Genetic differentiation and signatures of local adaptation revealed by RADseq for a highly-dispersive mud crab *Scylla olivacea* (Herbst, 1786) in the Sulu Sea

Michael John R. Mendiola*, Rachel Ravago-Gotanco

The Marine Science Institute, University of the Philippines Diliman, Quezon City 1101 Philippines

* Corresponding author: michaeljohnmendiola@gmail.com

ORCID

Mendiola, MJ 0000-0003-4329-7257

Ravago-Gotanco R 0000-0003-3570-2005

Abstract

Connectivity of marine populations is shaped by complex interactions of biological and physical processes across the seascape. The influence of environmental features on the genetic structure of populations has key implications to the dynamics and persistence of populations, and an understanding of spatial scales and patterns of connectivity is crucial for management and conservation. This study employed a seascape genetics approach combining larval dispersal modeling and population genomic analysis based on RADseq to examine environmental factors influencing patterns of genetic structure and connectivity for a highly-dispersive mud crab, *Scylla olivacea* (Herbst, 1796) in the Sulu Sea. Dispersal simulations reveal widespread but asymmetric larval dispersal influenced by persistent southward and westward surface circulation features in the Sulu Sea. Despite potential for widespread dispersal, significant genetic differentiation was detected among eight Sulu Sea populations based on 1,655 single-nucleotide polymorphism (SNP) markers (F_{ST} = 0.0057, p = 0.001) and a subset of 1,643 putatively neutral SNP markers (F_{ST} = 0.0042, p = 0.001). Oceanography influences genetic structure, as redundancy analysis (RDA) revealed significant contribution of asymmetric ocean currents to neutral genetic variation (R^2_{adj} = 0.133; p = 0.035). Genetic structure may also reflect demographic factors, with divergent populations characterized by low effective population sizes (N_e < 50). Pronounced latitudinal genetic structure was recovered for loci putatively under selection (F_{ST} = 0.2390, p = 0.001), significantly correlated with variability in mean sea surface temperatures during peak spawning months of *S. olivacea* (R^2_{adj} = 0.763; p = 0.041), suggesting putative signatures of selection and local adaptation of early life history stages to thermal clines. This study contributes to the growing body of literature documenting population genetic

structure and local adaptation for highly-dispersive marine species, and provides information useful for spatial management of the fishery resource.

Keywords: marine connectivity, population genomics, seascape genetics, RAD sequencing, mud crab, Sulu Sea

1 Introduction

Considering the spatial patterns and scales of dispersal, population connectivity has key implications for management and conservation (Moritz, 1994, Palumbi, 2003). For marine organisms, connectivity is primarily driven by complex interactions between life history characteristics and environmental conditions which influence the dynamics and persistence of populations (Cowen and Sponaugle, 2009). The absence of apparent physical barriers in the ocean combined with high dispersal potentials characteristic of most marine organisms shaped the paradigm of genetic homogeneity in the marine environment (Hauser and Carvalho, 2008). Advances in DNA sequencing technologies now provide the ability to interrogate genetic variation genome-wide, providing adequate resolution to study population genetic processes (Davey et al., 2011, Andrews et al., 2016). High-throughput genotyping of single nucleotide polymorphisms (SNPs) employing restriction-site associated DNA sequencing approaches (RAD-sequencing; Baird et al., 2008) is a widely-used approach having several variants, e.g ddRAD (Peterson et al., 2012), 2b-RAD (Wang et al., 2012), and ezRAD (Toonen et al., 2013). Genetic approaches, coupled with increased capabilities in numerical modeling to simulate and track larval dispersal in the marine environment (reviewed in Swearer et al., 2019), have resulted in considerable interest and wide use of seascape genomics approaches to examine the processes

57 that shape genetic variation in the marine environment (Riginos and Liggins, 2013, Selkoe et al.,
58 2016).

59 Seascape genomics studies have improved our understanding of the environmental
60 conditions influencing population connectivity and the spatial distribution of genetic variation in
61 the ocean. The ability to interrogate population diversity using thousands of SNPs, and to
62 identify loci which may be under the influence of selection versus neutral loci (Davey et al.,
63 2011, Gagnaire et al., 2015) allows the examination of environmental factors and their influence
64 on genetic structure. There is a growing body of literature documenting genetic structure either
65 due to neutral variation or adaptive polymorphisms, at finer spatial scales than expected from
66 species dispersal potentials. Recent studies have reported environmental factors influencing
67 neutral and adaptive genetic variation, among them: ocean currents (Benestan et al., 2016, Gilg
68 and Hilbish, 2003, Lal et al., 2017, Paterno et al., 2017, Riginos et al., 2019, Schunter et al.,
69 2011, Teske et al., 2016, Truelove et al., 2017, Van Wyngaarden et al., 2018, Xuereb et al.,
70 2018), temperature (Wang et al. 2013; Chu et al. 2014, Sandoval-Castillo et al. 2018, Hoey and
71 Pinsky 2018), and salinity (Sjöqvist et al., 2015).

72 Mud crabs (genus: *Scylla* de Haan, 1833) are commercially important species with a wide
73 distribution in mangrove areas throughout the Indo-West Pacific and other tropical and
74 subtropical regions (Alberts-Hubatsch et al., 2015). Three mud crab species occur in the
75 Philippines (*S. olivacea, S. tranquebarica*, and *S. serrata*), with *S. olivacea* (Herbst, 1796) being
76 the most abundant (Lebata et al., 2007). While adult mud crabs exhibit limited movement
77 (Hyland et al., 1984), larvae are thought to be highly dispersive due to their long pelagic larval
78 duration (PLD) of 20 to 30 days (Jantrarotai et al., 2002, Motoh et al., 1977, Thirunavukkarasu et
79 al., 2014), which can extend up to 75 days depending on environmental conditions (Baylon,

2010). Although ocean currents may play a huge role in larval dispersal and settlement success (Cowen and Sponaugle, 2009), survival and development of mud crab larvae strongly depends on water temperature and salinity (Baylon, 2011, Hamasaki, 2003, Hill, 1974, Nurdiani and Zeng, 2007). In the Sulu Sea basin, *S. olivacea* populations are distributed along regions influenced by temporally-varying environmental gradients (Oppo et al., 2003) and complex sea surface circulation such as the southward-flowing Sulu Sea throughflow from the South China Sea, the westward-flowing Bohol Sea current exiting via the Dipolog strait, and the southern Sulu Sea gyre (Han et al., 2009, Hurlburt et al., 2011). Oceanographic features in the Sulu Sea have been suggested to act as barriers to gene flow among populations for other taxa with relatively lower dispersal potentials than *S. olivacea* such as the seahorse *Hippocampus spinosissimus* (Lourie et al., 2005), damselfish *Dascyllus aruanus* (Raynal et al., 2014) and sea cucumber *Holuthuria scabra* (Ravago-Gotanco and Kim, 2019). There is limited information however, on the genetic structure of Philippine populations of *S. olivacea,* with one study reporting weak but significant genetic differentiation of Philippine populations based on microsatellite loci (Paran and Ravago-Gotanco, 2017). Moreover, there are no studies to date that explicitly examined the influence of asymmetric ocean currents and environmental heterogeneity on population connectivity and genetic structure of a highly dispersive species in the Sulu Sea.

This study examined patterns of connectivity among populations of the orange mud crab (*S. olivacea*) in the Sulu Sea basin. Using a seascape genomics approach, we aimed to (1) characterize genetic structure of *S. olivacea* across the Sulu Sea and examine spatial patterns of genetic connectivity using SNP markers generated by RADseq; (2) examine the influence of oceanographic circulation on genetic structure and connectivity of *S. olivacea* populations in the Sulu Sea; and (3) examine the SNP dataset for signatures of local adaptation which may be

103 correlated with other environmental factors. First we developed a biophysical model of larval
104 dispersal parameterized using the life-history characteristics of *S. olivacea*, to generate realistic
105 predictions of larval dispersal and connectivity in the Sulu Sea. We combined larval dispersal
106 estimates with empirical genetic observations at neutral loci to determine the influence of
107 asymmetric ocean currents on spatial patterns of connectivity. Second, we performed analyses to
108 recover loci putatively under selection to examine signatures of local adaptation. We assessed the
109 potential impact of environmental factors, specifically sea surface temperature and rainfall (as a
110 proxy for salinity) on adaptive divergence of *S. olivacea*. This study provides valuable insights
111 into the spatial scales of dispersal, patterns of genetic structure, and the influence of
112 environmental and evolutionary processes on population connectivity of *S. olivacea* in the Sulu
113 Sea basin. The results provide information useful to support the development of management and
114 conservation strategies for the fishery resource.

115

116 **2 Materials and methods**

117 *2.1 Larval dispersal simulation in the Sulu Sea basin*

118 A larval dispersal model was developed to examine connectivity of *S. olivacea* in the
119 Sulu Sea. Seven larval release sites were designated, coinciding with locations where samples
120 were collected for genetic analysis, with the exception of Coron (Table 1). Larval dispersal
121 simulations were performed using the Connectivity Modelling System (CMS; Paris et al., 2013).
122 The CMS is a model that couples Lagrangian-based descriptions of ocean circulation with
123 individual-based modeling to simulate the movement of particles each with individual behaviors
124 parameterized from known biological traits, e.g. larval duration, mortality, settlement behavior.

125 The model was configured using contemporary oceanographic data from the 3D Global Hybrid
126 Coordinate Ocean Model (HYCOM; Chassignet et al., 2007) with 1/25° (~4.4 km) horizontal
127 resolution, simulating realistic sea-surface (0 -10m) currents throughout the Sulu Sea. One year
128 of HYCOM outputs from March 2015 to February 2016 were used to run the model, covering
129 the reversing monsoon wind forcing in the region (Han et al., 2009), and the year-round
130 spawning season of *Scylla* in the Philippines (Arriola, 1940). A basin-scale habitat map was
131 included in the model using the Philippine mangrove Landsat data of Long and Giri (2011),
132 generating 159 larval settlement nodes along the boundaries of the basin. Particle release site
133 coordinates were adjusted up to 12 kilometers away from the coast to simulate the offshore
134 spawning migration reported for *S. olivacea* (Koolkalya et al., 2006, Moser et al., 2005). The
135 model was configured to release fifty thousand particles from each release site weekly over a
136 period of one year. This resulted in a total of 18.5 million larval particles released from seven
137 source nodes. Released particles were parameterized as competent to settle after 20 days of
138 passive dispersal, with a maximum duration in the water column up to 30 days based on the
139 pelagic larval duration of *Scylla* spp. (Jantrarotai et al., 2002, Motoh et al., 1977,
140 Thirunavukkarasu et al., 2014). To account for larval mortality, particles were set to be reduced
141 by half after 4 days of release based on the reported mortality of *Scylla* from larval stages zoea I
142 to III (Jantrarotai et al., 2002, Thirunavukkarasu et al., 2014). The resulting probability estimates
143 of larval dispersal (measured as percent settlement) were post-processed to generate a population
144 by population connectivity matrix. This was done by assigning settlement nodes to their
145 respective islands or nearby sampling locality (~75 km radius), resulting in a reduced population
146 source-sink dataset.
147

