A peer-reviewed version of this preprint was published in PeerJ on 7 August 2014.

View the peer-reviewed version (peerj.com/articles/511), which is the preferred citable publication unless you specifically need to cite this preprint.

Asymmetric connectivity of spawning aggregations of a commercially important marine fish using a multidisciplinary approach

Adrian Munguia-Vega\textsuperscript{1,2}, Brad Erisman\textsuperscript{3}, Alexis Jackson\textsuperscript{4}, Silvio Guido Marinone\textsuperscript{5}, Marcia Moreno-Baez\textsuperscript{3}, Alfredo Girón\textsuperscript{6}, Tad Pfister\textsuperscript{7}, Octavio Aburto-Oropeza\textsuperscript{3}, Jorge Torre\textsuperscript{1}

\textsuperscript{1} Comunidad y Biodiversidad A.C., Guaymas, Sonora, México
\textsuperscript{2} Conservation Genetics Laboratory, School of Natural Resources and the Environment, The University of Arizona, Tucson, AZ, USA
\textsuperscript{3} Marine Biology Research Division, Scripps Institution of Oceanography, University of California, San Diego, La Jolla CA, USA
\textsuperscript{4} Department of Ecology and Evolutionary Biology, University of California Santa Cruz, Santa Cruz, CA, USA
\textsuperscript{5} Departamento de Oceanografía Física, Centro de Investigación Científica y de Educación Superior de Ensenada, Ensenada, Baja California, México
\textsuperscript{6} Universidad Autónoma de Baja California, Ensenada, Baja California, México
\textsuperscript{7} School of Natural Resources and the Environment, Center for Latin American Studies, The University of Arizona, Tucson, AZ, USA

Corresponding Author: Adrian Munguia-Vega. Comunidad y Biodiversidad A.C., Isla del Peruano 215, Col. Lomas de Miramar, Guaymas, Sonora CP 85448, Mexico. Phone: +52 612-1232233. E-mail: airdrian@email.arizona.edu
Abstract

Understanding patterns of larval dispersal is key in determining whether no-take marine reserves are self-sustaining, what will be protected inside reserves and where the benefits of reserves will be observed. However, explicitly incorporating dispersal data into designing reserves for fisheries and conservation is still uncommon in many places around the world. We followed a multidisciplinary approach that merged detailed descriptions of fishing zones and spawning time at 17 sites distributed in the Midriff Island region of the Gulf of California (GC) with a biophysical oceanographic model that simulated larval transport at Pelagic Larval Duration (PLD) 14, 21 and 28 days for the most common and targeted predatory reef fish (leopard grouper *Mycteroperca rosacea*).

*M. rosacea* is endemic to the GC and considered ‘Vulnerable’ according to World Conservation Union. We described metapopulation dynamics using graph theory and employed empirical sequence data from a subset of 10 sites at two mitochondrial genes to verify the model predictions. Our approach made sense of seemingly chaotic patterns of genetic diversity and structure, and provided a mechanistic explanation of the location of fishing zones. Most of the connectivity patterns observed were strictly asymmetric, except for a small region in the Southeast. The best-supported gene flow model confirmed a pulse of larvae from the Baja Peninsula, across the GC and northward up the Sonoran coastline, in agreement with the cyclonic gyre present at the peak of spawning (May). We found support that genetic diversity increased in sink sites that concentrated larvae from many sources at the time of larval flexion (PLD 14 days), while diversity decreased at important gateways identified at PLD 28 days with high betweenness centrality that are key for multigenerational dispersal and population resilience. Heavily
targeted fished areas seem to be sustained by high levels of local retention, contribution of larvae from upstream sites and oceanographic patterns that concentrate larval density from all over the region. The general asymmetry in marine connectivity observed highlights that benefits from reserves are biased towards particular directions, that no-take areas need to be located upstream of targeted fishing zones, and that some fishing localities might not directly benefit from avoiding fishing within reserves located adjacent to their communities. We discuss the implications of marine connectivity for the current network of marine protected areas and no-take zones, and identify ways of improving it.
Introduction

Knowledge of patterns of larval dispersal is essential to implement fully-protected marine reserves (no-take zones), a tool frequently used to enhance the conservation of biodiversity and the recovery of fisheries (Gaines et al. 2010). Reserves must either be self-sufficient via local retention (larvae retained or returning to the reserve where they were produced), or need to be linked by a network of reserves for persistence (larval supply from reserves to other reserves/fished sites) (Hastings & Botsford 2006; White et al. 2010). However, the efficacy of networks of reserves has been hindered by a lack of knowledge regarding complex patterns of marine connectivity (Burgess et al. 2013; Sale et al. 2005). A multidisciplinary approach could best address the intricacy of connectivity by merging biophysical models of ocean currents that generate connectivity hypotheses using detailed biological information on the spatial and temporal distribution of propagules (larvae), followed by validation with empirical population genetics data (Alberto et al. 2011; Crandall et al. 2012; Feutry et al. 2013; Foster et al. 2012; Soria et al. 2012). A multi-prong approach could also help advance an increasing interest in incorporating genetic information into marine spatial planning, for instance, by identifying sites with high genetic diversity that hold evolutionary potential under future environmental change (Beger et al. 2013).

Marine connectivity within a single species is influenced by multiple biological and physical factors including: spawning time and location, pelagic larval duration (PLD), their interaction with ocean current speed and direction, as well as the distribution of suitable habitat for settlement (Cowen & Sponaugle 2009). Additionally, many commercially exploited species of invertebrates and fishes display meta populations that
are connected via larval dispersal (Cowen 2000). This is why relatively few attempts have been done to establish multidisciplinary approaches to understand marine connectivity, and the great challenge is to find a key species that can be a relevant case study, and which can be use to gather this information relatively easily and can be use as an umbrella species to design marine reserves.

The leopard grouper *Mycteroperca rosacea* (Streets, 1877), is a large predatory reef fish (Teleostei: Epinephelidae) endemic to the Gulf of California (GC) bioregion. It ranges from Bahía Magdalena in the Pacific coast of the Baja California Peninsula south to Bahía Banderas in Nayarit, Mexico, including all rocky reefs within the interior of GC (Hastings et al. 2010; Robertson & Cramer 2009; Thomson et al. 2000). Ecologically, it represents the most common and numerically abundant fish top predator on reefs in the entire GC. Individuals can reach 1 m in length and at least 22 years of age (Diaz-Uribe et al. 2001). Histological and population data indicate gonochorism, with no evidence of post-maturational sex change found in adults caught in the wild (Erisman et al. 2007b). Adults form spawning aggregations of hundreds of individuals during spring in the GC, with spawning occurring earlier in southern locations (Erisman et al. 2007a; Sala et al. 2003). Spawning occurs in the evening within groups of 6 to 40 individuals and is not correlated with the lunar cycle (Erisman et al. 2007a).

