Extensive Introgression among Strongylocentrotid Sea Urchins Revealed by Phylogenomics

Matthew Glasenapp[1] and Grant Pogson[1]

[1]University of California Santa Cruz

March 20, 2023

Abstract

Gametic isolation is thought to play an important role in the evolution of reproductive isolation in broadcast-spawning marine invertebrates. However, it remains uncertain whether gametic isolation typically develops early in the speciation process or accumulates after other reproductive barriers are already in place. It is also unknown whether gametic incompatibilities have effectively prevented introgression during later stages of divergence. Here, we use phylogenomic approaches to test whether the well-documented asymmetric gametic incompatibilities between strongylocentrotid urchins have been effective in preventing introgression. Despite a well-supported species tree, we found considerable phylogenetic discordance that cannot be explained by incomplete lineage sorting alone. There was strong support for introgression between at least four pairs of extant taxa: *S. pallidus* - *S. droebachiensis*, *S. intermedius* - *S. pallidus*, *S. purpuratus* - *S. fragilis*, and *M. franciscanus* - *P. depressus*. There was additional evidence for introgression on internal branches of the phylogeny. Although gametic incompatibilities may be important in species recognition and the maintenance of species boundaries in strongylocentrotid urchins, gametic isolation does not appear to have been an effective barrier to introgression. The continued divergence in the face of widespread introgression indicates that other reproductive isolating barriers likely exist and may have been more critical in establishing reproductive isolation early in speciation.

Extensive Introgression among Strongylocentrotid Sea Urchins Revealed by Phylogenomics

Matthew R. Glasenapp*[1] and Grant H. Pogson[1]

[1]Department of Ecology and Evolutionary Biology, University of California, Santa Cruz

*__Corresponding Author__: Matthew R. Glasenapp; mglasena at ucsc dot edu

Abstract

Gametic isolation is thought to play an important role in the evolution of reproductive isolation in broadcast-spawning marine invertebrates. However, it remains uncertain whether gametic isolation typically develops early in the speciation process or accumulates after other reproductive barriers are already in place. It is also unknown whether gametic incompatibilities have effectively prevented introgression during later stages of divergence. Here, we use phylogenomic approaches to test whether the well-documented asymmetric gametic incompatibilities between strongylocentrotid urchins have been effective in preventing introgression. Despite a well-supported species tree, we found considerable phylogenetic discordance that cannot be explained by incomplete lineage sorting alone. There was strong support for introgression between at least four pairs of extant taxa: *S. pallidus* ⇔ *S. droebachiensis*, *S. intermedius* ⇔ *S. pallidus*, *S. purpuratus* ⇔ *S. fragilis*, and *M. franciscanus* ⇔ *P. depressus*. There was additional evidence for introgression on internal branches of the phylogeny. Although gametic incompatibilities may be important in species recognition and the maintenance of species boundaries in strongylocentrotid urchins, gametic isolation does not appear to have been an effective barrier to introgression. The continued divergence in the face of widespread introgression indicates that other reproductive isolating barriers likely exist and may have been more critical in establishing reproductive isolation early in speciation.

1 Introduction

The new availability of genome-scale data has stimulated considerable investigation into the genomic architecture of speciation - the number, kind, location, and relative effect size of loci underlying reproductive isolation. Understanding the genetic basis of speciation requires identifying these so-called "barrier loci" and characterizing the selective agents responsible for their divergence (Orr, 2005). Although characters experiencing diversifying selection often exhibit pleiotropic effects on reproductive compatibility and produce isolating barriers, the connection between their divergence and reproductive isolation remains weak (Schluter & Rieseberg, 2022). Speciation genomics research typically proceeds by either characterizing genome-wide patterns of differentiation (Edelman et al., 2019; Malinsky et al., 2015; Poelstra et al., 2014) or mapping genes underlying known hybrid incompatibility phenotypes (Larson et al., 2018; Powell et al., 2020; Zuellig & Sweigart, 2018). One of the major outstanding questions concerns whether reproductive incompatibilities evolve more commonly from adaptive divergence or nonadaptive processes such as intragenomic conflict and divergent gene duplication resolution (Schluter & Rieseberg, 2022). Contrary to the recent enthusiasm for ecological speciation, hybrid incompatibility loci are often associated with nonadaptive processes, although research seeking to identify barrier loci has been historically biased towards postzygotic isolation (Campbell et al., 2018; Presgraves, 2010). Broader taxonomic representation is needed because most conclusions have been drawn from a limited number of taxa (Campbell et al., 2018).

Our understanding of speciation in marine species lags behind that of terrestrial groups. Studying marine speciation offers the unique opportunity to characterize the evolution of reproductive isolation in settings where physical geographic barriers are more obscure (Faria et

al., 2021). Broadcast-spawning marine invertebrates are a compelling group for speciation studies because their life histories and reproductive ecologies differ drastically from most animal speciation models. Broadcast spawners release their gametes into the water column, where fertilization occurs externally, resulting in large fecundities and population sizes. Due to long planktonic larval stages, they typically have broad geographic ranges and extensive dispersal potential. Their high levels of gene flow across large distances and the rarity of absolute geographic barriers should limit opportunities for population differentiation and the evolution of reproductive isolating barriers (Palumbi, 1994). Furthermore, broadcast spawners such as sea urchins lack behavioral drivers of reproductive isolation, including courtship and mate choice, and often show little morphological, ecological, or physiological divergence. Despite these constraints, species diversity in broadcast spawners appears high. One explanation for the high species richness observed in the absence of obvious physical barriers and ecological divergence is that the evolution of a small number of reproductive proteins may lead to the rapid evolution of reproductive isolation (Palumbi, 1992, 2009).

Many species of broadcast spawners exhibit species-specific fertilization mediated by gamete recognition proteins (GRPs) on the surfaces of sperm and egg cells (Metz et al., 1994; Summers & Hylander, 1975; Vacquier & Moy, 1977). These proteins often evolve rapidly under positive selection and have been implicated in contributing to reproductive isolation (Biermann, 1998; Lee et al., 1995; Lee & Vacquier, 1992; Metz & Palumbi, 1996; Swanson & Vacquier, 2002a, 2002b; Yang et al., 2000). Furthermore, Zigler et al. (2005) found that gametic compatibility in sea urchins correlates with *bindin* divergence but not mitochondrial *COI* divergence. These discoveries led to the hypothesis that speciation may occur when diversifying selection at GRPs produces intrinsic gametic incompatibility, the failure of heterospecific

gametes to bind and fuse (Lessios, 2011; Palumbi, 2009). Several mathematical models have shown that both allopatric and sympatric speciation are theoretically possible when sexual conflict mediated by polyspermy risk drives a coevolutionary chase between the sexes and causes GRP divergence (Gavrilets, 2000; Gavrilets & Hayashi, 2005; Gavrilets & Waxman, 2002; Van Doorn et al., 2001). Although gametic incompatibilities exist, it is unclear whether divergence at reproductive proteins was important early in speciation or instead accumulated after reproductive isolation evolved.

The strongylocentrotid sea urchin family represents a valuable research group for studying the evolution of reproductive isolation. Due to their translucent embryos, sea urchins became model organisms for fertilization studies during the late 19th century. Like many other marine species, sea urchins have large effective population sizes, broad geographic distributions, and limited population structure. The purple sea urchin, *Strongylocentrotus purpuratus* (Stimpson), is a member of the strongylocentrotid family and has a well-annotated reference genome in its fifth major revision. Due to the rarity of natural hybrids and well-documented gametic isolation, it was believed that strongylocentrotid congeners had not shared genes through introgression (Lessios, 2007; Strathmann, 1981). However, recent studies found that asymmetric introgression from *S. pallidus* (Sars) into *S. droebachiensis* (O. F. Müller) has occurred in the Northeast Pacific (Addison & Hart, 2005; Addison & Pogson, 2009; Harper et al., 2007; Pujolar & Pogson, 2011) and Northwest Atlantic (Addison & Hart, 2005; Harper et al., 2007), although only a handful of nuclear and mitochondrial loci were analyzed.

If gametic isolation was an important isolating barrier early in strongylocentrotid speciation, evidence of introgression should not be common and should be negatively correlated with phylogenetic distances and gametic incompatibilities. We tested these predictions using

whole-genome sequencing data and phylogenomic approaches. Given the documented susceptibility of *S. droebachiensis* to heterospecific sperm (Levitan, 2002b) and the previous finding of *S. pallidus* alleles in *S. droebachiensis* samples (Addison & Pogson, 2009), we expected to find a signal of introgression between *S. droebachiensis* and other congeners. Here, we document unexpected, widespread introgression across the strongylocentrotid family at different time scales. Our findings suggest that gametic incompatibilities have not been an effective barrier to introgression and that additional barriers must have been in place for speciation to occur.

2 Materials and Methods

2.1 Study System

The strongylocentrotid phylogeny comprises two major clades: Clade S includes *Strongylocentrotus* and *Hemicentrotus*; Clade M includes *Mesocentrotus* and *Pseudocentrotus*. Both *Hemicentrotus* and *Pseudocentrotus* are monotypic genera. The phylogeny is parsimoniously consistent with a western Pacific common ancestor and at least two independent eastern Pacific invasions (Kober & Bernardi, 2013). Four species are limited to the Northwest Pacific: *P. depressus* (A. Agassiz), *M. nudus* (A. Agassiz), *H. pulcherrimus* (A. Agassiz), and *S. intermedius* (A. Agassiz). Two species, *S. pallidus* and *S. droebachiensis*, are found in, but not limited to, the Northwest Pacific. Five species co-occur in the East Pacific with overlapping geographic ranges, depth preferences, and spawning seasons: *S. droebachiensis*, *S. fragilis* (Jackson), *S. pallidus*, *S. purpuratus*, and *M. franciscanus* (A. Agassiz). *S. droebachiensis* and *S. pallidus* have expanded their ranges, crossing the Bering Sea to colonize the Arctic Ocean and the west and east Atlantic. These two species show little differentiation between the Pacific and

Atlantic Oceans, likely due to stepping-stone populations that facilitate gene flow (Palumbi & Kessing, 1991).

2.2 Whole Genome Resequencing and Data Pre-Processing

Raw fastq sequencing reads for each strongylocentrotid species were downloaded from the NCBI Sequence Read Archive using fasterq-dump from the SRA toolkit (v2.11.2). All samples had been previously sequenced with the Illumina HiSeq 2500. Metadata for the downloaded accessions is available in Table S1.

Sequencing reads were pre-processed with Picard (Broad Institute, 2018) and GATK v4.2.6.1 following GATK's Best Practices (Van der Auwera et al., 2013). Adapter sequences were marked using Picard MarkIlluminaAdapters. Raw reads were mapped to the *S. purpuratus* reference genome (Spur_5.0) using bwa-mem2 v2.2.1 (Vasimuddin et al., 2019), and duplicate reads were marked with Picard MarkDuplicates. Reference mapping and alignment were evaluated using samtools flagstat (Danecek et al., 2021) and mosdepth v0.3.3 (Pedersen & Quinlan, 2018).

Variant calling and joint genotyping were performed using GATK's HaplotypeCaller and GenotypeGVCFs. Variant quality filtering was performed independently for each subset of species used in downstream analyses. Vcf files were hard-filtered for variants with skewed values across all samples following GATK recommendations. SNVs were filtered that had low quality (QUAL<30), low map quality (MQ<40), low quality by depth scores (QD<2), high fisher strand scores (FS > 60), high stand odds ratios (SOR>3), low mapping quality rank sum scores (MQRankSum<-12.5), or low read position rank sum scores (ReadPosRankSum<-8). Indels were filtered that had low quality (QUAL<30), low quality by depth scores (QD<2), high fisher strand

scores (FS>200), or low read position rank sum scores (ReadPosRankSum<-20.0). Individual genotypes with low quality (GQ<20) or low read depth (DP<3) were set to missing, and SNVs within three base pairs of an indel were filtered.

2.3 Phylogenetic Relationships and Concordance Factor Statistics

For phylogenetic inference, multiple sequence alignments were created for protein-coding single-copy orthologs inferred by filtering *S. purpuratus* nuclear gene models by coverage depth. Genes were filtered if any sample had a mean depth lower than 10x, a mean depth greater than double the sample's mean depth for *S. purpuratus* exons, or fewer than 75% of the bases in the gene covered by ten reads. To account for nonindependence among loci, genes were filtered so that there was a minimum of 20kb between included loci. Multiple sequence alignments of concatenated CDS were created for each gene passing filter by applying the hard-filtered SNVs and deletions to the *S. purpuratus* reference sequence using vcf2fasta (Sanchez-Ramirez, 2017/2023). Insertions were ignored to keep gene coordinates consistent with the *S. purpuratus* reference. After creating the fasta alignments, genes were excluded if they had no variant sites, no parsimony informative sites, or if their length was not a multiple of three.

