



CENTRO DE INVESTIGACIONES BIOLÓGICAS
DEL NOROESTE, S.C.

Programa de Estudios de Posgrado

GENÓMICA POBLACIONAL DEL GOBIO *Elacatinus*
puncticulatus (Ginsburg, 1938) EN EL GOLFO DE CALIFORNIA

TESIS

Que para obtener el grado de

Maestro en Ciencias

Uso, Manejo y Preservación de los Recursos Naturales

(Orientación Biología Marina)

Presenta

Alejandra Quintero Grijalva

La Paz, Baja California Sur, febrero de 2020.

ACTA DE LIBERACIÓN DE TESIS

En la Ciudad de La Paz, B. C. S., siendo las 17 horas del día 17 del Mes de febrero del 2020, se procedió por los abajo firmantes, miembros de la Comisión Revisora de Tesis avalada por la Dirección de Estudios de Posgrado y Formación de Recursos Humanos del Centro de Investigaciones Biológicas del Noroeste, S. C., a liberar la Tesis de Grado titulada:

"Genómica poblacional del gobio *Elacatinus puncticulatus* (Ginsburg, 1938) en el Golfo de California"

Presentada por el alumno:

Alejandra Quintero Grijalva

Aspirante al Grado de MAESTRO EN CIENCIAS EN EL USO, MANEJO Y PRESERVACIÓN DE LOS RECURSOS NATURALES CON ORIENTACIÓN EN **Biología Marina**

Después de intercambiar opiniones los miembros de la Comisión manifestaron su **APROBACIÓN DE LA TESIS**, en virtud de que satisface los requisitos señalados por las disposiciones reglamentarias vigentes.

LA COMISIÓN REVISORA



Dr. Fausto Valenzuela Quiñonez
Co-Director de Tesis



Dr. Pedro Cruz Henández
Co-Director de Tesis



Dr. Raúl Octavio Martínez Rincón
Co-Tutor



Dra. Gracia Alicia Gómez Anduro,
Directora de Estudios de Posgrado y
Formación de Recursos Humanos.

Confirmación de Comités

Comité Tutorial

Dr. Fausto Valenzuela Quiñonez
CONACyT - Centro de Investigaciones Biológicas del Noroeste, S.C.
Co-Director de tesis

Dr. Pedro Cruz Hernández
Centro de Investigaciones Biológicas del Noroeste, S.C.
Co-Director de tesis

Dr. Raúl Octavio Martínez Rincón
CONACyT - Centro de Investigaciones Biológicas del Noroeste, S.C.
Tutor

Comité Revisor de Tesis

Dr. Fausto Valenzuela Quiñonez
Dr. Pedro Cruz Hernández
Dr. Raúl Octavio Martínez Rincón

Jurado de Examen

Dr. Fausto Valenzuela Quiñonez
Dr. Pedro Cruz Hernández
Dr. Raúl Octavio Martínez Rincón

Suplente

Dr. David Arturo Paz García

Resumen

El estudio de las poblaciones marinas y la determinación de los grupos poblacionales es importante para la conservación y manejo de las especies. Sin embargo, la complejidad de las historias de vida de los organismos marinos y del ambiente marino, lo vuelven una tarea complicada. La genómica de poblaciones es una herramienta indirecta que nos permite inferir la estructura poblacional de especies en vida silvestre a través de variaciones en las frecuencias alélicas a lo largo del genoma. Algunas de sus ventajas son una mejor representación del genoma y la capacidad de distinguir entre marcadores bajo fuerzas neutrales y selección natural (*outliers*). En este sentido la marcada heterogeneidad ambiental y el patrón de corrientes del Golfo de California (GC) lo hacen un laboratorio natural para inferir el efecto de estas fuerzas sobre las poblaciones marinas. El objeto de estudio es el gobio cabeza roja (*Elacatinus puncticulatus*), el cual habita en arrecifes rocosos y coralinos a lo largo del GC. Su amplia distribución sobre un ambiente heterogéneo, su historia de vida ligada a la costa rocosa del GC y su baja movilidad durante etapa adulta, le confieren un fuerte potencial para ser utilizada como especie modelo en estudios de genómica poblacional. El objetivo de este trabajo fue el de determinar la estructura genómica del gobio *E. puncticulatus* a lo largo del GC. Mediante la técnica de secuenciación masiva ddRADseq se obtuvieron las secuencias genómicas de un total de 127 individuos distribuidos en 19 sitios de muestreo a lo largo del GC. Se identificaron un total de 4,979 marcadores, de los cuales 4,953 fueron neutrales y 26 *outliers*. La estructura inferida con marcadores neutrales nos sugiere la existencia de dos grupos poblacionales, dividiendo la región norte del sur, con un quiebre genético entre los sitios de Mulegé y Loreto, entre estas zonas se detectó una introgresión clinal, la cual nos sugiere una zona de mezcla en latitudes intermedias. Esto podría ser señal de contacto secundario, el cual ha sido mantenido por las dinámicas oceanográficas y ambientales contemporáneas del GC. Mientras que la estructura con marcadores *outliers* nos sugiere tres grupos con potencial adaptativo: 1) Bahía de los Ángeles, 2) Mulegé, y 3) Loreto, La Paz y Los Cabos, las cuales coinciden con bioregiones basadas en la composición ictiofaunística del GC. Nuevamente se observó una zona de transición entre las regiones de Mulegé y Loreto asociada a una introgresión génica y valores de diversidad genética superiores. Por lo anterior, se sugiere que la estructura genómica poblacional de *E. puncticulatus* es el reflejo de eventos no excluyentes de naturaleza neutral y adaptativa.

Palabras clave: Genómica poblacional, Golfo de California, ddRad-Seq, marcadores neutrales, marcadores *outliers*, heterogeneidad ambiental, gobio cabeza roja.



Vo.Bo.

Dr. Fausto Valenzuela Quiñonez
Co-Director de Tesis



Vo.Bo.

Dr. Pedro Cruz Hernández
Co-Director de Tesis

Summary

The study of marine populations and the determination of population groups is relevant for the conservation and management of species. However marine organisms present complex life histories, and combined with the marine environment characteristics, make it a complicated task. Population genomics is an indirect tool that enable to infer population structure of wildlife through variations in the allelic frequencies throughout the genome. Some of its advantages are a better representation of the genome, and discriminate between markers under neutral forces and natural selection (*outliers*). In this regard, environmental heterogeneity and oceanographic complex of the Gulf of California (GC) became a natural laboratory to infer the effect of this forces over marine populations. The study is focus in the red-headed goby (*Elacatinus puncticulatus* (Ginsburg, 1983)), which inhabits rocky and coral reefs along the GC. Its wide distribution over heterogeneous environment, its life history linked to the rocky coast of the GC and its low mobility during its adult stage, gives it a potential to be used as a model species in populations genomic studies. The aim of this work was to determine the genomic structure of the goby *E. puncticulatus* along the GC. Using the ddRADseq (Peterson *et al.*, 2012), sequences of a total of 127 individual distributed in 19 sampling site along the GC were obtained. A total of 4,979 SNP markers were identified, of which 4,953 were neutral, and 26 *outliers*. The structure inferred with neutral markers suggests the existence of two population groups, dividing the northern and southern region, with a genetic break between the sites in Mulegé and Loreto, between this areas a clinal introgression was detected, which suggests an area of admixture in intermediate latitudes. This could be a secondary contact signal, which has been maintained by the contemporary oceanographic and environmental dynamics of the GC. While structure of *outlier* markers suggest three groups with adaptive potential: 1) Angeles Bay, 2) Mulege, and 3) Loreto, La Paz, and Los Cabos. Again, a transition zone was observed between Mulege and Loreto regions, associated with a gene introgression and higher genetic diversity values. Therefore, it is suggested that the population genomic structure is the result of non-exclusive events of neutral and adaptive nature.

Key words: Populations genomics, Gulf of California, ddRAD-Seq, neutral markers, outlier markers, environmental heterogeneity, red-headed goby.



Vo.Bo.

Dr.Fausto Valenzuela Quiñonez
Co-Director de Tesis



Vo.Bo.

Dr.Pedro Cruz Hernández
Co-Director de Tesis

Dedicatoria

A Dios, mi familia, y amigos.

Gracias por su amor.

"Quizá ahora entiendas lo que es el amor: amor es una palabra que se refiere al encanto a nivel cósmico; a la fuerza elemental que crea agrupaciones de átomos, galaxias, estrellas, familias, naciones, personas, ecosistemas, océanos y estrellas. El amor da vida."

-Brian Swimme

Agradecimientos

Al Centro de Investigaciones Biológicas del Noroeste S.C. (CIBNOR), por recibirme en el programa de posgrado en el Uso, Manejo y Preservación de los Recursos Naturales. Así como a todo el personal docente y administrativo.

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada (633991)

Al proyecto Ciencia Básica 2015-01-257524 “Paisaje genómico en el medio marino: patrones de diversidad adaptativa” por el financiamiento de este trabajo.

A la Dirección del General de Ordenamiento Pesquero y Acuícola de la Comisión Nacional de Acuicultura y Pesca por el permiso de pesca de fomento con fines científicos: DGOPA-DAPA-02850/19.

A mi comité tutorial: Dr. Fausto Valenzuela Quiñonez y Dr. Pedro Cruz Hernández como co-directores, y el Dr. Raúl O. Martínez Rincón, como tutor; por sus valiosos aportes brindados para el desarrollo y conclusión de este trabajo. Gracias por su tiempo, paciencia y guía a lo largo de este viaje.

Al equipo que realizó el muestreo de los organismos utilizados en este trabajo: Dr. Fausto Valenzuela Quiñonez, M.C. Salwa El Khattabi Salazar, M.C. Camila Mac Loughlin Alemán, y M.C. Jorge Mares Mayagoitia.

Al laboratorio de genética molecular, en especial a la M.C. Delia Rojas Posadas, Dra. Crisalejandra Rivera y Dra. Hernández Saavedra, por espacio, material y apoyo brindado para llevar a cabo el procesamiento de las muestras.

A mis compañeros de laboratorio: Cecilia J. Hid Coronel, Juan. P. Mora Martínez, Paulina Mejía Ruiz, Camila Mac Loughlin Alemán, Jorge Mares Mayagoitia, y Salwa El Khattabi Salazar. Gracias por todo el apoyo brindado.

Al laboratorio de genómica por facilitarme el uso de su equipo de laboratorio, en especial a M.C. Gabriela Mendoza Carrión.

A M.C. Karina A. Vázquez Rojas, M.C. Victor López Flores, y Veronica Mendoza por su ayuda en la elaboración de librerías y análisis bioinformáticos.

A Perla Sol Cervantes Bernal, Tania Verónica Núñez Valdez, y Horacio Sandoval Gómez, por su ayuda y atención.

A mis compañeros y amigos de posgrado. Muy en especial a los científicos adultos M.C. Lucero B. Cervantes Montoya (y a tu familia), M.C. M. Fernanda Ochoa Espinosa, M.C. M. Yoshio Shiba Reyes, y M.C. L. Mariam Calixto Heredia; gracias por ser tan bonitas personas.

A mis compañeras de casa Carolina A. Larrinaga, Vainilla, Nori y Chiquita; por su amistad y apoyo.

A mi familia, por todo el apoyo brindado en este recorrido, y por ser un ejemplo de fuerza y amor. Gracias por ser parte de mi vida, los amo.

Contenido

Resumen	i
Summary	ii
Dedicatoria	iii
Agradecimientos.....	iv
Lista de figuras.....	viii
Lista de tabla.....	xi
Abreviaturas.....	xiii
1. INTRODUCCIÓN	1
2. ANTECEDENTES	3
2.1 Genómica poblacional y secuenciación masiva	3
2.2 Técnicas de RAD-seq	5
2.3 Estudios de estructura genómica poblacional para especies de ambientes marinos costeros ...	5
2.4 Características ambientales de la región.....	8
2.5 Biología del gobio cabeza roja (<i>Elacatinus puncticulatus</i>)	10
3. JUSTIFICACIÓN	13
4. HIPÓTESIS.....	14
5. OBJETIVOS	15
5.1 Objetivo general	15
5.2 Objetivos específicos.....	15
6. MATERIALES Y METODOS	16
6.1 Muestreo	16
6.2 Trabajo de laboratorio	19
6.2.1 Extracción de ADN	19
6.2.2 Elaboración de librerías ddRAD-Seq	19
6.3 Análisis de secuencias y obtención de genotipos	23
6.4 Detección de <i>outliers</i>	26
6.5 Prueba de equilibrio de Hardy-Weinberg	27
6.6 Diversidad genética	27
6.7 Estructura genética	28
6.7.1 F_{ST} pareado.....	28

6.7.2 Análisis de inferencia de <i>cluster</i>	28
6.7.3 Análisis Discriminante de Componentes Principales	28
6.7.4 AMOVA jerárquico	29
6.8 Migración.....	29
6.8 Aislamiento por distancia	29
7. RESULTADOS	30
7.1 Librería genómica ddRAD.....	30
7.3 Demultiplexeado	33
7.4 Ensamble <i>de novo</i> y parametrización	34
7.6 Filtrado de loci y SNPs.....	36
7.7 Detección de <i>outliers</i>	37
7.8 Diversidad genética	39
7.8 Aislamiento por distancia	42
7.9 Estructura poblacional con marcadores neutrales	45
7.10. Migración.....	55
7.11 Estructura poblacional con marcadores <i>outliers</i>	57
8. DISCUSIÓN	70
8.1 Estructura genómica poblacional neutral.....	70
8.1.2 Efecto de la distancia geográfica y la circulación oceanográficas en el Golfo de California sobre patrones estructurales de la población	72
8.1.2 Efecto de eventos vicariantes en el Golfo de California sobre patrones estructurales de la población	74
8.1.4 Efecto de la capacidad dispersiva de la especie en el Golfo de California sobre patrones estructurales de la población.....	77
8.2 Estructura genómica poblacional con marcadores <i>outliers</i>	78
8.2.1 Efecto de la heterogeneidad ambiental en el Golfo de California sobre patrones estructurales de la población.....	79
9. CONCLUSIONES.....	82
10. LITERATURA CITADA.....	84

Lista de figuras

Figura 1. Gobio cabeza roja (<i>Elacatinus puncticulatus</i>) (Fotografía: El Khatabbi-Salazar, 2019).	10
Figura 2. Puntos de muestreo de <i>E. puncticulatus</i> en el Golfo de California (INEGI <i>et al.</i> , 1990).	18
Figura 3. Muestreo de <i>E. puncticulatus</i> en la costa oriental de la península de Baja California, mediante la técnica de buceo autónomo, con el uso de aceite de clavo como anestésico y redes para su captura. Fotografía: Valenzuela-Quñonez, 2017.	19
Figura 4. Área de estudio (INEGI <i>et al.</i> , 1990). BAFPA: Bahía Algodones, Frailes y Punta Arenas; BOT: La Bota, CC: Cabeza de Caballo, IAV: Islote Agua Verde, ISAD: Isla San Damián, ISANC: Isla San Cosme, MAR: Marcial, PB: Punta Berrerndo, PCSMS: Punta Chivato, y San Marcos Sur, PD: Punta Diablo, PISF: El Portugués, e Isla San Francisquito; SALTI: Saltito Isla, SANJPS: San José Punta Sur, SMN: San Marcos Norte, y VEN: Ventana.....	26
Figura 5. Ejemplo de ADN genómico extraído de <i>Elacatinus puncticulatus</i> [25 ng]. M= Marcador de peso molecular 1kb (Thermo Fisher Scientific), 1= ISANC-099, 2= ISANC-099, 3= ISANC-100, 4= SMS-116, 5= SMS-117, 6= SMS-118, 8=SMS-118, 9= SMS-120, 10= SMS-121, 11= SMN-138, 12= SMN-134, 13= SMN-135, 14= SMN-137.	30
Figura 6. Verificación de ligación de adaptadores en <i>E. puncticulatus</i> . Se aprecia un barrido debido a la amplificación de fragmentos de ADN de distintos tamaños. Nota: M= Marcador de peso molecular GeneRuler 1kb, 1= ISANC-096, 2= ISANC-097, 3= ISANC-098, 4= ISANC-099, 5= ISANC-100, 6= SMS-116, 7= SMS-117, 8= SMS-118, 9= SMS-120, 10= SMS-121, 11=SMN-138, 12= SMN-134.	31
Figura 7. Pools de <i>E. puncticulatus</i> amplificados con <i>primers</i> indexados. El tamaño esperado después de la selección de fragmentos es de 475 pb. Nota: M= Marcador de peso molecular GeneRuler 1kb, 1= Pool 1, 2= Pool 2, 3= Pool 3, 4= Pool 4.	32
Figura 8. Electroferograma después de la purificación de la librería genómica de <i>Elacatinus puncticulatus</i> obtenido con Bioanalyzer Agilent 2100-H. a) Antes de la purificación. b) Después de la purificación.	33
Figura 9. Porcentaje de lecturas retenidas y eliminadas por el proceso de filtrado en 'process_radtags'.	34
Figura 10. Porcentaje de individuos de <i>Elacatinus puncticulatus</i> retenidos y descartados en el filtrado de los datos.	36
Figura 11. Número de loci, sitios variantes y máximo porcentaje de datos faltantes por individuo para <i>E. puncticulatus</i> para distintos valores de $-p$. En $-p14$ se muestran los datos utilizados en este trabajo.	37
Figura 12. Detección de <i>outliers</i> para <i>E. puncticulatus</i> utilizando modelo jerárquico de islas en Arlequin. Cada punto representa uno de los 4,979 marcadores, los puntos rojos con	

valores positivos de F_{ST} en el cuantil del 1% de los datos, fueron considerados como marcadores <i>outliers</i> . En total el programa detectó 104 <i>outliers</i>	38
Figura 13. Detección de <i>outliers</i> para <i>E. puncticulatus</i> utilizando método bayesiano implementado en Bayescan. Cada punto representa uno de los 4,979 marcadores, los puntos después de la línea vertical se encuentran en el cuantil del 1%, estos fueron considerados como marcadores <i>outliers</i> . En total el programa detectó 56 <i>outliers</i>	39
Figura 14. Heterocigosidad observada (H_o) y esperada (H_e) neutral para <i>E. puncticulatus</i> en el GC para cada sitio.....	40
Figura 15. Diversidad nucleotídica (π) neutral para <i>E. puncticulatus</i> para cada sitio en el GC.....	41
Figura 16. Heterocigosidad observada (H_o) y esperada (H_e) para <i>outliers</i> de <i>E. puncticulatus</i> para cada sitio en el GC.....	42
Figura 17. Diversidad nucleotídica (π) para <i>outliers</i> de <i>E. puncticulatus</i> para cada sitio en el GC.	42
Figura 18. Prueba de aislamiento por distancia para <i>Elacatinus puncticulatus</i> en el GC. Prueba de Mantel para marcadores neutrales: $Z= 19.544$, $r= 0.846$, $p<0.0001$. Prueba de Mantel para marcadores <i>outliers</i> : $Z= 218.348$, $r= 0.929$, $p<0.0001$	43
Figura 19. Prueba de Mantel para probar aislamiento por distancia con marcadores neutrales para <i>E. puncticulatus</i> en el GC. Se estimó densidad de kernel bidimensional.	44
Figura 20. Prueba de Mantel para probar aislamiento por distancia con marcadores <i>outliers</i> para <i>E. puncticulatus</i> en el GC. Se estimó densidad de kernel bidimensional.....	44
Figura 21. K más probable para marcadores neutrales de <i>E. puncticulatus</i> en el GC. a) Magnitud de ΔK para distintos valores de K (Evanno <i>et al.</i> , 2015). b) Pr(K) para cada K (Pritchard <i>et al.</i> , 2000).	50
Figura 22. Agrupamiento de poblaciones de <i>E. puncticulatus</i> con marcadores neutrales. a) Individuos con sus probabilidades de pertenencia agrupados en dos grupos principales (K=2) (azul: sur, y naranja: norte). b) Probabilidad de pertenencia de individuos agrupados en tres grupos (K=3) (azul: Los Cabos y La Paz; morado: Loreto y Muleg�; y naranja: Bah�a de los �ngeles). c) Individuos con sus probabilidades de pertenencia agrupados en cuatro grupos (K=4) (azul: Los Cabos y La Paz, morado: Loreto, verde: Muleg�, y naranja: Bah�a de los �ngeles). 1: BAFPA, 2: SALTI, 3: PD, 4: PISF, 5: SANJPS, 6: PB, 7: MAR, 8: IAV, 9: ISAD, 10: ISANC, 11: PCSMS, 12: SMN, 13: CC, 14: VEN, y 15: BOT.	51
Figura 23. ADCP con marcadores neutrales con 50 PCs retenidos. a) obtenido por localidad, con dos funciones discriminantes retenidas, y b) densidad de los individuos a partir de la K m�s probable obtenida con la funci�n <i>find.clusters</i> (K=2, BIC=830.057), con una funci�n discriminante retenida, en color azul las localidades del sur y en naranja las del norte.	52
Figura 24. Red de migraci�n relativa entre los 15 sitios para marcadores neutrales. La flecha indica la direcci�n del flujo, y su grosor la magnitud. Se presentan valores de $G_{ST}>0.3$. 1: BAFPA, 2: SALTI, 3: PD, 4: PISF, 5: SANJPS, 6: PB, 7: MAR, 8: IAV, 9: ISAD, 10: ISANC, 11: PCSMS, 12: SMN, 13: CC, 14: VEN, y 15: BOT.....	56

Figura 25. Proporción de la asignación de pertenencia de cada uno de los 15 sitios a los dos grupos de <i>E. puncticulatus</i> detectados en STRUCTURE (K=2) con marcadores neutrales. (INEGI <i>et al.</i> , 1990).	57
Figura 26. K más probable para marcadores <i>outliers</i> . a) Magnitud de ΔK para distintos valores de K (Evanno <i>et al.</i> , 2015). b) Pr(K) para cada K (Pritchard <i>et al.</i> , 2000).	59
Figura 27. Agrupamiento de poblaciones con marcadores <i>outliers</i> (26 SNPs) de <i>E. puncticulatus</i> en el GC. a) Individuos con sus probabilidades de pertenencia agrupados en dos grupos principales (K=2) (azul: sur, y naranja: norte). b) Individuos con sus probabilidades de pertenencia agrupados en cuatro grupos (K=3) (azul: Los Cabos, La Paz, y Loreto, morada: Mulegé, y naranja: Bahía de los Ángeles). c) Individuos con sus probabilidades de pertenencia agrupados en cuatro grupos (K=4) (azul: Los Cabos y La Paz, morado: Loreto, verde: Mulegé, y naranja: Bahía de los Ángeles). 1: BAFPA, 2: SALTI, 3: PD, 4: PISF, 5: SANJPS, 6: PB, 7: MAR, 8: IAV, 9: ISAD, 10: ISANC, 11: PCSMS, 12: SMN, 13: CC, 14: VEN, y 15: BOT.....	60
Figura 28. ADCP con marcadores <i>outliers</i> con 22 PCs retenidos. a) obtenido por localidad, con dos funciones discriminantes retenidas, y b) densidades individuales contra la primer función discriminante a partir de la K más probable obtenida con la función <i>find.clusters</i> (K=4, BIC=191.598), en color azul las localidades del sur y en naranja las del norte.....	61
Figura 29. Probabilidad de pertenencia de los 15 sitios a los tres grupos con potencial adaptativo de <i>E. puncticulatus</i> identificadas con STRUCUTRE (K=3) con marcadores <i>outliers</i> (INEGI <i>et al.</i> , 1990).	64
Figura 30. Probabilidad de pertenencia de los 15 sitios a los tres grupos con potencial adaptativo de <i>E. puncticulatus</i> identificadas con STRUCUTRE (K=3) con marcadores <i>outliers</i> , y su distribución sobre la media de temperatura superficial (SST) de invierno (1955-2017) en el Golfo de California (INEGI <i>et al.</i> , 1990 .Locarnini <i>et al.</i> , 2019).	65
Figura 31. Probabilidad de pertenencia de los 15 sitios a los tres grupos con potencial adaptativo de <i>E. puncticulatus</i> identificadas con STRUCUTRE (K=3) con marcadores <i>outliers</i> , y su distribución sobre la media de salinidad superficial (SSM) anual (1955-2017) en el Golfo de California (INEGI <i>et al.</i> , 1990; Zweng <i>et al.</i> , 2019).....	66

Lista de tabla

Tabla I. Número de individuos de <i>Elactinus puncticulatus</i> recolectados por sustrato en cada sitio a lo largo del Golfo de California.	17
Tabla II. Individuos de <i>E. puncticulatus</i> utilizados para la construcción de librerías por sitio.....	20
Tabla III. <i>Pools</i> indexados de <i>E. puncticulatus</i>	23
Tabla IV. Valores optimizados de los parámetros utilizados en <i>denovo_map.pl</i>	24
Tabla V. Individuos finales de <i>E. puncticulatus</i> utilizados en este estudio pertenecientes a los sitios reagrupados.	25
Tabla VI. Resumen del control de calidad de Novogene después de la purificación de la librería.	33
Tabla VII. Filtrado de lecturas crudas en el componente <i>process_radtags</i> en Stacks.	34
Tabla VIII. Pruebas de los parámetros principales (-M, -m, y -n) en <i>denovo_map.pl</i> . Los renglones sombreados indican el parámetro elegido por prueba.....	35
Tabla IX. Individuos descartados por filtro.	36
Tabla X. Diversidad genética global para <i>E. puncticulatus</i> para cada una de las tres bases de datos generadas (completa, neutral y <i>outliers</i>). H_O : heterocigosidad observada, H_S : diversidad genética dentro de la población, H_T : diversidad genética general, F_{ST} : medida de diferenciación genética de Wright, FIS: coeficiente de endogamia de Wright.	40
Tabla XI. F_{ST} pareados entre sitios para marcadores neutrales para <i>E. puncticulatus</i> en el GC. Valores de F_{ST} debajo de la diagonal, y valores de P sobre la diagonal. Valores en negrita resultaron significativos después de la corrección de Bonferroni (valor de P con la corrección de Bonferroni= 0.001)(Rice, 1989).	49
Tabla XII. AMOVA jerárquico con marcadores neutrales de <i>E. puncticulatus</i> para la hipótesis de panmixia.	53
Tabla XIII. AMOVA jerárquico con marcadores neutrales de <i>E. puncticulatus</i> para la hipótesis de dos grupos (1)norte, y 2) sur.	54
Tabla XIV. AMOVA jerárquico con marcadores neutrales de <i>E. puncticulatus</i> para la hipótesis de tres grupos (1) Los Cabos La Paz y Loreto, 2) Mulegé, y 3) Bahía de los Ángeles.....	54
Tabla XV. AMOVA jerárquico con marcadores neutrales de <i>E. puncticulatus</i> para la hipótesis de cuatro grupos (1) Los Cabos y La Paz, 2) Loreto, 3) Mulegé, y 4) Bahía de los Ángeles.....	55
Tabla XVI. F_{ST} pareados entre sitios para marcadores <i>outliers</i> de <i>E. puncticulatus</i> . Valores de F_{ST} debajo de la diagonal, y los valores de p sobre la diagonal. Valores en negrita resultaron significativos después de la corrección de Bonferroni (valor de p con la corrección de Bonferroni ($p \leq 0.00098$))(Rice, 1989).	58
Tabla XVII. Amova jerárquico con marcadores <i>outliers</i> para hipótesis de panmixia.	62
Tabla XVIII. AMOVA jerárquico con marcadores <i>outliers</i> para hipótesis de dos grupos: 1) norte, y 2) sur.....	62

Tabla XIX. AMOVA jerárquico con marcadores neutrales para hipótesis de tres grupos: 1) Bahía de los Ángeles, 2) Mulegé, y 3) La Paz.	63
Tabla XX. AMOVA jerárquico con marcadores <i>outliers</i> para hipótesis de cuatro grupos: 1) Bahía de los Ángeles, 2) Mulegé, 3) Loreto, y 4) La Paz norte y La Paz sur.	63

Abreviaturas

ADN	Ácido desoxirribonucleico
ADPC	Análisis discriminante de componentes principales
AMOVA	Análisis Molecular de Varianza
BA	Bahía Algodones
BAFPA	Bahía Algodones, Frailes, y Punta Arenas
BIC	Criterio de Inferencia Bayesiano
BOT	Bota
CC	Cabeza de Caballo
F_{CT}	Índice de fijación entre grupos (Wright, 1943)
F_{IS}	Índice de fijación intrapoblacional (Wright, 1943)
F_{IT}	Índice de fijación total (Wright, 1943)
FRAI	Frailes
F_{SC}	Índice de fijación entre poblaciones dentro de grupos (Wright, 1943)
F_{ST}	Índice de fijación entre poblaciones (Wright, 1943)
GC	Golfo de California
G_{ST}	Índice de fijación entre poblaciones (Nei, 1987)
H_e	Heterocigosidad esperada
H_o	Heterocigosidad observada
H_S	Diversidad genética dentro de la población
H_T	Diversidad genética general
IAV	Islote Agua Verde
ISAD	Isla San Damián
ISANC	Isla San Cosme
ISANF	Isla San Francisquito
Km	Kilometro
MAR	Marcial
PA	Punta Arenas
PB	Punta Berrendo
PC	Componente principal
PCH	Punta Chivato
PCSMS	Punta Chivato, y San Marcos Sur
PD	Punta Diablo
PISF	Portugués, e Isla San Francisquito
POR	Portugués
s.d.	Desviación estándar
SALTI	Saltito Isla
SANJPS	San José Punta Sur
SMN	San Marcos Norte
SNP	Polimorfismo de un solo nucleótido

VEN Ventana
II Diversidad nucleotídica

1. INTRODUCCIÓN

Investigadores en el área de la genética de poblaciones han trabajado por años con marcadores neutrales localizados en pocos loci en el genoma de las especies de interés (Luikart *et al.*, 2003; López-de Heredia, 2016), los cuales nos han permitido inferir grupos génicos y eventos demográficos (Oleksiak y Rajora, 2019) sin embargo, queda la pregunta sobre qué tan representativos son sobre la variación genómica total de los individuos, y qué tan importante es el efecto de la selección sobre la diversidad y estructuración de las mismas (Eguiarte *et al.*, 2013). La genómica de poblaciones impulsada por el desarrollo de las técnicas de secuenciación masiva se ha vuelto popular en los últimos años para responder estas preguntas (Luikart *et al.*, 2003).

