



PROYECTO FIPA N° 2023-19:

DIVERSIDAD MOLECULAR DE LAS POBLACIONES DE DOSIDICUS GIGAS EN EL PACÍFICO SUR, EN EL CONTEXTO DE LA OROP-PS

Informe Final

Requirente : Subsecretaría de Pesca y Acuicultura
Mandante : Fondo de Investigación Pesquera y de Acuicultura
Ejecutor : Universidad de Chile
Food Quality Research Center
Jefe(a) de proyecto : Dra. María Angélica Larraín Barth

Julio, 2025



INFORMACION DEL PROYECTO

Requirente	:	Subsecretaría de Pesca y Acuicultura
Mandante	:	Fondo de Investigación Pesquera y de Acuicultura
Supervisión	:	Consejo de Investigación Pesquera y de Acuicultura
Ejecutor	:	Universidad de Chile Food Quality Research Center
Jefe(a) de proyecto	:	Dra. María Angélica Larraín Barth
Autores	:	Dr. Cristian Araneda Tolosa Dra. Valentina Córdova Alarcón Dr. Alex DiGenova Dra. María Angélica Larraín Barth Dra. Carol Moraga Dr. Carlos Vargas Manríquez

Citar como: Larraín, M.A., Vargas, C.I., Córdova-Alarcón, V., DiGenova, A., Moraga, C., Araneda, C. 2025. Diversidad molecular de las poblaciones de *Dosidicus gigas* en el Pacífico Sur, en el contexto de la OROP-PS. Informe final FIPA 2023-19. 78 p.

I. Tabla de contenido

I.	Tabla de contenido	3
II.	Índice de Tablas	5
III.	Índice de Figuras	6
IV.	Índice de anexos	7
V.	Presentación de autores	9
VI.	Resumen ejecutivo.....	10
VII.	Abstract.....	12
1	Antecedentes.....	14
2	Objetivos.....	16
2.1	Objetivo General	16
2.2	Objetivos Específicos.....	16
3	Metodología de trabajo por objetivo.....	17
3.1	Obtención de muestras de <i>Dosidicus gigas</i>	17
3.2	OE1. Obtener genoma de referencia sin anotar de <i>Dosidicus gigas</i> para estudios genéticos en el Océano Pacífico Sur Este.	19
3.2.1	Extracción de ADN y construcción de librerías de ADN para secuenciación.....	19
3.2.2	Ensamble y anotación <i>de novo</i> del genoma de referencia	20
3.3	OE2. Analizar la existencia de diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de <i>Dosidicus gigas</i> en el Océano Pacífico Sur Este.....	21
3.3.1	Genotipado mediante tecnología DArT-seq	21
3.3.2	Filtrado de datos genotípicos.....	23
3.3.3	Identificación de marcadores putativamente neutrales y bajo selección	24
3.3.4	Caracterización de la diversidad genética de <i>Dosidicus gigas</i>	25
3.3.5	Descripción de la estructura genética poblacional.	25
3.3.6	Tamaño Efectivo Poblacional y Parentesco Genético	26
3.4	OE3. Conocer la asociación entre la variabilidad ambiental y las diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de <i>Dosidicus gigas</i> en el Océano Pacífico Sur Este.	26
3.4.1	Estructura espacio-temporal y asociación con variables ambientales de <i>Dosidicus gigas</i>	26
3.4.2	Identificación de loci candidatos de adaptación local en <i>Dosidicus gigas</i>	27
3.5	OE4. Asesorar a la Subsecretaría de Pesca y Acuicultura en el contexto de la OROP-PS relacionadas con la genética de <i>Dosidicus gigas</i>	27
4	Resultados.....	28