Small fisheries world-wide comprise most of the global catch, yet most lack formal assessments and are thought to continue to decline (Costello et al. 2012). *M. rosacea* is the most heavily targeted grouper by commercial, artisanal, and recreational fisheries in the GC (Craig et al. 2012; Sala et al. 2003). Due to increased fishing pressure and...
observed declines in fisheries landings, sizes of harvested fish, and population abundances in some areas of the GC over the past few decades (Sala et al. 2004), the World Conservation Union (IUCN) currently lists *M. rosacea* as ‘Vulnerable’ (Craig & Sadovy 2008). While commercial fishers are required to hold a finfish permit and record their landings of leopard grouper to their local fisheries offices, no specific regulations related to catch, size, or gear restrictions exist for this species. Currently, all *M. rosacea* catches are aggregated in the finfish group with other 270 fish species according to the National Fisheries Chart (CNP 2012). Marine Protected Areas (MPAs) represent the primary, current conservation and management strategy that has been implemented for this or any other reef fish in the GC, and are mainly concentrated in the western coast of the GC and contain a few small no-take zones (Fig. 1) (Aburto-Oropeza et al. 2011; Cudney-Bueno et al. 2009).

Our goal was to identify the potential larval connectivity of spawning aggregation sites and fishing zones for *M. rosacea* where larvae recruit in the Midriff Island region of the GC. We first determined the distinct spawning season for leopard grouper and identified the spatial distribution of spawning aggregation sites and fishing zones across the entire region. We modeled connectivity with a biophysical model and used graph theory to describe metapopulation dynamics. We then contrasted distinct measures derived from graph theory against empirical estimates of genetic diversity and differentiation to corroborate model expectations' and identify sites that are likely self-sustaining and important sources and sinks for leopard grouper larvae, including locations that may lie inside or outside the borders of existing MPAs. Results of this study provide insights on validating biophysical models with empirical genetic data, on the benefits and
limitations of the current network of MPAs to fishing communities in the Midriffs region that harvest *M. rosacea* and help identify areas that may serve as ideal locations for spawning or juveniles refuges of this economically important yet vulnerable species.

**Methods**

*Spawning sites, season and period*

We examined information acquired from underwater and fisheries surveys conducted at locations throughout the Midriffs Islands region in the GC, Mexico (Fig. 1) in order to identify representative sites to simulate the dispersal of leopard grouper eggs and larvae from spawning aggregation sites. Underwater surveys were performed at 33 sites throughout the Midriffs during the spawning season of *M. rosacea* (April to June) in 2008, 2009, and 2010. Evidence of the formation of spawning aggregations were based on standard protocols (Colin et al. 2003) and those adapted for leopard grouper (Erisman et al. 2007b). Direct evidence of spawning aggregations included observations of courtship or spawning behavior or the collection of females with hydrated or ovulated oocytes. Indirect evidence involved observations of putative females with enlarged abdomens indicative of imminent spawning, color patterns associated with courtship, the collection of males with ripe testes, and abundances and densities of fish that were markedly higher (e.g., 3-fold increases or greater) than observed during non-spawning months. Additional indirect evidence of spawning aggregations was acquired through interviews with commercial fishers at five fishing communities (Bahía de los Ángeles, Bahía de Kino, Desemboque Seri, Puerto Libertad and Punta Chueca) during 2005 and 2006 (Moreno-Báez et al. 2012; Moreno-Baez et al. 2010).
Spatial units (Fig. S1) were established to evaluate spatial connectivity by combining physical and political boundaries, as well as local knowledge from fishers (Moreno-Báez et al. 2012; Moreno-Baez et al. 2010). We incorporated coastline and bathymetry developed by the National Geophysical Data Center (http://www.ngdc.noaa.gov/mgg/shorelines/shorelines.html) and the marine protected areas in Mexico (www.conanp.gob.mx), respectively. We used the spatial union function to integrate the different boundaries and define the spatial units, under ArcGIS 10.1 (ESRI) with the Spatial Analyst Extension and Model Builder tools. The size of the spatial units varied from 13 to 812 km² (Fig. S1).

While the general spawning season for *M. rosacea* in the GC occurs from late April to June (Erisman et al. 2007a), it was necessary to collect empirical data to narrow the specific spawning season from the Midriff Islands region. We acquired gonad samples of adult female leopard groupers (*i.e.* > 30 cm TL; Erisman et al. 2007a) from commercial fishers on a monthly basis from December 2008 to June 2010. Fish were captured by gill nets or handlines at various sites at or near San Pedro Martir and Tiburon islands, and Bahia Kino (Fig. 1). We processed tissue taken from the central portion from one gonad lobe for each sample using standard histological techniques (Humason 1972) in order to determine sex and developmental stage. Classes of ovarian and testicular development were adapted from previous studies (Erisman et al. 2007b), and stages of gametogenesis followed previously established definitions (Wallace & Selman 1981).

We determined the duration of the spawning season using a combination of two methods. First, we examined the histological preparations of all gonad samples to identify the percentage of females capable of spawning or actively spawning during each month.
Females categorized as ‘spawning capable’ included those with ovaries dominated by oocytes in advanced stages of vitellogenesis (e.g., primary to tertiary yolk stage), whereas those categorized as ‘actively spawning’ contained ovaries with oocytes in the migratory nucleus, hydrated, or ovulated stage or post-ovulatory follicles were present. Data were pooled by month to estimate the monthly proportion of spawning females over a calendar year. Dates on which actively spawning females were collected were used as indicators to determine exact dates of spawning. A second estimate of the spawning season was obtained by calculating the mean monthly gonadosomatic index (GSI = 100* gonad weight/ total body weight) of female *M. rosacea* over the study period. Changes in monthly GSI were used to assess reproductive activity, associating elevated levels with gonadal development and spawning. This information was used in the release dates for larvae in the oceanographic model (see below).

 Supporting fishing knowledge

The central component to documenting the fishing grounds for the *M. rosacea* was captured through a series of interviews implemented across 17 fishing communities in the northern GC (Moreno-Báez et al. 2010, Moreno-Báez et al., 2012). The methodology entailed aggregating local knowledge of a representative set of individual fishers (captains) through semi-structured interviews conducted between December 2005 and July 2006 regarding what, where, when and how they fish. The interview included questions regarding the spatial and temporal distribution of fishing activities but also, their knowledge about spawning aggregations and juvenile sights. The maps were digitized, georeferenced, and integrated into a geographic information systems (GIS)
These interviews indicated that fishing activity frequently overlapped spatially with spawning aggregation sites in three main regions: 1) the north end of Angel de la Guarda Island; 2) on the south, western and northern edge of Tiburon Island, and on sites in mainland Sonora north of Tiburon, around Las Cuevitas, Puerto Libertad and Puerto Lobos (yellow areas Fig. 1). Other important fishing zones were identified around Puerto Peñasco in northern Sonora. According to the interviews, the principal fishing season starts in November and ends in June.