La secuenciación masiva nos permite la obtención de un gran número de marcadores a lo largo del genoma y de un gran número de individuos. Mientras que el desarrollo de nuevas herramientas bioinformáticas nos permite su análisis (Davey y Blaxter, 2010; Kirk y Freeland, 2011); así como la distinción entre marcadores bajo fuerzas neutrales y *outliers*, los cuales pueden ser considerados como candidatos a selección (Beaumont *et al.*, 1996, y Luikart *et al.*, 2003). Esto último nos permite comparar estructura y diversidad para cada tipo de marcador, y por ende inferir los tipos de procesos, ya sean neutrales o adaptativos, que están participando en el moldeado de la estructura espacial de las poblaciones (Riginos y Liggins, 2013, y Riginos *et al.*, 2016, Oleksiak y Rajora, 2019).

El Golfo de California (GC), ubicado al noroeste de México, entre el continente y la península de Baja California, es un ambiente altamente heterogéneo con una marcada variación espacial y temporal en su temperatura, productividad primaria y dirección de las corrientes (Lluch-Cota *et al.*, 2007; Ledesma *et al.* 2009; Marinone, 2012; Escalante *et al.* 2013), por lo cual representa un área de estudio ideal para observar los efectos de las fuerzas neutrales que limitan o promueven la dispersión de los organismos entre localidades, así como los efectos del ambiente sobre las poblaciones.

Por su parte, *Elacatinus puncticulatus* (Ginsburg, 1938), mejor conocido como el gobio de cabeza roja o cerillito, posee una amplia distribución dentro del GC. (Robertson, y Allen, 2015) Es un pez criptobentónico, que habita en los recovecos de los arrecifes rocosos y

coralinos. Durante su etapa adulta su capacidad dispersiva se ve reducida, sin embargo posee un estadio pelágico larval cuya dispersión se ve influenciada por la dirección de las corrientes (Galland, 2013; Robertson, y Allen, 2015). Estas características le confieren un fuerte potencial como especie modelo para estudios de genómica poblacional.

Este trabajo busca evaluar la estructura poblacional neutral y adaptativa de *E. puncticulatus* a lo largo del GC utilizando marcadores tipo SNP obtenidos con la técnica ddRADseq (Peterson *et al.*, 2012). Así como inferir los procesos que se ven involucrados en el moldeado de estos patrones en el espacio.

2. ANTECEDENTES

2.1 Genómica poblacional y secuenciación masiva

En genética, una población está conformada por un grupo de individuos, de una misma especie, los cuales tienen la posibilidad de reproducirse entre sí, y por lo tanto comparten el mismo acervo génico (Cabrero y Camacho, 2002). Las poblaciones son sistemas abiertos y dinámicos cuya evolución en el tiempo y el espacio puede ser observada a partir de cambios cuantificables en sus frecuencias alélicas y genotípicas (Cabrero y Camacho, 2002; y Hedgecock *et al.*, 2007). Entender cómo se estructuran las poblaciones es crítico para el manejo y conservación de las especies, y para ello las herramientas genéticas representan un método indirecto para estudiarlas (Hedgecock *et al.*, 2002).

A diferencia de la genética de poblaciones, que mide variación genética en pocos loci, la genómica de poblaciones busca polimorfismos a lo largo del genoma en individuos de poblaciones silvestres (Luikart *et al.*, 2003; Eguiarte *et al.*, 2013; López-de Heredia, 2016). Además, no solo permite la identificación y evaluación de efectos globales del genoma con marcadores neutrales, sino que posibilita la detección de efectos locus-específicos los cuales se denominan marcadores *outliers* o candidatos a selección (Luikart *et al.*, 2003; Oleksiak y Rajora, 2019). De esta manera la genómica de poblaciones nos permite conocer como el medio marino influye sobre la estructura poblacional y los procesos evolutivos de las especies, siendo este conocimiento crucial para la toma de decisiones y manejo de recursos naturales (Oleksiak, 2018; Oleksiak y Rajora, 2019).

Mientras que los marcadores neutrales han sido muy utilizados para evaluar la estructura genética y aspectos demográficos de las especies (Oleksiak y Rajora, 2019), los marcadores *outliers* pueden ser muy informativos, pues podrían estar asociados en una región del genoma que se encuentra bajo selección, o ligados a esta, por lo tanto nos brindan una idea de cómo la variación ambiental local moldea las poblaciones genéticas de las especies (Kirk y Freeland, 2011).

Existen dos visiones que explican la variabilidad genética intra-específica: la teoría evolucionista neutral que le atribuye los cambios moleculares a las fuerzas de evolución

neutrales (mutación, migración y deriva génica), y asume que los efectos de la selección natural serán poco frecuentes. Por otro lado, la visión seleccionista propone que la selección positiva, la cual permite la fijación de alelos benéficos, juega un papel significativo en las variaciones de las frecuencias alélicas de una población. Esta fuerza se presenta cuando un gen es influenciado por el ambiente o su nicho ecológico (Álvarez-Valín, 2000). Al analizar el genoma es posible observar que existen loci que varían en diferente grado con respecto a otros (Beaumont y Nichols, 1996; Eguiarte *et al.*, 2013). Se considera que regiones con una variación genética intermedia son genes neutros que responden a la historia evolutiva de la población, resultado de la deriva y migración (Eguiarte *et al.*, 2013). Por su parte loci con valores anómalos de diferenciación genética suelen asociarse a regiones que se encuentran bajo efecto de la selección natural (Beaumont y Nichols, 1996; Eguiarte *et al.*, 2013).

La secuenciación de nueva generación (NGS por sus siglas en inglés) o de alto rendimiento, son un conjunto de técnicas de secuenciación masiva que permiten la obtención de millones de fragmentos del ADN, los cuales posteriormente son analizados mediante softwares bioinformáticos (Behjati y Tarpey, 2013). Estas herramientas han impulsado el desarrollo de estudios de genómica poblacional; permitiendo un muestreo de fragmentos distribuidos a lo largo del genoma. Entre las ventajas que este tipo de herramientas nos ofrecen son precios accesibles, la posibilidad de obtener miles de marcadores distribuidos a lo largo del genoma en regiones que pueden o no ser neutras, y la posibilidad de trabajar con organismos no modelo pues no se requiere conocimiento previo del genoma de la especie (Eguiarte *et al.*, 2013). La técnica a emplear por cada investigador dependerá de los objetivos del trabajo, presupuesto, información previa del genoma, y características de la especie; pues cada técnica posee sus beneficios y limitantes (Andrew *et al.*, 2016).

Estudios de genómica poblacional pueden abordarse a partir de distintas aproximaciones, siendo una de ellas la secuenciación de una representación reducida del genoma, en donde se obtienen lecturas de aproximadamente 0.1-1% del total del genoma en un gran número de individuos (Oleksiak y Rajora, 2019), además de una alta cobertura de profundidad por locus. (Andrews *et al.*, 2016)

2.2 Técnicas de RAD-seq

Las técnicas de librerías de baja representación del genoma tipo la Secuenciación del ADN Asociado a Sitios de Restricción (RADseq por sus siglas en inglés) permiten la obtención de miles de marcadores a lo largo de los genomas de muchos individuos a precios reducidos, sin la necesidad de contar con un genoma de referencia; han ayudado en el desarrollo de estudios de genética de poblaciones. Las herramientas RADseq reducen la complejidad del genoma mediante el uso de enzimas de restricción, y con los datos obtenidos es posible identificar y genotipificar SNPs en un sólo paso (Davey y Blaxter, 2010; López-de Heredia, 2016; Andrews *et al.*, 2016). Esta técnica hace uso de la acción de una endonucleasa de restricción sobre el ADN para fragmentarlo al reconocer dianas de restricción poco frecuentes, cabe mencionar que el tamaño de los fragmentos generados puede variar, sin embargo, existen diferentes modificaciones de esta técnica (Davey y Blaxter, 2011).

Para este trabajo se utilizó la técnica de Secuenciación de Doble Digestión de ADN Asociada a Enzimas de Restricción (ddRADseq por sus siglas en inglés), desarrollada por Peterson *et al.* (2012), la cual es una variación del método de RADseq de Baird *et al.* (2008), la cual consiste en el uso de dos enzimas de restricción (una de corte frecuente y otra de corte no-frecuente) para la digestión de ADN, para evitar el corte aleatorio del ADN en los sitios adyacentes a la enzima. Los fragmentos distribuidos aleatoriamente a lo largo del genoma, son seleccionados por tamaño, de esta manera se obtiene una mejor representación de las regiones secuenciadas. Además, el uso de la combinación de dos *index* para identificar a los individuos favorece la secuenciación multiplex, por lo cual la elaboración de librerías puede llevarse a cabo en microplacas. Este método no requiere de un genoma de referencia, es relativamente económico, rápido y no necesita una gran cantidad de material genómico inicial (100 ng de ADN) (Peterson *et al.*, 2012).

2.3 Estudios de estructura genómica poblacional para especies de ambientes marinos costeros

Los paisajes marinos costeros son ambientes heterogéneos y altamente dinámicos debido a que son una interfaz entre la masa continental, oceánica y la atmosférica (Silva-Casarín *et al.*, 2017). Estas características los vuelven sitios interesantes para el estudio de

poblaciones, pues diversos estudios han encontrado que las características ambientales, junto con las características biológicas, e historia demográfica de las especies juegan un papel importante en la estructura poblacional (Riginos y Liggins, 2013; Riginos *et al.*, 2016), en donde corrientes oceánicas y gradientes ambientales pueden beneficiar el transporte de organismo o actuar como barreras al flujo génico.

Estudios alrededor del mundo han encontrado patrones de estructuración poblacional diferentes entre marcadores neutrales y *outliers*, inclusive en especies con un elevado flujo genético o panmixia se han detectado niveles de divergencia inesperados con *outliers* en ambientes costeros debido a la heterogeneidad ambiental entre localidades, superiores a los observados con marcadores neutrales (Sandoval-Castillo *et al.*, 2018; Wyngaarden *et al.*, 2016; Limborg *et al.*, 2012; DeFaveri *et al.*, 2013; Teske *et al.*, 2019). Estas diferencias muchas veces son asociadas a la capacidad de los organismos marinos para desplazarse grandes distancias, las características ambientales, y la permeabilidad de las barreras (Riginos y Liggins, 2013).

Diversos trabajos han encontrado que la estructura poblacional es el resultado no solo de variables que limitan la dispersión de los organismos, las cuales estarían actuando sobre marcadores neutrales, sino que las características locales del ambiente podrían jugar un rol importante actuando como barreras físicas en los océanos y modificando los patrones de variación genética de loci *outlier* (Riginos y Liggins, 2013; Oleksiak y Rajora, 2019).

Entre las variables más importantes y recurrentes que son consideradas en este tipo de estudios para explicar la estructura genética de las poblaciones, se puede mencionar la temperatura (Oleksiak, 2018), donde su influencia ha sido observada en especies como el bacalao del Atlántico (*Gadus morhua*), la langosta americana (*Homarus americanus*), la vieira del Atlántico (*Placopecten magellanicus*), el camarón del norte (*Pandalus borealis*) y el cangrejo verde europeo (*Carcinus maenas*), las cuales son todas especies encontradas en el Atlántico noroccidental, que poseen historias de vida distintas; sin embargo todas mostraron un quiebre genético, aproximadamente a la misma latitud, resultado de diferentes variables (productividad, límites entre ecorregiones, sistema de corrientes, historia de vida pelágica, y eventos vicariantes), pero principalmente asociado con el mínimo de temperatura de esta región (Stanley *et al.*, 2018).

Otro caso similar donde la temperatura jugaba un papel importante en la variación genética fue en la costa del sur de África, donde se observó un quiebre genético asociado a la temperatura para el gobio de arena de Knysna (*Psammogobius knysnaensis*). Para ello utilizaron genes bajo selección asociados a la temperatura, sugiriendo que la capacidad adaptativa de las especies a los gradientes termales juega un papel importante en la estructura genética poblacional, y posteriormente en la especiación (Teske *et al.*, 2019).

De igual manera se ha encontrado que otras variables además de la temperatura podrían poseer alguna influencia sobre la estructura poblacional de las especies marinas costeras, como lo son las clinas o gradientes de salinidad, oxígeno y profundidad (Oleksiak y Rajora, 2019). Sandoval-Castillo *et al.* (2018) encontraron una metapoblación de *Haliotis laevigata* a lo largo de la costa del Sur de Australia utilizando marcadores neutrales, pero marcadores candidatos a la selección indicaron la existencia de cinco grupos, lo cual se correlacionó a las diferencias en la temperatura superficial y concentración de oxígeno, además de coincidir con regiones ambientales (basadas en catorce variables oceanográficas).

Gasterostetus aculeatus, en el Mar Báltico, posee poblaciones diferenciadas genéticamente a través de gradientes, de salinidad y temperatura (Guo *et al.*, 2015); así como en el trabajo de do Prado *et al.* (2018) en el Mar Báltico y el Mar Negro, encontraron evidencia de que en regiones con características ambientales similares (temperatura y salinidad), poblaciones de rodaballo (*Scophthalmus maximus*), podían estar presentando una evolución paralela a pesar de la distancia entre las cuencas.

Además de los gradientes de factores fisicoquímicos, también se pueden observar asociaciones a gradientes de índole biológica, como lo es la productividad. Para individuos de *Amphiron bicinctus* y *A. omanensis*, en la región del Mar Rojo, Golfo de Adén y la costa de Omán, se encontraron poblaciones aisladas tanto por distancia y como por ambiente, asociado a un fuerte gradiente en la productividad primaria (Saenz-Agudelo *et al.*, 2015).

De ahí la importancia de considerar marcadores *outliers*, que podrían ser candidatos a selección, para la delimitación de *clusters* génicos y entender que variables poseen una influencia sobre la estructuración de poblaciones costeras (Riginos y Liggins, 2013; Teske *et al.*, 2019).

2.4 Características ambientales de la región

Para evaluar cómo se estructuran las poblaciones marinas es importante entender cómo se comportan las variables ambientales y oceanográficas de la zona de estudio. Las características de la dinámica oceánica pueden generar un transporte de individuos entre una localidad a otra favoreciendo el flujo génico y creando poblaciones abiertas bien conectadas, o por el contrario representar una barrera que limite parcial o totalmente el flujo génico, provocando cambios en las variaciones alélicas por medio de la deriva génica o la selección. A la vez cambios clinales en las frecuencias alélicas de las poblaciones marinas a lo largo del espacio puede ser indicio de que la selección está jugando un fuerte papel en la determinación de poblaciones, incluso en presencia de flujo genético (Hellberg *et al.*, 2002).

El GC, se caracteriza por ser una cuenca semicerrada formada hace 5.6 millones de años debido a la deriva y separación de la península del resto del continente, resultado del desplazamiento de la tectónica de placas (placa de Norte América y Placa del Pacífico). Mide aproximadamente 1400 km de largo, sin embargo es relativamente estrecha alcanzando los 150-200 km de ancho, con un máximo de 800 m de profundidad (Álvarez *et al.*, 2009; Lavín y Marinone, 2003).

Las fuerzas que dirigen la dinámica oceánica del Golfo de California son el forzamiento del Océano Pacífico en la boca del Golfo de California, influyendo en la distribución de las masas de agua dentro de la cuenca, así como generando una onda de Kelvin que recorre la costa del GC en sentido ciclónico (Lluch-Cota *et al.*, 2007); los flujos de calor con ganancias de calor neto a través del calentamiento solar en la superficie marina, y con pérdidas netas durante los meses de noviembre y diciembre en la región norte (Lavín y Marinone, 2003). Sin embargo, se sugiere que el forzamiento más importante es el que ejerce el viento, durante el verano estos son débiles y vienen desde el sureste, por lo cual el transporte del agua es hacia el norte y se generan surgencias del lado de la península. Mientras que durante el invierno se tienen nortes fuertes, lo cual genera surgencias del lado del continente (Molina-Cruz, 1984; Lluch-Cota *et al.*, 2007; Marinone, 2012).

El Sur del Golfo está dominado por fuertes corrientes en el lado del continente, lo cual genera una baja retención de partículas, pero esta corriente es débil del lado de la península y por el contrario es una zona de potencial autoreclutamiento (Molina-Cruz, 1984; Lluch-Cota *et al.*, 2007; Marinone, 2012). El GC es una región con características oceanográficas altamente dinámicas. Marinone (2012), realizó un estudio de conectividad en el GC mediante un modelo tridimensional numérico para calcular la trayectoria de partículas por medio de un campo de velocidades eulerianas. Se sugiere que existen dos principales zonas de retención de partículas: el Alto Golfo y Sonora, debido a la presencia de un giro de escala, y la región sur del GC en el lado península, ya que existen corrientes estacionales débiles. Mientras que la costa continental centro y sur de la península, cuando las corrientes estacionales son fuertes, podrían considerarse como áreas de dispersión alta.

Las características oceanográficas del golfo ejercen una fuerte influencia sobre las comunidades biológicas. Esta región se caracteriza por una alta diversidad de especies marinas, asociada a una alta productividad que se le atribuye a tres principales mecanismos: surgencias, mareas y el intercambio de aguas con el resto del Océano Pacífico. La productividad muestra una marcada estacionalidad teniendo altas concentraciones durante invierno, asociada a las surgencias costeras en el Alto Golfo, y también en la Región de las Grandes Islas, aunque la productividad en esta zona se asocia a otros procesos relacionados con las mareas y su geomorfología. En cambio, durante verano las concentraciones de clorofila α disminuyen (Lluch-Cota *et al.*, 2007). En la costa occidental del GC del muestreo las surgencias solo ocurren durante verano cuando los vientos cambian de dirección hacia el sur (Alvarez-Borrego, 2010; Marinone, 2012).

A diferencia del patrón de concentración de clorofila α , la TSM presenta los valores más bajos durante invierno (17-19°), y estos aumentan durante verano (28-31°C). Pero no solo existe una variación temporal, si no que la TSM se caracteriza por presentar valores que disminuyen de sur a norte; sin embargo se ha encontrado que la región con las temperaturas más bajas se observa en la Región de las Grandes Islas (RGI) (Ledesma *et al.*, 2009; Lluch *et al.*, 2007; Soto- Mardones *et al.*, 1999). La mayor variación de TMS es encontrada hacia el interior del GC, mientras que hacia la boca se vuelve menos variable (Soto-Mardones *et al.*, 1999)

2.5 Biología del gobio cabeza roja (*Elacatinus puncticulatus*)

Elacatinus puncticulatus (Ginsburg, 1938), también conocido como gobio cabeza roja o cerillito (Fig. 1), es un pez criptobentónico óseo arrecifal perteneciente a la familia de los Gobiidae (Robertson, y Allen, 2015). Se caracteriza por presentar un cuerpo alargado de aproximadamente 4.6 cm, con un incremento de tamaño en sitios más norteños; presenta una cabeza roja y un cuerpo translúcido con coloraciones negras, y durante la etapa de cortejo, se han registrado que machos presentan cambios desde su cabeza a la región oral a una coloración oscura (Pedrazzani *et al.*, 2014; Robertson, y Allen, 2015; Trujillo-García, 2018).



Figura 1 Gobio cabeza roja (*Elacatinus puncticulatus*) (Fotografía: El Khatabbi-Salazar, 2019).

Este pez habita en arrecifes de coral y arrecifes rocosos, vive principalmente dentro de grietas, a profundidades entre los 1-21 m, y se distribuye desde el Golfo de California hasta el norte de Perú (Robertson y Allen, 2015; Trujillo-García, 2018), con un centro de distribución a los 13.25°N (Galland, 2013). Juega un papel ecológico como carnívoro (se alimenta de gusanos y crustáceos sésiles) y eliminador de ectoparásitos facultativo (Pedrazzani *et al.*, 2014; Robertson y Allen, 2015; Quimbayo y Zapata, 2018).

Elacatinus puncticulatus es una especie hermafrodita (Shei *et al.*, 2017) e iterópara, con largas temporadas reproductivas (Trujillo-García, 2018). Produce huevos demersales y, una vez fertilizados, el macho no se alimenta y se dedica exclusivamente a su aireación (Pedrazzani *et al.*, 2014). Una pareja puede desovar, aproximadamente, entre 200-300

huevos por puesta, con una frecuencia de 7 días. Se han observado pérdidas de los huevos debido a conductas de limpieza del macho (hasta el 50%) y canibalismo parcial o total por parte de los padres (Pedrazzani et al., 2014, y Trujillo-García, 2018). Se trata de una especie con una historia de vida bipartita, con una etapa larval con duración de 36-42 días (Shei et al., 2017), y por lo tanto un fuerte potencial dispersivo. En su etapa adulta se encuentra asociado al bentos con una baja capacidad de dispersión. Cabe mencionar que estas características dispersivas son similares a la mayoría de los peces arrecifales en el GC (Bernardi et al., 2003).

El valor comercial de *Elacatinus puncticulatus* recae solamente sobre la pesca ornamental. Se tiene registrado que durante el 2010-2011 *E. puncticulatus* fue extraído por la cooperativa "Mujeres del Golfo", con una cuota de extracción de 1309 individuos anuales y un valor de 30.25 pesos por individuo a pie de playa (Germain et al., 2015).

Entre los estudios similares a este trabajo se puede mencionar el de Riginos (2005), que dirigieron su trabajo hacia la genética de poblaciones, utilizando como marcador la región control del ADN mitocondrial, encontrando diferencias genéticas significativas entre poblaciones de cinco especies de peces criptobentónicos del Golfo de California, asociado a dos grandes regiones biogeográficas (norte y centro del Golfo) diferencias que fueron atribuidas a las características ambientales y distancias geográficas.

Finalmente, un trabajo precedente de gran importancia es el de Sandoval-Huerta et al. (2019), quienes buscaron describir la historia evolutiva de *E. puncticulatus* a lo largo de su rango de distribución en el Pacífico Tropical occidental a través de un marcador mitocondrial y dos nucleares. Encontraron que la población más ancestral fue la del Golfo de Panamá y el Choco (Colombia), y descubrieron la existencia de tres especies alopátricas putativas. Los individuos del Golfo de California junto con individuos de la costa mexicana serían una de estas especies, pues fueron encontradas dentro de un mismo subclado al calcular un árbol filogenético para la especie. Es interesante mencionar que encontraron un Φ_{ST} significativo entre ambas poblaciones, y que aparentemente solo detectaron flujo génico de norte a sur, a pesar de que otros trabajos han encontrado que existe un flujo bidireccional entre el Golfo de California y el resto de la costa mexicana en el Pacífico Tropical. De existir un flujo bidireccional de larvas de *E. puncticulatus*, los autores

sugieren que las relaciones genéticas estarían asociadas fuertemente a la adaptación local. Este descubrimiento es de particular importancia, ya que la presente tesis busca evaluar la estructura poblacional de *E. puncticulatus* en la costa occidental del GC, así como inferir que procesos y mecanismos están involucrados en el moldeado de las mismas.

3. JUSTIFICACIÓN

El Golfo de California, al ser un ambiente altamente dinámico y heterogéneo, con una marcada estacionalidad, representa un área de estudio ideal para la evaluación de la importancia de las características ambientales y oceanográficas que moldean las diferencias genómicas entre poblaciones marinas, tanto en loci neutrales como *outliers*. Para ello se considera que *Elacatinus puncticulatus*, nuestra especie de estudio, posee un gran potencial para ser utilizada como especie modelo debido a su amplio rango de distribución a lo largo de un ambiente heterogéneo, siendo el Golfo de California su límite más norteño; su baja movilidad en su etapa adulta, pero con potencial de dispersión en su etapa larval; además de existir en esta zona numerosas especies con historias de vida similares.

4. HIPÓTESIS

Asociado a la heterogeneidad del paisaje y el amplio rango de la zona de estudio, se espera encontrar diferenciación genética en las poblaciones de *E. puncticulatus* a lo largo del Golfo de California, y que la estructura de loci neutrales se correlacione a variables que limiten su flujo génico, como la dirección de las corrientes, la capacidad de dispersión del organismo (en cualquier etapa de su vida), las distancias geográficas, y eventos de vicarianza; mientras que loci candidatos a selección podrían responder a variables que generen una presión selectiva, como la heterogeneidad del hábitat (ej. variaciones espaciales de la temperatura, clorofila α , y salinidad).

5. OBJETIVOS

5.1 Objetivo general

Determinar la diversidad y estructura genómica del gobio *Elacatinus puncticulatus* a lo largo del Golfo de California

5.2 Objetivos específicos

- 1) Identificar genes neutrales y candidatos a selección en el gobio *Elacatinus puncticulatus*.
- 2) Determinar estructura poblacional neutral y putativamente adaptativa.

6. MATERIALES Y METODOS

6.1 Muestreo

Se obtuvieron un total de 312 muestras del tejido de *E. puncticulatus* (Tabla I), colectadas en 20 sitios a lo largo de la costa oeste del Golfo de California y una del lado del Pacífico al sur de la península (Fig. 2). El muestreo se llevó a cabo de junio a noviembre de 2017, y de cada captura se registró el tipo de sustrato en el que se ocultaban los peces: 1) arrecife rocoso, 2) arrecife coralino, y 3) pared rocosa. Se realizaron de una a tres inmersiones por sitio y para la obtención de los organismos se utilizó una solución de aceite de clavo-etanol (3:1) como anestésico, el cual era vertido sobre el sustrato (1x1 m) donde los peces se refugiaban, para proseguir a la captura con redes (Fig. 3). Inmediatamente después de su colecta fueron preservados en etanol al 96%, para después ser almacenados a -20 °C.

Tabla I. Número de individuos de *Elactinus puncticulatus* recolectados por sustrato en cada sitio a lo largo del Golfo de California.

Sitio	AR ¹	AC ²	PR ³	Total
Frailes	2	0	0	2
Bahía Algodones	1	0	0	1
Punta Arenas	1	0	1	2
Saltito Isla	5	0	2	7
Punta Diablo	0	0	9	9
Portugués	1	0	0	1
Isla San Francisquito	10	0	0	10
Pardito	0	2	0	2
San José Punta Sur	4	0	0	4
Punta Berrendo	17	0	0	17
Marcial	10	0	0	10
Islote Agua Verde	5	0	0	5
Isla San Damián	7	5	8	20
Isla San Cosme	7	0	6	13
Punta Chivato	2	0	0	2
San Marcos Sur	15	0	2	17
San Marcos Norte	42	0	83	125
Cabeza de Caballo	29	0	0	29
Ventana	22	0	0	22
Bota	14	0	0	14
Total	194	7	111	312

Notas: ¹arrecife rocoso, ²arrecife coralino, ³pared rocosa.

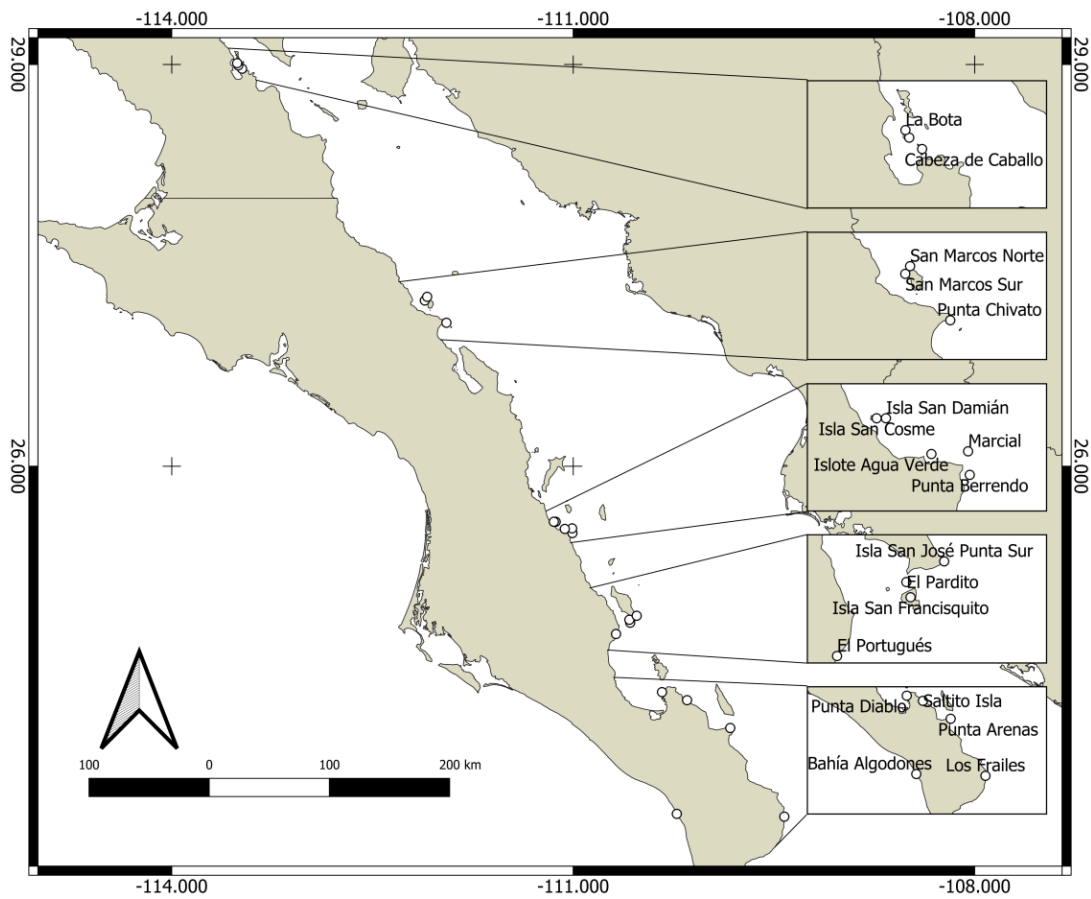


Figura 2. Puntos de muestreo de *E. puncticulatus* en el Golfo de California (INEGI *et al.*, 1990).



Figura 3. Muestreo de *E. puncticulatus* en la costa oriental de la península de Baja California, mediante la técnica de buceo autónomo, con el uso de aceite de clavo como anestésico y redes para su captura. Fotografía: Valenzuela-Quiñonez, 2017.

6.2 Trabajo de laboratorio

6.2.1 Extracción de ADN

A partir de aproximadamente 100 mg de tejido (de aleta caudal y/o músculo) se realizaron extracciones de ADN con el método de fenol-cloroformo alcohol isoamílico (24:1) (Sambrook y Russell, 2006). El ADN fue cuantificado mediante un espectrofotómetro NanoDrop 2000 (Thermo Scientific) y la pureza fue evaluada con los valores de relaciones de absorbancia (A260/A280). La integridad del ADN se evaluó con la presencia de una banda de alto peso molecular mediante corridas electroforéticas en geles de agarosa al 1%. Para la elaboración de librerías el ADN fue cuantificado mediante el fluorómetro Qubit (Thermo Fisher Scientific).