4.1	Obtención de muestras de <i>Dosidicus gigas</i>	28
4.2	OE1. Obtener genoma de referencia sin anotar de <i>Dosidicus gigas</i> para estudios genéticos en el Océano Pacífico Sur Este.	32
4.2.1	Extracción de ADN y construcción de librerías de ADN para secuenciación.....	32
4.2.2	Ensamble y anotación <i>de novo</i> del genoma de referencia	32
4.3	OE2. Analizar la existencia de diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de <i>Dosidicus gigas</i> en el Océano Pacífico Sur Este.....	39
4.4	Diversidad y Estructura Genética en poblaciones naturales.....	39
4.4.1	Secuenciación y construcción de base de datos	39
4.4.2	Identificación de loci putativamente bajo selección.....	40
4.4.3	Panel total de marcadores SNPs	41
4.4.4	Diferenciación poblacional.....	43
4.5	Estructura Poblacional	46
4.5.1	Análisis de Componentes Principales	46
4.5.2	Análisis Discriminante de Componentes Principales (DAPC)	47
4.5.3	Análisis de Admixture (Ancestría)	49
4.5.4	Tamaño Efectivo de la población y Parentesco genético.....	51
4.6	OE3. Conocer la asociación entre la variabilidad ambiental y las diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de <i>Dosidicus gigas</i> en el Océano Pacífico Sur Este.	54
4.6.1	Estructura espacio-temporal y asociación con variables ambientales de <i>Dosidicus gigas</i>	54
4.6.2	Identificación de loci candidatos de adaptación local en <i>Dosidicus gigas</i>	56
4.7	OE4. Asesorar a la Subsecretaría de Pesca y Acuicultura en el contexto de la OROP-PS relacionadas con la genética de <i>Dosidicus gigas</i>	59
4.7.1	Asesorías y Acuerdos OROP-PS	59
4.7.2	Difusión de resultados	61
5	Discusión	63
5.1	Genoma completo de <i>D. gigas</i>	63
5.2	Diversidad genética y estructura poblacional de <i>D. gigas</i> en la costa de Chile	64
5.3	Asociación ambiental con la variabilidad genética temporal de <i>D. gigas</i>	65
5.4	Recomendaciones	66
6	Conclusiones	68
7	Agradecimientos	69
8	Referencias bibliográficas	70
9	Anexos	79

II. Índice de Tablas

Tabla 1. Información de las muestras de Jibia (<i>Dosidicus gigas</i>) obtenidas por los proyectos FIPA 2020-20 y FIPA 2023-19 en la ZEE chilena y utilizadas para el análisis de diversidad genética y estructura poblacional.....	22
Tabla 2. Proceso de filtrado de calidad de datos DArT-Seq.	24
Tabla 3. Número de ejemplares y datos biométricos de <i>D. gigas</i> obtenidos por macrozona, región y temporada.....	29
Tabla 4. Estadísticas de ensamble para <i>Dosidicus gigas</i>	34
Tabla 5. Estadísticas de la anotación génica del ensamble de <i>D. gigas</i>	36
Tabla 6. Número de muestras por grupo/localidad.	39
Tabla 7. Resultados filtrado preliminar base de datos.....	40
Tabla 8. Diversidad genética global por localidad estimada con panel total de 1.313 SNPs.	42
Tabla 9. Diversidad genética neutral por localidad estimada con panel neutral de 1.148 SNPs.	42
Tabla 10. Diversidad genética adaptativa por localidad estimada con panel putativamente bajo selección de 165 SNPs.	42
Tabla 11. Valores de F_{ST} pareados estimados con panel total de 1.313 SNPs. F_{ST} pareados (bajo la diagonal) y sus p-values (sobre la diagonal).	44
Tabla 12. Valores de F_{ST} pareados estimados con panel neutral de 1.148 SNPs . F_{ST} pareados (bajo la diagonal) y sus p-values (sobre la diagonal).	44
Tabla 13. Valores de F_{ST} pareados estimados con panel putativamente bajo selección de 165 SNPs. F_{ST} pareados (bajo la diagonal) y sus p-values (sobre la diagonal).	45
Tabla 14. Estimadores del Tamaño efectivo de la población.	52
Tabla 15. Detalles de resultados del RDA y RDA parcial para cada variable ambiental estudiada. .	55
Tabla 16. Correlación de los SNPs candidatos a adaptación local con las variables ambientales predictoras seleccionadas para el modelo óptimo.....	58

III. Índice de Figuras

Figura 1. Identificación del sexo de los individuos de <i>D. gigas</i> muestreados. A la izquierda, una gónada femenina (Ovario) y a la derecha, una gónada masculina (Testículos).	18
Figura 2. Medición de parámetros biométricos durante la obtención de muestras de <i>D. gigas</i> . A la izquierda, medición del largo total del manto y a la derecha, peso de cada ejemplar.	18
Figura 3. Estadística pesquera de <i>Dosidicus gigas</i> de los últimos 5 años en cada macrozona estudiada en la ZEE de Chile (SUBPESCA 2024).	31
Figura 4. Evaluación BUSCO de 46 ensamblajes de moluscos usando el set mollusca_odb10 (n = 5 295 ortólogos de copia única).	35
Figura 5. Resumen de la secuenciación y control de calidad de lecturas cortas de 15 individuos de <i>D. gigas</i>	37
Figura 6. Genoma mitocondrial identificado en el ensamblaje de <i>Dosidicus gigas</i> . El genoma mitocondrial identificado corresponde al contig ptg000596l con una longitud de 19558 pares de bases.	38
Figura 7. Diagrama de Venn mostrando el resultado de la identificación de loci putativamente bajo selección.	41
Figura 8. Análisis de Componentes Principales obtenido con 165 SNPs. En el panel superior se encuentra la representación gráfica de los porcentajes de la varianza genética explicada por cada componente principal.	46
Figura 9. Análisis Discriminante de Componentes Principales realizado a partir del panel total de 1.313 SNPs.	48
Figura 10. Análisis Discriminante de Componentes Principales realizado a partir del panel putativamente bajo selección de 165 SNPs.	49
Figura 11. Gráfico de Admixture obtenido con el panel total de 1.313 SNPs.	50
Figura 12. Parentesco genético (R_{xy}) entre pares de individuos al interior de cada grupo estudiado. Las barras en la parte superior indican las diferencias significativas entre los grupos de muestreo estudiados.	53
Figura 13. Análisis de Redundancia de la variación genética utilizando un total de 1.313 SNPs y la variable ambiental “rango de concentración de fitoplancton a profundidad mínima” como variable explicativa de la variación genética de <i>D. gigas</i> . (RDA model $p=0.003$, $R^2_{adj}= 0.0839$).	56
Figura 14. Diagrama de Venn de los loci Outlier identificados por las cinco aproximaciones utilizadas.	57
Figura 15. Mapa de zonas de muestreo para el estudio poblacional de <i>D. gigas</i> acordado en la OROP-PS SC-12 Lima 2024.	61

