WILEY MOLECULAR ECOLOGY

Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*)

Amanda Xuereb[1] [iD] | Laura Benestan[2] | Éric Normandeau[2] | Rémi M. Daigle[1] | Janelle M. R. Curtis[3] | Louis Bernatchez[2] | Marie-Josée Fortin[1]

[1]Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada

[2]Institut de Biologie Intégrative et des Systèmes, Université Laval, Québec, QC, Canada

[3]Pacific Biological Station, Ecosystem Sciences Division, Fisheries and Oceans Canada, Nanaimo, BC, Canada

Correspondence
Amanda Xuereb, Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, ON, Canada.
Email: amanda.xuereb@mail.utoronto.ca

Funding information
Natural Sciences and Engineering Research Council of Canada, Grant/Award Number: 5134, D3-460408-2014, STPGP 430706-2012; Fisheries and Oceans Canada (DFO); Program for Aquaculture Regulatory Research (PARR)

Abstract

Marine populations are typically characterized by weak genetic differentiation due to the potential for long-distance dispersal favouring high levels of gene flow. However, strong directional advection of water masses or retentive hydrodynamic forces can influence the degree of genetic exchange among marine populations. To determine the oceanographic drivers of genetic structure in a highly dispersive marine invertebrate, the giant California sea cucumber (*Parastichopus californicus*), we first tested for the presence of genetic discontinuities along the coast of North America in the northeastern Pacific Ocean. Then, we tested two hypotheses regarding spatial processes influencing population structure: (i) *isolation by distance* (IBD: genetic structure is explained by geographic distance) and (ii) *isolation by resistance* (IBR: genetic structure is driven by ocean circulation). Using RADseq, we genotyped 717 individuals from 24 sampling locations across 2,719 neutral SNPs to assess the degree of population differentiation and integrated estimates of genetic variation with inferred connectivity probabilities from a biophysical model of larval dispersal mediated by ocean currents. We identified two clusters separating north and south regions, as well as significant, albeit weak, substructure within regions (F_{ST} = 0.002, p = .001). After modelling the asymmetric nature of ocean currents, we demonstrated that local oceanography (IBR) was a better predictor of genetic variation (R^2 = .49) than geographic distance (IBD) (R^2 = .18), and directional processes played an important role in shaping fine-scale structure. Our study contributes to the growing body of literature identifying significant population structure in marine systems and has important implications for the spatial management of *P. californicus* and other exploited marine species.

KEYWORDS
asymmetric eigenvector maps, marine connectivity, population genomics, RAD sequencing, sea cucumber, seascape genetics

1 | INTRODUCTION

Documenting the spatial patterns of population genetic structure and the features that shape marine connectivity is key to refining conservation efforts such as the design of marine reserve networks at appropriate scales (Almany et al., 2009; Cros, Toonen, Donahue, & Karl, 2017) and fisheries management (Bernatchez et al., 2017). In the marine environment, species have been characterized by a high propensity for dispersal, presumably resulting in widespread inter-population connectivity and homogenous genetic variation across space (Purcell, Cowen, Hughes, & Williams, 2006; Waples, 1998). Additionally, large effective population sizes typical for marine systems are predicted to dampen the effects of random genetic drift, further limiting population divergence at neutral loci (Waples, Punt, & Cope, 2008). As such, the potential for a high degree of genetic exchange combined with weak genetic drift is assumed to result in very low levels of genetic differentiation with an absence of clear spatial patterns of population genetic structure in marine systems.

Indeed, the vast majority of marine species undergo a pelagic larval stage, during which time dispersing individuals are subjected to the movement of water masses (Cowen & Sponaugle, 2009). Depending on the development time of larvae, (i.e., pelagic larval duration; PLD), they can remain in the water column for weeks or months and have the potential to be carried over great distances with prevailing ocean currents (Cowen, Gawarkiewicz, Pineda, Thorrold, & Werner, 2007; Kinlan, Gaines, & Lester, 2005). In this context, marine connectivity is largely defined by the exchange of larvae between natal and downstream sites (Treml et al., 2012). Despite the potential for long-distance dispersal, it is increasingly apparent that the scale of marine connectivity is often smaller, or more complex, than previously assumed, owing to a combination of seascape features and larval behaviours favouring local retention and restricted dispersal (D'Aloia et al., 2015; Selkoe et al., 2010).

While direct observations of larval dispersal between marine populations are often intractable, genomic tools provide a powerful means for quantifying marine connectivity (Gagnaire et al., 2015; Ovenden, Berry, Welch, Buckworth, & Dichmont, 2015). A coupled approach, combining estimates of genetic differentiation and oceanographic data, can provide valuable insights into the influence of ocean circulation on patterns of gene flow across seascapes (Gilg & Hilbish, 2003). This integration of genetic and oceanographic information forms the basis of a *seascape genetics* framework, which aims to identify physical and environmental features that shape spatial patterns of genetic variation (Riginos, Crandall, Liggins, Bongaerts, & Treml, 2016; Selkoe et al., 2016). Ocean circulation dynamics are known to play an important role in shaping patterns of connectivity among marine populations, either by facilitating the dispersal of individuals via advection of water masses or by circular forces such as eddies or gyres that increase local retention, thereby limiting dispersal to distant sites and restricting the scale of gene flow (Banks et al., 2007; Kough, Paris, & Butler, 2013; White et al., 2010). The highly variable and asymmetric nature of oceanic flow

regimes has important consequences on dispersal and gene flow in marine systems, resulting in highly anisotropic spatial patterns of connectivity. The use of biophysical models—combining oceanographic and biological parameters to simulate larval dispersal trajectories based on hydrodynamic processes—is a compelling approach to elucidate the asymmetric transport of larvae between natal and downstream sites (Cowen, Paris, & Srinivasan, 2006; Treml, Halpin, Urban, & Pratson, 2008). For example, Schunter et al. (2011) identified asymmetric gene flow across putative geographic barriers in the Mediterranean Sea contributing to observed spatial patterns of genetic variation among populations of the demersal fish *Serranus cabrilla*. As such, observed spatial patterns of genetic differentiation are often more complex than previously assumed given the potential for widespread gene flow and large-scale panmixia (Selkoe et al., 2010; White et al., 2010). Indeed, an increasing number of studies detect fine-scale patterns of genetic structure attributed to oceanographic features in many marine taxa including American lobster (*Homarus americanus*—Benestan, Quinn, et al., 2016): corals (*Montastraea annularis*—Foster et al., 2012; *Acanthophora spicifera*—Thomas et al., 2015; *Acropora hyacinthus*—Cros et al., 2017), oysters (*Pinctada margaritifera*—Lal, Southgate, Jerry, Bosserelle, & Zenger, 2017), shrimp (*Pandalus borealis*—Jorde et al., 2015) and spiny lobsters (*Jasus edwardsii*—Thomas & Bell, 2013; *Panulirus penicillatus*—Iacchei, Gaither, Bowen, & Toonen, 2016; *P. argus*—Truelove et al., 2017).

Here, we characterize spatial patterns of genetic variation and connectivity among populations of the giant California sea cucumber (*Parastichopus californicus*) in the northeastern Pacific Ocean. This region is dominated by complex oceanographic features, which have the potential to affect the degree of connectivity. Our objectives for this study were twofold. First, we determined whether *P. californicus* populations exhibit distinct spatial patterns of genetic structure in the northeastern Pacific Ocean along the west coast of North America despite the potential for widespread gene flow. A previous assessment of population genetic structure of this species using microsatellite markers identified significant genetic divergence between a sampled site on the west coast of Vancouver Island and other sampled areas, indicating genetic isolation (Nelson, 2003). Our study expands upon this early work using single nucleotide polymorphisms (SNPs) derived from restriction site-associated DNA sequencing (RADseq) to improve the resolution for detecting finer scale spatial patterns of genetic structure. Second, we tested two hypotheses regarding the spatial and oceanographic drivers of population genetic differentiation and marine connectivity: (i) isolation by distance (IBD), whereby genetic differentiation is a function of geographic distance and (ii) isolation by resistance (IBR), which predicts that ocean circulation dynamics shape spatial patterns of genetic variation. To this end, we integrated empirical genetic data with inferred dispersal probabilities from a biophysical model of larval dispersal based on the hydrographic environment of the British Columbia (BC) coastal region. We assessed the topology of a graph (i.e., network) based on the genetic covariance structure among all sampled locations to characterize the degree of genetic connectivity

among sites and incorporated a spatial filtering approach that explicitly models spatial genetic structure produced by directional current flow to determine the relative effect of asymmetric processes on marine connectivity.

## 2 | MATERIALS AND METHODS

2.1 | Study species

The Giant California sea cucumber, *Parastichopus californicus*, is a benthic invertebrate that occurs at a wide range of depths, from the intertidal zone up to 250 m below the surface (Lambert, 1997). Adult *P. californicus* inhabit rocky substrates, where daily movements are limited to a few metres while feeding on organic matter in the bottom sediment (Lambert, 1997). Spawning occurs annually between June and September via the broadcast of gametes by both sexes (Cameron & Fankboner, 1986). Although the species is relatively sedentary as an adult, *P. californicus* larvae have the potential to disperse over large distances with a PLD estimated between 30 and 120 days (Cameron & Fankboner, 1989; Lambert, 1997). The species is widely distributed along the Pacific coast of North America, from the Gulf of Alaska to Baja California, and is harvested for human consumption throughout most of its range (Hamel & Mercier, 2008). It is the only commercially exploited sea cucumber species in western Canada (DFO, 2016).

2.2 | Sampling location and collections

We focused our sampling of *P. californicus* in the northeastern Pacific coastal region. Here, the strong North Pacific Current (NPC) bifurcates into the northward-flowing Alaska Current and the southward-flowing California Current right around the central coast of BC (Foreman, Callendar, Masson, Morrison, & Fine, 2014; Masson & Fine, 2012) (Figure 1). Tissue samples were collected from 20 sites spanning coastal BC (Canada), including across the bifurcation zone of the NPC, as well as from an additional four sites in southeastern Alaska (USA) (Figure 1) to assess spatial patterns of genetic structure across an international boundary. Collections were made during the summer and early fall (June to October) in 2014 and 2015. Samples consisted of a small clip (~1 cm) from a single dorsal papilla from 30 to 41 individuals per site (*n* = 840 total clips collected; Table 1). Papillae clips were preserved in 95% ethanol in individual 1.5-ml microtubes for DNA extraction. Sampling equipment was sterilized with ethanol between all specimens. All samples were collected via SCUBA diving at a depth of ~20 m, and individuals were immediately returned to the capture location after processing. Collections were made by Fisheries and Oceans Canada (DFO) and the Alaska Department of Fish and Game (DFG).

2.3 | DNA extraction and sequencing

Tissues were lysed with metal beads using the TissueLyser (QIAGEN), followed by addition of 20 µl of proteinase K before
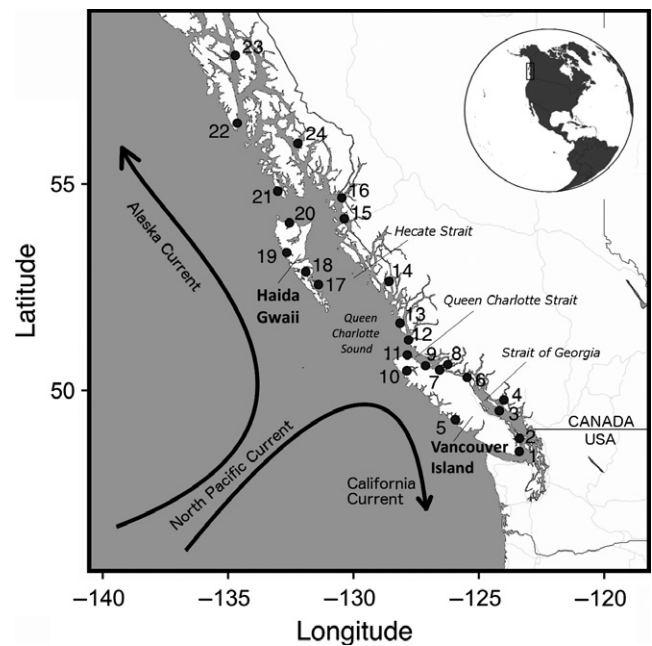
FIGURE 1 Map of sampling locations in coastal British Columbia (1–20) and southeastern Alaska (21–24). Site labels correspond with numbers in Table 1

incubating for 6–8 hr. Once the tissue was fully digested, whole genomic DNA was extracted and purified using the DNeasy blood and tissue spin-column protocol (QIAGEN, Toronto, ON, Canada). DNA was eluted in two steps, each with 40 µl of ddH$_2$O. The presence of DNA was verified on an agarose gel stained with ethidium bromide, and the concentration was measured using the NanoDrop 2000c (Thermo Scientific). All samples were desalted and concentrated using the Microcon DNA fast-flow filter system (EMD Millipore Corporation). DNA concentrations were quantified in 96-well plates using the AccuClear Ultra High Sensitivity dsDNA Quantitation Kit (Biotium, Inc., Fremont, CA, USA) and normalized to a final concentration of 20 ng/µl. Plates were prepared with a sample volume of 10 µl, for a total of 200 ng of DNA per individual. Libraries for restriction site-associated DNA sequencing (RADseq) were prepared using the single-end double-digest protocol (i.e., ddRAD) with restriction enzymes *MspI* and *PstI* described in Poland, Brown, Sorrells, and Jannink (2012). Single-end sequencing was performed using the Ion Proton™ Sequencer (Life Technologies, Grand Island, NY, USA) at the core sequencing facility at the Institut de Biologie Intégrative et des Systèmes at Université Laval (Québec, Canada).