6.2.2 Elaboración de librerías ddRAD-Seq

A partir de ADN de alta calidad de 144 individuos de *E. puncticulatus* (Tabla II), se construyeron librerías genómicas por la técnica de Secuenciación de ADN Asociado a Sitios de Restricción de Doble Digestión (ddRADseq, por sus siglas en inglés) (Peterson *et al.*, 2012). Las librerías se prepararon con un total de 500 ng iniciales de ADN en un

volumen de 20 μ l por individuo sobre dos placas de 96 pozos, utilizando una réplica por placa para estimar errores de secuenciación.

Tabla II. Individuos de *E. puncticulatus* utilizados para la construcción de librerías por sitio.

Sitio	Total
FRAI ¹	1
BA ²	1
PA ³	2
SALTI ⁴	6
PD ⁵	10
POR ⁶	1
ISANF ⁷	10
PAR ⁸	2
SANJPS ⁹	4
PB ¹⁰	10
MAR ¹¹	10
IAV ¹²	5
ISAD ¹³	10
ISANC ¹⁴	10
PCH ¹⁵	2
SMS ¹⁶	15
SMN ¹⁷	15
CC ¹⁸	10
VEN ¹⁹	10
BOT ²⁰	10
Total	144

Notas: ¹Frailles, ²Bahía Algodones, ³Punta Arenas, ⁴Saltito Isla, ⁵Punta Diablo, ⁶Portugués, ⁷Isla San Francisquito, ⁸Pardito, ⁹San José Punta Sur, ¹⁰Punta Berrendo, ¹¹Marcial, ¹²Islote Agua Verde, ¹³Isla San Damián, ¹⁴Isla San Cosme, ¹⁵Punta chivato, ¹⁶San Marcos Sur, ¹⁷San Marcos Norte, ¹⁸Cabeza de Caballo, ¹⁹Ventana, ²⁰Bota.

A continuación, se describen las etapas de la técnica de ddRADseq.

6.2.2.1 Doble digestión

Se realizó la digestión de ADN mediante el uso de dos enzimas de restricción de corte frecuente MspI (New England BioLabs®) y poco frecuente EcoRI-HF (New England BioLabs®). Para cada individuo se llevó a cabo la reacción en un volumen final de 30 μ l,

en donde se encontraban contenidos los 500 ng de ADN inicial, 1 X Buffer CutSmart (NEB), 10 U de EcoRI-HF (NEB) y 10 U de MspI (NEB), y se aforó con agua libre de endonucleasas. Las reacciones fueron incubadas a temperatura ambiente durante toda la noche.

El ADN digerido se purificó añadiendo 1.5X perlas magnéticas AMPure XP beads (Beckman Coulther). La placa se colocó sobre una placa magnética, se esperó a la separación de fases y se removió el sobrenadante. Posteriormente se realizaron dos lavados con 150 µl de etanol al 70%. Una vez que se removió el etanol, se esperó a que las perlas quedaran completamente secas, y el ADN se eluyó en 30 µl de TE 0.1X.

Ligación de adaptadores. Para identificar a los individuos, a cada uno de los productos digeridos se ligaron adaptadores P1(*barcode*), que contiene 1 de 48 barcodes, y es complementario al sitio de restricción de la enzima EcoRI; y P2 (*common*), que es común para todos los fragmentos, y es complementario al sitio de restricción de MspI. Cada reacción de ligación se llevó a cabo en un volumen final de 40 µl, en donde contenían 1X Buffer T4 ligasa (NEB), 100 U de T4 ADN ligasa (NEB), 0.1 µM del adaptador P2, y 0.1 µM del adaptador P1 correspondiente. Las reacciones se incubaron a temperatura ambiente durante toda la noche, finalizando con una incubación a 65 °C por 10 minutos en termociclador BioRad T100 (BioRad). Nuevamente se realizó una purificación de las reacciones con 1.5X de perlas magnéticas AMPure XP beads.

Verificación de ligación. Para comprobar la correcta ligación de los adaptadores, se llevó a cabo un PCR con adaptadores Illumina F (5'-AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACG-3') y R (5'-CAAGCAGAAGACGGCATAACGAGATCGTGATGTGACTGGAGTTCAGACGTGTGC-3'). El volumen final de cada reacción fue de 12.5 µl, en los cuales se encontraban contenidos 1X Go Taq (Promega), 0.8 µM de cada *primer*, 1 µl de ADN ligado, y se llevó al volumen final con agua libre de endonucleasas.

Las amplificaciones se llevaron a cabo en un termociclador T100 (BioRad) con las siguientes condiciones de PCR: una desnaturalización inicial de 94 °C/2 min y 20 ciclos de

94 °C/2 min, 60 °C/30 s, y 72 °C/45 s. Los productos se verificaron en un gel de agarosa al 2% y se reveló con SYBER Gold.

Una vez confirmada la ligación se realizó una cuantificación de ADN en Qubit (Thermo Fisher Scientific) de todos los individuos para conocer su concentración y poder realizar cuatro *pools* equimolares con un máximo de 48 barcodes únicos por *pool*. Para cada *pool* se utilizaron 12.5 ng de ADN. Posteriormente cada *pool* fue purificado con 1.5X perlas magnéticas y un único lavado con etanol al 70% para concentrarlos en un volumen final de 30 µl.

6.2.2.2 Selección de tamaño de fragmentos

A partir de los *pools* se realizó una selección de tamaño automatizada, mediante el uso del Pippin Prep (Sage Science). Ya que el tamaño de interés es de 376 pb (300 pb del fragmento de ADN, y 76 pb de los adaptadores), se programó para que la selección tuviera un rango amplio de 300 a 400 pb con un *target* de 350 pb. Finalmente los fragmentos de ADN recuperados de cada *pool* fueron eluidos en un volumen de 40 µl.

6.2.2.3 Indexado de pools.

Para optimizar reacciones multiplex, cada *pool* se amplificó a través de ocho reacciones de PCR utilizando *primers* Illumina indexados. Cada reacción se llevó a cabo en un volumen de 20 µl, el cual contenía 3 µl de cada *pool* de ADN, 1X Buffer Phusion HF (NEB), 0.4 µM dNTPs, 0.5 U de TAQ Phusion ADN polimerasa (NEB), y 0.5 µM de *Primer* Illumina F, y 0.5 µM *Primer* Illumina R con el index correspondiente (Tabla III). Las amplificaciones se llevaron a cabo en un termociclador T100 (BioRad) con las siguientes condiciones de PCR: desnaturalización inicial de 94 °C/2 min, seguida de 12 ciclos de 94 °C/30 s, 60 °C/30 s, y 72 °C/45 s.

Tabla III. *Pools* indexados de *E. puncticulatus*.

<i>Pool</i>	No. de individuos	Index	Secuencia
EPP1I7	32	7	ATCACG
EPP2I1	48	1	CAGATC
EPP3I8	48	8	ACTTGA
EPP4I9	17	9	GATCAG

Las ocho reacciones de cada *pool* se combinaron para crear el *pool* final, éste se volvió a purificar con 1.5X de perlas magnéticas, con dos lavados de etanol al 70% y se eluyó en 30 µl de agua libre de endonucleasas. Después se verificó la amplificación en un gel de agarosa 1%, y se reveló con el intercalante SYBER Gold (Thermo Fisher Scientific), en el cual se esperaba un tamaño de banda de 475 pb (300 pb del fragmento de ADN, 76 pb de los adaptadores, 53 pb *primer* F, 46 pb *primer* R).

Se cuantificó cada uno de los *pools* con Qubit (Thermo Fisher Scientific) para después formar la librería final la cual consistió en la mezcla de los cuatro *pools* con concentraciones equimolares y se envió a secuenciar a NOVOGENE© Illumina Hi Seq X con lecturas pareadas de 150 pb.

6.3 Análisis de secuencias y obtención de genotipos

Las lecturas crudas fueron procesadas en el pipeline STACKS V2 (Catchen *et al.*, 2013). Primero se hizo un filtrado y demultiplexado de las secuencias con '*process_radtags*'. Esto para corregir errores y eliminar aquellas cuya probabilidad de ser correctas sea menor al 90% (phred<10) o posean barcodes y sitios de corte RAD ambiguos. Para reducir posibles errores de secuenciación al extremo de las secuencias, las lecturas se cortaron a los 140 pb (-t 140). A partir de este primer filtrado, se eliminaron los individuos con < 950,000 lecturas.

Debido a que se carece de un genoma de referencia para la especie de estudio se realizó un ensamble *de novo* de las lecturas. Se utilizó el programa *denovo_map.pl* en STACKS V2.0, el cual ejecuta de manera consecutiva los diferentes componentes que integran el núcleo de STACKS para su filtrado, genotipificado, y obtención de loci y sitios variantes. Para ello

los valores de los parámetros utilizados por sus distintos componentes deben ser optimizados basándose en el número de loci ensamblados, el número de sitios variantes y la cobertura de profundidad (x) obtenidos con diferentes valores, así también se recomienda una cobertura superior a los 25x, ya que la tasa de error es alta para coberturas entre 5-10x (Paris *et al.*, 2017), por lo que para este trabajo se permitió una cobertura mínima de 15x.

El primer componente por el que pasan las lecturas es *ustacks*, en donde se realiza un alineamiento de las lecturas de cada individuo en alelos putativos (*stacks*), para determinarlos se realizaron pruebas con el parámetro *-m*, que determina el número de profundidad de cobertura mínima (entre 3-5) (Tabla IV). Después los alelos putativos son comparados para formar loci putativos, por lo que se realizaron pruebas con *-M*, que es el número máximo de diferencias entre alelos putativos (entre 3-5), finalmente usando un método de máxima verosimilitud el programa detectó SNPs. Con *cstacks* se buscó crear un catálogo de los loci, por lo que se optimizó el valor de *-n*, que es el número máximo de diferencias entre loci putativos ($-n = -M$ ó $-n = -M \pm 1$). Finalmente en *sstacks* se compararon los loci putativos contra los catálogos construidos y se buscaron coincidencias.

Tabla IV. Valores optimizados de los parámetros utilizados en *denovo_map.pl*.

Parámetro	Componente de STACKS	Valor optimo
<i>-m</i>	'ustacks'	4
<i>-M</i>	'ustacks'	3
<i>-n</i>	'cstacks'	3

En el módulo de *populations*, se realizó el filtrado de loci y sitios variantes. Los sitios retenidos fueron aquellos que presentaron una frecuencia alélica mínima de 0.05 (*--min-maf* 0.5), heterocigosidad máxima de 0.5 (*--max-obs-het*) para evitar la formación de loci parálogos, y solamente se permitió un SNP por locus (*--write-single-snp*) ya que estos tienen mayores posibilidades de tratarse de marcadores ligados. Solamente se consideraron loci presentes en más del 80% de los individuos (*-r* 0.8). Para el número de poblaciones mínimas en las que un locus debe encontrarse para ser considerado (*-p*), se realizaron una serie de pruebas.

Con VCF tools v.3.0 (Danecek *et al.*, 2011) se identificaron individuos con un porcentaje de missing data >25%, para después ser eliminados. De esta manera se buscaba obtener la mayor cantidad de loci y sitios variantes, con el mayor grado de información retenida para cada individuo.

Después del filtrado de loci e individuos, para este trabajo solo se retuvieron un total de 127 individuos (Tabla V), por lo cual sitios cercanos se reagruparon para aumentar el número de muestras por sitio (Figura 4), dejando un total de 15 sitios.

Tabla V. Individuos finales de *E. puncticulatus* utilizados en este estudio pertenecientes a los sitios reagrupados.

Sitio	n
BAFPA (BA, FRAI, PA) ¹	4
SALTI ²	5
PD ³	9
PISF ⁴	10
SANJPS ⁵	4
PB ⁶	10
MAR ⁷	4
IAV ⁸	10
ISAD ⁹	10
ISANC ¹⁰	8
PCSMS (PC, SMS) ¹¹	14
SMN ¹²	11
CC ¹³	9
VEN ¹⁴	10
BOT ¹⁵	9
Total	127

Notas: ¹Bahía Algodones, Frailes, Punta Arenas; ²Saltito Isla, ³Punta Diablo, ⁴Portugués, Isla San Francisquito; ⁵San José Punta Sur, ⁶Punta Berrendo, ⁷Marcial, ⁸Islote Agua Verde, ⁹Isla San Damián, ¹⁰Isla San Cosme, ¹¹Punta Chivato, San Marcos Sur; ¹²San Marcos Norte, ¹³Cabeza de Caballo, ¹⁴Ventana, ¹⁵Bota.

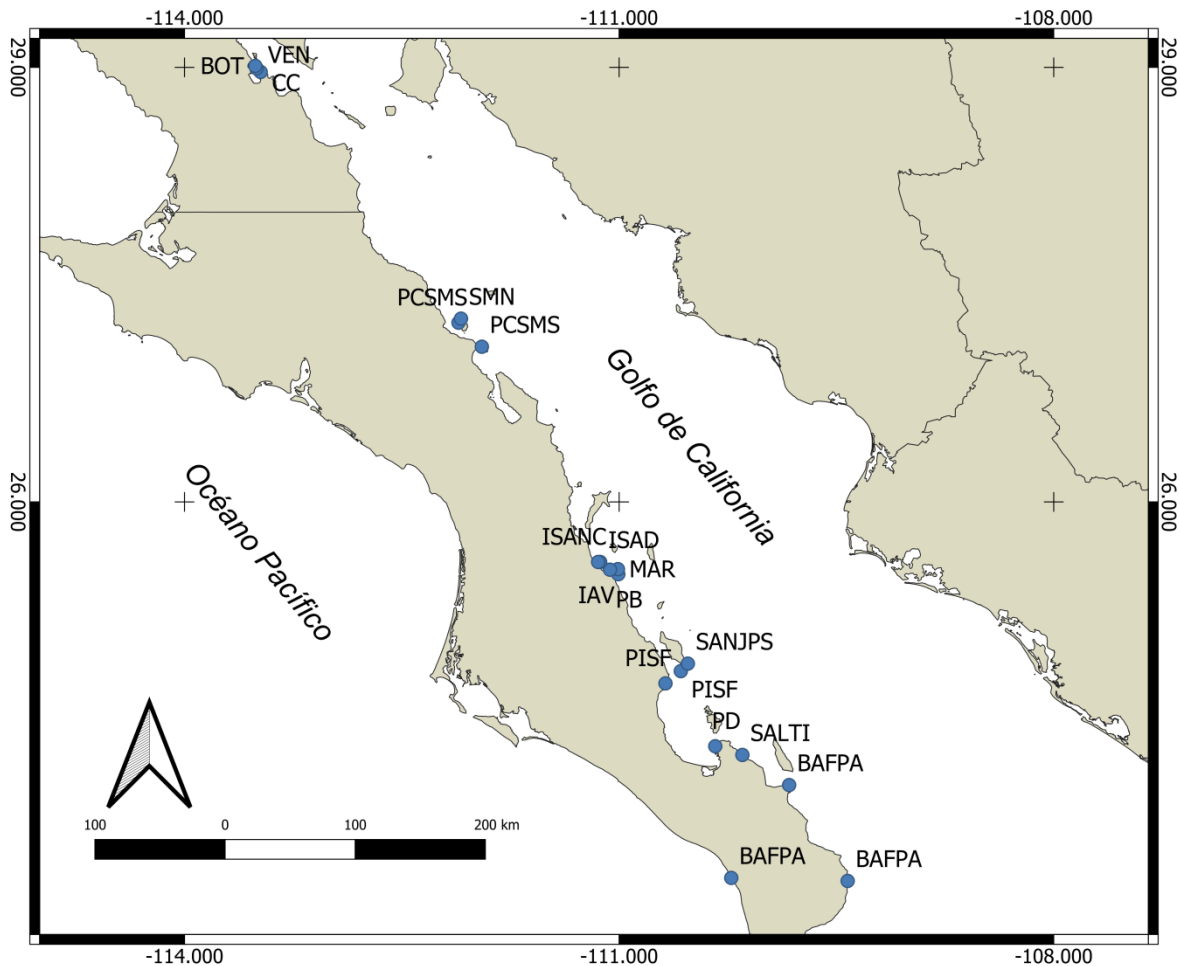


Figura 4. Área de estudio (INEGI *et al.*, 1990). BAFPA: Bahía Algodones, Frailes y Punta Arenas; BOT: La Bota, CC: Cabeza de Caballo, IAV: Islote Agua Verde, ISAD: Isla San Damián, ISANC: Isla San Cosme, MAR: Marcial, PB: Punta Berrerndo, PCSMS: Punta Chivato, y San Marcos Sur, PD: Punta Diablo, PISF: El Portugués, e Isla San Francisquito; SALT: Saltito Isla, SANJPS: San José Punta Sur, SMN: San Marcos Norte, y VEN: Ventana.

6.4 Detección de *outliers*

Los marcadores *outliers* pueden ser identificados como candidatos a estar bajo selección, por lo cual es importante su distinción de los marcadores neutrales. La detección de *outliers* se llevó a cabo a partir de dos métodos. El primero se basa en estadísticos de F bajo un modelo jerárquico de islas en ARLEQUIN v.3.5 (Excoffier y Lischer, 2010) Se corrieron 100,000 simulaciones coalescentes y 100 simulaciones por grupo (deme). Para este análisis solo se consideraron como *outliers* los marcadores con una $p \leq 0.01$.

Como una segunda aproximación a la detección de *outliers* se utilizó el método bayesiano implementado en Bayescan versión 2.1 (Foll y Gaggiotti, 2008), el cual consta de 20 corridas piloto, 10,000 iteraciones y 100,000 corridas de calentamiento, con una probabilidad anterior de 100. Se consideraron como marcadores *outliers* aquellos que presentaron una probabilidad posterior de 0.99. Los resultados se graficaron en el lenguaje de programación R versión 3.6.1 (R Core Team, 2019), utilizando la función *plot_bayescan* dentro del archivo *plot_R.r* proporcionado por Bayescan (Foll y Gaggiotti, 2008).

Para el set de marcadores *outliers* final, solo se tomaron en cuenta los marcadores detectado como *outliers* en ambos programas. De esta manera se pudieron distinguir tres sets de marcadores: 1) el completo (incluye todos los marcadores neutrales y *outliers*), 2) el de neutrales, y 3) el de *outliers*.

6.5 Prueba de equilibrio de Hardy-Weinberg

Se realizó la prueba de equilibrio de Hardy-Weinberg en Arlequin v.3.5 (Excoffier y Lischer, 2010) con 1,000,000 de iteraciones en la cadena de Markov y 100, 000 iteraciones de *dememorisation*. Se consideraron en desequilibrio, aquellos marcadores con una $p < 0.05$, y que fuera consistente en al menos el 50% de las localidades. Una vez identificados, se eliminaron de la base de datos.

6.6 Diversidad genética

Se calculó la diversidad genética global y por sitios de muestreo para los tres sets de marcadores (completos, neutrales, y *outliers*) mediante estadísticos de diversidad. Para evaluar y comparar la diversidad global se obtuvieron los valores de heterocigosidad observada (H_O), diversidad genética al interior de la población (H_S), diversidad genética general (H_T), la diferenciación genética definida por Wright (F_{ST}), y el coeficiente de endogamia definido por Wright (F_{IS}), calculados con el paquete *hierfstat* (Goudet y Jombart, 2015) en R v.3.6.1 (R Core team, 2019).

La diversidad por sitio se evaluó y comparó con los valores de heterocigosidad observada (H_o), heterocigosidad esperada (H_e), diversidad nucleotídica (π), obtenidos en Arlequin v.3.5.2.2 (Excoffier y Lischer, 2010).

6.7 Estructura genética

6.7.1 F_{ST} pareado

Para inferir las distancias genéticas entre sitios se utilizó el método de comparaciones pareadas de distancias genéticas (F_{ST}) implementado en Arlequin v.3.5.2.2 (Excoffier y Lischer, 2010). Se corrieron 10,000 permutaciones y para conocer el nivel de significancia de la prueba, a los valores obtenidos de p se le realizó una corrección con el método de Bonferroni (Rice, 1989).

6.7.2 Análisis de inferencia de *cluster*

Se infirió la estructura poblacional neutral y con marcadores *outliers* de *E. puncticulatus* con el pipeline StrAUTO (Chhatre y Emerson, 2017), el cual utiliza la computación paralela para correr STRUCTURE (Pritchard *et al.*, 2000), el método de la ΔK de Evanno (Evanno *et al.*, 2015) para la inferencia del número de K más probable, y STRUCTURE HARVESTER (Earl y vonHoldt, 2012) para la visualización grafica de los datos.

STRUCTURE utiliza un modelo de inferencia de *clusters* a partir de las frecuencias alélicas de datos genotípicos. Calcula la probabilidad de pertenencia de los individuos a una determinada población o *cluster* genético (K), para determinar el número más probable de K. El programa se corrió bajo un modelo de mezcla y utilizando el parámetro LOCPRIOR, el cual considera el sitio de donde proviene la muestra. Se corrieron 10,000 iteraciones de calentamiento, seguidas de 100,000 repeticiones para cada corrida. Se definió un número máximo de K de cinco, con cinco replicas para determinar el valor de K. Se utilizó STRUCTURE HARVESTER para encontrar el valor más probable de K, se consideró el valor más alto de $Pr(K)$ y el estadístico ΔK de Evanno (Evanno *et al.*, 2015). Los resultados se visualizaron como gráficos en CLUMPAK (Kopelman *et al.*, 2015).

6.7.3 Análisis Discriminante de Componentes Principales

Para la identificación de *clusters* genéticos y la relación entre ellos se corrió un Análisis Discriminante de Componentes Principales (ADPC) con la función *dapc* en el paquete

Adegenet v.2.1.1 (Jombart *et al.*, 2010) en R (R Development Core Team, 2019). Este análisis transforma los datos en componentes principales (PCs) mediante un Análisis de Componentes Principales. Para ello se consideró la información *a priori* (sitios de muestreo) y el número óptimo de grupos (K) sugerido por la función *find.clusters*, el cual es calculado mediante el criterio bayesiano de información (BIC). Con la función *cross-validation* se definió el número óptimo de PCs y funciones discriminantes a retener, para lo cual se corrieron 50 repeticiones para cada PC retenido.

6.7.4 AMOVA jerárquico

En Arlequin v.3.5.2.2 (Excoffier y Lischer, 2010) se evaluó la estructura génica en diferentes niveles jerárquicos (1) entre grupos, 2) entre poblaciones, 3) entre individuos dentro de las poblaciones, y 4) dentro de los individuos con un Análisis Molecular de Varianza (AMOVA) para cada locus, con 10,000 permutaciones. De acuerdo a los resultados obtenidos se evaluaron cuatro posibles estructuras de poblaciones: 1) panmixia, 2) dos grupos (norte y sur, con un quiebre genético ente los sitios de Loreto y Mulegé), 3) tres grupos (Los Cabos, La Paz y Loreto; Mulegé, y Bahía de los Ángeles), y 4) cuatro grupos (Los Cabos y La Paz; Loreto; Mulegé; y Bahía de los Ángeles)

6.8 Migración

Los patrones y magnitud de flujo génico entre sitios y grupos se estimaron con la función *divMigrate* implementada en el paquete *diveRsity* (Keenan *et al.*, 2013) en R v.3.6.1 (R Core team, 2019). Se calculó a partir de valores de G_{ST} , con un filtro de $G_{ST} > 0.3$ para marcadores neutrales, y uno menos estricto para marcadores *outliers* ($G_{ST} > 0.2$), y se generó un gráfico de la red de conexiones de los patrones de flujo génico.

6.8 Aislamiento por distancia

Se probó aislamiento por distancia mediante una prueba de Mantel en el software IBD (Bohonak, 2002). Para probar si este ocurría a manera de clina o como parches en el espacio, se analizó la densidad de los puntos en el espacio con el paquete ADEGENET v.2.1.1 (Jombart *et al.*, 2010) en R v.3.6.1 (R Core team, 2019).

7. RESULTADOS

7.1 Librería genómica ddRAD

Para la elaboración de la librería genómica se utilizó solamente el ADN de 144 individuos (y dos réplicas) de *E. puncticulatus* de los 19 sitios de muestreo (Tabla II). El ADN obtenido fue íntegro y de alto peso molecular (Fig. 5).

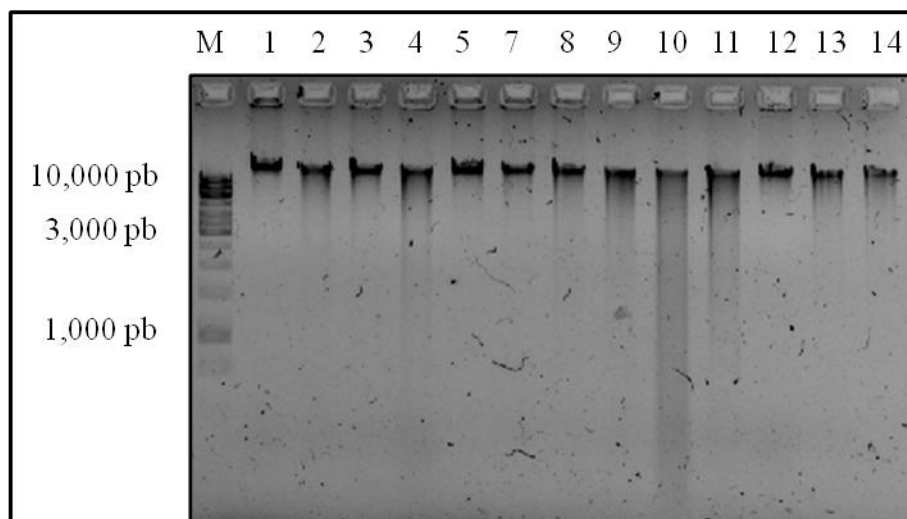


Figura 5. Ejemplo de ADN genómico extraído de *Elcatinus puncticulatus* [25 ng]. M= Marcador de peso molecular 1kb (Thermo Fisher Scientific), 1= ISANC-099, 2= ISANC-099, 3= ISANC-100, 4= SMS-116, 5= SMS-117, 6= SMS-118, 8=SMS-118, 9= SMS-120, 10= SMS-121, 11= SMN-138, 12= SMN-134, 13= SMN-135, 14= SMN-137.

Previo a la elaboración de los *pools* se verificó la correcta ligación de los barcodes esto se puede apreciar en la figura 6 donde se observa un barrido en los individuos donde los barcodes se ligaron exitosamente. Posteriormente, se elaboraron los *pools* y se llevó a cabo la selección de fragmentos de cada *pool* (350 pb). Finalmente, se agregó un índice específico a cada *pool* mediante PCR de tal forma que cada individuo tuviese una combinación única (barcode + índice) para su posterior identificación (Tabla III). La verificación de la correcta elaboración de los *pools* y selección de fragmentos se puede apreciar en figura 7 donde la amplificación de cada *pool* presenta una banda en el rango esperado de ~475 pb.

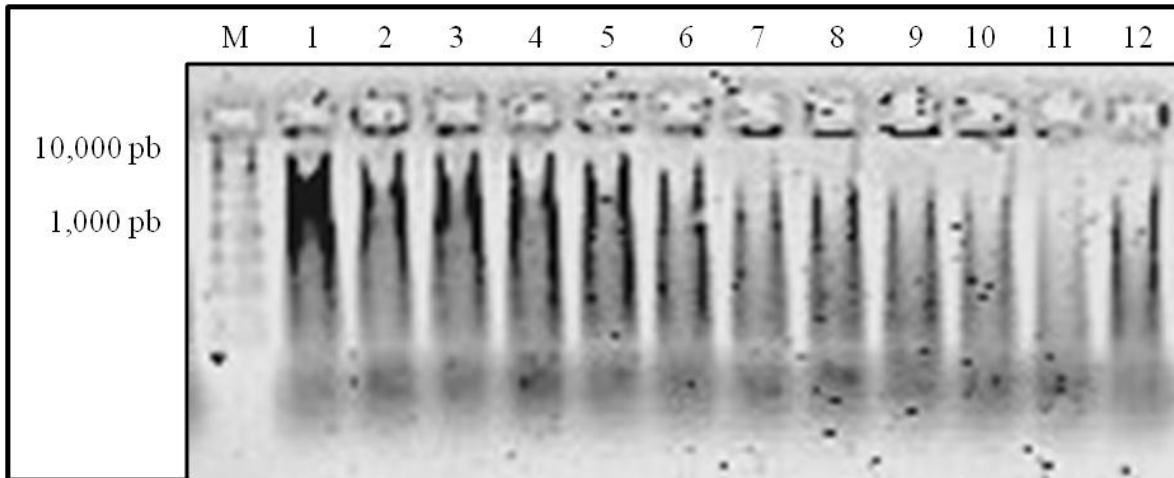


Figura 6. Verificación de ligación de adaptadores en *E. puncticulatus*. Se aprecia un barrido debido a la amplificación de fragmentos de ADN de distintos tamaños. Nota: M= Marcador de peso molecular GeneRuler 1kb, 1= ISANC-096, 2= ISANC-097, 3= ISANC-098, 4= ISANC-099, 5= ISANC-100, 6= SMS-116, 7= SMS-117, 8= SMS-118, 9= SMS-120, 10= SMS-121, 11=SMN-138, 12= SMN-134.

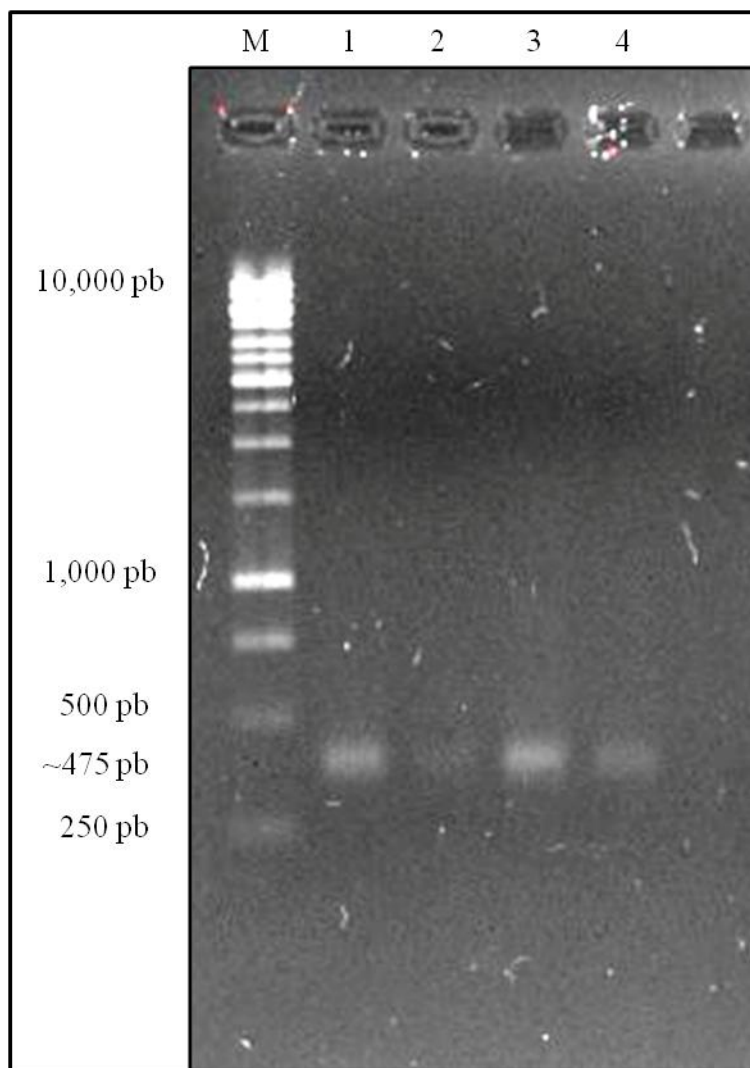


Figura 7. Pools de *E. puncticulatus* amplificados con *primers* indexados. El tamaño esperado después de la selección de fragmentos es de 475 pb. Nota: M= Marcador de peso molecular GeneRuler 1kb, 1= Pool 1, 2= Pool 2, 3= Pool 3, 4= Pool 4.

