Geographic Divergence in Population Genomics and Shell Morphology Reveal History of Glacial Refugia in a Coastal Dogwhelk

Emily K. Longman[1], Joaquin C.B. Nunez[1], Eric Sanford[2,3], & Melissa Pespeni[1]

[1] Department of Biology, University of Vermont, Burlington, VT, USA

[2] Bodega Marine Laboratory, University of California Davis, Bodega Bay, CA, USA

[3] Department of Evolution and Ecology, University of California Davis, Davis, CA, USA

Corresponding Author: Emily K. Longman, Emily.Longman@uvm.edu

Keywords: *Nucella canaliculata*, Population genomics, Geometric morphometrics, Phylogeography, Ecological genomics

Abstract:

Studying contemporaneous spatial patterns of genomic diversity can yield important insights into the evolutionary processes that structure populations and shape patterns of adaptation. In contrast to the large number of marine species with planktonic larvae, populations of marine taxa with low dispersal and deep evolutionary divergences offer an opportunity to reveal the phylogeographic histories of marine ecosystems. Here, we constructed a draft genome assembly for the low-dispersing marine dogwhelk, *Nucella canaliculata*, and studied patterns of genomic diversity and shell morphometrics in 19 populations distributed along ~1,500 km of the west coast of North America. We observed significant population structure with a strong phylogeographic break at Monterey Bay, which was matched with divergence in shell morphology. Genomic patterns, concomitant with computer simulations, suggest that there were at least two refugial populations during the last glacial maximum that subsequently experienced post-glacial expansion and admixture. Furthermore, by linking genotype to phenotype, we identified candidate loci associated with the molecular bases of variation in shell morphology. These findings highlight that historical events can play a persistent role in structuring the genomic diversity of modern coastal populations, with important consequences for spatial variation in the adaptive capacity and vulnerability of populations to future climatic changes.

Introduction:

Understanding the population structure and demography of a species is essential for interpreting patterns of adaptation and linking phenotypic variation to its genetic bases [1]. Spatially and temporally varying selection pressures [2], historical events [3,4], biogeographic barriers to dispersal [5], and life history patterns [6], can all shape the present-day geographic distribution of populations and the genetic diversity they harbor. However, our understanding of population structure and subsequently the molecular bases of adaptation in many species is significantly hindered by the lack of genomic resources for non-model organisms [7,8]. This demographic information is of pressing concern because an understanding of extant levels of genomic diversity is needed to predict a species' adaptive capacity and vulnerability to a rapidly changing climate [9-11].

In marine ecosystems, understanding population connectivity and genetic diversity is further complicated by extrinsic factors such as oceanographic processes and the open and three-dimensional nature of marine environments [12–15], as well as intrinsic factors including the wide range of life history strategies and developmental modes present in marine organisms [6,16]. The location of phylogeographic breaks and the level of population structure often reflect a species' reproductive strategy; species with direct development and low dispersal potential typically have much more defined genetic structure than species with planktonic larvae [13,14,17,18].

Using marine taxa with low dispersal to understand historical responses to climate change could be a useful tool to predict how marine species will respond during the modern era of rapid global change [19]. This is because the genomes of species with low dispersal often reflect deeper evolutionary divergence over longer timescales [19–21]. The low levels of gene flow in

55 taxa with direct development allow for the historical separation among populations to persist for
56 many generations [13,20]. Additionally, advances in genomics and single nucleotide
57 polymorphism (SNP) based approaches can provide much greater resolution of these
58 demographic patterns [12,22,23]. Thus, studying the phylogeographic structure of low-dispersing
59 species can lend insight into the processes shaping the evolutionary history of marine ecosystems
60 and the impact of past and possibly future climatic changes.

61 Fossil assemblages of calcifying marine species along the northeast Pacific from the late
62 Pleistocene show that the geographic distribution of many species shifted equatorially in
63 response to Pleistocene climatic fluctuations [24–26]. Classic genetic models of Pleistocene
64 glaciation suggest that these climatic changes should have impacts on present-day genetic
65 diversity, with greater genetic diversity in areas of glacial refugia than areas of population
66 expansion, and high dissimilarity between refugial populations due to the deeper divergence time
67 [3,4]. However, to date, relatively few phylogeographic studies along the west coast of North
68 America show evidence that historical events, such as Pleistocene glaciation, left a signature on
69 contemporary patterns of genetic diversity [13,19,20,27]. Instead, because most species studied
70 to date have planktonic larvae and display high levels of connectivity, their population structure
71 has been primarily attributed to oceanographic forces and vicariance, with Point Conception and
72 Cape Mendocino serving as common phylogeographic breaks [13,15,28]. In contrast, the few
73 studies on low-dispersing taxa suggest phylogeographical breaks at Monterey Bay and the long
74 stretches of sandy beaches around Los Angeles, California [13,19,28].

75 One such marine invertebrate with low dispersal potential is the channeled dogwhelk,
76 *Nucella canaliculata*, which is a species of predatory gastropod (Family Muricidae) that lives on
77 rocky shores from central California to Alaska and consumes mussel and barnacle prey via a

chemo-mechanical mode of feeding [29–30]. Females lay benthic egg capsules, and crawl-away young establish in the surrounding area [29]. Additionally, individuals reside only in mussel beds on wave-exposed headlands, thus populations are isolated due to intervening sandy beaches. Previous research has shown that this species displays a pattern of local adaptation with striking spatial variation in predatory drilling ability [30–33], foraging behavior [34], and heat tolerance [35]. Despite the wealth of ecological knowledge about *N. canaliculata*, there is a lack of genomic resources for the species, limiting our understanding of the molecular mechanisms generating these adaptive patterns. Genetic studies of the mitochondrial locus cytochrome oxidase I (*mtCOI*) suggest that populations display a pattern of isolation by distance and that the southern-most sampled population in central California was the most genetically diverged [30]. The biogeographic patterns of intraspecific trait variation and the species habitat specificity suggest that the population genomics of this low dispersing taxa may reveal spatially complex phylogeographic patterns.

Here, we present a high-quality draft genome assembly generated from a single *Nucella canaliculata* individual, a genome-wide panel of genetic variation from 19 populations spanning ~1,500km of the west coast of North America using pooled-sequencing [36], and analyses of shell morphometrics from dogwhelks collected across the 19 field sites. There is a long history of malacologists using geographic variation in shell morphology as an indicator of potential isolation and genetic divergence among gastropod populations [37,38]. Thus, our analyses explore whether spatial variation in morphological traits is concordant with the phylogeographic patterns revealed by the genomic data. Additionally, we integrate the genomic and morphological data to begin to link genetic and phenotypic variation. Specifically with these genomic and morphological datasets, we aimed to address the following questions: 1) How is

genetic diversity in *N. canaliculata* spatially structured along ~1,500km of the west coast of North America? 2) Do patterns of genetic diversity appear to be structured more strongly by contemporary or historical processes? 3) Does variation in shell morphology suggest a similar biogeographic pattern of divergence as the genomic data? 4) Lastly, what are the molecular bases for intraspecific variation in shell morphology? Overall, our analyses revealed strong population structure with congruence between the genomic and phenotypic datasets, as well as evidence of glacial refugia.

Methods:

Genome Sequencing and Assembly

Genome assembly was based on source DNA extracted from one adult female dogwhelk collected from Bodega Marine Reserve, California (38°19'8", -123°4'28") in June 2022. High molecular weight DNA extraction from flash frozen foot tissue, library preparation and sequencing (Oxford Nanopore PromethION and Element Bio AVITI 600 flow cell) occurred at the DNA Technologies and Expression Analysis Core at the UC Davis Genome Center (see Appendix S1: Section S1 for more details).

We assembled the genome using a hybrid assembly approach using `DBG2OLC` [39]. This approach combines the continuity of the Oxford Nanopore long reads and the quality of the AVITI short reads (see Appendix S1: Section S1 for more details). Genome quality and completeness was quantified using `QUAST` [40] and `BUSCO` [v5.7.1] [41]. Repeats and low complexity DNA sequences were masked using `RepeatMasker` [v4.1.6] with the model trained on the Pacific oyster, *Crassostrea gigas* [42]. Genome annotation of the softmasked genome occurred using `Augustus` [v3.5.0] trained on *Drosophila melanogaster* [43].

124 ***Field Collections and DNA Extractions***

125 We collected 20 adult (>18 mm in length) *N. canaliculata* from 19 populations (*N* = 380
126 total individuals) along the Pacific coast of North America (Table S1). Collections occurred
127 during low tide at the 19 rocky headlands from October 2023 to July 2024. Directly after field
128 collections, we sampled and preserved in 95% EtOH foot muscle tissue for each dogwhelk. We
129 extracted the DNA from each individual dogwhelk using an E.Z.N.A Mollusc & Insect DNA Kit
130 (Omega Bio-Tek, Norcross, Georgia) with minor modifications. We determined DNA quality
131 and quantity using gel electrophoresis, a Nanodrop, and a Qubit. For each population, DNA from
132 the 20 individuals was pooled in equimolar amount. Subsequent library preparation
133 (Watchmaker DNA library preparation kit [Watchmaker Genomics, Boulder, Colorado] with 6
134 PCR cycles) and sequencing (150 base pair paired-end sequencing on two lanes of NovaSeqX
135 25B) was performed at the DNA Technologies and Expression Analysis Core at the UC Davis
136 Genome Center.