2.4 | Raw data filtering and SNP identification

Adapter sequences were removed from raw sequencing reads using Cutadapt (Martin, 2011). Libraries were demultiplexed, and raw reads were trimmed to 70 base pairs using *process_radtags* in STACKS v.1.44 (Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013). Samples with <800,000 reads were removed, and samples with >6,000,000 reads were subsampled to 6,000,000 reads. Trimmed reads (average of 2.75 million reads per sample) were aligned to a

TABLE 1 Geographic coordinates, number of samples collected (N_s) and successfully genotyped (N_g), observed and expected heterozygosity (H_o and H_e, respectively), the measured inbreeding coefficient (G_{IS}), and the betweenness centrality calculated from the Population Graph of genetic covariance across sampling locations

No.	Site	Site code	Longitude	Latitude	N_s	N_g	H_o	H_e	G_{is}	Betweenness centrality
South region										
1	Ogden Point	OGD	−123.387	48.408	33	31	0.114	0.117	0.021	11
2	Southern Gulf Islands	SGI	−123.380	48.757	41	31	0.105	0.111	0.053	9
3	Lasqueti	LAS	−124.183	49.475	40	28	0.111	0.115	0.038	11
4	Jervis Inlet	JER	−124.001	49.753	41	32	0.110	0.115	0.041	4
5	Tofino	TOF	−125.938	49.248	30	20	0.105	0.110	0.049	12
6	Rock Bay	RBY	−125.467	50.330	41	32	0.108	0.113	0.044	2
7	Cracroft Island	CRA	−126.565	50.521	31	28	0.103	0.111	0.067	23
8	Shewell Island	SHE	−126.238	50.659	41	32	0.106	0.113	0.059	8
9	Malcolm Island	MAL	−127.128	50.627	41	33	0.11	0.115	0.043	5
10	Quatsino	QUA	−127.872	50.500	41	32	0.105	0.112	0.061	75
11	Hope Island	HOP	−127.851	50.899	41	34	0.11	0.115	0.044	13
12	Table Island	TBL	−127.804	51.272	41	32	0.108	0.114	0.051	6
North region										
13	Calvert Island	CAL	−128.143	51.690	30	30	0.112	0.114	0.013	96
14	Tolmie	TOL	−128.578	52.713	30	29	0.111	0.114	0.021	16
15	Prince Rupert	PRI	−130.366	54.197	31	30	0.114	0.115	0.012	0
16	Legace Bay	LEG	−130.464	54.682	31	28	0.115	0.116	0.01	3
17	Juan Perez	JUA	−131.396	52.632	34	30	0.111	0.113	0.017	3
18	Selwyn	SEL	−131.905	52.939	34	32	0.112	0.114	0.017	1
19	Rennell Sound	REN	−132.66	53.399	33	30	0.115	0.116	0.014	24
20	Mazarredo	MAZ	−132.553	54.100	35	30	0.113	0.114	0.011	11
21	Alaska 1	AK1	−133.019	54.835	30	28	0.111	0.112	0.013	35
22	Alaska 2	AK2	−134.633	56.364	30	27	0.113	0.114	0.006	3
23	Alaska 3	AK3	−134.715	57.828	30	28	0.112	0.115	0.022	16
24	Alaska 4	AK4	−132.217	55.914	30	30	0.112	0.115	0.022	10

reference genome of a closely related species (*Parastichopus parvimensis*) (Cameron, Samanta, Yuan, He, & Davidson, 2009) using GSNAP v.2015-12-31 (Wu & Nacu, 2010). The *P. parvimensis* genome was 771,595,166 bp, with 150,862 scaffolds and a N50 size of 9,587. On average, ~1.79 million reads per sample aligned with the reference genome. Stacks (loci) aligned to the reference genome were extracted using the *pstacks* module of STACKS with a minimum stack depth of four ($m = 4$). A catalog of ~1.81 million putative loci was built using the *cstacks* module, allowing for a maximum of three mismatches between loci ($n = 3$). To build this catalog, we used a subsample of eight individuals per site (i.e., one-third of the samples per site), selecting the individuals that showed the highest number of reads, to increase computational efficiency and based on developer recommendations (Julian Catchen, *pers. comm.*). We used the *populations* module of STACKS to first retain SNPs with a minimum stack depth of four ($m = 4$) and present in at least 16 sampling locations with at least 60% of data in each location (total number of SNPs retained = 94,842; see Table S1). Additional filtering steps (see Table S1) were performed to subsequently exclude

SNPs that were not genotyped in all sampling locations and in at least 70% of individuals within each sampling location. Markers with observed heterozygosity (H_o) >0.6 in at least one site were also excluded. We removed markers with an inbreeding coefficient (F_{IS}) >0.5 or <−0.5. This step excluded markers with either a deficiency or excess of heterozygosity. Finally, we retained SNPs with either a minor allele frequency (MAF) >0.01 across all sites or an MAF >0.1 in at least one site. We evaluated the amount of missing data across individuals using the package STACKR v.0.4.5 (Gosselin & Bernatchez, 2016) in R (R Core Team 2016) and excluded individual samples with >30% missing data across loci from subsequent analyses. All the filtering steps were performed according to the recommendations from Benestan, Ferchaud, et al. (2016). The filtered VCF file was converted into formats necessary for subsequent analyses using PLINK v.1.9 (Purcell et al., 2007) and PGDSPIDER v.2.0.5.0 (Lischer & Excoffier, 2012). We tested for linkage disequilibrium between all marker pairs using VCFTOOLS v.0.1.15 (Danecek et al., 2011) and excluded one marker from each pair with R^2 >0.8, following Larson et al. (2014).

2.5 | Identifying neutral SNPs

SNPs potentially under selection were identified and removed from our data set using a differentiation-based outlier detection approach to retain a panel of neutral SNPs for assessment of population structure and connectivity (Beaumont & Nichols, 1996). We used BayeScan, a Bayesian method for detecting candidate SNPs by estimating the posterior probability of SNP markers being under selection based on alternative models with and without selection (Foll & Gaggiotti, 2008). We ran BayeScan for 10,000 iterations and a burn-in of 200,000 steps. We set the prior odds of neutrality parameter (*pr_odds*) to 10,000 based on recommendations from Lotterhos and Whitlock (2014). This parameter represents the prior probability of a locus being under selection (Foll & Gaggiotti, 2008), and Lotterhos and Whitlock (2014) showed that prior odds equal to 10,000 resulted in lower error rates compared with prior odds equal to 10, 100 or 1,000. We defined a false discovery rate (FDR) *q*-value threshold of 0.01.

2.6 | Population genetic structure

Observed and expected heterozygosities (H_o and H_e, respectively) and inbreeding coefficients (G_{IS}) within all 24 sampling locations were estimated using GenoDive (Meirmans & Van Tienderen, 2004). Individual ancestries were estimated based on admixture proportions (Q) using ADMIXTURE v1.23 (Alexander, Novembre, & Lange, 2009), with the number of genetic clusters *K* ranging from 2 to 25. The optimal number of groups was selected based on the *K* value with the lowest fivefold cross-validation error (Alexander & Lange, 2011). We also performed a discriminant analysis of principal components (DAPC) using the R package ADEGENET v2.0.1 (Jombart, 2008). This analysis first transforms the data using a principal component analysis (PCA) followed by a discriminant analysis on the uncorrelated PCA variables and produces synthetic discriminant functions (axes) that maximize between-group variation while minimizing within-group variation (Jombart, Devillard, & Balloux, 2010). We performed DAPC first with no prior information, using the *find.clusters* function to determine the optimal number of groups. This function uses a sequential *k*-means clustering algorithm, and the best-supported model (i.e., number of clusters) is chosen using the Bayesian information criterion (BIC) (Jombart et al., 2010). We also performed DAPC including the sampling location of each individual as prior information. In both cases (with and without priors), the number of principal components (PCs) to retain was chosen according to the α-score using the *optim.a.score* function with 20 replicates to avoid issues of over-fitting with too many PCs. Genetic differentiation between sites was calculated based on pairwise estimates of F_{ST} (Weir & Cockerham, 1984) with 5,000 permutations using GenoDive and *p*-values were adjusted for multiple tests using the Benjamini–Hochberg method of controlling the false discovery rate (BH) (Benjamini & Hochberg, 1994). We also performed a hierarchical analysis of molecular variance (AMOVA) (Excoffier, Smouse, & Quattro, 1992) to test for population substructure within broad-scale groups using

GenoDive with 1,000 permutations. A hierarchical UPGMA clustering analysis based on pairwise F_{ST} was performed in R, and cluster support was determined using the PVCLUST v.2.0-0 package in R (Suzuki & Shimodaira, 2015) with 10,000 replications to compute approximately unbiased *p*-values based on multiscale bootstrap resampling.

2.7 | Assignment test

We performed an assignment test for the genetic structure observed between north and south regions (see Results). Assignment tests were not performed at a finer level (i.e., subpopulations) as too few individuals were genotyped per subpopulation to accurately calculate allele frequencies. Using all 2,719 SNPs, we performed a leave-one-out (LOO) test in GenoDive, which generates a null distribution of likelihood values with 5,000 permutations and a threshold calculated for each population separately based on a α-value of 0.05. Second, we tested how many markers were required to achieve high assignment success (i.e., >80%). We used the *assignment_ngs* function in the R package ASSIGNER v.0.4.1 (Gosselin, Anderson, & Bradbury, 2016) with the *gsi_sim* algorithm (Anderson, 2010). To avoid high-grading bias, we used the "training, holdout, leave-one-out" (THL) method with 10 iterations, and with ranked SNPs ranging from 10 to the total number of SNPs.

2.8 | Population Graphs

Population Graphs were developed as an approach to analyse patterns of population genetic structure and connectivity within a graph-theoretic framework, where *nodes* represent sample populations and *edges* represent multivariate genetic covariance measures between populations (Dyer & Nason, 2004). This approach differs from commonly used statistical methods for characterizing population genetic structure (e.g., *F*-statistics) in that it does not rely on *a priori* definition of an underlying population model. Instead, genetic relationships are quantified based on complex graphical topologies that represent the genetic covariance structure among all populations, whereby edge length is inversely proportional to the genetic covariance between a given pair of populations (Dyer & Nason, 2004; Dyer, Nason, & Garrick, 2010). Using statistical metrics of conditional independence, the smallest edge set (i.e., set of genetic covariance relationships) that adequately explains the population genetic covariance structure is identified, and edges that do not contribute to the overall covariance structure are removed.

To characterize the graph topology and identify the minimal edge set that explains genetic covariance among the 24 sites, we used the R packages GSTUDIO v.1.5.0 (Dyer, 2016) and POPGRAPH v.1.5.0 (Dyer, 2017). We determined whether the number of edges was significantly less than the total number of possible edges (i.e., in a saturated graph) using a binomial test. We calculated betweenness centrality, which indicates the number of shortest paths that pass through a given node (i.e., sampling location), as a metric of the

importance of each site for maintaining connectivity across the entire graph (Urban, Minor, Treml, & Schick, 2009).

