Evidence of Adaptation in Structural Variants among Wild Populations of the purple sea urchin, *Strongylocentrotus purpuratus*

Csenge Petak[†], Daniel E. Sadler[†], Melissa H. Pespeni*, Joaquin C. B. Nunez*

Biology Department, University of Vermont, Vermont, USA

†Equal contribution as Co-first authors
*Equal contribution as Co-senior authors

Corresponding Authors: Daniel E. Sadler; Daniel.Sadler@uvm.edu, Joaquin C.B. Nunez; Joaquin.Nunez@uvm.edu, Melissa H. Pespeni; Melissa.Pespeni@uvm.edu.

Abstract:

Structural variants (SVs) are important components of genetic architecture, expanding beyond traditionally used single nucleotide polymorphisms (SNPs). Though growing, our understanding of the evolutionary forces maintaining SVs in natural populations is limited. Inversions in particular can facilitate local adaptation in populations with high gene flow, including many marine species. The purple sea urchin (*Strongylocentrotus purpuratus*) spans a broad latitudinal range with diverse environmental conditions and is known for its high genetic diversity, making it an ideal system for studying inversion polymorphism, as well as being a crucial foundation species for marine ecosystems. We used low-coverage whole genome sequence data from 140 individuals from seven populations to identify structural variants using a range of methods including local PCA. We integrated Bayesian selection scans to test for local adaptation in putative inversions. We identified nine loci of interest as putative inversions, three of which overlap with areas of the genome associated with local adaptation scans. Additionally, we find functional enrichment within these regions associated with biomineralisation and development. Our results are the first instance of identifying putative inversions in the purple sea urchin and add to the genomic repertoire of model species. This study offers a valuable resource for future research and reinforces the growing evidence that chromosomal inversions represent a fundamental component of standing genetic variation in natural populations, with important implications for the study of local adaptation.

Significance statement:

Chromosomal inversions are an important part of the repertoire of standing genetic variation in wild populations and can facilitate adaptation despite strong gene flow. The purple sea urchin, a widely studied marine species, exhibits high gene flow, large population sizes, and extensive genetic diversity across diverse environmental conditions, making it an ideal model for evolutionary genomics. We identified nine putative inversions, three with signatures of selection, adding echinoderms to the growing list of phyla with putatively adaptive inversions. These findings provide new insights into structural variation in a highly dispersive marine species and highlight potential evolutionary mechanisms maintaining these polymorphisms.

Key words: Structural variants, Inversions, Population genetics, Whole genome sequencing, *Strongylocentrotus purpuratus,* Local adaptation.

INTRODUCTION

A central challenge in evolutionary biology is understanding the interplay between gene flow, which acts to homogenize populations, and natural selection, which acts to promote trait divergence (Maruki et al., 2022; Slatkin, 1987; Yeaman and Whitlock, 2011). Indeed, in classical population genetics theory, gene flow and selection can be viewed as antagonistic forces. Under this framework, adaptive evolution may be inhibited when migration overwhelms natural selection (Felsenstein, 1976), specifically when selection is too weak to preserve advantageous alleles attempting to become spatially structured (Haldane, 1930, *but see* Kottler et al., 2021). The balance between gene flow and selection is often described through the forces of gene swamping and migration load (Bolnick and Nosil, 2007; Lenormand, 2002). Both processes stem from the continual influx of maladapted alleles from other populations, which can obscure signals of loci maintaining local adaptation across complex spatial scales, particularly for traits with polygenic genetic architectures (Yeaman, 2015).

Despite these theoretical expectations, empirical studies in the era of genomics have highlighted that adaptive differentiation, and their underlying genetic basis, can be quantified even among populations connected by high levels of gene flow (i.e., weakly structured; Comeault et al., 2015; Joron et al., 2011; Laurent et al., 2016; Pespeni et al., 2012, 2010). In this context, chromosomal inversions are a prevalent form of genetic architecture that can facilitate the maintenance of local adaptation even in the face of high gene flow (Joron et al., 2011; Kirubakaran et al., 2016; Nunez et al., 2024; Twyford and Friedman, 2015). Adaptive inversions contribute to the maintenance of local adaptation by protecting beneficial mutations from gene swamping (Schaal et al., 2022). Furthermore, through the suppression of recombination between alternative karyotypic classes, inversions can maintain sets of co-adapted alleles in strong linkage, ensuring they will be inherited together and preventing the fitness loss that would otherwise arise in the offspring of locally adapted parents (Charlesworth, 2016; Dobzhansky and Wright, 1943; Kapun et al., 2023; Kapun and Flatt, 2019; Yeaman et al., 2016).

Natural populations of the purple sea urchin, *Strongylocentrotus purpuratus*, are a robust system to assess the role of chromosomal inversions on the interplay of local adaptation and high gene flow. *S. purpuratus* is an inter- and subtidal marine invertebrate that lives in coastal rocky reefs along a broad latitudinal gradient from the cold waters of Alaska to the warmer waters of Baja California, Mexico, presenting an opportunity for local adaptation. *S. purpuratus* are long-lived (> 50 years, Amir et al., 2020), dioecious broadcast spawners, releasing their eggs and sperm into the water column where fertilization occurs. Resulting feeding planktonic larvae can spend weeks to months traveling the coastal currents before recruiting to a suitable place to live. This high dispersal potential can swamp local adaptation but also suggests that selection on dispersing larvae is an important mechanism for limiting phenotype-environment mismatching (Levene, 1953; Marshall et al., 2010; M. H. Pespeni et al., 2013). As adults, *S. purpuratus* also has a disproportionate impact on its habitat as a voracious feeder on macroalgae. Indeed, their overgrazing of biodiverse kelp forests can result in the formation of desolate urchin barrens (Filbee-Dexter and Scheibling, 2014). *S. purpuratus* has also been a model organism in developmental biology for over a century resulting in a high quality, well annotated genome (Sodergren et al., 2006). Previous population genomic studies along the west coast of North America have revealed high levels of genetic diversity,

gene flow, and a lack of population structure across its distribution (Pespeni et al., 2012, 2010; Petak et al., 2023). Yet, despite this demographic structure, previous studies have reported signatures of local adaptation in specific regions of the genome along the species' range (Pespeni et al., 2012; M. H. Pespeni et al., 2013; Petak et al., 2023). Moreover, experimental selection and physiology studies have revealed ample functional phenotypic and genetic variation in the urchin genome, the raw material for spatially heterogeneous selection (Hedrick, 2006; Nunez et al., 2020; Melissa H. Pespeni et al., 2013b, 2013a). These patterns of adaptive differentiation in the context of the species' high levels of gene flow raise the question of whether urchin genomes harbor chromosomal inversions and whether such structural variants contribute to the patterns of local adaptation reported in this species.

In this study, we conducted the first genome-wide scan for structural variants including potentially adaptive inversions in natural populations of *S. purpuratus.* We leveraged an extensive panel of genomic variation sampled across a broad portion of the urchin's habitat, consisting of 137 phased genomes from single individuals. Specifically, we applied a local principal component analysis approach (i.e., Local PCA; Li and Ralph, 2019) to identify putative inverted regions across the genome and to pinpoint inversion-informative markers associated with these structural variants (Ma and Amos, 2012). Using these data, we asked: What is the spatial distribution of inversion markers across the species? What patterns of genetic variation do they harbor? And, do these potential inversions show signatures of natural selection across the urchin's habitat? We discovered nine loci containing potential inversions, which were also supported by patterns of linkage disequilibrium, nucleotide diversity, and F_{ST} statistics. We also investigated the relative age of genetic variation, and the distribution of genetic differentiation (using the X^TX statistic, (Günther and Coop, 2013), within and outside of loci of interest, as well as the frequency of the alternative orientations. These analyses, combined with gene functional class enrichment and SNP impact prediction, suggest that three of the nine loci are currently experiencing natural selection: one shows evidence of positive selection, while the other two exhibit signatures of balancing selection. Overall, these results suggest that putative chromosomal inversions in *S. purpuratus* may contribute to local adaptation in the presence of high gene flow.

RESULTS

Nine Potential Inversions Discovered in *S. purpuratus*

We used the genomic dataset from Petak et al., (2023) to identify structural variants including cosmopolitan inversions across the range of the species. The dataset contains 137 individuals, sequenced to an average depth of 6X, and harbors ~10 million SNPs across 21 chromosomes. The individuals originate from seven different locations, two in Oregon (Fogarty Creek (FOG) and Cape Blanco (CAP)), and five in California (Kibesilah Hill (KIB), Bodega Head (BOD), Terrace Point (TER), Lompoc Landing (LOM), and San Diego (SAN), from north to south). We applied the local PCA method (Li and Ralph, 2019) to each of the 21 chromosomes separately to find continuous genomic regions with PCA patterns different from the background. Results were compared across 4 different window sizes (500, 1000, 5,000, 10,000 SNPs) and genome-wide (see **Figures S1, S2**).

We identified nine regions across eight chromosomes exhibiting consistently high absolute MDS values (in either of two dimensions) across all window sizes, which also clustered individuals into three distinct groups (**Figure 1**). These loci were putatively identified as structural variants and had markers of being potential inversions. A summary of loci number, genomic location and frequency of karyotypes can be found in Table 1.