148 *2.2 Sample collection, species identification, and total DNA extraction*

149 Adult *Scylla olivacea* (n = 146) were collected from natural mangrove habitats from 8 sites in
150 the Sulu Sea and 2 outlier locations between 2016 to 2017 (Table 1). Species were identified
151 using morphological characters following the description of Keenan (1998). Tissue samples were
152 obtained from mud crab pereopods, preserved in salt saturated DMSO-EDTA (SSDE) solution
153 (Dawson et al., 1998) and stored at room temperature until analysis. Specimens were also
154 identified using species-diagnostic molecular markers following the protocol of Ma et al. (2012).
155 DNA was extracted using the GeneJET genomic DNA Purification Kit (Thermo Scientific),
156 following manufacturer instructions with some modifications. DNA concentration was quantified
157 using a Qubit® fluorometer. DNA quality was examined by agarose gel electrophoresis and
158 measurement of A260/280 using a NanoDrop™ spectrophotometer.
159

160 *2.3 Double-digest RAD (ddRAD) sequencing*

161 Double digest restriction site-associated (ddRAD) libraries were prepared according to
162 Peterson et al. (2012). DNA extracts were run on an agarose gel to check for DNA quality.
163 Samples with low molecular weight smears were further purified using paramagnetic beads
164 (SPRIselect; Beckman Coulter). Approximately 150 ng DNA from individual samples were
165 digested with *MluCI* and *MspI* and purified using AMPureXP (Beckman Coulter). Custom
166 barcoded adapters P1 and P2 (see Peterson et al. 2012 for sequences) were then ligated to ~50 ng
167 of DNA. The P1 adapter includes a 5bp inline unique sequence for individual barcoding. Groups
168 of 48 samples with unique barcodes were pooled (equal volumes of each sample), purified, and
169 size-selected using a BluePippin system (target insert size of 400 - 500 bp). Unique external
170 indices were added to each pool by PCR amplification. PCR products were purified, fragment

sizes verified, qPCR quantified, and further pooled in equimolar quantities. Sequencing of DNA libraries were performed using the Illumina NovaSeq™ 6000 Sequencing System with S4 flow cell type. Library construction and sequencing was performed at the Genomics Core Lab, Texas A&M University, Corpus Christi.

2.4 Read processing and SNP filtering

Sequence libraries were initially demultiplexed using the internal barcodes, while reads with low quality scores, uncalled bases, and sequences with intact adapters were removed using the module *process_radtags* in STACKs v2.2 (Rochette et al., 2019). Raw read quality scores (Phred33) and adapter contamination were examined through FastQC v0.10.1 (Andrews, 2010) and MultiQC v1.7 (Ewels et al., 2016). The STACKs pipeline module *denovo_map.pl* was used for the construction of stacks and generation of initial catalog of putative SNPs. Stack assembly parameters such as the minimum depth of coverage required to create a stack (-m) was set to 5 (default: 3), the maximum distance (in nucleotides) allowed between stacks (-M) and the number of mismatches allowed between sample tags when generating the catalogue (-n) were increased to 4 (default: 3) to increase the SNP calling confidence and to minimize missing data (Lal et al., 2016). Additional filtering steps were performed in the *populations* module with the following criteria: retain only the first SNP per locus, locus must be present in all populations (-p = 10), and excluding loci which were not present in at least 50% of the individuals for each population (-r = 0.5).

Post-processing of the SNP panel was done to exclude SNPs that were not genotyped in at least 70% of the individuals across the entire dataset. Loci with up to 30% missing data were excluded using *poppr* v2.8.3 (Kamvar et al., 2014), to minimize the effect of missing data on

194 population structure inference (Reeves et al., 2016). SNP markers with minor allele frequencies
195 (MAFs) less than 0.05 across all sites were excluded, to eliminate loci with lower power to detect
196 genetic variability (Ardlie et al., 2002, DeWoody and DeWoody, 2005). Hardy-Weinberg
197 equilibrium (HWE) tests at the population level were conducted using the package *pegas* v0.11
198 (Paradis, 2010), excluding loci which exhibited significant deviation from HWE ($p < 0.05$) after
199 correction for false discovery rate (FDR; Benjamini and Hochberg, 1995). To limit the influence
200 of non-independent loci, linkage disequilibrium (LD) was tested between all SNP pairs using the
201 package *genetics* v1.3.8.1.2 (Warnes et al., 2019), and SNPs at strong linkage disequilibrium (r^2
202 > 0.8) were removed (Lee et al., 2018, Tian et al., 2009). SNPs with high observed
203 heterozygosity ($H_o > 0.6$) were also dropped from the dataset (e.g. (Ackiss et al., 2018,
204 Hohenlohe et al., 2010, Van Wyngaarden et al., 2017) using *pegas* v0.11 (Paradis, 2010), to
205 eliminate loci exhibiting extremely high heterozygosity resulting from false SNP calls or
206 assembly errors (Lee et al., 2018). All R packages were run in R version 3.5.3 (R Core Team,
207 2019).
208
209 *2.5 Identifying non-neutral SNPs*
210 Loci were identified as being putatively neutral or under selection using two
211 differentiation-based (F_{ST}) outlier detection methods which use different underlying models:
212 BayeScan v2.1 (Foll and Gaggiotti, 2008) and Arlequin v3.5.2.2 (Excoffier and Lischer, 2010).
213 BayeScan identifies candidate outlier loci using a Bayesian likelihood approach to estimate the
214 posterior probability that each locus is under selection under the assumption that allele
215 frequencies within populations follows the multinomial-Dirichlet distribution (Foll and
216 Gaggiotti, 2008, Feng et al., 2015). Arlequin uses a hierarchical-island model which compares

observed locus-specific F_{ST} to the observed global F_{ST} value using coalescent simulations (Excoffier and Lischer, 2010). These two methods were chosen for outlier analysis due to relatively lower type I (false positive) and type II (false negative) error rates compared to other outlier tests (Lotterhos and Whitlock, 2014, Narum and Hess, 2011, Vilas et al., 2012). For BayeScan, we performed an outlier test using 100,000 iterations and a burn-in of 50,000 steps in 20 pilot runs. SNPs were then identified as outliers using a false discovery rate (FDR) *q*-value threshold of 0.05. For Arlequin, we used a total of 20,000 simulations consisting of 10 simulated groups with 100 demes per group to detect loci under selection. SNP markers with significant F_{ST} values at the 99% confidence interval (CI) limit were considered as putative outlier loci. Results of the analyses were used to generate three SNP datasets: (1) all loci; (2) putatively neutral loci only; and (3) outlier loci detected using both methods. For putative function annotation, consensus tags associated with candidate outlier loci were queried for sequence similarity against the NCBI nucleotide (nr/nt) collection using the BLAST algorithm BLASTN v2.6.1 (Morgulis et al., 2008, Zhang et al., 2000).

2.6 Population genetic structure and effective population size

The three SNP panels (all loci, putatively neutral loci, and outlier loci only) were used to examine genetic differentiation and infer connectivity among populations. We used Weir and Cockerham's F_{ST} (1984) to estimate genetic differentiation over all populations (global F_{ST}) and between populations (pairwise F_{ST}), calculated using the R packages *hierfstat* v0.04-22 (Goudet and Jombart, 2019) and *dartR* v1.1.11 (Gruber et al., 2018) respectively. Significance of F_{ST} values was tested with 10,000 bootstrap replicates, and p−values for population pairwise comparisons were adjusted for multiple tests using the false discovery rate (FDR). Pairwise F_{ST}

values were visualized using heatmaps with dendrograms generated from hierarchical clustering analysis performed using the function *heatmap.2* from the *gplots* v.3.1.0 package for R (Warnes et al., 2020).

Spatial patterns of genetic structure were examined using a discriminant analysis of principal components (DAPC; Jombart et al., 2010) implemented in the R package *adegenet* v2.1.3 (Jombart, 2008). DAPC was performed using sampling location as prior information. Cross-validation was performed using the function *xval.DAPC* with 100 replicates to determine the number of principal components to retain to avoid issues of overfitting (Jombart et al., 2010). For the outlier loci dataset, the spatially explicit clustering program GENELAND v4.0.8 (Guillot et al., 2005) was used to estimate the number of genetic clusters and infer genetic landscapes across the Sulu Sea. We used the correlated allele frequency model with the following recommended parameters: number of possible clusters (K) were initially set from 1 to 10 with 100,000 Markov Chain Monte Carlo (MCMC) iterations, thinning of 100, burn−in of 200, maximum rate of the Poisson process fixed to 146 (N = number of individuals), maximum number of nuclei in the Poisson−Voronoi tessellation process fixed to 438 (N multiplied by 3) and individual samples from each location were set to the same spatial coordinates. Post-processing of MCMC outputs generated a final estimate of K, which was used as the maximum number of populations in succeeding runs performed with 10 independent replicates. Individual runs were ranked, and the run having the highest average posterior probability was used to calculate individual membership coefficients, and maps of posterior probability of membership in each K cluster.

Hierarchical analysis of molecular variance (AMOVA; Excoffier et al., 1992) was performed to test for population structure inferred from F_{ST} analysis and DAPC on the three SNP

263 datasets. AMOVA was performed using *poppr* v2.8.3 (Kamvar et al., 2014), with significance
264 tested using 1,000 permutations. We examined the putatively neutral dataset for patterns of
265 isolation-by-distance using the *gl.ibd* function in *dartR* which performs a Mantel test (number of
266 permutations = 9,999) to assess correlation of log(geographic distance) and linearized genetic
267 distance (F_{ST}/(1-F_{ST})) matrices. Geographic distances were measured as the shortest distance
268 over water between all pairs of sites in the Sulu Sea using the *igraph* package for R (Csardi and
269 Nepusz, 2006). Effective population size was estimated for each population based on putatively
270 neutral loci only, using the linkage disequilibrium method (N_{eLD}) of NeEstimator v2.01 (Do et
271 al., 2014).
272
273 *2.7 Genetic structure and environmental factors*
274 We evaluated the influence of directional ocean currents, sea surface temperature (SST),
275 and rainfall on patterns of genetic structure of *S. olivacea* in the Sulu Sea basin based on
276 putatively neutral and adaptive variation. We used redundancy analysis (RDA), a direct gradient
277 analysis technique, to test for significant relationships between response and explanatory
278 variables (Legendre et al., 2011). Hellinger-transformed allele frequencies (Legendre and
279 Gallagher, 2001) from putatively neutral and outlier SNP datasets were used as the response
280 variable, with three environmental factors as the explanatory variables.
281 The effect of ocean currents on connectivity of populations was assessed using particle
282 dispersal estimates transformed into a set of synthetic variables known as asymmetric
283 eigenvector maps (AEMs), from which predicted spatial patterns of genetic connectivity were
284 generated (Blanchette et al., 2008, Riginos et al., 2019, Xuereb et al., 2018). Initially, AEM
285 eigenfunctions were generated by creating a site-by-edge (binary) matrix of connections between

all pairs of sites. Weight was then attributed to each edge, and when connectivity between a given pair of sites was greater than 0 in both directions, only the direction with the highest probability of dispersal was retained. AEM eigenfunctions were generated using *adespatial* v0.3-7 (Dray et al., 2020) in R v3.5.3. The contributions of individual AEMs on neutral genetic variation were calculated using redundancy analysis.

