs that were differentially expressed between ExE and HxH offspring reared in a common environment ('DEG-SNPs'; $n = 4168$). The microsatellite data set does not resolve population structure while the SNPs derived from the common garden experiments, particularly the DEG-SNPs, perform substantially better at resolving population structure between Monterey (expanded range) and Naples (historical range).

($F_{ST} = 0.04$) (Figure 2a). Additionally, DEG-SNPs also resulted in discrete population structure among the three adult collection sites: MON versus NAP ($F_{ST} = 0.02$), MON versus POL ($F_{ST} = 0.03$) and NAP versus POL ($F_{ST} = 0.04$) (Figure 2a). Although F_{ST} values of 0.02–0.04 could be considered low (especially for terrestrial species), these

values are an order of magnitude higher than previously identified for Kellet's whelks (Selkoe et al., 2010; White et al., 2010), comparatively high for marine species (Tepolt et al., 2009; Benestan et al., 2015; Rey et al., 2020), and are high enough to allow for accurate delineation of population structure (Christie et al., 2017; Latch

et al., 2006). Population structure between MON and NAP was also observed with a PCA of the differentially expressed genes based solely on their gene counts (Figure S4). For the all-SNPs and the DEG-SNPs data sets, STRUCTURE was able to identify population structure at $K=2$, which received the most support (Figure 2b; but not at $K=3$, Figure S5). Neighbour-joining trees revealed no population structure for the microsatellite data set (Figure S6 for the full microsatellite phylogeny) but revealed several discrete lineages for the all-SNPs and DEG-SNPs data sets (Figure 2c).

3.3 | Population assignment

Using both all-SNPs and DEG-SNPs data sets as reference, we assigned offspring from the *Additional Experimental Crosses* to MON, NAP and POL (Figure 3a, Figure S3). These additional HxH and ExE offspring were sequenced to validate the population structure identified in our main experimental cross. Population assignments were slightly more accurate overall when using DEG-SNPs as the

reference than using all-SNPs (see [Supporting Information Methods](#) for evaluation of the reference data sets). When assigning samples to our collection sites using all-SNPs, 81.5% and 92.6% of samples were correctly assigned back to MON and NAP, respectively; the accuracy increased to 92.6% and 94.5% when we used only DEG-SNPs (Figure 3a, Table S4). For POL, which had a much smaller number of samples in the reference data sets ($n=6$), correct assignment was much lower (15% and 27% for all-SNPs and DEG-SNPs, respectively, Figure 3a). Although assignment success for POL samples was lower than MON and NAP, this result is not surprising considering its smaller sample size in the reference. It is possible that an approach that subsamples larger populations to equalize each reference population could increase assignment success even for moderately small reference sizes (Puechmaile 2016; Wang 2017), but is unlikely to work for sample size as small as six brood capsules. Because we did not have any reference samples from DIC, we had 0% assignment rate back to that specific site (Figure 3a), but we did have 100% correct assignment to their broad geographic regions. Specifically, we further examined the accuracy of assignment tests when assigning

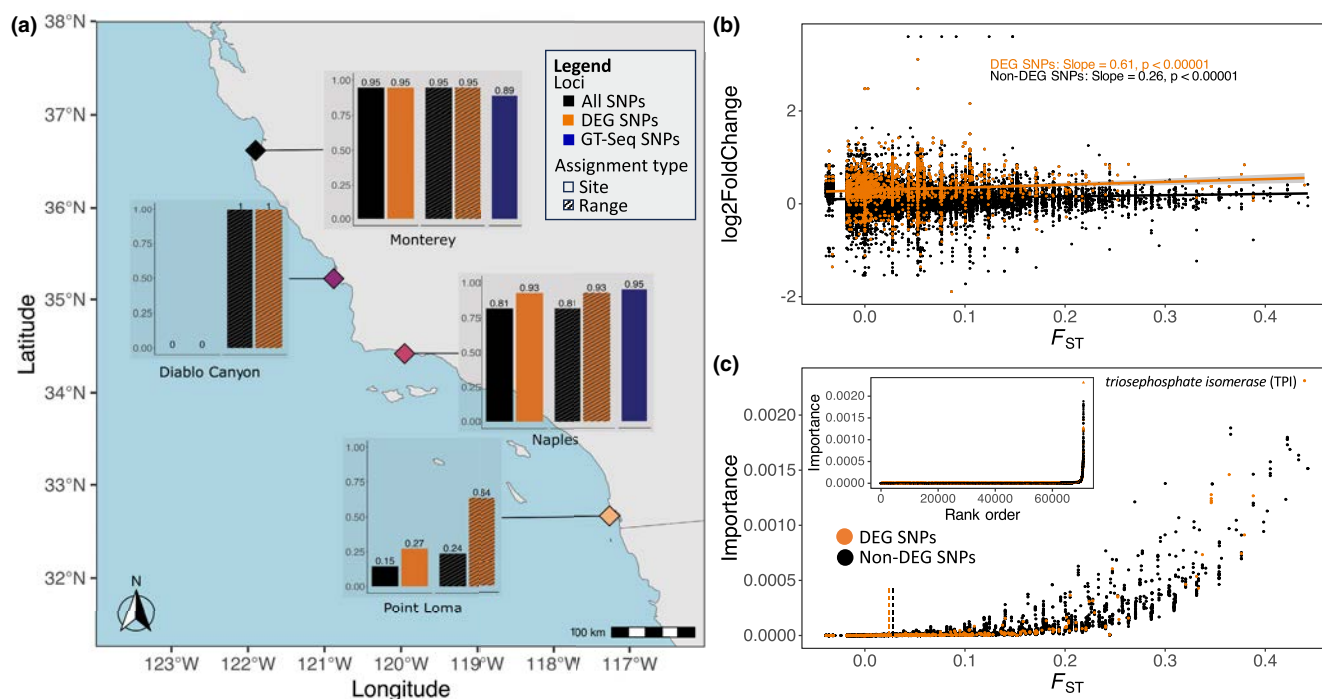


FIGURE 3 (a) Population assignment results for offspring reared in a common laboratory environment. The correct assignment percentages are shown with bar plots. Here, we assigned offspring from *Additional Experimental Crosses* back to the collection site of their parents (Monterey, Naples and Point Loma; shown in solid colours), or back to region (historical or expanded range; shown in dashed line pattern). Note that because Diablo Canyon was not included in this study, the model could not successfully assign individuals back to that site; however, we were able to assign individuals from DIC back to the region with 100% accuracy. We also illustrate the differences in assignment accuracy between all-SNPs (black) and DEG-SNPs (orange). Using 155 GT-Seq SNPs called directly from our differentially expressed genes, we assigned a completely independent set of 169 wild-caught adults collected from Monterey and Naples back to their collection sites with high accuracy (88.57% and 94.95%, respectively; blue bars). (b) Comparison between the F_{ST} values of every SNP (Monterey vs. Naples) versus the log-fold change in expression for every contig that the corresponding SNP was located. DEG-SNPs are marked in orange, and non-DEG-SNPs are black. Linear model trend lines are shown for each data set and the same trend persists after correcting for unequal variance (see [Results](#)). (c) The same pairwise F_{ST} values were next plotted against their importance values calculated from a random forest predictive model. DEG-SNPs are marked in orange, and non-DEG-SNPs in black. The inset shows RF importance values in rank order. Dashed vertical lines showing median F_{ST} values for each data set. The SNP with the highest importance value occurs on triosephosphate isomerase (TPI, top right of the figure).

samples to either the species' expanded (MON and DIC) or historical (NAP and POL) range. Whereas assignment accuracy increased from 0% to 100% for samples from DIC using both all-SNPs and DEG-SNPs (despite there being no samples from DIC in our reference data set), accuracy stayed the same for MON (94.5% for all-SNPs and DEG-SNPs) and NAP (81.5% for all-SNPs and 92.6% for DEG-SNPs), but for POL accuracy increased considerably (from 15% to 24% for all-SNPs, and from 27% to 64% for DEG-SNPs, summarized in Table S4). We obtained similar assignment results using two methods with very different underlying algorithms: conditional Bayesian approaches for genetic stock identification (Figure 3a) and random forest models that employ a machine-learning algorithm (Table S4). Overall, DEG-SNPs performed better than using all-SNPs, and the population assignment tests were effective in predicting the individual's range of origin, even when we had a small sample size (POL) or no samples (DIC) in our reference (Figure 3a).

