

UCH-FC

DOC-EBE

S 454

C.1



FACULTAD DE CIENCIAS  
UNIVERSIDAD DE CHILE

**Adaptación local y dinámicas espacio-temporales en el tunicado *Pyura chilensis* Molina 1782 (Urochordata: Ascidiacea) a lo largo del sistema de Corriente de Humboldt.**

Tesis entregada a la Universidad de Chile en cumplimiento parcial de los requisitos para optar al grado de Doctor en Ciencias con Mención en

Ecología y Biología Evolutiva

Facultad de Ciencias

Por

**Nicolás Ignacio Segovia Cortés**

Enero, 2018

Directores de Tesis: Dr. Elie Poulin,  
Dra. Pilar A. Haye (UCN-Coquimbo)

**FACULTAD DE CIENCIAS**  
**UNIVERSIDAD DE CHILE**  
**INFORME DE APROBACION**  
**TESIS DE DOCTORADO**

Se informa a la Escuela de Postgrado de la Facultad de Ciencias que la Tesis de Doctorado presentada por el candidato.

**Nicolás Ignacio Segovia Cortés**

Ha sido aprobada por la comisión de Evaluación de la tesis como requisito para optar al grado de Doctor en Ciencias Biológicas mención Ecología y Biología Evolutiva, en el examen de Defensa Privada de Tesis rendido el día 30 de noviembre de 2017.

Directores de Tesis:

Dr. Elie Poulin

Dr. Pilar A. Haye Molina



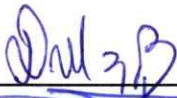


Comisión de Evaluación de la Tesis

Dr. David Veliz

Dr. Marco Mendez

Dra. Caren Vega

Dr. Cristian Gallardo-Escárate

  
\_\_\_\_\_  
  
\_\_\_\_\_  
  
\_\_\_\_\_  
  
\_\_\_\_\_  
  
\_\_\_\_\_  
\_\_\_\_\_





*“if it could be demonstrated that any complex organ existed which could not possibly have been formed by numerous, successive, slight modifications, my theory would absolutely break down.”*

*Charles Darwin*

## Resumen Biográfico



Soy un biólogo marino atípico. Casi no me meto al agua, casi no como mariscos, y casi no me gusta el mar. Mis intereses siempre han ido por el lado de intentar entender la diversidad biológica, la biología y ecología de los organismos marinos, pero usando herramientas moleculares y/o bioinformáticas. Me formé como biólogo marino en la Universidad Católica del Norte, en Coquimbo, ciudad a la cual llegué el año 1997 a la edad de 10 años y que adopté como mi hogar de forma casi instantánea. Antes de eso, viví en un pequeño pueblo llamado Potrerillos, el cual fue cerrado de forma definitiva el año 1998 por exceso de arsénico en el ambiente lo que hacía imposible su habitabilidad en el largo plazo, ¿a quién no le ha pasado?

Luego de terminar mi pregrado, tuve la suerte de seguir trabajando en el laboratorio que me cobijó, siendo asistente de investigación en múltiples proyectos orientados a la ciencia fundamental y al uso de herramientas moleculares para intentar responder preguntas de tipo ecológicas, evolutivas y/o filogenéticas en organismos marinos (desde algas hasta crustáceos importancia comercial). Todo ello dirigido por la incombustible Dra. Pilar Haye.

Esta experiencia me permite llegar el año 2012 al programa de Doctorado en Ciencias mención ecología y biología evolutiva de la Universidad de Chile bajo la tutela del Dr. Elie Poulin. Ahí, sufrí uno de los peores reveses que puede sufrir un estudiante de doctorado; que la tesis se desplome por completo. A pesar de esto, tengo la suerte de tener a dos directores de tesis todo terreno, quienes me apoyaron y permitieron que esto finalmente llegase a buen puerto.

And...here we are

## Agradecimientos

Aprovecho este espacio para agradecer a todos quienes de una u otra forma han sido parte fundamental, no solo para el desarrollo de esta tesis, si no para que la experiencia en el doctorado fuese una de las más reconfortantes a nivel tanto académico como personal. Parto por agradecer profundamente a mi directora de tesis, la Dra. Pilar Haye, con quién he tenido la suerte de trabajar durante casi 10 años. Llegué como un alumno con ganas de aprender todo lo que fuera posible, y ahora hemos logrado formar un equipo afiatado con ganas de seguir contribuyendo e investigando. Agradezco también a mi tutor, Dr. Elie Poulin, por la confianza depositada, por sus siempre iluminadoras reuniones y por el compromiso de siempre intentar mejorar y perfeccionar todo lo que llega a sus manos. Nunca olvidaré los «*Poulin's inflection points*».

Agradezco también al laboratorio de diversidad molecular (LADIMO), en especial a don Raúl Vera, un amigo de la vida, y una de las personas que más agradezco haber conocido en este viaje por la ciencia. Al Laboratorio de Ecología Molecular (TuLEM), uno de los lugares más gratos en donde me ha tocado estar, lleno de gente que quiere aprender, contribuir, y colaborar. No puedo dejar de mencionar a mis compañeros de «pocilga», un lugar donde las condiciones para poder hacer ciencia son de los más altos estándares internacionales. Ahí tuve la suerte de conocer a uno de mis grandes amigos, al Dr. Claudio González, un tremendo científico, una tremenda persona y un gran compañero.

Finalmente, en el aspecto académico, agradezco a mis compañeros de doctorado, Pamela Morales y Esteban Botero. Grandes amigos que quedarán para siempre. Vayan donde vayan, nuestros «paredones» en preparación para el examen de calificación serán unas las ~~peores~~ mejores experiencias durante el doctorado.

Agradezco también a mi familia, pilares fundamentales en cada etapa de mi vida, y a Natalie Ulloa, mi compañera de viaje, de vida y de locuras. Sin ella, este viaje habría sido mucho más pedregoso.

Esta tesis fue financiada por los proyectos FONDECYT 1140862 de la Dra, Pilar Haye, en colaboración con el proyecto FONDAP INCAR 1511027. Agradezco también a la beca de doctorado nacional CONICYT y al instituto de Ecología y Biodiversidad (IEB), quienes han sido fundamentales para el desarrollo tanto del doctorado como de la tesis.

## Índice

Índice.....	6
Resumen .....	8
Abstract.....	10
Introducción General .....	12
<b>Adaptación local</b> .....	<b>12</b>
<b>Aproximaciones para estudiar adaptación local</b> .....	<b>15</b>
<b>Sistema de la Corriente de Humboldt</b> .....	<b>20</b>
<b>Modelo y acercamiento al estudio</b> .....	<b>22</b>
CAPÍTULO I: Lineage divergence, local adaptation across a biogeographic break, and artificial transport shape the genetic structure in the ascidian <i>Pyura chilensis</i> .....	27
<b>Abstract</b> .....	<b>28</b>
<b>Introduction</b> .....	<b>29</b>
<b>Results</b> .....	<b>33</b>
SNP discovery and outlier detection.....	33
Population genetic diversity and population structure analyses.....	34
<b>Discussion</b> .....	<b>39</b>
<b>Methods</b> .....	<b>46</b>
Sample collection and DNA extraction.....	46
Genotyping by Sequencing and SNP discovery .....	47
Outlier detection.....	49
Population genetic diversity analysis and genetic structure .....	50
<b>References</b> .....	<b>52</b>
<b>Acknowledgements</b> .....	<b>60</b>
<b>Tables and Figures</b> .....	<b>62</b>
CAPÍTULO II: Interannual variation of COI haplotypes and haplogroups of the ascidian <i>Pyura chilensis</i> in the Southeast Pacific coast.....	74
Abstract.....	75
<b>INTRODUCTION</b> .....	<b>76</b>
<b>MATERIALS AND METHODS</b> .....	<b>80</b>
Data Analyses.....	81
<b>RESULTS</b> .....	<b>83</b>
<b>DISCUSSION</b> .....	<b>87</b>
Temporal genetic variation in <i>Pyura chilensis</i> .....	87
Dispersal potential and temporal genetic variation and mechanisms behind temporal genetic variation.....	89
<b>REFERENCES</b> .....	<b>93</b>
<b>Anexos</b> .....	<b>98</b>
Tables and Figures.....	98
CAPITULO III: Adaptación local y asociaciones genético ambientales en el pacífico sudeste en la ascidia <i>Pyura chilensis</i> .....	109

<b>Resumen .....</b>	<b>109</b>
<b>Introducción .....</b>	<b>110</b>
Adaptación local en el ambiente marino.....	110
Búsqueda de señales de selección .....	111
Sistema de Corriente de Humboldt.....	114
<b>Materiales y Métodos.....</b>	<b>117</b>
Toma de muestras y extracción de ADN .....	117
<i>Genotyping by Sequencing</i> (GBS) y genotipificación de SNPs.....	118
Búsqueda de loci candidatos a adaptación local .....	118
Datos ambientales.....	118
Aproximaciones basadas en análisis de diferenciación poblacional (ADP) .....	121
Estructura espacial y análisis de redundancia: Relación entre la estructura genética y variables ambientales. ....	123
Análisis de Asociaciones Ambientales (EAA) y análisis de estructura espacial de loci adaptativos.....	125
Ontología genética .....	129
<b>Resultados.....</b>	<b>130</b>
Detección de loci candidatos a selección natural.....	130
Análisis espacial de estructura genética .....	131
Análisis de asociaciones ambientales (EAA).....	133
Clinas multilocus de las frecuencias alélicas.....	135
Ontología genética .....	136
<b>Discusión .....</b>	<b>137</b>
Estructura genética adaptativa en <i>Pyura chilensis</i> .....	137
Aproximaciones basadas en análisis de diferenciación poblacional (ADP) versus las basadas en asociaciones genético-ambientales (EAA) .....	145
<b>Conclusión.....</b>	<b>147</b>
<b>Tablas y Figuras.....</b>	<b>149</b>
 Conclusión general.....	 167
Referencias.....	172

## Resumen

La adaptación local corresponde a aquellos procesos que surgen debido a que la selección natural en ambientes heterogéneos se contrapone a los efectos homogenizadores de procesos neutrales como el flujo genético. En hábitats marinos costeros, muchos animales cuentan con estadios larvales dispersivos que permiten la conectividad entre poblaciones locales, sin embargo, la conectividad no necesariamente está correlacionada con la capacidad de dispersión. Por otro lado, el transporte físico de individuos entre poblaciones locales no asegura flujo genético efectivo, ya que este requiere que los migrantes sobrevivan y se reproduzcan, lo que implica que la mortalidad selectiva de migrantes por efectos de presiones ambientales podría reducir el flujo genético entre poblaciones.

Pocos estudios filogeográficos han sido abordados usando marcadores moleculares que permitan detectar los efectos de la selección natural y contrastarlos con los efectos de fuerzas evolutivas neutrales. Para una mayor comprensión de los aportes de distintas fuerzas evolutivas en la estructuración filogeográfica, se utilizó como modelo de estudio el tunicado *Pyura chilensis*, un organismo sésil endémico del altamente heterogéneo sistema de corriente de Humboldt (HCS), con baja capacidad de dispersión larval, pero una alta capacidad de dispersión por efecto del transporte antropogénico en cubierta de barcos. Estudios previos, mostraron estructura genética moderada en la especie, y ausencia de diferenciación genética asociada a un quiebre biogeográfico ampliamente reportado a los 30°S, que si ha sido reportado para otros taxa marinos con baja capacidad de dispersión.

En el primer capítulo, se comparó la diversidad y diferenciación genética de seis poblaciones locales entre los 24° a los 42°S usando datos de 2251 loci SNPs neutrales y 81 loci putativamente sujetos a selección obtenidos mediante *Genotyping-by-Sequencing*. Los resultados encontrados con SNPs neutrales y SNPs putativamente bajo selección fueron contrastantes. Los loci neutrales, por un lado, reconocen un linaje evolutivo significativamente diferenciado en la población local de Los Molinos (39°S), con valores de  $F_{ST}$  un orden de magnitud mayores al resto de las comparaciones. Estos resultados son consistentes con los reportados en un estudio previo realizado con secuencias del gen mitocondrial Citocromo Oxidasa I (COI). Con marcadores SNPs neutrales no se detectó una discontinuidad genética asociada al quiebre biogeográfico de los 30°S. Por otro lado, los SNPs putativamente sujetos a selección arrojaron que la mayor parte de la divergencia se explica por la presencia dos grupos de poblaciones, uno al norte y otro al sur de los 30°S (que incluye a Los Molinos). Por lo tanto, la estructura filogeográfica de loci sujetos a selección estaría más influenciada por condiciones ambientales disímiles en el quiebre biogeográfico a los 30°S que por la historia de divergencia detectada con marcadores neutrales.



El segundo capítulo de esta tesis, se evaluó la estructura genética espacio-temporal en *Pyura chilensis*, utilizando 592 secuencias del gen mitocondrial COI en seis sitios (entre 24° y 42°S) y tres muestreos anuales (2012, 2014 y 2015). Las secuencias de COI correspondieron a tres linajes previamente descritos en la literatura, uno de los cuales es altamente divergente y está restringido a la zona de los 39°S. Los datos indican un patrón general de estabilidad temporal, aunque se observan diferencias significativas de un año a otro en ciertos sitios y al analizar estos tres linajes por separado. La estabilidad temporal detectada, podría deberse a restricciones al flujo genético, tal como se espera para especies con baja dispersión larval y susceptibles a cambios ambientales.

En el tercer capítulo, se utilizaron herramientas de genómica del paisaje, para asociar variables ambientales a la estructura genética adaptativa asociada al quiebre biogeográfico de los 30°S usando marcadores SNPs dispersos por el genoma en *Pyura chilensis* a lo largo del HCS. En este estudio analizamos la relación entre nueve variables oceanográficas y la estructura genética adaptativa usando 2902 SNPs en 149 individuos de 5 sitios entre los 24° y los 42°S. Se detectaron cientos de marcadores candidatos (entre 390-585), los que estuvieron significativamente asociados a la temperatura superficial del mar y a variables asociadas a surgencias y productividad como el fosfato, nitrato, fluorescencia y pH. La estructura genética espacial de los marcadores candidatos a cada variable ambiental muestra consistentemente el quiebre a los 30°S más allá de lo que se espera por la distancia entre sitios de estudio, lo que sugiere que estas variables podrían modelar la divergencia adaptativa entre el norte y sur de los 30°S en especies como *P. chilensis*. Adicionalmente, se identificaron 8 genes con funciones anotadas conocidas, las que están asociadas a procesos como filtración y ventilación (biogénesis de cilios), crecimiento y factores de transcripción.

La susceptibilidad del piure a cambios ambientales sugerida en este estudio se contraponen a reportes anteriores, donde se consideraba a la especie como un solo stock pesquero a lo largo de su distribución geográfica. Esto es fundamental tanto para la conservación del recurso, como para perspectivas futuras de estudios de resiliencia de organismos marinos al cambio climático.

En esta tesis, se ha mostrado como la heterogeneidad del paisaje puede contraponerse al efecto del flujo genético a lo largo de la corriente de Humboldt, lo que repercute en diferencias en la estructuración genética adaptativa de *P. chilensis* a través de procesos de adaptación local asociados a un quiebre biogeográfico asociado a los 30°s.

## Abstract

Local adaptation corresponds to the processes that arise because natural selection in heterogeneous environments is opposed to the homogenizing effects of neutral processes such as gene flow. In coastal marine habitats, many animals have dispersive larval stages that allow connectivity between local populations, yet connectivity is not necessarily correlated with dispersal capacity. On the other hand, the physical transport of individuals among local populations does not ensure effective gene flow, since this requires that the migrants survive and reproduce, implying that the selective mortality of migrants' due to environmental pressures could reduce gene flow between local populations.

Few phylogeographic studies have been approached using molecular markers to detect the effects of natural selection and contrast them with the effects of neutral evolutionary forces. For a better understanding of the contributions of different evolutionary forces in phylogeographic structuring, the tunicate *Pyura chilensis*, a sessile organism endemic to the highly heterogeneous Humboldt Current System (HCS), was used as the study model. This organism has a low capacity for larval dispersion, but a high dispersion capacity due to anthropogenic transport on deck of boats. Previous studies showed moderate genetic structure in the species and absence of genetic differentiation associated with a biogeographic break widely reported at 30°S, which has been reported for other marine taxa with low dispersal capacity.

In the first chapter, we compared the genetic diversity and differentiation of six local populations between 24° and 42°S using data from 2251 neutral SNPs and 81 putative under positive selection loci obtained through Genotyping-by-Sequencing. The results founded with neutral and putatively under selection loci were contrasting. Neutral loci, on the one hand, recognize a significantly differentiated evolutionary lineage in the local population of Los Molinos (39°S), with  $F_{ST}$  values of an order of magnitude higher than the rest of the pairwise comparisons. These results are consistent with those reported in a previous study using sequences of the mitochondrial gene Cytochrome Oxidase I (COI). Neutral SNP loci did not detect a genetic discontinuity associated with the biogeographic break at 30°S as main signal. On the other hand, SNPs putatively under selection showed that most of the divergence is explained by the presence of two groups of populations, one to the north and one to the south of the 30s (which includes Los Molinos). Therefore, the phylogeographic structure of loci under selection would be more influenced by dissimilar environmental conditions in the biogeographic break at 30°S than by the history of divergence detected with neutral markers.

The second chapter of this thesis, we evaluated the spatial-temporal genetic structure in *Pyura chilensis* using 592 partial sequences of the mitochondrial COI gene in six localities (between 24° and 42°S) and three annual sampling (2012, 2014 and 2015). The

COI sequences corresponded to three lineages previously described in the literature, one of which is highly divergent and restricted to the 39°S zone. The data indicate a general pattern of temporal stability, although significant differences were observed from year to year in certain geographical sites and when analyzing these three lineages separately. The detected temporal stability could be due to restrictions on the genetic flow, as expected for species with low larval dispersion and susceptible to environmental changes.

In the third chapter, seascape genomics tools were used to associate environmental variables with the adaptive genetic structure associated with the biogeographic break of the 30°S using SNPs scattered throughout the genome in *Pyura chilensis* along the HCS. In this study, we analyzed the relationship between nine oceanographic variables and the adaptive genetic structure using 2902 SNPs in 149 individuals from 5 sites between 24° and 42°S. Hundreds of candidate loci (between 390-585) were detected, which were significantly associated with sea surface temperature (SST) and variables associated with upwelling and productivity such as phosphate, nitrate, fluorescence and pH. The spatial genetic structure of candidate markers for each environmental variable consistently shows the break at 30°S beyond what is expected by the distance between study sites, suggesting that these variables could model the adaptive divergence between the north and south of 30°S in species such as *P. chilensis*. In addition, 5 genes with known annotated functions were identified, which are associated with processes such as filtration and ventilation (biogenesis of cilia), growth and transcription factors.

The susceptibility of the model species to environmental changes suggested in this study is in contrast to previous reports, where the species was considered as a single fishing stock along its geographical distribution. This is critical for both the conservation of the resource and future prospects for studies of the resilience of marine organisms to climate change.

In this thesis, it has been shown how the heterogeneity of the seascape can counteract the effect of gene flow along the Humboldt Current System, which has repercussions on the adaptive genetic structure of *P. chilensis* through associated local adaptation processes to a biogeographic break associated with the 30°S.

## Introducción General

### **Adaptación local**

Los ambientes naturales se caracterizan por ser altamente variables espacial y temporalmente tanto en factores bióticos como abióticos. Esta heterogeneidad ambiental implica que las presiones selectivas son también variables y que las poblaciones deben responder de forma dispar para poder adaptarse y enfrentar de forma óptima las condiciones locales (Nosil et al 2009, Sanford & Kelly 2011).

La selección natural tiende a favorecer a aquellos fenotipos que mejor se ajustan al ambiente local, provocando que los individuos tengan mayor adecuación biológica en su hábitat nativo que individuos foráneos provenientes de poblaciones emplazadas en condiciones ambientales disímiles (Kawecki & Ebert 2004). Tomando esto como base, el concepto de adaptación local se refiere a patrones observados entre poblaciones locales de la misma especie que se encuentran conectadas, al menos potencialmente por flujo genético, pero que están diferenciadas genéticamente por efecto de la activa elección de hábitats (Kawecki & Ebert 2004) y por desajustes entre los genotipos y los ambientes locales (Marshall et al 2010).

En hábitat marinos costeros, las especies que habitan en ambientes altamente heterogéneos, como roqueríos o playas de arena, tienen a estar estructuradas como metapoblaciones (Lipcius et al 2008, Carson et al 2011). En este tipo de sistemas, muchas de estas especies cuentan con estadíos larvales dispersivos que permiten la



conectividad entre poblaciones locales. La conectividad es un aspecto importante en la persistencia de las poblaciones, debido a que determina el grado de cohesión del acervo génico y la estructura geográfica de la diversidad genética (Palumbi 2003). Sin embargo, la existencia de estadíos larvales no está necesariamente correlacionada con la capacidad de dispersión efectiva (Swarer et al 2002; Thiel & Haye 2006; Teske et al 2006; Haye & Muñoz-Herrera 2013), ya que esta puede verse afectada por diversos factores. Entre estos factores están las barreras físicas y biológicas a la dispersión, el comportamiento larval, la ecología local, el transporte pasivo en objetos flotantes (rafting), la disponibilidad de hábitat, el transporte antropogénico, entre otros (Swearer et al 1999; Poulin et al 2002; Pineda et al 2007; Grosberg & Cunningham 2001; Gaines et al 2007; Shanks 2009; DeWitt et al 1998; Cowen et al 2000; Armsworth et al 2001; Largier 2003; Sotka et al 2004; Thiel & Haye 2006; Marshall et al 2008). Otro aspecto fundamental, es que el transporte físico de individuos entre poblaciones locales no asegura el flujo genético, ya que la conectividad efectiva (genética) requiere, por supuesto, que los migrantes sobrevivan y se reproduzcan (Pineda et al 2007; Hamilton et al 2008; Marshall et al 2010). Esto implica que la mortalidad selectiva de migrantes que dispersan a un nuevo hábitat debido a fuertes presiones ambientales previo a la reproducción podría reducir el flujo genético entre poblaciones (Marshall et al 2010). Por este motivo, la adaptación local podría ocasionar diferencias significativas entre el potencial de dispersión y la dispersión efectiva (Marshall et al 2010, Pineda et al 2007, Hamilton et al 2008), reduciendo ya sea el reclutamiento y/o establecimiento de individuos foráneos (Marshall et al 2010, Selkoe et al 2008, Cowen et al 2009), o la

adecuación biológica de juveniles que se establecen y persisten en la población receptora (Pechenick 2006).

En organismos marinos, se ha observado que, a pesar del efecto homogeneizador del flujo genético, la diferenciación genética puede prevalecer incluso entre poblaciones altamente conectadas (Hedgecock 1986, Slatkin 1987 Kawecki & Ebert 2004), o en especies con alto potencial de dispersión (e.g Peces: Conover 1998, Conover et al 2006, Nielsen et al 2009; Invertebrados: Grosberg & Cunningham 2001, Sotka 2005, Conover et al 2009). Lo anterior sugiere que en ambientes marinos la conectividad efectiva no depende exclusivamente del potencial de dispersión y que la adaptación local podría ser un proceso recurrente (Veliz et al 2005; Marshall et al 2010, Conover 1998, Conover et al 2006, Nielsen et al 2009, Grosberg & Cunningham 2001, Sotka 2005, Conover et al 2009).

Otro aspecto relevante es que el efecto homogeneizador del flujo genético podría no afectar de igual forma a todo el genoma. La idea de que unos pocos migrantes por generación son suficientes para mantener la homogeneidad genética entre poblaciones es relativamente común (e.g Wright 1932, Días-Viloria et al 2009, Goldsten et al 2009, Hellberg 2009). Sin embargo, esto podría no ser aplicable para loci sujetos a selección, sino solo para porciones del genoma que son neutrales, es decir no están sujetas a procesos selectivos (Hedgecock 1986, Slatkin 1987, Conover et al 2006) implicando que al solo considerar marcadores neutrales a la selección se podría sobrestimar la conectividad genética. Algunos estudios sugieren que poblaciones que experimentan baja diferenciación a nivel neutral pueden estar sujetas a fuertes presiones selectivas

debido principalmente a que estas regiones podrían estar ligadas a sectores del genoma sujetos a selección (Conover et al 2006). También se han observado correlaciones débiles en la diferenciación en marcadores neutrales y caracteres cuantitativos entre poblaciones (e.g. Leinonen et al 2008, Conover et al 2009). Por lo tanto, un aspecto importante a considerar es que es difícil inferir la arista adaptativa del cambio evolutivo a través del estudio de la variación genética neutral entre poblaciones (Nielsen et al 2009). Por esta razón, la inclusión de marcadores no neutrales o sujetos a selección es crucial para poner a prueba la hipótesis de que los procesos de adaptación local por efecto de presiones del ambiente podrían ocurrir a pesar de haber flujo genético activo entre poblaciones locales (Kawecki & Ebert 2004, Marshall et al 2008, Sanford & Kelly 2010), siendo la especificidad de los organismos a un ambiente en particular un objetivo fundamental para la conservación y manejo de recursos.

### **Aproximaciones para estudiar adaptación local**

La identificación señales de selección natural ha sido de gran interés para la biología evolutiva. Para poder identificar este tipo de señales se han utilizado dos tipos de aproximaciones (Baett & Hoekstra 2011); Aquellas denominadas *Top-Down*, las cuales incluyen asociaciones genómicas a amplio nivel (*Genome-Wide association o GWAS*, ver Korte & Farlow 2013), y mapeo de QTLs (ver Stinchcombe & Hoekstra 2008), utilizan la información de variables fenotípicas y las relacionan con datos genéticos y las *botton-up*, que incluyen a la genómica de poblaciones o la genética del paisaje, las que usan la información contenida en el genoma para poder identificar

señales de variación genética adaptativa y así intentar relacionarla con procesos evolutivos y/o con variación ambiental (Rellstab et al 2015).

En organismos marinos, la existencia de selección natural ha sido evaluada utilizando aproximaciones *top down* basadas en la identificación de diferencias en el rendimiento de distintos fenotipos/genotipos en distintos ambientes, ya sea mediante el estudio de las respuestas en ambientes controlados (jardines comunes) o a través de realización de trasplantes recíprocos (Conover 1998, Kawecki & Ebert 2004, Sanford & Kelly 2011, Conover et al 2006). Estos acercamientos experimentales buscan identificar el componente genotípico de la variación fenotípica (Sanford & Kelly 2011) minimizando la influencia de la variación ambiental. Una predicción asociada es que los individuos desplegarán mayor adecuación biológica en el hábitat de origen respecto a hábitats diferentes y, que estas diferencias a nivel poblacional pueden ser medibles genéticamente tanto a nivel genotípico como fenotípico (expresión génica; Sotka 2005). Estos estudios han mostrado que hay diferencias en la sobrevivencia de los organismos de diferentes orígenes en diferentes condiciones ambientales (Marshall et al 2010), así como mortalidad selectiva (genotipo dependiente) de reclutas y mortalidad post-colonización (Hedgecock 1986, Penney et al 2006), lo que se traduce en una alta diferenciación genética entre poblaciones. A pesar de los significativos y bien estudiados efectos de la adaptación local en la distribución geográfica y en la diversidad genética en poblaciones marinas, hay pocos estudios que hayan intentado abordar esta problemática desde una perspectiva filogeográfica, comparando la señales tanto espaciales como temporales de fuerzas tanto selectivas como neutrales a una



escala geográfica amplia (cubriendo gran parte de la distribución geográfica), permitiendo la evaluación de aquellos mecanismos determinantes de la distribución genética de una especie (e.j. flujo genético, deriva genética, selección natural).

Durante la última década, mediante el uso de técnicas de secuenciación masiva, ha sido posible evaluar diferenciación adaptativa mediante el estudio de marcadores putativamente sujetos a selección natural (e.g. Luikart et al 2003; Vasmägi & Primmer 2005; Nielsen 2001; Storz 2005; Joost et al 2007; Sanford & Kelly 2011). Dentro de estas aproximaciones están, por ejemplo, la caracterización de marcadores candidatos dispersos por todo el genoma, los cuales permiten evaluar la estructura genética adaptativa e inferir procesos de adaptación local a escalas temporales y espaciales entre poblaciones que habitan en condiciones ambientales disímiles (de Villermereuil & Gaggiotti 2015). En otras palabras, se intenta buscar e identificar marcadores con alta diferenciación genética (*outliers*) entre poblaciones basado en lo que se espera para un modelo neutral, lo que regularmente se asocia a la detección de selección positiva (Luikart et al 2003).

Una ventaja actual, es que las tecnologías de secuenciación masiva de última generación (NGS) permiten potencialmente mejorar significativamente la capacidad para estudiar procesos como adaptación local y selección natural a nivel genético, debido a la posibilidad de generar un gran volumen de datos extraídos desde varios individuos a lo largo de un gradiente ambiental (Sanford & Kelly, 2011). Por otra parte, en los últimos años, la cantidad de recursos genómicos en especies no modelo se ha incrementado significativamente (Ellegren 2014) pudiendo incluso hacer estudios a

nivel genómico a partir de organismos sin información genética previa (Hohenhohe et al 2011).

Para poder estudiar la estructura genética en marcadores sujetos selección, los llamados «escaneos genómicos» han demostrado ser una herramienta útil para evaluar la variabilidad genética usando un gran número de loci secuenciados en individuos que habitan en distintos ambientes (Pritchard & Di Rienzo 2010). La idea detrás de este tipo de aproximaciones es identificar las señales de adaptación local a un nivel genómico, con el objetivo de evaluar la estructura genética y la conectividad efectiva entre poblaciones en respuesta a la heterogeneidad ambiental (Nosil et al 2009; Schluter & Conte 2009; Manel et al 2016). En este contexto, las técnicas de Secuenciación de Representación Reducida (RRS por sus siglas en inglés), dentro de las cuales se encuentra RADSeq (Baird et al 2008) y Genotyping-by-Sequencing (GBS; Elshire et al 2011, DeDonato et al 2013), han sido ampliamente utilizadas, particularmente en especies no modelo. Este tipo de métodos se basan en la utilización de enzimas de restricción que reducen la complejidad del genoma, con la ventaja adicional de que pueden ser ejecutadas incluso sin recursos genómicos disponibles (i.e genoma de referencia). Estas técnicas permiten la secuenciación de un alto número de regiones genómicas anónimas cortas en varios individuos (Manel et al 2016), permitiendo la genotipificación de un alto número de Polimorfismos de Nucleótido Único o SNPs dispersos aleatoriamente por todo el genoma. El análisis de marcadores que se encuentran sujetos a selección (o al menos putativamente), permite identificar la

variación adaptativa a un nivel filogeográfico (Nosil et al 2009, Schluter & Conte 2009) para poder contrastarla con las señales de aquellos loci neutrales.

La gran parte de los estudios enfocados en evaluar la estructura genética espacial neutral y adaptativa en organismos marinos, han sido realizados principalmente en mamíferos marinos y peces (e.g Foll & Gaggiotti 2008, Moura et al 2016, Candy et al 2015). En estos estudios en general se observa que, como se espera por definición, aquellos loci putativamente sujetos a selección muestran una estructura genética mayor a lo observado a nivel neutral (Araneda et al 2016, Cahill & Levinton 2016), sin embargo, comparativamente, la señal tiende a ser congruente entre ambas. En el caso de invertebrados marinos bentónicos, estudios de este tipo son más bien escasos. En los últimos dos años estudios que evalúan la estructura genética neutral y adaptativa usando aproximaciones genómicas, los que han reportado señales consistentes a nivel neutral y putativamente adaptativo tanto en mitílidos (e.g Chorito; Araneda et al 2016, Xu et al 2016) como en ostras (e.g Lal et al 2016, Nayfa & Zenger, 2016).

Junto con los avances en la generación de datos genéticos de forma masiva, y en complemento al estudio de marcadores putativamente sujetos a selección detectados usando aproximaciones de diferenciación poblacional (*outliers*), está siendo recurrentemente siendo utilizada la identificación de variantes genéticas que se encuentren correlacionadas con factores ambientales específicos con herramientas de genética del paisaje a través de Análisis de Asociación Ambiental (EAAs por sus siglas en inglés), o Análisis genético-ambientales (e.g Lotherhos & Whitlock 2015). Este tipo de análisis tienen el potencial de detectar patrones inducidos por procesos adaptativos

los cuales no son detectados por aproximaciones genómicas tradicionales, o bien, pueden ser utilizadas en complemento a los resultados de estas (e.g Fischer et al 2013). El objetivo de las EAAs es integrar herramientas que van desde la genética del paisaje hasta la genómica poblacional con el fin de identificar los factores ambientales que están modelando la variación genética (adaptativa) y/o las variantes genéticas que están guiando a la adaptación local (Holderegger et al 2010, Manel et al 2010, Manel & Holderegger 2013, Sork et al 2013).

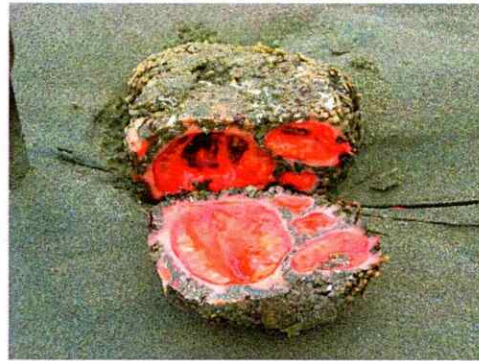
### **Sistema de la Corriente de Humboldt**

El sistema de la Corriente de Humboldt (HCS) en el pacífico sureste representa un escenario interesante para estudios de conectividad y estructura genética en organismos marinos. A pesar de ser un sistema lineal de aproximadamente 2600 km, y a la aparente ausencia de barreras geográficas/físicas al flujo genético, diversos estudios realizados en especies costeras del Sistema de Corriente de Humboldt han reportado diferenciación genética significativa entre poblaciones locales, particularmente asociada a un quiebre biogeográfico localizado entre 30°-32°S (e.g. Tellier et al 2009, Sánchez et al 2011, Montecinos et al 2012, Brante et al 2012, Haye et al 2014), zona que representa el límite de los rangos de distribución de muchas especies marinas (Thiel et al 2007, Rivadeneira et al 2012). En comparación al noreste del pacífico (Kelly & Palumbi 2010), al sureste australiano (Ayre et al 2009), y al sureste africano (Teske et al 2011), para la costa del Pacífico Sudeste en el Sistema de la Corriente de Humboldt, las características de la historia de vida y la oceanografía parecen estar

significativamente asociadas tanto a la diversidad como a la diferenciación genética de los organismos (Thiel et al 2007, Haye et al 2014). Según Haye et al (2014), el quiebre biogeográfico de los 30°S corresponde a una discontinuidad histórica que interrumpió el flujo genético entre ambas regiones (norte/sur de los 30°S) y solo los organismos con baja capacidad de dispersión han retenido esta señal, la cual ha sido borrada por el flujo genético en organismos con alta capacidad de dispersión. La permanencia de este quiebre podría deberse a diferencias ambientales al norte y al sur del quiebre de los 30°S, lo que permite hipotetizar que esta discontinuidad podría tener influencia contemporánea en los organismos que habitan este sistema. Dentro de estas diferencias ambientales, se observa un gradiente latitudinal de temperatura influenciado por flujo hacia al norte de aguas frías de origen subantártico y por aguas sub-superficiales de surgencias ricas en nutrientes de origen ecuatorial, provocando diferencias espaciales y temporales en la productividad a lo largo del sistema (Pizarro et al 1994, Arcos & Navarro 1986) lo que influencia significativamente a las comunidades costeras (Thiel et al 2007). Sumado a la influencia diferencial de aguas continentales (glaciares y descargas de ríos) (Figueroa 2002), la heterogeneidad del paisaje a nivel tanto espacial como temporal a lo largo de la HCS podría implicar que a pesar de haber flujo genético contemporáneo entre regiones podrían ocurrir procesos de adaptación local en organismos marinos.

## Modelo y acercamiento al estudio

*Pyura chilensis* Molina 1782 o piure, es un conspicuo habitante endémico del sistema de la Corriente de Humboldt. Es una ascidia intermareal sésil de alta demanda para consumo humano, con una amplia distribución geográfica habitando desde el intermareal bajo hasta el



Modelo de estudio: Piure chileno (*Pyura chilensis* Molina 1782)

submareal somero en la costa chilena y peruana, desde los 10°S hasta los 44°S (Lancelotti & Vásquez 2000). En ambientes intermareales rocosos, se pueden encontrar individuos y parches aislados adheridos al sustrato mediante una túnica cementada. En ambientes submareales, se observan amplias matrices que pueden contener hasta cientos de individuos. En el aspecto reproductivo, son organismos hermafroditas digónicos en donde, a pesar de que la reproducción sexual es la estrategia más común (más aun cuando hay conespecíficos cerca), también son capaces de autofecundarse (Manriquez & Castilla 2005). Es una especie con un bajo potencial de dispersión intrínseco, con una larva lecitotrófica de baja capacidad de dispersión (con una duración de 12-24 horas) que debe ser capaz de asentarse en ambientes heterogéneos a lo largo de la costa, lo que ocurre regularmente en sustratos duros o en la túnica de conespecíficos (Cea 1970, Manriquez & Castilla 2007). A pesar de su baja capacidad de dispersión larval, y en contraste a lo reportado para otras especies marinas

bentónicas de características biológicas similares (e.g. Haye et al 2014), *P. chilensis* muestra estructura genética moderada para marcadores mitocondriales (Citocromo oxidasa I. COI) y nucleares (factor de elongación alpha 1) (Haye & Muñoz-Herrera 2013). Adicionalmente, no se observa influencia de discontinuidad genética a los 30°S, siendo una excepción interesante tomando en cuenta que especies con larvas planctónicas de mayor duración muestran divergencia significativa a los 30°S (e.j. Sanchez et al 2011, Haye et al 2014).

La baja capacidad de dispersión larval es una característica extendida en ascídias. Sin embargo, este grupo se caracteriza por su carácter invasivo y por ser competidores dominantes del sustrato (Grosberg 1981, Nandakumar et al 1993, Rajbanshi & Pederson 2007). La mayoría de las introducciones ocurre cuando juveniles viables, adultos, o fragmentos de colonias son transportadas a nuevos hábitats ya sea en vectores marítimos (casquetes de barcos, cuerdas, o agua de lastre) (Hudson et al 2016, Ordoñez et al 2015, Pinneda et al 2011), por el movimiento en equipamientos de acuicultura, y como epibiontes adheridos a organismos bentónicos como crustáceos (Bernier et al 2008 para una revisión). A pesar de que esto no ha sido evaluado de forma explícita para *P. chilensis*, esta especie ha sido observada en los casquetes de navíos artesanales, en boyas, cuerdas, y en casi cualquier instalación marina como parte importante de las comunidades del *biofouling* (Valdivia et al 2005, Cifuentes et al 2007). Esto se suma reportes recientes de especies de ascidias de similares características biológicas introducidas en la Bahía de Coquimbo (Turon et al 2016).

La influencia del transporte antropogénico es una arista importante, tomando en cuenta que la actividad económica portuaria es una práctica común en toda la costa chilena (Thiel et al 2007), con una data de aproximadamente 150 años de actividad. Sin embargo, el flujo marítimo en la HCS es de menos de un tercio respecto otras zonas portuarias de Chile como Punta Arenas y puertos adyacentes (Keller et al 2010).

El transporte antropogénico es probablemente el agente dispersivo e invasivo más importante para este tipo de organismos (Pineda et al 2011, Hudson et al 2016, Ordoñez et al 2013; Pineda et al 2016; Bullard & Carman et al 2009, Evans et al 2013) y ha sido hipotetizado como uno de los mecanismos para explicar el bajo nivel de diferenciación genética de *P. chilensis* a lo largo de una escala espacial amplia, y que podría explicar la ausencia de discontinuidad genética a los 30°S (Haye & Muñoz-Herrera 2013).

El flujo genético entre regiones mediado por procesos antropogénicos influye fuertemente en que *P. chilensis* sea un competidor dominante que tiende a colonizar rápidamente sustratos libres (Valdivia et al 2005, Cifuentes et al 2007). A pesar de esto, la especie parece ser susceptible a perturbaciones y a cambios en factores abióticos (Valdivia et al 2005, Cifuentes et al 2007). Esto permite hipotetizar que la influencia de factores ambientales podría oponerse al efecto homogeneizador del flujo genético, lo que se traducirá una fuerte estructura genética adaptativa asociada al quiebre biogeográfico de los 30°S.

Otro aspecto interesante acerca del modelo de estudio es que la estructura genética de esta especie, basada en ADNmt, está caracterizada por la presencia de tres



haplogrupos, dos de los cuales son simpátricos y con una amplia distribución geográfica, y un tercero altamente divergente y geográficamente restringido, encontrándose primordialmente entre los 39°-42°s, siendo el haplogrupo más abundante en la localidad de Los Molinos ubicada a 39°S (Haye & Muñoz-Herrera 2013). La divergencia de este haplogrupos coincide con los bajos niveles del mar y períodos glaciales durante el Pleistoceno (Haye & Muñoz-Herrera, 2013), lo que podría ser un indicador de la huella histórica en la variación geográfica en la diversidad genética del *P. chilensis*. Sin embargo, la baja estructura encontrada en estos genes neutrales, los cuales son el foco de la mayoría de los estudios genético poblacionales, no representa necesariamente la estructura genética espacial de todo el genoma, ni es necesariamente un patrón estable en el tiempo.

Usando marcadores sujetos a selección, y comparando la estructura genética con marcadores neutrales tanto a nivel espacial como temporal, será posible explorar la señal putativamente adaptativa en presencia de un linaje altamente divergente (39°S), así como la influencia de un quiebre biogeográfico que implica desafíos importantes en términos fisiológicos para los individuos que logran migrar entre ambas regiones.

En este trabajo, se evaluó la estructura genética tanto espacial como temporal utilizando tanto marcadores SNPs dispersos por el genoma como secuencias parciales del gen mitocondrial *Citocromo Oxidasa I* (COI) en la ascidia *Pyura chilensis*, cubriendo una zona desde los 24 a los 42°S en el pacífico sudeste en la HCS, lo cual abarca gran parte del rango de distribución de la especie.

En el primer capítulo, utilizando tanto la variación adaptativa como neutral de SNPs dispersos por el genoma, se evaluó la influencia de fuerzas evolutivas contrastantes como la selección natural a lo largo de un ambiente heterogéneo (adaptación local), aislamiento y deriva genética (divergencia del linaje a los 39°S) y el transporte artificial en la estructura genética de *Pyura chilensis*. En el segundo capítulo, se evaluó la estructura genética inter-anual de los patrones observados para el gen *COI*, para evaluar la estructura tanto espacial como temporal de los tres linajes reportados por Haye & Muñoz-Herrera (2013). Finalmente, en el capítulo tres se utilizaron herramientas de genética del paisaje para intentar correlacionar la estructura genética adaptativa con variables ambientales en *P. chilensis* a lo largo de la HCS, e intentar identificar aquellos factores ambientales que podrían explicar la influencia de la heterogeneidad en la ocurrencia del quiebre de los 30°S.