Oceanographic model

In a computer simulation exercise, four thousand particles were released at a depth of 5 m at each of 17 spawning aggregation sites (see below). Particles were tracked for 28 days, which is close to the maximum pelagic larval duration (PLD) for *M. rosacea* (Aburto-Oropeza et al. 2007). We released particles during May 16th and May 25th of 2007, which based on our observations covers the peak spawning period around the Midriff Islands region (see Results). Since seasonal oceanographic regimes are consistent across years in the GC (Marinone 2003; Soria et al. 2013), the simulation year was chosen arbitrarily (2007) while the dates covered spring and neap tides. The phase of the springs-neaps cycle is simply shifted from year to year.

We used the velocity field from the GC implementation of the three-dimensional baroclinic Hamburg Shelf Ocean Model (HAMSOM) (Backhaus 1985) to calculate the particle trajectories. The model has been described in detail for the GC (Marinone 2003; Marinone 2008). Its domain has a mesh size of \(2.5'\times2.5'\) \((\sim1.31\times1.54\text{ km})\) in the horizontal and 12 layers in the vertical with nominal lower levels at 10, 20, 30, 60, 100,
150, 200, 250, 350, 600, 1000 and 4000 m. The model equations are solved semi-
implicitly with fully prognostic temperature and salinity fields, thus allowing time-
dependent baroclinic motions. The model is started from rest with a 300 s time step and
becomes periodically stable after three years. Results for this study were obtained from
the fourth year of the model when it adequately reflects the main seasonal signals of
surface temperature, heat balance, tidal elevation and tidal currents and surface
circulation in the NGC (Lavin et al. 1997; Marinone 2003). The forcing includes at the
open boundary model tidal components (M2, S2, N2, K2, K1, O1, P1, Ssa, and the Sa),
climatological hydrography historical data and at the sea surface climatological heat and
fresh water fluxes. We used the seasonal climatology constructed from QUICKSCAT
data as forcing for wind. The Lagrangian trajectories are due to the Eulerian velocity field
plus a random-walk contribution related to turbulent eddy diffusion processes (Proehl et
al. 2005; Visser 1997). We obtained values of the diffusivities from the numerical model.
A pseudo-advective term was introduced, since the vertical diffusivity is not constant, to
prevent particles from walking away from areas of high to low diffusivities. The velocity
at each particle position and the vertical eddy coefficients are calculated by bilinear
interpolation of the instantaneous Eulerian velocity fields and the eddy coefficient from
the numerical model, which were saved every hour. The horizontal diffusivity is taken as
a constant (100 m$^2$/s). We reasonably assumed that larvae are advected as passive
particles and do not migrate vertically downward to deep depths (Watson et al. 2010),
given that leopard grouper recruit to shallow Sargassum spp. beds of < 5 m deep
(Aburto-Oropeza et al. 2007).
Modeled connectivity

Hourly latitude and longitude data for each modeled particle were imported into MatLab (MATHWORKS). We estimated connectivity at different time intervals: 336 h (14 days), 504 h (21 days) and 672 h (28 days) respectively after the released dates. These PLDs were selected based on the average time of flexion in groupers that corresponds to the onset of larval behavior (14 days) (Cowen 2002; Gracia-Lopez et al. 2005) and the maximum PLD (Aburto-Oropeza et al. 2007). A selection by location function “inpolygon” was used to identify the intersection between particles and the recruitment areas (spatial units). We then generated connectivity matrices using the proportion of larvae that settled at each location relative to the total number of larvae released at each site. We constructed matrices averaging for the two spawning dates May 16th and May 25th within each PLD (i.e. day 14, 21, 28). The probability of local retention (i.e., diagonal in the connectivity matrix) was calculated as the proportion of particles produced locally that remained within the spatial unit at the end of the PLD (Burgess et al. 2013). The probabilities within each site were summarized with two statistics aimed at describing source-sink dynamics. Export probability was defined as the proportion of larvae produced within a site that successfully settled within any of the other 16 coastal areas left at the end of the PLD. Import probability was defined as the proportion of all larvae produced among the 17 sites that settled within each site. This later metric is identical to self-recruitment as defined by Burgess et al. 2013.

Marine connectivity patterns were displayed using graph theory and a spatial network approach (Treml et al. 2012) with the software NODEXL (Smith et al. 2010). We calculated four statistics that describe the relationships among elements in complex
networks (Newman 2003), including: 1) graph size (the total number of directed links within a graph); 2) in-degree (number of links that enter a node); 3) out-degree (number of links that leave a node); and 4) betweenness centrality, or the proportion of shortest paths between all node pairs that pass through a particular node, which highlights ‘most used’ dispersal pathways or stepping stones that act like gateways though which genes or individuals have to pass to spread to other nodes, emphasizing key sites for multigenerational connectivity (Andrello et al. 2013). Betweenness centrality can be viewed as a measure of resilience by measuring how many paths will get longer when a node is removed (Newman 2003).

Genetic connectivity

We collected tissue samples from the pectoral fins of *M. rosacea* from 10 sites included in our modeling exercise around the Midriff Islands region (Fig. 1). Samples were acquired in fish markets or directly from fishermen at harbors between 2009 and 2012 under IACUC protocol Berng1101. We interviewed both fish vendors and fishermen to determine the approximate localities where fish were collected. Immediately after collection, samples were stored in 95% ethanol and kept at -20°C in the laboratory. Genomic DNA was extracted using standard chloroform extraction protocols (Sambrook et al. 1989). We amplified a 787 bp fragment of mitochondrial marker cytochrome *b* using primers Gludgl and CB3H (Palumbi et al. 1991). Thermocycler parameters were as follows: initial hold at 94°C/5 min, 35 cycles of 94°C/45 sec, 45°C/45 sec, 72°C/45 sec, with a final extension of 72°C/7 min. We developed species-specific primers for *M. rosacea* (MYCROS Forward: TTCTCCCACTACCTGATT and
MYCROS Reverse: TACGTAGGCTTGATCATTG) to amplify a 726 bp fragment of mitochondrial marker ATPase. Thermocycler parameters were as follows: initial hold at 94°C/5 min, 35 cycles of 94°C/30 sec, 54°C/30 sec, 72°C/30 sec, with a final extension of 72°C/7 min. After purification of PCR products following ABI manufacturer’s protocols (ABI, Perkin-Elmer), we sequenced clean PCR products on an ABI 3730xl automated sequencer (Applied Biosystems, Foster City, CA).

We calculated molecular diversity indices including nucleotide diversity (π) and haplotype diversity (h). We corrected haplotype diversity using CONTRIB (Petit et al. 1998) to account for differences in sample size between sites based on rarefaction to a minimum sample size of n = 4. Theory predicts that genetic diversity levels observed whiting sites is highly dependent upon the amount of migration from source populations (Gaggiotti 1996), and that genetic diversity increases in sink sites that concentrate larvae from multiple sources (Kool et al. 2011). To compare directly the ocean model with the empirical genetic data, we performed simple linear regressions between corrected haplotype diversity and in-degree, out-degree and betweenness centrality estimates for networks at each PLD. To explore the possibility that patterns could be confounded if a single outlier site if it would have been modeled wrong, we repeated the regressions leaving out one site at a time.