3.4 | Assignment test validation and cross-evaluation

Because the subsequent experimental crosses used in our population assignment tests are closely related to the reference main crosses, we validated the population assignment tests using three different approaches: (1) an independent GT-Seq panel, (2) cross-validation and (3) an independent reciprocal cross (see Methods for details). First, using a Genotyping-in-Thousands (GT-Seq) panel of 155 SNPs called directly from our differentially expressed genes, we assigned a completely independent set of 169 wild-caught adults collected in 2015 and 2016 from MON and NAP back to their collection sites with high accuracy (88.57% and 94.95%, respectively; Figure 3a, Table S5). Here, we assigned wild-caught individuals to their correct populations based on loci identified in the lab: SNPs from all expressed genes (all-SNPs) and SNPs found within differentially expressed genes (DEG-SNPs). Second, we flipped the reference and assignment samples, using samples from the *Additional Experimental Crosses* as reference and samples from our main experimental crosses as the 'unknowns' to cross-validate the population assignment tests. By using a separate data set as reference, we avoid overfitting the model in our predictions (i.e. prevent covariance across samples due to environmental effects rather than population differences). Again, we had high assignment rates of samples back to MON, NAP and POL using both conditional Bayesian approaches and random forest models (96.7%–100%, Table S6, Supporting Information Results). Third, we further tested the accuracy of the assignment tests using offspring of the reciprocal crosses between adults from Naples and Monterey (MONxNAP in Figure 1c). Because the MONxNAP offspring were created with different sets of parents from the main crosses and raised in different tanks (Figure 1c, Table S2), we avoided assignment bias due to family structure and tank effects. On PC1, samples from the MONxNAP reciprocal cross separate from both MONxMON and NAPxNAP clusters (Figure S7). These samples are found between MON and NAP (or on the edge

of MON or NAP) on the PCA, consistent with the fact that there is some gene flow between the two populations (Figure 2). We assigned eight samples to Monterey and two samples were assigned to Naples, consistent with the visualization of these samples in PCA space (Supporting Information Methods, Figure S7). The high percentage of correct assignment of independent adult samples collected in different years, different reference data sets and independent crosses validates that loci identified via our common garden crosses can facilitate population assignment with high accuracy.

3.5 | Candidate loci driving population differences and adaptation

To investigate the roles of DEGs in describing population structure, we compared the $\log_2$ fold change of each DEG with the pairwise F_{ST} values (between MON and NAP) of each DEG-SNP. There was a slightly positive, significant relationship between F_{ST} and expression (measured as $\log_2$ fold change) (Figure 3b). This relationship was more positive for DEG-SNPs than all-SNPs (Chow test, $p < 0.0001$). Because our data sets have highly skewed right tails, we also corrected for unequal variance using a generalized least squares (GLS) model with a power of variance covariate. The same trend persisted after correcting for unequal variance for both non-DEGs SNPs (slope = 0.2636, p -value < 0.0000) and DEG SNPs (slope = 0.6467, p -value < 0.0000). Pairwise F_{ST} values between MON (ExE offspring) versus NAP (HxH offspring) were positively correlated with prediction importance in the RF models (Figure 3c), while F_{ST} values did not substantially differ between DEG SNPs (median $F_{ST} = 0.028$) and non-DEG SNPs (median $F_{ST} = 0.024$; Figure 3c, Figure S8). The genetic differences between locations are thus explained by relatively few SNPs with large effect sizes and high F_{ST} values.

The SNP with the highest importance value in our RF model occurs on the gene *triosephosphate isomerase* (*TPI*), a highly conserved enzyme fundamental to glycolysis and glycogenesis (RF importance = 0.0023; Figure 3c). *TPI* was highly differentiated between MON and NAP ($F_{ST} = 0.44$; Figure 3c) and significantly upregulated in MON (Figure 4) (FDR p -value = 5.56×10^{-6} ; $\log_2$ Fold Change = 0.4036). Not surprisingly, we observed a substantial allele frequency change between NAP and MON at this locus (allele frequency = 0 vs. 0.47, respectively) (SNP 1, Figure 4, Table S7). This SNP is a silent mutation that occurs within the *TPI* open reading frame (ORF). We identified a missense variant (SNP 3; Figure 4a, Table S7), two additional silent mutations in the ORF (SNPs 2 and 4; Figure 4a, Table S7), and four additional SNPs downstream of the ORF at this locus (SNPs 5–8; Figure 4a, Table S7). The missense variant changes leucine in the historical range to valine in the expanded range (Figure 4c). To investigate the effects of the missense mutation on *TPI*, we used an approach that combines structure prediction and mutation mapping. The predicted structure for the mutant was almost identical to the wild type (C α RMSD = 0.08 Å, Figure 4c) and does not result in functional changes to this highly conserved protein, and may be facilitating the differences in expression. However, because this SNP does

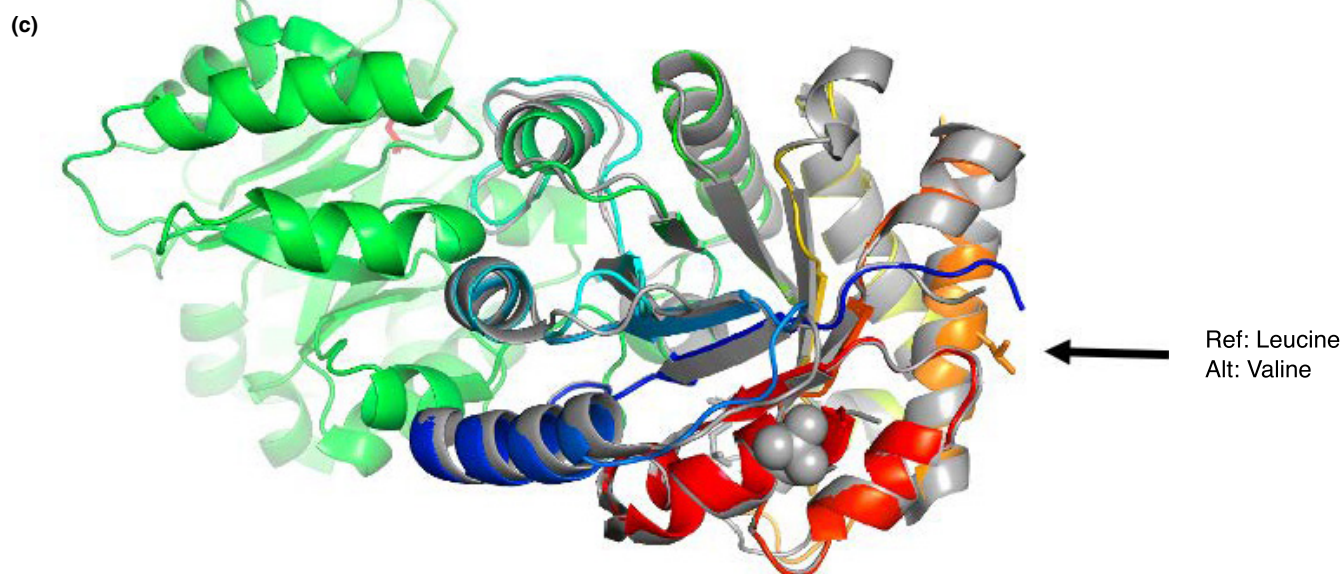
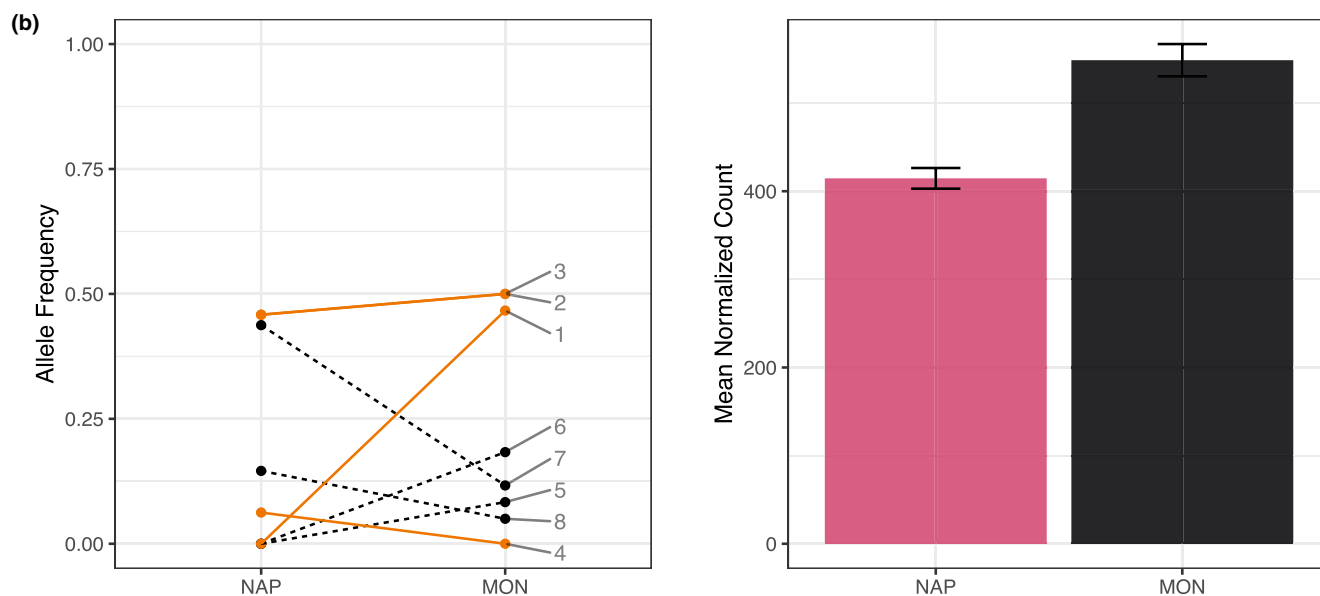
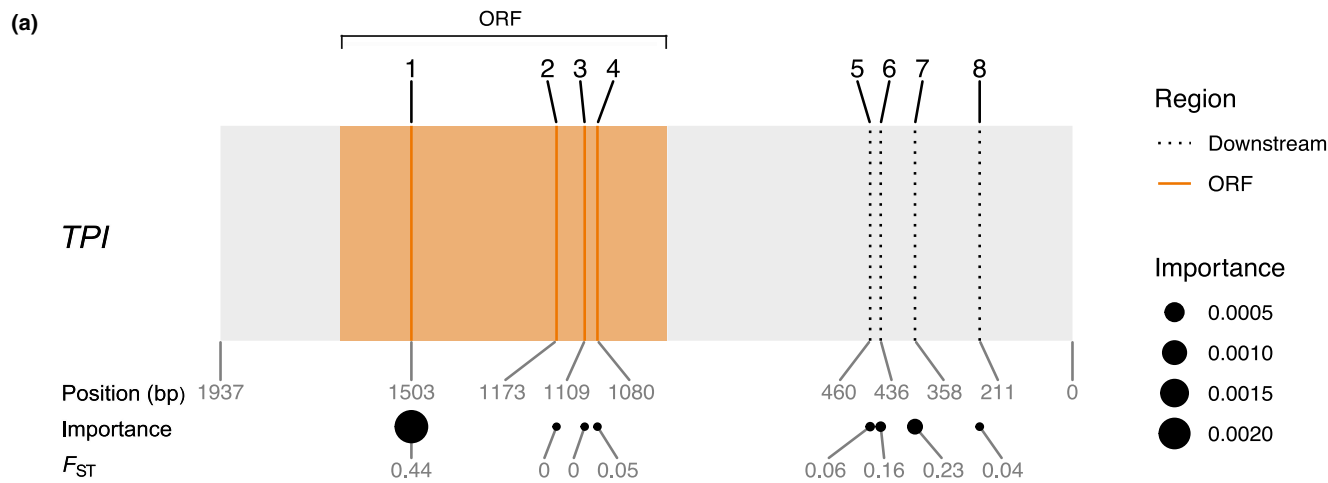