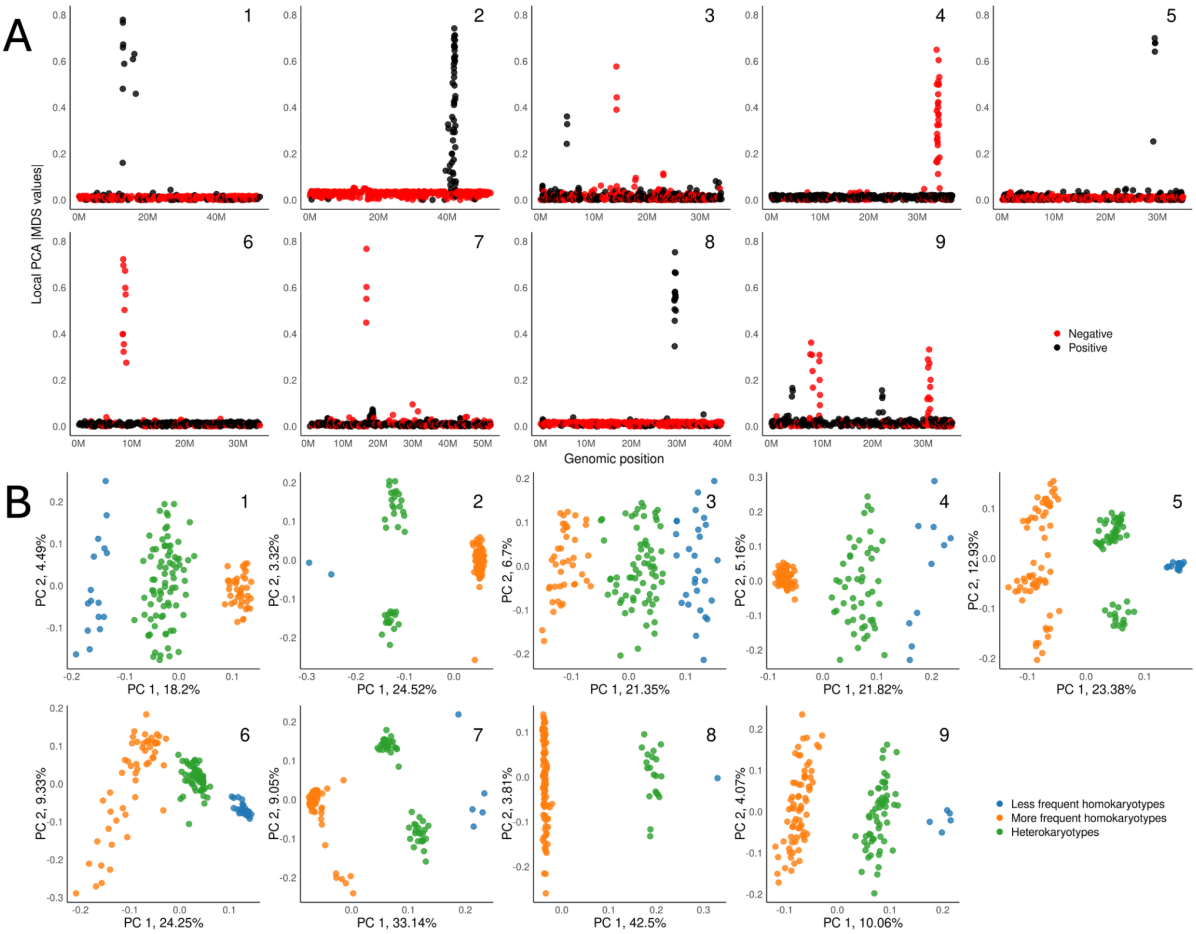

Figure 1. Nine regions with potential signatures of inversion polymorphisms. (A) The distribution of MDS values along the whole chromosome. (B) Each of the subplots show the PCA plot computed using genotype data restricted to the region with the MDS spike. Consistently high absolute MDS values indicate clustering along the first two principal components that is significantly different from the background (shown in A), and for these nine regions the PCA clustered the individuals into three separate groups (shown in B). The three colours correspond to the three karyotype groups: green = heterokaryotypes, blue = less frequent homokaryotypes, orange = more frequent homokaryotypes.

Locus	Scaffold ID	Start	Stop	Length (bp)	% homop/het./homoq
1	NW_022145594.1	12,702,886	16,793,794	4,090,908	29.2 / 59.1 / 11.7
2	NW_022145594.1	39,429,440	42,445,994	3,016,554	71.5 / 27.0 / 1.5
3	NW_022145597.1	14,219,351	14,298,524	79,173	29.9 / 49.6 / 20.4
4	NW_022145600.1	33,842,803	34,582,517	739,714	56.2 / 36.5 / 7.3
5	NW_022145601.1	28,950,395	29,566,247	615,852	40.1 / 47.4 / 12.4
6	NW_022145603.1	8,291,260	9,094,120	802,860	44.5 / 38.7 / 16.8
7	NW_022145606.1	16,610,900	16,737,625	126,725	62.8 / 33.6 / 3.6
8	NW_022145609.1	29,427,741	29,776,196	348,455	84.7 / 14.6 / 0.7
9	NW_022145610.1	30,779,143	31,460,853	681,710	41.6 / 54.7 / 3.6

Table 1: The Locus number, genetic position, length, and percentage of individuals belonging to each of the karyotype groups for each of the nine loci identified as putative inversion polymorphisms. Het. = heterokaryotype, homo(p) = homokaryotype to more frequent orientation, homo(q) = homokaryotype to the less frequent orientation. Coloured bars on the right show a visual representation of the karyotype percentages across individuals from all populations.

The evolutionary properties of potential inversion polymorphisms in *S. purpuratus*

To assess whether the nine regions identified through local PCA exhibited classical evolutionary signatures of structural variants including inversion polymorphisms, we characterized a range of standard population genetic statistics across all nine loci. In general, except for *locus 9,* we observed elevated levels of long-distance linkage disequilibrium (LD) relative to background LD, consistent with large blocks of suppressed recombination (**Figure 2**; a hallmark of inversions). Similarly, both homokaryotype groups, referred to here as "homokaryotype *p*" (i.e., major allele) and "homokaryotype *q*" (i.e., minor allele), show elevated levels of genetic differentiation (measured as F_{ST}). However, in four loci (i.e., *locus 2, locus 7, locus 8, and locus 9*), fewer than five homozygous inverted individuals were observed, which may reduce the statistical power to detect differentiation (**Table S1**). Notably, *locus 1*, *locus 3*, and *locus 6* exhibit the characteristic "hanging bridge" pattern of F_{ST} across the loci, with peaks centered over the putative inversion breakpoints, another hallmark of inversion polymorphisms (**Figure 3**). Conversely, the bridge pattern was absent in *locus 5*, indicating that this locus may represent a large, linked haplotype rather than an inversion. Additionally, for *locus 1* and *locus 6*, we detected F_{ST} peaks beyond the boundaries of the loci regions themselves, potentially reflecting linked variation near the breakpoints (**Figure 3**).

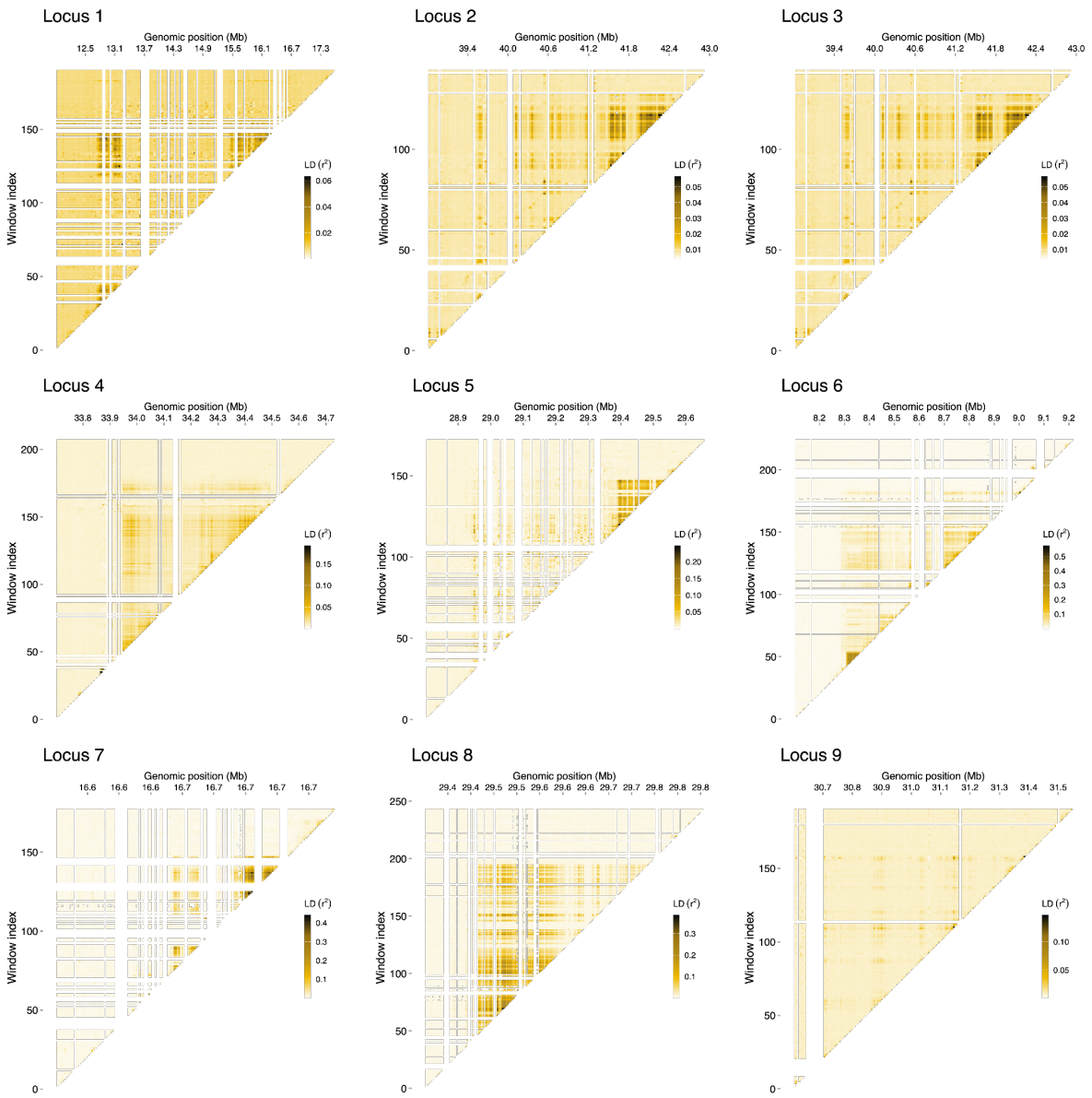

Figure 2: Linkage disequilibrium (LD) of loci (1-9). LD value (r^2) is coloured per SNP window from light yellow to black.

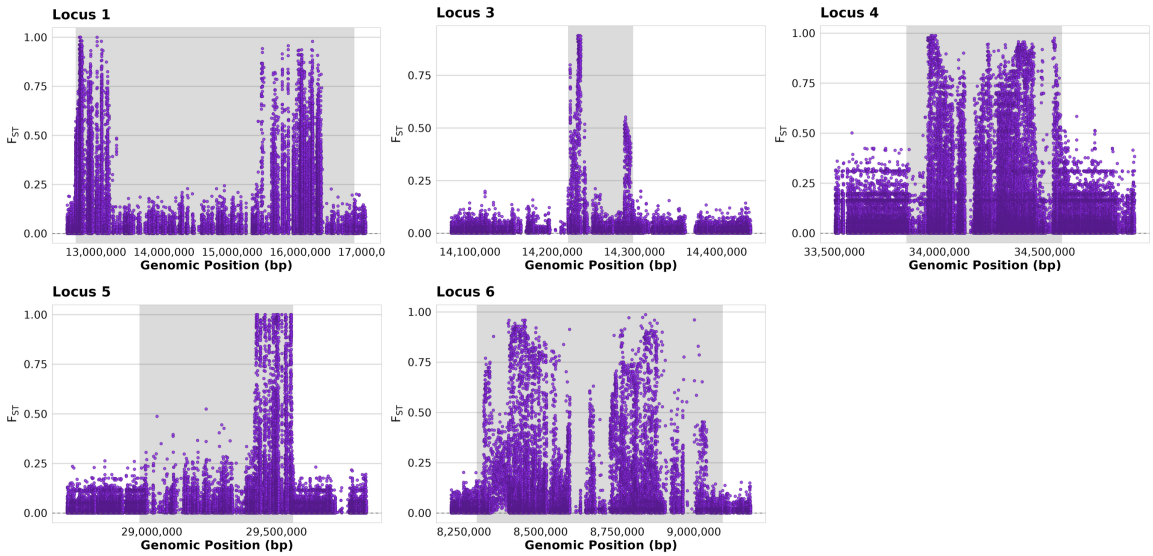

Figure 3: F_{ST} of loci +/- 1% of the loci region bases. Grey indicates potential inversion region. Plots not included for loci with alternative homozygote sample sizes below 10 (see table S1).