The influence of SST and rainfall variabilities on genetic structure were also evaluated through redundancy analysis using the outlier dataset, to examine signatures of local adaptation along environmental gradients. For SST, high resolution (0.25^{o} x 0.25^{o}) monthly mean data in the Sulu Sea domain (5-14^{o}N, 116-124^{o}E) from 1987-2005 were downloaded from the NOAA/OAR/ESRL PSL website at https://psl.noaa.gov/ (Reynolds et al., 2007). Similarly, fine-scale monthly mean rainfall data covering the Sulu Sea basin from 1998-2014 were obtained from the Tropical Rainfall Measuring Mission database (Huffman et al., 2007). SST and rainfall data were extracted for latitudes covering the range where samples for genetic analysis were collected.

All explanatory variables (AEMs, SST, and rainfall) were individually tested using a backward selection procedure with 999 permutations using *ordistep* in *vegan* v2.5-6 (Oksanen et al., 2019), to retain the most important explanatory variables. Selected variables were included in the model, and the adjusted coefficient of determination (R^2_{adj}) was calculated. Partial RDA was performed, and the global analysis of variance was determined using the *anova* function in *stats v3.5.3* with 999 permutations. Significance of individual RDA axes were also assessed using *anova* (permutations = 999) and selected environmental vectors were fitted into the ordination using *envfit* (permutations = 999).

3 Results

3.1 Larval dispersal estimates

Lagrangian simulations of larval dispersal using a parameterized biophysical model demonstrate ocean current-mediated connectivity among *Scylla olivacea* populations in the Sulu Sea basin (Figure 1a). Predicted particle settlement patterns indicate predominantly southward dispersal along the western boundary. A greater proportion of particles released from ROX settled in PPC (60.1%) compared to PPC particles dispersing northward and settling in ROX (27.3%) (Figure 1b, Table S1). predominantly. Particles from ROX and PPC settled in BAT (39.3%). Conversely, very low proportions of particles released from BAT were predicted to settle northward to ROX and PPC (0.004%). A small proportion of larvae released from BAT were transported to TWI (1.8%), and the majority of BAT particles were self-recruited (97.8%). Similar patterns of greater southward dispersal are observed along the eastern boundary populations (Figure 1b), although the predicted levels of particle connectivity were lower compared to western boundary populations. Particles released from MSJ were predicted to settle in ANT (12.70%) and NEG (9.21%), particles from ANT settled in NEG (10.4%) and TWI (2.70%), and particles from NEG settled in TWI (31.06%).

Across the Sulu Sea, dispersal simulations reveal a clear pattern of westward dispersal, with larval particles released from eastern boundary sites settling in western boundary sites (Figure 1b). In particular, particles from MSJ and ANT settled on three eastern sites: PPC (37.7% and 45.4%), ROX (11.7% and 8.10%) and BAT (5.00% and 7.80%). Particles released from NEG settled in PPC (17.6%) and BAT (26.2%). Self-recruitment was relatively higher for PPC (49.7%), while TWI exhibited complete self-recruitment (100%). There was no predicted settlement of particles released from western boundary sites to eastern sites. Clearly, the larval

dispersal model reveals asymmetric transport across the Sulu Sea basin, with larval dispersal predominantly southward, and from eastern to western boundary sites.

3.2 SNP filtering and identification of non-neutral loci

A total of 661,372,129 paired-end (PE) reads from 146 individual *S. olivacea* libraries were processed for SNP discovery and filtering using STACKs v2.2. Out of 777,762 loci, 491,192 (63.2%) were aligned with PE contigs, with an effective per-sample mean coverage of 20.8x. Following successive filtering steps (detailed in Table 2), a final dataset of 1,655 high quality, polymorphic SNP markers were recovered. BayeScan identified 12 putative outlier loci based on posterior probabilities at the 95% Bayes factor threshold, with F_{ST} values ranging from 0.0919 to 0.5308, and positive alpha values (range = 2.02 - 4.67), suggestive of diversifying selection (Foll, 2010, BayeScan v2.0 user manual). Arlequin identified 87 putative outlier loci at the 99% confidence interval limit with F_{ST} *p*-values values less than 0.0099. All 12 candidate outlier loci identified using BayeScan were also detected in Arlequin, thus a separate SNP panel consisting of these 12 outlier loci was generated and designated as the outlier loci dataset. Querying the contig sequences of putative outlier loci against public domain sequences using a BLAST search resulted in significant alignments of 8 RAD tags out of 12 outlier markers to known genomic regions of other crustacean species (*Portunus trituberculatus, Penaeus vannamei*) and fish (*Chanos chanos, Danio renio*). One tag had 87.76% identity match to a microsatellite region of Japanese blue swimming crab *P. trituberculatus*, while the rest were found to have 80%-100% identity to known and predicted genomic DNA regions (Table S2). There were no functional gene region matches for these 12 outlier loci.

353

354 *3.3 Genetic differentiation and population structure*

355 All *Scylla olivacea* populations exhibited significant genetic differentiation based on
356 global estimates of F_{ST} calculated using all loci (1655 SNPs; F_{ST} = 0.0070, p = 0.001) and
357 putatively neutral loci (1643 SNPs; F_{ST} = 0.0056, p = 0.001). Excluding outgroup sites, Sulu Sea
358 samples (n = 116 individuals; 8 sites) still exhibited significant genetic differentiation for all loci
359 (F_{ST} = 0.0057, p = 0.001) and putatively neutral loci (F_{ST} = 0.0042, p = 0.001). Pairwise F_{ST}
360 values using all loci revealed the most genetically divergent populations to be CRN (F_{ST} range =
361 0.0021 - 0.0219), PPC (F_{ST} range = 0.0014 - 0.0187), and one outgroup GSC (F_{ST} range = 0.0014
362 - 0.0219) (Table S3). The divergence of CRN, PPC and GSC was also evident for neutral loci,
363 although F_{ST} estimates were slightly lower (Table S4). Dendrograms of pairwise F_{ST} values
364 clearly separate CRN and PPC from the rest of the Sulu Sea populations using all loci and
365 putatively neutral loci; all p-values for pairwise comparisons were < 0.001 except for CRN-MSJ
366 and CRN-NEG for both marker sets (Figure 3a, b). Excluding CRN and PPC, further structure is
367 detected among the 6 Sulu Sea sites at all loci (F_{ST} = 0.0016, p = 0.002), with significant F_{ST}
368 between TWI-ROX at all loci (Table S3), but no further structure at neutral loci (Table S4).

369 The DAPC using putatively neutral loci shows separation of CRN, PPC, ANT and TWI
370 from the other Sulu Sea populations (Figure 4a). Retaining 55 PCs based on the cross-validation
371 results, the first three discriminant functions explained 28.0%, 26.3%, and 16.7% of the variance,
372 respectively. The first discriminant axis reveals the separation of CRN, PPC and ANT (Figure
373 4b), the second discriminant axis separates TWI (Figure 4c), and the third axis further separates
374 PPC (not shown). While pairwise F_{ST} p-values provide strong support for the separation of CRN
375 and PPC, and to a smaller extent TWI, there is no evident support for the separation of ANT at

neutral loci. An AMOVA testing a hypothesis of three genetic groups: (1) CRN, (2) PPC, and (3) the rest of the Sulu Sea populations showed significant differentiation between groups (F_{CT} = 0.017, p = 0.002; accounting for 1.7% of the total variance).

Mantel tests show no significant relationship between geographic distance and genetic distance for putatively neutral loci across all eight Sulu Sea sites (Mantel r = -0.086, p = 0.667). Examining eastern and western boundary populations separately, an emergent pattern of genetic distance increasing with geographic distance is observed for the eastern boundary populations (ANT, MSJ, NEG, TWI) although the relationship is not significant (Mantel r = 0.894, p = 0.083). Using the putatively neutral loci dataset, genetically divergent populations CRN, PPC and GSC were estimated to have small effective population size (N_e: CRN = 24.6-26.8, PPC = 10.7-11.2, GSC = 9.6-10.1) compared to other localities where N_e values range from 139.2 to very large (infinite) at 95% CI (Table S6).

Outlier loci revealed pronounced genetic differentiation across the Sulu Sea (12 SNPs; F_{ST} = 0.2390, p = 0.001). A dendrogram based on pairwise F_{ST} estimates suggests four genetic clusters in the Sulu Sea: (1) CRN-MSJ; (2) ROX-PPC-ANT; (3) BAT-TWI and (4) NEG (Figure 3c). Pairwise F_{ST} among populations within each of the 3 clusters are not significant (FDR adjusted p > 0.05), i.e. for ROX-PPC-ANT and BAT-TWI, while between-cluster comparisons are significant (FDR adjusted p < 0.05) (Table S5). The DAPC of outlier loci (11 PCs following cross-validation) suggests four genetic clusters exhibiting limited overlap in their 95% CI ellipses, and recovered the same spatial structure as pairwise F_{ST}, except that it clustered BAT and NEG, with TWI as a divergent population (Fig 4d). The first discriminant axis (58.8% of the total variance) establishes 3 groups: ROX-PPC-ANT-TWI, BAT-NEG and MSJ-CRN. The second discriminant axis (28.25% of total variance) separates TWI and establishes it as a fourth

genetic group. GENELAND recovered four genetic clusters consistent with the DAPC grouping (Figure 5). AMOVA provides further support for the concordant groupings recovered by DAPC and GENELAND, with significant differentiation among the four groups (F_{CT} = 0.209; p = 0.001) accounting for 20.9% of the total observed variance. No further structure is detected among samples within groups (F_{SC} = 0.0168 , p = 0.166). However, the discordance between pairwise F_{ST} versus groupings recovered by DAPC and GENELAND for the southern sites BAT, NEG and TWI populations may be influenced by missing data (19% over the outlier dataset). Thus, the clustering for these three populations should be approached with caution. Considering the small number of outlier loci, missing data may have a big impact on the spatial genetic structure recovered by F_{ST} and multivariate methods. To examine this further, we performed two separate analyses to handle missing data: (1) remove genotypes (individuals) with > 25 % missing data; (2) impute missing data based on population frequencies as implemented in GenoDive v3.0 (Meirmans 2020) (see Figure S1 for details). DAPC analysis of both datasets (genotypes removed and missing genotypes imputed), recovered the same four groups as the original dataset including missing data: CRN-MSJ, ANT-PPCR-ROX, BAT-NEG and TWI (Figure S1). The consistent recovery of CRN-MSJ (at 12^{o}N), ANT-PPC-ROX (at 9^{o}N - 11^{o}N), as genetically distinct groups from BAT-NEG and TWI (at 5^{o}N to 8^{o}N) by pairwise F_{ST}, DAPC and GENELAND, provides support for a pattern of latitudinal structure of Sulu Sea populations based on outlier loci.