2.9 | Biophysical model

We used a biophysical model of larval dispersal to estimate the probability of connectivity due to oceanographic circulation (see Foreman et al., 2014; Morrison, Callendar, Foreman, Masson, & Fine, 2014; Daigle, 2016 for a detailed description of the biophysical model). Briefly, the biophysical model used a coupled Lagrangian particle tracking (LTRANS v.2b) (North et al., 2011) and hydrodynamic model of ocean circulation over a ten-year period between 1998 and 2007 (Regional Ocean Modeling System; ROMS) (Masson & Fine, 2012) to simulate passive larval dispersal between patches. Patches were defined as 20×20 km grid cells ($n = 548$ grid cells) overlaid on the BC continental shelf up to a depth of 250 m. Dispersal probabilities between all pairs of patches (in both directions) were quantified as the proportion of larvae that moved from source patch i (i.e., one grid cell) to destination patch j after 120 days, as this is the upper estimate of the PLD for *P. californicus* (Lambert, 1997). The mean dispersal probabilities over the 10 years of the simulation were used to generate a connectivity matrix (Table S3). Because the extent of the biophysical model did not include Alaskan waters, we retained the dispersal probabilities between grid cells that contained each of the 20 sampling locations in BC, and we excluded the four sampling sites in Alaska from subsequent analyses.

2.10 | Isolation by distance and isolation by resistance

We tested for a correlation between genetic distance (linearized F_{ST}: $F_{ST}/(1-F_{ST})$) and the logarithm of both Euclidean (straight-line) and in-water geographic distance between all pairs of sampling locations (IBD) as well as between genetic distance and oceanographic distance based on biophysical simulations (IBR) using Mantel tests. In-water geographic distances between all pairs of sites were measured along the least-cost path (i.e., the shortest path excluding over land) using the R package IGRAPH v.1.1.2 (Csardi & Nepusz, 2006). Distances based on the local oceanography were calculated as the inverse of the probability of connectivity between all pairs of sites derived from the biophysical model of larval dispersal. Mantel tests were performed using the ECODIST v.1.2.9 package in R (Goslee & Urban, 2007) with 10,000 permutations.

We also used a spatial eigenfunction analysis (SEA) to evaluate the relative contribution of geographic distance and asymmetric processes driven by current circulation on multiscale spatial patterns of genetic variation. SEA is used to decompose spatial configuration into a set of synthetic vectors that can subsequently be used as predictor variables in a model (e.g., linear regression or multivariate canonical analysis) (Dray et al., 2012). We calculated two types of spatial variables: (i) asymmetric eigenvector maps (AEM) (Blanchet, Legendre, Maranger, Monti, & Pepin, 2011) and (ii) distance-based

Moran's eigenvector maps (dbMEM) (Dray, Legendre, & Peres-Neto, 2006), to represent both directional (e.g., current circulation) and symmetric (e.g., spatial distance) processes, respectively. AEM variables were generated by translating the presence of a connection (i.e., dispersal probability from the biophysical model >0) between all pairs of sites into a site-by-edge matrix. We attributed a weight to each edge, which was based on the probability of dispersal between each pair of sites derived from the biophysical model. When connectivity between a given pair of sites was >0 in both directions, we selected the direction with the highest probability of dispersal. The dbMEM variables were generated by decomposing Euclidean (geographic) distances between sites. We also computed dbMEM variables based on in-water distances between sites. For both AEMs and dbMEMs, the larger eigenvectors model broad-scale spatial structures, whereas the smaller eigenvectors model fine-scale spatial structures (Dray et al., 2012). We used the package ADESPATIAL v.0.0-8 (Dray et al., 2017) to generate both types of spatial variables.

We evaluated the effect of geography and directional currents on spatial patterns of neutral genetic variation using the derived dbMEM and AEM predictor variables using a redundancy analysis (RDA). We performed a principal component analysis (PCA) on the Hellinger-transformed allele frequencies (Legendre & Gallagher, 2001) and retained the PC axes that explained at least 5% of the total genetic variation as response variables in subsequent RDAs. The most important explanatory AEM and dbMEM variables were identified using a forward selection procedure with 10,000 permutations with the R package VEGAN v.2.4-4 (Oksanen et al., 2017). We used global and marginal analyses of variance (ANOVA) with 1,000 permutations to assess the significance of the model and evaluate the contribution of each spatial variable. Finally, we performed a partial RDA to partition the variance explained between the two sets of spatial eigenfunctions.

3 | RESULTS

3.1 | Individual and marker filtering and genetic diversity

After initial filtering steps, 4,340 SNPs were retained (Table S1). After removing one locus from each pair in linkage disequilibrium based on an R^2 >80%, we kept 3,699 SNPs successfully genotyped in a total of 717 individuals across 24 sampling locations (Table 2). Estimates of H_o and H_e over the filtered 3,699 loci were similar across sampling locations ($H_o = 0.103-0.115$; $H_e = 0.110-0.117$), and G_{IS} was low across all sites ($G_{IS} = 0.014-0.081$; Table 1).

3.2 | Detecting putatively neutral SNPs

BayeScan detected 925 SNPs potentially under balancing selection (25%) and 55 SNPs potentially influenced by divergent selection (0.015%) with a significance level of 1% (Figure S1). Therefore, the remaining 2,719 SNPs were retained as putatively neutral.

TABLE 2 Comparison of RDA results with spatial eigenfunction predictor variables including distance-based Moran's eigenvector map (dbMEM) and asymmetric eigenvector maps (AEMs). Only variables that were included in the best model by forward selection and were significant based on a marginal ANOVA are shown. The proportion of variation explained by each RDA axis is indicated under their headings with significant axes ($p < .05$) shown in bold text. R^2 represents the adjusted coefficient of determination; p-values were evaluated using an ANOVA with 999 permutations

	Significant variables	RDA1	RDA2	RDA3	R^2	p
dbMEM (Euclidean)	dbMEM1 dbMEM2	**0.235**	0.030	NA	.179	.001
dbMEM (in-water)	dbMEM2 dbMEM19	**0.226**	0.078	NA	.222	.001
AEM	AEM1 AEM2 AEM4 AEM8 AEM11 AEM17	**0.251**	**0.139**	**0.102**	.485	.001

3.3 | Population structure

Two genetic clusters ($K = 2$) were identified using ADMIXTURE based on the lowest cross-validation error (Figure 2a). The proportions of individual ancestry showed a split between sites north of the Queen Charlotte Sound and sites around Vancouver Island, which are hereafter referred to as the "north" and "south" regional groups (Figure 2b). We subsequently ran ADMIXTURE on each of the two putative clusters independently to determine whether subpopulations could be identified within each region. However, no clear substructure based on individual admixture proportions was revealed at this scale (Figure S2).

The DAPC with no prior information identified two genetic clusters based on the lowest BIC score (Figure S3). The two groups showed little overlap on one discriminant axis (Figure S3), which explained all of the between-group variation. More than 80% of all individuals were identified as belonging in a single cluster. Of the remaining individuals identifying with the second cluster, 78% were sampled from sites within the south region. When the DAPC included a priori sampling sites, 99 PCs were retained based on the α-score, and the first and second discriminant functions explained 38% and 9% of the variance, respectively (Figure 3). The first discriminant axis (DAPC 1) highlighted the dichotomy between the north and south regions. Furthermore, the second axis (DAPC 2) revealed a separation of Rennell Sound from all the other sites located in the north region. No apparent structure was observed on additional axes.

Although overall genetic differentiation (F_{ST}) was low (overall $F_{ST} = 0.0051$), 85% of the pairwise F_{ST} values remained significant after applying the FDR correction (Table S2), indicating the existence of significant population genetic structure. In agreement with the ADMIXTURE and DAPC analyses, hierarchical clustering based on pairwise F_{ST} values between all sampling locations revealed high

support for the discrimination between north and south regional populations (Figure 4a), and the AMOVA showed significant genetic differentiation between regions (AMOVA $F_{CT} = 0.0040$, $p = .001$). Additionally, a unique cluster corresponding to Rennell Sound (REN) was significantly differentiated from all other sites, as observed on the second DAPC axis. Finer scale population substructure was detected by the AMOVA, as indicated by significant genetic differentiation between sites within the north and south regional groups ($F_{ST} = 0.002$, $p = .001$). To resolve population substructure, we performed hierarchical clustering within the north and south regions independently. In the south region, two well-supported groups were, respectively, composed of neighbouring sites within the Queen Charlotte Strait (QCS) and Strait of Georgia (SOG) (Figure 4b) separated by the Johnstone Strait. Though geographically located within QCS, CRA clustered together with sites in SOG; however, this site shows significant genetic differentiation with all sampling locations except HOP, which was marginally significant ($F_{ST} = 0.0014$; $p = .06$) and LAS ($F_{ST} = 0.0005$; $p = .28$). QUA and TOF formed a third cluster, although these sites were also significantly differentiated from one another ($F_{ST} = 0.0032$; $p = .003$). In the north region, one highly supported group composed of sites within Hecate Strait, together with CAL and AK4 could be distinguished (HEC; Figure 4c). A cluster composed of the three remaining sites in Alaska (AK1, AK2 and AK3) formed a second group (AK) within the north region. The placement of LEG within either of the two putative groups was not well supported. However, LEG was not differentiated from sampling locations JUA, MAZ and PRI within HEC (Table S2). Taken together, these results clearly indicate hierarchical structure dominated by the split between the north and south regions, and local population substructure within each region. Although the exact number of genetically distinct populations must be interpreted cautiously, our results suggest the existence of at least seven subpopulations among the 24 sampling sites analysed: 1 – (QCS: TBL, MAL, HOP SHE); 2 – (SOG: RBY, SGI, JER, OGD, CRA, LAS); 3 – QUA; 4 – TOF; 5 – REN; 6 – (AK: AK1, AK2, AK3); 7 – (HEC: SEL, PRI, AK4, TOL, MAZ, CAL, JUA, LEG).

3.4 | Population assignment

The THL analysis performed in *assigner* showed that accurate inferences about population assignment (~80% assignment success) could be made with 100 SNPs (Figure S4). Using the full set of SNPs, individual sea cucumbers were assigned to the regional population (i.e., north/south) from which they were sampled with high success: 88% and 90% of individuals sampled from the north and south groups, respectively, were correctly assigned to their population of origin.

3.5 | Population Graph

The topology of the Population Graph (G) was composed of 24 nodes connected by 69 edges with significant conditional genetic covariance. The number of edges within the minimal edge set was significantly less than expected in a saturated graph ($p < .0001$), indicating the presence of significant population genetic structure. Two subgraphs were apparent within G, which represented north and south
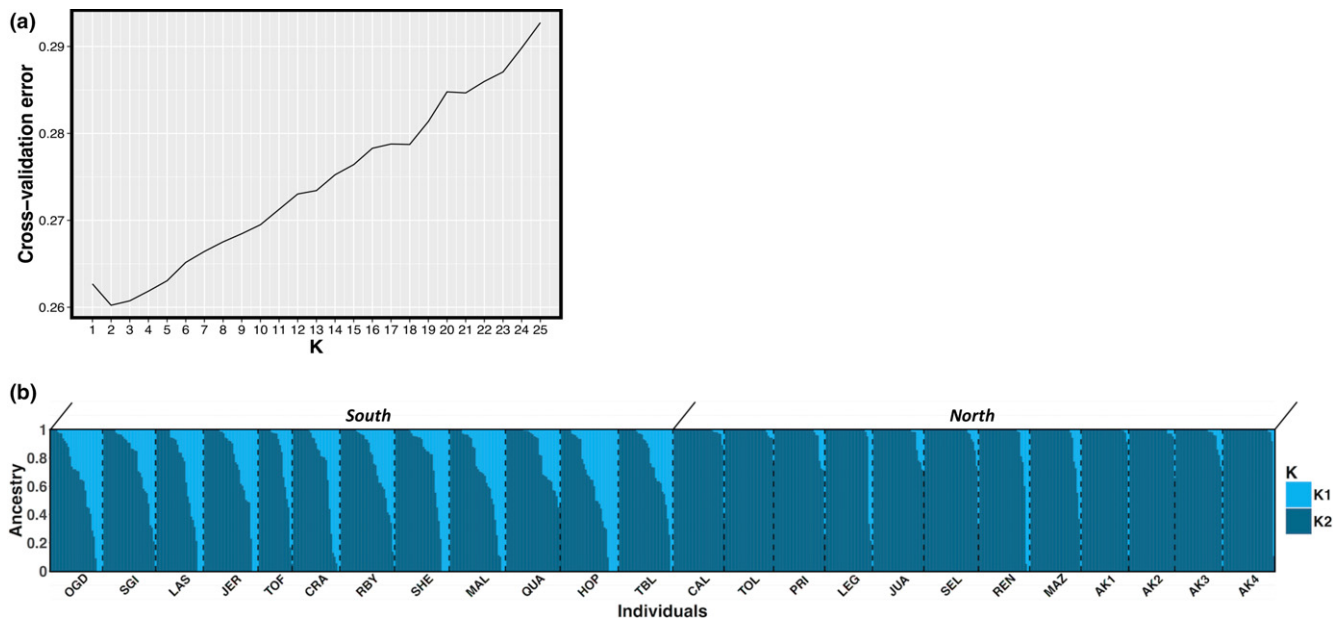
FIGURE 2 (a) Plot of cross-validation error for increasing *K* values from 2 to 25. The lowest cross-validation error indicates the optimal number of genetic clusters; (b) proportion of individual ancestry to each of *K* = 2 genetic clusters [Colour figure can be viewed at wileyonlinelibrary.com]
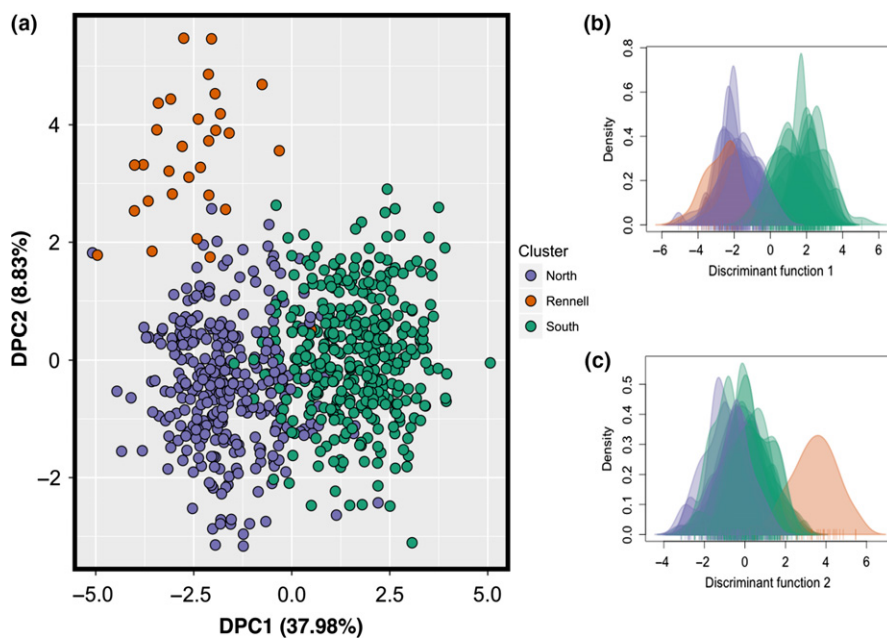
FIGURE 3 (a) Scatterplot showing the genetic clusters identified by DAPC with a priori sampling location information and density plots for (b) the first and (c) the second discriminant axes [Colour figure can be viewed at wileyonlinelibrary.com]

