

**Centro de Investigación Científica y de Educación  
Superior de Ensenada, Baja California**



---

**Doctorado en Ciencias  
en Ecología Marina**

---

**Integrative study of connectivity in the Gulf of California  
seascape**

Tesis  
para cubrir parcialmente los requisitos necesarios para obtener el grado de  
Doctora en Ciencias

Presenta:

**Tania Valdivia Carrillo**

Ensenada, Baja California, México  
2022

Tesis defendida por  
**Tania Valdivia Carrillo**

y aprobada por el siguiente Comité

---

Dr. Axayácatl Rocha Olivares  
Codirector de tesis

---

Dr. Adrián Munguía Vega  
Codirector de tesis

Miembros del comité

**Dra. Asunción Lago-Lestón**

**Dr. Silvio Guido Marinone-Moschetto**

**Dr. Héctor Reyes-Bonilla**

**Dr. Oscar Sosa-Nishisaki**



---

**Dr. Rafael Andrés Cabral Tena**  
Coordinador del Posgrado en Ecología Marina

---

**Dra. Pedro Negrete Regagnon**  
Director de Estudios de Posgrado

Resumen de la tesis que presenta **Tania Valdivia Carrillo** como requisito parcial para la obtención del grado de Doctor en Ciencias en Ecología Marina

### **Integrative study of connectivity in the Gulf of California seascape**

Resumen aprobado por:

---

Dr. Axayácatl Rocha Olivares  
Codirector de tesis

---

Dr. Adrián Munguía Vega  
Codirector de tesis

El Golfo de California (GC) es un *hotspot* de biodiversidad marina y una región prioritaria para la conservación. En sus ecosistemas de arrecifes rocosos habitan diversas poblaciones y comunidades de peces de importancia pesquera. Éstas se han visto amenazadas debido a factores como la sobrepesca y el cambio climático, con afectaciones potenciales como la modificación de sus patrones de conectividad. La conectividad es una propiedad del paisaje cuyo estudio se enfoca en la identificación de las características funcionales (biológicas) y estructurales (físicas) las cuales permiten la dispersión de los individuos a través del espacio. Su estudio ha sido fundamental en ecología y evolución ya que determina la prevalencia de la biodiversidad a través del tiempo. Conocer su estado actual, nos permite contrastar futuros cambios en estos ecosistemas y plantear soluciones para su conservación. Los objetivos de la presente tesis fueron i) evaluar los patrones de conectividad funcional entre poblaciones y comunidades de peces asociados a los arrecifes rocosos del GC, ii) evaluar la conectividad estructural del paisaje marino del GC; y iii) evaluar la relación entre estas dos conectividades. Esto se logró mediante el análisis de un conjunto de relaciones en las que la conectividad estructural juega un papel relevante en la dispersión de los peces y, por lo tanto, en sus valores de diversidad y diferenciación, a diferentes niveles taxonómicos. Para lograr este objetivo, se analizó la diversidad y diferenciación genómica poblacional de la cabrilla sardinera (*Mycteroperca rosacea*) utilizando polimorfismos de un solo nucleótido. Además, se estimó la diversidad alfa y beta en comunidades locales de peces arrecifales, caracterizándolas con dos métodos de monitoreo complementarios (censos visuales submarinos y metabarcoding de ADN ambiental). También se evaluó la conectividad potencial demográfica entre las poblaciones y comunidades locales, usando un modelo numérico oceanográfico (HAMSOM). Asimismo, se evaluó la conectividad estructural mediante la estimación de las distancias geográficas, ambientales y oceanográficas entre los sitios estudiados, empleando sistemas de información geográfica, bases de datos de variables ambientales, y el modelo oceanográfico. Finalmente, se integró la información sobre conectividad funcional y estructural utilizando métodos estadísticos espaciales y análisis de redes. Los resultados evidenciaron que la cabrilla sardinera presenta una alta conectividad funcional sin diferenciación genómica dentro del GC y una diversidad genómica media. Esto indica la presencia de una metapoblación dentro del golfo, mantenida por una alta conectividad demográfica y por la dinámica de corrientes marinas. En esta especie, los patrones de aislamiento por distancias geográficas, ambientales, y de resistencia ocurren de manera diferencial en cada región de GC. Además, su conectividad potencial demográfica se encuentra geográficamente regionalizada y presenta cambios estacionales, lo cual promueve la dispersión de larvas y el flujo de genes en un entorno oceanográficamente y geomorfológicamente complejo. A nivel de comunidad, los resultados mostraron que el GC presenta dos metacomunidades de peces arrecifales (Norte y Centro) con una composición específica determinada. En ellas, el decaimiento de la similitud entre comunidades se encuentra determinada por las distancias geográficas (en todo el GC y región Central), ambientales (en todo el GC y región Norte) y de resistencia (en todo el GC y en cada región). Finalmente, el análisis de redes mostró que la centralidad de los sitios dentro de la red está relacionada con la diversidad genómica y la diversidad alfa en poblaciones y comunidades de peces del GC.

**Palabras clave:** conectividad, genómica poblacional, ecología de comunidades, dispersión larvaria, modelo numérico oceanográfico, análisis de redes

Abstract of the thesis presented by **Tania Valdivia Carrillo** as a partial requirement to obtain the Doctor of Science degree in Marine Ecology

### **Integrative study of connectivity in the Gulf of California seascape**

Abstract approved by:

---

Dr. Axayácatl Rocha Olivares  
Codirector de tesis

---

Dr. Adrián Munguía Vega  
Codirector de tesis

The Gulf of California (GC) is a marine biodiversity hotspot and a priority region for conservation. Its rocky reef ecosystems are inhabited by diverse populations and communities of important fishing species. These have been threatened due to overfishing and climate change, with potential effects such as modifying their connectivity patterns. Connectivity is a seascape emergent property whose study focuses on identifying functional (biological) and structural (physical) characteristics that allow the dispersion of individuals through space. Its study has been fundamental in ecology and evolution since it determines the prevalence of biodiversity over time. Knowing their current state allows us to contrast future changes in these ecosystems and propose solutions for their conservation. The objectives of this thesis were: i) to evaluate functional connectivity patterns between populations and communities of fishes associated with the rocky reefs of the GC; ii) to evaluate the structural connectivity of the GC seascape; iii) to evaluate the relationship between these two connectivities. This was achieved by analyzing relationships in which structural connectivity plays a relevant role in fish dispersal and, therefore, in their diversity and differentiation values, at different taxonomic levels. To achieve this goal, the population genomic diversity and differentiation of the leopard grouper (*Mycteroperca rosacea*) were analyzed using single nucleotide polymorphisms. In addition, alpha and beta diversity in local reef fish communities was estimated, characterizing them with two complementary monitoring methods (underwater visual censuses and environmental DNA metabarcoding). Potential demographic connectivity between local populations and communities was assessed using a numerical oceanographic model (HAMSOM). Likewise, structural connectivity was evaluated by estimating the geographic, environmental, and oceanographic distances among the studied sites, using geographic information systems, databases of environmental variables, and the oceanographic model. Finally, functional and structural connectivity information was integrated using spatial statistical methods and network analysis. The results showed that the leopard grouper has high functional connectivity without genomic differentiation within the CG and a medium genomic diversity. This indicates the presence of a metapopulation within the Gulf, maintained by high demographic connectivity and by the dynamics of ocean currents. In this species, isolation patterns due to geographic, environmental, and resistance distances occur differentially in each GC region. In addition, its potential demographic connectivity is geographically regionalized and presents seasonal changes, which promotes larval dispersal and gene flow in an oceanographically and geomorphologically complex environment. The results showed that the CG presents two reef fish metacommunities (North and Center) with a specific composition. In them, the decay of the similarity between communities is determined by the geographical (in the entire GC and Central region), environmental (in the entire GC and North region), and resistance distances (in the entire GC and each region). Finally, network analysis showed that the site's centrality within the network is related to genomic diversity and alpha diversity in fish populations and communities in the GC.

**Keywords:** connectivity, population genomics, community ecology, larval dispersal, oceanographic numerical model, network analysis

## Dedication

*A Lulú, Rafael, Eréndira y Arturo*

*Gracias por todas las maneras en las que me ayudaron durante la realización  
de esta tesis.*

*Como siempre, han sido parte fundamental del camino.*

## Acknowledgements

Thanks to Consejo Nacional de Ciencia y Tecnología (CONACyT) for the Ph.D. scholarship granted (CVU 290839) and to the CICESE for the three-month support (October-December 2020) to finish this thesis.

I first want to thank my co-advisors: Dr. Axayácatl Rocha-Olivares and Dr. Adrián Munguia-Vega for their openness and curiosity in letting me pursue the ambitious objectives in this thesis, and for the time and support in all the intellectual and academic aspects for the development of this work. I also want to thank the members of my thesis committee for the support during the development of my doctoral studies and the completion of this thesis: Dra. Asunción Lago-Lestón, Dr. Guido Marinone Moschetto, Dr. Oscar Sosa Nishisaki, Dr. Héctor Reyes Bonilla. Thanks for your time, comments and orientation.

Thanks to the past and current Coordinator of the Marine Ecology postgraduate school, Dr. Rafael Cabral Tena and Dr. Jorge Rosales Casián for the administrative support during my doctorate; to the Postgraduate Studies Direction personnel: Dolores Sarracino, Norma Alicia Fuentes, and Citlali Romero; to the Library personnel, principally to Elizabeth Avilés; to the Telematic and videoconference technician José Francisco J. Magallanes; finally, thanks to the technician of the Molecular Ecology Lab at CICESE, M.C. Ivonne Martínez, for the support in the administrative procedures and logistics for buying some reagents and other aspects related to my stay in CICESE.

The second chapter might not have been possible without the help of the fishers from Bahía Kino and Puerto Libertad: Jaime Roberto Mejía Pereira, Rosario Solano, Samuel Gerardo Delgado, Mario Alberto Robles, Flaviano Pelayo Camacho, Jesús Alfredo López, José Gerardo Nieto Chávez, José Alfredo López Hermosillo, Jesús Rafael López Hemosillo, Ramón Camacho; from Puerto Lobos: Elvia and Luis Arturo; from Puerto Peñasco: Isidro; from Santa Rosalía: hermano del Cachas; from Loreto – La Paz corridor: Jaime de la Toba, Felipe González Amador, Ismael, Marcelo Amador Cuevas; and from Los Failes: Luis Castro, for the obtention of the leopard grouper tissue samples. Dr. Axayácatl Rocha, Dr. Héctor Reyes and Dr. Adrián Munguia for the support in reagent acquisition and sequencing for the cabrilla samples.

For helping the development of the third chapter of this thesis I thank the support from Consejo Nacional de Ciencia y Tecnología, México (institutional fund CONACyT Fronteras de la Ciencia 292/2016), The University of Arizona—CONACyT Binational Consortium for the Regional Scientific Development and Innovation (CAZMEX), Comunidad y Biodiversidad A.C., The Nature Conservancy Mexico, Pronatura Noroeste A.C. and Ecology Project International - La Paz. Adrian and I received a grant from Langebio - Biotech del Norte for Illumina sequencing of the eDNA metabarcoding samples.

I thank Manuel Alejandro Olán González, Damien Olivier, Jaime de la Toba, Omar Valencia, and José Torres for helping with the underwater visual censuses of fishes, and Eldridge Wisely who played a key role during seawater sampling and filtering and in the first steps of the metabarcoding adventure. Eric Ramírez helped with DNA extractions for the 12S reference database. Also, Francisco Domínguez for the edition of the 12S sequences that comprised the local reference database. I thank the invaluable logistical support in the field from the crew of the DV Quino el Guardian and appreciate the drawings of fish from the Gulf of California originally made by Juan Chuy and provided by Amy Hudson Weaver from Sociedad de Historia Natural Niparajá A.C.

Thanks to Dr. Fausto Valenzuela, Dr. Noé Díaz, Dr. Adrián Munguía, Dr. David Paz, Dra. Carolina Galván and Dra. Claudia Silva to help with the organization of the “Eukaryotic metabarcoding course” which occurred in La Paz just before the COVID lockdown. To the people from Latin America that I met during the Eukaryotic metabarcoding course in Costa Rica and the Metabarcoding School in Colombia who motivated me in pursuing metabarcoding knowledge, and promote the application of DNA technologies principally in Latin American countries.

Thanks to the former members of the Molecular Ecology Laboratory at CICESE principally Ana Castillo, Celia Bisbal, Alejandro Cisterna, and Pablo Guevara. Also to the 2016 generation of the Master in Marine Ecology principally Alesa Flores, Aurora Gaona, and Abril Heredia for making the courses more fun and for taking me birdwatching early in the mornings. Also, I want to thank Alesa Flores, Roberto Cortés, and Meliza Le for giving me shelter in their respective houses when I needed it, and for the food, beers, and friendship.

I also want to thank to the former and new members of the Applied genetics Lab in La Paz: Francisco Domínguez, Camila McLoughin, Blanca González, Karla Botello, Nicole Reguera, Manuel Morales, Veronica Mendoza, Paty Cerrillo, and the people who I met during the Quino trips, principally to Lazuli Piceno.

I want to thank my friends Fausto Valenzuela, Paulina Mejía, David Paz, Carina Gutierrez, Carolina Galván, Carmen Blázquez , Etna Sánchez and Andrea Cruz for their friendship and support.

# Table of contents

Abstract in spanish.....	ii
Abstract.....	iii
Dedication.....	iv
Acknowledgments.....	v
List of figures .....	xi
List of tables .....	xvi

## **Chapter. 1 Introduction..... 1**

1.1 Background .....	2
1.1.1 Contextualizing connectivity in the rocky reef ecosystems .....	2
1.1.2 Metapopulation and metacommunity connectivity .....	4
1.1.3 Direct and indirect methods for the study of connectivity in rocky reef ecosystems...6	
1.1.3.1 Seascape genomics: evaluation of the metapopulation's genomic connectivity and its relationship with the seascape environmental characteristics .....	6
1.1.3.2 Mechanisms of genomic isolation in metapopulations .....	8
1.1.3.3 Metapopulation demographic connectivity: estimation of potential larval dispersal among local populations using oceanographic models .....	10
1.1.3.4 Metacommunity connectivity: from species detection to species turnover evaluation .....	14
1.1.3.5 Mechanisms of isolation in metacommunities .....	15
1.1.3.6 Metacommunity demographic connectivity: estimation of potential larval dispersal of reproductive assemblages using oceanographic models .....	16
1.1.3.7 Graph theory and network analyses as a methodological approach for integrating functional and structural connectivity information .....	18
1.2 Justification .....	22
1.3 Hypotheses.....	23
1.4 Objectives.....	26
1.4.1 General objective .....	26
1.4.2 Specific objectives .....	26
1.5 Study objects.....	27
1.5.1 The Gulf of California seascape .....	27
1.5.2 The leopard grouper .....	28
1.5.3 Rocky reefs of the Gulf of California and its associated fishes.....	30

## **Chapter. 2 Methodology .....33**

2.1 Evaluation of the functional connectivity at population and community organization levels. ....	33
2.1.1 Population-level functional connectivity: evaluation of the genomic diversity and differentiation ( $F_{ST}$ ) in the leopard grouper ( <i>Mycteroperca rosacea</i> ) through the analysis of Single Nucleotide Polymorphisms (SNPs).....	33
2.1.1.1 Tissue sampling for population genomics analysis .....	33
2.1.1.2 Genomic DNA extraction and ddRADseq library construction.....	33

2.1.1.3	Bioinformatic analyses: SNP calling.....	36
2.1.1.4	Bioinformatic analyses: Data filtering and identification of neutral and candidate SNPs loci .....	36
2.1.1.5	Geographical distribution of the genomic diversity and evaluation of the population genomic differentiation .....	37
2.1.2	Community-level functional connectivity: estimation of the alpha (species richness) and beta diversities (species turnover) in communities of rocky reef fishes using complementary monitoring methods. ....	38
2.1.2.1	Underwater visual censuses and seawater sampling for eDNA metabarcoding analysis .....	38
2.1.2.2	Custom reference database for the ichthyofauna of the GC: primer design, PCR amplification, and sequencing .....	40
2.1.2.3	Mock community. ....	41
2.1.2.4	Environmental DNA extraction .....	42
2.1.2.5	Library preparation and high-throughput sequencing.....	42
2.1.2.6	Metabarcoding bioinformatic analyses.....	42
2.1.2.7	OTUs taxonomic assignment.....	43
2.1.2.8	Community-level ecological analyses.....	43
2.2	Structural connectivity: estimation of abiotic predictors and testing for mechanisms of isolation.....	44
2.2.1	Spatial and environmental predictors of functional connectivity .....	44
2.2.2	Oceanographic predictors of functional connectivity and network analyses .....	45
2.2.2.1	The HAMSOM numerical model, obtention of the potential larval connectivity matrices and the estimation of the resistance distance .....	45
2.2.3	Testing for mechanisms of isolation: integrating population genomic and community connectivity patterns and the seascape predictors .....	48
2.2.3.1	Mantel test and multiple regression on distance matrices (Link-level analyses)....	48
2.2.3.2	Network analysis of connectivity data (node-level analysis) .....	48
<b>Chapter. 3</b>	<b>Results .....</b>	<b>51</b>
3.1	Evaluation of the functional connectivity at population and community organization levels. ....	51
3.1.1	Population level functional connectivity: evaluation of the genomic diversity and differentiation ( $F_{ST}$ ) in the leopard grouper ( <i>Mycteroperca rosacea</i> ) through the analysis of Single Nucleotide Polymorphisms (SNPs).....	51
3.1.1.1	Bioinformatic analyses: SNP calling.....	51
3.1.1.2	Bioinformatic analyses: Data filtering and identification of neutral and candidate SNPs loci .....	51
3.1.1.3	Geographical distribution of the genomic diversity and evaluation of the population genomic differentiation .....	52
3.1.2	Community-level functional connectivity: estimation of the alpha (species richness) and beta diversities (species turnover) in communities of rocky reef fishes using complementary monitoring methods. ....	56
3.1.2.1	Underwater visual censuses species detection.....	56

3.1.2.2	Custom reference for the ichthyofauna of the GC: primer design, PCR amplification, and sequencing .....	57
3.1.2.3	eDNA metabarcoding sequencing statistics, OTUs identification in environmental samples, and taxonomic assignment .....	57
	Locality name .....	58
3.1.2.4	List of species/OTUs detected with both survey methods.....	59
3.1.2.5	Mock community and negative control .....	60
3.1.2.6	Community-level ecological analyses.....	61
3.2	Structural connectivity: estimation of abiotic predictors and testing for mechanisms of isolation.....	62
3.2.1	Spatial and environmental predictors of functional connectivity .....	62
3.2.2	Oceanographic predictors of functional connectivity and network analyses .....	66
3.2.3	Testing for mechanisms of isolation: integrating population genomic and community connectivity patterns and the seascape predictors .....	73
3.2.3.1	Mantel test and multiple regression on distance matrices (link-level analyses) ....	73
<b>Chapter. 4</b>	<b>Discussion .....</b>	<b>87</b>
4.1	High levels of functional connectivity in the leopard grouper local populations are related to the Euclidean, environmental, and resistance distances at different spatial scales. ....	87
4.2	Metacommunity characterization and the evaluation of connectivity and its seascape determinants.....	93
<b>Chapter. 5</b>	<b>Conclusions.....</b>	<b>99</b>
<b>Cited literature</b>	<b>.....</b>	<b>101</b>
<b>Annexes</b>	<b>.....</b>	<b>133</b>
Annex A	.....	133
Annex B	.....	134

## List of figures

Figure	Page
1	Some of the biological characteristics of marine fishes and the physical and environmental attributes of the seascape (bottom) influence demographic and genomic processes throughout the teleost fishes' lifecycle and determine the distribution of genomic variation within and among populations. Hypothetical source populations are shown as independent bands (top left) where bandwidth can be understood as the number of individuals or genomic variation. The bottom panel underscores the period in which the biological or seascape characteristics are most relevant. Reproductive output of the source population varies as a function of population size and timing of reproduction. During dispersal (from left to right), larvae can be advected by currents, mixed (crossed lines), and may become diluted or concentrated, represented by bandwidths. Larval dispersal potential is related to a species' early life-history traits, such as pelagic larval duration. Larval behaviors such as kin aggregation and active swimming can increase or counter physical oceanography to concentrate larvae. Bandwidths become smaller, depicting larval mortality due to the environment of the ocean (biotic and abiotic). Successful settlement of larvae into a hypothetical destination population (top right) is dependent on habitat suitability. Post-settlement survival in the population depends on the condition of the recruit, competition, and other factors (Figure taken from Liggins et al., (2019)). .....4
2	The conceptual figure shows how different organization levels are integrated within the connectivity framework across two axes: scale and organizational level. Populations are embedded within local communities, including regional metapopulations and regional metacommunities, respectively (Chase et al., 2020). .....5
3	Conceptual framework of the seascape genomics approach (Figure taken from Laura Benestan's Seascape genomics presentation September 2020). .....7
4	Comparison of open to closed population continuum. The degree to which a set of local populations (or nodes) might be connected to each other via larval dispersal can be expressed as a connectivity matrix where nodes are distributed along both the vertical and horizontal axes. In the fully open case (left panel), where all populations within a given domain contribute propagules equally to each other, all cells within the connectivity matrix are populated similarly ( $1/n$ where $n$ is the total number of local populations, i.e., source locations). This extreme case represents what might be expected if all propagules were equally mixed and then settled at the random back to the source sites. In the wholly closed case (right panel), all propagules return directly to their source location, all cells along the diagonal would be populated with a probability of 1 (i.e., 100% self-recruitment), and all other non-diagonal cells would be populated by zero. This case would be expected if all populations were utterly isolated from any different location through dispersal processes. The intermediate situation is represented by a matrix constructed from biophysical modeling data with some variation of probabilities along with both the diagonal and non-diagonal cells. The scale represents a range in cumulative abundance of successful recruits, where red/orange represents ecologically significant levels of exchange and green/blue represents very low levels of genetically relevant exchange (Image taken from Cowen et al. 2006). ....11
5	Larval dispersal in the GC for species spawning during summer (a, August) and fall (b, October) with PLD of 28 days, based on a three dimensional oceanographic model where larvae were released within each of eight coastal polygons shown in bold. Color shows the depth of virtual larvae (Taken from Munguia-Vega et al. (2018)). .....13

6	Top: a simple unweighted graph. Bottom: a weighted graph with nodes attributed with their sizes, arrows indicating the direction of the flow, and the line thickness indicating the magnitude of the fluxes (Image taken from Urban et al., 2009). .....	19
7	A. A graph-theoretic illustration of marine connectivity. Nodes represent reef habitats within the graph framework. When larvae from a source reef reach a downstream reef site, a dispersal connection is made. This dispersal connection and direction are represented by an arrow or 'edge' within the graph. The arrow's thickness reflects the strength of connection, i.e., the proportion of larvae that moves between the two nodes (from Treml et al., (2008)). B. Left panel: A at the node level, the allelic richness or species richness ( $Y$ ) at each sampling location $a$ , $b$ or $c$ are related to environmental conditions or landscape features ( $X$ ) observed at the exact location. Nodes of different sizes refer either to different richness values. Right panel: at the link level, genetic distance, or species spatial turnover $D_Y$ between pairs of sampling locations $ab$ , $ac$ , and $bc$ is related to distance-based landscape data $D_X$ describing the intervening matrix along each link (Modified from Wagner and Fortin, 2013). .....	20
8	Illustration for hypothesis 0. ....	24
9	Illustration for hypothesis 1. ....	24
10	Illustration for hypothesis 2. ....	25
11	Illustration for hypothesis 3. ....	25
12	Leopard grouper ( <i>Mycteroperca rosacea</i> ) life-cycle (Image from <a href="https://datamares.org/perfil-de-especie-cabrilla">https://datamares.org/perfil-de-especie-cabrilla</a> ).....	29
13	Faunistic regions: Northern, Central, and Southern are shown. Also, known rocky reefs (adult habitat of reef fishes) and sargassum beds (recruitment habitat for <i>M. rosacea</i> and other rocky reef fishes), and marine protected areas of the GC. Principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN). Data obtained from Munguia-Vega et al., 2018. ....	31
14	Leopard grouper sampling of sites for population analyses (see <b>Table 4</b> for site-specific details) in the GC (red dots) and the corresponding polygons from the HAMSOM model (yellow polygons) are represented. GC regions are labeled and delimited with dashed lines. Principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN). ....	34
15	Sampling sites for UVC and seawater collection for eDNA metabarcoding across the GC (red dots) and the corresponding polygons from the HAMSOM model (yellow polygons) are represented. Northern and Central regions are labeled and delimited with dashed lines. Principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN). ....	40
16	Fifty-nine polygons used in the HAMSOM numerical model are represented in green, and polygons with sampling sites at the population or community level are red. Principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN). ....	48

17	Boxplots showing genomic diversity metrics per GC region. Ar: rarefied allelic richness, He: expected heterozygosity, Ho: observed heterozygosity, and $F_{IS}$ : Wright's inbreeding coefficient. ....	53
18	Maps showing the geographical distribution of the genomic diversity of the leopard grouper A: Ho (observed heterozygosity), B: He (expected heterozygosity), C: A (allelic richness), D: $F_{IS}$ (inbreeding coefficient). Principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN).....	55
19	Discriminant Analysis of Principal Components (DAPC) results of the 985 SNP neutral dataset obtained from the leopard grouper population analysis. Figures show the scatterplot with a priori sampling location information (Panel A) and density plots for (Panel B) the first and (Panel C) the second discriminant axes. Map (Panel D) shows the geographical distribution of the sampling localities and the GC regions. The principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN).....	56
20	(A) Number of observations per taxonomic level. (B) Species accumulation curves with confidence intervals (95%), $S_{exp}$ is showed in dashed lines for each monitoring method. (C) Boxplot of the mean observed number of species ( $S_{obs}$ ) per site in each survey method. (D) Mean Sobs per method and per region of the Gulf of California. In all graphs, UVC is shown in yellow; eDNA is shown in blue, and both methods are shown in white.....	59
21	Families detected with each monitoring method. UVC: 1. Sparidae; 2. Zanclidae; 3. Scorpaenidae; 4. Labrisomidae; 5. Tetraodontidae; 6. Muraenidae; 7. Chaetodontidae; 8. Synodontidae; 9. Bleniidae; 10. Chaenopsidae; 11. Tripterygiidae; 12. Holocentridae. Both methods: 13. Carangidae; 14. Scaridae; 15. Lutjanidae; 16. Gobiidae; 17. Serranidae; 18. Fistulariidae; 19. Balistidae; 20. Kyphosidae; 21. Elopidae; 22. Sphyraenidae; 23. Mullidae; 24. Pomacanthidae; 25. Haemulidae; 26. Sciaenidae; 27. Acanthuridae; 28. Pomacentridae; 29. Diodontidae; 30. Cirrhitidae; 31. Labridae; 32. Apogonidae. eDNA: 33. Istiophoridae; 34. Scombridae; 35. Paralichthyidae; 36. Nomeidae; 37. Mugilidae; 38. Clupeidae (drawings by Juan Chuy).....	60
22	Geographical distribution of the number of species (Sobs) detected at each site with each survey method: UVC (left) and eDNA metabarcoding (right). Northern and Central regions are labeled and delimited with dashed lines. Principal cities shown including BA: Bahía de los Ángeles, SR: Santa Rosalía, L: Loreto, LP: La Paz, BK: Bahía de Kino. ....	62
23	Non-metric multidimensional scaling (nMDS) ordinations of community data using the Jaccard similarity index and $k=2$ for (A) underwater visual census (UVC) and (B) marine environmental DNA (eDNA). Red and Black dots indicate Northern and Central groups, respectively. ....	63
24	Heatmaps of the 14 environmental variables from the population (left panel) and community sampling sites (right panel). Columns and rows are separated per North, Central and South region with white lines.....	63
25	Above: PCA biplot of evaluating the 14 environmental variables of the population sampling sites. The color of the arrows indicates the variable contribution to each PC; the shape size indicates its contribution to each PC, the color of the shape indicates the region of the GC, and the shades ellipses were drawn to group the North and Central region localities. Below: Scree-plot of components contribution from the PCA analyses of the population sampling sites. ....	64

26	Above: PCA biplot of evaluating the 14 environmental variables of the community sampling sites. The color of the arrows indicates the variable contribution to each PC; the shape size indicates its contribution to each PC, the color of the shape indicates the region of the GC, and the shaded ellipses were drawn to group the North and Central region localities. Below: Scree-plot of components contribution from the PCA analyses of the community sampling sites. ....	65
27	March-June (Spring) (left) and June-September (Summer) (right) seasons. Nodes are colored according to the GC region they belong to North (orange), Central (green), or South (blue). Arrowheads indicate the direction of the larval dispersal, and the width of the arrow line means the proportion of larvae being dispersed. Below: Adjacency matrices of potential larval connectivity of the 22 leopard grouper's localities in the GC for four of the Spring (March-June) (left) and Summer (June-September) seasons (right). Columns indicate the source localities in these matrices, and the rows indicate the target localities. The colors of the cells show the amount of larval dispersal between the two adjacent localities. Diagonal represents the local retention of larvae. Columns and rows are separated per North, Central, and South regions with white lines. ....	68
28	Heatmaps of node network metrics for the 22 leopard groupers localities of the GC for two seasons: Spring (left) and Summer (right). Values were standardized and centered by column. ....	69
29	Circle network representation of potential larvae connectivity among the 23 rocky-reef fishes localities in the GC for four seasons: Winter (above left), Spring (above right), Summer (below left), and Fall (below right). Nodes are colored according to the GC region they belong: North (orange) and Central (green). Arrowheads indicate the direction of the larval dispersal, and the width of the arrow line means the proportion of larvae being dispersed.....	70
30	Adjacency matrices of potential larval connectivity of the 23 rocky-reef fishes localities in the GC for four seasons: Winter (above left), Spring (above right), Summer (below left) and Fall (below right). In this matrices columns indicate the source localities and the rows indicate the target localities. The colors of the cells indicates the amount (of larval dispersal between the two adjacent localities. Diagonal represents the local retention of larvae. ....	71
31	Heatmaps of node network metrics of the 23 rocky-reef fishes localities in the GC for four seasons: Winter (above left), Spring (above right), Summer (below left) and Fall (below right). Values were standardized and centered by column. ....	72
32	Leopard grouper's distance matrices of the pairwise genomic differentiation $F_{ST}$ (Weir and Cockerham, 1984), and Euclidean (km) and Environmental distances among the 22 localities in the GC. ....	74
33	Isolation by distance patterns of the 22 leopard grouper's populations relating the neutral genomic differentiation ( $F_{ST}$ , Weir and Cockerman 1984) to the Euclidean (above left panel) and the environmental distances (above right, below left and right panels). ....	76
34	Plots of the linear models between the network centrality metrics and the genomic diversity of the leopard grouper for the Spring season. The red line is the linear regression of the values. ....	79
35	Plots of the linear models between the network centrality metrics and the genomic diversity of the leopard grouper for the Summer season. The red line is the linear regression of the values. ....	81

36	Rocky reef-fishes distance matrices of the A) Species turnover (Beta diversity, Jaccard dissimilarity index), and Euclidean (km) and Environmental distances among the 23 localities in the GC. ....	83
37	Patterns of distance-decay of similarity resulting from the relationship between the communities (Jaccard index) and Euclidean and environmental distance matrices. Significance values of correlations are $p < 0.01$ . ....	84
38	Plots of the linear models between the network centrality metrics and the alpha diversity of the rocky reef fish communities for the Winter, Spring and Fall seasons. The red line is the linear regression of the values. ....	86
39	Description of the ddRAD protocol for SNP discovery. ....	134
40	Histogram of genetic distances at various taxonomic levels: intra-species, intra-genus, intra-family, and intra-order with the 12S rRNA barcode estimated from the custom reference database. ....	144
41	Admixture bar plots for K=2, K=3 and K=5. ....	145
42	Individual rarefaction curves for (A) eDNA metabarcoding and (B) UVC. Sampling sites are shown. ....	163
43	Pearson correlation tests among the environmental variables for population localities. $R^2$ values are shown and its statistical significance (p-value 0 ***, 0.001 **, 0.01 *, 0.05 ·)...	164
44	Pearson correlation tests among the environmental variables for community localities and the observed species. $R^2$ values are shown and its statistical significance (p-value 0 ***, 0.001 **, 0.01 *, 0.05 ·).....	165

## List of tables

Table	Page
1	Reported spawning periods and PLD inside parentheses for species of fishes in the region Loreto- La Paz. Grey lines indicate the reproductive period, and black lines reproductive peaks. Column names indicate the initial letter of each month).....18
2	Population, community, seascape and network connectivity data at node and link-levels. .21
3	Temporal variation in the reported leopard grouper's reproductive seasons in different areas of the GC. The numbers correspond to the number of each month (1: January ... 12: December). .....29
4	Leopard grouper's sampling localities grouped in their respective GC regions. Table shows the Locality number (Loc), Locality name, the Locality ID (ID), Number of genotyped samples (N), and the geographic location (Longitude, Latitude). .....35
5	Sampling sites for eDNA seawater collection and UVC.....39
6	Environmental layers included as predictors of functional connectivity. ....46
7	Hypothetical population-level reproductive seasons from which potential larval connectivity matrices were obtained. The numbers correspond to the number of each month (1: January ... 12: December).....49
8	Hypothetical community-level reproductive seasons from which potential larval connectivity matrices were obtained. The numbers correspond to the number of each month (1: January ... 12: December).....50
9	Bioinformatic analyses results including sequencing and the filtering steps. Results are shown in terms of obtained number of sequencing reads or obtained number of SNPs, and in the number of individuals remaining after each step. ....52
10	Table summarizing the results of genetic diversity analysis of local leopard grouper populations using 985 SNPs. The local population localities are grouped in their corresponding GC region. The table shows the Locality number (Loc), Locality name, the Locality ID (ID), Number of genotyped samples (N), the geographic location (Longitude, Latitude), the rarefied allelic richness (Ar), the observed (Ho) and expected heterozygosities (He), and the Wright's inbreeding coefficient ( $F_{IS}$ ).....54
11	Table summarizing the results of alpha diversity analysis of local rocky reef fish communities using UVC and eDNA metabarcoding. The table shows the Locality number (Loc), Locality name, Locality ID (ID), Collection date, the geographic location (Longitude, Latitude), and the alpha diversity estimated using eDNA (including or not the OTUs), UVC, and complementing both detection methods UVC + eDNA (including or not the OTUs). .....58
12	Results of the Mantel and MRM tests were applied to the leopard grouper populations comparing the effects of the logarithm of Euclidean distance and the environmental distance (PC1, PC1:PC2, and PC1:PC3) into the neutral genomic differentiation ( $F_{ST}$ , Weir, and Cockerman 1984). Results are presented for all the Gulf of California, per region (North and Central) and for the Santa Rosalía (SRO) to Ventana (VEN) region. Two statistics that evaluate the correlation between matrices are presented: Mantel r and p-value of the Mantel test and the Multiple regression of the distance matrices (MRM) $R^2$ and p-value of the regression. Red letters represent statistically significant tests. ....75

13	Results of the Mantel tests and MRM evaluated on the leopard grouper populations comparing effects of resistance distances (from 22x22 polygon probability matrix) into genomic differentiation ( $F_{ST}$ , Weir and Cockerman 1984) in the All the Gulf, North region and Central region. Red letters represents statistically significant tests. ....	77
14	Linear models of network metrics and population genomic diversity metrics for March-June (Spring) and June-September (Summer) season's. Red letters represents statistically significant tests. ....	78
15	Results from the Mantel tests applied to rocky reef fishes communities comparing effects of logarithm of Euclidean distance and the environmental distance, into the beta diversity (species turnover, Jaccard). Results are presented for all the Gulf of California and per region (North and Central). Two statistics that evaluate the correlation between matrices are presented: Mantel r and p-value of the Mantel test, and the Multiple regression of the distance matrices (MRM) $R^2$ and p-value of the regression. Red letters represents statistically significant tests. ....	82
16	Results of the Mantel tests and MRM evaluated on the rocky reef fishes communities comparing effects of resistance distances (from 23x23 polygon probability matrix per season) into the beta diversity (species turnover, Jaccard) in the All the Gulf, North region and Central region. Red letters represents statistically significant tests.....	82
17	Linear models of network metrics and community alpha diversity metrics for Winter, Spring, Summer and Fall seasons. Adjusted $R^2$ and p-value. Red letters represent statistically significant tests. ....	84
18	12S rRNA primers designed in this study and used for custom reference database of teleost fishes from the Gulf of California. The amplicon includes the "teleo" barcode from Valentini et al. 2016. A: Amplicon lenght (bp); B: %GC; C: Hair pin Tm ( $^{\circ}$ C); D: Pair Dimer Tm ( $^{\circ}$ C); E: Self Dimer Tm ( $^{\circ}$ C); F: Tm ( $^{\circ}$ C). ....	134
19	List of species in the custom reference database and assigned GenBank accession numbers. ....	135
20	Species included in the Mock community sample (n=22).....	137
21	Primers for library construction.....	139
22	Pipeline for 12S rRNA metabarcoding bioinformatic analysis (OBITOOLS, VSEARCH, SWARM). ....	141
23	Genetic distances at various taxonomic levels: intra-species, intra-genus, intra-family and intra-order with the 12S rRNA barcode (dmax, % maximum genetic distance) estimated from the custom reference database. ....	142
24	Taxonomic list of Actinopterygii detected with UVC. ....	146
25	Read count per bioinformatic step. ....	149
26	Summary of the taxonomic assignments for 542 OTUs identified from the eDNA metabarcoding analysis employing three different reference databases (see methods for details). GOC: gulf of California. We show the number (%) of unassigned/assigned OTUs, then the number of those OTUs that are present in the GOC and the final OTUs count after taxonomic collapse and minimal abundance filter. The last 5 columns describe the final taxonomic resolution of the OTUs based on thresholds of sequence identity percentage. ....	150
27	Taxonomic list of Actinopterygii class detected with eDNA metabarcoding. ....	151

28	Detection of Species/OTU with UVC and eDNA metabarcoding.....	154
29	Genera detection from UVC and eDNA metabarcoding. ....	159
30	Family detection from UVC and eDNA metabarcoding. ....	161
31	Order detection from UVC and eDNA metabarcoding.....	162
32	Class detection from UVC and eDNA metabarcoding. ....	162

## Chapter. 1 Introduction

---

In recent years we have started to recognize and study the sea as a highly interconnected system exhibiting intricate spatiotemporal patterning. A primary incentive for this change has been the technological and analytical advances that have allowed us to collect, integrate, analyze and visualize large quantities of data that reveal the prevalence of structural complexity and interconnectedness in the sea (Pittman, 2018). In this context, seascape ecology has developed as a multidisciplinary research area that has emphasized the importance of the interactions between the spatial patterns and the ecological processes and the consequences of the spatial heterogeneity, across scales, on the marine biodiversity patterns (Boström et al., 2011; Turner et al., 2001; Wedding et al., 2011).

The application of landscape ecology to marine systems came through recognition that concepts developed in this theory (e.g., patch-matrix and patch-mosaic) could apply to a range of environments: from plankton patches, seagrass beds to reef patches, among others (Boström et al., 2011; Pittman, 2018). In addition to patchiness, the marine environment also exhibits spatial variability in continuous multidimensional gradients without discrete patch boundaries, although discontinuities and ecotones may still be present (Cushman et al., 2010). Also, the integrity and functionality of marine ecosystems and the maintenance of its biodiversity are primarily possible due to the flow of organisms, materials, and energy through the space (Crooks and Sanjayan, 2006). Therefore, connectivity have prevailed as a central concept in the study of seascape ecology (Correa Ayram et al., 2015).

Connectivity can be defined as the degree to which the seascape facilitates (or impedes) the dispersion of organisms among habitat patches or gradients (Taylor et al., 1993). It can be approached through the evaluation of the *functional connectivity* and the *structural connectivity* (Baguette et al., 2013; Olds et al., 2016; Pittman and Olds, 2015; Taylor et al., 1993; Turner et al., 2001). *Functional connectivity* is a product of the organisms interacting with the seascape and its response to seascape heterogeneity. It can be described as *potential* and *realized connectivity* (Calabrese and Fagan, 2004b): the former uses secondary (indirect) information regarding movements and flow, and the latter measures movement and flow directly (Fletcher et al., 2016; Grober-Dunsmore et al., 2009). *Functional connectivity* is inherently species- and life-stage-specific, as it depends on the organism's behavior, life-history traits, and the spatiotemporal scales of their movements (Stuart et al., 2021). *Structural connectivity* refers to the physical characteristics of the seascape that allow the movement of organisms and is primarily determined by the distance between biological entities and the seascape spatial configuration (Calabrese and Fagan, 2004a; Crooks and Sanjayan, 2006; Grober-Dunsmore et al., 2009). This type of connectivity quantifies the

physical relationships of the seascape elements, e.g., the spatial positioning of a particular habitat, the geomorphological features of the seafloor, or the hydrodynamic features impacting dispersal (Selkoe et al., 2016). Structural connectivity creates the basis for functional connectivity (Selkoe et al., 2016). Seascape connectivity can influence individuals, populations, and communities through various mechanisms. Therefore, it is considered an essential attribute of natural ecosystems as it sustains biodiversity and ecosystem function (Loreau, Mouquet, and Gonzalez, 2003; Matisziw and Murray, 2009).

In the present thesis, I will study seascape connectivity on the rocky reefs of the Gulf of California (GC) and its effects on the associated bony fish biodiversity patterns. This will be achieved by evaluating hierarchical relationships, where structural connectivity influences organismal dispersal and, consequently, genetic variation and species diversity. While each of these responses can be impacted by factors other than connectivity, theory and concepts suggest that connectivity could have a substantial effect on these parameters (Fletcher et al., 2016; Hubbell, 2001; Leibold and Chase, 2018; Loreau, Mouquet, and Holt, 2003; MacArthur and Wilson, 1967; Vellend, 2010). Because of its relevance to understanding ecological processes and patterns, diverse approaches to connectivity quantification have been developed. Some of these will be further described in the following sections as the basis for the methodological approaches included in the present thesis in the context of the rocky reefs' ecosystems of the GC.

## **1.1 Background**

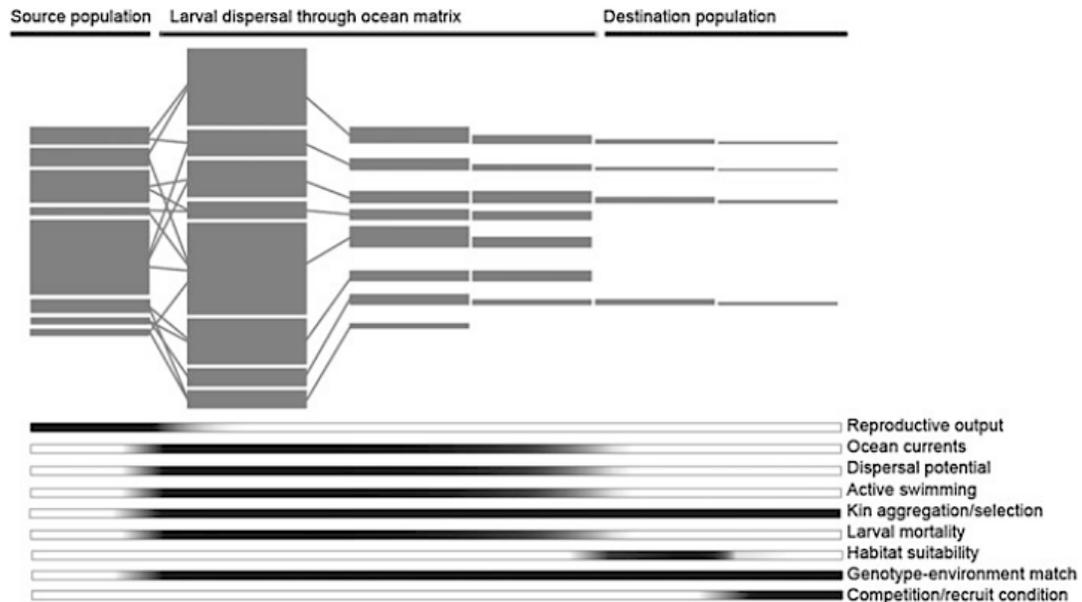
### **1.1.1 Contextualizing connectivity in the rocky reef ecosystems**

Rocky reefs are naturally fragmented ecosystems composed of submerged rocky structures that provide a substrate for the growth of marine life (Thomson et al., 2000). The patchy habitat structure of the rocky reefs is the basis where the development of reef fishes occurs and therefore is closely related to its ontogeny. In this sense, the life history of marine teleost fishes can be classified into five developmental stages: egg (embryonic), larvae, juvenile, adult, and senescence (Fuiman and Werner, 2002; Miller and Kendall, 2009). As a result of this complex development, fishes undertake ontogenetic habitat shifts via organismal dispersal during their lifetime. This means that the utilization of the seascape differs among life stages and, consequently, the dispersal capacity and barriers differ in each one of them (**Figure 1**).

Briefly, after fertilization, the egg stage develops (~1 week) from a single cell to a complex organism as part of the plankton (Miller and Kendall, 2009). Then hatching occurs, and fishes enter a larval stage that morphologically develops (Miller and Kendall, 2009). These early life stages of fishes are typically considered inert particles, including actively swimming larvae since the transport induced by currents generally spans over much larger scales than their swimming capacity. Depending on its specific pelagic larval duration (PLD), which can last from days to months, larvae can disperse great distances by water movements (to hundreds of kilometers) (Bandelj et al., 2020; Fuiman and Werner, 2002). The egg and larval periods have essential ecological and evolutionary functions as they represent an effective means of dispersal that can extend the range of a population and mix the gene pool (Fuiman and Werner, 2002). Nevertheless, they are also the most vulnerable stages since their dependence on external environmental factors and predation, causing high mortality rates (Fuiman and Werner, 2002; Helfman et al., 2009; Miller and Kendall, 2009).

Following the egg and planktonic larval stages, a metamorphosis to juveniles occurs in which a morphological and ecological transformation happens. Juveniles are more capable of directed swimming than larvae and thus are no longer considered planktonic. They can actively migrate to nursery areas, sometimes aided by prevailing currents. The settlement frequently occurs in nursery habitats with specific characteristics, leading to juvenile fishes competent to remain on a substratum, i.e., the reef itself or an intermediate habitat with marine vegetation (mangroves, *Sargassum* spp., and rhodoliths). The adult stage begins when gonads first mature. This stage is a period of reproduction when the gonads go through maturation cycles (annual or more frequent) and in which fishes often actively migrate to the reproductive areas (from meters to kilometers) to spawn. Spawning may be as diffuse as a region of the open ocean or as specific as a nest along the shore and depend upon physiological and ecological factors (i.e., temperature, photoperiod, latitude, and depth, among other factors). Still, juveniles and adults stay most of the time near their habitat and rarely move away from it (Green et al., 2015; Kritzer and Sale, 2006, Chapter 3). A period of senescence follows the adult stage.

Throughout the life cycle of the rocky reef fishes, some areas are critical for maintaining their populations, such as spawning aggregations and nursery habitats. The former usually takes place during spring and summer in the GC (Erismann et al., 2010b, 2012), and the latter is often used by multiple species (Sadovy and Colin, 2012). Consequently, these areas are critical for maintaining populations of focal species in the GC (e.g., *Lutjanus peru*, *L. argentiventris*; *Mycteroperca rosacea*, *Paralabrax auroguttatus*, *Paranthias colonus*; *Balistes polylepis*) (Munguia-Vega et al., 2018).

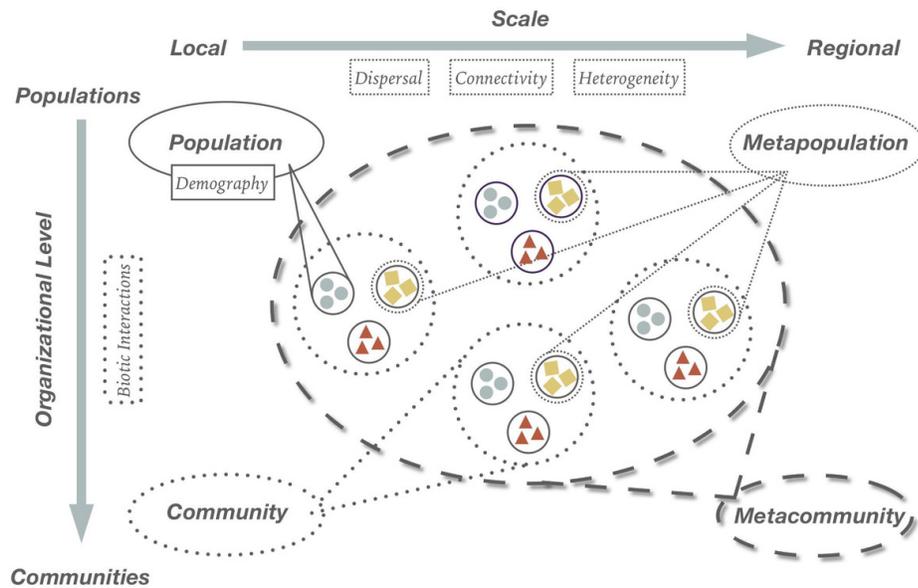


**Figure 1.** Some of the biological characteristics of marine fishes and the physical and environmental attributes of the seascape (bottom) influence demographic and genomic processes throughout the teleost fishes' lifecycle and determine the distribution of genomic variation within and among populations. Hypothetical source populations are shown as independent bands (top left) where bandwidth can be understood as the number of individuals or genomic variation. The bottom panel underscores the period in which the biological or seascape characteristics are most relevant. Reproductive output of the source population varies as a function of population size and timing of reproduction. During dispersal (from left to right), larvae can be advected by currents, mixed (crossed lines), and may become diluted or concentrated, represented by bandwidths. Larval dispersal potential is related to a species' early life-history traits, such as pelagic larval duration. Larval behaviors such as kin aggregation and active swimming can increase or counter physical oceanography to concentrate larvae. Bandwidths become smaller, depicting larval mortality due to the environment of the ocean (biotic and abiotic). Successful settlement of larvae into a hypothetical destination population (top right) is dependent on habitat suitability. Post-settlement survival in the population depends on the condition of the recruit, competition, and other factors (Figure taken from Liggins et al., (2019)).

### 1.1.2 Metapopulation and metacommunity connectivity

Theory indicates that in seascapes where habitats are fragmented, species are patchily distributed in *metapopulations*, composed of a set of *local populations* connected by dispersal of individuals (Hanski, 1998); and in species *metacommunities*, structured in a group of *local communities* connected by dispersal of multiple species (Leibold et al., 2004) (**Figure 2**). As a result of these ecological processes, we affirm that connectivity among rocky reefs is one of the most critical drivers for maintaining fish species' long-term population and community viability as it sustains their demographic processes (Braunisch et al., 2010; Cayuela et al., 2018; Cowen et al., 2007; Hanski and Gilpin, 1991; Melià et al., 2016; Uroy et al., 2021); allows individuals to move towards new habitat patches to shelter, feed or reproduce (Berkström et al.,

2020); allows genetic exchange between physically distant populations (Palumbi, 2003); and promotes community similarity (Moritz et al., 2013).



**Figure 2.** The conceptual figure shows how different organization levels are integrated within the connectivity framework across two axes: scale and organizational level. Populations are embedded within local communities, including regional metapopulations and regional metacommunities, respectively (Chase et al., 2020).

A significant difficulty in studying the functional connectivity of marine fishes is the direct observation of dispersion, mainly in its early life stages. This results from the small size of many species' propagules, the variability and complexity of the marine environment, the accessibility of the study sites, the costs associated with underwater research, and its associated detection limitations (e.g., visual vs. molecular methods), among other causes.

Indirect methods to investigate metapopulation connectivity on reef-fish species of the GC have been applied. In particular genetic methods (to estimate metapopulation genetic connectivity), numerical and biophysical models (to estimate demographic connectivity or the potential larval dispersal of a species), and network analysis (to resume complex connectivity patterns derived from oceanographic modeling) have provided relevant information on metapopulation connectivity estimations (Anadón et al., 2011; Avendaño-Ibarra et al., 2013; Cisneros-Mata et al., 2019; García-De León et al., 2018; Gutierrez et al., 2004; Marinone, 2012; Marinone et al., 2008; Munguia-Vega et al., 2014; Munguia-Vega et al., 2018; Reguera-Rouzaud et al., 2020; Santiago-Garcia et al., 2014; Sanvicente-Añorve et al., 2011; Soria et al., 2014). Additionally, direct methods evaluating larval assemblages' distribution across the GC have revealed connectivity routes and their relationship with oceanographic and environmental characteristics.

These investigations have recognized that circulation processes play an essential role in zooplankton dispersion, including fish eggs and larvae (Avenidaño-Ibarra et al., 2013; Camacho-Gastélum et al., 2020; Contreras-Catala et al., 2012; Inda-Díaz et al., 2010).

Indirect and direct methods applied to the study of functional connectivity of fish populations have contributed to understanding the connectivity patterns in the GC, facilitating and improving the design and implementation of marine protected areas and reserve networks. They have demonstrated the importance of considering spatially explicit connectivity patterns to evaluate potential changes in connectivity under climate-change scenarios (Álvarez-Romero et al., 2018). It is relevant to mention that much of the connectivity literature in the GC has focused on metapopulation processes. Conversely, at the metacommunity level, few studies investigate connectivity on a regional scale explicitly accounting for larvae or adult dispersal information or environmental and oceanographical information, although some approximations exist (Petatán, 2015; Ulate et al., 2016).

In the present thesis, I will approximate the study of the structural and functional seascape connectivity at metapopulation and metacommunity organization levels in the GC's rocky reef fishes. To achieve this, I will use direct and indirect methods to obtain connectivity information. Then, I will explore the potential relationship between both types of connectivity using a graph-theoretic approach, and finally, I will propose the processes determining the biodiversity patterns found.

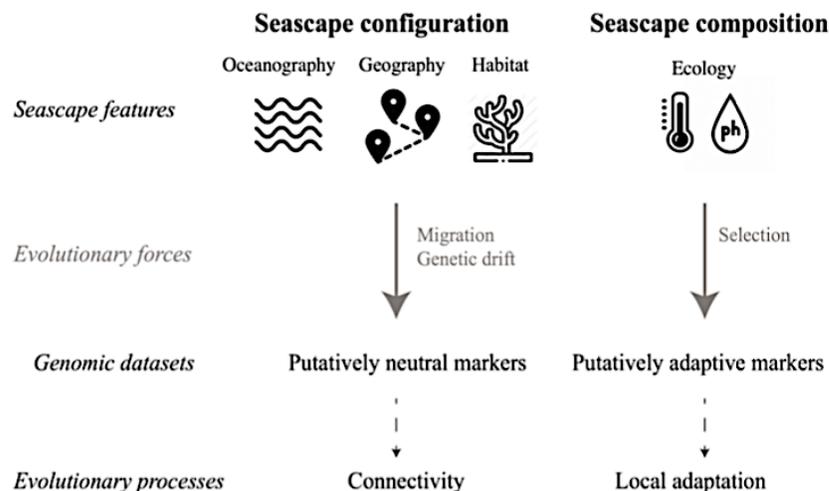
### **1.1.3 Direct and indirect methods for the study of connectivity in rocky reef ecosystems**

#### **1.1.3.1 Seascape genomics: evaluation of the metapopulation's genomic connectivity and its relationship with the seascape environmental characteristics**

Seascape genomics use seascape configuration and composition as statistical predictors of population genomic connectivity patterns (Liggins et al., 2019). It can be considered a part of the population genomics discipline, which is the population genetic analyses of a large number of loci that allow discrimination between locus-specific (selection) and genome-wide effects (migration and genetic drift) (Nielsen et al., 2009) (**Figure 3**). Recent advances in DNA sequencing technology have allowed this transition (from population genetics to population genomics) because high-throughput sequencing approaches can now produce hundreds to thousands of polymorphic markers across the genome, such as single nucleotide polymorphisms (SNPs). This allows for more significant discrimination of differentiation between populations (Oleksiak and Rajora, 2020). The genotyping and sequencing approaches applied to

population genomics are described in Holliday et al. (2019). A reduced representation sequencing approach (RADseq) was used in the present thesis. This approach allows sequencing of a reduced portion of the genome (0.1–1%) derived from specific restriction sites on hundreds of individuals (Oleksiak and Rajora, 2020, Chapter 1).

Genomic connectivity footprints the effective movement of genes, resulting from those individuals that successfully reproduce after dispersing. It is a population-level metric that reflects the cumulative impact of realized functional connectivity on the allele frequencies (Selkoe et al., 2016). There are two forms of evaluating the metapopulation genomic connectivity: using a species' neutral or adaptive genomic variation. Therefore, seascape genomics is operationally split into neutral and adaptive seascape genomics (Liggins et al., 2019; Riginos et al., 2016). Neutral genomic variation measures the movement of alleles that have no consequence on fitness and are related to demographic connectivity's long-term outcome (Selkoe et al., 2016). The adaptive genomic variation relates to the movement of alleles due to individual fitness and corresponds to the adaptive potential under natural selection pressures. So, describing the features that shape the population genomic connectivity using a seascape genomics approach requires, first, identifying the adaptive and neutral genomic variation and characterizing the population genomic structure of the studied species (i.e., the spatial distribution of the genomic variation), and second, recognizing the seascape variables affecting its spatial organization (**Figure 3**) (Dalongeville et al., 2018).



**Figure 3.** Conceptual framework of the seascape genomics approach (Figure taken from Laura Benestan's Seascape genomics presentation September 2020).

Traditional assessments of population genetic and genomic differentiation rely on the estimation of the number of migrants between populations ( $N_e m$ ) using the relationship described by Wright's island population model equation  $F_{ST} = 1/(4N_e m + 1)$ , where  $F_{ST}$  is an index of genetic differentiation,  $N_e$  the effective population size, a number of individuals contributing to reproduction, and  $m$  the migration rate (Wright, 1931). Estimators based on Wright's fixation index ( $F_{ST}$ ) integrate over longer time scales, possibly representing historical connectivity more than the present day because response lags depend on marker mutation rates and drift (Selkoe et al., 2016).

In marine fishes, detecting its population genomic differentiation is challenging because species have complex lifecycles, large population sizes, and pelagic larvae with high dispersal capacity, causing connectivity patterns to be not intuitive (Gaines et al., 2007; Hellberg, 2009; Nielsen et al., 2009; Waples et al., 2006).

### **1.1.3.2 Mechanisms of genomic isolation in metapopulations**

The most basic spatial description of a metapopulation involves a binary spatial structure in which there are suitable habitat patches with local populations distributed within a matrix of unsuitable non-patch space (Kritzer and Sale, 2006, Chapter 2). A patchy population (i.e., a metapopulation) that is not demographically subdivided will also be genetically panmictic. Nevertheless, changing ecological circumstances and environmental heterogeneity, leading to changing rates of migration, may provide interesting opportunities for metapopulation processes to shape genetic patterns (Harrison and Hastings, 1996). From this idea, one fundamental hypothesis underlying seascape genomics studies evaluating metapopulation connectivity is an association between the distribution of the genomic variation and the distance.

The *isolation-by-distance* hypothesis (IBD) (Rousset, 1997; Slatkin, 1994; Sewall Wright, 1943) describes the correlation between the geographic distance and genomic differentiation ( $F_{ST}$ ) between local populations to elucidate the effects of the spatially limited dispersal on the genomic structure of populations (i.e., stepping-stone model of dispersal) (Kimura and Weiss, 1964). After a random mating scenario of panmixia, IBD is one of the most fundamental forms for describing connectivity (Hedgecock et al., 2007). It assumes a structurally homogeneous yet discontinuous seascape and traditionally uses straight-line (Euclidean) geographic distances to predict population differentiation.

In addition to IBD (Euclidean), the genetic structure may also be due to patterns and gradients of environmental factors such as temperature (Benestan et al., 2016; Selmoni et al., 2021; Teske et al., 2019) and salinity (Guo et al., 2015), among others (García-De León et al., 2018; Sandoval-Castillo et al., 2018). This environmentally induced isolation has been termed *isolation by environment* (IBE) (Rodríguez-Zárate et al., 2018), which can also emerge when evaluating genomic patterns of population genomic structure in species that inhabit heterogeneous environments. In an IBE mode of isolation, the degree of genomic differentiation among sampling units is expected to increase with increasing environmental dissimilarity (Wang and Bradburd, 2014).

Connectivity is not only a function of the geographical distance between habitat patches but also varies at different spatial scales (Dalongeville et al., 2018; Xuereb et al., 2018) across a fragmented seascape (D'Aloia et al., 2014) and depends on the habitat matrix through which organisms disperse which determines the effective isolation of habitat patches (Hanski and Gaggiotti, 2004). In this sense, in the *isolation-by-resistance* (IBR) hypothesis (McRae, 2006), the distances are estimated using a graph-theoretic distance metric by weighting the cost (resistance) of crossing the seascape through organismal dispersion, e.g., in reef fishes with a pelagic larval stage, individuals are subjected to the circuitous movement of water masses (Cowen and Sponaugle, 2009; Riginos et al., 2011), which prevent mixing and diffusion of the larvae and decouple larval dispersal from a linear distance at determined spatial scales (Contreras-Catala et al., 2012; Sánchez-Velasco et al., 2013; Weersing and Toonen, 2009). Therefore, an IBR pattern may manifest because complex ocean circulation represents an essential barrier to dispersal, producing patterns of genomic structure that are difficult to interpret without integrating oceanographic information into the analyses of population genomic data. This has placed patterns of genetic patchiness into more realistic, ecologically relevant contexts (Selkoe et al., 2010; White et al., 2010).

In the GC, studies have demonstrated that complex genetic structure patterns are related to ocean currents evaluated explicitly (Munguia-Vega et al., 2014; Munguia-Vega et al., 2018). In such cases, asymmetric oceanographic distance (derived from oceanographic models) better captures the physical and biological processes influencing the genomic connectivity (gene flow). For instance, Reguera-Rozaud et al. (2021) studied two snappers species (*Lutjanus peru* and *L. argentiventris*) and identified the presence of a significant genetic structure in both species, with distinctive barriers to gene flow at different spatial scales: at large scales (>2500 km) IBD was the principal driver for the genetic structure, and at more minor scales (<250 km), habitat discontinuity for juveniles and adults and the environmental differences throughout the species ranges represented the main potential barriers.