Phylogenetic relationships among sequences were inferred from a haplotype network based on pairwise differences between haplotypes generated using Arlequin (Excoffier et al. 2005) and R software. To test for hierarchical population structure we performed an Analysis of Molecular Variance (AMOVA) in Arlequin. AMOVA significance was estimated using a permutation test of 10,000 replicates. The 10 sites
were clustered into three regions: Baja Peninsula (La Ventana, La Poma and San Francisquito), the Midriff Islands (San Pedro Martir, Salsipuedes, Datil, San Esteban) and the Sonoran coast (Puerto Libertad, El Tecomate, Puerto Lobos, Puerto Libertad). Chi-squared analyses were concurrently performed using DnaSP version 5.10 (Librado & Rozas 2009) to test for patterns of regional subdivision. Pairwise comparisons were made between each location to assess patterns of genetic differentiation.

We evaluated three different migration models using Migrate-n 3.2.16 (Beerli & Palczewski 2010). First, we tested an unrestricted full migration model between all sampling localities. Next, we considered two models with three population sizes comprised of a subsampling (n=30) from sampling localities in the Baja Peninsula, the Midriff Islands, and the Sonoran coast. One model assessed unidirectional gene flow from the Baja Peninsula, across the Midriff Islands, and northward up the Sonoran coast, while the other model tested gene flow in the reverse direction. The latter two models reflect seasonal differences in directionality of a cyclonic (May to September) and anticyclonic (October to April) gyres, respectively, present in the northern GC (Marinone 2003; Marinone 2012). Using a Bezier approximation, we chose the most appropriate model for our dataset by taking the natural log of the ratio of the marginal likelihoods (Baye’s factors) for each model (Beerli & Palczewski 2010). Running conditions for Migrate-n were as follows: 5,000,000 recorded steps, a burn-in of 2,500,000 steps, a static heating scheme using 20 temperatures, a tree swapping interval of 1, and an upper prior boundary for migration set to 7,500.
Results

Spawning sites, season and period

We chose seventeen sites representative of distinct spatial units for the release of virtual larvae in the simulation model based on direct and indirect evidence of the presence of spawning aggregations for leopard grouper (Table 1). Although some spatial units had evidence of multiple spawning aggregations, we assumed their close proximity along with the spatial resolution of the oceanographic model meant multiple aggregations within the same unit would disperse larvae in similar directions. Individual sites were distributed throughout the region and fulfilled a range of 3 to 7 criteria, with an average of 4.76±1.48 (SD). A marked increase in the abundance of adult groupers during the spawning season relative to the non-spawning season was the most common evidence (recorded at all 17 sites). Other types of indirect evidence such as the observation of gravid females with swollen abdomens, observations of fish exhibiting courtship coloration, the collection of running-ripe males, or elevated catch rates by fishers during the spawning season were recorded for the majority of sites. Direct evidence of spawning via the collection of hydrated females was recorded for 65% of the sites, whereas spawning was observed at only 35% of the sites.

A total of 162 samples of female *M. rosacea* were collected from commercial fishers over the study period, with an average of 14 samples collected each month (range = 8 to 27). Based on microscopic examinations of gonadal tissue samples, females in the spawning capable phase were collected from March through June, and actively spawning females were collected April to June. Similarly, the GSI of adult females showed elevated levels from April to June, with a peak during May (Fig. 2). When the results of
the gonadal phases and GSI were combined, they indicate that *M. rosacea* spawn from April to June in the Midriffs region, with peak spawning activity occurring in May. Actively spawning females were collected on three days in 2009 (14 May, 31 May, 25 June) and two days in 2010 (25 April, 7 May).

**Modeled connectivity**

From all simulated particles released (4,000 particles x 17 sites x two release dates = 136,000), coastal areas that are suitable for larval recruitment captured 41.84% (PLD 14 days), 35.24% (PLD 21 days) and 33.26% (PLD 28 days). Remaining particles did not reach any coastal habitat by the end of the PLD. Simulations of ocean currents produced the highest concentrations of larvae in coastal areas north of Tiburon Island after PLD 28 days (around Las Cuevitas Fig. 3), followed by the south end of Tiburon Island, Puerto Libertad, and the north end of Angel de la Guarda Island. Simulations at PLD 14 and 21 days indicated similar trends (Fig. 3).

The trajectories of particles released from each site at the two dates (Fig. 4a and 4b) showed sites in the Baja California Peninsula generally followed a southward direction (except those released from La Ventana that entered the Canal de Ballenas), while sites on mainland Sonora and the north edge of Tiburon Island followed a northward trajectory. Most locations around Angel de la Guarda Island and Tiburon Island had particles dispersing north and south of the release site. In all cases, the distance traveled by particles was directly proportional to the PLD. The graph size of the connectivity networks increased from 38 edges at PLD 14 days, to 59 at PLD 21 days and 67 at PLD 28 days, indicating a longer PLD is associated with more connected and...
complex networks (Fig. 5). Although few differences were observed between networks derived from the two release dates (see Fig. S2 for an example at PLD 28 days), connectivity patterns showed similar trends. Based on the directionality of the links in the networks, three main patterns were evident. First, northward links were prevalent along the eastern coast of the GC, in the Canal de Ballenas and between the southern end of Angel de la Guarda Island and San Lorenzo Island across the GC towards the northern end of Tiburon Island and mainland Sonora (Fig. 5). Second, southward links were present between the eastern coast of Angel de la Guarda Island towards southern locations in San Lorenzo Island and Baja California and across the GC to Tiburon Island and San Esteban Island. Third, bi-directional north-south links were evident only in a small area located between San Pedro Martir Island, San Esteban Island and the southern end of Tiburon Island. Overall, the strongest links (i.e., those showing the larger probabilities) were observed between the western and northern coasts of Tiburon Island towards northern localities situated in mainland Sonora and in the Canal de Ballenas (Fig.

The probability of local retention decreased with increasing PLD, except in Datil (Fig. 6). According to the ocean model, local retention was most likely in Puerto Libertad (range: 0.26 to 0.72 for PLD 28 and 14 days, respectively), the northern end of Tiburon Island (El Tecomate, range: 0.21 to 0.69), followed by the northern end of Angel de la Guarda Island (Punta Refugio, range: 0.28 to 0.46), the western coast of Tiburon Island (La Tordilla, range: 0.07 to 0.35), and the southern coast of Tiburon Island (Datil Island, range: 0.13 to 0.24). With one exception (Puerto Lobos at PLD 14), all other sites had probabilities below 0.07.
Our analyses about the performance of each site on the network using statistics describing the probabilities of local retention, export/import and the number of connections leaving and entering each site (degree) illustrated the spatial overlap of some criteria, highlighting a few sites that were important despite variation in PLD (Fig. 6, Table S1-S3). For instance, Tecomate on the north of Tiburon Island and Puerto Libertad on mainland Sonora had the largest probabilities of export, import and local retention. However, both sites exported larvae to only a few sites and showed relatively low betweenness centrality. The north end of Angel de la Guarda Island (Puerto Refugio) also had relatively large probability of import and local-retention, but low export probability to few sites. Some sites acted exclusively as sources (e.g., La Ventana) with high export probability to many sites despite low local retention and low import probabilities. Datil Island was identified as a sink with intermediate levels of larvae imported from many sites but relatively low export probabilities. Betweenness centrality was more sensitive to PLD variation than the other measures, and identified Las Cuevitas on mainland Sonora, the Southern end of Angel de la Guarda Island and La Ventana as key sites for multigenerational larval dispersal through the entire network.