FIGURE 4 (a) A gene schematic showing the assembled transcript that triosephosphate isomerase (*TPI*) occurs on. Each vertical line represents the positions (bp) of high-importance SNPs within the open reading frame (solid orange lines) and downstream of *TPI* (dotted black lines). The positions are also noted in grey below the schematic and listed in Table S7. For each SNP, the circles represent importance values calculated from the random forest predictive model. The pairwise F_{ST} values between Monterey and Naples are also depicted. (b) Differences in allele frequencies and expression values between Monterey and Naples at *TPI*. Allele frequency of each SNP on *TPI* (left) and the mean normalized count of *TPI* in NAP and MON (right). *TPI* is upregulated in MON and SNP 1 is highly differentiated between the two populations. (c) Structure modelling of the wild-type *TPI* in Kellet's whelk, the arrow points towards the non-synonymous mutation site. The conservative mutation from Leucine to Valine is exposed to the outside and does not occur on the catalytic site, interface of dimer or beta-strand. The structural evaluation on Missense3D was 'No structural damage detected'.

not occur in regulatory regions, it is also possible that it is neutral and not related to the expression or function of *TPI*.

Gene Ontology (GO) enrichment analysis of the DEGs between MON and NAP identified 161 overrepresented GO terms ($p < 0.001$). Of the top 20 overrepresented terms, four were associated with metabolic processes and two were associated with response to stress, including cold acclimation (GO: 0009631) (Figure S9, Table S9). In addition to GO enrichment analysis, we also performed a weighted gene correlation network analysis (WGCNA) where two modules of co-expressed genes significantly differed between our MON and NAP samples after sequential Bonferroni correction (corrected $p < 0.05$). The first of these modules contained 3574 genes that were associated with 253 overrepresented GO terms ($p < 0.001$) related to cellular processes and biological regulation of cilium, including regulation of cilium beat frequency, positive regulation of cilium assembly and ciliary basal body organization (top 20 shown in terminal nodes in Figure S10; Table S9). The second module contained 156 genes associated with 19 significantly overrepresented GO terms, of which four were associated with metabolism and two were associated with response to stimulus, including response to reactive oxygen species (Figure S10, Table S10).

4 | DISCUSSION

By calling SNPs from an RNA-Seq data set from a common garden experiment, we identified population structure that was previously unidentified by microsatellite data sets in Kellet's whelks (Figure 2; Selkoe et al., 2010; White et al., 2010) and comparatively high F_{ST} values for an open-coast marine species (Benestan et al., 2015; Rey et al., 2020; Tepolt et al., 2009). We further identified 2770 genes that were differentially expressed between the species' expanded and historical ranges. Because all offspring were reared in a common garden, and thus, environmental conditions were identical among sequenced samples, these differences in gene expression can largely be attributed to genetic and/or trans-generational epigenetic differences between populations (Christie et al., 2016; Fitz-James & Cavalli, 2022; Roberge et al., 2008). Furthermore, using both the all-SNPs and DEG-SNPs data sets, we assigned samples to their respective ranges with high accuracy (63.6%–100%, Figure 3a, Table S4). This level of assignment accuracy in an open coast marine species is unprecedented (Christie et al., 2017).

In this study, we identified population structure and conducted accurate population assignment not only on differentially expressed

genes across Kellet's whelk's expanded and historical range (DEG-SNPs) but also across the entire transcriptome (all-SNPs), potentially discovering the genetic loci that underpin the phenotypic and/or plastic variation among populations (Donelson et al., 2018). Although population structure was recovered in both the all-SNPs and DEG-SNPs data sets, the DEG-SNPs data set resulted in higher assignment success, suggesting that the expression level of DEGs must covary (at least partially) with genetic structure. Specifically, we show that DEG-SNPs have higher F_{ST} values than the genomic background (Figure 3b).

Because these analyses were primarily conducted on the offspring from a sample of parents from the field (HxH and ExE offspring), population structure may be confounded with family structure and associated with genetic drift (Allendorf & Phelps, 1981). To account for this potential bias, we validated our assignment models in three ways. (1) by assigning 169 completely independent wild-caught samples with 155 GT-Seq loci identified directly from DEGs. Because these samples were independently collected from the wild on different collection trips in different years, they represent an entirely independent set of samples useful for unbiased population assignment (Anderson, 2010). Here, we were able to assign these independent, wild-caught samples back to their population of origin based on SNPs identified in this study from our main experimental crosses conducted in a common laboratory environment, (2) cross-validation to avoid covariance across samples due to environmental effects rather than population differences and (3) independent crosses (reciprocal crosses between MON and NAP) which avoids family structure and tank bias in our assignment. By coupling transcriptomics and common garden experiments with a posteriori genotyping of independent samples, we present a novel approach to resolve population structure, identify genetic adaptation and assess population connectivity in high gene flow systems.