Tajima's D was also computed for each homokaryotype group individually across all loci (**Figure S3**). Notably, for *locus 6*, we observed elevated levels of Tajima's D within the homokaryotype *p* class relative to the alternative homokaryotype (**Figure S3**). We estimated the age of mutations within and outside the loci regions. SNPs located within loci were generally older than those outside. Notably, *locus 1* appears particularly ancient, with a time to the most recent common ancestor (TMRCA) 36.5% greater than the genome-wide average, exceeding that of all other chromosomes and loci regions. An exception to this pattern was observed in *locus 2* and *locus 6*, where SNPs inside the loci were comparatively younger (**Fig. 4A**).

Lastly, we examined genome annotations to assess the missense-to-silent mutation ratio within the identified loci, relative to the genomic background. We found that this ratio varied substantially across loci of interest (mean = 0.80 ± 0.48; **Fig. 4B**), in contrast to the relatively consistent background ratio observed across chromosomes (mean = 0.50 ± 0.08). Similarly, the proportion of SNPs with high, moderate, or low impact within loci varied (**Fig. 4C**).

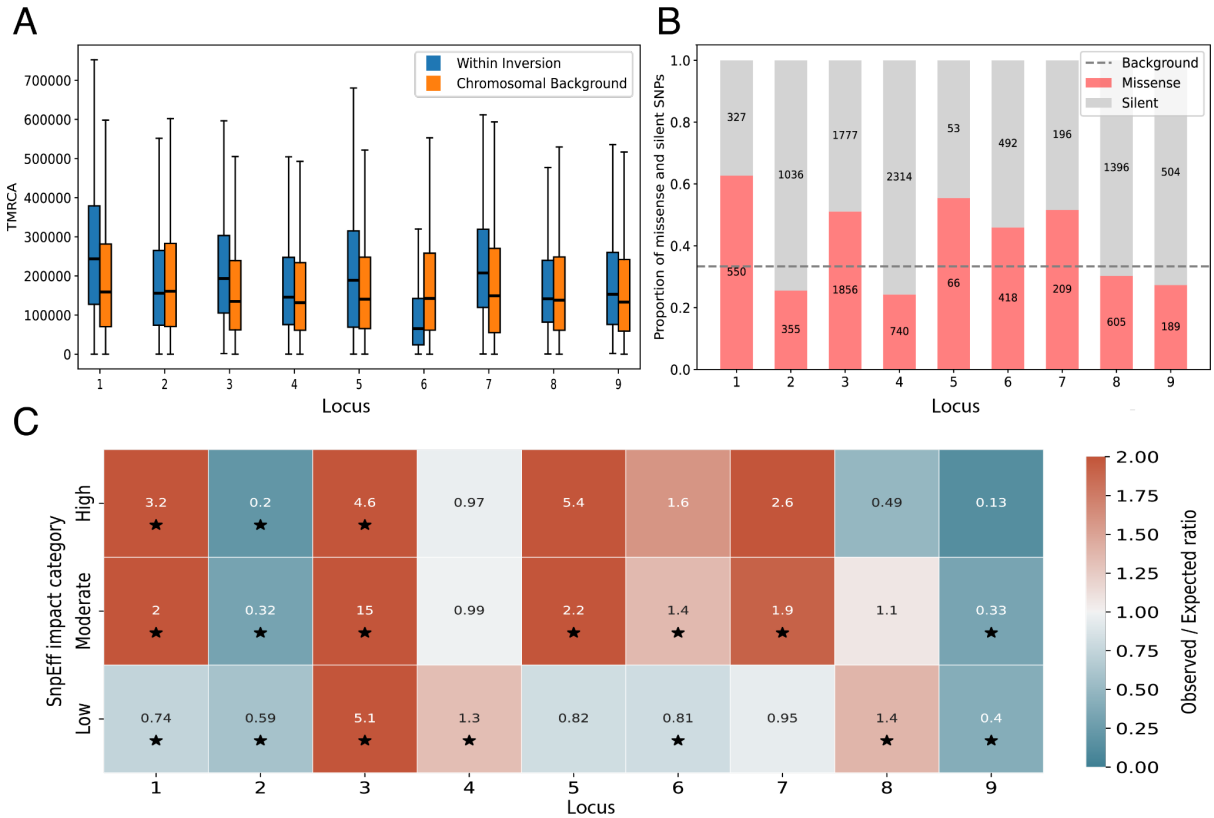

Figure 4: (A) Estimated Time to Most Recent Common Ancestor (TMRCA) for SNPs inside versus outside of the loci. All nine Mann–Whitney U tests showed significant differences (*P*<0.0001). SNPs in *locus 2* and *locus 6* are younger, whereas for the rest of the seven loci SNPs are older than expected based on the background. (B) Number of SNPs with missense or silent effect in each locus. The average missense / silent ratio across the chromosomes outside of loci was 0.5, meaning that the expectation is that there are twice as many silent SNPs than SNPs causing a change in amino acids (grey dashed line). *Locus 1* has a significant increase in missense SNPs, in contrast, the majority of SNPs in *locus 2* are silent. All ratios are significantly different from expectation based on two-sided binomial tests with Bonferroni correction for multiple testing. (C) Ratio of observed SNPs with high, moderate, or low impact to the expected number based on the chromosomal background (colour and value). SNP impacts were categorized using SnpEff. Red: more observed than expected, blue: less observed than expected. Stars indicate statistical significance based on two-sided binomial tests with Bonferroni correction for multiple testing.

Functional Gene Ontology Analyses Across Loci

Across the nine loci examined, two show significant patterns of functional enrichment, as determined by Gene Ontology (GO) analyses. Accordingly, *locus 1*, the largest potential inversion (~4 Mbp), contained 72 genes, with only the cell–cell adhesion category remaining significantly enriched after multiple-testing correction. *Locus 2*, the second largest loci, and potential inversion (~3 Mbp), harbored 76 genes with numerous enriched GO terms, most notably regulation of cellular response to hypoxia (GO:1900037) and regulation of telomere maintenance via telomere lengthening (GO:1904356). Two genes within the hypoxia category, LOC592339 and LOC105441271, both members of the *Hsp70* chaperone family, carried multiple variants including nonsynonymous and UTR changes.

None of the other loci showed significant GO enrichments. Nonetheless, we report other notable patterns. For example, *locus 6* showed a small F_{ST} spike adjacent to its right breakpoint (**Figure S4**), coinciding with clusters of 5S ribosomal RNA genes, suggesting a repetitive DNA region; excluding these, 21 genes mapped to the area, and are all associated with the carnitine biosynthetic process GO term (GO:0045329). Locus 7 spanned seven

genes, including LOC763759, annotated as zinc transporter ZIP14 (*Sp-Slc39a8*) and is potentially involved in bicarbonate transport during biomineralization.

Footprints of Natural Selection at Potential Inversion Polymorphisms

To evaluate whether any of the nine loci of interest are subject to natural selection, we performed a genome-wide scan using the Bayesian software *BayPass* to estimate the X^TX statistic while accounting for population structure. An elevated X^TX value relative to the genomic background indicates positive selection, whereas a reduced X^TX suggests a signature of balancing selection. Although X^TX is calculated on a per-locus basis, we complemented this analysis with a local score approach (implemented as a "Lindley Score"), which scans the genome for windows enriched in outlier X^TX values based $\xi = 3$ where only X^TX below *P-values* < 0.0001 contribute to the score. Regions showing significant enrichment of outliers were considered candidate loci evolving under natural selection. Our *BayPass* analysis revealed 351 significant peaks determined by Lindley Score, within these peaks we identified 2769 outlier SNPs. Although a comprehensive analysis of these peaks lies beyond the scope of this study, we examined the extent to which these signals of selection overlap with our candidate chromosomal inversions. Indeed, when we then mapped our potential inversions to regions experiencing significant selection (peaks; **Figure 5A**) we found *locus 2* (8 SNPs) *locus 6* (128 SNPs), *locus 8* (14 SNPs) and *locus 9* (105 SNPs) overlapped with significant peaks and SNPs (**Figures 5A, S4**). Within these loci, *locus 6*, *locus 7* and *locus 8* were significantly enriched for X^TX outliers (**Figure 5B; Table S2**).

Additionally, we found the X^TX scores produced by *BayPass* showed to be significantly lower in *locus 7* and *locus 8* than SNPs outside the loci within that chromosome (*P-value* < 0.001; **Figure 5C, Table S3**) consistent with signals of balancing selection, whilst *locus 6* is significantly higher than SNPs outside the locus within that chromosome suggesting there is a strong signal of local adaptation (*P-value* < 0.001; **Figure 5C, Table S3**). Lastly, we identified genes associated with overlapping SNPs across loci; although all were initially uncharacterized, BLASTp analysis of their protein sequences revealed that many SNPs within *locus 8* corresponded to *apolipoprotein B*, a gene implicated in lipid transport.