Los *pools* se combinaron equimolarmente y se purificó para eliminar fragmentos pequeños y disminuir la amplia distribución de los tamaños de los fragmentos (Fig. 8 y tabla VI). Finalmente se secuenciaron en un secuenciador Illumina Hi-Seq X.

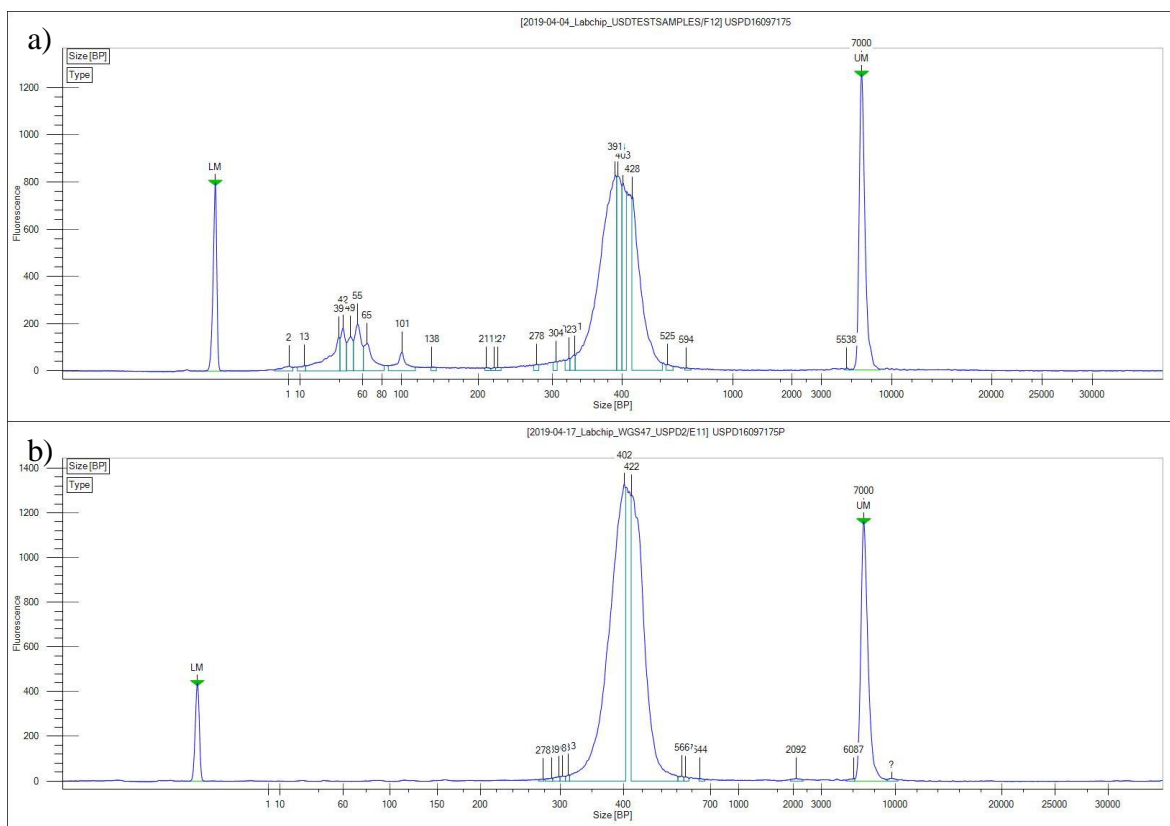


Figura 8. Electroferograma después de la purificación de la librería genómica de *Elacatinus puncticulatus* obtenido con Bioanalyzer Agilent 2100-H. a) Antes de la purificación. b) Después de la purificación.

Tabla VI. Resumen del control de calidad de Novogene después de la purificación de la librería.

Muestra	Librería	Index de secuenciación	Volumen de la librería (μL)	Tamaño de fragmento (bp)	Concentración (ng/μL)	Molaridad (nmol/L)
E-pun-L1	USPD16097175	ATCACG, CAGATC, ACTTGA, GATCAG	8	412	1.73	11.0007

7.3 Demultiplexeado

Se obtuvieron un total de 792,891,492 lecturas, de las cuales, en *process_radtags*, se eliminaron aquellas con barcodes ambiguos (1.9%) y RAD-Tags ambiguos (2%), y las de baja calidad (phred score<10, 0.1%) (Tabla VII y Fig. 9). Después de este primer filtrado se

retuvieron un total de 761,952,455 (96.1%) lecturas totales. Una vez limpias las secuencias, el componente realizó un demultiplexeo de las secuencias.

Tabla VII. Filtrado de lecturas crudas en el componente *process_radtags* en Stacks.

'process_radtags'	Número de lecturas	Porcentaje
Total de lecturas	792,891,492	100%
Barcodes ambiguos	14,943,620	1.90%
Baja Calidad	485,290	0.10%
RAD-Tags ambiguos	15,510,127	2.00%
Lecturas retenidas	761,952,455	96.10%

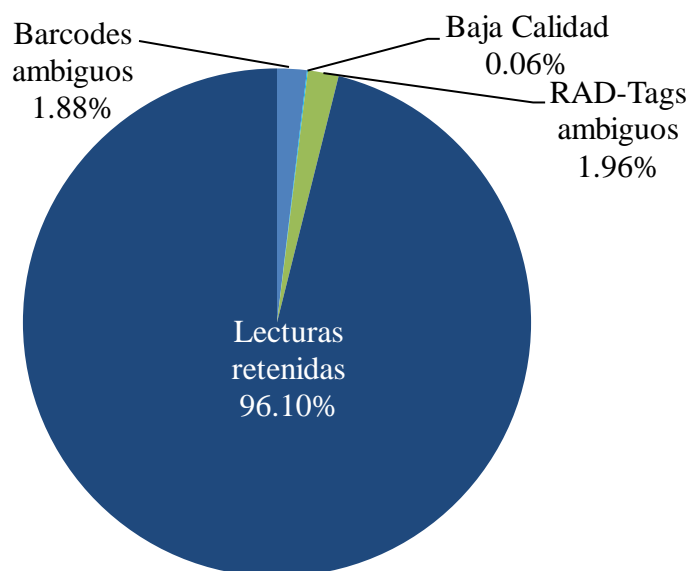


Figura 9. Porcentaje de lecturas retenidas y eliminadas por el proceso de filtrado en 'process_radtags'.

7.4 Ensamble *de novo* y parametrización

Para la elección de los parámetros en el componente *denovo.pl* se realizaron pruebas (Tabla VIII) con distintos valores para *-m*, *-M*, y *-n* y se optó por la combinación que encontrara la mayor variación genética, sin sacrificar la calidad de la información, en este sentido se eligió la combinación *-m 3*, *-M 4* y *-n 3* (Tabla VIII).

Tabla VIII. Pruebas de los parámetros principales (-M, -m, y -n) en *denovo_map.pl*. Los renglones sombreados indican el parámetro elegido por prueba.

Prueba	M	m	n	t	P	Loci retenidos	Sitios variantes retenidos	Profundidad de cobertura promedio
m3	3	3	3	140	18 de 20	6,916	14,638	70.42
m4	3	4	3	140	18 de 20	6,490	13,964	84.89
m5	3	5	3	140	18 de 20	6,060	12,912	95.88
M1	4	3	4	140	18 de 20	6,807	14,359	71.31
M2	5	3	5	140	18 de 20	6,284	13,393	86.50
M3	6	3	6	140	18 de 20	6,554	13,734	72.06
n1	4	4	4	140	18 de 20	6,420	13,832	85.96
n2	4	4	3	140	18 de 20	6,399	13,743	85.96
n3	4	4	5	140	18 de 20	6,414	13,874	85.96

7.5 Filtrado de individuos

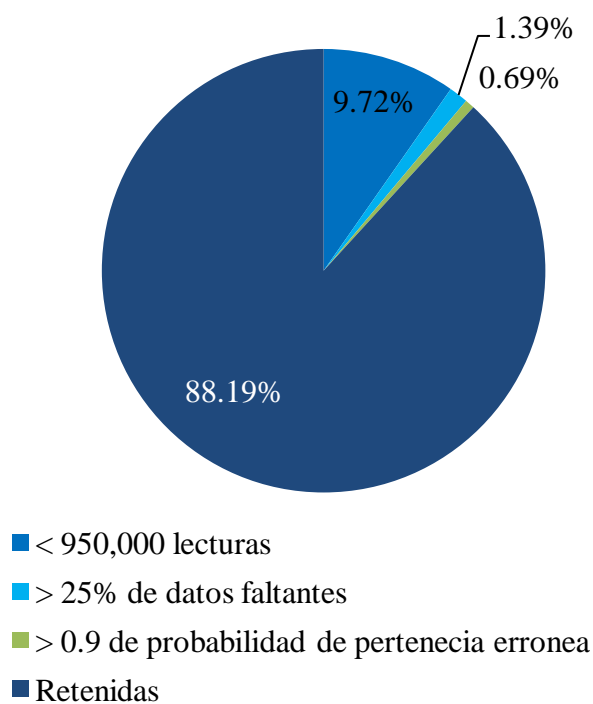
Se secuenciaron 144 individuos, sin embargo, los individuos que se retuvieron para este estudio fueron aquellos que pasaron tres filtros principales (Tabla IX y fig. 10). 1) De los resultados arrojados por *process_radtags.pl* se conservaron los individuos que poseían más de 950, 000 lecturas retenidas. Se descartaron un total de 16 individuos, incluidas las réplicas, por lo cual el set disminuyó a 130 individuos.

Los individuos restantes fueron reagrupados por sitios cercanos para aumentar el número de muestra por sitio, lo que redujó el número de sitios de 19 a 15 (Fig. 4). Con la nueva agrupación se realizó otro ensamble *de novo*, con los parámetros previamente establecidos (-m 4, -M 4, -n 3). 2) Para una -p de 14 en el módulo de *populations* se calculó el porcentaje de datos faltantes y se descartaron los individuos con más del 25%. Las dos muestras eliminadas en este filtro pertenecían a la localidad el Pardito.

Finalmente se eliminó un individuo debido a que presentaba más del 0.95 de probabilidad de pertenencia errónea a su población en los análisis de inferencia de *clusters* en Structure. De esta manera solo se retuvieron 127 individuos de 15 sitios (Tabla V).

Tabla IX. Individuos de *Elacatinus puncticulatus* descartados por filtro.

Filtro	Individuos descartados	Individuos retenidos
< 950,000 lecturas	14 (+2 réplicas)	130
> 25% de datos faltantes	2	128
> 0.95 de probabilidad de pertenencia errónea	1	127
Total	17 (+2 réplicas)	127

**Figura 10.** Porcentaje de individuos de *Elacatinus puncticulatus* retenidos y descartados en el filtrado de los datos.

7.6 Filtrado de loci y SNPs

El número de loci y sitios variantes se contrastó con el porcentaje de datos faltantes de las muestras, tratando de obtener la mayor cantidad de información sin sacrificar la calidad de la misma. Finalmente se optó por un valor de $-p$ 14, obteniendo como resultado un total de 6,513 loci, 1,722,337 sitios, 4,938 sitios variantes y un máximo de 20.6% de datos faltantes por individuo (Fig. 11).

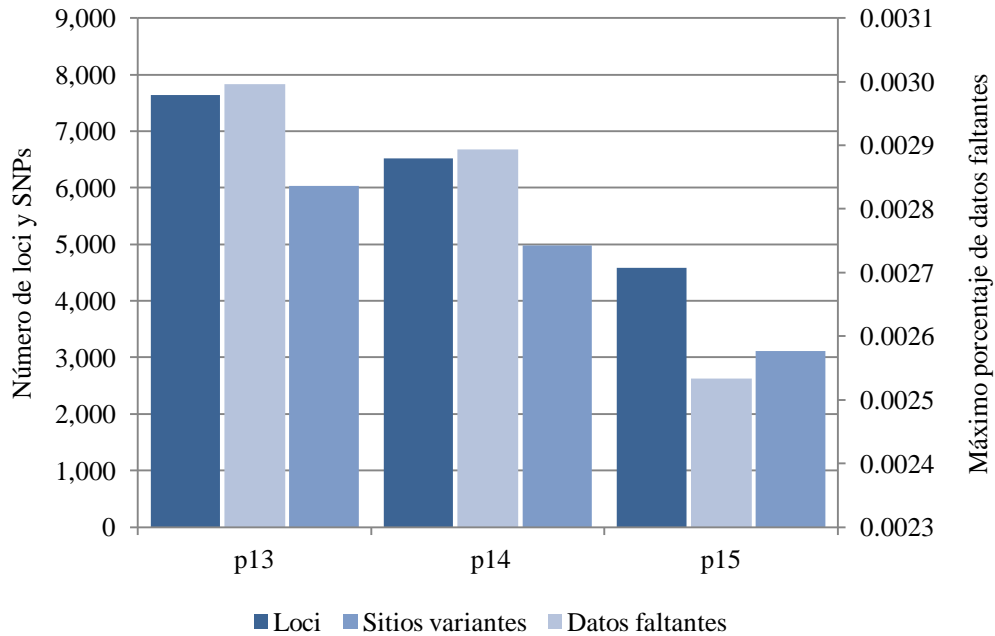


Figura 11. Número de loci, sitios variantes y máximo porcentaje de datos faltantes por individuo para *E. puncticulatus* para distintos valores de $-p$. En $-p14$ se muestran los datos utilizados en este trabajo.

A la base de marcadores completa se le realizó la prueba de equilibrio de Hardy-Weinberg en Arlequin, y se encontraron cuatro marcadores en desequilibrio reduciendo el número de sitios variables a 4,979.

7.7 Detección de *outliers*

Se realizó una detección de *outliers* con dos métodos distintos. Arlequin, utilizando los valores de F_{ST} , detectó la mayor cantidad de *outliers*, con 104 marcadores detectados (Fig. 12). Los resultados de Bayescan sugieren un total de 56 marcadores como *outliers* (Fig. 13). Solamente se consideraron como *outliers* solo aquellos marcadores que fueron identificados en ambos programas, siendo un total de 26 coincidencias.

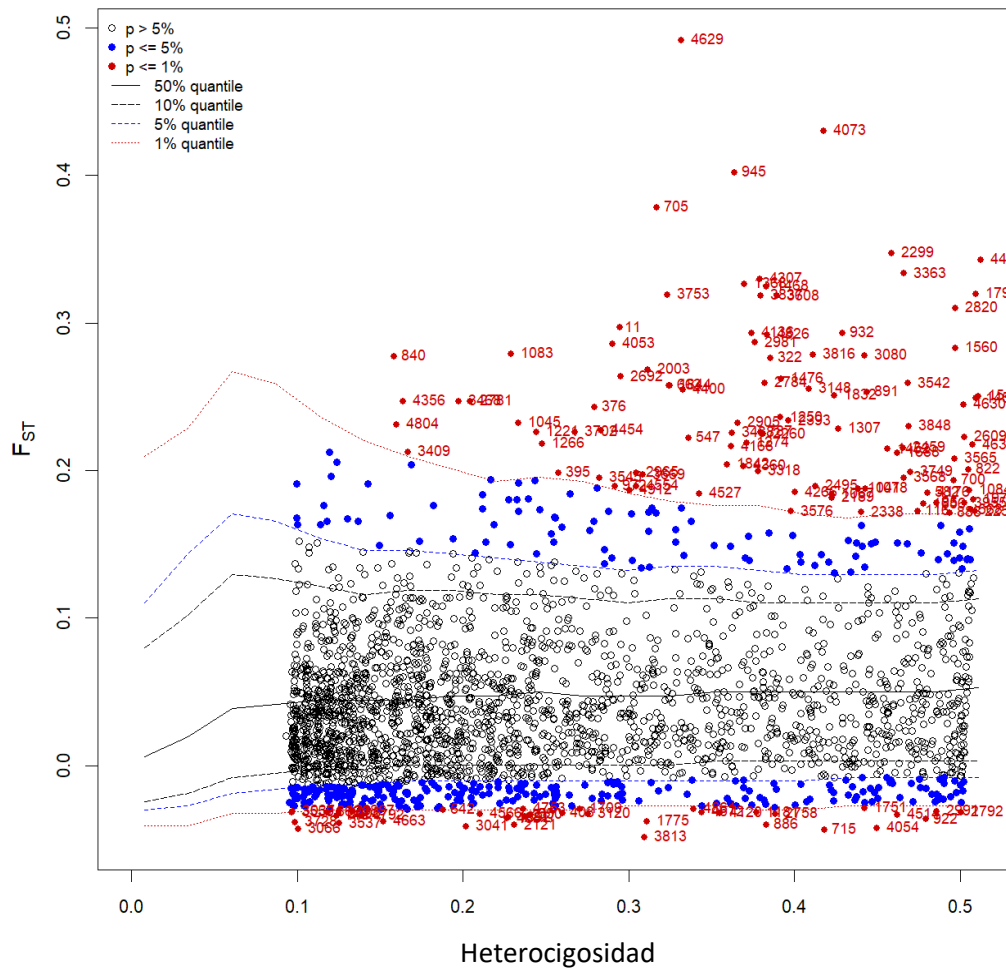


Figura 12. Detección de *outliers* para *E. puncticulatus* utilizando modelo jerárquico de islas en Arlequin. Cada punto representa uno de los 4,979 marcadores, los puntos rojos con valores positivos de F_{ST} en el cuantil del 1% de los datos, fueron considerados como marcadores *outliers*. En total el programa detectó 104 *outliers*.

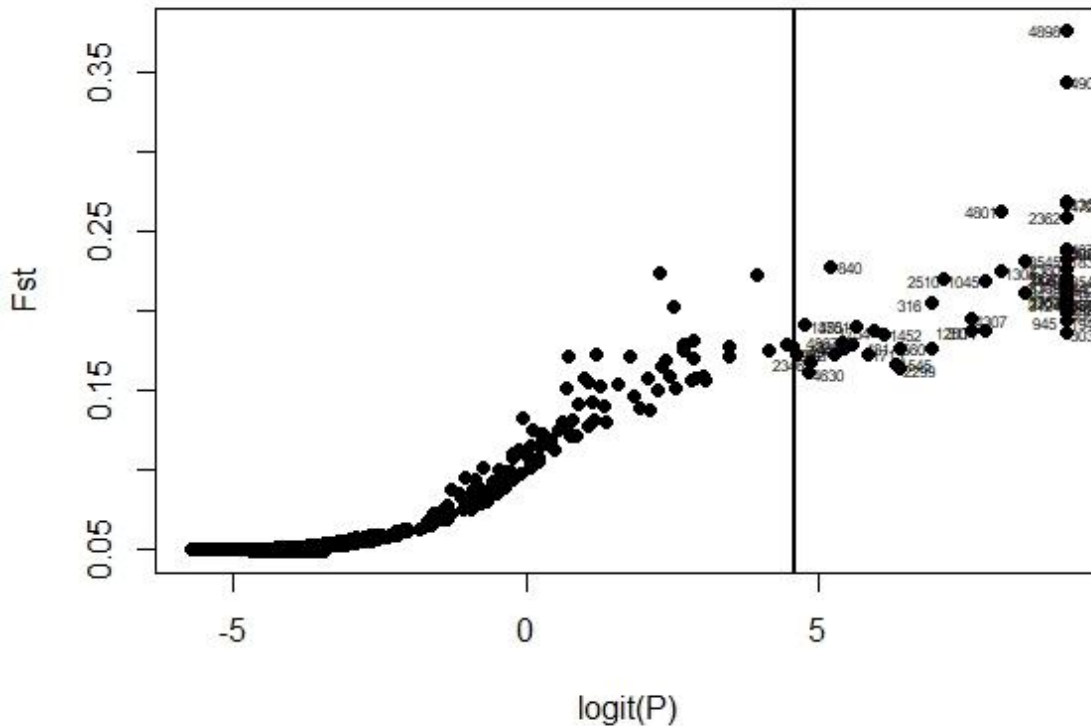


Figura 13. Detección de *outliers* para *E. puncticulatus* utilizando método bayesiano implementado en Bayescan. Cada punto representa uno de los 4,979 marcadores, los puntos después de la línea vertical se encuentran en el cuantil del 1%, estos fueron considerados como marcadores *outliers*. En total el programa detectó 56 *outliers*.

La detección de *outliers* permitió que los datos pudiesen separarse y ser analizados en tres sets de marcadores: la base completa con 4,979 SNPs, la base neutral con 4,953 SNPs, y la base de *outliers* con 26 SNPs.

7.8 Diversidad genética

Se calculó la diversidad genética global para cada set marcadores (Tabla X). Los valores de los estadísticos globales del set de datos completo y neutral fueron muy similares, únicamente el valor de F_{IS} fue ligeramente mayor en el set completo ($F_{IS}= 0.08$), por lo cual el trabajo solo considera la base de datos neutrales y de *outliers* en análisis posteriores.

A excepción de F_{IS} , todos los estadísticos presentaron valores mayores en el set de *outliers* que en el completo y el de neutrales. Las diferencias más grandes se encontraron para la

diversidad genética general (H_T) donde en *outliers* fue, aproximadamente, 1.5 veces mayor (0.4), y F_{ST} fue 5.4 veces más elevado.

Tabla X. Diversidad genética global para *E. puncticulatus* para cada una de las tres bases de datos generadas (completa, neutral y *outliers*). H_O : heterocigosidad observada, H_S : diversidad genética dentro de la población, H_T : diversidad genética general, F_{ST} : medida de diferenciación genética de Wright, F_{IS} : coeficiente de endogamia de Wright.

Marcadores	H_O	H_S	H_T	F_{ST}	F_{IS}
Completa	0.23	0.25	0.26	0.05	0.08
Neutrales	0.23	0.25	0.26	0.05	0.05
<i>Outliers</i>	0.28	0.29	0.4	0.27	0.05

Al analizar los estadísticos por sitio para los marcadores neutrales se observa que la H_e resultó ser mayor a la H_o en todas las localidades (Fig. 14). También, la heterocigosidad fue ligeramente mayor en las latitudes más extremas, principalmente en el sur, y disminuyó en las centrales, a excepción de SANJPS ($H_o = 0.359$, y $H_e = 0.387$) e IAV ($H_o = 0.359$, y $H_e = 0.389$), que presentaron los valores más altos. Los valores de diversidad nucleotídica (π) fueron relativamente homogéneos, aunque se presentaron valores sutilmente más altos en latitudes centrales (Fig. 15), con los valores más altos en IAV ($\pi = 0.255$) y PB ($\pi = 0.251$), mientras que hacia los extremos, estos valores disminuyen.

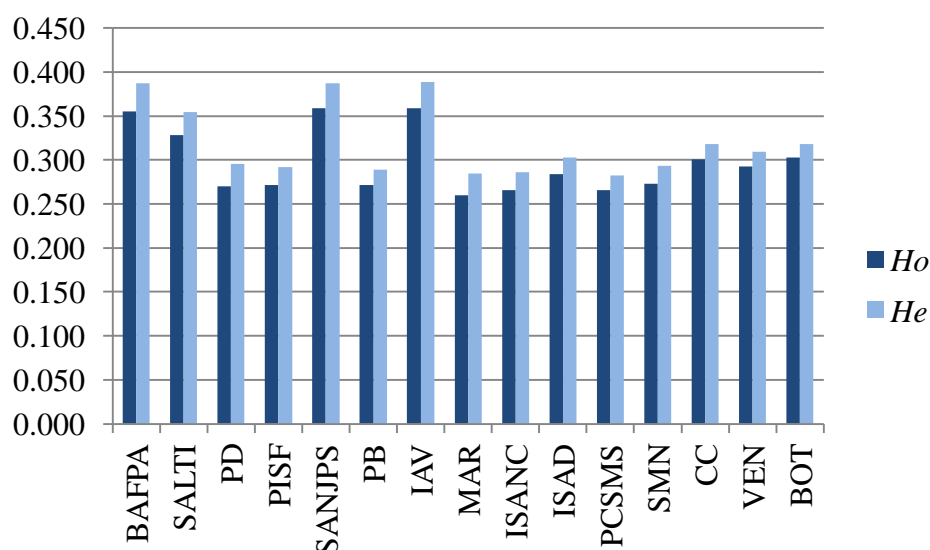


Figura 14. Heterocigosidad observada (H_o) y esperada (H_e) neutral para *E. puncticulatus* en el GC para cada sitio.

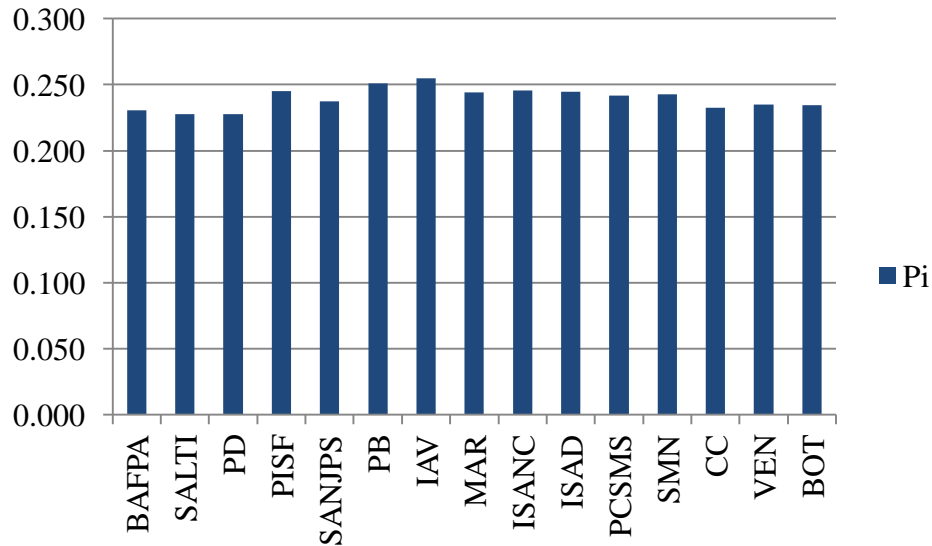


Figura 15. Diversidad nucleotídica (π) neutral para *E. puncticulatus* para cada sitio en el GC.

Los marcadores *outliers* mostraron un comportamiento más marcado de la disminución de H_e y H_o hacia el norte (Fig. 16), a excepción de BOT, donde los valores de H_e aumentan ($H_e=0.342$) e inclusive se observa un exceso de heterocigotos. Los valores más altos se presentaron en BAFPA ($H_o= 0.444$ y $H_e= 0.452$) y SANJPS ($H_o= 0.438$ y $H_e= 0.433$). Mientras que el comportamiento de π en la latitud fue similar al de los marcadores neutrales (Fig. 17), con los valores más altos en las latitudes intermedias, siendo IAV ($\pi=0.360$) y MAR ($\pi=0.350$) los mayores.

La H_o y H_e , así como la diversidad nucleotídica (π), por sitio, presentó valores más altos en marcadores *outliers* que en neutrales (Figuras 16 y 17).

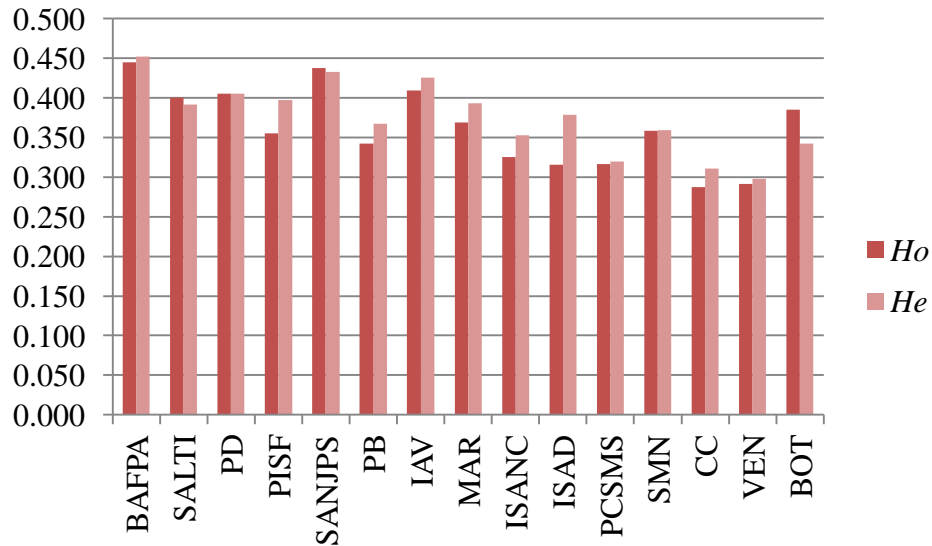


Figura 16. Heterocigosidad observada (H_o) y esperada (H_e) para *outliers* de *E. puncticulatus* para cada sitio en el GC.