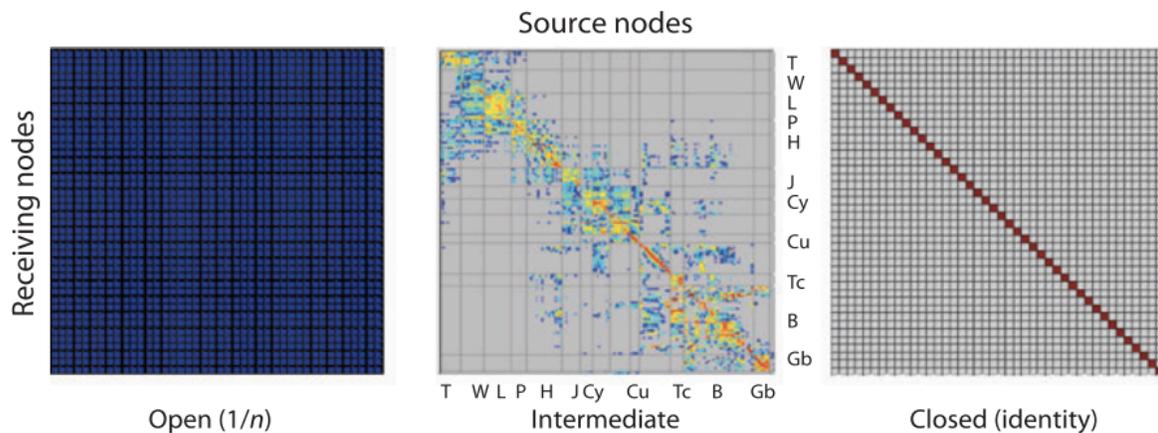
In the present thesis, I will evaluate the genomic connectivity of the leopard grouper, *Mycteroperca rosacea*, a conspicuous species of the rocky reefs in the GC, using a seascape genomics approach considering a reduced representation of the genome and SNPs. This species has been previously genetically evaluated in localities of the North and Central Gulf. First, Munguia-Vega et al. (2014) evaluated leopard grouper metapopulation genetic connectivity in 17 sites from the Midriff Islands region of the northern GC, using two mitochondrial DNA (cytochrome *b* and ATPase) markers, a biophysical oceanographic model, and a graph-based theoretical approach. Results suggested moderate levels of population genetic differentiation within the northern GC, insufficient gene flow to homogenize populations, and closely related haplotypes with some geographical structure despite occasional long-distance dispersion. Matrix-based analyses (Mantel tests) showed a lack of correlation between pairwise genetic differentiation against geographic distance but a significant correlation with the adjacency matrices and the graph distance matrices derived from the oceanographic model. Most of the observed connectivity patterns in downstream sites (Sonora coast) were asymmetric, while those between upstream sites (Baja California and the Midridges) were symmetric. This multidisciplinary approach allowed the authors to understand that the most significant fishing areas are sustained by high levels of local larval retention and high larvae contribution from upstream sites. Later, Jackson et al. (2015) analyzed 21 localities using the same mitochondrial DNA markers and 12 microsatellites loci, including localities from the North and Central GC. Results evidenced statistically significant genetic differentiation using mitochondrial DNA and microsatellites with multiple genetic groups, particularly between peninsular and mainland sites. Long-term migration rates suggested asymmetrical larval dispersal between the northern and central GC. The authors also found no evidence for IBD.

### **1.1.3.3 Metapopulation demographic connectivity: estimation of potential larval dispersal among local populations using oceanographic models**

Demographic connectivity is how the local population growth, extinction, and recolonization rates are affected by dispersal (Soria et al., 2012). Its evaluation requires the estimates of the relative contributions of self-recruitment (i.e., the proportion of locally settling individuals spawned by local parents) and immigration to these rates (see Lowe and Allendorf, 2010). To this end, oceanographic modeling of potential larval dispersal has been commonly used to indirectly predict the demographic connectivity among and within habitat patches (Kritzer and Sale, 2006; Lowe and Allendorf, 2010). These models can inform us about the direction, spatial scale, and magnitude of larval distribution in a given area. Besides, they can incorporate only oceanographic (i.e., numerical models), or oceanographic and

biological (i.e., biophysical models) factors, to provide insights into the demographic connectivity, allowing the elaboration of hypotheses about the connectivity dynamics.

Previously in the demographic research of fish populations, the notion of the larvae settling into a local population from a mixed larval pool (i.e., larvae from all potential sites mixed into a single source pool) led to the belief that marine populations were open, potentially over hundreds of kilometers. This perspective was supported by studies that found little genetic structure over large spatial scales (Benestan et al., 2021; Dalongeville et al., 2018; Xuereb et al., 2018). However, recent approaches incorporating complementary methods evaluating demographic connectivity inform us that some marine populations, even some with long PLD, exhibit some degree of self-recruitment of larvae (D'Aloia et al., 2013). If local larval retention is large enough for self-replacing populations and demographic isolation is sustained for many generations, subpopulations will become genetically differentiated. In such cases, the population's genetic structure will fail to reflect any spatial trend, and levels of genetic differentiation will be greater between adjacent sites than between distant ones. This evidence points to a complex metapopulation structure, in which differential dispersion patterns among locations exist, and indicates that marine populations may be less open than we thought (Cowen and Sponaugle, 2009) (**Figure 4**).

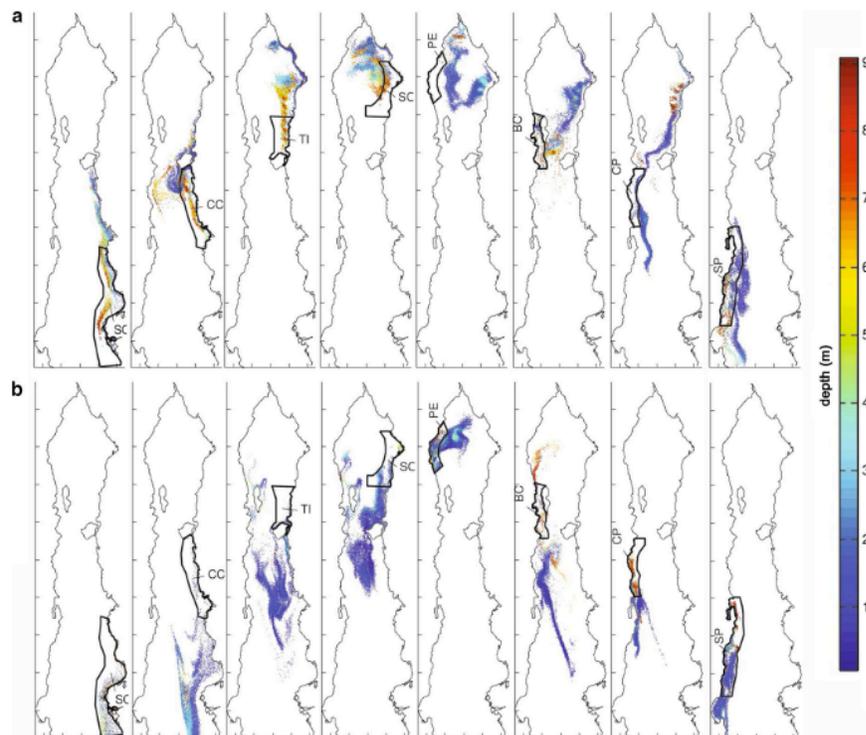


**Figure 4.** Comparison of open to closed population continuum. The degree to which a set of local populations (or nodes) might be connected to each other via larval dispersal can be expressed as a connectivity matrix where nodes are distributed along both the vertical and horizontal axes. In the fully open case (left panel), where all populations within a given domain contribute propagules equally to each other, all cells within the connectivity matrix are populated similarly ( $1/n$  where  $n$  is the total number of local populations, i.e., source locations). This extreme case represents what might be expected if all propagules were equally mixed and then settled at the random back to the source sites. In the wholly closed case (right panel), all propagules return directly to their source location, all cells along the diagonal would be populated with a probability of 1 (i.e., 100% self-recruitment), and all other non-diagonal cells would be populated by zero. This case would be expected if all populations were utterly isolated from any different location through dispersal processes. The intermediate situation is represented by a matrix constructed from biophysical modeling data with some variation of probabilities along with both the diagonal and non-diagonal cells. The scale represents a range in cumulative abundance of successful recruits, where red/orange represents ecologically significant levels of exchange and green/blue represents very low levels of genetically relevant exchange (Image taken from Cowen et al. 2006).

Species' demographic processes are also affected by the spatial and environmental heterogeneity, i.e., species reproductive success may increase, or survivorship may decrease in a habitat with particular oceanographic characteristics (e.g., transport conditions, food availability, season) (Bakun, 2006; Parrish et al., 1981; Selwyn et al., 2016). This habitat-specific survivorship and reproductive success is the basis for source-sink dynamics in metapopulations (Dunning et al., 1992), in which population growth rates are positive in some patches (*sources*) but are negative in others (*sinks*). In the GC reef fishes, systems of source-sink metapopulations have been identified. For example, in the leopard grouper (*Mycteroperca rosacea*), Munguia-Vega et al. (2014) used empirical genetic data, an oceanographic model describing metapopulation dynamics of potential larval dispersal, and graph distances of modeled networks to understand the mechanisms driving dispersal of the species in the Northern GC. In the Pacific Red snapper (*Lutjanus peru*), Munguia-Vega et al. (2018) described the spatial distribution of its genetic diversity with the modeled seasonal ocean circulation during spring and summer in the Central GC, based on expectations from metapopulation theory. Further, in the Yellow snapper (*L. argentiventis*), Reguera-Rouzaud et al. (2020) evidenced a metapopulation structure using genetic data and a ROMS oceanographic model that exposed a patchily distributed metapopulation interconnected by larval dispersal at different rates and directions.

In this context, models for the study of ocean circulation in the GC have been developed based on the equations of motion of fluid dynamics and seawater thermodynamics. Their purpose is to understand the complex transport processes in the marine system (Nihoul, 1973). As a result, our understanding of larval dispersal on the GC (Marinone, 2012; Marinone et al., 2011) has been principally based on a combination of these oceanographic models and population genetics (or other empirical methods) that have validated the modeled results (Munguia-Vega et al., 2014; Munguia-Vega et al., 2018). Oceanographic models have evidenced that the ocean currents in the Northern and Central GC are predominantly asymmetric due to the semi-enclosed conditions caused by the peninsula of Baja California. Also, a northward current is present along the eastern coast of the GC between the spring-summer seasons (e.g., August, **Figure 5**), transporting larvae to the north and producing a cyclonic gyre (counter-clockwise) in the Northern GC; while a predominantly southward current is present along the eastern coast of the GC moving larvae to the south during fall-winter (e.g., October, **Figure 5**) and producing an anticyclonic (clockwise) circulation phase in the Northern GC (Marinone et al., 2011). These oceanographic conditions have been determinants for larval connectivity. Biophysical models and genetic studies have indicated that larval dispersal kernels (i.e., the statistical distribution of dispersion distances in a spatially structured population, Cayuela et al., 2018) in the GC are not spatially symmetrical but are highly constrained in particular routes by the direction of the currents driven by the narrow shape of the Gulf, particularly for

species that spawn during a single season (Álvarez-Romero et al., 2018; Jackson et al., 2015; Lodeiros et al., 2016; Munguia-Vega et al., 2014; Munguia-Vega et al., 2018; Soria et al., 2012; Turk-Boyer et al., 2014). Thus, larvae spawned in summer on the eastern coast of the GC are more likely to move in a northerly direction, while those spawned in fall in the same location are more likely to move south (**Figure 5**). In such a highly asymmetric current system, it is, therefore, crucial that conservation strategies (e.g., location of marine reserves) consider the direction of the larval flow during the spawning season of target species since these areas act as larval sources to sustain metapopulations of those species (Álvarez-Romero et al., 2018; Green et al., 2014; Munguia-Vega et al., 2018).



**Figure 5.** Larval dispersal in the GC for species spawning during summer (a, August) and fall (b, October) with PLD of 28 days, based on a three dimensional oceanographic model where larvae were released within each of eight coastal polygons shown in bold. Color shows the depth of virtual larvae (Taken from Munguia-Vega et al. (2018)).

For the present thesis, it is important to emphasize what Lowe and Allendorf (2010) described: genetic methods provide information on population genetic connectivity and structure (i.e., the degree to which gene flow affects evolutionary processes within populations). Nevertheless, traditional population genetic analyses alone provide little information on demographic connectivity (i.e., the degree to which population growth and vital rates are affected by individual dispersion). The exception is the studies of connectivity that have incorporated kinship or parentage analyses to infer the trajectories and distances of larval movement and to obtain data on demographic connectivity, though they required great sampling

efforts and genotype datasets (Almany et al., 2017; D'Aloia et al., 2013; Melià et al., 2016; Saenz-Agudelo et al., 2011; Schunter et al., 2014). Here, I considered potential demographic connectivity as the degree to which larval dispersal promotes co-occurrence and interactions between individuals from different populations, and the genetic connectivity as its long-term result and indirect measure of the contribution of dispersal to local population growth and persistence. Genetic and demographic connectivity are complementary for understanding their effects on metapopulation processes and will be incorporated in the analysis of connectivity patterns in the GC.

#### **1.1.3.4 Metacommunity connectivity: from species detection to species turnover evaluation**

Studying the genomic and demographic connectivity in fish metapopulations has allowed the understanding of patterns of variation among geographically structured habitats in a complex seascape, such as rocky reefs. If we move to a broader organization level of analysis, differences among samples in terms of species composition affected by metacommunity connectivity can also be evaluated (Diniz-Filho and Bini, 2011). In this regard, the first step in understanding the spatial patterns in ecological communities is identifying the species components through monitoring methodologies. However, recognizing the complete species constituents in a particular ecosystem is arduous, especially in the marine realm.

In the GC, significant advances in the study and characterization of fish communities have been achieved through underwater visual censuses (UVC) (Aburto-Oropeza and Balart, 2001; Barjau-González et al., 2016; Brusca et al., 2005; Fernández-Rivera Melo et al., 2018; Olivier et al., 2018; Ramírez-Ortiz et al., 2017). This method relies on multiple trained observers to record individual fish's presence within a fixed area. It is a survey method biased against wary, highly mobile organisms or small or cryptic species, and the detection may be adversely affected by local conditions (e.g., currents, visibility, depth, among others) (Bozec et al., 2011; MacNeil et al., 2008). Alternatively, eDNA metabarcoding has emerged as a novel technique for fish surveys (Thomsen et al., 2012), as it takes advantage of high-throughput sequencing of a conserved standard genomic region (*barcode*) (Hebert, 2003), which is PCR-amplified from a complex environmental sample, to detect one or several species (Taberlet et al., 2012). In the present thesis, UVC and eDNA metabarcoding will be used to characterize and assess the communities of fishes associated with rocky reef ecosystems in the GC. The complementary information of both approaches will allow a comprehensive description of the local fish communities in the study area to evaluate patterns of metacommunity connectivity and identify its potential spatial or environmental determinants.

### 1.1.3.5 Mechanisms of isolation in metacommunities

The concept of metacommunity has marked a turning point in community ecology because it has shifted the focus from local interactions between functionally distinct species and their environments as the principal component determining ecological communities to a broader vision in spatial and temporal scales, including environmental filtering and dispersion limitation occurring in a heterogeneous environment at different scales (Holyoak et al., 2005; Hubbell, 2001; Leibold et al., 2004; Vellend, 2010), to understand patterns of species abundance, occurrence, composition, and diversity (Chase et al., 2020). It is now accepted that communities are structured by combining all these processes to different degrees (Moritz et al., 2013). Therefore, the focus is on understanding their relative importance, or what proportion of the overall variance in community structure can be explained by each process.

Among the indexes that have been used to describe community structure and composition are *alpha diversity* (species richness) and *beta diversity* (species turnover) (Magurran, 2004). The first is the fundamental and most intuitive expression of biological diversity and represents the number of coexisting species in a local community. The second evaluates the changes in species composition among sites within a geographical area of interest (Whittaker, 1960). Alpha and beta diversity have been assessed using UVC and eDNA metabarcoding (Boulanger et al., 2021; Cantera et al., 2021; Lamy et al., 2021; Li et al., 2018; Marwayana et al., 2022), and results have evidenced that the complementary use of both methods augments the resolution for examining biodiversity patterns.

Under the neutral theory of biodiversity and biogeography (Hubbell, 2001), the differences in species composition between local communities are expected to decrease exponentially with increasing geographic distance (Legendre and De Cáceres, 2013). This is empirically and theoretically analogous to the previously described pattern of IBD (Wright, 1943). This species' spatial turnover resulting from the decay of community similarity with geographic distance is known as the *distance–decay* relationship. This association has proven common in fish communities (Anderson et al., 2013; Leprieur et al., 2009; Navarrete et al., 2014) and is a valuable tool for understanding the variables driving community changes, such as dispersal limitation (Morlon et al., 2008).

The distance decay of similarity in ecological communities has been studied across a broad spectrum of marine organisms. This body of literature suggests that the decay process is caused by at least two, not necessarily mutually exclusive, mechanisms (Leprieur et al., 2009). First, the *environmental filtering* hypothesis predicts that changes in community composition are a result of species-specific niche

differences in adaptive responses that have evolved along environmental gradients (niche-based processes), i.e., sites that are located at short distances are more likely to be environmentally similar and thus suitable for the same species (Legendre and Legendre, 1998). Niche-based processes produce a decay of compositional similarity with geographical distance when environmental conditions are spatially autocorrelated, so nearby sites tend to be more similar in their environmental conditions than distant sites. Second, the *dispersal limitation* hypothesis (Hubbell, 2001) postulates that: (1) the differences in species dispersal capabilities produce patterns of distance decay in community similarity even in homogeneous environments (neutral processes), and (2) the spatial configuration of the landscape influences species turnover by controlling habitat patch connectivity and hence the permeability of the seascape to organism movement. A seascape with significant barriers to movement is expected to produce greater community similarity decay rates than a more homogeneous and well-connected seascape. These two different mechanisms will produce similar predictions regarding distance decay and are difficult to separate in empirical studies (Moritz et al., 2013). Therefore, we can identify three predictions that differ in terms of the effect of distance on community similarity: 1) when decay in similarity is observed for the environmental distance but not for the geographic distance; 2) when decay in similarity for the geographic distance but not for the environmental distance, and 3) a decay in similarity for both distances (Tornero et al., 2018).

While distance decay processes in community assembly are often quantified as geographic isolation (Euclidean), the growing understanding of the role of the seascape matrix in mediating metacommunity connectivity suggests that geographic isolation alone may often be insufficient for quantifying the role of this process (Fletcher et al., 2016), therefore, accounting for connectivity by the evaluation of seascape resistance to organismal dispersal can be necessary for providing better estimates of connectivity (and conversely of isolation) than geographic distance alone (McRae and Beier, 2007). In the present thesis, I will evaluate metacommunity connectivity by contrasting the distance decay of community similarity with the geographical (Euclidean), environmental, and resistance distance as predictors.

#### **1.1.3.6 Metacommunity demographic connectivity: estimation of potential larval dispersal of reproductive assemblages using oceanographic models**

Previous research developed in other fragmented marine ecosystems has shown that oceanographic models can capture the spatial complexity of oceanographically mediated dispersal among local communities and have provided insight into the potential scales of metacommunity connectivity.

Examples have been developed in mesophotic biogenic habitats in the Adriatic Sea, hydrothermal vents, polychaeta communities in the Mediterranean Sea, Brazilian coral reefs, nearshore marine species in the Southern California Bight, among others (Bandelj et al., 2020; Magris et al., 2016, 2018; Mullineaux et al., 2018; Schill et al., 2015; Watson et al., 2011). Results have highlighted that oceanographically mediated larval dispersal is an important determinant of the spatial patterns of community similarity (Watson et al., 2010) and that the incorporation of dispersal estimates in the analysis of metacommunity connectivity renders the study more realistic concerning the processes occurring in dispersive environments (Moritz et al., 2013).

The rocky reef fishes of the GC exhibit a wide range of reproductive strategies (including fecundity, spawning frequency, PLD, and larval dispersal capacities) (**Table 1**). Nevertheless, there is a general agreement that spring and summer are the primary spawning seasons (Aceves-Medina et al., 2004; Aceves-Medina et al., 2003; Ahern et al., 2018; Peguero-Icaza et al., 2008, 2011; Sala et al., 2003; Sánchez-Velasco et al., 2009). This evidence has been drawn from ichthyoplankton abundance and species composition data collected from the water column, which provided valuable information concerning the spawning activities, mainly of epipelagic fishes of the Gulf. Sánchez-Velasco et al. (2009) also found that the seasonally reversing central eddy and coastal current on the mainland continental shelf of the northern GC are related to the seasonal larval fish assemblages. They also propose that seasonal changes in hydrography and circulation trigger the species' spawning favored by each area's environmental conditions, which implies a close and predictable coupling between the environmental changes and the spawning periods.

Similarly, Peguero-Icaza et al. (2011) showed substantial seasonal differences in the connectivity routes for larval fish assemblage areas in the northern GC, related to circulation phases (cyclonic and anticyclonic). These seasonal changes have been proved to affect metapopulation connectivity between the northern and the southern GC, which has important implications for managing and conserving marine protected areas. Other indirect sources of information are the monthly trends in commercial landings (Erisman et al., 2010) which confirm the existence of seasonal patterns in larvae abundance along the GC, being spring and summer the ones that present higher values in contrast with fall and winter.

Even though the fish fauna of the GC is one of the best-studied in the eastern Pacific, knowledge about the reproductive aspects of each species remains scarce (Aceves-Medina et al., 2003). This is a significant limitation to our understanding of connectivity as a regional process structuring communities, and therefore we need to rely on broad approximation to understand the influence of larval dispersal on

the metacommunity organization (Magris et al., 2016). In the present thesis, I propose the estimation of potential larval connectivity of groups of species that shares reproductive strategies such as seasonal reproduction (Magris et al., 2018; Préau et al., 2021), i.e., grouping reef species into coarse *reproductive assemblages* according to their spawning seasons (Winter, Spring, Summer, and Fall) representative of the whole community. This approach would allow us to explore the potential relationship between the seasonal larval pulses of fishes with the variation in species composition of the local communities along the GC.

**Table 1.** Reported spawning periods and PLD inside parentheses for species of fishes in the region Loreto- La Paz. Grey lines indicate the reproductive period, and black lines reproductive peaks. Column names indicate the initial letter of each month)

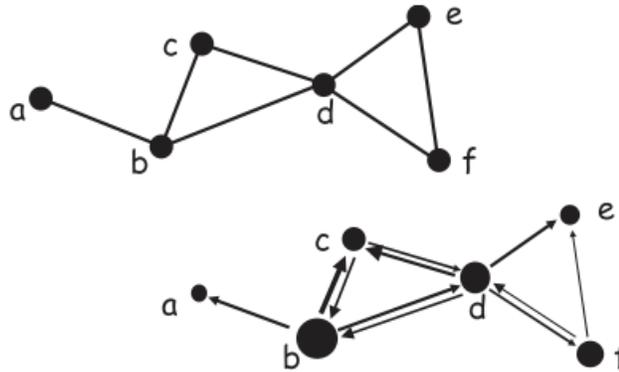
Species	J	F	M	A	M	J	J	A	S	O	N	D
<i>Balistes polylepis</i> (35)						■	■	■	■	■	■	
<i>Lutjanus argentiventris</i> (26)						■	■	■	■	■		
<i>Lutjanus peru</i> (32)				■	■	■	■	■	■	■	■	
<i>Mycteroperca rosacea</i> (28)			■	■	■	■	■					
<i>Paranthias colonus</i> (26)			■	■	■	■	■	■	■	■		
<i>Seriola lalandi</i> (30)			■	■	■	■	■					
<i>Caulolatilus affinis</i> (34)	■	■	■	■						■	■	■
<i>Caulolatilus princeps</i> (34)	■	■	■								■	■

### 1.1.3.7 Graph theory and network analyses as a methodological approach for integrating functional and structural connectivity information

One of the most used approaches in ecology to represent connectivity in spatially complex seascapes is the graph theory and the network analyses (Bunn et al., 2000; Guimarães, 2020; Urban and Keitt, 2001; Urban et al., 2009). A graph represents a seascape in the form of a network in which links connect a set of nodes representing habitat patches. A link between two nodes indicates a functional connection (**Figure 6**).

From an ecological point of view, the nodes typically represent habitat patches in which local populations or communities inhabit (**Figure 6**). This representation appeals to the metapopulation or metacommunity models of the habitat mosaic (Hanski and Gilpin, 1991; Urban and Keitt, 2001). The nodes can be annotated with attributes or information, e.g., their spatial coordinates, genomic or ecological

diversity, etc. On the other hand, the links represent functional connectivity between habitat patches (or local populations and communities). They can represent a distance (e.g., genomic differentiation or species turnover) or the likelihood or rate of dispersion (e.g., larval dispersal probability between local populations). In a graph in which the links represent distance or a rate of dispersal, the links are weighted (i.e., the link contains the quantitative value being represented) (Urban et al., 2009) (**Figure 6**).



**Figure 6.** Top: a simple unweighted graph. Bottom: a weighted graph with nodes attributed with their sizes, arrows indicating the direction of the flow, and the line thickness indicating the magnitude of the fluxes (Image taken from Urban et al., 2009).

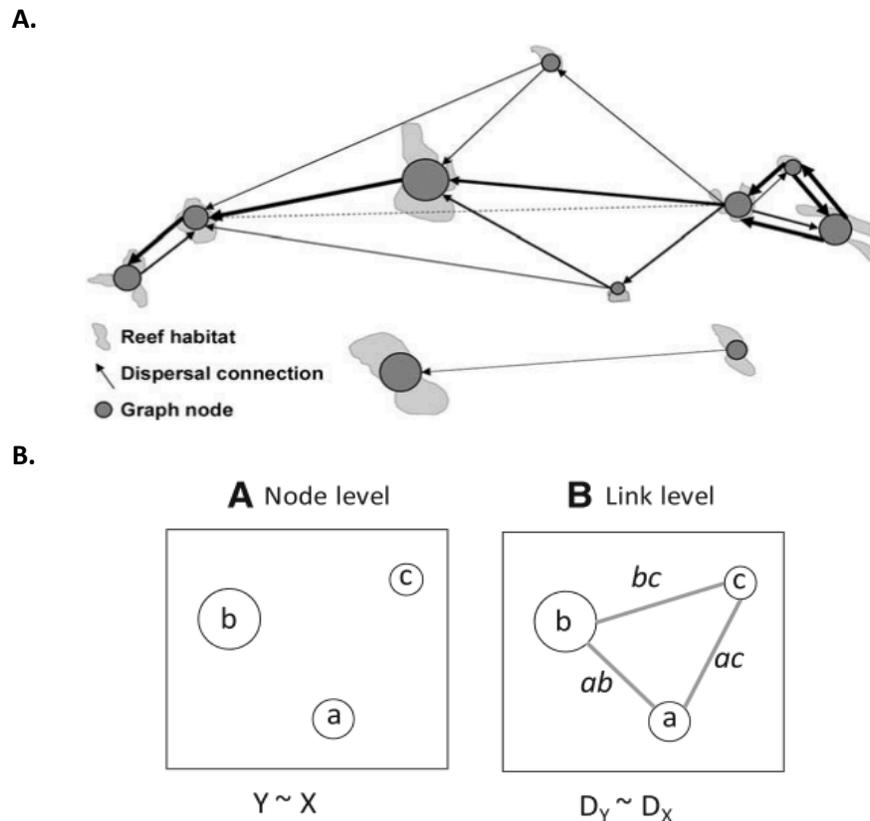
When constructing a network, we need to define two data structures describing its nodes and links (Hock and Mumby, 2015). First, the nodes are a set provided by an array or list. Additionally, we require a matrix that summarizes connections between nodes. here are three types of matrices:

- *distance matrix* (represent a functional distance, e.g., genomic differentiation)
- *probability matrix* (probability that an individual in node  $i$  will disperse to the node  $j$ , e.g., probability of larval dispersal)
- *adjacency matrix* (binary matrix of connected or non-connected elements)

Derived from the network configuration, multiple metrics are available to describe connectivity patterns at the node, link, and network levels (further described in the Methods section) (Costa et al., 2017), e.g., node centrality measures can be estimated to quantify local and regional node position and its importance in the seascape network (Borthagaray et al., 2015; Economo and Keitt, 2010).

There are two general classes of operations on networks, classified as *node* or *link* (**Figure 7**). In the present thesis, I will employ these two analytical levels to relate *functional connectivity* (i.e., empirical data) to *structural connectivity* (i.e., oceanographic modeling and spatial and environmental characteristics) (Wagner and Fortin, 2013) (**Figure 7**). *Node level* will relate the genomic diversity of local

populations and the alpha diversity of local communities to the node centrality metrics of the corresponding habitat patches obtained from the network analyses. This level of analysis addresses the question if there is a relationship between the diversity metrics and the centrality of the node network (**Figure 7**). On the other hand, *link-level* will relate pair-wise population genomic differentiation ( $F_{ST}$ ) between local populations and species turnover (beta diversity) to the seascape distances (e.g., Euclidean distance, resistance distance, environmental distance). This level of analysis will address the question if there is a relationship between the functional and structural connectivity's (**Figure 7**)



**Figure 7.** A. A graph-theoretic illustration of marine connectivity. Nodes represent reef habitats within the graph framework. When larvae from a source reef reach a downstream reef site, a dispersal connection is made. This dispersal connection and direction are represented by an arrow or 'edge' within the graph. The arrow's thickness reflects the strength of connection, i.e., the proportion of larvae that moves between the two nodes (from Trembl et al., (2008)). B. Left panel: A at the node level, the allelic richness or species richness ( $Y$ ) at each sampling location  $a$ ,  $b$  or  $c$  are related to environmental conditions or landscape features ( $X$ ) observed at the exact location. Nodes of different sizes refer either to different richness values. Right panel: at the link level, genetic distance, or species spatial turnover  $D_Y$  between pairs of sampling locations  $ab$ ,  $ac$ , and  $bc$  is related to distance-based landscape data  $D_X$  describing the intervening matrix along each link (Modified from Wagner and Fortin, 2013).

Most of the studies applying graph theory to evaluate seascape connectivity in the GC have focused on fish metapopulations. Their results have evidenced complex connectivity patterns related mainly to oceanographic characteristics. Consequently, besides analyzing metapopulation data, it would be informative to apply graph theory and network analyses to empirical data at the metacommunity level in the rocky-reef fishes to assess its association to the oceanography of the Gulf.

**Table 2.** Population, community, seascape and network connectivity data at node and link-levels.

Type of connectivity		Link-level		Node-level	
		<i>Seascape data</i>	<i>Biological data</i>	<i>Seascape data</i>	<i>Biological data</i>
Demographic	Functional/Potential	-	Movement (percentage of larvae dispersing/probability matrices)	-	Node centrality metrics
Genomic	Functional/Realized	-	Genomic differentiation ( $F_{ST}$ )	-	$H_e$ , $H_o$ , $A_r$ , $F_{IS}$
Community	Functional/Realized	-	Species turnover (beta diversity, Jaccard)	-	Alpha diversity
Euclidean	Structural	Euclidean distance	-		-
Environmental	Structural	Environmental distance	-		-
Oceanographic	Structural	Oceanographic distance	-		

## 1.2 Justification

Connectivity is a fundamental component of ecology and evolution and a target of conservation (Crooks and Sanjayan, 2006; Hanski and Ovaskainen, 2000; Kool et al., 2013) since it has been shown that the persistence of species, communities, and ecosystems can only be achieved if they are functionally connected (Resasco, 2019). In the ocean, the connectivity measurement is complicated because of practical difficulties in quantifying organisms' dispersal rates (Virtanen et al., 2020), consequently, direct and indirect have been used to characterize connectivity patterns across the sea.

The GC is a global marine biodiversity hotspot due to many species inhabiting it (Brusca et al., 2010). It's a very productive region supporting more than 40 % of total annual fisheries production in México (Cisneros-Mata, 2010). These fisheries are dependent on rocky reefs (Erisman et al., 2012). For several decades, fishing stocks worldwide have been declining or depleted due to overfishing (Pauly et al., 2005), with ecological consequences such as modifying the reef's general functioning. Also, there is increasing evidence that changes in species distribution and oceanographic patterns caused by climate change are occurring, which in turn will influence connectivity and, therefore, should receive more attention in future conservation planning and large-scale population and community evaluations (Andrello et al., 2015; Ayala-Bocos et al., 2015).

It's clear that maintaining biodiversity and ecosystem functioning is now the primary focus of conservation (Estes et al. 2011). In this sense, although advances in the study of connectivity of rocky reef fish populations and communities in the GC have been made, baseline information about the prevalent patterns in this critical group it's still scarce and needs to be further investigated. The present thesis represents a significant effort to integrate connectivity information on rocky reef ecosystems and their associated fishes at a regional scale, focusing on processes occurring across organization levels. This study expands upon early work of the leopard grouper (Jackson et al., 2015; Munguia-Vega et al., 2014) by using single nucleotide polymorphisms (SNPs) derived from restriction site-associated DNA sequencing to improve the statistical resolution for detecting population genomic structure, and including sampling sites from the north and central GC that were not considered in these previous investigations. Further, the evaluation of the potential larval connectivity of the leopard grouper and the reproductive assemblages among local communities is relevant because few species of the GC have been demographically evaluated to this extent with the incorporation of an oceanographic numerical model and network analysis. The resultant information informs the ecologically crucial regional connectivity processes for fish species

associated with rocky reefs. The integrative approach of the present investigation represents a significant effort to combine into one study information obtained from different sources (empirical data, oceanographic models, and graph theory). This gives us a novel view of the GC's ecological processes and provides information about the connectivity patterns that determine biodiversity at different spatial and temporal scales. Also, it helps to expand our insights into the complex relationship between community patterns and oceanography.

Furthermore, the application of a novel method for species detection as complement to a traditional method allowed provided critical information to evaluate the performance of eDNA in species-rich ecosystems and to investigate potential sources of bias, contributing to the proof-of-concept in the application of eDNA metabarcoding as a standardized high-throughput method for marine fish monitoring.

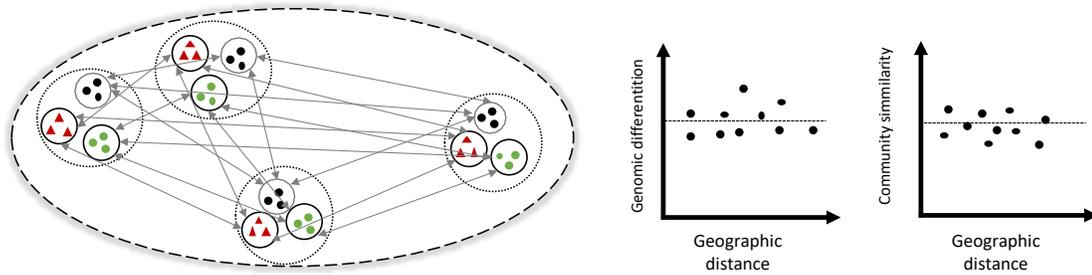
With this thesis, I provide a baseline for future conservation studies focused on implementing management strategies and designing connected networks of marine reserves to protect rocky reef ecosystems and their associated fishes in the GC.

### **1.3 Hypotheses**

Within the framework of this thesis, the hypothesis considered are:

The simplest way to understand functional connectivity is a scenario in which local populations and communities are connected by a constant dispersal of individuals. This means that there is a general homogeneity in the environmental conditions of the seascape matrix, which do not impose any barrier to dispersal. I assume that all local populations and communities are open and the abundance of individuals in each local population, and species in local communities are constant. This will serve as the null hypothesis for the present thesis.

**Hypothesis 0:** Functional connectivity is random in space due to unrestricted dispersal of individuals across the seascape matrix and to the homogeneity in the environmental condition of the GC (**Figure 8**).

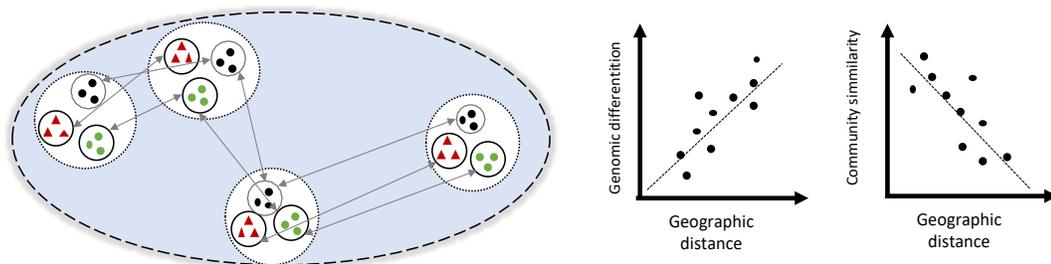


**Figure 8.** Illustration for hypothesis 0.

**Prediction 0:** At the *node* level, the genomic diversity metrics of the local populations, and the alpha diversity at local communities will be homogeneous in all localities. At the *link* level, I expect to find no-discernable functional connectivity patterns at metapopulation and metacommunity levels, i.e., I expect that measures of genomic differentiation (population) and species turnover (communities) will not be correlated with any structural seascape predictors.

The most straightforward step away from the null hypothesis is one in which the dispersal capacity of the organisms (individuals or species) and the habitat array will determine the functional connectivity, therefore dispersal will occur mainly between neighboring local populations and communities. In this sense, the genomic differentiation between local populations will increase with the geographic distance causing *isolation by distance pattern (IBD)*, or a *distance decay* of similarity for local communities.

**Hypothesis 1:** Functional connectivity will decrease with increasing geographic distance (**Figure 9**).

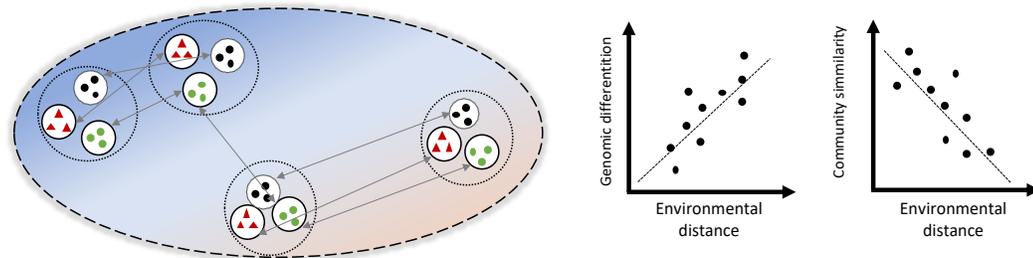


**Figure 9.** Illustration for hypothesis 1.

**Prediction 1:** At the *node* level, lower values of genomic diversity and alpha diversity will be found in local populations and communities that are more geographically isolated. At the *link* level, genomic differentiation and community dissimilarity will increase with geographic distance.

Dispersal potential may vary due to gradients of environmental factors. This environmentally induced isolation has been termed isolation by environment, IBE. The pattern of local genetic differences that can accumulate due to the local environment.

**Hypothesis 2:** Seascape environmental discontinuity restricts functional connectivity (**Figure 10**).

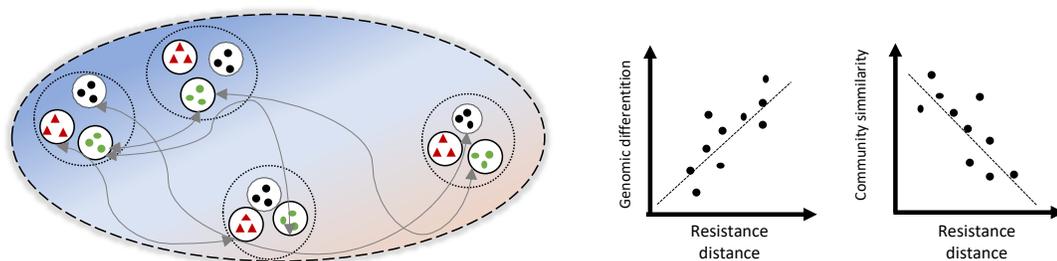


**Figure 10.** Illustration for hypothesis 2.

**Prediction 2:** At the *node* level, lower values of genomic diversity and alpha diversity will be found in local populations and communities that are more environmentally isolated. At the *link* level, genomic differentiation and community dissimilarity will increase with geographic distance.

Although IBD analyses traditionally use straight-line (Euclidean) geographic distances, other ecologically relevant distance measures, such as a resistance distance (derived from an oceanographic three-dimensional numerical model) may capture functional connectivity processes affecting gene exchange and, therefore, be better predictors of genetic differentiation than Euclidean geographic distance. Therefore, an *isolation by resistance* (IBR) pattern may manifest in organisms with a reproductive strategy that involves a pelagic larval stage subject to transport by ocean currents. Local genetic differences can accumulate due to the landscape resistance, which results from ocean currents filtering dispersal.

**Hypothesis 3:** Functional connectivity is determined by ocean currents (**Figure 11**).



**Figure 11.** Illustration for hypothesis 3.

**Prediction 3:** At the *node* level, lower values of genomic diversity and alpha diversity will be found in local populations and communities that less oceanographically connected (i.e., lower values of network centrality metrics). At the *link* level, genomic differentiation and community dissimilarity will increase with the resistance distance. I expect that measures of genomic differentiation and community similarity to be correlated with least-cost oceanographic distance, which is not necessarily correlated with Euclidean geographic distance.

## 1.4 Objectives

### 1.4.1 General objective

To identify patterns of functional connectivity at the population and community levels of organization of rocky-reef fishes to investigate its relationship with the structural seascape connectivity of the GC using a graph-theoretic approach at node and link levels.

### 1.4.2 Specific objectives

1. To evaluate the functional connectivity at population and community organization levels, I aimed to:
  - a. Evaluate the population genomic diversity and differentiation ( $F_{ST}$ ) in the leopard grouper (*Mycteroperca rosacea*) through the analysis of Single Nucleotide Polymorphisms (SNPs).
  - b. Estimate the alpha (species richness) and beta diversities (species turnover) in communities of rocky reef fishes by characterizing them with two complementary monitoring methods.
2. To evaluate the potential demographic connectivity among the leopard grouper local populations and the reproductive assemblages of rocky reef fishes using the HAMSOM oceanographic numerical model.
3. To evaluate the *structural connectivity* of the seascape by estimating geographic, environmental, and oceanographic distances using geographic information systems, public databases of environmental variables, and the output of an oceanographic model (HAMSOM).
4. To integrate the information about *functional* and *structural connectivity* of rocky reef fishes using spatial and multivariate statistical methods and a graph-theoretic approach with node and link-level analyses.

## 1.5 Study objects

### 1.5.1 The Gulf of California seascape

The GC is a marginal sea of the eastern Pacific Ocean, located between the continental part of Mexico and the Baja California peninsula. It is a semi-closed basin due to its meteorological and oceanographic characteristics since it is almost surrounded by an elevated topography and connected with the open ocean only at its southern end (Soria et al., 2014). These properties, together with the fact that it is a tropical-subtropical transition zone, result in a complex atmospheric and oceanographic environment with strong variations in physical and biological processes (Lluch-Cota, 2000; Lluch-Cota et al., 2007), making it a natural laboratory to study the effect of the environment on connectivity.

The GC is bathymetrically divided into a series of basins and trenches, with a maximum basin depth at the Gulf entrance of >3,000 m but as shallow as ~200 m in the northernmost region. The peninsular shore is primarily rocky, scattered with sandy stretches and a narrow shelf. On the other hand, the mainland shore is characterized by long sandy beaches, large coastal lagoons, and open muddy bays. In general, the GC is characterized by presenting a wide temporal variability in all physical-environmental characteristics such as temperature, ocean circulation, winds, upwelling, and productivity (Avendaño-Ibarra et al., 2013; Lavín et al., 2003), and based on its oceanographic characteristics it can be divided into three main regions: north GC, central GC and south GC (Soria et al., 2014).

Circulation in the GC is strongly seasonal due to the characteristics of the principal forcing's: the Pacific Ocean (Ripa, 1997; Ripa, 1990), the wind regime, and the heat flows (Beron-Vera and Ripa, 2000). These conditions impose characteristics that are particular to the Gulf, such as the circulation in the northern region, which is anticyclonic from November to March and cyclonic from June to September, as well as transition periods that occur from April to May and in October (Carrillo and Palacios-Hernández, 2002). In contrast, the flows are intense in the large island region due to water exchange between the northern and southern regions (Beier, 1997; Mateos et al., 2006). This region is distinguished by an intense tidal mixing (Argote et al., 1995) modulated by semi-diurnal, diurnal, and fortnightly frequencies (Lavín and Marinone, 2003) as well as a branching of deep flow that typically moves north. One branch flows toward the Ballenas–Salsipuedes Channel (BC) through the San Lorenzo sill, and the other flows through the San Esteban sill. The latter surrounds Ángel de la Guarda Island and converges with the other branch in the Ballenas–Salsipuedes Channel, thus producing a persistent upwelling (Marinone et al., 2008).

In the southern region, a train of eddies has been reported in both models and observations (Lavin et al., 2014; Marinone, 2003; Zamudio et al., 2008). Marinone (2003) identified a series of anticyclonic (April to May and October to November) and cyclonic (June to August) eddies in the southern region as well as a quasi-permanent eddy in the San Pedro Mártir Basin. Zamudio et al. (2008) studied the generation of eddies using a three-dimensional model with local and/or remote forcing. They concluded that the remote forcing is essential for the generation of eddies and that the mechanism of formation occurs by the interaction of the near-coastal poleward eastern boundary currents with coastline and topographic irregularities.

The dynamic and circulation characteristics in the GC imply various degrees of connection (i.e., connectivity) among the different regions. Quantifying this connectivity among the Gulf regions is helpful for understanding the transport of fish larvae between continental and peninsular coasts, as documented previously (Contreras-Catala et al., 2012; Sánchez-Velasco et al., 2013) due to the presence of eddies and dispersal patterns of some marine protected areas.

### 1.5.2 The leopard grouper

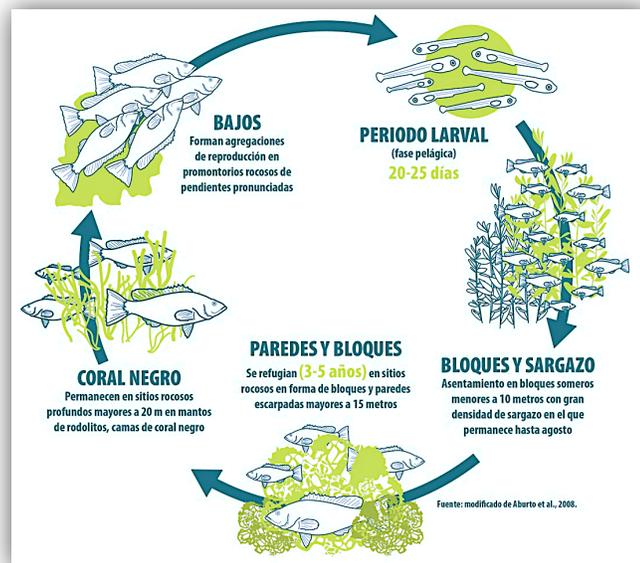
The leopard grouper, *Mycteroperca rosacea* (Streets 1877) (Serranidae), is an abundant top predator that, as an adult, inhabits patchy rocky reefs ecosystems from 1-50 meters depth (mean 16-18 m; Sala et al., 2003; Tinhán et al., 2014). It can be found over a latitudinal gradient of environmental characteristics (Escalante et al., 2013), from Bahía Magdalena on the Pacific coast of the Baja California Peninsula (Sala et al., 2003; Thomson et al., 2000) to the waters of the GC. Adults are most abundant on several microhabitats associated with rocky reefs, including boulder fields, black coral groves, rhodolite beds, and rock walls (Aburto-Oropeza and Balart, 2001). Individuals can reach 1 m in length and at least 22 years of age (Díaz-Urbe et al., 2001).

The species shows a group-spawning mating system in which site-specific aggregations of several hundred individuals (Estrada-Godínez et al., 2011) persist for extended periods during the reproductive season (Erisman et al., 2007) (**Figure 12**). There is evidence that seasonal patterns of marine productivity (i.e., high rates of primary productivity and plankton concentrations) influence the reproductive timing of the leopard grouper (Escalante et al., 2013; Kahru et al., 2004; Lluch-Cota et al., 2007). In this context, the spawning season of the leopard grouper has been shown to vary latitudinally (Pérez Olivas, 2016; Sala et

al., 2003) (**Table 3**): from May to July in Bahía de La Paz (Estrada-Godínez et al., 2011); from April to June in Loreto (Erisman et al., 2007); from March to May in Santa Rosalía (Pérez Olivas, 2016); from March to May in the Southern GC (the study doesn't specify localities) (Sala et al., 2003); only in May in the Midriff Islands (Sala et al., 2003); and from March to June in the Midriff Islands ( Munguia-Vega et al., 2014). After hatching, the larvae of leopard grouper possess a PLD of 24-28 days (Aburto-Oropeza et al., 2007; Adrian Munguia-Vega et al., 2014). Recruitment is positively related to the availability of suitable nursery habitat, i.e., *Sargassum* beds on shallow boulders preferentially during times of peak biomass (Aburto-Oropeza et al., 2007).

**Table 3.** Temporal variation in the reported leopard grouper's reproductive seasons in different areas of the GC. The numbers correspond to the number of each month (1: January ... 12: December).

1	2	3	4	5	6	7	8	9	10	11	12	Study area	Reference
				*								Midriff islands	Sala et al., 2003
		*	*	*	*							Midriff Islands	Munguía-Vega et al., 2014
		*	*	*								Santa Rosalía	Perez-Olivas 2016
			*	*	*							Loreto	Erisman et al., 2007
				*	*	*						Bahía de la Paz	Estrada-Godínez et al., 2011
		*	*	*								Southern GC	Sala et al., 2003



**Figure 12.** Leopard grouper (*Mycteroperca rosacea*) life-cycle (Image from <https://datamares.org/perfil-de-especie-cabrilla>).

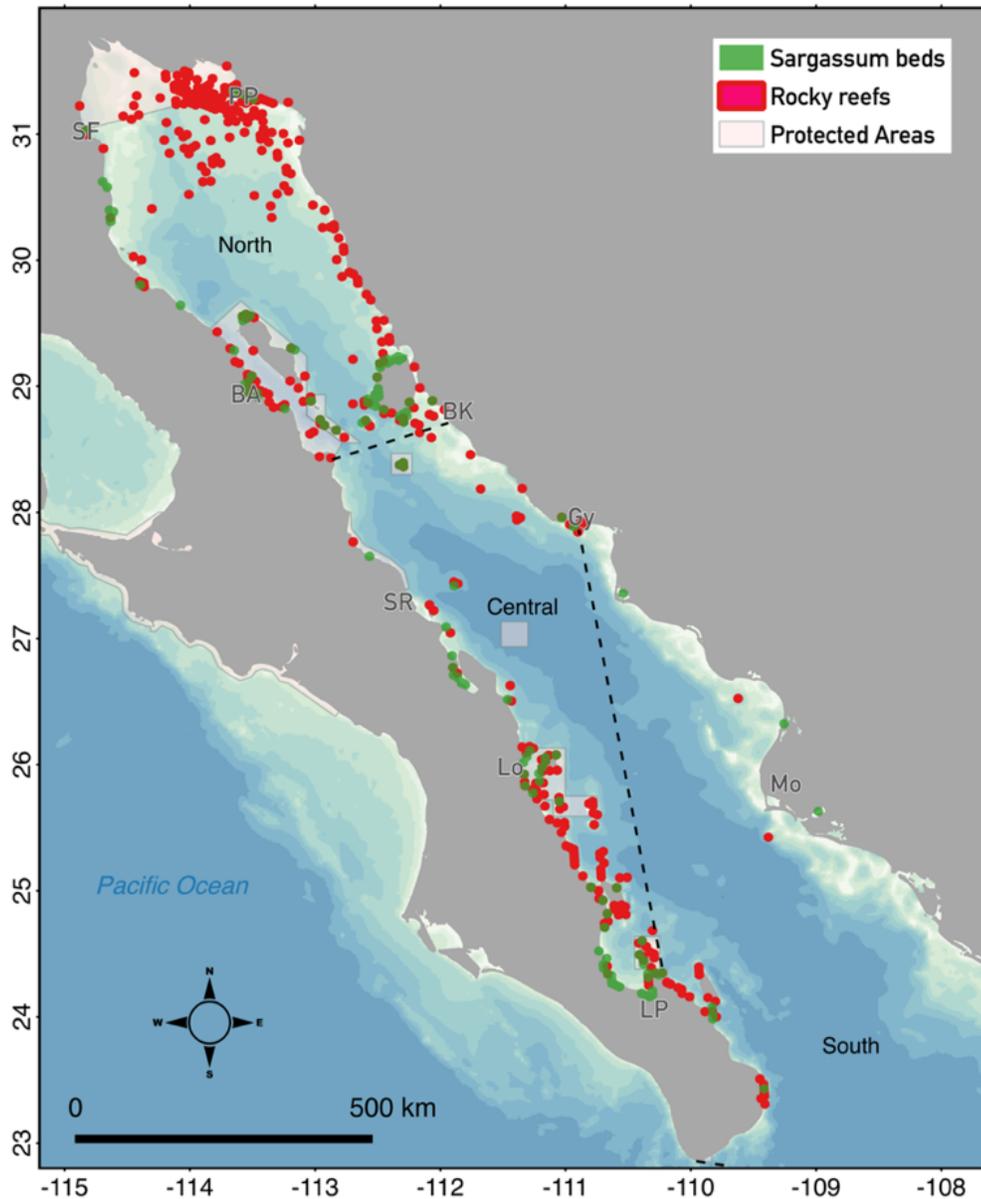
According to the IUCN (International Union for Conservation of Nature), this species shows a decreasing population trend. The IUCN also cautions that if unsustainable fishing pressure continues, it could become a significant threat to its global population, as no fishing regulations exist in Mexico.

### **1.5.3 Rocky reefs of the Gulf of California and its associated fishes**

In the GC, rocky reefs are a dominant feature in more than 900 islands and islets, along with shallow Gulf areas and submerged seamounts (**Figure 13**). These reefs are distributed irregularly along Gulf coasts and represent the main coverage field habitat type (Munguia-Vega et al., 2018). Their bathymetric distribution ranges from the intertidal zone to ~300 m although most of the habitat and its associated biodiversity is found in the first 20–30 m (Sánchez-Rodríguez et al., 2015). They support fisheries and provide multiple ecosystem services (Sánchez-Rodríguez et al., 2015).

One of the main biological components of the rocky reef ecosystems is the group of fishes that spend most of their time on reefs to feed, shelter, or reproduce (Thomson et al., 2000). According to Thomson et al. (2000), the GC hosts 821 fish species, of which 30% are bony fish species (Actinopterygii: Teleostei) (Brusca et al., 2005; Thomson, Findley, & Kerstitch, 2000) inhabit rocky reefs ecosystems.

Biogeographically, the GC lies at the intersection of the temperate and tropical faunal regions of the Eastern Pacific Ocean. Therefore its ichthyofauna represents a mixture of widespread and endemic species (Brusca et al., 2010, Chapter 5). It has been shown that the ichthyofauna distribution in the GC responds to a latitudinal gradient: (a) temperate species are more abundant in the northern GC, (b) tropical affinity species are more abundant in the south, and (c) an apparent mixture of temperate-tropical fauna has been found in the central GC zone (Avendaño-Ibarra et al., 2013). This pattern is determined by cold winter water temperatures of the northern GC, which limit the survivorship of some tropical fishes; also, it generally lacks habitats found in the central and southern parts, especially the deeper ocean basins; besides the absence of a direct connection with marine waters to the north and south, which precludes the easy movement of temperate fishes into the region (Brusca et al., 2005).



**Figure 13.** Faunistic regions: Northern, Central, and Southern are shown. Also, known rocky reefs (adult habitat of reef fishes) and sargassum beds (recruitment habitat for *M. rosacea* and other rocky reef fishes), and marine protected areas of the GC. Principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN). Data obtained from Munguia-Vega et al., 2018.

Ecologically, according to its species composition, the GC has three major subregions: northern, central, and southern GC (Brusca et al., 2005) that extend: from the Colorado River Delta southward to the Midriff Islands (northern GC); from the southern limit of the Midriff Islands to Guaymas (Sonora) and Punta Coyote (Baja California Sur) (central GC); and from the southern boundary of the central GC southward to Cabo Corrientes (Jalisco) on the mainland to Cabo San Lucas (Baja California Sur) (southern GC) (Brusca et

al., 2005) (Figure 13). Also, previous studies on communities of rocky reef fishes of the GC have found gradients in species composition and diversity, with the lower values at higher latitudes (Sala and Aburto-Oropeza, 2002). Viesca-Lobatón et al. (2008) found that the dominant fish species changed according to the latitude and that there is a seasonality of species composition across the GC. Ramírez-Ortiz et al. (2017) found high fish species biomass and richness and a decreasing pattern of species richness towards the tropics. Also, Fernández-Rivera Melo et al. (2018) showed a latitudinal variation in the community species composition of reef fishes along the western GC, with the highest values of species richness, diversity, taxonomic distinctness, and trophic level in the southern GC, probably due to a combination of environmental conditions, a more significant number of habitats and more functional diversity of the assemblages. Based on the species occurrence and their abundance, the authors identified that the northern GC presents the lowest values of species richness, abundance, and taxonomic distinctness; the central GC has intermediate ecological complexity and can be classified as a transition zone from tropical to temperate reefs; and the southern GC, with the highest values of all diversity indicators. Ultimately, Olivier et al. (2018) identified the Loreto-Cerralvo corridor as the host of the majority of the reef-fish diversity in the GC and emphasized that GC is not a homogenous inner sea and presents important oceanographic divergence along its latitudinal axis, which translates into a functional divergence between fish assemblages of the northern and southern GC, each zone favoring different ecological trait combinations.

## Chapter. 2 Methodology

---

### 2.1 Evaluation of the functional connectivity at population and community organization levels

#### 2.1.1 Population-level functional connectivity: evaluation of the genomic diversity and differentiation ( $F_{ST}$ ) in the leopard grouper (*Mycteroperca rosacea*) through the analysis of Single Nucleotide Polymorphisms (SNPs)

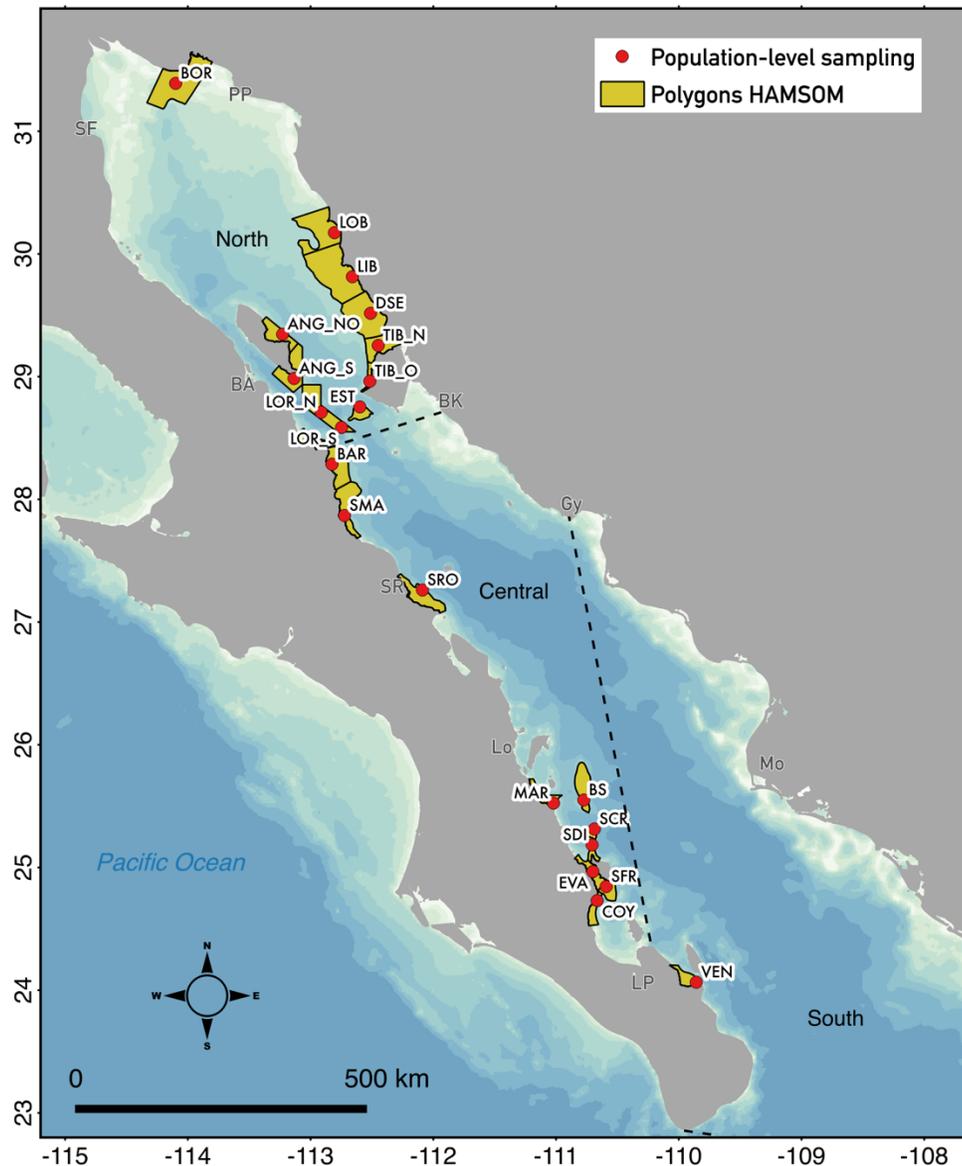
##### 2.1.1.1 Tissue sampling for population genomics analysis

Leopard grouper sampling was performed during 2016, 2017 and 2019, and was designed to maximize sampling coverage and balance for individual representation in the polygons of the oceanographic model. Individuals were collected with the help of small-scale fishers who contributed with taking the samples, and with the annotation of the spatial coordinates of each individual and its size (individuals size ranged varied from 25-86 cm). After individual collection, fin tissue was immediately stored in 96% ethanol until DNA extraction in the laboratory. Tissue was collected from 216 leopard grouper individuals across 22 sample localities along the GC (~10 individuals per locality). Each sampling locality corresponds to a polygon in the HAMSOM model (**Table 4** and **Figure 14**).