IV. Índice de anexos

- Anexo 1:** Detalle de muestreo de *D. gigas* en Macrozona Norte, temporada Otoño-Invierno 2024.
- Anexo 2:** Detalle de muestreo de *D. gigas* en Macrozona Centro, temporada Otoño-Invierno 2024.
- Anexo 3:** Estado reproductivo y madurez sexual determinada por histología de gónadas de individuos seleccionadas para complementar la secuenciación de genoma completo.
- Anexo 4:** Detalle de muestreo de *D. gigas* en Macrozona Norte, temporada Primavera-verano 2024.
- Anexo 5:** Detalle de muestreo de *D. gigas* en Macrozona Centro, temporada Primavera-verano 2024.
- Anexo 6:** Comparación estadística de tamaño de los individuos de *D. gigas* muestreados en el presente proyecto.
- Anexo 7:** Acta Reunión IFOP
- Anexo 8:** Protocolo toma de muestras de *D. gigas*.
- Anexo 9:** Detalle de muestreo de *D. gigas* por parte de INPESCA.
- Anexo 10:** Parámetros de calidad de ADN de 15 individuos de *D. gigas* para secuenciar por Illumina NovaSeq.
- Anexo 11:** Resumen de la secuenciación de lecturas largas de *D. gigas* muestra MO-01.
Confirmación de recepción de muestras a secuenciar.
- Anexo 12:** Resumen de calidad de dos bibliotecas Hi-C (MO-01 Hi-C y MO-02 Hi-C) generadas a 150 pb a partir del individuo MO-01.
- Anexo 13:** Comparación de contigüidad y completitud del ensamble Jibia_V01 con ensambles de especies cercanas disponibles en NCBI.
- Anexo 14:** Distribución de métricas de alineamiento para 15 individuos adicionales mapeados contra el genoma de referencia Jibia_V02.
- Anexo 15:** Análisis de Componentes Principales obtenido con el panel total 1.313 SNPs.
- Anexo 16:** Análisis de Componentes Principales obtenido con el panel neutral 1.148 SNPs.
- Anexo 17:** Análisis Discriminante de Componentes Principales realizado a partir del panel neutral de 1.148 SNPs.
- Anexo 18:** Error de validación cruzada para la selección del k óptimo en Admixture.
- Anexo 19:** Gráfico de Admixture obtenido con el panel neutral 1.148 SNPs.
- Anexo 20:** Gráfico de Admixture obtenido con el panel outlier 165 SNPs.
- Anexo 21:** Distribución parentesco genético entre pares de individuos de los grupos estudiados.
- Anexo 22:** Coeficientes e Correlación de Pearson de las variables ambientales seleccionadas en cada punto de muestreo para la estimación del modelo de regresión final.
- Anexo 23:** Resultados de búsqueda BLAST en NCBI de los loci candidatos de adaptación local.
- Anexo 24:** Acta Reunión de Coordinación Inicial FIPA.

Anexo 25: Acta Reunión de avance de resultados FIPA.

Anexo 26: Acta Reunión de Coordinación de Actividades: Coordinación OROP-PS.

Anexo 27: Acta Reunión de Coordinación de Actividades: Acuerdos OROP-PS.

Anexo 28: Preguntas SC12-Perú.

Anexo 29: Extracto del Final Report SPRFMO SC12. Avances y acuerdos obtenidos para *D. gigas*.

Anexo 30: Acta e invitación Taller de Capacitación y Difusión de Resultados FIPA.

Anexo 31: Presentación: Secuenciación del genoma completo del calamar de Humboldt (*Dosidicus gigas*), expositor Dr. Cristian Araneda T.