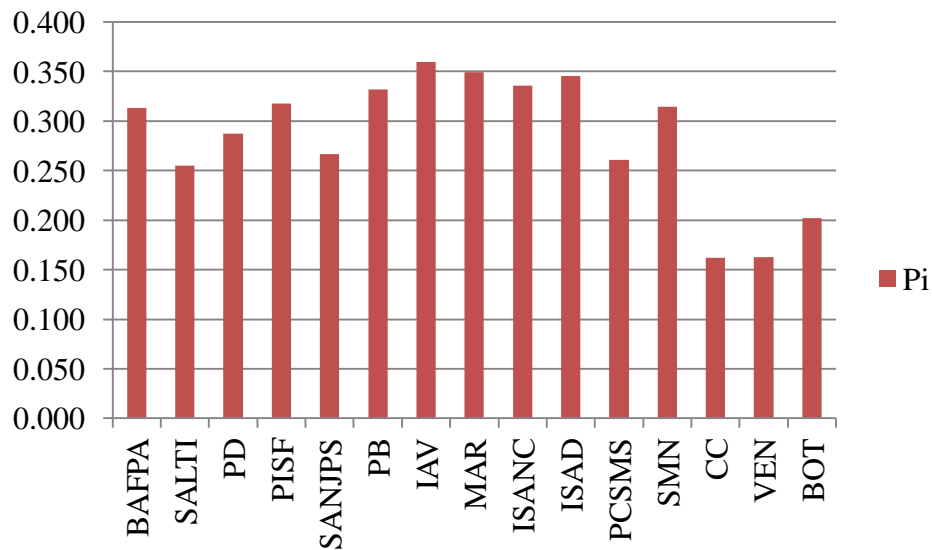


Figura 17. Diversidad nucleotídica (π) para *outliers* de *E. puncticulatus* para cada sitio en el GC.

7.8 Aislamiento por distancia

Como un primer acercamiento para comprender el comportamiento de la estructura poblacional de *E. puncticulatus* en el Golfo de California, en una prueba de Mantel se encontró un patrón de aislamiento por distancia con valores de p significativos ($p < 0.0001$)

en ambas bases de datos tanto neutral y *outliers* (Fig. 18). Es decir que individuos en sitios con un mayor aislamiento geográfico presentarán una mayor diferenciación genética, mientras que los más cercanos son más similares asociado a un mayor intercambio de material génico. La correlación fue mayor para marcadores *outliers* ($Z= 218.3476$) que para neutrales ($Z= 19.5436$), debido a que los valores de distancia genética de Slatkin (1995) fueron superiores en las comparaciones pareadas entre sitios para marcadores *outliers* (Fig. 18).

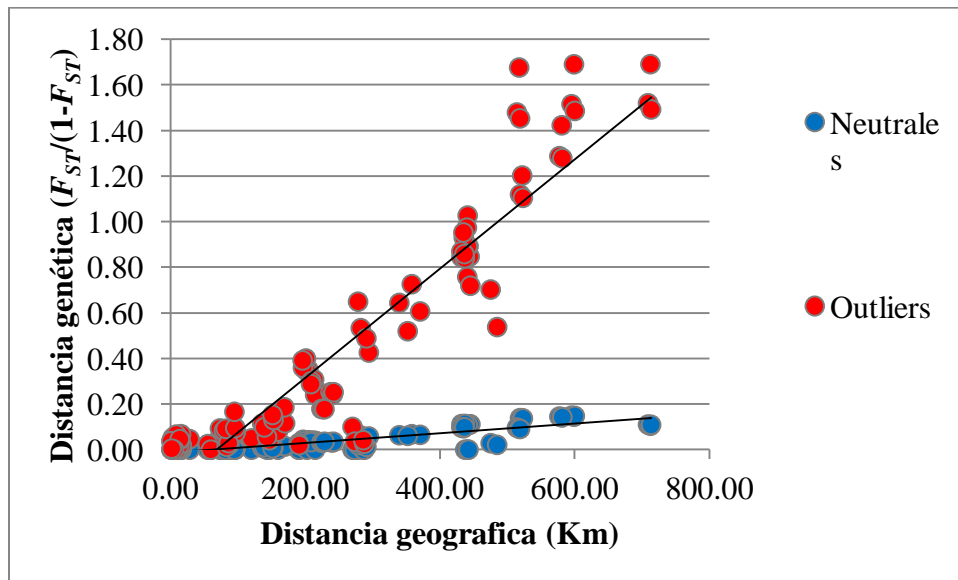


Figura 18. Prueba de aislamiento por distancia para *Elacatinus puncticulatus* en el GC. Prueba de Mantel para marcadores neutrales: $Z= 19.544$, $r= 0.846$, $p<0.0001$. Prueba de Mantel para marcadores *outliers*: $Z= 218.348$, $r= 0.929$, $p<0.0001$.

La estimación bidimensional de la densidad de kernel (Figs. 19 y 20), revela la diferenciación de dos parches tanto para marcadores neutrales como para *outliers*. En ambos sets de marcadores se encontró un parche grande (~300 km de ancho) y uno más pequeño (~150 km), lo cual podría ser un indicio de una autocorrelación espacial entre los sitios de muestreo.

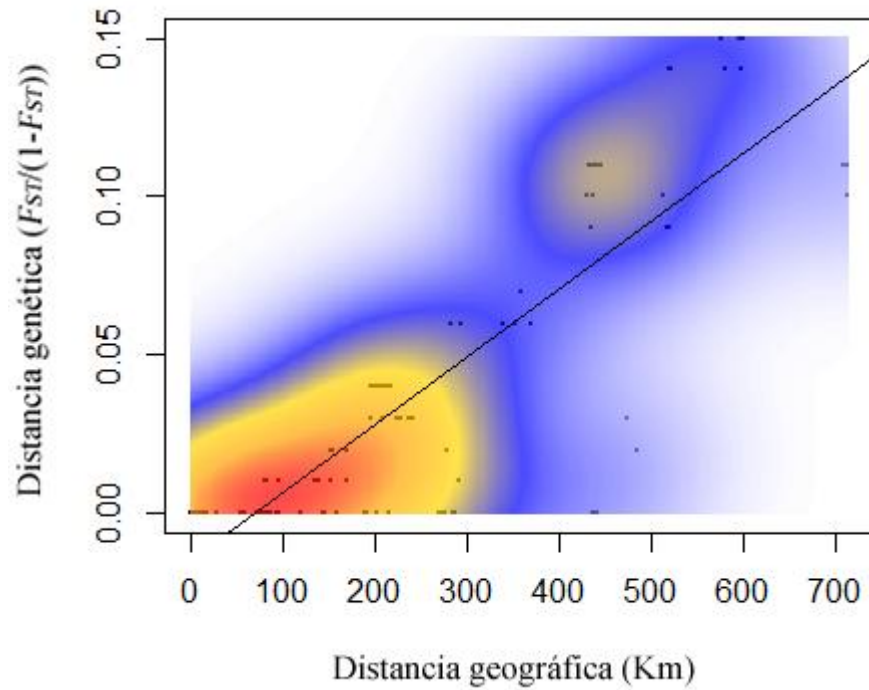


Figura 19. Prueba de Mantel para probar aislamiento por distancia con marcadores neutrales para *E. puncticulatus* en el GC. Se estimó densidad de kernel bidimensional.

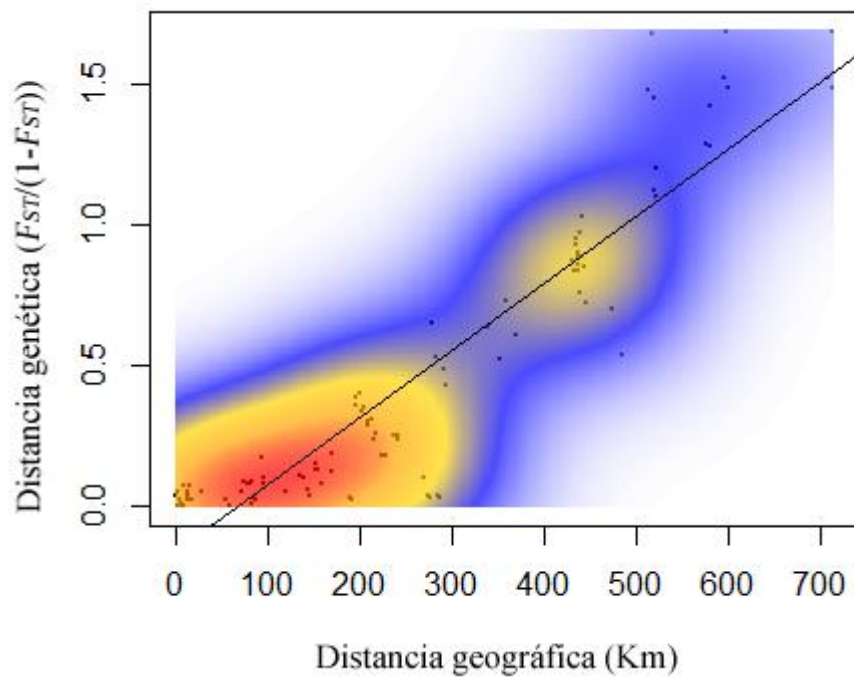


Figura 20. Prueba de Mantel para probar aislamiento por distancia con marcadores *outliers* para *E. puncticulatus* en el GC. Se estimó densidad de kernel bidimensional.

7.9 Estructura poblacional con marcadores neutrales

Se evaluó el nivel de diferenciación genética para marcadores neutrales entre sitios utilizando valores de F_{ST} pareados comparando las localidades se encuentran ordenadas de sur a norte (Tabla XI). Después de aplicar la corrección de Bonferroni, en una inspección visual robusta de la matriz de diferencias pareadas se encontró que los sitios más cercanos presentaban los valores más bajos de F_{ST} , y los valores de p son significativos ($p \leq 0.001$) conforme se incrementa la distancia, resultando en un evidente quiebre genético entre las regiones sur y norte, localizado entre los sitios ubicados en Loreto y Mulegé, con variación al interior de los grupos.

Tabla XI. FST pareados entre sitios para marcadores neutrales para *E. puncticulatus* en el GC. Valores de FST debajo de la diagonal, y valores de P sobre la diagonal. Valores en negrita resultaron significativos después de la corrección de Bonferroni (valor de P con la corrección de Bonferroni= 0.001)(Rice, 1989).

	BAFPA	SALTI	PD	PISF	SANJPS	PB	IAV	MAR	ISANC	ISAD	PCSMS	SMN	CC	VEN	BOT
BAFPA	---	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.000	0.001	0.001	0.001	0.001
SALTI	-0.041	---	1.000	0.712	1.000	0.001	1.000	0.000	0.000	0.001	0.000	0.000	0.001	0.000	0.001
PD	-0.049	-0.006	---	1.000	1.000	0.000	1.000	0.000	0.004	0.064	0.000	0.000	0.000	0.000	0.000
PISF	-0.042	0.003	-0.002	---	1.000	0.000	1.000	0.000	0.000	0.075	0.000	0.000	0.000	0.000	0.000
SANJPS	-0.090	-0.032	-0.044	-0.040	---	1.000	1.000	1.000	1.000	1.000	0.000	0.005	0.002	0.001	0.001
PB	-0.028	0.015	0.009	0.009	-0.032	---	1.000	0.874	0.964	1.000	0.000	0.000	0.000	0.000	0.000
IAV	-0.178	-0.107	-0.124	-0.123	-0.190	-0.135	---	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
MAR	-0.027	0.017	0.011	0.010	-0.026	0.001	-0.137	---	0.498	1.000	0.000	0.000	0.000	0.000	0.000
ISANC	-0.023	0.014	0.008	0.011	-0.032	0.000	-0.132	0.003	---	1.000	0.000	0.000	0.000	0.000	0.000
ISAD	-0.031	0.016	0.007	0.005	-0.034	-0.006	-0.136	-0.005	-0.004	---	0.000	0.000	0.000	0.000	0.000
PCSMS	0.028	0.063	0.060	0.058	0.020	0.040	-0.089	0.039	0.042	0.033	---	0.145	0.000	0.000	0.000
SMN	0.021	0.061	0.057	0.055	0.012	0.035	-0.090	0.038	0.038	0.032	0.003	---	0.000	0.000	0.000
CC	0.099	0.130	0.127	0.122	0.088	0.100	-0.019	0.099	0.101	0.089	0.032	0.030	---	0.983	1.000
VEN	0.095	0.126	0.125	0.120	0.084	0.099	-0.023	0.098	0.099	0.086	0.031	0.031	-0.001	---	0.922
BOT	0.098	0.129	0.125	0.121	0.083	0.100	-0.027	0.099	0.101	0.089	0.033	0.033	-0.005	0.000	---

Se identificaron los *clusters* genéticos infiriendo el número más probable de poblaciones (K) para los marcadores neutrales. El método de Evanno (Fig. 21a) sugiere que existen dos grupos genéticos, mientras que Pr(K) (Fig. 21b) muestra que el número de grupos más probable es cuatro.

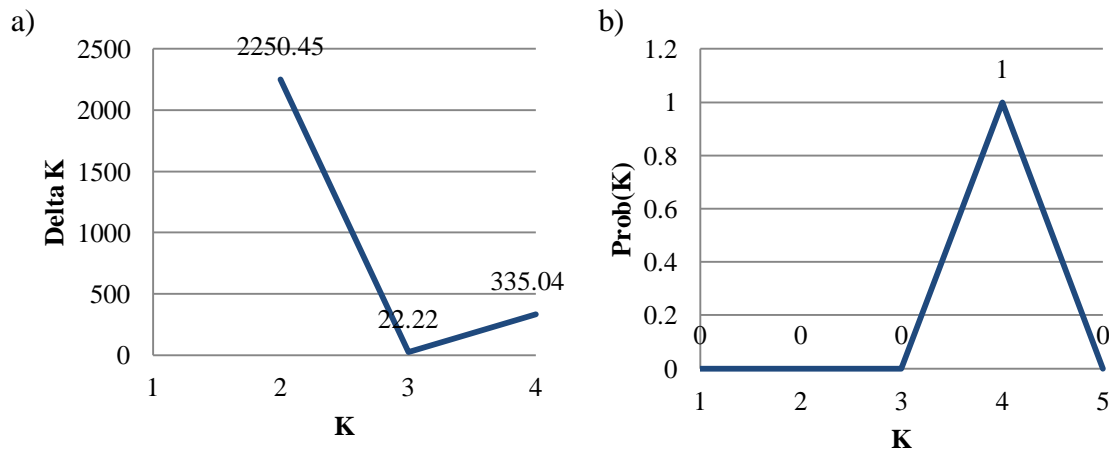


Figura 21. K más probable para marcadores neutrales de *E. puncticulatus* en el GC. a) Magnitud de ΔK para distintos valores de K (Evanno *et al.*, 2015). b) Pr(K) para cada K (Pritchard *et al.*, 2000).

Se presentan los gráficos de asignación de pertenencia generados en STRUCTURE, para cada una de las K más probables. En K=2 (Fig. 22a), el quiebre genético ocurre entre los sitios del 1) norte (BAFPA, SALTI, PD, PISF, SANJPS, PB, IAV, MAR, ISANC, e ISAD), y los del 2) y sur (PCSMS, SMN, CC, VEN y BOT). Es notoria una introgresión clinal de norte a sur conforme a la latitud de los sitios de muestreo (Fig. 22a). Para el caso de K=3 la asignación de los sitios con latitudes medias es menos precisa que las de los sitios en latitudes extremas, actuando como una zona de mezcla, no se aprecian de manera clara los probables cuatro clústers (Fig. 22b). Finalmente, el caso de K=4 es similar al de K=3, pues las latitudes medias parece no haber una clara distinción entre grupos; sin embargo, se observa quiebre genético ubicado entre Loreto y Mulegé (Fig. 22c).

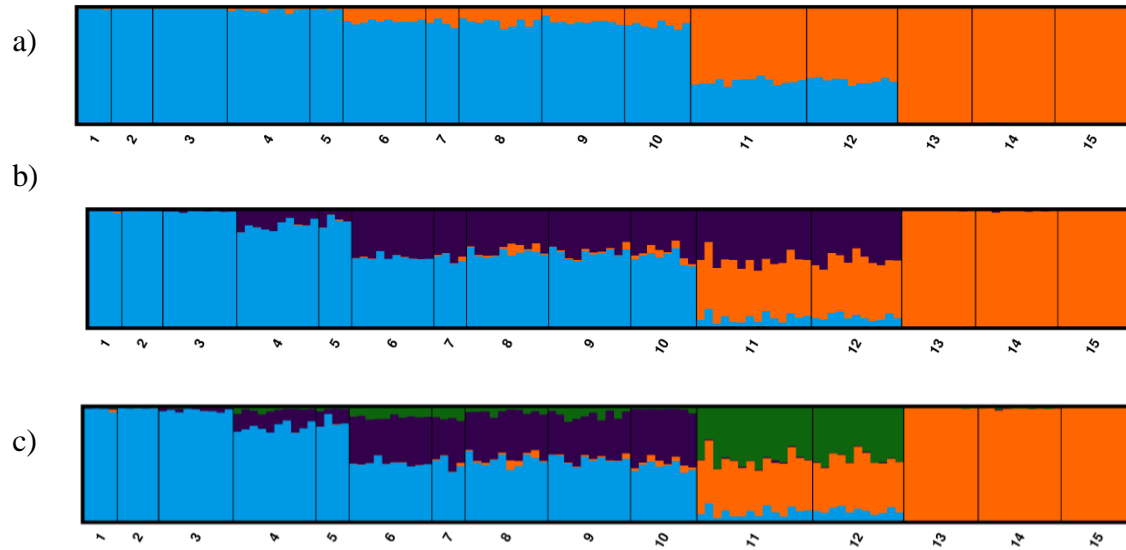


Figura 22. Agrupamiento de poblaciones de *E. puncticulatus* con marcadores neutrales. a) Individuos con sus probabilidades de pertenencia agrupados en dos grupos principales ($K=2$) (azul: sur, y naranja: norte). b) Probabilidad de pertenencia de individuos agrupados en tres grupos ($K=3$) (azul: Los Cabos y La Paz; morado: Loreto y Mulegé; y naranja: Bahía de los Ángeles). c) Individuos con sus probabilidades de pertenencia agrupados en cuatro grupos ($K=4$) (azul: Los Cabos y La Paz, morado: Loreto, verde: Mulegé, y naranja: Bahía de los Ángeles). 1: BAFPA, 2: SALTI, 3: PD, 4: PISF, 5: SANJPS, 6: PB, 7: MAR, 8: IAV, 9: ISAD, 10: ISANC, 11: PCSMS, 12: SMN, 13: CC, 14: VEN, y 15: BOT.

Los resultados anteriores se compararon con un ADCP (Fig. 23), que coincide con los resultados encontrados en el análisis anterior de inferencia de grupos genéticos. Al realizar el agrupamiento considerando los 15 sitios de muestreo (*a priori*) (Fig. 23a) sugiere la existencia de cuatro grupos: 1) Los Cabos y La Paz (BAFPA, SALTI, PD, PISF, y SANJPS), 2) Loreto (PB, IAV, MAR, ISANC, e ISAD), 3) Mulegé (PCSMS y SMN), y 4) Bahía de los Ángeles (CC, VEN, y BOT). Sin embargo, de acuerdo al valor mínimo del Criterio de Inferencia Bayesiano (BIC), como la K más probable se consideraron un total de 2 grupos ($K=2$, $BIC=830.057$). Considerando este agrupamiento, se graficó la densidad de los individuos sobre la primer función discriminante y se infirieron dos grupos: 1) sur (BAFPA, SALTI, PD, PISF, y SANJPS), y 2) norte (PCSMS, SMN, CC, VEN, y BOT) (Fig. 23b).

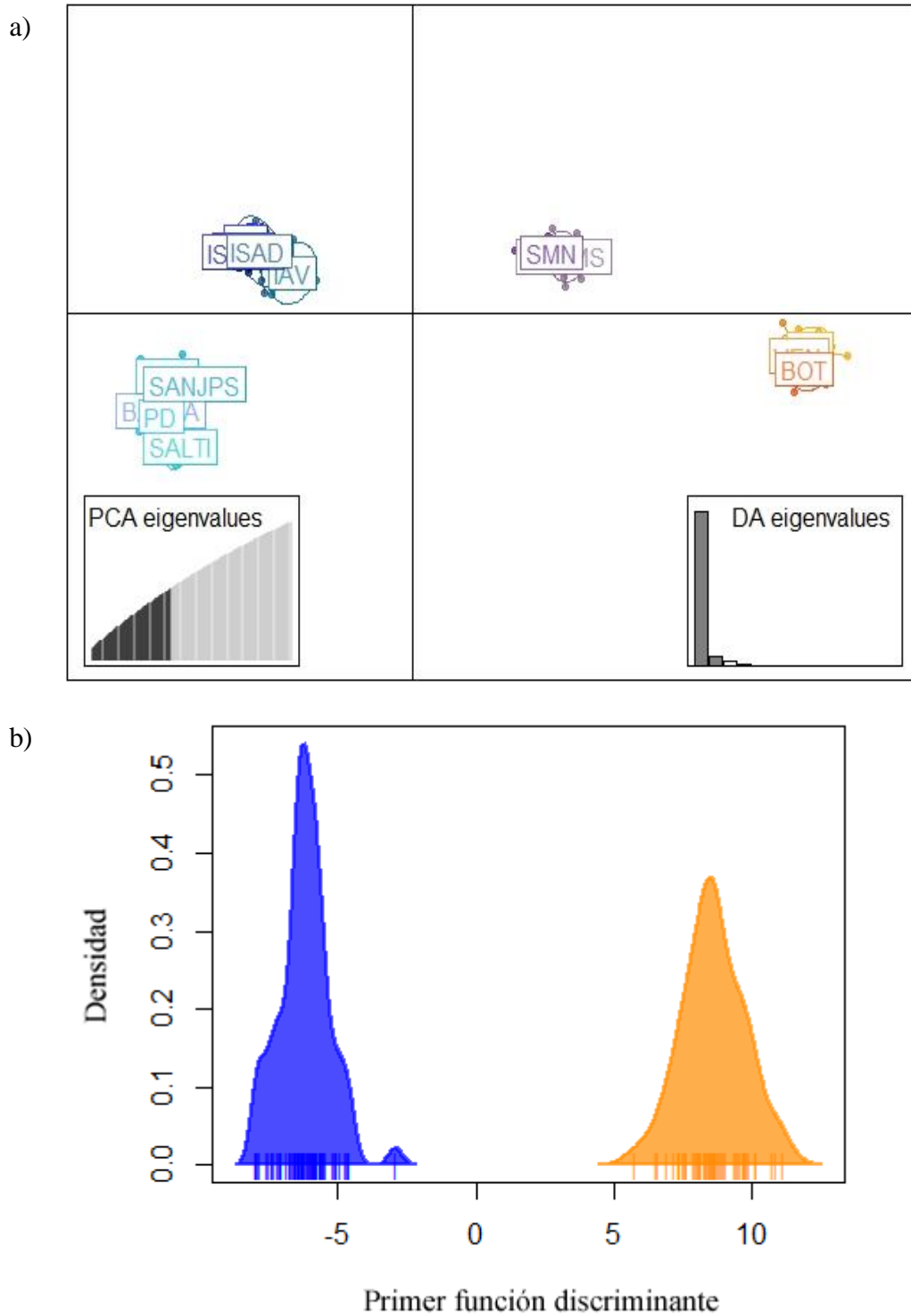


Figura 23. ADCP con marcadores neutrales con 50 PCs retenidos. a) obtenido por localidad, con dos funciones discriminantes retenidas, y b) densidad de los individuos a partir de la K más probable obtenida con la función *find.clusters* ($K=2$, $BIC=830.057$), con una función discriminante retenida, en color azul las localidades del sur y en naranja las del norte.

Como último análisis de estructura poblacional para los marcadores neutrales de *E. puncticulatus*, se probó la posibilidad de cuatro hipótesis de agrupamiento: 1) panmixia (Tabla), 2) norte y sur (Tabla), 3) Los Cabos, La Paz y Loreto; Mulegé, y Bahía de los Ángeles (Tabla); y 4) Los Cabos y La Paz, Loreto, Mulegé, y Bahía de los Ángeles (Tablas XVI-XIX).

Al observar los resultados obtenidos en la hipótesis 1 que considera las 15 localidades como un mismo grupo (Tabla XII), se observa que el mayor porcentaje de variación se encuentra dentro de los individuos (86.68%), mientras que el resto de la variación se divide entre las poblaciones (4.58%) y los individuos dentro de las poblaciones (5.74%). De igual manera el valor de F_{ST} (0.06012, $p < 0.001$) fue significativo. La diferenciación significativa entre poblaciones nos sugiere la existencia de más de un grupo, que además poseen una fuerte variación intrapoblacional.

La única hipótesis que mostró diferencias no significativas entre poblaciones al interior de los grupos fue en la que se sugieren cuatro grupos (Tabla XV), pero también se observó un menor porcentaje de varianza entre grupos (5.709%), así como un menor índice de fijación entre grupos ($F_{CT} = 0.05709$) en comparación con los valores encontrados para las hipótesis de dos (6.34% de variación entre grupos, $F_{CT} = 0.0634$) (Tabla XIII), y tres grupos (6.65% de variación entre grupos, $F_{CT} = 0.066$) (Tabla XIV), estas últimas dos hipótesis obtuvieron resultados similares.

Tabla XII. AMOVA jerárquico con marcadores neutrales de *E. puncticulatus* para la hipótesis de panmixia.

Fuente de variación	Suma de cuadrados	Componentes de varianza	Porcentaje de variación	Índices de fijación	<i>P-value</i>
Entre poblaciones	7927.831	14.68916	4.58079	$F_{ST} = 0.06012$	0
Entre individuos dentro de poblaciones	35521.202	18.39647	5.73691	$F_{IS} = 0.06012$	0
Dentro de los individuos	35845.5	287.58277	89.68229	$F_{IT} = 0.10318$	0
Total	79294.532	320.6684			

Tabla XIII. AMOVA jerárquico con marcadores neutrales de *E. puncticulatus* para la hipótesis de dos grupos (1)norte, y 2) sur.

Fuente de variación	Suma de cuadrados	Componentes de varianza	Porcentaje de variación	Índices de fijación	<i>P-value</i>
Entre grupos	2930.897	20.96729	6.34149	FCT= 0.06341	0
Entre poblaciones dentro de los grupos	4996.934	3.69028	1.11611	FSC= 0.01192	0
Entre individuos dentro de poblaciones	35521.202	18.39647	5.56395	FIS= 0.06012	0
Dentro de los individuos	35845.5	287.58277	86.97845	FIT= 0.13022	0
Total	79294.532	330.6368			

Tabla XIV. AMOVA jerárquico con marcadores neutrales de *E. puncticulatus* para la hipótesis de tres grupos (1) Los Cabos La Paz y Loreto, 2) Mulegé, y 3) Bahía de los Ángeles.

Fuente de variación	Suma de cuadrados	Componentes de varianza	Porcentaje de variación	Índices de fijación	<i>P-value</i>
Entre grupos	3810.688	21.87387	6.64786	FCT= 0.06648	0
Entre poblaciones dentro de los grupos	4117.143	1.18327	0.35962	FSC= 0.00385	0
Entre individuos dentro de poblaciones	35521.202	18.39647	5.59101	FIS= 0.06012	0
Dentro de los individuos	35845.5	287.58277	87.40151	FIT= 0.12598	0
Total	79294.532	330.6368			

Tabla XV. AMOVA jerárquico con marcadores neutrales de *E. puncticulatus* para la hipótesis de cuatro grupos (1) Los Cabos y La Paz, 2) Loreto, 3) Mulegé, y 4) Bahía de los Ángeles.

Fuente de variación	Suma de cuadrados	Componentes de varianza	Porcentaje de variación	Índices de fijación	<i>P-value</i>
Entre grupos	4377.301	18.52132	5.70945	FCT= 0.05709	0
Entre poblaciones dentro de los grupos	3550.53	-0.10324	-0.03183	FSC= -0.00034	0.47149
Entre individuos dentro de poblaciones	35521.202	18.39647	5.67097	FIS= 0.05709	0
Dentro de los individuos	35845.5	287.58277	88.6514	FIT= 0.11349	0
Total	79294.532	324.39731			

7.10. Migración

Se estimaron los patrones de flujo génico a partir de los marcadores neutrales entre los 15 sitios. En la figura 24 se puede apreciar como los 15 sitios se agrupan en tres *clusters* principales: 1) Los Cabos, La Paz y Loreto, 2) Mulegé, y 3) Bahía de los Ángeles, sugiriendo que dentro de estos grupos existe un fuerte nivel de retención de material génico. Donde los sitios en Mulegé reciben y aporta material génico al resto de los sitios, sugiriendo la posibilidad de que actúen como una zona de mezcla.

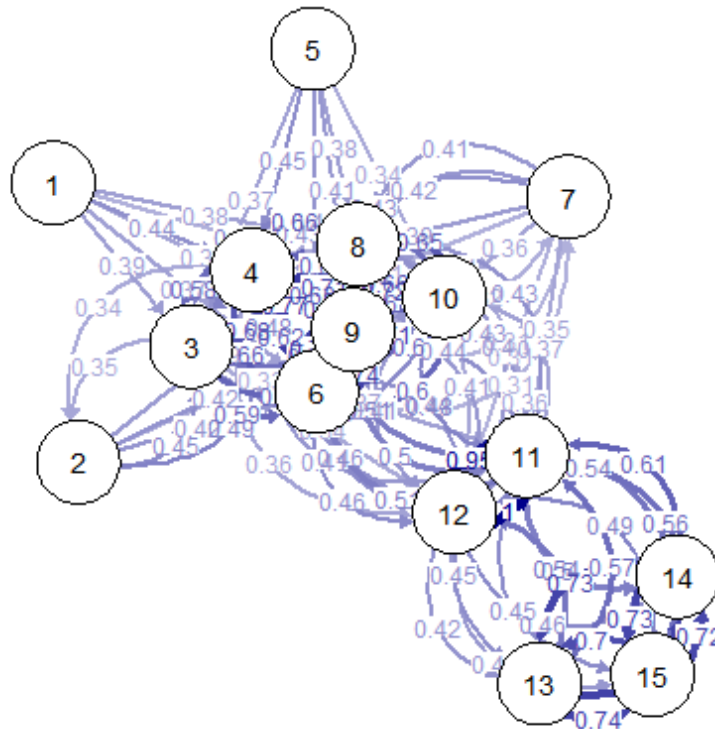


Figura 24. Red de migración relativa entre los 15 sitios para marcadores neutrales. La flecha indica la dirección del flujo, y su grosor la magnitud. Se presentan valores de $G_{ST} > 0.3$. 1: BAFPA, 2: SALTI, 3: PD, 4: PISF, 5: SANJPS, 6: PB, 7: MAR, 8: IAV, 9: ISAD, 10: ISANC, 11: PCSMS, 12: SMN, 13: CC, 14: VEN, y 15: BOT.

Con base en las evidencias de los análisis de estructura poblacional se muestra un mapa con las probabilidades de pertenencia de los sitios a dos grupos genéticos (q -value) de Structure de *E. puncticulatus* (Fig. 25), donde se puede observar como el color naranja se asocia más con los sitios del norte, mientras que el azul con los del sur. También se puede observar un gradiente, donde los sitios de Mulegé y Loreto actúan como una zona de transición o de mezcla entre las dos poblaciones.