##### 2.1.1.2 Genomic DNA extraction and ddRADseq library construction

In the laboratory, genomic DNA was extracted using a salting-out protocol (Aljanabi and Martinez, 1997) with RNase treatment (Thermo Scientific), and cleaned with 0.4X Ampure XP (Beckman Coulter). DNA was resuspended in molecular-grade water, and concentration was determined with HS assay kit using a Qubit 2.0 fluorometer (Life Technologies). DNA integrity was assessed by electrophoresis in 1.5% agarose gels. Restriction site-associated DNA (ddRAD, Annex B **Figure 39**) libraries were constructed using the Peterson and collaborators (2012) protocol with some modifications. 1) genomic DNA (300-400 ng) of each individual sample was double-digested in a 50  $\mu$ L reaction using 20 U of EcoRI-HF<sup>®</sup>, 20 U of MspI restriction enzymes (NEB), and 1X CutSmart Buffer 10X; 2) P1-EcoRI (sample barcode) and P2-MspI adapters (0.1  $\mu$ M final concentration each) were ligated with 100 U of T4 ligase (NEB) in a 40  $\mu$ L reaction;

3) individual libraries were pooled in groups of 12 individual samples for indexing (5 pools in total); 4) target fragments of 376 bp were selected in each pool using a Pippin Prep (Sage Science); 5) Illumina indexes (PCR\_primers) were incorporated to each pool (0.5  $\mu$ M each) in eight 20  $\mu$ L PCR reactions/pool adding 1X Buffer Phusion HF 5X, 0.4 mM dNTPs (Thermo) and 0.5 U of Phusion DNA polymerase (NEB) using ten cycles of amplification. Cleanup was performed after each step using 1.5X Ampure XP (Beckman Coulter). Finally, an equimolar combination of the indexed pools was made and sent to Novogene, CA (<https://en.novogene.com>) to be paired-end sequenced (2x150 bp) on an Illumina HiSeq system.



**Figure 14.** Leopard grouper sampling of sites for population analyses (see **Table 4** for site-specific details) in the GC (red dots) and the corresponding polygons from the HAMSOM model (yellow polygons) are represented. GC regions are labeled and delimited with dashed lines. Principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN).

**Table 4.** Leopard grouper's sampling localities grouped in their respective GC regions. Table shows the Locality number (Loc), Locality name, the Locality ID (ID), Number of genotyped samples (N), and the geographic location (Longitude, Latitude).

Loc	Locality name	ID	N	Lon	Lat
1	Puerto Peñasco (Borrascoso)	BOR	11	-114.100	31.390
2	Puerto Lobos	LOB	8	-112.806	30.173
3	Puerto Libertad	LIB	7	-112.660	29.814
4	Desemboque Seri (0-100m)	DSE	10	-112.512	29.516
5	Isla Tiburón (North)	TIB_N	8	-112.449	29.254
6	Isla Tiburón (West)	TIB_O	11	-112.518	28.963
7	Reserva Isla Ángel de la Guarda (South)	ANG_S	10	-113.135	28.985
8	Reserva Isla Ángel de la Guarda (Nortwest)	ANG_NO	10	-113.228	29.345
9	Reserva Isla San Lorenzo (South)	LOR_S	10	-112.748	28.586
10	Reserva San Lorenzo (North)	LOR_N	10	-112.914	28.710
11	Isla San Esteban	EST	10	-112.599	28.753
12	El Barril	BAR	6	-112.826	28.286
13	Santa Maria	SMA	5	-112.724	27.868
14	Santa Rosalía	SRO	7	-112.089	27.261
15	San Marcial	MAR	8	-110.774	25.549
16	Bajo Seco	BS	10	-110.707	25.183
17	Isla Santa Cruz	SCR	8	-111.021	25.524
18	Isla San Diego	SDI	10	-110.688	25.313
19	San Evaristo	EVA	8	-110.701	24.966
20	Isla San Francisco	SFR	7	-110.592	24.844
21	Coyote	COY	10	-110.666	24.730
22	Bahía Ventana	VEN	10	-109.858	24.064

### 2.1.1.3 Bioinformatic analyses: SNP calling

Demultiplexing of raw reads and quality filters were implemented using the *process\_radtags* program in STACKS v2.5 (Catchen et al., 2011). Read-pairs with the expected restriction sites and barcodes at both ends were identified, tolerating one error in each barcode (default value). The subsequent steps of the bioinformatic followed the *denovo\_map.pl* steps in STACKS v2.5. The parameter combination ( $M$ ,  $m$ ,  $n$ ) that maximized the number of loci generated was identified according to Paris et al. (2017). The final parameter combination to build loci was 3 for the minimum required read coverage depth to form a stack or group of identical reads ( $m$ ); 3 for the maximum number of mismatches allowed between stacks or groups of identical reads, to be considered as different alleles of the same locus ( $M$ ); and 4 for the maximum number of mismatches between loci from different individuals to be considered homologs ( $n$ ). Then, data were filtered with *populations* program, retaining loci present in 50% of the individuals ( $-r = 0.5$ ), and one SNP per loci with the option *write\_single\_SNP* to avoid linkage among loci.

### 2.1.1.4 Bioinformatic analyses: Data filtering and identification of neutral and candidate SNPs loci

Additional filtering steps were performed to reduce the calling of false SNPs due to artifacts of the ddRADseq approach (i.e., multicopy loci, paralogs, and sequencing errors). First, I excluded loci with > 20% of the missing data and retained loci with a maximum mean depth (max-meanDP) of 100 (to avoid possible paralogs since preliminary results without removing them yielded a high number of SNPs identified to be under balancing selection). This first dataset was analyzed to determine candidate loci under natural selection.

SNPs potentially under selection were identified using three independent approaches: BayeScan v2.1 (Foll and Gaggiotti, 2008), OutFLANK v0.2 (Michael C. Whitlock and Lotterhos, 2015), and *pcadapt* v4.3.3 (Luu et al., 2017) R packages. Briefly, Bayescan uses a Bayesian method based on a logistic regression model that separates locus-specific effects of selection ('adaptive' genetic variation) from population-specific effects of demography ('neutral' genetic variation). I ran 10,000 iterations and a burn-in of 200,000 steps, and I specified a prior odd of 100 to minimize false positives. I considered candidate SNPs those with a q-value below 0.05. Second, OutFLANK v0.2 calculates a likelihood based on a trimmed distribution of  $F_{ST}$  values to infer the distribution of  $F_{ST}$  for neutral markers. This test was ran using the

default options (LeftTrimFraction = 0.05, RightTrimFraction = 0.05, Hmin = 0.1), and identified outlier SNPs based on the q-threshold of 0.05. Finally, *pcadapt* performs genome scans to detect loci under selection assuming that candidate markers are outliers, concerning to how they are related to population structure when evaluated with principal component analysis using the Mahalanobis distance computed for each SNP. For this analysis, I applied the default settings and a min.maf value of 0.01. A SNP was considered a candidate for natural selection if two analyses identified it as an outlier. Using these results, I divided the data set into two categories: Neutral and Candidate.

As this thesis main objectives are related to the study of connectivity of leopard grouper populations, all the posterior analyses were made using only the Neutral dataset, which has great potential for investigating processes such as gene flow, migration, or dispersal. Hence, they allow us to empirically test the functional relevance of spatial indices used in seascape ecology (Holderegger et al., 2006).

#### **2.1.1.5 Geographical distribution of the genomic diversity and evaluation of the population genomic differentiation**

To evaluate the genomic diversity of the leopard grouper per sampling locality, I estimated the following parameters: rarefied allelic richness ( $A_r$ ) (to the minimum sampling number of 6 individuals), observed and expected heterozygosities ( $H_o$  and  $H_e$ , respectively), and Wright's inbreeding coefficient (FIS) with *hierfstat* v0.7-7 package (Goudet et al., 2020) in R v4.0.3 (R Core Team, 2021). Then, I used three approaches to evaluate the population genomic differentiation among the different sampled localities. First, using pairwise  $F_{ST}$  (Weir and Cockerham, 1984) implemented in *graph4lg* v1.2.0 package (Savary et al., 2021) in R v4.0.3 (R Core Team, 2021), and a hierarchical analysis of molecular variance (AMOVA) among the three GC regions: North, Central and South (20,000 permutations) with the software Arlequin v3.5.2.2 (Excoffier et al., 2005). Second, a maximum likelihood model with ADMIXTURE v1.3.0 (Alexander et al., 2015) with 2,000 bootstrap replicates. This algorithm assumes that each individual has ancestry from one or more K genetically distinct sources (admixture model, K-means) and that all members of the current population can breed. To determine the correct number of different populations (K) I used the cross-validation approach (Alexander et al., 2009). Finally, a multivariate discriminant analysis of principal components (DAPC) which is designed to identify and describe clusters of genetically related individuals and is implemented in the *adegenet* v2.1.3 package (Jombart et al., 2008) in R v4.0.3 (R Core Team, 2021). To assess the optimal number of genomic clusters, I used the Bayesian Information Criterion (BIC).

## **2.1.2 Community-level functional connectivity: estimation of the alpha (species richness) and beta diversities (species turnover) in communities of rocky reef fishes using complementary monitoring methods**

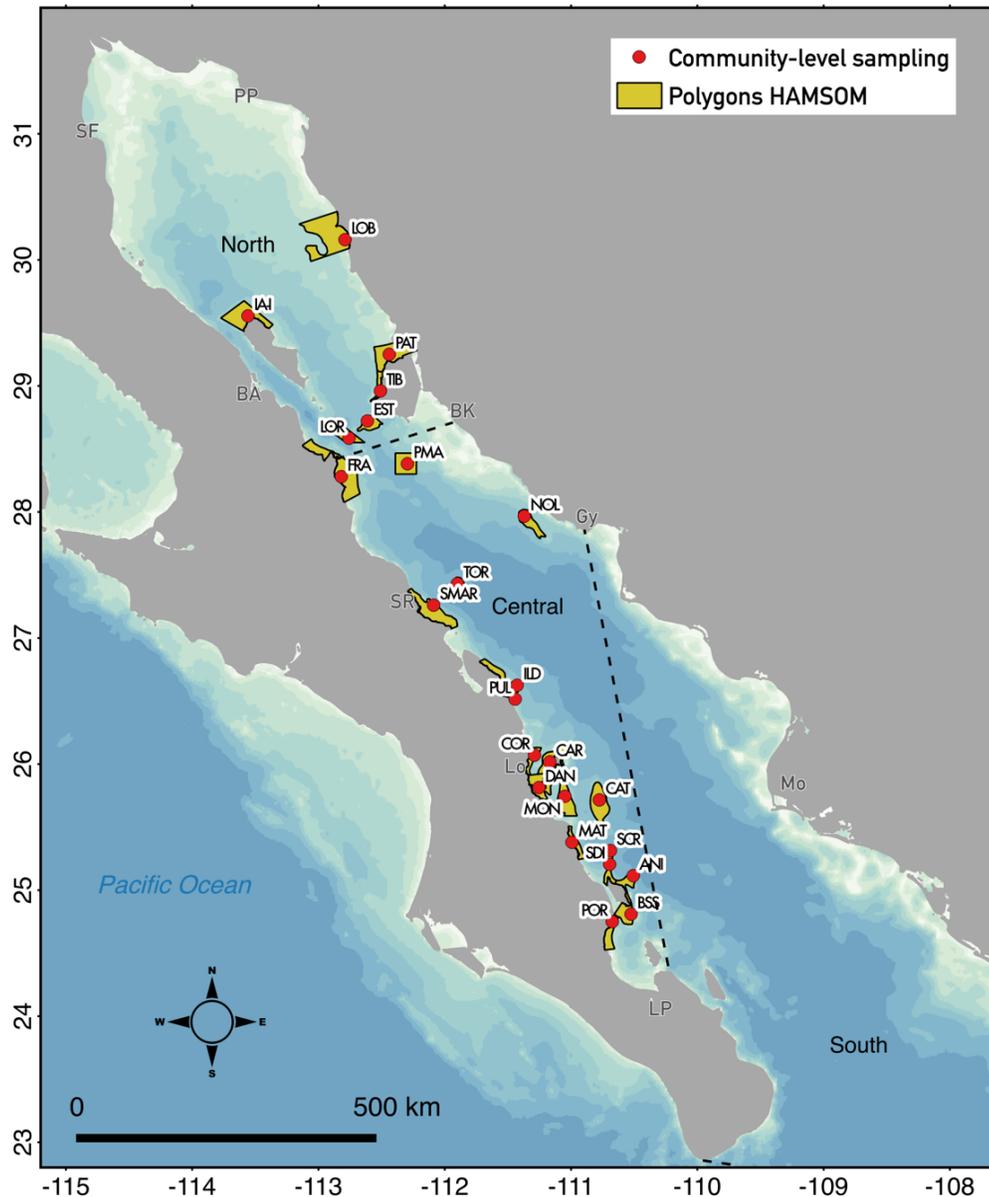
### **2.1.2.1 Underwater visual censuses and seawater sampling for eDNA metabarcoding analysis**

A scientific expedition was conducted from October 23 to November 15, 2016, on board a 90-foot vessel to perform UVC and to collect seawater samples for eDNA metabarcoding in rocky reefs across the GC. Our sampling approach was designed to maximize the biological signal while minimizing the sampling effort. To this end, we divided our study area into two non-overlapping strata based on known regions (north and central GC regions, hereafter) to sample 24 independent site (**Table 5** and **Figure 15**)

At each site, immersions were organized into two groups. The first group was composed of six divers that used SCUBA gear and extensive fish identification expertise to survey conspicuous fishes (adults > 5 cm total length) inside of 25 m-long transects, 4 m wide and 2 m high (i.e., 200 m<sup>3</sup> surveyed for each transect). Transects were placed at different depths ranging from 3 to 25 m, parallel to the coastline, or following the contour of the reefs (over seamounts), to cover as much habitat area as possible. The second group, composed of 2-3 divers, navigated underwater ~10 m away from the first group and sampled 1-L seawater using clean Nalgene™ Wide-Mouth HDPE Bottles (Thermo Scientific™). After the seawater collection, bottles were closed underwater and remained sealed until water filtration on board. In all cases, water was filtered < 4 h after collection, using 0.44 µm hydrophilic nitrocellulose Millipore® filters placed in a Millipore® Sterifil® filtration system connected to a manual vacuum pump. Each filter was removed from the filtration system, folded inwards, stored in 1.6 mL tubes (Neptune®) with STE sterile buffer (100 mM NaCl, 1 mM EDTA, 5 mM Tris/HCl pH 7.5) and preserved at room temperature until the end of the scientific expedition. Cross-contamination from diving equipment was prevented by handling eDNA samples in a dedicated area ~20 m away (on different decks of the research vessel) from where all the diving equipment and diving operations occurred. In addition, the water filtering step was performed in a decontaminated area using sterile material, cleaned properly between sampling sites, and between each water filtering step. Cleaning was accomplished by submerging the filtration systems and bottles in a 1% sodium hypochlorite solution for 2 hours and thoroughly rinsing with filtered fresh water.

**Table 5.** Sampling sites for eDNA seawater collection and UVC.

Site	Site Name	Site ID	Collection date	Lat	Long	Temp (°C)	Depth (m)	N transects
1	El Portugués	POR	10/29/16	24.749	110.674	28	6	14
2	Bajo Seco Sur	BSS	10/29/16	24.809	110.525	27	21	6
3	Bajo Las Ánimas	ANI	10/30/16	25.114	110.508	28	10	11
4	Isla San Diego	SDI	10/30/16	25.204	110.695	28	5	14
5	Isla Santa Cruz	SCR	10/30/16	25.313	110.689	28	11	10
6	Isla San Mateo	MAT	10/31/16	25.379	110.993	28	8	14
7	Isla Catalana	CAT	10/31/16	25.715	110.776	28	12	14
8	Isla Monserrat	MON	11/01/16	25.744	111.049	28	7	14
9	Isla Danzante	DAN	11/01/16	25.813	111.255	27	13	15
10	Isla Carmen	CAR	11/02/16	26.017	111.169	26	6	15
11	Isla Coronados	COR	11/02/16	26.174	111.262	27	9	10
12	Punta Púlpito	PUL	11/03/16	26.515	111.443	27	4	16
13	Isla San Ildefonso	ILD	11/03/16	26.625	111.427	27	9	18
14	Isla San Marcos	SMAR	11/04/16	27.260	112.089	25	8	16
15	Isla Tortuga	TOR	11/04/16	27.431	111.062	26	17	8
16	Isla San Pedro Nolasco	NOL	11/06/16	27.966	111.371	24	20	24
17	Isla San Pedro Mártir	PMA	11/07/16	28.382	112.296	21	13	16
18	Isla San Francisquito	FRA	11/08/16	28.441	112.268	23	10	8
19	Isla San Lorenzo	LOR	11/08/16	28.585	112.759	23	13	8
20	Isla Ángel de la Guarda	IA-I	11/09/16	29.555	113.559	24	11	16
21	Puerto Lobos	LOB	11/11/16	30.222	112.966	22	3	8
22	Isla Pato	PAT	11/12/16	29.266	112.464	19	9	8
23	Isla Tiburón	TIB	11/12/16	29.065	112.506	18	12	4
24	Isla San Esteban	EST	11/13/16	28.722	112.613	18	8	8



**Figure 15.** Sampling sites for UVC and seawater collection for eDNA metabarcoding across the GC (red dots) and the corresponding polygons from the HAMSOM model (yellow polygons) are represented. Northern and Central regions are labeled and delimited with dashed lines. Principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN).

#### 2.1.2.2 Custom reference database for the ichthyofauna of the GC: primer design, PCR amplification, and sequencing

I constructed a local reference database for the “teleo” 12S rRNA barcode (custom reference database, hereafter) (Valentini et al., 2016). Partial or complete teleost mtDNA sequences for the 12S rRNA gene were obtained from NCBI for species present in the GC and aligned along with the mitogenome of

*Cyprinus carpio* with the Muscle v3.8.31 plugin in Geneious Prime® version 2019.0.4. Primers Teleo12S\_1322-R, Teleo12S\_682-F, and Teleo12S\_792-F (Annex B **Table 18**) were designed with the Primer3 implementation in Geneious Prime® to amplify 530 - 640 bp regions of the 12S rRNA gene, surrounding the 65 bp region amplified with “teleo” primers (Valentini et al., 2016). DNA from 67 morphologically identified fish species (1-8 specimens per species) was extracted with a salt protocol (Aljanabi and Martinez, 1997) and then used for PCR amplification with the newly designed primers (Annex B **Table 20** for the species list). These specimens belong to commercial species collected in the GC over the last decade as part of the ecosystem-based project PANGAS (Munguía-Vega et al., 2015).

Amplifications were carried out on 25 µL reaction volumes containing 1x PCR Buffer Mg<sup>+</sup> free, 1.5 mM MgCl<sub>2</sub>, 0.8 mM dNTP mix, 0.3 µM of each primer, 1.5 U of DNA Taq polymerase (Invitrogen), 1% BSA and 2 µL of DNA template. The thermal cycling profile was denaturation at 95°C for 60 s; 30 cycles at 95°C for 30 s; annealing at 61°C for 15 s; and extension at 72°C for 15 s; and a final extension at 72°C for 5 min. Amplifications were visualized on 2% agarose gels stained with GelRed® Nucleic Acid Gel Stain (Biotum). After PCR amplification, products were sequenced (3' and 5' directions using the designed primers) on an Applied Biosystems 3730XL DNA Analyzer at the University of Arizona Genetics Core Laboratory. Sequences were edited using Chromas Pro v1.6 and aligned using MUSCLE multiple alignment tools implemented in Mega6 (Tamura et al., 2013). Annotation of the sequences was assisted using the MITOS web server (Bernt et al., 2013). The obtained sequences were used as a custom reference database for taxonomic assignment of sequence reads and registered in GenBank (Annex B **Table 19**).

### 2.1.2.3 Mock community

To test the detection sensitivity of metabarcoding in a sample of known composition, I included a mock community in which equal amounts (200 ng) of purified DNA from 22 reef fishes of the GC (Annex B **Table 20**) were pooled and subsequently used as a positive control. This sample was incorporated into the library construction process and sequenced parallel to eDNA samples.

#### 2.1.2.4 Environmental DNA extraction

eDNA extractions were completed with the Qiagen Blood and Tissue Kit (Qiagen, USA) with some modifications (Annex B **S1**). All eDNA extractions were performed in a hood dedicated solely for this purpose. The hood and pipettes were cleaned with 1% sodium hypochlorite solution and treated with UV light for 20 min before starting extractions. Filter tips were used in all pipetting to reduce the risk of cross-contamination. DNA concentrations were measured using a Qubit 3.0 Fluorometer (Invitrogen, CA, USA).

#### 2.1.2.5 Library preparation and high-throughput sequencing

I amplified via PCR ~ 65 bp from the “teleo” 12S rRNA barcode (Valentini et al., 2016). Library preparation involved two PCR steps and followed minor modifications from similar methods (Miya et al., 2015). Primer sequences, PCR methods, and reaction conditions are shown in Annex B **Table 21** and **S2**. In addition to the environmental samples, a positive control (mock community described above) and negative control (1 pooled negative control consisting of 4 negative controls processed along with the other eDNA samples during the two steps of PCR, including replicates) were incorporated during library preparation. A total of 26 samples (including positive and negative controls) were pooled into one 4 nM equimolar sample and sent to Genomic Services at Langebio-CINVESTAV. A single flow cell Illumina NextSeq 500 MID (35 Gb) v2 chemistry (2 × 150 bp paired-end) was used for sequencing.

#### 2.1.2.6 Metabarcoding bioinformatic analyses

Bioinformatic analyses were implemented with a UNIX shell script. The complete pipeline is described in Annex B **Table 22**. In the first step of the analysis, bcl2fastq v2.19 (Illumina) was used to demultiplex indexed sequences. Then, Obitools v1.2.11 (Boyer, Mercier, Bonin, Taberlet, & Coissac, 2014) was employed for applying robust primary filtering and selecting for high-quality, full-length sequences. Next, Vsearch v2.7.0 was used (Rognes et al., 2016) for chimera detection *de novo*, using the uchime\_ *de novo* algorithm (Edgar et al., 2018). A step-by-step aggregation analysis was implemented in Swarm v2.2.2 (Mahé et al., 2015) to cluster reads with a resolution  $d = 2$  (~3% genetic distance). This value was selected according to the results of genetic distances among species of the local ichthyofauna in our custom reference database, in which I found that all variation observed below the species level (i.e., intra-species)

was  $d < 2$ , while most variation observed above species level (e.g., 82 % at intra-genus and 86% at intra-family) was  $d > 2$  (Annex B **Table 23** and **Figure 40**). Finally, singletons were removed, and OTUs/site table conversion was performed using the `owi_recount_swarm` R script (Wangensteen, 2019).

### 2.1.2.7 OTUs taxonomic assignment

Taxonomic assignment of OTUs was performed using three approaches:

1. Using NCBI-GenBank
2. Using our custom reference database
3. Combining both (NCBI + custom reference database)

Basic Local Alignment Search Tool (BLAST) implemented in CLC Genomics Workbench v9 (Qiagen) was used with a word count value of 30, expected score of 10, and e-value of  $1 \times 10^{-20}$ . In any case, any OTU without hits in the Blast search was excluded from the taxonomic assignment. Conversely, in the resulting alignments, I established conservative thresholds of the percent identity of the hits to assign taxonomy as follows: 100-97% of identity were assigned to species; 97-94% to genus; 94-91% to family; 91-88% to order and  $< 88\%$  to class. When OTUs could not be assigned to species or genus level, I named the OTU with a consecutive number, followed by the taxonomic level achieved, e.g., OTU\_01 (Acanthuridae). I collapsed (grouped) OTUs taxonomically assigned to the same species or genus. Finally, a secondary filtering step was performed to avoid potential tag switching or false positives, which excluded singletons within each sample. Lists of species/OTUs, genus, families, orders, and class detected with UVC and eDNA metabarcoding were generated, and species/OTUs per site matrices were obtained to perform community-level ecological analyses.

### 2.1.2.8 Community-level ecological analyses

All the analyses were performed in R v4.0.0 and RStudio v1.2.1335 (RStudio Team, 2020) using *vegan* v2.5-6 (Oksanen et al., 2020). Individual rarefaction curves (*rarecurve*) and species accumulation curves per site (*speccaccum*) were computed for each survey method. Alpha diversity ( $S_{obs}$ ) and extrapolated species richness ( $S_{exp}$ , Chao) were also estimated (*specpool*). To check for significant differences in  $S_{obs}$  values between methods, I first tested for normality with a Shapiro-Wilk test and then used a Wilcoxon test to evaluate if the median of  $S_{obs}$  differed between methods. I also computed the

Spearman correlation between the  $S_{obs}$  from each method to assess the relationship of this parameter between survey methods.

I used the beta diversity (species turnover) using the Jaccard dissimilarity index based on species' presence/absence to assess differences in the overall community structure detected with both survey methods (UVC and eDNA metabarcoding). To achieve this, I estimated the multivariate homogeneity within-group (North and Central) covariance matrices with (*betadisper*), and then I evaluated differences in the data dispersion between groups with ANOVA (*anova*). Differences in species composition of communities between regions (North and Central) were tested statistically by permutational analysis of similarities (ANOSIM, *anosim*) and permutational multivariate analysis of variance (PERMANOVA, *adonis*). Dissimilarity values were ordinated using a non-metric Multidimensional Scaling (nMDS, *metaMDS*) to visualize the group discrimination of each survey method. A Shepard's curve was calculated (*stressplot*) to establish the k number of the ordination analysis.

Finally, I used the complementary detection data of reef-fish species obtained with UVC and eDNA metabarcoding to estimate the beta diversity (species turnover) using the Jaccard dissimilarity index based on species' presence/absence estimate the beta diversity (Jaccard) to evaluate its relationship with the geographical (Euclidean), environmental and resistance distances.

Derived from the previous results, a scientific article was published in the Molecular Ecology Resources Journal in March 2021 (**Annex A**).

## **2.2 Structural connectivity: estimation of abiotic predictors and testing for mechanisms of isolation**

### **2.2.1 Spatial and environmental predictors of functional connectivity**

To estimate the Euclidean distance among population and communities sampling localities (*Euclidean distance* from now on), I transformed the spatial coordinates of each sampling site into cartesian coordinates and create a distance matrix using *SoDA* v1.0-6.1, and *stats* v4.0.3 base packages in R v4.0.3 (R Core Team, 2021).

Then, to evaluate the effects of the spatial configuration and the mean annual environmental characteristics in which the leopard grouper and the rocky reef fishes inhabit into its connectivity, I

obtained the values of 14 environmental layers (\*.tif) (**Table 6**) from Copernicus Marine Environment Monitoring System (CMEMS) (<https://marine.copernicus.eu>). I clipped it onto the GC range (latitude max: 32.0953, latitude min: 20.3899, longitude max: -104.7695, and longitude min: -115.2966), using the *raster* v3.4-5 package in R v4.0.3. Then I extracted the values from the environmental layers for population and community sampling localities with the same R package. For each set of environmental information (population and community), I performed a principal components analysis (PCA) using *stats* v4.0.3 base package in R v4.0.3 (R Core Team, 2021) with the centered and scaled variables to summarize the patterns of environmental variation among sampling sites. Then I estimated the Euclidean environmental distance (*environmental distance* from now on) using the *stats* v4.0.3 base package in R v4.0.3 (R Core Team, 2021) of the three principal components of the PCA analysis (PC1, PC1:PC2, and PC1:PC3) which represents ~50, 60 and 80% of the environmental variation among sites, respectively. The removal of redundant variables with the creation of orthogonal synthetic predictors is a strategy to correct for multicollinearity.

## 2.2.2 Oceanographic predictors of functional connectivity and network analyses

### 2.2.2.1 The HAMSOM numerical model, obtention of the potential larval connectivity matrices and the estimation of the resistance distance

First, to establish the domain in which the population (leopard grouper) and community (rocky reef fishes) potentially distribute, 59 polygons were created using a geographic information system (ArcMap) combining shapefiles of the country, the marine protected areas, and the Gulf bathymetry (low limit -200 m depth isobath (**Figure 16**)). At each polygon, 4000 particles (or virtual larvae) were seeded for each release date that corresponds to each month of the year (January to December), at two starting moments (i.e., spring and neap tide). These particles were advected from an Eulerian velocity field obtained from the baroclinic three-dimensional numerical model HAMSOM developed by Backhaus (1985) adapted to the GC (Marinone, 2003, 2008) and tracked for 28 days or four weeks (mean PLD). The model used a mesh size of  $\sim 1.3 \times \sim 1.5$  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 meters, which depend on the site depth. Node by edges potential larval connectivity matrices (i.e., probability matrix) were obtained for each of the twelve months. Each matrix contained information about the proportion of larvae that settled at each polygon relative to the total number of larvae released.

**Table 6.** Environmental layers included as predictors of functional connectivity.

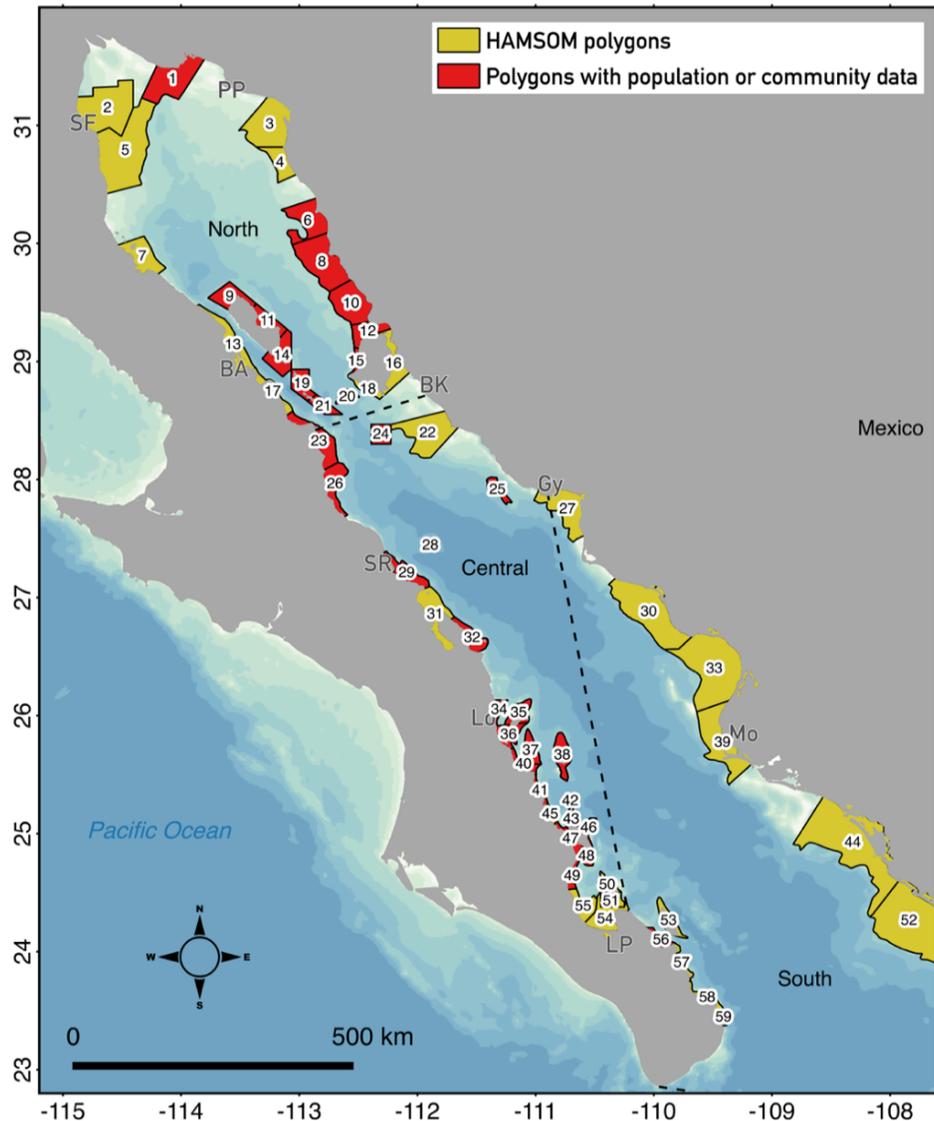
Name	Units	Spatial resolution	Source	Start year	End year
Bathymetry	meters	30 arcsecond	GEBCO	2016	2016
Sea surface temperature (SST) (annual mean)	Celsius	5 arcmin (9.2 km)	<a href="http://oceancolor.gsfc.nasa.gov/">http://oceancolor.gsfc.nasa.gov/</a>	2002	2009
Current velocity (annual mean)	m/s	0.25 arcdegree	Global Ocean Physics Reanalysis ECMWF ORAP5.0 (1979-2013)	2000	2014
Dissolved oxygen concentration (annual mean)	μmol kg <sup>-1</sup>	0.25 arcdegree	Global Ocean Biogeochemistry NON ASSIMILATIVE Hindcast (PISCES)	2000	2014
Iron concentration (annual mean)	μmol kg <sup>-1</sup>	0.25 arcdegree	Global Ocean Biogeochemistry	2000	2014
Nitrate concentration (annual mean)	μmol kg <sup>-1</sup>	0.25 arcdegree	Global Ocean Biogeochemistry NON ASSIMILATIVE Hindcast (PISCES)	2000	2014
Phosphate concentration (annual mean)	μmol kg <sup>-1</sup>	0.25 arcdegree	Global Ocean Biogeochemistry NON ASSIMILATIVE Hindcast (PISCES) URL: <a href="http://marine.copernicus.eu/">http://marine.copernicus.eu/</a>	2000	2014
Primary production (annual mean)	g/m <sup>2</sup> /day	0.25 arcdegree	Global Ocean Biogeochemistry NON ASSIMILATIVE Hindcast (PISCES) URL: <a href="http://marine.copernicus.eu/">http://marine.copernicus.eu/</a>	2000	2014
Sea surface salinity (annual mean)	PSS	0.25 arcdegree	Global Ocean Physics Reanalysis ECMWF ORAP5.0 (1979-2013) URL: <a href="http://marine.copernicus.eu/">http://marine.copernicus.eu/</a>	2000	2014
Silicate concentration (annual mean)	μmol kg <sup>-1</sup>	0.25 arcdegree	Global Ocean Biogeochemistry NON ASSIMILATIVE Hindcast (PISCES) URL: <a href="http://marine.copernicus.eu/">http://marine.copernicus.eu/</a>	2000	2014
East/West aspect	radians	30 arcsecond	SRTM30_PLUS V6.0 reference: Becker et al. 2009 URL: <a href="http://topex.ucsd.edu/WWW_html/srtm30_plus.html">http://topex.ucsd.edu/WWW_html/srtm30_plus.html</a>	2009	2009
North/South Aspect	radians	30 arcsecond	SRTM30_PLUS V6.0 reference: Becker et al. 2009 URL: <a href="http://topex.ucsd.edu/WWW_html/srtm30_plus.html">http://topex.ucsd.edu/WWW_html/srtm30_plus.html</a>	2009	2009
Distance to shore	kilometers	30 arcsecond	GSHHS v2.1 reference: Wessel and Smith 1996 URL: <a href="http://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html">http://www.ngdc.noaa.gov/mgg/shorelines/gshhs.html</a>	2009	2009
Concavity	degrees	30 arcsecond	SRTM30_PLUS V6.0 reference: Becker et al. 2009 URL: <a href="http://topex.ucsd.edu/WWW_html/srtm30_plus.html">http://topex.ucsd.edu/WWW_html/srtm30_plus.html</a>	2009	2009

Second, I averaged the corresponding probability matrices of the months within each hypothetical reproductive season (**Table 7** and **Table 8**). From each 59 x 59 matrix, I extracted the information for the 22 polygons sampled for population analyses and 23 polygons sampled for community analyses. As a result, I obtained 21 probability matrices (22 x 22 polygons) for the population hypothetical reproductive seasons and four probability matrices (23x23 polygons) for the community hypothetical reproductive seasons.

Population hypothetical reproductive seasons vary in amplitude (two to seven months), and the moment of the year that they occur. According to the variation in the literature, this happens mainly throughout spring and summer (**Table 3**). On the other hand, rocky reef fish's hypothetical reproductive seasons included Winter (December to March), Spring (March to June), Summer (June to September), and Fall (September to December). These hypothetical reproductive seasons encompass the full range of reproductive scenarios in which reef fishes can reproduce (**Table 1**), according to the variation in the literature.

Third, to determine the resistance distance (*resistance distance* for now on) derived from the probability matrices per hypothetical reproductive seasons, I estimated the shortest path (least cost path) using Dijkstra's algorithm weighted with the probability larval dispersal value with the *igraph* v1.2.6 package (Csardi and Nepusz, 2006) in R v4.0.3 (R Core Team, 2021). This allowed to account for the weight of the links (i.e., the proportion of larvae dispersed) as a proxy for seascape resistance to the movement of the larvae, i.e., links with high values of larval connectivity denote shortest resistance distances. Meanwhile, links with low values of larval connectivity represent the largest resistance distances. I obtained a resistance distance matrix for each hypothetical reproductive seasons evaluated.

I tested the linear relationship between the leopard grouper's genomic differentiation  $F_{ST}$  matrix and the community beta-diversity matrix (species turnover) with seascape predictors (i.e., Euclidean, environmental, and resistance distances). To achieve this, I applied a Mantel test in *vegan* v2.5-6 (Oksanen et al., 2020) package in R v4.0.3 (R Core Team, 2021) when including symmetrical matrices relationships with a Pearson correlation method, and a Mantel test implemented in *ape* v5.4-1 (Paradis et al., 2021) package in R v4.0.3 (R Core Team, 2021) when including the asymmetrical oceanographic matrices, using 9999. This test is a permutation-based statistical examination describing the correlation between two distance or dissimilarity matrices and allows matrix asymmetry. In addition, I performed multiple regression on distance matrices (MRM) using a permutation test (1000 permutations) of significance for regression coefficients and R-squared using the *ecodist* package in R v4.0.3 (R Core Team, 2021).



**Figure 16.** Fifty-nine polygons used in the HAMSOM numerical model are represented in green, and polygons with sampling sites at the population or community level are red. Principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN).

## 2.2.3 Testing for mechanisms of isolation: integrating population genomic and community connectivity patterns and the seascape predictors

### 2.2.3.1 Mantel test and multiple regression on distance matrices (Link-level analyses)

### 2.2.3.2 Network analysis of connectivity data (node-level analysis)

I implemented a network approach to analyze patterns of demographic connectivity at node and link levels for March-June (Spring) and June-September (Summer) population hypothetical reproductive seasons;

and Winter, Spring, Summer, and Fall community hypothetical reproductive seasons. First, from each probability matrix, I estimated different *node metrics*: *export probability* as the sum of larval export from sources *i* into location *j*; *import probability* as the sum of larval imports from sources *j* into location *i*; and *local retention* as the proportion of larvae released within a node that remains within the natal area at the end of the PLD (Burgess et al., 2014; Munguia-Vega et al., 2018). From these metrics, I derived the *export - local retention (net export)*, *import - local retention (net import)*, and *export-import (source-sink characteristic of the polygon)* (Munguia-Vega et al., 2014; Munguia-Vega et al., 2018) in which sources are populations in with net export of larvae (overall export is greater than imports), whereas the reverse are sinks (Cowen and Sponaugle, 2009).

**Table 7.** Hypothetical population-level reproductive seasons from which potential larval connectivity matrices were obtained. The numbers correspond to the number of each month (1: January ... 12: December).

Season	1	2	3	4	5	6	7	8	9	10	11	12
March-April			■	■								
April-May				■	■							
May-June					■	■						
June-July						■	■					
July-August							■	■				
August-September								■	■			
March-May			■	■	■							
April-June			■	■	■							
May-July				■	■	■						
June-August					■	■	■					
July-September						■	■	■				
March-June			■	■	■	■						
April-July			■	■	■	■						
May-August				■	■	■	■					
June-September					■	■	■	■				
March-July			■	■	■	■	■					
April-August			■	■	■	■	■					
May-September				■	■	■	■	■				
March-August			■	■	■	■	■	■				
April-September			■	■	■	■	■	■				
March-September			■	■	■	■	■	■	■			

**Table 8.** Hypothetical community-level reproductive seasons from which potential larval connectivity matrices were obtained. The numbers correspond to the number of each month (1: January ... 12: December).

Season	1	2	3	4	5	6	7	8	9	10	11	12
Winter												
Spring												
Summer												
Fall												

Then, I converted each probability matrix to an adjacent matrix and then to a graph object from which I obtained directed-weighted graphs (i.e., a network which links has a direction and weight) and the following node centrality metrics: *degree* (i.e., number of connections or links; “in” and “out”, and respectively); *eigenvector centrality* (i.e., an extension of the degree, but takes into account the importance of the nodes that the node has a connection with); *central degree* (i.e., node centrality in a graph); *betweenness* (i.e., centrality based in the number of links that pass through the node); *hub* (i.e., a node that contain a large number of outgoing links), and *authorities* (i.e., a node with a high number of incoming links from hubs) (Newman, 2018). The graph size, which indicates the number of links in a graph, was also estimated. Finally, I represented larval potential connectivity networks in a grid format to evaluate connectivity patterns in a visual form. All the steps for the network analyses were conducted with the *igraph* v1.2. 6 package (Csardi and Nepusz, 2006) in R v4.0.3 (R Core Team, 2021).

To evaluate significant associations among node metrics, I estimated a linear regression model among the leopard grouper’s (population) genomic diversity data and community alpha diversity (species richness) using *stats* v4.0.3 base package in R v4.0.3 (R Core Team, 2021).

## Chapter. 3      Results

---

### 3.1 Evaluation of the functional connectivity at population and community organization levels

#### 3.1.1 Population level functional connectivity: evaluation of the genomic diversity and differentiation ( $F_{ST}$ ) in the leopard grouper (*Mycteroperca rosacea*) through the analysis of Single Nucleotide Polymorphisms (SNPs)

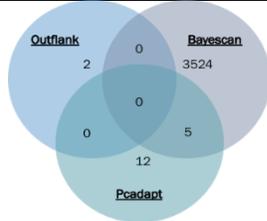
##### 3.1.1.1 Bioinformatic analyses: SNP calling

Sequencing results yielded 848.8 million raw sequence reads from the ddRAD sequencing, which was demultiplexed with *process\_radtags* retaining 802.3 million reads total, a mean of 34.8 million reads per sampling locality, and a mean of 3.7 million reads per individual. After *populations* filtering, 35,984 SNPs genotyped in 212 individuals remained in the dataset. Details on SNPs number retained after further filtering steps are shown in **Table 9**. The overall level of missing data was 6.86%.

##### 3.1.1.2 Bioinformatic analyses: Data filtering and identification of neutral and candidate SNPs loci

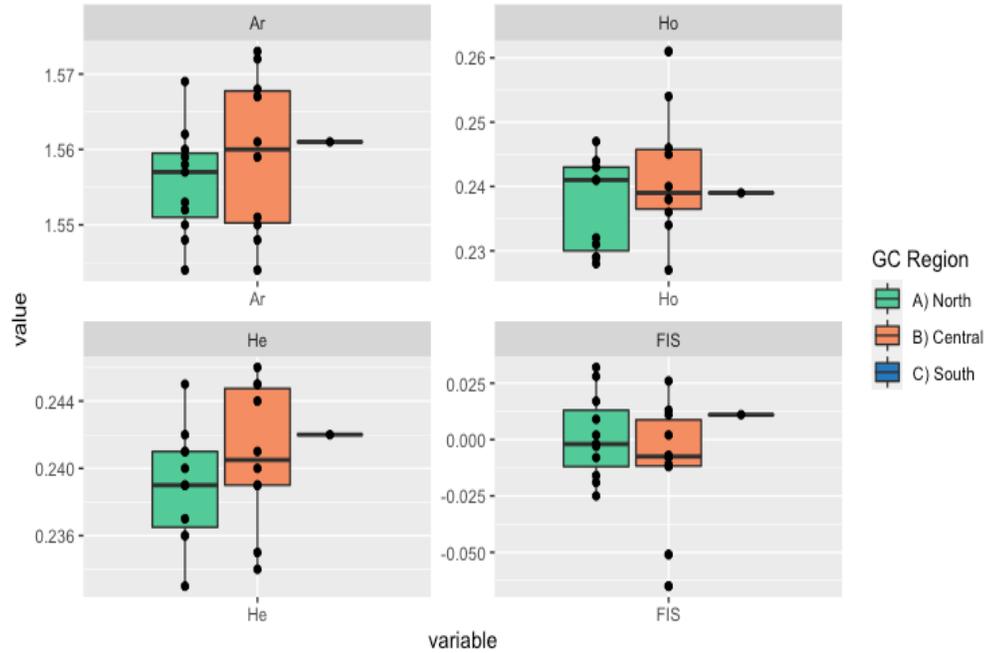
From the 6,589 SNPs, Bayescan detected 3,529 potentially under balancing selection, while Outflank detected 2 SNPs and *pcadapt* 17 SNPs. Five loci were shared between Bayescan and *pcadapt*. After removing the candidate loci, and the corresponding filtering steps for the Neutral loci dataset (MAF, HW, and allele number), 985 SNPs were retained (**Table 9**).

**Table 9.** Bioinformatic analyses results including sequencing and the filtering steps. Results are shown in terms of obtained number of sequencing reads or obtained number of SNPs, and in the number of individuals remaining after each step.

Sequencing step	Obtained Reads	Number of individuals
Total number of raw sequence reads	802,395,198	212
Filtering step 1	Number of SNPs	Number of individuals
denovo_map	147,841	212
Populations 50% Individuals, single-SNP	35,984	212
Remove loci > 20% missing data	11,852	212
Maximum mean depth/loci 100X	6,589	212
Filtering step 2 in the Neutral dataset	Number of SNPs	Number of individuals
MAF 0.04	1,204	212
HW ( $p < 0.05$ )	985	212
Number of alleles 2	985	212
Remove individuals > 20% missing data	985	201
Outlier dataset		212
Neutral dataset	985	203

### 3.1.1.3 Geographical distribution of the genomic diversity and evaluation of the population genomic differentiation

I found moderate levels of genomic diversity with low variation among localities. Allelic richness rarefied to  $N = 6$  ( $A_r$ ), varied from 1.544 (LIB and SRO) and 1.57 (SMA, EVA) (mean = 1.56). Observed heterozygosity ( $H_o$ ) fluctuated from 0.227 (BS) to 0.261 (EVA) (mean = 0.239), and expected heterozygosity ( $H_e$ ) from 0.233 (LIB) to 0.246 (EVA) (mean = 0.240).  $F_{IS}$  values varied from  $-0.065$  (SRO) to  $0.032$  (TIB\_N) (mean =  $-0.003$ ) (Table 10). Genomic diversity metrics were not statistically different among North, Central and South regions of the GC ( $A_r$ :  $F = 0.522$ ,  $p = 0.602$ ;  $H_o$ :  $F = 0.815$ ,  $p = 0.458$ ;  $H_e$ :  $F = 0.747$ ,  $p = 0.487$ ;  $F_{IS}$ :  $F = 0.823$ ,  $p = 0.454$ ), nevertheless the Central region presented the highest values of  $A_r$  (EVA),  $H_o$  (EVA) and  $H_e$  (EVA) (Figure 17 and Figure 18).

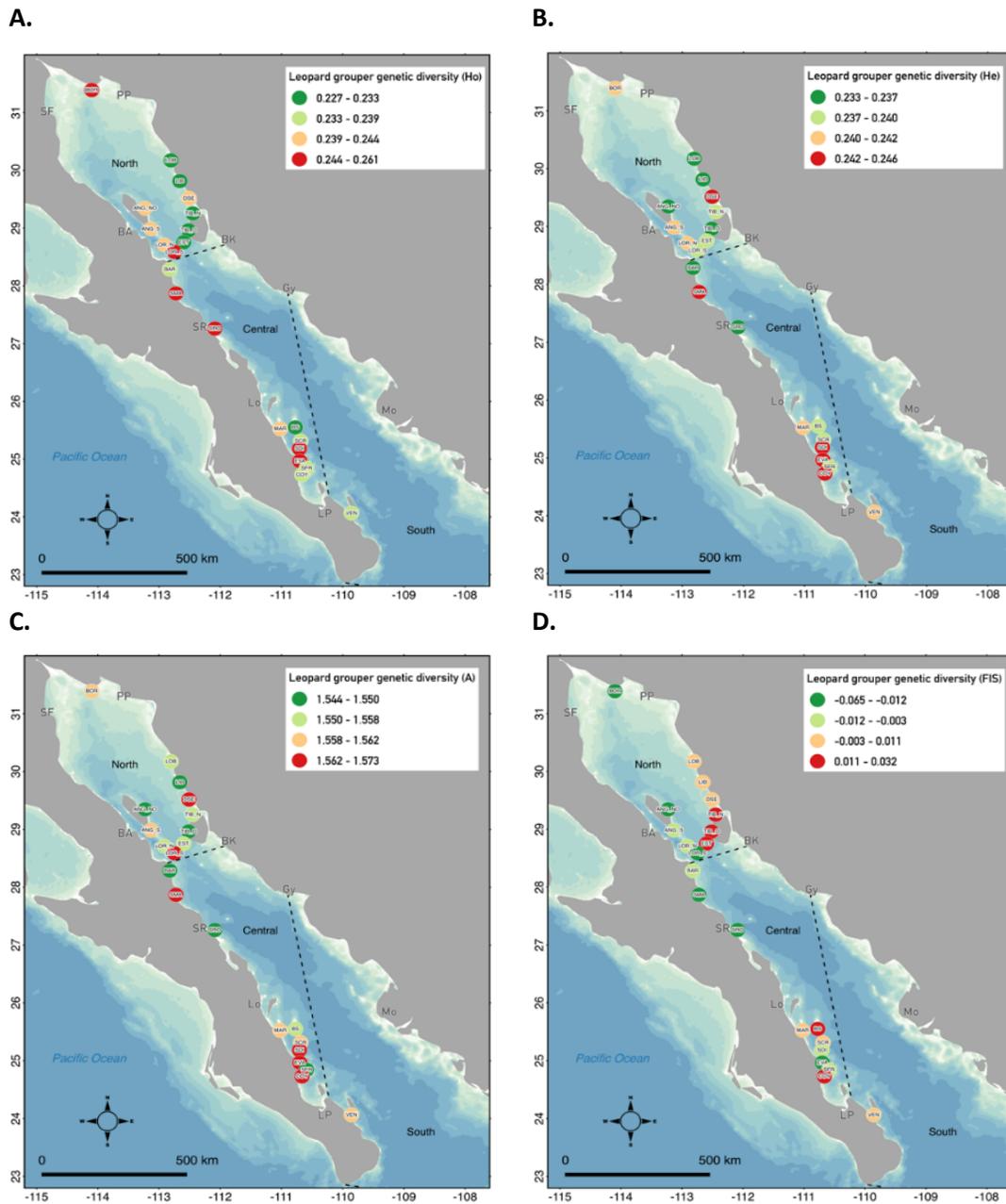


**Figure 17.** Boxplots showing genomic diversity metrics per GC region. Ar: rarefied allelic richness, He: expected heterozygosity, Ho: observed heterozygosity, and FIS: Wright's inbreeding coefficient.

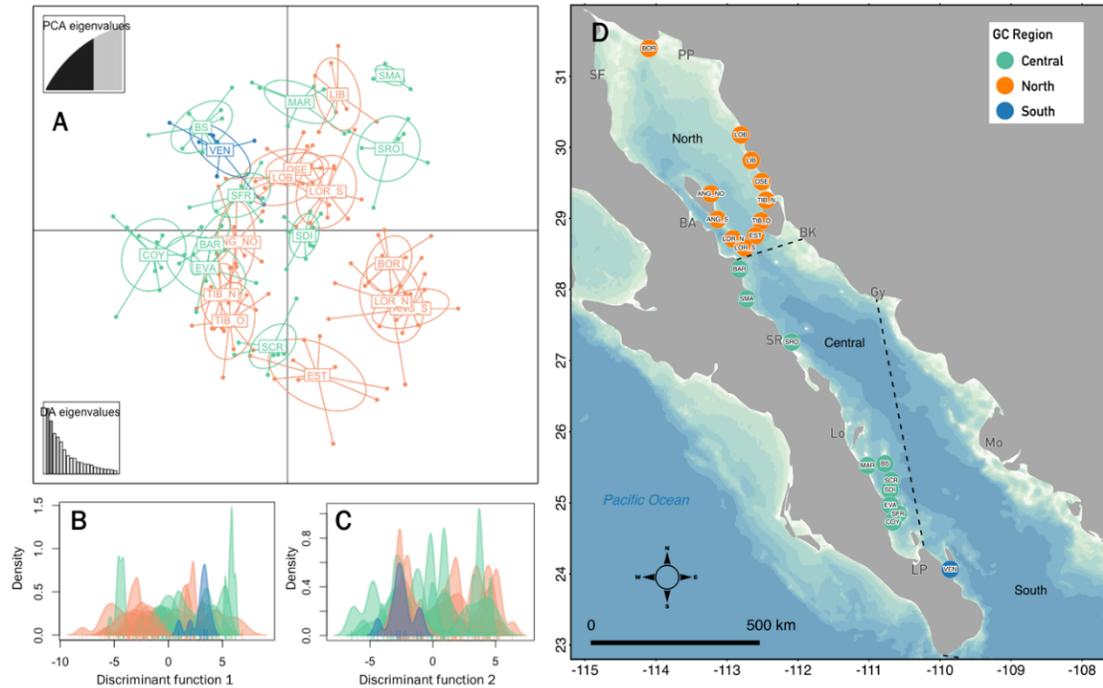
The analysis of the population differentiation in the study area using pairwise  $F_{ST}$  (Weir and Cockerham, 1984) evidenced low and non-significant values of genomic differentiation among all the sampling localities.  $F_{ST}$  values ranged from zero (~50 % of the paired comparisons), to 0.011 (ANG\_S-BS) (mean pairwise value = 0.001). The hierarchical AMOVA evidenced that the greatest percentage of genomic variation occurs within individuals (99%) and a non-significant population genomic structure ( $F_{ST}$  = 0.0012,  $p$  = 0.27). Admixture analysis evidenced the presence of one population ( $K = 1$ ) (Annex B **Figure 41**), and the multivariate analysis DAPC with no prior information concurs with this result when using the BIC criterion for assessing the optimal number of clusters. When DAPC included a priori sampling sites, 120 PCs, and 5 discriminant functions explaining 82% of the genomic variance, the scatter plot of the two first discriminant axes revealed no clear geographic structure on the genomic diversity (**Figure 19**).

**Table 10.** Table summarizing the results of genetic diversity analysis of local leopard grouper populations using 985 SNPs. The local population localities are grouped in their corresponding GC region. The table shows the Locality number (Loc), Locality name, the Locality ID (ID), Number of genotyped samples (N), the geographic location (Longitude, Latitude), the rarefied allelic richness ( $A_r$ ), the observed ( $H_o$ ) and expected heterozygosities ( $H_e$ ), and the Wright's inbreeding coefficient ( $F_{IS}$ ).

Loc	Locality name	ID	N	Lon	Lat	$A_r$	$H_o$	$H_e$	$F_{IS}$
<b>North GC region</b>									
1	Puerto Peñasco (Borrascoso)	BOR	11	-114.1	31.39	1.559	0.247	0.241	-0.019
2	Puerto Lobos	LOB	8	-112.806	30.173	1.552	0.231	0.237	0.009
3	Puerto Libertad	LIB	7	-112.66	29.814	1.544	0.232	0.233	-0.002
4	Desemboque Seri (0-100m)	DSE	10	-112.512	29.516	1.569	0.243	0.245	0.002
5	Isla Tiburón (North)	TIB_N	8	-112.449	29.254	1.553	0.228	0.239	0.032
6	Isla Tiburón (West)	TIB_O	11	-112.518	28.963	1.548	0.229	0.236	0.017
7	Reserva Isla Ángel de la Guarda (South)	ANG_S	10	-113.135	28.985	1.560	0.241	0.242	-0.003
8	Reserva Isla Ángel de la Guarda (Northwest)	ANG_NO	10	-113.228	29.345	1.550	0.243	0.236	-0.025
9	Reserva Isla San Lorenzo (South)	LOR_S	10	-112.748	28.586	1.562	0.244	0.240	-0.016
10	Reserva San Lorenzo (North)	LOR_N	10	-112.914	28.71	1.558	0.241	0.241	-0.008
11	Isla San Esteban	EST	10	-112.599	28.753	1.557	0.229	0.239	0.028
<b>Central GC region</b>									
12	El Barril	BAR	6	-112.826	28.286	1.548	0.234	0.234	-0.008
13	Santa María	SMA	5	-112.724	27.868	1.572	0.245	0.245	-0.012
14	Santa Rosalía	SRO	7	-112.089	27.261	1.544	0.254	0.235	-0.065
15	San Marcial	MAR	8	-110.774	25.549	1.561	0.240	0.241	0.002
16	Bajo Seco	BS	10	-110.707	25.183	1.551	0.227	0.239	0.026
17	Isla Santa Cruz	SCR	8	-111.021	25.524	1.559	0.236	0.240	0.011
18	Isla San Diego	SDI	10	-110.688	25.313	1.567	0.246	0.244	-0.011
19	San Evaristo	EVA	8	-110.701	24.966	1.573	0.261	0.246	-0.051
20	Isla San Francisco	SFR	7	-110.592	24.844	1.550	0.238	0.239	-0.007
21	Coyote	COY	10	-110.666	24.73	1.568	0.238	0.245	0.013
<b>South GC region</b>									
22	Bahía Ventana	VEN	10	-109.858	24.064	1.561	0.239	0.242	0.011



**Figure 18.** Maps showing the geographical distribution of the genomic diversity of the leopard grouper A:  $H_o$  (observed heterozygosity), B:  $H_e$  (expected heterozygosity), C: A (allelic richness), D:  $F_{IS}$  (inbreeding coefficient). Principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN).



**Figure 19.** Discriminant Analysis of Principal Components (DAPC) results of the 985 SNP neutral dataset obtained from the leopard grouper population analysis. Figures show the scatterplot with a priori sampling location information (Panel A) and density plots for (Panel B) the first and (Panel C) the second discriminant axes. Map (Panel D) shows the geographical distribution of the sampling localities and the GC regions. The principal cities shown include SF: San Felipe (BC), BA: Bahía de los Ángeles (BC), SR: Santa Rosalía (BCS), Lo: Loreto (BCS), LP: La Paz (BCS), BK: Bahía de Kino (SON), Gy: Guaymas (SON) and Mo: Los Mochis (SIN).

### 3.1.2 Community-level functional connectivity: estimation of the alpha (species richness) and beta diversities (species turnover) in communities of rocky reef fishes using complementary monitoring methods

#### 3.1.2.1 Underwater visual censuses species detection

UVC covered 295 transects (mean/site = 12, SD = 4.5) in 24 sites from the GC. In total, 43,647 individual teleost fishes were observed, representing 97 observed species, 66 genera, 32 families, and eight orders (Annex B **Table 24**).

### 3.1.2.2 Custom reference for the ichthyofauna of the GC: primer design, PCR amplification, and sequencing

The custom reference database included 112 12S rRNA gene sequences from 67 species and 32 families with a mean length of 552 bp (SD = 65 bp) (NCBI-GenBank accession numbers in Annex B **Table 19**). When analyzing the genetic variation among these species, I found high genetic variability taxonomic levels. Intra-family variation ranged from low genetic distance ( $d_{max}$ ) in Istiophoridae (0 – 2%) to high in Pomacentridae (42.9%) and Serranidae (42.7%). At the intra-genus level,  $d_{max}$  went from low values in Thunnus (0 – 1.4%), to high in Lutjanidae (14.5%). Intra-specific variation was found in the species *Kyphosus elegans* ( $d_{max}$  2.8%), *Lutjanus peru* ( $d_{max}$  1.4%), *Mycteroperca rosacea* ( $d_{max}$  1.4%), *Thunnus albacares* ( $d_{max}$  1.4%) and *Scomberomorus sierra* ( $d_{max}$  1.3%) (Annex B **Table 23**).

### 3.1.2.3 eDNA metabarcoding sequencing statistics, OTUs identification in environmental samples, and taxonomic assignment

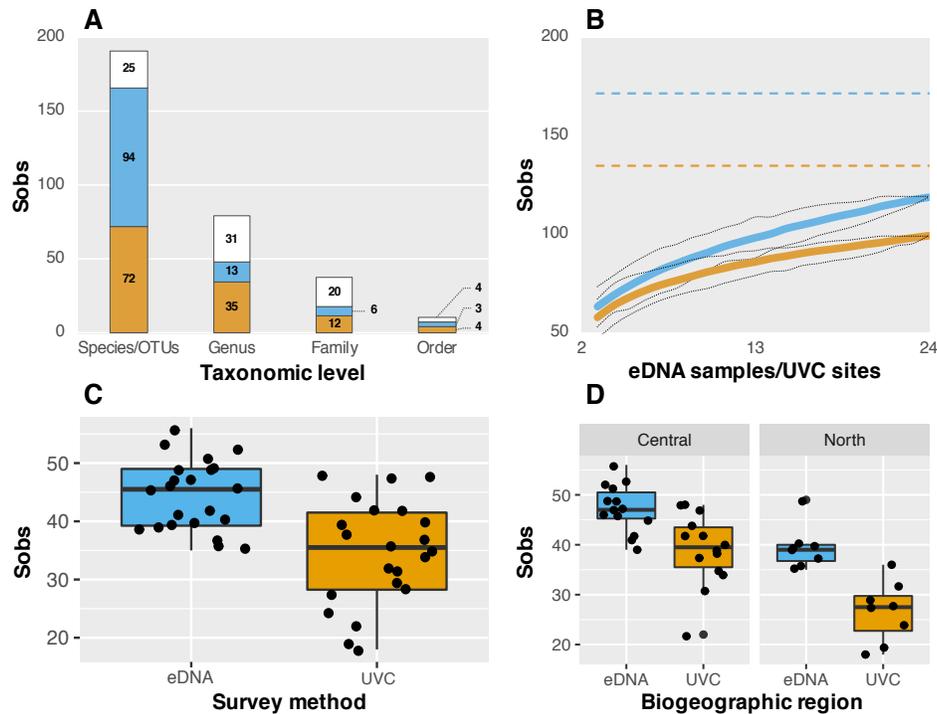
I obtained 5,429,682 raw paired-end reads in the 26 sequenced samples (mean = 208,833 reads/sample, SD = 43,538) from which I kept 2,897,810 (mean = 111,454 reads/sample, SD = 45,472) after quality filtering (Annex B **Table 25**). OTU clustering returned 542 OTUs that were used in the different taxonomic assignment approaches. The success and resolution of each of the taxonomic assignment methods varied depending on the reference database used (NCBI-GenBank, custom reference database, NCBI + custom reference database). Separately, the NCBI database and the custom reference database produced similar results in terms of the number of assigned OTUs (48% and 49% of 542 OTUs, respectively). On the other hand, the combined use of both databases outperformed their individual contributions in terms of the taxonomic resolution at species and genus levels and increased the identification success by ~12% (60% of 542 OTUs) (Annex B **Table 26**). Thus, I used the NCBI + custom reference database for the final taxonomic assignment of the OTUs, obtaining 119 unique Actinopterygii OTUs, of which 57 were taxonomically assigned above family level. The remaining 64 OTUs represent 64 species, 44 genera, 26 families, and seven orders (Annex B **Table 27**).