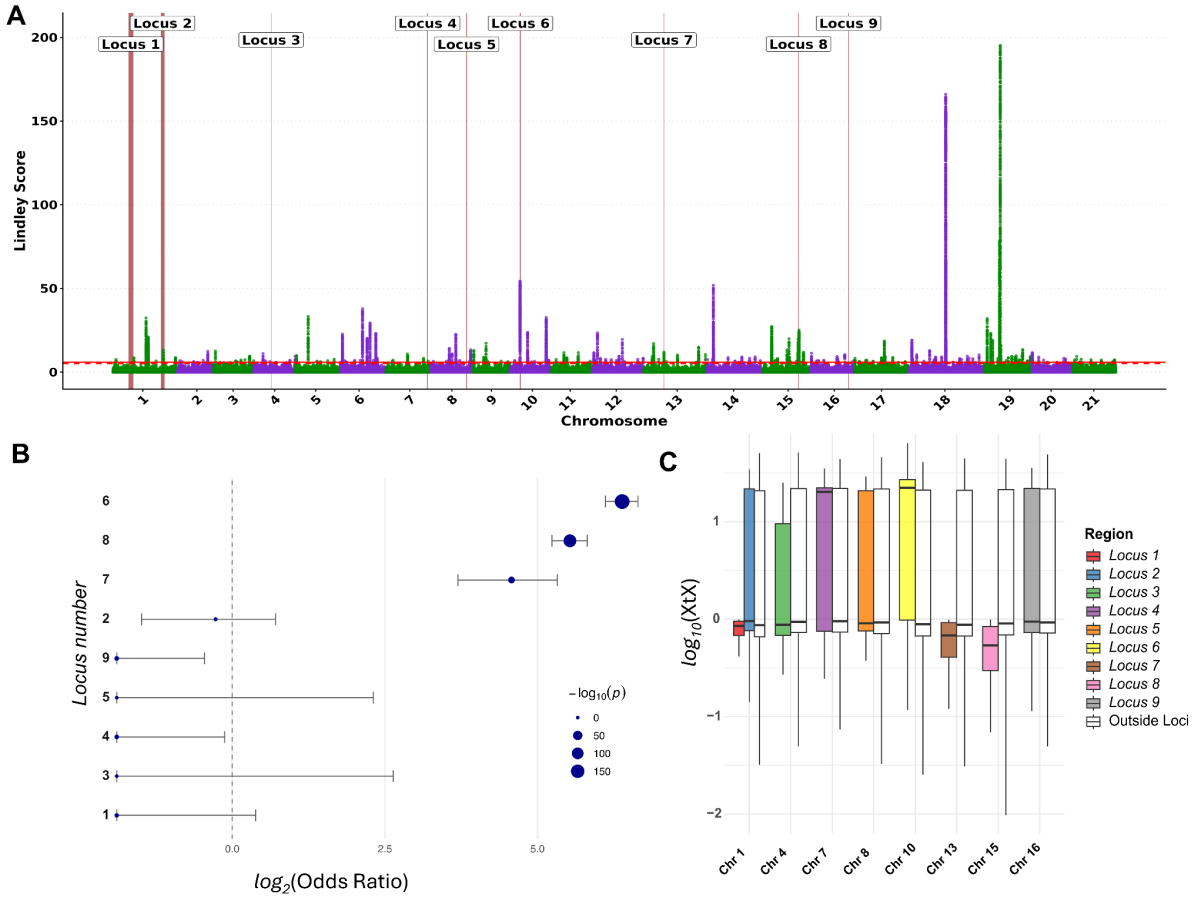

Figure 5: Comparisons across loci containing potential inversions for markers of selection. (a) Manhattan plot of SNPs displaying Lindley score values calculated from *localscore* grouped by chromosome and loci highlighted in red and numbered. (b) Fisher's exact test odd ratios across loci. Size of dot indicates -$\log_{10}$ *P*-value and error bar indicates ± SD. (c) Distribution of $\log_{10}(X^TX)$ within loci (coloured) versus outside (white) loci of interest per chromosome.

DISCUSSION

Here, we present a genome-wide assessment of loci that are potentially inversion polymorphisms in the model organism *Strongylocentrotus purpuratus*. To this end, we leveraged growing genomic resources for wild populations, including 137 diploid, fully phased, whole genomes sampled from seven sites along the west coast of North America. Using the local PCA method, which exploits patterns of SNP variation across the genome to identify highly diverged haplotypes, we detected nine genomic regions with putative signatures of inversion polymorphism. Local three-way clustering indicates the presence of two diverged haplotypes within populations (i.e., a common consequence of inversion polymorphism) but similar patterns can also arise from other processes, such as natural selection acting on linked alleles (Li & Ralph, 2019). Nevertheless, we conducted a series of follow-up analyses leveraging statistics of genetic differentiation and linkage. These analyses revealed several canonical signatures for most, but not all, inversion polymorphism, including the characteristic F_{ST} "bridge pattern," where breakpoints show elevated differentiation relative to the karyotype core (Guerrero et al., 2012). We also found evidence of long-range linkage disequilibrium (LD) among putative breakpoints. Because divergence between haplotype groups was often highest at the ends of these regions, and genome-wide LD in purple urchins is typically very low (Brennan et al., 2019; Petak et al., 2023), these patterns of long-range LD may indicate the presence of cosmopolitan structural variants.

Adaptive putative inversions in the genome of *S. purpuratus*

Since the early days of modern population genetics, structural variants including inversions have attracted considerable interest because of evidence that they can play key roles in facilitating local adaptation as well as speciation (Berdan et al., 2023; Dobzhansky and Wright, 1943; Hoffmann and Rieseberg, 2008; Wellenreuther and Bernatchez, 2018). With the advent of the genomic era, high-throughput sequencing has greatly expanded opportunities to study inversions across a wide range of species including, insects (Corbett-Detig and Hartl, 2012; Kapun et al., 2023, 2016; Lindtke et al., 2017; Nunez et al., 2024; Schaeffer, 2008; Wallberg et al., 2017, 2017; White et al., 2007), fishes (Barth et al., 2017; Jones et al., 2012; Leitwein et al., 2017), birds (Knief et al., 2017; Küpper et al., 2016; Lundberg et al., 2017; Pegan and Winger, 2025; Tuttle et al., 2016), mammals (Puig et al., 2015; Stefansson et al., 2005), and plants (Barb et al., 2014; Fransz et al., 2016; Lowry and Willis, 2010; Twyford and Friedman, 2015). For *S. purpuratus*, we identified nine loci that show evidence of being inversions distributed across the urchin genome. Given the species' complex landscape of selection and gene flow (Petak et al., 2023), we next asked whether any of these loci exhibit signatures of adaptation. To address this, we used a multipronged strategy that combined a suite of population genetic statistics with an agnostic scan for genomic differentiation, allowing us to evaluate patterns of variation within each potentially inverted locus.

First, we surveyed patterns of genetic variation. For example, our nucleotide diversity analyses varied considerably across the nine regions: it increased in *locus 2*, decreased in *locus 1* and *locus 5*, and in others showed either localized deviations or no change relative to background. Theory predicts that new inversions reduce diversity across the region,

329 followed by gradual recovery through gene flux, with diversity eventually peaking near the
330 breakpoints after long-term divergence (Navarro et al., 2000). Our results align with this
331 framework: *locus 1* appears older, showing elevated π and Tajima's D near the breakpoints;
332 *locus 4* shows reduced diversity at the breakpoints; and *locus 5* shows uniformly reduced
333 diversity. In terms of age, *locus 1*, *locus 3*, *locus 5*, and *locus 7* are older than the average
334 TMRCA of the genome, whereas *locus 6* is much younger (**Figure 4A**).
335
336 Next, we used *BayPass's* X^TX statistic (Gautier, 2015) to identify regions under
337 selection across the urchin genome. In marine species like sea urchins, high gene flow and
338 weak population structure is common (Petak et al., 2023), detecting selection is challenging
339 because divergence statistics such as F_{ST} are difficult to calibrate against demographic
340 expectations (Lotterhos and Whitlock, 2014). *BayPass* addresses this issue by incorporating
341 a genetic relatedness matrix (i.e., and Ω matrix) to account for population structure
342 (Bonhomme et al., 2010; Gautier, 2015). Our results identified multiple regions across the
343 urchin genome as likely targets of selection, a notable finding given the high levels of gene
344 flow in this species. Because our primary focus is on adaptation within inversions, a detailed
345 description of these results will be presented elsewhere. Nevertheless, an enrichment
346 analysis of the X^TX statistic highlighted three putative inversions, *locus 6, locus 7, and locus*
347 *8*, as being enriched for X^TX outliers. Taken together with the other lines of evidence
348 described above, these results suggest that the three loci may be subject to natural selection
349 in wild *S. purpuratus*.
350
351 **Positive and balancing selection in urchin loci**
352
353 Although all three candidate inversions, *locus 6, locus 7, and locus 8*, were enriched for X^TX
354 outliers, they displayed markedly different patterns of genetic variation across our analyses.
355 For example, *locus 6* showed the strongest enrichment of X^TX outliers, along with an
356 overrepresentation of missense mutations as well as mutations of low and medium effect in
357 SNPeff. At the same time, this locus appears to be the youngest of all putative inversion
358 polymorphisms and exhibits a median X^TX value higher than the chromosome-wide average.
359 Based on this evidence, we hypothesize that *locus 6* is experiencing positive selection in the
360 wild (Lewontin and Krakauer, 1973). In contrast, *locus 7* and *locus 8* both show X^TX values
361 consistently below the chromosome-wide average, a classical signature of balancing
362 selection (Brandt et al., 2018; Lewontin and Krakauer, 1973). Notably, *locus 7* is much older
363 than the genome-wide average's TMRCA, another hallmark of balancing selection, whereas
364 *locus 8* appears to be of similar age to the rest of the genome.
365
366 These findings allow us to hypothesize that both *locus 6* and *locus 7* may be
367 "supergene-like" loci associated with directional and balancing selection, respectively. Here,
368 supergenes refer to a class of inversions in which polymorphisms contain a suite of co-
369 adapted alleles that, shielded from recombination and migration load, confer increased
370 fitness in a given environment (Gutiérrez-Valencia et al., 2021; Schwander et al., 2014;
371 Stenløkk et al., 2022; Thompson and Jiggins, 2014). The elevated levels of missense
372 mutations observed at these loci are consistent with this interpretation and may represent
373 signatures of sheltered load, which is expected when processes that reduce recombination,
374 allows deleterious variants to persist so long as they do not substantially reduce haplotype
375 fitness (Uyenoyama, 2005). By contrast, the patterns for *locus 8* are less clear. Although this
376 locus shows several hallmarks of adaptive inversions (e.g., reduced X^TX, enrichment of

outlier SNPs), the evidence is inconclusive: it appears neither younger nor older than the genomic background, and no signatures of sheltered load were detected. One possibility is that the inversion is actually of average age and does not function as a supergene, but instead exerts its effects through breakpoint-mediated regulation of a nearby gene, disrupting *cis*- or *trans*-regulatory architecture, with missense variation playing little or no role (Harewood and Fraser, 2014; Lavington and Kern, 2017). Ultimately, functional follow-up experiments will be required to disentangle the mechanisms underlying these loci.

Whether the other candidate inversions are adaptive or functionally relevant remains open for follow-up studies. Inversion polymorphisms are a common form of genetic variation and can play diverse roles in genomic evolution: they may act as balanced inversions facilitating adaptation to heterogeneous environments, contribute to incipient speciation, regulate gene expression, harbor genetic load, or simply persist as neutral variants (Berdan et al., 2023; Faria et al., 2019; Harewood and Fraser, 2014; Westram et al., 2022). One possibility is that these loci influence traits in regions of the urchin's realized niche not captured by our sampling, such as different subtidal depths or geographic areas at the northern and southern range edges. And, since our *BayPass* analyses only considered variation within the sampled populations, any footprints of adaptation outside of this area would not be captured by the test. Further sampling and experimental work will be necessary to confirm whether these loci are indeed inversion polymorphisms, as well as to determine their functional roles.