137 ***Bioinformatic Processing and SNP Calling***

138 We assessed the quality of the raw pooled sequences using `FastQC` [44]. Adapters were
139 detected and low-quality reads were trimmed on a sliding window using a minimum phred score
140 of 20 with `fastp` [v0.23.4] [45]. Reads were mapped to the softmasked *N. canaliculata*
141 reference genome using the Burrows-Wheeler algorithm, `bwa-mem2` [2.2.1] [46]. We filtered
142 (quality threshold of 40), sorted and removed duplicate reads from the binary alignment map
143 (bam) files using `SAMtools` [v1.9] [47] and `PicardTools` [v1.9] [48]. The bam files from
144 the two sequencing lanes were merged and the final bam files were indexed using `SAMtools`
145 [v1.9] [47]. Mean coverage of the bam files was 33.13, with an average mapping rate of 57.25
146 (Table S2). We used `freebayes` [v1.3.6] to call variants from the 19 bam files [49] and

147 subsequently performed stringent filtering of the VCF file with `vcftools` [v0.1.16] and the
148 `poolfstat` R package [50] (Appendix S1: Section S2).

149 ***Population Structure and Genomic Diversity***

150 Population structure was analyzed using principal component analysis (PCA) and F_{ST}.
151 The VCF file was converted to a pooldata object using the *vcf2pooldata* function in the
152 `poolfstat` R package [50]. A PCA was performed using the *randomallele* function. We
153 calculated a global F_{ST} metric, a block-jackknife estimation of F_{ST}, and a multi-locus F_{ST} over a
154 sliding window of 1kb using the *computeFST* function. Additionally, we estimated pairwise-
155 population F_{ST} using the *compute.pairwiseFST* function using the ANOVA method.

156 Population demography was further characterized by analyzing the **Ω** matrix in
157 `BayPass` [v2.41] [51]. An **Ω** matrix was generated using 100 pilot runs, then converted into a
158 covariance matrix then into dissimilarity matrix, which was subsequently visualized as a
159 hierarchical clustering tree using the ape R package [52].

160 We tested for isolation by distance (IBD) by evaluating whether there was a relationship
161 between pairwise genetic distance, $F_{ST}/(1- F_{ST})$, and geographic distance. Given that *N.*
162 *canaliculata* lacks a planktonic larval stage, distance between sites was measured as the distance
163 along the coastline using the measure tool on Google Maps. We used Mantel tests (1,000
164 permutations) to determine the significance of the relationships using the `ecodist` R package.

165 We calculated nucleotide diversity (π), Watterson's estimator (θ), and Tajima's D for
166 each population using `npstat` [v1] [53]. To do so, the bam file for each population was
167 converted to pileup format using `SAMtools` [v1.9] [47], then calculations for each statistic were
168 performed on a sliding window with a length of 25kb with a minimum and maximum coverage
169 of 20 and 120, respectively, a minimum base quality of 30 and a minimum allele count of 5.

170 ***Demographic Inference***

171 We compared whether two post-glacial demography models fit our data using the
172 coalescent framework implemented in `moments` [v1.2.2] [54]. The models tested were: 1)
173 expansion from a single glacial refugium, and 2) a secondary encounter between two glacial
174 refugia. We used a pipeline similar to that deployed in [55] and [56] to fit `moments` to Pool-Seq
175 data (see Appendix S1: Section S3 for more details). For each model we report the Akaike
176 information criterion (AIC) for 10 optimization runs. To translate coalescent units into
177 chronological time, we scaled the output of the model (i.e., θ) by the factor $4L\mu$. Where L is the
178 length of sequence and μ is the mutation rate of *Drosophila melanogaster* 2.8×10^{-9} [57].

179 In addition to model inference, we also conducted forward genetic simulations using
180 `SLiM` [58]. The goal of these forward simulations was to contextualize the patterns of genetic
181 variation expected under two proposed scenarios. Our SLiM simulations consisted of seven
182 demes arranged in a stepping-stone fashion (see Appendix S1: Section S3 for more details). In
183 the one-refugium model, the first deme recolonizes the entire habitat after an extinction event. In
184 contrast, in the two refugia model, the habitat is recolonized by demes 1 and 7 post an extinction
185 event.

186 ***Dogwhelk Shell Geometric Morphometrics***

187 Shell morphology was characterized using landmark analysis. We took photographs of
188 the ventral surfaces of 380 field-collected *N. canaliculata* dogwhelks from the 19 populations
189 using a Canon DSLR camera with a Canon EF 28-136mm f/3.5-5.6 IS lens on a tripod. We
190 utilized `tpsUtil` (v1.83) and `tpsDig` (v2.31) to place 15 landmarks on predefined locations
191 on the ventral surface of the shell [59]. Dogwhelks with damaged shells were removed from the
192 analyses (n=10). We used the `geomorph` package in R [60] and the program `MorphoJ` [61] to

193 perform statistical analyses of the landmark coordinates. We performed a Procrustes
194 transformation to remove variation in dogwhelk orientation using the *gpagen* R function then
195 then analyzed the Procrustes coordinates using PCA. We also tested for intraspecific variation in
196 shell shape as a function of size (quantified as the log of the centroid size for each dogwhelk),
197 site or latitude, and their interaction with a Procrustes ANOVA using the R function *procD.lm*
198 with 1,000 permutations.

199 ***Detection of the Genomic Bases of Morphological Variation***

200 We identified outliers associated with shell morphology using `BayPass` [v2.41] [51].
201 The coefficient of variation for PC1 and PC2 of the morphological data were used as phenotypic
202 input data for `BayPass`. Outlier SNPs were identified using a Bonferroni threshold (alpha =
203 $0.05/8,277,206 = 6.04 \times 10^{-9}$), then subsequently annotated with `SnpEff` [62]. Putative functional
204 annotation of proteins was done using `blast` against Molluscan taxa within the UniProtKB
205 reference proteomes – Swiss-Prot database.

206

207 **<u>Results:</u>**

208 ***The First Draft Genome of* Nucella canaliculata**

209 We utilized a hybrid assembly approach using a combination of Oxford Nanopore
210 Technologies PromethION long reads and Element Bio AVITI short reads to generate the first
211 genome of *Nucella canaliculata*. The *de novo* assembly was very large (a total assembly length
212 of 1.57 Gb with an N50 of 146,860 bp) and consisted of numerous scaffolds (18,944 contigs)
213 (Table 1). BUSCO scores indicate that the draft genome was 91.0% and 80.7% complete for
214 Eukaryota and Mollusca, respectively (Table 2).

215

216 **Table 1.** *Nucella canaliculata* genome assembly metrics.

Assembly Length (bp)	1,571,681,309
Number of Contigs/Scaffolds (bp)	18,944
Largest Contig/Scaffold (bp)	2,196,136
GC %	42.52%
N50 (bp)	146,860
L50 (bp)	3,117

217

218 **Table 2.** Quality of *Nucella canaliculata* genome assembly.

	Complete BUSCOs	**Complete and Single Copy BUSCOs**	**Complete and Duplicated BUSCOs**	**Fragmented BUSCOs**	**Missing BUSCOs**
Eukaryota (n = 255)	91.0%	85.1%	5.9%	4.7%	4.3%
Mollusca (n = 5295)	80.7%	70.9%	9.8%	3.2%	16.1%

219

220 ***Patterns of Population Structure***

221 The Pool-Seq data were used to analyze patterns of population structure and spatial
222 patterns of genomic diversity. After stringent filtering, our final SNP list for the 19 pools
223 contained 8,277,206 SNPs. The principal component analysis (PCA) showed strong clustering of
224 populations based on geography (Figure 1; Figure S1). PC1 captured 30.54% of the variation and
225 strongly partitioned the populations into two clusters, one north and one south of Monterey Bay.
226 PC2 (8.83% variation), PC3 (7.49% variation), PC4 (5.4% variation), and PC5 (4.93% variation)
227 all further subdivide the seven populations in the southern cluster. It was not until PC6 (4.79%
228 variation) that there is notable separation among the 12 populations in the cluster north of
229 Monterey Bay.

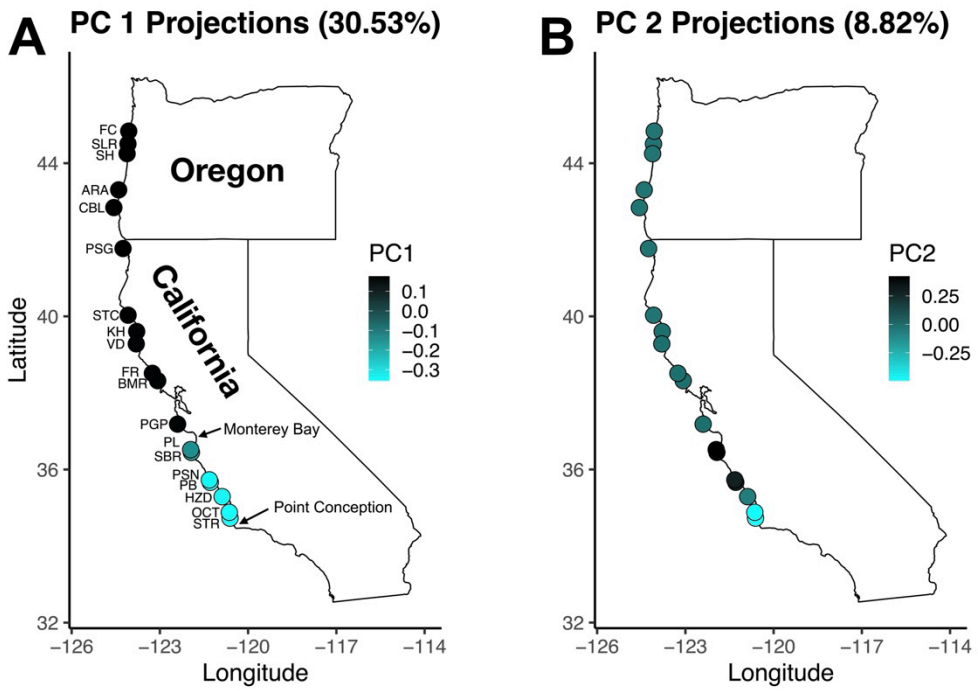

Figure 1. Population structure and genetic variation among the 19 *Nucella canaliculata* populations along the coast of California and Oregon, USA. Projections of the first two principal component analysis (PCA) loadings on the map of the collection sites: A) PC1 and B) PC2. The first map states the site codes and other biogeographically relevant landmarks. See Table S1 for an explanation of the population codes and coordinates.