**Table 11.** Table summarizing the results of alpha diversity analysis of local rocky reef fish communities using UVC and eDNA metabarcoding. The table shows the Locality number (Loc), Locality name, Locality ID (ID), Collection date, the geographic location (Longitude, Latitude), and the alpha diversity estimated using eDNA (including or not the OTUs), UVC, and complementing both detection methods UVC + eDNA (including or not the OTUs).

Loc	Locality name	ID	Long	Lat	Alpha diversity eDNA with OTUs	Alpha diversity eDNA without OTUs	Alpha diversity UVC	Alpha diversity UVC + eDNA with OTUs	Alpha diversity UVC + eDNA without OTUs
1	El Portugués	POR	110.674	24.749	45	27	32	55	37
2	Bajo Seco Sur	BSS	110.525	24.809	52	25	22	54	27
3	Bajo Las Ánimas	ANI	110.508	25.114	49	27	34	65	43
4	Isla San Diego	SDI	110.695	25.204	53	26	48	75	48
5	Isla Santa Cruz	SCR	110.689	25.313	47	24	39	63	40
6	Isla San Mateo	MAT	110.993	25.379	56	30	38	66	40
7	Isla Catalana	CAT	110.776	25.715	51	30	35	60	39
8	Isla Monserrat	MON	111.049	25.744	41	30	48	58	47
9	Isla Danzante	DAN	111.255	25.813	46	28	40	60	42
10	Isla Carmen	CAR	111.169	26.017	46	30	47	62	46
11	Isla Coronado	COR	111.262	26.174	47	24	43	68	45
12	Punta Púlpito	PUL	111.443	26.515	49	27	44	-	-
13	Isla San Ildefonso	ILD	111.427	26.625	42	27	42	61	46
14	Isla San Marcos	SMAR	112.089	27.260	39	24	39	58	43
15	Isla Tortuga	TOR	111.062	27.431	44	26	30	51	33
16	Isla San Pedro Nolasco	NOL	111.371	27.966	52	32	50	70	50
17	Isla San Pedro Mártir	PMA	112.296	28.382	39	23	37	56	40
18	Isla San Francisquito	FRA	112.268	28.441	40	24	29	51	35
19	Isla San Lorenzo	LOR	112.759	28.585	49	29	27	54	34
20	Isla Ángel de la Guarda	IA-I	113.559	29.555	35	22	29	45	32
21	Puerto Lobos	LOB	112.966	30.222	39	27	18	39	27
22	Isla Pato	PAT	112.464	29.266	36	22	24	43	29
23	Isla Tiburón	TIB	112.506	29.065	40	25	32	55	40
24	Isla San Esteban	EST	112.613	28.722	37	20	19	44	27

### 3.1.2.4 List of species/OTUs detected with both survey methods

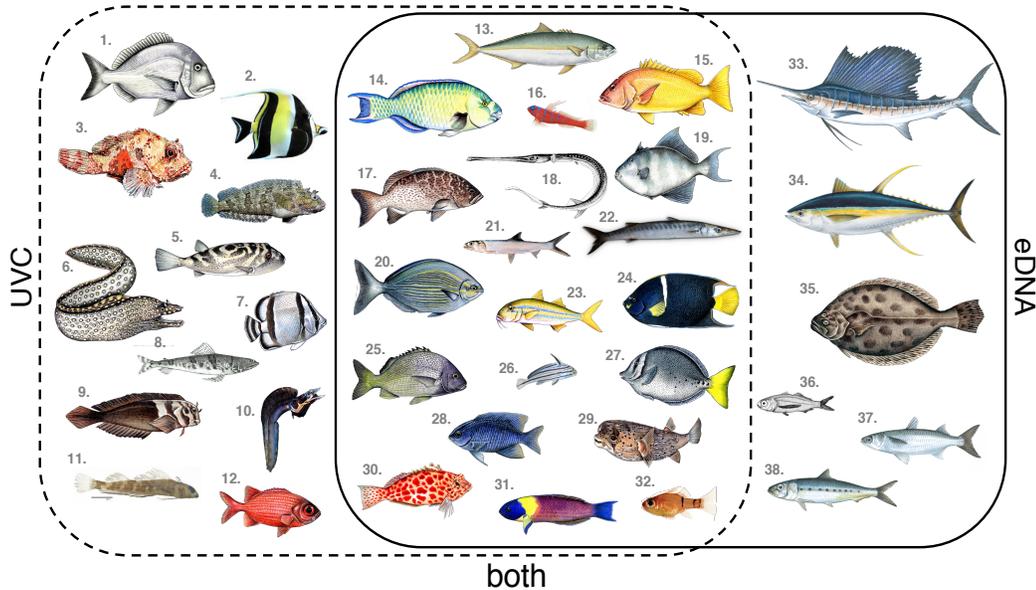
A total of 191 species/OTUs were detected with both survey methods. Of these, 13% were shared, 49% were identified only by eDNA, and 38% were observed only with UVC. The proportion of shared taxa between methods increased when higher taxonomic levels were considered: 39% for genus (31/79), 53% for family (20/38), and 36% for order (4/11) (Figure 20) (Annex B Table 29 - 32 for complete lists).



**Figure 20.** (A) Number of observations per taxonomic level. (B) Species accumulation curves with confidence intervals (95%),  $S_{exp}$  is showed in dashed lines for each monitoring method. (C) Boxplot of the mean observed number of species ( $S_{obs}$ ) per site in each survey method. (D) Mean  $S_{obs}$  per method and per region of the Gulf of California. In all graphs, UVC is shown in yellow; eDNA is shown in blue, and both methods are shown in white.

Even though UVC targeted only the class Actinopterygii, eDNA also recorded the class Chondrichthyes (*Carcharhinus leucas* three reads), Mammalia (*Homo sapiens* four reads), and Aves (*Gallus sonneratii* nine reads), which were removed for the ecological analyses. Furthermore, while all fishes registered in the UVC were reef-associated, eDNA metabarcoding detected additional taxa from surrounding pelagic, demersal, and estuarine habitats. From the 38 families jointly detected with both methods, six families were identified only with eDNA and included pelagic (*Istiophoridae*, *Scombridae*, *Clupeidae*), demersal (*Paralichthyidae*, *Nomeidae*), and estuarine (*Mugilidae*) taxa. Conversely, twelve reef-associated families were observed in UVC but were not recovered by the eDNA methods.

Interestingly, although they were also detected with UVC, eDNA metabarcoding detected three species of cryptic reef fish: *Apogon retrosella*, *Cirrhichthys oxycephalus*, and *Coryphopterus urospilus* (**Figure 21**).



**Figure 21.** Families detected with each monitoring method. UVC: 1. Sparidae; 2. Zanclidae; 3. Scorpaenidae; 4. Labrisomidae; 5. Tetraodontidae; 6. Muraenidae; 7. Chaetodontidae; 8. Synodontidae; 9. Bleniidae; 10. Chaenopsidae; 11. Tripterygiidae; 12. Holocentridae. Both methods: 13. Carangidae; 14. Scaridae; 15. Lutjanidae; 16. Gobiidae; 17. Serranidae; 18. Fistulariidae; 19. Balistidae; 20. Kyphosidae; 21. Elopidae; 22. Sphyraenidae; 23. Mullidae; 24. Pomacanthidae; 25. Haemulidae; 26. Sciaenidae; 27. Acanthuridae; 28. Pomacentridae; 29. Diodontidae; 30. Cirrhitidae; 31. Labridae; 32. Apogonidae. eDNA: 33. Istiophoridae; 34. Scombridae; 35. Paralichthyidae; 36. Nomeidae; 37. Mugilidae; 38. Clupeidae (drawings by Juan Chuy).

### 3.1.2.5 Mock community and negative control

Evaluation of the taxonomic identities in the mock community indicates that 20 of the 22 species included (90.9%) were identified in the sample (Annex BS **Table 20**). Two species, *Hoplopagrus guentheri* and *Kyphosus elegans* produced no sequencing reads in the mock community results despite having been included in the custom reference database. Nevertheless, *K. elegans* was detected in eDNA field samples. I detected one species that was excluded from the mock community sample (6 reads *Mycteroperca rosacea*). For the negative control (which contained four pooled PCR negative controls), only four reads were detected after filtering (two from *Mycteroperca rosacea* and two reads from *Thunnus albacares*). These results indicate negligible levels of foreign DNA contamination during PCR and library preparation.

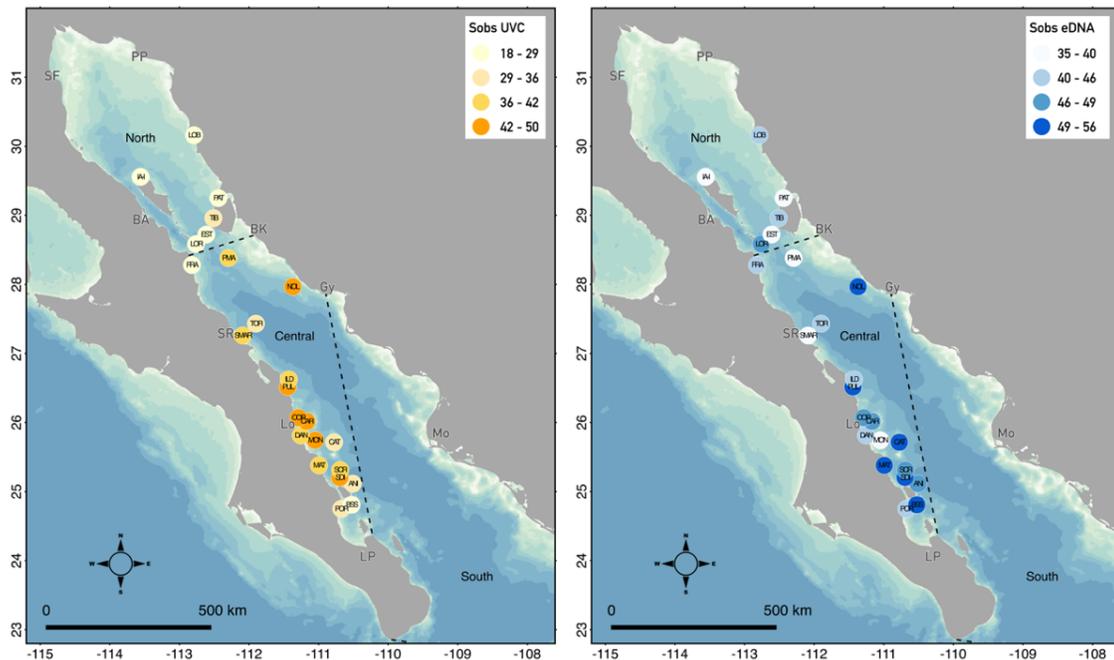
### 3.1.2.6 Community-level ecological analyses

Individual read rarefaction curves (eDNA) reached an asymptote in all sampling sites showing that sequencing depth was sufficient to capture the total components of the fish community. Conversely, individual fish rarefaction curves (UVC) did not show this pattern, indicating that the mean number of transects per site was insufficient to capture the complete species components (Annex B **Figure 42**). Species accumulation curves per site showed an incomplete asymptote in both methods, suggesting the need to increase sampling effort (**Figure 20B**) in terms of the number of sites visited. The total number of extrapolated species ( $S_{exp}$ ) estimated with Chao non-parametric estimator was higher in eDNA metabarcoding, than in UVC (172 and 135, respectively) (**Figure 20B**).

The alpha diversity or total number of observed species/OTUs per site were  $S_{obs}$  eDNA =  $45 \pm 6$  (range 35 – 56) and  $S_{obs}$  UVC  $35 \pm 9$  (range 18 – 50) and showed a heterogeneous geographical distribution among sampling sites (**Figure 22**). These values showed significant differences between their medians when they were compared between methods ( $p = 0.0004$ ) and were not normally distributed (eDNA  $p = 0.49$ , UVC  $p = 0.60$ ) (**Figure 20C**). Spearman's correlation of  $S_{obs}$  between methods was statistically significant ( $S = 1351$ ,  $\rho = 0.41$ ,  $p = 0.04$ ). When comparing mean  $S_{obs}$  between North and Central biogeographic regions, I found significant differences with both methods (eDNA:  $F = 11.28$ ,  $df = 1$ ,  $p = 0.002$ ; UVC  $F = 13.18$ ,  $df = 1$ ,  $p = 0.001$ ) (**Figure 20D**).

Community data showed multivariate heterogeneity within-groups (eDNA:  $F = 5.55$ ,  $Df = 1$ ,  $p = 0.02$ ; UVC:  $F = 6.77$ ,  $Df = 1$ ,  $p = 0.02$ ). Also, the evaluation of multivariate differences between groups with ANOSIM (eDNA:  $R = 0.202$ ,  $p = 0.035$ ; UVC:  $R = 0.605$ ,  $p = 0.0009$ ) and PERMANOVA (eDNA:  $F = 2.31$ ,  $Df = 1$ ,  $p = 0.003$ ; UVC:  $F = 7.54$ ,  $Df = 1$ ,  $p = 0.0001$ ) revealed significant differences in fish assemblages between the North and Central areas of the GC for both survey methods. This distinction was also discriminated by nMDS ordination analysis ( $k = 2$ ) (**Figure 23**) (eDNA: nMDS Stress = 0.217; UVC: nMDS Stress = 0.109). The Mantel tests between beta diversity (Jaccard) and geographic distances among sampling sites revealed a significant correlation for UVC ( $r = 0.62$ ,  $p = 0.0001$ ) but not for eDNA ( $r = 0.06$ ,  $p = 0.18$ ) when including all the species detected with each method separated.

Finally, I used the complementary detection data of reef-fish species obtained with UVC and eDNA metabarcoding to estimate the beta diversity (Jaccard) to evaluate its relationship with the geographical (Euclidean), environmental, and resistance distances.

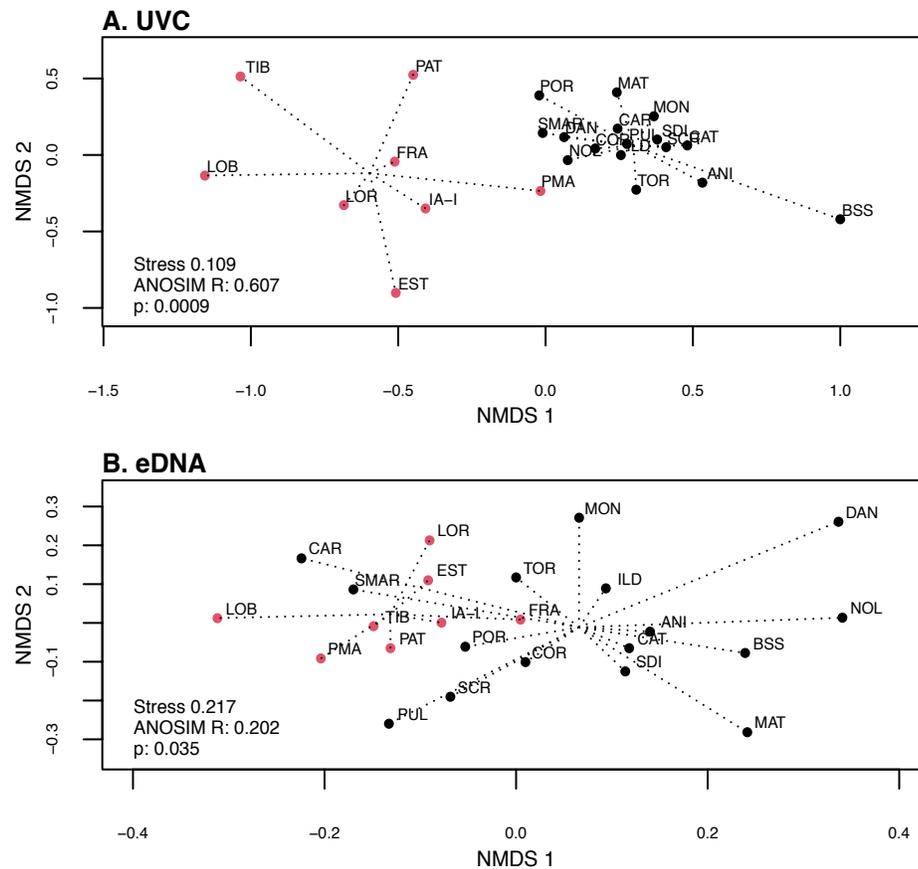


**Figure 22.** Geographical distribution of the number of species (Sobs) detected at each site with each survey method: UVC (left) and eDNA metabarcoding (right). Northern and Central regions are labeled and delimited with dashed lines. Principal cities shown including BA: Bahía de los Ángeles, SR: Santa Rosalía, L: Loreto, LP: La Paz, BK: Bahía de Kino.

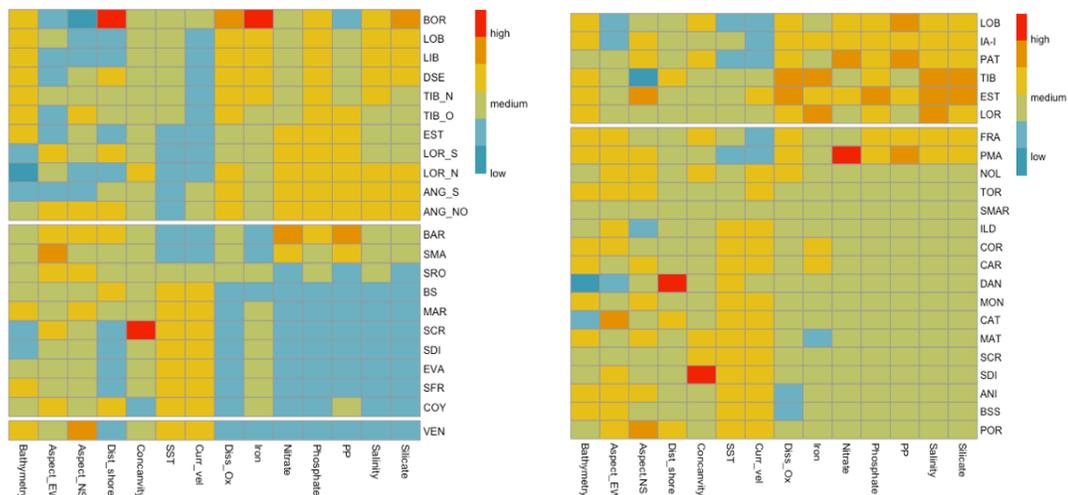
## 3.2 Structural connectivity: estimation of abiotic predictors and testing for mechanisms of isolation

### 3.2.1 Spatial and environmental predictors of functional connectivity

The 14 environmental variables for population and community sampling sites showed considerable irregularity along with the GC. Two groups of variables that varies differentially can be identified: the geomorphological and location-related variables (i.e., bathymetry, aspect East-West, aspect North-South, distance to shore and concavity), and the physicochemical variables (i.e., sea surface temperature, current velocity, dissolved oxygen, iron, nitrate, phosphate, primary productivity, salinity and silicate). The first didn't show a marked latitudinal variation compared with the latter, which evidenced a latitudinal change of environmental values from North to South, with a marked separation between the North and Central regions (Annex B **Figure 43 - 44**). Also there is evidence of some localities that present transitional environmental characteristics (BAR, SMA, and SRO in the population localities; and FRA and PMA for the community localities) (iError! No se encuentra el origen de la referencia.).

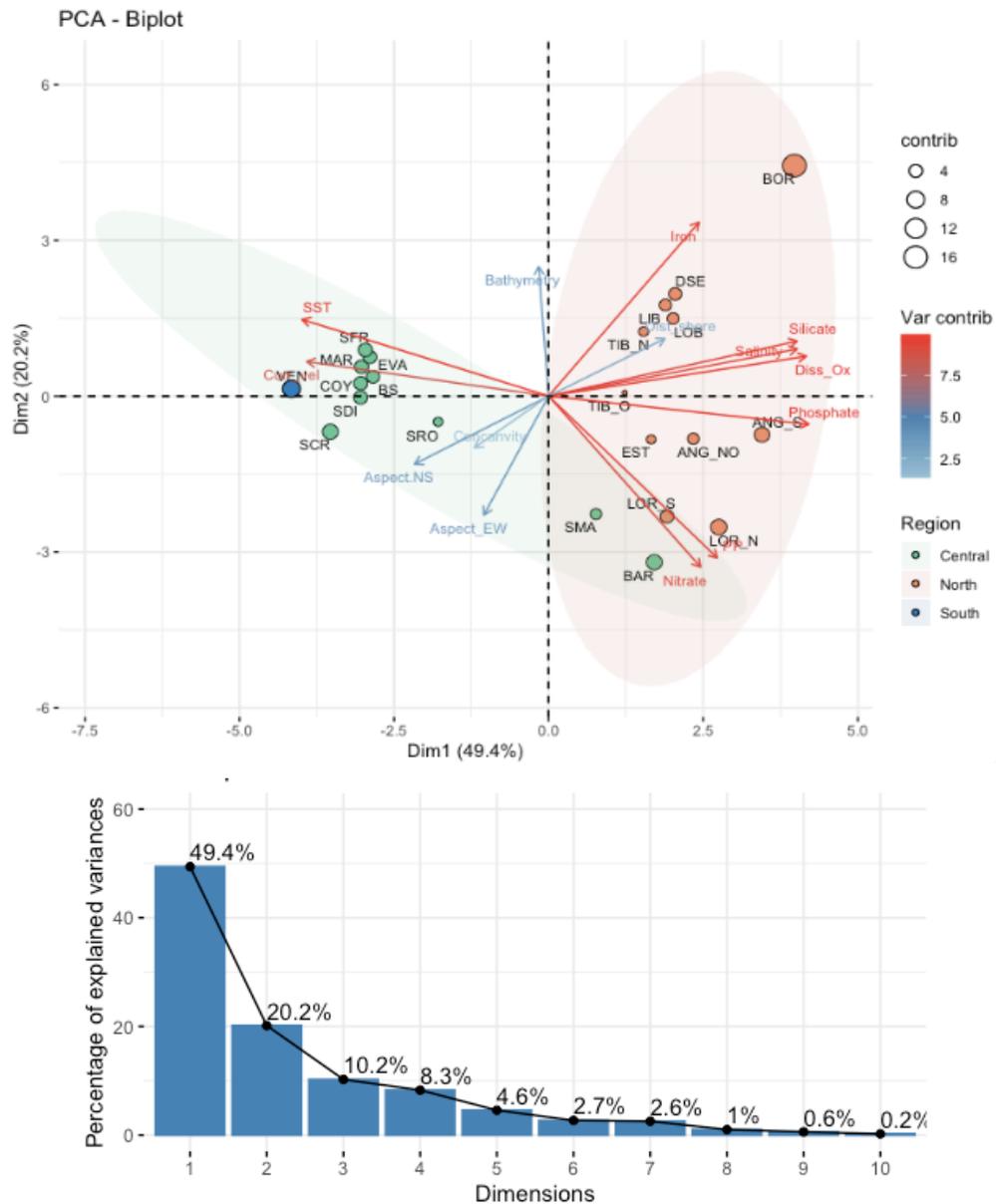


**Figure 23.** Non-metric multidimensional scaling (nMDS) ordinations of community data using the Jaccard similarity index and  $k=2$  for (A) underwater visual census (UVC) and (B) marine environmental DNA (eDNA). Red and Black dots indicate Northern and Central groups, respectively.

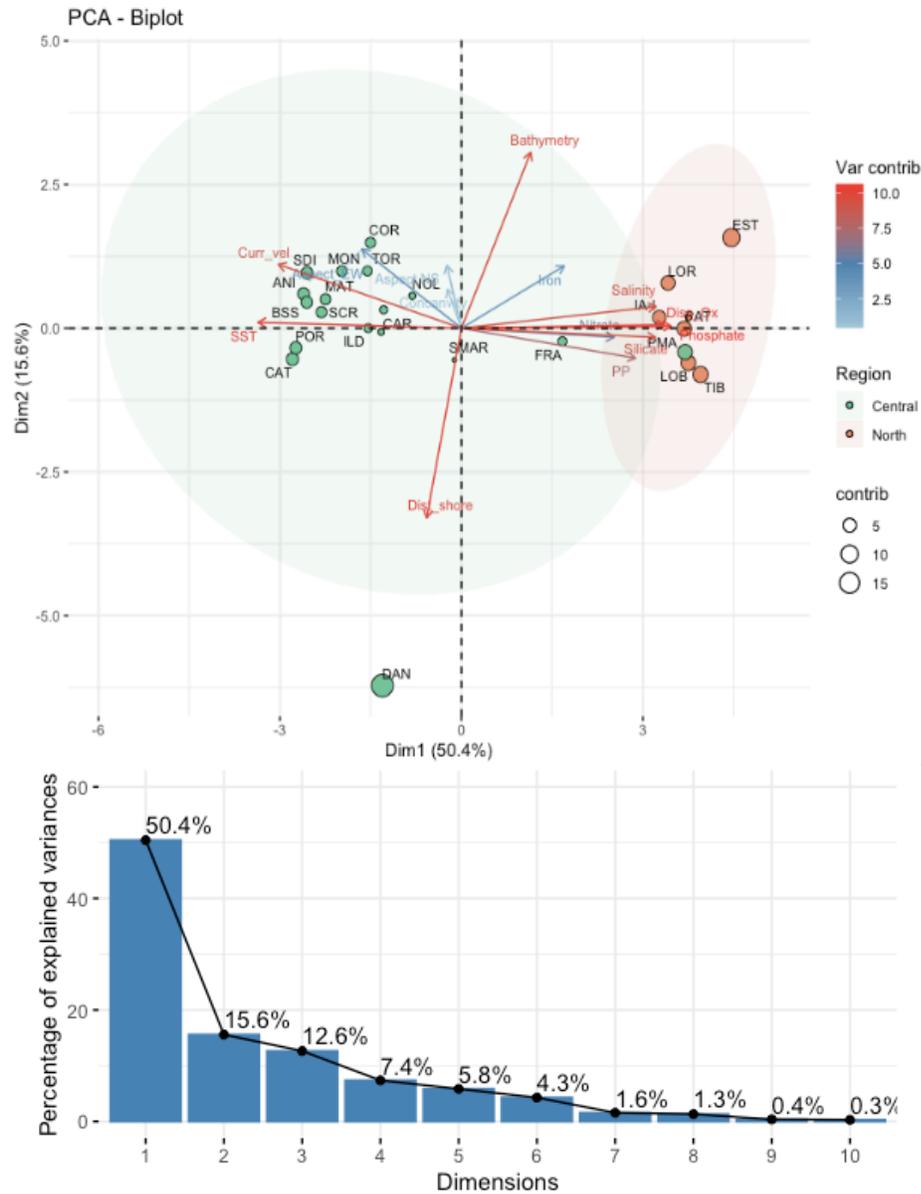


**Figure 24.** Heatmaps of the 14 environmental variables from the population (left panel) and community sampling sites (right panel). Columns and rows are separated per North, Central and South region with white lines.

When observing the PCA biplot, population and community localities were distributed along the first PC biplot axes in a latitudinal matter in both cases, being the localities from the Northern GC grouped in the right side of the graph, and the Central and Southern localities in the left side. In the second PC, we observed that the localities from the North region are distributed longitudinally. This not happen with the Central region localities (**Figure 25** above and **Figure 26** above).



**Figure 25.** Above: PCA biplot of evaluating the 14 environmental variables of the population sampling sites. The color of the arrows indicates the variable contribution to each PC; the shape size indicates its contribution to each PC, the color of the shape indicates the region of the GC, and the shades ellipses were drawn to group the North and Central region localities. Below: Scree-plot of components contribution from the PCA analyses of the population sampling sites.



**Figure 26.** Above: PCA biplot of evaluating the 14 environmental variables of the community sampling sites. The color of the arrows indicates the variable contribution to each PC; the shape size indicates its contribution to each PC, the color of the shape indicates the region of the GC, and the shaded ellipses were drawn to group the North and Central region localities. Below: Scree-plot of components contribution from the PCA analyses of the community sampling sites.

When evaluating the proportion of the variance explained by the eigenvalues in both analyses, the first three components in the population localities represent 49.4%, 69.5% and 79.8% of the variation, respectively; and in community localities 50.4%, 66.0% and 78.6% of the variation, respectively (**Figure 25** below and **Figure 26** below). When evaluating variable contribution results from each PCA, the five variables that contributed most to the first PC of the population localities were phosphates (13.9%),

dissolved oxygen (13%), sea surface temperature (13%), salinity (11.9%) and silicates (11%); for the second PC distance to shore (40.8%), bathymetry (35%), aspect East-West (7%), current velocity (4.6%) and aspect North-South (4.5%); and for the third PC iron (30%), nitrates (20%), aspect North-South (14.5%), concavity (7.5%), aspect East-West (7%). The five variables that contributed most to the first PC of the community localities were phosphates (13.7%), dissolved oxygen (13.4%), salinity (12.5%), silicates (12.5%) and sea surface temperature (12.3%); for the second PC iron (20.6%), nitrates (35%), primary productivity (18.4%), bathymetry (11.8%) and aspect East-West (9.9%); and for the third PC concavity (35.1%), aspect North-South (19.2%), bathymetry (17.6%), aspect East-West (15%) and distance to shore (11.5%).

### 3.2.2 Oceanographic predictors of functional connectivity and network analyses

For the 21 leopard grouper (population-level) hypothetical reproductive seasons, the main differences in the connectivity patterns are related to the intra-annual variability of ocean circulation of the GC, covering March-June (Spring) and June-September (Summer) seasons. In terms of graph sizes, the Spring network (171 links and density of 0.37) was more extensive than the Summer (146 links and density of 0.31), representing more complex relationships among nodes (**Figure 27**). As inferred from oceanic currents, patterns of larval dispersal changed directions from Spring, with an anticyclonic (clockwise) circulation that caused connections predominantly northward, to Summer with a cyclonic (anticlockwise) circulation with predominantly southward connections.

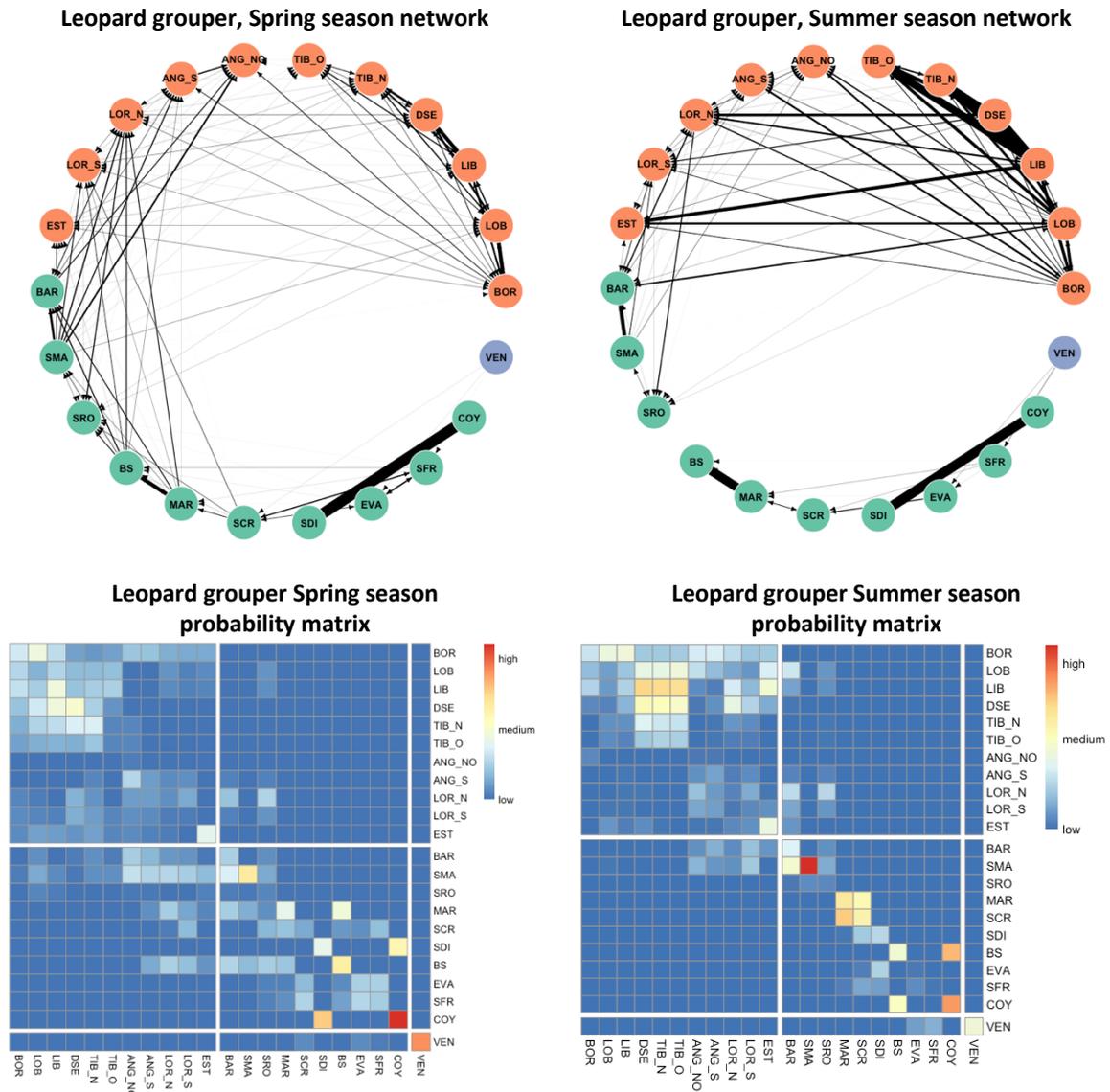
Network evidenced that complex associations characterize the Spring season with a marked North-Central-South regionalization in the connectivity patterns. In the North region of the GC, a group of high connectivity nodes occurs within North-East (BOR, LOB, LIB, DSE, and insular TIB\_N, and TIB\_O) (**Figure 27**). One of the main features of the Spring network is the northward flux of larvae that connects the Central and South region to the North region via four principal nodes: BAR, SMA, SRO, BS, MAR, SCR, and BS. In the Southern region, the VEN node exports its larvae to the Central GC nodes of SCR, EVA, and SFR, and imports 0% of larvae. VEN also presents the second-highest local retention value of all the nodes, with 30.3% of its larvae retained, preceded by COY (44% of larvae retention). Finally, it is of relevance to mention that in this season, there is no flux of larvae from the Northern region to the Southern region, nor from the Central region to the Southern region (**Figure 27**).

During the Summer season, the North-Central-South regionalization is still present. The number of connections in the Northern region is more numerous than in the Central (**Figure 27**). The mainland north-eastern nodes of BOR, LOB, LIB, and DSE present medium-high larval net export and net import values. Nodes on the Midriff Islands present low larval net export and medium net (**Figure 28**). The intense larval flux from the Central region to the North region characteristic of the Spring season is absent during summer, which produces a discontinuity of the network from the BS node to the South, and from the SMA node to the North. Moderate connectivity within the Central region. Finally, in the Southern region, the VEN node export 1.2% of its larvae to the Central nodes EVA and SFR and imports 0% of larvae. In this season, the SMA node presents the highest local retention value (37% of its larvae are retained) compared to all the resting nodes in the GC, followed by COY (23.4% of larval retention) (**Figure 27**).

On the other hand, the node centrality metrics estimated for Spring and Summer population-level hypothetical reproductive season (**Figure 28**) reflect the regional, seasonal variation described above, in which a latitudinal trend is present. The North region presents a more complex larval interchange configuration in both seasons and higher network centrality metrics values than the Central and South regions. This is reflected in their higher degree values (out, in, degree) hub and authority characteristics.

The Central region nodes BAR, SMA and SRO, which are crucial for establishing metapopulation connectivity between the North and the Central regions during Spring, present medium-high values of network centrality metrics (out, in, degree, central betweenness, hub, and authority). The Central and Southern nodes EVA, SFR, COY, and VEN showed medium-low node centrality values in both seasons, except for COY. During Spring, the North is mainly a sink and the Central region a source of larvae. In the Summer, the Central region changes to the opposite (sink). Also, the highest net import values occur in the North localities LOB and LIB during Spring, and DSE, TIB\_N, and TIB\_O in Summer. The net export peaks at various localities of the GC during Spring (BOR, DSE, SMA, and COY) and in LIB during Summer. VEN, COY presented high local retention during Spring and SMA during Summer (**Figure 28**).

The four rocky-reef fishes (community-level) hypothetical reproductive seasons showed a regionalization in the connectivity patterns. Graph size was larger during Fall (198 links and density 0.39), followed by Spring (187 links and density 0.37), Winter (168 links and density 0.33), and Summer (135 links and density 0.26). Oceanic currents changed directions from an anticyclonic (clockwise) circulation that caused connections predominantly northward to a cyclonic (anticlockwise) circulation with predominantly southward links from Winter to Fall (**Figure 29** and **Figure 30**).

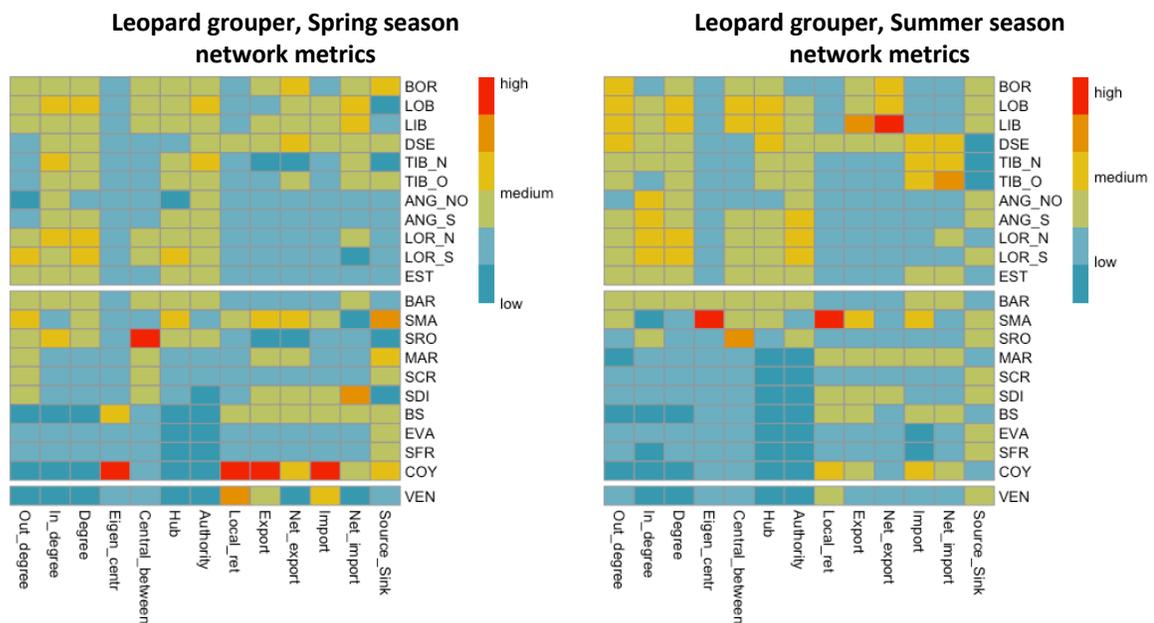


**Figure 27.** March-June (Spring) (left) and June-September (Summer) (right) seasons. Nodes are colored according to the GC region they belong to North (orange), Central (green), or South (blue). Arrowheads indicate the direction of the larval dispersal, and the width of the arrow line means the proportion of larvae being dispersed. Below: Adjacency matrices of potential larval connectivity of the 22 leopard grouper's localities in the GC for four of the Spring (March-June) (left) and Summer (June-September) seasons (right). Columns indicate the source localities in these matrices, and the rows indicate the target localities. The colors of the cells show the amount of larval dispersal between the two adjacent localities. Diagonal represents the local retention of larvae. Columns and rows are separated per North, Central, and South regions with white lines.

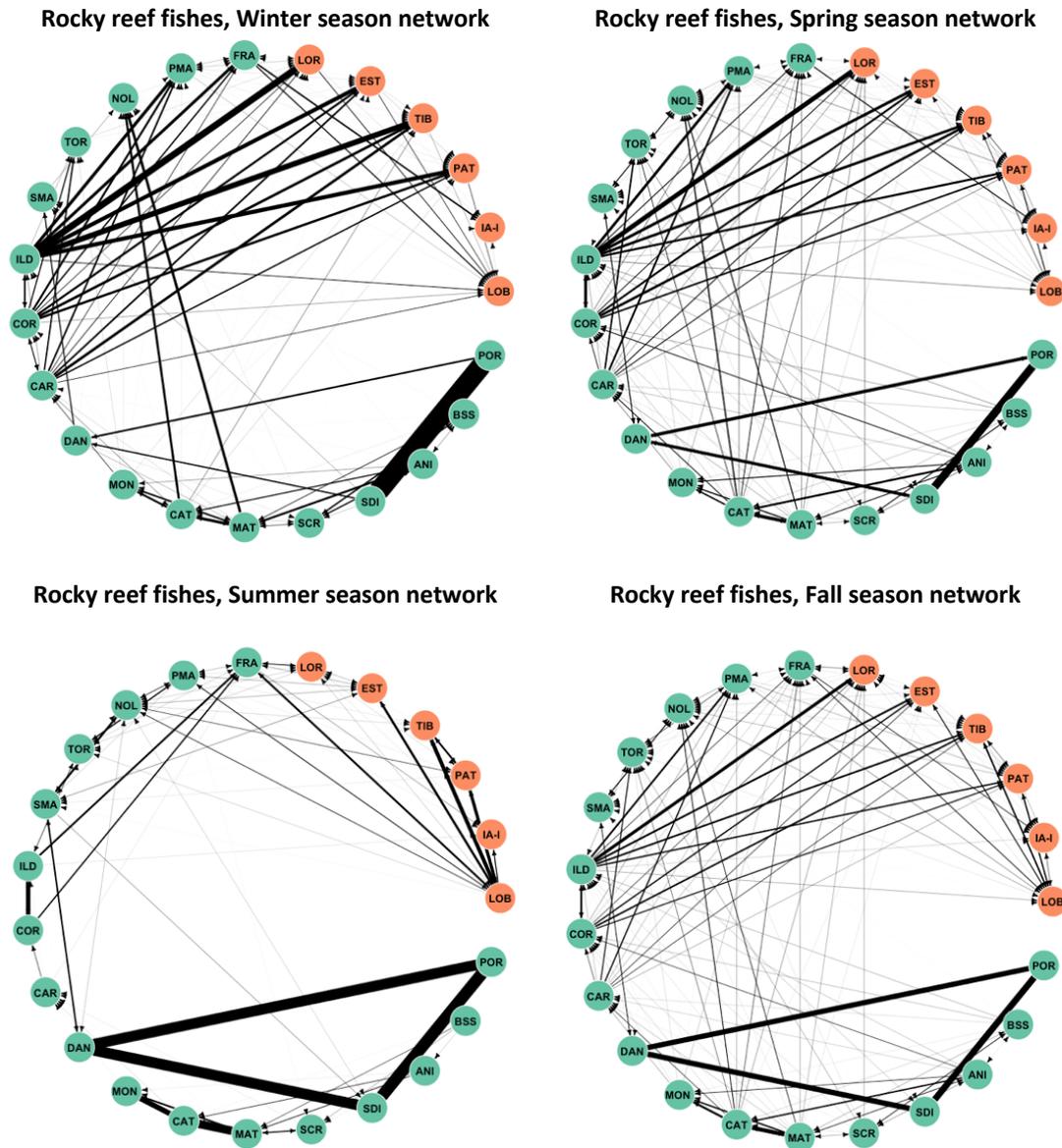
The northern region of the GC presents a set of nodes with high connectivity between them (LOB, IA-A, PAT, TIB, EST, and LOR). Then, to the south, a group of nodes that maintain connectivity between the Central and North regions and, in turn, present high connectivity between them (FRA, ILD, COR, and CAR). The DAN node presents low connectivity with all the other nodes, except with SDI, and COY to which it exports a large proportion of its larvae (towards the south), and NOL and TOR from which it receives a

proportion of larvae. This occurs in all seasons except Winter, in which connectivity to NOL and TOR is lost. Subsequently, a group of Center-South nodes that present high connectivity between them (MON, CAT, MAT, SCR, SDI, ANI, and COY), mainly with a southerly direction in all seasons, except in Winter, in which they present connections both to the north and to the south (**Figure 29** and **Figure 30**).

On the other hand, the node centrality metrics estimated per community level hypothetical reproductive seasons also reflect the regional seasonal variation described above, in which a latitudinal trend is present (**Figure 31**). Central localities present higher out-degree values during Winter and Fall, and Northern localities during Spring and Summer. Northern nodes present high in-degree values during Winter and Fall and in less degree in Spring and Summer. On the contrary, Central nodes show high in-degree values mainly during Spring. According to the hub and authority latitudinal distribution the prevalent connectivity during Winter is from Central-North nodes to the North; in Spring from the Central-North nodes to the Central-South nodes; in Summer the connectivity is restricted to the North and North-Central region, and in the Fall the connectivity goes from the Central nodes to the North and within. The net export peaks at various localities of the GC during Winter (ILD, LOT, and CAR), during Spring (ILD, SRO, and LOB), during Summer (DAN, LOB, and SDI) and in Fall (ILD, SAN, and SDI). mainly on the northern sites BOR, LOB, LIB, and DSE. VEN, COY, and SMA presented high local retention in both seasons (**Figure 31**).



**Figure 28.** Heatmaps of node network metrics for the 22 leopard groupers localities of the GC for two seasons: Spring (left) and Summer (right). Values were standardized and centered by column.



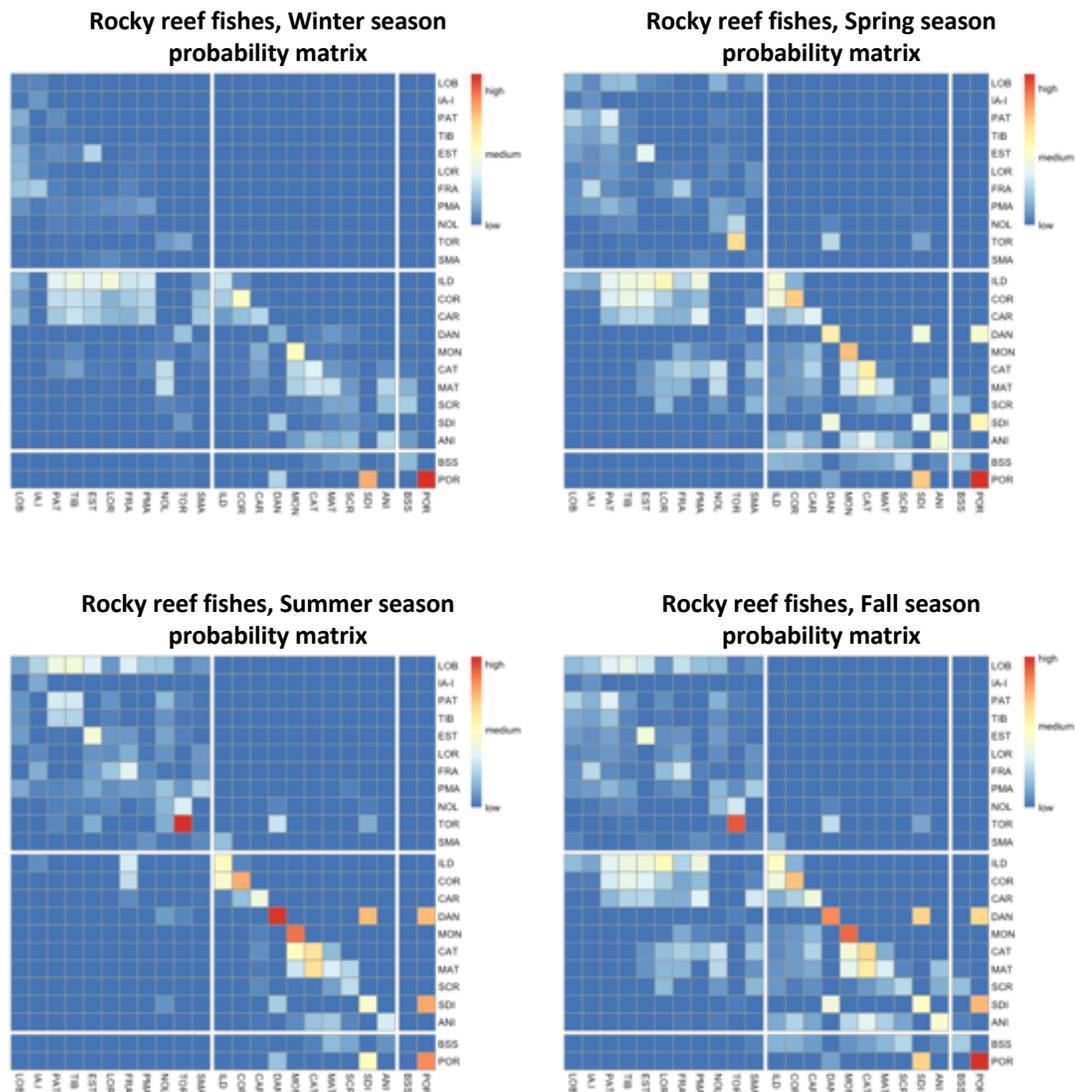
**Figure 29.** Circle network representation of potential larvae connectivity among the 23 rocky-reef fishes localities in the GC for four seasons: Winter (above left), Spring (above right), Summer (below left), and Fall (below right). Nodes are colored according to the GC region they belong: North (orange) and Central (green). Arrowheads indicate the direction of the larval dispersal, and the width of the arrow line means the proportion of larvae being dispersed.

### 3.2.3 Testing for mechanisms of isolation: integrating population genomic and community connectivity patterns and the seascape predictors

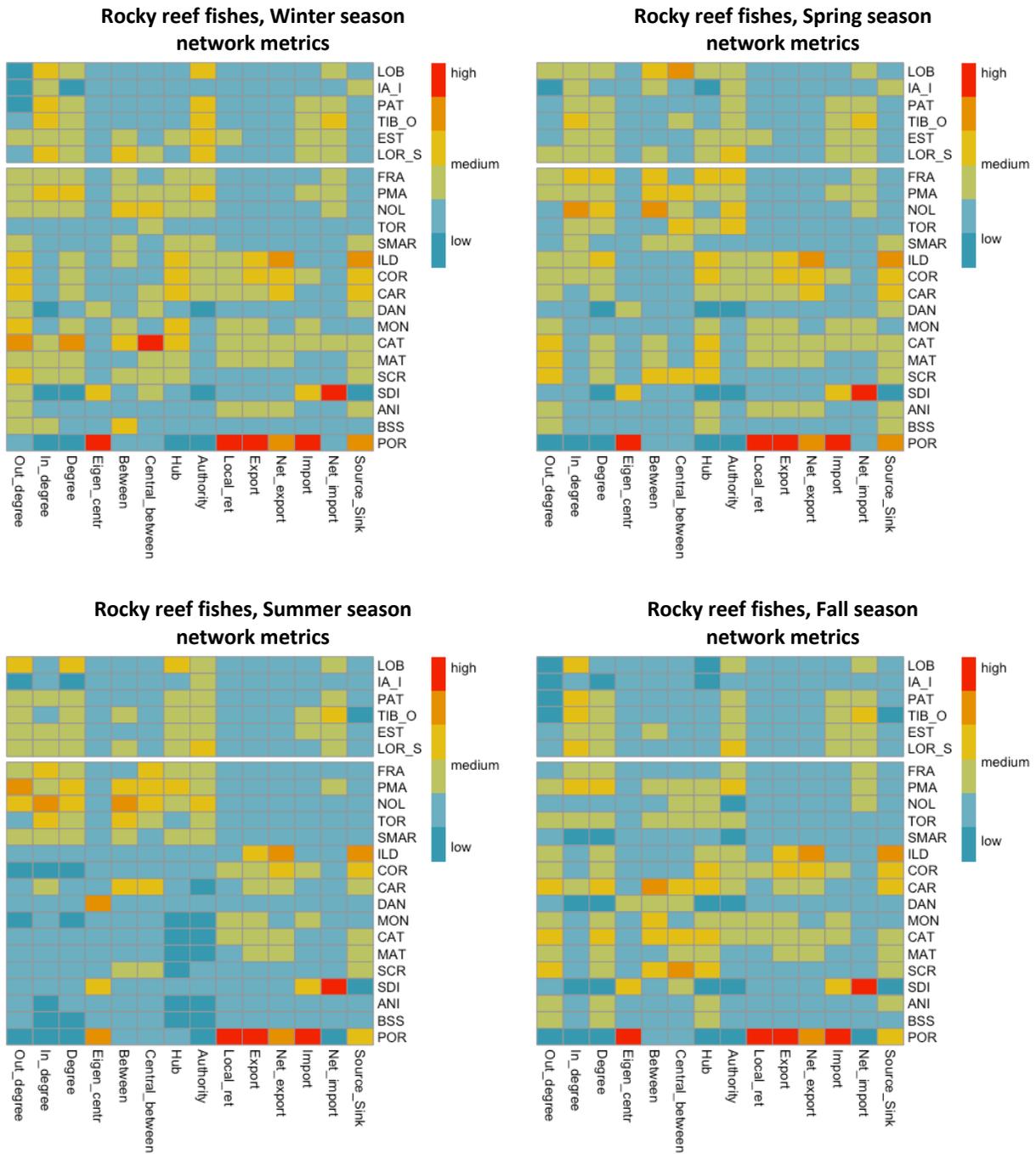
#### 3.2.3.1 Mantel test and multiple regression on distance matrices (link-level analyses)

The population genomic differentiation of the leopard grouper ( $F_{ST}$ , Weir and Cockerman 1984) and the logarithm of the Euclidean distance among 22 studied localities in the GC (Table 12, Figure 32 and

**Figure 33**), presented contrasting results when evaluating for an IBD pattern with Mantel test and MRM ( $r = 0.123$ ,  $p = 0.048$ ;  $R^2 = 0.015$ ,  $p = 0.097$ ). When comparing data from regions, neither North ( $r = -0.096$ ,  $p = 0.661$ ;  $R^2 = 0.009$ ,  $p = 0.483$ ) or Central ( $r = 0.102$ ,  $p = 0.261$ ;  $R^2 = 0.040$ ,  $p = 0.302$ ) presented significant correlations with the geographical distance, rejecting the null hypothesis of IBD. When evaluating localities from SRO to VEN, both tests showed statistically significant results ( $r = 0.482$ ,  $p = 0.042$ ;  $R^2 = 0.232$ ,  $p = 0.039$ ) demonstrating that for this subregion of the GC, the genomic differentiation of the leopard grouper present and IBD pattern in which the geographical distance among sites explains 23% of the genomic differentiation (**Table 12** and **Figure 33**).



**Figure 30.** Adjacency matrices of potential larval connectivity of the 23 rocky-reef fishes localities in the GC for four seasons: Winter (above left), Spring (above right), Summer (below left) and Fall (below right). In this matrices columns indicate the source localities and the rows indicate the target localities. The colors of the cells indicates the amount (of larval dispersal between the two adjacent localities. Diagonal represents the local retention of larvae.



**Figure 31.** Heatmaps of node network metrics of the 23 rocky-reef fishes localities in the GC for four seasons: Winter (above left), Spring (above right), Summer (below left) and Fall (below right). Values were standardized and centered by column.

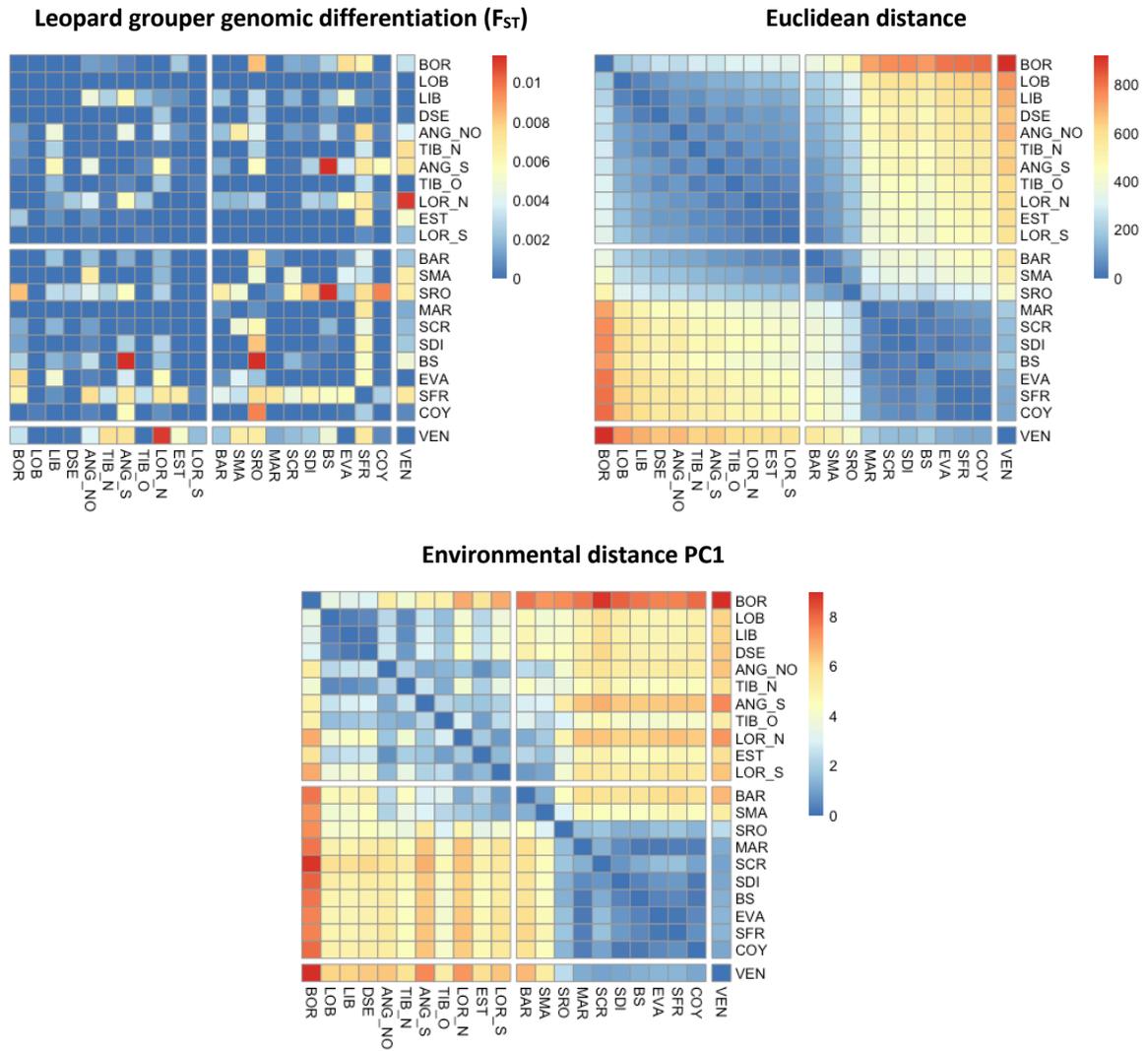
### 3.2.4 Testing for mechanisms of isolation: integrating population genomic and community connectivity patterns and the seascape predictors

#### 3.2.4.1 Mantel test and multiple regression on distance matrices (link-level analyses)

The population genomic differentiation of the leopard grouper ( $F_{ST}$ , Weir and Cockerman 1984) and the logarithm of the Euclidean distance among 22 studied localities in the GC (**Table 12**, **Figure 32** and **Figure 33**), presented contrasting results when evaluating for an IBD pattern with Mantel test and MRM ( $r = 0.123$ ,  $p = 0.048$ ;  $R^2 = 0.015$ ,  $p = 0.097$ ). When comparing data from regions, neither North ( $r = -0.096$ ,  $p = 0.661$ ;  $R^2 = 0.009$ ,  $p = 0.483$ ) or Central ( $r = 0.102$ ,  $p = 0.261$ ;  $R^2 = 0.040$ ,  $p = 0.302$ ) presented significant correlations with the geographical distance, rejecting the null hypothesis of IBD. When evaluating localities from SRO to VEN, both tests showed statistically significant results ( $r = 0.482$ ,  $p = 0.042$ ;  $R^2 = 0.232$ ,  $p = 0.039$ ) demonstrating that for this subregion of the GC, the genomic differentiation of the leopard grouper present and IBD pattern in which the geographical distance among sites explains 23% of the genomic differentiation (**Table 12** and **Figure 33**).

Mantel tests ( $r = 0.140$ ,  $p = 0.027$ ) and MRM ( $R^2 = 0.015$ ,  $p = 0.033$ ) between the genomic differentiation of the leopard grouper ( $F_{ST}$ , Weir and Cockerman 1984) and the environmental distance among 22 studied localities in the GC, presented significant results when assessing for an IBE pattern when including only the first PC component, evidencing that the environmental differences of the mainly contributing variables of the PC1 along the GC, explain the 1.5% of its genomic differentiation (**Table 12**, **Figure 32** and **Figure 33**). When comparing data from regions, neither North ( $r = -0.016$ ,  $p = 0.518$ ;  $R^2 = 0.000$ ,  $p = 0.923$ ) and Central ( $r = -0.078$ ,  $p = 0.546$ ;  $R^2 = 0.006$ ,  $p = 0.7744$ ) presented significant correlations in the Mantel tests or the MRM rejecting the hypothesis of IBE. When evaluating localities from SRO to VEN, tests didn't showed statistically significant results ( $r = -0.197$ ,  $p = 0.602$ ;  $R^2 = 0.038$ ,  $p = 0.514$ ) (**Table 12**). Mantel test didn't showed significant results when including the information of PC2 and PC3.