**CAPÍTULO I: Lineage divergence, local adaptation across a biogeographic break, and artificial transport shape the genetic structure in the ascidian *Pyura chilensis***

Nicolás I. Segovia<sup>1,2,3</sup>, Cristian Gallardo-Escárate<sup>2</sup>, Elie Poulin<sup>3</sup> & \*Pilar A. Haye<sup>1,2</sup>

<sup>1</sup>Departamento de Biología Marina, Facultad de Ciencias del Mar, Universidad Católica del Norte, 1781421, Coquimbo, Chile.

<sup>2</sup>Interdisciplinary Center for Aquaculture Research (INCAR), Departamento de Oceanografía, Universidad de Concepción, 4070386, Concepción, Chile.

<sup>3</sup>Instituto de Ecología y Biodiversidad, Departamento de Ciencias Ecológicas, Facultad de Ciencias, Universidad de Chile, 7800003, Santiago, Chile.

\***Corresponding author:** Pilar A. Haye, E-mail: [phaye@ucn.cl](mailto:phaye@ucn.cl)

**Key Words:**

Genotyping by Sequencing, GBS, diversifying selection, Southeast Pacific, Chile

**Formato: Scientific Reports (Nature).**

**Estado: Publicado**

## **Abstract**

Marine benthic organisms inhabit a heterogeneous environment in which connectivity between populations occurs mainly through dispersive larval stages, while local selective pressures acting on early life history stages lead to non-random mortality, shaping adaptive genetic structure. In order to test the influence of local adaptation and neutral processes in a marine benthic species with low dispersal, in this study we used Genotyping by Sequencing technology to compare the neutral and putatively selected signals (neutral and outlier loci, respectively) in SNPs scattered throughout the genome in six local populations of the commercially exploited ascidian *Pyura chilensis* along the southeast Pacific coast (24°-42°S). This species is sessile as an adult, has a short-lived larval stage, and may also be dispersed by artificial transport as biofouling. We found that the main signal in neutral loci was a highly divergent lineage present at 39°S, and a subjacent signal that indicated a separation at 30°S (north/south), widely reported in the area. North/south separation was the main signal in outlier loci, and the lineage divergence at 39°S was subjacent. We conclude that the geographic structure of the genetic diversity of outlier and neutral loci was established by different strengths of environmental, historical and anthropogenic factors.

## Introduction

Marine benthic species that inhabit patchy and heterogeneous substrates such as rock beds or sandy beaches are likely to be structured as metapopulations<sup>1,2</sup>. These species usually have dispersive larval stages that can act as propagules for connectivity between local populations. However, the amount of connectivity between their populations does not depend only on the presence and duration of dispersive larval stages<sup>3,4</sup>. Local adaptive pressures may differ drastically between local populations and may act as a settlement and survival sieve for early life stages of benthic marine organisms arriving to a local population<sup>5</sup>. Local adaptation can lead to strong differences between larval dispersal potential and realized dispersal<sup>5-7</sup>, by reducing recruitment and establishment of foreign individuals<sup>5,8,9</sup> or by reducing the fitness of juveniles that establish and persist in the receiving population<sup>10</sup>.

The consequences of local adaptation in marine benthic species have been assessed with experimental methods such as reciprocal transplants and common garden experiments. These studies have shown there is differential survival between organisms of different origins in different environmental conditions<sup>5,11</sup>, and that there is non-random, genotype-dependent, selective mortality of recruits and post-colonization mortality<sup>3,12</sup> that lead to high genetic differentiation among adult populations. In spite of the well-recognized effects of local adaptation on the geographic distribution of the genetic diversity of marine populations, there are few examples of phylogeographic studies that compare the spatial genetic signatures of natural selection and neutral evolutionary forces at a broad scale, allowing the

evaluation of gene flow, genetic drift, natural selection and other mechanisms that may determine the geographic distribution of the genetic diversity of a species.

One experimental approach to analyze neutral and putatively adaptive signals is the use of genome scans to detect thousands of Single Nucleotide Polymorphism (SNP) loci scattered throughout the genome, some of which can be identified as having signals of natural selection by analyzing several individuals from different environments<sup>13</sup>. Spatial analyses of loci putatively affected by natural selection (i.e. outlier loci) provide insights into how adaptive variation could affect phylogeographic structure<sup>14,15</sup> in contrast to the signal of loci that are neutral to natural selection. Studies focused on the assessment of phylogeographic signals of outlier and neutral loci in marine systems have been mostly developed for marine mammals and fishes<sup>16-18</sup>. These studies have generally shown that outlier loci possess a stronger signal of the genetic structure, as is expected by definition<sup>19,20</sup>. However, the main signal is mostly congruent between outlier and neutral loci. This kind of study in marine invertebrates is scarce; only recently a few studies have compared the phylogeographic signal of neutral and outlier loci using genome-wide procedures and have reported a mostly consistent structure between neutral and outlier loci in mussels<sup>19,21</sup> and oysters<sup>22,23</sup>.

In this study, we analyze neutral and adaptive spatial genetic structure of the benthic ascidian *Pyura chilensis* Molina 1782 along the Humboldt Current System (HCS) located in the southeast Pacific Ocean. The HCS is characterized by high productivity and regions with strong upwelling<sup>24</sup>. The wide geographic extent of the southeast Pacific coast and the cyclic variations of the HCS promote both temporal and spatial

changes in population sizes and distributions of marine benthic species<sup>24</sup>. Despite being an almost linear coastline with apparent absence of geographic barriers between 18°S and 42°S, within the HCS coastal area there is a well-described marine biogeographic boundary located at 30-32°S. This biogeographic break is the limit of the distribution range of several marine species<sup>24</sup>, and phylogeographic studies of benthic marine taxa have shown that species with low dispersal capability have strong genetic discontinuity across the 30°S biogeographic boundary<sup>25,26</sup>. Life history attributes of marine benthic species and the oceanographic conditions seem to be the main factors that contribute to the geographic structure of the genetic diversity and differentiation in this geographic area<sup>24,26</sup>.

*Pyura chilensis* is a conspicuous inhabitant of the HCS widely distributed in the intertidal and shallow sub-tidal from 10°S to 44°S in the southeast Pacific coast<sup>27</sup>, and has been intensively exploited for human consumption<sup>28</sup>. Settlement occurs mainly on rough substrates and on the tunic of conspecifics<sup>29</sup>; it lives either as solitary individuals or as massive clumps of individuals (Supplementary Fig. S1).

The species is hermaphroditic, with outcrossing as the main reproductive strategy<sup>28</sup>. *Pyura chilensis* has short-lived (12-24 hr) lecithotrophic tadpole-shaped larvae that provides a very low capacity to disperse between geographic areas along the HCS. In contrast to what has been reported for other sympatric benthic species of the HCS with low larval dispersal capacity<sup>26</sup>, *P. chilensis* did not show a genetic discontinuity at 30°S when analyzed with sequences of the mitochondrial gene Cytochrome Oxidase I (COI) and the nuclear gene Elongation Factor 1 alpha<sup>30</sup>. The lack of a genetic discontinuity at

30°S in the coast of Chile was novel for a benthic invertebrate with short-lived larvae. In fact, some species with planktonic larval durations of more than 2 days display significant genetic divergence across the 30°S biogeographic boundary<sup>26</sup>. By uncovering abundant neutral and adaptive genetic variation in *P. chilensis* it may be possible to detect variation at 30°S, as would be expected in a benthic species with a low dispersal potential.

An alternative means of dispersal for ascidians such as *P. chilensis* is transport in artificial structures such as ship and boat hulls, where the species can grow attached<sup>31-33</sup>. Although this has not been specifically assessed for *P. chilensis*, the species has often been seen on the hulls of boats of artisanal fisheries, on buoys, ropes and on any other marine facility that provides substratum as an important part of the fouling community<sup>34-35</sup>. This has been hypothesized as one of the mechanisms behind the low degree of genetic differentiation of the COI gene along a broad geographic area and across the 30°S biogeographic break<sup>30</sup>; similarly, anthropogenic transport has often been suggested as a main dispersal and invasion mechanism for other ascidians<sup>30-33</sup>.

Another relevant aspect about *P. chilensis* is the genetic structure of the species based on COI sequences. The COI phylogeographic structure of the species was characterized by three mitochondrial haplogroups, two of which were sympatric and widespread, and a third that was highly divergent from the other two and geographically restricted, predominantly found between 39°S and 42°S, being most abundant at 39°S in Los Molinos (LM)<sup>30</sup>. Based on COI sequences, LM was the most differentiated population because it harbored a high proportion of individuals of a



unique and divergent haplogroup. Here we evaluate the divergence of the LM population of *P. chilensis* using neutral and putatively adaptive SNP loci, and explore if natural selection follows the same signal as neutral sequence divergence, i.e. if LM is also highly differentiated with putatively adaptive loci. Even though physical transport of a few individuals per generation between local populations should be enough to maintain genetic homogeneity in the whole genome at both individual and population levels<sup>3,36</sup>, non-neutral genetic variation affects fitness, influencing recruitment and survival of individuals in the variable marine environment. Thus, lack of spatial genetic differentiation at neutral or nearly neutral loci does not imply that outlier loci will provide the same signal<sup>36</sup>.

Here we evaluate the spatial genetic structure of neutral and outlier SNP loci in six populations of *P. chilensis* covering from 24°S to 42°S, and across the 30°S biogeographic discontinuity in the southeast Pacific. We aim to assess the influence of contrasting evolutionary forces, including natural selection (local adaptation), isolation and drift (lineage divergence at 39°S), and gene flow (artificial transport), on the spatial genetic structure of the benthic ascidian *P. chilensis* in the HCS.

## **Results**

### ***SNP discovery and outlier detection***

Extracted DNA was of high quality, with average 260/280 and 260/230 ratios of  $1.96 \pm 0.01$  and  $2.15 \pm 0.03$ , respectively. Genotyping by Sequencing (GBS) of 90 individuals

of *P. chilensis* (15 of each of six local populations; Table 1) gave a total of 202,494,184 reads. Of these, 169,855,853 produced good quality barcode reads. The FastQC results showed a mean per-base quality score (*Phred* score) of 36 and a mean quality score of 33.12. After applying the trimming and demultiplexing approach, 163,692 tags were produced with a mean of 4.38 reads per individual. The UNEAK pipeline detected a total of 81,837 SNP. After filtering with a minor allele frequency (MAF) of 4.4% and a minimum call rate (mnC) of 90%, the resulting dataset consisted of 2,332 SNPs.

Using BAYESCAN with a False Discovery Rate (FDR) correction of *q-values* of 0.05, 81 of the 2,332 loci uncovered showed  $F_{ST}$  values greater than expected under a neutral distribution (*q-values* < 0.05) (Supplementary Fig. S2). These 81 loci were analyzed separately; thus, there were two datasets, one with the 81 outlier SNP loci and one with 2,251 neutral SNP loci.

### ***Population genetic diversity and population structure analyses***

Mean expected and observed heterozygosity ( $H_E$  and  $H_O$ , respectively) were mostly similar between neutral and outlier loci.  $H_E$  mean values ranged from 0.290 to 0.294 and 0.250 to 0.358 for neutral and outlier loci, respectively. Mean  $H_O$  values ranged from 0.169 to 0.304 and 0.143 to 0.345 for neutral and outlier loci, respectively (Table 2). The mean allelic richness ( $A_r$ ) within population ranged from 1.560 to 1.662 for neutral loci and 1.511 to 1.626 for outlier loci. Northern populations generally showed lower values of  $A_r$  in neutral loci and similar values between all sampled populations in

outlier loci; LH and LM were the sites with the lowest values. The exact test for HWE deviations showed that after FDR correction, only the LM site showed significant deviation (Table 2).

Most of the population-pairwise  $F_{ST}$  values for both neutral and outlier loci were significant, with the exception of PA and CP in neutral loci (Table 3), indicating that *P. chilensis* harbors neutral and adaptive population differentiation.  $F_{ST}$  values ranged between 0.003-0.417 and 0.027-0.692 for neutral and outlier loci, respectively (Table 3). For neutral loci  $F_{ST}$  values were generally low (0.003 to 0.049) with the exception of Los Molinos (LM) that was highly differentiated with values ranging between 0.393 and 0.417, which are an order of magnitude greater than all the rest of the values. In contrast, all  $F_{ST}$  values calculated with outlier loci were, as expected, relatively high; remarkably, in neutral loci LM did not show higher levels of differentiation than the other population pairwise comparisons. In the outlier loci population-pairwise differentiation matrix, the values between pairs from the northern (PA, 26°08'S; CP, 27°44'S; LH, 29°58'S) versus southern (TH, 36°38'S; LM, 39°50'S; AC, 41°52'S), *i.e.* north/south of 30°S, showed the highest differentiation values (Table 3).

Using AMOVA analysis the separation of populations north and south of 30°S explained 2.91% and 48.73% of the variation for neutral and outlier loci, being significant only for outlier loci. The analysis of the separation of LM from the rest of the populations explained 39.5% and 8.74% of the genetic variation for neutral and outlier loci, respectively, and was only significant only for neutral loci (Table 4).

Mantel tests did not detect a significant relationship between the linearized genetic distance and geographic distance for either neutral or outlier markers, with nearly significant *p-values* ( $r=0.034$ ,  $P = 0.089$  and  $r = 0.75$ ,  $P = 0.068$ , respectively) (Fig. 1a-b). Excluding LM from the analysis of neutral loci indicated a marginally significant relationship between genetic and geographic distance ( $r=0.85$ ,  $P=0.048$ ) (Fig. 1c). In the case of outlier loci, this led to virtually the same result as when LM was included in the analysis ( $r = 0.81$ ,  $P = 0.064$ ) (Fig. 1d).

In order to detect spatial population genetic structure, we used the GENELAND software, which revealed the presence of three groups in the neutral loci dataset. One corresponded to the three northern localities PA, CP and LH; the second included TH and AC and the third was LM (Fig. 2a). Five groups were detected using outlier loci; one group contained PA and CP, and each other local population was assigned to a single genetic group (Fig. 2b).

Cluster analyses confirmed the disparate signals of neutral and outlier loci. In the clustering approach of the neutral loci analysis using DAPC (considering the sampling site as a prior), the first principal component axis (vertical) separated LM from all the rest of the populations (Fig. 3a). The second axis (horizontal) separated the north/south of 30°S, with an overlap in the three populations of the north (PA, CP, LH) and an overlap of TH and AC in the south (LM did not show overlap) (Fig. 3a). Thus, the main signal derived from the DAPC analysis of neutral loci was the separation of LM from the rest of the populations and there was a weaker signal associated with the 30°S biogeographic boundary. In outlier loci, north/south of 30°S were well separated in the

horizontal axis, with an overlap of PA and CP in the north of TH and LM in the south. The vertical axis separated the PA/CP group from LH, and LM from AC and TH (Fig. 3b).

As with GENELAND analyses, DAPC performed without considering the information of the geographic origin of the samples, and based on the BIC, detected three and five groups as the optimal separation for neutral and outlier loci, respectively (Fig. 3c-d). For neutral loci, the three northern populations appeared in the same cluster, and in the south, TH and AC appeared in the same cluster but LM was clearly separated, with the exception of three individuals that clustered with TH and AC (Fig. 3c). In outlier loci the five groups were mostly consistent with the geographic origin with the exception of PA/CP, which formed a single cluster (Fig. 3d). The third group was composed mainly of individuals of LH (Fig. 3d).

Dendrograms constructed with Identity-by-Site (IBS) without incorporating geographic information detected three and four clusters for neutral and outlier loci, respectively, providing a similar pattern of differentiation as previous analyses (GENELAND, DAPC). With neutral loci LM appeared as a distant separate cluster, while the two other clusters were closer together. These are the three localities north of 30°S and the southern localities TH and AC (Fig. 3e). Again, three individuals of LM appeared together with the individuals sampled in TH and AC. For outlier loci, the north/south of 30°S were well separated (Fig. 3f). The analysis detected two groups in the northern cluster. The first group was composed mostly of individuals of LH and of individuals of PA and CP in lower frequency; the second group was composed

exclusively of individuals of PA and CP (Fig. 3f). In the southern cluster, most of the individuals from LM (except three) conformed a well-separated cluster and the rest of the southern individuals including those three individuals from LM formed a single cluster (Fig. 3f).

With Evanno's method,  $K = 2$  was determined as the optimal separation for both data sets in analyses based on STRUCTURE (Fig. 4a-b). For neutral loci one cluster included most of the individuals of LM (12 of 15 individuals), while the second included all other sampling sites plus the three remaining individuals of LM (Fig. 4c). A slight north/south of 30°S separation appeared as a subjacent signal when forcing  $K = 3$  and  $K = 4$  (Fig. 4c).

The optimal clustering separation for outlier loci,  $K = 2$ , corresponded to a separation between north/south of 30°S (Fig. 4d). These results indicate that the optimal  $K$  of neutral and outlier loci provide a different spatial genetic structure of *P. chilensis*. Forcing  $K = 3$  for outlier loci divided the northern cluster in two, one group with PA and CP with strong influence of a third group located mainly in LH. Only when forcing  $K = 4$  an additional cluster appeared in the south corresponding to 12 of the 15 individuals of LM (Fig. 4d). The LM cluster was evident with optimum  $K$  for neutral loci while it appeared with non-optimum  $K = 3-4$  for outlier loci. Exploring from  $K = 5$  to  $K = 7$ , there was no significant difference in the general pattern compared to  $K = 4$  in both datasets (results not shown). When LM was excluded from the analyses, the optimal separation also corresponded to  $K = 2$  (Fig. 4a-b). For both

data sets, the clusters formed corresponded to the north/south of 30°S separation (Fig. 4e-f).

## Discussion

The use of putative selected markers (outlier loci) in marine benthic invertebrates has usually increased the strength of detected phylogeographic structure with respect to neutral loci, as is expected, albeit the general patterns remain the same<sup>19,21,22,23</sup>. In contrast to this generalization, in the present study neutral and outlier loci showed contrasting spatial genetic structure patterns in the ascidian *Pyura chilensis* along 18° degrees of latitude in the Southeast Pacific coast. Neutral loci had a main signal associated with the lineage history and genetic differentiation of Los Molinos (LM) (39°S), and had a subjacent weaker signal of genetic structure across the 30°S biogeographic boundary. Inversely, outlier loci showed a stronger signal across 30°S and a weaker subjacent signal of lineage divergence at 39°S (LM). The two sets of loci showed the same two phylogeographic patterns, structure across 30°S and 39°S, but their strength was in reverse order. Additionally, the degree of neutral genetic differentiation detected was lower than expected from the short larval duration (2 days) and did not adjust to a pattern of isolation by distance (IBD), suggesting that artificial transport also contributes to the neutral genetic structure of *P. chilensis* along the Humboldt Current System (HCS). SNP sampling and detection of putatively adaptive loci allowed the evaluation of the contributions of several evolutionary mechanisms in the phylogeographic structure of *P. chilensis*. We detected a strong adaptive signal of

differentiation at the 30°S biogeographic break, a strong neutral signal associated with the divergent lineage present at 39°S, and the influence of artificial anthropogenic transport.

LM had pairwise genetic differentiation values with neutral loci one order of magnitude higher than the rest. This is explained by 12 of the 15 individuals from LM that were highly differentiated from all other individuals of all other sites, independent of whether the analytical approximation used was frequency or distance-based. These differentiated individuals belonged to the divergent mitochondrial DNA haplogroup previously identified in LM<sup>30</sup> (Supplementary Table S1). The extent of the geographic distribution of the divergent lineage that is present in LM needs to be further investigated with an intensive sampling between TH and LM and in the vicinity of LM towards the southern AC locality. It is likely that migration to LM from well-differentiated populations may account for the detected genetic differentiation. The apparent absence of admixture between individuals of LM belonging to the two different clusters as shown by STRUCTURE (Fig. 4e) suggests that these differentiated individuals found in LM may correspond to a reproductively isolated unit. Although these results clearly identified different genetic groups, the existence of a cryptic species in this area should be further investigated, including molecular and morphological systematic studies<sup>37</sup>, before drawing taxonomic conclusions.

As expected, outlier loci showed higher levels of differentiation between all local populations than neutral loci, and the genetic signal provided by outlier loci differed from the neutral signal. The main signal in outlier loci was a strong differentiation



north/south of the 30°S biogeographic break, explaining almost half of the total variation in the data. In contrast to the signal detected with neutral markers, outlier loci did not show a main signal associated with the genetic differentiation of LM. Still, the main signal of differentiation north/south of 30°S in outlier loci did not completely expunge the influence of the differentiation of LM; LM appeared as a distinct group when forcing a third suboptimum clustering ( $K = 4$ ) in STRUCTURE (Fig. 4f), and within the southern cluster the IBS dendrograms showed that the same 12 individuals from LM were clearly separated from the other individuals from south of 30°S (Fig. 4f). Additionally, Hardy-Weinberg Equilibrium (HWE) was prevalent; LM was the only population with significant deviations. We explored HWE excluding the 3 individuals of LM that are not highly divergent from the rest of the data set, and after the FDR correction, LM did not deviate from HWE, suggesting that lack of HWE in LM was likely due to a Wahlund effect, in agreement with presence of two divergent lineages in the site<sup>30</sup>. The fact that the 12 divergent individuals from LM did not appear as a separate and external cluster for outlier loci may reflect that natural selection acts globally on these taxa, responding more to environmental constraints than lineage evolution. In other words, the coherence between geographic localities and genetic distribution in outlier loci suggests that selection, independent of the evolutionary history of the lineage, maintains the genetic cohesion of the species mainly based on environmental constraints rather than on lineage history. As genetic variation in neutral markers may reflect historical and contemporary connectivity among populations<sup>38</sup>,

selected markers may reflect adaptive processes associated with differences in environment conditions experienced by the species along the HCS.

IBD including the all data set was nearly significant and it was marginally significant when LM was excluded from the analysis. However, STRUCTURE analyses showed that when LM was excluded, most of the genetic variation was explained by the differentiation to the north and south of 30°S, with optimum  $K = 2$ . Therefore, Mantel test results may just reflect a main genetic discontinuity between north/south of 30°S instead of a progressive decrease in migration rate among neighboring sites relative to geographic distance.

Phylogeographic patterns of marine invertebrates along the HCS have shown that the biogeographic boundary at 30°S is a historical discontinuity and that species with low dispersal potential retain the genetic signal and have a concordant phylogeographic break<sup>26</sup>. However, this structure had not been previously detected in *P. chilensis* using sequence data<sup>30</sup>. This is not surprising if we consider that neutral loci revealed that the structure at 30°S is subjacent to a stronger signal of lineage divergence at 39°S, which was the same signal detected with sequence data<sup>30</sup>. The clear differentiation detected by outlier loci between north/south of 30°S is likely linked to the historical and contemporary existence of a marine biogeographic boundary<sup>24</sup>. The contemporary influence of this discontinuity is likely attributed to the heterogeneous environmental conditions that differ markedly north and south of 30°S, mainly, the kinetic energy of the ocean<sup>39</sup>, differences and seasonality in upwelling-favorable winds<sup>40</sup>, and differences in the influence of freshwater on coastal waters<sup>24</sup>. The oceanographic differences

between north/south of 30°S could imply that in sessile organisms, such as *P. chilensis*, the effective connectivity (recruitment and survival of juveniles) is restricted, despite the genetic evidence that there is transport of individuals between the two areas<sup>30</sup>. Considering that neutral phylogeographic structure of benthic marine species across 30°S is strongly associated with life history<sup>26</sup>, *P. chilensis* showed phylogeographic structure across 30°S as expected, although it had less differentiation across 30°S than other species with similar dispersal potential based on larval duration.

Differentiation was greater than had been previously reported for *P. chilensis*<sup>30,41</sup>, and also greater than what has been reported for other ascidians that have intensive connectivity driven by anthropogenic transport<sup>31,32,33,42</sup>. The influence of anthropogenic transport was evident from the low, albeit significant differentiation detected between distant local populations, such as the southernmost site analyzed AC and all the northern sites, and general the lack of IBD. The single exception to the significant genetic differentiation detected with neutral loci was the population pairwise genetic differentiation between PA and CP, that was low and non-significant. A possible explanation is that PA and CP are separated by less geographic distance than any other pair of populations, permitting higher gene flow between them than with the rest of the sites. However, short larval life history alone is unlikely to maintain high connectivity between PA and CP, suggesting that artificial transport may be enhancing connectivity between these sites, in agreement, shipping trajectories in the area between PA and CP seem to be geographically continuous and intense (Supplementary Fig. S3).

Physical transport as biofouling may have allowed the latitudinal expansion of *P. chilensis* after periods of isolation during the Pleistocene<sup>30</sup>, but the invasive potential is likely less than in other invasive ascidians. In part, this may be explained by the susceptibility of the species to abiotic disturbances<sup>34</sup>, the low and strongly seasonal recruitment rate<sup>43</sup>, and, compared to colonial ascidian growth forms (e.g. *Diplosoma* sp.), a slower growth rate<sup>34</sup>. Higher growth and recruitment rate and rapid reproduction are important life history features in invasive ascidians that allow larvae to colonize open surfaces quickly<sup>44</sup>. Valdivia *et al*<sup>34</sup> suggested that *P. chilensis* is a good example of a strong competitor with relatively low colonizing ability.

Shipping routes (Supplementary Fig. S3) may also explain the greater differentiation of LM and of LH. LM is the most differentiated population, with a divergent group of individuals, and fittingly, there is very low maritime traffic in the area around LM. There is also low traffic inside La Herradura Bay (LH). The lower differentiation of LH than LM may be explained by the distance from the open coast, which is much smaller in LH, probably allowing slightly greater connectivity by maritime routes. Additionally, differentiation of LH was enhanced by the high proportion of individuals belonging to a distinct mitochondrial haplogroup<sup>30</sup> (Supplementary Table S1). Other sites had from 0 to 4 individuals of this haplogroup, while LH had 7 of the 15. The relatively high genetic differentiation of LH is likely shaped by the lower maritime traffic present in the area, enhancing genetic differentiation.

Although the results are robust and allow inferring the spatial genetic structure of loci affected by contrasting evolutionary forces, it is important to consider alternative explanations associated with the quality of information provided by the data. For example, identified outlier SNP loci may be linked to loci that are the direct target of natural selection instead of natural selection directly operating on them<sup>45</sup>. Additionally, spatial autocorrelation can cause random associations between the environment and the genetic structure of a species as a result of dispersal following IBD or demographic processes<sup>46</sup>. Some coding regions may have greater genetic-environmental associations than others due to deleterious mutations being selected against in all the study area (purifying selection), and not because advantageous mutations are being selected for in a particular environment (local adaptation<sup>47</sup>). Other factors that may cause loci to behave as outliers are recombination, sampling design, and locus-specific effects<sup>46,48</sup>.

Demonstrating that local adaptation has occurred involves not only the detection of variation at outlier loci, but also the understanding of the functional differences between alleles of individual candidate genes<sup>48</sup>. The knowledge of the physical linkage of SNPs allows the identification of genomic regions with unequal response to natural selection<sup>45</sup>. Since *P. chilensis* is a non-model organism, information on candidate genes is not currently available. In spite of the possible caveats, 3.5% (81) of the total retained SNPs (2,332) showed low and significant *q-values* and higher levels of  $F_{ST}$  for 90 individuals, allowing us to distinguish individuals genetically between sampling locations at a relatively small geographic scale, and suggesting that significant differences could be due to positive selective pressures<sup>49</sup> that differ north/south of 30°S.

The percentage of outlier loci detected is proportionally consistent with other studies that show high variation between local populations<sup>19,22,23,50,51</sup>. Further studies are needed to evaluate the genes involved in the putatively adaptive genetic structure detected here, and determine if putatively selected loci are actually under positive selection due to different environmental conditions north and south of the biogeographic boundary at 30°S in the southeast Pacific. Here we showed that even in the presence of a highly divergent lineage, the putatively selected signal could recover a pattern of genetic structure that masks the evolutionary history of the species.

In this study, we obtained several SNPs scattered throughout the genome and analyzed and compared the genetic structure of neutral and putatively adaptive outlier loci in *Pyura chilensis*. The comparison allowed us to conclude that both outlier and neutral loci were diverse and variable in space, and that both have shaped the genetic diversity of *P. chilensis*. The geographic structure of the genetic diversity of outlier and neutral loci was established by different strengths of environmental, historical and anthropogenic factors.

## **Methods**

### ***Sample collection and DNA extraction***

Samples of *Pyura chilensis* were obtained by local fisherman from six sites along the Humboldt Current System in the Southeast Pacific Ocean (Table 1). Fisherman sampled individuals from natural substrates in areas with abundant *P. chilensis*. In all

sites samples were taken in the shallow subtidal. Even though the degree of relatedness of individuals in a clump has not been assessed, in order to avoid possible effects of relatedness several clumps were obtained, all separated by at least 2 meters. Once in the fishing port, we randomly chose 40-45 clumps or isolated individuals in order to obtain one sample from each, and of those 40-45, we randomly picked 15 per site for this study. Mantle tissue (0.2 g) from each individual was used to extract DNA using the DNeasy Blood® & Tissue Kit (QIAGEN®, USA) according to the manufacturer's instructions. Quantity/purity of DNA was measured in Nanodrop 2,000 (Thermo, USA).

#### ***Genotyping by Sequencing and SNP discovery***

Samples were sequenced in the Biotechnology Center of the University of Wisconsin, USA using Genotyping by Sequencing (GBS)<sup>52</sup>, a method widely used in non-model species and that has proven to be a useful tool to distinguish between local populations at the genome level in marine benthic invertebrates<sup>20</sup>. This is a Reduced Representation Sequencing (RRS) method, which is based on the use of restriction enzymes that reduce the complexity of the genome, with the additional advantage that it can be performed without prior knowledge or reference genomes. This technique allows sequencing a large number of short genomic regions in several individuals<sup>45</sup>, permitting the genotyping of a large number of SNPs randomly located throughout the genome. Previous to sequencing, two additional enzymes were tested (*PstI*, *PstI/TaqI*) and finally *ApeKI*<sup>53</sup> was chosen due to the wide genome distribution of flanking regions described for other species and avoiding repetitive zones<sup>53</sup>.

Libraries were prepared for the DNA of 15 individuals of each population. Following the restriction enzyme digestion, DNA fragments were ligated to a unique barcode adaptor for each individual. The library was prepared in a 96-well plate with 6 wells as blank, using a Hiseq2000 (Illumina, USA) platform, which sequences reads of approximately 100 base pairs (bp). The reads were visualized in FastQC version 0.10.1<sup>54</sup> for quality checks.

All the data were prepared and analyzed using the pipeline Universal Network-Enabled Analysis Kit (UNEAK<sup>55</sup>) using TASSEL v.3<sup>56</sup> that is specially designed for species with no reference genome. TASSEL-UNEAK is a network based SNP detection algorithm that may be less flexible than other pipelines in aspects like read trimming and parameters for *de novo* locus identification<sup>57</sup>, reducing the potential number of total SNPs detected in the dataset. However, TASSEL-UNEAK has proven to be a useful, reliable and reproducible tool for demultiplexing and processing sequence data obtained through GBS for non-model species<sup>55</sup>. With the UNEAK pipeline, the dataset was demultiplexed and the reads were trimmed to 64 bp to remove the barcode sequence and the error-prone tail of the sequences. After the network filtering, identical reads were aligned as tags using an error tolerance rate of 0.03 in order to minimize the chance that real tags were discarded as sequencing errors and to remove potential paralogs before the SNP calling. This filtering is a goodness-of-fit  $\chi^2$  ( $\alpha=0.05$ ) based on the null hypothesis that the counts of two paired tags of a SNP are equal in all heterozygous individuals<sup>55</sup>.



Additionally, a minor allele frequency (MAF) of 0.044 was used to filter the number of SNP loci whose variability was well represented in the 90 analyzed individuals. This value was determined empirically based on the probability that at least 8 of the 180 possible allele calls (4.4% of the alleles detected in 90 individuals) correspond to the minor allele. Values of minimum call rate (mnC) from 75 to 90% were tested previous to the final SNP calling, finally choosing the most exigent value of mnC (90%) in order to remove all SNPs containing 90% or more “N” values, which designate those individuals for which no allele was assigned for the locus or that were covered by no more than one tag and because of the amount of missing data detected, preserving a reduced but robust set of SNPs with reliable genotypes without excluding individuals from the analysis. These restrictive filters were performed because all loci were sampled at a mean coverage of  $\geq 10x$  per individual. Under this condition, in theory each individual had in average 10 or more copies of a locus.

### ***Outlier detection***

To identify the SNPs putatively under diversifying selection we used the  $F_{ST}$  outlier approach, which is based on the estimation of those loci with greater values of  $F_{ST}$  (outliers) with respect to the expected values under a neutral distribution. We used the software BAYESCAN 2.1<sup>16</sup> for  $F_{ST}$  outlier detection, because it has been reported to have lower rates of false positives with respect to other similar software<sup>58</sup>. BAYESCAN uses a Bayesian framework to calculate the posterior probability that any

given locus is under selection<sup>16</sup>. A total of ten separate runs were performed, from 50,000 to 500,000 iterations with a 10% burn-in period to assure the convergence of the MCMC. After the runs, an FDR correction of *q-values* of 0.05 was applied in BAYESCAN to avoid the occurrence of false positives. The data set was then separated into two parts, one with the outlier SNPs, which are loci putatively under diversifying selection<sup>16</sup>, and one with neutral SNPs, which are all the rest of the loci that have  $F_{ST}$  values that match the expected under a neutral distribution.

#### ***Population genetic diversity analysis and genetic structure***

From here on, all analyses were performed separately for each of the data sets (neutral and outlier loci).

To evaluate the differences in genetic diversity across the study area in each data set, we calculated the expected heterozygosity, observed heterozygosity and allelic richness with rarefied allele counts, using the HIERFSTAT package version 0.04-22<sup>59</sup> in R v 3.22<sup>60</sup>. Exact tests for Hardy-Weinberg Equilibrium deviations for outlier loci were calculated using GENEPOP version 4.6<sup>61</sup>. Multiple comparisons of HWE were corrected using a FDR.

Pairwise  $F_{ST}$  and their significance were determined using a permutation test (10,000) in ARLEQUIN Version 3.5<sup>62</sup>. ARLEQUIN was also used to perform an Analysis of Molecular Variance (AMOVA) using the north/south of 30°S biogeographic boundary, and the separation of LM as two *a priori* groupings.

Additionally, a Mantel test with 10,000 permutations was carried out in ARLEQUIN 3.5 to test if there is a pattern of isolation by distance, using the relationship between the geographic distance (Km) and Slatkin's linearized genetic distance ( $F_{ST}$ ).

Spatial population genetic structure was assessed using the Bayesian clustering algorithm implemented in GENELAND v 3.1.4<sup>63</sup>. This analysis uses spatial coordinates to separate allele frequencies that can be detected with departure from Hardy-Weinberg and linkage equilibrium into  $K$  clusters. An independent model of allele frequencies was carried out for both data sets, with 10 independent runs with  $1e^6$  iterations and a 10% burn-in period, varying  $K$  from 2 to 6. The optimum  $K$  was inferred from the posterior individual membership probability of each genetic group. Posterior probabilities of membership were plotted with the *shapefiles* of the coastline available in the database of GEOdas (NOAA), filtering in the study area using GEOdas Coastline Extractor v 1.1.3.1 (<https://www.ngdc.noaa.gov/mgg/geodas/geodas.html>).

A clustering approach was used with a discriminant analysis of principal components (DAPC) in the package *adegenet*<sup>64</sup> in R v 3.22<sup>60</sup>. In DAPC, a principal components analysis of the multilocus genotypes of the individuals was calculated and then a discriminant analysis was carried out using the PC scores. These analyses are based on the detection of the number of clusters that minimizes between-group variation<sup>64</sup> using the k-means and a Bayesian Information Criterion (BIC) to identify the optimal number of clusters in the data, not assuming underlying structure in population subgroups or panmixia as other similar approaches do. The number of clusters in each data set was determined using  $10e^7$  iterations. To avoid unstable

assignments in each cluster, 30 PCs were retained ( $N_{\text{total}}/3$ ), using all the 6 discriminant functions. We determined the optimal number of PCs using 1,000 simulations to execute the final DAPC.

To infer SNP-based relatedness structure within populations, an analysis based on pairwise IBS (identity-by-state) was used. This approach uses the information of the genotypes to calculate the probability that two alleles in the same segment have the same ancestor. A cluster analysis of the matrix generated with the IBS pairwise coefficients between all possible pairs of individuals was performed using a permutation score of 10,000 to determine how many groups were present in each data set. This was performed in the package SNPRelate<sup>65</sup> in R environment v 3.22 (R Core Development Team 2015).

To determine the number of genetic groups we used STRUCTURE 2.3.4<sup>66</sup>. Clusters ( $K$ ) varied from one to seven, corresponding to the number of analyzed locations plus one. Ten replicate runs, each with 500,000 MCMC with a 10% burn-in period, were performed under the admixture and correlated frequency model. The most probable  $K$  value was inferred based on the delta  $K$  method proposed by Evanno *et al*<sup>67</sup>.

## References

1. Lipicius, R.N. *et al* Importance of metapopulation connectivity to restocking and restoration of marine species. *Rev. Fisheries. Sci.* **16**, 101-110 (2008).
2. Carson, H.S. *et al* Evaluating the importance of demographic connectivity in a marine metapopulation. *Ecology* **92**, 1972-1984 (2011).

3. Hedgecock, D. Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates? *Bull. Mar. Sci.* **39**, 550–64 (1986).
4. Cowen, R.K. *et al* Population connectivity in marine systems. *Oceanography* **20**, 14-21 (2007).
5. Marshall, D.J. *et al* Phenotype-environment mismatches reduce connectivity in the sea. *Ecol. Lett.* **13**, 128-140 (2010).
6. Pineda, J. *et al* Transport and dispersal in the coastal ocean and consequences for population connectivity. *Oceanography* **20**, 22-39 (2007).
7. Hamilton, S.L. *et al* Post settlement survival linked to larval life in a marine fish. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 1561-1566 (2008).
8. Selkoe, K.A. *et al* Seascape genetics and the spatial ecology of marine populations. *Fish. Fish.* **9**, 363-377 (2008).
9. Cowen, R.K. & Sponaugle, S. Larval dispersal and marine population connectivity. *Ann. Rev. Mar. Sci.* **1**, 443-466 (2009).
10. Pechenick, J.A. Larval experience and latent effects –metamorphosis is not a new beginning. *Integr. Compr. Biol.* **46**, 323-333 (2006).
11. Emllet, R.B. & Sadro, S.S. Linking stages of life history: how larval quality translates into juvenile performance for an intertidal barnacle (*Balanus glandula*). *Integr. Compr. Biol.* **46**, 334-346 (2006).
12. Penney, R.W. *et al* Genotype-dependent survival, growth, and reproduction in cultured blue mussels, *Mytilus* spp.: results of a reciprocal seed transfer experiment. *J. Shellfish. Res.* **25**, 515-525 (2006).

13. Pritchard, J.K. & Di Rienzo, A. Adaptation-not by sweeps alone. *Nat. Rev. Genet.* **11**, 665-667 (2010).
14. Nosil, P. *et al* Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* **18**, 375-402 (2009).
15. Schluter, D. & Conte, G.L. Genetics and ecological speciation. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 9955-9962 (2009).
16. Foll, M. & Gaggiotti, O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**, 977-993 (2008).
17. Moura, A.E. *et al* Population genomics of the killer whale indicates ecotype evolution in sympatry involving both selection and drift. *Mol. Ecol.* **23**, 5179-5192 (2014).
18. Candy, J. *et al* Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (*Thaleichthys pacificus*, Osmeridae), an anadromous Pacific smelt. *Mol. Ecol. Resour.* **15**, 1421-1434 (2015).
19. Araneda, C. *et al* Adaptive genetic variation distinguishes Chilean blue mussels (*Mytilus chilensis*) from different marine environments. *Ecol. Evol.* doi: 10.1002/ece3.2110 (2016).
20. Cahill, A. & Levinton, J.S. Genetic differentiation and reduced genetic diversity at the northern range edge of two species with different dispersal modes. *Mol. Ecol.* **25**, 515-526 (2016).
21. Xu, T. *et al* Genome-wide discovery of single nucleotide polymorphisms (SNPs) and single nucleotide variants (SNVs) in deep-sea mussels: Potential use in population

- genomics and cross-species application. *Deep-sea. Res. PT II*. doi: 10.1016/j.dsr2.2016.03.011 (2016).
22. Nayfa, M.G. & Zenger, M.R. Unravelling the effects of gene flow and selection in highly connected populations of the silver-lip pearl oyster (*Pinctada maxima*). *Mar. Genomics*. doi: 10.1016/j.margen.2016.02.005 (2016).
  23. Lal, M.M. *et al* A Parallel Population Genomic and Hydrodynamic Approach to Fishery Management of Highly-Dispersive Marine Invertebrates: The Case of the Fijian Black-Lip Pearl Oyster *Pinctada margaritifera*. *PLoS ONE* **11**, e0161390 (2016).
  24. Thiel, M. *et al* The Humboldt Current system of northern and central Chile oceanographic processes, ecological interactions and socioeconomic feedback. *Oceanogr. Mar. Biol. Annu. Rev.* **45**, 195-344 (2007).
  25. Sánchez, R. *et al* Spatial pattern of genetic and morphological diversity in the direct developer *Acanthina monodon* (Gastropoda: Mollusca). *Mar. Ecol. Prog. Ser.* **434**, 121–131 (2011).
  26. Haye, P.A. *et al* Phylogeographic Structure in Benthic Marine Invertebrates of the Southeast Pacific Coast of Chile with Differing Dispersal Potential. *PLoS ONE* **9**, e88613 (2014).
  27. Lancellotti, D. & Vásquez, J.A. Zoogeografía de macroinvertebrados bentónicos de la costa de Chile: contribución para la conservación marina. *Rev. Chil. Hist. Nat.* **73**, 99-129 (2000).