Genetic connectivity

We analyzed a 787 bp fragment of cytochrome \( b \) and a 726 bp fragment of ATPase for 235 individuals. We identified a total of 77 haplotypes, with adjacent haplotypes in the haplotype network separated by 1 to 4 bp (Fig. 7). There was limited evidence of geographic separation of haplotypes, and the three most frequent haplotypes were present in all locations, with the exception of the third most frequent which was
absent in San Francisquito on the Baja Peninsula. Corrected estimates of haplotype diversity were high, ranging from 0.844 (San Lorenzo Island) to 0.943 (San Esteban Island) (Table 2). Linear regression models were unable to find significant correlations between empirical observations (corrected haplotype diversity) and modeled connectivity measured with in-degree, out-degree and betweenness centrality at any PLD (all $R^2 \leq 0.147$, $P$ values $\geq 0.550$, Table S1-S3). However, upon elimination of one site at a time, we observed two significant trends when San Lorenzo Island was excluded from the analyses (Fig. 8). First, genetic diversity was significantly higher at sites with larger in-degree values estimated from the model at PLD 14 days ($R^2 = 0.489$, $P = 0.035$, Fig. 8a). Second, genetic diversity was significantly lower at locations showing higher betweenness centrality according to the model at PLD 28 days ($R^2 = 0.471$, $P = 0.041$, Fig. 8b). All other comparisons leaving out one site at a time were not significant ($R^2 < 0.170$, $P \geq 0.175$).

Statistically significant pairwise estimates of genetic structure were observed (Table 3). Pairwise $\phi_{ST}$ values suggest Puerto Libertad is genetically divergent from the majority of other sampling localities. Pairwise $F_{ST}$ values suggest greater overall genetic differentiation between most sampling localities, except between San Esteban Island and all the other sites. Global estimates of $F_{ST}$ and $\phi_{ST}$ suggest moderate levels of population structure within the northern Gulf ($\phi_{ST} = 0.04663$, $P = 0.0001$; $F_{ST} = 0.10836$, $P < 0.00001$). Regional genetic subdivision was also observed when sampling sites were clustered into the following groups – Baja Peninsula, Midriff Islands and the Sonoran coast. Genetic subdivision of regional groups was supported by a chi-squared test ($\chi^2 = 186.876$, d.f. = 152, $P = 0.0286$). Rankings of proposed larval dispersal models between
the aforementioned regions are listed in Table 4. The best-supported model was for unidirectional larval dispersal from the Baja Peninsula to the mainland. There was minor support for the full, unrestricted larval dispersal model and almost no support for unidirectional larval dispersal from the mainland to the Baja Peninsula.

Discussion

Our study contribute to a growing body of literature (Alberto et al. 2011; Crandall et al. 2012; Feutry et al. 2013; Foster et al. 2012; Galindo et al. 2010; Petitgas et al. 2012; Selkoe et al. 2010; Soria et al. 2012) highlighting the inherent value of verifying outputs of biophysical oceanographic models with empirical genetic data to inform larval dispersal patterns and marine connectivity. Concordance of genetic and biophysical modeling data for *M. rosacea* elucidate the role of oceanographic processes in driving patterns of larval dispersal, while models helped to explain seemingly chaotic patterns of genetic diversity and structure. The best supported gene flow model (based on genetic subdivisions) confirmed biophysical model outputs demonstrating a pulse of larvae from the Baja Peninsula, across the GC and northward up the Sonoran coastline, in agreement with the cyclonic gyre present from May to September. Our modeled connectivity networks mirrored these results. Furthermore, our study demonstrates that patterns of oceanographic circulation in the GC may be a powerful predictor of source-sink dynamics for *M. rosacea*. In our dataset we found support from empirical genetic data that genetic diversity increases in sink sites that concentrate larvae from many sources. However, the fact that patterns were significant only for PLD 14 days supports that larvae reaching suitable habitat are already able to settle early at the average time of flexion in
groupers. Our results could suggest that biophysical models in combination with graph theory could be used as a proxy for predicting genetic diversity, but further studies are needed to verify the validity of this relationship across markers and species. For instance, San Esteban Island, the site with the largest number of incoming sources of larvae at PLD 14 days and with the largest diversity of haplotypes, showed the lowest levels of genetic differentiation according to pairwise statistics, likely as a result of high levels of gene flow towards this site. In contrast, Puerto Libertad on mainland Sonora showed the largest genetic differences compared to other sites, which could be explained by a higher proportion of kin than expected by chance due to high levels of local retention (Iacchei et al. 2013), as supported by the model regardless of PLD. By coupling modeled and empirical connectivity approaches, we are able to better understand the mechanisms driving dispersal in the Gulf and to potentially inform spatially explicit management efforts for *M. rosacea* as well as marine organisms with similar life histories. However, validation of the passive dispersal model through subsequent studies such as those that use parentage analyses and highly polymorphic microsatellite loci are recommended and underway.

Our study had several limitations. Below we discuss some of the major ones. The relationship between patterns of genetic connectivity and modeled larval connectivity was significant only after removing San Lorenzo Island, where, according to observed levels of genetic variation, our ocean model seemed to have overestimated connectivity in terms of the number of sources supplying larvae to that site. San Lorenzo is a relatively thin and small island, and a detailed examination of the connectivity matrices indicated that the scenario where most of the larvae arrive from multiple sites to this one, has very...
small probabilities (< 0.01). These observations suggest that the large size of the spatial
unit in our study, relative to the small size of the island, coupled with a large bathymetric
profile around the island, likely biased real connectivity estimates, indicating room for
improving the spatial resolution of our analyses. While there is no direct evidence that
groupers larvae vertically migrate downward to escape advecting currents, there is a
growing body of evidence that suggests that local retention of larvae in groupers can be
quite high (Almany et al. 2013; Harrison et al. 2012). Thus our passive model could have
overestimated larval exchange rates and underestimated local retention (Cowen 2000).
However, the effects of vertical migration are comparable to those of reducing PLD
(Andrello et al. 2013). Investigations into larval behavior of groupers are warranted and
could greatly increase the precision and accuracy of the model. Our models did not
include an explicit description of the habitat for larval recruitment (Sargassum sp. beads),
neither considered larval mortality after settlement, thus we only assessed potential
connectivity, as opposed to realized connectivity.