From an ecological perspective, our ability to reliably assign individuals back to their range of origin (historical or expanded) at these loci may suggest self-recruitment at a regional level, even for a species with a relatively long pelagic duration of 40–60 days (Romero et al., 2012). Stated differently, our high-accuracy assignment results suggest that there is limited gene flow from the southern historical population to the northern expanded range that might slow or prevent a response to selection in the expanded range (Conover et al., 2006). This limited gene flow may be influenced by point conception, a previously identified biogeographic range boundary (Dong et al., 2009). There could be a combination of continuous

gene flow, high regional retention, large population sizes and selection that maintains genetic boundaries (Gould & Dunlap, 2017; Thia et al., 2021; Waple, 1998). Large effective population sizes and limited genetic drift may further minimize divergence at neutral loci (but see Hedgecock & Pudovkin, 2011). Another possibility is that ongoing gene flow is followed by subsequent viability selection that can be seen at the 'adaptive' loci. Lastly, the DEGs could be influenced, at least in part, by trans-generational epigenetic and/or maternal effects (Mousseau & Fox, 1998; Wolf & Wade, 2009). Nevertheless, the lack of population structure at putatively neutral microsatellite loci (Figure 2) suggests ongoing and moderate to high rates of gene flow between populations in the historical and expanded range, potentially driven by El Niño Southern Oscillation (ENSO) events (López et al., 2024; Zacherl et al., 2003), or by the relatively small number of generations that have elapsed since colonization since divergence (i.e. recent population expansion in the 1970s) (Hart & Marko, 2010; Marko & Hart, 2011).

We observed one missense mutation and three silent mutations (of which one had the single highest importance for RF model predictions) in the open reading frame and four additional mutations downstream of *triosephosphate isomerase* (*TPI*). *TPI* is an essential enzyme involved in glycolysis and glycogenesis that facilitates the conversion between glyceraldehyde 3-phosphate (GA3P) and dihydroxyacetone phosphate (DHAP). Like *TPI*, many metabolic enzymes that affect energy production and other aspects of organismal function are expected to undergo strong selection in novel environments (Marden, 2013). In fact, ecologically significant functional variations at these loci are often identified in transcriptome- and genome-wide analyses and implicated in population fitness (Marden, 2013). In addition to its important role in glycolysis and glucogenesis, previous studies have shown that *TPI* plays an important role in cold stress response. For example, mRNA level of *TPI* was higher in freeze-tolerant strains of baker's yeast *Saccharomyces cerevisiae* (Rodríguez-Vargas et al., 2002). Indeed, enzymes involved in carbohydrate metabolism (and specifically glycolysis) are known to increase in abundance when exposed to cold stress (Ralsler et al., 2007), possibly due to the need to alter metabolic rates under cold conditions (Tomanek, 2014; Vasquez et al., 2019). Furthermore, allele frequency variation of *TPI* is associated with body condition factors in lake whitefish (*Coregonus* spp.) (Bernatchez et al., 2010; Renaut et al., 2011). The upregulation of *TPI* observed in MON (ExE offspring) and the overrepresentation of metabolism genes among our DEGs in general suggest that whelks in their newly expanded range are adapting to the substantially colder environment (Figure 1) by altering their expression of metabolic pathways (Vasquez et al., 2019). Whether or not this response to selection is reversible remains unknown but may be important as ocean waters continue to warm (Lotterhos et al., 2021).

Although we do not explicitly relate population differentiation to differences in transcription (e.g. alternative splicing) in this study, we suspect that many of the loci identified using the experimental transcriptomic approach here would have been missed using a genomics approach. In genome-wide data, F_{ST} values follow a distribution where some values are larger simply as a result of variation and

may or may not be under genetic adaptation (Anderson, 2010). With more than 25,000 genes and lots of non-coding and untranslated DNA in the Kellet's whelk genome (Daniels, Andrasz, et al., 2023), we hypothesize that the experimental transcriptomic approach we applied here can uncover informative loci, which we illustrate here with assignment tests on independently caught wild samples. In the future, further investigations that directly compare genomic outlier approaches (Milano et al., 2014; Silliman, 2019) and experimental transcriptomic approaches (Harder et al., 2020; Yin et al., 2021) to identify putatively adaptive loci and population structure would be an important contribution to the field.

Using experimental transcriptomics, we uncovered putative genetic adaptation in a range-expanding population of a species with the potential for high gene flow. Identifying such distinct population structure is critical for the successful conservation and management of marine species (Reiss et al., 2009), including spatial management of fisheries species (such as our study organism; California Department of Fish and Wildlife, 2020), and species experiencing climate change across their geographic range (Selden & Pinsky, 2019). Expanded populations of Kellet's whelk are readily adapting to novel conditions in the expanded range. We provide strong evidence that genetic adaptation can occur in the ocean despite high gene flow (Brennan et al., 2022; Hughes et al., 2003; Wilder et al., 2020) and identify several putatively adaptive loci associated with cold tolerance and metabolic stress. These findings improve our understanding of how marine organisms are responding to their rapidly changing environment. Specifically, despite warming oceans, an adaptive response to cold stress may help marine species tolerate colder temperatures as they expand their ranges poleward (Sunday et al., 2012). Identification of genetic adaptation in marine systems has broad implications: marker panels can be designed to identify population connectivity in marine systems, facilitating the design and implementation of MPAs (Abesamis et al., 2017; Balbar & Metaxas, 2019), loci underlying traits responding to divergent or directional selection can be identified, and mechanistic frameworks can be developed for bridging micro- and macro-evolutionary patterns of diversification in the sea.

AUTHOR CONTRIBUTIONS

AL, CW, RT and MRC were involved in conceptualization. BD, CW and AL were involved in sample collection/preparation. All authors were involved in analysis and editing. AL and MRC wrote the manuscript with input from all authors. CW, RT and MRC were involved in funding.

ACKNOWLEDGEMENTS

We thank five anonymous reviewers for their useful comments that greatly improved the manuscript. Amanda Bird, Steve Lonhart and Grant Waltz assisted with field collections of wild adult whelks. Cody Mestice, Mari Ferlin and Jenna Nurge assisted with managing the experimental aquaria at the Cal Poly pier. The Rosen Center for Advanced Computing provided computational resources and support, with special thanks to YuCheng Zhang. Andy Lee

was supported in part by the Purdue Department of Biological Sciences Waser Fellowship and Frederick N. Andrews Fellowship. Daisuke Kihara acknowledges support from the National Institutes of Health (R01GM133840) and the National Science Foundation (DBI-2003635). This project was funded by the National Science Foundation grant (DEB-1856710) to C. White, R. Toonen and M. Christie.

CONFLICT OF INTEREST STATEMENT

None declared.

DATA AVAILABILITY STATEMENT

All raw data and sample metadata generated for this project are stored in the NCBI Short Read Archive (SRA) under project PRJNA1000198 (<https://www.ncbi.nlm.nih.gov/sra/PRJNA1000198>). VCF files for the all-SNPs and DEG-SNPs data sets are deposited in Dryad (<https://doi.org/10.5061/dryad.qbzk18s3>). The scripts used in this project are hosted in a public repository (https://github.com/ChristieLab/kellets_whelk_rnaseq).

ORCID

Andy Lee  <https://orcid.org/0009-0001-3128-737X>
 Benjamin N. Daniels  <https://orcid.org/0009-0002-1202-0765>
 William Hemstrom  <https://orcid.org/0000-0002-2408-9535>
 Cataixa López  <https://orcid.org/0000-0002-6877-9466>
 Robert J. Toonen  <https://orcid.org/0000-0001-6339-4340>
 Mark R. Christie  <https://orcid.org/0000-0001-7285-5364>