28. Manríquez P.H. & Castilla J.C. Self-fertilization as an alternative mode of fertilization in the solitary tunicate *Pyura chilensis*. *Mar. Ecol. Prog. Ser.* **305**, 113-125 (2005).
29. Cea, G. Contribución al conocimiento de algunos aspectos de la biología de *Pyura chilensis* Molina 1782 (Chordata, Tunicata, Ascidiacea). Tesis de Licenciatura en Biología, Universidad de Concepción, Concepción, Chile. 205 pp (1970).
30. Haye, P.A. & Munoz-Herrera, N.C. Isolation with differentiation followed by expansion with admixture in the tunicate *Pyura chilensis*. *BMC Evol. Biol.* **13**, 252 (2013).
31. Hudson, J. *et al* Anthropogenic transport of species across native ranges: unpredictable genetic and evolutionary consequences. *Biol. Lett.* **12**, doi: 10.1098/rsbl.2016.0620 (2016).
32. Ordoñez, V. *et al* Mixed but not admixed: a spatial analysis of genetic variation of an invasive ascidian on natural and artificial substrates. *Mar. Biol.* **160**, 1645-1660 (2013).
33. Pineda, M.C. *et al* The whereabouts of an ancient wanderer: global phylogeography of the solitary ascidian *Styela plicata*. *PLoS ONE* **6**, e25495 (2011).
34. Valdivia, N. *et al* Effects of disturbance on the diversity of hard-bottom macrobenthic communities on the coast of Chile. *Mar. Ecol. Prog. Ser.* **299**, 45-54 (2005).
35. Cifuentes, M. *et al* Effects of temporal variability of disturbance on the succession in marine fouling communities in northern-central Chile. *J. Exp. Mar. Biol. Ecol.* **352**, 280-294 (2007).
36. Conover, D.O. *et al* Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *J. Fish. Biol.* **69**, 21-47 (2006).



37. Liu, J. *et al* Cryptic diversity in the pen shell *Atrina pectinata* (Bivalvia: Pinnidae): high evidence and hybridization revealed by molecular and morphological data. *Mol. Ecol.* **20**, 4332-4345 (2011).
38. Conover, D.O. *et al* Reversal of evolutionary downsizing caused by selective harvest of large fish. *Proc. R. Soc. Long. Biol. Sci.* **276**, 2015–2020. (2009).
39. Hormazábal, S. *et al* Coastal transition zone off Chile. *J. Geophys. Res.* **109**, doi: 10.1029/2003JC001956 (2004).
40. Aravena, G. *et al* Twelve Years of Change in Coastal Upwelling along the Central-Northern Coast of Chile: Spatially Heterogeneous Responses to Climatic Variability. *PLoS ONE* **9**, e90276 (2014).
41. Astorga, M.O. & Ortiz, J.C. Genetic variability and population structure in the tunicate *Pyura chilensis* Molina, 1782, in the coast of Chile. *Rev. Chil. Hist. Nat.* **79**, 423–434 (2006).
42. Pineda, M.C. *et al* Stochasticity in space, persistence in time: genetic heterogeneity in harbour populations of the introduced ascidian *Styela plicata*. *PeerJ* **4**, e2158 (2016).
43. Ambler, R. & Cañete, J.I. Asentamiento y reclutamiento de *Pyura chilensis* Molina 1792 (Urochordata: Ascidiacea) sobre placas artificiales suspendidas en la Bahía La Herradura, Coquimbo, Chile. *Rev. Biol. Mar.* **26**, 403–413 (1991).
44. Bullard, S.G. & Carman, M.R. Current trends in invasive ascidian research. In: *Invasive Species: Detection, Impact and Control*. Editors C.P Wilcox & R.B Turpin. Nova Science Publishers, Inc (2009).
45. Manel, S. *et al* Genome Assemblies, Genomic Resources and their Influence on the

- Detection of the Signal of Positive Selection in Genome Scans. *Mol. Ecol.* **25**, 170-184 (2016).
46. Lotterhos, K.E. & Whitlock, M.C. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Mol. Ecol.* **24**, 1031–1046 (2015).
  47. Hoban, S. *et al* Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *Am. Nat.* **188**, doi: 10.1086/688018 (2016).
  48. Hassl, R.J. & Payseur, B.A. Fifteen years of genome-wide scans for selection: trends, lessons and unaddressed genetic sources of complication. *Mol. Ecol.* **25**, 5-23 (2016).
  49. Crawford, D.L & Oleksiak, M.F. Ecological population genomics in the marine environment. *Brief. Funct. Genomics* doi: 10.1093/bfgp/elw008 (2016).
  50. Drury, C. *et al* Genomic variation among populations of threatened coral: *Acropora cervicornis*. *BMC Genomics* doi: 10.1186/s12864-016-2583-8 (2016).
  51. Ravinet, M. *et al* Shared and non-shared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. *Mol. Ecol.* **25**, 287–305 (2016).
  52. De Donato, M. Genotyping-by-Sequencing (GBS): A Novel, Efficient and Cost-Effective Genotyping Method for Cattle Using Next-Generation Sequencing. *PLoS ONE* **8**, e62137 (2013).
  53. Elshire, R.J. *et al* A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* **6**, e19379 (2011).
  54. Andrews, S. FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc> (2010).



55. Lu, F. *et al* Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network- based SNP discovery protocol. *PLoS Genet.* **9**, e1003215 (2013).
56. Bradbury, P. J. *et al* TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* **23**, 2633–2635 (2007).
57. Andrews, K.R. *et al* Harnessing the power of RADseq for ecological and evolutionary genomics. *Nat. Rev. Genet.* **17**, 81-92 (2016).
58. Narum, S.R. & Hess, J.E. Comparison of FST outlier tests for SNP loci under selection. *Mol. Ecol. Resour.* **11**, 184–194 (2011).
59. Goudet, J. Hierfstat, a package for R to compute and test hierarchical F-statistics. *Mol. Ecol. Notes* **5**, 184–186 (2005).
60. R Development Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL <http://www.R-project.org>. (2015).
61. Rousset, F. Genepop'007: a complete reimplementaion of the Genepop software for Windows and Linux. *Mol. Ecol. Resour.* **8**, 103-106 (2008).
62. Excoffier, L. & Lischer, H.E. 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 (2010).
63. Guillot, G. *et al* 2005 A spatial statistical model for landscape genetics. *Genetics* **170**, 1261-1280 (2005).
64. Jombart, T. *et al* Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet.* **11**, 94 (2010).

65. Zheng, X. *et al* A High-performance Computing Toolset for Relatedness and Principal Component Analysis of SNP Data. *Bioinformatics* **28**, 3326-3328 (2012).
66. Pritchard, J.K. *et al* Inference of population structure using multilocus genotype data. *Genetics* **155**, 945–959 (2000).
67. Evanno, G. *et al* Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol. Ecol.* **14**, 2611–2620 (2005).

### **Acknowledgements**

Authors thank all fisherman that collected samples along the study area, as well as Natalia Muñoz, Francisca Gálvez and Raúl Vera for help with sample processing. This work was financially supported by Grant FONDECYT 1140862 (PAH & CGE), Grant FONDAP 15110027 (CGE & PAH). NIS was supported by the CONICYT Doctoral Thesis Scholarship 21120350.

### **AUTHOR CONTRIBUTIONS:**

PAH, CGE and NIS conceived the idea, designed the experiments, and performed the experiments. NIS and PAH analyzed the data and wrote the paper. NIS, PAH and EP contributed intellectually to the interpretation and discussion of results. All authors read and approved the final version of the manuscript.

## **ADDITIONAL INFORMATION**

### **COMPETING FINANCIAL INTERESTS**

The authors declare no competing financial interests

### **Data Accessibility**

SNP calls for all data set (2,332 SNPs) are available in a VCF file that has been deposited in DataDryad ([doi:10.5061/dryad.0s87v](https://doi.org/10.5061/dryad.0s87v)).

## Tables and Figures

**Table 1. Sampled local populations of *Pyura chilensis*.**

Sampling localities, coordinates, code, number of successfully analyzed individuals for genome scans for *Pyura chilensis* in the Southeast Pacific coast.

<b>Location</b>	<b>Coordinates</b>	<b>Code</b>	<b>N</b>
Pan de Azúcar	26°08'S; 70°39'W	PA	15
Caleta Pajonales	27°44'S; 71°02'W	CP	15
La Herradura	29°58'S; 71°21'W	LH	15
Talcahuano	36°38'S; 71°21'W	TH	15
Los Molinos	39°50'S; 73°23'W	LM	15
Ancud	41°52'S; 73°50'W	AC	15

**Table 2. Population genetic statistics.**

Proportion of polymorphic loci and allelic richness ( $A_r$ ) for neutral/outlier loci, observed ( $H_o$ ) and expected ( $H_e$ ) heterozygosity and deviations of Hardy-Weinberg Equilibrium (HWE) for outlier loci of *Pyura chilensis* in the Southeast Pacific coast. In bold, significant deviation of HWE after correction with a False Discovery Rate.

Site	Proportion of polymorphic loci		$A_r$		$H_o$		$H_e$		HWE (Outlier)			
	Neutral	Outlier	Neutral	Outlier	Neutral	Outlier	Neutral	Outlier	$\chi^2$	df	P-value	
PA	0.2	0.23	1.560 0.43	± 0.46	1.639 0.46	± 0.21	0.271 0.21	0.306 ± 0.20	0.294 ± 0.16	89.8	96	0.66
CP	0.59	0.59	1.579 0.42	± 0.46	1.619 0.46	± 0.23	0.304 0.23	0.318 ± 0.20	0.295 ± 0.16	118.1 7	90	0.02
LH	0.48	0.65	1.581 0.42	± 0.47	1.511 0.47	± 0.23	0.305 0.23	0.319 ± 0.26	0.291 ± 0.16	89.46	74	0.11
TH	0.37	0.4	1.605 0.40	± 0.45	1.666 0.45	± 0.22	0.285 0.22	0.334 ± 0.19	0.290 ± 0.16	69.14	102	0.99
LM	0.46	0.44	1.662 0.39	± 0.46	1.554 0.46	± 0.18	0.169 0.18	0.143 ± 0.15	0.288 ± 0.15	197.4 1	76	<b>&lt;0.001</b>
AC	0.57	0.61	1.604 0.41	± 0.46	1.626 0.46	± 0.23	0.298 0.23	0.345 ± 0.21	0.289 ± 0.16	57.08	92	0.998

**Table 3. Population-pairwise  $F_{ST}$  values for neutral and outlier SNP loci in *Pyura chilensis*.**

Genetic differentiation measures of pairwise comparisons calculated using 2251 neutral SNP loci (above the diagonal) and 81 outlier SNP loci (below the diagonal) in *Pyura chilensis*. Significant values after 10,000 permutations ( $P < 0.05$ ) are in bold.

	PA	CP	LH	TH	LM	AC
PA		0.003	<b>0.012</b>	<b>0.038</b>	<b>0.414</b>	<b>0.045</b>
CP	<b>0.027</b>		<b>0.015</b>	<b>0.038</b>	<b>0.417</b>	<b>0.047</b>
LH	<b>0.183</b>	<b>0.093</b>		<b>0.045</b>	<b>0.413</b>	<b>0.049</b>
TH	<b>0.524</b>	<b>0.548</b>	<b>0.643</b>		<b>0.399</b>	<b>0.011</b>
LM	<b>0.587</b>	<b>0.609</b>	<b>0.692</b>	<b>0.457</b>		<b>0.393</b>
AC	<b>0.564</b>	<b>0.585</b>	<b>0.672</b>	<b>0.142</b>	<b>0.376</b>	



**Table 4. Analysis of Molecular Variance (AMOVA)**

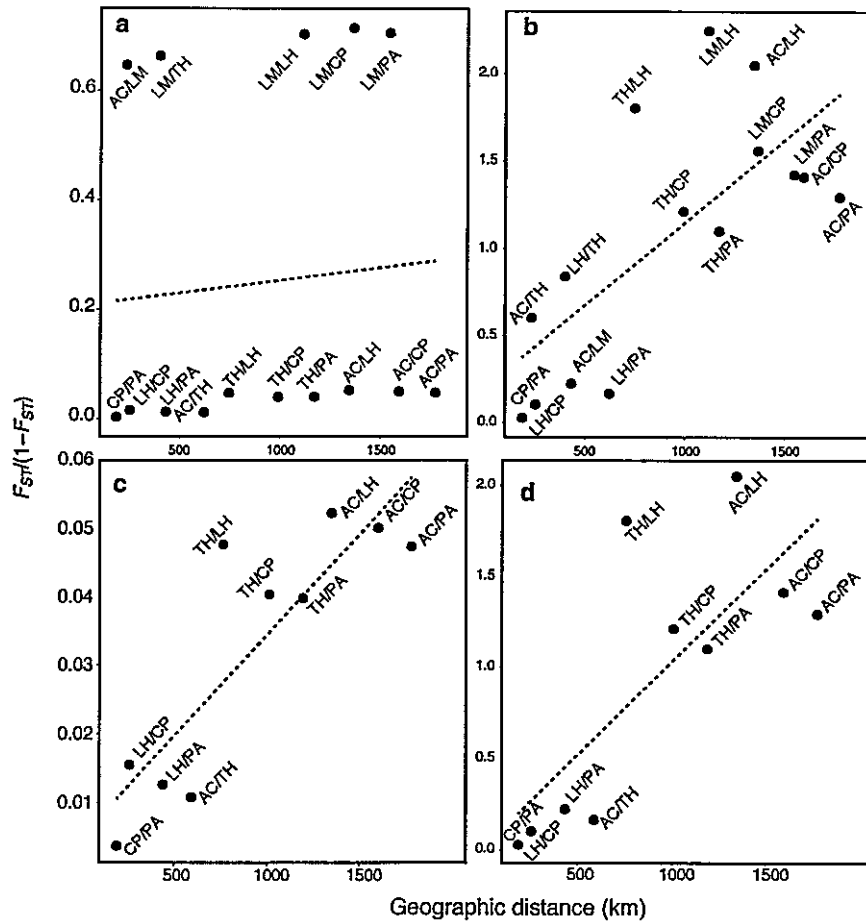
Percentage of variation explained by different hierarchical levels for neutral and outlier loci considering two groupings in *Pyura chilensis*. One separating North and South of 30°S (N/S) and the other one separating Los Molinos (LM) from the rest of the populations. Significant values in bold.

Source of variation	Neutral Loci		Outlier Loci	
	N/S	LM	N/S	LM
Among groups	2.91	<b>39.51</b>	<b>48.73</b>	8.74
Among populations within groups	<b>18.4</b>	<b>1.84</b>	<b>12.51</b>	<b>45.47</b>
Within populations	<b>78.69</b>	<b>58.65</b>	<b>38.77</b>	<b>45.79</b>

**Table S1: Mitochondrial haplogroup assignment**

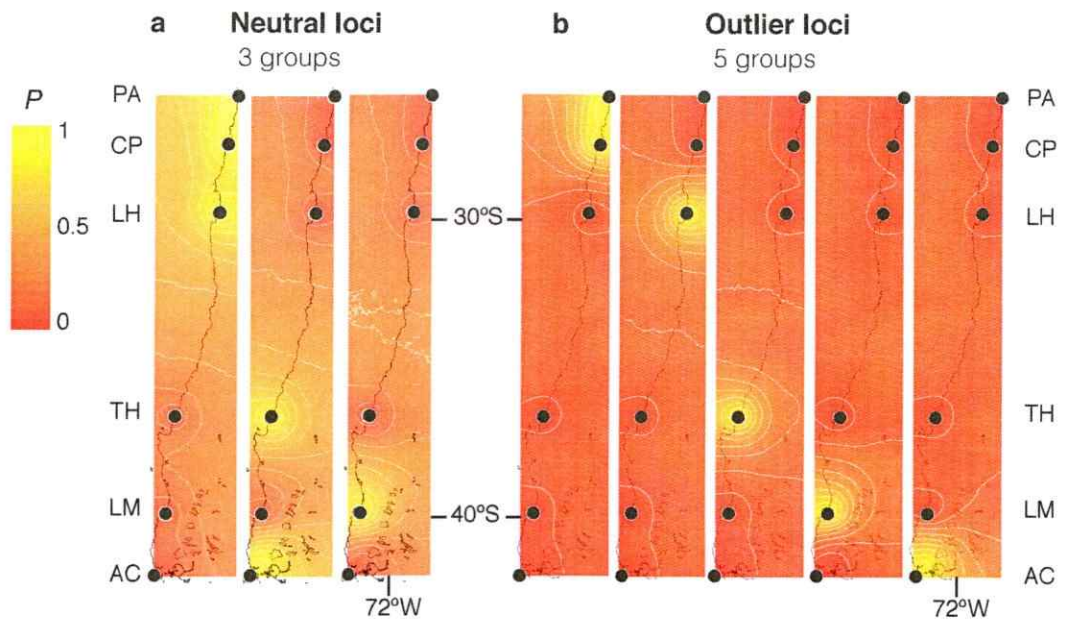
COI haplogroup assignment of each individual of *Pyura chilensis* analyzed with GBS data according to the haplogroups detected by Haye & Muñoz-Herrera<sup>30</sup>.

Site	Haplogroup 1	Haplogroup 2	Haplogroup 3
PA	15	0	0
CP	14	1	0
LH	8	7	0
TH	6	4	0
LM	0	3	12
AC	11	2	2



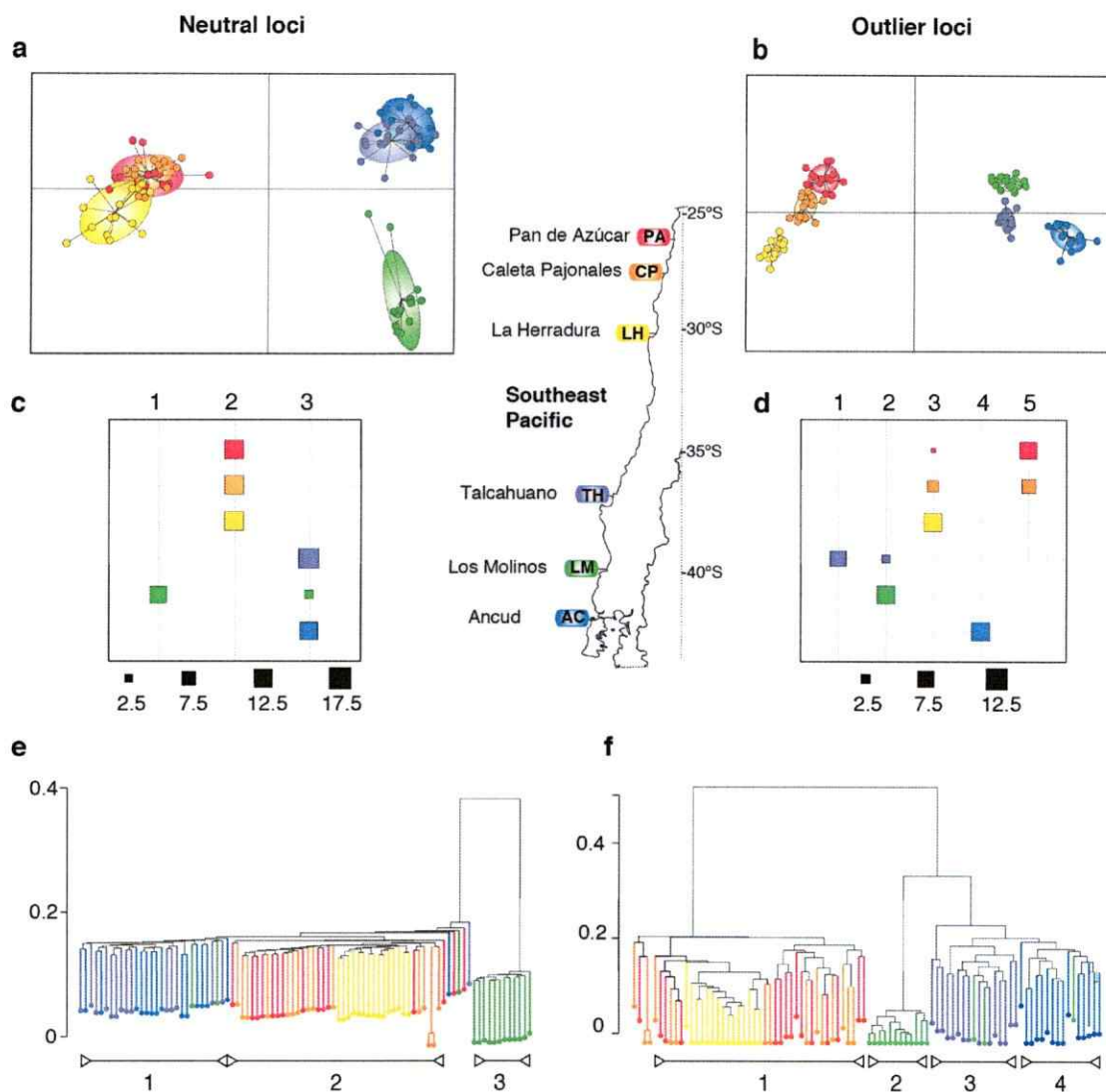
**Figure 1. Relationship between genetic differentiation and geographic distance for neutral and outlier loci of *Pyura chilensis*.**

Relationship between linearized genetic differentiation ( $F_{ST}$ ) and geographic distance (Km) along the Humboldt Current System. (a) Relationship using neutral markers, (b) for outlier markers and (c) and (d) show the relationship of neutral and outlier makers, respectively, excluding LM from the analyses. All the pairwise comparisons are marked in the graph.



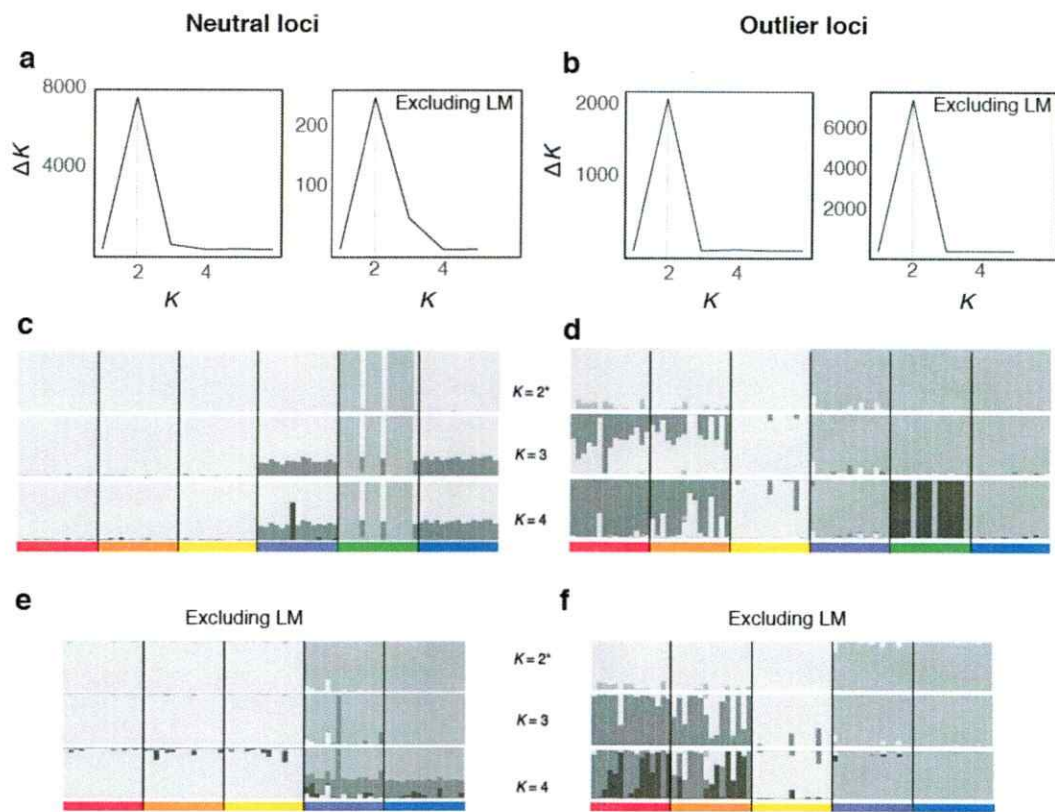
**Figure 2.** Spatial genetic structure of *Pyura chilensis* based on GENELAND analyses.

Each map indicates the posterior probability of belonging to one of the three groups found in neutral loci, or the five groups found in outlier loci. Black dots represent the coordinates of the sample sites, and the x and y-axes correspond to geographic coordinates.



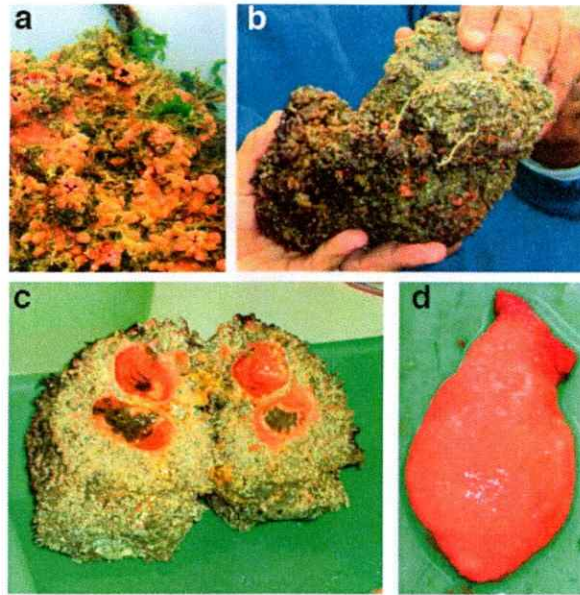
**Figure 3. Genetic Clustering of neutral and outlier SNP loci for *Pyura chilensis*.**

(a-b) Discriminant Analyses of Principal Components (DAPC) scatterplot showing the first 2 principal components for  $K = 3$  and  $K = 5$  (neutral and outliers, respectively). Colors represent the sampling location of each individual. (c-d) Proportion of individuals that belong to each cluster based on the DAPC analysis using a BIC criterion with geographic information, (e-f) Identity-by-state (IBS) pairwise relatedness dendrogram for all possible pairs of individuals. The dendrogram shows the three clusters found for neutral markers and the four clusters for outlier markers.



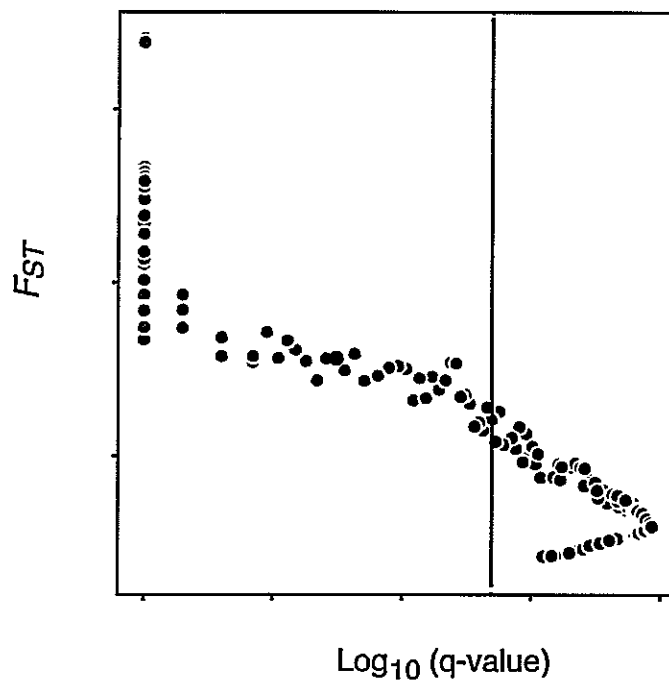
**Figure 4.** Bayesian clustering inferred with STRUCTURE for *Pyura chilensis*.

(a-b) Delta  $K$  results for neutral and outlier loci including and excluding Los Molinos (LM) to infer optimal separations according to Evanno *et al*<sup>67</sup>, (c-f) Bayesian assignment probabilities of individuals of inferred with STRUCTURE with the entire dataset and excluding LM from the analyses. Each column represents one individual, with the color representing the membership proportion of each of the clusters. The vertical line represents the delimitation of each sampled local population. Graphs show the results for  $K = 2-4$ . The optimum  $K$  was denoted by an asterisk.



**Figure S1.** *Pyura chilensis*.

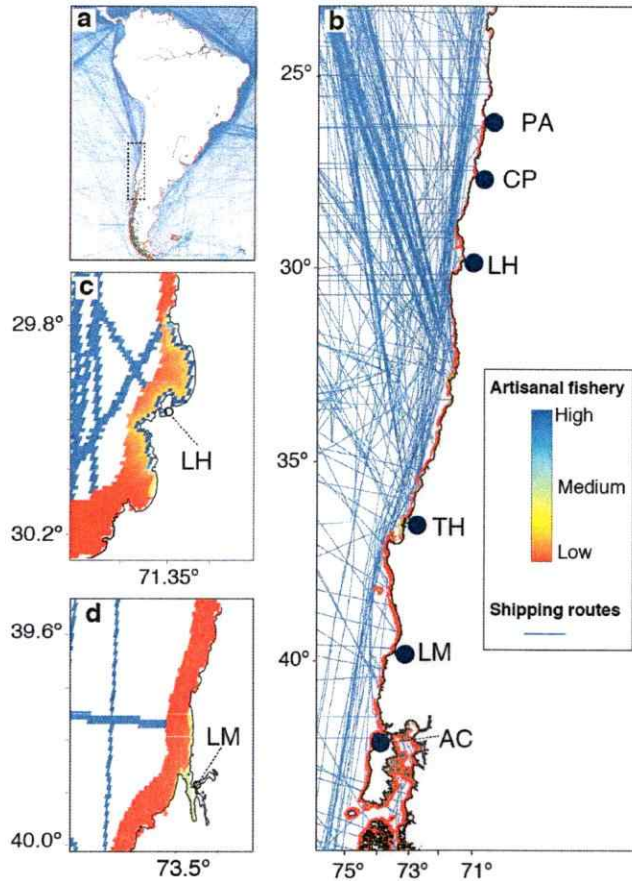
Clump of individuals of *P. chilensis* (a-b), and opened clump with two individuals (c), and an individual after extracting it from the tunic (d), showing its red mantle.



**Figure S2. SNP loci  $F_{ST}$  outlier analysis for *Pyura chilensis*.**

Results of Bayescan analysis to identify SNPs with outlier  $F_{ST}$  values for 2332 SNPs. The X-axis is the log of the posterior probabilities (q-value) and the solid line represents the threshold for posterior probabilities using a False Discovery Rate (FDR) of 0.05. Values of  $q < 0.05$  are considered as neutral markers, and  $q > 0.05$  as outlier markers.





**Figure S3. Shipping lines and artisanal fishing routes for the Study Area.**

Maps for marine commercial activity (blue lines) and artisanal fishery modeling (red to blue pixels) were obtained and modified from the Global Map of Cumulative Human Impact project data, available in <https://www.nceas.ucsb.edu/globalmarine/data> (Halpern *et al* 2015) and plotted in ArcMap 10.4 (ESRI 2011) with the base maps of the coastline of GEODAS, NG, NOAA (<https://www.ngdc.noaa.gov/mgg/shorelines/>), using Coastline Extractor Software to filter the study area (<https://www.ngdc.noaa.gov/mgg/geodas/geodas.html>). Cumulative influence of commercial and artisanal fishing routes in: (a) South America (data from 2008-2013), (b) detail of the study area, (c) close up of La Herradura (LH) site, and (d) close up of the Los Molinos (LM) site. Scale 1 km<sup>2</sup> per pixel.

- ESRI 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.
- Halpern, B. S. et al Spatial and temporal changes in cumulative human impacts on the world's ocean. *Nat. Commun.* 6:7615 doi: 10.1038/ncomms8615 (2015).

## CAPÍTULO II: Interannual variation of COI haplotypes and haplogroups of the ascidian *Pyura chilensis* in the Southeast Pacific coast

Pilar A. Haye<sup>1,2</sup>, Nicolás I. Segovia<sup>1,2</sup>

<sup>1</sup> Laboratorio de Diversidad Molecular, Departamento de Biología Marina, Facultad de Ciencias del Mar, Universidad Católica del Norte, Coquimbo, Chile

<sup>2</sup> Interdisciplinary Center for Aquaculture Research (INCAR), Universidad de Concepción, Concepción, Chile

Corresponding Author:

Pilar Haye<sup>1</sup>

Departamento de Biología Marina, Universidad Católica del Norte, Larrondo 1281, Coquimbo, Chile

Email address: [phaye@ucn.cl](mailto:phaye@ucn.cl)

Formato: PeerJ

## Abstract

Space and temporal variation of the environment could affect dispersal, recruitment and populations persistence of marine benthic species. Genetic studies including inter-annual comparisons in the genetic structure indicate both high and moderate temporal heterogeneity in marine invertebrates, suggesting a prevailing pattern. This suggests that temporal studies are necessary to understand the dynamics of marine metapopulations. In this study, we compared the spatial/temporal genetic structure of the ascidia *Pyura chilensis*, a low dispersal sessile marine benthic species endemic from the South East Pacific coast and highly demanded for human consumption. We used 592 sequences of the mitochondrial gen COI in six sites from 24° to 42°S sampled in three years (2012, 2014, 2015). Spatial structure of COI indicates the presence of three monophyletic lineages previously described in the species, being one of them highly divergent and geographically restricted (~39°S, Los Molinos). Inter-annual comparisons showed temporal stability in the general patterns of genetic structure of COI in *Pyura chilensis*, however, slightly variations were detected in both, specific geographic locations compared between years and analyzing each lineage per separate. Our results suggest temporal genetic homogeneity, which is expected for short-lived planktonic larval duration species.

## INTRODUCTION

Marine benthic coastal species inhabit an environment that is not only variable geographically but also temporarily (Menge et al. 2003). Heterogeneous marine environmental factors vary in space and also in time, exerting a strong influence on the dispersal, recruitment and population persistence of marine benthic species (Siegel et al. 2008, Marshall et al. 2010). Yet, how much does genetic structure vary in space versus time in marine benthic species?

Phylogeographic studies of marine benthic invertebrates have been highly focused on the spatial structure genetic structure using one or more molecular markers to infer genetic structure and with a skewed selection towards mitochondrial COI gene sequences (e.g. Lessios et al. 2001; Wares & Cunningham 2001; Barber et al. 2002; Bowen et al. 2013). This gene has been highly used because it is inherited as a maternal lineage, lacks recombination (in most taxa), can be sequenced using universal primers, and because it allows comparison with many other studies (e.g Kelly & Palumbi 2010, Haye et al. 2014). Additionally, COI mutational rates are often known allowing inferences of past demographic or divergence events. Phylogeographic patterns deduced with COI sequences among other markers such as microsatellite loci, have shown that benthic marine species harbor internal genetic structure in their range of distribution, with varying degrees of differentiation related to dispersal potential (e.g., Wilke & Davis 2000;

Weersing & Toonen, 2009; Haye et al. 2014; Villamor et al. 2014), coastal topography and oceanographic features (Kirkandale & Meyer, 2004; Collins et al. 2010, Barshis et al. 2011; Veale & Lavery 2011), biogeographic breaks (Goldstein et al. 2006; Zakas et al. 2009, Reuschel et al. 2010; Haye et al. 2014), colonization (e.g. Wares & Cunningham 2001; Wangensteen et al. 2012) or invasion/introduction history (Astanei et al. 2005; Lejeusne et al. 2011; Bouchemousse et al. 2016), effects of past climatic conditions on perish or persistence (existence of refugia) of local populations (Crandall et al. 2008; Provan & Bennett 2008; Haye & Muñoz-Herrera 2013), and cryptic diversity within lineages (Doellman et al. 2011, Rius & Teske 2011, Varela & Haye 2012). Genetic studies of marine benthic species that include inter-annual comparisons indicate that variation in the genetic structure of populations is a prevailing pattern in the highly active marine realm. Variability in the genetic structure of populations has been reported for a polychaete (Jolly et al. 2003), for a limpet (Johnson & Black 1984) and for a surf clam (David et al. 1997). Other studies have also shown evidence of high and moderate temporal heterogeneity in marine invertebrates (Colgan 1981; Hedgecock 1994; Li & Hedgecock 1998; Johnson & Wernham 1999; Moberg & Burton 2000; Lambert et al. 2003; Virgilio & Abbiati 2006; Barshis et al. 2011).

Metapopulations have been considered as dynamic systems of populations that are interconnected by reproduction (Hanski & Gilpin 1997).

Temporal studies are necessary to better understand the dynamics of local populations of marine metapopulations and how genetic diversity is affected by reproductive dynamics of local populations (Crowder et al. 2000; Cobben et al. 2011; Ludford et al. 2012). The degree of reproductive connectivity between populations is variable and stochastic in nature (Johnson 2004; Siegel et al. 2008), between highly connected populations, to populations that show variable degrees of isolation. In addition to dispersal between populations, other processes can promote or suppress local larval retention. Differential settlement and survival of arriving larvae has been highly documented in marine organisms (e.g., Pedersen et al. 2000; Berger et al. 2006; Penney et al. 2006; Vigliola et al. 2007). Larvae and recruits can also be strongly affected by oceanographic factors such as coastal circulation and upwelling (Roughgarden et al. 1988, Shanks & Brink 2005; Barshis et al. 2011).

Populations that undergo significant genetic variation in time are likely strongly affected by local conditions that preclude self-replenishment of populations. These areas can be considered sinks because they mostly depend on the immigration of propagules that were produced in source areas for their persistence in time (Pulliam 1988; Lipcius et al. 2008; Peery et al. 2008). Local populations that remain genetically stable over time are likely to have a net reproductive output greater than zero (self-sustained) and act as source areas that export propagules to other local populations.

Herein we studied the degree of temporal genetic variation of local populations of the exploited acidian *Pyura chilensis* in the Chilean coast between 24 and 42°S, covering the southern portion of the geographic range of distribution (from 10 to 44°S). Spatial structure of *P. chilensis* was studied by Haye & Muñoz-Herrera (2013) along the same geographic area and they showed that between genetic diversity of COI was structured in three lineages or haplogroups, two widely distributed and one restricted geographically restricted to southern area (39°-42°S). The frequency of each haplogroup showed strong geographic variation. On the northernmost location only one haplogroup was present, the most widespread one. In the center localities (between 26°S and 36°S) two haplogroups were present, the two most wide-spread. Finally, in the southern localities the two haplotypes from the center were present together with a third haplogroup. This last, southern haplogroup was particularly abundant in Los Molinos (39°S). The nuclear gene Elongation Factor 1 alpha (EF1a) did not show the same degree of structure as COI. Only some of the individuals of the more restricted third COI haplogroup were divergent with EL1a, suggesting that the three haplogroups are not reproductively isolated. The authors suggested that these haplogroups were formed during Pleistocene glacial cycles as a consequence of isolation during glaciations when sea-level falls. Following, in an interglacial, populations could interbreed and for a single species that has three divergent mitochondrial, yet not nuclear, lineages.

We chose to study interannual temporal variation using the gene COI as a first approach given that it has been the main marker used in marine phylogeography providing ample opportunity for comparison representing a good marker to reconcile data on temporal genetic variability with previous studies.

Does the genetic structure of *P. chilensis* vary between years? Are there populations that differ in their degree of temporal variation? Can areas with high temporal variation be detected (putative sinks)? To explore answers to these questions we obtained sequences of COI from six local populations of *P. chilensis* in three time points (2012, 2014 and 2015). The data was used to compare inter-annually the geographic distribution of the genetic diversity through estimations and comparison of the levels of genetic diversity, and genetic differentiation.

## **MATERIALS AND METHODS**

Samples were collected by local fishermen from the Fisheries Management Areas under their jurisdiction in six localities along the coast of Chile from Pan de Azúcar (26°08'S) to Ancud (41°52'S) (Table 1). Sampling locations were selected based on results of Haye & Muñoz-Herrera (2013), and include populations with different proportions of the three mitochondrial haplogroups of *P. chilensis*. Collections were performed in 2014 and 2015, and together with data available from samples collected in 2012 (see Haye &



Muñoz-Herrera 2013), lead to three time periods analyzed for comparison of inter-annual genetic structure.

Each year, sampled individuals had a maximum size of 5 cm in order to avoid overlapping generations between samples of different years (Cea 1973, Svane & Lundälv 1982, Valdivia et al 2005) and all individuals were obtained from clumps separated by at least 2 meters.

DNA extractions, PCR of mitochondrial *COI* gene, sequencing and alignment of sequences were performed following Haye & Muñoz-Herrera (2013).

Additionally, since *P. chilensis* is constituted by three COI haplogroups, analyses were performed both including all individuals and separated per haplogroup. The software DnaSP 5.10 (Librado & Rozas 2009) was used to determine average number of differences between haplogroups, frequency of haplogroups per location, genetic diversity indices (haplotypic and nucleotidic), and nearest-neighbor statistics  $S_{nn}$  values (Hudson 2005).

### **Data Analyses**

Non-random associations between haplogroups frequencies between pairwise temporal samples were carried out per location using a Fisher's exact test in R environment v 3.3.3 (R Development core team 2017). To compare both spatial and temporal variation in the genetic diversity, a nested

One-way ANOVA were performed on diversity indices (haplotype diversity, nucleotide diversity) alternating fixed and random effects. To evaluate spatial variation in genetic diversity, sampling site was used as fixed effect and year as random effect, and inversely, to evaluate temporal variation, year were used as fixed and location as random effect. All the analysis was carried out in R v 3.33 using all data set and haplogroups per separate.

Haplotype networks were constructed in the software PopArt v1.7 (<http://popart.otago.ac.nz>) for each sampling site including all temporal samples.

Population-pairwise and temporal population-pairwise  $\Phi_{ST}$  ( $F_{ST}$  analogue that accounts for haplotype frequency and divergence) was estimated with 10000 permutations (Weir & Cockerham 1984) using Arlequin 3.5 (Excoffier & Lischer 2010). Additionally, a single FASTA file was constructed for all years together in order to calculate the population-pairwise  $\Phi_{ST}$  not only between local populations but also between years in the same geographical location.

To examine temporal genetic structure, an ANOVA analysis was used setting «years» instead of geographic structure as *a priori* groups in Arlequin 3.5 considering all and per haplogroups data set. Mantel correlations were performed between population-pairwise  $F_{ST}$  for pairwise temporal samples in order to compare inter-annual matrices of genetic differentiation, assuming that temporal stability means higher correlations between pairwise  $F_{ST}$

matrices between years. Mantel tests were carried out with the package Vegan (Oksanen et al. 2017) in R environment (R Development Core Team, 2017) using all the data set and the three haplogroup pairwise differentiation matrices per separated.

## RESULTS

COI sequences were successfully obtained for a total of 592 individuals of *Pyura chilensis* from six localities (Table 1) and three time periods (2012, 2014 and 2015). Three haplogroups that had been detected in a previous study of *P. chilensis* based on COI (Haye & Muñoz-Herrera 2013) were found to have a temporally stable geographic distribution (Fig. 1A). Haplogroup 1 (Hg1) was found on all sampled locations, and had the widest distributional range of all haplogroups; Hg 2 was found from Caleta Pajonales (CP) to the south, only absent in the northern location. Pan de Azúcar (PA, 26°S). As expected, Hg3 was most restricted haplogroup, present only in the two southern locations and mainly in Los Molinos (LM) and at a much lower frequency in Ancud (AC).