Our multidisciplinary approach provided a mechanistic explanation of why some
areas in the Midriff Island region concentrate the fishing effort for leopard grouper in the
GC. Heavily targeted fished areas, including the north end of Angel de la Guarda Island,
the west, north and south edges of Tiburon Island and Las Cuevitas and Puerto Libertad
on mainland Sonora, showed the largest values of local retention of larvae, together with
a high probability of importing larvae from other spawning sites and for concentrating
larvae from all over the region. Notably, some of these are known to be sites that
historically have held huge spawning aggregations of leopard grouper that have been
harvested at high levels for decades, like the north end of Angel de la Guarda Island
Thus, the main fishing areas seem to depend both on local retention and contributions of larvae from upstream sites, coupled to oceanographic patterns that focus larval density towards these areas that sustain most of the fisheries.

A key result of our study is the observation that marine connectivity for *M. rosacea* around the Midriff island region is predominantly asymmetric. Other studies have previously shown the negative effects that asymmetric connectivity has on population persistence (Bode et al. 2008; Vuilleumier et al. 2010). In the presence of strong asymmetric currents, reserves can significantly outperform traditional quota based management strategies in terms of fisheries yield, with considerably less risk (Gaines et al. 2003). Asymmetry also constrains the notion that benefits of reserves in terms of larval input are proportional to their distance to the reserve (Almany et al. 2009; Buston et al. 2012). For example, one study using DNA parentage analyses found that reserves in the Great Barrier Reef, which accounted for 28% of the local reef area, produced approximately half of all juvenile recruitment of snappers and groupers to both reserve and fished reefs within 30 km of the source spawning site inside the reserve (Harrison et al. 2012), while a similar study in Papua New Guinea found that 50% of larvae in a coral grouper settled within 14 km of the spawning aggregation sites (Almany et al. 2013). In contrast, the benefits of reserves are completely biased towards one particular direction in the GC, highlighting that the spatial location of no-take zones is even more important in the Midriff Island region than in other systems. An exception to the general asymmetry in connectivity was detected in small area between the south edge of Tiburon Island, San Esteban Island, and San Pedro Martir Island.
A network of no-take zones within the Midriff region might have a very well defined zone of influence that does not include the eastern edge of Tiburon Island or any locality towards the south in mainland Sonora. This observation has important practical implications. For example, fishing localities on mainland Sonora South of Tiburon Island are restricted from fishing at no-take areas within MPAs in the Midriffs, yet according to this an other studies (Soria et al. 2013) they receive no benefit by not fishing there.

Conversely, fishing communities in mainland Sonora (Puerto Lobos, Puerto Libertad) seem to receive great benefits from San Pedro Martir, San Esteban and Tiburon Islands, even though they may not fish there. This brings up an important concept in highly advective systems like the GC where there may be a spatial disconnect and strong directionality between the location of no-take zones and the areas that benefit most from them, and highlights that, in order of reserves to be effective, they need to be located upstream of targeted fishing sites (Beger et al. 2010). Our analyses suggest that establishment of smaller no-take zones at the north end of Angel de la Guarda Island within the current MPA will likely boost local fisheries via local retention, while other current no-take zones within the Canal de Ballenas and San Lorenzo MPA could export larvae to fishing sites across the GC. The additional establishment of no-take zones adjacent to current heavily fished areas in the western and northern edges of Tiburon Island, and in the coast between Las Cuevitas-Puerto Lobos will likely increase not only local fisheries (via local retention) but fisheries at downstream fished sites on mainland Sonora as north as Puerto Peñasco (located ~300 km from Tiburon Island) via larval dispersal. Notably, except for San Francisquito on the coast of Baja California, current MPAs do not include those sink sites receiving larvae from multiple sources and that
harbor the largest genetic diversity and evolutionary potential (San Esteban and Tiburon Islands). In other hand, the fact that key gateway routes for multigenerational dispersal showing high betweenness centrality at PLD 28 days had low genetic variation (e.g. La Ventana on Baja California, Las Cuevitas on mainland Sonora), stress that important sites for long-term metapopulation persistence and resilience are not necessarily aligned spatially with other criteria for protection, such as preserving evolutionary potential via genetic variation.

Acknowledgements

We would like to acknowledge Rene Loiaza, Ivan Martinez, Angeles Cruz, Mario Rojo, Kimberly Tenggardjaja, Eva Salas, Gary Longo, Emmett Ziegler and Ana Liedke for their assistance with acquiring samples in the field. This a scientific contribution of the multidisciplinary PANGAS project (www.pangas.arizona.edu).

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

This work was funded by the David and Lucile Packard Foundation, National Geographic (Young Explorers Grant #8928-11), the University of California Office of the President (UC-HBCU Initiative grant), Marine Conservation Institute (Mia J. Tegner memorial research grant), Ecology Project International, CenTread, the Dr. Earl H. Myers and Ethel M. Myers Oceanographic and Marine Biology Trust, FCT (SFRH/BPD/26901/2006) and the Friends of Long Marine Lab. Additional funding for this research was provided by
Dr. Exequiel Ezcurra through the Pew Fellowship Program on Marine Conservation, and several grants from the Walton Family Foundation.

**Grant Disclosures**

The following grant information was disclosed by the authors:

- The David and Lucile Packard Foundation (grant award #2008-32210, #2013-39400)
- University of California Office of the President (UC-HBCU Initiative grant)
- Marine Conservation Institute (Mia J. Tegner memorial research grant)
- Ecology Project International
- CenTread
- Dr. Earl H. Myers and Ethel M. Myers Oceanographic and Marine Biology Trust
- FCT (SFRH/BPD/26901/2006)
- Friends of Long Marine Lab
- Pew Fellowship Program on Marine Conservation
- Walton Family Foundation

**Competing Interests**

The authors declare that they have no competing interests.

**Author Contributions**

Adrian Munguia-Vega, Brad Erisman, Alexis Jackson, Jorge Torre and Tad Pfister conceived and designed the experiments.
Adrian Munguia-Vega, Alexis Jackson, Silvio Guido Marinone, Marcia Moreno-Baez and Alfredo Giron performed the experiments and analyzed data.

Adrian Munguia-Vega, Brad Erisman, Alexis Jackson, Octavio Aburto-Oropeza and Jorge Torre wrote the paper.

Field Study Permissions

Scientific collection permits were acquired from SAGARPA (No. DGOPA.09151.260809.2885 and DAPA/2/020511/01197).

Supplemental Information

Table S1 Statistics describing the dynamics of larval dispersal in a network of 17 sites after PLD 14 days.

Table S2 Statistics describing the dynamics of larval dispersal in a network of 17 sites after PLD 21 days.

Table S3 Statistics describing the dynamics of larval dispersal in a network of 17 sites after PLD 28 days.

Figure S1 Spatial units of analyses for studying connectivity in the Gulf of California.