Anexo 32: Presentación: Marcadores Moleculares utilizados en el Estudio del Calamar de Humboldt (*Dosidicus gigas*), expositor Dr. Cristian Araneda T.

Anexo 33: Presentación: Diversidad genética y estructura poblacional de *D. gigas* estimada con SNPs genómicos, expositor Dr. Carlos I. Vargas Manríquez.

Anexo 34: Presentación: Asociación entre variabilidad ambiental genética de *D. gigas* en la costa de Chile, expositora Dra. Valentina Córdova Alarcón.

Anexo 35: Presentación: Asesoría a la SUBPESCA en el contexto de la OROP-PS relacionados con la genética de *D. gigas*, expositora Dra. María Angélica Larraín.

Anexo 36: Difusión de resultados, Simposio XLIV Congreso Ciencias del Mar 2025.

V. Presentación de autores

Nombre	Función FIPA2023-19	Afiliación
María Angélica Larraín Barth	Directora de Proyecto	Universidad de Chile Facultad de Ciencias Químicas y Farmacéuticas Departamento de Ciencias de los Alimentos y Tecnología Química
Cristian Araneda Tolosa	Director Alterno	Facultad de Ciencias Agronómicas Departamento de Producción Animal
Valentina Córdova Alarcón	Investigadora en Genómica y Biotecnología	Facultad de Ciencias Químicas y Farmacéuticas Departamento de Ciencias de los Alimentos y Tecnología Química
Carlos Ignacio Vargas Manríquez	Investigador en Genómica y Acuicultura	Facultad de Ciencias Químicas y Farmacéuticas Departamento de Ciencias de los Alimentos y Tecnología Química
Alex Génova	Di Investigador en Genómica	Universidad de O'Higgins Instituto de Ciencias de la Ingeniería
Carol Moraga	Investigadora en Genómica	Universidad de O'Higgins Instituto de Ciencias de la Ingeniería

VI. Resumen ejecutivo

Dosidicus gigas (d'Orbigny 1835), llamada comúnmente jibia, calamar rojo, calamar de Humboldt o calamar gigante, es una especie endémica del Océano Pacífico Oriental, con gran importancia comercial en México, Perú y Chile. Caracterizada por su rápido crecimiento y ciclo de vida de aproximadamente dos años, su dinámica poblacional está fuertemente influenciada por las condiciones ambientales, como la temperatura y disponibilidad de alimento. Además, su distribución está relacionada con grandes corrientes oceánicas, y su población responde a fenómenos como El Niño (ENSO).

Hasta la fecha se reconocen dos posibles unidades poblacionales de *D. gigas*: una en el hemisferio norte (México-EE.UU.) y otra en el hemisferio sur (Ecuador-Perú-Chile). Respecto a esta última, los estudios genéticos realizados sugieren un alto flujo genético entre Chile y Perú, por lo que se asume que ambas pesquerías forman parte de una única unidad genética. Sin embargo, la baja diversidad genética observada podría representar una vulnerabilidad ante cambios ambientales, lo que demanda un análisis más profundo mediante marcadores moleculares de mayor resolución y distribuidos a lo ancho del genoma, como son los (polimorfismos de un solo nucleótido (SNPs).

Para una gestión sostenible de la pesquería de la jibia, es esencial definir las unidades poblacionales y entender la estructura genética de la especie. En consecuencia, el objetivo de esta consultoría fue Identificar la diversidad genética de *D. gigas* en un contexto temporal y espacial en el Océano Pacífico Sur, mediante la generación de información genómica y la obtención de un genoma de referencia. Este genoma permitirá identificar la diversidad genética y las adaptaciones evolutivas clave, generando información de gran importancia para la mejora de la administración y modelamiento biológico-pesquero de esta pesquería. Chile ha avanzado en esta línea con el proyecto FIPA 2020-20, sin embargo, aún queda trabajo para responder las preguntas que emergen en la gestión pesquera a nivel internacional. Este proyecto (FIPA 2023-19) contribuye al conocimiento genómico de *D. gigas* para fortalecer la administración pesquera en el Pacífico Sur, permitiendo una explotación sostenible del recurso.