Haplogroup frequencies per location showed slightly variation between years (Fig. 1A-B). Overall there was a decrease in the frequency of Hg1, constituting 67.5% of the analyzed sequences in 2012, 62.2% in 2014 and 59.6% in 2015. Comparing 2012 and 2015 both Hg2 and Hg3 showed an

increase in frequency. Hg2 increased from 22.7% to 25.2% and Hg3 from 9.9% to 15% in 2015.

Site-by-site haplogroup frequency variation was most notable in La Herradura (LH) and Talcahuano (TH) with an increase of the proportion of Hg2 relative to Hg1 and in LM, that showed an important increase on the frequency of Hg3 from 2012 to 2014 which was maintained in 2015 (Fig 1A). Overall, both within haplogroups and in the total data set, the change in frequency was greater between 2012 and 2014 than between 2014 and 2015 (Fig. 1A). Fisher's exact probability test detected a non-random association between haplogroup frequencies only for LM and when comparing year 2012 with the other two ( $p = 0.0008$  and  $p = 0.032$ , respectively, Table 2).

Nucleotide and haplotype diversities presented some variation between years (Table S1 Supplement). However, One-way nested ANOVA of nucleotide and haplotype diversities resulted in significant F-values only for all dataset and for Hg1 and when the location was considered a fixed effect and the sampling year a random effect (Table 3, Fig 1 B) being non-significant for all comparisons using year as a fixed effect, suggesting temporal stability in genetic diversity.

Number of unique haplotypes were overall temporally stable. The number of unique haplotypes is lower in AC and higher in PA and LM (Fig 3). The number of unique haplotypes was slightly variable in LH and TH. The ratio between total haplotypes and unique haplotypes differed in AC,

while all other locations showed little variation in the ratio (Fig 3). In 2015 AC had proportionally less unique haplotypes than in previous time periods sampled

Haplotype networks for each site including samples from the three sampling times allowed the evaluation of the relationship between haplotypes detected in each sampling. The networks showed that haplotypes sampled in 2012 and 2014 are similar while those from 2015 often differed from those of 2012 and 2014 (Fig 2). Similarly, haplotype sharing was common between 2012 and 2014 and very uncommon for any comparison with 2015.

Most population-pairwise  $\Phi_{ST}$  values calculated separately for each year, were significant when analyzing all data together (Table S2 Supplement). Values for Hg1 had higher significance in 2012. In 2014 and 2015 most values were non-significant excepting most comparisons including AC. For Hg2, most comparisons are non-significant; the only significant values were from 2015 and from TH. Hg3, present in only the two southernmost locations, AC and LM, was characterized by one non-significant  $\Phi_{ST}$  each year (Table S2 Supplement).

Temporal population-pairwise  $\Phi_{ST}$  was used to reveal significant genetic differentiation between years for each location.  $\Phi_{ST}$  values were generally low. For all the dataset, significant  $\Phi_{ST}$  values were all associated to comparisons of 2015 with 2012 and 2014; only the local populations of PA

and CP did not have significant  $\Phi_{ST}$  values in any of the time-pairwise comparisons. Considering all data together the time period 2015 was significantly differentiated from 2012 and 2014 in all sampling sites excepting the two northernmost, PA and CP for haplogroup 1 (Table 4) as was shown by the haplotype networks (Fig 2). Hg2 had a similar tendency, with most 2015 comparisons yielding significant  $\Phi_{ST}$  values from LH to the south (keeping in mind that Hg2 was not detected in PA in any of the sampling times). Hg3 did not display significant differentiation based on  $\Phi_{ST}$  values for any of the time comparisons (Table 4).

AMOVA grouping sequences by years provided additional support for temporal stability. Considering all the dataset AMOVA between years variation among years was non-significant while variation among populations and within populations was significant (Table 5). Analyzing haplogroups separately showed that only Hg2 had significant variation between years and all had significant variation within populations (Table 5), suggesting that residual variation (non-explained by inter-annual variation) explained virtually all the molecular variance.

Mantel correlations between population-pairwise  $\Phi_{ST}$  for pairs of temporal samples were significant for all comparisons (2012/2014, 2012/2015, 2014/2015) when all data is analyzed together, while analyses per haplogroup gave mostly non-significant correlations (Fig 3). Here, the points in the upper right panel of the dot cloud of the correlation

corresponded to 5 points of each year analyzed, all of which involve comparisons with haplogroup 3, i.e. with LM (Figure 4a). Significant and higher mantel coefficient correlations was expected for situations in which population-pairwise values were similar between years.

## DISCUSSION

### Temporal genetic variation in *Pyura chilensis*

Genetic structure of *Pyura chilensis* has a greater spatial than temporal variation. Spatial genetic structure of COI of *P. chilensis* consisted of three COI haplogroups, one present throughout the analyzed geographic area, a second was only missing on the northernmost site and a third one restricted to the two southern sites, as shown previously by Haye & Muñoz-Herrera (2013). The proportion of haplogroups of COI of *P. chilensis* did not vary between years 2012, 2014 and 2015. Some temporal genetic variation was detected for COI haplogroup frequencies in the site LM, which harbors most individuals of the geographically restricted COI haplogroup mentioned above. Even though haplogroup frequency had low temporal variation, there are significant signals of variation in the genetic structure within haplogroups at the haplotype level, as revealed on the one hand by the significant differences in nucleotidic and haplotypic diversities within the most widely distributed haplogroup (Hg1), and on the other hand by the

significant degree of population genetic differentiation for many population pairs for the two most widely distributed haplogroups (Hg1 and Hg2). Most significant comparisons involved samples of 2015. Consistently, haplotype networks evidenced that in 2015 a different suite of haplotypes was encountered in all haplogroups. In the same way, one of the haplogroups of *P. chilensis*, Hg1, showed significant temporal variation with different analytical approaches, especially when year 2015 was compared. All together COI haplotype data revealed low to moderate temporal genetic differentiation of the gene COI.

Spatial variation of COI haplogroups and haplotypes in *P. chilensis* was greater than temporal variation, with an overall low to moderate temporal variation. This is also the case for other ascidia of similar biological features like *Styela plicata* (Pineda et al. 2016), as well as the polychaete *Pygospio elegans* (Kesäniemi et al. 2014); the molluscs *Cellana grata* (Ng et al. 2010), *Mactromeris polunyma* (Cassista & Hart 2007), *Placopecten magallanicus* (Kenchington et al. 2006), and direct developing *Littorina* spp. (Lee & Boulding 2009); the shrimp *Litopenaeus setiferus* (Ball & Chapman 2003); and the urchins *Paracentrotus lividus* (Calderón et al. 2009) and *Strogylocentrotus purpuratus* (Flowers et al. 2002). In opposition, temporal variation is greater than spatial variation in some marine invertebrates such as the barnacle *Balanus glandula* (Barshis et al. 2011), the oyster *Ostrea edulis* (Hedgecock et al. 2007), periwinkles



*Littorina* spp. with planktonic larval development (Lee & Boulding 2009), and the anomuran crab *Petrolisthes cinctipes* (Toonen & Grosberg 2011). For some systems, in the one hand it has been pointed that high variance in reproductive success could explain temporal variability in genetic diversity and structure (Turner et al. 2002; Pujolar et al. 2007), and in other other hand, reproductive variance can lead to a temporal chaotic genetic patchiness (Colgan 1981; Hedgecock 1994; Johnson & Black 1994).

#### **Dispersal potential and temporal genetic variation and mechanisms behind temporal genetic variation**

Temporal genetic variation may be determined by several factors. In marine invertebrates, the dispersal potential of species, generally determined by the duration of the planktonic larval life stages, seems to be a main factor contributing to temporal genetic heterogeneity.

For example, species of the genus *Littorina* vary in their degree of temporal variation in accordance with their dispersal potential; species with planktonic larval dispersal showed high temporal variation while species with direct-development showed low temporal variation (Lee & Boulding 2009). Most species with high temporal genetic structure have long-lived planktonic larval dispersal stages while species with low or moderate temporal variation have varying degrees of dispersal potential by planktonic larvae with most taxa exhibiting moderate to low dispersal potential. The moderate to low

temporal variation detected is consistent with the short-lived planktonic larval duration of *P. chilensis*.

Two main natural mechanisms that can explain the differentiation between temporal samples are natural selection with a local and temporary variable effect on larvae (Watts et al. 1990), and sweepstakes reproduction success (Calderón et al. 2009, Kesaniemi et al. 2014). Comparison of temporal studies with respect to molecular marker used indicate that generally studies based on allozyme loci show more temporal genetic structure than studies with putatively neutral molecular markers such as microsatellites or COI sequences. Lessios et al. 1994 using allozyme loci detected that populations of the isopod *Excirolana braziliensis* from Central America remained stable over time. The conservatism of allozyme loci together with the low dispersal potential of *Excirolana* that is a brooder species may account for this result. Along this line, Lee & Boulding (2009) analyzed genetic variation at mitochondrial and nuclear markers between 10 years in four marine snails of the genus *Littorina* and detected greatest variation in the two species that had planktonic larval dispersal while the two species of brooders had less temporal variation. Although at this point it seems that species with low dispersal capabilities tend to have low temporal variation.

Both mechanisms depend on the environmental characteristics and intrinsic characteristics of species. The marker used COI has been shown to



be neutral to natural selection in *P. chilensis* (Haye & Muñoz-Herrera 2013) and thus selection can be discarded as the main mechanism behind the (moderate) temporal variation detected.

The marine environment is highly unpredictable and along the southeast Pacific especially since there occur a series of common processes that have temporal variation such as seasonal winds and upwelling, and the strong influence of the El Niño Southern Oscillation, among others (Thiel et al. 2007). The species *Pyura chilensis* is a strong competitor for space but has a relatively low colonizing ability because it is highly susceptible to disturbance and recruitment is only high when the frequency of disturbance is low (Valdivia et al. 2005). Our results, based on genetic temporal variation also suggest that recruitment is variable and/or natural selection or other post-recruitment processes are associated to massive mortality of recruits and/or juveniles. In the barnacle *Balanus glandula*, Berger et al. (2006) described high variation in juvenile performance in space and time depending on environmental conditions. For the same species, the strength of coastal upwelling was linked to temporal variation in genetic structure (Barshis et al. 2011). The same mechanism could explain the temporal genetic variation detected in *P. chilensis*.

The sweepstakes reproductive success hypothesis establishes that which and how many adults contribute to the demographic continuity of a marine species is a random process (Hedgecock 1994). These are caused

primarily by stochastic changes in the environment acting over larval survival at the level of the individual, wherein genetic drift and recruitment dynamics may be temporarily coupled (Hedgecock & Pudovkin 2011). A prediction of the sweepstakes reproductive success hypothesis is that there will be variation in the frequency of haplotypes in time, which we encountered at moderate levels within haplogroups of *P. chilensis*. However, migration is also a possible explanation since the degree of temporal genetic variation detected in *P. chilensis* was less than variation in space.

High temporal variability in environmental conditions such as those found in the Southeast Pacific (Thiel et al. 2007) together with the high reproductive output of marine invertebrates, the existence of a planktonic larval phase (Lee & Boulding 2009), the high mortality of early life stages, and the intolerance of *P. chilensis* to disturbance (Valdivia et al. 2005) are all factors that contribute to the low to moderate degree of genetic temporal variation detected, while it may be moderated by the short larval period of *P. chilensis* which reduces the time period of larval mortality, with concomitantly to less temporal genetic structure than in species with long-lived larvae.

## REFERENCES

- Astanei I, Gosling E, Wilson J, Powell E (2005) Genetic variability and phylogeography of the invasive zebra mussel, *Dreissena polymorpha* (Pallas). *Mol Ecol* 14:1655–1666.
- Ayre DJ, Minchinton TE, Perrin C (2009) Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Mol Ecol* 18: 1887–1903.
- Barber PH, Palumbi SR, Erdmann MV, Moosa MK (2002) Sharp genetic breaks among populations of a benthic marine crustacean indicate limited oceanic larval transport: patterns, cause, and consequence. *Mol Ecol* 11: 659–674.
- Barshis DJ, Sotka EE, Kelly RP, Sivasundar A, Menge BA, Barth JA, Palumbi SR (2011) Coastal upwelling is linked to temporal genetic variability in the acorn barnacle *Balanus glandula*. *Mar Ecol Prog Ser* 439: 139–150.
- Berger MS, Darrah AJ, Emlet RB (2006) Spatial and temporal variability of the early post-settlement survivorship and growth in the barnacle *Balanus glandula* along an estuarine gradient. *J Exp Mar Biol Ecol* 336: 74–87.
- Bouchemousse S, Biship JDD, Viard F (2016) Contrasting global genetic patterns in two biologically similar, widespread and invasive *Ciona* species (Tunicata, Ascidiacea). *Scientific Reports* 6: 24875.
- Bowen BW, Chanjer K, Yasuda N, Malay MC, von der Heyden S, Paulay G, Rocha LA, Selkoe KA, Barber PH, Williams ST, Lessios HA, Crandall ED, Bernardi G, Meyer CP, Carpenter KE, Toonen RJ (2013) Phylogeography unplugged: comparative surveys in the genomic era. *Bull Mar Sci* 90:13–46.
- Cobben MM, Verboom J, Opdam PFM, et al. (2011) Projected climate change causes loss and redistribution of genetic diversity in a model metapopulation of a medium-good disperser. *Ecography* 34: 920–932.
- Colgan DJ (1981) Spatial and temporal variation in the genotypic frequencies of the mussel *Brachidontes rostratus*. *Heredity* 46: 197–208.
- Collins CJ, Fraser CI, Ashcroft A, Waters JM (2010) Asymmetric dispersal of southern bull-kelp (*Durvillaea antarctica*) adults in coastal New Zealand: testing an oceanographic hypothesis. *Mol Ecol* 19: 4572–4580.
- Crandall ED, Jones ME, Munoz MM, Akinronbi B, Erdmann MV, Barber PH. (2008) Comparative phylogeography of two seastars and their ectosymbionts within the Coral Triangle. *Mol Ecol* 17: 5276–5290.
- Crowder LB, Lyman SJ, Figueira WF, et al. (2000) Source-sink population dynamics and the problem of siting marine reserves. *Bull Mar Sci* 66: 799–820.
- David P, Perdieu MA, Pernot AF, et al. (1997) Fine-grained spatial and temporal population structure in the marine bivalve *Spisula ovalis*. *Evolution* 51: 1318–1322.
- Dawson MN (2012) Parallel phylogeographic structure in ecologically similar sympatric sister taxa. *Mol Ecol* 21: 987–1004.
- Doellman MM, Trussell GC, Grahame JW, Vollmer SV (2011) Phylogeographic

- analysis reveals a deep lineage split within North Atlantic *Littorina saxatilis*. *Proc Roy Soc London B* 278: 3175–3183.
- Excoffier L, Lischer HEL (2010) Arlequin ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows. *Mol Ecol Res* 10: 564–567.
- Goldstien SJ, Schiel DR, Gemmell NJ (2006) Comparative phylogeography of coastal limpets across a marine disjunction in New Zealand. *Mol Ecol* 15: 3259–3268.
- Hanski I, Gilpin M (Eds) (1997) *Metapopulation Biology: Ecology, Genetics and Evolution*. London: Academic Press.
- Haye PA, Muñoz-Herrera NC (2013) Isolation with differentiation followed by expansion with admixture in the tunicate *Pyura chilensis*. *BMC Evol Biol* 13: 252.
- Haye PA, Segovia NI, Muñoz-Herrera NC, Gálvez FE, Martínez A, Meynard A, Pardo-Gandarillas MC, Poulin E, Faugeron S (2014) Phylogeographic structure in benthic marine invertebrates of the southeast Pacific coast of Chile with differing dispersal potential. *PLoS ONE* 9: e88613.
- Hedgecock D (1994) Temporal and spatial genetic structure of marine animal populations in the California Current. *Calif Coop Ocean Fish Invest Rep* 35: 73–81.
- Hudson RR (2000) A new statistic for detecting genetic differentiation. *Genetics* 155: 2011–2014.
- Johnson DM (2004) Source-sink dynamics in a temporally heterogeneous environment. *Ecology* 85: 2037–2045.
- Johnson MS, Black R (1984) Pattern beneath the chaos: the effect of recruitment on genetic patchiness in an intertidal limpet. *Evolution* 38:1371–1383.
- Johnson MS, Wernham J (1999) Temporal variation of recruits as a basis of ephemeral genetic heterogeneity in the western rock lobster *Panulirus cygnus*. *Mar Biol* 135:133–139.
- Jolly MT, Viard F, Weinmayr G, et al. (2003) Does genetic structure of *Pectinaria koreni* (Polychaeta: Pectinariidae) conform to a source-sink metapopulation model at the scale of the Baie de Seine? *Helgol Mar Res* 56: 238–246.
- Kelly RP, Palumbi SR (2010) Genetic structure among 50 species of the Northeastern Pacific rocky intertidal community. *PLoS ONE* 5: e8594.
- Kesäniemi JE, Mustonen M, Boström C, Hansen BW, Knott KE (2014) Temporal genetic structure in a poecilognous polychaete: the interplay of developmental mode and environmental stochasticity. *BMC Evol Biol* 14:12.
- Kirkendale LA, Meyer CP (2004) Phylogeography of the *Patelloida profunda* group (Gastropoda: Lottidae): diversification in a dispersal-driven marine system. *Mol Ecol* 13: 2749–2762.
- Lee HJ, Boulding EG (2009) Spatial and temporal population genetic structure of four northeastern Pacific littorinid gastropods: the effect of mode of larval development on variation at one mitochondrial and two nuclear DNA markers. *Mol Ecol* 18: 2165–2184.

- Lejeusne C, Bock DG, Therriault TW, MacIsaac HJ, Crustescu ME (2011) Comparative phylogeography of two colonial ascidians reveals contrasting invasion histories in North America. *Biol Invasions* 13: 635-650.
- Lessios HA, Kessing BD, Pearse JS. 2001. Population structure and speciation in tropical seas: global phylogeography of the sea urchin *Diadema*. *Evolution* 55: 955-975.
- Lessios HA, Weinberg JR, Starczak VR (1994) Temporal variation in populations of the marine isopod *Exciorolana*: How stable are gene-frequencies and morphology? *Evolution* 48: 549-563.
- Li G, Hedgecock D (1998) Genetic heterogeneity, detected by PCR-SSCP, among samples of larval Pacific oysters (*Crassostrea gigas*) supports the hypothesis of large variance in reproductive success. *Can J Fish Aquat Sci* 55: 1025-1033.
- Librado P, Rozas J (2009) DnaSP v5: A software for comprehensive analysis of DNA polymorphism data. *Bioinformatics* 25: 1451-1452.
- Lipcius RN, Eggleston DB, Schreiber SJ, et al. (2008) Importance of metapopulation connectivity to restocking and restoration of marine species. *Rev Fisheries Sci* 16: 101-110.
- Ludford A, Cole VJ, Porri F, et al. (2012) Testing the source-sink theory: the spillover of mussel recruits beyond marine protected areas. *Landscape Ecol* 27: 859-868.
- Marshall DJ, Monro K, Bode M, et al. (2010) Phenotype-environment mismatches reduce connectivity in the sea. *Ecol Lett* 13: 128-140.
- Menge BA, Lubchenco J, Bracken MES, Chan F, Foley MM, Freidenburg TL, Gaines SD, Hudson G, Krenz C, Leslie H, et al. (2003) Coastal oceanography sets the pace of rocky intertidal community dynamics. *Proc Natl Acad Sci USA* 100:12229-12234.
- Moberg PE, Burton RS (2000) Genetic heterogeneity among adult and recruit red sea urchins, *Strongylocentrotus franciscanus*. *Mar Biol* 136:773-784.
- Ng W.C, Leung FCC, Chak STC, Slingsby G, Williams GA (2010) Temporal genetic variation in populations of the limpet *Cellana grata* from Hong Kong shores. *Mar Biol* 157: 325-337.
- Oksanen J, F. Guillaume Blanchet, Michael Friendly, Roeland Kindt, Pierre Legendre, Dan McGlenn, Peter R. Minchin, R. B. O'Hara, Gavin L. Simpson, Peter Solymos, M. Henry H. Stevens, Eduard Szoecs and Helene Wagner (2017). *vegan: Community Ecology Package*. R package version 2.4-3.
- Palsboll PJ, Peery MZ, Olsen MT, et al. (2013) Inferring recent historic abundance from current genetic diversity. *Mol Ecol* 22: 22-40.
- Pedersen EM, Hunt HL, Scheibling RE (2000) Temporal genetic heterogeneity within a developing mussel (*Mytilus trossulus* and *M. edulis*) assemblage. *J Mar Biol Ass UK* 80: 843-854.
- Peery MZ, Beissinger SR, House RF, et al. (2008) Characterizing source-sink dynamics with genetic parentage assignments. *Ecology* 89: 2746-2759.
- Penney RW, Hart MJ, Templeman ND (2006) Genotype-dependent survival,

- growth, and reproduction in cultured blue mussels, *Mytilus* spp.: results of a reciprocal seed transfer experiment. *J Shellfish Res* 25: 515-525.
- Pineda et al. (2016), Stochasticity in space, persistence in time: genetic heterogeneity in harbour populations of the introduced ascidian *Styela plicata*. *PeerJ* 4:e2158; DOI 10.7717/peerj.2158.
- Provan J, Bennett KD (2008) Phylogeographic insights into cryptic glacial refugia. *Trends in Ecology and Evolution*, 23: 564-571.
- Pujolar JM, Maes GE, Volckaert FAM (2007) Genetic and morphometric heterogeneity among recruits of the European eel, *Anguilla anguilla*. *Bull Mar Sci* 81: 297-308.
- Pulliam HR (1988) Sources, sinks, and population regulation. *Am Nat* 132: 652-661.
- R Development Core Team R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL <http://www.R-project.org>. (2017).
- Reuschel S, Cuesta JA, Schubart CD (2010) Marine biogeographic boundaries and human introduction along the European coast revealed by phylogeography of the prawn *Palaemon elegans*. *Mol Phylogenet Evol* 55: 765-775.
- Rius M, Teske PR (2011) A revision of the *Pyura stolonifera* species complex (Tunicata, Ascidiacea), with a description of a new species from Australia. *Zootaxa* 2754: 27-40.
- Roughgarden J, Gaines S, Possingham H (1988) Recruitment dynamics in complex life cycles. *Science* 241: 1460-1466.
- Shanks AL, Brink L (2005) Upwelling, downwelling, and cross-shelf transport of bivalve larvae: test of a hypothesis. *Mar Ecol Prog Ser* 302: 1-12.
- Siegel DA, Mitarai S, Costello CJ, et al. (2008) The stochastic nature of larval connectivity among nearshore marine populations. *PNAS* 105: 8974-8979.
- Singh RS, Tsagkogeorga G, Delsuc F, Blanquart S, Shenkar N, Loya Y, et al: Tunicate mitogenomics and phylogenetics: peculiarities of the *Herdmania monus* mitochondrial genome and support for the new chordate phylogeny. *BMC Genomics*. 2009, 10: 534-569. 10.1186/1471-2164-10-534.
- Svane, I. and Lunda "lv, T., 1982, Population dynamics and reproductive patterns of *Boltenia echinata* (Ascidiacea) on the Swedish West Coast, *Netherlands Journal of Sea Research*, 16, 105± 118.
- Thiel M, EC Macaya, E Acuña, WE Arntz, H Bastías, K Brokordt, P Camus, JC Castilla, LR Castro, M Cortés, CP Dumont, R Escribano, M Fernández, JA Gajardo, CF Gaymer, I Gomez, AE González, HE González, PA Haye, JE Illanes, JL Iriarte, DL Lancellotti, G Luna-Jorquera, C Luxoro, PH Manriquez, V Marín, P Muñoz, SA Navarrete, E Pérez, E Poulin, J Sellanes, HH Sepúlveda, W Stotz, F Tala, A Thomas, CA Vargas, JA Vásquez & JMA Vega. 2007. The Humboldt Current system of northern and central Chile oceanographic processes, ecological interactions and socioeconomic feedback. *Oceanography and Marine Biology: Annual Review* 45: 195-344
- Turner FT, Wares JP, Gold JR (2002) Genetic effective size is three orders of



- magnitude smaller than adult census size in an abundant, estuarine-dependent marine fish (*Sciaenops ocellatus*). *Genetics* 162: 1329-1339.
- Tsagkogeorga G, Cahais V, Galtier N: The population genomics of a fast evolver: high levels of diversity, functional constraint and molecular adaptation in the tunicate *Ciona intestinalis*. *Genome Biol Evol.* 2012, 4: 740-749. 10.1093/gbe/evs054.
- Valdivia, N. *et al.* Effects of disturbance on the diversity of hard-bottom macrobenthic communities on the coast of Chile. *Mar. Ecol. Prog. Ser.* **299**, 45-54 (2005).
- Varela AI, Haye PA (2012) The marine brooder *Excirrolana braziliensis* (Crustacea: Isopoda) is also a complex of cryptic species on the coast of Chile. *Rev Chil Hist Nat* 85: 495-502.
- Veale AJ, Lavery SD (2012) The population genetic structure of the waratah anemone (*Actinia tenebrosa*) around New Zealand. *N Z J Mar Freshw Res* 46: 523-536.
- Vigliola L, Doherty PJ, Meekan MG, et al. (2007) Genetic identity determines risk of postsettlement mortality of a marine fish. *Ecology* 88: 1263-1277.
- Villamor A, Costantini F, Abbiati M (2014) Genetic structuring across marine biogeographic boundaries in rocky shore invertebrates. *PLoS ONE* 9: e101135.
- Virgilio M, Abbiati M (2006) Temporal changes in the genetic structure of intertidal populations of *Hediste diversicolor* (Polychaeta: Nereididae). *J Sea Res* 56: 53-58.
- Wangensteen OS, Turon X, Pérez-Portela R, Palacín C (2012) Natural or naturalized? Phylogeography suggests that the abundant sea urchin *Arbacia lixula* is a recent colonizer of the Mediterranean. *PLoS ONE* 7: e45067.
- Wares JP, Cunningham CW (2001) Phylogeography and the historical ecology of the North American intertidal. *Evolution* 55: 2455-2469.
- Weersing K, Toonen RJ (2009) Population genetics, larval dispersal, and connectivity in marine systems. *Mar Ecol Prog Ser* 393: 1-12.
- Weir BS, Cockerham CC (1984) Estimating f-statistic for the analysis of population structure. *Evolution* 38:1358-1370.
- Wilke T, Davis GM (2000) Intraspecific mitochondrial sequence diversity in *Hydrobia ulvae* and *Hydrobia ventrosa* (Hydrobiidae: Rissooidea: Gastropoda): do their different life history life histories affect biogeographic patterns and gene flow? *Biol J Linn Soc* 70: 89-105.
- Zakas C, Binford J, Navarrete SA, Wares JP (2009) Restricted gene flow in Chilean barnacles reflects an oceanographic and biogeographic transition zone. *Mar Ecol Prog Ser* 394: 165-177.

## Anexos

### Tables and Figures

**Table 1.**

**Sampling locations of *Pyura chilensis* individuals collected from artisanal fisheries in the Chilean coast.** Sample location, location code, Latitude (only latitude given because of the north-south linear nature of the Chilean coast), and number of samples (N) analyzed per year. In total, sequences of the mitochondrial gene *COI* were obtained for 592 individuals: 203 from 2012, 196 from 2014, and 193 from 2015.

Location	Code	Latitude	N (2012)	N (2014)	N (2015)
Pan de Azúcar	PA	26°08'S	30	34	32
Caleta Pajonales	CP	27°44'S	31	32	31
La Herradura	LH	29°58'S	37	32	33
Talcahuano	TH	36°42'S	27	31	32
Los Molinos	LM	39°50'S	38	32	34
Ancud	AC	41°52'S	40	35	31
<b>TOTAL</b>			<b>203</b>	<b>196</b>	<b>193</b>

**Table 2.**

***P*- values of Fisher's exact test for a contingency table to determine non-random associations between *COI* haplogroup frequencies of *Pyura chilensis* between pairwise temporal samples per location. *P*-values indicate the probability of the observed array of haplogroup frequencies. Gray cells and bold numbers indicate significant values.**

	2012/2014	2012/2015	2014/2015
PA	0.99	0.99	1
CP	0.42	0.67	1
LH	0.13	0.14	1
TH	0.29	0.3	1
LM	<b>0.00084</b>	<b>0.032</b>	0.52
AC	0.3	0.76	0.39

**Table 3.**

**One-way Nested ANOVA of COI genetic diversity of *Pyura chilensis* using location and temporal sample (years 2012, 2014 and 2015) as fixed and random effects. Analyses were separately performed for all the COI dataset and for each haplogroup. Gray cells and bold numbers indicate significant values.**

Fixed effect	Random effect	Dataset	Nucleotide diversity		Haplotype diversity	
			F	<i>P</i>	F	<i>P</i>
Year	Location	All	0.074	0.791	0.016	0.902
		Hg1	0.289	0.62	0.525	0.485
		Hg2	0.114	0.744	1.114	0.319
		Hg3	0.056	0.835	0.13	0.753
Location	Year	All	4.8	<b>0.0143</b>	33.23	<b>2.79E-06</b>
		Hg1	4.652	<b>0.0187</b>	13.32	<b>0.000374</b>
		Hg2	0.129	0.968	0.928	0.489
		Hg3	1.556	0.338	1.298	0.373

**Table 4.**

**Temporal pairwise  $\phi_{ST}$  estimates per location based on COI sequences in *Pyura chilensis*.** Values represent genetic differentiation between year samples (2012, 2014 and 2015) and were calculated for each location including and were performed for all the COI dataset and for each haplogroup. Gray cells and bold numbers indicate significant values.

		TOTAL		Hg1		Hg2		Hg3	
		2014	2015	2014	2015	2014	2015	2014	2015
PA	2012	-0.0031	-0.0031	-0.0031	-0.0031				
	2014		-0.0083		-0.0083				
CP	2012	0.00649	0.00126	0.00830	0.00182	-0.1915	-0.0182		
	2014		0.00922		<b>0.01053</b>		0		
LH	2012	0.00111	<b>0.01685</b>	-0.0016	0.00902	-0.0091	0.06279		
	2014		<b>0.01730</b>		<b>0.01793</b>		<b>0.06765</b>		
TH	2012	0.00888	<b>0.05678</b>	0.01632	<b>0.08347</b>	-0.0404	<b>0.12313</b>		
	2014		<b>0.04186</b>		<b>0.01667</b>		<b>0.16164</b>		
LM	2012	0.00296	<b>0.01247</b>	0.07273	0.04660	0.15789	<b>0.32530</b>	-0.0134	0.00434
	2014		<b>0.00756</b>		0		0.34426		0.00666
AC	2012	-0.0069	<b>0.16386</b>	-0.0088	<b>0.43554</b>	0.05929	<b>0.16630</b>	0.33333	0
	2014		<b>0.15449</b>		<b>0.35826</b>		<b>0.29425</b>		0.20755

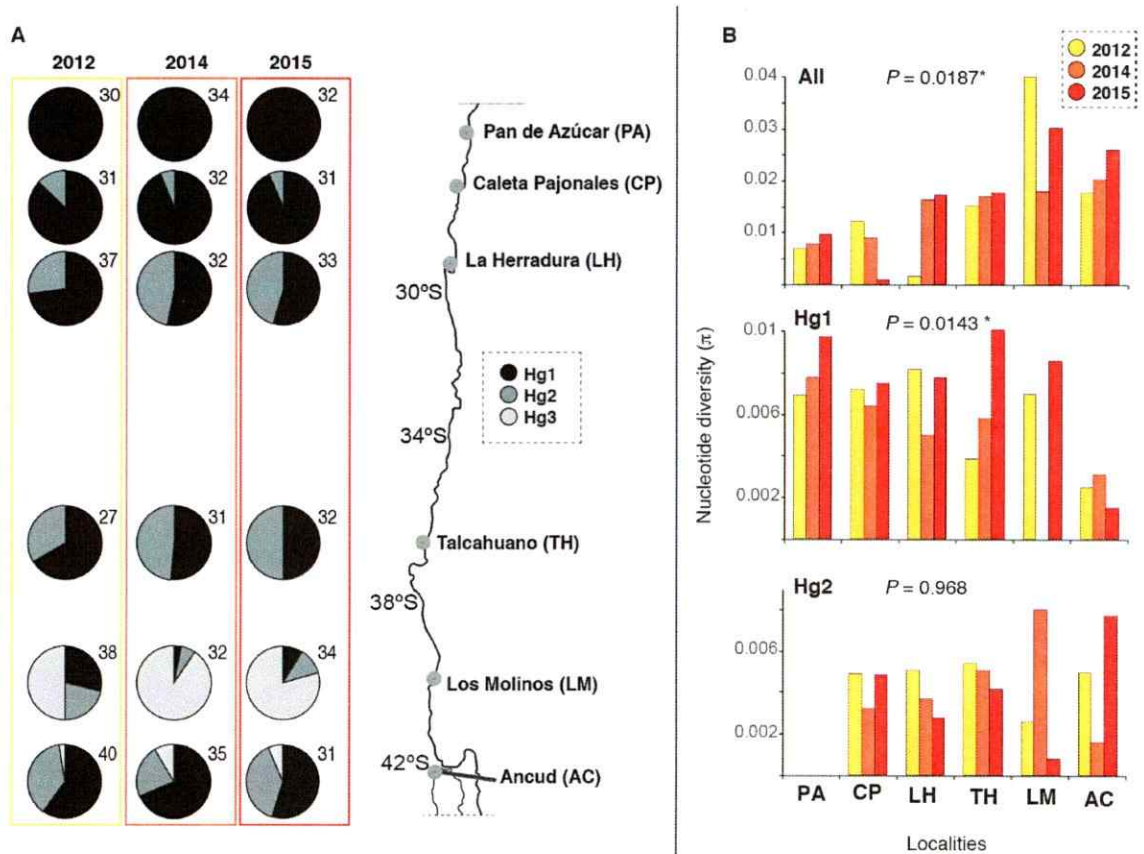
**Table 5.**

**Results from AMOVA and Snn of COI sequences of *Pyura chilensis* grouping by temporal sample (years 2012, 2014 and 2015).** Analyses were performed for all the COI dataset and for each haplogroup. Gray cells and bold numbers indicate significant values.

Data	AMOVA			Snn
	Among years	Among population within years	Within populations	
All	0.32	<b>3.08</b>	<b>96.6</b>	<b>0.51937</b>
Hg 1	-0.17	<b>6.43</b>	<b>93.74</b>	<b>0.50841</b>
Hg 2	<b>7.19</b>	3.03	<b>89.78</b>	<b>0.59105</b>
Hg 3	-4.35	5.24	<b>99.1</b>	0.343

**Figure 1.**

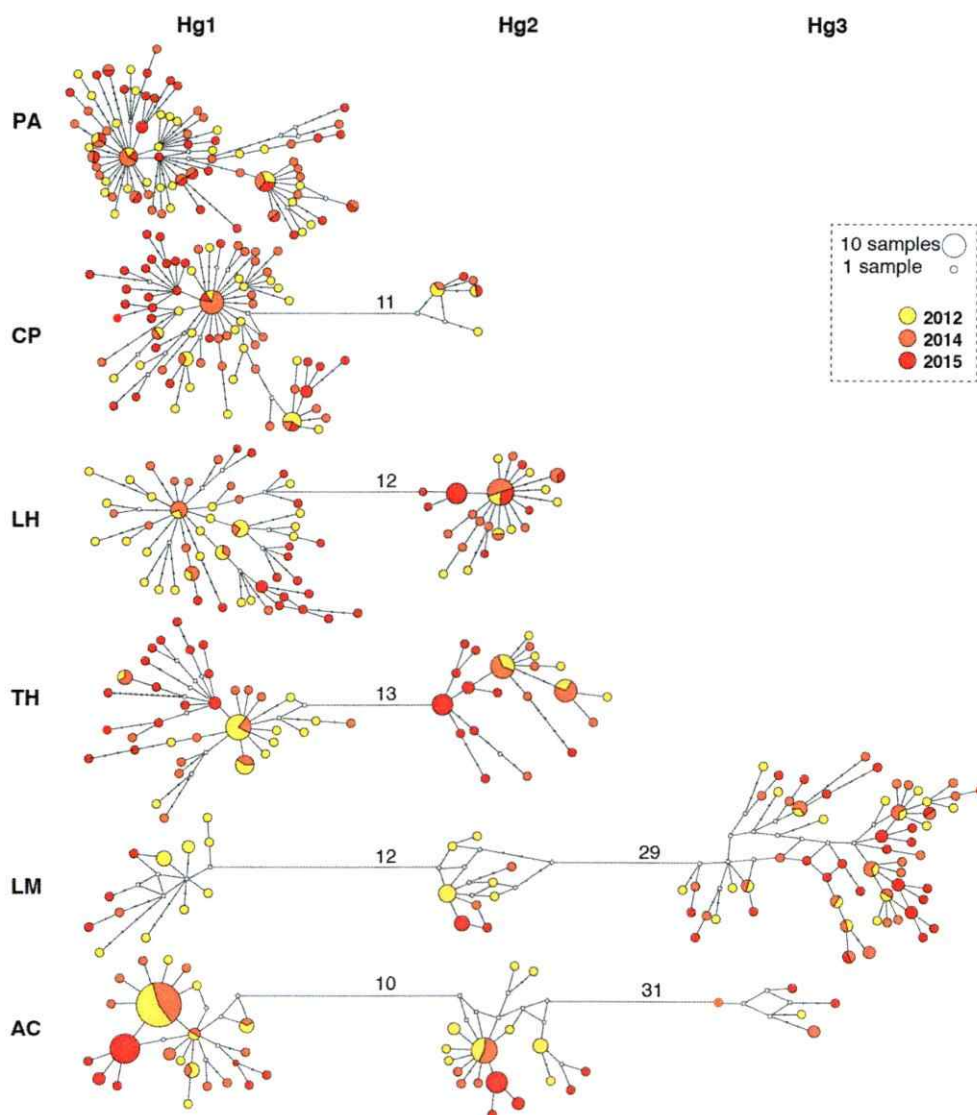
**Temporal variation of COI haplogroups of *Pyura chilensis*.** A. Geographic distribution and relative frequency of COI haplogroups for temporal samples (years 2012, 2014 and 2015). Pie charts represent the total number of sequences analyzed, colors represent COI haplogroups, and numbers by the pies indicate number of individuals analyzed. B. Histograms depicting the temporal variation in nucleotide diversity by location. Analyses were performed for all the COI dataset and for two widely distributed haplogroups 1 and 2 (Hg1 and Hg2 respectively). *P*-values correspond to nested ANOVA with locations as fixed effect (see Table 3). Asterisks indicate significant values.



**Figure 2**

**Haplotype networks per location of temporal COI haplotypes of *Pyura chilensis*.**

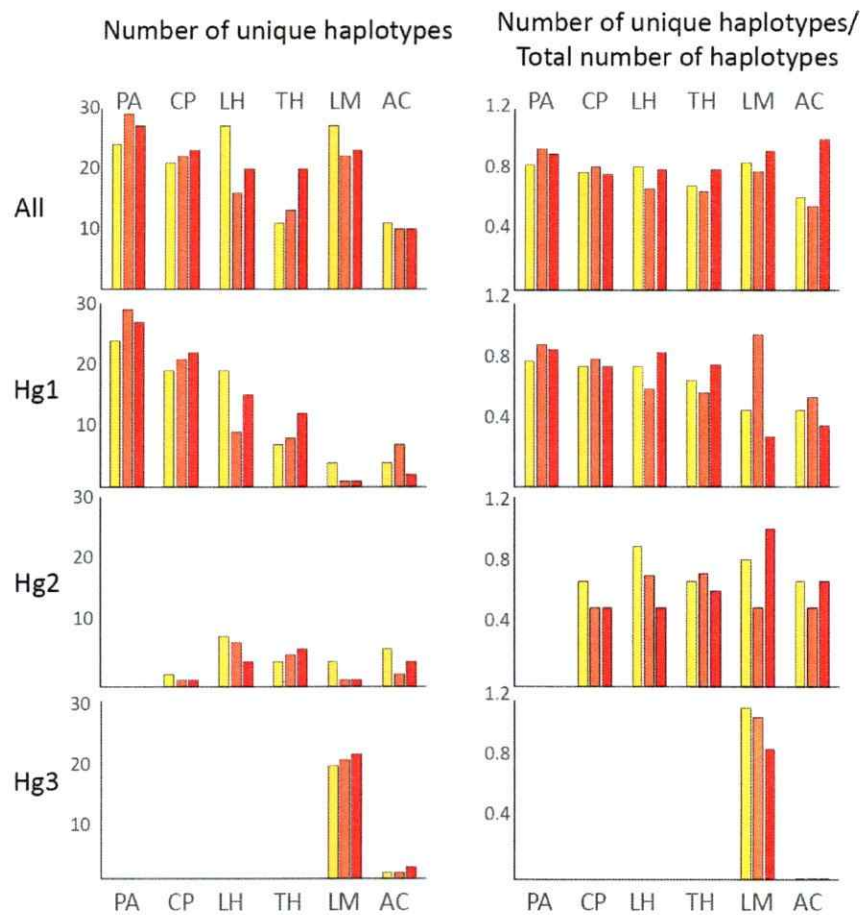
Each circle represents a sampled haplotype, and the size is proportional to the number of individuals with the haplotype. Colors represent temporal samples (years 2012, 2014, and 2015). Lines between haplotypes represent a mutational step, and small circles along the lines correspond to additional steps. The three COI haplogroups of *P. chilensis* (Hg1, Hg2 and Hg3) are shown.