regional populations (Figure 5). These two subgraphs (G_{NORTH} and G_{SOUTH}) are connected to each other by edges through two nodes (REN and CAL). CAL exhibited the highest betweenness centrality (Table 1), suggesting that this site is important for maintaining genetic connectivity between the north and south regional populations.

3.6 | Isolation by distance and isolation by resistance

A significant correlation between pairwise F_{ST} and Euclidean geographic distance (Mantel *r* = .67, *p* = .0001) and in-water geographic

distance (Mantel *r* = .41, *p* = .001) was detected by the Mantel test, indicating a significant effect of isolation by distance. We also identified a significant pattern of isolation by resistance based on distances derived from connectivity probabilities, although the correlation was lower than that of the relationship between F_{ST} and geographic distance (*r* = .29, *p* = .001). Within the north and south regional groups, we did not detect significant IBD (in-water – North: Mantel *r* = .02, *p* = .90; South: *r* = .25, *p* = .08) or IBR (North: *r* = .06, *p* = .69; South: *r* = .19, *p* = .10).

For the SEA, we retained the first seven axes from the PCA of Hellinger-transformed allele frequencies, which cumulatively
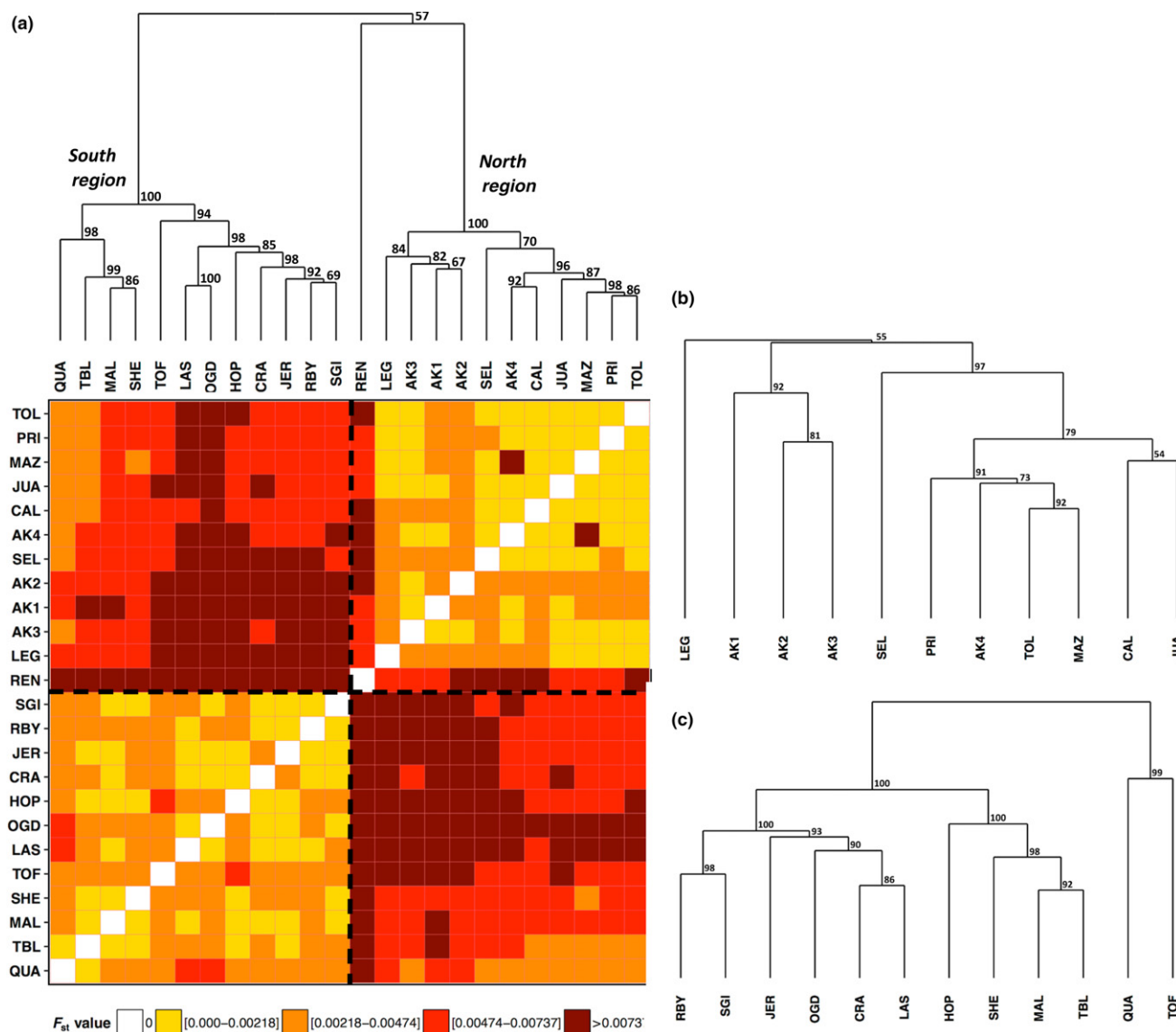
FIGURE 4 (a) Dendrogram and heat map of pairwise genetic differentiation (F_{ST}) between sampling locations showing a split between the north and south regions. Dendrograms for the north (b) and south (c) show hierarchical clustering of sites within regional groups. Numbers at dendrogram nodes indicate approximately unbiased probabilities based on multiscale bootstrap resampling [Colour figure can be viewed at wileyonlinelibrary.com]

explained 47% of the total neutral genetic variation. Results of the RDAs on seven PC axes using dbMEM or AEM eigenfunctions yielded considerable differences in their modelling of the spatial distribution of neutral genetic variation among sites. Forward selection of the Euclidean dbMEM variables identified two significant predictors (dbMEM1 and dbMEM2) representing broad-scale spatial structure. Two significant in-water dbMEM spatial predictors were also identified (dbMEM2 and dbMEM19) (Table 2). RDAs with only the selected dbMEM variables (both Euclidean and in-water) were significant ($p = .001$) with a coefficient of determination (adjusted R^2) of 17.9% and 22.2%, respectively. When AEM variables were used, six significant predictors that modelled directional spatial processes at a range of scales were selected (Table 2). Together, these AEM eigenfunctions explained 48.5% of the variation in the PC axes (adjusted

R^2; $p = .001$). The first three RDA axes were significant and explained 25%, 14% and 10% of the genetic variation summarized in the PC axes, respectively (Figure 6). The most significant AEM vectors modelled the separation between the north and south regional groups at the broadest spatial scale (AEM1; Figure 7a), as well as the separation of REN (AEM8; Figure 7b) and TOF (AEM11; Figure 7c) at finer spatial scales.

We combined the four significant dbMEM and six significant AEM eigenfunctions into a single model, which was globally significant ($R^2 = .495$, $p = .001$), but the dbMEM variables did not significantly contribute to the model ($p > .05$). When partitioning out the variation explained by the dbMEM variables, the partial model was significant ($p = .003$) and the fraction of variation explained by AEM variables alone was 0.24. The in-water dbMEM variables alone explained only
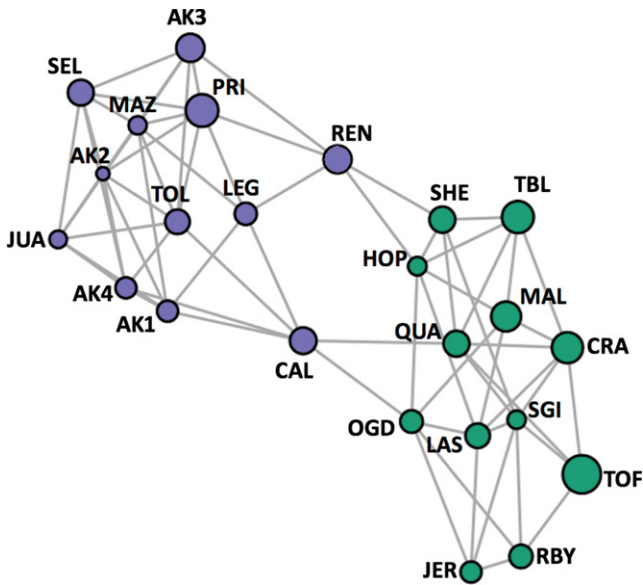
FIGURE 5 Population Graph showing the nodes (i.e., sampling locations) as filled circles, and edges between nodes representing significant genetic covariance between sites. The size of the nodes represents differences in genetic variance within sampling locations. Sampling locations from the north and south regions are shown in purple (subgraph G_{NORTH}) and green (subgraph G_{SOUTH}), respectively [Colour figure can be viewed at wileyonlinelibrary.com]

2% of the variation. In contrast, the fraction of the variation explained by Euclidean dbMEM variables was negative (adjusted $R^2 = -.013$) and thus explained less variation than would random variables (Legendre, 2008). The proportion of variance that could not be partitioned between the three types of spatial eigenfunctions because of collinearity (i.e., the joint effect) was 15%.