Se tomaron todas las muestras de *D. gigas* en la Zona Económica Exclusiva (ZEE) de Chile, desde las macrozonas Norte y Centro, en las temporadas otoño - invierno y primavera - verano y se recibieron la totalidad de datos desde los proveedores de servicios de secuenciación. Los resultados por objetivo son: **OE1:** Obtener genoma de referencia con caracterización preliminar de los genes que conforman una fracción de la secuencia genómica de *D. gigas* para estudios genéticos en el Océano Pacífico Sur Este. Se obtuvo el genoma de referencia de 3,63 Gb organizado en 29 cromosomas T2T (de telómero a telómero) y 17 “scaffolds” que pueden corresponder a brazos de cromosomas, con una caracterización preliminar de 40.072 genes. El ensamble del genoma obtenido supera el estándar de calidad VGP-16 comprometido y es adecuado para ser considerado dentro de Proyecto Genomas de Vertebrados (VGP), esfuerzo internacional para generar genomas de referencia completos y de alta calidad. **OE2:** Analizar la existencia de diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de *D. gigas* en el

Océano Pacífico Sur Este. En base al análisis de las muestras tomadas en la ZEE Chile en el contexto de este proyecto (FIPA 2023-19) y el anterior (FIPA 2020-20) no se detectaron diferencias genéticas espaciales, pero si temporales (2008/2021 vs 2024) en las poblaciones de *D. gigas* capturadas en la ZEE de Chile. **OE3:** Conocer la asociación entre la variabilidad ambiental y las diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de *D. gigas* en el Océano Pacífico Sur Este. Se encontró que las variables ambientales analizadas no evidenciaron una contribución significativa a la variación genética total de *D.gigas* en aguas nacionales (ZEE), sin embargo, se identificaron 10 loci bajo selección que contribuyen en la adaptación local asociados a la variación en los niveles de nutrientes (ej: concentración de fitoplancton). **OE4:** Asesorar a la Subsecretaría de Pesca y Acuicultura en el contexto de la OROP-PS relacionadas con la genética de *D. gigas*. Se prestó asesoría a la contraparte técnica de SUBPESCA en el contexto de la reunión del comité científico de la OROP-PS 2024 realizado en Lima, Perú; alcanzándose un acuerdo para trabajar en la genética poblacional de la jibia con las contrapartes de China, Perú y Ecuador en base al genoma obtenido en este proyecto.

VII. Abstract

Dosidicus gigas (d'Orbigny 1835), commonly known as the cuttlefish, red squid, Humboldt squid, or giant squid, is an endemic species of the eastern Pacific Ocean, with great commercial importance in Mexico, Peru, and Chile. Characterized by its rapid growth and a two-year life cycle, its population dynamic is strongly influenced by environmental conditions, such as temperature and food availability. Furthermore, its distribution is linked to major ocean currents, and its population responds to phenomena such as El Niño-Southern Oscillation (ENSO).

To date, two possible population units of *D. gigas* have been recognized: one in the Northern Hemisphere (Mexico-USA) and another in the Southern Hemisphere (Ecuador-Peru-Chile). Regarding the latter, genetic studies suggest high gene flow between Chile and Peru, leading to the assumption that both fisheries are part of a single genetic unit. However, the low genetic diversity observed could represent a vulnerability to environmental changes, which requires more in-depth analysis using advanced techniques such as SNPs (single nucleotide polymorphisms).

To achieve sustainable management of the cuttlefish fishery, it is essential to define population units and understand the genetic structure of the species. In this context, the objective of this consultancy was to identify the genetic diversity of *D. gigas* in a temporal and spatial context in the South Pacific Ocean by generating genomic information and obtaining the reference genome of the species. This genome will allow the identification of genetic diversity and key evolutionary adaptations, generating key information for improving the management and biological-fishery modeling of this fishery. Chile has made progress in this area with the FIPA 2020-20 project, but work remains to be done to answer the questions that emerge in fisheries management at the international level. This project (FIPA 2023-19) contributes to the genomic knowledge of *D. gigas* to strengthen fisheries management in the South Pacific, enabling sustainable resource exploitation.

All *D. gigas* samples were collected in the Exclusive Economic Zone (EEZ) of Chile from the Northern and Central macrozones during the fall-winter and spring-summer seasons, and all data have been received from the sequence service providers. The results by objective are: **SO1**: Obtain a reference genome with preliminary characterization of the genes that make up a fraction of the *D. gigas* genomic sequence for genetic studies in the Eastern South Pacific Ocean. A 3.6 Gb reference genome organized into 29 T2T (telomere-to-telomere) chromosomes and 17 scaffolds (presumptively chromosome arms) was obtained, with a preliminary characterization of 40,072 genes. The resulting genome assembly exceeds the VGP-16 quality standard and is suitable for consideration within the Vertebrate Genomes Project (VGP), an international effort to generate complete, high-quality reference genomes. **SO2**: Analyze the existence of temporal genetic differences using highly variable markers in *D. gigas* populations in the Eastern South Pacific Ocean. Based on the analysis of samples taken in Chile in the context of this project (FIPA 2023-19) and the previous one (FIPA 2020-20), no spatial genetic differences were detected but temporal differences were detected (2008/2021 vs 2024) in the populations of *D. gigas* captured in the Chilean EEZ.