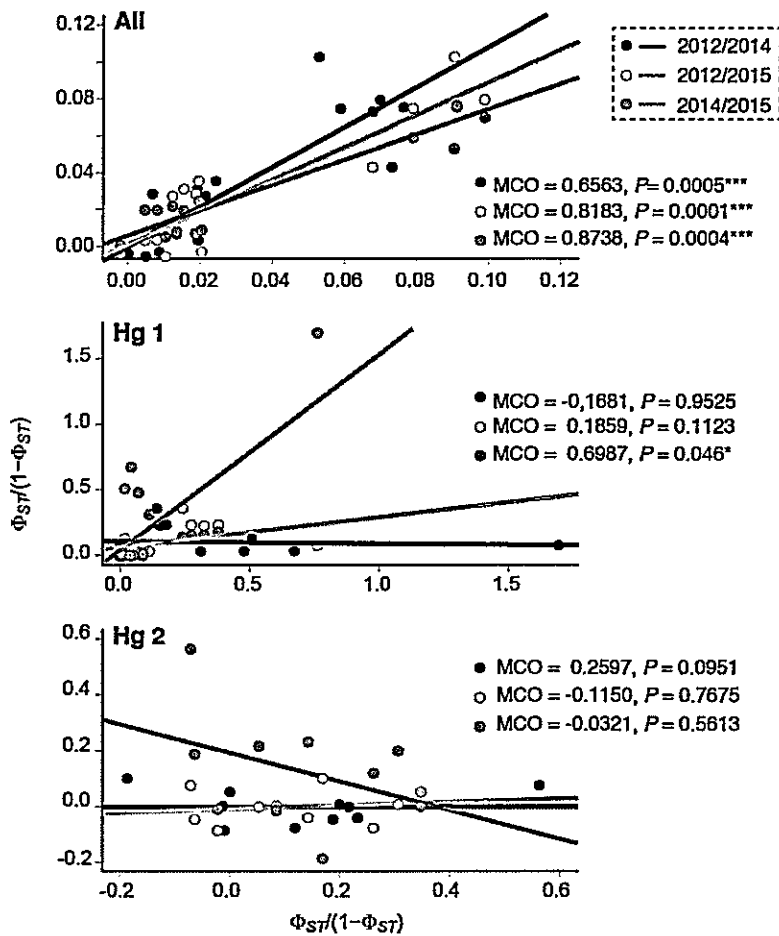
### Figure 3

Temporal variation in the number of unique COI haplotypes in *Pyura chilensis*. The total number of unique haplotypes and the ratio between unique haplotypes and total number of haplotypes per site are presented.



**Figure 4**

**Temporal population divergence of COI sequences in *Pyura chilensis*.** Scatterplots for all the COI dataset and for haplogroups 1 and 2 (Hg1 and Hg2 respectively) are presented. Population-pairwise  $\Phi_{ST}$  matrices were compared between temporal samples (years 2012, 2014 and 2015). MCO corresponds to the Mantel test correlation value; associated  $P$ -values are shown. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ .



## Supplementary Material

**Table S1**

**Genetic diversity measures for all COI data and for haplogroups 1 and 2 (Hg1 and Hg2 respectively) of *Pyura chilensis* from three temporal samples (years 2012, 2014 and 2015). Number of haplotypes (H), number of segregating sites (S), haplotype diversity ( $\pi$ ), nucleotide diversity ( $h$ ).**

		H			S			h			$\pi$		
		2012	2014	2015	2012	2014	2015	2012	2014	2015	2012	2014	2015
All	PA	29	31	30	38	53	59	0.9977	0.99287	0.99597	0.00692	0.00782	0.00977
	CP	27	27	30	52	55	50	0.989	0.96976	0.99785	0.01218	0.009	0.00102
	LH	33	24	25	60	45	48	0.99249	0.96976	0.96402	0.00163	0.01638	0.01724
	TH	16	20	25	37	41	51	0.92023	0.95054	0.96573	0.01517	0.01716	0.01784
	LM	32	28	29	78	76	79	0.98578	0.99194	0.97748	0.04013	0.01813	0.03035
	AC	18	18	14	58	63	57	0.85	0.86891	0.82366	0.01775	0.02047	0.02607
	Total	134	129	138	138	139	144	0.98015	0.98127	0.98257	0.02392	0.02834	0.02951
Hg 1	PA	29	31	30	38	53	59	0.9977	0.99287	0.99597	0.00692	0.00782	0.00977
	CP	24	25	28	39	44	38	0.9886	0.96552	0.99754	0.00723	0.0064	0.00749
	LH	24	14	17	44	24	29	0.9886	0.97059	0.99346	0.00821	0.00505	0.00781
	TH	10	13	15	18	22	34	0.84314	0.975	0.99167	0.00389	0.00582	0.0101
	LM	8	-	3	19	-	8	0.92727	-	1	0.00699	-	0.00863
	AC	8	12	5	9	17	7	0.61323	0.75725	0.50735	0.00251	0.00314	0.0015
	Total	87	82	88	91	100	99	0.96876	0.97052	0.98261	0.00636	0.00615	0.00846
Hg 2	PA	-	-	-	-	-	-	-	-	-	-	-	-
	CP	3	2	2	6	2	3	0.83333	1	1	0.00489	0.00323	0.00485
	LH	9	10	8	14	13	10	0.97778	0.89524	0.82857	0.00507	0.00372	0.0028
	TH	6	7	10	10	12	16	0.88889	0.80952	0.86667	0.00543	0.00507	0.00418
	LM	5	2	1	5	5	0	0.78571	1	0	0.00262	0.00806	0.00081
	AC	9	4	6	12	4	23	0.90476	0.64286	0.75758	0.00499	0.00161	0.00775
	Total	27	20	23	29	28	41	0.88792	0.84088	0.84864	0.00468	0.00433	0.00488
Hg 3	PA	-	-	-	-	-	-	-	-	-	-	-	-
	CP	-	-	-	-	-	-	-	-	-	-	-	-
	LH	-	-	-	-	-	-	-	-	-	-	-	-
	TH	-	-	-	-	-	-	-	-	-	-	-	-
	LM	20	25	24	32	37	36	1	0.99015	0.99145	0.00928	0.00815	0.00816
	AC	-	2	2	-	5	5	-	0.66667	1	-	0.00538	0.00809
	Total	20	26	26	32	38	38	1	0.98589	0.99261	0.00928	0.00798	0.00833

**Table S2**  
**Population-pairwise  $\Phi_{ST}$  matrices for all COI data and for each haplogroup of *Pyura chilensis*. Values are detailed separately for each temporal sample (years 2012, 2014 and 2015). Gray cells and bold numbers indicate significant values.**

All		Hg1					Hg2					Hg3							
	PA	CP	LH	TH	LM	PA	CP	LH	TH	LM	CP	LH	TH	LM	CP	LH	TH	LM	
2012	CP	-0.0043																	
	LH	-0.0032	-0.0058				-0.0065				-0.0187				-0.038	-0.0006			
	TH	<b>0.0301</b>	<b>0.0264</b>	<b>0.0276</b>		<b>0.0611</b>	<b>0.0626</b>				-0.0756	0.0173	-0.0057		-0.0106	0.0071	-0.0352	0.0202	
	LM	<b>0.0066</b>	0.0032	0.0031	<b>0.0342</b>	<b>0.0291</b>	<b>0.0302</b>	<b>0.1129</b>											
	AC	<b>0.0737</b>	<b>0.0702</b>	<b>0.0697</b>	<b>0.0933</b>	<b>0.0413</b>	<b>0.1839</b>	<b>0.1879</b>	<b>0.2627</b>	<b>0.0692</b>									<b>0</b>
2014	CP	0.0003				0.0012													
	LH	0.0086	0.0050			-0.0011	-0.0096				-0.1109				-0.0345	0.0411			
	TH	<b>0.0189</b>	<b>0.0211</b>	0.0068		-0.0027	0.0007	-0.0023			0	0.0783	-0.0345		-0.0932	0.0080	0.07415	0.2697	
	LM	<b>0.0076</b>	<b>0.0192</b>	<b>0.0192</b>	<b>0.0238</b>	0.0071	0.0345	0.0294	0.025										
	AC	<b>0.0653</b>	<b>0.0710</b>	<b>0.0556</b>	<b>0.0503</b>	<b>0.0683</b>	<b>0.1158</b>	<b>0.1278</b>	<b>0.1317</b>	<b>0.0882</b>	<b>0.2428</b>								<b>0.0924</b>
2015	CP	-0.002				-0.0022													
	LH	<b>0.0200</b>	<b>0.0104</b>			0.0053	-0.0013				-0.0976				0.0386	<b>0.0835</b>			
	TH	<b>0.0153</b>	<b>0.0123</b>	<b>0.0184</b>		-0.0017	-0.0033	0.0074			-0.0811	0.0049	<b>0.1881</b>		-0.1053	0.0079	<b>0.1315</b>	-0.0397	
	LM	<b>0.0133</b>	<b>0.0081</b>	0.0048	<b>0.0193</b>	0.0027	0.0016	0.0042	-0.0158										
	AC	<b>0.0900</b>	<b>0.0835</b>	<b>0.0733</b>	<b>0.0830</b>	<b>0.0635</b>	<b>0.2254</b>	<b>0.2276</b>	<b>0.2472</b>	<b>0.2184</b>	<b>0.1005</b>								<b>0.0063</b>

## CAPITULO III: Adaptación local y asociaciones genético ambientales en el pacífico sudeste en la ascidia *Pyura chilensis*.

### Resumen

Los estudios de genética del paisaje permiten mediante la asociación con variables ambientales, determinar la influencia del ambiente en los patrones de estructura genética adaptativa a lo largo de un paisaje heterogéneo. En este estudio, se realizaron análisis de asociaciones ambientales utilizando SNPs dispersos por el genoma obtenidos mediante *Genotyping-by-Sequencing* en la ascidia *Pyura chilensis*, un organismo marino bentónico endémico de la costa de Chile, con baja capacidad de dispersión larval, pero un alto potencial de dispersión por efecto del transporte antropogénico en cascos de embarcaciones. Estudios previos indican que esta especie presenta una moderada diferenciación genética entre poblaciones locales con loci neutrales. Con loci sujetos a selección, sin embargo, se detecta estructura genética significativa asociada a un quiebre biogeográfico ubicado a los 30°S, lo que sugiere un importante rol del ambiente en la estructura genética adaptativa asociada a este quiebre. Para determinar la influencia del ambiente en la estructuración al norte y sur de los 30°S, se utilizaron nueve variables ambientales las cuales se correlacionaron a la estructura genética de 2902 SNPs en 149 individuos de 5 sitios entre los 24° y los 42°S, usando dos familias de análisis de asociación ambiental; 1) tres métodos basados en análisis de diferenciación poblacional (ADP) (BayesScan, OUTFLANK, PCAdapt) para terminar loci candidatos a selección positiva y loci neutrales y luego análisis de redundancia (RDA), para correlacionar linealmente la estructura genética neutral y adaptativa a una o varias variables ambientales y 2) tres métodos de análisis de asociaciones genético-ambientales (EAA) (Bayenv2, LFMM, Sambada) los que permiten asociar directamente la estructura genética *locus-by-locus* a un set de variables ambientales controlando por estructura neutral y por espacio. Se detectaron cientos de loci candidatos (entre 75-585), los que estuvieron significativamente correlacionados a temperatura superficial del mar y a variables asociadas a surgencias y productividad como el fosfato, nitrato, fluorescencia y pH. La estructura genética espacial de los loci candidatos a cada variable ambiental muestra consistentemente el quiebre a los 30°S, y esta diferenciación se ajusta más a cada variable ambiental que lo que se espera de acuerdo a la distancia geográfica entre sitios. Esto sugiere que estas variables ambientales podrían modelar la divergencia adaptativa entre el norte y sur de los 30°S provocando procesos de adaptación local. Para los loci candidatos, se identificaron 12 genes de los cuales 8 tienen funciones anotadas conocidas, las que están asociadas a procesos como filtración y ventilación (biogénesis de cilios), crecimiento y factores de transcripción. Estos genes podrían ser genes candidatos interesantes para estudiar tanto la susceptibilidad a factores abióticos como la resiliencia de *Pyura chilensis* al cambio climático.

## Introducción

### *Adaptación local en el ambiente marino*

Establecer cómo el ambiente físico a lo largo de un paisaje heterogéneo influye en el flujo genético efectivo y en los patrones de distribución de la diversidad genética es un objetivo clave para la genética del paisaje. Para que el flujo genético entre poblaciones sea efectivo, los migrantes deben sobrevivir y reproducirse (Pineda et al 2007; Hamilton et al 2008; Marshall et al 2010). Los efectos de la selección natural repercuten en la mortalidad selectiva de migrantes por efecto de desajustes entre los genotipos y los ambientes locales (DeWitt 1998), lo que podría reducir el flujo genético a pesar de haber transporte físico activo de individuos entre poblaciones locales. Estos desajustes pueden ocasionarse por procesos de adaptación local entre poblaciones de la misma especie que se encuentran conectadas, o al menos potencialmente, pero que están diferenciadas genéticamente por efecto de la activa elección de hábitats (Kawecki & Ebert 2004).

En el caso del ambiente marino, definir barreras físicas al flujo genético en el ambiente marino es complejo (Palumbi et al 1994), sin embargo, este varía de forma significativa debido a factores tanto bióticos como abióticos. La mayoría de las especies marinas bentónicas costeras que habitan ambientes heterogéneos cuentan con estadíos larvales dispersivos que permiten el transporte de individuos entre poblaciones locales. Sin embargo, la capacidad de dispersión no está

necesariamente correlacionada con la conectividad (Swarer et al 2002; Thiel & Haye 2006; Teske et al 2006; Haye & Muñoz 2013), lo que implica que la estructura genética no siempre es consistente con la distancia geográfica entre poblaciones (Johnson & Black 2006; Hedgecock et al 2007; Schmidt et al 2008). Diversos estudios sugieren que los procesos de adaptación local podrían ocurrir en ambientes marinos a pesar del efecto homogenizador del flujo genético (e.j. Conover 1998; Grosberg & Cunningham 2001; Kawecki & Ebert 2004; Véliz et al 2005; Sotka 2005; Conover et al 2006; Nielsen et al 2009; Conover et al 2009; Marshall et al 2010, Sanford & Kelly 2011).

#### ***Búsqueda de señales de selección***

La identificación de señales de selección natural a nivel genético para poder investigar sus efectos, ha sido de gran interés para la biología evolutiva. Gracias a la implementación de las técnicas de secuenciación masiva de última generación, los estudios de genética del paisaje se han incrementado rápidamente en la última década, pero estos provienen en general de organismos terrestres (Manel & Holderegger 2013), soliendo ser menos numerosos en organismos marinos (e.j. Storfer et al 2006; Hansen & Hemmer-Hansen 2007; Riginos & Liggins 2013; Kershaw & Rosenbaum 2014, Benestan et al 2016, Selkoe et al 2016). Adicionalmente, en organismos marinos los estudios de genética del paisaje se han centrado primordialmente en evaluar procesos neutrales tales como patrones de conectividad genética respecto a la dispersión larval y a factores físicos como las

corrientes oceánicas (ver Selkoe et al 2016 para una revisión). Estudios que consideren los potenciales efectos de factores ambientales en la variación genética neutral y adaptativa siguen siendo escasos (e.j. Pujolar et al 2014, Tepolt & Palumbi 2015, Giles et al 2015, Benestan et al 2016). Destacan trabajos recientes como el de Benestan et al (2016), en que evaluaron la importancia relativa de las corrientes oceánicas y de la temperatura superficial del mar (SST) en los patrones de estructura genética neutral y adaptativa en la langosta *Homarus americanus*, detectando señales de selección asociadas a adaptación a los cambios de temperatura. Por otra parte, Tepold & Palumbi (2015), mediante el uso de Polimorfismos de Nucleótido Único (SNPs) extraídos desde análisis transcriptómicos del cangrejo invasor *Carcinus maenas*, destacan el rol de la temperatura como principal agente selectivo en poblaciones invasoras y la importancia del rol de los procesos no neutrales en la estructura genética. Ambos trabajos recalcan la importancia de la temperatura en los patrones de estructuración genética adaptativa para organismos marinos.

Durante la última década, una de las estrategias más comunes para la búsqueda de señales de selección ha sido a través del uso de herramientas basadas en análisis de diferenciación poblacional (ADP; Luikart et al 2003; Jensen et al 2016). El objetivo de este tipo de aproximaciones es identificar *loci* candidatos cuyos índices de diferenciación genética (e.j.  $F_{ST}$ ) sean significativamente más altos respecto a lo que se espera bajo una distribución neutral a la selección (Francois et al 2016, Jensen 2016) o bien, identificar aquellos *loci* que están «excesivamente»



relacionados a la estructura genética espacial (Luu et al 2017). Con este tipo de aproximaciones no se requiere conocimiento *a priori* sobre los factores ambientales que actúan potencialmente como fuerzas selectivas y, en general, tampoco es necesario contar con información genómica previa por lo que resultan muy convenientes para especies no modelo (Lotterhos & Whitlock 2015). La principal limitante de las ADP es que estas, al incluir solo la diferenciación genética como parámetro clave, pueden cubrir prácticamente cualquier fuerza evolutiva, lo que restringe la capacidad de identificar aquellos factores que modelan la estructura genética adaptativa (Lotterhos & Whitlock 2015). De esta forma, en complemento a las aproximaciones ADP, el uso de herramientas de genómica del paisaje, mediante la inclusión de variables ambientales en el análisis genético poblacional, resulta crucial para identificar potenciales agentes selectivos que influyen tanto la intensidad como las interacciones entre fuerzas evolutivas, así como la variación genética (adaptativa) que provoca procesos de adaptación local (Holderegger et al 2010, Manel et al 2010, Manel & Holderegger 2013, Sork et al 2013).

Una de estas herramientas son los análisis de asociaciones ambientales (o EAA por sus siglas en inglés), las que proponen tanto una alternativa como un complemento a los ADP. Los EAA integran herramientas de genética del paisaje y genómica poblacional para poner a prueba si un alelo o variante de un gen está significativamente asociada/o a un factor ambiental en el cual este ocurre de forma predominante (Holderegger et al. 2010, Manel et al 2010, Manel & Holderegger

2013, Sork et al 2013). A pesar de que la ocurrencia de falsos positivos es común en ambas aproximaciones (Frichot et al 2013; Rellstab et al 2015; de Villemereuil & Gaggiotti 2015; Francois et al 2016), al utilizar de forma explícita factores ambientales en el análisis, las EAA tienen el potencial de cubrir patrones inducidos por procesos adaptativos que no se detectan usando aproximaciones genómicas tradicionales (Pritchard et al 2010, Rellstab et al 2015, Francois et al 2016, Benestan et al 2016).

#### ***Sistema de Corriente de Humboldt***

La Corriente de Humboldt es un sistema lineal heterogéneo ambientalmente por lo que representa un modelo interesante para evaluar la estructura genética adaptativa en organismos marinos. Este sistema no parece presentar barreras geográficas/físicas evidentes al flujo genético a lo largo de los 2.600 km lineales de costa que lo componen, sin embargo, diversos estudios han mostrado diferencias genéticas significativas entre poblaciones locales cercanas en términos geográficos. Estas diferencias genéticas están asociadas principalmente a un quiebre biogeográfico ampliamente reportado localizado aproximadamente a los 30°S, el cual se observa principalmente a organismos con baja capacidad de dispersión (e.j. Haye et al 2014, Sánchez et al 2011). Haye et al (2014) indican que este quiebre filogeográfico podría ser histórico, ya que solo especies con baja capacidad de dispersión retienen la señal asociada mientras en organismos de alta capacidad de dispersión el efecto homogenizador del flujo genético ha logrado borrar esta señal. La mantención de este quiebre podría deberse a diferencias

ambientales al norte y al sur del quiebre de los 30°S, lo que permite hipotetizar que esta discontinuidad podría tener influencia contemporánea en los organismos que habitan el sistema.

Uno de los habitantes más conspicuos de este sistema es el piure, *Pyura chilensis* Molina 1782, una ascidia intermareal sésil de alta demanda comercial endémica de la costa del pacífico sureste. Habita tanto intermareal bajo como submareal somero entre los 10 y 44°S (Lancelotti & Vásquez 2000) y es de gran importancia comercial (Davis 1995) y ecológica debido a que es una especie bioingeniera donde se albergan más de 60 especies marinas bentónicas (Sepúlveda et al 2003, Zamorano & Moreno 1975). La especie presenta un estadio larval lecitotrófico breve (12-24 horas) y luego se asienta en sustratos duros o en la túnica de organismos conespecíficos (Cea 1970, Davis 1996, Manriquez & Castilla 2007), por lo que presenta un bajo potencial de dispersión intrínseco. Sin embargo, estudios genéticos previos muestran estructura genética moderada y no detectan evidencia genética del quiebre filogeográfico a los 30°S (Haye & Muñoz-Herrera 2013, Astorga & Ortiz 2006), el cual ha sido reportado para especies de similar o incluso mayor capacidad de dispersión larval (Haye & Muñoz-Herrera 2013, Haye et al 2014), entre las que se incluyen incluso otras ascidias de similares características biológicas (*Ciona*) (Bouchemousse et al 2016).

Si bien la baja capacidad de dispersión larval es una característica común en ascidias, estos organismos son capaces de invadir diversos ambientes a través del transporte en vectores marítimos como casquetes de barcos o cuerdas (Hudson et

al 2016, Ordoñez et al 2015, Pinneda et al 2011). A pesar de que este tipo de transporte antropogénico no ha sido evaluado de forma explícita en *P. chilensis*, la especie ha sido recurrentemente reportada en navíos artesanales, boyas e instalaciones marinas como parte importante del *biofouling* (Valdivia et al 2005, Cifuentes et al 2007, Pérez-Valdez, 2017), lo que podría explicar la ausencia del quiebre filogeográfico de los 30°S (Haye & Muñoz-Herrera 2013).

Segovia et al (2017) evaluaron la estructura genética neutral y putativamente adaptativa de *P. chilensis* utilizando SNPs dispersos por el genoma. En este estudio, en contraposición a lo previamente reportado, se observaron diferencias significativas al norte y sur de los 30°S, particularmente en loci putativamente sujetos a selección. El hecho de que a nivel neutral la influencia de este quiebre biogeográfico sea inexistente (Haye & Muñoz-Herrera 2013) o se observe solo como una señal subyacente a la señal principal (Segovia et al 2017), sugiere que la estructura genética adaptativa en *P. chilensis* podría estar significativamente influenciada por diferencias ambientales al norte y sur del quiebre biogeográfico de los 30°S. Dentro de estas diferencias ambientales, destaca principalmente un gradiente natural en la temperatura superficial del mar (SST) asociado a la latitud (Aravena et al 2014) y la presencia de distintas zonas de surgencia influenciando significativamente a organismos marinos (Thiel et al 2007, Riquelme-Bugueño et al 2016).

La heterogeneidad del paisaje a lo largo de la HCS, y particularmente las diferencias ambientales entre ambas regiones (norte/sur de los 30°) podrían tener

un rol fundamental en mantener la diferenciación genética a pesar de haber transporte activo de individuos entre poblaciones locales, generando procesos de adaptación local. El objetivo de este estudio es, mediante la utilización de herramientas de genética del paisaje, identificar variables ambientales que podrían modelar la estructura genética adaptativa en la ascidia *P. chilensis*, y establecer la influencia del ambiente en la distribución espacial de la diversidad genética tanto neutral como adaptativa al norte y al sur de los 30°S.

## **Materiales y Métodos**

### ***Toma de muestras y extracción de ADN***

Individuos de *Pyura chilensis* fueron extraídos por pescadores locales desde cinco localidades (30 por localidad) a lo largo del HCS entre los 24°S a los 42°S. El muestreo se realizó en sustratos naturales desde el submareal somero, ya sea tanto de coipás como de individuos aislados separados por al menos 2 metros para evitar sesgos por efecto del parentesco. De cada individuo, se extrajo una pequeña muestra de manto (0,2 g) la que fue usada para extraer el ADN usando el kit DNeasy Blood® & Tissue (QIAGEN), siguiendo las instrucciones del fabricante. La pureza y cantidad de ADN fueron medidas en un Nanodrop 2000 (Thermo, USA)

### ***Genotyping by Sequencing (GBS) y genotipificación de SNPs.***

Las muestras obtenidas fueron secuenciadas mediante Genotyping-by-Sequencing (GBS). Tanto parámetros del *demultiplexing* como del SNP *calling* fueron hechos de acuerdo a lo descrito en Segovia et al (2017) usando el pipeline Universal Network-Enabled Analysis Kit (UNEAK, Lu et al 2013) diseñado para trabajar con especies sin genoma de referencia, en la plataforma TASSEL V3 (Bradbury et al 2007).

### **Búsqueda de loci candidatos a adaptación local.**

#### ***Datos ambientales***

Para analizar los efectos de factores ambientales en la estructura genética del piure, se obtuvieron datos de variables oceanográficas relevantes para los 5 sitios estudiados. Se seleccionaron 9 variables cuyos datos fueron extraídos desde dos bases de datos distintas (Tabla 1). Un primer set (Set A1) de cinco variables oceanográficas, 1) concentración de *clorofila a*, 2) Fluorescencia, 3) carbono orgánico particulado (PIC), 4) Carbono Inorgánico particulado (PIC) y 5) Temperatura superficial del mar (SST), fueron obtenidas desde la base de datos satelitales de Aqua-MODIS de la NASA (disponibles en GIOVANI; <https://giovanni.gsfc.nasa.gov/giovanni/>) con una resolución espacial de 4 km.

Para cada variable del Set A1 se utilizó el promedio de los últimos 10 años hasta la fecha del muestreo (2005-2015).

El segundo set de datos ambientales (Set A2), contiene 5 variables oceanográficas; 1) Concentración de Nitrato, 2) Concentración de fosfato, 3) Concentración de silicato, 4) pH 5) Temperatura superficial del mar (SST), las que fueron obtenidas desde la base de datos Bio-Oracle (Ocean Rasters for Analyses of Climate and Environment) (Tyberghein et al 2012) con una resolución de 5 arcmin (9.2 km).

La concentración de clorofila a (*Chl-a*) ( $\text{mg m}^{-3}$ ) y la fluorescencia normalizada (*Flu*) medida en  $\text{mW cm}^{-2} \mu\text{m}^{-1} \text{sr}^{-1}$  se utilizan como *proxis* de actividad fotosintética y biomasa fitoplanctónica. La concentración de *clorofila a* es un factor selectivo importante, ampliamente utilizado tanto para estudios ambientales (e.j. Blanco et al 2001; Yuras et al 2005, Iriarte et al 2007, Correa-Ramirez et al 2007, Demarcq 2009, Vergara et al. 2017) como genéticos (e.j. Galindo et al 2010, Nuñez-Acuña et al 2012, Amaral et al 2012, Giles et al 2015). Sin embargo, se usó también la fluorescencia en complemento a la *clorofila a* debido a que esta representa un *proxi* directo de la actividad fotosintética (Leeuw et al 2013). Se decidió complementar con esta variable ambiental debido a que, dada la naturaleza del dato satelital, los datos de concentración de *clorofila a* podrían incluir además de clorofila, mediciones de otras sustancias ópticamente activas que son detectadas en el mismo espectro de colores que la clorofila (Leeuw et al 2013). También se incluyó concentración de carbono particulado tanto orgánico ( $\text{mol m}^{-3}$ , POC) como inorgánico ( $\text{mol m}^{-3}$ , PIC, i.e Carbonado de calcio o calcita). La concentración de

Carbono Orgánico Particulado (POC) incluye organismos tanto autotróficos como heterotróficos, así como partículas de detritos de origen biológico suspendidas en la columna de agua. Generalmente es ampliamente aceptado que el flujo vertical del POC es un reflejo de los procesos biológicos (es la vía principal de formación de carbono via fotosíntesis en la superficie), que ocurren en aguas superficiales (Wong et al 1999; Lampitt & Antia, 1997). Este tipo de compuestos particulados carbonatados (PIC, POC), han sido utilizados recientemente en estudios de genómica del paisaje marino (Giles et al 2015). Finalmente, se incluyó el promedio mensual de la temperatura superficial del mar (SST) (Tabla 1) debido a su importancia en la estructura biogeográfica de los organismos marinos tanto a nivel regional como global, siendo una de las variables ambientales más estudiadas (Roy et al 1994, Blanchette et al 2008, Fenberg et al 2015).

Para el set A2, se utilizó concentración de nutrientes como variables asociadas directamente a surgencias (Nitrato, Silicato y Fosfato), pH como proxy de la influencia de acidificación y nuevamente se consideró la SST. Esta última, se incluyó debido a que el intervalo de tiempo y resolución espacial es desigual en ambas bases de datos, por lo que se utilizó como control para determinar diferencias y similitudes con los resultados obtenidos con el set A1.

Para identificar loci candidatos a adaptación local asociados a estas variables ambientales se utilizaron en forma paralela dos familias de análisis; 1) Análisis basados en diferenciación poblacional y 2) Análisis de Asociaciones Ambientales (EAA) (Figura 1).



### ***Aproximaciones basadas en análisis de diferenciación poblacional (ADP)***

Para detectar loci putativamente sujetos a selección positiva se utilizaron tres aproximaciones independientes basadas en análisis de diferenciación poblacional (ADP). El uso de aproximaciones múltiples para la detección de SNPs putativamente sujetos a selección reduce la probabilidad de considerar falsos positivos en el set de datos final (Lotterhos & Whitlock 2015; Rellstab et al 2015). Para esto, se utilizaron los siguientes programas que implementan diferentes acercamientos: BayeScan 2.1 (Foll & Gaggiotti 2008), OUT-Flank (Whitlock & Lotterhos 2015) y PCAdapt (Luu et al 2017) (Figura 1).

BayeScan 2.1 implementa un método bayesiano basado en un modelo de regresión logística que identifica loci putativamente afectados por la selección (variación genética putativamente adaptativa locus-específica) de aquellos afectos a procesos demográficos población-específicos (modelo nulo o variación genética neutral). Este análisis permite detectar señales neutrales a la selección y señales de selección tanto positiva como balanceadora. Aquellos bajo selección balanceadora fueron eliminados del set de datos para posteriores análisis debido a que al ser loci con valores de diferenciación menor a la esperada bajo un régimen neutral, podrían enmascarar la estructura neutral. Se utilizó un modelo *a priori* (parámetro *pr\_odds*) incluyendo un total de 100.000 iteraciones y un periodo de *burn-in* de 10.000 pasos. Los resultados fueron corregidos usando una *False Discovery Rate* de 0,05 sobre el logaritmo de los valores de *q*. Los loci candidatos a selección

positiva considerados fueron aquellos con evidencia fuerte o muy fuerte de estar sujetos a selección de acuerdo al criterio de Jeffrey's (Jeffreys 1961) basado en los valores del *Bayes Factor* ( $bf > 10$ ).

OUT-Flank, otro de los acercamientos de ADP, estima una probabilidad basada en una distribución ajustada de los valores de  $F_{ST}$  para inferir la distribución de  $F_{ST}$  de loci putativamente neutrales. Esta distribución es utilizada para asignar valores de  $q$  a cada locus y detectar loci con valores *outlier* de  $F_{ST}$ . Para esto se eliminó una proporción de loci con valores extremos de  $F_{ST}$  en ambas direcciones usando una *LeftTrimFraction* y una *RightTrimFraction* de 0,05, además de una heterocigosidad mínima (*Hmin*) de 0,1. Los *outlier* fueron identificados en los 5 sitios de muestreo utilizando un intervalo de valores de  $q$  de 0,05.

Finalmente, PCadapt utiliza un análisis de componentes principales (CP) para determinar el porcentaje de varianza genética explicado por cada componente. Para esto, se ejecutó un primer análisis de CP usando un valor de número de agrupamientos (*K*) igual al número de poblaciones estudiadas (5). Luego, basado en este resultado se definió el valor óptimo de *K* basado en el valor más alto de *K* luego de que la estabilización de la curva de proporción de varianza explicada en función del número de componentes principales se estabilice (Luu et al 2017). De esta forma para un SNP dado, PCAdapt utiliza un test distancia de *Mahalanobis* para definir si un marcador es un *outlier* o no basado en los *loadings*, los que se definen como la correlación entre ese SNP y cada componente principal. Los gráficos *Manhattan* para definir los *outliers* fueron hechos usando el  $\log_{10}$  de los

valores de  $p$ . Adicionalmente, se corrigió la distribución de los valores de  $p$  usando una False Discovery Rate de 0,05 para obtener finalmente una lista de loci candidatos a selección (*outliers*).

Todos los ADP para la búsqueda de *outliers* fueron hechos en el set de datos completo de SNPs. Basado en los resultados de las tres aproximaciones, luego de remover los loci sujetos a selección balanceadora detectados por BayeScan, el set de datos fue dividido en loci putativamente neutrales y putativamente sujetos a selección positiva o loci candidatos a adaptación local (Figura 1).

***Estructura espacial y análisis de redundancia: Relación entre la estructura genética y variables ambientales.***

La estructura espacial fue estimada a través de las coordenadas geográficas de los sitios de muestreo mediante el uso de vectores de distancia basados en mapas de autovalores de Moran, dbMEMs (*distance-based Moran's eigenvector maps*) (Figura 1). Los dbMEMs corresponden a variables independientes que resumen la estructura espacial asociada a una matriz de distancia, en este caso de distancia geográfica (Borcard & Legendre 2010) y representan la distribución espacial en un plano cartesiano (proporcionales al índice de autocorrelación de Moran) (Dray et al 2006; Legendre & Legendre 2012). Para determinar los dbMEMs, se convirtieron tanto latitud como longitud a coordenadas cartesianas utilizando el paquete *SoDa* en R 3.22. Luego, se calculó una matriz de distancia euclidiana de estas coordenadas cartesianas (usando la función *dist* en *Vegan*) y finalmente los dbMEMs fueron estimados usando la función *create.dbMEM.model* en el paquete

*adespatial*, creando una matriz espacial en datos rectangulares con dos vectores, uno asociado a longitud (dbMEM1) y otro a latitud (dbMEM2), los que fueron utilizados posteriormente en los análisis de redundancia (RDA).

Los análisis de redundancia (RDA) fueron realizados usando los vectores espaciales dbMEMs, las variables ambientales y los genotipos asociados a loci neutrales y candidatos a selección positiva detectados por las ADP (Figura 1). El RDA es un análogo multivariado de la regresión lineal simple (Legendre & Gallagher 2001), que permite examinar qué tanta variación de los factores ambientales estudiados explica la variación a nivel genotípico (neutral y adaptativo), suponiendo una relación lineal entre variables ambientales/espaciales y frecuencias alélicas. Con estos análisis, se determinó la contribución relativa de cada variable ambiental a la estructura genética de ambos sets de loci (neutral y adaptativa).

Se ejecutaron tres RDA distintos (Figura 1). Primero un «RDA total» en el que se utilizaron tanto variables ambientales como vectores espaciales (dbMEMs) como factores fijos, sin controlar por ambiente y espacio, y dos «RDA parciales» los que permitieron condicionar por los efectos del espacio («RDA parcial-espacial») y del ambiente («RDA parcial-ambiental»). En el «RDA parcial-espacial» se utilizaron los vectores dbMEMs como covariables para controlar el efecto de la distribución espacial en la estructura genética y el «RDA parcial-ambiental» se usaron los vectores dbMEMs como factores fijos y los factores ambientales como variables condicionales (covariables) para determinar cuánto de la variación genética

(neutral y adaptativa) depende solo de la geografía. Para este análisis, se utilizó la función *rda* en el paquete *Vegan* en R v 3.22. Para los tres tipos de RDA ejecutados, se determinó el modelo óptimo respecto a los factores ambientales y/o espaciales que mejor explican la variabilidad genética usando la función *ordistep* en el paquete *Vegan* en R 3.22. Esta función selecciona las variables que constituyen al modelo óptimo de acuerdo a su significancia y al valor del criterio de información de Akaike (AIC). Para cada modelo se utilizó un ANOVA marginal usando 1000 permutaciones para evaluar la significancia de cada factor fijo considerado.

#### ***Análisis de Asociaciones Ambientales (EAA) y análisis de estructura espacial de loci adaptativos***

Para la identificación de loci candidatos a adaptación local y con el propósito de evitar la ocurrencia de falsos positivos en los análisis de asociaciones ambientales (EAA), se usó la señal de loci neutrales como control (DeMita et al 2013, Lotterhos & Whitlock 2015, de Villemereuil et al 2014, Frichot et al 2015). Para esto se utilizó un análisis discriminante de componentes principales (DAPC) de los SNP neutrales determinados con BayeScan (Figura 1), para establecer el número de grupos genéticos presentes en los datos usando el paquete *adegenet* en R 3.22. Para detectar el número de *clusters* que minimiza la variación entre grupos se utilizó un valor de K basado en un criterio de información Bayesiano (BIC). Esto permite identificar el número óptimo de *clusters* relajando los supuestos de

estructura genética subyacente o panmixia dentro de los grupos como hacen otras aproximaciones. Este valor de K fue posteriormente utilizado en dos de las tres aproximaciones de EAA (ver siguiente sección).

Una vez determinada la estructura neutral, para buscar asociaciones ambiente-loci SNPs se utilizaron tres aproximaciones de EAA (Figura 1) implementadas en: 1) LFMM v.1.4 (Frichot et al 2013), 2) Bayenv2 (Guenther & Coop 2013) y 3) Samβada (Stucki et al 2016).

En primer lugar, el análisis de Modelamiento Mixto de Factores Latentes (LFMM) permite determinar asociaciones entre variables ambientales y SNPs utilizando variables ambientales como efecto fijo, y la estructura genética neutral como factor latente o factor no observado (Frichot et al 2013). El programa fue ejecutado utilizando un número de factores latentes basado en los resultados de DAPC (K óptimo de la estructura neutral) usando 200.000 iteraciones, seguido de un *burn-in* de 10.000. Para cada análisis, se hicieron 5 corridas independientes para computar los distintos parámetros tal como recomienda Frichot et al (2013) y se utilizó la media de los *z-scores* para los resultados finales. Todo esto fue ejecutado en el paquete LEA en R 3.22 y la significancia de los valores asociados a cada SNP fue calculada usando un valor de alfa de 0,01 corregido con una *False Discovery Rate* (FDR).

La segunda aproximación de EAA utilizada fue Bayenv2 (Guenther & Coop 2013). Este programa utiliza un marco Bayesiano y permite evaluar correlaciones entre marcadores moleculares y factores ambientales poniendo a prueba si el

modelo que incluye el factor ambiental, en donde se evalúa cada variable ambiental por separado, se ajusta mejor a los datos respecto a un modelo nulo. Este modelo nulo corresponde a una matriz de covarianza genética (análoga a una matriz de  $F_{ST}$  entre pares de poblaciones) calculada a partir de las estimaciones de las frecuencias alélicas de loci neutrales. Para estimar esta matriz se hicieron cinco corridas independientes usando 100.000 iteraciones para evitar sesgos producto del número de iteraciones del MCMC y, una vez que se obtuvieron matrices no significativamente distintas entre sí, se generó una matriz promedio a partir de la última iteración de cada una de las 5 corridas (Frichot et al 2013). Luego, el set completo de SNPs fue individualmente puesto a prueba con las 10 variables ambientales previamente estandarizadas (media=0, varianza=1) usando una cadena de MCMC de 80.000 iteraciones con un intervalo de muestreo cada 500 iteraciones. Para identificar SNPs candidatos con evidencia de selección positiva a cada variable ambiental, se consideraron *bayes factor* >10 (Kass & Raftery 1995), de acuerdo al criterio de Jeffreys's (Jeffreys 1961).

Finalmente, se usó el programa *Samβada* (Stucki et al 2016) para poner a prueba asociaciones entre las frecuencias alélicas y parámetros ambientales entre sitios considerando la estructura neutral tanto basada en el número de grupos presentes (*k*-óptimo determinado por DAPC) como en autocorrelación espacial a partir de la información de las coordenadas geográficas de los sitios estudiados. *Samβada* implementa un análisis univariado de regresiones logísticas múltiples, en que los valores de *p* se calculan a partir de regresiones de todas las combinaciones

genético ambientales posibles de acuerdo a un análisis llamado *Wald Score Test*. Estos resultados son comparados a una distribución de Chi cuadrado con 1 grado de libertad. Los loci con valores de  $p$  menores a 0,05, los que fueron corregidos usando una FDR, fueron considerados como asociaciones significativas entre dicho marcador y la variable ambiental correspondiente.

Una vez determinados aquellos loci candidatos a adaptación local detectados en común por Bayenv2, LFMM, Samβada (Figura 1) para cada variable ambiental, éstos se extrajeron del set de datos total y fueron analizados de forma independiente en análisis posteriores.

A partir de los loci identificados en común por los tres métodos de EAA, se realizaron análisis espaciales de componentes principales (sPCA) utilizando el paquete *adegenet* (Jombart 2008) en R 3.22. sPCA es un análisis multivariado que busca agrupaciones genéticas discretas y/o clínicas en base a las frecuencias alélicas de los loci candidatos (Jombart et al 2008). Esta aproximación incluye tanto la información espacial, basada en las coordenadas geográficas, como los datos genéticos, y permite inferir si la variación en los SNPs candidatos está significativamente asociada a cada variable ambiental por sobre lo que se espera basado solo en la proximidad geográfica de cada sitio por separado.

Para ejecutar el sPCA, se usó un *network* espacial de proximidad entre localidades construido utilizando el método de *neighbourhood-by-distance* utilizando las coordenadas geográficas de cada sitio estudiado. A partir de esto se extrajeron los *lagged scores* (asociados a las primeras dos componentes principales) de cada



localidad desde el sPCA, los cuales reflejan la variabilidad genética que está ligada a la estructura espacial entre sitios. Estos *lagged scores* fueron usados para transformar la variación genética de los SNPs candidatos en clinas geográficas multilocus. Para evaluar si la variación de los SNPs candidatos se ajustó mejor a las variables ambientales o a las espaciales, a partir de las clinas multilocus, se realizaron regresiones lineales tanto con las variables ambientales como espaciales usando vectores dbMEMs en lugar de las coordenadas geográficas, evitando la violación de supuestos de independencia de los datos.

### **Ontología genética**

De forma exploratoria, se extrajeron las secuencias asociadas a los loci candidatos detectados (*tags*) para cada variable ambiental para identificar potenciales genes con funciones conocidas disponibles en la base de datos SWISS-PROT (Bairoch & Apweiler 2000) y en GeneBank (<https://www.ncbi.nlm.nih.gov/genbank/>). Debido al reducido tamaño de los *tags* (64 bp), se filtraron los resultados de acuerdo a su *e-value* tomando como valor mínimo un valor de  $1 \times 10^{-3}$  y un mínimo de 75% de homología para evitar falsos positivos. Las secuencias (*tags*) asociadas a los loci candidatos fueron usados como *query sequences* para los *BLAST* ejecutados en la base de datos de genes anotados de SWISS-PROT. Para el caso de la SST, debido a su importancia como factor ambiental, se exploraron también aquellos loci candidatos detectados por al menos dos aproximaciones y no solo aquellos detectados por las tres aproximaciones en común.