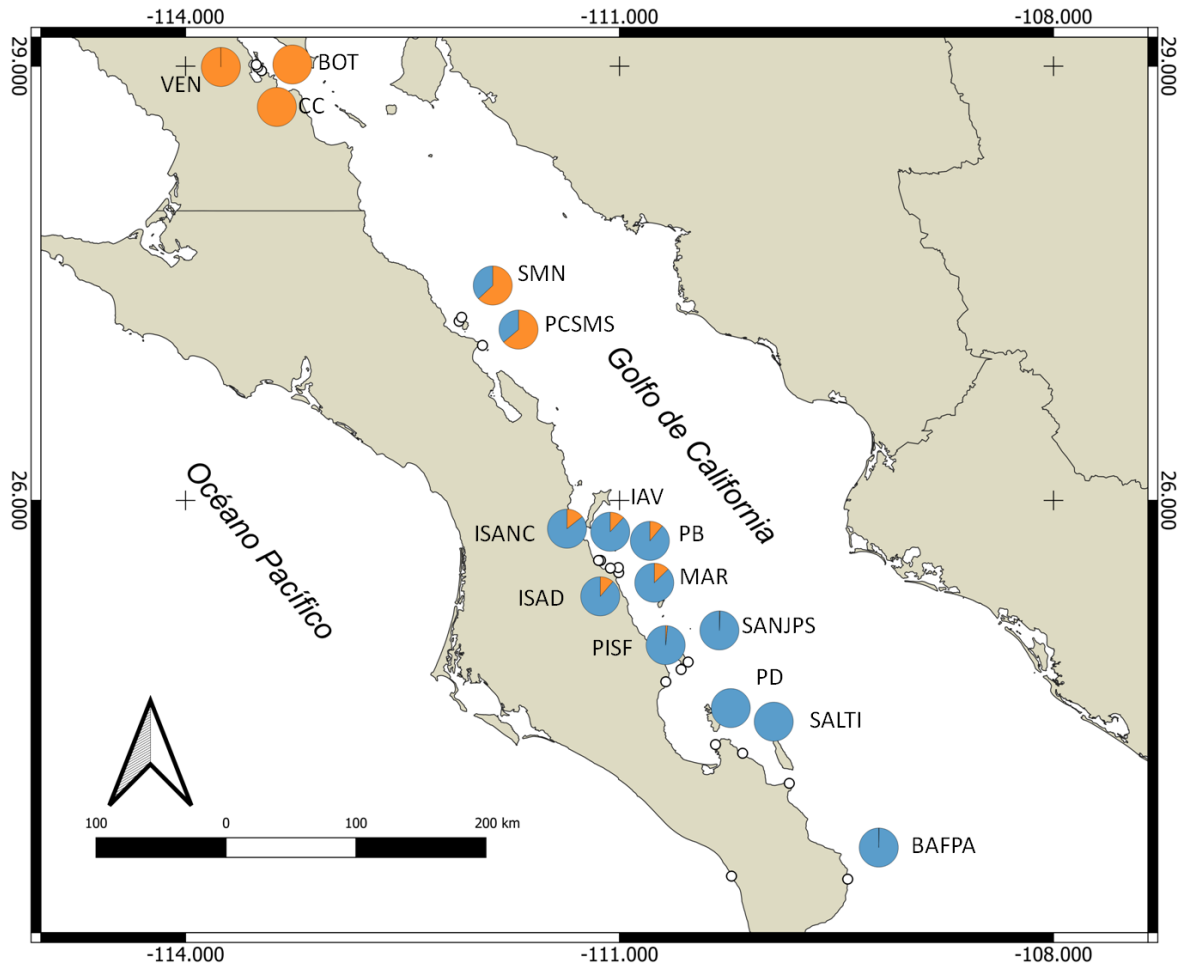


Figura 25. Proporción de la asignación de pertenencia de cada uno de los 15 sitios a los dos grupos de *E. puncticulatus* detectados en STRUCTURE (K=2) con marcadores neutrales. (INEGI *et al.*, 1990).

7.11 Estructura poblacional con marcadores *outliers*.

Se determinó la diferenciación genética entre sitios mediante valores pareados de F_{ST} (Tabla XVI) y al igual que los marcadores neutrales se observa como localidades cercanas se agrupan, mientras que las lejanas se diferencian. Sin embargo, de forma general, se observan valores de F_{ST} mayores que los encontrados para marcadores neutrales, sugiriendo un mayor grado de divergencia genética entre los sitios de muestreo. Después de aplicar la corrección de Bonferroni ($p \leq 0.00098$), al igual que con marcadores neutrales se observa un quiebre genético entre los sitios ubicados en la región sur y norte a partir de la zona entre Loreto y Mulegé. Los marcadores *outliers* también indican variación al interior de las regiones.

Tabla XVI. F_{ST} pareados entre sitios para marcadores *outliers* de *E. puncticulatus*. Valores de F_{ST} debajo de la diagonal, y los valores de p sobre la diagonal. Valores en negrita resultaron significativos después de la corrección de Bonferroni (valor de p con la corrección de Bonferroni ($p \leq 0.00098$))(Rice, 1989).

	BAFPA	SALTI	PD	PISF	SANJPS	PB	IAV	MAR	ISANC	ISAD	PCSMS	SMN	CC	VEN	BOT
BAFPA	---	0.134	0.112	0.229	0.261	0.009	0.229	0.166	0.220	0.218	0.001	0.001	0.001	0.001	0.002
SALTI	0.047	---	0.056	0.060	0.007	0.001	0.048	0.001	0.005	0.001	0.000	0.000	0.001	0.001	0.001
PD	0.040	0.044	---	0.137	0.806	0.000	0.080	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.000
PISF	0.028	0.047	0.024	---	0.065	0.001	0.261	0.270	0.004	0.000	0.000	0.000	0.000	0.000	0.000
SANJPS	0.020	0.085	-0.030	0.046	---	0.020	0.293	0.005	0.014	0.011	0.000	0.001	0.001	0.001	0.002
PB	0.091	0.115	0.103	0.071	0.077	---	0.471	0.012	0.202	0.009	0.000	0.000	0.000	0.000	0.000
IAV	0.031	0.074	0.053	0.023	0.028	0.010	---	0.276	0.884	0.076	0.000	0.001	0.001	0.001	0.001
MAR	0.034	0.116	0.089	0.013	0.084	0.050	0.019	---	0.153	0.054	0.000	0.000	0.000	0.000	0.000
ISANC	0.025	0.104	0.115	0.071	0.087	0.020	-0.027	0.020	---	0.096	0.000	0.000	0.000	0.000	0.000
ISAD	0.040	0.157	0.132	0.089	0.143	0.064	0.062	0.042	0.036	---	0.000	0.000	0.000	0.000	0.000
PCSMS	0.413	0.421	0.392	0.348	0.394	0.260	0.255	0.286	0.264	0.281	---	0.004	0.000	0.000	0.000
SMN	0.350	0.378	0.342	0.299	0.328	0.206	0.193	0.235	0.231	0.223	0.044	---	0.000	0.000	0.000
CC	0.603	0.602	0.563	0.528	0.597	0.431	0.474	0.470	0.458	0.465	0.201	0.152	---	0.716	0.902
VEN	0.628	0.628	0.587	0.546	0.626	0.459	0.507	0.493	0.482	0.488	0.194	0.151	-0.012	---	0.295
BOT	0.599	0.598	0.561	0.525	0.592	0.418	0.472	0.472	0.457	0.462	0.201	0.151	-0.026	0.005	---

El análisis de *clusters* de marcadores *outliers* sugiere que la K más probable por el método de Evanno es $K=2$ (Fig. 26a), pero nuevamente el método de $\text{Pr}(K)$ sugiere una $K=4$ (Fig. 27b). Cabe mencionar que la segunda estructura más probable sugerida por el método de Evanno ($K=3$) no coincide con la obtenida con $\text{Pr}(K)$.

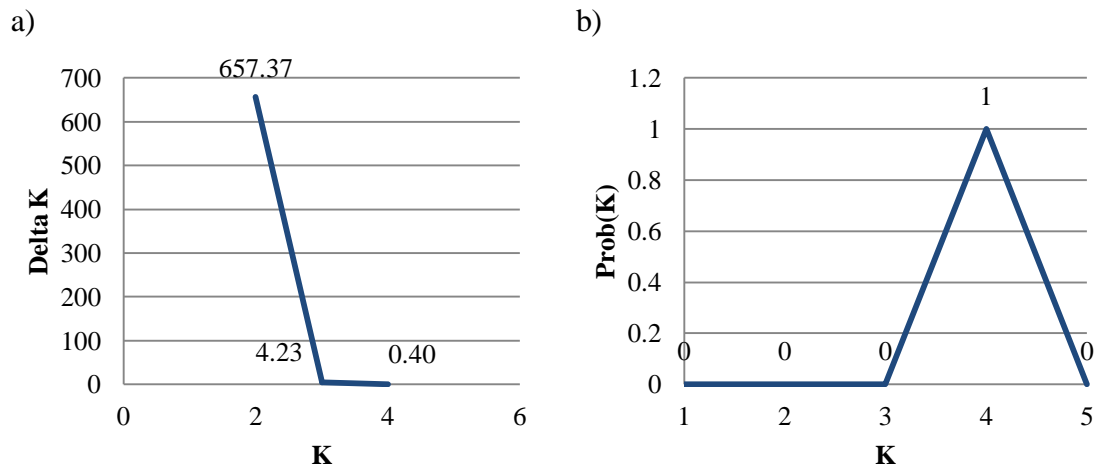


Figura 26. K más probable para marcadores *outliers*. a) Magnitud de ΔK para distintos valores de K (Evanno *et al.*, 2015). b) $\text{Pr}(K)$ para cada K (Pritchard *et al.*, 2000).

Debido a los resultados obtenidos, se presentan tres análisis de *cluster* para distintas K (Fig. 27). El número de poblaciones más probable sugerido por el método de Evanno fue $K=2$ e indica que las poblaciones se dividen en norte y sur, con un quiebre genético ubicado entre los sitios de Loreto y Mulegé, agrupando los mismo sitios que con los marcadores neutrales (Fig. 28a). La agrupación $K=3$ (Fig. 27b), separa a los sitios de las latitudes extremas de los de las medias: 1) Los Cabos y La Paz (BAFPA, SALTI, PD, PISF, y SANJPS), 2) Loreto y Mulegé (PB, IAV, MAR, ISANC, ISAD, PCSMS y SMN), y 3) Bahía de los Ángeles (CC, VEN, y BOT). Finalmente, $K=4$ agrupa los sitios de la siguiente manera: 1) Los Cabos y La Paz (BAFPA, SALTI, PD, PISF, y SANJPS), 2) Loreto (PB, IAV, MAR, ISANC, e ISAD), 3) Mulegé (PCSMS y SMN), y 4) Bahía de los Ángeles (CC, VEN, y BOT) (Fig. 27c). La agrupación para cuatro grupos no es tan clara para los sitios de Mulegé, pues se observa una fuerte mezcla.

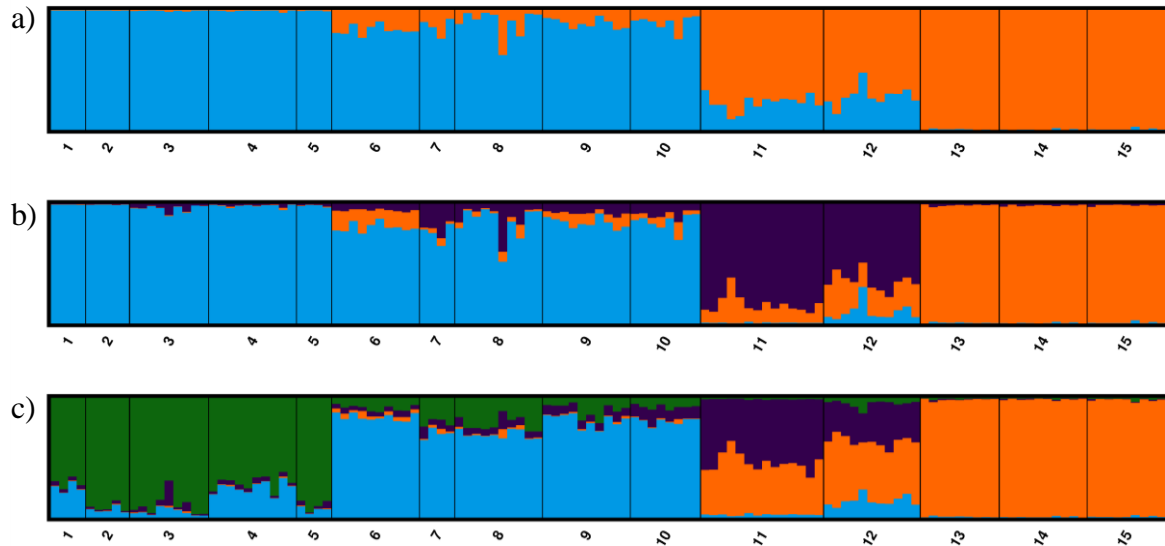


Figura 27. Agrupamiento de poblaciones con marcadores *outliers* (26 SNPs) de *E. puncticulatus* en el GC. a) Individuos con sus probabilidades de pertenencia agrupados en dos grupos principales (K=2) (azul: sur, y naranja: norte). b) Individuos con sus probabilidades de pertenencia agrupados en cuatro grupos (K=3) (azul: Los Cabos, La Paz, y Loreto, morada: Mulegé, y naranja: Bahía de los Ángeles). c) Individuos con sus probabilidades de pertenencia agrupados en cuatro grupos (K=4) (azul: Los Cabos y La Paz, morado: Loreto, verde: Mulegé, y naranja: Bahía de los Ángeles). 1: BAFPA, 2: SALTI, 3: PD, 4: PISF, 5: SANJPS, 6: PB, 7: MAR, 8: IAV, 9: ISAD, 10: ISANC, 11: PCSMS, 12: SMN, 13: CC, 14: VEN, y 15: BOT.

Por su parte, el ADCP que considera los 15 sitios de muestreo para la inferencia de clusters (Fig. 28a) coincide con la existencia de los mismos tres grupos identificados visualmente encontrados en el análisis de STRUCTURE para una K=3. Sin embargo, al considerar los grupos sugeridos por la función *find.clusters* determinado por el criterio de inferencia bayesiano (K=4, BIC=191.598) (Fig. 28b), solo una función discriminante es retenida, y sugiere que los individuos se diferencian en dos grupos 1) sur (BAFPA, SALTI, PD, PISF, y SANJPS), y 2) norte (PCSMS, SMN, CC, VEN, y BOT). Estos grupos coinciden con los encontrados para marcadores neutrales.

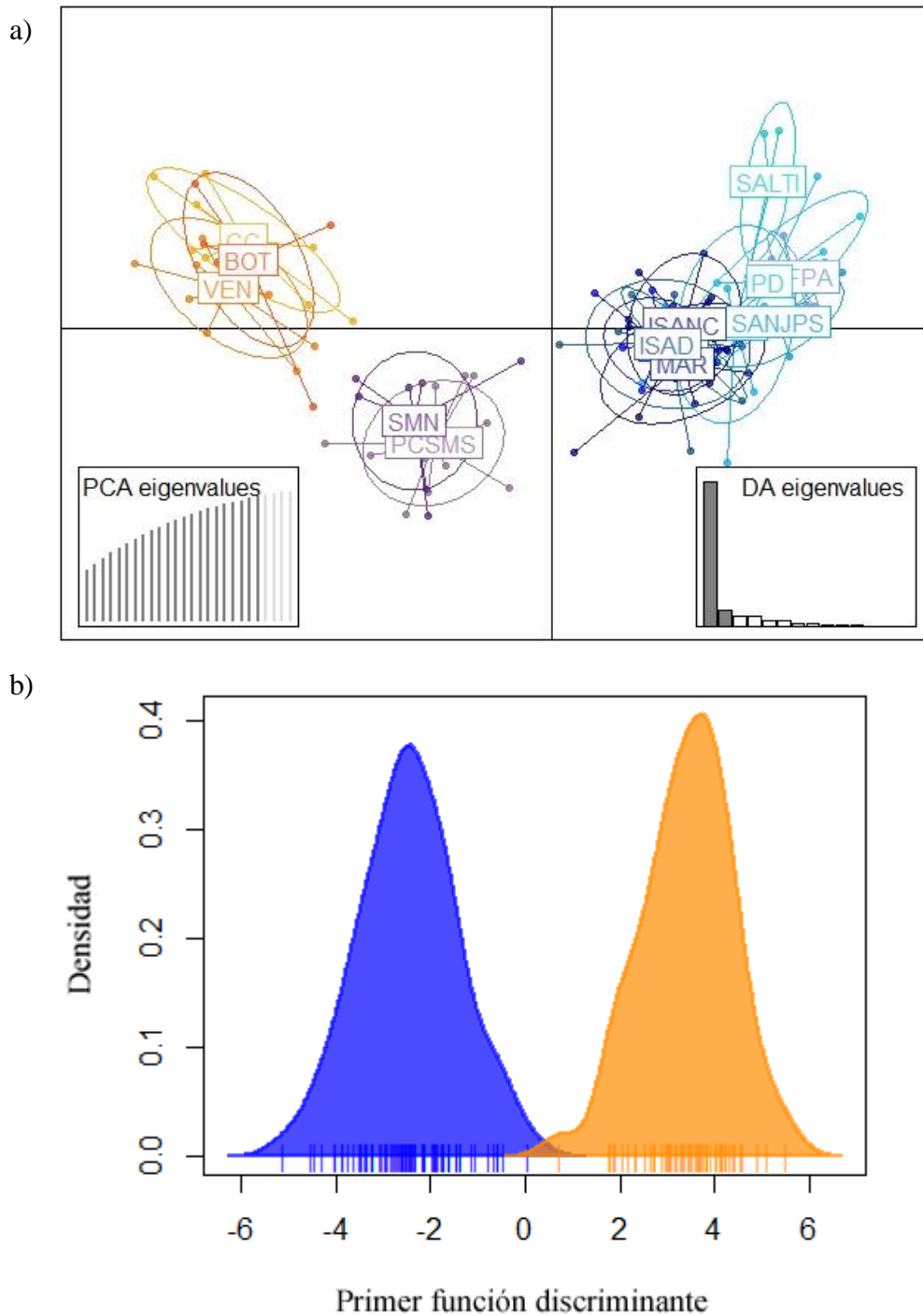


Figura 28. ADCP con marcadores *outliers* con 22 PCs retenidos. a) obtenido por localidad, con dos funciones discriminantes retenidas, y b) densidades individuales contra la primer función discriminante a partir de la K más probable obtenida con la función *find.clusters* (K=4, BIC=191.598), en color azul las localidades del sur y en naranja las del norte.

Se probaron cuatro hipótesis de estructura poblacional con diferentes AMOVAs jerárquicos: 1) panmixia, 2) norte y sur, 3) Los Cabos, La Paz y Loreto; Mulegé, y Bahía de

los Ángeles; y 4) Los Cabos y La Paz, Loreto, Mulegé y Bahía de los Ángeles (Tablas XVII-XX).

La prueba para panmixia sugiere una alta variación entre poblaciones (28% de variación, $F_{ST}= 0.28794$) (Tabla XVII). La mayor variación entre grupos se observó para la hipótesis de dos grupos (35.56%) (Tabla XVIII) pero también la mayor diferencia al interior de los mismos (5.27%). Al probar el AMOVA de tres grupos (Tabla XIX), disminuyó ligeramente la variación entre grupos (35.04%) y al interior de los mismos (3.29%). Para los cuatro grupos (Tabla XX) se observó el mismo patrón y se encontró el valor más bajo de variación al interior de los grupos (1.9%) sin embargo, el índice de fijación para este nivel jerárquico siguió presentando diferencias significativas ($F_{SC}= 0.03$, $p\text{-value}= 0.0059$).

Tabla XVII. Amova jerárquico con marcadores *outliers* para hipótesis de panmixia.

Fuente de variación	Suma de cuadrados	Componentes de varianza	Porcentaje de variación	Índices de fijación	$p\text{-value}$
Entre poblaciones	406.109	1.51909	28.79445	$F_{ST}= 0.28794$	0
Entre individuos dentro de poblaciones	432.021	0.18803	3.56414	$F_{IS}= 0.05005$	0.00941
Dentro de los individuos	445.5	3.56852	67.64141	$F_{IT}= 0.32359$	0
Total	1283.63	5.27564			

Tabla XVIII. AMOVA jerárquico con marcadores *outliers* para hipótesis de dos grupos: 1) norte, y 2) sur.

Fuente de variación	Suma de cuadrados	Componentes de varianza	Porcentaje de variación	Índices de fijación	$p\text{-value}$
Entre grupos	283.829	2.25753	35.56073	$F_{CT}= 0.35561$	0
Entre poblaciones dentro de los grupos	122.28	0.33429	5.2658	$F_{SC}= 0.08172$	0
Entre individuos dentro de poblaciones	432.021	0.18803	2.96188	$F_{IS}= 0.05005$	0.00911
Dentro de los individuos	445.5	3.56852	56.21159	$F_{IT}= 0.43788$	0
Total	1283.63	6.34837			

Tabla XIX. AMOVA jerárquico con marcadores neutrales para hipótesis de tres grupos: 1) Bahía de los Ángeles, 2) Mulegé, y 3) La Paz.

Fuente de variación	Suma de cuadrados	Componentes de varianza	Porcentaje de variación	Índices de fijación	<i>p-value</i>
Entre grupos	320.447	2.135	35.04517	$F_{CT} = 0.35045$	0
Entre poblaciones dentro de los grupos	85.662	0.20059	3.29263	$F_{SC} = 0.05069$	0
Entre individuos dentro de poblaciones	432.021	0.18803	3.08645	$F_{IS} = 0.05005$	0.01
Dentro de los individuos	445.5	3.56852	58.57574	$F_{IT} = 0.41424$	0
Total	1283.63	6.0921			

Tabla XX. AMOVA jerárquico con marcadores *outliers* para hipótesis de cuatro grupos: 1) Bahía de los Ángeles, 2) Mulegé, 3) Loreto, y 4) La Paz norte y La Paz sur.

Fuente de variación	Suma de cuadrados	Componentes de varianza	Porcentaje de variación	Índices de fijación	<i>p-value</i>
Entre grupos	344.022	1.76741	31.38652	$F_{CT} = 0.31387$	0
Entre poblaciones dentro de los grupos	62.088	0.10741	1.90269	$F_{SC} = 0.02773$	0.0059
Entre individuos dentro de poblaciones	432.021	0.18803	3.33916	$F_{IS} = 0.05005$	0.00871
Dentro de los individuos	445.5	3.56852	63.37163	$F_{IT} = 0.36628$	0
Total	1283.63	5.6311			

Los análisis de estructura presentados anteriormente, dan indicios de la existencia de tres grupo con potencial adaptativo de *E. puncticulatus* a lo largo del GC. El mapa de la figura 29, nos muestra la probabilidad de pertenencia promedio de cada uno de los sitios de muestreo a los tres grupos encontrados. Donde el grupo de Los Cabos y La Paz, es representado por una predominancia del color azul, el cual disminuye conforme la latitud aumenta. En los sitios del grupo de Mulegé destaca el morado, seguido del color naranja; el cual representa el grupo de Bahía de los Ángeles, en donde la mezcla es casi nula.

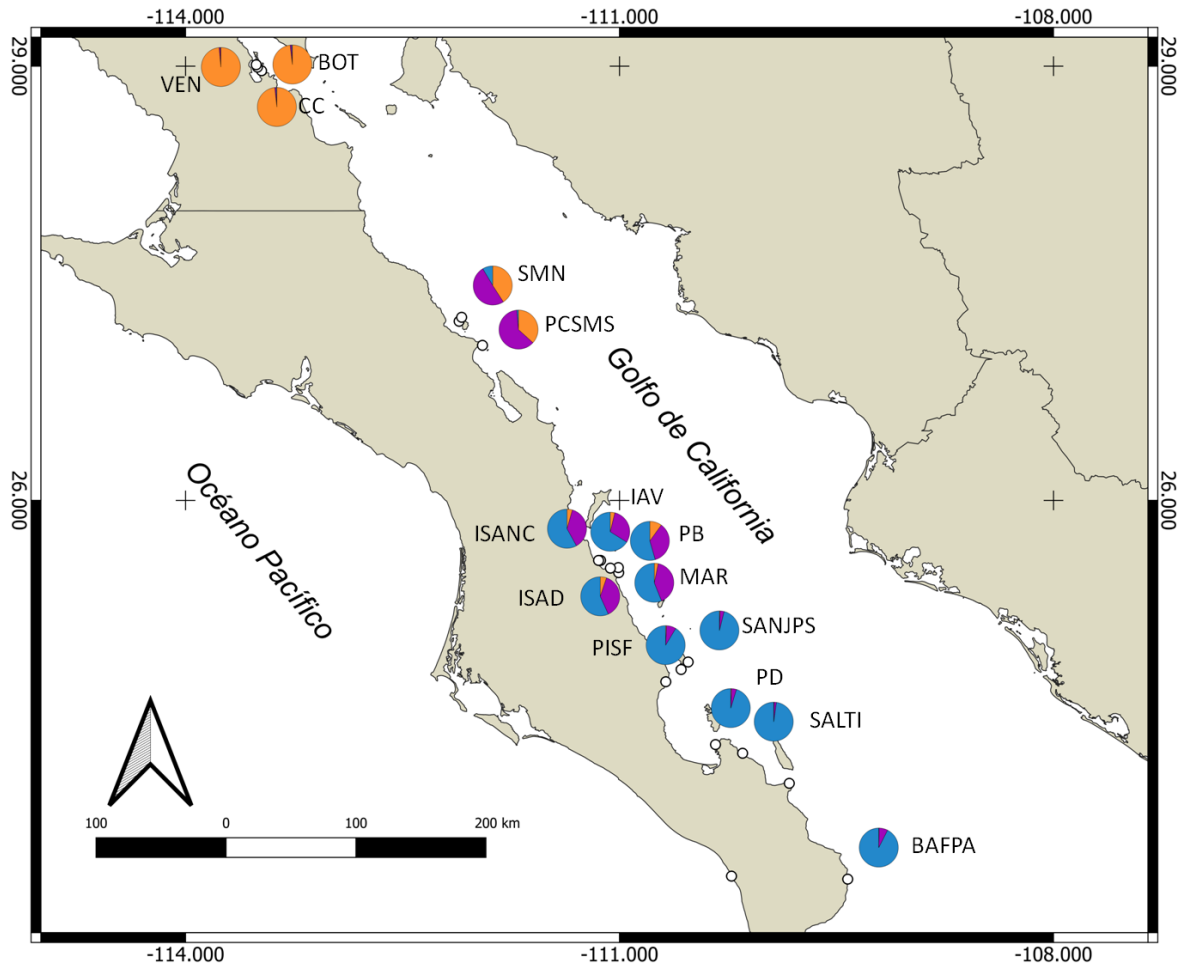


Figura 29. Probabilidad de pertenencia de los 15 sitios a los tres grupos con potencial adaptativo de *E. puncticulatus* identificadas con STRUCUTRE (K=3) con marcadores *outliers* (INEGI *et al.*, 1990).

Ya que los marcadores *outliers* podrían ser considerados como marcadores candidatos a selección, en las figuras 30 y 31 se muestran las probabilidades de asignación de cada sitio a los tres grupos detectados, y como cada uno de estos grupos se distribuyen sobre un interpolado de la media anual de temperatura y salinidad.

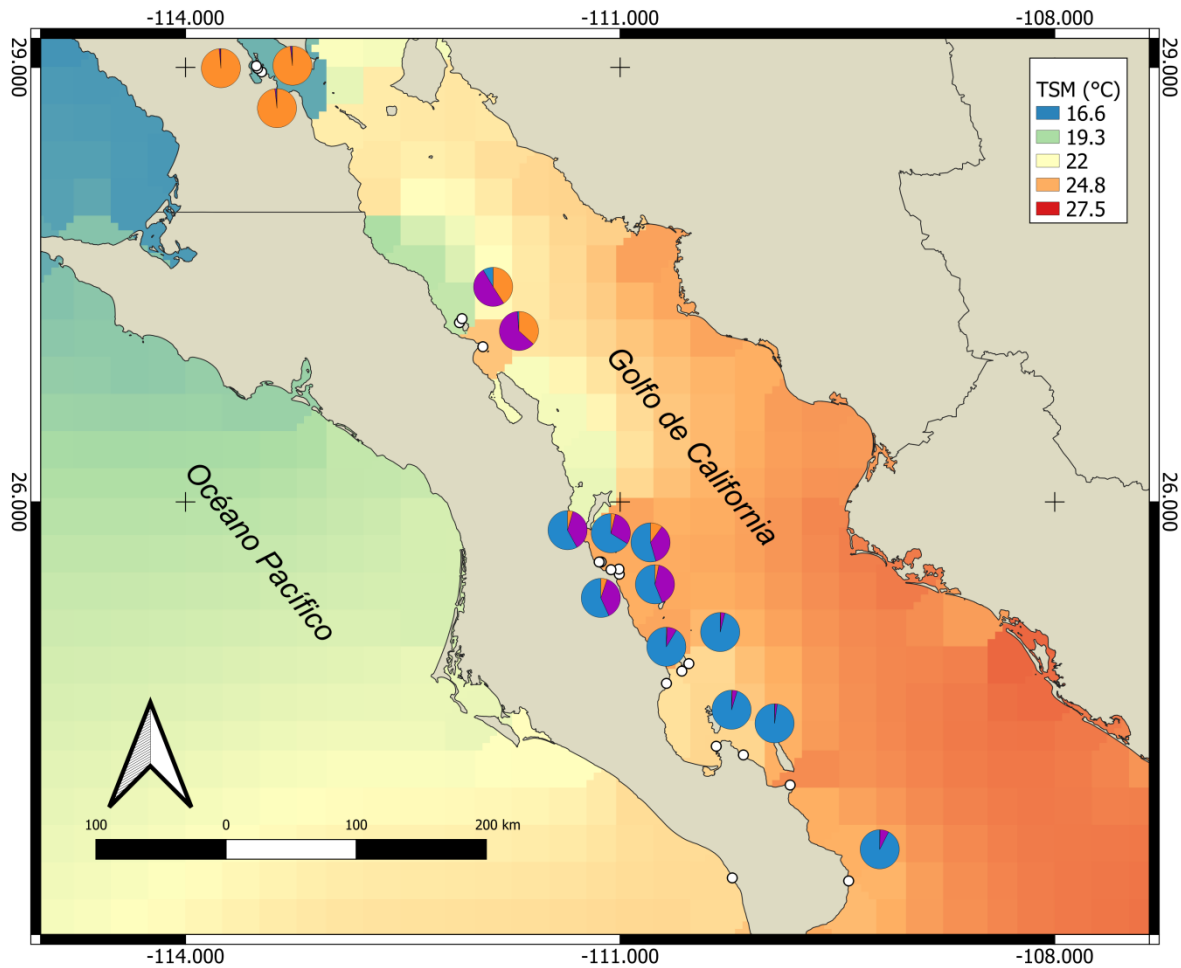


Figura 30. Probabilidad de pertenencia de los 15 sitios a los tres grupos con potencial adaptativo de *E. puncticulatus* identificadas con STRUCUTRE (K=3) con marcadores *outliers*, y su distribución sobre la media de temperatura superficial (SST) de invierno (1955-2017) en el Golfo de California (INEGI *et al.*, 1990 .Locarnini *et al.*, 2019).