4 | DISCUSSION

In this study, we aimed to assess the population genetic structure in a highly dispersive marine invertebrate and evaluate the oceanographic drivers of spatial patterns of genetic variation. We focused on the giant California sea cucumber (*Parastichopus californicus*), which has the potential for long-distance dispersal due to its long-lived pelagic larval stage. As a result, populations are expected to exhibit very low differentiation. The results presented here provided evidence of population genetic structure, despite the potential for widespread gene flow. In particular, we demonstrated hierarchical population structure dominated by broad-scale differentiation between northern and southern regional groups, with significant, albeit weaker, local substructure, delineating at least seven subpopulations among sites within both regions. While we did find a stronger correlation between geographic distance and genetic distance among sampled sites in BC
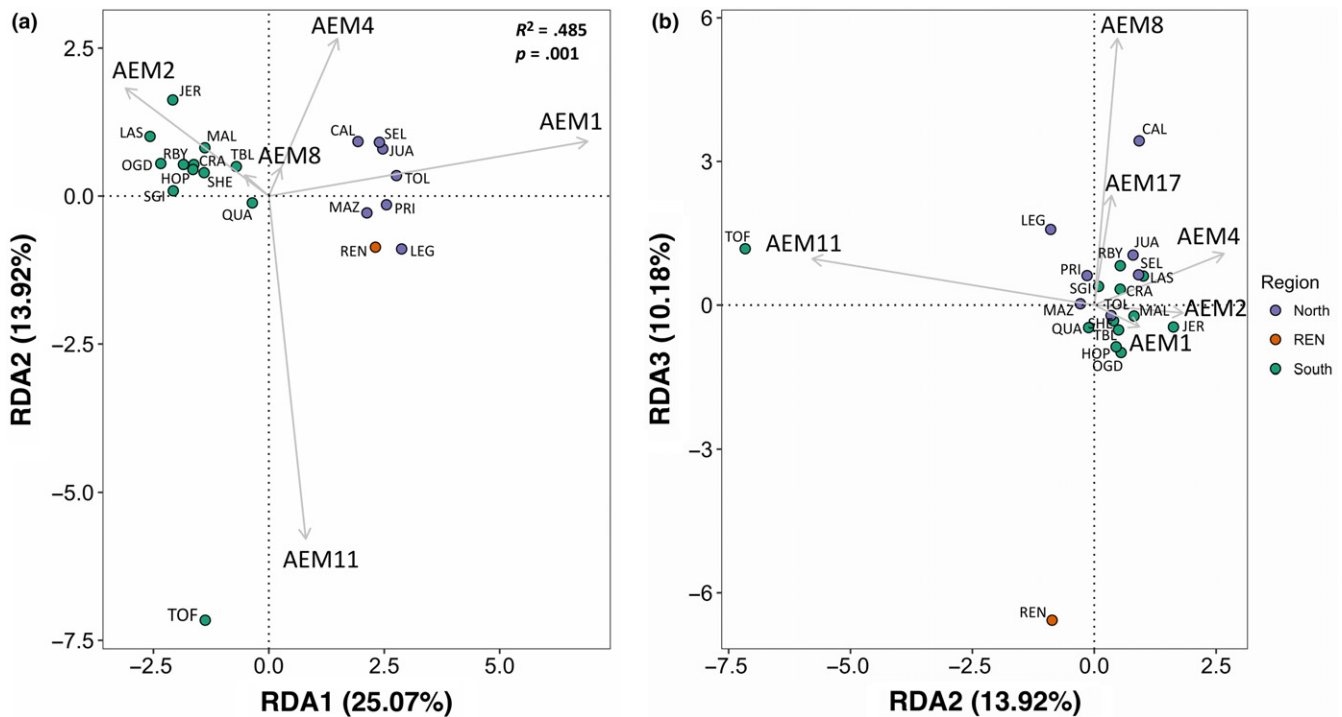
FIGURE 6 RDA biplot showing the significant AEM predictor variables (arrows) on (a) RDA axes 1 and 2, and (b) RDA axes 1 and 3. The dots represent sampling locations and are coloured based on inferred genetic clusters identified in previous analyses of population structure [Colour figure can be viewed at wileyonlinelibrary.com]
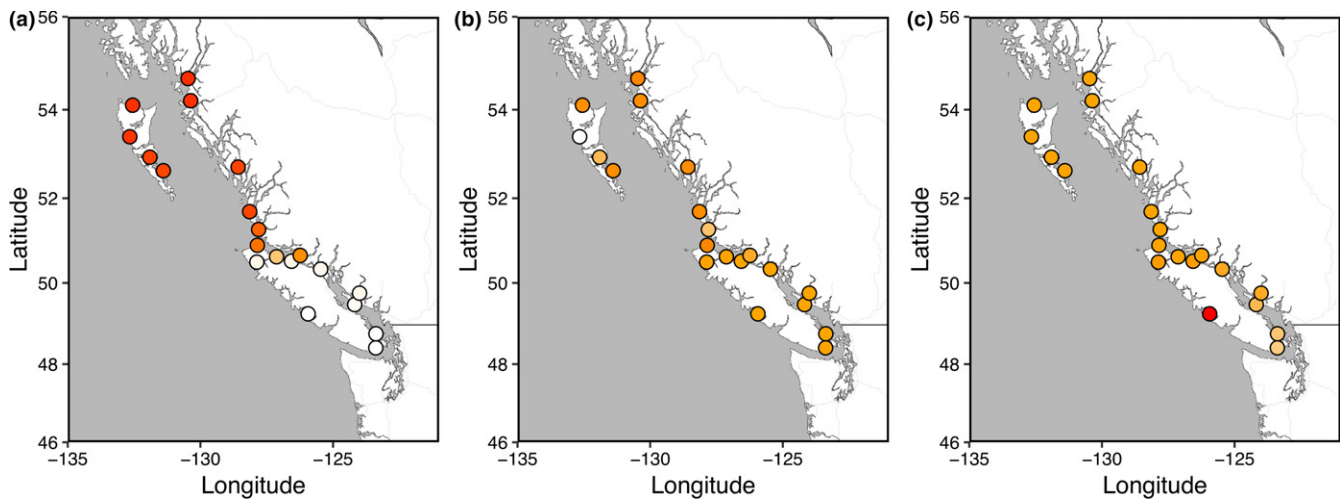
FIGURE 7 Visualization of the asymmetric eigenvector maps (a) AEM1, (b) AEM8 and (c) AEM11 at each sampling location. Similar colours represent similar AEM values [Colour figure can be viewed at wileyonlinelibrary.com]

(i.e., IBD) compared to the strength of IBR using a Mantel test, spatial predictors that accounted for asymmetry and variable strength in connectivity between sites had more than twice the explanatory power ($R^2 = .485$) compared to geography alone ($R^2 = .179$ [Euclidean] and .222 [in-water]), implying a significant effect of isolation by resistance in this region. Moreover, the spatial eigenfunction analysis revealed finer-scale spatial genetic substructure that could not be explained by geographic distance.

4.1 | Broad-scale population structure and connectivity

At the broadest spatial scale, the presence of two distinct genetic groups was corroborated by multiple analyses including estimates of individual ancestry (ADMIXTURE), discriminant analysis of principal components (DAPC) and hierarchical clustering based on pairwise genetic differentiation (F_{ST}). The location of the genetic break was also consistent across analyses, delineating a split between regions north and south of Queen Charlotte Sound. Similar patterns of regional population genetic structure have previously been reported by other studies in this region, highlighting an important genetic boundary for multiple marine species (Rocha-Olivares & Vetter, 1999; Sunday, Popovic, Palen, Foreman, & Hart, 2014). In one of the earliest studies to detect genetic discontinuities along the BC continental shelf, Rocha-Olivares and Vetter (1999) identified two genetic groups in the rosethorn rockfish (*Sebastes helvomaculatus*) with a subdivision occurring between Haida Gwaii and Vancouver Island. The complex glaciation history of the BC and southeastern Alaskan shoreline likely played an important role in the evolution of population genetic structure for coastal species in this region. Using coalescent-based approaches, McGovern, Keever, Saski, Hart, and Marko (2010) demonstrated that an observed split between Vancouver Island and Haida Gwaii in another marine echinoderm, the sea star

Pisaster miniata, was attributed to historical vicariance caused by Pleistocene glaciation events. Moreover, Sunday et al. (2014) employed an integrative approach using biophysical models and simulations of gene flow to show that physical oceanographic forces maintain this divergence by limiting contemporary gene flow and promoting high levels of self-recruitment. The striking similarity in the geographic distribution of genetic variation in *P. miniata* and *P. californicus* suggests that similar forces may be driving the evolution of population genetic structure in multiple species. Strong directional surface currents dominate the northeastern Pacific coastal region and conceivably facilitate the advection of dispersing larvae. Notably, the northward-flowing Alaska Current and the southward-flowing California current diverge from the offshore North Pacific Current (NPC) at the location of the observed genetic break (Masson & Fine, 2012; Morrison et al., 2014) (Figure 1), potentially carrying dispersing larvae in opposite directions and separating populations across the bifurcation zone.

Our findings also highlight two important aspects of connectivity between the two regions. First, 89% of individuals were correctly assigned to the region from which they were sampled, indicating restricted connectivity between regions (Saenz-Agudelo, Jones, Thorrold, & Planes, 2009; Waples & Gaggiotti, 2008). Although demographic independence of regional groups must be interpreted cautiously (Lowe & Allendorf, 2010), the observation that population assignment could be made with relatively few markers (~100 SNPs) implies that gene flow is indeed limited. Second, our evaluation of genetic connectivity based on Population Graphs of genetic covariance characterized two subgraphs corresponding to the same genetic groups. Relatively fewer edges between sites in either group compared to the number of edges between sites within groups indicate vicariance (Dyer & Nason, 2004). Furthermore, Calvert Island (CAL) displayed the highest betweenness centrality compared to all other sampling locations,

indicating that this particular site is important for maintaining genetic connectivity throughout the sampled network. Betweenness centrality can highlight critical dispersal pathways and identify specific locations that may serve as stepping stones for connecting distant populations (Treml et al., 2008; Urban et al., 2009). The importance of this region as a connectivity node was also demonstrated by Sunday et al. (2014) based on simulated dispersal patterns of *P. miniata*, suggesting that the southern central coast may represent a key intermediate site for connectivity across multiple species.

4.2 | Within-region population substructure

An increasing number of studies provide evidence of finer-scale spatial patterns of genetic structure in marine populations that do not match expectations of high gene flow (e.g., American lobster: Benestan et al., 2015; bat star: Sunday et al., 2014; reef-building corals: Truelove et al., 2017; spiny lobsters: Iacchei et al., 2013). Our study substantiates these findings, suggesting that along-shore connectivity between populations of *P. californicus* may be restricted within regions despite the long PLD and may warrant further investigation into the extent to which observed genetic structure represents the relevant scales of demographic connectivity in the region.

Previous work on *P. californicus* described weaker genetic differentiation between sampled sites within Hecate Strait (north region) compared to relatively higher genetic divergence between the west coast of Vancouver Island and other sampled locations (south region), likely owing to the more complex oceanographic environment within Hecate Strait (Nelson, 2003). The results presented here support this assertion of differential levels of genetic differentiation within each region, as shown by the greater number of significant pairwise comparisons of F_{ST} between sites in the south group compared to the north group, and the more well-supported subpopulations in the south region delineated by hierarchical clustering. For example, in the south, two distinct genetic clusters separated sites within the Strait of Georgia from sites within the Queen Charlotte Strait. These two large straits between Vancouver Island and mainland BC are separated by a narrow passage (the Johnstone Strait) that may play a role in limiting movement between both regions. In the north region, distinct population substructure was less clearly defined, potentially owing to the relative lack of physical boundaries north of Vancouver Island and the largely open water of the Hecate Strait. Although Legace Bay (LEG) was not clearly placed within any genetic cluster, it could not be differentiated from neighbouring sites based on pairwise F_{ST} values. However, we did observe significant genetic differentiation between the three outermost Alaska sites (AK1 − 3) and the other sites in the north region.

A distinct genetic cluster composed of individuals from Rennell Sound (REN) on the west coast of Haida Gwaii that was well separated from the other sites in the northern regional group was detected by DAPC, but not by the analysis of individual ancestry (ADMIXTURE). This outcome is in agreement with previous studies showing that DAPC is more efficient than other clustering methods at detecting genetic subdivisions when overall genetic differentiation is weak (Benestan et al., 2015; Hotaling et al., 2017; Jombart et al., 2010). Using hierarchical clustering of pairwise F_{ST}, we showed that REN was significantly differentiated from other sampling locations, thus further corroborating the DAPC results, as were the two sites sampled on the west coast of Vancouver Island (TOF and QUA). Significant F_{ST} was also observed between TOF and QUA, suggesting that they may represent isolated subpopulations. Interannual variability in the break point of the NPC may play a role in the genetic divergence between QUA and TOF. The NPC is known to split roughly between 45° and 50°N latitude, but the exact location of where it breaks into the northward Alaska Current and southward California Current varies from year to year (Morrison et al., 2014; Ware & McFarlane, 1989), such that if the NPC bifurcates along the west coast of Vancouver Island, connectivity between QUA and TOF may be restricted as larvae are carried in opposite directions. Like REN, these sites occurring on the west coast of large landmasses (i.e., Haida Gwaii and Vancouver Island) are exposed to retentive hydrodynamic forces that may entrain larvae within embayments and contribute to genetic isolation (see below).

4.3 | Isolation by distance and isolation by resistance

Testing for IBD using a Mantel test indicated a significant effect of geographic distance on genetic differentiation (F_{ST}) between sampling locations in coastal BC (Euclidean: Mantel $r = .67$; in-water: Mantel $r = .41$), indicating that IBD is an important spatial driver of observed patterns of genetic structure. Significant IBD was not detected within the north and south regional groups, indicating that IBD only plays an important role at a broad spatial scale in this system. This result suggests that this effect may be driven by hierarchical clustering between the north and south regional groups (Meirmans, 2012). Mantel tests also revealed an important effect of isolation by resistance, albeit a lower correlation (Mantel $r = .29$), only when all pairwise comparisons were analysed, suggesting that local oceanography also plays an important role in shaping spatial genetic structure at the scale of the entire BC coast, but not within regional groups.