## Resultados

### *Detección de loci candidatos a selección natural.*

En total fueron genotipificados 2.902 SNPs para los 149 individuos analizados en 5 localidades lo largo del pacífico sureste (Tabla 2). Utilizando BayeScan, se detectaron un total de 897 loci neutrales, 1786 loci putativamente sujetos a selección balanceadora y 219 loci candidatos a selección diversificadora. De estos últimos, 179 muestran evidencia fuerte o muy fuerte de estar sujetos a selección diversificadora de acuerdo al criterio de Jeffreys (Figura 2). A su vez, OUT-Flank y PCAdapt detectaron un total de 94 y 124 loci putativamente sujetos a selección diversificadora, respectivamente. Hubo 75 loci candidatos a selección diversificadora detectados por los tres métodos (Figura 3) los que se usaron para analizar la estructura genética adaptativa. Por otra parte, los 897 loci neutrales determinados por BayeScan fueron utilizados tanto para analizar la estructura neutral, como para corregir los resultados de los EAA. Los 1786 loci putativamente sujetos a selección balanceadora detectados por BayeScan no fueron considerados.

A nivel neutral, DAPC identificó dos grupos genéticos: un grupo norte de los 30° (PA, 26°08'S; CP, 27°44'S; LH, 29°58'S) y un grupo al sur (TH, 36°38'S; AC, 41°52'S).

### ***Análisis espacial de estructura genética***

Incorporando tanto variables ambientales como espaciales como factores fijos («RDA total») para los 75 loci candidatos a selección diversificadora el modelo óptimo, y que por lo tanto incluye las variables más importantes para explicar la estructura genética adaptativa, es para el set A1 aquel que incluye *clorofila a*, fluorescencia y ambos vectores espaciales (Tabla 3) y para el set A2, el que incluye pH, nitrato y silicato (Tabla 4). A su vez, para los 897 loci neutrales, el modelo óptimo para el set A1 incluyó *clorofila a*, SST y ambos vectores espaciales dbMEMs (latitud y longitud) (Tabla 3) y el set de datos A2 incluyó SST, pH, fosfato y silicato (Tabla 4).

Los modelos óptimos controlando por el efecto del espacio («RDA parcial-espacial»), incluyeron solo dos variables ambientales de las 5 consideradas para ambos sets de datos ambientales. Para el set A1 este modelo incluyó *clorofila a* y SST y el set A2 incluyó pH y Silicato. Respecto a loci neutrales, el modelo parcial-espacial óptimo detectado para el set A1 rescató las mismas variables ambientales asociadas a loci *outlier* (SST y *clorofila a*), aunque solo el valor asociado a la *clorofila a* fue estadísticamente significativo ( $p=0,09$  y  $p=0,005$ , respectivamente. Tabla 3). Por otro lado, para el set A2 el modelo óptimo usando la estructura neutral difirió del modelo para loci candidatos, incluyendo como variables más importantes y significativas a la SST y al nitrato (Tabla 4).

Los «RDA parciales-ambientales», controlando por los efectos del ambiente, muestran que tomando en cuenta la longitud del vector, el dbMEM2 (latitud) es el que tiene mayor contribución relativa a la estructura genética neutral y adaptativa para ambos sets de datos ambientales (Figura 3c, 4c). Los modelos óptimos asociados a los «RDA parciales-ambientales» para loci *outlier* incluyen ambos vectores espaciales (latitud y longitud) para el set A1 (Tablas 3 y 4) y solo el vector asociado a latitud (dbMEM2) para el set A2 (Tabla 4). Para loci neutrales, ningún valor de *p* asociado a los vectores espaciales fue estadísticamente significativo (Tabla 3 y 4).

En resumen, para el set A1 al corregir por espacio, los análisis de RDA muestran señales coincidentes para loci *outlier* y neutrales, lo que se traduce en que tanto la SST como *clorofila a* sean los factores que más explican la variación genética tanto adaptativa como neutral. Lo anterior, sumado a valores significativos del vector espacial asociado a la latitud (Tabla 3, Figura 4C), sugiere que la contribución relativa de estas dos variables ambientales a la estructura genética adaptativa estaría altamente influenciada tanto por el espacio como por la estructura genética neutral. Este patrón también es posible observarlo en el modelo total, donde al incluir tanto variables ambientales como espaciales como factores fijos, el modelo óptimo incluyó *clorofila a* y ambos vectores espaciales tanto para loci *outlier* como neutrales. Por otro lado, el set de datos A2, los RDA parciales controlados por espacio muestran modelos disímiles para loci *outlier* y neutrales. Para loci neutrales, se observó una gran influencia de la SST en la estructura

genética, siendo el único factor ambiental que aparece tanto en el modelo total como en el parcial-espacial. Para este set de datos, también se observó influencia significativa del vector asociado a la latitud (dbMEM2), pero solo a para loci candidatos. Para el set A2 el análisis sugiere que al controlar la influencia del espacio, tanto pH como silicato fueron las variables que más explicaron la estructura genética adaptativa.

### ***Análisis de asociaciones ambientales (EAA)***

Para el set A1 de variables ambientales, los métodos de EAA detectaron un rango de 470-585 loci candidatos a adaptación local, mientras que para el set A2 se detectaron entre 375-666 loci candidatos, siendo ambas cifras sustancialmente mayores que las obtenidas con las aproximaciones basadas en ADP (Figura 5, 6).

Par el set de datos A1, la variable ambiental para la que se detectaron más loci candidatos fue la SST (406), y la menor cantidad total fue para *clorofila a* (235). Las tres aproximaciones en su conjunto detectaron un total de 34 loci en común, siendo la fluorescencia la variable con más loci detectados en común por los tres métodos de forma independiente (16) seguido de la SST (11) y la concentración de *clorofila a* y el carbono inorgánico particulado (PIC) las variables con menos loci detectados en común (2 loci cada una) (Figura 6). De estos 34 loci en total, obviando aquellos detectados para más de una variable ambiental, se detectaron un total de 27 loci candidatos a adaptación local. De estos, el 79.4% (24 loci) están contenidos en los 75 loci *outlier* detectados por las ADP.

Para el set A2, tanto el pH como la concentración de fosfato son las variables con una cantidad mayor de loci significativamente asociados de forma independiente por los tres métodos (402 y 378 respectivamente) detectando 28 y 26 loci en común, respectivamente (Figura 7). Para el silicato se detectaron 132 loci candidatos en total. Sin embargo, ninguno de estos fue detectado en común por los tres métodos (Figura 7). Para la SST, utilizada como variable control entre set de datos ambientales, se detectaron 10 loci en común de los cuales el 100% está contenido en los 11 loci detectados usando la base de datos A1, lo que sugiere que ambos sets de datos ambientales generan resultados consistentes para aproximaciones EAA.

De un total de 75 loci detectados en común para las variables ambientales del set A2, 48 loci se encuentran asociados a al menos dos variables, 18 a tres y 3 loci a cuatro de las cinco variables estudiadas (Figura 7). En total, 38 loci candidatos a adaptación local fueron detectados para las 5 variables ambientales de los cuales el 86.8% (33 loci) aparecen también contenidos en el set de 75 loci *outlier* detectados por las aproximaciones ADP.

En resumen, las variables ambientales para las que más loci candidatos fueron detectados por los tres métodos en común fueron, pH (28), fosfato (26), fluorescencia (16), SST (11) y nitrato (6).

### ***Clinas multilocus de las frecuencias alélicas***

Los análisis espaciales de componentes principales (sPCA) fueron realizados con los datos de los loci candidatos a adaptación local detectados en común por los tres métodos de EAA para cada variable ambiental. Por razones estadísticas se retuvieron solo sets con más de 5 loci. De las nueve variables ambientales estudiadas, las que cumplieron con este requisito, y por lo tanto aquellas variables ambientales detectadas como más relevantes para la estructura genética adaptativa de *Pyura chilensis* fueron: SST, pH, nitrato, fosfato y fluorescencia.

Las regresiones lineales de las clinas multilocus (*lagged scores*) extraídas desde los resultados del sPCA versus las variables ambientales y variables espaciales (dbMEMs) indicaron que las variables que mejor se ajustaron a la estructura genética adaptativa son el pH ( $R^2$  ajustado= 0.7464,  $p < 0.001$ , tabla 5, Figura 8) y a la SST ( $R^2$  ajustado: 0.7218,  $p < 0.001$ , Tabla 5, Figura 9). El ajuste lineal de ambas variables a la estructura genética adaptativa se explica en parte debido a que las clinas multilocus asociadas a cada sitio mostraron para ambos casos dos grupos geográficos con valores de *lagged scores* disímiles; uno al norte de los 30°S conformado por PA, CP y LH, donde fue posible observar diferencias leves entre sitios y otro al sur conformado por TH y AC con *lagged scores* muy similares entre sí (Figuras 8-9). Esto dos grupos de *lagged scores* fueron también detectados para el resto de las variables ambientales de importancia (fosfato, nitrato, fluorescencia) en donde los modelos de regresión asociados fueron también

significativos, pero con un menor ajuste lineal (Tabla 5, Figura 10-12). Interesantemente, los vectores espaciales (dbMEMs) mostraron correlaciones significativas con las clinas multilocus, pero los valores de  $R^2$  ajustado fueron cercanos a 0 (Tabla 5), observándose en todos los casos ajustes no lineales respecto a la latitud (Figuras 8-12).

Para todas las variables ambientales relevantes para la estructura genética adaptativa en el piure, salvo el fosfato, se observó un gradiente latitudinal (Figuras 8-12), y los SNPs candidatos a cada variable ambiental mostraron asociaciones más intensas con las variables ambientales que con los vectores espaciales. Esto sugiere que la estructura genética adaptativa del piure está más influenciada por diferencias ambientales al norte y sur de los 30°S que por la distancia entre los sitios estudiados.



### ***Ontología genética***

Se utilizó para la búsqueda de genes funcionales un total de 64 loci candidatos, los cuales correspondieron al total de loci detectados en común por las tres aproximaciones de EAA para ambas bases de datos ambientales sumado a aquellos loci significativamente asociados a SST detectados por al menos dos métodos de EAA. De estos, cinco loci arrojaron resultados significativos para genes anotados en la base de datos UniProt y en GeneBank, involucrados en diversas funciones biológicas. De estos loci, el TP53299 aparece con asociaciones significativas para



cinco variables ambientales (Tabla 6). Los genes asociados a estos loci muestran funciones importantes en la biogénesis de cilios y embriogénesis (gen CC2D2A, proteína HfTCL2), a procesos metabólicos asociados a funciones tanto en crecimiento como señales celulares (Tabla 6). El resto de los genes anotados encontrados corresponden a factores de transcripción, sistema nervioso, embriogénesis y desarrollo corporal en adultos.

## Discusión

### *Estructura genética adaptativa en Pyura chilensis*

Diversos estudios han mostrado diferencias a los 30°S para la costa Pacífico Sureste a nivel genético (Tellier et al 2009, Sanchez et al 2011, Brante et al 2012, Montecinos et al 2012, Haye et al 2014, Segovia et al 2017), ecológico/fisiológico (Thiel et al 2007, Gaitán-Espitia et al 2014, Gaitán-Espitia et al 2017, Tapia & Gallardo-Escárate 2015), y ambiental (Thomas et al 2001, Yuras et al 2005, Correa-Ramirez et al 2007, Torres et al 2011, Aravena et al 2014, Tapia et al 2014). Sin embargo, no hay estudios que hayan intentado evaluar de forma integrativa aquellos factores que podrían explicar la influencia de este de esta zona a nivel genético en los procesos de adaptación local de organismos marinos bentónicos. Como aproximación, este estudio corresponde al primer trabajo en donde se intentan identificar las variables ambientales que podrían modelar la

estructura genética adaptativa en invertebrados marinos en respuesta este quiebre biogeográfico.

Nuestros resultados indican que la estructura genética adaptativa en el piure se explica por las condiciones ambientales disímiles que se dan al norte y al sur del quiebre biogeográfico de los 30°S por sobre la distancia geográfica entre sitios. En este estudio se sugiere que los factores ambientales más importantes para la estructura genética adaptativa del piure son la temperatura superficial del mar (SST), el pH, la actividad fotosintética medida a través de los niveles de fluorescencia y los nutrientes nitrato y fosfato. Los análisis EAA permitieron por separado identificar una gran cantidad de loci candidatos a cada variable ambiental estudiada, siendo entre 0 a 28 los loci detectados en común por las tres aproximaciones utilizadas. Estos análisis determinaron por separado una cantidad ampliamente superior de loci candidatos que los análisis de ADP, lo que es consistente con lo reportado por otros estudios (Pritchard et al 2010, Rellstab et al 2015, Francois et al 2016, Benestan et al 2016). Esto es debido a que este tipo de aproximaciones buscan, en términos generales, correlacionar la estructura genética locus-por-locus a una variable ambiental dada, lo que permite una búsqueda específica en respuesta al ambiente. Las variables ambientales con menos loci significativamente asociados fueron el silicato (0), la concentración de clorofila a (2) y el carbono orgánico e inorgánico (2 y 3 loci, respectivamente). En el caso particular de la clorofila a, el bajo número de loci significativamente asociados detectados por los tres métodos de EAA puede explicarse debido a que nuestros

resultados indican que esta variable tiene una alta influencia tanto a nivel neutral como adaptativo (Tabla 3), lo que implica que gran parte de estos loci podrían corresponder a asociaciones ambientales que no son causales.

Los genotipos asociados a los loci candidatos muestran un ajuste significativo a las variables ambientales más allá de lo que se espera por simple proximidad entre sitios. Por otra parte, la estructura genética espacial de los loci candidatos a adaptación local muestran que, más que gradientes o clinas multilocus, los genotipos tienden a establecer consistentemente dos grupos; uno al norte y otro al sur de los 30°S, a pesar de que gran parte de las variables relevantes para la estructura genética del piure muestran un gradiente latitudinal. Sin embargo, se observó una leve variabilidad en las clinas multilocus hacia el norte de los 30°S (PA, CP, LH), particularmente en Coquimbo (LH). Esto puede deberse a que este sitio está emplazado entre dos centros importantes de surgencia (Punta Lengua de Vaca y Punta de Choros) que contribuyen a una alta productividad primaria en comparación con otras localidades más al norte (Thomas et al 2001). A pesar de estas leves variaciones, a nivel intragrupo (norte v/s sur) ambas regiones muestran un importante grado de cohesión (Figuras 6-12).

El rol del ambiente en la diferenciación genética al norte y sur de los 30°S ya había sido sugerido por Segovia et al (2017) para *P. chilensis* basado en las diferencias genéticas norte/sur usando marcadores sujetos a selección. Estas diferencias pueden explicarse debido a la susceptibilidad a perturbaciones

abióticas de la especie (Valdivia et al 2005, Cifuentes et al 2007), lo que limita su capacidad para colonizar ambientes a pesar de ser un competidor dominante por el sustrato (Ambler & Cañete 1991, Valdivia et al 2005). Su limitada capacidad colonizadora puede explicarse debido a que el asentamiento de las larvas es poco selectivo y depende primordialmente del flujo de agua y de la presencia de conoespecíficos, lo que se traduce en estas se asientan rápidamente en prácticamente cualquier sustrato duro (Manriquez & Castilla 2007). Esto implica que el rol del transporte antropogénico en el transporte físico de individuos entre regiones en el piure, podría provocar la ocurrencia de reclutas en condiciones sub-óptimas para su desarrollo.

Las diferencias ambientales, particularmente en la temperatura, repercuten de forma significativa en la actividad reproductiva tanto estacional como espacialmente en esta especie (Cea 1970, Cancino et al 1998). A pesar de que *P. chilensis* presenta ciclos reproductivos extensos (Cea 1969, Perez-Valdés et al 2017) tal como ocurre en general en otras ascidias (Wong et al 2011, Pineda et al 2013), y suelen estar reproductivos todo el año, la madurez sexual de los individuos tiende a ocurrir a temperaturas bajas. Esto coincide con lo reportado en la Bahía la Herradura en el piure donde la mayor ocurrencia de gametos maduros suele ser durante los meses del año con temperaturas más frías (Junio/Julio) (Ambler & Cañete 1991) y desde junio a septiembre para localidades al sur de los 30° (Perez-Valdés et al 2017, Cea 1970). En este contexto, estudios recientes sobre el ciclo reproductivo del piure sugieren que dado su amplio rango geográfico (10°-

44°S) (Lancelloti & Vásquez 2000) la especie podría tener un amplio rango de tolerancia, particularmente a cambios de temperatura (Perez-Valdés et al 2017), a pesar de que nuestros resultados sugieren fuertemente la ocurrencia de procesos de adaptación local asociados las condiciones ambientales al norte y sur de los 30°S. Esto también ha sido sugerido para el anómuro *Petrolisthes violaceus*, en donde el flujo genético entre regiones ha borrado la señal del quiebre filogeográfico a los 30°S (Haye et al 2014), pero sus amplios rangos fisiológicos le permiten tolerar cambios ambientales a lo largo de la HCS, no evidenciando señales de adaptación a regímenes locales de temperatura (Gaitán-Espitia et al 2014, Gaitán-Espitia et al 2017).

Otro efecto de la temperatura en la estructura genética adaptativa del piure podría tener relación con procesos fundamentales como alimentación y filtración mediada por cilios (Riisgard & Larsen 2001; Gili & Coma 1998). Esto es debido a que uno de los loci candidatos detectado en este estudio aparece asociado significativamente a un gen llamado *Coiled-Coil & C2 Domain Containing 2A* (CC2D2A) que tiene gran relevancia en la biogénesis de cilios (Gorden et al 2008; Bachmann-Gagescu 2011). La respuesta de este tipo de funciones a cambios ambientales ha sido ampliamente estudiada en el pasado. Pettersen et al (1999) sugiere que en ascídias hay una relación positiva entre la tasa de filtración y la temperatura, reflejándose también en cambios en la actividad ciliar, lo que también había sido observado en el chorito *Mytilus edulis* (Aiellio 1960; Yoneda 1962; Stefano et al 1977). A su vez, también observa una relación negativa de la

actividad ciliar con la concentración de células en el medio (Petersen et al 1999). Esto sugiere que cambios tanto en la concentración de partículas como en la temperatura, podrían repercutir de forma significativa en la tasa de filtración y ventilación en *P. chilensis*. Debido a que este tipo de respuestas son más bien fisiológicas que físico-mecánicas (Pettersen et al 1999), las funciones mediadas por cilios podrían responder de forma diferencial a variables ambientales relacionadas a temperatura y biomasa fitoplanctónica, lo que podría ser un factor primordial en la mantención de la diferenciación genética al norte y sur de los 30°.

Adicional a los efectos de la temperatura, la importancia de variables ambientales asociadas a surgencias en la mantención de la diferenciación genética adaptativa entre regiones en el piure podría deberse a que la zona de estudio se caracteriza por la presencia de dos zonas principales; Una desde los 30°S hasta los 18°S, caracterizada por vientos favorables para surgencia durante todo el año (Pizarro et al 1994), y una región central, desde los 30 hasta los 41°S, la cual muestra marcada estacionalidad para los períodos de surgencia, particularmente para la primavera y verano austral (Arcos & Navarro 1986, Aravena et al 2014) lo a su vez provoca la ocurrencia de aguas frías cerca de la costa desde los 30°S hacia el sur (Aravena et al 2014). Se necesitan estudios adicionales para establecer como los regímenes estacionales de surgencia afectan a *Pyura chilensis*, sin embargo, una primera aproximación indica que este tipo de cambios podrían ser fundamentales durante tanto en estadíos tempranos como adultos. Esto se sugiere debido a que uno de los genes que se ha identificado asociado significativamente a una variable

ambiental ligada a surgencias (concentración de clorofila a) corresponde a la proteína *Trunk lateral cell-specific* (HrTLC2). Esta proteína tiene diversos roles durante la embriogénesis y una importante labor en el crecimiento (desde juveniles a adultos) en la familia Pyuridae (*Halocynthia roretzi*) (Takahashi & Satoh 2001). La importancia de este tipo de proteínas para las ascidias radica en que son las únicas fuentes de células celómicas (células sanguíneas) para individuos juveniles, y también dan origen a hendiduras branquiales, epitelios ciliados en individuos adultos, y a células musculares longitudinales del manto y del sifón oral (Hirano & Nishida 1997, Takahashi & Sato 2001).

A pesar de que la identificación de esta proteína y es el único antecedente en el piure, la influencia de variables ambientales asociadas a temperatura y productividad que afectan a especies intermareales o submareales someras, han sido ampliamente reportadas tanto para la costa noreste el pacífico como para la costa sudeste (e.j. Menge 2000, Broitman et al 2001, Blanchette et al 2008, Espinosa et al 2017). En este tipo de estudios se sugiere justamente que los regímenes de surgencias influyen a las comunidades de forma subyacente tanto en tasas de reclutamiento como en biomasa y crecimiento individual de los organismos (Menge 2000, Menge & Menge 2013). De forma más específica, aquellas variables que funcionan como suministros ecológicos para la productividad primaria (e.j. nutrientes que permiten la ploriferación de *blooms* de fitoplancton), muestran que su variación respecto a la tasa de crecimiento de las especies varía a lo largo del gradiente de surgencias, en cuyo caso la respuesta más

alta ocurre cuando estos niveles son intermitentes (e.j. estacionales) (Menge & Menge 2013, Fenberg et al 2015).

Finalmente, adicional al rol de las surgencias, la influencia del pH, y de nutrientes como el fosfato, y de la actividad fotosintética (fluorescencia) podrían indicar influencia de aguas continentales, particularmente de descargas de ríos (Vargas et al 2017), en la estructura genética adaptativa del piure. A partir de la influencia de aguas continentales, también es posible dividir la zona de estudio dos; una zona con poca influencia (al norte de los 32°S), una zona estrecha de no más de 10km que se va ensanchando hacia los 36°S, y luego una zona bastante ancha, de unos 70km con bastante influencia de aguas continentales desde los 36°S hasta los 42°S, particularmente en estaciones más frías (Figuroa 2002, Thiel et al 2007). Para estas variables ambientales, las clinas multilocus asociadas a los loci candidatos a adaptación local muestran valores similares entre los dos sitios al sur de los 30°S, lo que sugiere que las frecuencias alélicas varían menos que lo esperado para la distancia entre estos sitios (~580 km). Esta «homogeneidad» en la estructura genética adaptativa hacia el sur de los 30°S, puede deberse a que la influencia de aguas continentales (glaciares y descargas de ríos) ocurre principalmente al sur de los 30° (Vargas et al 2017). Adicionalmente, la importancia del pH y el fosfato en la estructura genética adaptativa del piure podría sugerir la influencia de actividades antropogénicas. Gradientes en ambas variables pueden ser afectadas por actividades humanas (e.j. contaminación, descargas agrícolas, eutrofización) particularmente influenciado por rutas marítimas y



agricultura (Breitburg et al 2009). Si bien la influencia de actividades humanas ha mostrado tener efectos a nivel ecológico (Doney 2010), ecológico-fisiológico (Fabry et al 2008) y a nivel etológico y comunitario (Vargas et al 2017), se debe evaluar con estudios adicionales el rol eventual este tipo de factores en el piure.

***Aproximaciones basadas en análisis de diferenciación poblacional (ADP) versus las basadas en asociaciones genético-ambientales (EAA)***

La detección de adaptación local, particularmente para paisajes heterogéneos, aún no está optimizada para ningún método por si solo (Rellstab et al 2015). Por tal razón, se utilizaron dos familias de aproximaciones para la búsqueda de loci candidatos a adaptación local para las que se obtuvieron resultados contrastantes. Primero, en lugar de usar regresiones lineales entre distancias euclidianas y/o ambientales y valores de diferenciación genética ( $F_{ST}$ ) como es común en estudios de genética del paisaje marino (White et al 2010; Godhe et al 2013; Jorde et al 2015), se utilizaron análisis de redundancia (RDA), los que han sido ampliamente utilizados para correlacionar la estructura genética adaptativa con factores ambientales (e.j. Lasky et al 2012; Bourret et al 2012; McGaughan et al 2014; Benestan et al 2016; Gugger et al 2016). Sin embargo, dependiendo del set de datos ambientales utilizado, los resultados de los RDA tendieron a ser consistentes al comparar los modelos óptimos obtenidos usando la estructura genética neutral respecto a loci candidatos. Esto limita bastante la sensibilidad del análisis en la

detección de variables ambientales importantes para la estructura genética tanto adaptativa como neutral debido a que supone que todos los loci responden a una misma fuerza selectiva, no pudiendo poner a prueba ninguna hipótesis específica sobre la naturaleza de la selección. Estudios que controlen estos resultados comparándolos con estructura genética a nivel neutral o que comparen estos resultados con otras aproximaciones para su validación son escasos (ver Benestan et al 2016 para una excepción), lo que a la luz de estos resultados es altamente recomendable.

Ambas familias de análisis (APD v/s EEA) resultaron útiles para aproximarse a la búsqueda de factores ambientales relevantes para la estructura genética adaptativa de *P. chilensis* a lo largo de la zona de estudio. Sin embargo, los resultados basados en EEA parecen ser más robustos que los APD a la hora de asociar la estructura genética a la variación ambiental y por lo tanto para la detección de procesos de adaptación local. A pesar de esto, sobre el 85% de los loci detectados por las aproximaciones EEA fueron detectados también en por las aproximaciones APD, lo que sugiere que gran parte de la variación genética adaptativa detectada a partir de los análisis de diferenciación genética se debe a respuestas a factores ambientales. Esto sugiere no descartar *a priori* la utilización de aproximaciones basadas ADP para estudios de genética del paisaje marino, debido a que la combinación de estos resultados ha sido tanto en este estudio como en otros ejemplos en la literatura, un método eficiente tanto para determinar estructura genética neutral, identificar patrones asociados a adaptación local (Ganaire et al

2015, Rellstab et al 2015) como para manejar de forma óptima la ocurrencia de falsos positivos (Villemereuil et al 2014, Francois et al 2016).

### **Conclusión**

Los resultados de este estudio muestran una gran influencia de la temperatura, así como de las variables asociadas a productividad y surgencias en la estructura genética adaptativa de *P. chilensis*. Esta especie muestra un ejemplo interesante de cómo la heterogeneidad del paisaje puede contraponerse los efectos del flujo genético y provocar procesos de adaptación local.

En este estudio, hemos mostrado cómo a pesar del importante rol del transporte antropogénico en el flujo genético entre regiones en la estructura genética neutral, esta especie parece ser sensible a cambios ambientales particularmente a aquellos que influyen en sus funciones básicas (reclutamiento, filtración, crecimiento, etc), mostrando una fuerte estructura genética asociada al quiebre biogeográfico de los 30°S. Este tipo de susceptibilidad a variación ambiental resulta fundamental tomando en cuenta los potenciales impactos de actividades humanas en organismos marinos y la importancia de la especie a nivel ecológico (especie bioingeniera).

La identificación parcial de genes con función conocida reportados en este estudio, abre una nueva perspectiva en el estudio de esta especie como modelo, no solo para investigar procesos de adaptación local sino también, dada la susceptibilidad de *P. chilensis* a factores abióticos, para investigar la resiliencia de

organismos marinos bentónicos ante un continuo y progresivo calentamiento y acidificación de la HCS.

Si bien este estudio no está diseñado para determinar la ubicación del quiebre filogeográfico de los 30-33°S, las variables ambientales utilizadas podrían ser un primer paso para identificar aquellos factores selectivos que mantienen la estructura genética contemporánea al norte y sur de los 30°S para otras especies marinas bentónicas.

## Tablas y Figuras

**Tabla 1: Variables ambientales incluidas en los análisis de genética del paisaje marino. Se indica la variable ambiental (con su acrónimo en paréntesis), la fuente de datos de origen, la resolución temporal (en caso de estar disponible) y espacial, la fecha de inicio y término del rango de datos considerado y la unidad de medida de la variable estudiada. Para los datos obtenidos desde Aqua-MODIS (Set A1), se escogió un intervalo de 10 años hasta la fecha de muestreo (2005-2015). Para el set de datos extraído desde Bio-Oracle (Set A2), el intervalo temporal viene previamente determinado.**

Set	Variable	Fuente	Res. Tem	Res. Esp	Fecha Inicio	Fecha término	Unidad
Set A1	Sea Surface Temperature at 4 microns (SST)	Aqua-MODIS	Mensual	4 km	2005	2015	C
	Chlorophyll a concentration (Chl-a)	Aqua-MODIS	Mensual	4 km	2005	2015	mg m <sup>-3</sup>
	Concentration of Particulate Organic Carbon (POC)	Aqua-MODIS	Mensual	4 km	2005	2015	mg m <sup>-3</sup>
	Normalized fluorescence line height (Flu)	Aqua-MODIS	Mensual	4 km	2005	2015	mW cm <sup>-2</sup> um <sup>-1</sup> sr <sup>-1</sup>
	Concentration of Particulate Inorganic Carbon (PIC)	Aqua-MODIS	Mensual	4 km	2005	2015	mol m <sup>-3</sup>
	Sea Surface Temperature (SST) /Control	Bio-Oracle/Aqua-MODIS	Mensual	5 arcmin (9.2km)	2002	2009	°C
Set A2	Nitrate	Bio-Oracle/World Ocean Data BASE	-	Medido In Situ	1928	2008	umol/l
	pH	Bio-Oracle/World Ocean Data BASE	-	Medido In Situ	1910	2007	Unitless
	Phosphate	Bio-Oracle/World Ocean Data BASE	-	Medido In Situ	1910	1986	umol/l
	Silicate	Bio-Oracle/World Ocean Data BASE	-	Medido In Situ	1930	2008	umol/l

**Tabla 2. Localidades de muestreo en la ascidia *Pyura chilensis* en el pacífico sureste. Nombre de cada localidad, coordenadas, código de cada localidad y número de individuos genotipificados (N).**

Sitio	Coordenadas	Código	N
Pan de Azúcar	26°08'S; 70°39'W	PA	30
Caleta Pajonales	27°44'S; 71°02'W	CP	29
La Herradura	29°58'S; 71°21'W	LH	30
Talcahuano	36°38'S; 71°21'W	TH	30
Ancud	41°52'S; 73°50'W	AC	30
<b>TOTAL</b>			<b>149</b>

**Tabla 3: Resultados de los análisis de redundancia (RDA) para determinar la contribución relativa de cada variable tanto ambiental (set de datos A1) como espacial (vectores dbMEMs) a la estructura genética adaptativa y neutral. Se muestran los modelos óptimos, basado en el valor de p y en el criterio de información de Akaike (AIC), obtenidos para los análisis totales (variables ambientales y espaciales como factores fijos), parcial-espacial (controlando el efecto del espacio) y parcial-ambiental (controlando por efecto del ambiente). Valores significativos asociados a cada F-ratio (F) en negrita. Los acrónimos de las variables ambientales corresponden a los mismos utilizados en la tabla 1.**

SNPs	RDA	Análisis	Variable	AIC	F	p value
75 outlier	RDA Total	Modelo óptimo	<i>Genotipos ~ Chl-a + Flu + dbMEM1 + dbMEM2</i>			
			dbMEM1	838.36	1.6157	0.09
			Flu	838.63	1.8784	<b>0.035</b>
			Chl-a	839.66	2.8859	<b>0.005</b>
	RDA Parcial	Modelo óptimo Espacial	dbMEM1	839.98	3.204	<b>0.005</b>
			<i>Genotipos ~ Chl-a + SST + Condicion (dbMEM1 + dbMEM2)</i>			
			SST	838.63	1.965	<b>0.04</b>
			Chl-a	839.84	3.1516	<b>0.005</b>
			<i>Genotipos ~ dbMEM1 + dbMEM2</i>			
			dbMEM1	839.05	1.8043	<b>0.02</b>
856 neutral	RDA Total	Modelo óptimo	<i>Genotipos ~ Chl-a + dbMEM1 + dbMEM2 + SST</i>			
			SST	1231.6	1.5121	0.07
			dbMEM1	1231.8	1.7265	<b>0.04</b>
			dbMEM2	1232.6	2.5205	<b>0.015</b>
	RDA Parcial	Modelo óptimo Espacial	Chl-a	1233.7	3.5921	<b>0.005</b>
			<i>Genotipos ~ Chl-a + SST + Condicion (dbMEM1 + dbMEM2)</i>			
			SST	1231.6	1.5121	0.09
			Chl-a	1233.7	3.5921	<b>0.005</b>
			<i>Genotipos ~ dbMEM2</i>			
			dbMEM2	1232.4	1.4214	0.08

Tabla 3: Resultados de los análisis de redundancia (RDA) para determinar la contribución relativa de cada variable tanto ambiental (set de datos A2) como espacial (vectores dbMEMs) a la estructura genética adaptativa y neutral. Se muestran los modelos óptimos, basado en el valor de p y en el criterio de información de Akaike (AIC), obtenidos para los análisis totales (variables ambientales y espaciales como factores fijos), parcial-espacial (controlando el efecto del espacio) y parcial-ambiental (controlando por efecto del ambiente). Valores significativos asociados a cada F-ratio (F) en negrita. Los acrónimos de las variables ambientales corresponden a los mismos utilizados en la tabla 1.

SNPs	RDA	Análisis	Variable	AIC	F	p value	
75 <i>Outlier</i>	RDA Total	Modelo óptimo	<i>Genotipos ~ Nitrato + pH + Silicato + SST</i>				
			pH	839.54	2.5785	<b>0.015</b>	
	RDA Parcial	Modelo óptimo Espacial	SST	840.31	3.3355	<b>0.01</b>	
			Silicato	840.21	3.2329	<b>0.005</b>	
			Nitrato	840.41	3.4334	<b>0.005</b>	
			<i>Genotipos ~ pH + Silicato + Condicion (dbMEM1 + dbMEM2)</i>				
	856 <i>neutral</i>	RDA Total	Modelo óptimo	<i>Genotipos ~ dbMEM2</i>			
				dbMEM2	839.05	12.7456	<b>0.005</b>
		RDA Parcial	Modelo óptimo Ambiental	dbMEM1	849.75	1.6701	0.065
				<i>Genotype ~ pH + fosfato + Silicato + SST</i>			
SST				1234.5	3.9859	<b>0.01</b>	
Fosfato				1234.3	3.8559	<b>0.005</b>	
856 <i>neutral</i>	RDA Total	Modelo óptimo	<i>Genotipos ~ Nitrato + SST + Condicion (dbMEM1 + dbMEM2)</i>				
			Nitrato	1233.6	3.334	<b>0.005</b>	
	RDA Parcial	Modelo óptimo Espacial	SST	1233.7	3.4018	<b>0.005</b>	
			<i>Genotipos ~ dbMEM1 + dbMEM2</i>				
			dbMEM1	1232.9	0.7135	0.885	
			dbMEM2	1233.5	1.2612	0.14	



**Tabla 5: Resultados de los modelos de regresión usando los *lagged scores* del análisis espacial de componentes principales (sPCA) y los valores asociados a cada variable ambiental y/o espacial (usando el vector dbMEM2, asociado a la latitud).**

Variable ambiental	Modelo	R <sup>2</sup> adj	<i>p value</i>
Sea Surface Temperature (SST)	Ambiental	0.7218	<0.001
	Espacial	0.0538	0.003
Fluorescence	Ambiental	0.3821	<0.001
	Espacial	0.0860	<0.001
pH	Ambiental	0.7474	<0.001
	Espacial	0.0612	0.001
Phosphate	Ambiental	0.6522	<0.001
	Espacial	0.0610	0.001
Nitrate	Ambiental	0.5958	<0.001
	Espacial	0.0370	0.01

**Tabla 6:** Caracterización de los *match* obtenidos mediante análisis de BLAST tanto en GeneBank como en la base de datos de UniProt para loci candidatos a adaptación local en *Pyura chilensis*. Se especifica el locus, el número y variables ambientales asociados, e-values, homología de las secuencias, parámetros taxonómicos (especie, familia clase) del mejor *match*, el nombre del gen/proteína y su función biológica general en caso de estar disponible. \*SST genes identificados solo para temperatura superficial del mar (SST) en loci detectados por al menos dos métodos de EAA.

Locus	Número de variables ambientales asociadas	Variables ambientales asociadas	UniProt ID	e-value	Ident.	Especie, familia, clase	Proteína	Nombre del gen	Función
TP53299	5	pH, phosphate, nitrate, SST, PIC	K7F5L5	1.00E-03	0.762	<i>Petodiscus sinensis</i> / Trionychidae / Testudines	Coiled-coil and C2 domain containing 2A	CC2D2A	Biogénesis y degradación de cilios
TP77542	3	pH, phosphate, fluorescence	H0YM91	2.00E-04	0.944	<i>Homo sapiens</i> / Homínidos / Mammalia	Brachyury protein	T	Factor de transcripción
TP2255	3	pH, phosphate, fluorescence	A0A1W2WCM2	7.00E-02	0.75	<i>Ciona savignyi</i> / Cionidae / Ascidiacea	cyclin-dependent kinase 5 activator 1-like	CDK5R1	Ritmos biológicos y desarrollo del sistema nervioso central.
TP73388	1	Chlorophyll a	Q9GR85	2.00E-05	0.857	<i>Halocynthia roretzi</i> / Pyuridae / Ascidiacea	HFTLC2 protein	RNA binding	Embriogénesis y formación del cuerpo.
TP57884	1	SST*	F7A786	5.00E-02	0.75	<i>Ciona intestinalis</i> / Cionidae / Ascidiacea	Phospholipid scramblase	PLSCR	Actividad fosfolipídica enzimática (escramblasa)

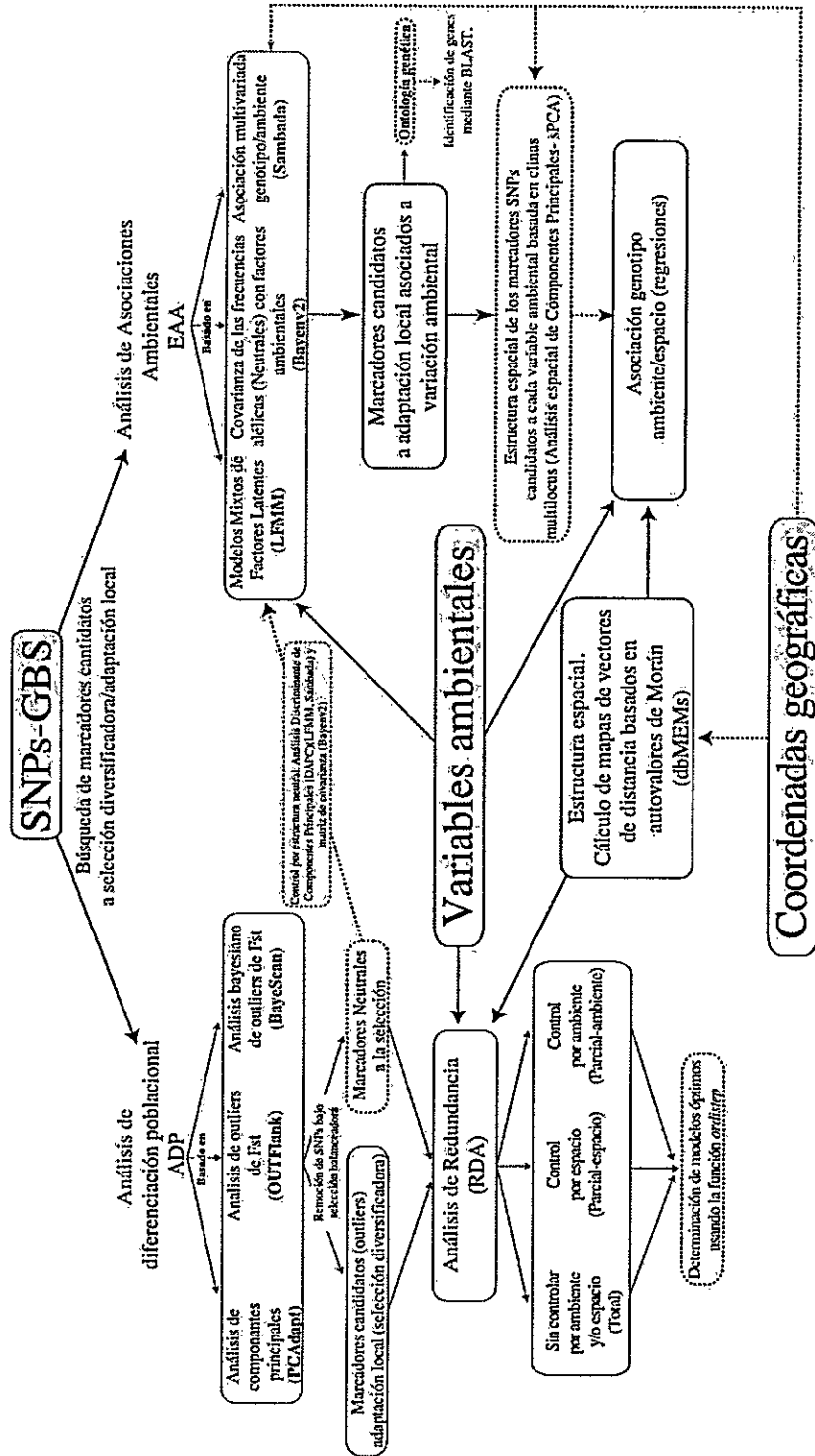


Figura 1: Diagrama de flujo del análisis de datos de genética del paisaje marino de SNPs GBS para *Pyura chilensis*. Se muestran los programas utilizados para cada aproximación (ADP y EAA) para la determinación de loci candidatos a adaptación local y cómo se relacionan entre sí a través del control por estructura genética neutral. Las celdas en color corresponden a los tres tipos de datos utilizados: SNPs obtenidos mediante Genotyping-By-Sequencing (GBS), Variables ambientales (obtenidas desde dos bases de datos) y las coordenadas geográficas de cada sitio estudiado.