The genome wide mean F_{ST} was 0.582, a high value for a marine species (Table S3). Multi-locus F_{ST} indicated high levels of F_{ST} distributed evenly across the genome (Figure S2). Pair-wise F_{ST} among the 19 populations ranged from 0.002 to 0.8 (Table S4). The greatest differentiation occurred between populations south versus north of Monterey Bay (top right corner of Figure 2A). There was little genetic differentiation among the populations north of Monterey Bay. In contrast, there was highly variable amounts of genetic differentiation among the populations south of Monterey Bay (bottom right corner of Figure 2A).

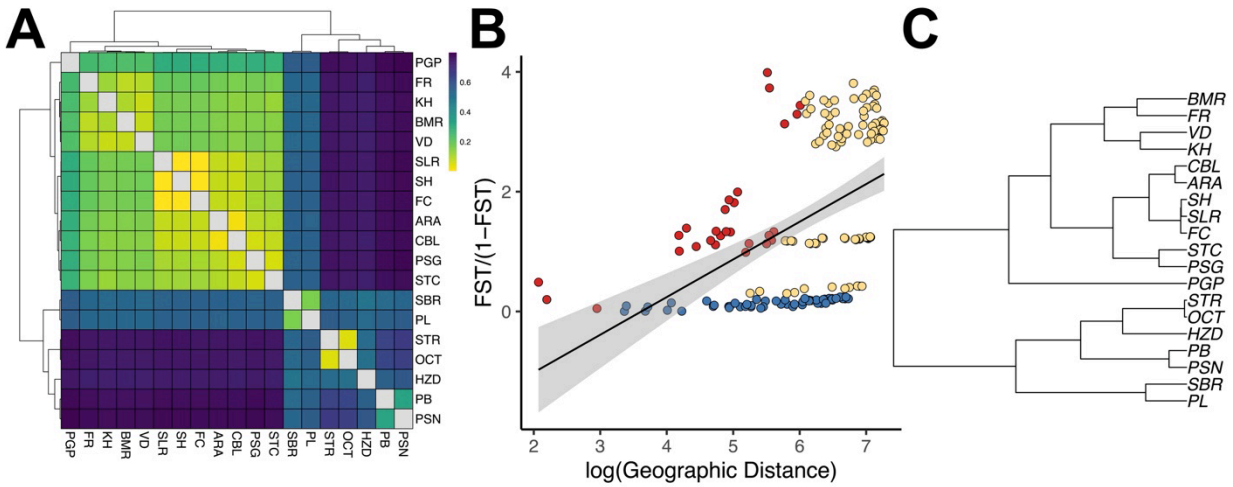

Figure 2. A) Heatmap of pairwise F_{ST} for the 19 *Nucella canaliculata* populations. Blue and yellow indicate high and low levels of genetic differentiation, respectively. B) The log of the geographic distance in km between sites and pairwise F_{ST}. Color indicates which sites are being compared — both in the southern cluster (red), both in the northern cluster (blue) and a comparison between a northern and southern site (yellow). C) Hierarchical clustering tree based on the omega matrix from `BayPass`. See Table S1 for definitions of population codes.

Across the 19 populations, *Nucella canaliculata* displayed a pattern of isolation by distance (Mantel tests, R^2=0.480, *P*-value=0.001; Figure 2B). However, this pattern is only present when contrasting populations across the southern versus northern genetic cluster, and within the southern cluster (Figure 2B). Despite their greater geographic distance, there was no pattern of isolation by distance among the populations north of Monterey Bay (blue dots in Figure 2B). Hierarchical clustering based on the `BayPass` population covariance matrix recapitulated the clustering of the populations into two genetically distinct groups (Figure 2C).

All 19 populations had very low levels of genetic variation, heterozygosity and Tajima's D (Figure 3; Figure S3; Table S4). The two populations just south of Monterey Bay, Soberanes Point and Point Lobos, had the highest nucleotide diversity and heterozygosity (Figure 3; Figure S3). Additionally, despite mean nucleotide diversity being similar across most of the populations (Table S4), the distribution and median nucleotide diversity strongly differed by geography (Figure 3A). The majority of populations south of Monterey Bay had a broader range of

nucleotide diversity and low median π, while the populations north of Monterey Bay have a very narrow range of nucleotide diversity and higher π. Almost all populations, except the three populations closest to Monterey Bay (SBR, PL and PGP) had negative median Tajima's D values indicating an excess of rare alleles in each population (Figure 3B; Table S4). The 11 most northern populations had mean Tajima's D values that were more negative than the five most southern populations. This pattern is indicative of an excess of rare alleles in the northern populations and is consistent with recent population expansion.

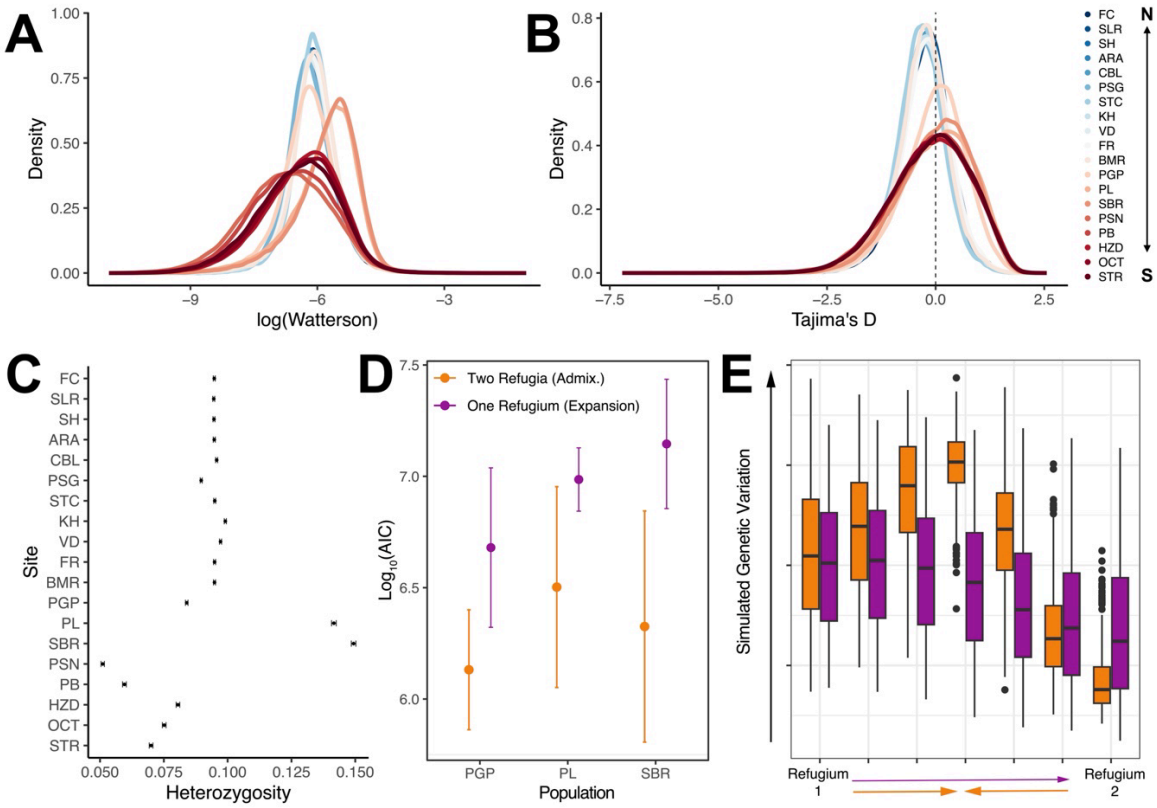

Figure 3. A) Nucleotide diversity ($Log_{10}(\pi)$), and B) Tajima's D across the 19 *Nucella canaliculata* populations. Calculations were performed on a sliding window with a window size of 25kb. The dashed vertical line for Tajima's D indicates 0, the expected value under a neutral model. The colors match the 19 populations indicated in Table S1 and are ordered from north to south. C) Within-population block-jackknife estimate ± standard error of heterozygosity from `poolfstat`. Populations are ordered north to south from top to bottom. D) Distribution of AIC values for the two demographic models tested using `moments`. The mean ± standard deviation of AIC among all model replicates is shown (the best model is reported in the text). E) Hypothetical levels of genetic variation derived from simulating the two proposed demographic scenarios using `SLiM`.

283

284 ***Genomic Evidence of Post-Glacial Demography***

285 We performed coalescent demographic inference with `moments` [54] to compare
286 whether our data matched either of two phylogeographic models: (1) a single southern refugium
287 expanding northward, and (2) a secondary contact (admixture) between two glacial refugia.
288 These models were tested on three central California populations (PGP, PL, and SBR) that
289 showed potential admixture signals in the genomic diversity analyses (Figure 3A). Our results
290 indicate that, for all three populations, the admixture model provides a better fit based on AIC
291 values (Figure 3D). The ΔAIC (Log_{10} transformed) values for the admixture model, relative to
292 the expansion model, were: PGP = 5.40; PL = 6.62; and SBR = 6.83.