Functional candidates inside the putative inversions

When assessing functional gene candidates within the loci of interest (i.e., identified as under selection from *BayPass*), we found in *locus 6* GO term associated with carnitine biosynthetic process as significantly enriched, this is involved in fatty-acid transportation important for mitochondrial activity, which may be important for energetic efficiency. Whilst, in *locus 7* zinc transporter ZIP14 was significantly enriched, which is hypothesised to be involved in bicarbonate transport potentially during biomineralisation. This may be relevant for local adaptation to regions of pH variation (Petak et al., 2023) due to the biomineral skeletal structure of the larvae. Within *locus 8,* there were three significant GO terms associated with lipid transportation and localisation which may be important for essential fatty acid distribution, key for growth and development. Additionally, we looked for genes that were under selection and within a potential inversion, (both an X^TX outlier and within a loci), finding only one characterised gene meeting this criteria. The characterised gene was within *locus 8*; apolipoprotein B, which is associated with the transport of key nutrients to the eggs from the digestive system via the coelomic fluid (Yuhi et al., 2022). Although we find some interesting functional genes within these loci and that are potentially under selection, future studies should assess genotype to phenotype interactions within these newly identified inversions.

Conclusion

We identified nine genomic regions in the purple sea urchin that exhibit reduced recombination and strong linkage, likely corresponding to highly polymorphic inversions that produced divergent haplotypes. For some of our candidate loci (e.g., *locus 6*) we found robust evidence supporting their classification as putative adaptive inversions. In contrast, other loci (e.g., *locus 5*) remain inconclusive and require further study. Indeed, future studies comparing gene expression and phenotypic differences between inversion orientations, as well as

experiments testing fitness effects under environmental stressors such as hypoxia or low pH, will be essential to further uncover the ecological and evolutionary significance of these loci. Nevertheless, these findings broaden our understanding of genome evolution in natural populations and add to the genetic repertoire of *S. purpuratus*, a key model system in developmental and evolutionary biology.

MATERIALS AND METHODS

Data acquisition and genome subsampling

Sequencing reads for all 140 *S. purpuratus* individuals were downloaded from NCBI Sequence Read Archive BioProject accession number PRJNA674711. See (Petak et al., 2023) for detail on sample collection and sequencing. Whilst the original dataset had 140 samples, here we used 137 removing outliers based on the genome-wide PCA (Petak et al., 2023). The reads were mapped to Spur ver. 5.0 (scaffold N50 ~37 Mbp), the same way as in (Petak et al., 2023). Next, the bcftools (v. 1.10.2) mpileup algorithm was used to call variants for the largest 21 scaffolds (more detail on this in the Results section) (Danecek et al., 2021). The output bcf files were filtered by the following criteria: posterior genotype probability (QUAL >= 40), read depth across all individuals (DP > 560), Mann-Whitney U test of Mapping Quality Bias (MQB > -3), Mann-Whitney U test of Read Position Bias (-3 < RPB < 3) and total number of alleles in called genotypes (AN > 238). Variants with allele counts of less than 14 were discarded (5% MAF of 137 diploid genomes) and only biallelic single nucleotide polymorphisms (SNPs) were kept. Lastly, all SNPs that fell into repetitive regions according to the dataset published by Panyushev et al., (2021) using RepeatModeler 2.0.1 were excluded from the dataset. The reference genome Spur v5.0 consists of 870 scaffolds. However, sorting all of the scaffolds by size, it is clear that 21 scaffolds are much larger than the rest. The largest 21 scaffolds all have length greater than ~30Mb and together they make up 89.87% of the data in the assembly. We focused our analysis on these scaffolds and since *S. purpuratus* is known to have 21 chromosomes (Eno et al., 2009), here, we refer to these scaffolds as chromosomes. We found the same three outlier individuals as (Petak et al., 2023), as many of the windows clustered them separately from the rest of the individuals along PC1. Thus, we excluded these three individuals from further analysis.

Finding putative inversions

We used the lostruct package in R to perform local PCA along the genome (Li and Ralph, 2019). The local PCA method of identifying genomic regions with population structure different from the background consists of the following steps. First, the genome of interest is divided into fixed sized windows (here we used a certain number of SNPs, but the results were unchanged when defining windows according to the number of base pairs). All chromosome-specific local PCA analyses were repeated for four different window sizes, 500, 1000, 5,000 and 10,000 SNPs. Then, for each window a PCA is done using the SNP values of the sequenced individuals, following which the first N (two in this study) principal components are used to compare each window with every other window. This matrix is then reduced to two dimensions using a multidimensional scaling (MDS). Windows that cluster together along these two MDS dimensions are windows where the SNP values of the sequenced individuals position the individuals similarly along the first two principal components. For example, if there is a large genomic region where there is population structure, i.e., SNP values cluster the

individuals according to the geographical location where they were collected, windows spanning this region will be closer together on the MDS plot than window from a genomic region where individuals from different collection sites overlap on the PCA. In this study, all MDS plots displayed the distribution of windows either in an L or a T shape (**Fig S2**), spreading windows along only one of two dimensions. This meant that we could plot the two MDS dimension values separately along the genome to find regions with a continuous series of outlier windows (i.e., windows where the individuals were clustered on the PCA the most different compared to the background). Three-way clustering was determined visually for each MDS outlier region on each chromosome, followed by analytical confirmation by the elbow method following Scikit-learn's (Pedregosa et al., 2011) (v.1.4.2) agglomerative clustering with different k-values. A specific MDS outlier region was selected for further analysis if 1) the windows with MDS spikes were consecutive and 2) the elbow method showed clear three-way clustering. Both of these criteria needed to be met consistently across the four window sizes for the region to be selected. The breakpoints were determined using Scikit-learn's K-means clustering of MDS values above |0.1|.

From regions to signals of selection

Each loci of interest (i.e., putative inversion) was extracted from the original filtered vcf file. These files were then converted to the gds file format to be used as an input to a principal component analysis (snpgdsPCA R package SNPRelate, (Zheng et al., 2012)). Based on the first two principal components, individuals were assigned to three groups using agglomerative clustering (Scikit-learn). Individuals were genotyped as heterokaryotypes if they belonged to the middle group along PC1, and as homokaryotypes if they belonged to groups on the right or left of the middle group. Following the standard convention of p for the more frequent and q for the less frequent allele, the homokaryotype group with more individuals was named homop and the group with less individuals as homoq.

Deviation from the Hardy-Weinberg equilibrium was determined using a chi-squared test, and the correlation between genotype and allele frequency with geographical latitude (from (Petak et al., 2023) was assessed using simple linear regression. Per site F_{ST} between the homokaryotype groups was calculated using vcftools (v.0.1.14) (Danecek et al., 2011) “weir-fst-pop” flag which estimates F_{ST} in accordance to (Weir and Cockerham, 1984). Vcftools was also used to calculate π diversity using the “window-pi” flag, and Tajima's D using the “TajimaD” flag, both with a window size of 100. Linkage disequilibrium was calculated from the gds files in windows (window size dependent on the size of the region of interest such that around 200 windows are compared) using the SNPRelate snpgdsLDMat function with “corr" option for returning the correlation coefficient. GO enrichment was conducted using GO terms retrieved from UniProt (The UniProt Consortium, 2025), followed by enrichment calculation with topGO (Alexa and Rahnenführer, 2022) and the GO enrichment tool on https://geneontology.org/ using the Bonferroni correction for multiple testing (Aleksander et al., 2023; Ashburner et al., 2000; Thomas et al., 2022). Finally, the SnpEff software (Cingolani et al., 2012) was used to categorize the effects of SNPs inside and outside of loci of interest (e.g., nonsynonymous/synonymous coding variants). A two-sided binomial test was used to determine the significance of differences between the number SNPs inside the loci with certain effects compared to the expectation given SNPs outside of the loci. Expectation was scaled by the frequency of variants in intergenic regions to account for potential biases in gene density.

Allele Age Estimates

Time to Most Recent Common Ancestor (TMRCA) was computed using the Genealogical Estimation of Variant Age (GEVA) package available on GitHub (Albers and McVean, 2020). First, the variant call files were phased using shapit5 (Hofmeister et al., 2023). Since *S. purpuratus* has no known mutation rate, recombination rate, or effective population size, we used parameters from *Drosophila melanogaster* and code adapted from (Nunez et al., 2024). Thus, the TMRCA values calculated here should not be converted into years, rather, they are meant to reflect relative differences in ages between SNPs. Mann–Whitney U tests were used to compare TMRCA values inside and outside of loci.

Identifying regions under selection

To identify regions experiencing selection we ran an analysis with *BayPass* v. 2.41 due to the lack of population structure (Petak et al., 2023) we ran the core *BayPass* model, randomly sampling and mixing the genome over 1000 chunks, then reassembling them in the correct order. *BayPass* was set to seven populations, with the rest of the settings as default. Peaks were then identified from the X^TX output using the package *localscore* (Robelin et al., 2025). Additionally, we ran a Fisher's exact test to identify loci that were significantly enriched for X^TX outliers found from the Baypass analysis.

Data Availability

All analyses in this study were based on publicly available datasets.

Acknowledgements

This work was supported by National Science Foundation awards IOS 1943396 to M.H.P. This work was also supported by start-up funds from the University of Vermont to J.C.B.N. D.E.S salary was supported by the Office of the Vice President for Research at the University of Vermont. The authors acknowledge the Vermont Advanced Computing Center (VACC; https://www.uvm.edu/vacc) for providing computational resources that contributed to this publication.