The advantage of using a spatial eigenfunction approach is that it allows the direct comparison of the relative contribution of both symmetric (i.e., geography) and asymmetric processes (i.e., current flow) on spatial patterns of genetic variation across multiple spatial scales (Dray et al., 2012). Additionally, unlike Mantel correlations, the spatial eigenfunction framework is more powerful at detecting fine-scale spatial structures (Diniz-Filho et al., 2013). The results of this study suggest that spatial variables that account for the directional movement of current flow (i.e., AEMs) are better predictors of spatial genetic structure than variables derived from geographic distances (i.e., dbMEMs) between sites (Table 2). Several studies have shown that spatial patterns of genetic structure among marine populations often do not conform to predictions based on the classical stepping-stone model of isolation by distance (D'Aloia, Bogdanowicz, Harrison, & Buston, 2014; Selkoe et al., 2010; Truelove et al., 2017). Comparisons of dispersal estimates from biophysical models with

observed patterns of genetic structure have demonstrated that isolation by resistance is indeed a better predictor of population genetic structure for many marine species, especially in regions with complex hydrodynamic environments (Sunday et al., 2014; Thomas et al., 2015; Truelove et al., 2017). For example, distances derived from biophysical connectivity explained patterns of genetic structure in the spiny lobster (*Panulirus argus*) better than an isolation by geographic distance model, likely due to physical boundaries created by oceanographic features (Truelove et al., 2017). Thomas et al. (2015) also showed that oceanographic resistance was significantly correlated with genetic differentiation in broadcast spawning corals (*Acropora spicifera*), while geographic distance was not. While our Mantel tests did not show a stronger effect of oceanographic resistance compared to geographic distance at a broad spatial scale, our spatial eigenfunction analysis demonstrated that accounting for directional processes can improve inferences of genetic connectivity among marine populations and the effects of ocean circulation at local scales (Benestan et al., 2015; Schunter et al., 2011).

Our analysis detected a significant contribution of AEM predictors at a range of spatial scales, whereas dbMEMs were predominantly influential at broad scales (dbMEM1 and dbMEM2). Broadscale spatial predictors were important in distinguishing between the north and south regional groups identified in previous analyses of population structure. When we partitioned the variance between both types of spatial eigenvectors, a substantial fraction of the total genetic variation was explained by the joint (i.e., collinear) contribution of AEMs and dbMEMs (R^2 = .15) and the dbMEM variables were not significant predictors in the partial model. These results corroborate the significant effects of both IBD and IBR at the scale of the entire coastal region detected by Mantel tests. The movement of surface currents in opposite directions potentially leads to an observed pattern of genetic structure that resembles IBD (and thus can also be explained by geographic distance).

In contrast, AEM variables representing finer scale asymmetric processes remained significant predictors of genetic variation after controlling for the effect of geography highlighting an important effect of directional connectivity on spatial genetic structure in *P. californicus* populations that cannot be explained by geographic distance. Fine-scale AEM predictors were important drivers of the separation between Tofino (TOF) from other sites in the southern region and Rennell Sound (REN) from other sites in the northern region, in agreement with the observation that REN and TOF form unique genetic clusters based on DAPC and F_{ST} results. Examination of the connectivity matrix (Table S3) reveals important patterns of larval export and recruitment from other sites that lend additional support to the hypothesis of IBR in these regions. Tofino exhibits considerably higher recruitment (0.013) compared to export (0.004), highlighting clear asymmetry in particle transport. Moreover, the sites that contribute a substantial proportion of particles to Tofino are located in the Strait of Georgia, suggesting that larvae are transported around the southern tip of Vancouver Island and travel northward along the western coastline. In the northern region, self-recruitment in Rennell Sound is higher than recruitment from any

other site, although a large proportion of particles that were released from sites on the eastern shore of Haida Gwaii also settle in Rennell Sound, indicating the transport of larvae out of the Hecate Strait and northward along the west side of Haida Gwaii. Seasonal hydrodynamic forces are known to occur on the western coasts of both landmasses and potentially contribute to the directional and/or restricted genetic connectivity between sites in these regions, including a northerly near-shore surface current (the Vancouver Island Coastal Current) occurring close to the west coast of Vancouver Island during the summer months (Masson & Cummins, 2000), and retentive mesoscale eddies that form around Haida Gwaii and at the southwestern tip of Vancouver Island (Di Lorenzo, Foreman, & Crawford, 2005; Morrison et al., 2014). Although high retention at these locales is expected in the context of local current dynamics, additional sampling on the exposed shorelines of the major islands is needed to further explore the population genetic structure in this region. Moreover, assessing temporal change in genetic integrity will also be required to draw more accurate demographic inferences.

Other factors could further contribute to explaining the observed spatial patterns of genetic variation. In particular, larval behaviours such as active swimming by muscular larvae, natal homing and settlement cues have been implicated as drivers of enhanced local retention and restricted connectivity in several marine fish species (Gerlach, Atema, Kingsford, Black, & Miller-Sims, 2007; Montgomery, Jeffs, Simpson, Meekan, & Tindle, 2006; Rocha-Olivares & Vetter, 1999; Thorrold, Latkoczy, Swart, & Jones, 2001). Truelove et al. (2017) also demonstrated that ontogenetic vertical migration coupled with retentive physical hydrodynamic forces increases local retention and drives genetic differentiation between populations of *P. argus* in the Caribbean. Little is known with respect to larval behaviour of *P. californicus*; so for the purpose of our study, we assumed that larvae were neutrally buoyant, passively dispersing larvae. This assumption holds for other species with small ciliated larvae including some echinoderm species, for which surface current flow can provide an accurate model of dispersal (Pedrotti & Fenaux, 1992). Also, we estimated connection probabilities between sites based on the predicted larval settlement patterns after 120 days, which is the upper estimate of the pelagic larval duration period for *P. californicus* in this region (Cameron & Fankboner, 1989). It is possible that some proportion of larvae would metamorphose and settle earlier than 120 days, potentially reducing overall dispersal distances. The biophysical model that we used did not incorporate larval mortality, which could have an effect of overestimating larval exchange (Pineda, Hare, & Sponaugle, 2007). Nonetheless, even in the case of maximum dispersal potential, our results clearly refute the hypothesis of a panmictic population of *P. californicus*. They also show that existing genetic discontinuities can be explained, at least in part, by the oceanographic environment of the region.

4.4 | Implications for management and future directions

Our findings have important implications for the spatial management of *P. californicus* populations. This species is the only commercially

exploited sea cucumber on the west coast of Canada, and informed management practices are needed to ensure the sustainability of the fishery (Hajas, Hand, Duprey, Lochead, & Deault, 2011; Hand et al., 2008). Significant genetic differentiation and restricted connectivity observed between regional groups indicate that populations north and south of the Queen Charlotte Sound may warrant independent management (Funk, McKay, Hohenlohe, & Allendorf, 2012) and further investigation into the demographic independence of populations. Our results also support the presence of significant fine-scale population substructure within regions, primarily delineating unique genetic clusters on the west coasts of the major islands and separating sites within the Queen Charlotte Strait and the Strait of Georgia. This information could potentially inform fisheries management in the region as well as the optimization of shellfish transfer zone (STZ) boundaries for managing the translocation of cultured shellfish (e.g., Miller, Supernault, Li, & Withler, 2006). Furthermore, high population assignment success suggests that using genomic technologies such as RADseq on marine organisms may be a relevant step towards developing DNA markers to reliably identify the origin of harvested individuals (Bernatchez et al., 2017; Stockstad, 2010).

Understanding the spatial patterns of marine connectivity is also directly relevant to the establishment of marine protected area (MPA) networks (Almany et al., 2009; Kininmonth et al., 2011). Characterizing directional gene flow that maintains genetic connectivity provides insight into the location of valuable sources of larval export (Beger et al., 2014). Furthermore, the high betweenness centrality of Calvert Island (CAL) in the Population Graph highlights an important dispersal pathway in the central coast region of BC for linking the north and south regional groups that has previously been identified as an important connectivity node for other echinoderm species (Sunday et al., 2014). Given the congruent evidence across studies, we recommend that. The central coast region warrants protection as a key stepping stone for dispersal that facilitates connectivity throughout a larger MPA network (Kininmonth et al., 2011; Watson et al., 2011), not just for *P. californicus*, but also for other species that share a similar distribution and life history.

In our study, we were unable to extend analysis of the effect of connectivity mediated by ocean currents into the Alaska region due to the limited extent of the biophysical model. Incorporating biophysical models that extend beyond the Canadian Pacific coastal region would elucidate spatial patterns of gene flow across a much broader geographic region, and additional sampling across the entire range of *P. californicus* would provide guidance for both local and international management of sea cucumber fisheries. Finally, understanding the role of the oceanographic environment on marine connectivity allows predictions to be made regarding how shifts in ocean circulation dynamics as a result of climate change may affect marine connectivity in the future (Watson, Kendall, Siegel, & Mitarai, 2012). Thus, understanding how marine connectivity may change in the future can help inform the design of MPA networks with appropriate size and spacing between protected sites to ensure that they will withstand changes to seascape conditions in the face of climate change (Gerber, Del Mark Mancha-Cisneros, O'Connor, & Selig, 2014).

DATA ACCESSIBILITY

Raw demultiplexed sequences are available on NCBI SRA (BioProject Accession # PRJNA436919). Filtered data sets and R scripts for the analyses are available from the Dryad Digital Repository: https://doi.org/10.5061/dryad.db6177b

AUTHOR CONTRIBUTION

The study was designed by A.X., M.J.F., J.M.R.C., and L.Ber. The bioinformatics pipeline was developed by E.N. R.M.D. provided the biophysical model. The analysis was performed by A.X. with contributions from L.Ben. The manuscript was written by A.X., and all authors contributed to editing and revising the manuscript.

ORCID

Amanda Xuereb (iD) http://orcid.org/0000-0002-3975-2299

REFERENCES

Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, *12*, 246. https://doi.org/10.1186/1471-2105-12-246

Alexander, D., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, *19*, 1655–1664. https://doi.org/10.1101/gr.094052.109

Almany, G. R., Connolly, S. R., Heath, D. D., Hogan, J. D., Jones, G. P., McCook, L. J., ... Williamson, D. H. (2009). Connectivity, biodiversity conservation and the design of marine reserve networks for coral reefs. *Coral Reefs*, *28*, 339–351. https://doi.org/10.1007/s00338-009-0484-x

Anderson, E. C. (2010). Assessing the power of informative subsets of loci for population assignment: Standard methods are upwardly

biased. *Molecular Ecology Resources*, *10*, 701–710. https://doi.org/10.1111/j.1755-0998.2010.02846.x

Banks, S. C., Piggott, M. P., Williamson, J. E., Bové, U., Holbrook, N. J., & Beheregaray, L. B. (2007). Oceanic variability and coastal topography shape genetic structure in a long-dispersing sea urchin. *Ecology*, *88*, 3055–3064. https://doi.org/10.1890/07-0091.1

Beaumont, M. A., & Nichols, R. A. (1996). Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B: Biological Sciences*, *263*, 1619–1626. https://doi.org/10.1098/rspb.1996.0237

Beger, M., Selkoe, K. A., Treml, E., Barber, P. H., Crandall, E. D., Toonen, R. J., & Riginos, C. (2014). Evolving coral reef conservation with genetic information. *Bulletin of Marine Science*, *90*, 159–185. https://doi.org/10.5343/bms.2012.1106

Benestan, L. M., Ferchaud, A. L., Hohenlohe, P. A., Garner, B. A., Naylor, G. J. P., Baums, I. B., ... Luikart, G. (2016). Conservation genomics of natural and managed populations: Building a conceptual and practical framework. *Molecular Ecology*, *25*, 2967–2977. https://doi.org/10.1111/mec.13647

Benestan, L., Gosselin, T., Perrier, C., Sainte-Marie, B., Rochette, R., & Bernatchez, L. (2015). RAD genotyping reveals fine-scale genetic structuring and provides powerful population assignment in a widely distributed marine species, the American lobster (*Homarus americanus*). *Molecular Ecology*, *24*, 3299–3315. https://doi.org/10.1111/mec.13245

Benestan, L., Quinn, B., Larporte, M., Maaroufi, H., Rochette, R., & Bernatchez, L. (2016). Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology*, *25*, 5073–5092. https://doi.org/10.1111/mec.13811

Benjamini, Y., & Hochberg, Y. (1994). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, *1*, 289–300.