A maximum-likelihood species tree was inferred using the edge-linked partition model of IQTREE (Chernomor et al., 2016; Nguyen et al., 2015) on the concatenated single-copy ortholog fasta alignments. Branch supports were obtained using ultrafast bootstrap with 1,000 replicates (Hoang et al., 2018). Single locus trees were reconstructed for each single-copy ortholog fasta alignment using IQ-TREE's ModelFinder (Kalyaanamoorthy et al., 2017).

Gene concordance factor (gCF) and site concordance factor (sCF) statistics (Minh et al., 2020) were calculated for each branch in the species tree to quantify the amount of phylogenetic

discordance present in the data. For each branch in the species tree, the gCF measures the proportion of gene trees containing that branch, and sCF measures the proportion of informative sites concordant with that branch. The sCFs were calculated by randomly sampling 300 quartets around each internal branch in the phylogeny using an updated version of sCF based on maximum likelihood implemented in IQ-TREE v2.2.2 (Mo et al., 2023). In addition to the gCF and sCF values, IQ-TREE also calculates the frequencies of the two discordant trees produced by nearest-neighbor interchanges (NNI) around each branch. Coalescent theory predicts that the two discordant trees should be equally observed if the discordance is caused by ILS only, but that one tree may become more common if introgression has occurred. To test for introgression, chi-square tests were used to compare counts of the two discordant NNI trees for each branch in the species tree.

2.4 Mitochondrial Phylogenetics

To investigate the relationship between mitochondrial genomes and look for signs of introgression, all complete mitochondrial genome assemblies publicly available for the strongylocentrotid family on NCBI (Table S2) were downloaded and aligned with Clustal Omega v1.2.3 (Sievers et al., 2011; Sievers & Higgins, 2018). A maximum likelihood tree was created with IQTREE using ModelFinder. Branch supports were obtained with ultrafast bootstrap with 10,000 replicates.

2.5 Tests for Introgression

Patterson's *D* Statistic

Patterson's *D* statistic, or the ABBA-BABA test, is the most widely used summary statistic in introgression studies and is robust in a wide parameter space (Kong & Kubatko, 2021; Zheng & Janke, 2018). Patterson's *D* statistic tests for a genome-wide imbalance in the counts of the biallelic site patterns consistent with the two possible discordant topologies in a rooted triplet (Durand et al., 2011; Green et al., 2010). Significance for *D* is calculated using a block jackknife approach that accounts for nonindependence among sites in the data. Patterson's *D* statistic was calculated for all phylogenetically relevant triplets using the genome-wide genotype call set and the Dsuite Dtrios program (Malinsky et al., 2021) with a block-jackknife size of 1 Mb. For comparisons within the S clade, separate tests were run with *M. nudus*, *M. franciscanus*, and *P. depressus* as the outgroup. For comparisons within the M clade, *S. purpuratus* and *S. fragilis* were used as the outgroup. A recent addition to the D statistic, D_P, can approximate the genome-wide introgression proportion (Hamlin et al., 2020) and was calculated for each triplet using the Dsuite output.

Δ Statistic

The Δ statistic is an alternative approach to Patterson's *D* that uses counts of discordant gene tree topologies rather than site patterns (Huson et al., 2005). Δ is less sensitive to the assumption of Patterson's *D* that there have not been multiple substitutions per site (Hahn, 2018) and was therefore used as a secondary measure to confirm significant Patterson's *D* statistic tests where introgression must have occurred between extant taxa. Δ was estimated using gene tree topologies for single-copy orthologs inferred for three different quartets: ((((*M. nudus*, *M. franciscanus*), *P. depressus*), *S. purpuratus*); ((((*S. droebachiensis*, *S. pallidus*), *S. intermedius*),

206 *M. franciscanus*); (((*S. fragilis*, *S. droebachiensis*), *S. pallidus*), *M. franciscanus*). Significance
207 was assessed by calculating Δ for 10,000 pseudoreplicate datasets created by resampling gene
208 tree topologies with replacement (Vanderpool et al., 2020).
209
210 ## PhyloNet
211 The PhyloNet software package implements a powerful set of likelihood methods based
212 on the multispecies network coalescent (MSNC) model (Meng & Kubatko, 2009) that can be
213 used to formally test for introgression (Than et al., 2008; Wen et al., 2018). PhyloNet programs
214 can identify introgression on the internal branches of a phylogeny and reliably infer the direction
215 of introgression (Hibbins & Hahn, 2022). To further characterize the history of introgression
216 within the strongylocentrotid family, we ran PhyloNet's InferNetwork_ML program (Yu et al.,
217 2014) with reconstructed gene tree topologies to infer phylogenetic networks with reticulation
218 edges representing discrete introgression events. A smaller subset of species was used in the
219 PhyloNet analysis due to computational constraints and the requirement that the gene trees be
220 rooted. A new set of single-copy orthologs was inferred for *M. franciscanus*, *H. pulcherrimus*,
221 and the five *Strongylocentrotus* taxa (Table S10). Gene trees were estimated with IQ-TREE2,
222 and 100 bootstrap trees were generated for each locus using standard nonparametric bootstrap to
223 account for uncertainty in gene tree reconstruction. InferNetwork_ML was run to infer
224 phylogenetic networks with 0, 1, 2, and 3 reticulations.

3 Results

Data Pre-processing

The results of the reference genome mapping are summarized in Table 1. The read mapping percentage per sample ranged from 76% to 98%. Mean genome-wide coverage depth typically ranged from 18x - 32x, except for *S. purpuratus* and *S. pallidus*. Coverage depth for *S. pallidus* (12x) was lower because of a reduced library complexity resulting from the early developmental phase of automated library preparation protocols (Kober & Pogson, 2017). *S. purpuratus* was sequenced at a higher depth (91x) for reference genome assembly. Mean coverage depth increased to >38x for protein-coding single-copy orthologs, except for *S. pallidus* (15x). Additional coverage metrics are presented in tables S3-S5.

Phylogenetic Discordance among Strongylocentrotids

Although the inferred strongylocentrotid species relationships agreed with previous studies, the gene and site concordance factor statistics revealed significant phylogenetic discordance. The maximum likelihood species tree inferred from alignments of inferred single copy orthologs agreed with the topology produced by Kober & Bernardi (2013), and all branches had 100% bootstrap support (Figure 1).

The concordance factor analysis revealed extensive phylogenetic discordance on most species tree branches (Figure 1, Table S6). On the two oldest branches (A, B), the concordant resolution was supported by the majority of genes and sites, and the two NNI discordant topologies were nearly equally frequent. The three internal branches relating the *Strongylocentrotus* species all had very low gCF and sCF values. These branches are short, and the lower gCF values than sCF values signal that error in gene tree reconstruction likely

contributed to the signal of phylogenetic discordance. However, the low sCF values suggest that there is not overwhelming support for any single resolution of these branches, implying considerable ILS and introgression.

Although the low gCF values may be partially explained by error in gene tree reconstruction, some nonrandom patterns in the frequencies of the different discordant trees suggest introgression (Table S6). For the branch in the species tree placing *S. purpuratus* as the outgroup to the rest of the *Strongylocentrotus* species (Branch C), the discordant resolution placing *S. intermedius* as the first diverging member of *Strongylocentrotus* (15.9% gene trees, 34.5% sites) was observed more frequently than the other NNI discordant resolution (13.3% gene trees, 29.7% sites, p=0.0015). This pattern may indicate introgression between *S. purpuratus* and one or more of *S. pallidus*, *S. droebachiensis*, *S. fragilis*, or an ancestral lineage. For the branch separating *S. intermedius* from *S. pallidus*, *S. droebachiensis*, and *S. fragilis* (Branch D), the discordant resolution grouping *S. intermedius* and *S. pallidus* as sister taxa (14.7% gene trees, 35.5% sites) was significantly more frequent than the second NNI resolution (8.4% gene trees, 27.1% sites, p<0.0001), consistent with introgression between *S. intermedius* and *S. pallidus*. For the branch placing *S. pallidus* as the outgroup to *S. droebachiensis* and *S. fragilis* (Branch E), the discordant resolution grouping *S. pallidus* and *S. droebachiensis* as sister taxa (21.7% gene trees, 32.7% sites) occurred more frequently than the resolution grouping *S. fragilis* and *S. pallidus* as sister taxa (14.0% gene trees, 29.3% sites, p<0.0001), implying *S. pallidus* ⇔ *S. droebachiensis* introgression. On the branch separating *P. depressus* from the two *Mesocentrotus* species (Branch F), the discordant resolution grouping *P. depressus* with *M. franciscanus* (28.0% gene trees, 31.4% sites) occurred more frequently than the discordant

resolution grouping *P. depressus* and *M. nudus* as sister taxa (20.3% gene trees, 28.07% sites, $p<0.0001$), indicating introgression between *P. depressus* and *M. franciscanus*.

Mitochondrial Introgression

The phylogeny of the mitochondrial genome accessions did not recover the true species relationships and revealed several discordant patterns (Figure 2). The three *M. franciscanus* accessions cluster the two *P. depressus* accessions rather than the two *M. nudus* accessions, consistent with introgression. The tree also reveals introgression and possible mitochondrial capture between *S. droebachiensis* and *S. pallidus*. The three *S. droebachiensis* accessions cluster with the *S. pallidus* accessions rather than its true sister taxa, *S. fragilis*. Additionally, neither the *S. droebachiensis* nor *S. pallidus* accessions are reciprocally monophyletic. The *S. droebachiensis* accession from Svalbard, Norway (EU054306.1) is more closely related to the two *S. pallidus* accessions than it is to the other two *S. droebachiensis* accessions and is placed sister to the *S. pallidus* accession from Norway (NC_009941.1) with 98% bootstrap support, making both *S. droebachiensis* and *S. pallidus* paraphyletic. The short branch lengths are consistent with recent or ongoing hybridization and introgression from *S. pallidus* into *S. droebachiensis* off the coast of Norway. The last source of discordance in the mitochondrial tree is the placement of *S. purpuratus* and *S. intermedius*. In the tree, the positions of *S. purpuratus* and *S. intermedius* are swapped with 62% bootstrap support, consistent with gene flow between *S. purpuratus* and one or more of *S. pallidus*, *S. droebachiensis*, *S. fragilis*, or an ancestral lineage of some or all of the three species. This pattern is consistent with the concordance factor analysis finding that the discordant topology placing *S. intermedius* as the outgroup to the rest of the *Strongylocentrotus* taxa was overrepresented.

Introgression Tests

Patterson's *D* Statistic

Ten species pairs had significant Patterson's *D* statistics, revealing extensive introgression across the phylogeny (Figure 3, Table 2). When testing for introgression in the S clade, tests run with *M. nudus*, *M. franciscanus*, and *P. depressus* as the outgroup returned qualitatively similar results. When testing for introgression in the M clade, tests with *S. fragilis* and *S. purpuratus* as the outgroup were also consistent. The results with *M. nudus* and *S. purpuratus* are displayed (Figure 3, Table 2), and the full results are provided in Tables S7-9. In the M clade, there was support for introgression between *P. depressus* and *M. franciscanus*. In the S clade, there was evidence for introgression between *H. pulcherrimus* and each of *S. intermedius*, *S. pallidus*, *S. droebachiensis*, and *S. fragilis*. There was also support for introgression between *S. purpuratus* and each of *S. pallidus*, *S. fragilis*, and *S. droebachiensis*. Two additional species pairs were implicated in introgression: *S. intermedius* and *S. pallidus*, and *S. pallidus* and *S. droebachiensis*. In cases where a taxon shows introgression with several species that form a monophyletic group, it may be more parsimonious to assume that introgression occurred between that taxon and the MRCA of the monophyletic group, an internal branch in the phylogeny (Suvorov et al., 2022). For example, the significant tests indicating introgression between *H. pulcherrimus* and each of *S. intermedius*, *S. pallidus*, *S. droebachiensis*, and *S. fragilis* could be explained by a single introgression event between *H. pulcherrimus* and the MRCA of *S. intermedius*, *S. pallidus*, *S. droebachiensis*, and *S. fragilis*. Likewise, the significant tests involving *S. purpuratus* could have been produced by a single introgression event between *S. purpuratus* and the MRCA of *S. pallidus*, S. *droebachiensis*, and *S. fragilis*. This would reduce the total number of introgression events from ten to five, a conservative

number because introgression could have occurred both on the internal and terminal branches. Hypotheses of introgression on internal branches are directly testable with phylogenetic network software.