293 For PGP, the best-fitting model estimated an admixture event around 10,500 years ago,
294 with a very low southern ancestry contribution (1%), indicating that most of its genetic ancestry
295 derives from northern populations. This aligns with its position as the southern boundary of the
296 hypothetical northern refugia. In contrast, PL and SBR exhibit strong signatures of admixture
297 between northern and southern populations. For PL, the best model estimated an admixture event
298 approximately 10,231 years ago, with 57% of its ancestry derived from southern populations.
299 Similarly, for SBR, admixture time was estimated at 7,451 years ago, with 59% southern
300 ancestry. These results suggest that PL and SBR are situated at the "core" of the admixture zone
301 between the hypothetical northern and southern glacial refugia.

302 To further investigate these patterns, we conducted forward genetic simulations using
303 `SLiM` [58] to compare the expected levels of genetic variation under our proposed demographic
304 scenarios. Among the two simulated models, the "two refugia" scenario produced patterns of
305 genetic variation most consistent with our empirical data. In this model, the core admixture

306 populations, i.e., SBR and PL, exhibited the highest levels of genetic variation (Figure 3E),
307 mirroring the observed data (Figure 3A). Additionally, in the simulations, genetic variation
308 decreased monotonically away from the admixture zone at a rate proportional to the simulated
309 population size of each refugium (in our simulations, refugium 2 had a smaller population size
310 relative to refugium 1). In contrast, the one refugium model produced a different pattern, in
311 which genetic variation steadily declined with range expansion, lacking the pronounced peaks in
312 diversity seen in the admixed populations.

313 ***Shell Geometric Morphometrics Show Patterns of Spatial Differentiation***

314 Given the strong signature of genetic divergence among populations, we explored
315 whether shell morphology displayed a similar pattern of divergence. We performed landmark
316 analysis on images of the ventral surface of the dogwhelks (Figure 4A). PC1 (21.60% of the
317 variance) and PC2 (19.65%) explained most of the shape variation and geographically separated
318 the populations based on shell shape (Figure 4B; Figure S4). PC1 described a shell shape with a
319 lack of elongation of the spire and a rotundness of the main body whorl and aperture (wireframe
320 image for PC1 on Figure 4B). PC2 described shape changes in the widest point of the outer lip
321 and the angle of the siphonal canal. Dogwhelks from populations south of Monterey Bay occupy
322 a greater amount of PC space than the northern dogwhelks despite there being fewer populations
323 analyzed (Figure 4B). When comparing pairwise Procrustes distances, the greatest divergence in
324 shape was between the populations south of Monterey Bay (except HZD) with those from
325 northern California and southern Oregon (dark blue/teal section in the top right of Figure 4C).
326 Size (measured as centroid size) and population explained 5.5% and 21.6% of the dogwhelk
327 shape variation (Figure 4D; Table S7). Similarly, size, latitude and their interaction all explained
328 a significant amount of the variance in dogwhelk shape (Figure 4D; Table S7).

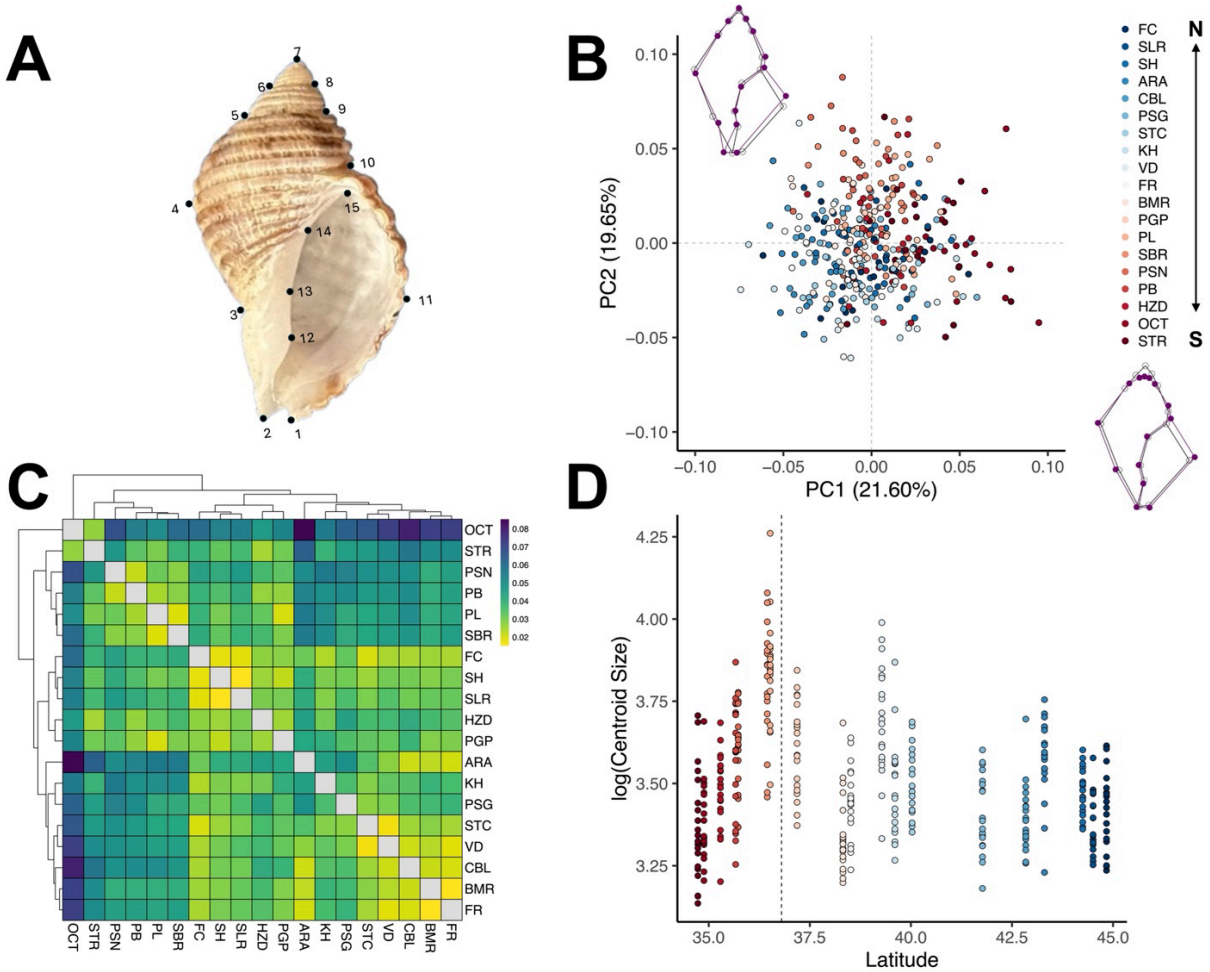

Figure 4. A) Diagram showing the position of the 15 landmarks used in the geometric morphometric analysis of *Nucella canaliculata* shells from 19 populations. B) Principal component 1 and 2 of the Procrustes coordinates. Wireframes display the shape variation for each PC (purple) in relation to the mean shape (black). Dogwhelks are colored by population. See Table S1 for an explanation of population codes. C) Heatmap of Procrustes distances among the 19 populations. D) The centroid size (log transformed) as a function of latitude. The dotted line indicates the location of Monterey Bay. Colors indicate the 19 populations as specified in B.

A Shell Matrix Protein Emerges as a Genomic Outlier of Local Adaptation

We used `BayPass` to identify outlier SNPs associated with morphological variation. Seven Bonferroni corrected outlier SNPs were identified associated with the coefficient of variation for PC1 (Figure 5A; Table S8), and six outlier SNPs were identified associated with the coefficient of variation for PC2 (Figure 5B; Table S8). Four candidate loci were in the same protein (gene g26813), with one SNP (position 41601) identified in both analyses (Table S8).

Most of the populations south of Monterey Bay were fixed at these four loci, while the populations north of Monterey Bay were polymorphic with variable allele frequencies (Figure 5C). Three of the loci show strong associations with morphology — when the allele frequency of reference allele approaches 0 the morphometric space becomes negative (Figure S5). To identify the function of this protein of interest, we ran `blastp`. This protein has been identified as a shell matrix protein in multiple gastropod species including the zebra mussel, *Dreissena polymorpha* (UniProt entry: A0A9D4NEM1; percent identity 26.2%, positive matches = 52.4%, E-value = 0.0002) [63], and the lettuce slug, *Elysia crispata* (UniProt entry: A0AAE1DT89; percent identity 24.3%, positive matches = 45.0%, E-value = 0.00044) [64].