SO3: Understand the association between environmental variability and temporal genetic differences using highly variable markers in *D. gigas* populations in the Eastern South Pacific Ocean. It was found that the environmental variables analyzed did not show a significant contribution to the total genetic variation of *D. gigas* in national waters (EEZ). However, 10 loci under selection were identified that contribute to local adaptation associated with variation in nutrient levels (e.g., phytoplankton concentration). **SO4:** Advise the Subsecretaría de Pesca y Acuicultura in the context of the SPRFMO related to the genetics of *D. gigas*. Advice was provided to the technical counterpart of SUBPESCA in the context of the 2024 SPRFMO scientific committee meeting held in Lima, Peru; An agreement was reached to work on the population genetics of cuttlefish with counterparts from China, Peru, and Ecuador based on the genome obtained in this project.

1 Antecedentes

Dosidicus gigas, (d'Orbigny 1835), llamada comúnmente jibia, calamar rojo, calamar de Humboldt o calamar gigante, es un molusco cefalópodo, endémico de la región este del Océano Pacífico, que se distribuye verticalmente entre la superficie y los 1.200 metros de profundidad. Su rango de distribución geográfica comprende desde Alaska hasta el sur de Chile (Ibáñez et al., 2016b; Markaida, 2001). Debido a su amplia distribución en el sector del Pacífico Oriental, varios países (i.e., México, Perú, Chile) explotan comercialmente esta especie, y en la última década, se ha transformado en un recurso pesquero de alta relevancia, tanto en las zonas económicas exclusivas (ZEE) de los países ribereños al Pacífico, como para otros países que operan en aguas internacionales (Ibáñez and Ulloa, 2014).

Entre las características de *D. gigas* destacan, un rápido crecimiento, madurez temprana y un ciclo de vida no mayor a dos años. Sin embargo, estos rasgos son altamente variables, con una fuerte dependencia de las condiciones ambientales, por lo que, tanto la temperatura del mar y la disponibilidad de alimento controlarían las tasas de crecimiento e inicio del proceso de madurez, delimitando la longevidad y longitud máxima de esta especie (Argüelles and Tafur, 2010a). Por otra parte, se ha descrito que posee múltiples eventos de desove cerca de la superficie (Rocha et al., 2001), por lo que es considerado un desovante de tipo asincrónico oceánico. Por tanto, es probable observar ejemplares inmaduros en las cercanías de las costas (Argüelles and Tafur, 2010b) (Ibáñez et al., 2019; Nigmatullin et al., 2001; Staaf et al., 2008).

En su ciclo de vida, al igual que en otros calamares que tienden a estar sincronizados con los movimientos de las masas de agua (Bakun and Weeks, 2008), *D. gigas* muestra una tendencia a distribuirse con relación a grandes sistemas de corrientes oceánicas. El sistema de corrientes donde esta especie desarrolla su ciclo de vida está sujeto al sistema del Fenómeno del Niño o ENSO (El Niño-Southern Oscillation), uno de los fenómenos oceanográficos de gran escala más significativos. La influencia local que el ENSO tiene sobre la presencia de *D. gigas* varía latitudinalmente y se ha descrito, desde un punto de vista pesquero, como una influencia tanto negativa (Nevárez-Martínez et al., 2000) como positiva (Anderson and Rodhouse, 2001).

En cuanto a la migración de *D. gigas*, (Ibáñez et al., 2016a), plantean que las complejas fluctuaciones de su población podrían ser causadas por la interacción de procesos oceanográficos, biológicos y antropogénicos. La identificación y comprensión de los procesos y variables que influyen sobre la dinámica de la población de *D. gigas* es de vital importancia para el establecimiento de una administración efectiva de este recurso y su pesquería, tanto en las áreas comprendidas en la ZEE como en aguas internacionales.

Los estudios genéticos en *D. gigas* han sugerido la existencia de dos unidades poblacionales; una en el hemisferio Norte (México-EE.UU.) y otra en el hemisferio Sur (Perú-Chile) (Sandoval-Castellanos et al., 2010; D. J. Staaf et al., 2010). En el caso particular del Pacífico Sur, la información actualizada sugiere una baja diferenciación y un alto flujo genético entre

las poblaciones de Chile y Perú, por lo que formarían una misma unidad genética (Ibáñez et al., 2011; Ibáñez and Poulin, 2014; Sanchez et al., 2016a). Todos los estudios realizados hasta la fecha son concordantes al encontrar una baja diversidad haplotípica lo que sugiere una expansión demográfica entre el último período glacial y el actual. Estas investigaciones se han realizado con secuencias de genes mitocondriales o microsatélites (Ibáñez and Poulin, 2014; Sanchez et al., 2020a, 2016a; Sandoval-Castellanos et al., 2010).