REFERENCES

- Abesamis, R. A., Saenz-Agudelo, P., Berumen, M. L., Bode, M., Jadloc, C. R. L., Solera, L. A., Villanoy, C. L., Bernardo, L. P. C., Alcala, A. C., & Russ, G. R. (2017). Reef-fish larval dispersal patterns validate no-take marine reserve network connectivity that links human communities. *Coral Reefs*, 36(3), 791–801. <https://doi.org/10.1007/s00338-017-1570-0>
- Alexa, A., & Rahnenführer, J. (2023). *topGO: Enrichment analysis for gene ontology* (2.52.0) [computer software]. Bioconductor version: Release (3.17). <https://doi.org/10.18129/B9.bioc.topGO>
- Alexa, A., Rahnenführer, J., & Lengauer, T. (2006). Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics*, 22(13), 1600–1607. <https://doi.org/10.1093/bioinformatics/btl140>
- Allendorf, F. W., & Phelps, S. R. (1981). Use of allelic frequencies to describe population structure. *Canadian Journal of Fisheries and Aquatic Sciences*, 38(12), 1507–1514. <https://doi.org/10.1139/f81-203>
- Altmann, A., Tološi, L., Sander, O., & Lengauer, T. (2010). Permutation importance: A corrected feature importance measure. *Bioinformatics*, 26(10), 1340–1347. <https://doi.org/10.1093/bioinformatics/btq134>
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 3, 403–410. [https://doi.org/10.1016/S0022-2836\(05\)80360-2](https://doi.org/10.1016/S0022-2836(05)80360-2)
- Anderson, E. C. (2010). Assessing the power of informative subsets of loci for population assignment: Standard methods are upwardly biased. *Molecular Ecology Resources*, 10, 701–710. <https://doi.org/10.1111/j.1755-0998.2010.02846.x>
- Balbar, A. C., & Metaxas, A. (2019). The current application of ecological connectivity in the design of marine protected areas. *Global Ecology and Conservation*, 17, e00569. <https://doi.org/10.1016/j.gecco.2019.e00569>
- Barth, J. M. I., Berg, P. R., Jonsson, P. R., Bonanomi, S., Corell, H., Hemmer-Hansen, J., Jakobsen, K. S., Johannesson, K., Jorde, P. E., Knutsen, H., Moksnes, P.-O., Star, B., Stenseth, N. C., Svedäng, H., Jentoft, S., & André, C. (2017). Genome architecture enables local adaptation of Atlantic cod despite high connectivity. *Molecular Ecology*, 26(17), 4452–4466. <https://doi.org/10.1111/mec.14207>
- Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R., & Bernatchez, L. (2015). RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). *Molecular Ecology*, 24, 3299–3315. <https://doi.org/10.1111/mec.13245>
- Benestan, L., Quinn, B. K., Maaroufi, H., Laporte, M., Clark, F. K., Greenwood, S. J., Rochette, R., & Bernatchez, L. (2016). Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology*, 25(20), 5073–5092. <https://doi.org/10.1111/mec.13811>
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B: Methodological*, 57(1), 289–300.
- Bernatchez, L., Renaut, S., Whiteley, A. R., Derome, N., Jeukens, J., Landry, L., ... St-Cyr, J. (2010). On the origin of species: Insights from the ecological genomics of lake whitefish. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 365(1547), 1783–1800.
- Bierne, N. (2010). The distinctive footprints of local hitchhiking in a varied environment and global hitchhiking in a subdivided population. *Evolution*, 64(11), 3254–3272. <https://doi.org/10.1111/j.1558-5646.2010.01050.x>
- Brennan, R. S., deMayo, J. A., Dam, H. G., Finiguerra, M., Baumann, H., Buffalo, V., & Pespeni, M. H. (2022). Experimental evolution reveals the synergistic genomic mechanisms of adaptation to ocean warming and acidification in a marine copepod. *Proceedings of the National Academy of Sciences of the United States of America*, 119(38), e2201521119. <https://doi.org/10.1073/pnas.2201521119>
- California Department of Fish and Wildlife. (2020). Kellet's Whelk, *Kelletia kelletii*. Enhanced Status Report.
- Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., & Madden, T. L. (2009). BLAST+: Architecture and applications. *BMC Bioinformatics*, 10(1), 421. <https://doi.org/10.1186/1471-2105-10-421>
- Campbell, N. R., Harmon, S. A., & Narum, S. R. (2015). Genotyping-in-thousands by sequencing (GT-seq): A cost effective SNP genotyping method based on custom amplicon sequencing. *Molecular Ecology Resources*, 15(4), 855–867. <https://doi.org/10.1111/1755-0998.12357>
- Cantalapiedra, C. P., Hernández-Plaza, A., Letunic, I., Bork, P., & Huerta-Cepas, J. (2021). eggNOG-mapper v2: Functional annotation, Orthology assignments, and domain prediction at the metagenomic scale. *Molecular Biology and Evolution*, 38(12), 5825–5829. <https://doi.org/10.1093/molbev/msab293>
- Christie, M. R., Marine, M. L., Fox, S. E., French, R. A., & Blouin, M. S. (2016). A single generation of domestication heritably alters the expression of hundreds of genes. *Nature Communications*, 7(1), 10676. <https://doi.org/10.1038/ncomms10676>
- Christie, M. R., Meirmans, P. G., Gaggiotti, O. E., Toonen, R. J., & White, C. (2017). Disentangling the relative merits and disadvantages of parentage analysis and assignment tests for inferring population connectivity. *ICES Journal of Marine Science*, 74(6), 1749–1762. <https://doi.org/10.1093/icesjms/fsx044>
- Christie, M. R., Tissot, B. N., Albins, M. A., Beets, J. P., Jia, Y., Ortiz, D. M., Thompson, S. E., & Hixon, M. A. (2010). Larval connectivity in an effective network of Marine protected areas. *PLoS One*, 5(12), e15715. <https://doi.org/10.1371/journal.pone.0015715>