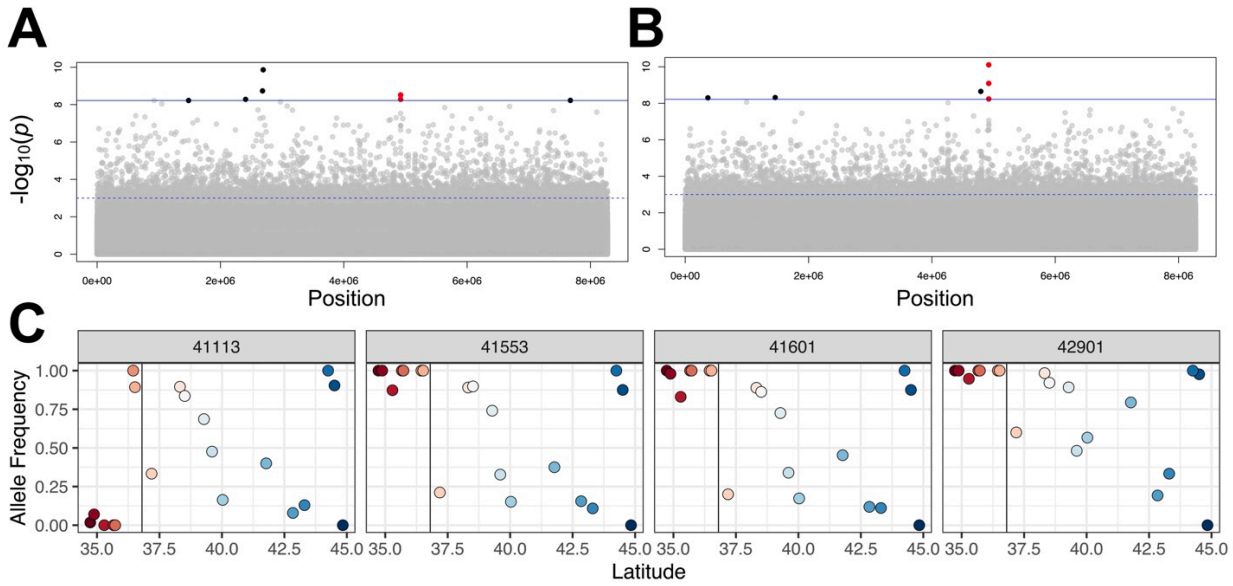

Figure 5. Signatures of selection from `BayPass` for the coefficient of variation A) for PC1 and B) for PC2 of the *Nucella canaliculata* shell morphology. The solid blue line indicates Bonferroni corrected outliers and the dotted blue line indicates significance of SNPs at a value of 0.001. Bonferroni corrected outliers are identified in black, with the four candidate loci on gene g26813 in red. C) Allele frequencies of the four candidate loci on gene g26813 as a function of latitude. The solid black line indicates the location of Monterey Bay.

Discussion:

362 The present-day genomic diversity and population structure of a species reflects a
363 combination of contemporary and historical processes [5]. In contrast to most marine species
364 which have high population connectivity due to planktonic larvae, the genomics of marine taxa
365 with direct development and low dispersal provide an especially powerful opportunity to analyze
366 deeper evolutionary histories and the impacts of historical events [19–21]. Our pooled
367 sequencing of 19 populations of the low-dispersing channeled dogwhelk, *Nucella canaliculata*,
368 across ~1,500 km of the west coast of North America indicated a high level of population
369 structure, with a phylogeographic split at Monterey Bay, and a pattern of genomic diversity that
370 revealed a history of being structured by past geological events. The contemporary genomic
371 patterns and accompanying computer simulations suggest that during the last glacial maximum
372 there were at least two refugial populations, north and south of Monterey Bay, that experienced
373 post-glacial expansions and admixture. This biogeographic pattern of divergence was further
374 supported by spatial variation in shell morphology. Additionally, we identified putatively
375 adaptive outlier loci that may be the molecular drivers of variation in shell morphology. Our
376 findings have important management implications; if the range of this species shifts poleward, as
377 predicted with climate change, there is risk of losing the unique genetic diversity that is primarily
378 maintained in the populations in the southern portion of the species' range limit [20,27,65].

379 ***Phylogeography and Glacial Refugia***

380 Pleistocene glaciation drove changes in species distributions, with most temperate marine
381 species in the northern hemisphere retreating southward to escape glacial ice sheets [24,25].
382 Geological assemblages from Pleistocene terraces along the west coast of North America
383 indicate that many species, including *N. canaliculata*, shifted their ranges due to climatic
384 fluctuations in accord with shifting habitat conditions [24–26]. This historical event is assumed

385 to have left a signature on the current genetic composition of populations. Genetic models
386 suggest high genetic diversity within areas of glacial refugia and high dissimilarity between
387 refugial populations, and low genetic diversity in the area of subsequent population expansion
388 [3]. In marine environments, particularly the Northeastern Pacific, there has been mixed
389 evidence that this historical event has left imprints on the genetics of extant marine populations
390 [13,15,19,27]. The high dispersal potential present in most marine species [6] and the importance
391 of oceanographic processes and vicariant barriers in structuring genetic diversity typically
392 overwhelms any remnant signature of Pleistocene glaciation [15,66]. Additionally, genomic
393 advancements in SNP based approaches are increasing our phylogeographic resolution and
394 power to detect the persistent influence of glaciation [12,27].

395 The present-day phylogeographic structure of *N. canaliculata* showed clear evidence of
396 glacial refugia, followed by population expansion and admixture (Figure 2; Figure 3). Our results
397 revealed two genetically distinct clusters with Monterey Bay being the phylogeographic
398 breakpoint (Figure 1). All populations north of Monterey Bay were genetically similar and have
399 a negative Tajima's D value, indicating population expansion (Figure 3B). In contrast, the
400 populations between Monterey Bay and the species southern range limit of Point Conception
401 contained variable levels of nucleotide diversity and were highly diverged from each other
402 despite their spatial proximity (Figure 2B; Figure 3A). These patterns closely match those
403 predicted from genetic models of Pleistocene glaciation [3,7]. Additionally, the populations
404 nearest Monterey Bay show a unique signature of higher nucleotide diversity and heterozygosity
405 compared to the populations to the north and south, suggesting potential admixture in this region
406 (Figure 3).

407 Our `moments` analyses and `SLiM` simulation results provided greater clarity to this
408 glacial pattern as well as unique insights about the fine scale demography of the species,
409 particularly about the evolutionary histories of the populations nearest Monterey Bay. These
410 analyses suggest that at least one northern population refugium persisted through the cycles of
411 glaciation, rather than the entire contemporary species being derived from southerly populations
412 (Figure 3D; Figure 3E). These refugial populations expanded post-glacially and experienced a
413 secondary contact near Monterey Bay. Further, the admixture proportions show that the two
414 populations just south of Monterey Bay (Point Lobos [PL] and Soberanes Point [SBR]) are the
415 'true' admixed zone, while the region just north of Monterey Bay (i.e., Pigeon Point [PGP]) is
416 the likely southern edge of the northern refugia. The drastic drop in admixture proportions
417 between Point Lobos and Pigeon Point, in the context of the F_{ST} findings (Figure 2), suggests
418 that admixture was relatively recent, and that migration is low. An important caveat of our
419 analysis is that several key genetic parameters for *Nucella* remain uncharacterized; thus, we
420 relied on published rates from the model system *Drosophila*. Consequently, parameters related to
421 population divergence and admixture timing should be considered approximations and
422 interpreted with caution.
423 Genetic variation in the northern populations indicates that the phylogeographic history
424 of *N. canaliculata* may be more complex than suggested by our models of glacial refugia. Under
425 a simple scenario of a "large" southern refugia and a "small" (or remanent) northern refugia, we
426 should expect to see a sharp decline in the levels of standing genetic variation towards the
427 northern edge (Figure 3D). Yet, in our real data, we see a peak in variation at Soberanes Point
428 (SBR) and Point Lobos (PL), with variable variation in the southern populations and stable levels
429 of genetic variation in the northern populations that are, on average, higher than the southern

430 sites (Figure 3A). One possible hypothesis to explain this departure from expectations is that the
431 southern refugia was smaller than the northern refugia; this would be unusual, but not impossible
432 (see [4]). For example, strong competition with other snails in the refugia or limitations in prey
433 availability could have kept populations small [67]. An alternative hypothesis could be that there
434 were multiple northern refugia that experienced sequential mixing post-glaciation, giving rise to
435 the elevated variation. To clarify these patterns and determine the exact location and size of the
436 northern refugia, we would need thorough sampling of populations near the species northern
437 range limit, as well as ancient samples.

438 Understanding the present-day phylogeographic structure of a species is particularly
439 important as species are being forced to either rapidly adapt or shift their ranges in response to
440 anthropogenic climate change [9–11]. Most of the unique genetic diversity of *N. canaliculata*
441 (i.e., suggested by high F_{ST}) was in the populations south of Monterey Bay. Given that
442 populations at the warmer, trailing edge of a range are more likely to go extinct in response to
443 climate change, this could mean losing large amounts of endemic diversity [20,27], particularly
444 for species with low dispersal. In accord, a recent study that surveyed several coastal marine
445 species, including *N. canaliculata*, across their ranges in the northeast Pacific found that
446 populations near their equatorward edge are demographically vulnerable to climate change and
447 environmental stress [68]. Population density of *N. canaliculata* declined latitudinally from
448 Oregon to central California in association with increasing aerial and water temperature stress,
449 and a reduction in rocky habitat availability [68]. Small populations are highly vulnerable to
450 local extirpation due to Allee effects and stochastic events [69]. There is accumulating evidence
451 that marine heatwaves and other oceanic extreme events are becoming increasingly common

452 [70]. Thus, there is substantial risk that the genetically unique, but demographically small,
453 populations of *N. canaliculata* near their trailing edge could face local extirpation.