Δ Statistic

Δ was significantly positive for each of the three quartets tested, signaling introgression between *P. depressus* and *M. franciscanus*, *S. intermedius* and *S. pallidus*, and *S. pallidus* and *S. droebachiensis* (Table 3). All three test results were consistent with the estimated Patterson's *D* statistics (Figure 3, Table 2).

PhyloNet

The PhyloNet analysis revealed similar patterns of introgression to the Patterson's *D* and the Δ statistic. Conditioning on the species tree backbone, the one-reticulation edge phylogenetic network with the highest likelihood score implied introgression from *S. purpuratus* into *S. fragilis* (Figure 4b). This is consistent with the *D* statistic with the highest magnitude from the triplet ((*S. intermedius, S. fragilis*), *S. purpuratus*). A similar network with introgression between *S. purpuratus* and the *S. droebachiensis*-*S. fragilis*-*S. pallidus* MRCA had the next highest likelihood (Figure 4c). The best network with two reticulation edges had an additional edge implying introgression from *S. intermedius* into *S. pallidus* (Figure 4d), and the network with three reticulation edges added a third edge indicating introgression from the MRCA of *S. intermedius, S. pallidus, S. droebachiensis*, and *S. fragilis* into *H. pulcherrimus* (Figure 4e).

4 Discussion

Widespread Introgression among Strongylocentrotids

Our study is the first to describe genome-wide patterns of introgression in sea urchins. It is currently believed that limited introgression has occurred among sea urchins, but the results of our study indicate that it may be common, at least within *Strongylocentrotidae*. The ubiquity of introgression among strongylocentrotid taxa suggests that gametic isolation has not been an effective barrier to introgression and may not have played a major role in speciation. More work is needed to characterize additional reproductive barriers in broadcast spawners to allow comparisons to other marine and terrestrial groups.

Our tests for introgression suggest that there have been at least five discrete introgression events within the strongylocentrotid family. The introgression patterns are clear and consistent regardless of the methodology used. Additionally, the mitochondrial genome phylogeny was discordant with the species tree, and the relationships recapitulated the results of the statistical introgression tests. Further introgression events may have gone undetected because it was not possible to test for introgression between the M and S clade members without high-quality sequence data from a close outgroup. Despite considerable phylogenetic discordance in the underlying data, there was strong support for all branches in the strongylocentrotid species tree. This is unsurprising given that these species are well-diverged, with the youngest pair of sister taxa evolving 4-6 million years ago (Kober & Bernardi, 2013). Incomplete lineage sorting is expected to be pervasive in species with high levels of polymorphism, and the five *Strongylocentrotus* taxa speciated relatively rapidly 4-9 mya (Kober & Bernardi, 2013), which resulted in short internal branches. However, incomplete lineage sorting alone is insufficient to explain the discordance patterns.

The *D*, Δ, and gCF/sCF statistics implied introgression between at least three pairs of extant taxa: *S. pallidus* ⇔ *S. droebachiensis*, *S. intermedius* ⇔ *S. pallidus*, and *P. depressus* ⇔ *M. franciscanus*. Introgression between *S. purpuratus* and *S. fragilis* is also likely to have occurred, but the signal could also be explained by introgression on an internal branch. The mitochondrial phylogeny supported the *S. pallidus* ⇔ *S. droebachiensis* and *P. depressus* ⇔ *M. franciscanus* introgression events. The PhyloNet analysis supported introgression between *S. intermedius* and *S. pallidus*, and *S. purpuratus* and *S. fragilis*. Although introgression tests using a single sequence per species typically detect ancient introgression, the mitochondrial genome phylogeny suggests that introgression between *S. pallidus* and *S. droebachiensis* is ongoing in the East Atlantic, evidenced by the paraphyly of both *S. droebachiensis* and *S. pallidus*.

The signal of introgression between *P. depressus* and *M. franciscanus* was unexpected, given that each currently resides on either side of the North Pacific. The M clade phylogeny is consistent with a West Pacific common ancestor, followed by the colonization of the East Pacific by *M. franciscanus*. Introgression must have occurred at a time of range overlap in the distant past, implying that populations of *M. franciscanus* existed in the West Pacific following the speciation of *M. nudus* and *M. franciscanus* 5.6 – 8.1 mya (Kober & Bernardi, 2013).

It was similarly unexpected to find support for introgression between *S. intermedius* and *S. pallidus*, given their current distributions. Although *S. intermedius* and *S. pallidus* co-occur in the Sea of Japan, the *S. pallidus* sample used in this study was from coastal Washington State, indicating that the signal of introgression is ancient. The net direction of gene flow inferred by PhyloNet was from *S. intermedius* into *S. pallidus*, implying that introgression must have occurred before *S. pallidus* expanded its range into the East Pacific. Whether introgression is ongoing between S. intermedius and S. pallidus in the Sea of Japan is unknown.

In addition to introgression between extant taxa, introgression also likely occurred between extant taxa and ancestral lineages (i.e., internal branches). While the optimal phylogenetic network with one reticulation edge implied introgression from *S. purpuratus* into *S. fragilis*, a second network with a similar likelihood supported introgression from the *S. droebachiensis*-*S. fragilis*-*S. pallidus* MRCA into *S. purpuratus*. Both networks are consistent with the Patterson's *D* statistic results, given that there was support for introgression between *S. purpuratus* and each of *S. droebachiensis*, *S. fragilis*, and *S. pallidus*. Both the mitochondrial phylogeny and the concordance factor analysis were also consistent with introgression on an internal branch. In the mitochondrial phylogeny, *S. purpuratus* is pulled down as a sister to the *S. droebachiensis*-*S. fragilis*-*S. pallidus* MRCA and the concordance factor analysis revealed that this topology was overrepresented. A similar potential case of introgression on an internal branch was evidenced by the optimal phylogenetic network with three reticulation edges, which implied introgression between *H. pulcherrimus* and the MRCA of *S. intermedius*, *S. pallidus*, *S. fragilis*, and *S. droebachiensis*. The results of the phylogenetic network analyses underscore the importance of sampling all species of the focal genus or family when testing for introgression. By only sampling a subset of the taxa, introgression may be incorrectly attributed to extant taxa in cases where it occurred on internal branches of the phylogeny. If introgression did occur on an internal branch, there should be considerable overlap in the location of introgressed DNA in each species descendent from that branch.

There are several limitations in the approaches we used to test for introgression. First, it is difficult to quantify the proportion of the genome that is introgressed in each scenario without polymorphism data or population samples of species that are known a priori to have not experienced introgression. Second, evaluating whether introgression is ongoing is difficult

because the methods applied here mainly test for ancient introgression. Third, the geographic history of speciation, hybridization, and introgression remains unclear given the old divergence times of this group, its limited fossil record, and the fact that current ranges of the extant taxa may not be representative of their past distributions. Furthermore, the geographic pattern of hybridization and introgression may be especially complex for marine organisms with high dispersal potential because hybrid zones are more ambiguous.

The introgression signal we observed was likely conservative because only protein-coding regions were used in the gCF/sCF, Δ, and PhyloNet analyses. Introgression is expected to be more common in noncoding regions where it is more likely to be selectively neutral or slightly deleterious. An even stronger signal might have been observed if gene trees from intronic and intergenic regions were also included. These regions were excluded from gene tree-based analyses to reduce false positives from reference alignment and genotyping errors.

Our study adds further representation of marine invertebrates to the rapidly growing evidence for hybridization and introgression and will facilitate investigations into how patterns of introgression vary across different organismal groups. Introgression had long been recognized as a significant evolutionary force in plants (Anderson & Hubricht, 1938; Anderson & Stebbins, 1954) but was only recently appreciated in animals (Hedrick, 2013). Historically, it was thought that introgression between marine taxa was rare (Arnold & Fogarty, 2009) and had not occurred among sea urchins (Lessios, 2007). However, reticulate evolution in marine systems may be as common as that of non-marine taxa (Gardner, 1997), but the difficulty in collecting and observing marine organisms has limited its detection (Arnold & Fogarty, 2009). Although hybridization has been detected in at least five genera of sea urchins (*Diadema*: Lessios & Pearse, 1996; *Pseudoboletia*: Zigler et al., 2012; *Arbacia*: Lessios et al., 2012; *Lytechinus*:

(Zigler & Lessios, 2004); *Strongylocentrotus*: Addison & Pogson, 2009), this is the first study that has tested for introgression among sea urchins with genome-scale data.

A growing body of work has suggested that introgression may be common among other broadcast spawners. Introgression has been detected in reef-building *Acropora* corals (Mao et al., 2018), *Mytilus* mussels (Fraïsse et al., 2016; Popovic et al., 2021; Saarman & Pogson, 2015; Simon et al., 2021; Vendrami et al., 2020), *Ophioderma* brittle stars (Weber et al., 2019), *Asterias* sea stars (Harper & Hart, 2007), Western Pacific *Haliotis* abalone (Hirase et al., 2021), and *Ciona* sea squirts (Nydam et al., 2017; Nydam & Harrison, 2011). Recent research has revealed that speciation with gene flow may be more common in marine than terrestrial environments, underscoring the importance of including more marine organisms in speciation research (Faria et al., 2021; Hirase et al., 2021; Potkamp & Fransen, 2019).

On the relative importance of gametic isolation

It is currently believed that the rapid evolution of gamete recognition proteins (GRPs) is a major contributor to reproductive isolation among broadcast spawners. Although reproductive proteins evolve rapidly under positive selection in a wide variety of taxa (Swanson & Vacquier, 2002a), it remains unclear how often this rapid evolution causes reproductive isolation and speciation (Turner & Hoekstra, 2008). Gametic compatibility can sometimes be maintained for up to 5 million years and is rarely a bi-directional barrier to sea urchin hybridization (McCartney & Lessios, 2004; Zigler et al., 2005). Asymmetric gamete incompatibilities alone cannot prevent gene flow between incipient species (Addison & Pogson, 2009; Lessios, 2011; McCartney & Lessios, 2004), suggesting the importance of additional barriers. Furthermore, *bindin* is not one of the fastest-evolving sea urchin genes and only shows evidence of positive selection in three of

452 the seven sea urchin genera studied to date (Geyer et al., 2020). The drivers of selection at *bindin*
453 are poorly understood and vary across the three groups exhibiting positive selection at this locus.
454 Positive selection at bindin has been observed in the sea urchin genera *Echinometra* (Geyer &
455 Palumbi, 2003; McCartney & Lessios, 2004; Metz & Palumbi, 1996), *Strongylocentrotus*
456 (Biermann, 1998; Pujolar & Pogson, 2011), and *Heliocidaris* (Zigler et al., 2003). However, the
457 driver of selection might be reinforcement in some cases, while in other examples, it's not clear
458 that the selection at *bindin* has established sufficient reproductive isolation for the formation of
459 new species.

460 Within *Strongylocentrotidae*, gametic compatibility *between species* is likely determined
461 by variation in the selective pressures acting on gamete traits *within species* because intraspecific
462 density-dependent selection acting on gametes to maximize fecundity also influences
463 susceptibility to heterospecific fertilization (Levitan, 2002b). Species that more commonly
464 experience sperm-limiting conditions are selected for high fertilization rates and produce eggs
465 that are more readily fertilized by both conspecific and heterospecific sperm (Levitan, 2002a).
466 Conversely, species with higher population densities and high sperm availability likely evolve
467 under sexual conflict and produce faster, more competitive sperm and more sperm-resistant eggs
468 (Levitan, 2002a). This density-dependent selection has likely led to the asymmetric gametic
469 incompatibilities observed between *S. droebachiensis* and other congeners (Hagström &
470 Lönning, 1967; Levitan, 2002b; Strathmann, 1981) and may have also resulted in asymmetric
471 introgression. Under the scenario of density-dependent selection on sperm and egg traits,
472 reproductive isolation between populations should only be strengthened in times or locations of
473 high spawning density. When spawning density is low, and populations experience sperm

limitation, purifying selection to maximize mating opportunities should favor more easily fertilized eggs and prevent divergence of GRPs.