- Conover, D. O., Clarke, L. M., Munch, S. B., & Wagner, G. N. (2006). Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *Journal of Fish Biology*, 69(sc), 21–47. <https://doi.org/10.1111/j.1095-8649.2006.01274.x>
- Cornwell, B. H. (2020). Gene flow in the anemone *Anthopleura elegantissima* limits signatures of local adaptation across an extensive geographic range. *Molecular Ecology*, 29(14), 2550–2566. <https://doi.org/10.1111/mec.15506>
- Cowen, R. K., Gawarkiewicz, G., Pineda, J., Thorrold, S. R., & Werner, F. E. (2007). Population connectivity in Marine systems: An overview. *Oceanography*, 20(3), 14–21. <https://doi.org/10.5670/oceanog.2007.26>
- D'Aloia, C. C., Bogdanowicz, S. M., Andrés, J. A., & Buston, P. M. (2022). Population assignment tests uncover rare long-distance marine larval dispersal events. *Ecology*, 103(1), e03559. <https://doi.org/10.1002/ecy.3559>
- Daniels, B., Andras, C. L., Zarate, N., Lee, A., López, C., Christie, M. R., Toonen, R. J., White, C., Anderson, P., & Davidson, J. (2023). De novo genome and transcriptome assembly of *Kelletia kelletii*, a coastal gastropod and fisheries species exhibiting a northern range expansion. *Frontiers in Marine Science*, 10, 1278131. <https://doi.org/10.3389/fmars.2023.1278131>
- Daniels, B. N., Nurge, J., Sleeper, O., Lee, A., Lopez, C., Christie, M. R., Toonen, R. J., Davidson, J., & White, C. (2023). Genomic DNA extraction optimization and validation for genome sequencing using the marine gastropod *Kellet's* whelk. *PeerJ*, 11, e16510. <https://doi.org/10.7717/peerj.16510>
- Donelson, J., Salinas, S., Munday, P. L., & Shama, L. N. S. (2018). Transgenerational plasticity and climate change experiments: Where do we go from here? *Global Change Biology*, 24, 13–34. <https://doi.org/10.1111/gcb.13903>
- Dong, C., Idica, E. Y., & McWilliams, J. C. (2009). Circulation and multiple-scale variability in the Southern California bight. *Progress in Oceanography*, 82(3), 168–190. <https://doi.org/10.1016/j.pocean.2009.07.005>
- DuBois, K., Pollard, K. N., Kauffman, B. J., Williams, S. L., & Stachowicz, J. J. (2022). Local adaptation in a marine foundation species: Implications for resilience to future global change. *Global Change Biology*, 28(8), 2596–2610. <https://doi.org/10.1111/gcb.16080>
- Evanno, G., Regnaut, S., & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Molecular Ecology*, 14(8), 2611–2620. <https://doi.org/10.1111/j.1365-294X.2005.02553.x>
- Fitz-James, M. H., & Cavalli, G. (2022). Molecular mechanisms of transgenerational epigenetic inheritance. *Nature Reviews Genetics*, 23, 325–341. <https://doi.org/10.1038/s41576-021-00438-5>
- Foster, Z. S. L., Sharpton, T. J., & Grünwald, N. J. (2017). Metacoder: An R package for visualization and manipulation of community taxonomic diversity data. *PLoS Computational Biology*, 13(2), e1005404. <https://doi.org/10.1371/journal.pcbi.1005404>
- Francis, R. M. (2017). Pophelper: An R package and web app to analyse and visualize population structure. *Molecular Ecology Resources*, 17(1), 27–32. <https://doi.org/10.1111/1755-0998.12509>
- Gagnaire, P.-A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., Arnaud-Haond, S., & Bierne, N. (2015). Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications*, 8(8), 769–786. <https://doi.org/10.1111/eva.12288>
- Gagnaire, P.-A., Normandeau, E., Côté, C., Møller Hansen, M., & Bernatchez, L. (2012). The genetic consequences of spatially varying selection in the Panmictic American eel (*Anguilla rostrata*). *Genetics*, 190(2), 725–736. <https://doi.org/10.1534/genetics.111.134825>
- Gaylord, B., & Gaines, S. D. (2000). Temperature or transport? Range limits in Marine species mediated solely by flow. *The American Naturalist*, 155(6), 769–789. <https://doi.org/10.1086/303357>
- Gil, N., & Ulitsky, I. (2020). Regulation of gene expression by cis-acting long non-coding RNAs. *Nature Reviews Genetics*, 21, 102–117. <https://doi.org/10.1038/s41576-019-0184-5>
- Gkanogiannis, A. (2023). *fastreeR: Phylogenetic, distance and other calculations on VCF and Fasta files* (1.4.0) [computer software]. Bioconductor Version: Release (3.17). <https://doi.org/10.18129/B9.bioc.fastreeR>
- Gosnell, J. S., Clare, X. S., Whited, D., Chiu, J., Huie, S., Zacherl, D., & White, C. (2024). Length-weight relationship of the kelp forest gastropod and emerging fisheries species, kellet's whelk, *Kelletia kelletii*. *Journal of Shellfish Research*, 42(3), 501–506.
- Gould, A. L., & Dunlap, P. V. (2017). Genomic analysis of a cardinal-fish with larval homing potential reveals genetic admixture in the Okinawa Islands. *Molecular Ecology*, 26, 3870–3882. <https://doi.org/10.1111/mec.14169>
- Haenel, Q., Roesti, M., Moser, D., MacColl, A. D. C., & Berner, D. (2019). Predictable genome-wide sorting of standing genetic variation during parallel adaptation to basic versus acidic environments in stickleback fish. *Evolution Letters*, 3(1), 28–42. <https://doi.org/10.1002/evl3.99>
- Harder, A. M., Willoughby, J. R., Ardren, W. R., & Christie, M. R. (2020). Among-family variation in survival and gene expression uncovers adaptive genetic variation in a threatened fish. *Molecular Ecology*, 29(6), 1035–1049. <https://doi.org/10.1111/mec.15334>
- Hart, M. W., & Marko, P. B. (2010). It's about time: Divergence, demography, and the evolution of developmental modes in Marine invertebrates. *Integrative and Comparative Biology*, 50(4), 643–661. <https://doi.org/10.1093/icb/icq068>
- Hedgecock, D., & Pudovkin, A. I. (2011). Sweepstakes reproductive success in highly fecund Marine fish and shellfish: A review and commentary. *Bulletin of Marine Science*, 87(4), 971–1002. <https://doi.org/10.5343/bms.2010.1051>
- Hemstrom, W., Grummer, J. A., Luikart, G., & Christie, M. R. (2024). Next-generation data filtering in the genomics era. *Nature Reviews Genetics*, 1–18. <https://doi.org/10.1038/s41576-024-00738-6>
- Hemstrom, W., & Jones, M. (2022). snpR: User friendly population genomics for SNP datasets with categorical metadata. *Molecular Ecology Resources*, 23, 962–973. <https://doi.org/10.1111/1755-0998.13721>
- Hermisson, J., & Pennings, P. S. (2005). Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics*, 169(4), 2335–2352. <https://doi.org/10.1534/genetics.104.036947>
- Herrlinger, T. (1981). Range extension of *Kelletia kelletii*. *The Veliger*, 24(78), 1–530. <https://www.biodiversitylibrary.org/item/134152>
- Hughes, T. P., Baird, A. H., Bellwood, D. R., Card, M., Connolly, S. R., Folke, C., Grosberg, R., Hoegh-Guldberg, O., Jackson, J. B. C., Kleypas, J., Lough, J. M., Marshall, P., Nyström, M., Palumbi, S. R., Pandolfi, J. M., Rosen, B., & Roughgarden, J. (2003). Climate change, human impacts, and the resilience of coral reefs. *Science* (New York, N.Y.), 301(5635), 929–933. <https://doi.org/10.1126/science.1085046>
- Ittisoponpisan, S., Islam, S. A., Khanna, T., Alhuzimi, E., David, A., & Sternberg, M. J. E. (2019). Can predicted protein 3D structures provide reliable insights into whether missense variants are disease associated? *Journal of Molecular Biology*, 431(11), 2197–2212. <https://doi.org/10.1016/j.jmb.2019.04.009>
- Jakobsson, M., & Rosenberg, N. A. (2007). CLUMPP: A cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. *Bioinformatics*, 23(14), 1801–1806. <https://doi.org/10.1093/bioinformatics/btm233>
- Janes, J. K., Miller, J. M., Dupuis, J. R., Malenfant, R. M., Gorrell, J. C., Cullingham, C. I., & Andrew, R. L. (2017). The K=2 conundrum. *Molecular Ecology*, 26(14), 3594–3602. <https://doi.org/10.1111/mec.14187>