454 ***Adaptive consequences of low dispersal potential***

455 The phylogeographic pattern of defined population structure and low levels of genetic
456 diversity that we observed has likely impacted the adaptive potential and evolutionary trajectory
457 of the *N. canaliculata* populations. Previous research studying intraspecific variation in
458 predatory drilling ability in *N. canaliculata* showed that the species displays a pattern of local
459 adaptation [30–33]. However, rather than a smooth latitudinal gradient, the adaptive landscape
460 appears to be a spatial mosaic, with specific sites deviating from regional patterns [32,33].
461 Despite their geographic separation, the two populations with the greatest drilling ability are
462 Soberanes Point (SBR) and Van Damme (VD). Our genomic results provide important
463 demographic context to this phenotypic pattern and suggest that these two populations have
464 strikingly different genetic backgrounds, with Soberanes Point (SBR) being an admixed
465 population and Van Damme (VD) being in the northern genetic cluster. Thus, this phenotypic
466 pattern of parallel adaptation could be due to the populations evolving strong drilling using
467 independent molecular pathways either via the same or different morphological mechanism. This
468 hypothesis is further supported by the fact that these two populations display very different
469 behavioral patterns when exposed to predator cues [34]. Further research that links genotype to
470 phenotype would be needed to elucidate the molecular mechanisms behind this adaptive
471 landscape.

472 For many species, linking phenotype to genotype and understanding patterns of
473 adaptation are significantly hindered by the lack of genomic resources for non-model organisms
474 [7,8]. The genome assembly presented here is the first for the species, and the second assembly

475 for the ecologically important genus *Nucella* [71]. Extraction of high-quality molluscan DNA is
476 a challenge due to the presence of mucopolysaccharides and other secondary metabolites [72].
477 Additionally, molluscan genomes are large, have a high proportion of repetitive regions, and
478 typically harbor high heterozygosity, thus genome assembly is computationally difficult and
479 relatively few genomes have been assembled, despite there being more than 100,000 living
480 molluscs and incredibly high morphological diversity [73,74]. Like previously published
481 molluscan genomes [71,73], the *N. canaliculata* assembly presented here was very large and
482 consisted of numerous scaffolds. Despite this, the assembly was smaller than the estimated
483 genome size for the species of 2.54pg±0.057 [75], a pattern not uncommon in molluscan
484 genomes [74].

485 ***Shell Morphology and Species Taxonomy***

486 The study of gastropod shell morphology predates Linnaean taxonomy [37,38]. Given
487 their ubiquity and longevity, the calcareous shells of gastropods have long been used to infer
488 taxonomy and evolutionary patterns [76]. Despite this, few modern phylogeographic studies of
489 marine gastropods incorporate morphological data alongside genetic data (but see [19,77]). We
490 found evidence that *N. canaliculata* dogwhelks show spatial variation in shell morphology, with
491 the greatest difference between populations south versus north of Monterey Bay (Figure 4C),
492 consistent with patterns of geographic divergence suggested by the genomic data. Dogwhelks
493 from southern populations were typically more rotund and had a shorter spire, compared to
494 dogwhelks from northern populations (Figure 4B). Additionally, the size of the dogwhelks varied
495 latitudinally with the largest dogwhelks present at Soberanes Point (SBR) and Point Lobos (PL)
496 (Figure 4D). Shell morphology in *Nucella* is a result of both phenotypic plasticity [34] and
497 adaptive evolution [78]. Consequently, the morphometric data (Figure 4) showed a less stark

498 pattern of divergence among the populations than the genomic data (Figure 2). Notwithstanding,
499 the concordance of our datasets increases our confidence in the biogeographic pattern.

500 Our genomic and morphometric results both indicate a strong spatial signal of divergence
501 with a breakpoint at Monterey Bay. This biogeographic pattern in *N. canaliculata* matches an
502 overlooked taxonomic hypothesis, based on shell morphology, that there may be a northern and
503 southern subspecies divided at Monterey Bay [79]. In this context, what had previously been
504 hypothesized as two morphological subspecies, are according to our data, two
505 phylogeographically distinct genetic clusters whose morphological differences reflect divergent
506 histories of demography and selection. Future research with more intensive sampling,
507 particularly around Monterey Bay and at the species northern range limit, as well as individual
508 sequencing is necessary to resolve the taxonomy of the species.

509 To test for signatures of selection and begin to link genotype to phenotype, we ran
510 `BayPass` on the coefficient of variation for PC1 and PC2 of the shell morphology dataset. This
511 framework allowed us to identify outlier loci while controlling for population structure. We
512 identified a handful of SNPs associated with both coefficients of variation (Figure 5). Notably,
513 one protein stood out in both analyses. The four candidate loci on this protein display a spatial
514 pattern such that they are fixed in the populations south of Monterey Bay and polymorphic in the
515 populations north of Monterey Bay (Figure 5C). This protein has been identified as a shell matrix
516 protein in other gastropod species, including the zebra mussel, *Dreissena polymorpha* [63], and
517 the lettuce slug, *Elysia crispata* [64]. We hypothesize that this protein is likely a candidate for
518 morphological variation and local adaptation and should be studied further.

519 Overall, this research highlights the power of genomic approaches to advance our
520 understanding of how historical events have shaped contemporary demography in coastal

521 species. The geographic distribution of genomic diversity in *Nucella canaliculata* reveals a
522 persistent signature of glacial events in the northeast Pacific, with important implications for
523 adaptive capacity and conservation in an era of rapid global change.
524
525 **Data Accessibility Statement:** The *Nucella canaliculata* genome assembly can be found at
526 BioProject accession number PRJNA1282036 and the raw reads for the assembly can be found
527 under the BioProject accession number PRJNA1282048. Raw pooled-sequencing reads can be
528 found under the SRA BioProject accession number PRJNA1276871, and the individual
529 BioSample accession numbers can be found in Table S1. A GitHub repository with the code for
530 analyzing the data can be found at https://github.com/emily-
531 longman/Nucella_can_Pop_Genomics.
532 **Ethics:** Collections in California were performed under CA Fish & Wildlife Scientific Collecting
533 Permits S-191200004-19122-001 and S-190980002-23226-001, and specific use amendments S-
534 190980003-23226-001-01 and S-190980003-23226-001-02, and CA State Parks permit 24-820-
535 015 and amendment 24-820-015A. Collections in Oregon were performed under the taking
536 permit 27976.
537 **Author's contributions:** E.K.L.: conceptualization, data curation, formal analysis, funding
538 acquisition, investigation, methodology, project administration, resources, software, supervision,
539 validation, visualization, writing – original draft, writing – review & editing; J.C.B.N.: formal
540 analysis, software, visualization, writing – review & editing; E.S.: funding acquisition,
541 investigation, writing – review & editing; M.P.: conceptualization, methodology, resources,
542 supervision, writing – review & editing. All authors gave final approval for publication and
543 agreed to be held accountable for the work performed therein.

Funding: This research was supported by an NSF postdoctoral fellowship (OCE-2307933) awarded to E.K.L. and the National Science Foundation grant OCE-1851462 awarded to E.S.

Acknowledgements: We thank the Vandenberg Space Force Base, the UC Reserve System and the Abbott family for access to field sites. We thank N. Longman and J. Fosnight for assistance with the field collections, and R. A. Bay, B. B. Cameron, and E. S. Nielsen for help with optimizing the extraction protocol and discussions about sequencing approaches for the genome.

References:

[1] Hughes, A.R., Inouye, B.D., Johnson, M.T., Underwood, N., & Vellend, M. (2008). Ecological consequences of genetic diversity. *Ecology Letters*, *11*(6), 609–623. (https://doi.org/10.1111/j.1461-0248.2008.01179.x)

[2] Saccheri, I., & Hanski, I. (2006). Natural selection and population dynamics. *Trends in Ecology & Evolution*, *21*(6), 341–347. (https://doi.org/10.1016/j.tree.2006.03.018)

[3] Hewitt, G.M. (1996). Some genetic consequences of ice ages, and their role in divergence and speciation. *Biological journal of the Linnean Society*, *58*(3), 247–276. (https://doi.org/10.1111/j.1095-8312.1996.tb01434.x)

[4] Maggs, C.A., Castilho, R., Foltz, D., Henzler, C., Jolly, M.T., Kelly, J., Olsen, J., Perez, K.E., Stam, W., Väinölä, R., et al. (2008). Evaluating signatures of glacial refugia for North Atlantic benthic marine taxa. *Ecology, 89*(sp11), S108–S122. (https://doi.org/10.1890/08-0257.1)

[5] Avise, J.C. (2000). *Phylogeography: the history and formation of species*. Harvard University Press. (https://doi.org/10.2307/j.ctv1nzfgj7)

[6] Thorson, G. (1950). Reproductive and larval ecology of marine bottom invertebrates. *Biological Reviews of the Cambridge Philosophical Society, 25*, 1–45. (https://doi.org/10.1111/j.1469-185X.1950.tb00585.x)

[7] Savolainen, O., Lascoux, M., & Merilä, J. (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*, *14*(11), 807–820. (https://doi.org/10.1038/nrg3522)

[8] Da Fonesca, R.R. Albrechtsen, A., Themudo, G.E., Ramos-Madrigal, J., Sibbesen, J.A., Maretty, L., Zepeda-Mendoza, M.L., Campos, P.F., Heller, R., & Pereira, R.J. (2016). Next generation biology: sequencing and data analysis approaches for non-model organisms. *Marine genomics, 30*, 3–13. (https://doi.org/10.1016/j.margen.2016.04.012)

[9] Waldvogel, A.M., Feldmeyer, B., Rolshausen, G., Exposito-Alonso, M., Rellstab, C., Kofler, R., Mock, T., Schmid, K., Schmitt, I., Bataillon, T., et al. (2020). Evolutionary genomics can improve prediction of species' responses to climate change. *Evolution Letters*, *4*(1), 4–18. (https://doi.org/10.1002/evl3.154)