References

Albers, P.K., McVean, G., 2020. Dating genomic variants and shared ancestry in population-scale sequencing data. PLOS Biol. 18, e3000586. https://doi.org/10.1371/journal.pbio.3000586

Aleksander, S.A., Balhoff, J., Carbon, S., Cherry, J.M., Drabkin, H.J., Ebert, D., Feuermann, M., Gaudet, P., Harris, N.L., Hill, D.P., Lee, R., Mi, H., Moxon, S., Mungall, C.J., Muruganugan, A., Mushayahama, T., Sternberg, P.W., Thomas, P.D., Van Auken, K., Ramsey, J., Siegele, D.A., Chisholm, R.L., Fey, P., Aspromonte, M.C., Nugnes, M.V., Quaglia, F., Tosatto, S., Giglio, M., Nadendla, S., Antonazzo, G., Attrill, H., dos Santos, G., Marygold, S., Strelets, V., Tabone, C.J., Thurmond, J., Zhou, P., Ahmed, S.H., Asanitthong, P., Luna Buitrago, D., Erdol, M.N., Gage, M.C., Ali Kadhum, M., Li, K.Y.C., Long, M., Michalak, A., Pesala, A., Pritazahra, A., Saverimuttu, S.C.C., Su, R., Thurlow, K.E., Lovering, R.C., Logie, C., Oliferenko, S., Blake, J., Christie, K., Corbani, L., Dolan, M.E., Drabkin, H.J., Hill, D.P., Ni, L., Sitnikov, D., Smith, C., Cuzick, A., Seager, J., Cooper, L., Elser, J., Jaiswal, P., Gupta, P., Jaiswal, P., Naithani, S., Lera-Ramirez, M., Rutherford, K., Wood, V., De Pons, J.L., Dwinell, M.R., Hayman, G.T., Kaldunski, M.L., Kwitek, A.E., Laulederkind, S.J.F., Tutaj, M.A., Vedi, M., Wang, S.-J.,

D'Eustachio, P., Aimo, L., Axelsen, K., Bridge, A., Hyka-Nouspikel, N., Morgat, A., Aleksander, S.A., Cherry, J.M., Engel, S.R., Karra, K., Miyasato, S.R., Nash, R.S., Skrzypek, M.S., Weng, S., Wong, E.D., Bakker, E., Berardini, T.Z., Reiser, L., Auchincloss, A., Axelsen, K., Argoud-Puy, G., Blatter, M.-C., Boutet, E., Breuza, L., Bridge, A., Casals-Casas, C., Coudert, E., Estreicher, A., Livia Famiglietti, M., Feuermann, M., Gos, A., Gruaz-Gumowski, N., Hulo, C., Hyka-Nouspikel, N., Jungo, F., Le Mercier, P., Lieberherr, D., Masson, P., Morgat, A., Pedruzzi, I., Pourcel, L., Poux, S., Rivoire, C., Sundaram, S., Bateman, A., Bowler-Barnett, E., Bye-A-Jee, H., Denny, P., Ignatchenko, A., Ishtiaq, R., Lock, A., Lussi, Y., Magrane, M., Martin, M.J., Orchard, S., Raposo, P., Speretta, E., Tyagi, N., Warner, K., Zaru, R., Diehl, A.D., Lee, R., Chan, J., Diamantakis, S., Raciti, D., Zarowiecki, M., Fisher, M., James-Zorn, C., Ponferrada, V., Zorn, A., Ramachandran, S., Ruzicka, L., Westerfield, M., 2023. The Gene Ontology knowledgebase in 2023. Genetics 224, iyad031. https://doi.org/10.1093/genetics/iyad031

Alexa, A., Rahnenführer, J., 2022. topGO: enrichment analysis for gene ontology. R package version 2.44.0.

Amir, Y., Insler, M., Giller, A., Gutman, D., Atzmon, G., 2020. Senescence and Longevity of Sea Urchins. Genes 11, 573. https://doi.org/10.3390/genes11050573

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., Sherlock, G., 2000. Gene Ontology: tool for the unification of biology. Nat. Genet. 25, 25–29. https://doi.org/10.1038/75556

Barb, J.G., Bowers, J.E., Renaut, S., Rey, J.I., Knapp, S.J., Rieseberg, L.H., Burke, J.M., 2014. Chromosomal Evolution and Patterns of Introgression in Helianthus. Genetics 197, 969–979. https://doi.org/10.1534/genetics.114.165548

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.Chr., Svedäng, H., Jentoft, S., André, C., 2017. Genome architecture enables local adaptation of Atlantic cod despite high connectivity. Mol. Ecol. 26, 4452–4466. https://doi.org/10.1111/mec.14207

Berdan, E.L., Barton, N.H., Butlin, R., Charlesworth, B., Faria, R., Fragata, I., Gilbert, K.J., Jay, P., Kapun, M., Lotterhos, K.E., Mérot, C., Durmaz Mitchell, E., Pascual, M., Peichel, C.L., Rafajlović, M., Westram, A.M., Schaeffer, S.W., Johannesson, K., Flatt, T., 2023. How chromosomal inversions reorient the evolutionary process. J. Evol. Biol. 36, 1761–1782. https://doi.org/10.1111/jeb.14242

Bolnick, D.I., Nosil, P., 2007. NATURAL SELECTION IN POPULATIONS SUBJECT TO A MIGRATION LOAD. Evolution 61, 2229–2243. https://doi.org/10.1111/j.1558-5646.2007.00179.x

Bonhomme, M., Chevalet, C., Servin, B., Boitard, S., Abdallah, J., Blott, S., SanCristobal, M., 2010. Detecting Selection in Population Trees: The Lewontin and Krakauer Test Extended. Genetics 186, 241–262. https://doi.org/10.1534/genetics.110.117275

Brandt, D.Y.C., César, J., Goudet, J., Meyer, D., 2018. The Effect of Balancing Selection on Population Differentiation: A Study with HLA Genes. G3 Bethesda Md 8, 2805–2815. https://doi.org/10.1534/g3.118.200367

Brennan, R.S., Garrett, A.D., Huber, K.E., Hargarten, H., Pespeni, M.H., 2019. Rare genetic variation and balanced polymorphisms are important for survival in global change conditions. Proc. R. Soc. B Biol. Sci. 286, 20190943. https://doi.org/10.1098/rspb.2019.0943

Charlesworth, D., 2016. The status of supergenes in the 21st century: recombination suppression in Batesian mimicry and sex chromosomes and other complex adaptations. Evol. Appl. 9, 74–90. https://doi.org/10.1111/eva.12291

Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., Ruden, D.M., 2012. A program for annotating and predicting the effects of single

nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w^{1118}; iso-2; iso-3. Fly (Austin) 6, 80–92. https://doi.org/10.4161/fly.19695

Comeault, A.A., Flaxman, S.M., Riesch, R., Curran, E., Soria-Carrasco, V., Gompert, Z., Farkas, T.E., Muschick, M., Parchman, T.L., Schwander, T., Slate, J., Nosil, P., 2015. Selection on a Genetic Polymorphism Counteracts Ecological Speciation in a Stick Insect. Curr. Biol. 25, 1975–1981. https://doi.org/10.1016/j.cub.2015.05.058

Corbett-Detig, R.B., Hartl, D.L., 2012. Population Genomics of Inversion Polymorphisms in *Drosophila melanogaster*. PLOS Genet. 8, e1003056. https://doi.org/10.1371/journal.pgen.1003056

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 1000 Genomes Project Analysis Group, 2011. The variant call format and VCFtools. Bioinformatics 27, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., Li, H., 2021. Twelve years of SAMtools and BCFtools. GigaScience 10, giab008. https://doi.org/10.1093/gigascience/giab008

Dobzhansky, T., Wright, S., 1943. Genetics of Natural Populations. X. Dispersion Rates in Drosophila Pseudoobscura. Genetics 28, 304–340. https://doi.org/10.1093/genetics/28.4.304

Eno, C.C., Böttger, S.A., Walker, C.W., 2009. Methods for karyotyping and for localization of developmentally relevant genes on the chromosomes of the purple sea urchin, *Strongylocentrotus purpuratus*. Biol. Bull. 217, 306–312. https://doi.org/10.1086/BBLv217n3p306

Faria, R., Johannesson, K., Butlin, R.K., Westram, A.M., 2019. Evolving Inversions. Trends Ecol. Evol. 34, 239–248. https://doi.org/10.1016/j.tree.2018.12.005

Felsenstein, J., 1976. The theoretical population genetics of variable selection and migration. Annu. Rev. Genet. 10, 253–280. https://doi.org/10.1146/annurev.ge.10.120176.001345

Filbee-Dexter, K., Scheibling, R.E., 2014. Sea urchin barrens as alternative stable states of collapsed kelp ecosystems. Mar. Ecol. Prog. Ser. 495, 1–25. https://doi.org/10.3354/meps10573

Fransz, P., Linc, G., Lee, C.-R., Aflitos, S.A., Lasky, J.R., Toomajian, C., Ali, H., Peters, J., van Dam, P., Ji, X., Kuzak, M., Gerats, T., Schubert, I., Schneeberger, K., Colot, V., Martienssen, R., Koornneef, M., Nordborg, M., Juenger, T.E., de Jong, H., Schranz, M.E., 2016. Molecular, genetic and evolutionary analysis of a paracentric inversion in *Arabidopsis thaliana*. Plant J. 88, 159–178. https://doi.org/10.1111/tpj.13262

Gautier, M., 2015. Genome-Wide Scan for Adaptive Divergence and Association with Population-Specific Covariates. Genetics 201, 1555–1579. https://doi.org/10.1534/genetics.115.181453

Guerrero, R.F., Rousset, F., Kirkpatrick, M., 2012. Coalescent patterns for chromosomal inversions in divergent populations. Philos. Trans. R. Soc. B Biol. Sci. 367, 430–438. https://doi.org/10.1098/rstb.2011.0246

Günther, T., Coop, G., 2013. Robust Identification of Local Adaptation from Allele Frequencies. Genetics 195, 205–220. https://doi.org/10.1534/genetics.113.152462

Gutiérrez-Valencia, J., Hughes, P.W., Berdan, E.L., Slotte, T., 2021. The Genomic Architecture and Evolutionary Fates of Supergenes. Genome Biol. Evol. 13, evab057. https://doi.org/10.1093/gbe/evab057

Haldane, J.B.S., 1930. A mathematical theory of natural and artificial selection. (Part VI, Isolation.). Math. Proc. Camb. Philos. Soc. 26, 220–230. https://doi.org/10.1017/S0305004100015450

Harewood, L., Fraser, P., 2014. The impact of chromosomal rearrangements on regulation of gene expression. Hum. Mol. Genet. 23, R76–R82. https://doi.org/10.1093/hmg/ddu278

Hedrick, P.W., 2006. Genetic Polymorphism in Heterogeneous Environments: The Age of Genomics. Annu. Rev. Ecol. Evol. Syst. 37, 67–93. https://doi.org/10.1146/annurev.ecolsys.37.091305.110132

Hoffmann, A.A., Rieseberg, L.H., 2008. Revisiting the Impact of Inversions in Evolution: From Population Genetic Markers to Drivers of Adaptive Shifts and Speciation? Annu. Rev. Ecol. Evol. Syst. 39, 21–42. https://doi.org/10.1146/annurev.ecolsys.39.110707.173532

Hofmeister, R.J., Ribeiro, D.M., Rubinacci, S., Delaneau, O., 2023. Accurate rare variant phasing of whole-genome and whole-exome sequencing data in the UK Biobank. Nat. Genet. 55, 1243–1249. https://doi.org/10.1038/s41588-023-01415-w

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M.C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M.C., Myers, R.M., Miller, C.T., Summers, B.R., Knecht, A.K., Brady, S.D., Zhang, H., Pollen, A.A., Howes, T., Amemiya, C., Lander, E.S., Di Palma, F., Lindblad-Toh, K., Kingsley, D.M., 2012. The genomic basis of adaptive evolution in threespine sticklebacks. Nature 484, 55–61. https://doi.org/10.1038/nature10944