Recientemente, (Xu et al., 2021) desarrollaron marcadores genómicos más variables como son SNPs (*Single Nucleotide Polymorphisms*) obtenidos mediante la técnica de secuenciación de bibliotecas de representación reducida del genoma llamada genotipado por secuenciación o GBS (*Genotyping-by-Sequencing*) para ejemplares de *D. gigas* recolectados frente a Ecuador, Perú y Chile. Dicho estudio logró el genotipado de un bajo número de SNPs (101), mostrando niveles variados de diversidad. Por lo tanto, se hace imperativo realizar nuevos estudios genéticos con marcadores genómicos para obtener un alto número de SNPs, y así poder conocer la diversidad genética y como ésta se estructura espacialmente, con el objetivo de mejorar la administración pesquera de este recurso en su área de distribución en el Pacífico Sur.

En el contexto descrito, para poder identificar la estructura genética de *D. gigas* de forma más detallada, detectar variaciones genéticas asociadas a cambios ambientales, e identificar stocks pesqueros de forma más precisa (Lu and Luo, 2020) es necesario constituir un genoma de referencia para la especie. Hasta la fecha, sólo se conocen once genomas de nueve especies de cefalópodos, entre los que no estaba *D. gigas* (Albertin et al., 2015; Da Fonseca et al., 2020)¹. En consecuencia, es de crucial importancia aumentar la información genética y genómica existente para comprender la evolución y adaptación de estos cefalópodos.

Previo a las técnicas de secuenciación de segunda generación, “Next Generation Sequencing” (NGS) o “short reads”, los marcadores moleculares más utilizados, como las secuencias de ADN nuclear, ADN mitocondrial y microsatélites, sólo analizaban una pequeña fracción de todo el genoma que permitía sólo responder preguntas acotadas. En general, se han utilizado marcadores neutrales, por lo que no se consideraba el rol de la selección natural sobre la diversidad genética poblacional. En este contexto, al contar con un genoma de referencia, se pueden realizar análisis más detallados y acabados, los que al hacerlos a escala de todo el genoma (en lugar de usar uno o unos pocos genes), permitirán comprender de forma más detallada la historia demográfica de la especie. Adicionalmente, contar con un genoma de referencia, no sólo podría aumentar las capacidades para hacer inferencias sobre los procesos poblacionales, sino que también ofrece un medio para estudiar el “paisaje genómico” y cómo éste está moldeado dentro y entre poblaciones (Ellegren, 2014; Seehausen et al., 2014).

¹ <https://www.ncbi.nlm.gov/data-hub/taxonomy/6605>

2 Objetivos

2.1 Objetivo General

Identificar la variabilidad genética temporal de *Dosidicus gigas* en un contexto espacial en el Océano Pacífico Sur Este.

2.2 Objetivos Específicos

- i. **OE1:** Obtener genoma de referencia con caracterización preliminar de los genes que conforman una fracción de la secuencia genómica de *Dosidicus gigas* para estudios genéticos en el Océano Pacífico Sur Este.
- ii. **OE2:** Analizar la existencia de diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de *Dosidicus gigas* en el Océano Pacífico Sur Este.
- iii. **OE3:** Conocer la asociación entre la variabilidad ambiental y las diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de *Dosidicus gigas* en el Océano Pacífico Sur Este.
- iv. **OE4:** Asesorar a la Subsecretaría de Pesca y Acuicultura en el contexto de la OROP-PS relacionadas con la genética de *Dosidicus gigas*.

3 Metodología de trabajo por objetivo

3.1 Obtención de muestras de *Dosidicus gigas*

De acuerdo con la propuesta técnica, para la obtención de muestras se definieron tres macrozonas dentro de la Zona Económica Exclusiva de Chile (Macrozonas Norte – Centro y Sur) y dos temporadas de muestreo (otoño - invierno y primavera – verano). Para cada una de las macrozonas, se propuso obtener un total de 100 individuos, 50 correspondientes a la temporada de primavera-verano y 50 a la temporada de otoño-invierno. El muestreo de cada temporada se realizó en un período no superior a un mes, con el objetivo de no obtener muestras de un mismo “grupo genético” (en caso de existir) por el desplazamiento de individuos entre zonas geográficas.

Durante los muestreos se buscó obtener igual proporción de sexos en las muestras (1:1), para esto se determinó el sexo de cada individuo a partir de la observación macroscópica de las gónadas (**Figura 1**). En caso de no alcanzar la proporción deseada, se continuó con el muestreo de tejido hasta obtener un número similar de machos y hembras ($n \sim 25$ de cada sexo). Para cada individuo muestreado, se registró el peso (Kg), longitud del manto (cm) y sexo (**Figura 2**). Finalmente, para la obtención de ADN se obtuvo tejido del cuarto brazo de cada individuo, por duplicado. Una porción se almacenó en tubo de 15 mL con etanol al 90% y la otra en tubo de 1,5 mL con RNAlater (Thermo Fisher Scientific). Los tubos se trasladaron dentro de neveras con hielo al laboratorio, dónde fueron almacenados a -80°C hasta la extracción de ADN. Adicionalmente, en coordinación con la pesca artesanal, se obtuvo las coordenadas geográficas correspondientes a cada muestreo (Datum WGS84).