[10] Urban, M.C., Swaegers, J., Stoks, R., Snook, R.R., Otto, S.P., Noble, D.W., Moiron, M., Hällfors, M.H., Gómez-Llano, M., Fior, S., et al. (2024). When and how can we predict adaptive responses to climate change? *Evolution Letters*, *8*(1), 172–187. (https://doi.org/10.1093/evlett/qrad038)

[11] Bernatchez, L., Ferchaud, A.L., Berger, C.S., Venney, C.J., & Xuereb, A. (2024). Genomics for monitoring and understanding species responses to global climate change. *Nature Reviews Genetics*, *25*(3), 165–183. (https://doi.org/10.1038/s41576-023-00657-y)

[12] Crawford, D.L., & Oleksiak, M.F. (2016). Ecological population genomics in the marine environment. *Briefings in Gunctional Genomics*, *15*(5), 342–351. (https://doi.org/10.1093/bfgp/elw008)

[13] Pelc, R.A., Warner, R.R., & Gaines, S.D. (2009). Geographical patterns of genetic structure in marine species with contrasting life histories. *Journal of Biogeography*, 36(10), 1881–1890. (https://doi.org/10.1111/j.1365-2699.2009.02138.x)

[14] Pascual, M., Rives, B., Schunter, C., & Macpherson, E. (2017). Impact of life history traits on gene flow: A multispecies systematic review across oceanographic barriers in the Mediterranean Sea. *PLoS One*, *12*(5), e0176419. (https://doi.org/10.1371/journal.pone.0176419)

[15] Kelly, R.P., & Palumbi, S.R. (2010). Genetic structure among 50 species of the northeastern Pacific rocky intertidal community. *PloS one*, *5*(1), e8594. (https://doi.org/10.1371/journal.pone.0008594)

[16] Grosberg R.K. & C.W. Cunningham. (2001). Genetic structure in the sea: from populations to communities. In Marine Community Ecology, ed. M.D. Bertness, S.D. Gaines, M.E. Hay, pp. 61–84. Sunderland, MA: Sinauer

[17] Riginos, C., Douglas, K.E., Jin, Y., Shanahan, D.F., & Treml, E.A. (2011). Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, *34*(4), 566–575. (https://doi.org/10.1111/j.1600-0587.2010.06511.x)

[18] Arranz, V., Fewster, R.M., & Lavery, S.D. (2024). Concordance and drivers of marine spatial structure determined using genogeographic clustering. *Journal of Biogeography*, *51*(10), 2034–2046. (https://doi.org/10.1111/jbi.14967)

[19] Hellberg, M.E., Balch, D.P., & Roy, K. (2001). Climate-driven range expansion and morphological evolution in a marine gastropod. *Science, 292*(5522), 1707–1710. (https://doi.org/10.1126/science.1060102)

[20] Fenberg, P.B., Posbic, K., & Hellberg, M.E. (2014). Historical and recent processes shaping the geographic range of a rocky intertidal gastropod: phylogeography, ecology, and habitat availability. *Ecology and Evolution*, *4*(16), 3244–3255. (https://doi.org/10.1002/ece3.1181)

[21] Jacobs, D.K., Haney, T.A., & Louie, K.D. (2004). Genes, diversity, and geologic process on the Pacific coast. *Annual Review of Earth and Planetary Sciences*, *32*(1), 601–652. (https://doi.org/10.1146/annurev.earth.32.092203.122436)

[22] McGaughran, A., Liggins, L., Marske, K.A., Dawson, M.N., Schiebelhut, L.M., Lavery, S.D., Knowles, LL., Moritz, C. & Riginos, C. (2022). Comparative phylogeography in the genomic age: Opportunities and challenges. *Journal of Biogeography*, *49*(12), 2130–2144. (https://doi.org/10.1111/jbi.14481)

[23] Bringloe, T.T., Fort, A., Inaba, M., Sulpice, R., Ghriofa, C.N., Mols-Mortensen, A., Filbee-Dexter, K., Vieira, C., Kawai, H., Hanyuda, T., et al. (2022). Whole genome population structure of North Atlantic kelp confirms high-latitude glacial refugia. *Molecular Ecology*, *31*(24), 6473–6488. (https://doi.org/10.1111/mec.16714)

[24] Addicott, W.O. (1966). *Late Pleistocene marine paleoecology and zoogeography in central California*. US Government Printing Office. (https://doi.org/10.3133/pp523C)

[25] Fields, P.A., Graham, J.B., Rosenblatt, R.H. & Somero, G.N. (1993). Effects of expected global climate change on marine faunas. *Trends in Ecology & Evolution, 8*(10), 361–367. (https://doi.org/10.1016/0169-5347(93)90220-J)

[26] Roy, K., Valentine, J.W., Jablonski, D., & Kidwell, S.M. (1996). Scales of climatic variability and time averaging in Pleistocene biotas: implications for ecology and

633 evolution. *Trends in ecology & evolution*, *11*(11), 458–463.
634 (https://doi.org/10.1016/0169-5347(96)10054-9)
635 [27] Nielsen, E.S., Walkes, S., Sones, J.L., Fenberg, P.B., Paz-García, D.A., Cameron, B.B.,
636 Grosberg, R.K., Sanford, E. & Bay, R.A. (2024). Pushed waves, trailing edges, and
637 extreme events: Eco-evolutionary dynamics of a geographic range shift in the owl limpet,
638 *Lottia gigantea*. *Global Change Biology*, *30*(7), e17414.
639 (https://doi.org/10.1111/gcb.17414)
640 [28] Dawson, M.N. (2001). Phylogeography in coastal marine animals: a solution from
641 California? *Journal of Biogeography, 28*(6), 723–736. (https://doi.org/10.1046/j.1365-
642 2699.2001.00572.x)
643 [29] Marko, P.B., Moran, A.L., Kolotuchina, N.K., & Zaslavskaya, N.I. (2014). Phylogenetics of
644 the gastropod genus Nucella (Neogastropoda: Muricidae): species identities, timing of
645 diversification and correlated patterns of life-history evolution. *Journal of Molluscan*
646 *Studies, 80*(4), 341–353. (https://doi.org/10.1093/mollus/eyu024)
647 [30] Sanford, E., Roth, M.S., Johns, G.C., Wares, J.P., & Somero, G.N. (2003). Local selection
648 and latitudinal variation in a marine predator-prey interactions. *Science, 300*(5622),
649 1135–1137. (https://doi.org/10.1126/science.1083437)
650 [31] Sanford, E. & Worth, D.J. (2010). Local adaptation along a continuous coastline: prey
651 recruitment drives differentiation of a predatory snail. *Ecology, 91*(3), 891–901.
652 (https://doi.org/10.1890/09-0536.1)
653 [32] Sanford, E. & Worth, D.J. (2009). Genetic differences among populations of a marine snail
654 drive geographic variation in predation. *Ecology, 90*(11), 3108–3118.
655 (https://doi.org/10.1890/08-2055.1)

[33] Longman, E.K. & Sanford, E. (2025). Biogeographic variation in mussel shell thickness and drilling predation on rocky shores. *Oecologia, 207*, 126. (https://doi.org/10.1007/s00442-025-05760-x)

[34] Neylan, I.P., Longman, E.K., Sanford, E., Stachowicz, J.J., & Sih, A. (2024). Long-term anti-predator learning and memory differ across populations and sexes in an intertidal snail. *Proceedings B, 291*(2032), 20240944. (https://doi.org/10.1098/rspb.2024.0944)

[35] Kuo, E.S., & Sanford, E. (2009). Geographic variation in the upper thermal limits of an intertidal snail: implications for climate envelope models. *Marine Ecology Progress Series*, *388*, 137–146. (https://doi.org/10.3354/meps08102)

[36] Schlötterer, C., Tobler, R., Kofler, R., & Nolte, V. (2014). Sequencing pools of individuals—mining genome-wide polymorphism data without big funding. *Nature Reviews Genetics*, *15*(11), 749–763. (https://doi.org/10.1038/nrg3803)

[37] Fretter, V. & Graham, A. (1962). British Prosobranch Molluscs; Their functional anatomy and ecology. London: Ray Soc. 755 pp.

[38] Bieler, R. (1992). Gastropod phylogeny and systematics. *Annual Review of Ecology and Systematics*, 311–338.

[39] Ye, C., Hill, C. M., Wu, S., Ruan, J., & Ma, Z. (2016). DBG2OLC: efficient assembly of large genomes using long erroneous reads of the third generation sequencing technologies. *Scientific Reports*, 6(1), 31900. (https://doi.org/10.1038/srep31900)

[40] Gurevich, A., Saveliev, V., Vyahhi, N., & Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics, 29*(8), 1072–1075. (https://doi.org/10.1093/bioinformatics/btt086)

[41] Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A., Zdobnov, E.M. (2021). BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution, 38*(10), 4647–4654. (https://doi.org/10.1093/molbev/msab199)

[42] Smit, A.F.A., Hubley, R., Green P. (2013-2015). RepeatMasker Open-4.0. http://www.repeatmasker.org

[43] Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., & Morgenstern, B. (2006). AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Research, 34*, W435–W439. (https://doi.org/10.1093/nar/gkl200)

[44] Andrews, S. (2010). *FastQC: a quality control tool for high throughput sequence data.* http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

[45] Chen, S., Zhou, Y., Chen, Y., & Gu, J. (2018). Fastp: An ultra-fast all-in-one FASTQ processor. *Bioinformatics, 34*(17), i884–i890. (https://doi.org/10.1093/bioinformatics/bty560)

[46] Vasimuddin, M. Sanchit, M., Henge, L., Srinivas A. (2019). Efficient architecture-aware acceleration of BWA-MEM for multicore systems. *IEEE Parallel and Distributed Processing Symposiums (IPDPS)*. (https://doi:10.1109/IPDPS.2019.00041)