Bernatchez, L., Wellenreuther, M., Araneda, C., Ashton, D. T., Barth, J. M. I., Beacham, T. D., ... Withler, R. E. (2017). Harnessing the power of genomics to secure the future of seafood. *Trends in Ecology and Evolution*, *32*, 665–680. https://doi.org/10.1016/j.tree.2017.06.010

Blanchet, F. G., Legendre, P., Maranger, R., Monti, D., & Pepin, P. (2011). Modelling the effect of directional spatial ecological processes at different scales. *Oecologia*, *166*, 357–368. https://doi.org/10.1007/s00442-010-1867-y

Cameron, J. L., & Fankboner, P. V. (1986). Reproductive biology of the commercial sea cucumber *Parastichopus californicus* (Stimpson) (Echinodermata: Holothuroidea). I. Reproductive periodicity and spawning behavior. *Canadian Journal of Zoology*, *64*, 168–175. https://doi.org/10.1139/z86-027

Cameron, J. L., & Fankboner, P. V. (1989). Reproductive biology of the commercial sea cucumber *Parastichopus californicus* (Stimpson) (Echinodermata: Holothuroidea). II. Observations on the ecology of development, recruitment, and the juvenile life stage. *Journal of Experimental Marine Biology and Ecology*, *127*, 43–67. https://doi.org/10.1016/0022-0981(89)90208-6

Cameron, R. A., Samanta, M., Yuan, A., He, D., & Davidson, E. (2009). SpBase: The sea urchin genome database and web site. *Nucleic Acids Research*, *37*, D750–D754. https://doi.org/10.1093/nar/gkn887

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A., & Cresko, W. A. (2013). Stacks: An analysis tool set for population genomics. *Molecular Ecology*, *22*, 3124–3140. https://doi.org/10.1111/mec.12354

Cowen, R. K., Gawarkiewicz, G., Pineda, J., Thorrold, S. R., & Werner, F. E. (2007). Population connectivity in marine systems. *Oceanography*, *20*, 14–21. https://doi.org/10.5670/oceanog

Cowen, R. K., Paris, C. B., & Srinivasan, A. (2006). Scaling of connectivity in marine populations. *Science*, *311*, 522–527. https://doi.org/10.1126/science.1122039

Cowen, R. K., & Sponaugle, S. (2009). Larval dispersal and marine population connectivity. *Annual Review of Marine Science*, *1*, 443–466. https://doi.org/10.1146/annurev.marine.010908.163757

Cros, A., Toonen, R. J., Donahue, M. J., & Karl, S. A. (2017). Connecting Palau's marine protected areas: A population genetic approach to conservation. *Coral Reefs*, *36*, 735–748. https://doi.org/10.1007/s00338-017-1565-x

Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *InterJournal, Complex Systems*, *1695*, 1–9.

Daigle, R. (2016). cuke_biophysical_particle_tracking: publication version. Zenodo. 10.5281/zenodo.51411. Retrieved from https://github.com/remi-daigle/cuke_biophysical_particle_tracking

D'Aloia, C. C., Bogdanowicz, S. M., Francis, R. K., Majoris, J. E., Harrison, R. G., & Buston, P. M. (2015). Patterns, causes, and consequences of marine larval dispersal. *Proceedings of the National Academy of Sciences*, *112*, 13940–13945. https://doi.org/10.1073/pnas.1513754112

D'Aloia, C. C., Bogdanowicz, S. M., Harrison, R. G., & Buston, P. M. (2014). Seascape continuity plays an important role in determining patterns of spatial genetic structure in a coral reef fish. *Molecular Ecology*, *23*, 2902–2913. https://doi.org/10.1111/mec.12782

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., ... 1000 Genomes Project Analysis Group. (2011). The variant call format and VCFtools. *Bioinformatics*, *27*, 2156–2158. https://doi.org/10.1093/bioinformatics/btr330

DFO. (2016). Integrated Fisheries Management Plan Summary. Sea Cucumber (Parastichopus californicus) By Dive. Pacific Region 2016/2017.

Di Lorenzo, E., Foreman, M. G. G., & Crawford, W. R. (2005). Modelling the generation of Haida Eddies. *Deep Sea Research*, *52*, 853–873. https://doi.org/10.1016/j.dsr2.2005.02.007

Diniz-Filho, J. A., Diniz, J. V., Rangel, T. F., Soares, T. N., de Campos Telles, M. P., Collevatti, R. G., & Bini, L. M. (2013). A new eigenfunction spatial analysis describing population genetic structure. *Genetica*, *141*, 479–489. https://doi.org/10.1007/s10709-013-9747-0

Dray, S., Blanchet, G., Borcard, D., Clappe, S., Guenard, G., Jombart, T., ... Wagner, H. H. (2017). adespatial: Multivariate Multiscale Spatial Analysis. R package version 0.0-9.

Dray, S., Legendre, P., & Peres-Neto, P. R. (2006). Spatial modelling: A comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, *196*, 483–493. https://doi.org/10.1016/j.ecolmodel.2006.02.015

Dray, S., Pélissier, R., Couteron, P., Fortin, M. J., Legendre, P., Peres-Neto, P. R., ... Wagner, H. H. (2012). Community ecology in the age of multivariate spatial analysis. *Ecological Monographs*, *82*, 257–275. https://doi.org/10.1890/11-1183.1

Dyer, R. J. (2016). gstudio: Tools Related to the Spatial Analysis of Genetic Marker Data. R package version 1.5.0.

Dyer, R. J. (2017). popgraph: This is an R package that constructs and manipulates population graphs. R package version 1.5.0.

Dyer, R. J., & Nason, J. D. (2004). Population Graphs: The graph theoretic shape of genetic structure. *Molecular Ecology*, *13*, 1713–1727. https://doi.org/10.1111/j.1365-294X.2004.02177.x

Dyer, R. J., Nason, J. D., & Garrick, R. C. (2010). Landscape modelling of gene flow: Improved power using conditional genetic distance derived from the topology of population networks. *Molecular Ecology*, *19*, 3746–3759. https://doi.org/10.1111/j.1365-294X.2010.04748.x

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*, *132*, 479–491.

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. https://doi.org/10.1534/genetics.108.092221

Foreman, M. G. G., Callendar, W., Masson, D., Morrison, J., & Fine, I. (2014). A model simulation of future oceanic conditions along the British Columbia Continental shelf. Part II: Results and analyses. *Atmosphere-Ocean*, *52*, 20–38. https://doi.org/10.1080/07055900.2013.873014

Foster, N. L., Paris, C. B., Kool, J. T., Baums, I. B., Stevens, J. R., Sanchez, J. A., ... Mumby, P. J. (2012). Connectivity of Caribbean coral populations: Complementary insights from empirical and modelled gene flow. *Molecular Ecology*, *21*, 1143–1157. https://doi.org/10.1111/j.1365-294X.2012.05455.x

Funk, W. C., McKay, J. K., Hohenlohe, P. A., & Allendorf, F. W. (2012). Harnessing genomics for delineating conservation units. *Trends in Ecology and Evolution*, *27*, 489–496. https://doi.org/10.1016/j.tree.2012.05.012

Gagnaire, P.-A., Broquet, T., Aurelle, D., Viard, F., Souissi, A., Bonhomme, F., ... Bierne, N. (2015). Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications*, *8*, 769–786. https://doi.org/10.1111/eva.12288

Gerber, L. R., Del Mark Mancha-Cisneros, M., O'Connor, M., & Selig, E. R. (2014). Climate change impacts on connectivity in the ocean: Implications for conservation. *Ecosphere*, *5*, 1–18.

Gerlach, G., Atema, J., Kingsford, M. J., Black, K. P., & Miller-Sims, V. (2007). Smelling home can prevent dispersal of reef fish larvae. *Proceedings of the National Academy of Sciences*, *104*, 858–863. https://doi.org/10.1073/pnas.0606777104

Gilg, M. R., & Hilbish, T. J. (2003). The geography of marine larval dispersal: Coupling genetics with fine-scale physical oceanography. *Ecology*, *84*, 2989–2998. https://doi.org/10.1890/02-0498

Goslee, S. C., & Urban, D. L. (2007). The ecodist package for dissimilarity-based analysis of ecological data. *Journal of Statistical Software*, *22*, 1–19.

Gosselin, T., Anderson, E. C., & Bradbury, I. (2016). assigner: Assignment Analysis with GBS/RAD Data using R. R package version 0.4.1.

Gosselin, T., & Bernatchez, L. (2016). stackr: GBS/RAD data exploration, manipulation and visualization using R. R package version 0.2.1.

Hajas, W., Hand, C., Duprey, N., Lochead, J., & Deault, J. (2011). Using production models with new and developing fisheries: A case study using the sea cucumber *Parastichopus californicus* in British Columbia, Canada. *Fisheries Research*, *110*, 421–434. https://doi.org/10.1016/j.fishres.2011.05.009

Hamel, J., & Mercier, A. (2008). Population status, fisheries and trade of sea cucumbers in temperate areas of the Northern Hemisphere. In V. Toral-Granda, A. Lovatelli & M. Vasconcellos (Eds.), *Sea cucumbers. A global review of fisheries and trade* (pp. 257–291). FAO Fisheries and Aquaculture Technical Paper. No. 516. Rome: FAO.

Hand, C. M., Hajas, W., Duprey, N., Lochead, J., Deault, J., & Caldwell, J. (2008). An evaluation of fishery and research data collected during the Phase 1 sea cucumber fishery in British Columbia, 1998–2007. Canadian Science Advisory Secretariat (CSAS) (Vol. 2008/065).

Hotaling, S., Muhlfeld, C. C., Giersch, J. J., Ali, O. A., Jordan, S., Miller, M. R., ... Weisrock, D. W. (2017). Demographic modelling reveals a history of divergence with gene flow for a glacially tied stonefly in a changing post-Pleistocene landscape. *Journal of Biogeography*, *45*, 1–14.

Iacchei, M., Ben-Horin, T., Selkoe, K. A., Bird, C. E., García Rodríguez, F., & Toonen, R. J. (2013). Combined analyses of kinship and F_{ST} suggest potential drivers of chaotic genetic patchiness in high gene-flow populations. *Molecular Ecology*, *22*, 3476–3494. https://doi.org/10.1111/mec.12341

Iacchei, M., Gaither, M. R., Bowen, B. W., & Toonen, R. J. (2016). Testing dispersal limits in the sea: Range-wide phylogeography of the pronghorn spiny lobster *Panulirus penicillatus*. *Journal of Biogeography*, *43*, 1032–1044. https://doi.org/10.1111/jbi.12689

Jombart, T. (2008). adegenet: A R package for the multivariate analysis of genetic markers. *Bioinformatics*, *24*, 1403–1405. https://doi.org/10.1093/bioinformatics/btn129

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*, 94. https://doi.org/10.1186/1471-2156-11-94

Jorde, P. E., Søvik, G., Westgaard, J. I., Albretsen, J., André, C., Hvingel, C., ... Jørstad, K. E. (2015). Genetically distinct populations of northern shrimp, *Pandalus borealis*, in the North Atlantic: Adaptation to different temperatures as an isolation factor. *Molecular Ecology*, *24*, 1742–1757. https://doi.org/10.1111/mec.13158

Kininmonth, S., Beger, M., Bode, M., Peterson, E., Adams, V. M., Dorfman, D., ... Possingham, H. P. (2011). Dispersal connectivity and reserve selection for marine conservation. *Ecological Modelling*, *222*, 1272–1282. https://doi.org/10.1016/j.ecolmodel.2011.01.012

Kinlan, B. P., Gaines, S. D., & Lester, S. E. (2005). Propagule dispersal and the scales of marine community process. *Diversity and Distributions*, *11*, 139–148. https://doi.org/10.1111/j.1366-9516.2005.00158.x

Kough, A. S., Paris, C. B., & Butler, M. J. IV (2013). Larval connectivity and the international management of fisheries. *PLoS ONE*, *8*, e64970. https://doi.org/10.1371/journal.pone.0064970

Lal, M. M., Southgate, P. C., Jerry, D. R., Bosserelle, C., & Zenger, K. R. (2017). Swept away: Ocean currents and seascape features influence genetic structure across the 18,000 km Indo-Pacific distribution of a marine invertebrate, the black-lip pearl oyster *Pinctada margaritifera*. *BMC Genomics*, *18*, 66. https://doi.org/10.1186/s12864-016-3410-y

Lambert, P. (1997). *Sea cucumbers of British Columbia. Southeast Alaska and Puget sound*. Vancouver, BC: UBC Press.