Field experiments on *S. droebachiensis* in the Barkley Sound have demonstrated that conspecific sperm precedence (CSP) is not an effective barrier to hybrid matings when spawning females are closer to heterospecific males than conspecific males (Levitan, 2002b). Hybrid fertilizations readily occur when *S. droebachiensis* eggs are swamped by heterospecific sperm, suggesting that some spatial or temporal isolation during spawning is required for CSP to prevent hybridization. Work in other broadcast spawner groups has shown that reproductive isolation commonly evolves without gamete recognition barriers. For example, in Western Pacific abalones, Hirase et al. (2021) found that ecological divergence evolved before GRP divergence and maintains species barriers despite ongoing hybridization and introgression. In another case, strong reproductive isolation has evolved between the Australian sea urchin species *Pseudoboletia indiana* and *P. maculata* despite only a single amino acid substitution at bindin (Zigler et al., 2012).

Given that introgression has been common among strongylocentrotid urchins, it is unlikely that gametic isolation alone could have been a sufficient barrier to allow speciation to proceed. Other barriers must have been in place and played an important role early in speciation. Lessios (2007) reviewed isolating barriers in sea urchins and concluded that each prezygotic barrier alone appeared incapable of preventing gene flow between sympatric species. Unfortunately, the relative strength of different isolating barriers has rarely been quantified in pairs of sea urchin sister taxa (Palumbi, 2009).

Possible Alternative Isolating Mechanisms

Postzygotic Isolation

How does speciation proceed in high gene flow marine invertebrates with minimal population structure and ecological divergence when geographic barriers are seemingly limited? One possibility is that some postzygotic isolation evolved in allopatry before gametic isolation. There are well-documented cases of hybrid sterility and inviability in interspecific crosses of strongylocentrotids. For example, the *M. nudus* ♀ x *S. intermedius* ♂ cross is lethal (Ding et al., 2007). Although the reciprocal cross produces viable offspring, hybrid larval survival, metamorphosis rates, and juvenile survival are significantly lower than conspecific controls. Furthermore, the surviving juveniles produce very few or no mature gamete cells, a pattern also observed in the *Hemicentrotus pulcherrimus* ♀ x *S. intermedius* ♂ cross (Liu et al., 2020).

In crosses of *S. droebachiensis* x *S. pallidus*, Hagström & Lönning (1967) found that chromosomal abnormalities were frequent during mitosis in embryos of F1 hybrids. Strathmann (1981) performed ten separate reciprocal crosses between *S. droebachiensis* and *S. pallidus*, but only four hybrids survived to the three-year mark when spawning was induced, and all were female. The female hybrids were successfully backcrossed in both directions, although backcross fertilization success was much higher with *S. pallidus* males than with *S. droebachiensis* males. Reduced survival of hybrid juveniles has also been found in crosses of female *S. droebachiensis* with male *S. purpuratus* and *M. franciscanus* (Levitan, 2002b) and the cross between *S. purpuratus* and *M. franciscanus* (Newman, 1923). Postzygotic isolation may be even stronger than these studies suggest because intrinsic postzygotic isolation may not appear until generations beyond the F1 if the alleles that cause intrinsic postzygotic isolation are partially recessive in hybrids (Coyne & Orr, 2004). Reproductive barriers may also result from extrinsic

(i.e., ecological) postzygotic isolation produced by a mismatch between hybrid individuals and their environment.

Chemical Barriers and Carbohydrate-Based Gamete Recognition

The possibility that chemical barriers contribute to reproductive isolation has received limited attention. The egg jelly of broadcast spawners often serves as a chemoattractant to guide conspecific sperm towards the egg, a process called sperm chemotaxis. Conspecific chemoattractant preference has been demonstrated in the abalone species *H. rufescens* and *H. fulgens* (Riffell et al., 2004), although the interaction of gamete recognition proteins is a better predictor of fertilization success in these species (Evans & Sherman, 2013). Sperm chemotaxis has also been described in the sea urchins *Arbacia puctulata* (Ward et al., 1985), *Lytechinus pictus* (Guerrero et al., 2010), and *S. purpuratus* (Ramírez-Gómez et al., 2020).

In sea urchin fertilization, the acrosome reaction is a precondition for the binding of sperm to the egg and may also be species-specific in some cases. Alves et al. (1997) found that sulfated polysaccharides in the egg jelly induce the acrosome reaction in a conspecific manner, although the three species tested were quite divergent (*Echinometra lucunter*, *Arbacia lixula*, and *Lytechinus variegatus*). Biermann et al. (2004) similarly found that the jelly coat of *S. droebachiensis* eggs only induces the acrosome reaction in conspecific sperm due to the rapid evolutionary change in the *S. droebachiensis* egg-jelly fucan. Furthermore, *S. droebachiensis* sperm react with *S. pallidus* and *S. purpuratus* eggs at considerably lower rates than with conspecific eggs. However, the acrosome reaction is not species-specific between *S. purpuratus*, *M. franciscanus*, and *S. pallidus* (Biermann et al., 2004) or between *Echinometra mathaei* and *Echinometra oblonga* (Metz et al., 1994).

Habitat and Temporal Isolation

While differences in habitat preference or spawning time could prevent most heterospecific gamete encounters, sea urchin species' ranges commonly overlap, and it is believed that the cues of spawning cycles are too spatially or temporally variable for spawning asynchrony to be an effective barrier (Lessios, 2007). However, species often show depth zonation in areas of range overlap (Lessios, 2007), and slight differences in the timing and location of gamete release among congeners could prevent heterospecific fertilization as sperm rapidly age, disperse, and become diluted following release (Levitan, 1993; Levitan et al., 2004; Pennington, 1985). A short gap in peak spawning times is an effective reproductive barrier for a pair of Panamanian *Montastraea* reef-building corals (Knowlton et al., 1997) and a pair of Australian subspecies of *Heliocidaris erythrogramma* (Binks et al., 2012). Genetic differences in habitat preference isolate two *Mytilus* mussel species in a contact zone in southern France (Bierne et al., 2003).

Conclusions

Although gametic incompatibilities may help maintain species boundaries in strongylocentrotid urchins, gametic isolation does not appear to have been an effective barrier to introgression. The long persistence of gametic compatibility between divergent taxa and evidence of extensive introgression within the family are inconsistent with the rapid evolution of gametic isolation being an important mode of speciation in this family. Additional isolating barriers likely evolved earlier and were more critical in establishing reproductive isolation. The continued divergence of the strongylocentrotid species in the face of significant introgression emphasizes the importance of postzygotic isolation in maintaining species integrities.

Acknowledgments

We thank UC Santa Cruz for access to the Hummingbird computational cluster. We thank Kord Kober for insightful discussions and Matthew Kustra for his thoughtful comments on the manuscript.

References

Addison, J. A., & Hart, M. W. (2005). Colonization, dispersal, and hybridization influence phylogeography of North Atlantic sea urchins (Strongylocentrotus droebachiensis). *Evolution; International Journal of Organic Evolution*, *59*(3), 532–543.

Addison, J. A., & Pogson, G. H. (2009). Multiple gene genealogies reveal asymmetrical hybridization and introgression among strongylocentrotid sea urchins. *Molecular Ecology*, *18*(6), 1239–1251. https://doi.org/10.1111/j.1365-294X.2009.04094.x

Alves, A.-P., Mulloy, B., Diniz, J. A., & Mourão, P. A. S. (1997). Sulfated Polysaccharides from the Egg Jelly Layer Are Species-specific Inducers of Acrosomal Reaction in Sperms of Sea Urchins. *Journal of Biological Chemistry*, *272*(11), 6965–6971. https://doi.org/10.1074/jbc.272.11.6965

Anderson, E., & Hubricht, L. (1938). Hybridization in Tradescantia. III. The Evidence for Introgressive Hybridization. *American Journal of Botany*, *25*(6), 396–402.

Anderson, E., & Stebbins, G. L. (1954). Hybridization as an Evolutionary Stimulus. *Evolution*, *8*(4), 378–388. https://doi.org/10.2307/2405784

Arnold, M., & Fogarty, N. (2009). Reticulate Evolution and Marine Organisms: The Final Frontier? *International Journal of Molecular Sciences*, *10*(9), 3836–3860. https://doi.org/10.3390/ijms10093836

Biermann, C. H. (1998). The molecular evolution of sperm bindin in six species of sea urchins (Echinoida: Strongylocentrotidae). *Molecular Biology and Evolution*, *15*(12), 1761–1771. https://doi.org/10.1093/oxfordjournals.molbev.a025902

Biermann, C. H., Marks, J. A., Vilela-Silva, A.-C. E. S., Castro, M. O., & Mourao, P. A. S. (2004). Carbohydrate-based species recognition in sea urchin fertilization: Another avenue for speciation? *Evolution and Development*, *6*(5), 353–361. https://doi.org/10.1111/j.1525-142X.2004.04043.x

Bierne, N., Bonhomme, F., & David, P. (2003). Habitat preference and the marine-speciation paradox. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1522), 1399–1406. https://doi.org/10.1098/rspb.2003.2404

Binks, R. M., Prince, J., Evans, J. P., & Kennington, W. J. (2012). MORE THAN BINDIN DIVERGENCE: REPRODUCTIVE ISOLATION BETWEEN SYMPATRIC SUBSPECIES OF A SEA URCHIN BY ASYNCHRONOUS SPAWNING: REPRODUCTIVE BARRIERS BETWEEN SEA URCHIN SUBSPECIES. *Evolution*, *66*(11), 3545–3557. https://doi.org/10.1111/j.1558-5646.2012.01700.x

Broad Institute. (2018). *Picard Tools—By Broad Institute*. http://broadinstitute.github.io/picard/

Campbell, C. R., Poelstra, J. W., & Yoder, A. D. (2018). What is Speciation Genomics? The roles of ecology, gene flow, and genomic architecture in the formation of species. *Biological Journal of the Linnean Society*, *124*(4), 561–583. https://doi.org/10.1093/biolinnean/bly063

Chernomor, O., von Haeseler, A., & Minh, B. Q. (2016). Terrace Aware Data Structure for Phylogenomic Inference from Supermatrices. *Systematic Biology*, *65*(6), 997–1008. https://doi.org/10.1093/sysbio/syw037

Coyne, J. A., & Orr, H. A. (2004). *Speciation*. Oxford University Press.

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*(2), giab008. https://doi.org/10.1093/gigascience/giab008

Ding, J., Chang, Y., Wang, C., & Cao, X. (2007). Evaluation of the growth and heterosis of hybrids among three commercially important sea urchins in China: Strongylocentrotus nudus, S. intermedius and Anthocidaris crassispina. *Aquaculture*, *272*(1–4), 273–280. https://doi.org/10.1016/j.aquaculture.2007.07.231

Durand, E. Y., Patterson, N., Reich, D., & Slatkin, M. (2011). Testing for Ancient Admixture between Closely Related Populations. *Molecular Biology and Evolution*, *28*(8), 2239–2252. https://doi.org/10.1093/molbev/msr048

Edelman, N. B., Frandsen, P. B., Miyagi, M., Clavijo, B., Davey, J., Dikow, R., García-Accinelli, G., Van Belleghem, S. M., Patterson, N., Neafsey, D. E., Challis, R., Kumar, S., Moreira, G. R. P., Salazar, C., Chouteau, M., Counterman, B. A., Papa, R., Blaxter, M., Reed, R. D., … Mallet, J. (2019). Genomic architecture and introgression shape a butterfly radiation. *Science (New York, N.Y.)*, *366*(6465), 594–599. https://doi.org/10.1126/science.aaw2090

Evans, J. P., & Sherman, C. D. H. (2013). Sexual Selection and the Evolution of Egg-Sperm Interactions in Broadcast-Spawning Invertebrates. *The Biological Bulletin*, *224*(3), 166–183. https://doi.org/10.1086/BBLv224n3p166

Faria, R., Johannesson, K., & Stankowski, S. (2021). Speciation in marine environments: Diving under the surface. *Journal of Evolutionary Biology*, *34*(1), 4–15. https://doi.org/10.1111/jeb.13756

Fraïsse, C., Belkhir, K., Welch, J. J., & Bierne, N. (2016). Local interspecies introgression is the main cause of extreme levels of intraspecific differentiation in mussels. *Molecular Ecology*, *25*(1), 269–286. https://doi.org/10.1111/mec.13299

Gardner, J. P. A. (1997). Hybridization in the Sea. In *Advances in Marine Biology* (Vol. 31, pp. 1–78). Elsevier. https://doi.org/10.1016/S0065-2881(08)60221-7

Gavrilets, S. (2000). Rapid evolution of reproductive barriers driven by sexual conflict. *Nature*, *403*, 886–889. https://doi.org/10.1038/35002564

Gavrilets, S., & Hayashi, T. I. (2005). Speciation and Sexual Conflict. *Evolutionary Ecology*, *19*(2), 167–198. https://doi.org/10.1007/s10682-004-7916-4

Gavrilets, S., & Waxman, D. (2002). Sympatric speciation by sexual conflict. *Proceedings of the National Academy of Sciences*, *99*(16), 10533–10538. https://doi.org/10.1073/pnas.152011499

Geyer, L. B., & Palumbi, S. R. (2003). REPRODUCTIVE CHARACTER DISPLACEMENT AND THE GENETICS OF GAMETE RECOGNITION IN TROPICAL SEA URCHINS. *Evolution*, *57*(5), 1049–1060. https://doi.org/10.1111/j.0014-3820.2003.tb00315.x

Geyer, L. B., Zigler, K. S., Tiozzo, S., & Lessios, H. A. (2020). Slow evolution under purifying selection in the gamete recognition protein bindin of the sea urchin Diadema. *Scientific Reports*, *10*(1), 9834. https://doi.org/10.1038/s41598-020-66390-2

[dataset] Glasenapp, M.R., and GH Pogson. 2023. Data from: Extensive Introgression among Strongylocentrotid Sea Urchins Revealed by Phylogenomics. Dryad.