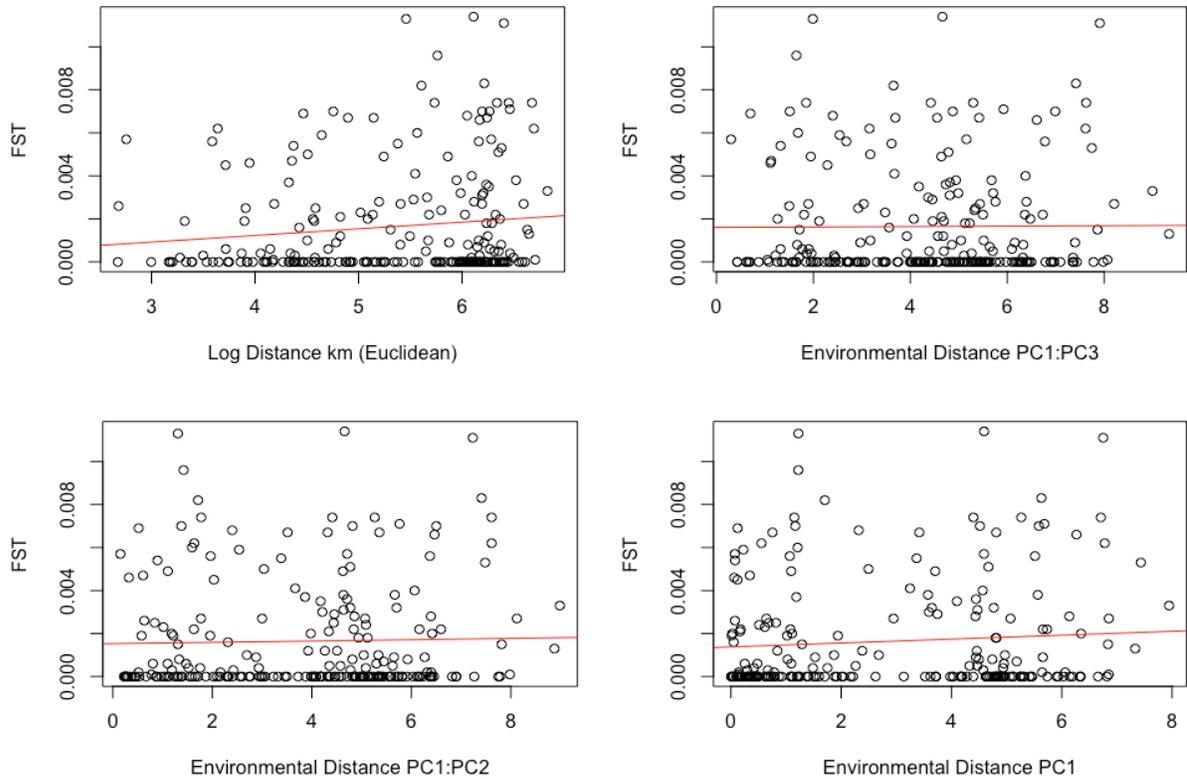
Mantel tests between genomic differentiation ( $F_{ST}$ , Weir and Cockerman 1984) and resistance distance matrices for all the GC and the North and Central regions didn't show statistically significant results for 21 Mantel tests, except for the June-July ( $R^2 = 0.147$ ,  $p = 0.034$ ) and the April-September seasons ( $R^2 = 0.097$ ,  $p = 0.045$ ) seasons in the Northern GC, indicating that the IBR hypothesis can be rejected only for all the GC and Central region (**Table 13**).



**Figure 32.** Leopard grouper's distance matrices of the pairwise genomic differentiation  $F_{ST}$  (Weir and Cockerham, 1984), and Euclidean (km) and Environmental distances among the 22 localities in the GC.

**Table 12.** Results of the Mantel and MRM tests were applied to the leopard grouper populations comparing the effects of the logarithm of Euclidean distance and the environmental distance (PC1, PC1:PC2, and PC1:PC3) into the neutral genomic differentiation ( $F_{ST}$ , Weir, and Cockerman 1984). Results are presented for all the Gulf of California, per region (North and Central) and for the Santa Rosalía (SRO) to Ventana (VEN) region. Two statistics that evaluate the correlation between matrices are presented: Mantel  $r$  and  $p$ -value of the Mantel test and the Multiple regression of the distance matrices (MRM)  $R^2$  and  $p$ -value of the regression. Red letters represent statistically significant tests.

Predictor	Mantel $r$ ( $p$ -value)	MRM $R^2$ ( $p$ -value)
<b>Euclidean distance (logarithm)</b>		
All the Gulf	0.123 (0.048)	0.015 (0.097)
North region	-0.096 (0.661)	0.009 (0.483)
Central region	0.102 (0.261)	0.040 (0.302)
SRO-VEN	0.482 (0.042)	0.232 (0.039)
<b>Environmental distance PC1:PC3</b>		
All the Gulf	0.064 (0.238)	0.004 (0.492)
North region	0.087 (0.336)	0.001 (0.854)
Central region	-0.123 (0.637)	0.015 (0.638)
SRO-VEN	-0.052 (0.481)	0.002 (0.874)
<b>Environmental distance PC1:PC2</b>		
All the Gulf	0.064 (0.218)	0.004 (0.468)
North region	-0.042 (0.518)	0.001 (0.825)
Central region	-0.136 (0.686)	0.018 (0.609)
SRO-VEN	-0.004 (0.416)	0.000 (0.987)
<b>Environmental distance PC1</b>		
All the Gulf	0.140 (0.027)	0.015 (0.033)
North region	-0.016 (0.518)	0.000 (0.923)
Central region	-0.078 (0.546)	0.006 (0.774)
SRO-VEN	-0.197 (0.602)	0.038 (0.514)



**Figure 33.** Isolation by distance patterns of the 22 leopard grouper's populations relating the neutral genomic differentiation ( $F_{ST}$ , Weir and Cockerman 1984) to the Euclidean (above left panel) and the environmental distances (above right, below left and right panels).

Linear models among the network metrics of Spring and Summer seasons and the genomic diversity metrics of the leopard grouper presented a statistically significant relationship with several genomic diversity metrics: in degree presented a negative relationship with  $H_e$  in Spring and with  $A_r$  and  $H_e$  in Summer; degree a negative association with  $H_e$  in Spring and negative with  $H_e$  in Summer; central betweenness a negative correlation with  $F_{IS}$  in Spring and negative with  $A_r$  and  $H_e$  in Summer; authority a negative association with  $A_r$  and  $H_e$  in Spring and Summer; local retention a positive relationship with  $H_e$  in Spring and with  $A_r$  and  $H_e$  in Summer; export a positive relationship with  $A_r$  and  $H_e$  in Spring; source-sink a positive relationship with  $H_o$  in Spring, and negative with  $F_{IS}$  in Summer; import a positive correlation with  $F_{IS}$  in Summer; and net import a positive relationship with  $F_{IS}$  during Summer (**Table 14**, **Figure 34** and **Figure 35**).

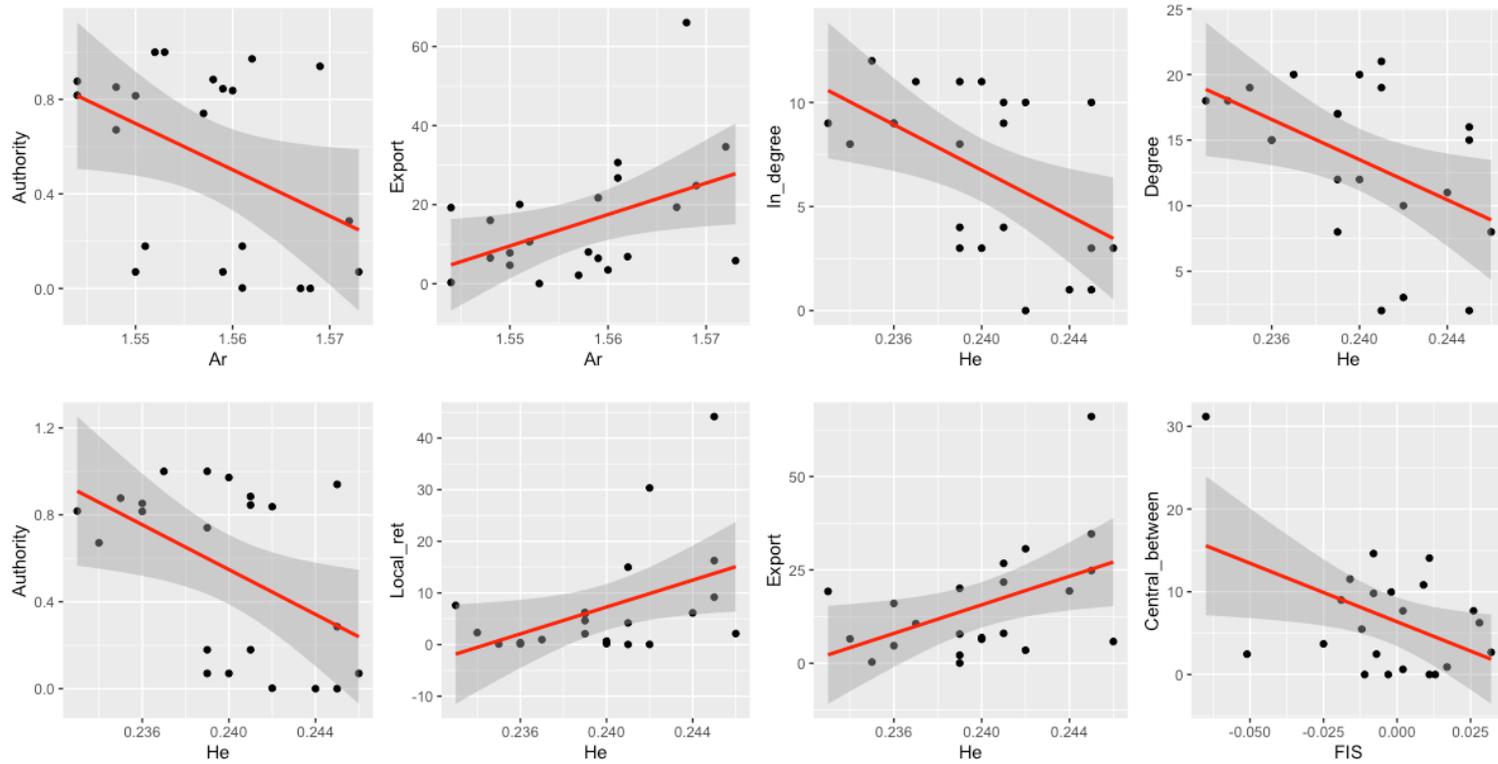
**Table 13.** Results of the Mantel tests and MRM evaluated on the leopard grouper populations comparing effects of resistance distances (from 22x22 polygon probability matrix) into genomic differentiation ( $F_{ST}$ , Weir and Cockerman 1984) in the All the Gulf, North region and Central region. Red letters represents statistically significant tests.

Season	All the Gulf (BOR-VEN)		North (BOR-LOR_S)		Central (BAR-COY)	
	Mantel ape ( <i>p</i> -value)	MRM R <sup>2</sup> ( <i>p</i> -value)	Mantel ape ( <i>p</i> -value)	MRM R <sup>2</sup> ( <i>p</i> -value)	Mantel ape ( <i>p</i> -value)	MRM R <sup>2</sup> ( <i>p</i> -value)
March-April	0.404	0.008 (0.390)	0.880	0.044 (0.107)	0.605	0.037 (0.387)
April-May	0.850	0.006 (0.539)	0.696	0.000 (0.964)	0.818	0.013 (0.685)
May-June	0.307	0.003 (0.676)	0.282	0.004 (0.821)	0.483	0.038 (0.337)
June-July	0.361	0.003 (0.604)	0.078	0.147 (0.034)	0.810	0.001 (0.815)
July-August	0.264	0.012 (0.346)	0.864	0.024 (0.474)	0.558	0.009 (0.586)
August-September	0.127	0.031 (0.088)	0.274	0.019 (0.435)	0.909	0.000 (0.938)
March-May	0.916	0.008 (0.531)	0.957	0.003 (0.734)	0.940	0.007 (0.804)
April-June	0.613	0.006 (0.587)	0.690	0.015 (0.514)	0.494	0.032 (0.517)
May-July	0.298	0.008 (0.552)	0.396	0.023 (0.299)	0.720	0.017 (0.676)
June-August	0.245	0.013 (0.282)	0.867	0.018 (0.566)	0.094	0.041 (0.117)
July-September	0.342	0.007 (0.451)	0.624	0.003 (0.972)	0.144	0.022 (0.286)
March-June	0.667	0.007 (0.596)	0.620	0.012 (0.467)	0.409	0.031 (0.540)
April-July	0.679	0.008 (0.598)	0.572	0.003 (0.748)	0.473	0.033 (0.500)
May-August	0.845	0.009 (0.417)	0.082	0.025 (0.354)	0.128	0.048 (0.212)
June-September	0.205	0.013 (0.264)	0.548	0.000 (0.891)	0.967	0.002 (0.781)
March-July	0.592	0.008 (0.575)	0.667	0.000 (0.975)	0.484	0.033 (0.532)
April-August	0.668	0.008 (0.595)	0.759	0.026 (0.350)	0.296	0.040 (0.321)
May-September	0.667	0.008 (0.574)	0.245	0.097 (0.052)	0.350	0.038 (0.351)
March-August	0.661	0.008 (0.584)	0.263	0.000 (0.994)	0.315	0.027 (0.335)
April-September	0.720	0.008 (0.599)	0.244	0.097 (0.045)	0.691	0.031 (0.548)
March-September	0.700	0.008 (0.581)	0.280	0.054 (0.203)	0.709	0.320 (0.572)

**Table 14.** Linear models of network metrics and population genomic diversity metrics for March-June (Spring) and June-September (Summer) season's. Red letters represents statistically significant tests.

Network metric	March-June season (Spring)				June-September season (Summer)			
	Allelic richness R <sup>2</sup> (p-value)	Ho R <sup>2</sup> (p-value)	He R <sup>2</sup> (p-value)	F <sub>IS</sub> R <sup>2</sup> (p-value)	Allelic richness R <sup>2</sup> (p-value)	Ho R <sup>2</sup> (p-value)	He R <sup>2</sup> (p-value)	F <sub>IS</sub> R <sup>2</sup> (p-value)
<b>Out degree</b>	-0.026 (0.502)	-0.045 (0.770)	0.021 (0.240)	-0.039 (0.664)	0.000 (0.329)	0.035 (0.198)	0.056 (0.148)	-0.021 (0.464)
<b>In degree</b>	-0.030 (0.546)	-0.039 (0.653)	<b>0.220 (0.016)</b>	-0.030 (0.546)	<b>0.157 (0.038)</b>	-0.046 (0.792)	<b>0.246 (0.011)</b>	-0.008 (0.377)
<b>Degree</b>	0.114 (0.068)	-0.038 (0.643)	<b>0.194 (0.022)</b>	-0.020 (0.456)	0.106 (0.075)	-0.007 (0.370)	<b>0.207 (0.019)</b>	-0.040 (0.678)
<b>Eigenvector centrality</b>	0.036 (0.195)	-0.049 (0.903)	0.050 (0.162)	-0.020 (0.456)	0.070 (0.124)	-0.030 (0.549)	0.019 (0.249)	-0.040 (0.678)
<b>Central between</b>	0.096 (0.087)	0.026 (0.225)	0.111 (0.070)	<b>0.241 (0.011)</b>	<b>0.148 (0.043)</b>	-0.026 (0.511)	<b>0.248 (0.010)</b>	<b>0.172 (0.031)</b>
<b>Hub</b>	0.018 (0.253)	-0.013 (0.408)	0.103 (0.079)	-0.049 (0.951)	0.008 (0.288)	0.034 (0.201)	0.082 (0.104)	-0.027 (0.520)
<b>Authority</b>	<b>0.138 (0.049)</b>	-0.019 (0.450)	<b>0.191 (0.024)</b>	-0.048 (0.871)	<b>0.150 (0.042)</b>	-0.034 (0.588)	<b>0.236 (0.023)</b>	-0.03 (0.606)
<b>Local retention</b>	0.135 (0.051)	-0.047 (0.844)	<b>0.156 (0.038)</b>	0.028 (0.217)	<b>0.231 (0.013)</b>	-0.049 (0.966)	<b>0.196 (0.021)</b>	-0.008 (0.377)
<b>Source-Sink</b>	-0.043 (0.736)	-0.038 (0.643)	0.078 (0.111)	0.054 (0.153)	-0.039 (0.657)	<b>0.186 (0.025)</b>	-0.041 (0.693)	<b>0.222 (0.015)</b>
<b>Export</b>	<b>0.179 (0.028)</b>	-0.045 (0.763)	<b>0.195 (0.022)</b>	0.046 (0.169)	-0.029 (0.534)	-0.021 (0.462)	-0.042 (0.716)	-0.010 (0.388)
<b>Import</b>	0.020 (0.246)	0.087 (0.098)	0.014 (0.266)	0.222 (0.015)	0.020 (0.245)	0.087 (0.098)	0.014 (0.267)	<b>0.222 (0.015)</b>
<b>Net export</b>	0.103 (0.079)	-0.042 (0.710)	0.102 (0.080)	0.013 (0.269)	-0.030 (0.545)	-0.012 (0.398)	-0.012 (0.402)	-0.036 (0.618)
<b>Net import</b>	0.020 (0.245)	0.048 (0.167)	0.037 (0.192)	-0.010 (0.388)	-0.039 (0.657)	<b>0.186 (0.025)</b>	-0.041 (0.693)	<b>0.222 (0.015)</b>

Leopard grouper, linear models of the genomic diversity and the network centrality metrics for the Spring season



**Figure 34.** Plots of the linear models between the network centrality metrics and the genomic diversity of the leopard grouper for the Spring season. The red line is the linear regression of the values.

Leopard grouper, linear models of the genomic diversity and the network centrality metrics for the Summer season

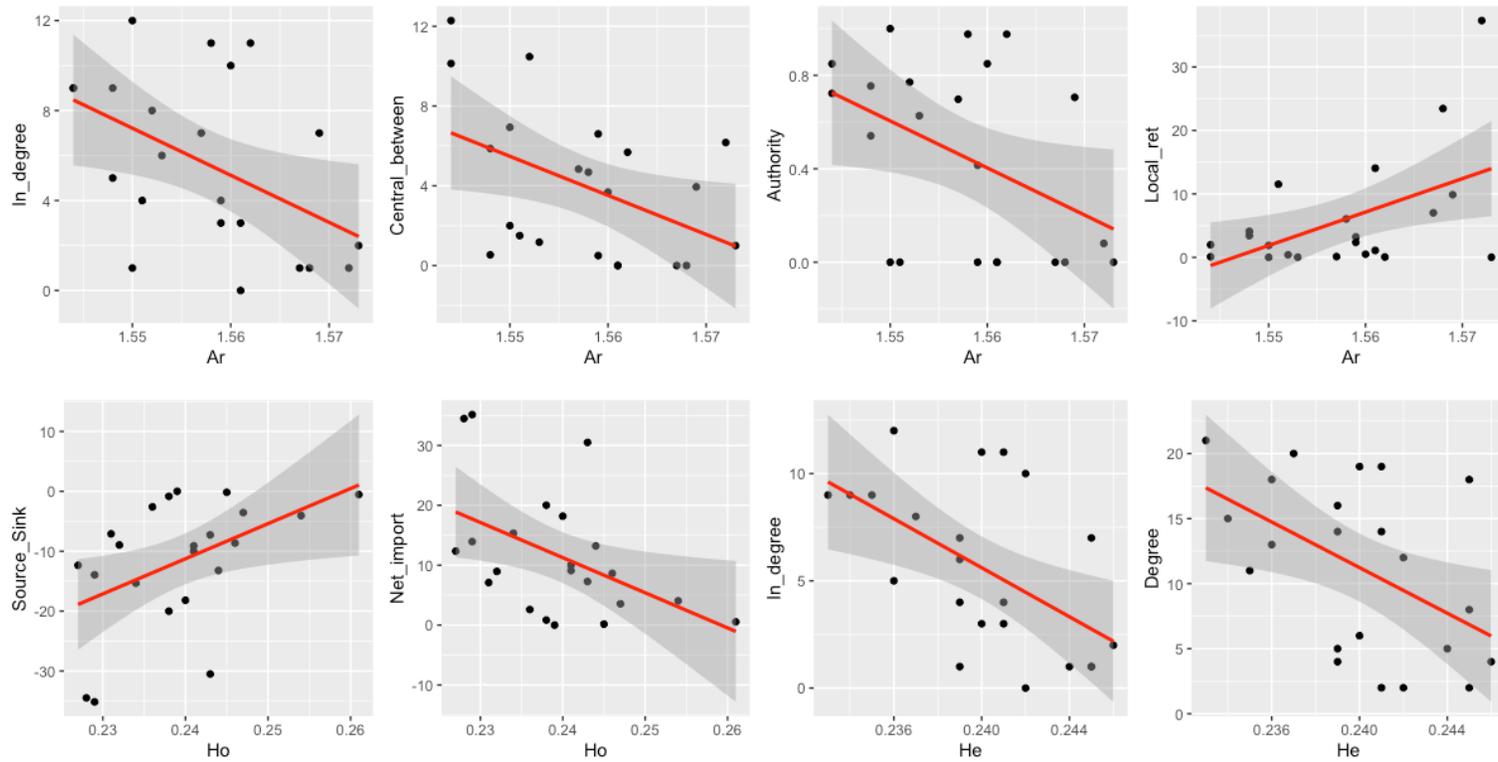
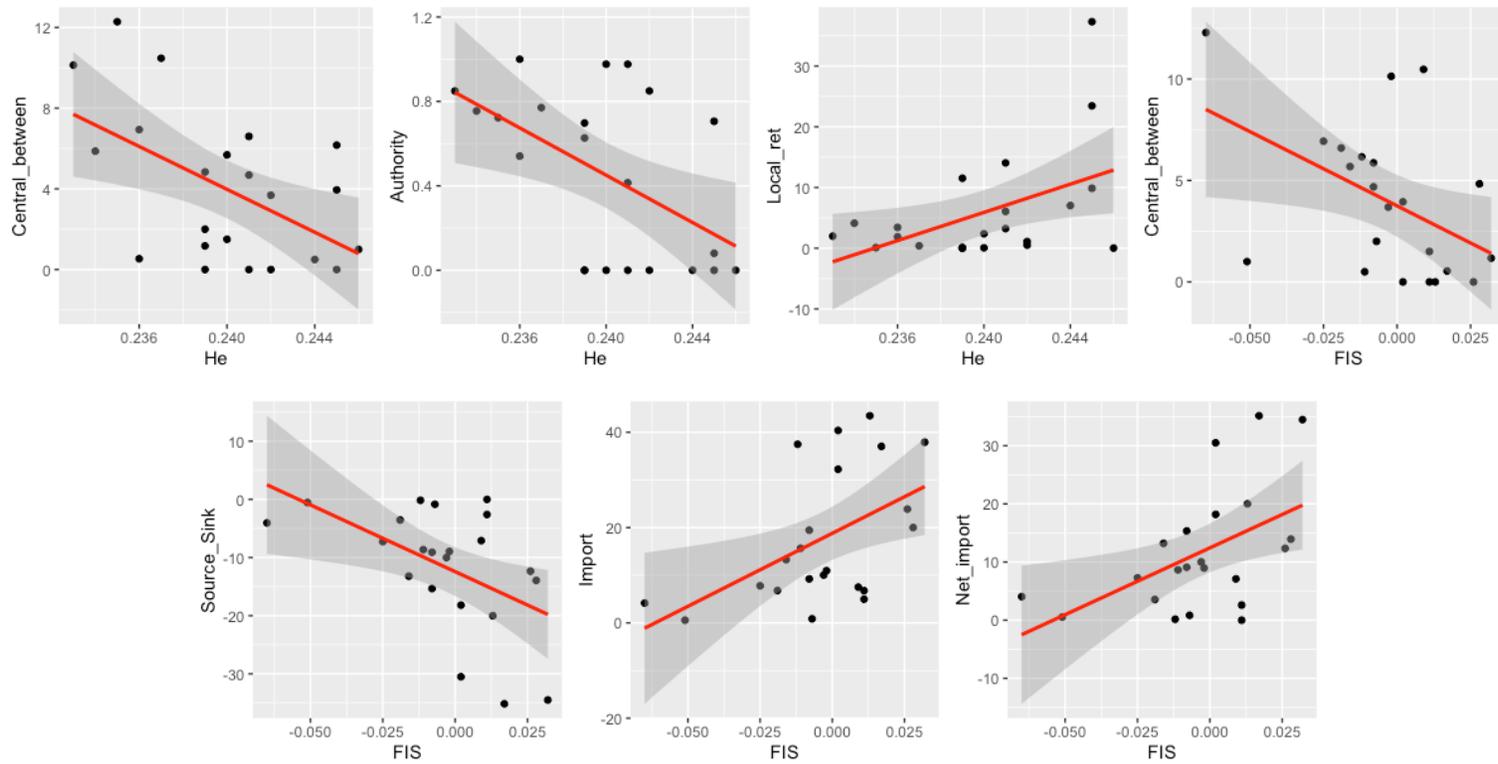


Figure continue in the next page...

Figure continue from the previous page...



**Figure 35.** Plots of the linear models between the network centrality metrics and the genomic diversity of the leopard grouper for the Summer season. The red line is the linear regression of the values.

The community beta diversity (Jaccard index) presented a significant relationship with the logarithm of the Euclidean distance among 23 studied localities in the GC when evaluating with Mantel test and MRM ( $r$  value = 0.541,  $p$  = 0.00;  $R^2$  = 0.293,  $p$  = 0.00). When relating beta diversity from each GC region, to the Euclidean distance, North ( $r$  = -0.354,  $p$  = 0.835;  $R^2$  = 0.125,  $p$  = 0.375) didn't showed significant results, but Central did ( $r$  = 0.371,  $p$  = 0.006;  $R^2$  = 0.138,  $p$  = 0.011) presented significant correlations (**Table 15** and **Figure 37**).

Mantel tests ( $r$  = 0.697,  $p$  = 0.000) and MRM ( $R^2$  = 0.486,  $p$  = 0.00) between the beta diversity and the environmental distance including the first PC (PC1:PC3) among 23 studied localities in the GC, presented significant results when assessing for an IBE pattern evidencing that the environmental differences along the GC explain the 69% of the species turnover. When comparing data from regions, the Northern region didn't showed a significant relationship ( $r$  = -0.282,  $p$  = 0.813);  $R^2$  = 0.079,  $p$  = 0.0374), but the Central region did ( $r$  = 0.484,  $p$  = 0.013;  $R^2$  = 0.233,  $p$  = 0.011) (**Table 15** and **Figure 37**).

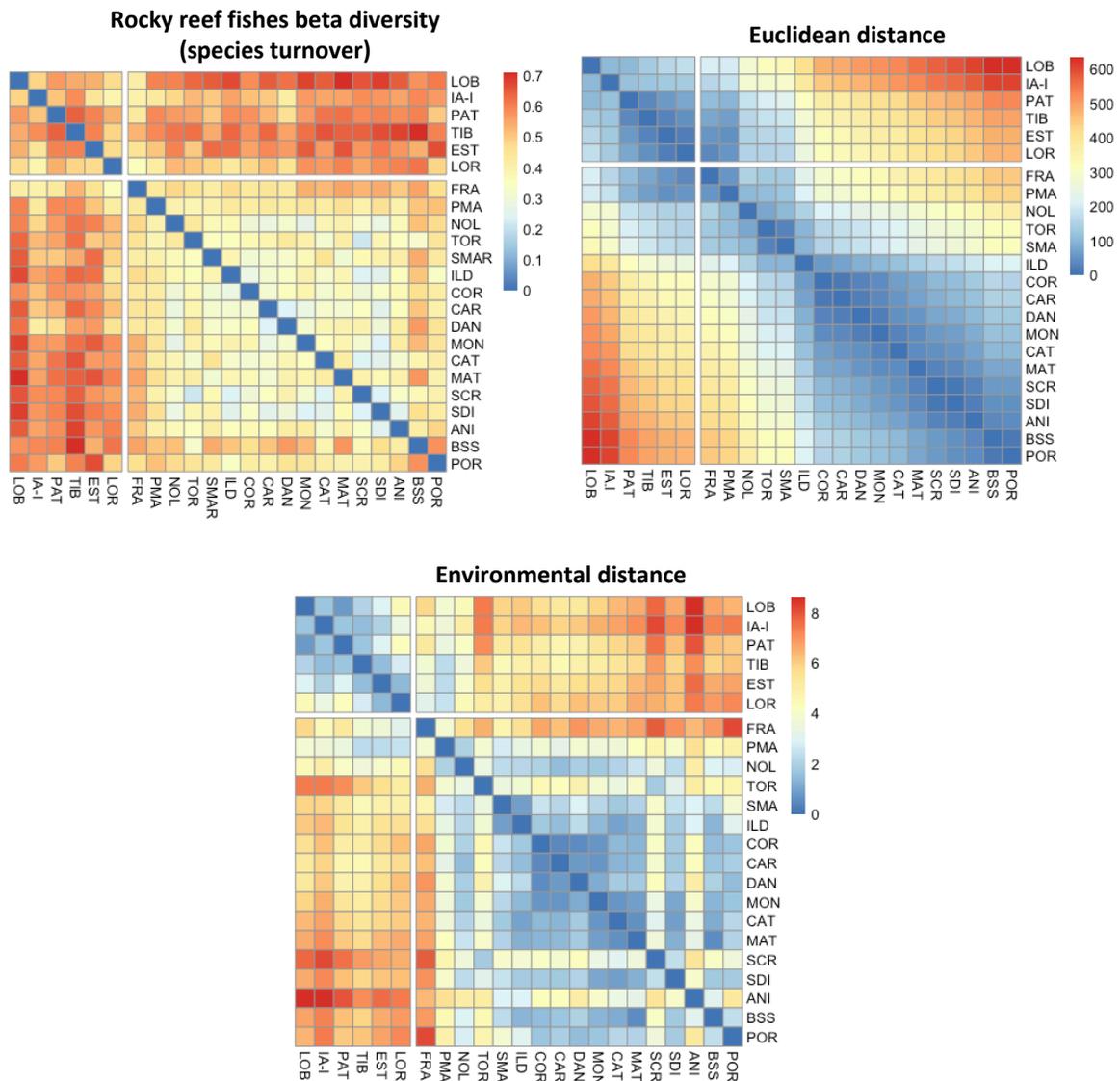
**Table 15.** Results from the Mantel tests applied to rocky reef fishes communities comparing effects of logarithm of Euclidean distance and the environmental distance, into the beta diversity (species turnover, Jaccard). Results are presented for all the Gulf of California and per region (North and Central). Two statistics that evaluate the correlation between matrices are presented: Mantel  $r$  and  $p$ -value of the Mantel test, and the Multiple regression of the distance matrices (MRM)  $R^2$  and  $p$ -value of the regression. Red letters represents statistically significant tests.

Seascape predictor	All the Gulf (BOR-VEN)		North (BOR-LOR_S)		Central (PMA-ANI)	
	Mantel $r$ ( $p$ -value)	MRM $R^2$ ( $p$ -value)	Mantel $r$ ( $p$ -value)	MRM $R^2$ ( $p$ -value)	Mantel $r$ ( $p$ -value)	MRM $R^2$ ( $p$ -value)
Log Euclidean	0.541 (0.00)	0.293 (0.00)	-0.354 (0.835)	0.125 (0.375)	0.371 (0.006)	0.138 (0.011)
Environmental PC1:PC3	0.697 (0.00)	0.486 (0.00)	-0.282 (0.813)	0.079 (0.374)	0.483 (0.013)	0.233 (0.011)

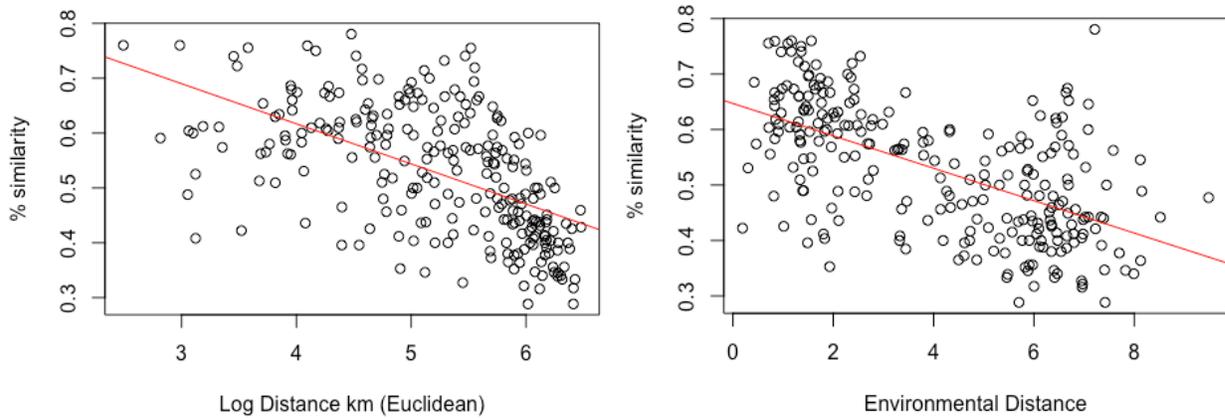
**Table 16.** Results of the Mantel tests and MRM evaluated on the rocky reef fishes communities comparing effects of resistance distances (from 23x23 polygon probability matrix per season) into the beta diversity (species turnover, Jaccard) in the All the Gulf, North region and Central region. Red letters represents statistically significant tests.

Season	All the Gulf (BOR-VEN)		North (BOR-LOR_S)		Central (PMA-ANI)	
	Mantel $ade4$ ( $p$ -value)	MRM $R^2$ ( $p$ -value)	Mantel $ade4$ ( $p$ -value)	MRM $R^2$ ( $p$ -value)	Mantel $ade4$ ( $p$ -value)	MRM $R^2$ ( $p$ -value)
Winter	0.428	0.057 (0.001)	0.771	0.068 (0.404)	0.752	0.004 (0.482)
Spring	0.814	0.019 (0.347)	0.035	0.418 (0.056)	0.487	0.000 (0.965)
Summer	0.282	0.001 (0.577)	0.678	0.024 (0.566)	0.025	0.047 (0.032)
Fall	0.061	0.086 (0.003)	0.323	0.022 (0.632)	0.855	0.021 (0.302)
Year	0.690	0.002 (0.603)	0.169	0.077 (0.302)	0.132	0.055 (0.137)

Mantel tests between beta diversity and resistance distance matrices evaluating all the GC localities showed statistically significant results for Winter seasons ( $p = 0.428$ ;  $R^2 = 0.057$ ,  $p = 0.001$ ), and for Fall season ( $p = 0.061$ ;  $R^2 = 0.061$ ,  $p = 0.006$ ). For the Northern region, beta-diversity showed a significant relationship with the all the Spring resistance distance ( $p = 0.035$ ;  $R^2 = 0.418$ ,  $p = 0.056$ ). The Central region, beta diversity presented a significant relationship with the Summer resistance distance ( $p = 0.025$ ;  $R^2 = 0.047$ ,  $p = 0.032$ ). These results indicate that the IBR hypothesis cannot be rejected for the rocky reef fishes of the GC (Table 16).



**Figure 36.** Rocky reef-fishes distance matrices of the A) Species turnover (Beta diversity, Jaccard dissimilarity index), and Euclidean (km) and Environmental distances among the 23 localities in the GC.



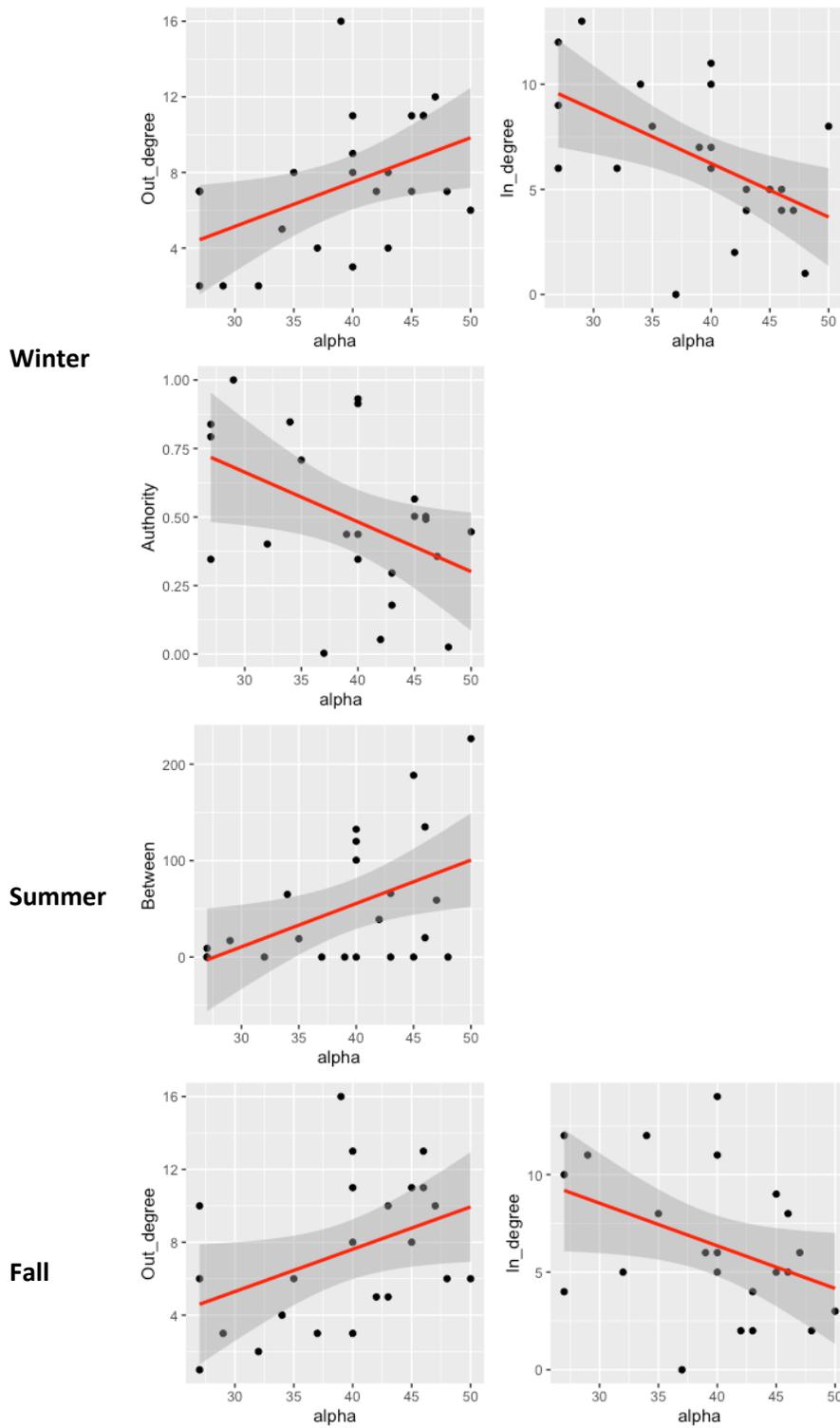
**Figure 37.** Patterns of distance-decay of similarity resulting from the relationship between the communities (Jaccard index) and Euclidean and environmental distance matrices. Significance values of correlations are  $p < 0.01$ .

**Table 17.** Linear models of network metrics and community alpha diversity metrics for Winter, Spring, Summer and Fall seasons. Adjusted  $R^2$  and p-value. Red letters represent statistically significant tests.

Network metric	Community			
	Winter $R^2$ (p-value)	Spring $R^2$ (p-value)	Summer $R^2$ (p-value)	Fall $R^2$ (p-value)
<b>Out degree</b>	<b>0.183 (0.023)</b>	-0.039 (0.686)	0.033 (0.197)	<b>0.139 (0.044)</b>
<b>In degree</b>	<b>0.254 (0.008)</b>	-0.046 (0.898)	-0.028 (0.534)	<b>0.126 (0.053)</b>
<b>Degree</b>	-0.046 (0.890)	-0.045 (0.834)	-0.038 (0.674)	-0.046 (0.905)
<b>Eigenvector centrality</b>	-0.044 (0.793)	-0.043 (0.786)	-0.023 (0.493)	-0.044 (0.794)
<b>Between</b>	-0.045 (0.824)	-0.042 (0.746)	<b>0.187 (0.022)</b>	0.016 (0.256)
<b>Central between</b>	0.241 (0.242)	-0.045 (0.856)	0.065 (0.126)	0.092 (0.086)
<b>Hub</b>	0.120 (0.058)	-0.046 (0.887)	0.041 (0.177)	0.077 (0.106)
<b>Authority</b>	<b>0.168 (0.029)</b>	-0.041 (0.887)	-0.003 (0.345)	0.036 (0.191)
<b>Local ret</b>	-0.045 (0.850)	0.020 (0.513)	0.098 (0.144)	0.097 (0.146)
<b>Source-Sink</b>	-0.001 (0.331)	0.005 (0.726)	0.005 (0.738)	0.049 (0.309)
<b>Export</b>	0.033 (0.406)	0.002 (0.824)	0.051 (0.297)	0.149 (0.068)
<b>Import</b>	0.000 (0.889)	0.000 (0.970)	0.081 (0.188)	0.055 (0.280)
<b>Net export</b>	0.075 (0.203)	0.002 (0.813)	0.001 (0.853)	0.095 (0.151)
<b>Net import</b>	0.000 (0.993)	0.050 (0.304)	0.018 (0.533)	0.001 (0.843)

Linear models among the network metrics of the Winter, Spring, Summer, and Fall community hypothetical reproductive seasons and the alpha diversity indicate a significant positive relationship between out-degree with the alpha diversity for Winter and Fall seasons. The in degree network metric showed a negative significant relationship with the alpha diversity in Winter and Fall. The authority metric also showed a negative significant relationship during Winter, and betweenness a positive relationship with alpha diversity during Summer (**Table 17** and **Figure 38**).

### Rocky reef fishes linear models of the alpha diversity and the network centrality metrics



**Figure 38.** Plots of the linear models between the network centrality metrics and the alpha diversity of the rocky reef fish communities for the Winter, Spring and Fall seasons. The red line is the linear regression of the values.

## Chapter. 4 Discussion

---

In the present thesis, an integrative approach including population genomics, eDNA metabarcoding, community ecology, oceanographic modeling, and network analysis was applied to evaluate functional and structural connectivity in the seascape of the GC. The independent estimation of functional connectivity within a species (*Mycteroperca rosacea*) and in rocky reef fish communities and the structural connectivity permitted contrasting hypotheses focused on identifying the abiotic components that allow metapopulation and metacommunity connectivity and diversity using a network approach at link and node levels. In the following sections, I'll discuss the main findings.

### **4.1 High levels of functional connectivity in the leopard grouper local populations are related to the Euclidean, environmental, and resistance distances at different spatial scales**

The population genomic characterization of the leopard grouper (*Mycteroperca rosacea*) evidenced no-significant population genomic structure among the 22 studied localities evaluated with 985 neutral SNPs markers. Also, moderate levels of genomic diversity with low variation and no significant differences among the GC regions. Furthermore, although there are no genomic differences, the demographic connectivity of the species reveals high levels of larval dispersal among the studied localities. These results indicate that the leopard grouper presents one connected metapopulation across the sampled area, conformed by local populations highly connected by larval dispersal.

Marine fish populations are often very large with high dispersal potential due to their planktonic larval stage (e.g., Hauser and Carvalho, 2008). When populations are well-connected, wide-scale genetic homogeneity is expected (Lowe and Allendorf, 2010; Waples et al., 2006), and high rates of gene flow coincide with low or nonexistent genetic structure as measured by traditional F-statistics (Crandall et al., 2019; D'Aloia et al., 2020; Junge et al., 2019; Riginos and Victor, 2001; Riginos and Liggins, 2013). It is relevant to mention that, besides being the result in species with high vagility, this non-significant population genomic structure is partially due to the non-linear relationship between  $F_{ST}$  and gene flow. This is because flows greater than  $\sim 10$  migrants/generation cannot be statistically distinguished from  $F_{ST} = 0$ , therefore, being unable to reject the null hypothesis that sampled individuals are part of a single,

randomly mating population (i.e., panmictic). This can happen even when sampling localities are separated by hundreds of kilometers (Crandall et al., 2019; Hauser and Carvalho, 2008).

Recently, Munguia-Vega et al. (2022) showed that the leopard grouper presents strong connectivity between Bahía de Kino and Puerto Libertad (in the Midriff Island Region) evaluated with microsatellites, even when they are geographically distant (150 km apart). Besides, the network analysis based on larval dispersal showed that the Bahía de Kino fishing zones are a source of larvae for Puerto Libertad due to oceanic currents' cyclonic circulation during the species' spawning season in the study region. Also, the analysis of migrants suggested that larvae, juvenile, and adult fish from the fishing zones of Puerto Libertad migrate to Bahía de Kino fishing grounds. Extrapolating these results to our study area allows us to infer that few generations of the leopard grouper are needed to connect distant local populations across the GC.

Additionally, previous research showed that leopard groupers' mitochondrial lineages had diverged since the last glacial maximum (LGM), when there was a dramatic drop in sea levels of 100–150 m, and revealed patterns of population expansion in some of the northernmost local populations, particularly from sites located on the large extended shelf that would have been exposed in the upper GC during the LGM, in contrast to deep-water areas that would have remained demographically stable in the long-term despite sea-level fluctuations. Evidence of population expansion in leopard grouper subpopulations is similar to patterns observed for other marine invertebrates and fishes occupying shallow waters in the GC (Pfeiler et al., 2005, 2008). In this sense, the evolutionary history of the leopard grouper in terms of a population expansion from deep island refuges following rapid population expansion, also determines its shallow genomic differentiation.

On the other hand, Jackson et al. (2015) evidenced that microsatellite pairwise  $F_{ST}$  estimates in the leopard grouper did not show clear geographic trends (only 6 out of 210  $F_{ST}$  comparisons were statistically significant) and a non-significant IBD when evaluating 21 localities across the GC (from Puerto Peñasco to Isla Cerralvo). This low genetic differentiation mostly concurs with what I found using SNP- markers in which zero of 242  $F_{ST}$  comparisons were statistically significant. Further, this pattern also happens in other teleost fishes with restricted geographical distribution, but high vagility (to the Tropical Eastern Pacific or the GC), such as *Epinephelus labriformis* (Craig et al., 2006), *Albula* sp. (Pfeiler et al., 2008), and *Hyporthodus acanthistius* (Beldade et al., 2014) and has been attributed to the biological aspects of these species such as their dispersal ability (i.e., PLD, spawning aggregations) and the complex oceanographic

characteristics that promote organismal dispersal, among other causes. Also, despite the long-standing hypothesis and, in some cases, empirical evidence that biogeographical breaks (North-Central-South in the GC) drive population differentiation (Brusca et al., 2005), it has been shown that, in certain species, some of these breaks do not have a barrier effect. Thus the role of commonly recognized biogeographical barriers in shaping population differentiation and, ultimately, driving speciation appears not to be occurring in the studied species (Craig et al., 2006).

Even minor genomic differences can expose biologically meaningful information when seascape predictors are included in the seascape genomic analyses. When performing Mantel tests and MRM between the genomic differentiation and the Euclidean distance among the 21 leopard grouper local populations, significant results were obtained, accepting the hypothesis of an IBD pattern at the GC extension and indicating that the geographic distance explains ~1.5% of the genomic differentiation among local populations. Although significant, the low predictive potential of the geographical distance at the GC extension into the leopard grouper's genomic differentiation highlights the species' vast dispersal potential throughout generations (900 km is the maximum distance between sites in all the studied areas) due to larval dispersal and the availability of finding suitable habitat patches across the GC. When performing these analyses within regions, neither North nor Central presented significant correlations with the Euclidean distance (max. distances 581 and 451 km, respectively), rejecting the null hypothesis of IBD. Yet, when excluding two bordering localities between the North and Central regions (SRO to VEN, max. distance 318 km), the genomic differentiation of the leopard grouper presents a significant IBD pattern in which the Euclidean distance among sites explains 23% of the genomic differentiation. In addition, the maximum correlation distance (results not shown) between the genomic differentiation and the geographical distance is ~350 km, which corresponds to the distance at which pairwise genetic distances are no longer significantly correlated or, in other words, the distance at which individuals have more similar allelic frequencies among them than with the total. This scale might inform about the genetic neighborhood size of the species (Sewall Wright, 1946). The higher predictive power of the Euclidean distance in the lower extension of the Central region can also be related to the seascape spatial configuration and the ocean currents patterns, i.e., the central SRO-VEN zone is primarily a linear coast with N-S larval transport, causing the Euclidean distance to be the most important predictor for the genomic differentiation.

On the contrary, the Northern region is located on the top of the semi-closed Gulf. It presents non-linear coast spatial arrangement in which the studied localities are distributed near the peninsular or

mainland coasts and across the Midriff Islands. Furthermore, the numerical model and environmental analysis results evidenced that this region exhibits higher seasonal oceanographic connectivity (non-linear). All these factors make the Euclidean distance between sites of the Northern part a poor predictor of genomic differentiation. Nevertheless, the influence of the seascape spatial configuration on the population genomic connectivity of fish species of the GC must be further investigated.

The model of IBD indicates that gene flow is more common among close local populations and diminishes in distant local populations (Hellberg, 2006; Palumbi, 2003) and that genetic differences between local populations should accumulate if dispersal is geographically restricted. Therefore, in marine fishes, this relationship can be moderated by several factors that restrict gene flow: ocean circulation patterns, discontinuity of habitats, and duration of the pelagic larvae, among others. In this sense, low or null levels of genetic structure across large geographic areas are not necessarily incompatible with localized dispersal, i.e., in the ecological timescale, dispersal occurring during one or two generations may typically occur within tens of kilometers (Puebla et al., 2012). However, at evolutionary timescales, the genetic population differentiation is related to the effective dispersal integrated over extended periods and the gene flow increasing the distances in a stepping stone manner through generations (i.e., as in a metapopulation) (Hauser and Carvalho, 2008).

On the other hand, elucidating the environmental determinants of population structure in marine ecosystems is a worthy enterprise needed to answer important questions of relevance facing marine conservation and management (Benestan et al., 2016; Selkoe et al., 2008, 2016). When evaluating for an IBE pattern, Mantel and MRM tests revealed that the environmental differences along the GC explained 1.5% of the genomic differentiation among leopard grouper local populations in all the GC but not within regions. Nevertheless, the Euclidean and environmental distances are correlated ( $r = 0.9$ ,  $p < 0.001$ ). Therefore, the outcome or the relationship among the genomic differentiation of the species and the predictor distances was partialled-out (partial Mantel test) by controlling the effect of one matrix over the other (Balkenhol et al., 2009). After this correction, IBD ( $r = 0.102$ ,  $p=0.153$ ) and IBE ( $r = -0.043$ ,  $p=0.66$ ) patterns lost its statistical significance.

When evaluating the environmental similarity of the leopard grouper study localities using 14 variables using PCA, the first PC explained ~60% of the environmental variation among the 22 population sites, and the principal variables contributing were phosphates, dissolved oxygen, sea surface temperature, salinity, and silicates (mean annual values). Further, the PCA showed that the localities were

grouped in two clusters (Northern and Southern localities) according to their environmental characteristics. This is an important result because we can infer that the environmental distinction between the North and Central regions of the GC doesn't represent a dispersal barrier for the leopard grouper gene flow. Accordingly, the effect of this environmental arrangement on the population genomic structure of the adaptive SNP markers must be further evaluated concerning its potential impact on individual adaptation to particular environmental features.

On the other hand, White et al. (2010) highlighted the benefits of using oceanographic data to advance our ability to interpret species' population structure with pelagic larval stages and high gene flow. They showed that ocean currents better explained genetic patterns than geographic distance. Additionally, recent studies in marine populations of the GC have highlighted that the presence of low  $F_{ST}$  values and a metapopulation structure with moderate levels of asymmetric migration are not mutually exclusive scenarios (Reguera-Rouzaud et al., 2020). For example, non-significant  $F_{ST}$  values have been reported in metapopulation structure in an asymmetric gene flow in the source-sink dynamics (Cisneros-Mata et al., 2019). These departures from the assumptions of the island model (i.e., dispersal is equally likely within all subpopulations within a species) occur within marine populations to an extent determined by the biological attributes (e.g., reproductive season) and the oceanographic characteristics in which species and populations live (Hellberg, 2006).

In the absence of precise information about the reproductive season of the leopard grouper to evaluate the potential larval dispersal of the species, the observed genomic patterns were contrasted against 21 hypothetical reproductive seasons representing a metapopulation connected by larval dispersal and informed by the modeled ocean circulation (HAMSOM) to investigate the demographic connectivity of the leopard grouper. The probability matrices of potential larval connectivity allowed the estimation of the resistance distances (Dijkstra's algorithm), contrasted with the genomic differentiation matrix to test for IBR. This method of indirectly inferring representative reproductive periods based on the fit between seasonal modeled patterns of larval dispersal and empirical genetic data based on expectations from metapopulation theory has been used in other reef species of the GC where information about the reproductive timing and duration is conflicting or currently unavailable (Cisneros-Mata et al., 2019; Munguia-Vega et al., 2014; Munguia-Vega, Marinone, et al., 2018; Reguera-Rouzaud et al., 2020).

For the leopard grouper, the Mantel and MRM test results showed that resistance distances of the hypothetical reproductive seasons have no significant relationship with its genomic differentiation at all

the extent of the GC and in the Central region, only a significant relationship within the Northern region for the June-July and the April-September seasons explaining the 14.7% and 9.7% of the genomic variation within local Northern populations, respectively. This result suggests that a more extended reproductive season than previously reported can be occurring in the leopard grouper's Northern distribution. Thus, IBR is rejected for all the GC and the Central region but not for the Northern region. This concurs with the results of a previous study of the same species in the Northern region of GC. The use of an oceanographic model describing metapopulation dynamics of larval dispersal helped explain genetic differences between sites (Munguia-Vega et al., 2014).

Different explanations can be proposed to elucidate the absence of the relationship between the ocean currents and the genomic differentiation of the leopard grouper in all the GC. First, the numerical model estimates the larvae' potential dispersal but does not consider differential mortality, recruitment (Pineda et al., 2007), or other critical biological attributes of larvae (Cowen et al., 2007), which can reduce the amount of potential connectivity. Second, there is a temporal mismatch between the demographic and the molecular estimates because the numerical model uses oceanographic information collected in recent years. In contrast, genomic patterns arise from long-term processes (Riginos et al., 2011, 2019), i.e., population allele frequencies do not only reflect gene flow resulting from a dispersal event (from one generation to the next) but are the accumulated result of the dispersal over several generations, therefore are also molded by changes in population sizes, range expansions, colonization and so forth (Whitlock and McCauley, 1999). Lastly, spatially limited sampling effort can cause mismatches between connectivity estimates; e.g., in the present thesis, I extracted the information from the 59x59 polygons of the oceanographic model only for the 22 sampled polygons for populations, leaving a percentage of dispersal information out of the analysis. If the leopard grouper distribution is broader than the extent of the polygons included here, I am sub-sampling the patches occupied by a population, i.e., underestimating dispersal rates, and the inferred demographic connectivity can be biased.

Incongruent genomic and demographic connectivity patterns may also result from genomic differentiation indices. Regarding genomic approaches, it is essential to keep in mind, as explained before, that  $F_{ST}$  cannot be considered a direct estimate of the effective dispersal (Marko and Hart, 2011; Whitlock and McCauley, 1999). Gene flow and genetic drift interact as opposing forces, the former decreasing and the latter increasing genetic variability among populations (Slatkin, 1989). Hence, mismatches between demographic and genetic estimates of dispersal could partly be due to the contribution of genetic drift to

genetic differentiation, especially when the spatial extent of the genetic studies exceeds the maximal dispersal distance of surveyed species.

On the other hand, theory predicts that the distribution of genomic diversity in a seascape depends on the connectivity of the metapopulation and the dispersal of individuals between patches. Network analysis and linear model results showed a significant negative relationship between the in-degree, central betweenness, and authority with the allelic richness and the expected heterozygosity of local populations in both seasons, indicating that high levels of connections into a node (larval influx) decrease the genetic diversity. On the other hand, larvae export and source-sink presented a significant positive relationship with the allelic richness and the expected heterozygosity in both seasons, indicating that nodes that export few larvae and import much present low diversity. There was also a significant negative linear relationship with  $F_{IS}$  for the central betweenness metric in both seasons.

A previous study using this approach in *Lutjanus peru* found important relationships between the source-sink role and  $F_{IS}$  and the effective number of alleles with the in-degree value of the nodes. Authors described that net sinks (that received more larvae than they exported) showed higher levels of relatedness than net sources (Munguia-Vega et al., 2018). This result suggests that local population dynamics in strongly advective systems like the GC could be more influenced by the role of the site as a net sink or source within the metapopulation (i.e., external influences) and less by the process of local larval retention itself (Munguia-Vega et al., 2018). Our results support this idea since the source-sink dynamics were significantly correlated with the allelic richness and expected heterozygosity values for the Spring season, whereas larval retention was not. Alternatively, a previous study using mitochondrial DNA found no relation between haplotypic or nucleotide diversity and network metrics (Munguia-Vega et al., 2014). Our results demonstrate that network centrality metrics reflect local leopard group populations' genomic diversity patterns. Nevertheless, this relationship must be further investigated.

## **4.2 Metacommunity characterization and the evaluation of connectivity and its seascape determinants**

I used simultaneous UVC and eDNA metabarcoding to characterize rocky reef fish communities in the GC and to evaluate their functional connectivity among 23 localities across the North and Central regions of the GC. Results derived from the use of two detection methods indicated that eDNA

metabarcoding detects biodiversity over a broader spatial scale, which depends on physical conditions responsible for its transport across the seascape, such as ocean currents (Yamamoto et al., 2017), i.e., the species assemblages detected with eDNA integrate additional species inhabiting adjacent marine habitats (e.g., pelagic, demersal, soft bottom), which may be sampled at any stage of their life cycle (i.e., larvae, juveniles, and adults). Furthermore, UVC and eDNA metabarcoding community data evaluated separately identified significant differences in fish community composition between North and Central GC regions. Besides, mean values and geographical patterns of alpha diversity of both UVC and eDNA data confirmed that the Central region between Loreto and La Paz localities is a hotspot of alpha diversity, in contrast to the lower alpha diversity observed in the Northern GC (Morzaria-Luna et al., 2018; Olivier et al., 2018).

The complementary use of UVC with eDNA metabarcoding surveys increased the number of species detected in a complex fish community such as the GC (Valdivia-Carrillo et al., 2021), i.e., a total of 191 species/OTUs were detected with both survey methods; of these, 13% were shared, 49% were identified only by eDNA, and 38% were observed only with UVC. This is a relevant result because species records in the GC are likely still lacking from many taxa (Brusca et al., 2005; Morzaria-Luna et al., 2018), e.g., before this study, only 5% of the ~800 teleost species from the GC possessed reference 12S rRNA gene DNA sequences in NCBI-GenBank. Considering this insufficient taxonomic coverage of the barcode used for the species of interest, I developed a custom reference database for 67 additional species, increasing the taxonomic coverage to 13%. Despite the modest size of this reference database, the use of both NCBI +custom reference databases almost tripled the fraction of taxonomically identifiable OTUs. Due to this previous information, I combined the species detections of both methods. I obtained a site x species matrix including only reef fish species (94 species in total) to estimate the beta diversity (species turnover) among the 23 sampling localities in the GC to have more complete information about the species composition of the studied rocky reefs.

Among the patterns that have been used to explain community composition, community similarity (i.e., beta diversity) has been widely documented (Cottenie, 2005; Diniz-Filho and Bini, 2011; Tuomisto and Ruokolainen, 2006). In particular, the decrease in community similarity with geographic distance has proven to be common to different groups of organisms in which closely-located communities are generally more similar in terms of species composition than those located further apart (Borthagaray et al., 2009; Leprieur et al., 2009; Maloney and Munguia, 2011; Moritz et al., 2013; Papadopoulou et al., 2011). This pattern is expected if dispersal is an essential limiting factor in structuring ecological communities. For the rocky reef fishes of the GC, when evaluating a distance decay of community similarity, the Mantel test and

MRM results showed that the beta diversity presented a significant relationship with the Euclidean distance among 23 studied localities in the GC. This denotes a distance decay in which the distance among the sampling sites explains 29% of the species turnover along the GC (max. distance between sites = 651 km).

On the contrary, the lower spatial extension of the Northern region results didn't show significant results when evaluating a distance decay with the Euclidean distance. Yet, the distance between similarity and geographical distance explains 14% of the species turnover for the Central region. This result is supported by previous research that showed a latitudinal variation in the community structure of reef fish along with the western GC, with one of the highest values in richness, diversity, taxonomic distinctness, and trophic level in this Central region (Santa Rosalía to La Paz), probably due to a combination of environmental conditions, a greater number of habitats, and more functional diversity of the assemblages (Fernández-Rivera Melo et al., 2018).

When evaluating for a distance decay in community similarity with the environmental distance including PC1:PC3 among 23 studied localities, the Mantel test presented significant results evidencing that the environmental differences explain the 69% of the species spatial turnover in all the study sites of the GC. The Central region also showed a significant environmental distance decay relationship explaining 48% of the species. Nevertheless, the Northern region did not follow this pattern. This decay within specific sub-scales demonstrates that complex community dynamics need to be investigated at a large regional scale and within specific environments to account for structural factors of community composition. Besides, I found that a higher number of variables are related to the community structure compared to population-level comparisons: phosphates, dissolved oxygen, salinity, silicates, and sea surface temperature (PC1); nitrates, primary productivity, bathymetry, and aspect East-West (PC2); and concavity, aspect North-South, bathymetry, aspect East-West and distance to shore (PC3).

As in the population level analyses, community localities' geographical and environmental distances are correlated ( $r = 0.842$ ,  $p = 0.000$ ). Therefore, Mantel tests had to be partialled-out. When performing a partial Mantel tests controlling for spatial ( $r = 0.377$ ,  $p=0.004$ ) or environmental ( $r = 0.193$ ,  $p=0.02$ ) factors, the distance decay pattern remains when including all the GC localities. Also, for the Central region Mantel tests controlling for spatial ( $r = 0.286$ ,  $p=0.073$ ) or environmental ( $r = 0.217$ ,  $p=0.048$ ) factor evidence that only the environmental variation remain as a factor determining species turnover in this GC region.