Larson, W. A., Seeb, L. W., Everett, M. V., Waples, R. K., Templin, W. D., & Seeb, J. E. (2014). Genotyping by sequencing resolves shallow population structure to inform conservation of Chinook salmon (*Oncorhynchus tshawytscha*). *Evolutionary Applications*, *7*, 355–369. https://doi.org/10.1111/eva.12128

Legendre, P. (2008). Studying beta diversity: Ecological variation partitioning by multiple regression and canonical analysis. *Journal of Plant Ecology*, *1*, 3–8. https://doi.org/10.1093/jpe/rtm001

Legendre, P., & Gallagher, E. D. (2001). Ecologically meaningful transformations for ordination of species data. *Oecologia*, *129*, 271–280. https://doi.org/10.1007/s004420100716

Lischer, H. E. L., & Excoffier, L. (2012). PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs. *Bioinformatics*, *28*, 298–299. https://doi.org/10.1093/bioinformatics/btr642

Lotterhos, K. E., & Whitlock, M. C. (2014). Evaluation of demographic history and neutral parameterization on the performance of F_{ST} outlier tests. *Molecular Ecology*, *23*, 2178–2192. https://doi.org/10.1111/mec.12725

Lowe, W. H., & Allendorf, F. W. (2010). What can genetics tell us about population connectivity? *Molecular Ecology*, *19*, 3038–3051. https://doi.org/10.1111/j.1365-294X.2010.04688.x

Martin, M. (2011). Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*, *17*, 10–12. https://doi.org/10.14806/ej.17.1.200

Masson, D., & Cummins, P. F. (2000). Fortnightly modulation of the estuarine circulation in the Juan de Fuca Strait. *Journal of Marine Resources*, *58*, 439–463. https://doi.org/10.1357/002224000321511106

Masson, D., & Fine, I. (2012). Modeling seasonal to interannual ocean variability of coastal British Columbia. *Journal of Geophysical Research*, *117*, 1–14.

McGovern, T. M., Keever, C. C., Saski, C. A., Hart, M. W., & Marko, P. B. (2010). Divergence genetics analysis reveals historical population genetic processes leading to contrasting phylogeographic patterns in co-distributed species. *Molecular Ecology*, *19*, 5043–5060. https://doi.org/10.1111/j.1365-294X.2010.04854.x

Meirmans, P. G. (2012). The trouble with isolation by distance. *Molecular Ecology*, 21, 2839–2846. https://doi.org/10.1111/j.1365-294X.2012.05578.x

Meirmans, P. G., & Van Tienderen, P. H. (2004). GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, 4, 792–794. https://doi.org/10.1111/j.1471-8286.2004.00770.x

Miller, K. M., Supernault, K. J., Li, S., & Withler, R. E. (2006). Population structure in two marine invertebrate species (*Panopea abrupta* and *Strongylocentrotus franciscanus*) targeted for aquaculture and enhancement in British Columbia. *Journal of Shellfish Research*, 25, 33–42. https://doi.org/10.2983/0730-8000(2006)25[33:PSITMI]2.0.CO;2

Montgomery, J. C., Jeffs, A., Simpson, S. D., Meekan, M., & Tindle, C. (2006). Sound as an orientation cue for the pelagic larvae of reef fishes and decapod crustaceans. *Advances in Marine Biology*, 51, 143–196. https://doi.org/10.1016/S0065-2881(06)51003-X

Morrison, J., Callendar, W., Foreman, M. G. G., Masson, D., & Fine, I. (2014). A model simulation of future oceanic conditions along the British Columbia Continental shelf. Part I: Forcing fields and initial conditions. *Atmosphere - Ocean*, 52, 1–19. https://doi.org/10.1080/07055900.2013.868340

Nelson, R. J. (2003). CFAR Fisheries Development Sea Cucumber Project. Final Report. SeaStar Biotech Inc.

North, E., Adams, E., Schlag, S., Sherwood, C., He, R., & Socolofsky, S. (2011). Simulating oil droplet dispersal from the deepwater horizon spill with a Lagrangian approach. In Y. Liu, A. Macfadyen, Z.-G. Ji & R. Weisberg (Eds.), *AGU book series: Monitoring and modeling the deepwater horizon oil spill: A record breaking enterprise* (pp. 217–226). Washington, DC: Geophysical Union. https://doi.org/10.1029/GM195

Oksanen, J., Blanchet, F. G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., ... Wagner, H. H. (2017). vegan: Community Ecology Package. Package version 2.4-4.

Ovenden, J. R., Berry, O., Welch, D. J., Buckworth, R. C., & Dichmont, C. M. (2015). Ocean's eleven: A critical evaluation of the role of population, evolutionary and molecular genetics in the management of wild fisheries. *Fish and Fisheries*, 16, 125–159. https://doi.org/10.1111/faf.12052

Pedrotti, M. L., & Fenaux, L. (1992). Dispersal of echinoderm larvae in a geographical area marked by upwelling (Ligurian Sea, NW Mediterranean). *Marine Ecology Progress Series*, 86, 217–227. https://doi.org/10.3354/meps086217

Pineda, J., Hare, J. A., & Sponaugle, S. (2007). Larval transport and dispersal in the coastal ocean and consequences for population connectivity. *Oceanography*, 20, 22–39. https://doi.org/10.5670/oceanog

Poland, J. A., Brown, P. J., Sorrells, M. E., & Jannink, J. L. (2012). Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach. *PLoS ONE*, 7, e32253. https://doi.org/10.1371/journal.pone.0032253

Purcell, J. F., Cowen, R. K., Hughes, C. R., & Williams, D. A. (2006). Weak genetic structure indicates strong dispersal limits: A tale of two coral reef fish. *Proceedings of the Royal Society B: Biological Sciences*, 273, 1483–1490. https://doi.org/10.1098/rspb.2006.3470

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., ... Sham, P. C. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *The American Journal of Human Genetics*, 81, 559–575. https://doi.org/10.1086/519795

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Riginos, C., Crandall, E. D., Liggins, L., Bongaerts, P., & Treml, E. A. (2016). Navigating the currents of seascape genomics: How spatial analyses can augment population genomic studies. *Current Zoology*, 62, 581–601. https://doi.org/10.1093/cz/zow067

Rocha-Olivares, A., & Vetter, R. D. (1999). Effects of oceanographic circulation on the gene flow, genetic structure, and phylogeography of the rosethorn rockfish (*Sebastes helvomaculatus*). *Canadian Journal of Fisheries and Aquatic Sciences*, 56, 803–813. https://doi.org/10.1139/f99-004

Saenz-Agudelo, P., Jones, G. P., Thorrold, S. R., & Planes, S. (2009). Estimating connectivity in marine populations: An empirical evaluation of assignment tests and parentage analysis under different gene flow scenarios. *Molecular Ecology*, 18, 1765–1776. https://doi.org/10.1111/j.1365-294X.2009.04109.x

Schunter, C., Carreras-Carbonell, J., MacPherson, E., Tintoré, J., Vidal-Vijande, E., Pascual, A., ... Pascual, M. (2011). Matching genetics with oceanography: Directional gene flow in a Mediterranean fish species. *Molecular Ecology*, 20, 5167–5181. https://doi.org/10.1111/j.1365-294X.2011.05355.x

Selkoe, K. A., D'Aloia, C. C., Crandall, E. D., Iacchei, M., Liggins, L., Puritz, J. B., ... Toonen, R. J. (2016). A decade of seascape genetics: Contributions to basic and applied marine connectivity. *Marine Ecology Progress Series*, 554, 1–19. https://doi.org/10.3354/meps11792

Selkoe, K. A., Watson, J. R., White, C., Horin, T. Ben., Iacchei, M., Mitarai, S., ... Toonen, R. J. (2010). Taking the chaos out of genetic patchiness: Seascape genetics reveals ecological and oceanographic drivers of genetic patterns in three temperate reef species. *Molecular Ecology*, 19, 3708–3726. https://doi.org/10.1111/j.1365-294X.2010.04658.x

Stockstad, E. (2010). To fight illegal fishing, forensic DNA gets local. *Science*, 330, 1468–1469. https://doi.org/10.1126/science.330.6010.1468

Sunday, J. M., Popovic, I., Palen, W. J., Foreman, M. G. G., & Hart, M. W. (2014). Ocean circulation model predicts high genetic structure observed in a long-lived pelagic developer. *Molecular Ecology*, 23, 5036–5047. https://doi.org/10.1111/mec.12924

Suzuki, R., & Shimodaira, H. (2015). pvclust: Hierarchical Clustering with P-Values via Multiscale Bootstrap Resampling. R package version 2.0-0.

Thomas, L., & Bell, J. J. (2013). Testing the consistency of connectivity patterns for a widely dispersing marine species. *Heredity*, 111, 345–354. https://doi.org/10.1038/hdy.2013.58

Thomas, L., Kennington, W. J., Stat, M., Wilkinson, S. P., Kool, J. T., & Kendrick, G. A. (2015). Isolation by resistance across a complex coral reef seascape. *Proceedings of the Royal Society B: Biological Sciences*, 282, 20151217. https://doi.org/10.1098/rspb.2015.1217

Thorrold, S. R., Latkoczy, C., Swart, P. K., & Jones, C. M. (2001). Natal homing in a marine fish metapopulation. *Science*, 291, 297–299. https://doi.org/10.1126/science.291.5502.297

Treml, E. A., Halpin, P. N., Urban, D. L., & Pratson, L. F. (2008). Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. *Landscape Ecology*, 23, 19–36. https://doi.org/10.1007/s10980-007-9138-y

Treml, E. A., Roberts, J. J., Chao, Y., Halpin, P. N., Possingham, H. P., & Riginos, C. (2012). Reproductive output and duration of the pelagic larval stage determine seascape-wide connectivity of marine populations. *Integrative and Comparative Biology*, 52, 525–537. https://doi.org/10.1093/icb/ics101

Truelove, N. K., Kough, A. S., Behringer, D. C., Paris, C. B., Box, S. J., Preziosi, R. F., & Butler, M. J. (2017). Biophysical connectivity explains population genetic structure in a highly dispersive marine species. *Coral Reefs*, 36, 233–244. https://doi.org/10.1007/s00338-016-1516-y

Urban, D. L., Minor, E. S., Treml, E. A., & Schick, R. S. (2009). Graph models of habitat mosaics. *Ecology Letters*, 12, 260–273. https://doi.org/10.1111/j.1461-0248.2008.01271.x

Waples, R. S. (1998). Separating the wheat from the chaff: Patterns of genetic differentiation in high gene flow species. *The Journal of Heredity*, *89*, 438–450. https://doi.org/10.1093/jhered/89.5.438

Waples, R. S., & Gaggiotti, O. (2008). What is a population? An empirical evaluation of genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, *15*, 1419–1439.

Waples, R. S., Punt, A. E., & Cope, J. M. (2008). Integrating genetic data into management of marine resources: How can we do it better? *Fish and Fisheries*, *9*, 423–449. https://doi.org/10.1111/j.1467-2979.2008.00303.x

Ware, D. M., & McFarlane, G. A. (1989). Fisheries production domains in the Northeast Pacific Ocean. In R. J. Beamish & G. A. McFarlane (Eds.), *Effects of ocean variability on recruitment and an evaluation of parameters used in stock assessment models* (pp. 359–379). Canadian Special Publication of Fisheries and Aquatic Science 108. Ottawa: Fisheries and Oceans.

Watson, J. R., Kendall, B. E., Siegel, D. A., & Mitarai, S. (2012). Changing seascapes, stochastic connectivity, and marine metapopulation dynamics. *The American Naturalist*, *180*, 99–112. https://doi.org/10.1086/665992

Watson, J. R., Siegel, D. A., Kendall, B. E., Mitarai, S., Rassweiller, A., & Gaines, S. D. (2011). Identifying critical regions in small-world marine metapopulations. *Proceedings of the National Academy of Sciences*, *108*, E907–E913. https://doi.org/10.1073/pnas.1111461108

Weir, B., & Cockerham, C. C. (1984). Estimating F-statistics for the analysis of population structure. *Evolution*, *38*, 1358–1370.

White, C., Selkoe, K. A., Watson, J., Siegel, D. A., Zacherl, D. C., & Toonen, R. J. (2010). Ocean currents help explain population genetic structure. *Proceedings of the Royal Society B: Biological Sciences*, *277*, 1685–1694. https://doi.org/10.1098/rspb.2009.2214

Wu, T. D., & Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics*, *26*, 873–881.

SUPPORTING INFORMATION

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

How to cite this article: Xuereb A, Benestan L, Normandeau É, et al. Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). *Mol Ecol*. 2018;27:2347–2364. https://doi.org/10.1111/mec.14589