- Jombart, T. (2008). Adegnet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, 24(11), 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Židek, A., Potapenko, A., Bridgland, A., Meyer, C., Kohl, S. A. A., Ballard, A. J., Cowie, A., Romera-Paredes, B., Nikolov, S., Jain, R., Adler, J., ... Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature*, 596(7873), 583–589. <https://doi.org/10.1038/s41586-021-03819-2>
- Korneliussen, T. S., Albrechtsen, A., & Nielsen, R. (2014). ANGSD: Analysis of next generation sequencing data. *BMC Bioinformatics*, 15(1), 356. <https://doi.org/10.1186/s12859-014-0356-4>
- Kovaka, S., Zimin, A. V., Pertea, G. M., Razaghi, R., Salzberg, S. L., & Pertea, M. (2019). Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biology*, 20(1), 278. <https://doi.org/10.1186/s13059-019-1910-1>
- Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, 9(1), 559. <https://doi.org/10.1186/1471-2105-9-559>
- Langmead, B., & Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nature Methods*, 9(4), 357–359. <https://doi.org/10.1038/nmeth.1923>
- Latch, E. K., Dharmarajan, G., Glaubitz, J. C., & Rhodes, O. E., Jr. (2006). Relative performance of Bayesian clustering software for inferring population substructure and individual assignment at low levels of population differentiation. *Conservation Genetics*, 7, 295–302. <https://doi.org/10.1007/s10592-005-9098-1>
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., & 1000 Genome Project Data Processing Subgroup. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics*, 25(16), 2078–2079. <https://doi.org/10.1093/bioinformatics/btp352>
- Liao, Y., Smyth, G. K., & Shi, W. (2014). featureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics*, 30(7), 923–930. <https://doi.org/10.1093/bioinformatics/btt656>
- Liu, L., Zhang, X., Sun, D., Gao, T., & Song, N. (2019). Population genetic structure of *Liza affinis* (eastern Keelback mullet), reveals high gene flow inferred from microsatellite analysis. *Ocean Science Journal*, 54(2), 245–256. <https://doi.org/10.1007/s12601-019-0013-y>
- López, C., Danies, B. N., Freel, E. B., Lee, A., Davidson, M., White, C., Christie, M. R., & Toonen, R. J. (2024) [Preprint]. Climate-driven range expansion via long-distance larval dispersal. *Research Square*. <https://doi.org/10.21203/rs.3.rs-4670567/v1>
- Lotterhos, K. E., Láruson, Á. J., & Jiang, L.-Q. (2021). Novel and disappearing climates in the global surface ocean from 1800 to 2100. *Scientific Reports*, 11(1), 15535. <https://doi.org/10.1038/s41598-021-94872-4>
- Love, M. I., Huber, W., & Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology*, 15(12), 550. <https://doi.org/10.1186/s13059-014-0550-8>
- Mackay, T. F. C., Stone, E. A., & Ayroles, J. F. (2009). The genetics of quantitative traits: Challenges and prospects. *Nature Reviews Genetics*, 10(8), 565–577. <https://doi.org/10.1038/nrg2612>
- Maes, S. M., Christiansen, H., Mark, F. C., Lucassen, M., Van de Putte, A., Volckaert, F. A. M., & Flores, H. (2021). High gene flow in polar cod (*Boreogadus saida*) from West-Svalbard and the Eurasian Basin. *Journal of Fish Biology*, 99(1), 49–60. <https://doi.org/10.1111/jfb.14697>
- Marden, J. H. (2013). Nature's inordinate fondness for metabolic enzymes: Why metabolic enzyme loci are so frequently targets of selection. *Molecular Ecology*, 22, 5743–5764. <https://doi.org/10.1111/mec.12534>
- Marko, P. B., & Hart, M. W. (2011). The complex analytical landscape of gene flow inference. *Trends in Ecology & Evolution*, 26(9), 448–456. <https://doi.org/10.1016/j.tree.2011.05.007>
- McGuigan, K., Nishimura, N., Currey, M., Hurwit, D., & Cresko, W. A. (2011). Cryptic genetic variation and body size evolution in threespine stickleback. *Evolution*, 65(4), 1203–1211. <https://doi.org/10.1111/j.1558-5646.2010.01195.x>
- Meyer, L., Barry, P., Riquet, F., Foote, A., Der Sarkissian, C., Cunha, R. L., ... Gagnaire, P. A. (2024). Divergence and gene flow history at two large chromosomal inversions underlying ecotype differentiation in the long-snouted seahorse. *Molecular Ecology*, e17277.
- Milano, I., Babbucci, M., Cariani, A., Atanassova, M., Bekkevold, D., Carvalho, G. R., ... Bargelloni, L. (2014). Outlier SNP markers reveal fine-scale genetic structuring across European hake populations (*Merluccius merluccius*). *Molecular Ecology*, 23(1), 118–135. <https://doi.org/10.1111/mec.12568>
- Moran, B. M., & Anderson, E. C. (2019). Bayesian inference from the conditional genetic stock identification model. *Canadian Journal of Fisheries and Aquatic Sciences*, 76(4), 551–560. <https://doi.org/10.1139/cjfas-2018-0016>
- Mousseau, T. A., & Fox, C. W. (1998). The adaptive significance of maternal effects. *Trends in Ecology & Evolution*, 13(10), 403–407. [https://doi.org/10.1016/S0169-5347\(98\)01472-4](https://doi.org/10.1016/S0169-5347(98)01472-4)
- Nelson, J., Grande, T., & Wilson, M. (2006). *Fishes of the world* (4th ed.). Wiley.
- O'Donnell, D. R., Hamman, C. R., Johnson, E. C., Kremer, C. T., Klausmeier, C. A., & Litchman, E. (2018). Rapid thermal adaptation in a marine diatom reveals constraints and trade-offs. *Global Change Biology*, 24(10), 4554–4565. <https://doi.org/10.1111/gcb.14360>
- Paaby, A. B., & Rockman, M. V. (2014). Cryptic genetic variation: Evolution's hidden substrate. *Nature Reviews Genetics*, 15(4), 247–258. <https://doi.org/10.1038/nrg3688>
- Paradis, E., Claude, J., & Strimmer, K. (2004). APE: Analyses of Phylogenetics and evolution in R language. *Bioinformatics*, 20(2), 289–290. <https://doi.org/10.1093/bioinformatics/btg412>
- Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and Eigenanalysis. *PLoS Genetics*, 2(12), e190. <https://doi.org/10.1371/journal.pgen.0020190>
- Pinsky, M. L., Selden, R. L., & Kitchel, Z. J. (2019). Climate-driven shifts in marine species ranges: scaling from organisms to communities. *Annual Review of Marine Science*, 12, 153–179.
- Pinsky, M. L., Worm, B., Fogarty, M. J., Sarmiento, J. L., & Levin, S. A. (2013). Marine taxa track local climate velocities. *Science (New York, N.Y.)*, 341(6151), 1239–1242. <https://doi.org/10.1126/science.1239352>
- Poloczanska, E. S., Brown, C. J., Sydeman, W. J., Kiessling, W., Schoeman, D. S., Moore, P. J., Brander, K., Bruno, J. F., Buckley, L. B., Burrows, M. T., Duarte, C. M., Halpern, B. S., Holding, J., Kappel, C. V., O'Connor, M. I., Pandolfi, J. M., Parmesan, C., Schwing, F., Thompson, S. A., & Richardson, A. J. (2013). Global imprint of climate change on marine life. *Nature Climate Change*, 3(10), 919–925. <https://doi.org/10.1038/nclimate1958>
- Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. *Nature Genetics*, 38(8), 904–909. <https://doi.org/10.1038/ng1847>
- Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, 155(2), 945–959. <https://doi.org/10.1093/genetics/155.2.945>
- Puechmaille, S. J. (2016). The program structure does not reliably recover the correct population structure when sampling is uneven: subsampling and new estimators alleviate the problem. *Molecular Ecology Resources*, 16(3), 608–627.
- Quintero-Galvis, J. F., Ocampo-Zuleta, K., Castro, L. R., & Narváez-Barandica, J. C. (2023). Limited population genetic structure in the littoral gastropod *Nerita tessellata* (*Neritimorpha*, *Neritidae*) suggests high dispersal across the Caribbean Sea. *Journal of Experimental Marine Biology and Ecology*, 568, 151942. <https://doi.org/10.1016/j.jembe.2023.151942>