3.4 Genetic structure and environmental factors

Directional estimates of modelled larval dispersal generated seven asymmetric eigenvector maps (AEMs) representing predicted patterns of spatial genetic connectivity in the

422 Sulu Sea. Backward selection of the AEM variables identified two significant predictors (AEM6
423 and AEM7, with $p > 0.05$) (Table 4). Together, these two AEM eigenfunctions explained 13.3%
424 of neutral genetic variation among sites (adjusted R^2; $p = 0.035$). The first RDA axis (RDA1)
425 constituted the highest proportion of genetic variation in the response data (60.9%) which is only
426 significant at the 10% level ($p = 0.065$), whereas RDA2 accounted for 39.1% of the total genetic
427 variation ($p = 0.292$). Although AEM6 and AEM7 vectors were both selected to construct the
428 model, individual testing of explanatory variables revealed that only AEM6 was significant ($p <$
429 0.001). The AEM6 eigenvector modelled Puerto Princesa (PPC) and Tawi-Tawi (TWI) as
430 separate units (Figure 2b).

431 Environmental data on mean SST and rainfall exhibited different levels of contribution to
432 genetic differentiation and potential latitudinal adaptation of *S. olivacea* in the Sulu Sea. For
433 SST, four independent vectors were identified consisting of months mostly during the wet season
434 (June, and August through October; Table 4) contributing 76.3% of the total genetic variation
435 among sites ($R^2_{adj} = 0.763$; $p = 0.041$). RDA1 explained the highest fraction of genetic variation
436 comprising 88.0% ($p = 0.048$), while RDA2 accounted for 10.2% ($p = 0.341$). Moreover, rainfall
437 data explained a lower proportion of the variation ($R^2_{adj} = 0.656$) despite having a similar number
438 of significant variables to contribute with the genetic variation. A less significant RDA model
439 was constructed using the rainfall vectors ($p = 0.089$), which suggests that mean SST (Figure S2)
440 is a stronger predictor of the observed latitudinal genetic variation than rainfall.

4 Discussion

In this study, we employed a seascape genomics approach to examine environmental factors influencing genetic structure of *Scylla olivacea* populations in the Sulu Sea. Putatively neutral markers revealed weak yet significant genetic differentiation significantly correlated with genetic structure predicted from particle dispersal simulations, indicating the influence of ocean currents on gene flow. Geographic distance was not a significant predictor of genetic structure. Meanwhile, outlier loci revealed a pattern of latitudinal genetic structure suggesting local adaptation to latitudinal environmental gradients. Mean SST is a stronger predictor of adaptive divergence along environmental gradients than rainfall. The results presented here provided evidence of basin-scale genetic differentiation of *S. olivacea* populations in the Sulu Sea, which may be used to devise spatially explicit management and conservation interventions.

4.1 Genetic structure and connectivity in the Sulu Sea

Broad dispersal of *S. olivacea* may be inferred from life history features such as offshore spawning migration and a relatively long pelagic larval duration which may reach up to 75 days. Biophysical modeling predicts that larvae have the potential to disperse widely across the Sulu Sea, an area spanning 800 km north to south and 600 km maximum east to west. Surface circulation features modify the directionality of dispersal, which simulations show to be greater southward and eastward across the domain. However, the recovery of weak yet significant genetic differentiation among *S. olivacea* populations across the Sulu Sea demonstrates that dispersive life history features may not necessarily lead to widespread connectivity and genetic homogeneity. This study demonstrates the greater resolution afforded by SNP loci generated

463 from RAD-sequencing approaches to detect genetic differences. Using a panel of 1,655 SNPs
464 and a reduced set of 1,643 putatively neutral SNPs excluding outlier loci, genetic structure was
465 detected over a relatively smaller geographic area compared to a previous study based on
466 mitochondrial DNA sequences reporting panmixia of *S. olivacea* from geographically disjunct
467 sites along the western and eastern coasts of peninsular Malaysia (Strait of Malacca and South
468 China Sea, respectively) (Rosly et al., 2013). While weak genetic differentiation was reported
469 for *S. olivacea* based on microsatellite loci, geographical coverage was broader across the
470 Philippine archipelago and was not limited to the Sulu Sea (Paran and Ravago-Gotanco, 2017)

471 This study adds to the growing body of literature reporting significant genetic
472 differentiation for populations of marine organisms despite the potential for broad dispersal
473 (Hauser and Carvalho, 2008). Weak yet significant genetic differentiation, with comparable
474 estimates of low F_{ST} values over sampling scales of hundreds of kilometers have been reported
475 for broadly-dispersing species such as highly mobile fish (Atlantic cod, F_{ST} = 0.004; Knutsen et
476 al., 2003), or invertebrates with extensive larval duration periods such as red rock lobsters (F_{ST} =
477 0.004; Iacchei et al., 2013) and spiny lobsters (F_{ST} = 0.0016; Truelove et al., 2017). In the
478 Adriatic Sea (800 km long, 200 km wide), a semi-enclosed ocean basin comparable in area to the
479 Sulu Sea (800 km long, 600 km wide), significant genetic differentiation was also reported for a
480 range of organisms with similar bipartite life histories and broad dispersal potentials, such as the
481 anchovy *Engraulis encrasicolus* (Bembo et al., 1996), shore crab *Carcinus aestuarii* (Schiavina
482 et al., 2014), and peacock wrasse *Symphodus tinca* (Carreras et al., 2017), suggesting apparent
483 barriers to dispersal even across distances of several hundred kilometers.

484 Genetic differentiation across the Sulu Sea does not appear to be a function of geographic
485 distance alone, but is influenced by oceanographic features based on the significant contributions

of AEM predictors on the neutral genetic variation among Sulu Sea populations ($R^2_{adj} = 0.133$, $p = 0.035$). In particular, the predicted genetic pattern from AEM6, which identifies PPC and TWI as separate populations, is significantly correlated to the empirical allelic frequencies based on putatively neutral loci. This pattern is broadly consistent with the genetic analyses, i.e. F_{ST}-based approaches reveal PPC as divergent from other Sulu Sea populations (Figure 2a, 3a, 3b). The separation of TWI, while not supported by pairwise F_{ST} p-values after table-wide FDR adjustment, is emergent in the DAPC plot (Figure 4d). The divergence of PPC and TWI may be due to self-recruitment. TWI is modeled to have 100% self-recruitment likely due to the southern Sulu gyre which might promote entraintment, while self-recruitment for PPC (49.7%) is relatively higher compared to the other Sulu Sea sites (22-24%, except for BAT at 97.8%). Self-recruitment estimates are not available for CRN, but we hypothesize high rates of self-recruitment considering the deep embayment of CRN, which may preclude larval dispersal offshore. Excluding CRN, PPC and TWI, the rest of the Sulu Sea populations do not exhibit further genetic differentiation ($F_{ST} = 0.0004$, $p = 0.094$), suggesting no apparent barriers to *S. olivacea* larval dispersal across the Sulu Sea. Genetic studies for other species with similar limited adult movement, but shorter pelagic larval durations than *S. olivacea* indicate limited gene flow between eastern and western boundary populations. Genetic structure for the seahorse *Hippocampus spinosissimus* (Lourie et al., 2005), damselfish *Dascyllus aruanus* (Raynal et al., 2014), and sea cucumber *Holothuria scabra* (Ravago-Gotanco and Kim, 2019), attributed limited dispersal across the Sulu Sea to a combination of oceanographic circulation features such as the Sulu Sea throughflow, the geographic distance across the Sulu Sea, and the absence of stepping stone reef habitats across the basin, as barriers to dispersal between eastern and western boundary populations. In contrast, larval dispersal simulations for three model organisms with

509 varied dispersal potentials: a broadcast-spawning coral (*Acropora millepora*), sea urchin
510 (*Tripneustes gratilla*), and a reef fish (*Epinephelus* sp.) recovered three clusters in the Sulu Sea
511 domain (North, Central, and Southern), but did not appear to indicate restricted dispersal
512 between eastern and western boundary populations (Pata and Yniguez, 2019).

513 Population allele frequencies, while largely influenced by gene flow, may also reflect
514 demographic changes (Whitlock and McCauley, 1999). The two most divergent populations,
515 CRN and PPC are characterized by low estimates of effective population size (CRN N_e = 24.6-
516 26.8, PPC N_e = 10.7-11.2; Table S6), indicating the possible influence of genetic drift on allele
517 frequency of small populations which may lead to neutral divergence (Hare et al., 2011, Waples,
518 2010). Genetic divergence associated with low effective population sizes have been previously
519 reported for other marine taxa, e.g. red cusk-eel *Genypterus chilensis* due to high fishing
520 pressure (Córdova-Alarcón et al., 2019), and population bottlenecks for *Gadus* morhua (Andreev
521 et al., 2015) and *Epinephelus marginatus* (Buchholz-Sørensen and Vella, 2016). The possible
522 causes of low effective population sizes in CRN and PPC are not known. While high exploitation
523 rates or diminished suitable habitat area may be underlying reasons for low population sizes,
524 additional information from fishery data and habitat surveys (e.g. mangrove cover) are needed to
525 make a more conclusive determination.

526

527 *4.2 Latitudinal adaptation of S. olivacea in the Sulu Sea*

528 Environmental conditions can be agents of selection shaping the genotypic composition
529 of local populations, with environmental heterogeneity resulting in increased adaptive potential
530 i.e. an increased average fitness of organisms in their local environment than elsewhere (Hoban

et al., 2016, Sanford and Kelly, 2011). *Scylla olivacea* populations exhibit pronounced genetic differentiation at outlier loci, demonstrating fine-scale latitudinal genetic structure across the Sulu Sea. Four genetic clusters were identified using multiple genetic approaches (F_{ST}, DAPC, GENELAND), with AMOVA indicating significant differentiation among groups accounting for 20.9% of the total variance (F_{CT} = 0.209; p = 0.001): (1) CRN-MSJ, (2) ROX-PPC-ANT, (3) BAT-NEG, and (4) TWI. Two environmental variables, monthly mean SST and rainfall can explain the latitudinal genetic structure of *S. olivacea,* with SST (R^2_{adj} = 0.763, p = 0.041) as a stronger predictor of genetic variation than rainfall (R^2_{adj} = 0.656, p = 0.089). For both SST and rainfall data, most of the selected variables (months) that were included in the model covers the wet season (June - November) where the latitudinal thermal cline was steepest. Moreover, the significant association between latitudinal genetic structure and environmental variation during the wet season coincides with the reported peak spawning season of *Scylla* species in the Philippines (Arriola, 1940, Lebata et al., 2007), suggesting a biological response to environmental clines. Fine scale genetic structure recovered by adaptive polymorphisms likely reflects the influence of temporally-variable latitudinal variations in environmental variables on *S. olivacea* during their early life stages, despite the potential for widespread connectivity.