Joron, M., Frezal, L., Jones, R.T., Chamberlain, N.L., Lee, S.F., Haag, C.R., Whibley, A., Becuwe, M., Baxter, S.W., Ferguson, L., Wilkinson, P.A., Salazar, C., Davidson, C., Clark, R., Quail, M.A., Beasley, H., Glithero, R., Lloyd, C., Sims, S., Jones, M.C., Rogers, J., Jiggins, C.D., ffrench-Constant, R.H., 2011. Chromosomal rearrangements maintain a polymorphic supergene controlling butterfly mimicry. Nature 477, 203–206. https://doi.org/10.1038/nature10341

Kapun, M., Fabian, D.K., Goudet, J., Flatt, T., 2016. Genomic Evidence for Adaptive Inversion Clines in *Drosophila melanogaster*. Mol. Biol. Evol. 33, 1317–1336. https://doi.org/10.1093/molbev/msw016

Kapun, M., Flatt, T., 2019. The adaptive significance of chromosomal inversion polymorphisms in *Drosophila melanogaster*. Mol. Ecol. 28, 1263–1282. https://doi.org/10.1111/mec.14871

Kapun, M., Mitchell, E.D., Kawecki, T.J., Schmidt, P., Flatt, T., 2023. An Ancestral Balanced Inversion Polymorphism Confers Global Adaptation. Mol. Biol. Evol. 40, msad118. https://doi.org/10.1093/molbev/msad118

Kirubakaran, T.G., Grove, H., Kent, M.P., Sandve, S.R., Baranski, M., Nome, T., De Rosa, M.C., Righino, B., Johansen, T., Otterå, H., Sonesson, A., Lien, S., Andersen, Ø., 2016. Two adjacent inversions maintain genomic differentiation between migratory and stationary ecotypes of Atlantic cod. Mol. Ecol. 25, 2130–2143. https://doi.org/10.1111/mec.13592

Knief, U., Forstmeier, W., Pei, Y., Ihle, M., Wang, D., Martin, K., Opatová, P., Albrechtová, J., Wittig, M., Franke, A., Albrecht, T., Kempenaers, B., 2017. A sex-chromosome inversion causes strong overdominance for sperm traits that affect siring success. Nat. Ecol. Evol. 1, 1177–1184. https://doi.org/10.1038/s41559-017-0236-1

Kottler, E.J., Dickman, E.E., Sexton, J.P., Emery, N.C., Franks, S.J., 2021. Draining the Swamping Hypothesis: Little Evidence that Gene Flow Reduces Fitness at Range Edges. Trends Ecol. Evol. 36, 533–544. https://doi.org/10.1016/j.tree.2021.02.004

Küpper, C., Stocks, M., Risse, J.E., dos Remedios, N., Farrell, L.L., McRae, S.B., Morgan, T.C., Karlionova, N., Pinchuk, P., Verkuil, Y.I., Kitaysky, A.S., Wingfield, J.C., Piersma, T., Zeng, K., Slate, J., Blaxter, M., Lank, D.B., Burke, T., 2016. A supergene determines highly divergent male reproductive morphs in the ruff. Nat. Genet. 48, 79–83. https://doi.org/10.1038/ng.3443

Laurent, S., Pfeifer, S.P., Settles, M.L., Hunter, S.S., Hardwick, K.M., Ormond, L., Sousa, V.C., Jensen, J.D., Rosenblum, E.B., 2016. The population genomics of rapid adaptation: disentangling signatures of selection and demography in white sands lizards. Mol. Ecol. 25, 306–323. https://doi.org/10.1111/mec.13385

Lavington, E., Kern, A.D., 2017. The Effect of Common Inversion Polymorphisms In(2L)t and In(3R)Mo on Patterns of Transcriptional Variation in *Drosophila melanogaster*. G3 GenesGenomesGenetics 7, 3659–3668. https://doi.org/10.1534/g3.117.1133

Leitwein, M., Garza, J.C., Pearse, D.E., 2017. Ancestry and adaptive evolution of anadromous, resident, and adfluvial rainbow trout (*Oncorhynchus mykiss*) in the San

Francisco bay area: application of adaptive genomic variation to conservation in a highly impacted landscape. Evol. Appl. 10, 56–67. https://doi.org/10.1111/eva.12416

Lenormand, T., 2002. Gene flow and the limits to natural selection. Trends Ecol. Evol. 17, 183–189. https://doi.org/10.1016/S0169-5347(02)02497-7

Levene, H., 1953. Genetic Equilibrium When More Than One Ecological Niche is Available. Am. Nat. 87, 331–333. https://doi.org/10.1086/281792

Lewontin, R.C., Krakauer, J., 1973. DISTRIBUTION OF GENE FREQUENCY AS A TEST OF THE THEORY OF THE SELECTIVE NEUTRALITY OF POLYMORPHISMS. Genetics 74, 175–195. https://doi.org/10.1093/genetics/74.1.175

Li, H., Ralph, P., 2019. Local PCA Shows How the Effect of Population Structure Differs Along the Genome. Genetics 211, 289–304. https://doi.org/10.1534/genetics.118.301747

Lindtke, D., Lucek, K., Soria-Carrasco, V., Villoutreix, R., Farkas, T.E., Riesch, R., Dennis, S.R., Gompert, Z., Nosil, P., 2017. Long-term balancing selection on chromosomal variants associated with crypsis in a stick insect. Mol. Ecol. 26, 6189–6205. https://doi.org/10.1111/mec.14280

Lotterhos, K.E., Whitlock, M.C., 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. Mol. Ecol. 23, 2178–2192. https://doi.org/10.1111/mec.12725

Lowry, D.B., Willis, J.H., 2010. A Widespread Chromosomal Inversion Polymorphism Contributes to a Major Life-History Transition, Local Adaptation, and Reproductive Isolation. PLOS Biol. 8, e1000500. https://doi.org/10.1371/journal.pbio.1000500

Lundberg, M., Liedvogel, M., Larson, K., Sigeman, H., Grahn, M., Wright, A., Åkesson, S., Bensch, S., 2017. Genetic differences between willow warbler migratory phenotypes are few and cluster in large haplotype blocks. Evol. Lett. 1, 155–168. https://doi.org/10.1002/evl3.15

Ma, J., Amos, C.I., 2012. Investigation of Inversion Polymorphisms in the Human Genome Using Principal Components Analysis. PLOS ONE 7, e40224. https://doi.org/10.1371/journal.pone.0040224

Marshall, D.J., Monro, K., Bode, M., Keough, M.J., Swearer, S., 2010. Phenotype–environment mismatches reduce connectivity in the sea. Ecol. Lett. 13, 128–140. https://doi.org/10.1111/j.1461-0248.2009.01408.x

Maruki, T., Ye, Z., Lynch, M., 2022. Evolutionary Genomics of a Subdivided Species. Mol. Biol. Evol. 39, msac152. https://doi.org/10.1093/molbev/msac152

Navarro, A., Barbadilla, A., Ruiz, A., 2000. Effect of Inversion Polymorphism on the Neutral Nucleotide Variability of Linked Chromosomal Regions in *Drosophila*. Genetics 155, 685–698. https://doi.org/10.1093/genetics/155.2.685

Nunez, J.C.B., Flight, P.A., Neil, K.B., Rong, S., Eriksson, L.A., Ferranti, D.A., Rosenblad, M.A., Blomberg, A., Rand, D.M., 2020. Footprints of natural selection at the mannose-6-phosphate isomerase locus in barnacles. Proc. Natl. Acad. Sci. U. S. A. 117, 5376–5385. https://doi.org/10.1073/pnas.1918232117

Nunez, J.C.B., Lenhart, B.A., Bangerter, A., Murray, C.S., Mazzeo, G.R., Yu, Y., Nystrom, T.L., Tern, C., Erickson, P.A., Bergland, A.O., 2024. A cosmopolitan inversion facilitates seasonal adaptation in overwintering *Drosophila*. Genetics 226, iyad207. https://doi.org/10.1093/genetics/iyad207

Panyushev, N., Okorokova, L., Danilov, L., Adonin, L., 2021. Pattern of Repetitive Element Transcription Segregate Cell Lineages during the Embryogenesis of Sea Urchin *Strongylocentrotus purpuratus*. Biomedicines 9, 1736. https://doi.org/10.3390/biomedicines9111736

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res. 12, 2825–2830.

Pegan, T.M., Winger, B.M., 2025. Large Inversion Polymorphisms are Widespread in North American Songbirds. Genome Biol Evol. 17, evaf205. https://doi.org/10.1093/gbe/evaf205.

Pespeni, Melissa H., Barney, B.T., Palumbi, S.R., 2013a. DIFFERENCES IN THE REGULATION OF GROWTH AND BIOMINERALIZATION GENES REVEALED THROUGH LONG-TERM COMMON-GARDEN ACCLIMATION AND EXPERIMENTAL GENOMICS IN THE PURPLE SEA URCHIN. Evolution 67, 1901–1914. https://doi.org/10.1111/evo.12036

Pespeni, M. H., Chan, F., Menge, B.A., Palumbi, S.R., 2013. Signs of Adaptation to Local pH Conditions across an Environmental Mosaic in the California Current Ecosystem. Integr. Comp. Biol. 53, 857–870. https://doi.org/10.1093/icb/ict094

Pespeni, M.H., Garfield, D.A., Manier, M.K., Palumbi, S.R., 2012. Genome-wide polymorphisms show unexpected targets of natural selection. Proc. R. Soc. B Biol. Sci. 279, 1412–1420. https://doi.org/10.1098/rspb.2011.1823

Pespeni, M.H., Oliver, T.A., Manier, M.K., Palumbi, S.R., 2010. Restriction Site Tiling Analysis: accurate discovery and quantitative genotyping of genome-wide polymorphisms using nucleotide arrays. Genome Biol. 11, R44. https://doi.org/10.1186/gb-2010-11-4-r44

Pespeni, Melissa H., Sanford, E., Gaylord, B., Hill, T.M., Hosfelt, J.D., Jaris, H.K., LaVigne, M., Lenz, E.A., Russell, A.D., Young, M.K., Palumbi, S.R., 2013b. Evolutionary change during experimental ocean acidification. Proc. Natl. Acad. Sci. 110, 6937–6942. https://doi.org/10.1073/pnas.1220673110

Petak, C., Frati, L., Brennan, R.S., Pespeni, M.H., 2023. Whole-Genome Sequencing Reveals That Regulatory and Low Pleiotropy Variants Underlie Local Adaptation to Environmental Variability in Purple Sea Urchins. Am. Nat. 202, 571–586. https://doi.org/10.1086/726013

Puig, M., Casillas, S., Villatoro, S., Cáceres, M., 2015. Human inversions and their functional consequences. Brief. Funct. Genomics 14, 369–379. https://doi.org/10.1093/bfgp/elv020