Figure S2 Modeled networks of larval connectivity for PLD 28 days for larvae released at two distinct dates.
References


Gaines SD, Gaylord B, and Largier JL. 2003. Avoiding current oversights in marine 

networks for both conservation and fisheries management. Proceedings of 
the National Academy of Sciences of the United States of America 107:18286-
18293.

Galindo HM, Pfeiffer-Herbert AS, McManus MA, Chao Y, Chai F, and Palumbi SR. 
2010. Seascape genetics along a steep cline: using genetic patterns to test 

Gracia-Lopez V, Kiewek-Martinez M, Maldonado-Garcia M, Monsalvo-Spencer P, 
Portillo-Clark G, Civera-Cerecedo R, Linares-Aranda M, Robles-Mungaray M, 
and Mazon-Suastegui JM. 2005. Larvae and juvenile production of the 
leopard grouper, Mycteroperca rosea (Streets, 1877). Aquaculture 
Research 36:110-112.

Harrison HB, Williamson DH, Evans RD, Almany GR, Thorrold SR, Russ GR, Feldheim 
KA, van Herwerden L, Planes S, Srinivasan M et al. 2012. Larval export from 
marine reserves and the recruitment benefit for fish and fisheries. Current 

Hastings A, and Botsford LW. 2006. Persistence of spatial populations depends on 
returning home. Proceedings of the National Academy of Sciences of the United 
States of America 103:6067-6072.

Hastings PA, Findley LT, and Van der heiden AMV. 2010. Fishes of the Gulf of 
California. In: Brusca RC, ed. The Gulf of California Biodiversity and 

and Company.

Iacchei M, Ben-Horin T, Selkoe KA, Bird CE, Garcia-Rodriguez FJ, and Toonen RJ. 
2013. Combined analyses of kinship and FST suggest potential drivers of 
chaotic genetic patchiness in high gene-flow populations. Molecular ecology 
22:3476-3494.

Kool JT, Paris CB, Barber PH, and Cowen RK. 2011. Connectivity and the 
development of population genetic structure in Indo-West Pacific coral reef 

Lavin MF, Durazo R, Palacios E, Argote ML, and Carrillo L. 1997. Lagrangian 
observations of the circulation in the northern Gulf of California. Journal of 

Librado P, and Rozas J. 2009. DnaSP v5: a software for comprehensive analysis of 

Marinone SG. 2003. A three-dimensional model of the mean and seasonal circulation 

Marinone SG. 2008. On the three-dimensional numerical modeling of the deep 
circulation around Ángel de la Guarda Island in the Gulf of California. 
Estuarine, Coastal and Shelf Science 80:430-434.


Table 1 Study sites in the Gulf of California and selection criteria. Each site was selected based on seven criteria to define where spawning aggregations might act like source of larvae.

<table>
<thead>
<tr>
<th>Site #</th>
<th>Site Name</th>
<th>High Abundance of Fish</th>
<th>Elevated Catch Rates</th>
<th>Hydrated Females Collected</th>
<th>Running-Ripe Males Collected</th>
<th>Courtship or Spawning Observed</th>
<th>Gravid Females Observed</th>
<th>Courtship Coloration Observed</th>
<th># Criteria observed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Punta Refugio</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>7</td>
</tr>
<tr>
<td>2</td>
<td>Punta Diablo</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>3</td>
<td>La Ventana</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>7</td>
</tr>
<tr>
<td>4</td>
<td>Chorros</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>5</td>
<td>San Pedro Martir Island</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>6</td>
</tr>
<tr>
<td>6</td>
<td>San Lorenzo Island</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>4</td>
</tr>
<tr>
<td>7</td>
<td>Las Cuevitas</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>4</td>
</tr>
<tr>
<td>8</td>
<td>Puerto Libertad</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>4</td>
</tr>
<tr>
<td>9</td>
<td>Datil Island</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>7</td>
</tr>
<tr>
<td>10</td>
<td>La Tordilla</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>4</td>
</tr>
<tr>
<td>11</td>
<td>El Tecomate</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>4</td>
</tr>
<tr>
<td>12</td>
<td>San Esteban Island</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>3</td>
</tr>
<tr>
<td>13</td>
<td>Punta Roja</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>6</td>
</tr>
<tr>
<td>14</td>
<td>Los Machos</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>6</td>
</tr>
<tr>
<td>15</td>
<td>Puerto Lobos</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>4</td>
</tr>
<tr>
<td>16</td>
<td>La Poma</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>3</td>
</tr>
<tr>
<td>17</td>
<td>San Francisquito</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✓</td>
<td>4</td>
</tr>
<tr>
<td># Sites</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>17 13 11 12 6 12 11</td>
</tr>
</tbody>
</table>
Table 2 Molecular diversity. Number of samples (n), number of haplotypes (nH), haplotype diversity (h), corrected haplotype diversity (hilda) and nucleotide diversity (π).

<table>
<thead>
<tr>
<th>Location</th>
<th>n</th>
<th>nH</th>
<th>h     ± 0.035</th>
<th>hilda ± 0.001</th>
<th>π     ± 0.001</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. La Ventana*</td>
<td>52</td>
<td>23</td>
<td>0.870</td>
<td>0.885</td>
<td>0.002</td>
</tr>
<tr>
<td>5. San Pedro Martir Island*</td>
<td>17</td>
<td>9</td>
<td>0.860</td>
<td>0.881</td>
<td>0.001</td>
</tr>
<tr>
<td>6. San Lorenzo Island*</td>
<td>11</td>
<td>4</td>
<td>0.673</td>
<td>0.844</td>
<td>0.001</td>
</tr>
<tr>
<td>8. Puerto Libertad</td>
<td>55</td>
<td>21</td>
<td>0.874</td>
<td>0.887</td>
<td>0.002</td>
</tr>
<tr>
<td>9. Datil Island</td>
<td>20</td>
<td>13</td>
<td>0.947</td>
<td>0.919</td>
<td>0.002</td>
</tr>
<tr>
<td>11. El Tecomate</td>
<td>26</td>
<td>16</td>
<td>0.935</td>
<td>0.914</td>
<td>0.002</td>
</tr>
<tr>
<td>12. San Esteban Island</td>
<td>4</td>
<td>4</td>
<td>1.000</td>
<td>0.943</td>
<td>0.001</td>
</tr>
<tr>
<td>15. Puerto Lobos</td>
<td>11</td>
<td>7</td>
<td>0.909</td>
<td>0.902</td>
<td>0.002</td>
</tr>
<tr>
<td>16. La Poma</td>
<td>16</td>
<td>12</td>
<td>0.950</td>
<td>0.920</td>
<td>0.002</td>
</tr>
<tr>
<td>17. San Francisquito*</td>
<td>23</td>
<td>17</td>
<td>0.949</td>
<td>0.920</td>
<td>0.002</td>
</tr>
</tbody>
</table>

* Indicates sites within MPAs.
Table 3  Pairwise $F$ statistics between sites. Pairwise $F_{ST}$ values are above diagonal and pairwise $\phi_{ST}$ values are below diagonal. Values in bold are statistically significant ($P < 0.05$).