Temperature and salinity play a significant role in the seasonal occurrence and abundance of mudcrabs. Changes brought about by increased freshwater flow during the rainy season are thought to influence mud crab abundance through enhanced productivity of coastal areas and estuaries (Alberts-Hubatsch et al., 2015, Butcher et al., 2002, Meynecke et al., 2008). Moreover, latitudinal variations in temperature and salinity can be expected to drive variability in reproductive characteristics. Size-at-maturity in mud crabs has been shown to vary with latitude, with smaller size at maturity in tropical regions hypothesized to be due to faster maturation in

554 warmer waters (Alberts-Hubatsch et al., 2015, Quinn and Kojis, 1987, Robertson and Kruger,
555 1994). Latitudinal variation in reproductive characteristics was also reported for a closely related
556 taxa, the burrowing mud crab, *Helise crasa* (Grapsidae) where the maximum crab size, size of
557 maturity of females, and numbers of eggs carried per female increased significantly with
558 increased latitude (Jones and Simons, 1981). Latitudinal variation in peak spawning of *S.*
559 *olivacea*, which occurs from July to November at latitudes between 9°N to 11°N (Koolkalya et
560 al., 2006, Viswanathan et al., 2019), and from March to September at higher latitudes (Ali et al.,
561 2020, Ogawa et al., 2012) suggests a potential for adaptive variation across broad latitudinal
562 scales. Moreover, water temperature and salinity are known key factors influencing larval
563 development, growth and survival of mud crabs (Baylon, 2011, Hill, 1974, Nurdiani and Zeng,
564 2007). Thus, latitudinal variability in temperature and salinity are expected to significantly
565 impact reproduction, larval survival, and development, and ultimately, the dynamics, genetic
566 structure, and persistence of populations. Patterns of genetic differentiation associated with
567 latitudinal gradients of temperature and salinity have been reported for several marine organisms
568 across varying spatial scales. For instance, two major latitudinal clades were recovered in the
569 North Atlantic snail (*Nucella lapillus*) along midcoastal Maine (between 43°N and 44°N; with
570 water and air difference reaching up to 5-10 °C), in which some of the genes involved in the
571 genetic structure were associated with heat stress tolerance (Chu et al., 2014). A pattern of
572 population structure was also found for a high gene-flow marine fish (*Larimichthys polyactis*) in
573 the Northwest Pacific marginal seas by using an outlier locus (e.g. heat shock protein), which is
574 linked to local adaptation relating to seasonal variability in temperature between two regions
575 separated by 1-2°C thermal difference between sites (Wang et al., 2013). For a marine diatom
576 *Skeletonema marinoi*, genetic break was found between the low-salinity Baltic Sea and high-

577 salinity North Baltic Sea populations, despite the potential for migration between
578 metapopulations based on oceanographic connectivity (Sjöqvist et al., 2015). This study is the
579 first report of latitudinal adaptive divergence for the mudcrab. Assessment of local adaptation
580 and underlying factors influencing genetic differentiation is essential to understand the dynamics
581 of populations associated with environmental factors.

582 *4.3 Implications to Management*

583 Understanding of spatial patterns of connectivity and local adaptation of populations
584 represents key considerations for the design of effective, resilience-based management
585 interventions for fishery resources. For *S. olivacea*, basin-scale genetic differentiation was
586 detected at both the putatively neutral and outlier loci, reflecting the influence of evolutionary
587 (e.g. genetic drift) and environmental processes (e.g. ocean currents, temperature, salinity) on
588 genotypic composition of populations in the Sulu Sea. The assessment of genetic diversity and
589 connectivity of marine populations inferred from both neutral and outlier loci provides more
590 holistic genetic information for fisheries management of populations (Sandoval-Castillo et al.,
591 2018, Carreras et al., 2017, Gagnaire et al., 2015, Nayfa and Zenger, 2016, Van Wyngaarden et
592 al., 2017). In this study, we provide genetic resources (neutral and adaptive) to support the
593 development of policy recommendations for management and conservation of *S. olivacea*.

594 From the perspective of neutral loci, *S. olivacea* populations in the Sulu Sea can be
595 considered as a metapopulation with two divergent populations (Coron and Puerto Princesa)
596 likely influenced by genetic drift as a consequence of small effective population sizes.
597 Populations with low N_e are particularly vulnerable to continued loss of genetic diversity, and
598 may need to be prioritized in restoration and conservation plans such as stock enhancement
599 programs aimed at increasing yields beyond levels supported by natural recruitment (Bell et al.,

2005). Stock enhancement programs initiated for the depleted mud crab (*S. paramamosain*) fishery in Japan report promising results towards increasing catch and population sizes after more than a decade of restoration efforts (Obata et al., 2006). Thus, this study recommends the development of management and conservation plans for vulnerable populations of *S. olivacea*, in Coron and Puerto Princesa which are potentially facing higher rates of local extinction due to small effective population size.

Management strategies employing translocation of individuals should also be conducted with caution, with the view to maintain localized adaptive divergence among populations. In this context, evaluation of genetic variation using outlier markers is important, to detect signatures of local adaptation. In *S. olivacea*, we detected a pattern of genetic structure associated with environmental gradients such as temperature and salinity. These findings are important to consider in aquaculture practices and resource management interventions that rely on translocation of individuals across geographic locations. Successful adaptation is predicted to produce genotype-phenotype-environment associations, and translocation of locally-adapted individuals may result in genetic-environment mismatch, and have significant impacts on fitness traits particularly beyond the limits of phenotypic plasticity (e.g. (Kvingedal et al., 2010, Nayfa and Zenger, 2016). Thus, genetic information from this study can be used to identify sources of broodstock which are potentially adapted to similar local environments. For example, individuals to be used in restocking the Coron population may be sourced from Mindoro, a nearby locality which is genetically similar to Coron. Likewise, the Puerto Princesa population may be restocked using individuals from an adjacent population in Roxas, Palawan or a population across the basin, in Antique. This process may reduce outbreeding of genetically mismatched individuals that are locally adapted to different environmental conditions, which also limits the adverse

623 effects on fitness and survival of these populations (Edmands, 2007, Edmands and Timmerman,
624 2003, Gharrett et al., 1999).

625 Overall, the results of this study can contribute to improve existing management and
626 conservation plans for *S. olivacea* in the Philippines. *Scylla olivacea* was among the species
627 included in a recent fisheries ordinance establishing guidelines limiting catch, trade, and
628 transport of crablets, juvenile, and gravid individuals across the Philippines (Fisheries
629 Administrative Order (FAO) 264 s 2020; (BFAR, 2020). While not a priority species for
630 aquaculture because of its aggressive behavior and smaller size than *S. serrata*, *S. olivacea* is the
631 more abundant species and represents an important fishery resource that should be maintained
632 and protected as a source of livelihood for small-scale fishers across the Philippine archipelago.
633 It is essential to augment genetics-based approaches with other assessments of the fishery
634 resource, to provide further insight into spatial distributions, genetic boundaries, and local
635 adaptation in a rapidly changing marine environment, which are critical towards the design of
636 management and conservation strategies.

637 **Literature cited**

638 Ackiss, A. S., Bird, C. E., Akita, Y., Santos, M. D., Tachihara, K. & Carpenter, K. E. 2018.
639 Genetic patterns in peripheral marine populations of the fusilier fish Caesio cuning within
640 the Kuroshio Current. *Ecol Evol,* 8**,** 11875-11886.

641 Alberts-Hubatsch, H., Lee, S. Y., Meynecke, J.-O., Diele, K., Nordhaus, I. & Wolff, M. 2015.
642 Life-history, movement, and habitat use of Scylla serrata (Decapoda, Portunidae): current
643 knowledge and future challenges. *Hydrobiologia,* 763**,** 5-21.

644 Ali, M. Y., Hossain, M. B., Sana, S., Rouf, M. A., Yasmin, S. & Sarower, M. G. 2020.
645 Identifying peak breeding season and estimating size at first maturity of mud crab (Scylla
646 olivacea) from a coastal region of Bangladesh. *Heliyon,* 6**,** e04318.

647 Andreev, V., Fokin, M., Mugue, N. & Strelkov, P. 2015. Long-term persistence and evolutionary
648 divergence of a marine fish population with a very small effective population size (Kildin
649 cod Gadus morhua kildinensis). *Marine Biology,* 162**,** 979-992.

650 Andrews, K. R., Good, J. M., Miller, M. R., Luikart, G. & Hohenlohe, P. A. 2016. Harnessing
651 the power of RADseq for ecological and evolutionary genomics. *Nature Reviews*
652 *Genetics,* 17**,** 81.

653 Andrews, S. 2010. FastQC: a quality control tool for high throughput sequence data. Babraham
654 Bioinformatics, Babraham Institute, Cambridge, United Kingdom.

655 Ardlie, K. G., Lunetta, K. L. & Seielstad, M. 2002. Testing for population subdivision and
656 association in four case-control studies. *Am J Hum Genet,* 71**,** 304-11.

657 Arriola, F. J. 1940. *A Preliminary Study of the Life History of Scylla Serrata (Forskal)*,
658 Philippine Journal of Science.

659 Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., Selker, E.
660 U., Cresko, W. A. & Johnson, E. A. 2008. Rapid SNP discovery and genetic mapping
661 using sequenced RAD markers. *PloS one,* 3**,** e3376.
662 Baylon, J. C. 2010. Effects of Salinity and Temperature on Survival and Development of Larvae
663 and Juveniles of the Mud Crab, Scylla serrata (Crustacea: Decapoda: Portunidae).
664 *Journal of the World Aquaculture Society,* 41**,** 858-873.
665 Baylon, J. C. 2011. Survival and development of larvae and juveniles of the mud crab Scylla
666 olivacea Forskal (Crustacea: Decapoda: Portunidae) at various temperatures and
667 salinities. *The Philippine Agricultural Scientist,* 94.
668 Bell, J. D., Munro, J. L., Nash, W. J., Rothlisberg, P. C., Loneragan, N. R., Ward, R. D. &
669 Andrew, N. L. B. T. a. I. M. B. 2005. Restocking and stock enhancement of marine
670 invertebrate fisheries. Academic Press.
671 Bembo, D. G., Carvalho, G. R., Cingolani, N., Arneri, E., Giannetti, G. & Pitcher, T. J. 1996.
672 Allozymic and morphometric evidence for two stocks of the European anchovy Engraulis
673 encrasicolus in Adriatic waters. *Marine Biology,* 126**,** 529-538.
674 Benestan, L., Quinn, B. K., Maaroufi, H., Laporte, M., Clark, F. K., Greenwood, S. J., Rochette,
675 R. & Bernatchez, L. 2016. Seascape genomics provides evidence for thermal adaptation
676 and current-mediated population structure in American lobster (Homarus americanus).
677 *Molecular ecology,* 25**,** 5073-5092.
678 Benjamini, Y. & Hochberg, Y. 1995. Controlling the False Discovery Rate: A Practical and
679 Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society. Series B*
680 *(Methodological),* 57**,** 289-300.

Bfar 2020. Regulation on the catching, possession, transporting, selling, trading and exporting of Mangrove Crablets, Juvenile Mangrove Crabs, and Gravid Mangrove Crabs (Scylla Spp.). *In:* RESOURCES, B. O. F. A. A. (ed.) *FAO 264 s 2020.*

Blanchette, C. A., Melissa Miner, C., Raimondi, P. T., Lohse, D., Heady, K. E. K. & Broitman, B. R. 2008. Biogeographical patterns of rocky intertidal communities along the Pacific coast of North America. *Journal of Biogeography,* 35**,** 1593-1607.

Buchholz-Sørensen, M. & Vella, A. 2016. Population Structure, Genetic Diversity, Effective Population Size, Demographic History and Regional Connectivity Patterns of the Endangered Dusky Grouper, Epinephelus marginatus (Teleostei: Serranidae), within Malta's Fisheries Management Zone. *PLOS ONE,* 11**,** e0159864-e0159864.