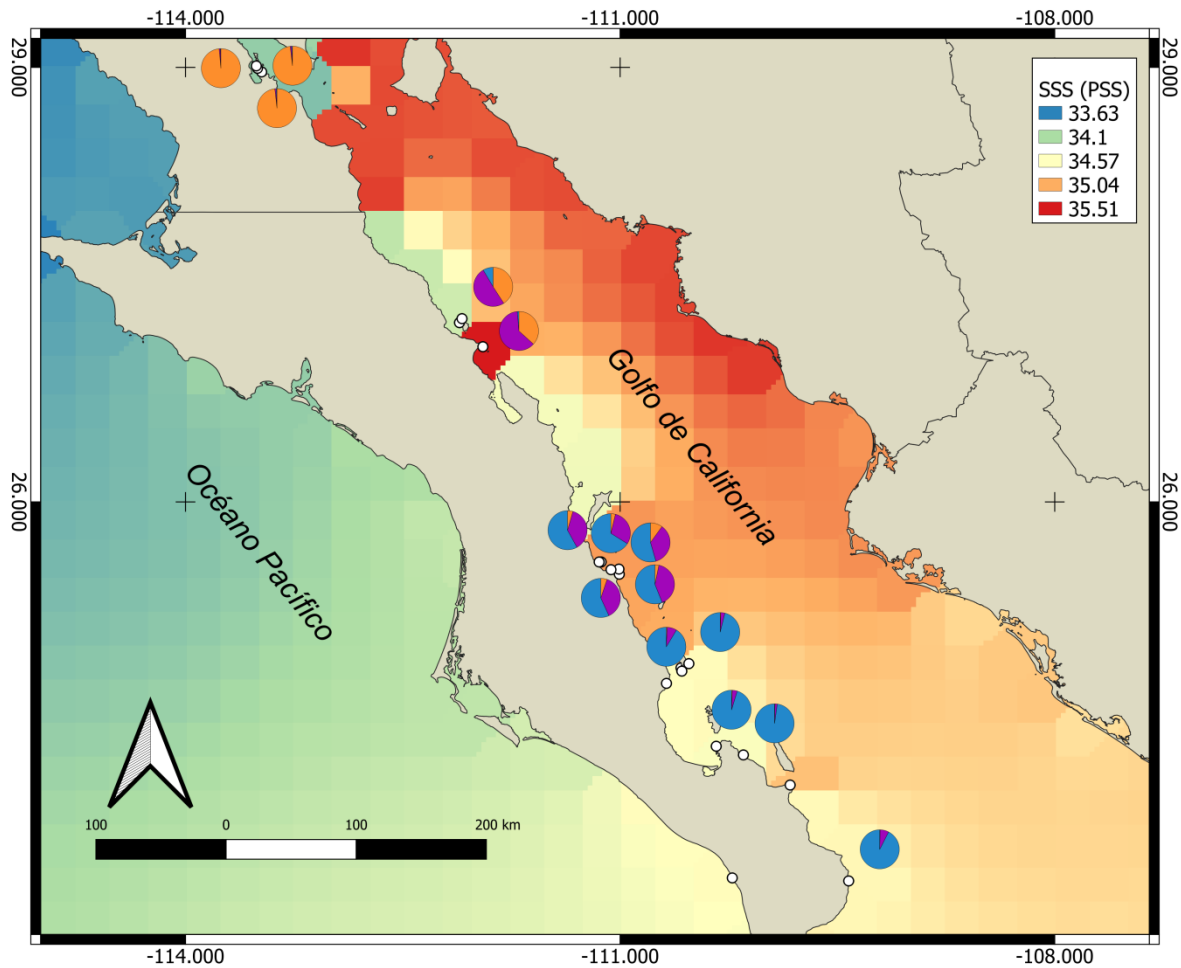


Figura 31. Probabilidad de pertenencia de los 15 sitios a los tres grupos con potencial adaptativo de *E. puncticulatus* identificadas con STRUCUTRE (K=3) con marcadores *outliers*, y su distribución sobre la media de salinidad superficial (SSM) anual (1955-2017) en el Golfo de California (INEGI *et al.*, 1990; Zweng *et al.*, 2019).

8. DISCUSIÓN

El presente trabajo permitió la identificación de un total de 4,979 marcadores, de los cuales 4953 loci fueron marcadores neutrales y 26 fueron marcadores *outliers* (candidatos a selección), sin embargo se recomienda realizar más pruebas para evaluar si se encuentran bajo algún tipo de selección local. Ambos grupos de marcadores mostraron evidencia de una diferenciación en la estructura poblacional, la cual podría ser el resultado de una combinación de procesos históricos y contemporáneos que moldean la diversidad genómica. Para el caso de los neutrales; se sugiere la existencia de dos poblaciones, y se asume que la distancia geográfica, la circulación oceanográfica, y eventos de vicarianza, han jugado un papel importante en el moldeo de las poblaciones encontradas. Mientras que para *outliers*, se encontraron tres grupos con potencial adaptativo, y si su potencial adaptativo es real, se sugiere que la heterogeneidad ambiental podría estar explicando la división de los grupos genéticos encontrados.

8.1 Estructura genómica poblacional neutral

La estructura genética inferida para *E. puncticulatus*, indicó la existencia de dos poblaciones con un quiebre genético localizado entre los sitios de Mulegé y Loreto, dividiendo las localidades norteñas (BOT, VEN, CC, SMN, y PCSMS) de las del sur (ISANC, ISAD, IAV, MAR, PB, SANJPS, PISF, PD, SALTI, y BAFPA). Otros autores coinciden en la existencia de dos poblaciones al interior del GC con peces criptobentónicos: *Axoclinus nigricaudus*, *Coralliozetus micropes*, *Malacoctenus hubbsi*, *Girella nigricans* y *Paralabrax maculatofasciatus*, sin embargo difieren en que este quiebre se encuentra más al norte, dividiendo a Bahía de los Ángeles, del resto de las localidades de la región centro, lo cual coincide con la bioregionalización del golfo propuesta por Walker (1960) (Riginos y Nachman, 2001, y Riginos, 2005). Cabe mencionar que para los peces en estos estudios, solamente se secuenció la región control mitocondrial, mientras que el presente trabajo utilizó casi 5,000 marcadores nucleares distribuidos a lo largo del genoma, por lo cual podría haber un cambio en la ventana temporal de los datos, así como una mejor representación del genoma. No obstante, si se encontró en todos los análisis de inferencia

de *clusters* y estadísticos de *F*, que la región más divergente del resto, fue Bahía de los Ángeles.

Por el contrario, la estructura genómica recientemente encontrada para el pez *Lyprinus dalli* perteneciente a la misma familia (Gobiidae) que *E. puncticulatus* y colectada simultáneamente con las muestras del presente estudio, sugiere la existencia de una población panmíctica, sin diferencias entre los marcadores neutrales y *outliers* (Mac Loughlin-Aleman, 2019). A pesar de esto coincide en que los individuos de Bahía de los Ángeles, presentan el menor grado de parentesco con respecto al resto de los sitios. Estas coincidencias en el aislamiento de esta región del resto de los sitios en el sur, nos podría estar hablando de la existencia de algún factor extrínseco a las especies que propicie la divergencia de estos individuos. Otras discrepancias en la estructura poblacional ya han sido reportadas para especies con historias de vida similares en esta zona. Para la especie de pez críptico *Acanthemblemaria crockeri*, dentro del GC, utilizando dos marcadores mitocondriales, se encontraron dos linajes monofiléticos, sugiriendo una especiación parapátrica, la cual podría ser explicada mediante la existencia de un canal marino hipotético. Sin embargo, al compararlo con *Acanthemblemaria macrospilus*, encontraron que esta muestra niveles muy bajos de divergencia para individuos de la misma región, esto sugiere que otros mecanismos contemporáneos, como las características ambientales, podrían estar jugando un fuerte rol sobre los patrones de estructura genética de las especies (Lin *et al.*, 2009).

Cabe destacar que el quiebre genético para los grupos genéticos de *E. puncticulatus* encontrado entre los sitios norte y sur no fue abrupto, sino que se observó una introgresión genética hacia las localidades centrales, Loreto (PB, IAV, MAR, ISANC, e ISAD) y Mulegé (PCSMS, y SMN). Al observar la asignación individual para más de dos grupos ($K > 2$), se encontró una mezcla en los individuos de esta región. Siendo esta una posible zona de mezcla o transición. Para lo que ya antes otros autores habían sugerido un flujo predominante norte a sur en esta zona (Riginos, 2005; Mac Loughlin-Aleman, 2019) sin embargo los grupos encontrados en STRUCTURE nos indican que la posibilidad de que esta introgresión hacia los sitios intermedios también esté ocurriendo de sur a norte, lo cual coincide con lo encontrado en la evaluación de flujo génico.

8.1.2 Efecto de la distancia geográfica y la circulación oceanográficas en el Golfo de California sobre patrones estructurales de la población

Se observó una fuerte correlación entre la distancia genética y la geográfica, lo cual indica aislamiento por distancia entre los sitios. Con una autocorrelación entre aquellos geográficamente más cercanos, es importante destacar que el rango geográfico utilizado en este trabajo es relativamente amplio (>700 km), por lo tanto se sugiere evaluar el grado en que ocurre la autocorrelación debido a que podría ser señal de autoreclutamiento y un fuerte potencial de reproducción entre sitios cercanos en una amplia extensión geográfica, lo cual no es extraño en ambientes marinos, donde las barreras son comúnmente permeables (Riginos y Liggins, 2013).

El aislamiento no es clinal, sino que ocurre a manera de parches, lo cual podría sugerir que no solo la distancia está produciendo este patrón, sino que un proceso distinto está ocasionando una diferenciación abrupta en los sitios marginales, esto ya ha sido reportado en otros trabajos y descrito como resultado del efecto de poblaciones periféricas, donde en poblaciones aisladas se ha observado un mayor efecto de deriva génica, caracterizado por menores niveles de diversidad y una mayor diferenciación (Eckert *et al.*, 2008; van Strien *et al.*, 2015), lo cual puede observarse en este trabajo para los sitios en Bahía de los Ángeles. Las poblaciones que presentan los valores más alto de diversidad nucleotídica, son las de Loreto y Mulegé, sugiriendo un fuerte flujo génico de las poblaciones de los extremos al centro, esto coincide con el análisis de migración, donde estas poblaciones son colocadas en el centro del esquema y reciben un aporte de material genético del resto de las poblaciones, a excepción de Loreto que se encuentra aislada de Bahía de los Ángeles. Estos altos niveles de diversidad genética podrían deberse a los grandes flujos genéticos que convergen en la región central del GC (Eguiarte *et al.*, 2013).

La dispersión de organismos marinos costeros asociados al bentos ocurre principalmente en su estadio larval pelágico, y durante esta etapa dependen en gran manera de variables biofísicas como las características propias de las larvas, variables ambientales que puedan afectar su sobrevivencia, y la circulación que propicia su transporte (Cowen y Sponaugle, 2009). En el GC el sistema de corrientes posee una marcada estacionalidad, en el sur el

flujo cambia su dirección dos veces al año, con un giro ciclónico en los meses de febrero y agosto, generando un flujo hacia el sur en la costa peninsular; y anticiclónico en mayo y noviembre (Marinone, 2003). Mientras que el flujo en el Canal de Ballenas, ubicado frente a los sitios de muestreo en Bahía de los Ángeles, se ha reportado un flujo superficial hacia el norte durante verano, el cual se revierte en invierno (Marinone, 2003). Cabe mencionar que existe una fuerte retención de partículas debida a débiles corrientes residuales generadas por una fuerte mezcla de mareas (Marinone, 2012). En este trabajo se encontró un flujo génico asimétrico y bidireccional a manera de *stepping stone*, con agrupaciones discretas y aislamiento por distancia. Esto difiere de lo encontrado en otros trabajos para diversos peces criptobentónicos donde se ha observado un flujo predominante de norte a sur. Sandoval-Huerta *et al.* (2018) para *E. punctulatus* encontraron una dispersión de la provincia del Mar de Cortez a la mexicana, pero este resultado se obtuvo a partir de la variación obtenida en solo tres genes (dos mitocondriales y uno nuclear) y 26 individuos, cabe mencionar que este trabajo no alcanzó a detectar diferencias en la estructura poblacional al interior del GC, por lo cual, para este propósito, este trabajo posee datos a más fina escala. De igual manera se tiene el trabajo de Riginos (2005) quien recopiló secuencias de la región control para *Axoclinus nigricaudus*, *Coralliozetus micropes*, *Malacoctenus hubbsi*, *Girella nigricans* y *Paralabrax maculatofasciatus*, y encontró un patrón asimétrico con dirección hacia el sur. Finalmente se tiene el trabajo de Mac Loughlin-Aleman (2019), quien encontró que el flujo génico para *Lythrypnus dalli*, ocurre predominantemente de norte a sur, el cual lo asoció principalmente a la circulación primaveral (abril-junio).

La bidireccionalidad en el flujo génico encontrado podría estar asociado a que la especie tenga dos picos reproductivos en temporadas distintas del año; o por el contrario que el asentamiento larval ocurra cuando las corrientes de Bahía de los Ángeles y la región sur del GC, fluyan hacia el sur y norte, respectivamente. Otros trabajos han reportado en el sur del GC una corriente costera del lado de la península que viaja hacia el norte, que coincide con un flujo hacia el sur desde el Canal de Ballenas, aproximadamente de mayo-junio y de noviembre a enero (Marinone, 2003; Marinone, 2012). Para comprender más sobre su conectividad se sugiere probar la correlación entre la información genética y la circulación del GC en diferentes temporadas.

El quiebre genético observado entre Mulegé y Loreto nos habla de que existe una combinación tanto de aislamiento por distancia como por la circulación oceanográfica. Marinone (2012) en un estudio por describir la conectividad superficial en el GC, realizó una regionalización de la zona, describiendo la región desde la bahía de La Paz, hasta Bahía Concepción como una zona de débiles corrientes costeras, y menciona que en esta región el autoreclutamiento es alto. Por lo cual se sugiere que el transporte larval de los sitios en Mulegé, se ven limitados, de manera clinal, por una débil corriente costera que favorece al autoreclutamiento y una amplia distancia geográfica. Los altos niveles de diferenciación de Bahía de los Ángeles para todas las pruebas de inferencia de estructura genética, nos habla de la existencia de una fuerte barrera permeable en esta zona, que sumada al aislamiento por distancia, provoca la divergencia de este sitio del resto. Es posible que el alto dinamismo de la región, causado por mezcla de mareas, genera débiles corrientes residuales, esto aumenta el potencial de autoreclutamiento de la zona (Marinone, 2012). Sin embargo, el grado de flujo génico que mantiene con los sitios más cercanos (Mulegé), es suficiente para ser agrupados dentro de la misma población. A pesar de la fuerte diferenciación, Bahía de los Ángeles, no se encuentra completamente aislado, se sugiere que esto se debe en gran parte a que el área San Francisquito, ubicada entre los sitios de Bahía de los Ángeles y Mulegé, actúa como una zona de tránsito, es decir, no retiene partículas si no que favorecen su dispersión, pues en ese punto se encuentra una corriente de retorno, y una frontera de eddy. (Marinone, 2012).

8.1.2 Efecto de eventos vicariantes en el Golfo de California sobre patrones estructurales de la población

A pesar de que las localidades del sur y del norte presentan una fuerte diferenciación genética, se alcanza a ver una ligera introgresión la cual disminuye con la distancia hacia los extremos norte y sur este patrón podría ser evidencia de contacto secundario, aunque bajo el efecto de aislamiento por distancia, las clinas de mezcla deben ser consideradas con cuidado, pues podría confundirse (Hedgecock *et al.*, 2007; Bertl *et al.*, 2018); y para este estudio los efectos del aislamiento por distancia estuvieron presentes. Otras características de contacto secundario después de un evento de vicarianza son el aumento en la diversidad genética y desequilibrio de ligamiento (Gompert y Buerkle, 2016; Bertl *et al.*, 2018). En la

zona de contacto se observa un aumento de diversidad nucleotídica, sin embargo este efecto no se observó en el grado de heterocigosis. Por otro lado se buscó reducir el riesgo de desequilibrio de ligamiento en nuestros datos al permitir únicamente un SNP por loci. Por lo tanto se cuenta con evidencia parcial de este efecto en la zona central del GC.

Existe evidencia de que en el pasado (Mioceno medio-Mioceno tardío) la península pudo haber sido fragmentada por un canal marino al sur de la Sierra San Pedro Mártir, creando una conexión entre el GC y el Océano Pacífico, actuando como barrera entre las poblaciones del norte y del sur, estudios geológicos, basados en evidencia de fósiles y la distribución de depósitos marinos y no marinos (Helenes y Carreño, 1999 en Ledesma *et al.*, 2009; Ledesma-Vázquez, 2002 en Ledesma *et al.*, 2009; Carreño y Smith 2007 en Ledesma *et al.*, 2009). Se sugiere que este canal cruzó desde El Barril en el GC, hasta San Ignacio en el Pacífico (Ledesma *et al.*, 2009). Sin embargo este es un tema que no ha sido completamente esclarecido, pues aún quedan muchas dudas alrededor de esta teoría, y evidencias filogenéticas difieren de las geológicas, principalmente en tanto a la temporalidad del evento.

Upton y Murphy (1997) propusieron la existencia de esta barrera en el desierto de Vizcaíno para explicar la divergencia entre poblaciones de lagartos pertenecientes al género *Uta*, sugiriendo que esta ocurrió hace 1 Ma. Riginos (2005), encontró un patrón de vicarianza paralelo al de otras especies terrestres, para cinco especies de peces criptobentónicos, encontrando un quiebre que separaba a los individuos de Bahía de los Ángeles del resto de los sitios en la región centro del GC, pues supone que para especies costeras con un potencial de dispersión bajo, esta pudo significar una importante barrera. A partir de relojes moleculares coincidió que esta divergencia pudo haber ocurrido alrededor del Plio-Pleistoceno. Sin embargo, el trabajo de Hurtado *et al.* (2010), difiere de otros estudios genéticos, sugiriendo que la divergencia entre clados de isópodos del género *Ligia* ocurrió durante el Mioceno, y propone dos escenarios distintos: uno en que la divergencia entre clados ocurrió dentro del GC, y esta divergencia pudo ocurrir por la formación del canal marino, o como hipótesis alternativa, por un puente de tierra que separaba las cuencas del norte y del sur, pues Ledesma-Vázquez (2002) en Hurtado *et al.* (2010), propone que las cuencas marinas se formaron independientemente, siendo la de la región norte la primera en

formarse. El segundo escenario sugiere que la divergencia ocurrió en el Pacífico, antes de que los isópodos entraran al GC.

En nuestro caso a pesar de la fuerte diferenciación encontrada para los sitios en Bahía de los Ángeles, los análisis de inferencia de clúster sugieren que el quiebre ocurre más al sur de donde se especuló el quiebre genético y geográfico, es posible que el flujo génico actual, y aspectos de la biología de la especie hayan borrado parte de esa señal. Cabe mencionar que Riginos (2005) sugiere que el desfase del quiebre génico con el canal marino, podría ser ocasionado por una transición al límite entre las bioregiones propuestas por Walker (1960) quien sugiere la existencia de un límite biogeográfico entre la región norte y central, al sur de Bahía de los Ángeles; El Khattabi-Salazar (2019), ha encontrado evidencia en la composición de peces crípticos, de que la región central propuesta por Walker (1960), puede ser dividida en centro y sur, coincidiendo este límite, con el quiebre génico encontrado en este trabajo.

Poblaciones disyuntas entre el Pacífico y el GC, se suman a la evidencia de la existencia de un canal marino (*Lypthripnus dalli*, *Paralabrax maculatofasciatus*, *Chaenopsis alepidota*, *Girella nigricans*, *Leuresthes sardina/tenuis*, *Anisotremus davidsonii*, *Hypsoblennius jenkinsi*, *Gillichthys mirabilis* (Bernardi *et al.*, 2003, y Bernardi, 2014), y *Hypsoblennius jenkinsi* (Bernardi, 2014)), esto hace que la hipótesis de vicarianza sea más parsimoniosa para explicar la divergencia de estos peces entre el Golfo y Pacífico. Sin embargo, al interior del Golfo el mantenimiento o erosión de la huella de vicarianza depende en mayor grado de otros factores como el patrón de corrientes, así como ecológicos como la preferencia de hábitat y características ambientales, los cuales denotan el efecto de la selección natural pero han sido poco explorados para una mejor comprensión de los patrones de estructuración contemporáneos en especies costeras de la región (Riginos, 2005).

De igual manera se ha observado un quiebre genético para poblaciones del lado del Pacífico, en Punta Eugenia, al norte de donde se supone existió el canal a media península, entre ellos los peces que mostraron poblaciones disyuntas en el trabajo de Bernardi *et al.* (2003), *Paralabrax maculatofasciatus* (Tranah y Allen, 1999 en Riginos, 2005), *Embiotoca*

jacksoni (Bernardi, 2000, en Riginos, 2005), *Fundulus parvipinnis* (Bernardi y Crane, 2000, en Riginos, 2005), *Girella nigricans* (Terry *et al.*, 2000, en Riginos, 2005), *Gillichthys mirabilis* (Huang y Bernardi, 2001, en Riginos, 2005), y en la estructura poblacional obtenida con marcadores *outliers* del abulón *Haliotis corrugata* (Mares-Mayagoitia, 2019), los cuales coinciden en ser especies con historias de vida ligadas a ambientes rocoso costeros.

A pesar de la evidencia que apunta a considerar el efecto de eventos vicariantes sobre los patrones encontrados en la estructura de *E. puncticulatus* en el GC, se sugiere que los efectos contemporáneos como las corrientes oceanográficas pueden ser el mecanismo preponderante en el patrón neutral observado aunado a posibles efectos selectivos ambientales.

8.1.4 Efecto de la capacidad dispersiva de la especie en el Golfo de California sobre patrones estructurales de la población

Elacatinus puncticulatus es un pez críptico bentónico, que como muchas otras especies, su historia de vida se encuentra fuertemente asociada a la línea de costa. Tanto su etapa como huevo, y adulto, se encuentra fuertemente ligada los ambientes rocosos costeros; y su capacidad de dispersión se ve restringida a su etapa larval. Riginos (2005), encontró que para las especies criptobentónicas *C. micropes* and *A. nigricaudus* con baja dispersión larval y que mantienen sus larvas cerca de la costa una mayor divergencia genética entre las poblaciones centro y norte del GC, que para *M. hubbsi*, *G. nigricans*, y *P. maculatofasciatus*, cuyas larvas se han encontrado alejadas de la costa.

Sin embargo, en un trabajo que buscó evaluar la estructura poblacional de *Lyprinus dalli*, un pez con una baja capacidad dispersiva en su etapa adulto se encontró una población panmíctica, pero se logró detectar diferencias en los niveles de parentesco de la región norte, contra la región sur (Mac-Loughlin-Aleman, 2019). De igual manera, en un estudio realizado con tres especies de blénidos se encontró una disyunción para dos de las especies entre poblaciones del Pacífico y el GC (Bernardi, 2014), pero la reconstrucción filogenética parece no mostrar diferencias evidentes entre la región sur y norte del GC. Esto podría sugerir la importancia de las características intrínsecas de las especies sobre la

estructuración de las poblaciones marinas costeras, las cuales deberían ser evaluadas para determinar sus efectos. Bernardi (2014), ya había hecho un énfasis en que el hábitat preferente de las especie durante su etapa adulta, juega un papel incluso más importante que la duración larval pelágica.

El hecho de que encontremos resultados diferentes en el flujo génico de *Elacatinus puncticulatus* contra otras especies, podría estar asociado principalmente a factores intrínsecos de la biología de la especie que limiten su sobrevivencia y asentamiento en otras localidades. Desafortunadamente se desconoce la época de desove para esta especie en vida silvestre, aunque, se sabe que bajo condiciones de laboratorio la reproducción de esta especie posee una temperatura óptima de 26 °C (Burgees *et al.*, 2000 en Trujillo-García, 2018), y se sabe que para especies demersales en aguas someras del GC se ha observado que el mayor número de especies larvales y abundancia ocurre durante verano (Aceves-Medina *et al.*, 2003). Es durante la época de verano que en la región sur el flujo ocurre principalmente de norte a sur, mientras que en la región norte este patrón es contrario por la presencia de un giro anticiclónico (Marinone, 2003), y el potencial de retención larval es alto en la zona sur, favoreciendo al autoreclutamiento (Marinone, 2012), este patrón de corrientes podría dificultar el flujo encontrado hacia la región de Mulegé. Aunque se conoce que la circulación en el GC posee una estacionalidad marcada, donde existe un cambio en la dirección general de las corrientes (Marinone, 2003).

8.2 Estructura genómica poblacional con marcadores *outliers*

La estructura de marcadores *outliers* sugiere la existencia de tres grupos potencialmente adaptativos, dividiéndolos espacialmente en 1) norte (Bahía de los Ángeles), 2) centro (Mulegé), y 3) sur (Los Cabos y La Paz). Estas zonas se pueden caracterizar por su heterogeneidad ambiental marcada espacial y estacional. A lo largo del año es posible observar un gradiente donde las temperaturas más bajas se tienen en la parte norte, con un máximo en Bahía de los Ángeles, y estas van aumentando hacia el sur, por la influencia de aguas subtropicales en la entrada de la boca (Soto-Mardones *et al.*, 1999, Lavin y Marinone, 2003, Lluch-Cota *et al.*, 2007, y Ledesma *et al.*, 2009). En tanto a la variabilidad

de la temperatura, esta posee un patrón contrario al del promedio, ya que aumenta desde la boca hacia el interior del GC (Soto-Mardones *et al.*, 1999).

8.2.1 Efecto de la heterogeneidad ambiental en el Golfo de California sobre patrones estructurales de la población

La heterogeneidad ambiental ha sido muy utilizada en diversos trabajos para explicar los patrones de estructura genética con marcadores *outliers*, pues estas pueden actuar como barreras físicas en el océano, siendo la temperatura una de las más importantes (Cure *et al.*, 2017; Stanley *et al.*, 2018; Teske *et al.*, 2019; Oleksiak y Rajora, 2019).

Teske *et al.* (2019), a partir de la determinación de poblaciones del gobio *Psammogobius knysnaensis* en la costa Sudafricana, caracterizada por un fuerte gradiente termal, encontraron un mayor número de *clusters* para marcadores asociados a la temperatura (K=3-4), que con marcadores neutrales (K=1); por lo que destacaron la importancia de la temperatura sobre la estructura de las poblaciones y especiación causada por un gradiente ecológico, y no una barrera absoluta al flujo génico como tal. En el GC es posible observar un gradiente termal y diferencias en la variación, los cuales podrían estar afectando la divergencia de los grupos genéticos encontrados, por lo cual se recomienda, evaluar la posibilidad de una divergencia parapátrica contra una alopátrica para *E. puncticulatus*.

Todos los sitios mostraron una heterocigosidad observada menor a la esperada. Además de que análisis de diversidad genética, nos muestran que los valores de heterocigosidad esperada mantienen una correlación negativa con la latitud; y se encontró un aumento de la diversidad nucleotídica en los sitios intermedios (Loreto y Mulegé). Esto nos podría sugerir un aumento de mezcla en la parte central, lo cual coincide con el flujo génico asimétrico encontrado donde la mayor parte del material génico se acumula en los sitios intermedios. O de ser el caso si los marcadores *outliers* poseen un verdadero potencial adaptativo, el cambio podría estar asociado a una respuesta a la heterogeneidad entre regiones, pues ya otros trabajos han asociado la región central a una zona de transición entre bioregiones que a su vez se encuentran moldeadas por la heterogeneidad ambiental (Aceves-Medina *et al.*, 2004; Galland, 2013; Ulate *et al.*, 2016; El Khattabi-Salazar, 2019).

Esta heterogeneidad en el paisaje puede ser descrita por variables físicas. En la parte sur del GC, la influencia del Pacífico, la atmosfera, y eventos de surgencia dan origen al Agua Superficial Tropical (AST) ($S < 35$, $T \geq 18$ °C). Aproximadamente entre los 25-26° de latitud, se encuentra con el Agua Superficial del Golfo (ASG) ($S \geq 35$, $T > 12$), donde la salinidad aumenta, resultado de las altas tasas de evaporación (Lavín y Marinone, 2003). Soto-Mardones *et al.* (1999) dividió el GC en regiones: norte, islas, central, y sur, y encontró diferencias espaciales en la climatología de la temperatura superficial marina (TSM). Donde la temperatura disminuye hacia el interior del GC, con una variación anual, teniendo un mínimo de temperatura en la región de las islas a lo largo del año, y un gradiente del promedio de la TSM de $\sim 3^{\circ}\text{C}$ de norte a sur durante los meses de invierno, las bajas temperaturas en la región norte se le atribuyen a que esta es una zona somera y la fuerte influencia de las mareas. Mientras que en verano las diferencias de la TSM promedio mensual era menor entre regiones. También encontraron diferencias en la variación de la TSM, la cual aumenta hacia el interior del GC. Estas regiones (a excepción de la región norte, pues de esta región no se obtuvieron muestras) coinciden con los tres grupos genéticos encontrados: Bahía de los Ángeles; Mulegé y Loreto; La Paz y Los Cabos. Sugiriendo que la temperatura podría estar jugando un papel importante en la estructura espacial de las poblaciones.