In this sense, it is known that the distribution of marine organisms is influenced by their physical environment (e.g., depth., temperature, currents, among others). The environmental analyses showed that the GC is a heterogeneous dynamic seascape with a continuously variable gradient of environmental features for the rocky reefs. The geographic proximity and physicochemical characteristics influence fish dispersal among habitat patches within the seascape. I identified that the mean annual environmental heterogeneity (0-5 m depth) varies differentially regarding the type of variable investigated: a significant latitudinal variation was found in the physicochemical environmental characteristics of the Gulf, and a non-latitudinal variation in geomorphological and location-related variables. This characteristic of non-randomly environmental variation and its determinants had been described before (Fernández-Rivera Melo et al., 2018; Salvador E. Lluch-Cota et al., 2007; Soto-Mardones et al., 1999).

Furthermore, when evaluating the multivariate environmental relationships among the community sites, I found two main clusters grouping Northern and Central localities. In a geographical context, these groups are delimited by the Midriff islands region, which includes the largest islands in the Gulf (i.e., Ángel de la Guarda and Tiburón), considered the central physiographic element in the GC and that separates the North and Central biogeographic regions of the GC. The boundary zone between the Northern and the Central GC regions is characterized by an abrupt change in the tidal range and tidal currents. This barrier also corresponds to the separation between North and Central clusters of local communities in the GC identified before.

On the other hand, the significant relationship of the environmental characteristics in structuring communities of the Central GC region summed to its high values of alpha diversity can be related to the fact that reef habitats with high structural complexity generally have greater species diversity compared with less complex environments, because habitats provide a greater spectrum of resources like food, shelter and reproduction sites (Sánchez-Caballero et al., 2017). Fernández-Rivera Melo et al. (2018) described a latitudinal variation of reef fish assemblages along with the western GC in which the central portion of the Gulf is characterized by tropical water and a higher number of habitats (rocky reefs, islands, mangrove areas, rhodolite beds, coralline algae patches, submarine mountains, and rocky walls) that are considered adequate habitats for reef fish over different stages of life (Aburto-Oropeza and Balart, 2001); whereas in the north, the temperate water and the occurrence of reef corals and mangroves are practically absent because of the cold temperature in Winter (Brusca et al., 2005). Therefore, studying seascapes with varying habitat configurations and distances between habitat patches can increase our understanding of potential thresholds of connectivity (Berkström et al., 2020).

In marine systems, dispersal is likely to be influenced not only by geographic or environmental distances but also by its direction influenced by oceanic currents and, to some lesser extent, by larval behavior (Cowen et al., 2007). In these systems, the directionality of dispersal influences connectivity between communities, which prevents the latter from fully overlapping with environmental gradients. In such cases, comparing changes in community composition with geographic distance does not account for directional dispersal (Moritz et al., 2013). Instead, a connectivity measure based on oceanographic properties should be more relevant to test the relative effects of dispersal and environmental filtering in shaping community structure and composition. In the present thesis, I proposed an approximation to evaluate the contribution of reproductive fish assemblages which reproduces during each station (Winter, Spring, Summer, and Fall) to the metacommunity connectivity. As a first-order analysis, significant Mantel tests between the beta diversity and the resistance distance matrices obtained from the four hypothetical reproductive seasons of reef fishes indicate a pattern of the resistance distance decay of community similarity. In all the studied extensions of the GC, Winter, and Fall, the resistance distance explains 2.4% and 2.9% of the species' spatial turnover. For the Northern region, beta diversity showed a significant relationship with the Spring resistance distance, indicating that the oceanographic connectivity explains 6.5% of the species turnover. The beta diversity presented a significant relationship with the Summer resistance distance in the Central region, indicating that the oceanographic connectivity in this season explains 2.1% of the species turnover. Overall, these results support the idea that the metacommunity structure across the GC and within its regions is partly a result of the seasonal oceanographic connectivity patterns (i.e., resistance distance decay). Numerical simulations have been applied to different areas of the GC and have helped explain connectivity patterns for several local populations (Cisneros-Mata et al., 2019; Marinone et al., 2008; Adrian Munguia-Vega et al., 2014; Adrian Munguia-Vega, Marinone, et al., 2018; Peguero-Icaza et al., 2008; Reguera-Rouzaud et al., 2020; Soria et al., 2012). Here, connectivity was determined using a numerical model to study dispersal at the metacommunity regional scale of the GC (Marinone, 2003). Because the coupling with a species-specific larval behavior may not explain the different behaviors and PLDs of all the fish species present in the rocky reefs of the GC, I estimated the potential larval connectivity using a four-week PLD and the liberation of 4000 particles per polygon as in the population analyses. Incorporating individual biological differences of each species larvae in the model is almost an impossible task, i.e., to account for these differences implies incorporating proportions of larvae for each species, which would require running separate simulations, species by species. As a result, the approximation applied in the present study implied underestimating the potential demographic connectivity among local communities. Besides, Munguia-Vega et al. (2018) predicted that the average distance traveled by passive larvae is a function of their PLD varying from >20 to 80 km (PLD 7 days) up to

>200 km (PLD 28–60 days). Therefore, using a unique PLD of 28 days for all the species implied an overestimated dispersal distance for some species with shorter PLDs. A solution for these pitfalls could be using species-specific data from umbrella species that represent the variety of dispersal abilities within the community (Calabrese and Fagan, 2004a).

Previous research has shown that the GC fishes are strongly influenced by asymmetric dispersal and water flows, creating highly complex structures favoring retention in some areas. In such complex systems, connectivity between local communities will thus not be a linear function of geographic distance but, somewhat, will be influenced by the characteristics of the medium transporting organisms (Cowen et al., 2000). Although the importance of local biotic and abiotic conditions in the maintenance of marine organisms has also been demonstrated, quantifying the relative effects of the oceanographic connectivity versus the environmental filtering in fishes has rarely been achieved. Here, I accounted for the oceanographic, environmental, and spatial distances to better understand the spatial structure of a marine rocky reef fish metacommunity across the GC, where the three contribute to shaping the local community species composition.

Finally, the most critical connection between metacommunity ecology and seascape ecology is the equilibrium theory of island biogeography, emphasizing the essential roles of habitat size and habitat isolation (and thus dispersal limitation) in maintaining biodiversity in island-like habitats (Chase et al., 2020). According to the island biogeography, species richness should be negatively correlated with the degree of isolation of the habitat. Many studies have found adverse effects of patch isolation on species richness. In this sense, centrality (i.e., connectivity) measures are relevant since they are typically used as an indirect method to assess dispersal at the metacommunity level (Jacobson and Peres-Neto, 2010). Moreover, the results presented here showed a relationship between the centrality metrics (connectivity) and the observed species in the local communities. In other words, relationships between linear models among the network metrics of the Winter, Spring, Summer, and Fall community hypothetical reproductive seasons and the alpha diversity indicate a significant positive relationship between out-degree and degree metrics with the alpha diversity for Winter and Fall seasons. The in-degree network metric showed a significant negative association with the alpha diversity in Winter and Fall but a contrary significant relationship during Spring. Hub and authority metrics also showed a positive and negative significant relationship, respectively, during Winter.

## Chapter. 5      Conclusions

---

At the link level:

The leopard grouper metapopulation presents high genomic connectivity and homogeneity in its genomic diversity patterns with no differences among the GC regions. The biological characteristics of the species (i.e., PLD) and the oceanographic connectivity of the GC allow it to disperse across the GC in a few generations. The outcome of its high potential larval dispersal is observed in the low genomic differentiation among the local populations within the species metapopulation.

For the leopard grouper, the seascape structural connectivity in terms of the geographical, environmental, and resistance distances influences the species differentially at different spatial scales, depending on the region of the Gulf, its spatial configuration, and dominant abiotic characteristics, i.e., in the Central GC region, an IBD pattern explains 23% of the genomic differentiation among local populations. In the Northern, an IBR pattern describes 14.7% (June-July season) and 9.7% (April-September season) of the genomic differentiation among local populations. The environmental heterogeneity did not represent a barrier to dispersal and gene flow for the species.

In rocky reef fishes, local communities show a North and Central ecological differentiation in terms of species composition in which the latter present higher values of alpha diversity. The division between the two local community clusters is located southern Midriff Islands and coincides with the limit between the Northern and Central biogeographical regions.

For the rocky reef fish communities, the seascape structural connectivity presents barriers to dispersal in terms of geographical, environmental, and resistance distances. It influences the communities differentially at different spatial scales. A spatial distance decay was a critical determinant of the species beta diversity at all the GC extent (explaining 19.3% of the species turnover) and in the Central region (explaining 21.7% of the species turnover). An environmental distance decay was also an essential determinant for the species turnover in all the GC (37.7%). Additionally, the beta diversity of the rocky reef fishes metacommunity showed a resistance distance decay of community similarity, indicating that the oceanographic connectivity and early life-history dispersal of fishes (demographic connectivity) were significantly related to the differences in the species composition of the local communities. In all the GC, the Winter and Fall oceanographic connectivity explain 2.4% and 2.9% of the species turnover; in the

Northern region, the Spring season explains 6.5% of the species turnover within the region; and in the Central region, the Summer season explains 2.1% of the species turnover.

At node level:

Higher centrality metrics are significantly related to the leopard grouper's genomic diversity and the alpha diversity of the rocky reefs fish communities. This relationship has to be further investigated.

## Cited literature

---

- Aburto-Oropeza, O. and Balart, E. F. 2001. Community Structure of Reef Fish in Several Habitats of a Rocky Reef in the Gulf of California. *Marine Ecology*, 22(4), pp. 283–305. doi:10.1046/j.1439-0485.2001.t01-1-01747.x
- Aburto-Oropeza, O., Dominguez-Guerrero, I., Cota-Nieto, J. and Plomozo-Lugo, T. 2009. Recruitment and ontogenetic habitat shifts of the yellow snapper (*Lutjanus argentiventris*) in the Gulf of California. *Marine Biology*, 156(12), pp. 2461–2472. doi:10.1007/s00227-009-1271-5
- Aburto-Oropeza, O., Sala, E., Paredes, G. A., Mendoza, A., Ballesteros, E., Gustavo, P., Abraham, M. and Enric, B. 2007. Predictability of reef fish recruitment in a highly variable nursery habitat. *Ecology*, 88(9), pp. 2220–2228. doi:10.1890/06-0857.1
- Aceves-Medina, G., Jiménez-Rosenberg, S. P. A., Hinojosa-Medina, A., Funes-Rodríguez, R., Saldierna-Martínez, R. J. and Smith, P. E. 2004. Fish larvae assemblages in the Gulf of California. *Journal of Fish Biology*, 65(3), pp. 832–847. doi:10.1111/j.0022-1112.2004.00490.x
- Aceves-Medina, G., Jimenez Rosenberg, P., Hinojosa-Medina, A., Funes, R., Saldierna, R. J., Lluch-Belda, D., Smith, P. E. and Watson, W. 2003. Fish larvae from the Gulf of California. *Scientia Marina*, 67(1), pp. 1–11.
- Ahern, A. L. M., Gómez-Gutiérrez, J., Aburto-oropeza, O., Saldierna-martínez, R. J., Johnson, A. F., Harada, A. E., Erisman, B., Arvizú, D. I. C. and Burton, R. S. 2018. DNA sequencing of fish eggs and larvae reveals high species diversity and seasonal changes in spawning activity in the southeastern Gulf of California Table S1 . List of 157 Operational Taxonomic Units (OTU) found from ichthyoplankton collections taken . 592, pp. 1–6.
- Alberdi, A., Aizpurua, O., Gilbert, M. T. P. and Bohmann, K. 2018. Scrutinizing key steps for reliable metabarcoding of environmental samples. *Methods in Ecology and Evolution*, 2018(May 2017), pp. 134–147. doi:10.1111/2041-210X.12849
- Alexander, D. H., Novembre, J. and Lange, K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), pp. 1655–1664. doi:10.1101/gr.094052.109
- Alexander, D. H., Shringarpure, S. S., Novembre, J. and Lange, K. 2015. *Admixture 1.3* (p. 12).
- Aljanabi, S. M. and Martinez, I. 1997. Universal and rapid salt-extraction of high quality genomic DNA for PCR-based techniques. *Nucleic Acids Research*, 25(22), pp. 4692–4693. doi:10.1093/nar/25.22.4692
- Almany, G. R., Planes, S., Thorrold, S. R., Berumen, M. L., Bode, M., Saenz-Agudelo, P., Bonin, M. C., Frisch, A. J., Harrison, H. B., Messmer, V., Nanninga, G. B., Priest, M. A., Srinivasan, M., Sinclair-Taylor, T., Williamson, D. H. and Jones, G. P. 2017. Larval fish dispersal in a coral-reef seascape. *Nature Ecology and Evolution*, 1(6). doi:10.1038/s41559-017-0148

- Álvarez-Romero, J. G., Munguía-Vega, A., Beger, M., del Mar Mancha-Cisneros, M., Suárez-Castillo, A. N., Gurney, G. G., Pressey, R. L., Gerber, L. R., Morzaria-Luna, H. N., Reyes-Bonilla, H., Adams, V. M., Kolb, M., Graham, E. M., VanDerWal, J., Castillo-López, A., Hinojosa-Arango, G., Petatán-Ramírez, D., Moreno-Baez, M., Godínez-Reyes, C. R. and Torre, J. 2018. Designing connected marine reserves in the face of global warming. *Global Change Biology*, 24(2), pp. e671–e691. doi:10.1111/gcb.13989
- Anadón, J. D., D'Agrosa, C., Gondor, A. and Gerber, L. R. 2011. Quantifying the spatial ecology of wide-ranging marine species in the Gulf of California: Implications for marine conservation planning. *PLoS ONE*, 6(12), pp. 1–10. doi:10.1371/journal.pone.0028400
- Andersen, J. C., Oboyski, P., Davies, N., Charlat, S., Ewing, C., Meyer, C., Krehenwinkel, H., Lim, J. Y., Suzuki, N., Ramage, T., Gillespie, R. G. and Roderick, G. K. 2019. Categorization of species as native or non-native using DNA sequence signatures without a complete reference library. *Ecological Applications*, 29(5), p. e01914. doi:10.1002/eap.1914
- Anderson, M. J., Tolimieri, N. and Millar, R. B. 2013. Beta diversity of demersal fish assemblages in the North-Eastern Pacific: Interactions of Latitude and Depth. *PLoS ONE*, 8(3), pp. 1–15. doi:10.1371/journal.pone.0057918
- Andrello, M., Mouillot, D., Somot, S., Thuiller, W. and Manel, S. 2015. Additive effects of climate change on connectivity between marine protected areas and larval supply to fished areas. *Diversity and Distributions*, 21(2), pp. 139–150. doi:10.1111/ddi.12250
- Andruszkiewicz, E. A., Koseff, J. R., Fringer, O. B., Ouellette, N. T., Lowe, A. B., Edwards, C. A. and Boehm, A. B. 2019. Modeling environmental DNA transport in the coastal ocean using lagrangian particle tracking. *Frontiers in Marine Science*, 6(August), pp. 1–14. doi:10.3389/fmars.2019.00477
- Argote, M. L., Amador, A. and Lavín, M. F. 1995. Tidal dissipation and stratification in the Gulf of California. *Journal of Geophysical Research*, 100(95), pp. 16,103-16,118.
- Avendaño-Ibarra, R., Godínez-Domínguez, E., Aceves-Medina, G., González-Rodríguez, E. and Trasviña, A. 2013. Fish Larvae Response to Biophysical Changes in the Gulf of California, Mexico (Winter-Summer). *Journal of Marine Biology*, 2013, pp. 1–17. doi:10.1155/2013/176760
- Ayala-Bocos, A., Reyes-Bonilla, H., Calderón-Aguilera, L. E., Herrero-Perezrul, M. D. and González-Espinosa, P. C. 2015. Proyección de cambios en la temperatura superficial del mar del Golfo de California y efectos sobre la abundancia y distribución de especies arrecifales. *Revista Ciencias Marinas y Costeras*, 8(1), pp. 29–40. doi:10.15359/revmar.8-1.2
- Bachelor, N. M., Gerald, N. R., Burton, M. L., Muñoz, R. C. and Kellison, G. T. 2017. Comparing relative abundance, lengths, and habitat of temperate reef fishes using simultaneous underwater visual census, video, and trap sampling. *Marine Ecology Progress Series*, 574, pp. 141–155. doi:10.3354/meps12172

- Backhaus, J. O. 1985. A Three-Dimensional Model for the Simulation of Shelf Sea Dynamics. *Deutsche Hydrografische Zeitschrift*, 38(4), pp. 165–187.
- Baguette, M., Blanchet, S., Legrand, D., Stevens, V. M. and Turlure, C. 2013. Individual dispersal, landscape connectivity and ecological networks. *Biological Reviews*, 88(2), pp. 310–326. doi:10.1111/brv.12000
- Bailey, L. L., Mackenzie, D. I. and Nichols, J. D. 2014. Advances and applications of occupancy models. *Methods in Ecology and Evolution*, 5(12), pp. 1269–1279. doi:10.1111/2041-210X.12100
- Bakun, A. 2006. Fronts and eddies as key structures in the habitat of marine fish larvae: opportunity, adaptive response and competitive advantage. *Scientia Marina*, 70S2(October), pp. 105–122. doi:10.3989/scimar.2006.70s2105
- Bálint, M., Bahram, M., Eren, A. M., Faust, K., Fuhrman, J. A., Lindahl, B., O’Hara, R. B., öpik, M., Sogin, M. L., Unterseher, M. and Tedersoo, L. 2016. Millions of reads, thousands of taxa: Microbial community structure and associations analyzed via marker genes. *FEMS Microbiology Reviews*, 40(5), pp. 686–700. doi:10.1093/femsre/fuw017
- Balkenhol, N., Waits, L. P. and Dezzani, R. J. 2009. Statistical approaches in landscape genetics: An evaluation of methods for linking landscape and genetic data. *Ecography*, 32(5), pp. 818–830. doi:10.1111/j.1600-0587.2009.05807.x
- Bandelj, V., Solidoro, C., Laurent, C., Querin, S., Kaleb, S., Gianni, F. and Falace, A. 2020. Cross-scale connectivity of macrobenthic communities in a patchy network of habitats: the mesophotic biogenic habitats of the Northern Adriatic Sea. *Estuarine, Coastal and Shelf Science*, 245. doi:10.1016/j.ecss.2020.106978
- Barbosa-Ortega, W., Rivera-Camacho, A., Avila-Poveda, O., Ceballos-Vázquez, B. and Arellano-Martínez, M. 2015. *Biología reproductiva de Lutjanus peru y Lutjanus argentiventris (Perciformes: Lutjanidae) en la costa sur-occidental del Golfo de California*. La Paz, B.C.S.
- Barjau-González, E., Rodríguez-Romero, J., Galván-Magaña, F. and Maldonado-García, M. 2016. Variación estacional de la diversidad taxonómica de los peces de arrecifes rocosos en el suroeste del Golfo de California. *Revista de Biología Marina y Oceanografía*, 51(1), pp. 11–19. doi:10.4067/S0718-19572016000100002
- Barnes, M. A. and Turner, C. R. 2016. The ecology of environmental DNA and implications for conservation genetics. *Conservation Genetics*, Springer Netherlands, 17(1), pp. 1–17. doi:10.1007/s10592-015-0775-4
- Beentjes, K. K., Speksnijder, A. G. C. L., Schilthuizen, M., Hoogeveen, M. and van der Hoorn, B. B. 2019. The effects of spatial and temporal replicate sampling on eDNA metabarcoding. *PeerJ*, 7, p. e7335. doi:10.7717/peerj.7335

- Beier, E. 1997. A numerical investigation of the annual variability in the Gulf of California. *Journal of Physical Oceanography*, 27(5), pp. 615–632. doi:10.1175/1520-0485(1997)027<0615:ANIOTA>2.0.CO;2
- Beldade, R., Jackson, A. M., Cudney-Bueno, R., Raimondi, P. T. and Bernardi, G. 2014. Genetic structure among spawning aggregations of the Gulf coney *Hyporthodus acanthistius*. *Marine Ecology Progress Series*, 499, pp. 193–201. doi:10.3354/meps10637
- Benestan, L., Fietz, K., Loiseau, N., Guerin, P. E., Trofimenko, E., Rühs, S., Schmidt, C., Rath, W., Biastoch, A., Pérez-Ruzafa, A., Baixauli, P., Forcada, A., Arcas, E., Lenfant, P., Mallol, S., Goñi, R., Velez, L., Höppner, M., Kininmonth, S., ... Manel, S. 2021. Restricted dispersal in a sea of gene flow. *Proceedings of the Royal Society B: Biological Sciences*, 288(1951). doi:10.1098/rspb.2021.0458
- Benestan, L., Quinn, B. K., Maaroufi, H., Laporte, M., Clark, F. K., Greenwood, S. J., Rochette, R. and Bernatchez, L. 2016. Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology*, 25(20), pp. 5073–5092. doi:10.1111/mec.13811
- Berkström, C., Eggertsen, L., Goodell, W., Cordeiro, C. A. M. M., Lucena, M. B., Gustafsson, R., Bandeira, S., Jiddawi, N. and Ferreira, C. E. L. 2020. Thresholds in seascape connectivity: the spatial arrangement of nursery habitats structure fish communities on nearby reefs. *Ecography*, 43(6), pp. 882–896. doi:10.1111/ecog.04868
- Bernt, M., Donath, A., Jühling, F., Externbrink, F., Florentz, C., Fritzsche, G., Pütz, J., Middendorf, M. and Stadler, P. F. 2013. MITOS: Improved *de novo* metazoan mitochondrial genome annotation. *Molecular Phylogenetics and Evolution*, Elsevier Inc., 69(2), pp. 313–319. doi:10.1016/j.ympev.2012.08.023
- Beron-Vera, F. J. and Ripa, P. 2000. Three-dimensional aspects of the seasonal heat balance in the Gulf of California. *Journal of Geophysical Research*, 105(C5), pp. 441–457.
- Bessey, C., Jarman, S. N., Berry, O., Olsen, Y. S., Bunce, M., Simpson, T., Power, M., McLaughlin, J., Edgar, G. J. and Keesing, J. 2020. Maximizing fish detection with eDNA metabarcoding. *Environmental DNA*, pp. 1–12.
- Borthagaray, A. I., Berazategui, M. and Arim, M. 2015. Disentangling the effects of local and regional processes on biodiversity patterns through taxon-contingent metacommunity network analysis. *Oikos*, 124(10), pp. 1383–1390. doi:10.1111/oik.01317
- Borthagaray, A. I., Brazeiro, A. and Giménez, L. 2009. Connectivity and patch area in a coastal marine landscape: Disentangling their influence on local species richness and composition. *Austral Ecology*, 34(6), pp. 641–652. doi:10.1111/j.1442-9993.2009.01969.x

- Boström, C., Pittman, S. J., Simenstad, C. and Kneib, R. T. 2011. Seascape ecology of coastal biogenic habitats: Advances , gaps, and challenges. *Marine Ecology Progress Series*, 427(May 2014), pp. 191–217. doi:10.3354/meps09051
- Boulanger, E., Loiseau, N., Valentini, A., Arnal, V., Boissery, P., Dejean, T., Deter, J., Guellati, N., Holon, F., Juhel, J. B., Lenfant, P., Manel, S. and Mouillot, D. 2021. Environmental DNA metabarcoding reveals and unpacks a biodiversity conservation paradox in Mediterranean marine reserves. *Proceedings of the Royal Society B: Biological Sciences*, 288(1949). doi:10.1098/rspb.2021.0112
- Boussarie, G., Bakker, J., Wangensteen, O. S., Mariani, S., Bonnin, L., Juhel, J.-B., Kiszka, J. J., Kulbicki, M., Manel, S., Robbins, W. D., Vigliola, L. and Mouillot, D. 2018. Environmental DNA illuminates the dark diversity of sharks. *Science Advances*, 4(5), p. eaap9661. doi:10.1126/sciadv.aap9661
- Boyer, F., Mercier, C., Bonin, A., Taberlet, P. and Coissac, E. 2014. obitools: a Unix-inspired software package for DNA metabarcoding. *Molecular Ecology Resources*, 16(1), pp. 176–182. doi:10.1111/1755-0998.12428
- Bozec, Y. M., Kulbicki, M., Laloë, F., Mou-Tham, G. and Gascuel, D. 2011. Factors affecting the detection distances of reef fish: Implications for visual counts. *Marine Biology*, 158(5), pp. 969–981. doi:10.1007/s00227-011-1623-9
- Braunisch, V., Segelbacher, G. and Hirzel, A. H. 2010. Modelling functional landscape connectivity from genetic population structure: a new spatially explicit approach. *Molecular Ecology*, 19(17), pp. 3664–3678. doi:10.1111/j.1365-294X.2010.04703.x
- Brown, E. A., Chain, F. J. J., Crease, T. J., Macisaac, H. J. and Cristescu, M. E. 2015. Divergence thresholds and divergent biodiversity estimates: Can metabarcoding reliably describe zooplankton communities? *Ecology and Evolution*, 5(11), pp. 2234–2251. doi:10.1002/ece3.1485
- Brusca, R. C., Conte, C., and Dimmitt, M. A. (Eds.). 2010. *The Gulf of California: Biodiversity and conservation* (1st ed.). University of Arizona Press, USA.
- Brusca, R. C., Findley, L. T., Hastings, P. a, Hendrickx, M. E., Torre-Cosío, J. and Van Der Heiden, A. M. 2005. Macrofaunal Diversity in the Gulf of California. *Biodiversity, Ecosystems, and Conservation in Northern Mexico*. 1a, Issue January .
- Bunn, A. G., Urban, D. L. and Keitt, T. H. 2000. Landscape connectivity: A conservation application of graph theory. *Journal of Environmental Management*, 59(June), pp. 265–278. doi:10.1006
- Burgess, S. C., Nickols, K. J., Griesemer, C. D., Barnett, L. A. K., Dedrick, A. G., Satterthwaite, E. V., Yamane, L., Morgan, S. G., White, J. W. and Botsford, L. W. 2014. Beyond connectivity: How empirical methods can quantify population persistence to improve marine protected-area design. *Ecological Applications*, 24(2), pp. 257–270. doi:10.1890/13-0710.1

- Calabrese, J. M. and Fagan, W. F. 2004a. A comparison-shopper's guide to connectivity metrics. *Frontiers in Ecology and the Environment*, 2(10), pp. 529–536. doi:10.1890/1540-9295(2004)002[0529:ACGTCM]2.0.CO;2
- Calabrese, J. M. and Fagan, W. F. 2004b. A comparison-shopper's guide to connectivity metrics. *Frontiers in Ecology and the Environment*, 2(10), pp. 529–536. doi:10.1890/1540-9295(2004)002[0529:ACGTCM]2.0.CO;2
- Calderon-Aguilera, L. E., Marinone, S. G. and Aragón-Noriega, E. A. 2003. Influence of oceanographic processes on the early life stages of the blue shrimp (*Litopenaeus stylirostris*) in the Upper Gulf of California. *Journal of Marine Systems*, 39(2003), pp. 507–513. doi:10.1016/S
- Calderón-Sanou, I., Münkemüller, T., Boyer, F., Zinger, L. and Thuiller, W. 2019. From environmental DNA sequences to ecological conclusions: How strong is the influence of data curation strategies? *Journal of Biogeography*, July, pp. 1–14. doi:10.1111/jbi.13681
- Callahan, B. J., McMurdie, P. J. and Holmes, S. P. 2017. Exact sequence variants should replace operational taxonomic units in marker-gene data analysis. *ISME Journal*, Nature Publishing Group, 11(12), pp. 2639–2643. doi:10.1038/ismej.2017.119
- Camacho-Gastélum, R., Sánchez-Velasco, L., Jiménez-Rosenberg, S. P. A., Godínez, V. M. and Tenorio-Fernandez, L. 2020. Interannual variations of environmental factors and the effects on larval fish habitats in the Upper Gulf of California during early summer. *Continental Shelf Research*, 195(November 2019). doi:10.1016/j.csr.2020.104058
- Cantera, I., Cilleros, K., Valentini, A., Cerdan, A., Dejean, T., Iribar, A., Taberlet, P., Vigouroux, R. and Brosse, S. 2019. Optimizing environmental DNA sampling effort for fish inventories in tropical streams and rivers. *Scientific Reports*, 9(1), p. 3085. doi:10.1038/s41598-019-39399-5
- Cantera, I., Decotte, J. B., Dejean, T., Murienne, J., Vigouroux, R., Valentini, A. and Brosse, S. 2021. Characterizing the spatial signal of environmental DNA in river systems using a community ecology approach. *Molecular Ecology Resources*, September, pp. 1–10. doi:10.1111/1755-0998.13544
- Carrillo, L. and Palacios-Hernández, E. 2002. Seasonal evolution of the geostrophic circulation in the Northern Gulf of California. *Estuarine, Coastal and Shelf Science*, 54(2), pp. 157–173. doi:10.1006/ecss.2001.0845
- Cartron, J.-L. E., Ceballos, G. and Felger, R. S. 2005. Biodiversity, Ecosystems, and Conservation in Northern Mexico. J. E. Cartron, G. Ceballos, and R. S. Felger (Eds.), *Biodiversity, Ecosystems, and Conservation in Northern Mexico*. Oxford University Press, New York, NY, 1a, p. 513 .
- Castañeda, R. A., Nynatten, A. Van, Crookes, S. and Weyl, O. L. F. 2020. Detecting Native Freshwater Fishes Using Novel Non-invasive Methods. *Frontiers in Environmental Science*, 8(March), p. 16 pp. doi:10.3389/fenvs.2020.00029

- Catchen, J. M., Amores, a., Hohenlohe, P., Cresko, W., Postlethwait, J. H. and De Koning, D.-J. 2011. Stacks: Building and Genotyping Loci De Novo From Short-Read Sequences. *Genes|Genomes|Genetics*, 1(3), pp. 171–182. doi:10.1534/g3.111.000240
- Cayuela, H., Rougemont, Q., Prunier, J. G., Moore, J. S., Clobert, J., Besnard, A. and Bernatchez, L. 2018. Demographic and genetic approaches to study dispersal in wild animal populations: A methodological review. *Molecular Ecology*, 27(20), pp. 3976–4010. doi:10.1111/mec.14848
- Chase, J. M., Jeliaskov, A., Ladouceur, E. and Viana, D. S. 2020. Biodiversity conservation through the lens of metacommunity ecology. *Annals of the New York Academy of Sciences*, 1469(1), pp. 86–104. doi:10.1111/nyas.14378
- Chen, Q., Zobel, J. and Verspoor, K. 2017. Duplicates, redundancies and inconsistencies in the primary nucleotide databases: A descriptive study. *Database*, 2017(1), pp. 1–16. doi:10.1093/database/baw163
- Cilleros, K., Valentini, A., Allard, L., Dejean, T., Etienne, R., Grenouillet, G., Iribar, A., Taberlet, P., Vigouroux, R. and Brosse, S. 2018. Unlocking biodiversity and conservation studies in high-diversity environments using environmental DNA (eDNA): A test with Guianese freshwater fishes. *Molecular Ecology Resources*, March, pp. 1–20. doi:10.1111/1755-0998.12900
- Cisneros-Mata, M. A. 2010. The importance of fisheries in the Gulf of California and ecosystem-based sustainable co-management for conservation. R. C. Brusca (Ed.), *The Gulf of California: biodiversity and conservation*. University of Arizona Press, 10, pp. 119–134 .
- Cisneros-Mata, M. Á., Munguía-Vega, A., Rodríguez-Félix, D., Aragón-Noriega, E. A., Grijalva-Chon, J. M., Arreola-Lizárraga, J. A. and Hurtado, L. A. 2019. Genetic diversity and metapopulation structure of the brown swimming crab (*Callinectes bellicosus*) along the coast of Sonora, Mexico: Implications for fisheries management. *Fisheries Research*, Elsevier, 212(December 2018), pp. 97–106. doi:10.1016/j.fishres.2018.11.021
- Closek, C. J., Santora, J. A., Starks, H. A., Schroeder, I. D., Andruszkiewicz, E. A., Sakuma, K. M., Bograd, S. J., Hazen, E. L., Field, J. C. and Boehm, A. B. 2019. Marine Vertebrate Biodiversity and Distribution Within the Central California Current Using Environmental DNA (eDNA) Metabarcoding and Ecosystem Surveys. *Frontiers in Marine Science*, 6(December), pp. 1–17. doi:10.3389/fmars.2019.00732
- Collins, R. A., Bakker, J., Wangensteen, O. S., Soto, A. Z., Corrigan, L., Sims, D. W., Genner, M. J. and Mariani, S. 2019. Non-specific amplification compromises environmental DNA metabarcoding with COI. *Methods in Ecology and Evolution*, 10(11), pp. 2041–210X.13276. doi:10.1111/2041-210X.13276
- Contreras-Catala, F., Sánchez-Velasco, L., Lavín, M. F. and Godínez, V. M. 2012. Three-dimensional distribution of larval fish assemblages in an anticyclonic eddy in a semi-enclosed sea (Gulf of California). *Journal of Plankton Research*, 34(6), pp. 548–562. doi:10.1093/plankt/fbs024

- Cordier, T., Forster, D., Dufresne, Y., Martins, C. I. M., Stoeck, T. and Pawlowski, J. 2018. Supervised machine learning outperforms taxonomy-based environmental DNA metabarcoding applied to biomonitoring. *Molecular Ecology Resources*, June. doi:10.1111/1755-0998.12926
- Correa Ayram, C., Mendoza, M. E., Etter, A. and Salicrup, D. R. P. 2015. Habitat connectivity in biodiversity conservation: A review of recent studies and applications. *Progress in Physical Geography*, pp. 1–32. doi:10.1177/0309133315598713
- Costa, A., Petrenko, A. A., Guizien, K. and Doglioli, A. M. 2017. On the calculation of betweenness centrality in marine connectivity studies using transfer probabilities. *PLoS ONE*, 12(12). doi:10.1371/journal.pone.0189021
- Cottenie, K. 2005. Integrating environmental and spatial processes in ecological community dynamics. *Ecology Letters*, 8(11), pp. 1175–1182. doi:10.1111/j.1461-0248.2005.00820.x
- Cowen, R.K., Gawarkiewicz, G., Pineda, J., Thorrold, S. R. and Werner, F. 2007. Population Connectivity in Marine Systems, an overview. *Oceanography*, 20(3), pp. 14–21. doi:10.1126/science.1122039
- Cowen, Robert K, Lwiza, K. M. M., Sponaugle, S., Paris, C. B. and Olson, D. B. 2000. Connectivity of Marine Populations : Open or Closed 287(February), pp. 857–860.
- Cowen, Robert K and Sponaugle, S. 2009. Larval Dispersal and Marine Population Connectivity. *Annual Review of Marine Science*, 1(1), p. 26. doi:10.1146/annurev.marine.010908.163757
- Craig, M. T., Hastings, P. A., Pondella, D. J., Ross Robertson, D. and Rosales-Casián, J. A. 2006. Phylogeography of the flag cabrilla *Epinephelus labriformis* (Serranidae): Implications for the biogeography of the Tropical Eastern Pacific and the early stages of speciation in a marine shore fish. *Journal of Biogeography*, 33(6), pp. 969–979. doi:10.1111/j.1365-2699.2006.01467.x
- Crandall, E. D., Toonen, R. J. and Selkoe, K. A. 2019. A coalescent sampler successfully detects biologically meaningful population structure overlooked by F-statistics. *Evolutionary Applications*, 12(2), pp. 255–265. doi:10.1111/eva.12712
- Cristescu, M. E. and Hebert, P. 2018. Uses and misuses of environmental DNA in biodiversity science and conservation. *Annual Reviews*, 49, pp. 209–230.
- Crooks, K. R. and Sanjayan, M. 2006. *Connectivity conservation* (1st ed.). Cambridge University Press, United States of America.
- Csardi, G. and Nepusz, T. 2006. The *igraph* software package for complex network research (p. 1695). *InterJournal*.
- Cushman, S. A., Gutzweiler, K., Evans, J. S. and McGarigal, K. 2010. The gradient paradigm: a conceptual and analytical framework for landscape ecology. *Spatial Complexity, Informatics, and Wildlife Conservation*. Vol. 9784431877. pp. 1–458. doi:10.1007/978-4-431-87771-4

- D'Aloia, C. C., Andrés, J. A., Bogdanowicz, S. M., McCune, A. R., Harrison, R. G. and Buston, P. M. 2020. Unraveling hierarchical genetic structure in a marine metapopulation: A comparison of three high-throughput genotyping approaches. *Molecular Ecology*, 29(12), pp. 2189–2203. doi:10.1111/mec.15405
- D'Aloia, C. C., Bogdanowicz, S. M., Harrison, R. G. and 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(12), pp. 2902–2913. doi:10.1111/mec.12782
- D'Aloia, C. C., Bogdanowicz, S. M., Majoris, J. E., Harrison, R. G. and Buston, P. M. 2013. Self-recruitment in a Caribbean reef fish: A method for approximating dispersal kernels accounting for seascape. *Molecular Ecology*, 22(9), pp. 2563–2572. doi:10.1111/mec.12274
- Dalongeville, A., Andrello, M., Mouillot, D., Lobreaux, S., Fortin, M. J., Lasram, F., Belmaker, J., Rocklin, D. and Manel, S. 2018. Geographic isolation and larval dispersal shape seascape genetic patterns differently according to spatial scale. *Evolutionary Applications*, 11(8), pp. 1437–1447. doi:10.1111/eva.12638
- Deiner, K., Bik, H. M., Elvira, M., Seymour, M., Creer, S., Bista, I., Pfrender, M. E., Bernatchez, L., Altermatt, F., Lodge, D. M. and Vere, N. De. 2017. Environmental DNA metabarcoding : Transforming how we survey animal and plant communities. *Molecular Ecology*, September, pp. 1–24. doi:10.1111/mec.14350
- Diana, A., Matechou, E., Griffin, J. E., Buxtron, A. S. and Griffiths, R. A. 2020. An Rshiny app for modelling environmental DNA data: accounting for false positive and false negative observation error. *BioRxiv*, 2019, p. doi.org/10.1101/2020.12.09.417600.
- Díaz-Urbibe, J. G., Elorduy-Garay, J. F. and González-Valdovinos, M. T. 2001. Age and growth of the Leopard grouper, *Mycteroperca rosacea*, in the Southern Gulf of California, Mexico. *Pacific Science*, 55(2), pp. 171–182. doi:10.1353/psc.2001.0012
- Diniz-Filho, J. A. F. and Bini, L. M. 2011. Geographical Patterns in Biodiversity: Towards an Integration of Concepts and Methods from Genes to Species Diversity. *Natureza & Conservação*, 9(2), pp. 179–187. doi:10.4322/natcon.2011.023
- Doi, H., Fukaya, K., Oka, S., Ichiro, Sato, K., Kondoh, M. and Miya, M. 2019. Evaluation of detection probabilities at the water-filtering and initial PCR steps in environmental DNA metabarcoding using a multispecies site occupancy model. *Scientific Reports*, Springer US, 9(1), pp. 1–8. doi:10.1038/s41598-019-40233-1
- Dorazio, R. M. and Erickson, R. A. 2017. ednaoccupancy: An r package for multiscale occupancy modelling of environmental DNA data. *Molecular Ecology Resources*, 18(2), pp. 368–380.
- Dunning, J. B., Danielson, B. J. and Pulliam, H. R. 1992. Ecological populations affect processes that in complex landscapes. *Nordic Society Oikos*, 65(1), pp. 169–175.

- Economo, E. P. and Keitt, T. H. 2010. Network isolation and local diversity in neutral metacommunities. *Oikos*, 119(8), pp. 1355–1363. doi:10.1111/j.1600-0706.2010.18272.x
- Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C. and Knight, R. 2018. UCHIME improves sensitivity and speed of chimera detection. *27(16)*, pp. 2194–2200. doi:10.1093/bioinformatics/btr381
- Elorduy-Garay, J. F. and Caraveo-Patiño, J. 1994. Feeding habits of the ocean whitefish, *Caulolatilus princeps* Jenyns 1842 (Pisces: Branchiostegidae), in La Paz Bay, B.C.S., Mexico. *Ciencias Marinas*, 20(2), pp. 199–218.
- Erisman, B., Aburto-Oropeza, O., Gonzalez-Abraham, C., Mascareñas-Osorio, I., Moreno-Báez, M. and Hastings, P. A. 2012. Spatio-temporal dynamics of a fish spawning aggregation and its fishery in the Gulf of California. *Scientific Reports*, 2, p. 284. doi:10.1038/srep00284
- Erisman, B. E., Buckhorn, M. L. and Hastings, P. A. 2007. Spawning patterns in the leopard grouper, *Mycteroperca rosacea*, in comparison with other aggregating groupers. *Marine Biology*, 151(5), pp. 1849–1861. doi:10.1007/s00227-007-0623-2
- Erisman, B., Mascarenas, I., Paredes, G., Sadovy de Mitcheson, Y., Aburto-Oropeza, O. and Hastings, P. 2010. Seasonal, annual, and long-term trends in commercial fisheries for aggregating reef fishes in the Gulf of California, Mexico. *Fisheries Research*, Elsevier B.V., 106(3), pp. 279–288. doi:10.1016/j.fishres.2010.08.007
- Escalante, F., Valdez-Holguín, J. E., Álvarez-Borrego, S. and Lara-Lara, J. R. 2013. Temporal and spatial variation of sea surface temperature, chlorophyll a, and primary productivity in the Gulf of California. *Ciencias Marinas*, 39(2), pp. 203–215. doi:10.7773/cm.v39i2.2233
- Estrada-Godínez, J. A., García-Maldonado, M., Gracia-López, V. and Carrillo, M. 2011. Reproductive cycle of the leopard grouper *Mycteroperca rosacea* in La Paz Bay, Mexico. *Ciencias Marinas*, 37(4), pp. 425–441.
- Evans, N. T. and Lamberti, G. A. 2018. Freshwater fisheries assessment using environmental DNA: A primer on the method, its potential, and shortcomings as a conservation tool. *Fisheries Research*, Elsevier, 197(February 2017), pp. 60–66. doi:10.1016/j.fishres.2017.09.013
- Evans, N. T., Olds, B. P., Renshaw, M. A., Turner, C. R., Li, Y., Jerde, C. L., Mahon, A. R., Pfrender, M. E., Lamberti, G. A. and Lodge, D. M. 2016. Quantification of mesocosm fish and amphibian species diversity via environmental DNA metabarcoding. *Molecular Ecology Resources*, 16(1), pp. 29–41. doi:10.1111/1755-0998.12433
- Evans, N. T., Olds, B. P., Turner, C. R., Renshaw, M. A., Li, Y., Deiner, K., Lodge, D. M., Jerde, C. L., Pfrender, M. E. and Lamberti, G. A. 2016. Fish community assessment with eDNA metabarcoding: effects of sampling design and bioinformatic filtering. *Canadian Journal of Fisheries and Aquatic Sciences*, 74(9), pp. 1362–1374. doi:10.1139/cjfas-2016-0306

- Excoffier, L., Laval, G. and Schneider, S. 2005. Arlequin (version 3.0): an integrated software package for population genetics data analysis. *Evolutionary Bioinformatics Online*, 1, pp. 47–50. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2658868&tool=pmcentrez&rendertype=abstract>
- Fernández-Rivera Melo, F. J., Reyes-Bonilla, H., Ramírez-Ortiz, G. and Alvarez-Filip, L. 2018. Latitudinal variation in structure and function of conspicuous reef fish assemblages along the western Gulf of California. *Revista Mexicana de Biodiversidad*, 89(4), pp. 1154–1166. doi:10.22201/ib.20078706e.2018.4.2231
- Ficetola, G. F., Taberlet, P. and Coissac, E. 2016. How to limit false positives in environmental DNA and metabarcoding? *Molecular Ecology Resources*, 16(3), pp. 604–607. doi:10.1111/1755-0998.12508
- Ficetola, G., Pansu, J., Bonin, A., Coissac, E., Giguët-Covex, C., De Barba, M., Gielly, L., Lopes, C. M., Boyer, F., Pompanon, F., Rayé, G. and Taberlet, P. 2014. Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data. *Molecular Ecology Resources*, 15(3), pp. 543–556. doi:10.1111/1755-0998.12338
- Fletcher, R. J., Burrell, N. S., Reichert, B. E., Vasudev, D. and Austin, J. D. 2016. Divergent Perspectives on Landscape Connectivity Reveal Consistent Effects from Genes to Communities. *Current Landscape Ecology Reports*, *Current Landscape Ecology Reports*, 1(2), pp. 67–79. doi:10.1007/s40823-016-0009-6
- Foll, M. and Gaggiotti, O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180(2), pp. 977–993. doi:10.1534/genetics.108.092221
- Fuiman, L. A. and Werner, R. G. 2002. *Fishery Science : The Unique contributions of early life stages (1st ed.)*. Gran Bretaña.
- Gaines, S. D., Gaylord, B., Gerber, L. R., Hastings, A. and Kinlan, B. P. 2007. Connecting places: The ecological consequences of dispersal in the sea. *Oceanography*, 20(SPL.ISS. 3), pp. 90–99. doi:10.5670/oceanog.2007.32
- Garcés-Rodríguez, Y., Sánchez-Velasco, L., Parés-Sierra, A., Jiménez-Rosenberg, S. P. A., Márquez-Artavia, A. and Flores-Morales, A. L. 2021. FISH larvae distribution and transport on the thermal fronts in the Midriff Archipelago region, Gulf of California. *Continental Shelf Research*, 218(February), pp. 1–13. doi:10.1016/j.csr.2021.104384
- García-De León, F. J., Galvan-Tirado, C., Sanchez Velasco, L., Silva-Segundo, C. A., Hernandez-Guzman, R., Barriga-Sosa, I. de los A., Jaimes, P. D., Canino, M. and Cruz-Hernandez, P. 2018. Role of oceanography in shaping the genetic structure in the North Pacific hake *Merluccius productus*. *PLoS ONE*, 13(3), pp. 1–26. doi:https://doi.org/10.1371/journal.pone.0194646

- Glenn, T. C., Nilsen, R. A., Kieran, T. J., Jr, J. W. F., Pierson, T. W., Bentley, K. E., Hoffberg, S. L., Louha, S., León, F. J. G.-D., Portilla, M. A. D. R., Reed, K. D., Anderson, J. L., Meece, J. K., Aggr, S. E. and Faircloth, B. C. 2016. Adapterama I: Universal stubs and primers for thousands of dual-indexed Illumina libraries (iTru & iNext). *BioRxiv*, p. 049114. doi:10.1101/049114
- Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., Spear, S. F., McKee, A., Oyler-McCance, S. J., Cornman, R. S., Laramie, M. B., Mahon, A. R., Lance, R. F., Pilliod, D. S., Strickler, K. M., Waits, L. P., Fremier, A. K., Takahara, T., Herder, J. E., ... Gilbert, M. 2016. Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, 7(11), pp. 1299–1307. doi:10.1111/2041-210X.12595
- Goudet, J., Jombart, T., Kamvar, Z. N., Archer, E. and Hardy, O. J. 2020. Package ‘hierfstat’ (p. 67).
- Green, A. L., Fernandes, L., Almany, G., Abesamis, R., McLeod, E., Aliño, P. M., White, A. T., Salm, R., Tanzer, J. and Pressey, R. L. 2014. Designing Marine Reserves for Fisheries Management, Biodiversity Conservation, and Climate Change Adaptation. *Coastal Management*, 42(2). doi:10.1080/08920753.2014.877763
- Griffin, J. E., Matechou, E., Buxton, A. S., Bormpoudakis, D. and Griffiths, R. A. 2020. Modelling environmental DNA data; Bayesian variable selection accounting for false positive and false negative errors. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, pp. 377–392. doi:10.1111/rssc.12390
- Grober-Dunsmore, R., Pittman, S. J., Kendall, M. S. and Frazer, T. K. 2009. A Landscape Ecology Approach for the Study of Ecological Connectivity Across Tropical Marine Seascapes. In: Nagelkerke (Ed.), *Ecological Connectivity among Tropical Coastal Ecosystems*. Springer Science+Business Media, 1st ed., pp. 493–530. doi:10.1007/978-90-481-2406-0
- Guardiola, M., Wangensteen, O. S., Taberlet, P., Coissac, E., Uriz, M. J. and Turon, X. 2016. Spatio-temporal monitoring of deep-sea communities using metabarcoding of sediment DNA and RNA. *PeerJ*, 4, p. e2807. doi:10.7717/peerj.2807
- Guillera-Arroita, G., Lahoz-Monfort, J. J., van Rooyen, A., Weeks, A. R. and Tingley, R. 2017. Dealing with false-positive and false-negative errors about species occurrence at multiple levels. *Methods in Ecology and Evolution*, 8(1), pp. 1081–1091.
- Guimarães, P. R. 2020. The Structure of Ecological Networks across Levels of Organization. *Annual Review of Ecology, Evolution, and Systematics*, 51, pp. 433–460. doi:10.1146/annurev-ecolsys-012220-120819
- Günther, B., Knebelsberger, T., Neumann, H., Laakmann, S. and Martínez Arbizu, P. 2018. Metabarcoding of marine environmental DNA based on mitochondrial and nuclear genes. *Scientific Reports*, 8(1), pp. 1–13. doi:10.1038/s41598-018-32917-x

- Guo, B., DeFaveri, J., Sotelo, G., Nair, A. and Merilä, J. 2015. Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biology*, 2015(13), p. 19. doi:10.1186/s12915-015-0130-8
- Gutierrez, O., Marinone, S. G. and Parés-Sierra, A. 2004. Lagrangian surface circulation in the Gulf of California from a 3D numerical model. *Deep Sea Research II*, 51, pp. 659–672. doi:10.1016/j.dsr2.2004.05.016
- Hansen, B. K., Bekkevold, D., Clausen, L. W. and Nielsen, E. E. 2018. The sceptical optimist: challenges and perspectives for the application of environmental DNA in marine fisheries. *Fish and Fisheries*, 19(5), pp. 1–18. doi:10.1111/faf.12286
- Hanski, I. 1998. Metapopulation dynamics. *Nature*, 396(6706), pp. 41–49. doi:10.1038/23876
- Hanski, I., and Gaggiotti, O. E. (Eds.). 2004. *Ecology, Genetics, and Evolution of Metapopulations* (1<sup>st</sup>ed). Elsevier Inc., United States of America.
- Hanski, I. and Gilpin, M. 1991. Metapopulation dynamics : brief history and conceptual domain. *Biological Journal of the Linnean Society*, 42, pp. 3–16. doi:10.1111/j.1095-8312.1991.tb00548.x
- Hanski, I. and Ovaskainen, O. The metapopulation capacity of a fragmented landscape. *Nature*, 404(6779), pp. 755–758.
- Harrison, S. and Hastings, A. 1996. Genetic and evolutionary consequences of metapopulation structure. *Trends in Ecology and Evolution*, 11(4), pp. 180–183. doi:10.1016/0169-5347(96)20008-4
- Hastings, P., Findley, L. T. and Van Der Heiden, A. M. 2010. *Fishes of the Gulf of California*. C. Brusca, Richard C. (Ed.), *The Gulf of California Biodiversity and Conservation*. The University of Arizona Press, USA, 1st ed. .
- Hauser, L. and Carvalho, G. R. 2008. Paradigm shifts in marine fisheries genetics: Ugly hypotheses slain by beautiful facts. *Fish and Fisheries*, 9(4), pp. 333–362. doi:10.1111/j.1467-2979.2008.00299.x
- Hebert, P. D. N., Cywinska, A., Ball, S. L. and DeWaard, J. R. 2003. Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences*, 270(1512), pp. 313–321. doi:10.1098/rspb.2002.2218
- Hedgecock, D., Barber, P. H. and Edmands, S. 2007. Genetic Approaches to Measuring Connectivity. *Oceanography*, 20(3), pp. 70–79. doi:10.1073/pnas.0401921101
- Helfman, G., Collette, B. B., Facey, D. E. and Bowen, B. W. 2009. *The diversity of fishes biology, evolution and ecology* (2a). Wiley-Blackwell, United Kingdom.

- Hellberg, M. E. 2006. Genetic Approaches to Understanding Marine Metapopulation Dynamics. Marine Metapopulations. Elsevier Inc. doi:10.1016/B978-012088781-1/50016-9
- Hellberg, M. E. 2009. Gene flow and isolation among populations of marine animals. Annual Review of Ecology, Evolution, and Systematics, 40, pp. 291–310. doi:10.1146/annurev.ecolsys.110308.120223
- Hines, J. 2006. PRESENCE- Software to estimate patch occupancy and related parameters.
- Hock, K. and Mumby, P. J. 2015. Quantifying the reliability of dispersal paths in connectivity networks. Journal of the Royal Society, Interface, 12(February), p. 20150013. doi:10.1098/rsif.2015.0013
- Holderegger, R., Kamm, U. and Gugerli, F. 2006. Adaptive vs. neutral genetic diversity: implications for landscape genetics. Landscape Ecology, 21(6), pp. 797–807. doi:10.1007/s10980-005-5245-9
- Holliday, J. A., Hallerman, E. M. and Haak, D. C. 2019. Genotyping and sequencing technologies in population genetics and genomics. O. P. Rajora (Ed.), Population genomics: concepts, approaches and applications. Springer Nature Switzerland AG, pp. 83–126 .
- Holyoak, M., Leibold, M. A., and Holts, R. (Eds.). 2005. Metacommunities, Spatial Dynamics and Ecological Communities (1o). The University of Chicago Press, United States of America.
- Hubbell, S. P. 2001. The unified neutral theory of biodiversity and biogeography (1°). Princeton University Press. doi:10.1111/j.1939-7445.2005.tb00163.x
- Hunter, M. E., Dorazio, R. M., Butterfield, J. S. S., Meigs-Friend, G., Nico, L. G. and Ferrante, J. A. 2017. Detection limits of quantitative and digital PCR assays and their influence in presence–absence surveys of environmental DNA. Molecular Ecology Resources, 17(2), pp. 221–229. doi:10.1111/1755-0998.12619
- Inda-Díaz, E. A., Sánchez-Velasco, L. and Lavín, M. F. 2010. Three-dimensional distribution of small pelagic fish larvae (*Sardinops sagax* and *Engraulis mordax*) in a tidal-mixing front and surrounding waters (Gulf of California). Journal of Plankton Research, 32(9), pp. 1241–1254. doi:10.1093/plankt/fbq051
- Jackson, A. M., Munguia-Vega, A., Beldade, R., Erisman, B. E. and Bernardi, G. 2015. Incorporating historical and ecological genetic data for leopard grouper (*Mycteroperca rosacea*) into marine reserve design in the Gulf of California. Conservation Genetics, 16(4), pp. 811–822. doi:10.1007/s10592-015-0702-8
- Jacobson, B. and Peres-Neto, P. R. 2010. Quantifying and disentangling dispersal in metacommunities: How close have we come? How far is there to go? Landscape Ecology, 25(4), pp. 495–507. doi:10.1007/s10980-009-9442-9
- Jeunen, G. J., Knapp, M., Spencer, H. G., Lamare, M. D., Taylor, H. R., Stat, M., Bunce, M. and Gemmill, N. J. 2019. Environmental DNA (eDNA) metabarcoding reveals strong discrimination among diverse

- marine habitats connected by water movement. *Molecular Ecology Resources*, 19(2), pp. 426–438. doi:10.1111/1755-0998.12982
- Jombart, T., Devillard, S., Dufour, A.-B. and Pontier, D. 2008. Revealing cryptic spatial patterns in genetic variability by a new multivariate method. *Heredity*, 101(1), pp. 92–103. doi:10.1038/hdy.2008.34
- Juhel, J. B., Utama, R. S., Marques, V., Vimono, I. B., Sugeha, H. Y., Kadarusman, Pouyaud, L., Dejean, T., Mouillot, D. and Hocdé, R. 2020. Accumulation curves of environmental DNA sequences predict coastal fish diversity in the coral triangle. *Proceedings. Biological Sciences*, 287(1930), p. 20200248. doi:10.1098/rspb.2020.0248
- Junge, C., Donnellan, S. C., Huvneers, C., Bradshaw, C. J. A., Simon, A., Drew, M., Duffy, C., Johnson, G., Cliff, G., Braccini, M., Cutmore, S. C., Butcher, P., McAuley, R., Peddemors, V., Rogers, P. and Gillanders, B. M. 2019. Comparative population genomics confirms little population structure in two commercially targeted carcharhinid sharks. *Marine Biology*, Springer Berlin Heidelberg, 166(2), pp. 1–15. doi:10.1007/s00227-018-3454-4
- Kahru, M., Marinone, S., Lluch-Cota, S., Parés-Sierra, A. and Mitchell, B. 2004. Ocean-color variability in the Gulf of California: scales from days to ENSO. *Deep Sea Research II*, 51, pp. 139–146.
- Kelly, R. P., Shelton, A. O. and Gallego, R. 2019. Understanding PCR processes to draw meaningful conclusions from environmental DNA Studies. *BioRxiv*, p. 660530. doi:10.1101/660530
- Kimura, M. and Weiss, G. H. 1964. The stepping stone model of population structure and the decrease of genetic correlation with distance. *Genetics*, 49(480), pp. 561–576.
- Kool, J. T., Moilanen, A. and Treml, E. A. 2013. Population connectivity: Recent advances and new perspectives. *Landscape Ecology*, 28(2), pp. 165–185. doi:10.1007/s10980-012-9819-z
- Kritzer, J. P., and Sale, P. F. (Eds.). 2006. *Marine metapopulations (1a)*. Elsevier Academic Press, United States of America.
- Lafferty, K. D., Garcia-Vedrenne, A. E., McLaughlin, J. P., Childress, J. N., Morse, M. F. and Jerde, C. L. 2020. At Palmyra Atoll, the fish-community environmental DNA signal changes across habitats but not with tides. *Journal of Fish Biology*, May. doi:10.1111/jfb.14403
- Lahoz-Monfort, J. J., Guillera-Arroita, G. and Tingley, R. 2016. Statistical approaches to account for false-positive errors in environmental DNA samples. *Molecular Ecology Resources*, 16(3), pp. 673–685. doi:10.1111/1755-0998.12486
- Lamy, T., Pitz, K. J., Chavez, F. P., Yorke, C. E. and Miller, R. J. 2021. Environmental DNA reveals the fine-grained and hierarchical spatial structure of kelp forest fish communities. *Scientific Reports*, Nature Publishing Group UK, 11(1), pp. 1–13. doi:10.1038/s41598-021-93859-5

- Lavin, M. F., Castro, R., Beier, E., Cabrera, C., Godínez, V. M. and Amador-Buenrostro, A. 2014. Surface circulation in the Gulf of California in summer from surface drifters and satellite images (2004–2006). *Journal of Geophysical Research*, pp. 4278–4290. doi:10.1002/2013JC009345. Received
- Lavín, M. F. and Marinone, S. G. 2003. An overview of the physical oceanography of the Gulf of California. O. U. Velasco-Fuentes, J. Sheinbaum, and J. Ochoa (Eds.), *Nonlinear Processes in Geophysical Fluid Dynamics*. Kluwer Academic Publishers, The Netherlands, 1st ed., pp. 173–204 . doi:10.1029/RG021i005p01105
- Lavín, M., Hernández-Palacios, E. and Cabrera, C. 2003. Sea surface temperature anomalies in the Gulf of California. *Geofísica Internacional*, 42(3), pp. 363–375. [http://www.raugm.org.mx/publicaciones/GeofisicaInternacional/GI\\_2003\\_V.42/03/lavin.pdf](http://www.raugm.org.mx/publicaciones/GeofisicaInternacional/GI_2003_V.42/03/lavin.pdf)
- Lecaudey, L., Schletterer, M., Kuzovlev, V. V, Hahn, C. and Weiss, S. J. 2019. Fish diversity assessment in the headwaters of the Volga River using environmental DNA metabarcoding. *Aquatic Conservation: Marine and Freshwater Ecosystems*, September 2018, pp. 1–16. doi:10.1002/aqc.3163
- Legendre, P. and De Cáceres, M. 2013. Beta diversity as the variance of community data: Dissimilarity coefficients and partitioning. *Ecology Letters*, 16(8), pp. 951–963. doi:10.1111/ele.12141
- Legendre, P. and Legendre, L. 1998. *Numerical ecology* (2nd ed.). Elsevier, Amsterdam.
- Leibold, M. A. and Chase, J. M. 2018. *Metacommunity ecology* (1st ed.). Princeton University Press, United States of America.
- Leibold, M. A., Holyoak, M., Mouquet, N., Amarasekare, P., Chase, J. M., Hoopes, M. F., Holt, R. D., Shurin, J. B., Law, R., Tilman, D., Loreau, M. and Gonzalez, A. 2004. The metacommunity concept: A framework for multi-scale community ecology. *Ecology Letters*, 7(7), pp. 601–613. doi:10.1111/j.1461-0248.2004.00608.x
- Leprieur, F., Olden, J. D., Lek, S. and Brosse, S. 2009. Contrasting patterns and mechanisms of spatial turnover for native and exotic freshwater fish in Europe. *Journal of Biogeography*, 36(10), pp. 1899–1912. doi:10.1111/j.1365-2699.2009.02107.x
- Li, Y., Evans, N. T., Renshaw, M. A., Jerde, C. L., Olds, B. P., Shogren, A. J., Deiner, K., Lodge, D. M., Lamberti, G. A. and Pfrender, M. E. 2018. Estimating fish alpha- And beta-diversity along a small stream with environmental DNA metabarcoding. *Metabarcoding and Metagenomics*, 2, pp. 1–11. doi:10.3897/mbmg.2.24262
- Liggins, L., Treml, E. A. and Riginos, C. 2019. Seascape Genomics: Contextualizing Adaptive and Neutral Genomic Variation in the Ocean Environment. M. F. Oleksiak and O. P. Rajora (Eds.), *Population genomics: concepts, approaches and applications*. Springer Nature Switzerland AG, 1st ed., Issue December. pp. 171–218 . doi:10.1007/13836\_2019\_68

- Lluch-Cota, S. E. 2000. Coastal upwelling in the eastern Gulf of California. *Oceanologica Acta*, 23, pp. 731–740. doi:10.1016/S0399-1784(00)00121-3
- Lluch-Cota, Salvador E., Aragón-Noriega, E. A., Arreguín-Sánchez, F., Aurióles-Gamboa, D., Jesús Bautista-Romero, J., Brusca, R. C., Cervantes-Duarte, R., Cortés-Altamirano, R., Del-Monte-Luna, P., Esquivel-Herrera, A., Fernández, G., Hendrickx, M. E., Hernández-Vázquez, S., Herrera-Cervantes, H., Kahru, M., Lavín, M., Lluch-Belda, D., Lluch-Cota, D. B., López-Martínez, J., ... Sierra-Beltrán, A. P. 2007. The Gulf of California: Review of ecosystem status and sustainability challenges. *Progress in Oceanography*, 73(1), pp. 1–26. doi:10.1016/j.pocean.2007.01.013
- Lodeiros, C., Soria, G., Valentich-Scott, P., Munguia-Vega, A., Cabrera, J. S., Cudney-Bueno, R., Loor, A., Márquez, A. and Sonnenholzner, S. 2016. Spondylids of Eastern Pacific Ocean. *Journal of Shellfish Research*, 35(2), pp. 279–293. doi:10.2983/035.035.0203
- Loreau, M., Mouquet, N. and Gonzalez, A. 2003. Biodiversity as spatial insurance in heterogeneous landscapes. *Proceedings of the National Academy of Sciences of the United States of America*, 100(22), pp. 12765–12770. doi:10.1073/pnas.2235465100
- Loreau, M., Mouquet, N. and Holt, R. D. 2003. Meta-ecosystems: A theoretical framework for a spatial ecosystem ecology. *Ecology Letters*, 6(8), pp. 673–679. doi:10.1046/j.1461-0248.2003.00483.x
- Lowe, W. H. and Allendorf, F. W. 2010. What can genetics tell us about population connectivity? *Molecular Ecology*, 19(1), pp. 3038–3051. doi:10.1111/j.1365-294X.2010.04688.x
- Luiz, O. J., Allen, A. P., Robertson, D. R., Floeter, S. R., Kulbicki, M., Vigliola, L., Becheler, R. and Madin, J. S. 2013. Adult and larval traits as determinants of geographic range size among tropical reef fishes. *Proceedings of the National Academy of Sciences of the United States of America*, 110(41), pp. 16498–16502. doi:10.1073/pnas.1304074110
- Luu, K., Bazin, E. and Blum, M. G. B. 2017. *pcadapt*: an R package to perform genome scans for selection based on principal component analysis. *Molecular Ecology Resources*, 17(1), pp. 67–77. doi:10.1111/1755-0998.12592
- MacArthur, R. H. and Wilson, E. O. 1967. *The theory of island biogeography* (1st ed.). Princeton University Press, Princeton, New Jersey.
- Mackenzie, D. I., Nichols, J. D., Lachman, G., Droege, S., Royle, J. A. and Langtimm, C. 2002. Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, 83(8), pp. 2248–2255.
- MacNeil, M. A., Tyler, E. H. M., Fonnesebeck, C. J., Rushton, S. P., Polunin, N. V. C. and Conroy, M. J. 2008. Accounting for detectability in reef-fish biodiversity estimates. *Marine Ecology Progress Series*, 367, pp. 249–260. doi:10.3354/meps07580

- Magris, R. A., Andrello, M., Pressey, R. L., Mouillot, D., Dalongeville, A., Jacobi, M. N. and Manel, S. 2018. Biologically representative and well-connected marine reserves enhance biodiversity persistence in conservation planning. *Conservation Letters*, 11(4), pp. 1–10. doi:10.1111/conl.12439
- Magris, R. A., Treml, E. A., Pressey, R. L. and Weeks, R. 2016. Integrating multiple species connectivity and habitat quality into conservation planning for coral reefs. *Ecography*, June, pp. 1–16. doi:10.1111/ecog.01507
- Magurran, A. E. 2004. *Measuring biological diversity* (1st ed.). Blackwell Publishing.
- Mahé, F., Rognes, T., Quince, C., de Vargas, C. and Dunthorn, M. 2015. Swarm v2: highly-scalable and high-resolution amplicon clustering. *PeerJ*, 3, p. e1420. doi:10.7717/peerj.1420
- Maloney, K. O. and Munguia, P. 2011. Distance decay of similarity in temperate aquatic communities: effects of environmental transition zones, distance measure, and life histories. *Ecography*, 34(2), pp. 287–295.
- Marinone, S. G. 2003. A three-dimensional model of the mean and seasonal circulation of the Gulf of California. *Journal of Geophysical Research*, 108(C10), pp. 1–68. doi:10.1029/2002JC001720
- Marinone, S. G. 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(3), pp. 430–434. doi:10.1016/j.ecss.2008.09.002
- Marinone, S. G. 2012. Seasonal surface connectivity in the Gulf of California. *Estuarine, Coastal and Shelf Science*, Elsevier Ltd, 100, pp. 133–141. doi:10.1016/j.ecss.2012.01.003
- Marinone, S. G., Lavín, M. F. and Parés-Sierra, A. 2011. A Quantitative characterization of the seasonal lagrangian circulation of the Gulf of California from a three-dimensional numerical model. *Continental Shelf Research*, 31(14), pp. 1420–1426. doi:10.1016/j.csr.2011.05.014
- Marinone, S. G., Ulloa, M. J., Parés-Sierra, A., Lavín, M. F. and Cudney-Bueno, R. 2008. Connectivity in the northern Gulf of California from particle tracking in a three-dimensional numerical model. *Journal of Marine Systems*, 71(1–2), pp. 149–158. doi:10.1016/j.jmarsys.2007.06.005
- Marko, P. B. and Hart, M. W. 2011. The complex analytical landscape of gene flow inference. *Trends in Ecology & Evolution*, 26(9), pp. 448–456. doi:10.1016/j.tree.2011.05.007
- Martel, C. M., Sutter, M., Dorazio, R. M. and Kinziger, A. P. 2020. Using environmental DNA and occupancy modelling to estimate rangewide metapopulation dynamics. *Molecular Ecology*, October, pp. 1–15. doi:10.1111/mec.15693
- Martínez-Montaño, E., González-Álvarez, K., Lazo, J. P., Audelo-Naranjo, J. M. and Vélez-Medel, A. 2014. Morphological development and allometric growth of yellowtail kingfish *Seriola lalandi* V. larvae under culture conditions. *Aquaculture Research*, 47(4), pp. 1277–1287.