Butcher, P., Boulton, A. J. & Smith, S. 2002. Mud crab (Scylla serrata: Portunidae) populations as indicators of the effectiveness of estuarine marine protected areas. *World Congress on Aquatic Protected Areas Proceedings***,** 421-427.

Carreras, C., Ordonez, V., Zane, L., Kruschel, C., Nasto, I., Macpherson, E. & Pascual, M. 2017. Population genomics of an endemic Mediterranean fish: differentiation by fine scale dispersal and adaptation. *Sci Rep,* 7**,** 43417.

Chassignet, E. P., Hurlburt, H. E., Smedstad, O. M., Halliwell, G. R., Hogan, P. J., Wallcraft, A. J., Baraille, R. & Bleck, R. 2007. The HYCOM (HYbrid Coordinate Ocean Model) data assimilative system. *Journal of Marine Systems,* 65**,** 60-83.

Chu, N. D., Kaluziak, S. T., Trussell, G. C. & Vollmer, S. V. 2014. Phylogenomic analyses reveal latitudinal population structure and polymorphisms in heat stress genes in the North Atlantic snail Nucella lapillus. *Mol Ecol,* 23**,** 1863-73.

703 Córdova-Alarcón, V. R., Araneda, C., Jilberto, F., Magnolfi, P., Toledo, M. I. & Lam, N. 2019.
704 Genetic Diversity and Population Structure of Genypterus chilensis, a Commercial
705 Benthic Marine Species of the South Pacific. *Frontiers in Marine Science,* 6**,** 748-748.
706 Cowen, R. K. & Sponaugle, S. 2009. Larval dispersal and marine population connectivity. *Ann*
707 *Rev Mar Sci,* 1**,** 443-66.
708 Csardi, G. & Nepusz, T. 2006. The igraph software package for complex network research.
709 *InterJournal, complex systems,* 1695**,** 1-9.
710 Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M. & Blaxter, M. L. 2011.
711 Genome-wide genetic marker discovery and genotyping using next-generation
712 sequencing. *Nature Reviews Genetics,* 12**,** 499-510.
713 Dawson, M. N., Raskoff, K. A. & Jacobs, D. K. 1998. Field preservation of marine invertebrate
714 tissue for DNA analyses. *Molecular marine biology and biotechnology,* 7**,** 145-152.
715 Dewoody, Y. D. & Dewoody, J. A. 2005. On the Estimation of Genome-wide Heterozygosity
716 Using Molecular Markers. *Journal of Heredity,* 96**,** 85-88.
717 Do, C., Waples, R. S., Peel, D., Macbeth, G. M., Tillett, B. J. & Ovenden, J. R. 2014.
718 NeEstimator v2: re-implementation of software for the estimation of contemporary
719 effective population size (Ne) from genetic data. *Mol Ecol Resour,* 14**,** 209-14.
720 Dray, S., Bauman, D., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., Jombart, T., Larocque,
721 G., Legendre, P. & Madi, N. 2020. adespatial: Multivariate multiscale spatial analysis. R
722 package version 0.3-7. Available in: https://CRAN. R-project. org/package= adespatial.
723 Edmands, S. 2007. Between a rock and a hard place: evaluating the relative risks of inbreeding
724 and outbreeding for conservation and management. *Mol Ecol,* 16**,** 463-75.

Edmands, S. & Timmerman, C. C. 2003. Modeling Factors Affecting the Severity of Outbreeding Depression. *Conservation Biology,* 17**,** 883-892.

Ewels, P., Magnusson, M., Lundin, S. & Kaller, M. 2016. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics,* 32**,** 3047-8.

Excoffier, L. & Lischer, H. E. 2010. Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Resour,* 10**,** 564-7.

Excoffier, L., Smouse, P. E. & Quattro, J. M. 1992. Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics,* 131**,** 479-491.

Feng, X. J., Jiang, G. F. & Fan, Z. 2015. Identification of outliers in a genomic scan for selection along environmental gradients in the bamboo locust, Ceracris kiangsu. *Sci Rep,* 5**,** 13758.

Foll, M. & Gaggiotti, O. 2008. A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant and Codominant Markers: A Bayesian Perspective. *Genetics,* 180**,** 977-993.

Gagnaire, P. A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., Arnaud-Haond, S. & Bierne, N. 2015. Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications,* 8**,** 769-786.

Gharrett, A. J., Smoker, W. W., Reisenbichler, R. R. & Taylor, S. G. 1999. Outbreeding depression in hybrids between odd- and even-broodyear pink salmon. *Aquaculture,* 173**,** 117-129.

Gilg, M. R. & Hilbish, T. J. 2003. The Geography of Marine Larval Dispersal: Coupling Genetics with Fine-Scale Physical Oceanography. *Ecology,* 84**,** 2989-2998.

747 Goudet, J. & Jombart, T. 2019. hierfstat: Estimation and Tests of Hierarchical F-Statistics. R
748 package version 0.04–22. 2015.
749 Gruber, B., Unmack, P. J., Berry, O. F. & Georges, A. 2018. dartr: An r package to facilitate
750 analysis of SNP data generated from reduced representation genome sequencing. *Mol*
751 *Ecol Resour,* 18**,** 691-699.
752 Guillot, G., Mortier, F. & Estoup, A. 2005. Geneland: a computer package for landscape
753 genetics. *Molecular Ecology Notes,* 5**,** 712-715.
754 Hamasaki, K. 2003. Effects of temperature on the egg incubation period, survival and
755 developmental period of larvae of the mud crab Scylla serrata (Forskål) (Brachyura:
756 Portunidae) reared in the laboratory. *Aquaculture,* 219**,** 561-572.
757 Han, W., Moore, A. M., Levin, J., Zhang, B., Arango, H. G., Curchitser, E., Di Lorenzo, E.,
758 Gordon, A. L. & Lin, J. 2009. Seasonal surface ocean circulation and dynamics in the
759 Philippine Archipelago region during 2004–2008. *Dynamics of Atmospheres and Oceans,*
760 47**,** 114-137.
761 Hare, M. P., Nunney, L., Schwartz, M. K., Ruzzante, D. E., Burford, M., Waples, R. S., Ruegg,
762 K. & Palstra, F. 2011. Understanding and estimating effective population size for
763 practical application in marine species management. *Conserv Biol,* 25**,** 438-49.
764 Hauser, L. & Carvalho, G. R. 2008. Paradigm shifts in marine fisheries genetics: ugly hypotheses
765 slain by beautiful facts. *Fish and Fisheries,* 9**,** 333-362.
766 Hill, B. J. 1974. Salinity and temperature tolerance of zoeae of the portunid crab Scylla serrata.
767 *Marine Biology,* 25**,** 21-24.

768 Hoban, S., Kelley, J. L., Lotterhos, K. E., Antolin, M. F., Bradburd, G., Lowry, D. B., Poss, M.
769 L., Reed, L. K., Storfer, A. & Whitlock, M. C. 2016. Finding the Genomic Basis of Local
770 Adaptation: Pitfalls, Practical Solutions, and Future Directions. *Am Nat,* 188**,** 379-97.
771 Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A. & Cresko, W. A. 2010.
772 Population genomics of parallel adaptation in threespine stickleback using sequenced
773 RAD tags. *PLoS Genet,* 6**,** e1000862.
774 Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y.,
775 Bowman, K. P. & Stocker, E. F. 2007. The TRMM multisatellite precipitation analysis
776 (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales.
777 *Journal of hydrometeorology,* 8**,** 38-55.
778 Hurlburt, H. E., Metzger, J. E., Sprintall, J., Riedlinger, S. N., Arnone, R. A., Shinoda, T. & Xu,
779 X. 2011. Circulation in the Philippine Archipelago simulated by 1/12° and 1/25° Global
780 HYCOM and EAS NCOM. *Oceanography,* 24**,** 28-47.
781 Hyland, S., Hill, B. & Lee, C. 1984. Movement within and between different habitats by the
782 portunid crab Scylla serrata. *Marine Biology,* 80**,** 57-61.
783 Iacchei, M., Ben-Horin, T., Selkoe, K. A., Bird, C. E., García-Rodríguez, F. J. & Toonen, R. J.
784 2013. Combined analyses of kinship and FST suggest potential drivers of chaotic genetic
785 patchiness in high gene-flow populations. *Molecular Ecology,* 22**,** 3476-3494.
786 Jantrarotai, P., Taweechuer, K. & Pripanapong, S. 2002. Salinity Levels on Survival Rate and
787 Development of Mud Crab (Scylla olivacea) from Zoea to Megalopa and from
788 Megalopa to Crab Stage. *Kasetsart Journal (Natural Science),* 36**,** 278-284.
789 Jombart, T. 2008. adegenet: a R package for the multivariate analysis of genetic markers.
790 *Bioinformatics,* 24**,** 1403-5.

791 Jombart, T., Devillard, S. & Balloux, F. 2010. Discriminant analysis of principal components: a
792 new method for the analysis of genetically structured populations. *BMC Genet,* 11**,** 94.
793 Jones, M. B. & Simons, M. J. 1981. Habitat preferences of two estuarine burrowing crabs Helice
794 crassa Dana (Grapsidae) and Macrophthalmus hirtipes (Jacquinot) (Ocypodidae). *Journal*
795 *of Experimental Marine Biology and Ecology,* 56**,** 49-62.
796 Kamvar, Z. N., Tabima, J. F. & Grunwald, N. J. 2014. Poppr: an R package for genetic analysis
797 of populations with clonal, partially clonal, and/or sexual reproduction. *PeerJ,* 2**,** e281.
798 Knutsen, H., Jorde, P. E., André, C. & Stenseth, N. C. 2003. Fine-scaled geographical population
799 structuring in a highly mobile marine species: The Atlantic cod. *Molecular Ecology,* 12**,**
800 385-394.
801 Koolkalya, S., Thapanand, T., Tunkijjanujij, S., Havanont, V. & Jutagate, T. 2006. Aspects in
802 spawning biology and migration of the mud crab Scylla olivacea in the Andaman Sea,
803 Thailand. *Fisheries Management and Ecology,* 13**,** 391-397.
804 Kvingedal, R., Evans, B. S., Lind, C. E., Taylor, J. J. U., Dupont-Nivet, M. & Jerry, D. R. 2010.
805 Population and family growth response to different rearing location, heritability estimates
806 and genotype×environment interaction in the silver-lip pearl oyster (Pinctada maxima).
807 *Aquaculture,* 304**,** 1-6.
808 Lal, M. M., Southgate, P. C., Jerry, D. R., Bosserelle, C. & Zenger, K. R. 2017. Swept away:
809 Ocean currents and seascape features influence genetic structure across the 18,000 Km
810 Indo-Pacific distribution of a marine invertebrate, the black-lip pearl oyster Pinctada
811 margaritifera. *BMC Genomics,* 18.