Green, R. E., Krause, J., Briggs, A. W., Maricic, T., Stenzel, U., Kircher, M., Patterson, N., Li, H., Zhai, W., Fritz, M. H.-Y., Hansen, N. F., Durand, E. Y., Malaspinas, A.-S., Jensen, J. D., Marques-Bonet, T., Alkan, C., Prüfer, K., Meyer, M., Burbano, H. A., … Pääbo, S. (2010). A Draft Sequence of the Neandertal Genome. *Science*, *328*(5979), 710–722. https://doi.org/10.1126/science.1188021

Guerrero, A., Nishigaki, T., Carneiro, J., Yoshiro Tatsu, Wood, C. D., & Darszon, A. (2010). Tuning sperm chemotaxis by calcium burst timing. *Developmental Biology*, *344*(1), 52–65. https://doi.org/10.1016/j.ydbio.2010.04.013

Hagström, B. E., & Lönning, S. (1967). Experimental studies of Strongylocentrotus droebachiensis and S. pallidus. *Sarsia*, *29*(1), 165–176. https://doi.org/10.1080/00364827.1967.10411077

Hahn, M. W. (2018). *Molecular Population Genetics*. Oxford University Press.

Hamlin, J. A. P., Hibbins, M. S., & Moyle, L. C. (2020). Assessing biological factors affecting postspeciation introgression. *Evolution Letters*, *4*(2), 137–154. https://doi.org/10.1002/evl3.159

Harper, F. M., Addison, J. A., & Hart, M. W. (2007). Introgression Versus Immigration in Hybridizing High-Dispersal Echinoderms. *Evolution*, *61*(10), 2410–2418. https://doi.org/10.1111/j.1558-5646.2007.00200.x

Harper, F. M., & Hart, M. W. (2007). Morphological and phylogenetic evidence for hybridization and introgression in a sea star secondary contact zone: Hybridization between Asterias sea stars. *Invertebrate Biology*, *126*(4), 373–384. https://doi.org/10.1111/j.1744-7410.2007.00107.x

Hedrick, P. W. (2013). Adaptive introgression in animals: Examples and comparison to new mutation and standing variation as sources of adaptive variation. *Molecular Ecology*, *22*(18), 4606–4618. https://doi.org/10.1111/mec.12415

Hibbins, M. S., & Hahn, M. W. (2022). Phylogenomic approaches to detecting and characterizing introgression. *Genetics*, *220*(2), iyab173. https://doi.org/10.1093/genetics/iyab173

Hirase, S., Yamasaki, Y. Y., Sekino, M., Nishisako, M., Ikeda, M., Hara, M., Merilä, J., & Kikuchi, K. (2021). Genomic Evidence for Speciation with Gene Flow in Broadcast Spawning Marine Invertebrates. *Molecular Biology and Evolution*, *38*(11), 4683–4699. https://doi.org/10.1093/molbev/msab194

Hoang, D. T., Chernomor, O., von Haeseler, A., Minh, B. Q., & Vinh, L. S. (2018). UFBoot2: Improving the Ultrafast Bootstrap Approximation. *Molecular Biology and Evolution*, *35*(2), 518–522. https://doi.org/10.1093/molbev/msx281

Huson, D. H., Klöpper, T., Lockhart, P. J., & Steel, M. A. (2005). Reconstruction of Reticulate Networks from Gene Trees. In S. Miyano, J. Mesirov, S. Kasif, S. Istrail, P. A. Pevzner, & M. Waterman (Eds.), *Research in Computational Molecular Biology* (pp. 233–249). Springer. https://doi.org/10.1007/11415770_18

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K. F., von Haeseler, A., & Jermiin, L. S. (2017). ModelFinder: Fast model selection for accurate phylogenetic estimates. *Nature Methods*, *14*(6), 587–589. https://doi.org/10.1038/nmeth.4285

Knowlton, N., Maté, J. L., Guzmán, H. M., Rowan, R., & Jara, J. (1997). Direct evidence for reproductive isolation among the three species of the Montastraea annularis complex in Central America (Panamá and Honduras). *Marine Biology*, *127*(4), 705–711. https://doi.org/10.1007/s002270050061

Kober, K. M., & Bernardi, G. (2013). Phylogenomics of strongylocentrotid sea urchins. *BMC Evolutionary Biology*, *13*(1), 88. https://doi.org/10.1186/1471-2148-13-88

Kober, K. M., & Pogson, G. H. (2017). Genome-wide signals of positive selection in strongylocentrotid sea urchins. *BMC Genomics*, *18*(1), 555. https://doi.org/10.1186/s12864-017-3944-7

Kong, S., & Kubatko, L. S. (2021). Comparative Performance of Popular Methods for Hybrid Detection using Genomic Data. *Systematic Biology*, *70*(5), 891–907. https://doi.org/10.1093/sysbio/syaa092

Larson, E. L., Vanderpool, D., Sarver, B. A. J., Callahan, C., Keeble, S., Provencio, L. L., Kessler, M. D., Stewart, V., Nordquist, E., Dean, M. D., & Good, J. M. (2018). The Evolution of Polymorphic Hybrid Incompatibilities in House Mice. *Genetics*, *209*(3), 845–859. https://doi.org/10.1534/genetics.118.300840

Lee, Y. H., Ota, T., & Vacquier, V. (1995). Positive selection is a general phenomenon in the evolution of abalone sperm lysin. *Molecular Biology and Evolution*, *12*(2), 231–238. https://doi.org/10.1093/oxfordjournals.molbev.a040200

Lee, Y. H., & Vacquier, V. D. (1992). The Divergence of Species-Specific Abalone Sperm Lysins is Promoted by Positive Darwinian Selection. *The Biological Bulletin*, *182*(1), 97–104. https://doi.org/10.2307/1542183

Lessios, H. A. (2007). REPRODUCTIVE ISOLATION BETWEEN SPECIES OF SEA URCHINS. *BULLETIN OF MARINE SCIENCE*, *81*(2), 191–208.

Lessios, H. A. (2011). Speciation Genes in Free-Spawning Marine Invertebrates. *Integrative and Comparative Biology*, *51*(3), 456–465. https://doi.org/10.1093/icb/icr039

Lessios, H. A., Lockhart, S., Collin, R., Sotil, G., Sanchez-Jerez, P., Zigler, K. S., Perez, A. F., Garrido, M. J., Geyer, L. B., Bernardi, G., Vacquier, V. D., Haroun, R., & Kessing, B. D. (2012). Phylogeography and bindin evolution in Arbacia, a sea urchin genus with an unusual distribution. *Molecular Ecology*, *21*(1), 130–144. https://doi.org/10.1111/j.1365-294X.2011.05303.x

Lessios, H. A., & Pearse, J. S. (1996). Hybridization and introgression between Indo-Pacific species of Diadema. *Marine Biology*, *126*(4), 715–723. https://doi.org/10.1007/BF00351338

Levitan, D. R. (1993). The Importance of Sperm Limitation to the Evolution of Egg Size in Marine Invertebrates. *The American Naturalist*, *141*(4), 517–536. https://doi.org/10.1086/285489

Levitan, D. R. (2002a). DENSITY-DEPENDENT SELECTION ON GAMETE TRAITS IN THREE CONGENERIC SEA URCHINS. *Ecology*, *83*(2), 464–479. https://doi.org/10.1890/0012-9658(2002)083[0464:DDSOGT]2.0.CO;2

Levitan, D. R. (2002b). THE RELATIONSHIP BETWEEN CONSPECIFIC FERTILIZATION SUCCESS AND REPRODUCTIVE ISOLATION AMONG THREE CONGENERIC SEA URCHINS. *Evolution*, *56*(8), 1599–1689. https://doi.org/10.1111/j.0014-3820.2002.tb01472.x

Levitan, D. R., Fukami, H., Jara, J., Kline, D., McGovern, T. M., McGhee, K. E., Swanson, C. A., & Knowlton, N. (2004). MECHANISMS OF REPRODUCTIVE ISOLATION AMONG SYMPATRIC BROADCAST-SPAWNING CORALS OF THE MONTASTRAEA ANNULARIS SPECIES COMPLEX. *Evolution*, *58*(2), 308–323. https://doi.org/10.1111/j.0014-3820.2004.tb01647.x

Liu, L., Sun, J., Zhan, Y., Zhao, T., Zou, Y., Yan, H., Zhang, W., & Chang, Y. (2020). Gonadal traits and nutrient compositions of novel sea urchin hybrids of Hemicentrotus pulcherrimus (♀) and Strongylocentrotus intermedius (♂). *Aquaculture Reports*, *18*, 100439. https://doi.org/10.1016/j.aqrep.2020.100439