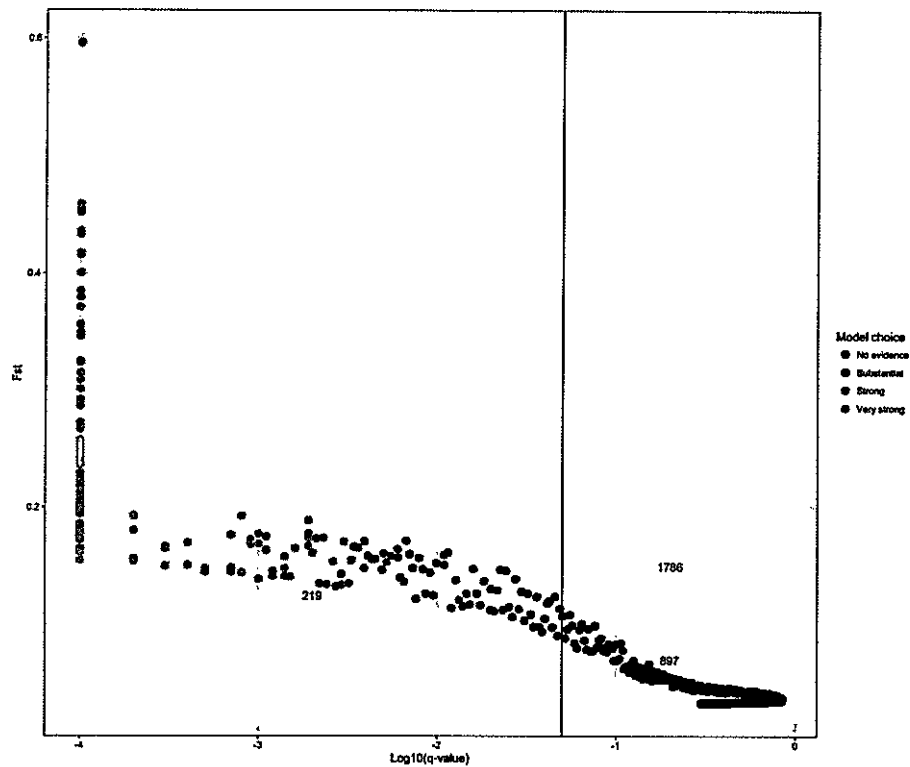
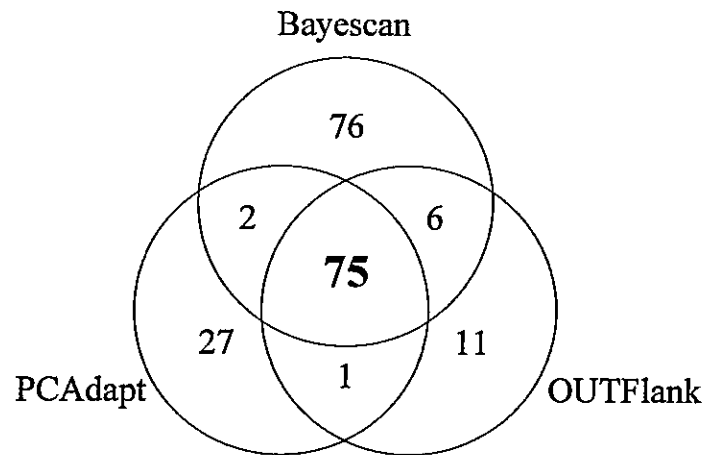
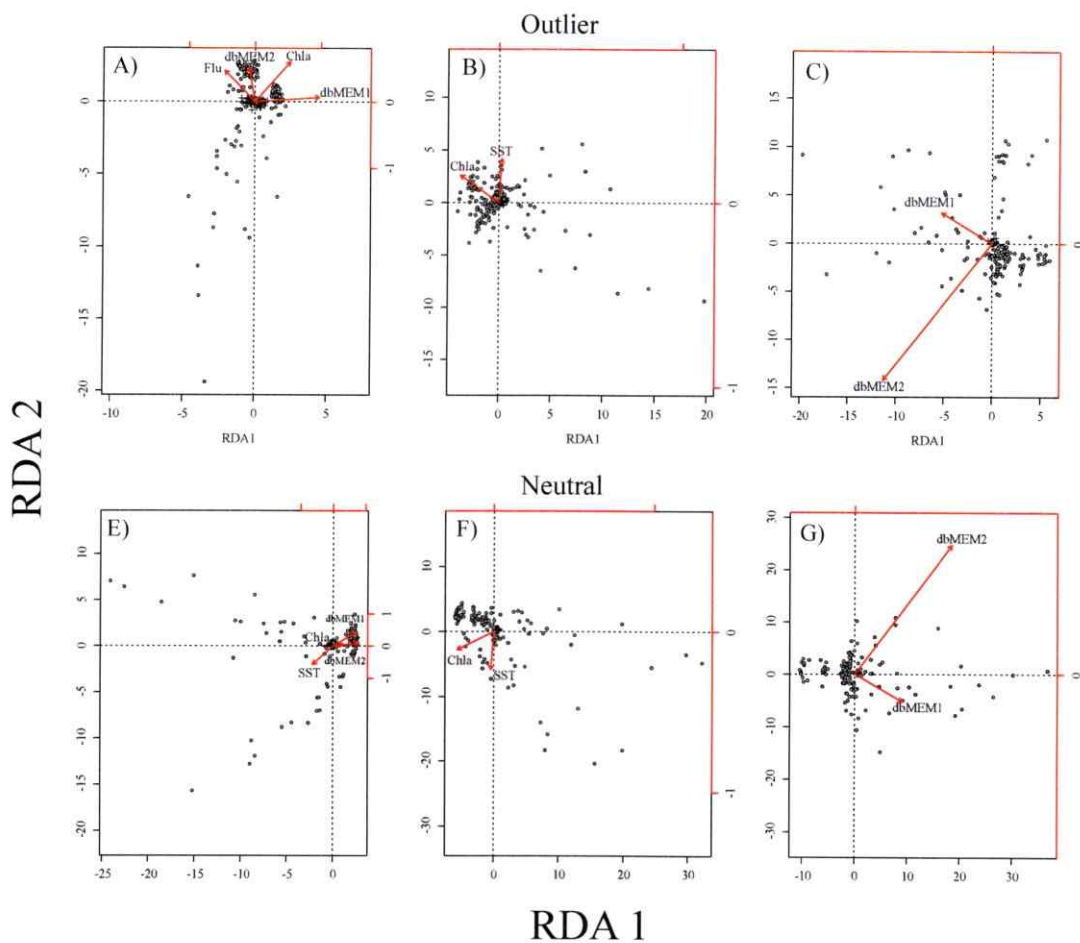


Figura 2: Resultados de análisis en BayeScan 2.0 para la identificación de SNPs *outlier* dentro de los 2902 loci analizados. El eje X corresponde al  $\log$  de la probabilidad posterior (valores de  $q$ ) y la línea vertical representa el límite de la probabilidad posterior usando una *False Discovery Rate* (FDR) de 0,05. Aquellos valores con  $q < 0,05$  fueron considerados como putativamente neutrales, y aquellos con  $q > 0,05$  como *outlier* en tres categorías de acuerdo su *bayes factor* (Jeffreys 1961): evidencia sustancial ( $3 < bf < 10$ ), evidencia fuerte ( $10 < bf < 30$ ) y muy fuerte ( $bf > 30$ ). Los 1786 loci sujetos a selección balanceadora no fueron considerados para análisis posteriores.



**Figura 3: Diagrama de Venn del número de loci candidatos a selección diversificadora detectados mediante tres métodos de análisis de diferenciación poblacional (ADP): BayeScan, OUTFlank y PCAapt, indicando los exclusivos de cada método y los comunes entre dos y tres métodos (valor en negrita).**



**Figura 4:** Análisis de redundancia (RDA) usando el paquete *vegan* para determinar la contribución relativa de variables tanto ambientales (set A1) como espaciales en la estructura genética adaptativa (*outlier*, fila superior) como neutral (fila inferior) en *Pyura chilensis*. A) y E) corresponden al modelo RDA total óptimo utilizando tanto variables ambientales como espaciales como factores fijos para loci *outlier* y neutrales, respectivamente B) y F) corresponde al modelo RDA parcial-espacial óptimo usando variables ambientales como efecto fijo controlando por los vectores espaciales dbMEMs para latitud y longitud en *outlier* y neutrales, respectivamente. C) y G) corresponden a los modelos parciales-ambientales usando vectores espaciales como factores fijos y variables ambientales como covariables condicionales para *outlier* y neutrales, respectivamente. Los biplot representan a los autovalores y longitudes de los vectores obtenidos para cada variable. Las siglas utilizadas son las mismas que la tabla 3.

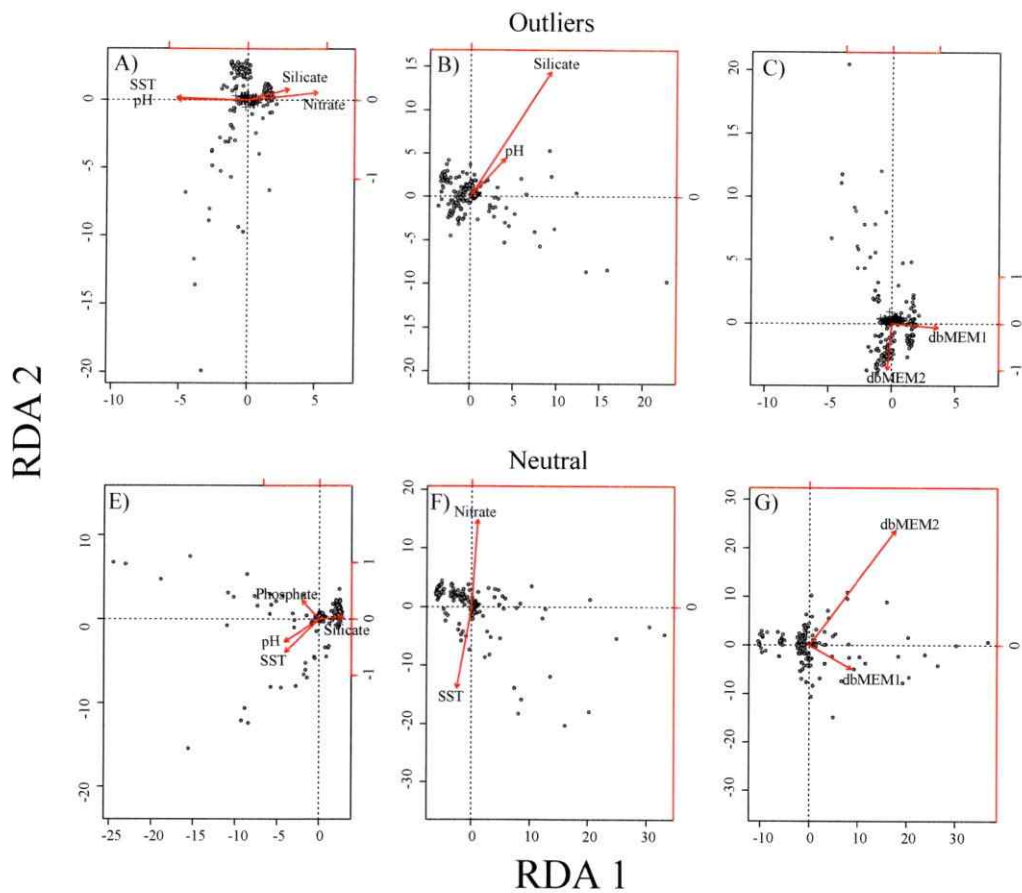


Figura 5: Análisis de redundancia (RDA) usando el paquete *vegan* para determinar la contribución relativa de variables tanto ambientales (set A2) como espaciales en la estructura genética adaptativa (*outlier*, fila superior) como neutral (fila inferior) en *Pyura chilensis*. A) y E) corresponden al modelo RDA total óptimo utilizando tanto variables ambientales como espaciales como factores fijos para loci *outlier* y neutrales, respectivamente B) y F) corresponde al modelo RDA parcial-espacial óptimo usando variables ambientales como efecto fijo controlando por los vectores espaciales dbMEMs para latitud y longitud en *outlier* y neutrales, respectivamente. C) y G) corresponden a los modelos parciales-ambientales usando vectores espaciales como factores fijos y variables ambientales como covariables condicionales para *outlier* y neutrales, respectivamente. Los biplot representan a los autovalores y longitudes de los vectores obtenidos para cada variable. Las siglas utilizadas son las mismas que la tabla 3.

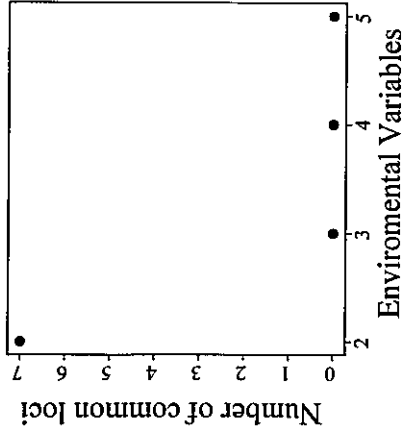
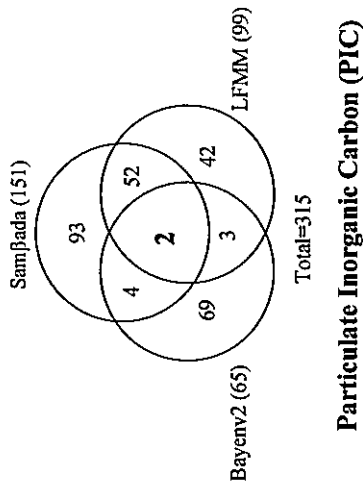
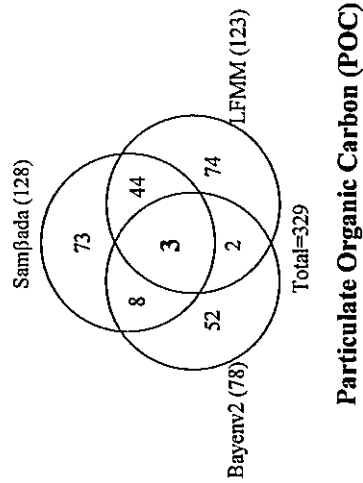
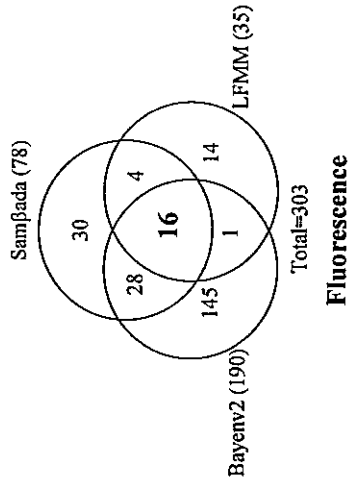
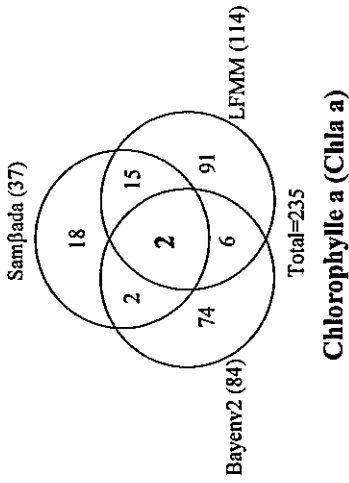
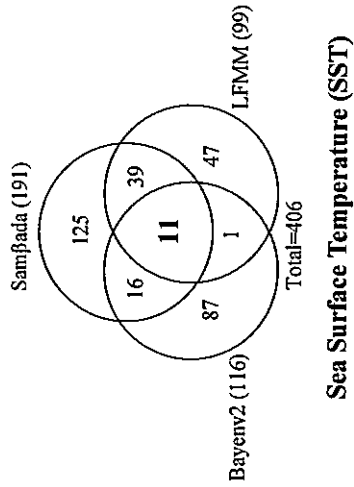


Figura 6: Diagrama de Venn del número de loci candidatos a adaptación local detectados por LFM, Bayenv2 y Samβada para *Pyura chilteris* usando el set de datos ambiental A1 (Aqua-Modis/NASA). El valor en negrita corresponde al número de loci candidatos detectados en común por los tres métodos. El gráfico de frecuencias (esquina inferior derecha) muestra el número de loci candidatos que aparecen significativamente asociados a más de una variable ambiental.



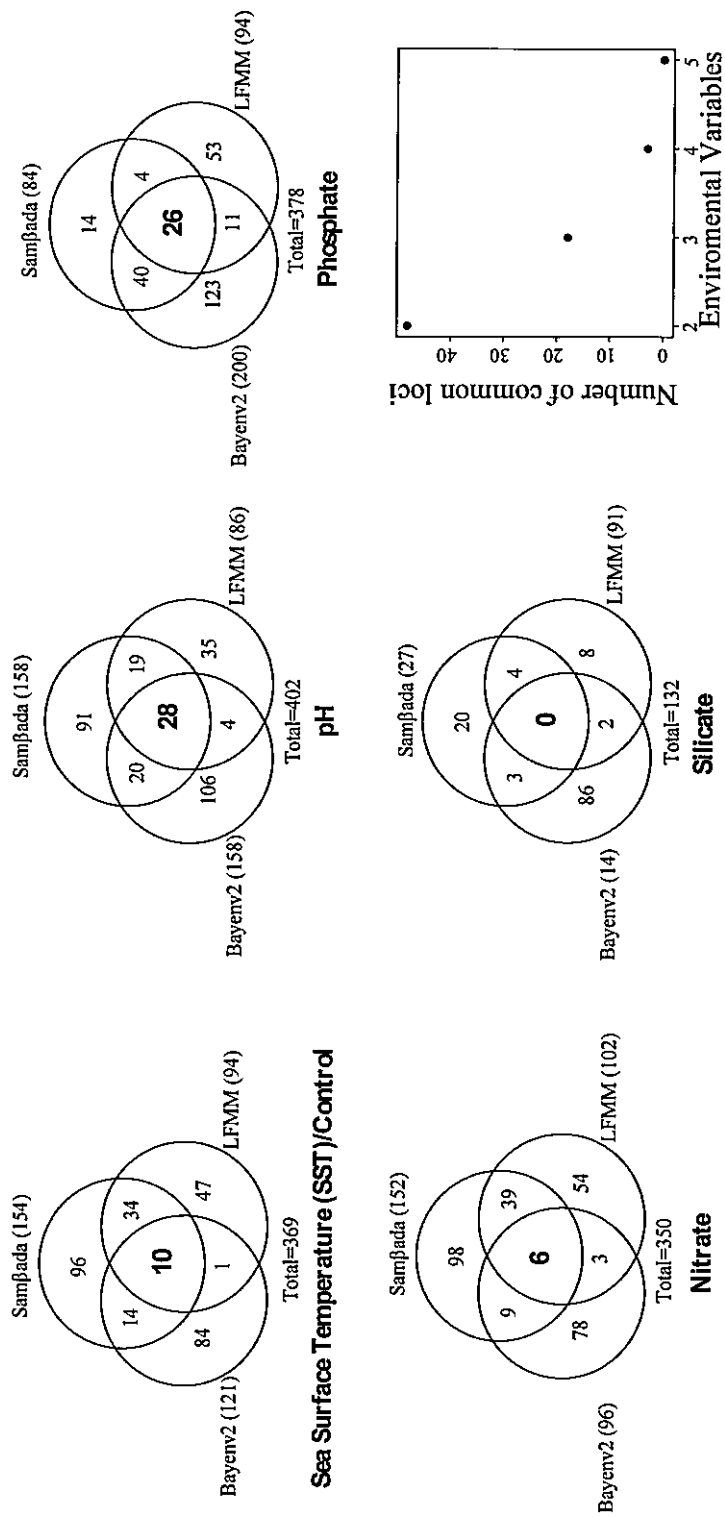


Figura 7: Diagrama de Venn del número de loci candidatos a adaptación local detectados por LFM, Bayenv2 y Samβada para *Pyura chilensis* usando el set de datos ambiental A2 (Bio-Oracle). El valor en negrita corresponde al número de loci candidatos detectados en común por los tres métodos. El gráfico de frecuencias (esquina inferior derecha) muestra el número de loci candidatos que aparecen significativamente asociados a más de una variable ambiental.

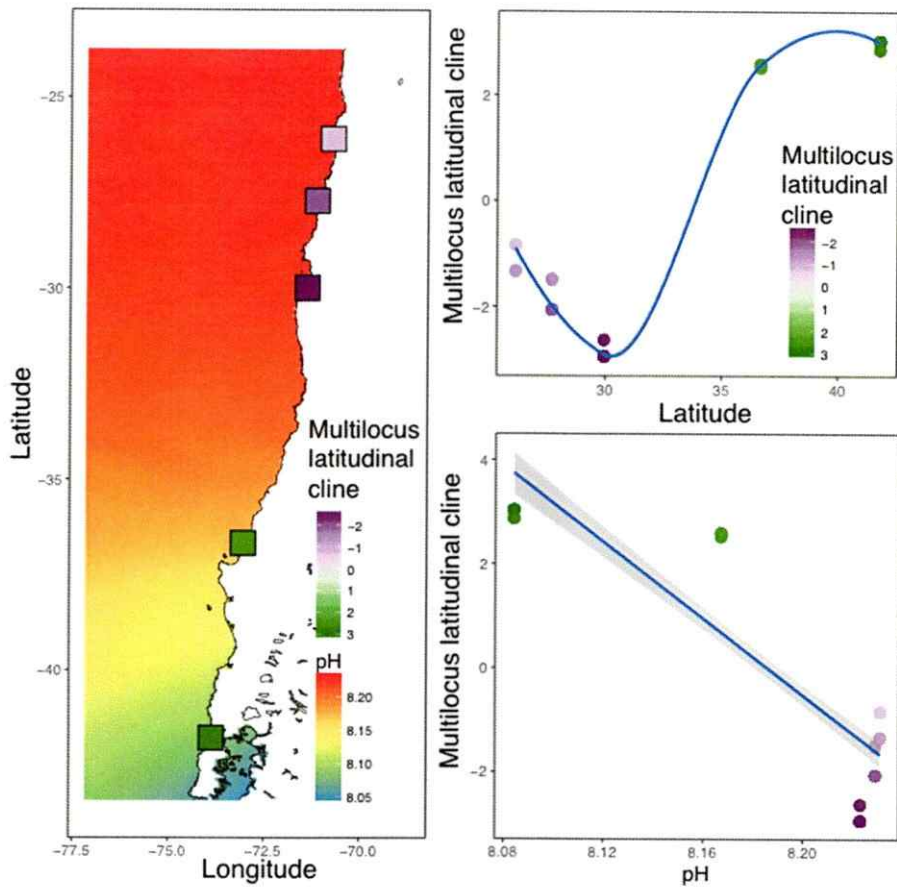


Figura 8: Variación genética adaptativa espacial de los 28 loci candidatos a adaptación local para la variable ambiental pH en los 5 sitios estudiados en *Pyura chilensis*. A) Estructura espacial del pH a lo largo de la zona estudio (El color del agua representa a la variación del pH) y valores de las clinas multilocus para cada sitio medidas mediante los *lagged scores* del análisis de sPCA, los que reflejan la variabilidad genética que está ligada a la estructura espacial entre sitios. Con fines gráficos, se utilizó la misma escala de colores para los valores de los *lagged scores* y el pH. B) Gráfico de dispersión de las clinas multilocus determinadas en el análisis sPCA para cada localidad considerando la relación entre la latitud y los dos primeros ejes de los *lagged scores*. El gradiente de colores es el mismo usado para el panel A). C) Ajuste lineal entre las clinas multilocus de los loci candidatos (*lagged scores*) respecto a la variación de los valores del pH en cada sitio.

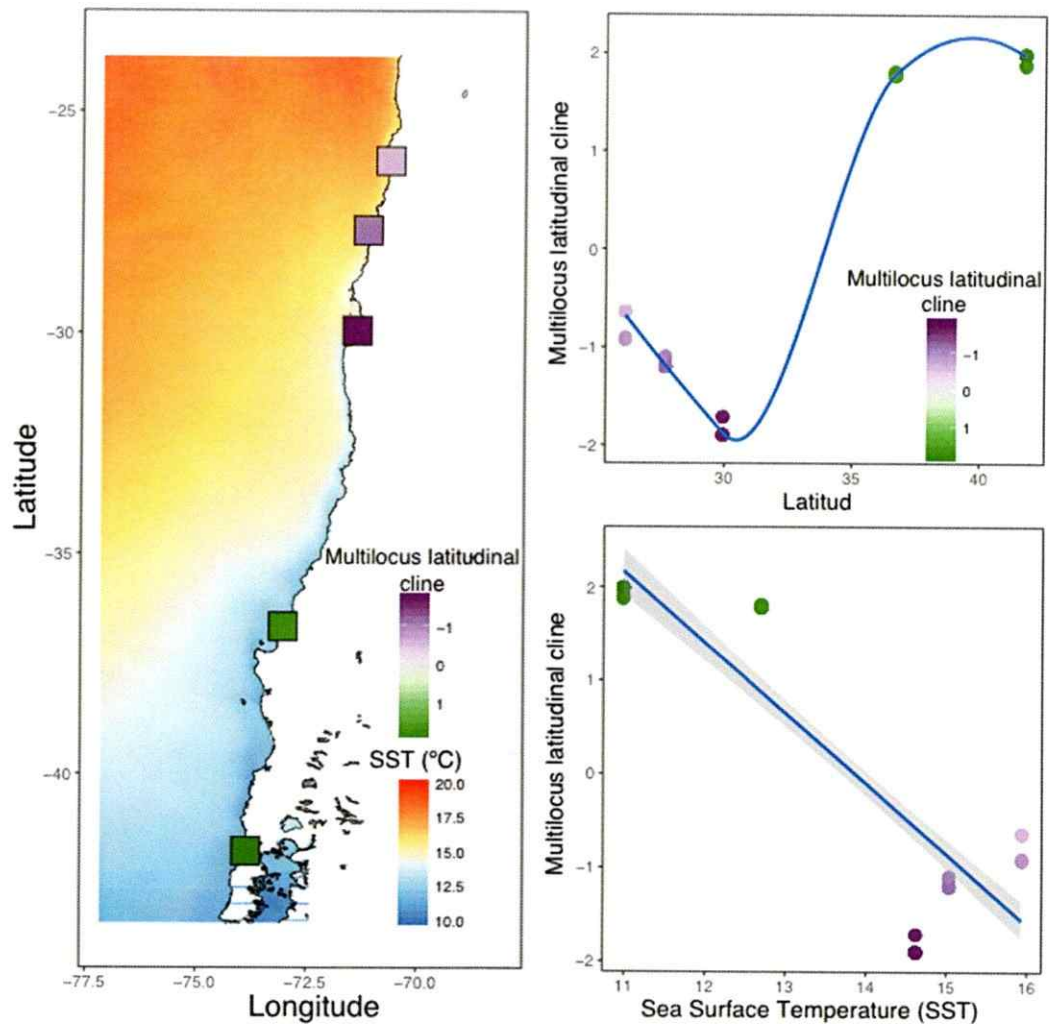


Figura 9: Variación genética adaptativa espacial de los 11 loci candidatos a adaptación local para la variable ambiental SST para los 5 sitios estudiados en *Pyura chilensis*. A) Estructura espacial de la SST a lo largo de la zona estudio (El color del agua representa a la variación de la SST) y valores de las clinas multilocus para cada sitio medidas mediante los *lagged scores* del análisis de sPCA, los que reflejan la variabilidad genética que está ligada a la estructura espacial entre sitios. Con fines gráficos, se utilizó la misma escala de colores para los valores de los *lagged scores* y la SST. B) Gráfico de dispersión de las clinas multilocus determinadas en el análisis sPCA para cada localidad considerando la relación entre la latitud y los dos primeros ejes de los *lagged scores*. El gradiente de colores es el mismo usado para el panel A). C) Ajuste lineal entre las clinas multilocus de los loci candidatos (*lagged scores*) respecto a la variación de la SST en cada sitio.

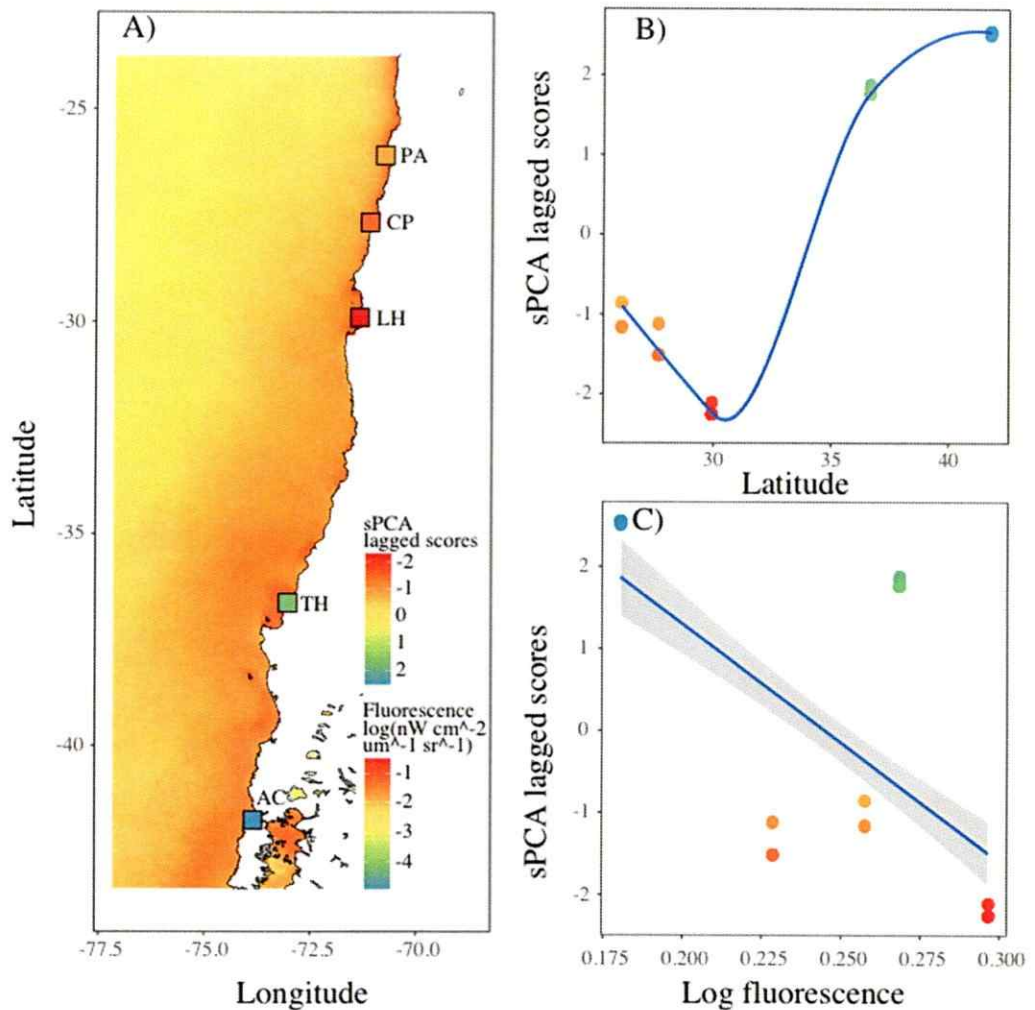


Figura 10: Variación genética adaptativa espacial de los 16 loci candidatos a adaptación local para la variable ambiental fluorescencia para los 5 sitios estudiados en *Pyura chilensis*. A) Estructura espacial de la fluorescencia a lo largo de la zona estudio (El color del agua representa a la variación de la fluorescencia) y valores de las clinas multilocus para cada sitio medidas mediante los *lagged scores* del análisis de sPCA, los que reflejan la variabilidad genética que está ligada a la estructura espacial entre sitios. Con fines gráficos, se utilizó la misma escala de colores para los valores de los *lagged scores* y fluorescencia. B) Gráfico de dispersión de las clinas multilocus determinadas en el análisis sPCA para cada localidad considerando la relación entre la latitud y los dos primeros ejes de los *lagged scores*. El gradiente de colores es el mismo usado para el panel A). C) Ajuste lineal entre las clinas multilocus de los loci candidatos (*lagged scores*) respecto a la variación de la fluorescencia en cada sitio.

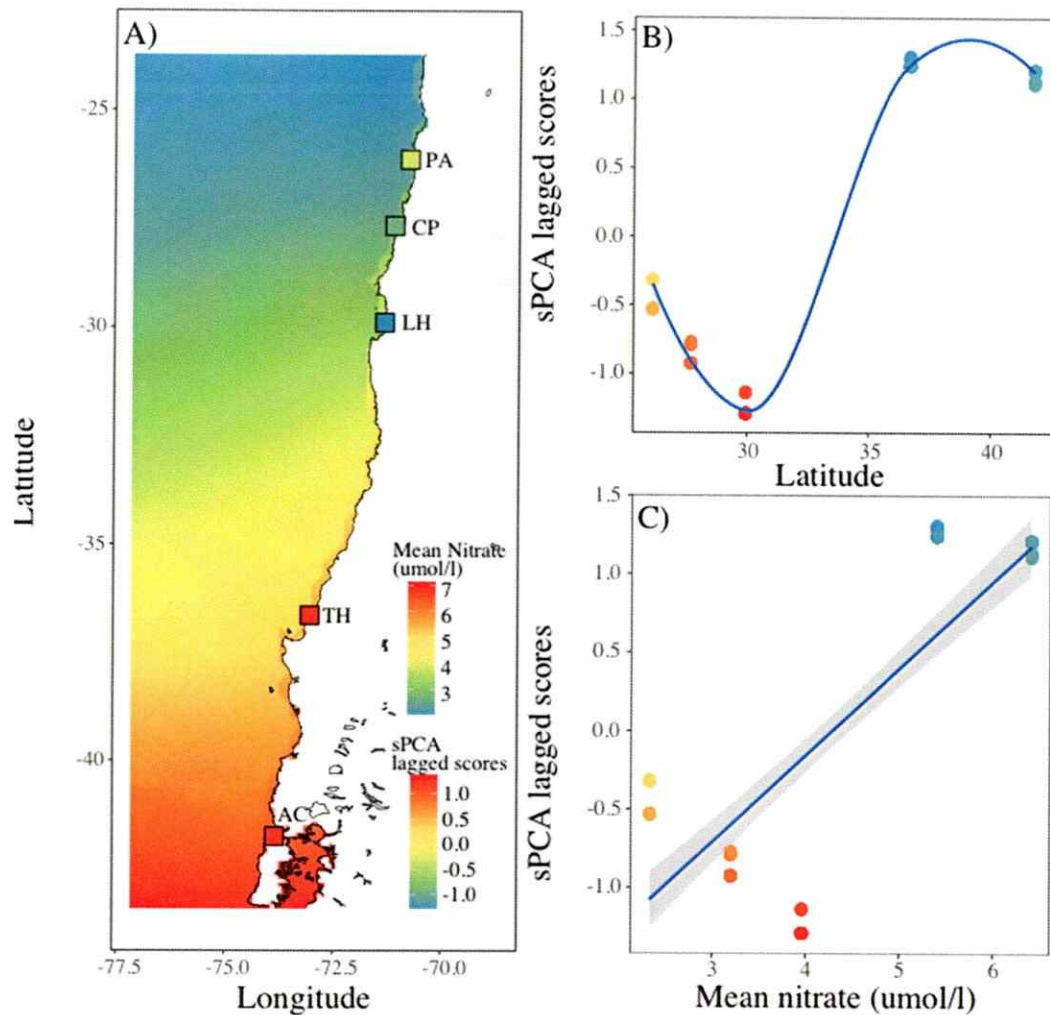


Figura 11: Variación genética adaptativa espacial de los 6 loci candidatos a adaptación local para la variable ambiental nitrato para los 5 sitios estudiados en *Pyura chilensis*. A) Estructura espacial del nitrato a lo largo de la zona estudio (El color del agua representa a la variación del nitrato) y valores de las clinas multilocus para cada sitio medidas mediante los *lagged scores* del análisis de sPCA, los que reflejan la variabilidad genética que está ligada a la estructura espacial entre sitios. Con fines gráficos, se utilizó la misma escala de colores para los valores de los *lagged scores* y nitrato, B) Gráfico de dispersión de las clinas multilocus determinadas en el análisis sPCA para cada localidad considerando la relación entre la latitud y los dos primeros ejes de los *lagged scores*. El gradiente de colores es el mismo usado para el panel A). C) Ajuste lineal entre las clinas multilocus de los loci candidatos (*lagged scores*) respecto a la variación del nitrato en cada sitio.

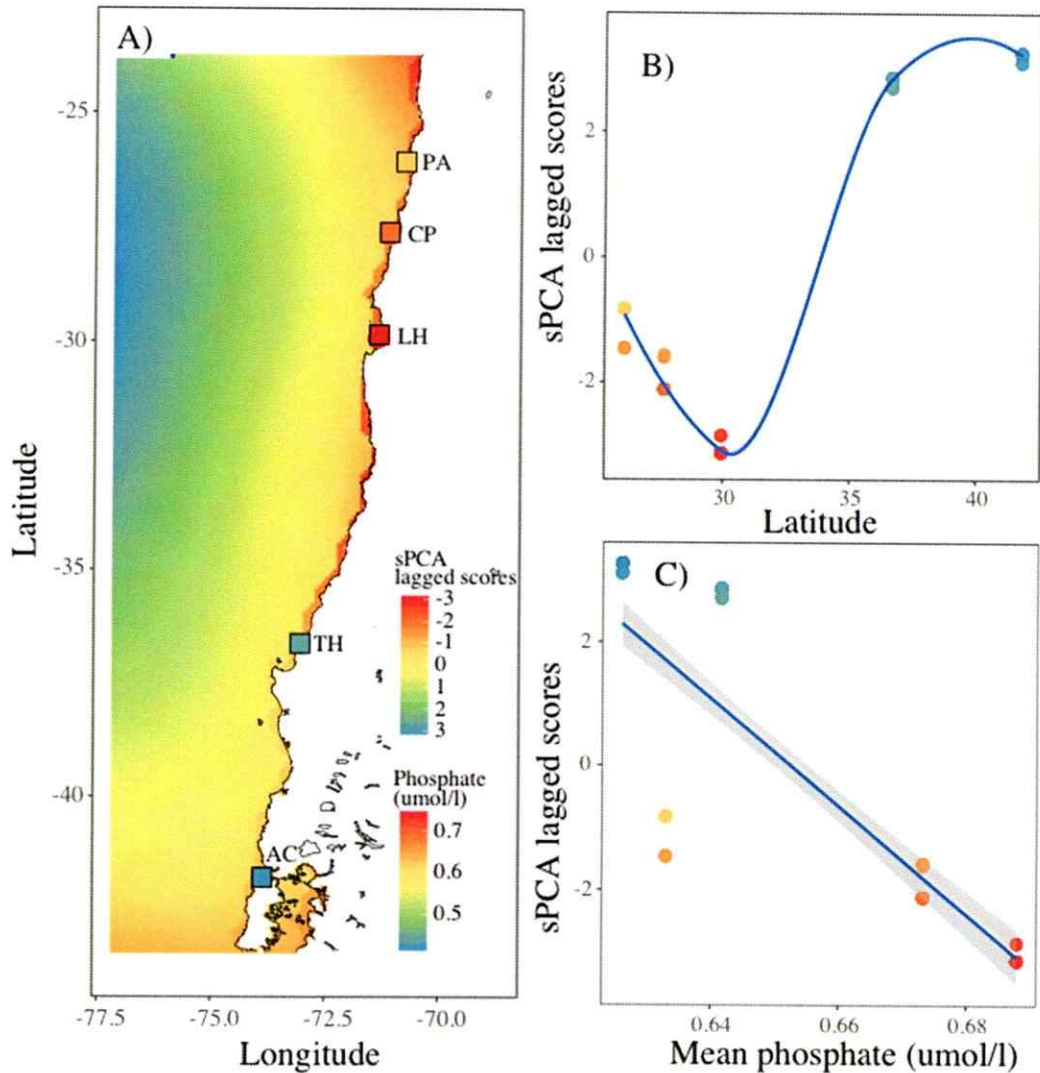


Figura 12: Variación genética adaptativa espacial de los 16 loci candidatos a adaptación local para la variable ambiental fosfato para los 5 sitios estudiados en *Pyura chilensis*. A) Estructura espacial del fosfato a lo largo de la zona estudio (El color del agua representa a la variación del fosfato) y valores de las clinas multilocus para cada sitio medidas mediante los *lagged scores* del análisis de sPCA, los que reflejan la variabilidad genética que está ligada a la estructura espacial entre sitios. Con fines gráficos, se utilizó la misma escala de colores para los valores de los *lagged scores* y fosfato, B) Gráfico de dispersión de las clinas multilocus determinadas en el análisis sPCA para cada localidad considerando la relación entre la latitud y los dos primeros ejes de los *lagged scores*. El gradiente de colores es el mismo usado para el panel A). C) Ajuste lineal entre las clinas multilocus de los loci candidatos (*lagged scores*) respecto a la variación del fosfato en cada sitio.

## Conclusión general

*Pyura chilensis* es una especie endémica de la corriente de Humboldt con características biológicas que la hacen un modelo relevante para estudiar procesos de adaptación local. Son organismos sésiles que viven tanto aislados como en agregaciones, y cuentan con una fase larval breve (12-24 horas), la cual le otorga una baja capacidad de dispersión intrínseca. Sin embargo, estudios previos indicaban que esta especie presentaba estructura genética moderada, menos que lo esperado dado su bajo potencial de dispersión, lo que se explicaba principalmente por la influencia del transporte antropogénico que permitía el flujo genético entre poblaciones locales. Este flujo genético mediado por actividades antropogénicas sería el responsable de haber borrado la huella genética de un quiebre biogeográfico ampliamente reportado en la zona, particularmente a los 30-33°S.

En el primer capítulo de esta tesis, utilizando SNPs dispersos por el genoma, mostramos que loci putativamente sujetos a selección muestran una estructura genética correspondiente al quiebre a los 30°S, evidenciando de forma general dos grupos genéticos asociados a ambas regiones. A nivel neutral, se detectó un linaje divergente que explica gran parte de la varianza genética el cual se encuentra restringido geográficamente primordialmente a la localidad de Los Molinos (39°S). Este linaje corresponde a un haplogrupo detectado previamente con el marcador mitocondrial COI. De forma interesante, a nivel neutral, la influencia de los 30°S también apareció, pero como estructura adyacente a la señal principal asociada al linaje divergente presente en Los Molinos, lo que quiere decir que el número de agrupaciones óptimas para loci

neutrales y putativamente sujetos a selección es el mismo, pero estos muestran patrones de estructuración distintos, siendo una excepción interesante a lo regularmente reportado en la literatura para estudios genéticos similares en organismos marinos.

Utilizando loci putativamente sujetos a selección la divergencia del linaje a los 39°S no sobrepasó a la diferenciación observada entre el resto de las poblaciones locales estudiadas, lo que sugiere que la selección natural mantiene la cohesión geográfica de la diversidad genética a pesar de los procesos neutrales que afectan a la especie. Los resultados obtenidos en este capítulo, nos permitieron hipotetizar que el ambiente, y particularmente las diferencias entre el norte y sur de los 30°S, podrían ser relevantes para mantener la diferenciación genética entre regiones y provocar procesos de adaptación local en el piure.