<table>
<thead>
<tr>
<th>Location</th>
<th>3</th>
<th>5</th>
<th>6</th>
<th>8</th>
<th>9</th>
<th>11</th>
<th>12</th>
<th>15</th>
<th>16</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>3. La Ventana</td>
<td>-</td>
<td>0.1342</td>
<td>0.2089</td>
<td>0.1278</td>
<td>0.0939</td>
<td>0.0988</td>
<td>0.0837</td>
<td>0.1130</td>
<td>0.0935</td>
<td>0.0929</td>
</tr>
<tr>
<td>5. San Pedro Martir Island</td>
<td>-0.0122</td>
<td>-</td>
<td>0.2262</td>
<td>0.1321</td>
<td>0.0955</td>
<td>0.1006</td>
<td>0.0868</td>
<td>0.1167</td>
<td>0.0951</td>
<td>0.0943</td>
</tr>
<tr>
<td>6. San Lorenzo Island</td>
<td>0.0381</td>
<td>0.0869</td>
<td>-</td>
<td>0.2064</td>
<td>0.1776</td>
<td>0.1800</td>
<td>0.2039</td>
<td>0.2091</td>
<td>0.1803</td>
<td>0.1747</td>
</tr>
<tr>
<td>8. Puerto Libertad</td>
<td>0.0912</td>
<td>0.1189</td>
<td>-0.0094</td>
<td>-</td>
<td>0.0920</td>
<td>0.0969</td>
<td>0.0812</td>
<td>0.1109</td>
<td>0.0915</td>
<td>0.0910</td>
</tr>
<tr>
<td>9. Datil Island</td>
<td>-0.0001</td>
<td>-0.0018</td>
<td>-0.0090</td>
<td>0.0408</td>
<td>-</td>
<td>0.0587</td>
<td>0.0323</td>
<td>0.0706</td>
<td>0.0513</td>
<td>0.0520</td>
</tr>
<tr>
<td>11. El Tecomate</td>
<td>0.0101</td>
<td>-0.0094</td>
<td>0.0621</td>
<td>0.1234</td>
<td>0.0169</td>
<td>-</td>
<td>0.0401</td>
<td>0.0767</td>
<td>0.0576</td>
<td>0.0581</td>
</tr>
<tr>
<td>12. San Esteban Island</td>
<td>-0.0421</td>
<td>-0.0391</td>
<td>0.0772</td>
<td>0.0386</td>
<td>-0.0605</td>
<td>0.0203</td>
<td>-</td>
<td>0.0542</td>
<td>0.0915</td>
<td>0.0317</td>
</tr>
<tr>
<td>15. Puerto Lobos</td>
<td>0.0567</td>
<td>0.0555</td>
<td>-0.0353</td>
<td>0.0122</td>
<td>-0.0035</td>
<td>0.0433</td>
<td>0.0193</td>
<td>-</td>
<td>0.0696</td>
<td>0.0697</td>
</tr>
<tr>
<td>16. La Poma</td>
<td>-0.0124</td>
<td>-0.0196</td>
<td>0.0725</td>
<td>0.0891</td>
<td>-0.0169</td>
<td>0.0038</td>
<td>-0.0578</td>
<td>0.0554</td>
<td>-</td>
<td>0.0507</td>
</tr>
<tr>
<td>17. San Franciscohito</td>
<td>-0.0045</td>
<td>-0.0055</td>
<td>0.0193</td>
<td>0.0601</td>
<td>-0.0129</td>
<td>0.0092</td>
<td>-0.0455</td>
<td>0.0193</td>
<td>-0.0176</td>
<td>-</td>
</tr>
</tbody>
</table>
Table 4 Probability of three larval dispersal models. Bayes factors and marginal log likelihoods for proposed larval dispersal models estimated in Migrate-n version 3.2.16 using Bayesian approximation and thermal integration.

<table>
<thead>
<tr>
<th>Model</th>
<th>Bezier lML</th>
<th>Harmonic lML</th>
<th>Choice (Bezier)</th>
<th>Model probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>Full Matrix</td>
<td>-2673.20</td>
<td>-2514.33</td>
<td>2</td>
<td>0.00001</td>
</tr>
<tr>
<td>Baja to</td>
<td>-2658.64</td>
<td>-2513.07</td>
<td>1</td>
<td>0.99999</td>
</tr>
<tr>
<td>Mainland</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mainland to</td>
<td>-2974.33</td>
<td>-2529.21</td>
<td>3</td>
<td>0.00000</td>
</tr>
<tr>
<td>Baja</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


Figure 1 Area of study around the Midriff Islands in the Gulf of California. Study sites, in which modeled and genetic connectivity was measured (Release sites with genetic samples) and those only for modeled connectivity (Release sites only), including marine protected areas, no-take zones and fishing zones identified by interviews with fishers (Moreno-Báez et al. 2012, 2010).
**Figure 2** Spawning season and period. Monthly proportion of actively spawning and spawning capable females (left $y$ axis) and female gonadosomatic index (GSI, right $y$ axis) for the Midriff Islands collected in 2009.
Figure 3 Probability of larval density. Density is shown for larvae only in coastal areas after PLD 14 days (A), 21 days (B) and 28 days (C) in each spatial unit of analysis.
Figure 4 Larval dispersal from each spawning site. Maps showing the spatial distribution of larvae released in May 16th (A) and May 25th (B) at each of 17 sites ("X" in each box) for PLD 14, 21 and 28 days. Scale in kilometers.
Figure 5 Modeled networks of larval connectivity. Spatial networks of larval dispersal between sites for PLD 14 days (A), 21 days (B) and 28 days (C), showing dispersal events (links) between sites (nodes). Line width is proportional to probability, according to the scale to the right. The direction of the larval dispersal events is indicated by different colors: northward (red), southward (blue) or both simultaneously (green). Open nodes are sites within MPAs, solid nodes are fished sites.
Figure 6 Performance of 17 sites for different aspects of marine connectivity.

Performance was measured with export probability, import probability, local-retention probability, out-degree, in-degree and betweenness centrality, as estimated by an oceanographic model at PLD 14 and 28 days.
Figure 7 Minimum spanning network among haplotypes. The network shows the relationships among 77 haplotypes found in *M. rosacea*. Circles are sized proportionally to the number of individuals that possess each haplotype and colors indicate their geographic distribution in 10 sites shown to the left. All haplotypes are separated by one to four mutation steps as denoted by scaling provided.
**Figure 8** Correlations between genetic diversity and characteristics of modeled spatial networks derived from graph theory. Correlation between corrected haplotype diversity and properties of each site calculated from the modeled networks of larval dispersal, including: in-degree (A) and betweenness centrality (B). In both, the coefficient of correlation ($R^2$) and the respective P values are shown for all sites (excluding San Lorenzo Island). Numbers refers to localities as in Table 1.