718 Malinsky, M., Challis, R. J., Tyers, A. M., Schiffels, S., Terai, Y., Ngatunga, B. P., Miska, E. A., Durbin, R.,
719 Genner, M. J., & Turner, G. F. (2015). Genomic islands of speciation separate cichlid ecomorphs in an East
720 African crater lake. *Science*, *350*(6267), 1493–1498. https://doi.org/10.1126/science.aac9927
721 Malinsky, M., Matschiner, M., & Svardal, H. (2021). Dsuite—Fast D-statistics and related admixture evidence from
722 VCF files. *Molecular Ecology Resources*, *21*(2), 584–595. https://doi.org/10.1111/1755-0998.13265
723 Mao, Y., Economo, E. P., & Satoh, N. (2018). The Roles of Introgression and Climate Change in the Rise to
724 Dominance of Acropora Corals. *Current Biology*, *28*(21), 3373-3382.e5.
725 https://doi.org/10.1016/j.cub.2018.08.061
726 McCartney, M. A., & Lessios, H. A. (2004). Adaptive Evolution of Sperm Bindin Tracks Egg Incompatibility in
727 Neotropical Sea Urchins of the Genus Echinometra. *Molecular Biology and Evolution*, *21*(4), 732–745.
728 https://doi.org/10.1093/molbev/msh071
729 Meng, C., & Kubatko, L. S. (2009). Detecting hybrid speciation in the presence of incomplete lineage sorting using
730 gene tree incongruence: A model. *Theoretical Population Biology*, *75*(1), 35–45.
731 https://doi.org/10.1016/j.tpb.2008.10.004
732 Metz, E. C., Kane, R. E., Yanagimachi, H., & Palumbi, S. R. (1994). Fertilization Between Closely Related Sea
733 Urchins Is Blocked by Incompatibilities During Sperm-Egg Attachment and Early Stages of Fusion. *The*
734 *Biological Bulletin*, *187*(1), 23–34. https://doi.org/10.2307/1542162
735 Metz, E. C., & Palumbi, S. R. (1996). Positive selection and sequence rearrangements generate extensive
736 polymorphism in the gamete recognition protein bindin. *Molecular Biology and Evolution*, *13*(2), 397–406.
737 https://doi.org/10.1093/oxfordjournals.molbev.a025598
738 Minh, B. Q., Hahn, M. W., & Lanfear, R. (2020). New Methods to Calculate Concordance Factors for
739 Phylogenomic Datasets. *Molecular Biology and Evolution*, *37*(9), 2727–2733.
740 https://doi.org/10.1093/molbev/msaa106
741 Mo, Y. K., Lanfear, R., Hahn, M. W., & Minh, B. Q. (2023). Updated site concordance factors minimize effects of
742 homoplasy and taxon sampling. *Bioinformatics*, *39*(1), btac741.
743 https://doi.org/10.1093/bioinformatics/btac741
744 Newman, H. H. (1923). Hybrid vigor, hybrid weakness, and the chromosome theory of heredity. An experimental
745 analysis of the physiology of heredity in the reciprocal crosses between two closely associated species of
746 sea-urchins, Strongylocentrotus purpuratus and S. franciscanus. *Journal of Experimental Zoology*, *37*(2),
747 169–205. https://doi.org/10.1002/jez.1400370203
748 Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A Fast and Effective Stochastic
749 Algorithm for Estimating Maximum-Likelihood Phylogenies. *Molecular Biology and Evolution*, *32*(1),
750 268–274. https://doi.org/10.1093/molbev/msu300
751 Nydam, M. L., & Harrison, R. G. (2011). INTROGRESSION DESPITE SUBSTANTIAL DIVERGENCE IN A
752 BROADCAST SPAWNING MARINE INVERTEBRATE: INTROGRESSION AND DIVERGENCE IN
753 CIONA INTESTINALIS. *Evolution*, *65*(2), 429–442. https://doi.org/10.1111/j.1558-5646.2010.01153.x
754 Nydam, M. L., Yanckello, L. M., Bialik, S. B., Giesbrecht, K. B., Nation, G. K., & Peak, J. L. (2017). Introgression
755 in two species of broadcast spawning marine invertebrate. *Biological Journal of the Linnean Society*,
756 *120*(4), 879–890. https://doi.org/10.1093/biolinnean/blw012
757 Orr, H. A. (2005). The genetic basis of reproductive isolation: Insights from Drosophila. *Proceedings of the*
758 *National Academy of Sciences*, *102*(suppl_1), 6522–6526. https://doi.org/10.1073/pnas.0501893102
759 Palumbi, S. R. (1992). Marine speciation on a small planet. *Trends in Ecology & Evolution*, *7*(4), 114–118.
760 https://doi.org/10.1016/0169-5347(92)90144-Z
761 Palumbi, S. R. (1994). Genetic Divergence, Reproductive Isolation, and Marine Speciation. *Annual Review of*
762 *Ecology and Systematics*, *25*(1), 547–572. https://doi.org/10.1146/annurev.es.25.110194.002555
763 Palumbi, S. R. (2009). Speciation and the evolution of gamete recognition genes: Pattern and process. *Heredity*,
764 *102*(1), 66–76. https://doi.org/10.1038/hdy.2008.104
765 Palumbi, S. R., & Kessing, B. D. (1991). Population Biology of the Trans-Arctic Exchange: MtDNA Sequence
766 Similarity between Pacific and Atlantic Sea Urchins. *Evolution*, *45*(8), 1790–1805.
767 Pedersen, B. S., & Quinlan, A. R. (2018). Mosdepth: Quick coverage calculation for genomes and exomes.
768 *Bioinformatics*, *34*(5), 867–868. https://doi.org/10.1093/bioinformatics/btx699
769 Pennington, J. T. (1985). THE ECOLOGY OF FERTILIZATION OF ECHINOID EGGS: THE CONSEQUENCES
770 OF SPERM DILUTION, ADULT AGGREGATION, AND SYNCHRONOUS SPAWNING. *The*
771 *Biological Bulletin*, *169*(2), 417–430. https://doi.org/10.2307/1541492

Poelstra, J. W., Vijay, N., Bossu, C. M., Lantz, H., Ryll, B., Müller, I., Baglione, V., Unneberg, P., Wikelski, M., Grabherr, M. G., & Wolf, J. B. W. (2014). The genomic landscape underlying phenotypic integrity in the face of gene flow in crows. *Science*, *344*(6190), 1410–1414. https://doi.org/10.1126/science.1253226

Popovic, I., Bierne, N., Gaiti, F., Tanurdžić, M., & Riginos, C. (2021). Pre-introduction introgression contributes to parallel differentiation and contrasting hybridization outcomes between invasive and native marine mussels. *Journal of Evolutionary Biology*, *34*(1), 175–192. https://doi.org/10.1111/jeb.13746

Potkamp, G., & Fransen, C. H. J. M. (2019). Speciation with gene flow in marine systems. *Contributions to Zoology*, *88*(2), 133–172. https://doi.org/10.1163/18759866-20191344

Powell, D. L., García-Olazábal, M., Keegan, M., Reilly, P., Du, K., Díaz-Loyo, A. P., Banerjee, S., Blakkan, D., Reich, D., Andolfatto, P., Rosenthal, G. G., Schartl, M., & Schumer, M. (2020). Natural hybridization reveals incompatible alleles that cause melanoma in swordtail fish. *Science*, *368*(6492), 731–736. https://doi.org/DOI: 10.1126/science.aba521

Presgraves, D. C. (2010). The molecular evolutionary basis of species formation. *Nature Reviews Genetics*, *11*(3), 175–180. https://doi.org/10.1038/nrg2718

Pujolar, J. M., & Pogson, G. H. (2011). Positive Darwinian selection in gamete recognition proteins of Strongylocentrotus sea urchins. *Molecular Ecology*, *20*(23), 4968–4982. https://doi.org/10.1111/j.1365-294X.2011.05336.x

Ramírez-Gómez, H. V., Jimenez Sabinina, V., Velázquez Pérez, M., Beltran, C., Carneiro, J., Wood, C. D., Tuval, I., Darszon, A., & Guerrero, A. (2020). Sperm chemotaxis is driven by the slope of the chemoattractant concentration field. *ELife*, *9*, e50532. https://doi.org/10.7554/eLife.50532

Riffell, J. A., Krug, P. J., & Zimmer, R. K. (2004). The ecological and evolutionary consequences of sperm chemoattraction. *Proceedings of the National Academy of Sciences*, *101*(13), 4501–4506. https://doi.org/10.1073/pnas.0304594101

Saarman, N. P., & Pogson, G. H. (2015). Introgression between invasive and native blue mussels (genus *Mytilus*) in the central California hybrid zone. *Molecular Ecology*, *24*(18), 4723–4738. https://doi.org/10.1111/mec.13340

Sanchez-Ramirez, S. (2023). *Vcf2fasta* [Python]. https://github.com/santiagosnchez/vcf2fasta (Original work published 2017)

Schluter, D., & Rieseberg, L. H. (2022). Three problems in the genetics of speciation by selection. *Proceedings of the National Academy of Sciences*, *119*(30), e2122153119. https://doi.org/10.1073/pnas.2122153119

Sievers, F., & Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sequences. *Protein Science*, *27*(1), 135–145. https://doi.org/10.1002/pro.3290

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J. D., & Higgins, D. G. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Molecular Systems Biology*, *7*(1), 539. https://doi.org/10.1038/msb.2011.75

Simon, A., Fraïsse, C., El Ayari, T., Liautard-Haag, C., Strelkov, P., Welch, J. J., & Bierne, N. (2021). How do species barriers decay? Concordance and local introgression in mosaic hybrid zones of mussels. *Journal of Evolutionary Biology*, *34*(1), 208–223. https://doi.org/10.1111/jeb.13709

Strathmann, R. R. (1981). On barriers to hybridization between Strongylocentrotus droebachiensis (O.F. Müller) and S. Pallidus (G.O. Sars). *Journal of Experimental Marine Biology and Ecology*, *55*(1), 39–47. https://doi.org/10.1016/0022-0981(81)90091-5

Summers, R. G., & Hylander, B. L. (1975). Species-specificity of acrosome reaction and primary gamete binding in echinoids. *Experimental Cell Research*, *96*(1), 63–68. https://doi.org/10.1016/S0014-4827(75)80037-1

Swanson, W. J., & Vacquier, V. D. (2002a). The rapid evolution of reproductive proteins. *Nature Reviews Genetics*, *3*(2), 137–144. https://doi.org/10.1038/nrg733

Swanson, W. J., & Vacquier, V. D. (2002b). Reproductive Protein Evolution. *Annual Review of Ecology and Systematics*, *33*(1), 161–179. https://doi.org/10.1146/annurev.ecolsys.33.010802.150439

Than, C., Ruths, D., & Nakhleh, L. (2008). PhyloNet: A software package for analyzing and reconstructing reticulate evolutionary relationships. *BMC Bioinformatics*, *9*(1), 322. https://doi.org/10.1186/1471-2105-9-322

Turner, L. M., & Hoekstra, H. E. (2008). Causes and consequences of the evolution of reproductive proteins. *The International Journal of Developmental Biology*, *52*(5–6), 769–780. https://doi.org/10.1387/ijdb.082577lt

Vacquier, V. D., & Moy, G. W. (1977). Isolation of bindin: The protein responsible for adhesion of sperm to sea urchin eggs. *Proceedings of the National Academy of Sciences*, *74*(6), 2456–2460. https://doi.org/10.1073/pnas.74.6.2456

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., Banks, E., Garimella, K. V., Altshuler, D., Gabriel, S., & DePristo, M. A. (2013). From FastQ Data to High-Confidence Variant Calls: The Genome Analysis Toolkit Best Practices Pipeline. *Current Protocols in Bioinformatics*, *43*(1). https://doi.org/10.1002/0471250953.bi1110s43

Van Doorn, G. S., Luttikhuizen, P. C., & Weissing, F. J. (2001). Sexual selection at the protein level drives the extraordinary divergence of sex–related genes during sympatric speciation. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *268*(1481), 2155–2161. https://doi.org/10.1098/rspb.2001.1780

Vanderpool, D., Minh, B. Q., Lanfear, R., Hughes, D., Murali, S., Harris, R. A., Raveendran, M., Muzny, D. M., Hibbins, M. S., Williamson, R. J., Gibbs, R. A., Worley, K. C., Rogers, J., & Hahn, M. W. (2020). Primate phylogenomics uncovers multiple rapid radiations and ancient interspecific introgression. *PLOS Biology*, *18*(12), e3000954. https://doi.org/10.1371/journal.pbio.3000954

Vasimuddin, Md., Misra, S., Li, H., & Aluru, S. (2019). Efficient Architecture-Aware Acceleration of BWA-MEM for Multicore Systems. *2019 IEEE International Parallel and Distributed Processing Symposium (IPDPS)*, 314–324. https://doi.org/10.1109/IPDPS.2019.00041

Vendrami, D. L. J., De Noia, M., Telesca, L., Brodte, E., & Hoffman, J. I. (2020). Genome-wide insights into introgression and its consequences for genome-wide heterozygosity in the Mytilus species complex across Europe. *Evolutionary Applications*, *13*(8), 2130–2142. https://doi.org/10.1111/eva.12974

Ward, G. E., Brokaw, C. J., Garbers, D. L., & Vacquier, V. D. (1985). Chemotaxis of Arbacia punctulata spermatozoa to resact, a peptide from the egg jelly layer. *The Journal of Cell Biology*, *101*(6), 2324–2329. https://doi.org/10.1083/jcb.101.6.2324

Weber, A. A.-T., Stöhr, S., & Chenuil, A. (2019). Species delimitation in the presence of strong incomplete lineage sorting and hybridization: Lessons from Ophioderma (Ophiuroidea: Echinodermata). *Molecular Phylogenetics and Evolution*, *131*, 138–148. https://doi.org/10.1016/j.ympev.2018.11.014

Wen, D., Yu, Y., Zhu, J., & Nakhleh, L. (2018). Inferring Phylogenetic Networks Using PhyloNet. *Systematic Biology*, *67*(4), 735–740. https://doi.org/10.1093/sysbio/syy015

Yang, Z., Swanson, W. J., & Vacquier, V. D. (2000). Maximum-Likelihood Analysis of Molecular Adaptation in Abalone Sperm Lysin Reveals Variable Selective Pressures Among Lineages and Sites. *Molecular Biology and Evolution*, *17*(10), 1446–1455. https://doi.org/10.1093/oxfordjournals.molbev.a026245

Yu, Y., Dong, J., Liu, K. J., & Nakhleh, L. (2014). Maximum likelihood inference of reticulate evolutionary histories. *Proceedings of the National Academy of Sciences*, *111*(46), 16448–16453. https://doi.org/10.1073/pnas.1407950111