- Ralser, M., Wamelink, M. M., Kowald, A., Gerisch, B., Heeren, G., Struys, E. A., Klipp, E., Jakobs, C., Breitenbach, M., Lehrach, H., & Krobisch, S. (2007). Dynamic rerouting of the carbohydrate flux is key to counteracting oxidative stress. *Journal of Biology*, 6(4), 10. <https://doi.org/10.1186/jbiol61>
- Reiss, H., Hoarau, G., Dickey-Collas, M., & Wolff, W. J. (2009). Genetic population structure of marine fish: Mismatch between biological and fisheries management units. *Fish and Fisheries*, 10(4), 361–395. <https://doi.org/10.1111/j.1467-2979.2008.00324.x>
- Renaut, S., Nolte, A. W., Rogers, S. M., Derome, N., & Bernatchez, L. (2011). SNP signatures of selection on standing genetic variation and their association with adaptive phenotypes along gradients of ecological speciation in lake whitefish species pairs (*Coregonus* spp.). *Molecular Ecology*, 20, 545–559. <https://doi.org/10.1111/j.1365-294X.2010.04952.x>
- Rey, C., Darnaude, A., Ferraton, F., Guinand, B., Bonhomme, F., Bierne, N., & Gagnaire, P. A. (2020). Within-generation polygenic selection shapes fitness-related traits across environments in juvenile sea bream. *Genes*, 11(4), 398.
- Roberge, C., Normandeau, É., Einum, S., Guderley, H., & Bernatchez, L. (2008). Genetic consequences of interbreeding between farmed and wild Atlantic salmon: Insights from the transcriptome. *Molecular Ecology*, 17(1), 314–324. <https://doi.org/10.1111/j.1365-294X.2007.03438.x>
- Rodriguez-Vargas, S., Estruch, F., & Randez-Gil, F. (2002). Gene expression analysis of cold and freeze stress in Baker's yeast. *Applied and Environmental Microbiology*, 68(6), 3024–3030. <https://doi.org/10.1128/AEM.68.6.3024-3030.2002>
- Rombel, I. T., Sykes, K. F., Rayner, S., & Johnston, S. A. (2002). ORF-FINDER: A vector for high-throughput gene identification. *Gene*, 282(1), 33–41. [https://doi.org/10.1016/S0378-1119\(01\)00819-8](https://doi.org/10.1016/S0378-1119(01)00819-8)
- Romero, M. R., Walker, K. M., Cortez, C. J., Sanchez, Y., Nelson, K. J., Ortega, D. C., Smick, S. L., Hoese, W. J., & Zacherl, D. C. (2012). Larval diel vertical migration of the Marine gastropod *Kelletia kelletii* (Forbes, 1850). *Journal of Marine Sciences*, 2012, e386575. <https://doi.org/10.1155/2012/386575>
- Rosenthal, R. J. (1970). Observations on the reproductive biology of the Kellet's whelk, *Kelletia Kelletii*. *Veliger*, 12(3), 319–324.
- Sanford, E., & Kelly, M. W. (2011). Local adaptation in Marine invertebrates. *Annual Review of Marine Science*, 3(1), 509–535. <https://doi.org/10.1146/annurev-marine-120709-142756>
- Schaal, S. M., Haller, B. C., & Lotterhos, K. E. (2022). Inversion invasions: When the genetic basis of local adaptation is concentrated within inversions in the face of gene flow. *Philosophical Transactions of the Royal Society B*, 377(1856), 20210200.
- Schraiddt, C. E., Ackiss, A. S., Larson, W. A., Rowe, M. D., Höök, T. O., & Christie, M. R. (2023). Dispersive currents explain patterns of population connectivity in an ecologically and economically important fish. *Evolutionary Applications*, 16(7), 1284–1301. <https://doi.org/10.1111/eva.13567>
- Selden, R., & Pinsky, M. (2019). Chapter 20—Climate change adaptation and spatial fisheries management. In A. M. Cisneros-Montemayor, W. W. L. Cheung, & Y. Ota (Eds.), *Predicting future oceans* (pp. 207–214). Elsevier. <https://doi.org/10.1016/B978-0-12-817945-1.00023-X>
- Selkoe, K. A., Watson, J. R., White, C., Horin, T. B., Iacchei, M., Mitarai, S., Siegel, D. A., Gaines, S. D., & Toonen, R. J. (2010). Taking the chaos out of genetic patchiness: Seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology*, 19(17), 3708–3726. <https://doi.org/10.1111/j.1365-294X.2010.04658.x>
- Signor, S. A., & Nuzhdin, S. V. (2018). The evolution of gene expression in cis and trans. *Trends in Genetics*, 34(7), 532–544. <https://doi.org/10.1016/j.tig.2018.03.007>
- Silliman, K. (2019). Population structure, genetic connectivity, and adaptation in the Olympia oyster (*Ostrea lurida*) along the west coast of North America. *Evolutionary Applications*, 12(5), 923–939. <https://doi.org/10.1111/eva.12766>
- Singh, S. P., Groeneveld, J. C., & Willows-Munro, S. (2019). Between the current and the coast: Genetic connectivity in the spiny lobster *Panulirus homarus rubellus*, despite potential barriers to gene flow. *Marine Biology*, 166(3), 36. <https://doi.org/10.1007/s00227-019-3486-4>
- Skotte, L., Korneliussen, T. S., & Albrechtsen, A. (2012). Association testing for next-generation sequencing data using score statistics. *Genetic Epidemiology*, 36(5), 430–437. <https://doi.org/10.1002/gepi.21636>
- Sunday, J. M., Bates, A. E., & Dulvy, N. K. (2012). Thermal tolerance and the global redistribution of animals. *Nature Climate Change*, 2(9), 686–690. <https://doi.org/10.1038/nclimate1539>
- Takeuchi, T., Masaoka, T., Aoki, H., Koyanagi, R., Fujie, M., & Satoh, N. (2020). Divergent northern and southern populations and demographic history of the pearl oyster in the western Pacific revealed with genomic SNPs. *Evolutionary Applications*, 13(4), 837–853. <https://doi.org/10.1111/eva.12905>
- Tepolt, C. K., Darling, J. A., Bagley, M. J., Geller, J. B., Blum, M. J., & Grosholz, E. D. (2009). European green crabs (*Carcinus maenas*) in the northeastern Pacific: Genetic evidence for high population connectivity and current-mediated expansion from a single introduced source population. *Diversity and Distributions*, 15(6), 997–1009.
- Thia, J. A., McGuigan, K., Liggins, L., Figueira, W. F., Bird, C. E., Mather, A., Evans, J. L., & Riginos, C. (2021). Genetic and phenotypic variation exhibit both predictable and stochastic patterns across an intertidal fish metapopulation. *Molecular Ecology*, 30, 4392–4414. <https://doi.org/10.1111/mec.15829>
- Thorson, G. (1950). Reproductive and larval ecology of marine bottom invertebrates. *Biological Reviews*, 25(1), 1–45. <https://doi.org/10.1111/j.1469-185X.1950.tb00585.x>
- Thorstensen, M. J., Baerwald, M. R., & Jeffries, K. M. (2021). RNA sequencing describes both population structure and plasticity-selection dynamics in a non-model fish. *BMC Genomics*, 22(1), 273. <https://doi.org/10.1186/s12864-021-07592-4>
- Tigano, A., & Friesen, V. L. (2016). Genomics of local adaptation with gene flow. *Molecular Ecology*, 25(10), 2144–2164. <https://doi.org/10.1111/mec.13606>
- Tomanek, L. (2014). Proteomics to study adaptations in marine organisms to environmental stress. *Journal of Proteomics*, 105, 92–106. <https://doi.org/10.1016/j.jprot.2014.04.009>
- Van der Auwera, G., & O'Connor, B. D. (2020). *Genomics in the cloud: Using Docker, GATK, and WDL in Terra*. O'Reilly Media.
- Vandamme, S. G., Maes, G. E., Raeymaekers, J. A. M., Cottenie, K., Imsland, A. K., Hellemans, B., Lacroix, G., Mac Aoidh, E., Martinsohn, J. T., Martínez, P., Robbins, J., Vilas, R., & Volckaert, F. A. M. (2014). Regional environmental pressure influences population differentiation in turbot (*Scophthalmus maximus*). *Molecular Ecology*, 23(3), 618–636. <https://doi.org/10.1111/mec.12628>
- Vasquez, M. C., Lippert, M. R., White, C., Walter, R. K., & Tomanek, L. (2019). Proteomic changes across a natural temperature gradient in a marine gastropod. *Marine Environmental Research*, 149, 137–147.
- Vendetti, J. E. (2020). Early development in *Kelletia kelletii* (Forbes, 1850) (Gastropoda: Buccinidae), an eastern Pacific gastropod with planktonic larvae. *Ciencias Marinas*, 46(4), 269–282. <https://doi.org/10.7731/cm.v46i4.3109>
- Wang, J. (2017). The computer program structure for assigning individuals to populations: Easy to use but easier to misuse. *Molecular Ecology Resources*, 17(5), 981–990.
- Waple, R. S. (1998). Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *Journal of Heredity*, 89(5), 438–450. <https://doi.org/10.1093/jhered/89.5.438>
- White, C., Selkoe, K. A., Watson, J., Siegel, D. A., Zacherl, D. C., & Toonen, R. J. (2010). Ocean currents help explain population genetic

- structure. *Proceedings. Biological sciences*, 277(1688), 1685–1694. <https://doi.org/10.1098/rspb.2009.2214>
- White, C., Tett, P., Beas, R., Lorda, J., Kushner, D. J., & Zacherl, D. C. (2022). *Tracking cohorts over time using size-frequency population survey data to derive growth functions, with Kellet's whelk case study*. Western Society of Naturalists.
- White, C., & Toonen, R. J. (2008). Isolation and characterization of polymorphic microsatellite and COI loci from the whelk *Kelletia kelletii*. *Molecular Ecology Resources*, 8(4), 881–883. <https://doi.org/10.1111/j.1755-0998.2008.02097.x>
- Whitlock, M. C., & McCauley, D. E. (1999). Indirect measures of gene flow and migration: $F_{ST} \approx 1/(4Nm+1)$. *Heredity*, 82(2), 117–125. <https://doi.org/10.1046/j.1365-2540.1999.00496.x>
- Wilder, A. P., Palumbi, S. R., Conover, D. O., & Therkildsen, N. O. (2020). Footprints of local adaptation span hundreds of linked genes in the Atlantic silverside genome. *Evolution Letters*, 4(5), 430–443.
- Wolf, J. B., & Wade, M. J. (2009). What are maternal effects (and what are they not)? *Philosophical Transactions of the Royal Society, B: Biological Sciences*, 364(1520), 1107–1115. <https://doi.org/10.1098/rstb.2008.0238>
- Wright, M. N., & Ziegler, A. (2017). Ranger: A fast implementation of random forests for high dimensional data in C++ and R. *Journal of Statistical Software*, 77, 1–17. <https://doi.org/10.18637/jss.v077.i01>
- Yin, X., Martinez, A. S., Perkins, A., Sparks, M. M., Harder, A. M., Willoughby, J. R., Sepúlveda, M. S., & Christie, M. R. (2021). Incipient resistance to an effective pesticide results from genetic adaptation and the canalization of gene expression. *Evolutionary Applications*, 14(3), 847–859. <https://doi.org/10.1111/eva.13166>
- Yu, G., Smith, D. K., Zhu, H., Guan, Y., & Lam, T. T.-Y. (2017). Ggtree: An R package for visualization and annotation of phylogenetic trees with their covariates and other associated data. *Methods in Ecology and Evolution*, 8(1), 28–36. <https://doi.org/10.1111/2041-210X.12628>
- Zacherl, D., Gaines, S. D., & Lonhart, S. I. (2003). The limits to biogeographical distributions: Insights from the northward range extension of the marine snail, *Kelletia kelletii* (Forbes, 1852). *Journal of Biogeography*, 30(6), 913–924. <https://doi.org/10.1046/j.1365-2699.2003.00899.x>

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Lee, A., Daniels, B. N., Hemstrom, W., López, C., Kagaya, Y., Kihara, D., Davidson, J. M., Toonen, R. J., White, C., & Christie, M. R. (2024). Genetic adaptation despite high gene flow in a range-expanding population. *Molecular Ecology*, 00, e17511. <https://doi.org/10.1111/mec.17511>