812 Lal, M. M., Southgate, P. C., Jerry, D. R. & Zenger, K. R. 2016. Fishing for divergence in a sea
813 of connectivity: The utility of ddRADseq genotyping in a marine invertebrate, the black-
814 lip pearl oyster Pinctada margaritifera. *Mar Genomics,* 25**,** 57-68.
815 Lebata, M., Hazel, J., Vay, L. L., Primavera, J. H., Walton, M. E. & Biñas, J. B. 2007. Baseline
816 assessment of fisheries for three species of mud crabs (Scylla spp.) in the mangroves of
817 Ibajay, Aklan, Philippines. *Bulletin of Marine Science,* 80**,** 891-904.
818 Lee, S. R., Jo, Y. S., Park, C. H., Friedman, J. M. & Olson, M. S. 2018. Population genomic
819 analysis suggests strong influence of river network on spatial distribution of genetic
820 variation in invasive saltcedar across the southwestern United States. *Mol Ecol,* 27**,** 636-
821 646.
822 Legendre, P. & Gallagher, E. D. 2001. Ecologically meaningful transformations for ordination of
823 species data. *Oecologia,* 129**,** 271-280.
824 Legendre, P., Oksanen, J. & Ter Braak, C. J. F. 2011. Testing the significance of canonical axes
825 in redundancy analysis. *Methods in Ecology and Evolution,* 2**,** 269-277.
826 Long, J. B. & Giri, C. 2011. Mapping the Philippines' mangrove forests using Landsat imagery.
827 *Sensors (Basel),* 11**,** 2972-81.
828 Lotterhos, K. E. & Whitlock, M. C. 2014. Evaluation of demographic history and neutral
829 parameterization on the performance of FST outlier tests. *Mol Ecol,* 23**,** 2178-92.
830 Lourie, S. A., Green, D. M. & Vincent, A. C. 2005. Dispersal, habitat differences, and
831 comparative phylogeography of Southeast Asian seahorses (Syngnathidae:
832 Hippocampus). *Mol Ecol,* 14**,** 1073-94.

833 Ma, H., Ma, C. & Ma, L. 2012. Molecular identification of genus Scylla (Decapoda: Portunidae)
834 based on DNA barcoding and polymerase chain reaction. *Biochemical Systematics and*
835 *Ecology,* 41**,** 41-47.
836 Meynecke, J.-O., Lee, S. Y., Grubert, M., Brown, I., Montgomery, S., Johnston, D. & Gillson, J.
837 2008. Evaluating the environmental drivers of mud crab (Scylla serrata) catches in
838 Australia. *Fisheries Research and Development Corporation and Griffith University***,** 1-
839 94.
840 Morgulis, A., Coulouris, G., Raytselis, Y., Madden, T. L., Agarwala, R. & Schaffer, A. A. 2008.
841 Database indexing for production MegaBLAST searches. *Bioinformatics,* 24**,** 1757-64.
842 Moritz, C. 1994. Defining 'Evolutionarily Significant Units' for conservation. *Trends in Ecology*
843 *& Evolution,* 9**,** 373-375.
844 Moser, S., Macintosh, D., Laoprasert, S. & Tongdee, N. 2005. Population ecology of the mud
845 crab Scylla olivacea: a study in the Ranong mangrove ecosystem, Thailand, with
846 emphasis on juvenile recruitment and mortality. *Fisheries Research,* 71**,** 27-41.
847 Motoh, H., De La Peña, D. & Tampos, E. 1977. Laboratory breeding of the mud crab Scylla
848 serrata (Forskal) through the zoea and megalopa stages to the crab stage. *SEAFDEC*
849 *Aquaculture Department Quarterly Research Report,* 1**,** 14-18.
850 Narum, S. R. & Hess, J. E. 2011. Comparison of F(ST) outlier tests for SNP loci under selection.
851 *Mol Ecol Resour,* 11 Suppl 1**,** 184-94.
852 Nayfa, M. G. & Zenger, K. R. 2016. Unravelling the effects of gene flow and selection in highly
853 connected populations of the silver-lip pearl oyster (Pinctada maxima). *Mar Genomics,*
854 28**,** 99-106.

855 Nurdiani, R. & Zeng, C. 2007. Effects of temperature and salinity on the survival and
856 development of mud crab, Scylla serrata (Forsskål), larvae. *Aquaculture Research,* 38**,**
857 1529-1538.
858 Obata, Y., Imai, H., Kitakado, T., Hamasaki, K. & Kitada, S. 2006. The contribution of stocked
859 mud crabs Scylla paramamosain to commercial catches in Japan, estimated using a
860 genetic stock identification technique. *Fisheries Research,* 80**,** 113-121.
861 Ogawa, C. Y., Hamasaki, K., Kitada, S., Obata, Y. & Dan, S. 2012. Species composition,
862 reproduction, and body size of mud crabs, Scylla spp., caught in Urado Bay, Japan.
863 *Journal of Crustacean Biology,* 32**,** 762-768.
864 Oksanen, J., Blanchet, F., Friendly, M., Kindt, R., Legendre, P., Mcglinn, D., Minchin, P.,
865 O'hara, R., Simpson, G. & Solymos, P. 2019. vegan: Community Ecology Package. R
866 package version 2.5-6. 2019.
867 Oppo, D. W., Linsley, B. K., Rosenthal, Y., Dannenmann, S. & Beaufort, L. 2003. Orbital and
868 suborbital climate variability in the Sulu Sea, western tropical Pacific. *Geochemistry,*
869 *Geophysics, Geosystems,* 4**,** 1-20.
870 Palumbi, S. R. 2003. Population Genetics, Demographic Connectivity, and the Design of Marine
871 Reserves. *Ecological Applications,* 13**,** 146-158.
872 Paradis, E. 2010. pegas: an R package for population genetics with an integrated-modular
873 approach. *Bioinformatics,* 26**,** 419-20.
874 Paran, F. J. M. & Ravago-Gotanco, R. J. Genetic diversity and stock delineation of Philippine
875 populations of the orange mud crab, Scylla olivacea. Philippines: In the forefront of the
876 mud crab industry development: proceedings of the 1st National Mud Crab Congress, 16-

18 November 2015, Iloilo City, Philippines, 2017. Aquaculture Department, Southeast Asian Fisheries Development Center, 138.

Paris, C. B., Helgers, J., Van Sebille, E. & Srinivasan, A. 2013. Connectivity Modeling System: A probabilistic modeling tool for the multi-scale tracking of biotic and abiotic variability in the ocean. *Environmental Modelling & Software,* 42**,** 47-54.

Pata, P. R. & Yniguez, A. T. 2019. Larval connectivity patterns of the North Indo-West Pacific coral reefs. *PLoS One,* 14**,** e0219913.

Paterno, M., Schiavina, M., Aglieri, G., Ben Souissi, J., Boscari, E., Casagrandi, R., Chassanite, A., Chiantore, M., Congiu, L., Guarnieri, G., Kruschel, C., Macic, V., Marino, I. a. M., Papetti, C., Patarnello, T., Zane, L. & Melià, P. 2017. Population genomics meet Lagrangian simulations: Oceanographic patterns and long larval duration ensure connectivity among Paracentrotus lividus populations in the Adriatic and Ionian seas. *Ecology and Evolution,* 7**,** 2463-2479.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. & Hoekstra, H. E. 2012. Double Digest RADseq: An Inexpensive Method for De Novo SNP Discovery and Genotyping in Model and Non-Model Species. *PloS one,* 7**,** e37135.

Quinn, N. & Kojis, B. 1987. Reproductive Biology of Scylla Spp. (Crustacea: Portunidae) from the Labu Estuary in Papua New Guinea. *Bulletin of Marine Science,* 41**,** 234-241.

Ravago-Gotanco, R. & Kim, K. M. 2019. Regional genetic structure of sandfish Holothuria (Metriatyla) scabra populations across the Philippine archipelago. *Fisheries Research,* 209**,** 143-155.

Raynal, J. M., Crandall, E. D., Barber, P. H., Mahardika, G. N., Lagman, M. C. & Carpenter, K. E. 2014. Basin isolation and oceanographic features influencing lineage divergence in the

900 humbug damselfish <I>(Dascyllus aruanus)</I> in the Coral Triangle. *Bulletin of Marine*
901 *Science,* 90**,** 513-532.
902 Reeves, P. A., Bowker, C. L., Fettig, C. E., Tembrock, L. R. & Richards, C. M. 2016. Effect of
903 error and missing data on population structure inference using microsatellite data.
904 *bioRxiv***,** 80630-80630.
905 Reynolds, R. W., Smith, T. M., Liu, C., Chelton, D. B., Casey, K. S. & Schlax, M. G. 2007.
906 Daily high-resolution-blended analyses for sea surface temperature. *Journal of Climate,*
907 20**,** 5473-5496.
908 Riginos, C., Hock, K., Matias, A. M., Mumby, P. J., Van Oppen, M. J. H. & Lukoschek, V.
909 2019. Asymmetric dispersal is a critical element of concordance between biophysical
910 dispersal models and spatial genetic structure in Great Barrier Reef corals. *Diversity and*
911 *Distributions,* 25**,** 1684-1696.
912 Riginos, C. & Liggins, L. 2013. Seascape Genetics: Populations, Individuals, and Genes
913 Marooned and Adrift. *Geography Compass,* 7**,** 197-216.
914 Robertson, W. D. & Kruger, A. 1994. Size at Maturity, Mating and Spawning in the Portunid
915 Crab Scylla serrata (Forskål) in Natal, South Africa. *Estuarine, Coastal and Shelf*
916 *Science,* 39**,** 185-200.
917 Rochette, N. C., Rivera-Colon, A. G. & Catchen, J. M. 2019. Stacks 2: Analytical methods for
918 paired-end sequencing improve RADseq-based population genomics. *Mol Ecol,* 28**,**
919 4737-4754.
920 Rosly, H. A., Nor, S. A., Yahya, K. & Naim, D. M. 2013. Mitochondrial DNA diversity of mud
921 crab Scylla olivacea (Portunidae) in Peninsular Malaysia: a preliminary assessment. *Mol*
922 *Biol Rep,* 40**,** 6407-18.

923 Sandoval-Castillo, J., Robinson, N. A., Hart, A. M., Strain, L. W. S. & Beheregaray, L. B. 2018.
924 Seascape genomics reveals adaptive divergence in a connected and commercially
925 important mollusc, the greenlip abalone (Haliotis laevigata), along a longitudinal
926 environmental gradient. *Mol Ecol,* 27**,** 1603-1620.
927 Sanford, E. & Kelly, M. W. 2011. Local adaptation in marine invertebrates. *Annual review of*
928 *marine science,* 3**,** 509-535.
929 Schiavina, M., Marino, I. a. M., Zane, L. & Melià, P. 2014. Matching oceanography and genetics
930 at the basin scale. Seascape connectivity of the Mediterranean shore crab in the Adriatic
931 Sea. *Molecular Ecology,* 23**,** 5496-5507.
932 Schunter, C., Carreras-Carbonell, J., Macpherson, E., Tintoré, J., Vidal-Vijande, E., Pascual, A.,
933 Guidetti, P. & Pascual, M. 2011. Matching genetics with oceanography: Directional gene
934 flow in a Mediterranean fish species. *Molecular Ecology,* 20**,** 5167-5181.
935 Selkoe, K. A., D'aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., Von Der
936 Heyden, S. & Toonen, R. J. 2016. A decade of seascape genetics: contributions to basic
937 and applied marine connectivity. *Marine Ecology Progress Series,* 554**,** 1-19.
938 Sjöqvist, C., Godhe, A., Jonsson, P. R., Sundqvist, L. & Kremp, A. 2015. Local adaptation and
939 oceanographic connectivity patterns explain genetic differentiation of a marine diatom
940 across the North Sea-Baltic Sea salinity gradient. *Molecular Ecology,* 24**,** 2871-2885.
941 Swearer, S. E., Treml, E. A. & Shima, J. S. 2019. A Review of Biophysical Models of Marine
942 Larval Dispersal. *In:* HAWKINS, S. J., ALLCOCK, A. L., BATES, A. E., FIRTH, L. B.,
943 SMITH, I. P., SWEARER, S. E. & TODD, P. A. (eds.) *Oceanography and Marine*
944 *Biology: An Annual Review.* Taylor & Francis.