Zheng, Y., & Janke, A. (2018). Gene flow analysis method, the D-statistic, is robust in a wide parameter space. *BMC Bioinformatics*, *19*(1), 10. https://doi.org/10.1186/s12859-017-2002-4

Zigler, K. S., Byrne, M., Raff, E. C., Lessios, H. A., & Raff, R. A. (2012). NATURAL HYBRIDIZATION IN THE SEA URCHIN GENUS PSEUDOBOLETIA BETWEEN SPECIES WITHOUT APPARENT BARRIERS TO GAMETE RECOGNITION. *Evolution*, *66*(6), 1695–1708. https://doi.org/10.1111/j.1558-5646.2012.01609.x

Zigler, K. S., & Lessios, H. A. (2004). SPECIATION ON THE COASTS OF THE NEW WORLD: PHYLOGEOGRAPHY AND THE EVOLUTION OF BINDIN IN THE SEA URCHIN GENUS LYTECHINUS. *Evolution*, *58*(6), 1225–1241. https://doi.org/10.1111/j.0014-3820.2004.tb01702.x

Zigler, K. S., McCartney, M. A., Levitan, D. R., & Lessios, H. A. (2005). Sea Urchin Bindin Divergence Predicts Gamete Compatibility. *Evolution*, *59*(11), 2399–2404. https://doi.org/10.1111/j.0014-3820.2005.tb00949.x

Zigler, K. S., Raff, E. C., Popodi, E., Raff, R. A., & Lessios, H. A. (2003). ADAPTIVE EVOLUTION OF BINDIN IN THE GENUS HELIOCIDARIS IS CORRELATED WITH THE SHIFT TO DIRECT DEVELOPMENT. *Evolution*, *57*(10), 2293–2302. https://doi.org/10.1111/j.0014-3820.2003.tb00241.x

Zuellig, M. P., & Sweigart, A. L. (2018). Gene duplicates cause hybrid lethality between sympatric species of Mimulus. *PLOS Genetics*, *14*(4), e1007130. https://doi.org/10.1371/journal.pgen.1007130

Data Accessibility and Benefit-Sharing

Data Accessibility Statement

The data and code that support the findings of this study are available on Dryad (https://doi.org/10.7291/D1BT34). Raw sequence reads are available in the NCBI SRA (BioProject PRJNA391452).

Author Contributions

Matthew R. Glasenapp and Grant H. Pogson designed the research. Matthew R, Glasenapp performed the research, analyzed the data, and wrote the manuscript.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Funding

Funding for data collection for the study was provided by the National Science Foundation (DEB-1011061), the STEPS Foundation, Friends of Long Marine Lab, and the Myers Trust. The funding bodies did not participate in research design, sample collection, data analysis, or manuscript writing.

Tables and Figures

Table 1. Summary of genomic DNA sequencing, reference mapping, and coverage.

Species	Reference Mapping			% Bases Covered			Mean Coverage Depth		
	Raw Reads	Mapped%	Proper Pair %	Whole Genome[a]	Coding[b]	Single Copy Orthologs 10x[c]	Whole Genome[d]	Coding[e]	Single Copy Orthologs[f]
Sdro	3.04E+08	91.74%	78.11%	78%	92%	0.97	24.7x	41.5x	42.5x
Sfra	3.97E+08	89.87%	78.21%	81%	93%	0.97	32.1x	46.8x	48.2x
Spal	1.50E+08	91.82%	72.39%	78%	91%	0.97	11.9x	15x	15.5x
Sint	4.01E+08	84.24%	73.06%	77%	91%	0.97	28.3x	44.2x	50.3x
Spur	6.21E+08	98.11%	89.04%	99%	100%	0.99	91.3x	100.3x	108.2x
Hpul	3.76E+08	82.71%	68.67%	69%	86%	0.95	24.5x	44.3x	53.3x
Mnud	3.82E+08	77.00%	63.08%	58%	82%	0.92	21.1x	40.5x	45.3x
Mfra	3.39E+08	80.36%	64.30%	60%	84%	0.93	19.9x	33.8x	38.3x
Pdep	3.28E+08	76.17%	60.79%	50%	77%	0.89	18.1x	47.5x	53.5x

[a]Percentage of bases in the *S. purpuratus* reference genome covered with at least one read

[b]Percentage of coding bases in the *S. purpuratus* reference genome covered with at least one read

[c]Percentage of single copy ortholog coding bases covered at 10x depth

[d]Mean genome-wide coverage depth of the *S. purpuratus* reference genome

[e]Mean coverage depth for 246,202 unique exons in the *S. purpuratus* genome assembly

[f]Mean coverage depth of coding bases for 4,497 single-copy orthologs

Species abbreviations: *Sfra - S. fragilis*; *Sdro - S. droebachiensis*; *Spal - S. pallidus*; *Spur - S. purpuratus*; *Hpul - H. pulcherrimus*; *Mnud - M. nudus*; *Mfra - M. franciscanus*; *Pdep - P. depressus*.

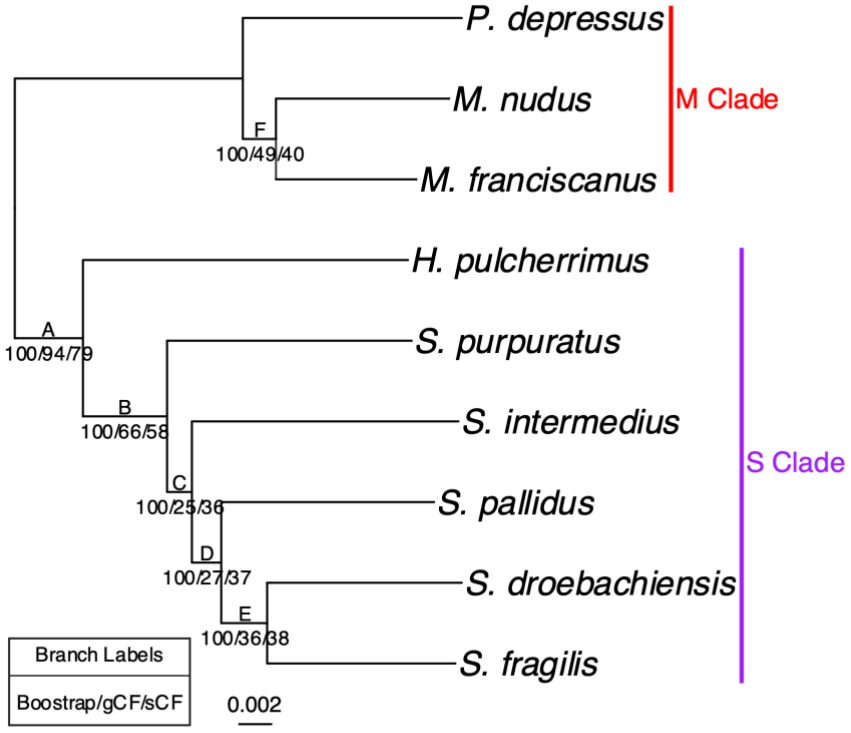

Figure 1. Phylogeny of nine species of strongylocentrotid urchins included in the study. A maximum-likelihood species tree was inferred using the edge-linked partition model of IQTREE (Nguyen et al. 2015; Chernomor et al. 2016) on 4,497 concatenated single-copy ortholog alignments. Gene concordance factor (gCF) and site concordance factor (sCF) statistics (Minh et al., 2020; Mo et al., 2022) were calculated using IQ-TREEv2.2.2. For each branch in the species tree, the gCF measures the proportion of gene trees containing that branch, and sCF measures the proportion of informative sites concordant with that branch (Minh et al., 2020).

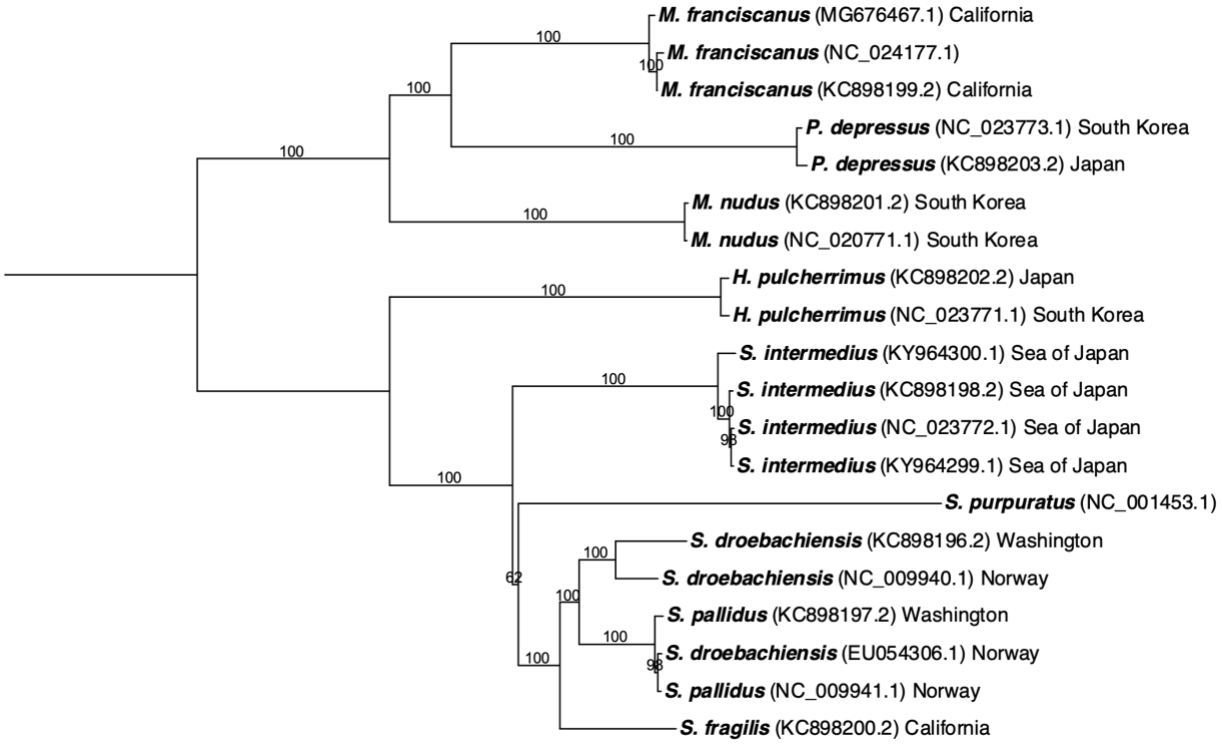

Figure 2. A maximum likelihood tree was inferred from alignments of the strongylocentrotid mitochondrion genome assemblies available on NCBI. The tree is rooted at the midpoint. The branches are labeled with the species name followed by the NCBI accession number. The sampling location is provided for samples with known locations. The mitochondrial genomes were aligned using Clustal Omega v1.2.3, and a maximum likelihood tree was constructed using IQTREE (Nguyen et al., 2015) and ModelFinder (Kalyaanamoorthy et al., 2017). Branch supports were obtained using ultrafast bootstrap (Hoang et al., 2018) with 1,000 replicates. Relative to the true species relationships (Figure 1), the placements of the following are swapped: (i) *M. nudus* and *P. depressus*, (ii) *S. purpuratus* and *S. intermedius*, and (iii) *S. pallidus* and *S. fragilis*.