Trabajos que buscan caracterizar el GC en bioregiones, basados en diversidad y abundancia de peces, también coinciden con los grupos genéticos con potencial adaptativo encontrados. La bioregionalización propuesta por Walker (1960) nos marca que la población norte se encuentra dentro de esta región del Golfo, hasta el sur de la región de las islas; mientras que el resto dentro de la parte central. Para este trabajo, Bahía de los Ángeles fue el sitio más divergente del resto. Aceves-Medina (2004), al estudiar la composición larvaria de la ictiofauna en el GC, indica que estas varían con la temporadas, donde en invierno y verano, el GC se divide en dos regiones, norte y sur, separando nuevamente a Bahía de los Ángeles del resto de los sitios. Aunque también encontró que durante los periodos de transición, verano y otoño, el golfo se divide en tres regiones: norte, centro, y sur, donde la región centro actúa como una zona de transición, y se observa una mezcla de especies de ambas regiones. Nuevamente, estas regiones coinciden con la encontradas en este trabajo.

Aceves-Medina *et al.*, (2004), asocia esta división con a la dinámica de las masas de agua y la circulación general del GC.

El trabajo de El Khattabi-Salazar (2019), quien propone una bioregionalización basada en la diversidad de peces crípticos, divide al GC en tres regiones, las cuales coinciden con los grupos encontrados en este trabajo, es de particular interés, pues estudia la composición de especies criptobentónicas con historias de vida similares. Encuentra diferencias en diversidad y riqueza taxonómicas en un gradiente latitudinal, con un incremento en los valores de norte a sur, mientras que índices de diversidad tradicionales indicaron que la mayor riqueza y diversidad se encontraba en la región centro; además de que diferencias en la composición de las especies por zonas se correlacionan a la variabilidad ambiental. Donde las especies encontradas en la parte norte se asocian a las bajas temperaturas y alta concentración de clorofila, mientras que el sur presenta un patrón inverso, finalmente el centro lo describe como una zona de transición.

9. CONCLUSIONES

A lo largo del GC la estructura poblacional neutral de *E. punctiluatatus*, difiere de la encontrada con marcadores *outliers*. Los marcadores neutrales detectaron dos poblaciones y una zona de mezcla, mientras que los marcadores *outliers* sugieren la existencia de tres grupos con potencial adaptativo. Para ambas estructuras se encontró un aislamiento por distancia a manera de parches. Lo que sugiere una diferenciación no clinal e isotrópica, sino que se sugiere la influencia de otros procesos y mecanismos de estructuración poblacional.

Para el caso de los marcadores neutrales, se observó el quiebre genético entre las dos poblaciones ocurre entre los sitios de Loreto y Mulegé. Esta divergencia podría estar siendo explicada por la circulación oceanográfica general del GC. Ya que en esta región se observan débiles corrientes de retorno, lo que favorece el autoreclutamiento. Además cerca de los sitios de Mulegé se encuentra uno de los límites inferiores para el cambio de dirección en la corriente de retorno.

En la zona de quiebre en el Golfo Central se observa una introgresión genética, que podría estar hablándonos de una zona de mezcla, la cual es consistente con una alta diversidad nucleotídica y una disminución de heterocigotos. Aunque también existe la posibilidad de que esta estructura se deba a una introgresión por contacto secundario, posiblemente asociado a un evento de vicarianza en el pasado, producido por la apertura de un canal marino en la península de Baja California, y que condiciones contemporáneas han contribuido al mantenimiento de este patrón. Ambos eventos no son excluyentes el uno del otro, mas se desconoce el grado con el que puedan estar afectando el patrón estructural de *E. puncticulatus*.

Los sitios encontrados en Bahía de los Ángeles, fueron los que presentaron mayores diferencias contra el resto, más no lo suficiente para ser considerado significativamente distinto de los sitios en Mulegé. Se sugiere que esta diferenciación podría estar dada por el alto dinamismo de la circulación oceanográfica en esta zona, que se ha reportado podría propiciar su potencial de autoreclutamiento; así como su lejanía al resto de las localidades lo cual podría favorecer la deriva génica.

La estructura encontrada con marcadores *outliers*, revelaron la existencia de tres poblaciones: 1) Bahía de los Ángeles, 2) Mulegé, 3) Loreto, La Paz, y Los Cabos. Esta diferenciación de grupos, es consistente con propuestas de bioregiones de diferentes autores para el GC basadas en índices de diversidad y abundancia de peces, la cuales se sugiere que pueden estar asociadas características ambientales.

Al igual que con los marcadores neutrales, se observa una introgresión secundaria entre los potenciales ecotipos, la cual pareciera ocurrir de sur a norte, lo cual se podría explicar porque individuos de regiones norteñas se han tenido que adaptar a las condiciones locales, las cuales son más extremas y heterogéneas, comparadas con el sur, lo cual limitaría la dispersión de individuos del sur al norte.

Por lo anterior se sugiere que la estructura genética de *E. puncticulatus* en el GC es el resultado de variables espaciales que limitan su dispersión y establecimiento entre sitios, como la circulación oceanográfica, el aislamiento geográfico, y la heterogeneidad espacial. Aunque no se descarta la posibilidad de que eventos históricos hayan limitado su dispersión en el pasado. El grado con el que cada una de estas variables afecta las poblaciones de esta especie aun permanece incierto, por lo cual se recomienda continuar con estudios que puedan elucidar esta relación.

10. LITERATURA CITADA

Aceves-Medina, G., S. P. A. Jimenez-Rosenberg, A. Hinojos-Medina, R. Funes-Rodríguez, R. J. Saldierna, D. Lluch-Belda, P. E. Smith, and W. Watson. 2003. Fish larvae from the Gulf of California. *Sci. Mar.* 67:1–11.

Aceves-Medina, G., S. P. A. Jimenez-Rosenberg, A. Hinojos-Medina, R. Funes-Rodríguez, R. J. Saldierna-Martínez, and P. E. Smith. 2004. Fish larvae assemblages in the Gulf of California. *J. Fish Biol.* 65:832–847.

Álvarez, L. G., F. Suárez-Vidal, R. Mendoza-Borunda, A. González-Escobar. 2009. Bathymetry and active geological structures in the Upper Gulf of California. *Bol. Soc. Geol. Mex.* 61(1): 129-141.

Alvarez-Borrego, S. 2010. Chapter 2: Physical, Chemical, and Biological Oceanography of the Gulf of California. En: Brusca, R. C. *The Gulf of California: Biodiversity and conservation.* University of Arizona Press. Tucson, Arizona. pp. 24-28.

Álvarez-Valín, F. 2009. Evolución molecular: neutralismo y seleccionismo. En: Altuna, C., M. Ubilla (eds.). *El prisma de la Evolución.* DIRAC. Primera reedición. Montevideo, Uruguay. pp. 243-263.

Andrews, K. R., J. M. Good, M. R. Miller, G. Luikart, P.A. Hohenlohe. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Rev. Genet.* 17: 81-92.

Baird, N., P. D. Etter, T. S. Atwood, M. C. Currey, A. L. Shiver, Z.A. Lewis, E. U. Selker, W. A. Cresko, E. A. Johnson. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLOS ONE.* 3(10): 1-7.

Beaumont, M. A., R. A. Nichols. 1996. Evaluating loci for use in the genetic analysis of population structure. *P. Roy. Soc. Lond. B. Bio.* 263(1377):1619-1626.

Behjati, S., P. S. Tarpey. 2013. What is next generation sequencing? *Arch. Dis. Child. Educ. Pract.* 98:236–238.

Bernardi, G. 2000. Barriers to gene flow in *Embiotoca jacksoni*, a marine fish lacking a pelagic larval stage. *Evolution* 54:226–237. En: Riginos, C. 2005. Cryptic Vicariance in Gulf of California Fishes Parallels Vicariant Patterns Found in Baja California Mammals and Reptiles. *Evolution*. 59(12):2678-2690.

Bernardi, G. 2014. Baja California disjunctions and phylogeographic patterns in sympatric California blennies. *Front. Ecol. Evol.* 2(53): 1-9.

Bernardi, G., L. Findley, A. Rocha-Oliveras. 2003. Vicariance and Dispersal across Baja California in Disjunct Marine Fish Populations. *Evolution*. 57(7):1599-1609.

Bernardi, G., N. L. Crane. 2000. Genetic evidence for limited dispersal in the coastal California killifish, *Fundulus parvipinnis*. *J. Exp. Mar. Biol. Ecol.* 255:187–199. En: Riginos, C. 2005. Cryptic Vicariance in Gulf of California Fishes Parallels Vicariant Patterns Found in Baja California Mammals and Reptiles. *Evolution*. 59(12):2678-2690.

Bertl, J., H. Ringbauer, M. G. B. Blum. 2018. Can secondary contact following range expansion be distinguished from barriers to gene flow? *PeerJ*. 6: 1-18.

Bohonak, A. J. 2002. IBD (Isolation by Distance): A Program for Analyses of Isolation by Distance. *J. Hered.* 93(2): 153-154.

Burgess, W. E., H. R. Exelrod, R. E. Hunziker. 2000. Dr Burgess's atlas of marine aquarium fishes. T.F.H. Publications. Tercera edición. Neptune, New Jersey. 736p. En: Trujillo-García, M. 2018. Fecundidad por desove y comportamiento reproductivo en el gobio de cabeza roja *Elacatinus puncticulatus*. Tesis (Maestría en Ciencias). La Paz, Baja California Sur, Centro de Investigaciones Biológicas del Noroeste, S.C. 85 p.

Cabrero, J., J. P. Camacho. 2002. Capítulo 6: Fundamentos de genética de poblaciones. En: Soler-Cruz, M. (ed.) EVOLUCIÓN. La base de la Biología. Proyecto Sur de Ediciones, S.L. España. pp. 83-126.

Carreño, A. L., J. T. Smith. 2007. Stratigraphy and correlation for the ancient Gulf of California and Baja California Peninsula, Mexico. *Bull. Am. paleontol.* (371):1-146.

- Catchen, J. M., P. A. Hohenlohe, S. Bassham, A. Amores, W. A. Cresko. 2013. Stacks: An analysis tool set for population genomics. *Mol. Ecol.* 22:3124-3140.
- Chhatre, V. E., K. J. Emerson. 2017. StrAuto: automation and parallelization of STRUCTURE analysis. *BMC Bioinformatics.* 18(192): 1-5.
- Cowen, R. K., S. Sponaugle. 2009. Larval dispersal and marine population connectivity. *Annu. Rev. Mar. Sci.* 1: 443-466.
- Cure, K., L. Thomas, J. P. A. Hobbs, D. V. Fairclough, W. J. Kennington. 2017. Genomic signatures of local adaptation reveal source-sink dynamics in a high gene flow fish species. *Sci. Rep.* 7: 1-10.
- Danecek, P., A. Auton, G. Abecasis, C. A. Albers, E. Banks, M. A. DePristo, R. E. Handsaker, G. Lunter, G. T. Marth, S. T. Sherry, G. McVean, R. Durbin, 1000 Genomes Project Analysis Group. 2011. The variant call format and VCFtools. *Bioinformatics.* 27(15):2156-2158.
- Davey, J. W., M. L. Blaxter. 2010. RADseq: next-generation population genetics. *Brief. Funct. Genom.* 9(5): 416-423.
- DeFaveri J., P. R. Jonsson, J. Merila. 2013. Heterogeneous genomic differentiation in marine threespine sticklebacks: adaptation along an environmental gradient. *Evolution.* 67:2530–2546.
- do Prado, F. D., M. Vera, M. Hermida, C. Bouza, B. G. Pardo, R. Vilas, A. Blanco, C. Fernández, F. Maroso, G. E. Maes, C. Turan, F. A. M. Volckaert, J. B. Taggart, A. Carr, R. Ogden, E. E. Nielsen, The Aquatrace Consortium, P. Martínez. 2018. Parallel evolution and adaptation to environmental factors in a marine flatfish: implications for fisheries and aquaculture management of the turbot (*Scophthalmus maximus*). *Evol. Appl.* 1-20.
- Earl, D. A., B. M. von Holdt. 2012. STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Conservation Genet. Resour.* 4: 359-361.

Eckert, C. G., K. E. Samis, C. Lougheed. 2008. Genetic variation across species' geographical ranges: the central-marginal hypothesis and beyond. *Mol Ecol.* 17: 1170-1188.

Eguiarte, L. E., J. A. Aguirre-Liguori, L. Jardón-Barbolla, E. Aguirre-Planter, V. Souza. 2013. Genómica de poblaciones: Nada en evolución va a tener sentidos si no es a la luz de la genómica, y nada en genómica tendrá sentido si no es a la luz de la evolución. *TIP Rev. Esp. Cienc. Quím. Biol.* 16(1): 42-56.

El Khattabi-Salazar, S. 2019. Estructura de la comunidad y preferencia de hábitat de peces crípticos en la costa oeste del Golfo de California. Tesis (Maestría en Ciencias). La Paz, Baja California Sur, Centro de Investigaciones Biológicas del Noroeste, S.C. 90 p.

Evanno, G., S. Regnaut, J. Goudet. 2005. Detecting the number of clusters of individuals using the software STRUCTURE: A simulation study. *Mol. Ecol.* 14(8):2611-2620.

Excoffier, L., H. E. L. Lischer. 2010. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10:564-567.

Foll, M., O. Gaggiotti. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics.* 180:977-993.

Galland, G. R. 2013. The conservation and ecology of cryptobenthic fishes on rocky reefs in the Gulf of California, Mexico. Tesis (Doctorado en Biología Marina). San Diego, California. Universidad de California. 190 p.

Germain, N., H. J. Hartmann, F. J. Fernández-Rivera Melo, H. Reyes-Bonilla. 2015. Ornamental reef fish fisheries: new indicators of sustainability and human development at a coastal community level. *Ocean Coast Manag.* 104:136–149.

Ginsburg, I. 1938. Two new gobioid fishes of the genus *Gobiosoma* from Lower California. *Stanford Ichthyol. Bull.* 1(2): 57-59.

Gompert, Z., A. Buerkle. 2016. What , if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. *Evol. Appl.* 9: 909-923.

Goudet, J., & Jombart, T. (2015). hierfstat: Estimation and tests of hierarchical F-statistics

Guo, B., J. DeFaveri, G. Sotelo, A. Nair, J. Merila. 2015. Population genomic evidence for adaptive differentiation in Baltic Sea three-spined sticklebacks. *BMC Biol.* 13:19.

Hedgecock, D., P. H. Barber, S. Edmands. 2007. Genetic approaches to measuring connectivity. *Oceanography.* 20(3):70-79.

Hedgecock, D., P. H., Barber, y S., Edmands. 2007. Genetic approaches to measuring connectivity. *Oceanography.* 20(3): 70-79.

Helenes, J., A. L. Carreño. 1999. Neogene sedimentary evolution of Baja California in relation to regional tectonics. *J. S. Am. Earth. Sci.* 12: 589-605. En: Ledesma-Vázquez, J., M. E. Johnson, O. Gonzalez-Yajimovich, E. Santamaría-de-Angel. 2009. Gulf of California Geography, Geological Origins, Oceanography, and Sedimentation Patterns. En: Johnson, M. E., J. Ledesma-Vázquez (eds.). *Atlas of Coastal Ecosystems in the Western Gulf of California: Tracking Limestone Deposits on the Margin of a Young Sea.* University of Arizona Press. Tucson, Arizona. pp. 1-10.

Hellberg, M. E., R. S. Burton, J. E. Neigel, S. R. Palumbi. 2002. Genetic assesment of connectivity among marine populations. *Bull. Mar. Sci.* 70(1): 273-290.

Huang, D., G. Bernardi. 2001. Disjunct Sea of Cortez-Pacific Ocean *Gillichthys mirabilis* populations and the evolutionary origin of their Sea of Cortez endemic relative, *Gillichthys seta*. *Mar. Biol.* 138:421–428. En: Riginos, C. 2005. Cryptic Vicariance in Gulf of California Fishes Parallels Vicariant Patterns Found in Baja California Mammals and Reptiles. *Evolution.* 59(12):2678-2690.

Hurtado, L. A., M. Mateos, C. A. Santamaria. 2010. Phylogeography of supralittoral rocky intertidal *Ligia* isopods in the Pacific region from Central California to Central Mexico. *PLOS ONE.* 5(7):1-13.

Instituto Nacional de Estadística, Geografía e Informática (INEGI), L. Ita-Rubio de, I. Escamilla-Herrera, C. García-de León, M. C. Soto-Núñez. 1990. División Política Estatal 1:4000000, I.1.2. Atlas Nacional de México. Vol. I. Escala 1:4000000. Instituto de Geografía, UNAM. México.

Jombart T. Devillard S., Balloux F. 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genetics*. 11(94): 1-15.

Keenan, K., P. McGinnity, T. F. Cross, W. W. Crozier, P. A. Prodöhl. (2013). *diveRsity*: An R package for the estimation of population genetics parameters and their associated errors. *Methods Ecol. Evol.* 4(8): 782-788

Kirk, H., J. R. Freeland. 2011. Applications and implications of neutral versus non-neutral markers in molecular ecology. *Int. J. Mol. Sci.* 12: 3966-3988.

Kopelman, N. M., J. Mayzel, M. Jakobsson, N. A. Rosenberg, I. Mayrose. 2015. Clumpak: a program for identifying clustering modes and packaging population structure inferences across K. *Mol. Ecol. Resour.* 15(5): 1179-1191.

Lavín, M. F. y S. G. Marinone. 2003. An Overview of the Physical Oceanography of the Gulf of California. En: Velasco-Fuentes, O. U., J. Sheinbaum, J. Ochoa (eds.). *Nonlinear Processes in Geophysical Fluid Dynamics*. Springer. Dordrecht. pp. 173-204.

Ledesma-Vázquez, J. 2002. A gap in the pliocene invasion of seawater to the Gulf of California. *Rev. Mex. Cienc. Geol.* 19(3): 145-151. En: Ledesma-Vázquez, J., M. E. Johnson, O. Gonzalez-Yajimovich, E. Santamaría-de-Angel. 2009. *Gulf of California Geography, Geological Origins, Oceanography, and Sedimentation Patterns*. En: Johnson, M. E., J. Ledesma-Vázquez (eds.). *Atlas of Coastal Ecosystems in the Western Gulf of California: Tracking Limestone Deposits on the Margin of a Young Sea*. University of Arizona Press. Tucson, Arizona. pp. 1-10.

Ledesma-Vázquez, J. 2002. A gap in the pliocene invasion of seawater to the Gulf of California. *Rev. Mex. Cienc. Geol.* 19(3): 145-151. En: Hurtado, L. A., M. Mateos, C. A. Santamaria. 2010. Phylogeography of supralittoral rocky intertidal *Ligia* isopods in the Pacific region from Central California to Central Mexico. *PLOS ONE*. 5(7):1-13.

Ledesma-Vázquez, J., M. E. Johnson, O. Gonzalez-Yajimovich, E. Santamaría-de-Angel. 2009. Gulf of California Geography, Geological Origins, Oceanography, and Sedimentation Patterns. En: Johnson, M. E., J. Ledesma-Vázquez (eds.). Atlas of Coastal Ecosystems in the Western Gulf of California: Tracking Limestone Deposits on the Margin of a Young Sea. University of Arizona Press. Tucson, Arizona. pp. 1-10.

Limborg M. T., S. J. Helyar, M. De Bruyn, M. I. Taylor, E. E. Nielsen, R. Ogden, G. R. Carvalho, FPT Consortium, D. Bekkevold. 2012. Environmental selection on transcriptome-derived SNPs in a high gene flow marine fish, the Atlantic herring (*Clupea harengus*). *Mol. Ecol.* 21: 3686-3703.

Lin, H-C., C. Sánchez-Ortiz, P. A. Hastings. 2009. Colour variation is incongruent with mitochondrial lineages: Cryptic speciation and subsequent diversification in a Gulf of California reef fish (Teleostei: Blennioidei) *Mol. Ecol.* 18: 2476-2488.

Lluch-Cota, S. E., E. A. Aragón-Noriega, F. Arreguín-Sánchez, D. Aurióles-Gamboa, J. J. Bautista-Romero, R. C. Brusca, R. Cervantes-Duarte, R. Cortés-Altamirano, P. Del-Monte-Luna, A. Esquivel-Herrera, G. Fernández, M. E. Hendrickx, S. Hernández-Vázquez, H. Herrera-Cervantes, M. Kahru, M. Lavín, D. Lluch-Belda, D. B. Lluch-Cota, J. López-Martínez, S. G. Marinone, M. O. Nevárez-Martínez, S. Ortega-García, E. Palacios-Castro, A. Parés-Sierra, G. Ponce-Díaz, M. Ramírez-Rodríguez, C. A. Salinas-Zavala, R. A. Schwartzlose, A. P. Sierra-Beltrán. 2007. The Gulf of California: Review of ecosystem status and sustainability challenges. *Prog. Oceanogr.* 73(1): 1-26.

Locarnini, R. A., A. V. Mishonov, O. K. Baranova, T. P. Boyer, M. M. Zweng, H. E. Garcia, J. R. Reagan, D. Seidov, K. Weathers, C. R. Paver, I. Smolyar. 2018. *World Ocean Atlas 2018. Volume 1: Temperature*. En: Mishonov, A. (ed.). NOAA Atlas NESDIS 81. Silver Spring, Maryland. pp 1-52-

López-de Heredia, U. 2016. Las técnicas de secuenciación masiva en el estudio de la diversidad biológica. *Munibe, Cienc. nat.* 64:7-31.

Luikart, G., P. R. England, D. Tallmon, S. Jordan, P. Taberlet. 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Rev. Genet.* 4:981-994.

Mac Loughlin-Aleman, C. 2019. Paisaje marino genómico del gobio *Lythrypnus dalli* (Gilbert 1890) en el Golfo de California. Tesis (Maestría en Ciencias). La Paz, Baja California. Centro de Investigaciones Biológicas del Noroeste, S.C. 105 p.

Mares-Mayagoitia. 2019. Genómica poblacional del abulón amarillo (*Haliotis corrugata*) en la costa occidental de Baja California. Tesis (Maestría en Ciencias). La Paz, Baja California. Centro de Investigaciones Biológicas del Noroeste, S.C. 84 p.

Marinone, S. G. 2003. A three-dimensional model of the mean and seasonal circulation of the Gulf of California. *J. Geophys. Res.* 108(10):3325.

Marinone, S. G. 2012. Seasonal surface connectivity in the Gulf of California. *Estuar. Coast. Shelf S.* 100:133-141

Molina-Cruz, A. 1984. Evolución oceanográfica de la boca del Golfo de California. *An. Inst. Cienc. del Mar y Limnol. Univ. Nal. Autón. México.* 3(2): 95–120.

Oleksiak, M. F. 2018. Adaptation without boundaries: Population genomics in marine systems. En Rajora, O. P. (ed.). *Population genomics: Concepts, approaches and applications.* Springer. pp. 587–612.

Oleksiak, M. F., O. P. Rajora. 2019. Marine Population Genomics: Challenges and Opportunities. En: Oleksiak, M., O. P. Rajora (eds.) *Population genomics: Marine organisms.* Springer. pp. 3-35.

Paris, J. R., J. R. Stevens, J. M. Catchen. 2017. Lost in parameter space: A road map for stacks. *Methods Ecol. Evol.* 8(10):1360-1373

Pedrazzani, A. S., N. K. Pham, J. Lin, A. O. Neto. 2014. Reproductive behavior, embryonic and early larval development of the red head goby, *Elacatinus puncticulatus*. *Anim. Reprod. Sci.* 145(1–2):69–74.

Peterson, B. K., J.N. Weber, E. H. Kay, H. S. Fisher, H. E. Hoekstra. 2012. Double digest RADseq: An inexpensive method for de novo SNP discovery and genotyping in model and non-model species. PLOS ONE. 7(5):1-11.

Pritchard J. K., M. Stephens, P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics. 155(2):945-959.

Quimbayo, J. P., Zapata, F. A. 2018. Cleaning interactions by gobies on a tropical eastern Pacific coral reef. J. Fish Biol. 92: 1110-1125.

R Core team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.

Rice, W. R. 1989. Analyzing tables of statistical tests. Evolution. 43(1): 223-2255.

Riginos C., E. D. Crandall, L. Liggins, P. Bongaerts, E. A. Trembl. 2016. Navigating the currents of seascape genomics: how spatial analyses can augment population genomic studies. Curr. Zool.62(6): 1–21.

Riginos, C. 2005. Cryptic Vicariance in Gulf of California Fishes Parallels Vicariant Patterns Found in Baja California Mammals and Reptiles. Evolution. 59(12):2678-2690.

Riginos, C., L. Liggins. 2013. Seascape genetics: Populations, individuals, and genes. Geogr. compass. 7: 197-216.

Riginos, C., M. W. Nachman. 2001. Population subdivision in marine environments: The contributions of biogeography, geographical distance and discontinuous habitat to genetic differentiation in a blennioid fish, *Axoclinus nigricaudus*. Mol. Ecol. 10(6):1439-1453.

Robertson, D. R., G. R. Allen. 2015. Peces costeros del Pacifico Oriental Tropical: sistema de información en línea. Versión 2.0 Instituto Smithsonian de Investigaciones Tropicales, Balboa, Republica de Panamá. <https://biogeodb.stri.si.edu/sftep/es/pages>

Saenz-Agudelo, P., J. D. DiBattista, M. J. Piatek, M. R. Gaither, H. B. Harrison, G. B. Nanninga, M. L. Berumen. 2015. Seascape genetics along environmental gradients in the

Arabian Peninsula: Insights from ddRAD sequencing of anemonefishes. *Mol. Ecol.* 24:6241-6255.

Sandoval-Castillo, J., N. A. Robinson, A. M. Hart, L. W. S. Strain, L. B. Beheregaray. 2018. Seascape genomics reveals adaptive divergence in a connected and commercially important mollusc, the greenlip abalone (*Haliotis laevis*), along a longitudinal environmental gradient. *Mol. Ecol.* 27: 1603-1620.

Sandoval-Huerta, E. R., R. G. Beltrán-López, C. R. Pedraza-Marrón, M. A. Paz-Velásquez, A. Angulo, D. R. Robertson, E. Espinoza, O. Domínguez-Domínguez. 2018. The evolutionary history of the goby *Elacatinus puncticulatus* in the tropical eastern pacific: effects of habitat discontinuities and local environmental variability. *Mol. Phylogenet. Evol.* 130: 269-285.

Shei, M., M. Mies, I. Olivotto. 2017. Other Demersal spawners and mouthbrooders. En: Calado, R., I. Olivotto, M. Planas-Oliver, J. Holt. *Marine Ornamental Species Aquaculture*. Editorial Wiley Blackwell. Primera edición. Sussex occidental, Reino Unido. pp 223-250.

Silva, R., M. L. Martínez, P. Moreno-Casasola, E. Mendoza, J. López-Portillo, D. Lithgow, G. Vázquez, R. E. Martínez-Martínez, R. Monroy-Ibarra, J. I. Cáceres-Pulg, A. Ramírez-Hernández, M. Boy-Tamborell. 2017. Aspectos generales de la zona costera. UNAM; INECOL. 9-12.

Soto-Mardones, L., S. G. Marinone, A. Parés-Sierra. 1999. Variabilidad espaciotemporal de la temperatura superficial del mar en el Golfo de California. *Ciencias Marinas.* 25(1): 1-30.

Stanley, R. R. E., C. DiBacco, B. Lowen, R. G. Beiko, M. Van Wyngaarden, N. W. Jeffery, P. Bentzen, L. Benestan, L. Bernatchez, C. Johnson, P. V. R. Snelgrove, B. F. Wringe, I. R. Bradbury. 2018. A climate-associated multi-species cryptic genetic cline in the northwest Atlantic. *Sci. Adv.* 4: 1-7.

Terry, A., G. Bucciarelli, G. Bernardi. 2000. Restricted gene flow and incipient speciation in disjunct Pacific Ocean and Sea of Cortez populations of a reef fish species, *Girella nigricans*. *Evolution* 54:652–659. En: Riginos, C. 2005. Cryptic Vicariance in Gulf of

California Fishes Parallels Vicariant Patterns Found in Baja California Mammals and Reptiles. *Evolution*. 59(12):2678-2690.

Teske, P. R., J. Sandoval-Castillo, T. R. Golla, A. Emami-Khoyi, M. Tine, S. von der Heyden, L. B. Beheregaray. 2019. Thermal selection as a driver of marine ecological speciation. *Proc. R. Soc. B*. 286: 1-7.

Tranah, G. J., L. G. Allen. 1999. Morphologic and genetic variation among six populations of the spotted sand bass, *Paralabrax maculatofasciatus*, from southern California to the upper Sea of Cortez. *Bull. South. Calif. Acad. Sci.* 98:103–118. En: Riginos, C. 2005. Cryptic Vicariance in Gulf of California Fishes Parallels Vicariant Patterns Found in Baja California Mammals and Reptiles. *Evolution*. 59(12):2678-2690.

Trujillo-García, M. 2018. Fecundidad por desove y comportamiento reproductivo en el gobio de cabeza roja *Elacatinus puncticulatus*. Tesis (Maestría en Ciencias). La Paz, Baja California Sur, Centro de Investigaciones Biológicas del Noroeste, S.C. 85 p.

Ulate, K., C. Sánchez, A. Sánchez-Rodríguez, D. Alonso, O. Aburto-Oropeza & L. Huato-Soberanis (2016). Latitudinal regionalization of epibenthic macroinvertebrate communities on rocky reefs in the Gulf of California. *Mar. Biol. Res.* 12(4): 1-13.

Upton, D. E., R. W. Murphy. 1997. Phylogeny of the side- blotched lizards (Phrynosomatidae: *Uta*) based on mtDNA sequences: support for a midpeninsular seaway in Baja California. *Mol. Phy. Evol.* 8: 104-113.

van Strien, M. J., R. Holderegger, H. J. Van Heck. 2015. Isolation-by-distance in landscapes: considerations for landscape genetics. *Heredity*. 114: 27-37.

Van Wyngaarden, M., P. V. R. Snelgrove, C. DiBacco, L. C. Hamilton, N. Rodríguez-Ezpeleta, N. W. Jeffery, R. R. E. Stanley, I. R. Bradbury. 2016. Identifying patterns of dispersal, connectivity and selection in the sea scallop, *Placopecten magellanicus*, using RADseq-derived SNPs. *Evol. Appl.* 10: 102-117.

Walker, B. W. 1960. The Distribution and Affinities of the Marine Fish Fauna of the Gulf of California. *Syst. Zool.* 9(3):123-133.

Zweng, M. M., J. R. Reagan, D. Seidov, T. P. Boyer, R. A. Locarnini, H. E. Garcia, A. V. Mishonov, O. K. Baranova, K. Weathers, C. R. Paver, I. Smolyar. 2018. World Ocean Atlas 2018, Volume 2: Salinity. Mishonov, A. (ed.). NOAA Atlas NESDIS 82 Silver Spring, Maryland. pp 1-50.