945 Teske, P. R., Sandoval-Castillo, J., Van Sebille, E., Waters, J. & Beheregaray, L. B. 2016.
946 Oceanography promotes self-recruitment in a planktonic larval disperser. *Scientific*
947 *Reports,* 6.
948 Thirunavukkarasu, N., Sheeba Anitha Nesacumari, C. & A. Shanmugam, A. 2014. Larval rearing
949 and seed production of mud crab Scylla tranquebarica (Fabricius , 1798). *International*
950 *Journal of Fisheries and Aquatic Studies,* 2**,** 19-25.
951 Tian, C., Kosoy, R., Nassir, R., Lee, A., Villoslada, P., Klareskog, L., Hammarstrom, L.,
952 Garchon, H. J., Pulver, A. E., Ransom, M., Gregersen, P. K. & Seldin, M. F. 2009.
953 European population genetic substructure: further definition of ancestry informative
954 markers for distinguishing among diverse European ethnic groups. *Mol Med,* 15**,** 371-83.
955 Toonen, R. J., Puritz, J. B., Forsman, Z. H., Whitney, J. L., Fernandez-Silva, I., Andrews, K. R.
956 & Bird, C. E. 2013. ezRAD: a simplified method for genomic genotyping in non-model
957 organisms. *PeerJ,* 1**,** e203.
958 Truelove, N. K., Kough, A. S., Behringer, D. C., Paris, C. B., Box, S. J., Preziosi, R. F. & Butler,
959 M. J. I. 2017. Biophysical connectivity explains population genetic structure in a highly
960 dispersive marine species. *Coral Reefs,* 36**,** 233-244.
961 Van Wyngaarden, M., Snelgrove, P. V., Dibacco, C., Hamilton, L. C., Rodriguez-Ezpeleta, N.,
962 Jeffery, N. W., Stanley, R. R. & Bradbury, I. R. 2017. Identifying patterns of dispersal,
963 connectivity and selection in the sea scallop, Placopecten magellanicus, using RADseq-
964 derived SNPs. *Evol Appl,* 10**,** 102-117.
965 Van Wyngaarden, M., Snelgrove, P. V. R., Dibacco, C., Hamilton, L. C., Rodriguez-Ezpeleta,
966 N., Zhan, L., Beiko, R. G. & Bradbury, I. R. 2018. Oceanographic variation influences

967 spatial genomic structure in the sea scallop, Placopecten magellanicus. *Ecol Evol,* 8**,**
968 2824-2841.
969 Vilas, A., Perez-Figueroa, A. & Caballero, A. 2012. A simulation study on the performance of
970 differentiation-based methods to detect selected loci using linked neutral markers. *J Evol*
971 *Biol,* 25**,** 1364-76.
972 Viswanathan, C., Pravinkumar, M., Suresh, T. V., Elumalai, V. & Raffi, S. M. 2019.
973 Reproductive biology of the orange mud crab Scylla olivacea (Herbst, 1796) from the
974 Pichavaram mangroves of south-east India. *Indian Journal of Fisheries,* 66.
975 Wang, L., Liu, S., Zhuang, Z., Guo, L., Meng, Z. & Lin, H. 2013. Population genetic studies
976 revealed local adaptation in a high gene-flow marine fish, the small yellow croaker
977 (Larimichthys polyactis). *PLoS One,* 8**,** e83493.
978 Wang, S., Meyer, E., Mckay, J. K. & Matz, M. V. 2012. 2b-RAD: a simple and flexible method
979 for genome-wide genotyping. *Nat Methods,* 9**,** 808-10.
980 Waples, R. S. 2010. Spatial-temporal stratifications in natural populations and how they affect
981 understanding and estimation of effective population size. *Mol Ecol Resour,* 10**,** 785-96.
982 Warnes, G., Bolker, B., Bonebakker, L., Gentleman, R., Liaw, W., Lumley, T., Maechler, M.,
983 Magnusson, A., Moeller, S. & Schwartz, M. 2020. gplots: Various R Programming Tools
984 for Plotting Data. R package version 3.1.0.
985 Warnes, G., Gorjanc, G., Leisch, F. & Man, M. 2019. genetics: Population genetics. R package
986 version 1.3. 8.1. 2. .
987 Weir, B. S. & Cockerham, C. C. 1984. Estimating F-Statistics for the Analysis of Population
988 Structure. *Evolution,* 38**,** 1358-1370.

Whitlock, M. C. & Mccauley, D. E. 1999. Indirect measures of gene flow and migration: FST not equal to 1/(4Nm + 1). *Heredity (Edinb),* 82 (Pt 2)**,** 117-25.

Xuereb, A., Benestan, L., Normandeau, E., Daigle, R. M., Curtis, J. M. R., Bernatchez, L. & Fortin, M. J. 2018. Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (Parastichopus californicus). *Mol Ecol,* 27**,** 2347-2364.

Zhang, Z., Schwartz, S., Wagner, L. & Miller, W. 2000. A greedy algorithm for aligning DNA sequences. *J Comput Biol,* 7**,** 203-14.

998 **Tables (separate file)**

999 Table 1. Summary information of *S. olivacea* sample location and genetic diversity estimates.

1000

1001 Table 2. Filtering processes and specific parameters used to filter the final SNP panel of *S.*
1002 *olivacea*. Each filtering step indicates the number of SNPs that were removed and retained.
1003 Filtering tools and programs are indicated in the methods section.

1004

1005 Table 3. Summary of the redundancy analysis (RDA) of genetic data (neutral and outlier loci-
1006 only) and environmental variables including the asymmetric eigenvector maps (AEMs), sea
1007 surface temperature (SST), and rainfall. Only those variables retained by backward selection and
1008 were significant ($p < 0.05$) are included in the final model. R^2_{adj} represents the adjusted
1009 coefficient of determination with p-values calculated using the analysis of variance (*anova*;
1010 permutations = 999). The proportion of constrained RDA axes were presented in columns and
1011 values in bold and with asterisk (*) indicate significant axes at 5% and 10% level, respectively.

1012

1013 **Figure legends**

1014 Figure 1. Estimated larval dispersal and connectivity of *S. olivacea* populations in the Sulu Sea
1015 basin from Lagrangian simulations. (a) Node connections plot generated from the larval dispersal
1016 estimates resulted in six population clusters indicated by the colored nodes and connections, gray
1017 links indicating connections between clusters. Site codes represent particles release sites and
1018 settlement nodes were labelled numerically. (b) Population connectivity matrix showing the

1019 proportion of larvae successfully settled from the source locations (y-axis) to settlement areas (x-
1020 axis). Sulu Sea populations were segregated according to their boundary location (east or west).
1021 See Table 1 for location codes.
1022
1023 Figure 2. Spatial patterns of *S. olivacea* connectivity based on (a) empirical genetic estimates
1024 (pairwise F_{ST}) from the putatively neutral loci (1,643 SNPs); and (b) asymmetric larval dispersal
1025 estimates (AEM6) from the particle dispersal simulations in the Sulu Sea. Sampling points are
1026 colored according to their cluster assignment. Persistent surface currents in the Sulu Sea are
1027 shown.
1028
1029 Figure 3. Heatmap of pairwise genetic differentiation (F_{ST}) of *S. olivacea* between sampling
1030 locations in the Sulu Sea, based on Weir and Cockerham weighted estimates using (a) all
1031 markers (1,655 loci), (b) neutral-only markers (1,643 SNPs), and (c) outlier only (12 SNPs). All
1032 points were clustered by pairwise F_{ST} values, based on classification and hierarchical clustering
1033 method. Site combinations (below diagonal) with boxed bold lines indicate significant genetic
1034 differentiation following FDR correction for multiple tests ($p < 0.05$).
1035
1036 Figure 4. (a) Scatterplot of genetic clusters of Sulu Sea populations identified by DAPC based on
1037 neutral loci (1,643 SNPs) and sampling location information as a prior, with density plots for (b)
1038 the first and (c) the second discriminant axes. Genetic clusters identified using outlier loci (12
1039 SNPs) and sampling location information as a prior (d).
1040

Figure 5. Distribution of genetic clusters of *S. olivacea* in the Sulu Sea based on outlier loci and Geneland analysis. Posterior probability isoclines illustrate putative genetic landscapes for the Sulu Sea domain, where sites are represented by black dots, and darker colors indicate higher probabilities of membership to each of the six clusters identified across all sampling sites. Isoclines for the outgroup populations representing two genetic clusters are not shown.

Data Accessibility Statement

Raw demultiplexed sequence libraries in fastq format were archived in NCBI SRA (BioProject Accession # PJRNA662443) and will be released upon publication. The corresponding filtered datasets and R scripts for analyses have been deposited in Dryad Digital Repository: https://doi.org/10.5061/dryad.3xsj3txdz

Author Contribution Section

MJM: Conceptualization (equal), Data Curation, Formal Analysis (lead), Investigation, Methodology (equal), Visualization, Writing-Original Draft (equal)

RRG: Conceptualization (equal), Formal Analysis (supporting), Funding Acquisition, Methodology (equal), Project Administration, Supervision, Writing-Original Draft (equal)

Acknowledgements

This project was funded by the Department of Science and Technology - Philippine Council for Agriculture, Aquatic, and Natural Resources Research and Development (DOST-

PCAARRD) and implemented by the University of the Philippines Marine Science Institute (UP-MSI). We are deeply thankful to Dr. Evangeline Magdaong, Jeniffer De Maligaya, and Benedict Castro of the Physical Oceanography Laboratory, UP-MSI headed by Dr. Cesar Villanoy for the larval dispersal biophysical modelling. We are grateful to Angela Camille Aguila and Simon Alcantara for laboratory assistance, Bhenjamin Ona for the remote sensing data, Dr. Din Matias for analysis recommendations, Von Yip for QGIS assistance, Dr. Richard Mualil and Yunadzmal Ong of Mindanao State University (MSU Tawi-Tawi) for sample collection. We thank Sharon Magnuson and Chris Bird (Genomics Core Lab, Texas A&M University, Corpus Christi) for RAD sequencing.

Ethical Approval

Scylla olivacea is a commercially harvested species and was not considered as a regulated species during the time of sampling. For Palawan, collection and local transport were covered under Palawan Council for Sustainable Development (PCSD) local transport permits and Wildlife Gratuitous Permit (GP) no. 2016-23. Export of material for DNA sequencing was covered under the Bureau of Fisheries and Aquatic Resources GP no. 2018-0005.

Conflict of Interest

None declared.