- Martinez-Takeshita, N., Purcell, C. M., Chabot, C. L., Craig, M. T., Paterson, C. N., Hyde, J. R. and Allen, L. G. 2015. A Tale of Three Tails: Cryptic Speciation in a Globally Distributed Marine Fish of the Genus *Seriola*. *Copeia*, 103(2), pp. 357–368. doi:10.1643/CI-124-224
- Marwayana, O. N., Gold, Z., Meyer, C. P. and Barber, P. H. 2022. Environmental DNA in a global biodiversity hotspot: Lessons from coral reef fish diversity across the Indonesian archipelago. *Environmental DNA*, 4(1), pp. 222–238. doi:10.1002/edn3.257
- Mateos, E., Marinone, S. G. and Lavín, M. F. 2006. Role of tides and mixing in the formation of an anticyclonic gyre in San Pedro Mártir Basin, Gulf of California. *Deep-Sea Research Part II: Topical Studies in Oceanography*, 53(1–2), pp. 60–76. doi:10.1016/j.dsr2.2005.07.010
- Matisziw, T. C. and Murray, A. T. 2009. Connectivity change in habitat networks. *Landscape Ecology*, 24(1), pp. 89–100. doi:10.1007/s10980-008-9282-z
- McClenaghan, B., Compson, Z. G. and Hajibabaei, M. 2020. Validating metabarcoding-based biodiversity assessments with multi-species occupancy models: A case study using coastal marine eDNA. *PLoS ONE*, 15(3), pp. 1–17. doi:10.1371/journal.pone.0224119
- McElroy, M. E., Dressler, T. L., Titcomb, G. C., Wilson, E. A., Deiner, K., Dudley, T. L., Eliason, E. J., Evans, N. T., Gaines, S. D., Lafferty, K. D., Lamberti, G. A., Li, Y., Lodge, D. M., Love, M. S., Mahon, A. R., Pfrender, M. E., Renshaw, M. A., Selkoe, K. A. and Allan, E. A. 2020. Calibrating Environmental DNA Metabarcoding to Conventional Surveys for Measuring Fish Species Richness. *Frontiers in Ecology and Evolution*, 8(August), pp. 1–12. doi:10.3389/fevo.2020.00276
- McRae, B. H. 2006. Isolation by resistance. *Evolution*, 60(8), pp. 1551–1561. <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1558-5646.1975.tb00853.x>
- McRae, B. H. and Beier, P. 2007. Circuit theory predicts gene flow in plant and animal populations. *Proceedings of the National Academy of Sciences of the United States of America*, 104(50), pp. 19885–19890. doi:10.1073/pnas.0706568104
- Melià, P., Schiavina, M., Rossetto, M., Gatto, M., Fraschetti, S. and Casagrandi, R. 2016. Looking for hotspots of marine metacommunity connectivity: A methodological framework. *Scientific Reports*, Nature Publishing Group, 6(November 2015), pp. 1–11. doi:10.1038/srep23705
- Mendoza-Zepeda, M. L., Sicheritz-Pontén, T. and Thomas Gilbert, M. P. 2014. Environmental genes and genomes: Understanding the differences and challenges in the approaches and software for their analyses. *Briefings in Bioinformatics*, 16(5), pp. 745–758. doi:10.1093/bib/bbv001
- Miller, B. S. and Kendall, A. W. J. 2009. *Early life history of marine fishes*. University of California Press, United States of America.
- Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., Minamoto, T., Yamamoto, S., Yamanaka, H., Araki, H., Kondoh, M. and Iwasaki, W. 2015. MiFish, a set of universal PCR primers for

metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society Open Science*, 2(7), p. 150088. doi:10.1098/rsos.150088

- Moritz, C., Meynard, C. N., Devictor, V., Guizien, K., Labrune, C., Guarini, J. M. and Mouquet, N. 2013. Disentangling the role of connectivity, environmental filtering, and spatial structure on metacommunity dynamics. *Oikos*, 122(10), pp. 1401–1410. doi:10.1111/j.1600-0706.2013.00377.x
- Morlon, H., Chuyong, G., Condit, R., Hubbell, S., Kenfack, D., Thomas, D., Valencia, R. and Green, J. L. 2008. A general framework for the distance-decay of similarity in ecological communities. *Ecology Letters*, 11(9), pp. 904–917. doi:10.1111/j.1461-0248.2008.01202.x
- Morzaria-Luna, H. N., Cruz-Piñón, G., Brusca, R. C., López-Ortiz, A. M., Moreno-Báez, M., Reyes-Bonilla, H. and Turk-Boyer, P. 2018. Biodiversity hotspots are not congruent with conservation areas in the Gulf of California. *Biodiversity and Conservation*, 7(14), pp. 3819–3842. doi:10.1007/s10531-018-1631-x
- Mullineaux, L. S., Metaxas, A., Beaulieu, S. E., Bright, M., Gollner, S., Grupe, B. M., Herrera, S., Kellner, J. B., Levin, L. A., Mitarai, S., Neubert, M. G., Thurnherr, A. M., Tunnicliffe, V., Watanabe, H. K. and Won, Y. J. 2018. Exploring the ecology of deep-sea hydrothermal vents in a metacommunity framework. *Frontiers in Marine Science*, 4(FEB). doi:10.3389/fmars.2018.00049
- Munguía-Vega, A., Suarez-Castillo, A. N., Espinosa-Romero, M. J., Torre, J., Green, A. L., Aburto-Oropeza, O., Giron-Nava, A., Moreno-Báez, M., Cisneros-Montemayor, A. M., Cruz-Piñón, G., Reyes-Bonilla, H., Danemann, G., Gonzalez-Cuellar, O., Weaver, A. H., Lasch, C., del Mar Mancha-Cisneros, M., Marinone, S. G., Morzaria-Luna, H. N., Turk-Boyer, P. and Walther, M. 2018. Supplementary: Ecological guidelines for designing networks of marine reserves in the unique biophysical environment of the Gulf of California. *Reviews in Fish Biology and Fisheries*, 28.
- Munguía-Vega, A., Torre Cosío, J., Turk-Boyer, P. J., Marinone, S. G., Lavín, M. F., Pfister, T., Shaw, W. W., Danemann, G. D., Rainndi, P., Castillo López, A., Cinti, A., Duberstein, J. N., Moreno-Báez, M., Rojo, M., Soria, G., Sánchez Velasco, L., Morzaria-Luna, H. N., Bourillón, L., Rowell, K. and Cudney-Bueno, R. 2015. PANGAS: An interdisciplinary Ecosystem-Based research framework for small-scale fisheries in the Northern Gulf of California. *Journal of the Southwest*, 57(2–3), pp. 337–390. doi:10.1353/jsw.2015.0003
- Munguía-Vega, Adrian, Green, A. L., Suarez-Castillo, A. N., Espinosa-Romero, M. J., Aburto-Oropeza, O., Cisneros-Montemayor, A. M., Cruz-Piñón, G., Danemann, G., Giron-Nava, A., Gonzalez-Cuellar, O., Lasch, C., Mancha-Cisneros, M. del M., Marinone, S. G., Moreno-Báez, M., Morzaria-Luna, H. N., Reyes-Bonilla, H., Torre, J., Turk-Boyer, P., Walther, M. and Weaver, A. H. 2018. Ecological guidelines for designing networks of marine reserves in the unique biophysical environment of the Gulf of California. *Reviews in Fish Biology and Fisheries*, 28(4), pp. 749–776. doi:10.1007/s11160-018-9529-y
- Munguía-Vega, Adrian, Jackson, A., Marinone, S. G., Erisman, B., Moreno-Baez, M., Girón-Nava, A., Pfister, T., Aburto-Oropeza, O. and Torre, J. 2014. Asymmetric connectivity of spawning aggregations of a commercially important marine fish using a multidisciplinary approach. *PeerJ*, 2(January 2016), p. e511. doi:10.7717/peerj.511

- Munguia-Vega, A., Marinone, S. G., Paz-Garcia, D. A., Giron-Nava, A., Plomozo-Lugo, T., Gonzalez-Cuellar, O., Weaver, A. H., García-Rodríguez, F. J. and Reyes-Bonilla, H. 2018. Anisotropic larval connectivity and metapopulation structure driven by directional oceanic currents in a marine fish targeted by small-scale fisheries. *Marine Biology*, Springer Berlin Heidelberg, 165(1), pp. 1–16. doi:10.1007/s00227-017-3267-x
- Navarrete, A. H., Lagos, N. A. and Ojeda, F. P. 2014. Latitudinal diversity patterns of Chilean coastal fishes: searching for causal processes. *Revista Chilena de Historia Natural*, 87(1), pp. 1–11. doi:10.1186/0717-6317-87-2
- Newman, M. 2018. *Networks (2a)*. Oxford University Press, United Kingdom.
- Nichols, J. D., Bailey, L. L., O’Connell, A. F., Talancy, N. W., Campbell Grant, E. H., Gilbert, A. T., Annand, E. M., Husband, T. P. and Hines, J. E. 2008. Multi-scale occupancy estimation and modelling using multiple detection methods. *Journal of Applied Ecology*, 45(5), pp. 1321–1329. doi:10.1111/j.1365-2664.2008.01509.x
- Nielsen, E. E., Hemmer-Hansen, J., Larsen, P. F. and Bekkevold, D. 2009. Population genomics of marine fishes: Identifying adaptive variation in space and time. *Molecular Ecology*, 18(15), pp. 3128–3150. doi:10.1111/j.1365-294X.2009.04272.x
- Nihoul, J. (Ed.). 1973. *Modelling of marine systems (1st ed.)*. Elsevier, The Netherlands.
- O’Donnell, J. L., Kelly, R. P., Shelton, A. O., Samhouri, J. F., Lowell, N. C. and Williams, G. D. 2017. Spatial distribution of environmental DNA in a nearshore marine habitat. *PeerJ*, 5, p. e3044. doi:10.7717/peerj.3044
- O’Donnell, J. L., Kelly, R. P., Shelton, A. O., Samhouri, J. F., Lowell, N. C., Williams, G. D., O’Donnell, J. L., Kelly, R. P., Shelton, A. O., Samhouri, J. F., Lowell, N. C. and Williams, G. D. 2017. Spatial distribution of environmental DNA in a nearshore marine habitat. *PeerJ*, 5, p. e3044. doi:10.7717/peerj.3044
- Oksanen, J., Blanchet, G., Kindt, R., Legendre, P., Michin, P. R., O’hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H. and Wagner, H. 2020. *vegan* community ecology package (2.5.6). R package.
- Olds, A. D., Connolly, R. M., Pitt, K. A., Pittman, S. J., Maxwell, P. S., Huijbers, C. M., Moore, B. R., Albert, S., Rissik, D., Babcock, R. C. and Schlacher, T. A. 2016. Quantifying the conservation value of seascape connectivity: A global synthesis. *Global Ecology and Biogeography*, 25(1), pp. 3–15. doi:10.1111/geb.12388
- Oleksiak, M. F. and Rajora, O. P. 2020. Population Genomics: Marine Organisms. <http://link.springer.com/10.1007/978-3-030-37936-0>
- Olivier, D., Loiseau, N., Petatán-Ramírez, D., Millán, O. T., Suárez-Castillo, A. N., Torre, J., Munguia-Vega, A. and Reyes-Bonilla, H. 2018. Functional-biogeography of the reef fishes of the islands of the Gulf of

- California: Integrating functional divergence into marine conservation. *Global Ecology and Conservation*, 16, p. e00506. doi:10.1016/j.gecco.2018.e00506
- Palumbi, S. R. 2003. Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications*, 13(1 SUPPL.), pp. 146–158. doi:10.1890/1051-0761(2003)013[0146:pgdcat]2.0.co;2
- Papadopoulou, A., Anastasiou, I., Spagopoulou, F., Stalimerou, M., Terzopoulou, S., Legakis, A. and Vogler, A. P. 2011. Testing the species-genetic diversity correlation in the Aegean archipelago: Toward a haplotype-based macroecology? *American Naturalist*, 178(2), pp. 241–255. doi:10.1086/660828
- Paradis, E., Bolmber, S., Bolker, B., Brown, J., Claramunt, S., Claude, J. and Sien, H. 2021. 'ape' package. R package. (p. 294).
- Paris, J. R., Stevens, J. R. and Catchen, J. M. 2017. Lost in parameter space: a road map for stacks. *Methods in Ecology and Evolution*, 8(10), pp. 1360–1373. doi:10.1111/2041-210X.12775
- Parrish, R. H., Nelson, C. S., Bakun, A., Parrish, R. H., Nelson, C. S., Transport, A. B., Parrish, R. H. and Nelson, C. S. 1981. Transport Mechanisms and Reproductive Success of Fishes in the California Current. *Biological Oceanography*, 1(2), pp. 175–203. doi:10.1080/01965581.1981.10749438
- Pauly, D., Watson, R. and Alder, J. 2005. Global trends in world fisheries: Impacts on marine ecosystems and food security. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1453), pp. 5–12. doi:10.1098/rstb.2004.1574
- Pawlowski, J., Kelly-quinn, M., Altermatt, F., Apothéoz-perret-gentil, L., Beja, P., Boggero, A., Borja, A., Bouchez, A., Cordier, T., Domaizon, I., Joao, M., Filipa, A., Fornaroli, R., Graf, W., Herder, J., Hoorn, B. Van Der, Jones, J. I., Sagova-mareckova, M., Moritz, C., ... Kahlert, M. 2018. The future of biotic indices in the ecogenomic era: Integrating (e) DNA metabarcoding in biological assessment of aquatic ecosystems. *Science of the Total Environment*, Elsevier B.V., 637–638(1), pp. 1295–1310. doi:10.1016/j.scitotenv.2018.05.002
- Peguero-Icaza, M., Sánchez-Velasco, L., Lavin, M. F. and Marinone, S. G. 2008. Larval fish assemblages, environment and circulation in a semi enclosed sea (Gulf of California, Mexico). *Estuarine, Coastal and Shelf Science*, 79(1), pp. 277–288. doi:10.1016/j.ecss.2008.04.008
- Peguero-Icaza, M., Sánchez-Velasco, L., Lavín, M. F., Marinone, S. G., Beier, E. and Miguel, F. 2011. Seasonal changes in connectivity routes among larval fish assemblages in a semi-enclosed sea (Gulf of California). *Journal of Plankton Research*, 33(3), pp. 517–533. doi:10.1093/plankt/fbq107
- Pérez Olivas, A. 2016. Biología reproductiva de la cabrilla sardinera (*Mycteroperca rosacea*, Streets 1877) en la zona costera de Santa Rosalía, BCS, México. *CICIMAR*, pp. 73 .
- Petatán, D. 2015. Propuesta de zonación del Golfo de California con base en variables oceanográficas y distribución de macroinvertebrados . *UABCS* . pp. 106 . doi:10.1145/3132847.3132886

- Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S. and Hoekstra, H. E. 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS ONE*, 7(5). doi:10.1371/journal.pone.0037135
- Pfeiler, E., Hurtado, L. A., Knowles, L. L., Torre-Cosío, J., Bourillón-Moreno, L., Márquez-Farías, J. F. and Montemayor-López, G. 2005. Population genetics of the swimming crab *Callinectes bellicosus* (Brachyura: Portunidae) from the eastern Pacific Ocean. *Marine Biology*, 146(3), pp. 559–569. doi:10.1007/s00227-004-1463-y
- Pfeiler, E., Watts, T., Pugh, J. and Van Der Heiden, A. M. 2008. Speciation and demographic history of the Cortez bonefish, *Albula* sp. a (Albuliformes: Albulidae), in the Gulf of California inferred from mitochondrial DNA. *Journal of Fish Biology*, 73(2), pp. 382–394. doi:10.1111/j.1095-8649.2008.01892.x
- Pineda, J., Hare, J. and Sponaugle, S. 2007. Larval transport and dispersal in the coastal ocean and consequences for population connectivity. *Oceanography*, 20(3), pp. 22–39.
- Piñol, J., Mir, G., Gomez-Polo, P. and Agustí, N. 2015. Universal and blocking primer mismatches limit the use of high-throughput DNA sequencing for the quantitative metabarcoding of arthropods. *Molecular Ecology Resources*, 15(4), pp. 819–830. doi:10.1111/1755-0998.12355
- Pittman, S. J. 2018. *Seascape ecology* (S. J. Pittman (Ed.)). John Wiley and Sons.
- Pittman, S. J. and Olds, A. D. 2015. Seascape ecology of fishes on coral reefs. *Ecology of Fishes on Coral Reefs*, May, pp. 274–282. doi:10.1017/CBO9781316105412.036
- Port, J. A., O'Donnell, J. L., Romero-Maraccini, O. C., Leary, P. R., Litvin, S. Y., Nickols, K. J., Yamahara, K. M. and Kelly, R. P. 2016. Assessing vertebrate biodiversity in a kelp forest ecosystem using environmental DNA. *Molecular Ecology*, 25(2), pp. 527–541. doi:10.1111/mec.13481
- Préau, C., Dubos, N., Lenormand, M., Denelle, P., Le Louarn, M., Alleaume, S. and Luque, S. 2021. Dispersal-based species pools as sources of connectivity area mismatches. *Landscape Ecology*, 0123456789. doi:10.1007/s10980-021-01371-y
- Puebla, O., Bermingham, E. and McMillan, W. O. 2012. On the spatial scale of dispersal in coral reef fishes. *Molecular Ecology*, 21(23), pp. 5675–5688. doi:10.1111/j.1365-294X.2012.05734.x
- R Core Team. 2021. R: A language and environment for statistical computing. R Foundation for Statistical Computing. <https://www.r-project.org/>
- Ramírez-Ortiz, G., Calderon-Aguilera, L. E., Reyes-Bonilla, H., Ayala-Bocos, A., Hernández, L., Fernández Rivera-Melo, F., López-Pérez, A. and Dominici-Arosamena, A. 2017. Functional diversity of fish and invertebrates in coral and rocky reefs of the Eastern Tropical Pacific. *Marine Ecology*, 38(4). doi:10.1111/maec.12447

- Reguera-Rouzaud, N., Díaz-Viloria, N., Pérez-Enríquez, R., Espino-Barr, E., Rivera-Lucero, M. I. and Munguía-Vega, A. 2021. Drivers for genetic structure at different geographic scales for Pacific red snapper (*Lutjanus peru*) and yellow snapper (*Lutjanus argentiventris*) in the tropical eastern Pacific. *Journal of Fish Biology*, 98(5), pp. 1267–1280. doi:10.1111/jfb.14656
- Reguera-Rouzaud, N., Díaz-Viloria, N., Sánchez-Velasco, L., Flores-Morales, A. L., Parés-Sierra, A., Aburto-Oropeza, O. and Munguía-Vega, A. 2020. Yellow snapper (*Lutjanus argentiventris*) connectivity in the Southern Gulf of California. *Marine Biodiversity*, 50(4). doi:10.1007/s12526-020-01070-y
- Resasco, J. 2019. Meta-analysis on a Decade of Testing Corridor Efficacy: What New Have we Learned? *Current Landscape Ecology Reports*, 4(3), pp. 61–69. doi:10.1007/s40823-019-00041-9
- Riginos, C and Victor, B. C. 2001. Larval spatial distributions and other early life-history characteristics predict genetic differentiation in eastern Pacific blennioid fishes. *Proceedings of the Royal Society B*, 268(1479), pp. 1931–1936. doi:10.1098/rspb.2001.1748
- Riginos, Cynthia, Crandall, E. D., Liggins, L., Bongaerts, P. and Trembl, E. A. 2016. Navigating the currents of seascape genomics: how spatial analyses can augment population genomic studies. *Current Zoology*, in press(July), p. doi: 10.1093/cz/zow067. doi:10.1093/cz/zow067
- Riginos, Cynthia, Douglas, K. E., Jin, Y., Shanahan, D. F. and Trembl, E. A. 2011. Effects of geography and life history traits on genetic differentiation in benthic marine fishes. *Ecography*, 34(4), pp. 566–575. doi:10.1111/j.1600-0587.2010.06511.x
- Riginos, Cynthia, Hock, K., Matias, A. M., Mumby, P. J., van Oppen, M. J. H. and Lukoschek, V. 2019. Asymmetric dispersal is a critical element of concordance between biophysical dispersal models and spatial genetic structure in Great Barrier Reef corals. *Diversity and Distributions*, 25(11), pp. 1684–1696. doi:10.1111/ddi.12969
- Riginos, Cynthia and Liggins, L. 2013. Seascape Genetics: Populations, Individuals, and Genes Marooned and Adrift. *Geography Compass*, 7(3), pp. 197–216. doi:10.1111/gec3.12032
- Ripa, P. 1990. Seasonal circulation in the Gulf of California. *Annales Geophysicae*, 8, p. 559.564.
- Ripa, P. 1997. Toward a physical explanation of the seasonal dynamics and thermodynamics of the Gulf of California. *Journal of Physical Oceanography*, 27(5), pp. 597–614. doi:10.1175/1520-0485(1997)027<0597:TAPEOT>2.0.CO;2
- Roberts, C. M., McClean, C. J., Veron, J. E. N., Hawkins, J. P., Allen, G. R., McAllister, D. E., Mittermeier, C. G., Schueler, F. W., Spalding, M., Wells, F., Vynne, C. and Werner, T. B. 2002. Marine biodiversity hotspots and conservation priorities for tropical reefs. *Science*, 295(5558), pp. 1280–1284. doi:10.1126/science.1067728

- Rodríguez-Zárate, C. J., Sandoval-Castillo, J., van Sebille, E., Keane, R. G., Rocha-Olivares, A., Urteaga, J. and Beheregaray, L. B. 2018. Isolation by environment in the highly mobile olive ridley turtle (*Lepidochelys olivacea*) in the eastern pacific. *Proceedings of the Royal Society B: Biological Sciences*, 285(1878). doi:10.1098/rspb.2018.0264
- Rognes, T., Flouri, T., Nichols, B., Quince, C. and Mahé, F. 2016. Vsearch: A versatile open source tool for metagenomics. *PeerJ*, 2016(10), pp. 1–22. doi:10.7717/peerj.2584
- Rousset, F. 1997. Genetic differentiation and estimation of gene flow from F statistics under isolation by distance. *Genetics*, 145(April), pp. 1219–1228. doi:10.1177/000992287301200933
- Sadovy, de M. Y., and Colin, P. L. (Eds.). 2012. Reef fish spawning aggregations: Biology, research and management (1st ed.). Springer, United States of America.
- Saenz-Agudelo, P., Jones, G. P., Thorrold, S. R. and Planes, S. 2011. Connectivity dominates larval replenishment in a coastal reef fish metapopulation. *Proceedings of the Royal Society B: Biological Sciences*, 278(1720), pp. 2954–2961. doi:10.1098/rspb.2010.2780
- Sala, E., Aburto-Oropeza, O., Paredes, G. and Thompson, G. 2003. Spawning aggregations and reproductive behavior of reef fishes in the Gulf of California. *Bulletin of Marine Science*, 72(1), pp. 103–121. doi:10.1126/science.1075284
- Sala, E. and Aburto-Oropeza, O. 2002. A general model for designing networks of marine reserves. *Science*, 298(1991). doi:10.1126/science.1075284
- Sánchez-Caballero, C. A., Borges-Souza, J. M., De La Cruz-Agüero, G. and Ferse, S. C. A. 2017. Links between fish community structure and habitat complexity of a rocky reef in the Gulf of California threatened by development: Implications for mitigation measures. *Ocean and Coastal Management*, 137, pp. 96–106. doi:10.1016/j.ocecoaman.2016.12.013
- Sánchez-Rodríguez, A., Aburto-Oropeza, O., Erisman, B., Jiménez-Esquivel, V. M. and Hinojosa-Arango, G. 2015. Rocky Reefs: Preserving Biodiversity for the Benefit of the Communities in the Aquarium of the World. *Ethnobiology of Corals and Coral Reefs*. Springer International Publishing, Switzerland, 1st ed., pp. 177–208 .
- Sánchez-Velasco, L., Lavín, M. F., Jiménez-Rosenberg, S. P. A., Godínez, V. M., Santamaría-del-Angel, E. and Hernández-Becerril, D. U. 2013. Three-dimensional distribution of fish larvae in a cyclonic eddy in the Gulf of California during the summer. *Deep-Sea Research Part I: Oceanographic Research Papers*, 75, pp. 39–51. doi:10.1016/j.dsr.2013.01.009
- Sánchez-Velasco, L., Lavín, M. F., Peguero-Icaza, M., León-Chávez, C. A., Contreras-Catala, F., Marinone, S. G., Gutiérrez-Palacios, I. V. and Godínez, V. M. 2009. Seasonal changes in larval fish assemblages in a semi-enclosed sea (Gulf of California). *Continental Shelf Research*, 29(14), pp. 1697–1710. doi:10.1016/j.csr.2009.06.001

- Sanchez, A. 2015. Ethnobiology of Corals and Coral Reefs (Issue April). doi:10.1007/978-3-319-23763-3
- Sandoval-Castillo, J., Robinson, N. A., Hart, A. M., Strain, L. W. S. and Beheregaray, L. B. 2018. Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (*Haliotis laevis*), along a longitudinal environmental gradient. *Molecular Ecology*, December 2017, pp. 1–18. doi:10.1111/mec.14526
- Santiago-Garcia, M. W., Marinone, S. G. and Velasco-Fuentes, O. U. 2014. Three-dimensional connectivity in the Gulf of California based on a numerical model. *Progress in Oceanography*, 123, pp. 64–73. doi:10.1016/j.pocean.2014.02.002
- Sanvicente-Añorve, L., Sánchez-Ramírez, M., Ocaña-Luna, A., Flores-Coto, C. and Ordóñez-López, U. 2011. Metacomunity structure of estuarine fish larvae: The role of regional and local processes. *Journal of Plankton Research*, 33(1), pp. 179–194. doi:10.1093/plankt/fbq098
- Savary, P., Foltête, J. C., Moal, H., Vuidel, G. and Garnier, S. 2021. *graph4lg*: A package for constructing and analysing graphs for landscape genetics in R. *Methods in Ecology and Evolution*, 12(3), pp. 539–547. doi:10.1111/2041-210X.13530
- Schietteakte, N. M. D., Brandl, S. J. and Casey, J. M. 2019. fishualize: Color Palettes Based on Fish Species (0.1.0; p. 9). <https://cran.r-project.org/package=fishualize>
- Schill, S. R., Raber, G. T., Roberts, J. J., Tremblay, E. A., Brenner, J. and Halpin, P. N. 2015. No Reef Is an Island: Integrating Coral Reef Connectivity Data into the Design of Regional-Scale Marine Protected Area Networks. *PLoS One*, 10(12), pp. 1–24. doi:10.1371/journal.pone.0144199
- Schmelzle, M. C. and Kinziger, A. P. 2016. Using occupancy modelling to compare environmental DNA to traditional field methods for regional-scale monitoring of an endangered aquatic species. *Molecular Ecology Resources*, 16(4), pp. 895–908. doi:10.1111/1755-0998.12501
- Schmidt, B. R., Kéry, M., Ursenbacher, S., Hyman, O. J. and Collins, J. P. 2013. Site occupancy models in the analysis of environmental DNA presence/absence surveys: a case study of an emerging amphibian pathogen. *Methods in Ecology and Evolution*, John Wiley & Sons, Ltd, 4(7), pp. 646–653. doi:10.1111/2041-210X.12052
- Schunter, C., Pascual, M., Garza, J. C., Raventos, N. and Macpherson, E. 2014. Kinship analyses identify fish dispersal events on a temperate coastline. *Proceedings of the Royal Society B: Biological Sciences*, 281(1785). doi:10.1098/rspb.2014.0556
- Selkoe, K. A., D'Aloia, C. C., Crandall, E. D., Iacchi, M., Liggins, L., Puritz, J. B., Von Der Heyden, S. and Toonen, R. J. 2016. A decade of seascape genetics: Contributions to basic and applied marine connectivity. *Marine Ecology Progress Series*, 554(August), pp. 1–19. doi:10.3354/meps11792
- Selkoe, K. A., Henzler, C. M. and Gaines, S. D. 2008. Seascape genetics and the spatial ecology of marine populations. *Fish and Fisheries*, 9, pp. 363–377.

- Selkoe, K. A., Watson, J. R., White, C., Horin, T. Ben, Iacchi, M., Mitarai, S., Siegel, D. a, Gaines, S. D. and 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(17), pp. 3708–3726. doi:10.1111/j.1365-294X.2010.04658.x
- Selmoni, O., Lecellier, G., Magalon, H., Vigliola, L., Oury, N., Benzoni, F., Peignon, C., Joost, S. and Berteaux-Lecellier, V. 2021. Seascape genomics reveals candidate molecular targets of heat stress adaptation in three coral species. *Molecular Ecology*, 30(8), pp. 1892–1906. doi:10.1111/mec.15857
- Selwyn, J. D., Hogan, J. D., Downey-Wall, A. M., Gurski, L. M., Portnoy, D. S. and Heath, D. D. 2016. Kin-aggregations explain chaotic genetic patchiness, a commonly observed genetic pattern, in a marine fish. *PLoS ONE*, 11(4), pp. 1–11. doi:10.1371/journal.pone.0153381
- Oka, S., Doi, H., Miyamoto, K., Hanahara, N., Sado, T. and Masaki, M. 2020. Environmental DNA metabarcoding for biodiversity monitoring of a highly-diverse tropical fish community in a coral-reef lagoon: Estimation of species richness and detection of habitat segregation. *BioRxiv*, May, pp. 1–27. doi:10.1002/edn3.132
- Shu, L., Ludwig, A. and Peng, Z. 2020. Standards for methods utilizing environmental DNA for detection of fish species. *Genes*, 11(3). doi:10.3390/genes11030296
- Slatkin, M. 1989. Detecting small amounts of gene flow from phylogenies of alleles. pp. 3.
- Slatkin, M. 1994. Gene flow and population structure. L. A. Real (Ed.), *Ecological Genetics*. Princeton University Press, Princeton, New Jersey, 1 st, pp. 3–17 .
- Soria, G., Munguia-Vega, A., Marinone, S. G., Moreno-Báez, M., Martínez-Tovar, I. and Cudney-Bueno, R. 2012. Linking bio-oceanography and population genetics to assess larval connectivity. *Marine Ecology Progress Series*, 463, pp. 159–175. doi:10.3354/meps09866
- Soria, G., Torre-Cosío, J., Munguia-Vega, A., Marinone, S. G., Lavinn, M. F., Cinti, A. and Moreno-Baez, M. 2014. Dynamic connectivity patterns from an insular marine protected area in the Gulf of California. *Journal of Marine Systems*, The Authors, 129, pp. 248–258. doi:10.1016/j.jmarsys.2013.06.012
- Soto-Mardones, L., Marinone, S. G. and Parés-Sierra, A. 1999. Variabilidad espaciotemporal de la temperatura superficial del mar en el Golfo de California. *Ciencias Marinas*, 25, pp. 1–30.
- Stat, M., John, J., DiBattista, J. D., Newman, S. J., Bunce, M. and Harvey, E. S. 2019. Combined use of eDNA metabarcoding and video surveillance for the assessment of fish biodiversity. *Conservation Biology*, 33(1), pp. 196–205. doi:10.1111/cobi.13183
- Stuart, C. E., Wedding, L. M., Pittman, S. J. and Green, S. J. 2021. Habitat suitability modeling to inform seascape connectivity conservation and management. *Diversity*, 13(465), pp. 1–20.

- Taberlet, P., Bonin, A., Zinger, L. and Coissac, E. 2018. Environmental DNA For Biodiversity Research and Monitoring (1st ed.). Oxford University Press, United Kingdom.
- Taberlet, P., Coissac, E., Pompanon, F., Brochmann, C. and Willerslev, E. 2012. Towards next-generation biodiversity assessment using DNA metabarcoding. *Molecular Ecology*, 21(8), pp. 2045–2050. doi:10.1111/j.1365-294X.2012.05470.x
- Tamura, K., Stecher, G., Peterson, D., Filipowski, A. and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*, 30(12), pp. 2725–2729. doi:10.1093/molbev/mst197
- Taylor, P. D., Fahrig, L., Henein, K. and Merriam, G. 1993. Connectivity is a vital element of landscape structure. *Oikos*, 68(3), pp. 571–573. <http://www.jstor.org/stable/3544927>
- Teske, P. R., Sandoval-Castillo, J., Golla, T. R., Emami-Khoyi, A., Tine, M., Von Der Heyden, S. and Beheregaray, L. B. 2019. Thermal selection as a driver of marine ecological speciation. *Proceedings of the Royal Society B: Biological Sciences*, 286(1896). doi:10.1098/rspb.2018.2023
- Thomsen, P. F., Kielgast, J., Iversen, L. L., Moller, P. R., Rasmussen, M. and Willerslev, E. 2012. Detection of a Diverse Marine Fish Fauna Using Environmental DNA from Seawater Samples. *PLoS ONE*, 7(8), pp. 1–9. doi:10.1371/journal.pone.0041732
- Thomson, D A, Findley, L. and Kerstitch, A. 2000. Reef fishes of the Sea of Cortez (1st ed.). University of Texas, USA.
- Thomson, Donald A., Findley, L. T. and Kerstitch, A. 2000. Reef fishes of the Sea of Cortez (Donald A Thomson, L. T. Findley, and A. Kerstitch (Eds.); 1st ed.). University of Texas, USA.
- Tingley, M. W., Nadeau, C. P. and Sandor, M. E. 2020. Multi species occupancy models as robust estimators of community richness. *Methods in Ecology and Evolution*, pp. 0–2. doi:10.1111/2041-210x.13378
- Tingley, R., Coleman, R., Gecse, N., van Rooyen, A. and Weeks, A. 2020. Accounting for false positive detections in occupancy studies based on environmental DNA: A case study of a threatened freshwater fish (*Galaxiella pusilla*) . *Environmental DNA*, 3(July), pp. 1–10. doi:10.1002/edn3.124
- Tornero, I., Boix, D., Bagella, S., Pinto-Cruz, C., Caria, M. C., Belo, A., Lumbreras, A., Sala, J., Compte, J. and Gascón, S. 2018. Dispersal mode and spatial extent influence distance-decay patterns in pond metacommunities. *PLoS ONE*, 13(8), pp. 1–17. doi:10.1371/journal.pone.0203119
- Treml, E. A., Halpin, P. N., Urban, D. L. and Pratson, L. F. 2008. Modeling population connectivity by ocean currents, a graph-theoretic approach for marine conservation. *Landscape Ecology*, 23(SUPPL. 1), pp. 19–36. doi:10.1007/s10980-007-9138-y

- Tuomisto, H. and Ruokolainen, K. 2006. Analyzing or Explaining Beta Diversity? Understanding the Targets of Different Methods of Analysis. *Ecology*, 87(11), pp. 2697–2708.
- Turk-Boyer, P. J., Morzaria-Luna, H., Martínez-Tovar, I., Downton-Hoffman, C. and Munguia-Vega, A. 2014. Ecosystem-Based Fisheries Management of a Biological Corridor Along the Northern Sonora Coastline (NE Gulf of California). F. Amezcua and B. Bellgraph (Eds.), *Fisheries Management of Mexican and Central American Estuaries*. Springer Science+Business Media, 1st ed., Issue February 2015. pp. 125–154 . doi:10.1007/978-94-017-8917-2
- Turner, M. G., Gardner, R. H. and O'Neill, R. V. 2001. *Landscape Ecology in Theory and Practice*. Pattern and Process. Springer. doi:10.1007/b97434
- Ulate, K., Sánchez, C., Sánchez-Rodríguez, A., Alonso, D., Aburto-Oropeza, O. and Huato-Soberanis, L. 2016. Latitudinal regionalization of epibenthic macroinvertebrate communities on rocky reefs in the Gulf of California. *Marine Biology Research*, Taylor & Francis, 12(4), pp. 389–401. doi:10.1080/17451000.2016.1143105
- Urban, D. and Keitt, T. 2001. Landscape connectivity: a graph-theoretic perspective. *Ecology*, 82(5), pp. 1205–1218. doi:10.1002/sml.200801002
- Urban, D. L., Minor, E. S., Treml, E. A., Schick, R. S., Robert, S. and Schick, R. S. 2009. Graph models of habitat mosaics. *Ecology Letters*, 12(3), pp. 260–273. doi:10.1111/j.1461-0248.2008.01271.x
- Uroy, L., Alignier, A., Mony, C., Foltête, J. C. and Ernoult, A. 2021. How to assess the temporal dynamics of landscape connectivity in ever-changing landscapes: a literature review. *Landscape Ecology*, 9, pp. 2487–2504. doi:10.1007/s10980-021-01277-9
- Valdivia-Carrillo, T., Rocha-Olivares, A., Reyes - Bonilla, H., Dominguez-Contreras, J. F., Munguía-Vega, A., Valdivia - Carrillo, T., Rocha - Olivares, A., Reyes - Bonilla, H., Francisco Domínguez - Contreras, J. and Munguia - Vega, A. 2021. Integrating eDNA metabarcoding and simultaneous underwater visual surveys to describe complex fish communities in a marine biodiversity hotspot. *Molecular Ecology Resources*, February, pp. 1–17. doi:10.1111/1755-0998.13375
- Valentini, A., Taberlet, P., Miaud, C., Civade, R. R., Herder, J., Thomsen, P. F., Bellemain, E., Besnard, A. A., Coissac, E., Boyer, F. F., Gaboriaud, C., Jean, P., Poulet, N., Roset, N., Copp, G. H., Geniez, P., Pont, D., Argillier, C., Baudoin, J. M., ... Dejean, T. 2016. Next-generation monitoring of aquatic biodiversity using environmental DNA metabarcoding. *Molecular Ecology*, 25(4), pp. 929–942. doi:10.1111/mec.13428
- Vellend, M. 2010. Conceptual synthesis in community ecology. *Quarterly Review of Biology*, 85(2), pp. 183–206. doi:10.1086/652373
- Viesca-Lobatón, C., Balart, E. F., Mascareñas-Osorio, I., Reyes-Bonilla, H. and Torreblanca, E. 2008. Los peces de arrecife de bahía de los Ángeles, Golfo de California. G. Danemann and E. Ezcurra (Eds.),

Bahía de Los Ángeles: recursos naturales y comunidad. Instituto Nacional de Ecología, Mexico City, 1st ed., pp. 385–428 .

- Virgilio, M., Backeljau, T., Nevado, B. and De Meyer, M. 2010. Comparative performances of DNA barcoding across insect orders. *BMC Bioinformatics*, 11. doi:10.1186/1471-2105-11-206
- Virtanen, E. A., Moilanen, A. and Viitasalo, M. 2020. Marine connectivity in spatial conservation planning: analogues from the terrestrial realm. *Landscape Ecology*, Springer Netherlands, 35(5), pp. 1021–1034. doi:10.1007/s10980-020-00997-8
- Wagner, H. H. and Fortin, M.-J. 2013. A conceptual framework for the spatial analysis of landscape genetic data. *Conservation Genetics*, 14(2), pp. 253–261. doi:10.1007/s10592-012-0391-5
- Wang, I. J. and Bradburd, G. S. 2014. Isolation by environment. *Molecular Ecology*, 23(23), pp. 5649–5662. doi:10.1111/mec.12938
- Wangensteen, O. S. 2019. Metabarpark. R Scripts for Reformatting Metabarcoding Databases. [https://github.com/metabarpark/R\\_scripts\\_metabarpark](https://github.com/metabarpark/R_scripts_metabarpark)
- Waples, R. S., Gaggiotti, O. and Review, I. 2006. What is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Molecular Ecology*, 15(6), pp. 1419–1439. doi:10.1111/j.1365-294X.2006.02890.x
- Watson, J. R., Mitarai, S., Siegel, D. A., Caselle, J. E., Dong, C. and McWilliams, J. C. 2010. Realized and potential larval connectivity in the southern California bight. *Marine Ecology Progress Series*, 401, pp. 31–48. doi:10.3354/meps08376
- Watson, James R., Hays, G. C., Raimondi, P. T., Mitarai, S., Dong, C., McWilliams, J. C., Blanchette, C. A., Caselle, J. E. and Siegel, D. A. 2011. Currents connecting communities: nearshore community similarity and ocean circulation. *Ecology*, 92(6), pp. 1193–1200.
- Watson, W. and Brogan, M. W. 1996. Lutjanidae: Snappers. H. G. Moser (Ed.), *The Early Life Stages of Fishes in the California Current Region*. Allen Press, Lawrence, Kansas, pp. 977–989 .
- Wedding, L. M., Christopher, L. A., Pittman, S. J., Friedlander, A. M. and Jorgensen, S. 2011. Quantifying seascape structure: Extending terrestrial spatial pattern metrics to the marine realm. *Marine Ecology Progress Series*, 427(Gustafson 1998), pp. 219–232. doi:10.3354/meps09119
- Weersing, K. and Toonen, R. J. 2009. Population genetics, larval dispersal, and connectivity in marine systems. *Marine Ecology Progress Series*, 393(May 2014), pp. 1–12. doi:10.3354/meps08287
- Weigand, H., Beermann, A. J., Čiampor, F., Costa, F. O., Csabai, Z., Duarte, S., Geiger, M. F., Grabowski, M., Rimet, F., Rulik, B., Strand, M., Szucsich, N., Weigand, A. M., Willassen, E., Wyler, S. A., Bouchez, A., Borja, A., Čiamporová-Zat'ovičová, Z., Ferreira, S., ... Ekrem, T. 2019. DNA barcode reference libraries

for the monitoring of aquatic biota in Europe: Gap-analysis and recommendations for future work. *BioRxiv*, p. 576553. doi:10.1101/576553

- West, K. M., Stat, M., Harvey, E. S., Skepper, C. L., DiBattista, J. D., Richards, Z. T., Travers, M. J., Newman, S. J. and Bunce, M. 2020. eDNA metabarcoding survey reveals fine-scale coral reef community variation across a remote, tropical island ecosystem. *Molecular Ecology*, 29(6), pp. 1069–1086. doi:10.1111/mec.15382
- West, K., Travers, M. J., Stat, M., Harvey, E. S., Richards, Z. T., DiBattista, J. D., Newman, S. J., Harry, A., Skepper, C. L., Heydenrych, M. and Bunce, M. 2021. Large-scale eDNA metabarcoding survey reveals marine biogeographic break and transitions over tropical north-western Australia. *Diversity and Distributions*, July 2020, pp. 1–16. doi:10.1111/ddi.13228
- White, C., Selkoe, K. A., Watson, J., Siegel, D. A., Zacherl, D. C. and Toonen, R. J. 2010. Ocean currents help explain population genetic structure. *Proceedings The Royal Society*, 277(1688), pp. 1685–1694. doi:10.1098/rspb.2009.2214
- Whitlock, M. C and McCauley, D. E. 1999. Indirect measures of gene flow and migration:  $F_{ST}$  not equal to  $1/(4Nm+1)$ . *Heredity*, 82 (Pt2) (November 1998), pp. 117–125. <http://www.ncbi.nlm.nih.gov/pubmed/10098262>
- Whitlock, Michael C. and Lotterhos, K. E. 2015. Reliable detection of loci responsible for local adaptation: Inference of a null model through trimming the distribution of  $F_{ST}$ . *American Naturalist*, 186(October), pp. S24–S36. doi:10.1086/682949
- Whittaker, R. H. 1960. Vegetation of the Siskiyou mountains, Oregon and California. *Ecological Monographs*, 30(1), pp. 279–338.
- Wilcox, T. M., Zarn, K. E., Piggott, M. P., Young, M. K., McKelvey, K. S. and Schwartz, M. K. 2018. Capture enrichment of aquatic environmental DNA: A first proof of concept. *Molecular Ecology Resources*, 18(6), pp. 1392–1401. doi:10.1111/1755-0998.12928
- Willoughby, J. R., Wijayawardena, B. K., Sundaram, M., Swihart, R. K. and DeWoody, J. A. 2016. The importance of including imperfect detection models in eDNA experimental design. *Molecular Ecology Resources*, 16(4), pp. 837–844. doi:10.1111/1755-0998.12531
- Wood, S. A., Pochon, X., Laroche, O., Ammon, U. von, Adamson, J. and Zaiko, A. 2019. A comparison of droplet digital polymerase chain reaction (PCR), quantitative PCR and metabarcoding for species - specific detection in environmental DNA. *Molecular Ecology Resources*, 19(6), pp. 1407–1419.
- Wright, S. 1931. Evolution in Mendelian Populations. *Genetics*, 16(2), pp. 97–159. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1201091&tool=pmcentrez&rendertype=abstract>

- Wright, Sewall. 1943. Isolation by distance. *Genetics*, 28(March), pp. 114–138. <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1209196&tool=pmcentrez&rendertype=abstract>
- Wright, Sewall. 1946. Isolation by distance under diverse systems of mating. *Genetics*, 31(January), pp. 39–59. doi:10.1093/genetics/31.1.39
- Xuereb, A., Benestan, L., Normandeau, É., Daigle, R. M., Curtis, J. M. R., Bernatchez, L. and Fortin, M. J. 2018. Asymmetric oceanographic processes mediate connectivity and population genetic structure, as revealed by RADseq, in a highly dispersive marine invertebrate (*Parastichopus californicus*). *Molecular Ecology*, 27(10), pp. 2347–2364. doi:10.1111/mec.14589
- Yamamoto, S., Masuda, R., Sato, Y., Sado, T., Araki, H., Kondoh, M., Minamoto, T. and Miya, M. 2017. Environmental DNA metabarcoding reveals local fish communities in a species-rich coastal sea. *Scientific Reports*, 7(January), p. 40368. doi:10.1038/srep40368
- Zamudio, L., Hogan, P. and Metzger, E. J. 2008. Summer generation of the Southern Gulf of California eddy train. *Journal of Geophysical Research: Oceans*, 113(6), pp. 1–21. doi:10.1029/2007JC004467

## Annexes

## Annex A

Article published in *Molecular Ecology Resources*

<https://doi.org/10.1111/1755-0998.13375>



Received: 17 October 2019 | Revised: 13 February 2021 | Accepted: 2 March 2021  
DOI: 10.1111/1755-0998.13375

RESOURCE ARTICLE

MOLECULAR ECOLOGY  
RESOURCES WILEY

## Integrating eDNA metabarcoding and simultaneous underwater visual surveys to describe complex fish communities in a marine biodiversity hotspot

Tania Valdivia-Carrillo<sup>1,4</sup> | Axayácatl Rocha-Olivares<sup>1</sup> | Héctor Reyes-Bonilla<sup>2</sup> | José Francisco Domínguez-Contreras<sup>2</sup> | Adrian Munguia-Vega<sup>3,4</sup>

<sup>1</sup>Laboratorio de Ecología Molecular, Departamento de Oceanografía Biológica, Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE), Ensenada, Baja California, México

<sup>2</sup>Laboratorio de Sistemas Arrecifales, Universidad Autónoma de Baja California Sur (UABCS), La Paz, Baja California Sur, México

<sup>3</sup>Conservation Genetics Laboratory & Desert Laboratory on Tumamoc Hill, The University of Arizona, Tucson, AZ, USA

<sup>4</sup>Lab Applied Genomics, La Paz, Baja California Sur, México

### Correspondence

Adrian Munguia-Vega, Conservation Genetics Laboratory & Desert Laboratory on Tumamoc Hill, The University of Arizona, 1311 East 4th Street, BSE-317, Tucson, 85721 AZ, USA.  
Email: airdrian@email.arizona.edu

### Funding information

The University of Arizona - CONACYT, Cazmex Consortium; David and Lucille Packard Foundation, Grant/Award Number: 2015-62798; COBI (Comunidad y Biodiversidad A.C.); Langebio - Illumina - Biotech del Norte; Consejo Nacional de Ciencia y Tecnología, Grant/Award Number: 292/2016; Ecology Project International; Pronatura Noroeste A.C.; Nature Conservancy

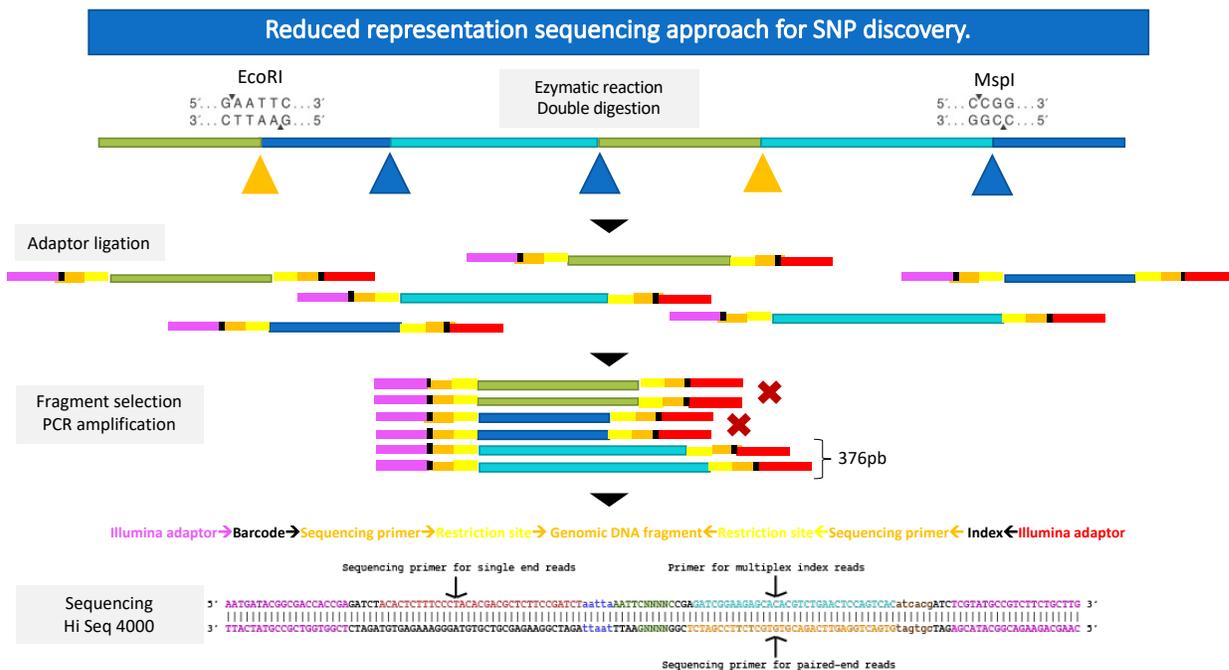
### Abstract

Marine biodiversity can be surveyed using underwater visual censuses and recently with eDNA metabarcoding. Although a promising tool, eDNA studies have shown contrasting results related to its detection scale and the number of species identified compared to other survey methods. Also, its accuracy relies on complete reference databases used for taxonomic assignment and, as other survey methods, species detection may show false-negative and false-positive errors. Here, we compared results from underwater visual censuses and simultaneous eDNA metabarcoding fish surveys in terms of observed species and community composition. We also assess the effect of a custom reference database in the taxonomic assignment, and evaluate occupancy, capture and detection probabilities, as well as error rates of eDNA survey data. We amplified a 12S rRNA fish barcode from 24 sampling sites in the gulf of California. More species were detected with eDNA metabarcoding than with UVC. Because each survey method largely detected different sets of species, the combined approach doubled the number of species registered. Both survey methods recovered a known biodiversity gradient and a biogeographic break, but eDNA captured diversity over a broader geographic and bathymetric scale. Furthermore, the use of a modest-sized custom reference database significantly increased taxonomic assignment. In a subset of species, occupancy models revealed eDNA surveys provided similar or higher detection probabilities compared to UVC. The occupancy value of each species had a large influence on eDNA detectability, and in the false positive and negative error. Overall, these results highlight the potential of eDNA metabarcoding in complementing other established ecological methods for studies of marine fishes.

### KEYWORDS

biodiversity monitoring, detection probability, eDNA metabarcoding, fish community, occupancy model, reference database, underwater visual census

## Annex B



**Figure 39.** Description of the ddRAD protocol for SNP discovery.

**Table 18.** 12S rRNA primers designed in this study and used for custom reference database of teleost fishes from the Gulf of California. The amplicon includes the “teleo” barcode from Valentini et al. 2016. A: Amplicon length (bp); B: %GC; C: Hair pin Tm (°C); D: Pair Dimer Tm (°C); E: Self Dimer Tm (°C); F: Tm (°C).

Name	A	Sequence	B	C	D	E	F
<b>1,322 R</b>	640	CTTTCAGCTTTCCCTTGCGG	55	None	None	None	59.8
<b>682 F</b>		CGTTCAACCTCACCTTCTCT	55	None	None	None	59.6
<b>792 F</b>	530	CAGGTCGAGGTGTAGCGYATG	60	None	None	11.5	60.3 - 63.4

**Table 19.** List of species in the custom reference database and assigned GenBank accession numbers.

RefDB	Species	Accession number
1	<i>Abudefduf troschelli</i>	MK902806, MK902816
2	<i>Acanthocybium solandri</i>	MK902889
3	<i>Acanthurus xanthopterus</i>	MK902826
4	<i>Anisotremus interruptus</i>	MK902835, MK902845, MK902894
5	<i>Atractoscion nobilis</i>	MK902914
6	<i>Balistes polylepis</i>	MK902859
7	<i>Bodianus diplotaenia</i>	MK902867
8	<i>Chaetodon humeralis</i>	MK902807, MK902817
9	<i>Chanos chanos</i>	MK902919
10	<i>Cirrhitus rivulatus</i>	MK902836
11	<i>Coryphaena hippurus</i>	MK902902, MK902913
12	<i>Cynoscion reticulatus</i>	MK902846
13	<i>Cynoscion xanthulus</i>	MK902915
14	<i>Diapterus brevirostris</i>	MK902906
15	<i>Elacatinus puncticulatus</i>	MK902879, MK902886
16	<i>Epinephelus acanthistius</i>	MK902884
17	<i>Epinephelus labriformis</i>	MK902907
18	<i>Gnathanodon speciosus</i>	MK902899
19	<i>Haemulon sexfasciatum</i>	MK902852, MK902860
20	<i>Haemulopsis leuciscus</i>	MK902868
21	<i>Holacanthus passer</i>	MK902808, MK902818
22	<i>Hoplopagrus guentheri</i>	MK902827, MK902837
23	<i>Hypopgthalamichthys molitrix</i>	MK902871
24	<i>Istiompax indica</i>	MK902861
25	<i>Istiophorus platypterus</i>	MK902869, MK902909
26	<i>Johnrandallia nigrirostris</i>	MK902809, MK902819
27	<i>Kajikia audax</i>	MK902828, MK902838
28	<i>Katsuwonus pelamis</i>	MK902847, MK902853
29	<i>Kyphosus elegans</i>	MK902810, MK902862, MK902870
30	<i>Lobotes pacificus</i>	MK902896
31	<i>Lutjanus aratus</i>	MK902811, MK902863, MK902872, MK902888
32	<i>Lutjanus argentiventris</i>	MK902840, MK902849
33	<i>Lutjanus colorado</i>	MK902821, MK902830
34	<i>Lutjanus novemfasciatus</i>	MK902854, MK902855, MK902901
35	<i>Lutjanus peru</i>	MK902820, MK902829, MK902839
36	<i>Lutjanus viridis</i>	MK902848
37	<i>Microlepidotus inornatus</i>	MK902864
38	<i>Mugil cephalus</i>	MK902878
39	<i>Mulloidichthys dentatus</i>	MK902873
40	<i>Mycteroperca jordani</i>	MK902812, MK902813, MK902822, MK902831, MK902841, MK902856, MK902865, MK902874

41	<i>Mycteroperca rosacea</i>	MK902823, MK902832, MK902842, MK902866
42	<i>Paralabrax auroguttatus</i>	MK902824, MK902833
43	<i>Paralabrax maculatofasciatus</i>	MK902875
44	<i>Paralabrax nebulifer</i>	MK902843, MK902850, MK902857
45	<i>Paranthias colonus</i>	MK902908, MK902911
46	<i>Peprilus snyderi</i>	MK902917
47	<i>Phthanophaneron harveyi</i>	MK902912
48	<i>Prionurus punctatus</i>	MK902876
49	<i>Rachycentron canadum</i>	MK902895
50	<i>Rypticus bicolor</i>	MK902880, MK902918
51	<i>Scarus ghobban</i>	MK902891
52	<i>Scarus perrico</i>	MK902897, MK902903
53	<i>Scomberus sierra</i>	MK902814, MK902877
54	<i>Sebastes macdonaldi</i>	MK902904
55	<i>Seriola dumerili</i>	MK902905
56	<i>Seriola lalandi</i>	MK902844
57	<i>Seriola rivoliana</i>	MK902851, MK902858
58	<i>Sphoeroides anulatus</i>	MK902882
59	<i>Sphyraena ensis</i>	MK902890
60	<i>Stegastes rectifaenum</i>	MK902815
61	<i>Sufflamen verres</i>	MK902825, MK902834, MK902881
62	<i>Tetrapturus audax</i>	MK902887
63	<i>Thunnus albacares</i>	MK902892, MK902898, MK902900
64	<i>Thunnus thynnus</i>	MK902885
65	<i>Totoaba macdonaldi</i>	MK902883
66	<i>Xiphias gladius</i>	MK902916
67	<i>Xystreurus liolepis</i>	MK902910

**Table 20.** Species included in the Mock community sample (n=22).

Class	Order	Family	Genus	Species	Detection in mock sample after sequencing	Sequence in Custom reference database?	Sequence in NCBI?
Actinopterygii	Perciformes	Balistidae	Balistes	<i>Balistes polylepis</i>	*	*	
Actinopterygii	Tetraodontiformes	Balistidae	Sufflamen	<i>Sufflamen verres</i>	*	*	
Actinopterygii	Perciformes	Carangidae	Seriola	<i>Seriola rivoliana</i>	*	*	
Actinopterygii	Perciformes	Chaetodontidae	Johnrandalia	<i>Johnrandalia nigrirostris</i>	*	*	*
Actinopterygii	Perciformes	Cirrhitidae	Cirrhitus	<i>Cirrhitus rivulatus</i>	*	*	
Actinopterygii	Perciformes	Haemulidae	Anisotremus	<i>Anisotremus interruptus</i>	*	*	
Actinopterygii	Perciformes	Haemulidae	Haemulon	<i>Haemulon sexfasciatum</i>	*	*	
Actinopterygii	Perciformes	Kyphosidae	Kyphosus	<i>Kyphosus elegans</i>		*	*
Actinopterygii	Perciformes	Labridae	Bodianus	<i>Bodianus diplotaenia</i>	*	*	*
Actinopterygii	Perciformes	Labridae	Thalassoma	<i>Thalassoma lucasanum</i>	*		*
Actinopterygii	Perciformes	Lutjanidae	Lutjanus	<i>Lutjanus argentiventris</i>	*	*	
Actinopterygii	Perciformes	Lutjanidae	Lutjanus	<i>Lutjanus novemfasciatus</i>	*	*	
Actinopterygii	Perciformes	Lutjanidae	Lutjanus	<i>Lutjanus peru</i>	*	*	*
Actinopterygii	Perciformes	Pomacanthidae	Holacanthus	<i>Holacanthus passer</i>	*	*	
Actinopterygii	Perciformes	Pomacanthidae	Hoplopagrus	<i>Hoplopagrus guentheri</i>		*	
Actinopterygii	Perciformes	Pomacentridae	Abudefduf	<i>Abudefduf troschelii</i>	*	*	*
Actinopterygii	Perciformes	Pomacentridae	Stegastes	<i>Stegastes rectifraenum</i>	*	*	
Actinopterygii	Perciformes	Scaridae	Scarus	<i>Scarus ghobban</i>	*	*	*
Actinopterygii	Perciformes	Serranidae	Epinephelus	<i>Epinephelus labriformis</i>	*	*	
Actinopterygii	Perciformes	Serranidae	Paralabrax	<i>Paralabrax auroguttatus</i>	*	*	
Actinopterygii	Tetraodontiformes	Serranidae	Paranthias	<i>Paranthias colonus</i>	*	*	*
Actinopterygii	Perciformes	Serranidae	Semicossiphus	<i>Semicossyphus pulcher</i>	*		*

**S1.** eDNA extraction protocol modified from the Blood and Tissue kit (Qiagen).

eDNA extractions followed the spin-column protocol for purification of total DNA from animal tissues of the Qiagen Blood and Tissue Kit (Qiagen, USA) with some modifications as follows.