Robelin, D., Déjean, S., Mercier, S., 2025. localScore: an R package to highlight optimal and suboptimal segments in a sequence with associated p-values computation. Peer Community J. 5. https://doi.org/10.24072/pcjournal.650

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. Philos. Trans. R. Soc. B Biol. Sci. 377, 20210200. https://doi.org/10.1098/rstb.2021.0200

Schaeffer, S.W., 2008. Selection in Heterogeneous Environments Maintains the Gene Arrangement Polymorphism of *Drosophila pseudoobscura*. Evolution 62, 3082–3099. https://doi.org/10.1111/j.1558-5646.2008.00504.x

Schwander, T., Libbrecht, R., Keller, L., 2014. Supergenes and complex phenotypes. Curr. Biol. CB 24, R288-294. https://doi.org/10.1016/j.cub.2014.01.056

Slatkin, M., 1987. Gene flow and the geographic structure of natural populations. Science 236, 787–792. https://doi.org/10.1126/science.3576198

Sodergren, E., Weinstock, G.M., Davidson, E.H., Cameron, R.A., Gibbs, R.A., Angerer, R.C., Angerer, L.M., Arnone, M.I., Burgess, D.R., Burke, R.D., Coffman, J.A., Dean, M., Elphick, M.R., Ettensohn, C.A., Foltz, K.R., Hamdoun, A., Hynes, R.O., Klein, W.H., Marzluff, W., McClay, D.R., Morris, R.L., Mushegian, A., Rast, J.P., Smith, L.C., Thorndyke, M.C., Vacquier, V.D., Wessel, G.M., Wray, G., Zhang, L., Elsik, C.G., Ermolaeva, O., Hlavina, W., Hofmann, G., Kitts, P., Landrum, M.J., Mackey, A.J., Maglott, D., Panopoulou, G., Poustka, A.J., Pruitt, K., Sapojnikov, V., Song, X., Souvorov, A., Solovyev, V., Wei, Z., Whittaker, C.A., Worley, K., Durbin, K.J., Shen, Y., Fedrigo, O., Garfield, D., Haygood, R., Primus, A., Satija, R., Severson, T., Gonzalez-Garay, M.L., Jackson, A.R., Milosavljevic, A., Tong, M., Killian, C.E., Livingston, B.T., Wilt, F.H., Adams, N., BelleÌ, R., Carbonneau, S., Cheung, R., Cormier, P., Cosson, B., Croce, J., Fernandez-Guerra, A., GenevieÌ€re, A.-M., Goel, M., Kelkar, H., Morales, J., Mulner-Lorillon, O., Robertson, A.J., Goldstone, J.V., Cole, B., Epel, D., Gold, B., Hahn, M.E., Howard-Ashby, M., Scally, M., Stegeman, J.J., Allgood, E.L., Cool, J., Judkins, K.M., McCafferty, S.S., Musante, A.M., Obar, R.A., Rawson, A.P., Rossetti, B.J., Gibbons, I.R., Hoffman, M.P., Leone, A., Istrail, S., Materna, S.C., Samanta, M.P., Stolc, V., Tongprasit, W., Tu, Q., Bergeron, K.-F.,

Brandhorst, B.P., Whittle, J., Berney, K., Bottjer, D.J., Calestani, C., Peterson, K., Chow, E., Yuan, Q.A., Elhaik, E., Graur, D., Reese, J.T., Bosdet, I., Heesun, S., Marra, M.A., Schein, J., Anderson, M.K., Brockton, V., Buckley, K.M., Cohen, A.H., Fugmann, S.D., Hibino, T., Loza-Coll, M., Majeske, A.J., Messier, C., Nair, S.V., Pancer, Z., Terwilliger, D.P., Agca, C., Arboleda, E., Chen, N., Churcher, A.M., HallboÌˆoÌˆk, F., Humphrey, G.W., Idris, M.M., Kiyama, T., Liang, S., Mellott, D., Mu, X., Murray, G., Olinski, R.P., Raible, F., Rowe, M., Taylor, J.S., Tessmar-Raible, K., Wang, D., Wilson, K.H., Yaguchi, S., Gaasterland, T., Galindo, B.E., Gunaratne, H.J., Juliano, C., Kinukawa, M., Moy, G.W., Neill, A.T., Nomura, M., Raisch, M., Reade, A., Roux, M.M., Song, J.L., Su, Y.-H., Townley, I.K., Voronina, E., Wong, J.L., Amore, G., Branno, M., Brown, E.R., Cavalieri, V., Duboc, V., Duloquin, L., Flytzanis, C., Gache, C., Lapraz, F., Lepage, T., Locascio, A., Martinez, P., Matassi, G., Matranga, V., Range, R., Rizzo, F., RoÌˆttinger, E., Beane, W., Bradham, C., Byrum, C., Glenn, T., Hussain, S., Manning, G., Miranda, E., Thomason, R., Walton, K., Wikramanayke, A., Wu, S.-Y., Xu, R., Brown, C.T., Chen, L., Gray, R.F., Lee, P.Y., Nam, J., Oliveri, P., Smith, J., Muzny, D., Bell, S., Chacko, J., Cree, A., Curry, S., Davis, C., Dinh, H., Dugan-Rocha, S., Fowler, J., Gill, R., Hamilton, C., Hernandez, J., Hines, S., Hume, J., Jackson, L., Jolivet, A., Kovar, C., Lee, S., Lewis, L., Miner, G., Morgan, M., Nazareth, L.V., Okwuonu, G., Parker, D., Pu, L.-L., Thorn, R., Wright, R., 2006. The Genome of the Sea Urchin *Strongylocentrotus purpuratus*. Science 314, 941–952. https://doi.org/10.1126/science.1133609

Stefansson, H., Helgason, A., Thorleifsson, G., Steinthorsdottir, V., Masson, G., Barnard, J., Baker, A., Jonasdottir, A., Ingason, A., Gudnadottir, V.G., Desnica, N., Hicks, A., Gylfason, A., Gudbjartsson, D.F., Jonsdottir, G.M., Sainz, J., Agnarsson, K., Birgisdottir, B., Ghosh, S., Olafsdottir, A., Cazier, J.-B., Kristjansson, K., Frigge, M.L., Thorgeirsson, T.E., Gulcher, J.R., Kong, A., Stefansson, K., 2005. A common inversion under selection in Europeans. Nat. Genet. 37, 129–137. https://doi.org/10.1038/ng1508

Stenløkk, K., Saitou, M., Rud-Johansen, L., Nome, T., Moser, M., Árnyasi, M., Kent, M., Barson, N.J., Lien, S., 2022. The emergence of supergenes from inversions in Atlantic salmon. Philos. Trans. R. Soc. B Biol. Sci. 377, 20210195. https://doi.org/10.1098/rstb.2021.0195

The UniProt Consortium, 2025. UniProt: the Universal Protein Knowledgebase in 2025. Nucleic Acids Res. 53, D609–D617. https://doi.org/10.1093/nar/gkae1010

Thomas, P.D., Ebert, D., Muruganujan, A., Mushayahama, T., Albou, L., Mi, H., 2022. PANTHER: Making genome-scale phylogenetics accessible to all. Protein Sci. Publ. Protein Soc. 31, 8–22. https://doi.org/10.1002/pro.4218

Thompson, M.J., Jiggins, C.D., 2014. Supergenes and their role in evolution. Heredity 113, 1–8. https://doi.org/10.1038/hdy.2014.20

Tuttle, E.M., Bergland, A.O., Korody, M.L., Brewer, M.S., Newhouse, D.J., Minx, P., Stager, M., Betuel, A., Cheviron, Z.A., Warren, W.C., Gonser, R.A., Balakrishnan, C.N., 2016. Divergence and Functional Degradation of a Sex Chromosome-like Supergene. Curr. Biol. CB 26, 344–350. https://doi.org/10.1016/j.cub.2015.11.069

Twyford, A.D., Friedman, J., 2015. Adaptive divergence in the monkey flower *Mimulus guttatus* is maintained by a chromosomal inversion. Evolution 69, 1476–1486. https://doi.org/10.1111/evo.12663

Uyenoyama, M.K., 2005. Evolution under tight linkage to mating type. New Phytol. 165, 63–70. https://doi.org/10.1111/j.1469-8137.2004.01246.x

Wallberg, A., Schöning, C., Webster, M.T., Hasselmann, M., 2017. Two extended haplotype blocks are associated with adaptation to high altitude habitats in East African honey bees. PLoS Genet. 13, e1006792. https://doi.org/10.1371/journal.pgen.1006792

Weir, B.S., Cockerham, C.C., 1984. Estimating F-Statistics for the Analysis of Population Structure. Evolution 38, 1358–1370. https://doi.org/10.2307/2408641

Wellenreuther, M., Bernatchez, L., 2018. Eco-Evolutionary Genomics of Chromosomal Inversions. Trends Ecol. Evol. 33, 427–440. https://doi.org/10.1016/j.tree.2018.04.002

Westram, A.M., Faria, R., Johannesson, K., Butlin, R., Barton, N., 2022. Inversions and parallel evolution. Philos. Trans. R. Soc. B Biol. Sci. 377, 20210203. https://doi.org/10.1098/rstb.2021.0203

White, B.J., Hahn, M.W., Pombi, M., Cassone, B.J., Lobo, N.F., Simard, F., Besansky, N.J., 2007. Localization of candidate regions maintaining a common polymorphic inversion (2La) in *Anopheles gambiae*. PLoS Genet. 3, e217. https://doi.org/10.1371/journal.pgen.0030217

Yeaman, S., 2015. Local Adaptation by Alleles of Small Effect. Am. Nat. 186, S74–S89. https://doi.org/10.1086/682405

Yeaman, S., Aeschbacher, S., Bürger, R., 2016. The evolution of genomic islands by increased establishment probability of linked alleles. Mol. Ecol. 25, 2542–2558. https://doi.org/10.1111/mec.13611

Yeaman, S., Whitlock, M.C., 2011. THE GENETIC ARCHITECTURE OF ADAPTATION UNDER MIGRATION–SELECTION BALANCE. Evolution 65, 1897–1911. https://doi.org/10.1111/j.1558-5646.2011.01269.x

Yuhi, T., Nishimiya, O., Ohno, K., Takita, A., Inoguchi, T., Ura, K., Takagi, Y., 2022. Cloning of cDNA encoding a newly recognized apolipoprotein-like protein and its expression in the northern sea urchin *Mesocentrotus nudus*. Fish. Sci. 88, 259–273. https://doi.org/10.1007/s12562-022-01584-3

Zheng, X., Levine, D., Shen, J., Gogarten, S.M., Laurie, C., Weir, B.S., 2012. A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinforma. Oxf. Engl. 28, 3326–3328. https://doi.org/10.1093/bioinformatics/bts606