En el segundo capítulo de esta tesis, abordamos la estructura genética temporal desde una perspectiva filogeográfica. *Pyura chilensis* se caracteriza por la presencia de tres linajes monofiléticos, dos de los cuales son simpátricos en gran parte de la zona de estudio (desde los 26° hasta los 42°S), y un tercero con un alto nivel de divergencia restringido geográficamente (Los Molinos-39°S). Utilizando análisis clásicos de genética de poblaciones, pero intercambiando efectos fijos y aleatorios entre el espacio geográfico y el tiempo (comparaciones inter-anales) se pudo determinar que la estructura genética en el piure es mayor a nivel espacial que temporal, lo que es esperable para especies con bajo potencial de dispersión larval y sensibles a perturbaciones ambientales como el piure. A pesar de este esquema general, se detectaron diferencias en las frecuencias haplotípicas dentro de algunos sitios de un año a otro, así como diferencias temporales al analizar los linajes (haplogrupos) por



separado. Este estudio, realiza la idea de considerar la arista temporal en los análisis filogeográficos, los cuales regularmente suelen agrupar muestreos efectuados en años distintos pudiendo desestimar la influencia del tiempo en las frecuencias alélicas.

En el tercer capítulo, nuestra hipótesis de que la heterogeneidad ambiental a lo largo del HCS es la responsable de mantener la divergencia genética entre regiones (norte y sur 30°S) en el piure fue puesta a prueba utilizando herramientas del estado del arte en la emergente disciplina de genética del paisaje marino. Con este trabajo, se abren diversas perspectivas tanto a nivel de estudios de expresión genética (hemos logrado identificar loci sujetos a selección que podrían ser genes candidatos óptimos), de genética del paisaje marino en el área (es el primero ejecutado íntegramente con análisis de asociaciones genético-ambientales para la HCS) así como estudios de resiliencia ante el cambio climático y conservación de recursos marinos. Nuestros resultados indican que esta especie es bastante susceptible a cambios ambientales, debido a que a pesar del flujo genético activo entre regiones mediado por el transporte antropogénico (capítulo I y II), variables ambientales asociadas a surgencias y temperatura son relevantes para mantener la diferenciación genética al norte y sur de los 30°S. La susceptibilidad del piure a cambios ambientales que sugiere este estudio es fundamental en términos de conservación del recurso, debido a que esta especie es regularmente considerada como un solo stock pesquero a lo largo de su distribución geográfica, primordialmente debido a los resultados obtenidos en estudios anteriores utilizando genes neutrales a la selección que muestran un importante flujo genético (probablemente sobrestimado), estudios ecológicos y sobre ciclos reproductivos, indican una gran capacidad colonizadora y una amplia tolerancia a gradientes ambientales.

Uno de los objetivos fundamentales de este estudio fue intentar entender cómo la heterogeneidad del paisaje puede provocar diferencias entre la capacidad de dispersión y el flujo genético efectivo determinado de forma indirecta mediante el estudio de la estructura genética particularmente en porciones del genoma que están sujetas a los efectos de la selección natural. Hemos utilizado herramientas analíticas actuales que están disponibles para este tipo de estudios, tales como escaneos genómicos que permiten analizar miles de SNPs dispersos por el genoma en varios individuos a lo largo de la zona de estudio, herramientas de genética de poblaciones y herramientas de genética del paisaje. A pesar de la falta de información genómica relevante en la especie (ej. genoma de referencia), la baja resolución y disponibilidad de los datos ambientales, las limitantes metodológicas de las aproximaciones genéticas utilizadas y de las bases de datos de genes y proteínas, hemos podido delinear diversas perspectivas para este tipo de estudios en la costa sudeste del pacífico.

Primero, se recomienda continuar con esta especie como modelo de estudio para poder generar recursos genómicos que permitan identificar la posición física de los marcadores SNPs sujetos a selección a lo largo del genoma, lo que permitirá eventualmente mejorar la detección de genes candidatos a adaptación local. Segundo, se recomienda utilizar mediciones *in-situ* de variables ambientales para evitar sesgos por efecto de la resolución espacial de los datos, a pesar de que datos satelitales de un intervalo de tiempo de 10 años parecen ser apropiados de forma transitoria para identificar patrones adaptativos en respuesta al ambiente, y tercero, resulta fundamental utilizar múltiples aproximaciones en la detección de loci candidatos a adaptación local debido a la alta densidad de falsos positivos que produce cada método por separado. El uso de un set de

loci candidatos detectados en común resultó ser una herramienta útil para poder determinar la estructura espacial de la diversidad genética adaptativa en *Pyura chilensis*. Finalmente, tal como se propuso al inicio de este trabajo, pudimos mostrar como la diferencia de condiciones ambientales a lo largo de la corriente de Humboldt participa en la estructuración espacial de la diversidad genética del piure *P. chilensis* a través de procesos de adaptación local.

## Referencias

- Aiello, E.L. 1960. Factors affecting ciliary activity on the gill of the mussel *Mytilus edulis*. *Physiol. Zool.* 33: 120-135.
- Amaral, A.R. et al. 2012. Seascape Genetics of a Globally Distributed, Highly Mobile Marine Mammal: The Short-Beaked Common Dolphin (Genus *Delphinus*). *PLoS ONE* 7(2): e31482. doi:10.1371/journal.pone.0031482
- Ambler RP & JI Cañete. 1991. Asentamiento y reclutamiento de *Pyura chilensis* Molina, 1782 (Urochordata: Ascidiacea) sobre placas artificiales suspendidas en Bahía La Herradura, Coquimbo, Chile. *Revista de Biología Marina* 26: 403-413
- Arcos, D. & Navarro, N. 1986. Análisis de un índice de surgencia para la zona de Talcahuano, Chile (Lat. 37° S). *Invest. Pesq.* 33: 91-98.
- Araneda, C. et al. 2016. Adaptive genetic variation distinguishes Chilean blue mussels (*Mytilus chilensis*) from different marine environments. *Ecol. Evol.* doi: 10.1002/ece3.2110
- Aravena, G. et al. 2014. Twelve Years of Change in Coastal Upwelling along the Central-Northern Coast of Chile: Spatially Heterogeneous Responses to Climatic Variability. *PLoS ONE* 9: e90276.
- Armsworth, P. R et al. 2001. When to press on or turn back: dispersal strategies for reef fish larvae. *American Naturalist* 157:434–450.
- Arsenault, G., & Davidson, J. 2008. Temporal and spatial development of an infestation of *Styela clava* on mussel farms in Malpeque Bay, Prince Edward Island, Canada. *Aquatic Invasions*, 4: 189-194.
- Astorga, M.O. & Ortiz, J.C. 2006. Genetic variability and population structure in the tunicate *Pyura chilensis* Molina, 1782, in the coast of Chile. *Rev. Chil. Hist. Nat.* 79: 423-434.
- Ayre D.J., et al. 2009. Does life history predict past and current connectivity for rocky intertidal invertebrates across a marine biogeographic barrier? *Mol Ecol* 18: 1887–1903.
- Bachmann-Gagescu, R. et al. 2011. The ciliopathy gene *cc2d2a* controls zebrafish photoreceptor outer segment development through a role in Rab8-dependent

- vesicle trafficking. *Hum Mol Genet.* 20(20):4041-55.
- Baird N.A., et al. 2008. Rapid SNP Discovery and Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* 3(10): e3376
- Bairoch, A & Apweiler, R. 2000. The SWISS-PROT protein sequence database and its supplement TREMBL in 2000. *Nucleic Acids Research.* 28: 45–48.
- Benestan, L. et al. 2016. Seascape genomics provides evidence for thermal adaptation and current-mediated population structure in American lobster (*Homarus americanus*). *Molecular Ecology.* 25: 5073–5092.
- Bernier, R. Y, et al. 2008. Lobsters and crabs as potential vectors for tunicate dispersal in the southern Gulf of St. Lawrence, Canada. *Aquatic Invasions,* 4 :105-110.
- Bird, C.E. 2007. Contrasting phylogeography in three endemic Hawaiian limpets (*Cellana* spp.) with similar life histories. *Mol Ecol* 16: 3173–3186.
- Blanco, J.L. et al. 2001. Satellite-measured chlorophyll and temperature variability off northern Chile during the 1996–1998 La Niña and El Niño. *J Geophys Res.* 106: 899–915.
- Blanchette, C. A. et al. 2008. Biogeographical patterns of rocky intertidal communities along the Pacific coast of North America. *J. Biogeogr.* 35 : 1593–1607
- Borcard, D. & Legendre, P. 2002. All-scale spatial analysis of ecological data by means of principal coordinates of neighbour matrices. *Ecological Modelling.* 153: 5
- Bouchemousse, S. et al. 2016. Contrasting global genetic patterns in two biologically similar, widespread and invasive *Ciona* species (Tunicata, Ascidiacea). *Scientific Reports.* 6: 24875.
- Bourque, D. et al. 2007. Reproduction and early life history of an invasive ascidian *Styela clava* Herdman in Prince Edward Island, Canada. *Journal of Experimental Marine Biology and Ecology,* 342: 78-84.
- Bourret, V. et al. 2013. Landscape genomics in Atlantic salmon (*Salmo salar*): searching for gene-environment interactions driving local adaptation. *Evolution.* 67: 3469–3487.
- Bradbury, P. J. et al. 2007. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635.

- Brante, A. et al. 2012. Phylogeography and biogeography concordance in the marine gastropod *Crepidatella dilatata* (Calyptraeidae) along the Southeastern Pacific coast. *J Hered.* 103: 630–63.
- Breitburg, D. et al. 2009. Hypoxia, nitrogen, and fisheries: Integrating effects across local and global landscapes, *Annu. Rev. Mar. Sci.* 1: 329–349.
- Broitman, B. R. et al. 2001. Geographic variation of southeastern Pacific intertidal communities. *Mar. Ecol. Prog. Ser.* 224: 21–34.
- Bullard, S.G. & Carman, M.R. 2009. Current trends in invasive ascidian research. In: *Invasive Species: Detection, Impact and Control*. Editors C.P Wilcox & R.B Turpin. Nova Science Publishers, Inc.
- Bullard, S. G., et al. 2007. Fragmentation of colonial ascidians: differences in reattachment capability among species. *Journal of Experimental Marine Biology and Ecology*, 342:166-168.
- Cahill, A. & Levinton, J.S. 2016. Genetic differentiation and reduced genetic diversity at the northern range edge of two species with different dispersal modes. *Mol. Ecol.* 25: 515-526.
- Cancino, J. et al. 1998. Estudios del ciclo vital del piure y picoroco en la VII Región. Informe final, Fondo de Investigación Pesquera (FIP), Proyecto N° 96-49: 1-165.
- Candy, J. et al. 2015. Population differentiation determined from putative neutral and divergent adaptive genetic markers in Eulachon (*Thaleichthys pacificus*, Osmeridae), an anadromous Pacific smelt. *Mol. Ecol. Resour.* 15: 1421–1434.
- Carson, H.S. et al. 2011. Evaluating the importance of demographic connectivity in a marine metapopulation. *Ecology.* 92, 1972-1984.
- Cea, G. 1969. Estadios primarios de desarrollo y metamorfosis de *Pyura chilensis* Molina, 1782 (Tunicata, Ascidiacea, Pyuridae). *Bol Soc Biol Concepción*, 42:317–331.
- Cea, G. Contribución al conocimiento de algunos aspectos de la biología de *Pyura chilensis* Molina 1782 (Chordata, Tunicata, Ascidiacea). Tesis de Licenciatura en Biología, Universidad de Concepción, Concepción, Chile. 205 pp (1970).
- Cifuentes, M. et al. 2007. Effects of temporal variability of disturbance on the succession in marine fouling communities in northern-central Chile. *J. Exp. Mar. Biol. Ecol.* 352: 280–294.

- Conover, D.O. 1998. Local adaptation in marine fishes: evidence and implications for stock enhancement. *Bull. Mar. Sci.* 62:477–93,
- Conover, D.O. et al. 2006. Spatial and temporal scales of adaptive divergence in marine fishes and the implications for conservation. *J. Fish Biol.* 69:21–47.
- Conover, D.O. et al. 2009. Reversal of evolutionary downsizing caused by selective harvest of large fish. *Proc. R. Soc. Long. Biol. Sci.* 276, 2015–2020.
- Correa-Ramirez, M.A. et al. 2007. Mesoscale eddies and high chlorophyll concentrations of central Chile (29–39°S). *Geophys. Res. Lett.* 34 (12). DOI: <http://dx.doi.org/10.1029/2007GL029541>.
- Cowen, R.K., et al. 2000. Connectivity of marine populations: open or closed? *Science*, 287: 857–859.
- Cowen, R.K. & Sponaugle, S. 2009. Larval dispersal and marine population connectivity. *Ann. Rev. Mar. Sci.* 1: 443–466.
- Davis, A. R., & Butler, A. J. 1989. Direct observations of larval dispersal in the colonial ascidian *Podoclavella moluccensis* Sluiter: evidence for closed populations. *Journal of Experimental Marine Biology and Ecology*, 127: 189–203.
- Davis, A.R. 1995. Over-exploitation of *Pyura chilensis* (Ascidiacea) in southern Chile: the urgent need to establish marine reserves. *Rev Chil Hist Nat.* 68:107–116.
- Davis, A.R. 1996. Association among ascidians: facilitation of recruitment in *Pyura spinifera*. *Mar Biol* 126:35–41.
- Doney, S. C. 2010. The growing human footprint on coastal and open-ocean biogeochemistry, *Science*. 328(5985): 1512–1516.
- Demarcq, H. 2009. Trends in primary production, sea surface temperature and wind in upwelling systems (1998–2007). *Progress in Oceanography* 83: 376–385.
- De Mita, S. et al. 2013. Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*. 22: 1383–1399.
- De Donato, M. 2013. Genotyping-by-Sequencing (GBS): A Novel, Efficient and Cost-Effective Genotyping Method for Cattle Using Next-Generation Sequencing. *PLoS ONE* 8, e62137.

- DeWitt, et al. 1998. Costs and limits of phenotypic plasticity. *Trends Ecol. Evol.* 13: 77–81.
- Díaz-Viloria, N. 2009. Genetic connectivity among pink abalone *Haliotis corrugata* populations. *J. Shellfish Res.* 28:599–608.
- Dray, S. et al. 2006. Spatial modelling: a comprehensive framework for principal coordinate analysis of neighbour matrices (PCNM). *Ecological Modelling*, 96: 483–493.
- Ellegren, H. 2014. Genome sequencing and population genomics in non-model organisms. *Trends in Ecology & Evolution*, 29(1): 51–63.
- Elshire, R.J. et al. 2011. A Robust, Simple Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* 6, e19379.
- Espinoza, P. 2017. Trophic structure in the northern Humboldt Current system: new perspectives from stable isotope analysis. *Mar Biol.* 164:86.
- Evans J, et al. 2013. First record of *Herdmania momus* (Ascidiacea: Pyuridae) from the central Mediterranean Sea. *Marine Biodiversity Records* 6(e134): 1–4.
- Fabry, V. J. et al. 2008. Impacts of ocean acidification on marine fauna and ecosystem processes. *ICES J. Mar. Sci.* 65: 414–432.
- Fenberg, P.B. et al. 2015. Biogeographic structure of the northeastern Pacific rocky intertidal: The role of upwelling and dispersal to drive patterns. *Ecography*. 38:93–95.
- Figueroa, D. 2002. Forcing of physical exchanges in the nearshore Chilean ocean. In *the Oceanography and Ecology of the Nearshore and Bays in Chile*, J.C. Castilla & J.L. Largier (eds). Santiago, Chile: Ediciones Universidad Católica de Chile, 31–43.
- Fischer, M.C., et al. 2013. Population genomic footprints of selection and associations with climate in natural populations of *Arabidopsis halleri* from the Alps. *Molecular Ecology*, 22: 5594–5607.
- Foll, M. & Gaggiotti, O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics*. 180: 977–993.
- Francois, O. et al. 2016. Controlling false discoveries in genome scans for selection. *Molecular Ecology*. 25: 454–469.



- Frichot, E. et al. 2013. Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. *Molecular Biology and Evolution*. 30 (7):1687–1699.
- Frichot E, et al. 2015. Detecting adaptive evolution based on association with ecological gradients: orientation matters! *Heredity*. 115: 22–28.
- Gaines, S.D. et al. 2007. Connecting places: the ecological consequences of dispersal in the sea. *Oceanography*, 20: 90–99.
- Galindo, H.M. et al. 2010. Seascape genetics along a steep cline: using genetic patterns to test predictions of marine larval dispersal. *Mol. Ecol.* 19: 3692–3707.
- Gaitán-Espitia, J.D. et al. 2014. Exploring physiological plasticity and local thermal adaptation in an intertidal crab along a latitudinal cline. *Journal of Thermal Biology*. doi:10.1242/jeb.108217
- Gaitán-Espitia, J.D. et al. 2017. Geographic variation in thermal physiological performance of the intertidal crab *Petrolisthes violaceus* along a latitudinal gradient. *Journal of Experimental Biology*. doi.org/10.1016/j.jtherbio.2017.02.011
- Gagnaire, P.A. et al. 2015. Using neutral, selected, and hitchhiker loci to assess connectivity of marine populations in the genomic era. *Evolutionary Applications*. 8: 769–786.
- Godhe, A. et al. 2013. Seascape analysis reveals regional gene flow patterns among populations of a marine planktonic diatom. *Proceedings of the Royal Society of London B: Biological Sciences*. 280: 20131599.
- Goldstein, S.J. et al. 2009. Colonization and connectivity by intertidal limpets among New Zealand, Chatham and Sub-Antarctic Islands. I. Genetic connections. *Mar. Ecol. Prog. Ser.* 388:111–19
- Giles E.C. et al. 2015. Exploring seascape genetics and kinship in the reef sponge *Stylissa carteri* in the Red Sea. *Ecology and Evolution*, 5(13): 2487–2502
- Gili, J.O. & Coma, R. 1998. Benthic suspension feeders: Their paramount role in littoral marine food webs. *TREE*. 13: 316– 321.
- Gómez-Gutiérrez, J. et al. 2012. Seasonal growth, molt, and egg production of *Nyctiphanes simplex* (Crustacea: Euphausiacea) juveniles and adults in the Gulf of California. *Mar. Ecol. Prog. Ser.* 455, 173–194.

- Gorden, N.T. et al. 2008. CC2D2A Is mutated in Joubert syndrome and interacts with the ciliopathy-associated basal body protein CEP290. *Am J Hum Genet.* 83(5):559–71.
- Grosberg, R.K. 1981. Competitive ability influences habitat choice in marine invertebrates. *Nature*, 290: 700-702.
- Grosberg, R.K. 1987. Limited dispersal and proximity-dependent mating success in the colonial ascidian *Botryllus schlosseri*. *Evolution*, 41, 372-384.
- Grosberg, R.K. & Cunningham, C. 2001. Genetic structure in the sea. Pp. 61–84. in M. D. Bertness, S. Gaines and M. E. Hay, eds. *Marine community ecology*. Sinauer, Sunderland, MA.
- Guenther, T & Coop, G. 2013. Robust identification of local adaptation from allele frequencies. *Genetics*. 195: 205–220.
- Gugger, P.F. et al. 2016. Species-wide patterns of DNA methylation variation in *Quercus lobata* and their association with climate gradients. *Mol. Eco.* 25: 1665–1680.
- Hamilton, S.L. et al. 2008. Post settlement survival linked to larval life in a marine fish. *Proc. Natl. Acad. Sci. U.S.A.* 105: 1561-1566.
- Hansen MM & Hemmer-Hansen J. 2007. Landscape genetics goes to sea. *J Biol* 6: 6.
- Hassl, R.J. & Payseur, B.A. 2016. Fifteen years of genome-wide scans for selection: trends, lessons and unaddressed genetic sources of complication. *Mol. Ecol.* 25: 5-23.
- Haye P.A & Muñoz-Herrera N.C. 2013. Isolation with differentiation followed by expansion with admixture in the tunicate *Pyura chilensis*. *BMC Evolutionary Biology*. 13:252.
- Haye, P.A. et al. 2014. Phylogeographic structure in benthic marine invertebrates of the southeast Pacific coast of Chile with differing dispersal potential. *PLoS ONE* 9: e88613.
- Havenhand, J. N. 1991. Fertilization and the potential for dispersal of gametes and larvae in the solitary ascidian *Ascidia mentula* Mueller. *Ophelia*, 33: 1-15.
- Hedgecock, D. 1986. Is gene flow from pelagic larval dispersal important in the adaptation and evolution of marine invertebrates? *Bull. Mar. Sci.* 39: 550–64.

- Hedgecock, D. et al. 2007. Genetic approaches to measuring connectivity. *Oceanography*. 20: 70–79.
- Hellberg, M.E. 2009. Gene flow and isolation among populations of marine animals. *Annu. Rev. Ecol. Evol. Syst.* 40: 291–310.
- Hellberg, M.E, et al. 2002. Genetic assessment of connectivity among marine populations. *Bull Mar Sci.* 70: 273–290.
- Hendry, A.P. 2004. Selection against migrants contributes to the rapid evolution of ecologically dependent reproductive isolation. *Evol. Ecol. Res.*, 6: 1219–1236.
- Hirano, T. & Nishida, H. 1997. Developmental fates of larval tissues after metamorphosis in ascidian *Halocynthia roretzi*. I. Origin of mesodermal tissues of the juvenile. *Dev Biol.* 192: 199–210
- Hohenlohe, P.A. et al. 2011. Next-generation RAD sequencing identifies thousands of SNPs for assessing hybridization between rainbow and west slope cutthroat trout. *Molecular Ecology Resources.* 11: 117–122.
- Holderegger R, et al. 2010. Landscape genetics of plants. *Trends in Plant Science.* 15: 675–683.
- Hudson, J. et al. 2016. Anthropogenic transport of species across native ranges: unpredictable genetic and evolutionary consequences. *Biol. Lett.* 12, doi: 10.1098/rsbl.2016.0620 (2016).
- Iriarte, J.L. et al. 2007. Spatial and temporal variability of chlorophyll and primary productivity in surface waters of southern Chile. *Estuarine. Coast. Shelf Sci.* 74: 471e480.
- Jeffreys, H. 1961. *The Theory of Probability*, 3rd edn. Oxford University Press, Oxford, UK
- Jensen, J.D. et al. 2016. The past, present and future of genomic scans for selection. *Molecular Ecology.* 25: 1–4.
- Johnson, M.S. & Black, R. 2006. Islands increase genetic subdivision and disrupt patterns of connectivity of intertidal snails in a complex archipelago. *Evolution.* 60: 2498–2506.
- Jombart, T. 2008. ADEGENET: An R package for the multivariate analysis of genetic markers. *Bioinformatics.* 24: 1403–1405.
- Joost S, Bonin A, Bruford W et al (2007) A spatial analysis method (SAM) to

detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, 16, 3955–3969.

Jorde, P.E. et al. 2015. Genetically distinct populations of northern shrimp, *Pandalus borealis*, in the North Atlantic: adaptation to different temperatures as an isolation factor. *Molecular Ecology*. 24: 1742–1757.

Kawecki, T.J. & Ebert, D. 2004. Conceptual issues in local adaptation. *Ecol. Lett.* 7:1225–41.

Keller RP, Drake JM, Drew MB, Lodge DM (2010) Linking environmental conditions and ship movements to estimate species transport across the global shipping network. *Diversity and Distributions* 17: 93–102, <http://dx.doi.org/10.1111/j.1472-4642.2010.00696.x>

Kelly, R.P. & Palumbi, S.R. 2010. Genetic structure among 50 species of the Northeastern Pacific rocky intertidal community. *PLoS ONE* 5:e8594. doi: 10.1371/journal.pone.0008594.

Kershaw, F. & Rosenbaum, H.C. 2014. Ten years lost at sea: response to Manel and Holderegger. *Trends in Ecology and Evolution*. 29: 69–70.

Korte, A & Farlow, A. 2013. The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods*, 9: 29.

Lal, M.M. et al. 2016. A Parallel Population Genomic and Hydrodynamic Approach to Fishery Management of Highly-Dispersive Marine Invertebrates: The Case of the Fijian Black-Lip Pearl Oyster *Pinctada margaritifera*. *PLoS ONE* 11, e0161390.

Largier, J.L. 2003. Considerations in estimating larval dispersal distances from oceanographic data. *Ecological Applications* 13: S71–S89.

Lasky, J.R. et al. 2012. Characterizing genomic variation of *Arabidopsis thaliana*: the roles of geography and climate. *Mol. Ecol.* 21: 5512–5529.

Lampitt, R. S. & Antia A.N. 1997. Particle flux in deep seas: Regional characteristics and temporal variability, *Deep Sea Res., Part I*. 44: 1377–1403.

Lancellotti, D. & Vásquez, J.A. 2000. Zoogeografía de macroinvertebrados bentónicos de la costa de Chile: contribución para la conservación marina. *Rev. Chil. Hist. Nat.* 73, 99-129.

Leinonen, T. et al. 2008. Comparative studies of quantitative trait and neutral marker divergence: a meta-analysis. *Journal of Evolutionary Biology*, 21: 1–17.

- Legendre, P. & Gallagher, E.D. 2001. Ecologically meaningful transformations for ordination of species data. *Oecologia*. 129: 271–280.
- Legendre, P & Legendre, L. 2012. *Numerical Ecology*. Elsevier, Amsterdam, The Netherlands.
- Lipicius, R.N. et al. 2008. Importance of metapopulation connectivity to restocking and restoration of marine species. *Rev. Fisheries. Sci.* 16: 101-110.
- Lotterhos, K.E. & Whitlock, M.C. 2015. The relative power of genome scans to detect local adaptation depends on sampling design and statistical method. *Molecular Ecology*. 24: 1031– 1046.
- Luikart, G. et al. 2003. The power and promise of population genomics: from genotyping to genome typing. *Nature Reviews Genetics*. 4: 981– 94.
- Lu, F. et al. 2013. Switchgrass genomic diversity, ploidy, and evolution: novel insights from a network- based SNP discovery protocol. *PLoS Genet.* 9, e1003215.
- Luu, K. et al. 2017. pcadapt: an R package to perform genome scans for selection based on principal component analysis. *Mol. Ecol. Res.* 17(1):67-77.
- Leeuw, T. 2013. In situ Measurements of Phytoplankton Fluorescence Using Low Cost Electronics. *Sensors*. 13: 7872-7883.
- McGaughran, A. et al. 2014. Environmental Variables Explain Genetic Structure in a Beetle-Associated Nematode. *PLoS ONE*9(1): e87317.
- Manel, S. & Holderegger, R. 2013. Ten years of landscape genetics. *TREE*. 28, 614–621.
- Manel, S. et al. 2010. Perspectives on the use of landscape genetics to detect genetic adaptive variation in the field. *Molecular Ecology*. 19: 3760–3772.
- Manríquez P.H. & Castilla J.C. 2005. Self-fertilization as an alternative mode of fertilization in the solitary tunicate *Pyura chilensis*. *Mar. Ecol. Prog. Ser.* 305: 113-125.
- Manríquez P & J Castilla. 2007. Role of larval behaviour and microhabitat traits in determining spatial aggregations in the ascidian *Pyura chilensis*. *Marine Ecology Progress Series*. 332: 155-165.
- Marshall, D.J. 2008. Transgenerational plasticity in the sea: context-dependent maternal effects across the life-history. *Ecology*, 89: 418-427.

- Marshall, D.J. et al. 2010. Phenotype-environment mismatches reduce connectivity in the sea. *Ecol. Lett.* 13: 128-140.
- Menge, B. A. 2000. Top-down and bottom-up community regulation in marine rocky intertidal habitats. *J. Exp. Mar. Biol. Ecol.* 250: 257-289.
- Menge, B. A. & Menge, D. N. L. 2013. Dynamics of coastal meta-ecosystems: the intermittent upwelling hypothesis and a test in rocky intertidal regions. *Ecol. Monogr.* 83: 283-310.
- Montecino, V. & Quiroz, D. 2000. Specific primary production and phytoplankton cell size structure in an upwelling area off the coast of Chile (30°S). *Aquatic Sciences*, 62:364-380.
- Montecino, V. et al. 2004. Bio-optical characteristics and primary productivity during upwelling and non-upwelling conditions in a highly productive coastal ecosystem off central Chile (~36°S). *Deep-Sea Research II* 51: 2413-2426.
- Montecinos, A. et al. 2012. Species replacement along a lineal coastal habitat: phylogeography and speciation in the red alga *Mazzaella laminarioides* along the south-east Pacific. *BMC Evol Biol.* 12: 97.
- Moraga, J., et al. 2001. Condiciones oceanográficas durante la fase de relajación de un evento de surgencia invernal frente a Punta Lengua de Vaca, Coquimbo. *Investigaciones Marinas, Valparaiso* 29: 59-71
- Moura, A.E. et al. 2014. Population genomics of the killer whale indicates ecotype evolution in sympatry involving both selection and drift. *Mol. Ecol.* 23: 5179-5192.
- Nandakumar, K. et al. 1993. Interspecific competition among fouling organisms in Tomioka Bay, Japan. *Marine Ecology Progress Series*, 94: 43-50.
- Nayfa, M.G. & Zenger, M.R. 2016. Unravelling the effects of gene flow and selection in highly connected populations of the silver-lip pearl oyster (*Pinctada maxima*). *Mar. Genomics*. doi: 10.1016/j.margen.2016.02.005.
- Nielsen, R. 2001. Statistical tests of selective neutrality in the age of genomics. *Journal of Heredity*, 86: 641-647.
- Nielsen, E.E. et al. 2009. Population genomics of marine fishes: identifying adaptive variation in space and time. *Molecular Ecology*. 19:3128-3150.

- Nosil, P. et al. 2009. Divergent selection and heterogeneous genomic divergence. *Mol. Ecol.* 18, 375–402.
- Nosil, P., et al. 2005. Perspective: reproductive isolation caused by natural selection against immigrants from divergent habitats. *Evolution.* 59: 705–719.
- Núñez-Acuña, G. et al. 2012. Gene expression analysis in *Mytilus chilensis* populations reveals local patterns associated with ocean environmental conditions. *Journal of Experimental Marine Biology and Ecology.* 420: 56-64.
- Ordoñez, V. et al. 2013. Mixed but not admixed: a spatial analysis of genetic variation of an invasive ascidian on natural and artificial substrates. *Mar. Biol.* 160, 1645–1660.
- Palumbi, S.R. 1994. Genetic divergence, reproductive isolation, and marine speciation. *Annual Review of Ecology and Systematics.* 25: 547–572.
- Palumbi, S.R. 2003. Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications,* 13: S146-S158.
- Pechenick, J.A. 2006. Larval experience and latent effects –metamorphosis is not a new beginning. *Integr. Compr. Biol.* 46: 323-333.
- Pelc, R.A. et al. 2009. Geographical patterns of genetic structure in marine species with contrasting life histories. *J Biogeo* 36: 1881–1890.
- Penney, R.W. et al. 2006. Genotype-dependent survival, growth, and reproduction in cultured blue mussels, *Mytilus* spp.: results of a reciprocal seed transfer experiment. *J. Shellfish. Res.* 25: 515-525.
- Pérez-Valdés, M. 2017. Reproductive cycle of sea squirt *Pyura chilensis* (Urochordata: Ascidiacea) originating from aquaculture mussel systems. *Revista de biología marina y oceanografía,* 52(2): 333-342.
- Petersen, J.K. et al. 1999. Beat frequency of cilia in the branchial basket of the ascidian *Ciona intestinalis* in relation to temperature and algal cell concentration. *Marine Biology.* 133: 185-192.
- Pineda, J. et al. 2007. Transport and dispersal in the coastal ocean and consequences for population connectivity. *Oceanography.* 20: 22-39.
- Pineda, M.C. et al. 2011. The whereabouts of an ancient wanderer: global phylogeography of the solitary ascidian *Styela plicata*. *PLoS ONE.* 6: e25495.

- Pineda, M.C. et al. 2013. Year-round reproduction in a seasonal sea: biological cycle of the introduced ascidian *Styela plicata* in the Western Mediterranean. *Marine Biology*. 160: 221-230.
- Pizarro, O. et al. 1994. Variabilidad del viento, nivel del mar y temperatura en la costa norte de Chile. *Invest. Mar., Valparaíso*. 22: 85-101.
- Poulin, E., et al. 2002. Avoiding offshore transport of competent larvae during upwelling events: the case of the gastropod *Concholepas concholepas* in Central Chile. *Limnology and Oceanography* 47: 1248–1255.
- Pritchard, J.K. et al. 2010. The genetics of human adaptation: hard sweeps, soft sweeps, and polygenic adaptation. *Current Biology*. 20: R208–R215.
- Pritchard, J.K. & Di Rienzo, A. 2010. Adaptation — not by sweeps alone. *Nature Reviews Genetics*, 11: 665–667.
- Pujolar, J.M. et al. 2014. Genome-wide single-generation signatures of local selection in the panmictic European eel. *Molecular Ecology*. 23: 2514–2528.
- Rajbanshi, R. & Pederson, J. 2007. Competition among invading ascidians and a native mussel. *Journal of Experimental Marine Biology and Ecology*. 342: 163-165.
- Rellstab, C. et al. 2015. A practical guide to environmental association analysis in landscape genomics. *Molecular Ecology*. 24: 4348–4370.
- Riginos, C. & Liggins, L. 2013. Seascape genetics: populations, individuals, and genes Marooned and Adrift. *Geography Compass*. 7: 197–216.
- Riquelme-Bugueño, R. et al. 2016. Growth of the Humboldt Current krill in the upwelling zone off central Chile. *Journal of Marine Systems*. 163:1-11.
- Rivadeneira, M et al. 2002. Latitudinal trends of species diversity in rocky intertidal herbivore assemblages: spatial scale and the relationship between local and regional species richness. *Mar. Ecol. Prog. Ser.* 245: 123-131.
- Riisgård, H.U. & Larsen, P.S. 2001. Minireview: Ciliary filter feeding and bio-fluid mechanics—present understanding and unsolved problems. *Limnology and Oceanography*. 46(4):882–891.
- Roy, K. et al. 1994. Eastern Pacific Molluscan provinces and latitudinal diversity gradient – no evidence for “Rapoport’s rule”. *Proc. Natl Acad. Sci. USA* 91: 8871 – 8874



- Sánchez, et al. 2011. Spatial pattern of genetic and morphological diversity in the direct developer *Acanthina monodon* (Gastropoda: Mollusca). *Mar Ecol Prog Ser.* 434: 121–131.
- Sanford, E. & Kelly, M.W. 2011. Local adaptation in marine invertebrates. *Annual Review of Marine Science.* 3: 509–535.
- Schmidt, P.S. et al. 2008. Ecological genetics in the north Atlantic: environmental gradients and adaptation at specific loci. *Ecology.* 89: S91–S107.
- Shanks, A.L. 2009. Pelagic larval duration and dispersal distance revisited. *Biol. Bull.*, 216: 373–385.
- Segovia, N. I. et al. 2017. Lineage divergence, local adaptation across a biogeographic break, and artificial transport, shape the genetic structure in the ascidian *Pyura chilensis*. *Sci. Rep.* 7: 44559 doi: 10.1038/srep44559.
- Selkoe, K.A. et al. 2008. Seascape genetics and the spatial ecology of marine populations. *Fish. Fish.* 9: 363–377.
- Selkoe, K.A. et al. 2016. A decade of seascape genetics: Contributions to basic and applied marine connectivity. *Marine Ecology Progress Series.* 554: 1–19.
- Sepúlveda, R. et al. 2003. The peracarid epifauna associated with the ascidian *Pyura chilensis* (Molina, 1782) (Ascidiacea: Pyuridae). *J Nat Hist.* 37: 1555 – 1569.
- Sork, V.L. et al. 2013. Putting the landscape into the genomics of trees: approaches for understanding local adaptation and population responses to changing climate. *Tree Genetics & Genomes.* 9: 901–911.
- Sotka, E. et al. 2004. Strong genetic clines and geographical variation in gene flow in the rocky intertidal barnacle *Balanus glandula*. *Mol. Ecol.* 13:2143–2156.
- Sotka, E.E. 2005. Local adaptation in host use among marine invertebrates. *Ecol. Lett.* 8:448–59.
- Stefano, G.B. et al. 1977. Temperature dependent ciliary rhythmicity in *Mytilus edulis* and the effects of mono amnergic agents on its manifestation. *Biol Bull mar biol Lab. Woods Hole.* 153: 618–629.
- Stinchcombe J.R. & Hoekstra, H.E. 2008 Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. *Heredity*, 100: 158–170.

- Storfer, A. et al. 2006. Putting the “landscape” in landscape genetics. *Heredity*. 98: 128–142.
- Storz, J.F. 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology*, 14: 671–688.
- Stucki, S. et al. 2016. High performance computation of landscape genomic models integrating local indices of spatial association. *Mol. Ecol. Res.* DOI: 10.1111/1755-0998.12629
- Swearer, S.E et al. 1999. Larval retention and recruitment in an island population of a coral reef fish. *Nature* 402: 799–802.
- Swearer, S.E. et al. 2002. Evidence of self-recruitment in demersal marine populations. *Bull. Mar. Sci.* 70: 251.
- Takahashi, H. & Satoh, N. 2001. Trunk Lateral Cell-Specific Genes of the Ascidian *Halocynthia roretzi*. *Zoological Science*. 18(3):361-366.
- Tapia, F.J. et al. 2014. Latitudinal Discontinuity in Thermal Conditions along the Nearshore of Central-Northern Chile. *PLoS ONE*9(10): e110841.
- Tapia F.J & Gallardo-Escárate, C. 2015. Spatio-temporal transcriptome analysis in the marine snail *Tegula atra* along central-northern Chile (28–31°S). *Marine Genomics*. <http://dx.doi.org/10.1016/j.margen.2015.05.005>
- Tellier, F. et al. 2009. Phylogeographic analyses of the 30°S south-east Pacific biogeographic transition zone establish the occurrence of a sharp genetic discontinuity in the kelp *Lessonia nigrescens*: vicariance or parapatry? *Mol Phylogenet Evol.* 53: 679–693.
- Tepolt, C.K. & Palumbi, S.R. 2015. Transcriptome sequencing reveals both neutral and adaptive genome dynamics in a marine invader. *Molecular Ecology*. 24: 4145–4158.
- Teske, P.R. et al. 2006. Impacts of marine biogeographic boundaries on phylogeographic patterns of three South African estuarine crustaceans. *Marine Ecology Progress Series*. 314: 283–293.
- Teske, P.R. et al. 2011. Climate-driven genetic divergence of limpets with different life histories across a southeastern African marine biogeographic disjunction. *Mol. Ecol.* 20: 5025–5041.
- Thiel, M & Haye, P.A. 2006. The Ecology of rafting in the marine environment.

III. Biogeographical and evolutionary consequences. *Oceanography and Marine Biology: An Annual Review*. 44: 323-429.

Thiel, M. et al. 2007. The Humboldt Current System of Northern and Central Chile: oceanographic processes, ecological interactions and socioeconomic feedback. *Annu. Rev. Oceanogr. Mar. Biol.* 45: 195–344.

Thomas, A.C. et al. 2001. Satellite-measured chlorophyll and temperature variability off northern Chile during the 1996–1998 La Niña and El Niño. *Journal of Geophysical Research-Oceans*. 106: 899–915.

Torres, R. et al. 2011. Air-sea CO<sub>2</sub> fluxes along the coast of Chile: From CO<sub>2</sub> outgassing in central northern upwelling waters to CO<sub>2</sub> uptake in southern Patagonian fjords. *J Geophys Res.* 116: DOI: 10.1029/2010JC006344.

Turon, X. et al. 2016. Too cold for invasions? Contrasting patterns of native and introduced ascidians in subantarctic and temperate Chile. *Management of Biological Invasions*. 7:77-86.

Tyberghein, L., et al. 2012. Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography*. 21: 272–281. doi:10.1111/j.1466-8238.2011.00656.x

Valdivia, N. et al. 2005. Effects of disturbance on the diversity of hard-bottom macrobenthic communities on the coast of Chile. *Mar. Ecol. Prog. Ser.* 299. 45–54.

Vargas, C.A. et al. 2017. Species-specific responses to ocean acidification should account for local adaptation and adaptive plasticity. *Nature Ecology & Evolution*. 1: DOI:10.1038/s41559-017-0084

Vasemägi, A. & Primmer, C.R. 2005. Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. *Molecular Ecology*, 14: 3623–3642.

Véliz, D. et al. 2005. Stable genetic polymorphism in heterogeneous environments: balance between asymmetrical dispersal and selection in the acorn barnacle. *J. Evol. Biol.* 19: 589-599.

Vergara, O. et al. 2017. Controlling factors of the seasonal variability of productivity in the southern Humboldt Current System (30–40°S): A biophysical modeling approach. *Continental Shelf Research*. DOI: <http://dx.doi.org/10.1016/j.csr.2017.08.013>

de Villemereuil, P. et al. 2014. Genome scan methods against more complex

- models: When and how much should we trust them? *Molecular Ecology*. 23: 2006–2019.
- de Villemereuil, P & Gaggiotti, O.E. 2015. A new FST-based method to uncover local adaptation using environmental variables. *Methods in Ecology and Evolution*. 6(11): 1248–1258.
- Wares, J.P. et al. 2001. A comparative study of asymmetric migration events across a marine biogeographic boundary. *Evolution*. 55 : 295–306.
- White, C. et al. 2010. Ocean currents help explain population genetic structure. *Proceedings Biological sciences/The Royal Society*. 277: 1685–1694.
- Whitlock, M.C & Lotterhos, K.E. 2015. Reliable detection of loci responsible for local adaptation: inference of a null model through trimming the distribution of FST. *The American Naturalist*, 186 (Suppl 1): S24–36.
- Wright, S .1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. *Proceedings of the VI International Congress of Genetics*: 1 pp 356-366.
- Wong, C. S. et al. 1999. Seasonal and interannual variability in particle fluxes of carbon, nitrogen and silicon from time series of sediment traps at Ocean Station P, 1982–1993: Relationship to changes in subarctic primary productivity, *Deep Sea Res., Part II*. 46: 2735–2760.
- Wong, N.A. et al. 2011. The reproductive ecology of the invasive ascidian, *Styela clava*, in Auckland Harbour, New Zealand. *Marine Biology*. 158: 2775-2785.
- Yoneda, M. 1962. Force exerted by a single cilium of *Mytilus edulis*. II. Free motion. *J exp Biol*. 39: 307-317.
- Yuras, G. et al. 2005. On the annual cycle of coastal and open ocean satellite chlorophyll off Chile (18°–40°S). *Geophys. Res Lett*. 32(23): DOI 10.1029/2005GL023946.
- Xu, T. et al. 2016. Genome-wide discovery of single nucleotide polymorphisms (SNPs) and single nucleotide variants (SNVs) in deep-sea mussels: Potential use in population genomics and cross-species application. *Deep-sea. Res. PT II*. doi: 10.1016/j.dsr2.2016.03.011.
- Zamorano, J.H. & Moreno, C.A. 1975. Comunidades bentónicas del sublitoral rocoso de Bahía Corral. I. Area mínima de muestreo y descripción cuantitativa de la asociación *Pyura chilensis* Molina. *Medio Ambiente*. 1:58–67.