1. Each filter was cut finely (approximately 1 mm<sup>2</sup>) with the help of sterile stainless-steel scissors and tweezers over a clean aluminum foil paper.
2. The total STE buffer (~200 µL) contained in each tube and the pieces of the filter, were divided into 2 x 1.5 mL tubes per sample, to proceed with DNA extraction.
3. I added 250 µL of ATL buffer and 27 µL of proteinase K to each 1.5 mL tube, and incubate at 56°C for 4 hours, vortexing occasionally during this time.
4. Then, I added to each tube 280 µL of Buffer AL and 280 µL of ethanol (96%) and vortex for 20 seconds.
5. After this, I pipetted the liquid phase of the mixture from step 4 into a DNeasy Mini spin column placed in a 2mL collection tube and centrifuged at 8000rpm for 1 min, and discarded the flow-through and collection tube. This was repeated, and in the second centrifugation, I include the pieces of the filter. Centrifugations were repeated until the mixture of the 2 tubes obtained in the digestion from 1 sample was finished.
6. After the last centrifugation step, I removed pieces of the filter located inside the DNeasy Mini spin with sterile tweezers.
7. I placed the DNeasy Mini spin column in a new 2 mL collection tube and added 500 µL of Buffer AW1, and centrifuged for 1 min at 8000 rpm.
8. I placed the DNeasy Mini spin column in a new 2 mL collection tube and added 500 mL of Buffer AW2 and centrifuged for 3 minutes at 14000 rpm to dry the DNeasy membrane. Then, discarded flow-through and collection tube.
9. I placed the DNeasy Mini spin column in a clean 1.5 mL tube, and pipetted 50 µL of the Buffer AE, directly onto the column membrane. Incubated at room temperature for 1 min, and then centrifuged for 1 min at 8000 rpm to elute (This will be elution 1).
10. Finally, I placed the DNeasy Mini spin column in a second clean 1.5 mL tube, and pipetted 100 µL of the Buffer AE, directly onto the column membrane. Incubated at room temperature for 1 min, and then centrifuged for 1 min at 8000 rpm to elute (This will be elution 2).
11. All samples were quantified by fluorescence with an HS assay kit for a Qubit 3.0 fluorometer (Invitrogen, CA, US).

**Table 21.** Primers for library construction.

**First round primer pairs**

Primer Forward	Adapter used for 2 <sup>nd</sup> PCR	3' end of Illumina adapter	Universal 12S Forward Primer teleo_F
eDNA12SV-F	TCGTCGGCAGCGTC	AGATGTGTATAAGAGACAG	ACACCGCCCGTCACTCT
Primer Reverse	Adapter used for 2 <sup>nd</sup> PCR	3' end of Illumina adapter	Universal 12S Reverse Primer teleo_R
eDNA12SV-R	GTCTCGTGGGCTCGG	AGATGTGTATAAGAGACAG	CTCCGGTACACTTACCATG

**Primers with Illumina adapters and MID for second-round PCR**

Forward	Index 2	i5 MID	Adapter used for 2 <sup>nd</sup> PCR
eDNA2F-A	AATGATACGGCGACCACCGAGATCTACAC	GACACAGT	TCGTCGGCAGCGTC
eDNA2F-B	AATGATACGGCGACCACCGAGATCTACAC	GCATAACG	TCGTCGGCAGCGTC
eDNA2F-C	AATGATACGGCGACCACCGAGATCTACAC	ACAGAGGT	TCGTCGGCAGCGTC
eDNA2F-D	AATGATACGGCGACCACCGAGATCTACAC	CCACTAAG	TCGTCGGCAGCGTC
eDNA2F-E	AATGATACGGCGACCACCGAGATCTACAC	TGTTCCGT	TCGTCGGCAGCGTC
eDNA2F-F	AATGATACGGCGACCACCGAGATCTACAC	GATACCTG	TCGTCGGCAGCGTC
eDNA2F-G	AATGATACGGCGACCACCGAGATCTACAC	AGCCGTAA	TCGTCGGCAGCGTC
eDNA2F-H	AATGATACGGCGACCACCGAGATCTACAC	CTCCTGAA	TCGTCGGCAGCGTC
Reverse	Index 1	i7 MID	Adapter used for 2 <sup>nd</sup> PCR
eDNA2R-01	CAAGCAGAAGACGGCATAACGAGAT	TCACCTAG	GTCTCGTGGGCTCGG
eDNA2R-02	CAAGCAGAAGACGGCATAACGAGAT	CAAGTCGT	GTCTCGTGGGCTCGG
eDNA2R-03	CAAGCAGAAGACGGCATAACGAGAT	CTGTATGC	GTCTCGTGGGCTCGG
eDNA2R-04	CAAGCAGAAGACGGCATAACGAGAT	AGTTCGCA	GTCTCGTGGGCTCGG
eDNA2R-05	CAAGCAGAAGACGGCATAACGAGAT	ATCGGAGA	GTCTCGTGGGCTCGG
eDNA2R-06	CAAGCAGAAGACGGCATAACGAGAT	AAGTCCTC	GTCTCGTGGGCTCGG
eDNA2R-07	CAAGCAGAAGACGGCATAACGAGAT	TGGATGGT	GTCTCGTGGGCTCGG
eDNA2R-08	CAAGCAGAAGACGGCATAACGAGAT	AGGTGTTG	GTCTCGTGGGCTCGG
eDNA2R-09	CAAGCAGAAGACGGCATAACGAGAT	GACGAACT	GTCTCGTGGGCTCGG
eDNA2R-10	CAAGCAGAAGACGGCATAACGAGAT	GTTCTTCG	GTCTCGTGGGCTCGG
eDNA2R-11	CAAGCAGAAGACGGCATAACGAGAT	TTCGCCAT	GTCTCGTGGGCTCGG
eDNA2R-12	CAAGCAGAAGACGGCATAACGAGAT	CAACTCCA	GTCTCGTGGGCTCGG

## S2. Methods for PCR 1 and PCR 2 in library preparation.

### **First PCR step (PCR 1).**

The first step was performed to amplify ~ 65 bp from the 12S rRNA gene using “teleo” primers reported previously in Valentini et al., 2016. Primers included a standard Illumina sequencing adapter, the gene-specific primer and an adapter used for the second PCR, listed in S7. Three PCR replicates per sample were performed in order to offset the variability in individual PCR replicates, maximize diversity detection and minimize the potential of false negatives. Two PCRs used as template the DNA from the first elution (50 µL), and one with eDNA from the second elution (100 µL).

Each PCR1 reaction contained PCR 1x Buffer (5X Buffer Thermo), 0.2 mM dNTPs, 0.2 µM of each “teleo” F and R primers, 2 µM human blocking primer, 3% DMSO, 0.6 U Phusion polymerase (Thermo) and 1-2 µL eDNA in a total 12 µL reaction. The parameters for the thermocycling were: 98 ° C x 10min, 35 cycles of 98 ° C x 30 s, 61 ° C x 30 s, 72 ° C x 30 s, and a final extension of 72 ° C x 5 min. For all tests, negative (nuclease-free water, NEB) and positive controls (tissue-derived DNA of *Myxeroperca rosacea*) were used. Successful PCR amplifications were verified via 2% agarose gels.

### **Second PCR step (PCR 2).**

The second PCR step was performed in order to incorporate dual molecular identifiers (MID) to individual samples on the same flow cell. Primer pairs used contain the appropriate 8-nt index sequence (Glenn et al., 2016), the adapter to bind to the first PCR, and the sequencing primer sites. Following to the second PCR, the three experimental replicates per sample were combined and purified using 1.8X volume of AmpureXP beads (Beckman and Coulter). After cleaning, each sample was quantified by fluorescence with the HS assay kit for Qubit. Finally, a total of 26 samples (24 sampling sites, one Mock community, and one pooled negative control) were pooled into one 4 nM equimolar sample and sent to Genomic Services at Langebio-CINVESTAV. A single flow cell Illumina NextSeq 500 MID (35 Gb) v2 chemistry (2x150 bp paired-ends) was used for sequencing.

PCR amplifications contained 1x PCR Buffer (5X Buffer Thermo), 0.2 mM dNTPs, 0.2 µM F and R primers, 3% DMSO, 0.6 U Phusion (Thermo), 1-2 µL eDNA, in reactions of 12 µL. The parameters for the thermocycling were 98 ° C x 5min, 8 cycles of 98 ° C x 30 s, 61 ° C x 30 s, 72 ° C x 15 s, and a final extension of 72 ° C x 5min. I performed three PCR 2, one for each PCR 1 replicate.

**Table 22.** Pipeline for 12S rRNA metabarcoding bioinformatic analysis (OBITOOLS, VSEARCH, SWARM).

<b>1. Original files:</b>
sample_R1_001.fastq (323,473 reads), sample_R2_001.fastq (323,473 reads)
<b>2. Merge paired-end data (archive order must be reverse R2 – forward R1).</b>
illumina paired-end -r sample_R2_001.fastq sample_R1_001.fastq > sample.fastq
<b>3. Evaluate quality of the match between merging reads with score &gt;40.</b>
obigrep -p 'score>40.00' -p 'mode!="joined"' sample.fastq > sample.ali.fastq
<b>4. Label sample and experiment with <i>ngsfilter</i>.</b>
ngsfilter -t S1_ngsfilter.txt -u sample_unidentified.fastq sample.ali.fastq > sample.ali.assigned.fasta
<b>5. Obistats. To evaluate mean read length.</b>
obistats -c sample -mean seq_length sample.ali.assigned.fastq > sample.stat.ali.assigned.fastq
<b>6. Filter by length and with no Ns.</b>
obigrep -p 'seq_length<66' -p 'seq_length>59' -s '^[ACGT]+\$' sample.ali.assigned.fastq > sample.ali.filtered.fastq
<b>7. Concatenate all samples.</b>
cat *.ali.filtered.fastq > samples.cat.filtered.fastq
<b>8. Dereplicate reads into unique sequences.</b>
obiuniq -m sample samples.cat.filtered.fastq > samples.cat.unique.fasta
<b>9. Sort by abundance and change identifier of the sequence.</b>
obiannotate --seq-rank samples.cat.unique.fasta   obiannotate --set-identifier "'Gulf' %09d" % seq_rank' samples.cat.rank.fasta > samples.cat.new.fasta
<b>10. Change format fasta to vsearch.</b>
Rscript owi_obifasta2vsearch -l samples.cat.new.fasta -o samples.cat.vsearch.fasta
<b>11. Remove chimaeras of all samples together with vsearch.</b>
vsearch --uchime_denovo samples.cat.vsearch.fasta --sizeout --minh 0.90 --nonchimeras cat.nonchimeras.fasta --chimeras cat.chimeras.fasta --uchimeout cat.uchime_out.txt
<b>12. Change format from vsearch to obitools.</b>
Rscript owi_obifasta2vsearch -l cat.nonchimeras.fasta -o cat.nochimeras.vsearch.fasta
<b>13. Cluster using SWARM.</b>
swarm -d 2 -z -t 40 -o cat_SWARM1nc_output -s cat_SWARM1nc_stats -w cat_SWARM1nc_seeds.fasta cat.nochimeras.vsearch.fasta
<b>14. Recount the abundances.</b>
obitab -o cat.new.fasta > cat.new.tab
<b>15. Recount after SWARM to generate OTU table per sampling site and removes singletons after clustering</b>
Rscript owi_recount_swarm cat_SWARM_2_output cat.new.tab obigrep -p 'size>1' samples_seeds.fasta > samples_nonsingleton.fasta

**Table 23.** Genetic distances at various taxonomic levels: intra-species, intra-genus, intra-family and intra-order with the 12S rRNA barcode ( $d_{max}$ , % maximum genetic distance) estimated from the custom reference database.

**INTRA-SPECIES**

<b>Species included</b>	<b><math>d_{max}</math></b>
<i>Abudefduf troschelli</i>	0
<i>Anisotremus interruptus</i>	0
<i>Chaetodon humeralis</i>	0
<i>Coryphaena hyppurus</i>	0
<i>Elacatinus puncticulatus</i>	0
<i>Haemulon sexfasciatum</i>	0
<i>Holacanthus passer</i>	0
<i>Hoplopagrus guentheri</i>	0
<i>Istiophorus platypterus</i>	0
<i>Johnrandalia nigrirostris</i>	0
<i>Kajikia audax</i>	0
<i>Katsuwomis pelamis</i>	0
<i>Lutjanus aratus</i>	0
<i>Lutjanus argentiventris</i>	0
<i>Lutjanus colorado</i>	0
<i>Lutjanus novemfasciatus</i>	0
<i>Mycteroperca jordani</i>	0
<i>Paralabrax aurogutatus</i>	0
<i>Paralabrax nebulifer</i>	0
<i>Paranthias colonus</i>	0
<i>Rypticus bicolor</i>	0
<i>Scarus perrico</i>	0
<i>Seriola rivoliana</i>	0
<i>Sufflamen verres</i>	0
<i>Kyphosus elegans</i>	0 - 2.8
<i>Lutjanus peru</i>	0 - 1.4
<i>Myceroperca rosacea</i>	0 - 1.4
<i>Thunnus albacares</i>	0 - 1.4
<i>Scomberomorus sierra</i>	0 - 1.3

**INTRA-GENUS**

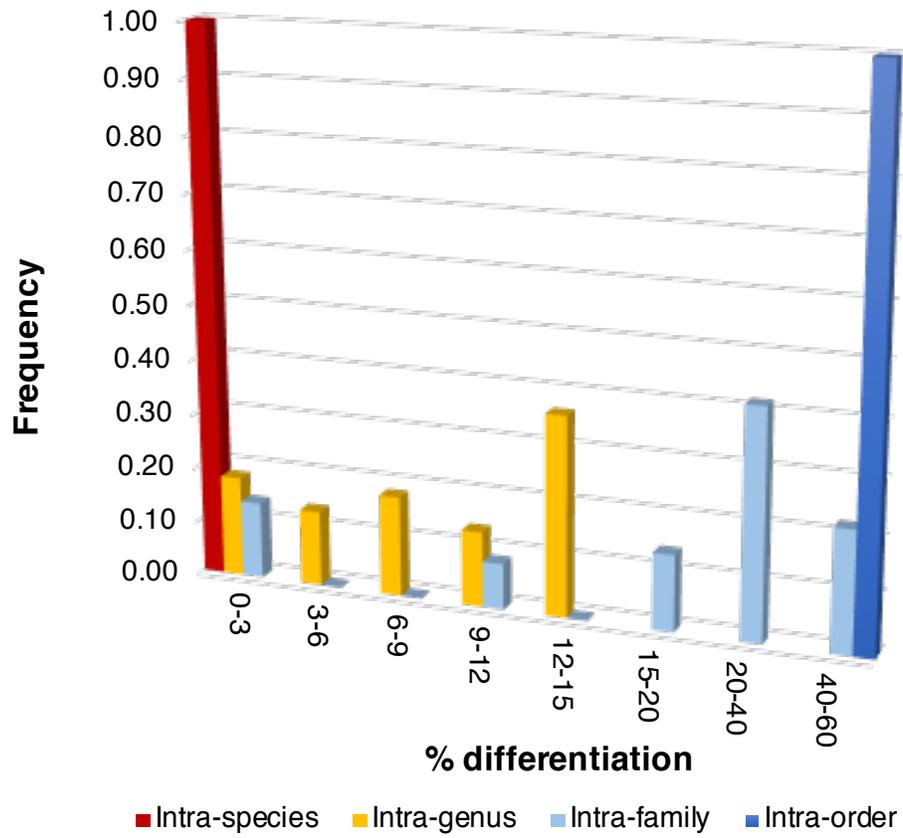
Genus	Species included	d <sub>max</sub>
Cynoscion	<i>C. reticulatus</i> , <i>C. xanthurus</i>	6.8
Epinephelus	<i>E. acanthistius</i> , <i>E. labrifformis</i>	12.8
Lutjanus	<i>L. aratus</i> , <i>L. argentiventris</i> , <i>L. colorado</i> , <i>L. novemfasciatus</i> , <i>L. peru</i>	14.5
Mycteroperca	<i>M. rosacea</i> , <i>M. j ordani</i>	8.1
Paralabrax	<i>P. auroguttatus</i> , <i>P. maculatofasciatus</i> , <i>P. nebulifer</i>	10.3
Scarus	<i>S. ghobban</i> , <i>S. perrico</i>	0
Seriola	<i>S. dumerili</i> , <i>S. rivoliana</i> , <i>S. lalandi</i>	5.4
Thunnus	<i>T. albacares</i> , <i>T. thynnus</i>	1.4

**INTRA-FAMILY**

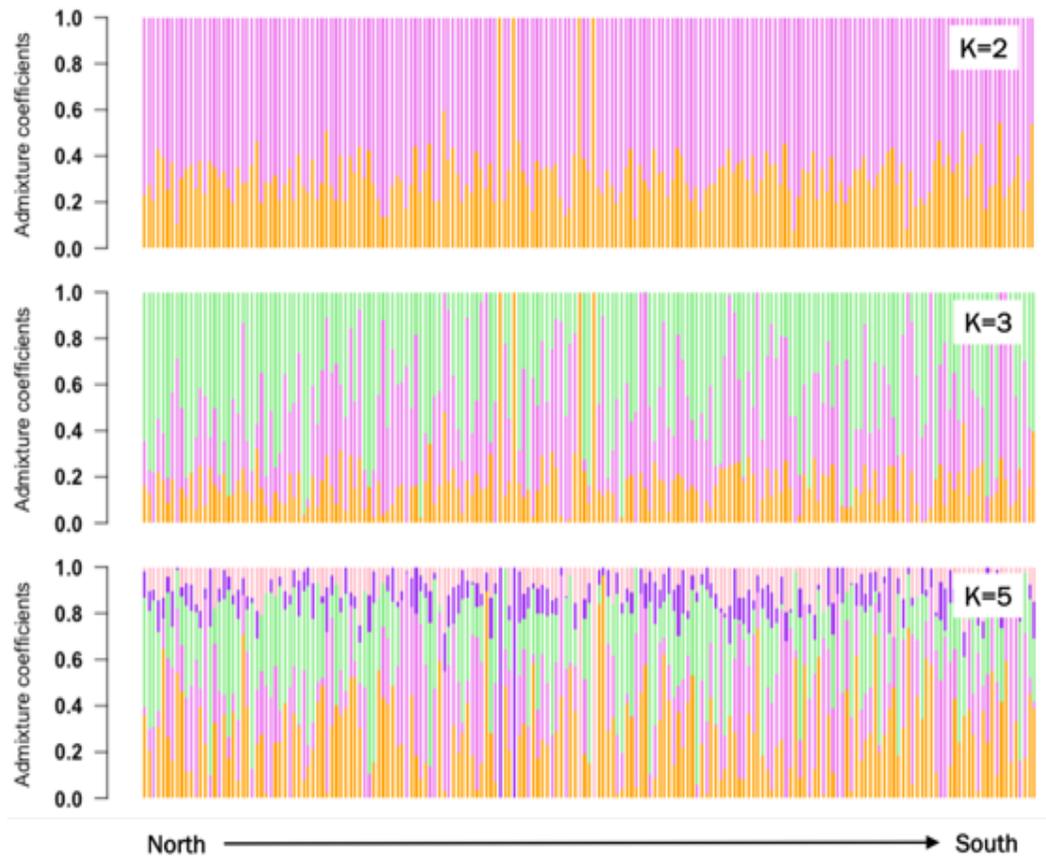
Family	Genus included	d <sub>max</sub>
Scianidae	<i>Cynoscion</i> , <i>Atractoscion</i> , <i>Totoaba</i>	9.6
Istiophoridae	<i>Istiompax</i> , <i>Istiophorus</i> , <i>Kajikia</i> , <i>Tetrapturux</i>	2
Scombridae	<i>Katsuwomis</i> , <i>Thunnus</i> , <i>Acanthocybium</i>	16.7
Pomacentridae	<i>Stegastes</i> , <i>Abudefduf</i>	42.9
Haemulidae	<i>Anisotremus</i> , <i>Haemulon</i> , <i>Haemulopsis</i> , <i>Microlepidotus</i>	28.8
Chaetodontidae	<i>Chaetodon</i> , <i>Johnrandalia</i>	34.6
Balistidae	<i>Balistes</i> , <i>Sufflamen</i>	18.9
Serranidae	<i>Epinephelus</i> , <i>Mycteroperca</i> , <i>Paralabrax</i> , <i>Paranthias</i> , <i>Rypticus</i>	42.7
Carangidae	<i>Seriola</i> , <i>Gnathanodon</i>	25.3
Lutjanidae	<i>Lutjanus</i> , <i>Hoplopagrus</i>	20.3

**INTRA-ORDER**

Order	Families included	d <sub>max</sub>
Perciformes	Acanthuridae, Carangidae, Chaetodontidae, Cirrhitidae, Coryphaenidae, Gerreidae, Gobiidae, Haemulidae, Istiophoridae, Kyphosidae, Labridae, Lobotidae, Lutjanidae, Mugilidae, Mullidae, Pomacanthidae, Pomacentridae, Scaridae, Sciaenidae, Scianidae, Scombridae, Serranidae, Sphyrnaenida, Stromateidae, Xiphiidae	54.7



**Figure 40.** Histogram of genetic distances at various taxonomic levels: intra-species, intra-genus, intra-family, and intra-order with the 12S rRNA barcode estimated from the custom reference database.



**Figure 41.** Admixture bar plots for K=2, K=3 and K=5.

**Table 24.** Taxonomic list of Actinopterygii detected with UVC.

Class	Order	Family	Genus	Species		
Actinopterygii	Anguilliformes	Muraenidae	Gymnothorax	<i>Gymnothorax castaneus</i>		
			Muraena	<i>Muraena lentiginosa</i>		
	Aulopiformes	Synodontidae	Synodus	<i>Synodus lacertinus</i>		
	Beryciformes	Holocentridae	Myripristis	<i>Myripristis leiognathus</i>		
			Neoniphon	<i>Neoniphon suborbitalis</i>		
	Elopiformes	Elopidae	Elops	<i>Elops affinis</i>		
	Perciformes	Acanthuridae	Acanthurus	Acanthurus	<i>Acanthurus nigricans</i>	
				Acanthurus	<i>Acanthurus triostegus</i>	
				Acanthurus	<i>Acanthurus xanthopterus</i>	
			Prionurus	<i>Prionurus punctatus</i>		
			Apogonidae	Apogon	<i>Apogon pacificus</i>	
				Apogon	<i>Apogon retrosella</i>	
			Blenniidae	Ophioblennius	<i>Ophioblennius steindachneri</i>	
				Plagiotremus	<i>Plagiotremus azaleus</i>	
			Carangidae	Carangoides	Carangoides	<i>Carangoides orthogrammus</i>
					Caranx	<i>Caranx caballus</i>
					Caranx	<i>Caranx sp</i>
					Gnathanodon	<i>Gnathanodon speciosus</i>
					Seriola	<i>Seriola lalandi</i>
					Trachinotus	<i>Trachinotus rhodopus</i>
					Chaenopsis	<i>Chaenopsis alepidota</i>
			Chaetodontidae	Chaetodon	Chaetodon	<i>Chaetodon humeralis</i>
					Johnrandallia	<i>Johnrandallia nigrirostris</i>
			Cirrhitidae	Cirrhitichthys	Cirrhitichthys	<i>Cirrhitichthys oxycephalus</i>
					Cirrhitus	<i>Cirrhitus rivulatus</i>
			Gobiidae	Coryphopterus	Coryphopterus	<i>Coryphopterus urosphilus</i>
					Lythrypnus	<i>Lythrypnus dalli</i>
			Haemulidae	Anisotremus	Anisotremus	<i>Anisotremus davidsonii</i>
					Anisotremus	<i>Anisotremus interruptus</i>
					Anisotremus	<i>Anisotremus taeniatus</i>
				Haemulon	Haemulon	<i>Haemulon maculicauda</i>
					Haemulon	<i>Haemulon scudderii</i>
					Haemulon	<i>Haemulon sexfasciatum</i>
	Haemulon	<i>Haemulon steindachneri</i>				
	Microlepidotus	Microlepidotus		<i>Microlepidotus inornatus</i>		
	Kyphosidae	Girella		Girella	<i>Girella simplicidens</i>	
				Kyphosus	<i>Kyphosus azurea</i>	
			Kyphosus	<i>Kyphosus elegans</i>		

		<i>Kyphosus ocyurus</i>
		<i>Kyphosus vaigiensis</i>
Labridae	Bodianus	<i>Bodianus diplotaenia</i>
	Halichoeres	<i>Halichoeres chierchiae</i>
		<i>Halichoeres dispilus</i>
		<i>Halichoeres melanotis</i>
		<i>Halichoeres nicholsi</i>
		<i>Halichoeres notospilus</i>
		<i>Halichoeres semicinctus</i>
	Semicossyphus	<i>Semicossyphus pulcher</i>
	Thalassoma	<i>Thalassoma lucasanum</i>
Labrisomidae	Labrisomus	<i>Labrisomus xanti</i>
	Malacoctenus	<i>Malacoctenus sp</i>
Lutjanidae	Lutjanus	<i>Lutjanus argentiventris</i>
		<i>Lutjanus guttatus</i>
		<i>Lutjanus novemfasciatus</i>
		<i>Lutjanus viridis</i>
Mullidae	Mulloidichthys	<i>Mulloidichthys dentatus</i>
Pomacanthidae	Holacanthus	<i>Holacanthus passer</i>
	Hoplopagrus	<i>Hoplopagrus guentherii</i>
	Pomacanthus	<i>Pomacanthus zonipectus</i>
Pomacentridae	Abudefduf	<i>Abudefduf troschelii</i>
	Chromis	<i>Chromis atrilobata</i>
		<i>Chromis limbaughi</i>
	Microspathodon	<i>Microspathodon bairdii</i>
		<i>Microspathodon dorsalis</i>
	Stegastes	<i>Stegastes acapulcoensis</i>
		<i>Stegastes flavilatus</i>
		<i>Stegastes rectifraenum</i>
Scaridae	Nicholsina	<i>Nicholsina denticulata</i>
	Scarus	<i>Scarus compressus</i>
		<i>Scarus ghobban</i>
		<i>Scarus perrico</i>
		<i>Scarus rubroviolaceus</i>
Sciaenidae	Pareques	<i>Pareques sp</i>
Serranidae	Alphestes	<i>Alphestes immaculatus</i>
	Cephalopholis	<i>Cephalopholis colonus</i>
		<i>Cephalopholis panamensis</i>
	Epinephelus	<i>Epinephelus labriformis</i>
	Mycteroperca	<i>Mycteroperca jordani</i>
		<i>Mycteroperca prionura</i>

			<i>Mycteroperca rosacea</i>
		Paralabrax	<i>Paralabrax maculatofasciatus</i>
			<i>Paralabrax sp</i>
		Rypticus	<i>Rypticus bicolor</i>
		Serranus	<i>Serranus psittacinus</i>
	Sparidae	Calamus	<i>Calamus brachysomus</i>
	Sphyrænidae	Sphyræna	<i>Sphyræna lucasana</i>
	Tripterygiidae	Crocodilichthys	<i>Crocodilichthys gracilis</i>
	Zanclidae	Zanclus	<i>Zanclus cornutus</i>
Scorpaeniformes	Scorpaenidae	Scorpaena	<i>Scorpaena mystes</i>
Sygnathiformes	Fistulariidae	Fistularia	<i>Fistularia commersonii</i>
Tetraodontiformes	Balistidae	Balistes	<i>Balistes polylepis</i>
		Pseudobalistes	<i>Pseudobalistes naufragium</i>
		Sufflamen	<i>Sufflamen verres</i>
	Diodontidae	Diodon	<i>Diodon holocanthus</i>
	Tetraodontidae	Arothron	<i>Arothron meleagris</i>
		Canthigaster	<i>Canthigaster punctatissima</i>
		Sphoeroides	<i>Sphoeroides annulatus</i>

**Table 25.** Read count per bioinformatic step.

Site ID	illuminapairedend	obigrep	ngsfilter	obigrep	obiuniq	vsearch	Swarm (OTUs)
POR	323,473	316,642	303,473	218,991	10,555	10,553	542
BSS	228,163	223,001	216,396	168,903	9,727	9,721	542
ANI	207,120	202,009	195,084	153,536	9,172	9,168	542
SDI	243,323	237,758	231,105	157,864	9,656	9,650	542
SCR	236,665	230,230	224,202	131,085	7,419	7,413	542
MAT	229,032	223,873	215,868	163,328	9,385	9,377	542
CAT	197,082	192,269	184,964	125,589	9,115	9,110	542
MON	198,737	194,233	188,297	142,806	8,431	8,428	542
DAN	231,739	225,811	217,330	167,751	9,874	9,869	542
CAR	248,964	242,700	220,371	132,021	8,188	8,182	542
COR	221,945	216,209	208,510	164,326	9,368	9,360	542
PUL	150,982	146,025	140,560	98,291	6,167	6,164	542
ILD	121,385	117,339	113,328	89,139	5,921	5,919	542
SMAR	170,012	165,282	159,290	103,654	6,376	6,374	542
TOR	214,539	208,221	206,297	158,662	8,105	8,102	542
NOL	198,148	193,252	184,141	132,260	7,702	7,694	542
PMA	212,156	206,831	196,502	109,218	6,634	6,631	542
FRA	121,642	116,636	112,236	75,776	6,017	6,015	542
LOR	206,876	201,315	193,542	110,516	7,049	7,046	542
IA-I	181,361	176,382	168,824	85,520	5,395	5,392	542
LOB	176,736	171,860	165,237	70,014	5,085	5,085	542
PAT	202,884	197,733	189,627	135,112	7,673	7,668	542
TIB	211,946	206,273	199,588	146,645	8,050	8,048	542
EST	206,776	201,297	191,224	148,228	8,297	8,287	542
MOCK	197,205	192,561	187,960	184,857	7,606	7,605	542
NEG	290,791	258,785	6	6	6	6	4

**Table 26.** Summary of the taxonomic assignments for 542 OTUs identified from the eDNA metabarcoding analysis employing three different reference databases (see methods for details). GOC: gulf of California. We show the number (%) of unassigned/assigned OTUs, then the number of those OTUs that are present in the GOC and the final OTUs count after taxonomic collapse and minimal abundance filter. The last 5 columns describe the final taxonomic resolution of the OTUs based on thresholds of sequence identity percentage.

	Unassigned OTUs	Assigned OTUs	Present in GOC	Final number of OTUs	Species (100-97%)	Genus (97-94%)	Family (94-91%)	Order (91-88%)	Class (< 88%)
<b>1. NCBI-Genebank</b>	281 (52%)	261(48%)	91	45	24	7	8	4	2
<b>2. Custom reference database</b>	275 (51%)	267 (49%)	267	122	30	15	37	29	11
<b>3. NCBI+ Custom reference databa:</b>	216 (40%)	326 (60%)	326	119	38	26	28	13	16

**Table 27.** Taxonomic list of Actinopterygii class detected with eDNA metabarcoding.

Class	Order	Family	Genre	Species	
Actinopterygii	Clupeiformes	Clupeidae	Dorosoma	<i>Dorosoma</i> sp	
			Sardinops	<i>Sardinops sagax</i>	
	Elopiformes	Elopidae	Elops	<i>Elops</i> sp	
	Mugilidiformes	Mugilidae	Mugil	<i>Mugil cephalus</i>	
	Perciformes	Acanthuridae	Acanthurus	Acanthurus	<i>Acanthurus</i> sp
				Acanthurus	<i>Acanthurus xanthopterus</i>
				Prionurus	<i>Prionurus punctatus</i>
		Apogonidae	Apogon	<i>Apogon retrosella</i>	
		Balistidae	Balistes	Balistes	<i>Balistes polylepis</i>
				Balistes	<i>Balistes</i> sp
		Carangidae	Carangoides	Carangoides	<i>Carangoides</i> sp
				Seriola	<i>Seriola lalandi</i>
		Cirrhitidae	Cirrhitichthys	<i>Cirrhitichthys oxycephalus</i>	
		Gobiidae	Coryphopterus	<i>Coryphopterus urospilus</i>	
		Haemulidae	Haemulon	Haemulon	<i>Haemulon sexfasciatum</i>
				Haemulopsis	<i>Haemulopsis leuciscus</i>
		Istiophoridae	Istiophorus	Istiophorus	<i>Istiophorus platypterus</i>
					OTU_01 (Istiophoridae)
					OTU_02 (Istiophoridae)
					OTU_03 (Istiophoridae)
					OTU_04 (Istiophoridae)
					OTU_05 (Istiophoridae)
		Kyphosidae	Hermosilla	Hermosilla	<i>Hermosilla azurea</i>
				Kyphosus	<i>Kyphosus</i> sp
	Kyphosus			<i>Kyphosus elegans</i>	
	Sectator	Sectator	Sectator	<i>Sectator</i> sp	
	Labridae	Bodianus	Bodianus	<i>Bodianus diplotaenia</i>	
			Bodianus	<i>Bodianus</i> sp	
			Semicossyphus	<i>Semicossyphus pulcher</i>	
			Thalassoma	<i>Thalassoma lucasanum</i>	
				OTU_06 (Labridae)	
	Lutjanidae	Lutjanus	Lutjanus	<i>Lutjanus novemfasciatus</i>	
Lutjanus			<i>Lutjanus argentiventris</i>		
Lutjanus			<i>Lutjanus peru</i>		
Lutjanus			<i>Lutjanus</i> sp		
Lutjanus			<i>Lutjanus viridis</i>		
			OTU_07 (Lutjanidae)		
			OTU_08 (Lutjanidae)		
	OTU_09 (Lutjanidae)				

		OTU_10 (Lutjanidae)
		OTU_11 (Lutjanidae)
		OTU_12 (Lutjanidae)
		OTU_13 (Lutjanidae)
		OTU_14 (Lutjanidae)
		OTU_15 (Lutjanidae)
Mullidae	Mulloidichthys	<i>Mulloidichthys dentatus</i>
Nomeidae	Cubiceps	<i>Cubiceps</i> sp
Pomacanthidae	Holacanthus	<i>Holacanthus</i> sp
	Microspathodon	<i>Microspathodon dorsalis</i>
	Pomacanthus	<i>Pomacanthus</i> sp
Pomacentridae	Abudefduf	<i>Abudefduf</i> sp
		<i>Abudefduf troschelli</i>
	Chromis	<i>Chromis</i> sp
		<i>Chromis viridis</i>
	Stegastes	<i>Stegastes flavilatus</i>
		<i>Stegastes rectifraenum</i>
Scaridae	Scarus	<i>Scarus perrico</i>
		<i>Scarus</i> sp
		OTU_17 (Scaridae)
		OTU_18 (Scaridae)
		OTU_19 (Scaridae)
		OTU_20 (Scaridae)
Sciaenidae	Atractoscion	<i>Atractoscion nobilis</i>
	Cynoscion	<i>Cynoscion</i> sp
		<i>Cynoscion xanthulus</i>
		OTU_21 (Sciaenidae)
Scombridae	Katsuwomis	<i>Katsuwomis pelamis</i>
		<i>Katsuwomis</i> sp
	Thunnus	<i>Thunnus albacares</i>
		<i>Thunnus</i> sp
		OTU_22 (Scombridae)
Serranidae	Cephalopholis	<i>Cephalopholis</i> sp
	Epinephelus	<i>Epinephelus</i> sp
	Hyporthodus	<i>Hyporthodus</i> sp
	Mycteroperca	<i>Mycteroperca rosacea</i>
		<i>Mycteroperca</i> sp
	Paralabrax	<i>Paralabrax nebulifer</i>
		<i>Paralabrax</i> sp
	Rypticus	<i>Rypticus</i> sp
		<i>Rypticus bicolor</i>

			OTU_23 (Serranidae)
			OTU_24 (Serranidae)
			OTU_25 (Serranidae)
			OTU_26 (Serranidae)
			OTU_27 (Serranidae)
			OTU_28 (Serranidae)
	Sphyraenidae	Sphyraena	<i>Sphyraena ensis</i>
			<i>Sphyraena</i> sp
			OTU_29 (Perciformes)
			OTU_30 (Perciformes)
			OTU_31 (Perciformes)
			OTU_32 (Perciformes)
			OTU_33 (Perciformes)
			OTU_34 (Perciformes)
			OTU_35 (Perciformes)
			OTU_36 (Perciformes)
			OTU_37 (Perciformes)
			OTU_38 (Perciformes)
			OTU_39 (Perciformes)
			OTU_40 (Perciformes)
			OTU_41 (Perciformes)
			OTU_42 (Perciformes)
Pleuronectiformes	Paralichthyidae	Paralichthys	<i>Paralichthys</i> sp
			OTU_16 (Paralichthyidae)
Sygnathiformes	Fistulariidae	Fistularia	<i>Fistularia commersonii</i>
Tetraodontiformes	Diodontidae	Diodon	<i>Diodon liturosus</i>
			OTU_43 (Actinopterygii)
			OTU_44 (Actinopterygii)
			OTU_45 (Actinopterygii)
			OTU_46 (Actinopterygii)
			OTU_47 (Actinopterygii)
			OTU_48 (Actinopterygii)
			OTU_49 (Actinopterygii)
			OTU_50 (Actinopterygii)
			OTU_51 (Actinopterygii)
			OTU_52 (Actinopterygii)
			OTU_53 (Actinopterygii)
			OTU_54 (Actinopterygii)
			OTU_55 (Actinopterygii)
			OTU_56 (Actinopterygii)
			OTU_57 (Actinopterygii)

Table 28. Detection of Species/OTU with UVC and eDNA metabarcoding.

Species/OTU	UVC	eDNA
<i>Abudefduf</i> sp		*
<i>Abudefduf troschelii</i>	*	*
<i>Acanthurus nigricans</i>	*	
<i>Acanthurus</i> sp		*
<i>Acanthurus triostegus</i>	*	
<i>Acanthurus xanthopterus</i>	*	*
<i>Alphestes immaculatus</i>	*	
<i>Anisotremus davidsonii</i>	*	
<i>Anisotremus interruptus</i>	*	
<i>Anisotremus taeniatus</i>	*	
<i>Apogon pacificus</i>	*	
<i>Apogon retrosella</i>	*	*
<i>Arothron meleagris</i>	*	
<i>Atractoscion nobilis</i>		*
<i>Balistes polylepis</i>	*	*
<i>Balistes</i> sp		*
<i>Bodianus diplotaenia</i>	*	*
<i>Bodianus</i> sp		*
<i>Calamus brachysomus</i>	*	
<i>Canthigaster punctatissima</i>	*	
<i>Carangoides orthogrammus</i>	*	
<i>Carangoides</i> sp		*
<i>Caranx caballus</i>	*	
<i>Caranx</i> sp	*	
<i>Cephalopholis panamensis</i>	*	
<i>Cephalopholis</i> sp		*
<i>Chaenopsis alepidota</i>	*	
<i>Chaetodon humeralis</i>	*	
<i>Chromis atrilobata</i>	*	
<i>Chromis limbaughi</i>	*	
<i>Chromis</i> sp		*
<i>Chromis viridis</i>		*
<i>Cirrhitichthys oxycephalus</i>	*	*
<i>Cirrhitus rivulatus</i>	*	
<i>Coryphopterus uropsilus</i>	*	*
<i>Crocodilichthys gracilis</i>	*	
<i>Cubiceps</i> sp		*
<i>Cynoscion</i> sp		*
<i>Cynoscion xanthulus</i>		*

<i>Diodon holocanthus</i>	*	
<i>Diodon liturosus</i>		*
<i>Dorosoma</i> sp		*
<i>Elops affinis</i>	*	
<i>Elops</i> sp		*
<i>Epinephelus labriformis</i>	*	
<i>Epinephelus</i> sp		*
<i>Fistularia commersonii</i>	*	*
<i>Girella simplicidens</i>	*	
<i>Gnathanodon speciosus</i>	*	
<i>Gymnothorax castaneus</i>	*	
<i>Haemulon maculicauda</i>	*	
<i>Haemulon scudderii</i>	*	
<i>Haemulon sexfasciatum</i>	*	*
<i>Haemulon steindachneri</i>	*	
<i>Haemulopsis leuciscus</i>		*
<i>Halichoeres chierchiae</i>	*	
<i>Halichoeres dispilus</i>	*	
<i>Halichoeres melanotis</i>	*	
<i>Halichoeres nicholsi</i>	*	
<i>Halichoeres notospilus</i>	*	
<i>Halichoeres semicinctus</i>	*	
<i>Hermosilla azurea</i>		*
<i>Holacanthus passer</i>	*	
<i>Holacanthus</i> sp		*
<i>Hoplopagrus guentherii</i>	*	
<i>Hyporthodus</i> sp		*
<i>Istiophorus platypterus</i>		*
<i>Johnrandallia nigrirostris</i>	*	
<i>Katsuwomis pelamis</i>		*
<i>Katsuwomis</i> sp		*
<i>Kyphosus azurea</i>	*	
<i>Kyphosus elegans</i>	*	*
<i>Kyphosus ocyurus</i>	*	
<i>Kyphosus</i> sp		*
<i>Kyphosus vaigiensis</i>	*	
<i>Labrisomus xanti</i>	*	
<i>Lutjanus argentiventris</i>	*	*
<i>Lutjanus guttatus</i>	*	
<i>Lutjanus novemfasciatus</i>	*	*
<i>Lutjanus peru</i>		*

<i>Lutjanus</i> sp		*
<i>Lutjanus viridis</i>	*	*
<i>Lythrypnus dalli</i>	*	
<i>Malacoctenus</i> sp	*	
<i>Microlepidotus inornatus</i>	*	
<i>Microspathodon bairdii</i>	*	
<i>Microspathodon dorsalis</i>	*	*
<i>Mugil cephalus</i>		*
<i>Mulloidichthys dentatus</i>	*	*
<i>Muraena lentiginosa</i>	*	
<i>Mycteroperca jordani</i>	*	
<i>Mycteroperca prionura</i>	*	
<i>Mycteroperca rosacea</i>	*	*
<i>Mycteroperca</i> sp		*
<i>Myripristis leiognathus</i>	*	
<i>Neoniphon suborbitalis</i>	*	
<i>Nicholsina denticulata</i>	*	
<i>Ophioblennius steindachneri</i>	*	
OTU_01 (Istiophoridae)		*
OTU_02 (Istiophoridae)		*
OTU_03 (Istiophoridae)		*
OTU_04 (Istiophoridae)		*
OTU_05 (Istiophoridae)		*
OTU_06 (Labridae)		*
OTU_07 (Lutjanidae)		*
OTU_08 (Lutjanidae)		*
OTU_09 (Lutjanidae)		*
OTU_10 (Lutjanidae)		*
OTU_11 (Lutjanidae)		*
OTU_12 (Lutjanidae)		*
OTU_13 (Lutjanidae)		*
OTU_14 (Lutjanidae)		*
OTU_15 (Paralichthyidae)		*
OTU_16 (Scaridae)		*
OTU_17 (Scaridae)		*
OTU_18 (Scaridae)		*
OTU_19 (Scaridae)		*
OTU_20 (Sciaenidae)		*
OTU_21 (Scombridae)		*
OTU_22 (Serranidae)		*
OTU_23 (Serranidae)		*

OTU_24 (Serranidae)		*
OTU_25 (Serranidae)		*
OTU_26 (Serranidae)		*
OTU_27 (Serranidae)		*
OTU_28 (Perciformes)		*
OTU_29 (Perciformes)		*
OTU_30 (Perciformes)		*
OTU_31 (Perciformes)		*
OTU_32 (Perciformes)		*
OTU_33 (Perciformes)		*
OTU_34 (Perciformes)		*
OTU_35 (Perciformes)		*
OTU_36 (Perciformes)		*
OTU_37 (Perciformes)		*
OTU_38 (Perciformes)		*
OTU_39 (Perciformes)		*
OTU_40 (Perciformes)		*
OTU_41 (Actinopterygii)		*
OTU_42 (Actinopterygii)		*
OTU_43 (Actinopterygii)		*
OTU_44 (Actinopterygii)		*
OTU_45 (Actinopterygii)		*
OTU_46 (Actinopterygii)		*
OTU_47 (Actinopterygii)		*
OTU_48 (Actinopterygii)		*
OTU_49 (Actinopterygii)		*
OTU_50 (Actinopterygii)		*
OTU_51 (Actinopterygii)		*
OTU_52 (Actinopterygii)		*
OTU_53 (Actinopterygii)		*
OTU_54 (Actinopterygii)		*
OTU_55 (Actinopterygii)		*
<i>Paralabrax maculatofasciatus</i>	*	
<i>Paralabrax nebulifer</i>		*
<i>Paralabrax</i> sp	*	*
<i>Paralichthys</i> sp		*
<i>Paranthias colonus</i>	*	
<i>Pareques</i> sp	*	
<i>Plagiotremus azaleus</i>	*	
<i>Pomacanthus</i> sp		*
<i>Pomacanthus zonipectus</i>	*	

<i>Prionurus punctatus</i>	*	*
<i>Pseudobalistes naufragium</i>	*	
<i>Rypicus</i> sp		*
<i>Rypticus bicolor</i>	*	*
<i>Sardinops sagax</i>		*
<i>Scarus compressus</i>	*	
<i>Scarus ghobban</i>	*	
<i>Scarus perrico</i>	*	*
<i>Scarus rubroviolaceus</i>	*	
<i>Scarus</i> sp		*
<i>Scorpaena mystes</i>	*	
<i>Sectator</i> sp		*
<i>Semicossyphus pulcher</i>	*	*
<i>Seriola lalandi</i>	*	*
<i>Serranus psittacinus</i>	*	
<i>Sphoeroides annulatus</i>	*	
<i>Sphyraena ensis</i>		*
<i>Sphyraena lucasana</i>	*	
<i>Sphyraena</i> sp		*
<i>Stegastes acapulcoensis</i>	*	
<i>Stegastes flavilatus</i>	*	*
<i>Stegastes rectifraenum</i>	*	*
<i>Sufflamen verres</i>	*	
<i>Synodus lacertinus</i>	*	
<i>Thalassoma lucasanum</i>	*	*
<i>Thunnus albacares</i>		*
<i>Thunnus</i> sp		*
<i>Trachinotus rhodopus</i>	*	
<i>Zanclus cornutus</i>	*	

**Table 29.** Genera detection from UVC and eDNA metabarcoding.

<b>Genus</b>	<b>UVC</b>	<b>eDNA</b>
<i>Abudefduf</i>	*	*
<i>Acanthurus</i>	*	*
<i>Alphestes</i>	*	
<i>Anisotremus</i>	*	
<i>Apogon</i>	*	*
<i>Arothron</i>	*	
<i>Atractosion</i>		*
<i>Balistes</i>	*	*
<i>Bodianus</i>	*	*
<i>Calamus</i>	*	
<i>Canthigaster</i>	*	
<i>Carangoides</i>	*	*
<i>Caranx</i>	*	
<i>Cephalopholis</i>	*	*
<i>Chaenopsis</i>	*	
<i>Chaetodon</i>	*	
<i>Chromis</i>	*	*
<i>Cirrhichthys</i>	*	*
<i>Cirrhitis</i>	*	
<i>Coryphopterus</i>	*	*
<i>Crocodilichthys</i>	*	
<i>Cubiceps</i>		*
<i>Cynoscion</i>		*
<i>Diodon</i>	*	*
<i>Dorosoma</i>		*
<i>Elops</i>	*	*
<i>Epinephelus</i>	*	*
<i>Fistularia</i>	*	*
<i>Girella</i>	*	
<i>Gnathanodon</i>	*	
<i>Gymnothorax</i>	*	
<i>Haemulon</i>	*	*
<i>Halichoeres</i>	*	
<i>Hermosilla</i>		*
<i>Holacanthus</i>	*	*
<i>Hoplopagrus</i>	*	
<i>Hyporthodus</i>		*
<i>Istiophorus</i>		*
<i>Johnrandallia</i>	*	

<i>Katsuwonus</i>		*
<i>Kyphosus</i>	*	*
<i>Labrisomus</i>	*	
<i>Lutjanus</i>	*	*
<i>Lythrypnus</i>	*	
<i>Malacoctenus</i>	*	
<i>Microlepidotus</i>	*	
<i>Microspathodon</i>	*	*
<i>Mugil</i>		*
<i>Mulloidichthys</i>	*	*
<i>Muraena</i>	*	
<i>Mycteroperca</i>	*	*
<i>Myripristis</i>	*	
<i>Neoniphon</i>	*	
<i>Nicholsina</i>	*	
<i>Ophioblennius</i>	*	
<i>Paralabrax</i>	*	*
<i>Paralichthys</i>		*
<i>Pareques</i>	*	
<i>Plagiotremus</i>	*	
<i>Pomacanthus</i>	*	*
<i>Prionurus</i>	*	*
<i>Pseudobalistes</i>	*	
<i>Rypticus</i>	*	*
<i>Sardinops</i>		*
<i>Scarus</i>	*	*
<i>Scorpaena</i>	*	
<i>Sectator</i>		*
<i>Semicossyphus</i>	*	*
<i>Seriola</i>	*	*
<i>Serranus</i>	*	
<i>Sphoeroides</i>	*	
<i>Sphyraena</i>	*	*
<i>Stegastes</i>	*	*
<i>Sufflamen</i>	*	
<i>Synodus</i>	*	
<i>Thalassoma</i>	*	*
<i>Thunnus</i>		*
<i>Trachinotus</i>	*	
<i>Zanclus</i>	*	

**Table 30.** Family detection from UVC and eDNA metabarcoding.

Family	UVC	eDNA
Acanthuridae	*	*
Apogonidae	*	*
Balistidae	*	*
Blenniidae	*	
Carangidae	*	*
Chaenopsidae	*	
Chaetodontidae	*	
Cirrhitidae	*	*
Clupeidae		*
Diodontidae	*	*
Elopidae	*	*
Fistulariidae	*	*
Gobiidae	*	*
Haemulidae	*	*
Holocentridae	*	
Istiophoridae		*
Kyphosidae	*	*
Labridae	*	*
Labrisomidae	*	
Lutjanidae	*	*
Mugilidae		*
Mullidae	*	*
Muraenidae	*	
Nomeidae		*
Paralichthyidae		*
Pomacanthidae	*	*
Pomacentridae	*	*
Scaridae	*	*
Sciaenidae	*	*
Scombridae		*
Scorpaenidae	*	
Serranidae	*	*
Sparidae	*	
Sphyraenidae	*	*
Synodontidae	*	
Tetraodontidae	*	
Tripterygiidae	*	
Zanclidae	*	

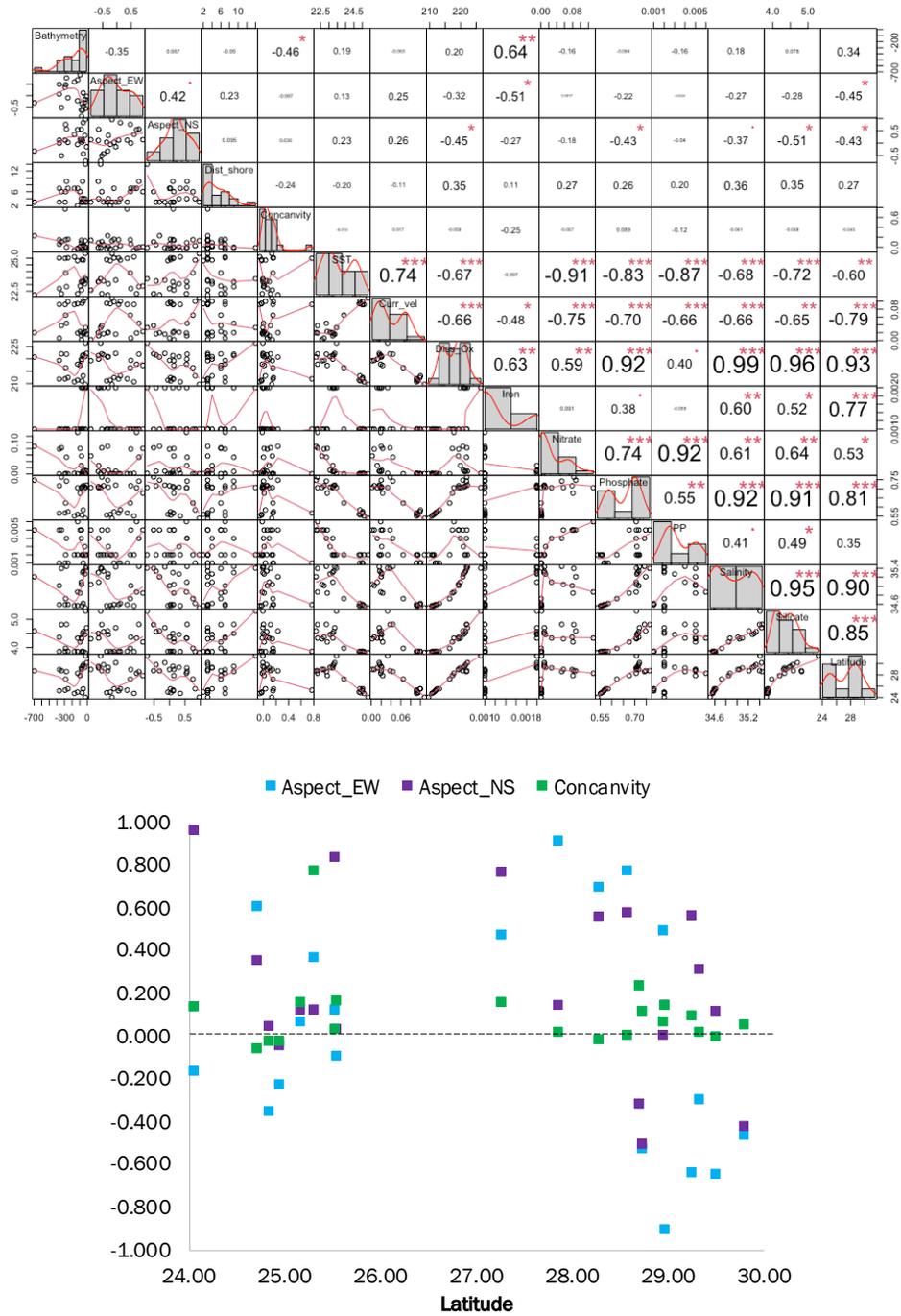
**Table 31.** Order detection from UVC and eDNA metabarcoding.

Order	UVC	eDNA
Anguilliformes	*	
Beryciformes	*	
Clupeiformes		*
Elopiformes	*	*
Mugilidiformes		*
Myliobatiformes	*	
Perciformes	*	*
Pleuronectiformes		*
Scorpaeniformes	*	
Sygnathiformes	*	*
Tetraodontiformes	*	*

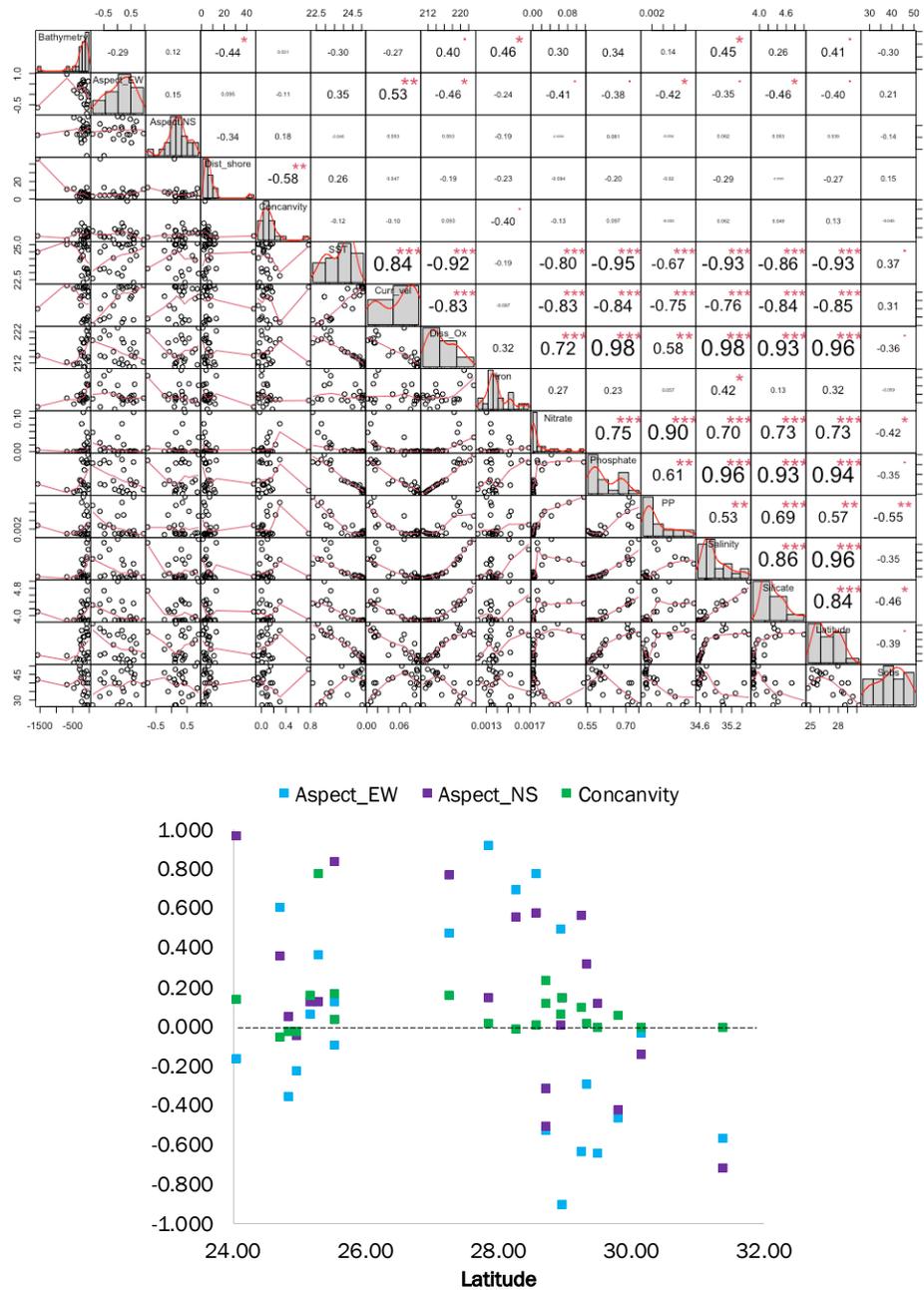
**Table 32.** Class detection from UVC and eDNA metabarcoding.

Class	UVC	eDNA
Actinopterygii	*	*





**Figure 43.** Pearson correlation tests among the environmental variables for population localities.  $R^2$  values are shown and its statistical significance (p-value 0 \*\*\*, 0.001 \*\*, 0.01 \*, 0.05 .).



**Figure 44.** Pearson correlation tests among the environmental variables for community localities and the observed species.  $R^2$  values are shown and its statistical significance (p-value 0 \*\*\*, 0.001 \*\*, 0.01 \*, 0.05 .).