[47] Li, J., 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)

[48] Broad Institute. (2019). Picard Toolkit. GitHub Repository. https://broadinstitute.github.io/picard/

[49] Garrison, E., & Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. *arXiv*, 1207.3907. (https://doi.org/10.48550/arXiv.1207.3907)

[50] Gautier, M., Vitalis, R., Flori, L., & Estoup, A. (2022). f-Statistic estimation and admixture graph construction with Pool-seq or allele count data using the R package poolfstat. *Molecular Ecology Resources, 22*, 1394–1416. (https://doi.org/10.1111/1755-0998.13557)

[51] Gautier, M. (2015). Genome-wide scan for adaptive divergence and association with population-specific covariates. *Genetics, 201*(4), 1555–1579. (https://doi.org/10.1534/genetics.115.181453)

[52] Paradis, E., & Schliep, K. (2019). Ape 5.0: an environment for modern phylogenetics and evolutionary analyses in R. *Bioinformatics, 35*, 526–528. (https://doi.org/10.1093/bioinformatics/bty633)

[53] Ferretti, L., Ramos-Onsins, S.E., & Pérez-Enciso, M. (2013). Population genomics from pool sequencing. *Molecular ecology*, *22*(22), 5561–5576. (https://doi.org/10.1111/mec.12522)

[54] Jouganous, J., Long, W., Ragsdale, A.P., & Gravel, S. (2017). Inferring the joint demographic history of multiple populations: beyond the diffusion approximation. *Genetics*, *206*(3), 1549–1567. (https://doi.org/10.1534/genetics.117.200493)

[54] Kapun, M., Nunez, J. C., Bogaerts-Márquez, M., Murga-Moreno, J., Paris, M., Outten, J., Coronado-Zamora, M., Tern, C., Rota-Stabelli, O., García Guerreiro, M.P., et al. (2021). Drosophila evolution over space and time (DEST): a new population genomics

723 resource. *Molecular biology and evolution*, *38*(12), 5782–5805.
724 (https://doi.org/10.1093/molbev/msab259)
725 [56] Nunez, J.C.B., Coronado-Zamora, M., Gautier, M., Kapun, M., Steindl, S., Ometto, L.,
726 Hoedjes, K.M., Beets, J., Wiberg, R.A.W., Mazzeo, G.R., et al. (*in press*) Footprints of
727 worldwide adaptation in structured populations of *D. melanogaster* through the expanded
728 DEST 2.0 genomic resource. *Molecular biology and evolution.*
729 [57] Keightley, P.D., Ness, R.W., Halligan, D.L., & Haddrill, P.R. (2014). Estimation of the
730 spontaneous mutation rate per nucleotide site in a *Drosophila melanogaster* full-sib
731 family. *Genetics, 196*(1), 313–320. (https://doi.org/10.1534/genetics.113.158758)
732 [58] Haller, B.C., & Messer, P.W. (2023). SLiM 4: multispecies eco-evolutionary modeling. *The*
733 *American Naturalist*, *201*(5), E127–E139. (https://doi.org/10.1086/723601)
734 [59] Rohlf, F.J. (2015). The tps series of software. *Hystrix, 26*(1), 9–12.
735 (https://doi.org/10.4404/hystrix-26.1-11264)
736 [60] Adams, D.C., Collyer, M., Kaliontzopoulou, A., & Baken, E. (2025). *Geomorph: Geometric*
737 *morphometric analyses of 2D and 3D landmark data*. R package version 4.0.10.
738 (https://doi.org/10.32614/CRAN.package.geomorph)
739 [61] Klingenberg, C.P. (2011). MorphoJ: an integrated software package for geometric
740 morphometrics. *Molecular ecology resources*, *11*(2), 353-357.
741 (https://doi.org/10.1111/j.1755-0998.2010.02924.x)
742 [62] Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., &
743 Ruden, D.M. (2012). A program for annotating and predicting the effects of single
744 nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster*
745 strain w1118; iso-2; iso-3. *fly*, *6*(2), 80–92. (https://doi.org/10.4161/fly.19695)

[63] McCartney, M.A., Auch, B., Kono, T., Mallez, S., Zhang, Y., Obille, A., Becker, A., Abrahante, J.E., Garbe, J., Badalamenti, J.P. et al. (2022). The genome of the zebra mussel, *Dreissena polymorpha*: a resource for comparative genomics, invasion genetics, and biocontrol. *G3*, *12*(2), jkab423. (https://doi.org/10.1093/g3journal/jkab423)

[64] Eastman, K.E., Pendleton, A.L., Shaikh, M.A., Suttiyut, T., Ogas, R., Tomko, P., Gavelis, G., Widhalm, J.R. & Wisecaver, J.H. (2023). A reference genome for the long-term kleptoplast-retaining sea slug *Elysia crispata* morphotype clarki. *G3: Genes, Genomes, Genetics*, *13*(12), jkad234. (https://doi.org/10.1093/g3journal/jkad234)

[65] Zarzyczny, K.M., Hellberg, M.E., Lugli, E.B., MacLean, M., Paz-García, D.A., Rius, M., Ross, E.G., Treviño Balandra, E.X., Vanstone, J., Williams, S.T. & Fenberg, P.B. (2024). Opposing genetic patterns of range shifting temperate and tropical gastropods in an area undergoing tropicalisation. *Journal of Biogeography*, *51*(2), 246–262. (https://doi.org/10.1111/jbi.14744)

[66] Wares, J.P., & Cunningham, C.W. (2001). Phylogeography and historical ecology of the north Atlantic intertidal. *Evolution*, *55*(12), 2455–2469. (https://doi.org/10.1111/j.0014-3820.2001.tb00760.x)

[67] Spight, T.M., & Emlen, J. (1976). Clutch sizes of two marine snails with a changing food supply. *Ecology*, *57*(6), 1162–1178. (https://doi.org/10.2307/1935042)

[68] Gravem, S.A., Bachhuber, S., Bignami, S., Chiachi, A.E., Field, L.C., Gaddam, R.N., Raimondi, P.T., & Menge, B.A. (2025). Biogeographic Patterns in Density, Recruitment, Body Size and Zonation of Rocky Intertidal Predators Suggest Increased Population Vulnerability Near Southern Range Limits. *Journal of Biogeography*, *52*(2), 257–273. (https://doi.org/10.1111/jbi.15029)

[69] Stephens, P.A., & Sutherland, W.J. (1999). Consequences of the Allee effect for behaviour, ecology and conservation. *Trends in ecology & evolution*, *14*(10), 401–405. (https://doi.org/10.1016/S0169-5347(99)01684-5)

[70] Smith, K.E., Aubin, M., Burrows, M.T., Filbee-Dexter, K., Hobday, A. J., Holbrook, N.J., King, N.G., Moore, P.J., Sen Gupta, A., Thomsen, M. et al. (2024). Global impacts of marine heatwaves on coastal foundation species. *Nature Communications*, *15*(1), 5052. (https://doi.org/10.1038/s41467-024-49307-9)

[71] Kadiwala, J., Hesketh, A., De Weerd, H., Ritch, H., & Shakur, R. (2025). The first genome Assembly of the Dogwhelk *Nucella lapillus*, a bioindicator species for the marine environment. *Scientific Data*, *12*(1), 1–8. (https://doi.org/10.1038/s41597-025-04764-9)

[72] Adema, C. M. (2021). Sticky problems: extraction of nucleic acids from molluscs. *Philosophical Transactions of the Royal Society B*, *376*(1825), 20200162. (https://doi.org/10.1098/rstb.2020.0162)

[73] Sun, J., Li, R., Chen, C., Sigwart, J.D., & Kocot, K.M. (2021). Benchmarking Oxford Nanopore read assemblers for high-quality molluscan genomes. *Philosophical Transactions of the Royal Society B*, *376*(1825), 20200160. (https://doi.org/10.1098/rstb.2020.0160)

[74] Chen, Z., Baeza, J.A., Chen, C., Gonzalez, M.T., González, V.L., Greve, C., Kocot, K.M., Arbizu, P.M., Moles, J., Schell, T., et al. (2025). A genome-based phylogeny for Mollusca is concordant with fossils and morphology. *Science, 387*(6737), 1001–1007. (https://doi.org/10.1126/science.ads0215)

[75] Pascoe, P.L. Jha, A.N., & Dixon, D.R. (2004). Variation of karyotype composition and genome size in some muricid gastropods from the northern hemisphere. *Journal of Molluscan Studies, 70*(4), 389–398. (https://doi.org/10.1093/mollus/70.4.389)

[76] Vermeij, G. J. (1993). *Evolution and escalation: an ecological history of life*. Princeton University Press. (https://doi.org/10.2307/j.ctv18zhf8b)

[77] Tirado, T., Saura, M., Rolán-Alvarez, E., & Quesada, H. (2016). Historical biogeography of the marine snail Littorina saxatilis inferred from haplotype and shell morphology evolution in NW Spain. *PLoS One*, *11*(8), e0161287. (https://doi.org/10.1371/journal.pone.0161287)

[78] McIntire, L.C., Kota, T., Flagor, T.E., & Bourdeau, P.E. (2021). Weak directional selection on inducible defensive shell traits in a marine gastropod by the inducing shell-breaking predator. *Marine Biology, 168*, 1–12. (https://doi.org/10.1007/s00227-021-03860-9)

[79] McLean, J.H. (2006). Hypothesis for the recognition of *Nucella analoga* (Forbes, 18520 in the northeastern Pacific. *Festivus, 38*, 17–21.