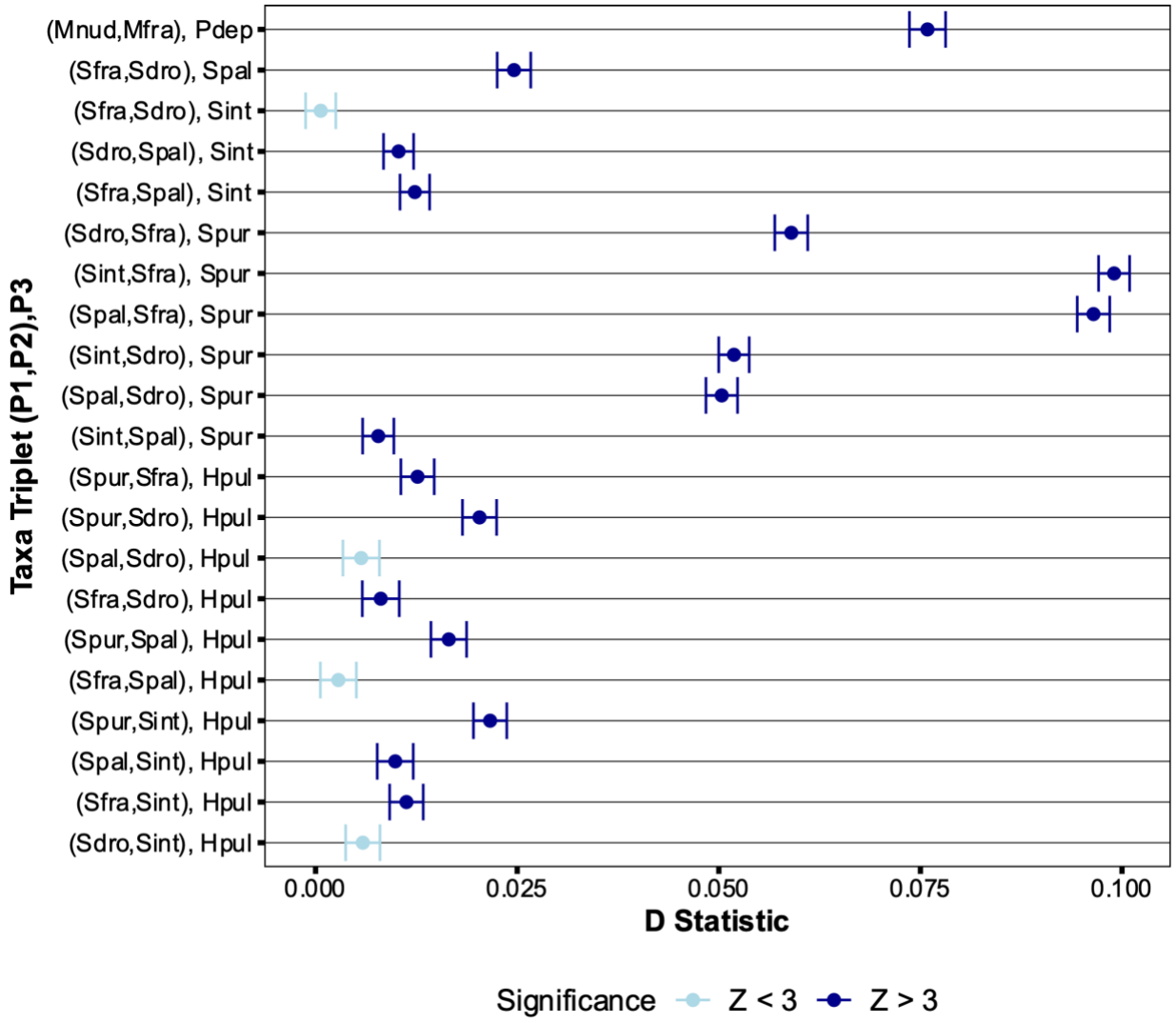

Figure 3. Results of ABBA-BABA tests. *M. nudus* was used as the outgroup. Equal numbers of ABBA and BABA sites are expected under the null hypothesis of no introgression (D = 0). A positive D statistic indicates introgression between P3 and P2. Significance was assessed using a block jackknife size of 1Mb. Species abbreviations: *Sfra - S. fragilis*; *Sdro - S. droebachiensis*; *Spal - S. pallidus*; *Spur - S. purpuratus*; *Hpul - H. pulcherrimus*; *Mnud - M. nudus*; *Mfra - M. franciscanus*; *Pdep - P. depressus*.

Table 2. Results of ABBA BABA tests with Dsuite. Tests are grouped by P3 taxa. Equal numbers of ABBA and BABA sites are expected under the null hypothesis of no introgression (D = 0). A positive D statistic indicates introgression between P3 and P2. Significance was assessed using a block jackknife size of 1Mb. Species abbreviations: *Sfra - S. fragilis*; *Sdro - S. droebachiensis*; *Spal - S. pallidus*; *Spur - S. purpuratus*; *Hpul - H. pulcherrimus*; *Mnud - M. nudus*; *Mfra - M. franciscanus*; *Pdep - P. depressus*.

Samples			Dsuite						
P1	**P2**	**P3**	**D**	**z**	**p**	**D_P**	**BBAA**	**ABBA**	**BABA**
Mnud	Mfra	Pdep	0.076	33.8	0.000	0.040	240,218	144,747	124,331
Sfra	Sdro	Spal	0.025	11.8	0.000	0.013	319,896	185,499	176,591
Sfra	Sdro	Sint	0.001	0.3	0.735	0.000	427,693	185,058	184,824
Sdro	Spal	Sint	0.010	5.5	0.000	0.006	249,986	187,513	183,693
Sfra	Spal	Sint	0.012	6.7	0.000	0.007	250,248	194,472	189,743
Sdro	Sfra	Spur	0.059	28.9	0.000	0.026	490,027	200,788	178,420
Sint	Sfra	Spur	0.099	51.5	0.000	0.062	289,884	271,623	222,678
Spal	Sfra	Spur	0.096	47.9	0.000	0.055	292,707	210,001	173,050
Sint	Sdro	Spur	0.052	27.5	0.000	0.032	278,541	239,301	215,697
Spal	Sdro	Spur	0.050	25.7	0.000	0.028	297,221	189,217	171,072
Sint	Spal	Spur	0.008	4.0	0.000	0.005	251,450	194,590	191,590
Spur	Sfra	Hpul	0.013	6.1	0.000	0.005	443,234	162,520	158,463
Spur	Sdro	Hpul	0.020	9.6	0.000	0.009	406,457	159,147	152,805
Spal	Sdro	Hpul	0.006	2.5	0.013	0.002	411,339	115,830	114,528
Sfra	Sdro	Hpul	0.008	3.5	0.000	0.002	608,640	119,046	117,138
Spur	Spal	Hpul	0.017	7.5	0.000	0.007	342,870	139,011	134,494
Sfra	Spal	Hpul	0.003	1.3	0.206	0.001	414,614	118,974	118,304
Spur	Sint	Hpul	0.022	10.5	0.000	0.010	406,767	172,255	164,957
Spal	Sint	Hpul	0.010	4.4	0.000	0.004	370,005	128,140	125,634
Sfra	Sint	Hpul	0.011	5.4	0.000	0.005	436,461	156,898	153,403
Sdro	Sint	Hpul	0.006	2.8	0.006	0.002	417,256	149,052	147,317

Table 3. Results of Δ analysis

Samples	Δ Analysis						
Quartet	**Trees**[†]	**Concordant**[‡]	**Discordant 1**[§]	**Discordant 2**[¶]	**Δ**	**SE**	**z**
(((Sfra,Sdro),Spal),Mfra)	2,085	974	639	472	0.15	0.03	5.04
(((Sdro,Spal,),Sint),Mfra)	2,107	1,104	550	453	0.10	0.03	3.06
(((Mnud,Mfra),Pdep),Spur)	2,416	1,187	683	546	0.11	0.03	3.94

[†]Total number of gene trees reconstructed from single copy orthologs

[‡]Number of gene trees that were concordant with the species tree relationships (((P1,P2),P3),O)

[§]Number of gene trees that had the discordant relationship (((P2,P3),P1),O)

[¶]Number of gene threes that had the discordant relationship (((P1,P3),P2),O)

Species abbreviations: *Sfra - S. fragilis*; *Sdro - S. droebachiensis*; *Spal - S. pallidus*; *Spur - S. purpuratus*; *Mnud - M. nudus*; *Mfra - M. franciscanus*; *Pdep - P. depressus*.

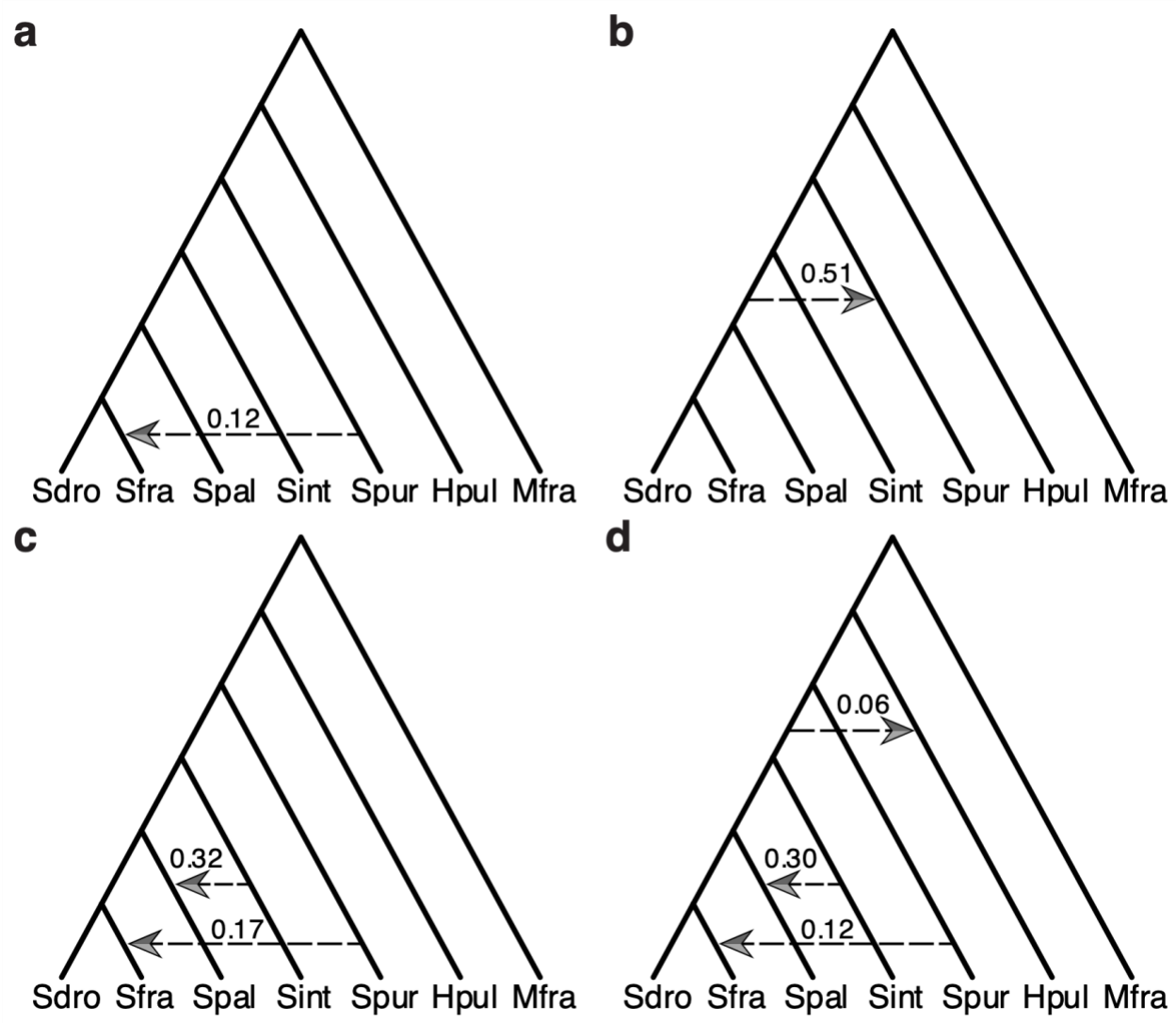

Figure 4. Phylogenetic Networks with reticulation edges and inheritance probabilities inferred by PhyloNet InferNetwork_ML. The inheritance probabilities represent the proportion of sampled genes inherited through gene flow. The network with zero reticulation edges recovered the species relationships and had a log likelihood of -11,054 (not shown). **a**. The network with one reticulation edges had a log likelihood of -10,966. **b**. A second network with one reticulation edge had a log likelihood of -10,976 and implies introgression between *S. purpuratus* and the MRCA of *S. pallidus*, *S. droebachiensis*, and *S. fragilis*. **c**. The network inferred with two reticulation edges had a log likelihood of -10,929 and implies two independent introgression events: (i) between *S. purpuratus* and *S. fragilis*, and (ii) between *S. intermedius* and *S. pallidus*. **d**. The network inferred with three reticulation edges had a log likelihood of -10,903. This network was similar to the two-reticulation edge network, with an added reticulation edge between *H. pulcherrimus* and the MRCA of *S. intermedius*, *S. pallidus*, *S. droebachiensis* and *S. fragilis*. Species abbreviations: *Sfra - S. fragilis*; *Sdro - S. droebachiensis*; *Spal - S. pallidus*; *Spur - S. purpuratus*; *Hpul - H. pulcherrimus*; *Mfra - M. franciscanus*.