Figura 1. Identificación del sexo de los individuos de *D. gigas* muestreados. A la izquierda, una gónada femenina (Ovario) y a la derecha, una gónada masculina (Testículos).



Figura 2. Medición de parámetros biométricos durante la obtención de muestras de *D. gigas*. A la izquierda, medición del largo total del manto y a la derecha, peso de cada ejemplar.

3.2 OE1. Obtener genoma de referencia sin anotar de *Dosidicus gigas* para estudios genéticos en el Océano Pacífico Sur Este.

3.2.1 Extracción de ADN y construcción de librerías de ADN para secuenciación

El proceso de obtención del genoma completo de *D. gigas* se llevó a cabo a partir de la secuenciación de ADN genómico de un ejemplar hembra proveniente de la Macrozona Centro de la Zona Económica Exclusiva de Chile (MO-01), usando dos aproximaciones metodológicas: Primero, secuenciación de fragmentos largos en la plataforma PacBio HiFi Revio y con la tecnología Oxford Nanopore PromethION. Segundo, secuenciación Hi-C con tecnología Illumina NovaSeq para identificar y fasear los cromosomas. Adicionalmente se realizaron lecturas cortas “short reads” del individuo MO-01 más 14 individuos extras (8 hembras y 6 machos) obtenidas por secuenciación con tecnología MGI Tech.

La extracción de ADN de alto peso molecular para secuenciación de fragmentos largos en plataforma PacBio HiFi Revio y Oxford Nanopore PromethION, se realizó desde el individuo MO-01. Éste también fue utilizado para secuenciación Hi-C. Adicionalmente, se obtuvo ADN para la secuenciación short reads con tecnología MGI desde los individuos: MO-01, 02, 04, 05, 06, 08, 10, 11, 12, 13, 18, 28, 30, 31 y 32). Se usó el kit HMW (High Molecular Weight) Monarch® para tejido. El kit emplea un método de captura usando perlas de vidrio, en lugar de columnas de sílica, recomendado para la extracción de ADN en fragmentos de tamaño superiores a 100 Kb. Este proceso incluye la lisis del tejido y la eliminación de proteínas, lo que se realiza mediante homogeneización y agitación controlada para ajustar el tamaño de los fragmentos de ADN, alcanzando tamaños ideales (50–250 Kb) para las tecnologías de secuenciación. El ADN capturado en las perlas se libera en un solo paso de elución, lo que asegura una extracción pura y eficiente, optimizando el rendimiento en secuenciación de lecturas largas al evitar fragmentación y contaminación, factores importantes para obtener secuenciaciones de gran calidad. Adicionalmente, para mejorar la calidad y pureza del ADN resultante, se utilizó el Monarch® Spin gDNA Extraction Kit ya que ha demostrado buenos resultados en una variedad de muestras biológicas incluyendo sangre y tejidos. Para determinar la calidad del ADN extraído, previo a la secuenciación se comprobó la longitud y calidad de los fragmentos de ADN mediante electroforesis capilar en los equipos TapeStation (Agilent Technologies Inc.) y Q-Seq1 Bio-Fragment Analyzer (BioOptic Inc.). El TapeStation es un sistema de electroforesis automatizado que mide el tamaño y la concentración de los fragmentos mediante cartuchos de gel y marcadores moleculares. Por otro lado, el Q-Seq1 cuantifica y verifica el tamaño de los fragmentos usando tecnologías de fluorescencia, garantizando que las muestras estén en el rango óptimo para la secuenciación. También se cuantificó la concentración por fluorometría en un equipo Qubit (Thermo Fischer) y se determinaron las razones 260/280 y 260/230 para verificar la pureza del ADN extraído.

La **construcción de librerías de ADN** para las secuenciaciones PacBio HiFi y Oxford Nanopore PromethION, se realizó utilizando el kit Ligation Sequencing Kit V14. Este kit permite la ligación eficiente de adaptadores de secuenciación específicos, optimizando la integridad y calidad de las lecturas en las secuencias de ADN. El objetivo fue obtener

aproximadamente 60X de cobertura de lecturas PacBio HiFi (N50 > 15Kb) y 30X de lecturas Oxford Nanopore (N50 > 30Kb), para lo cual se utilizaron 3 flow cells. Para la secuenciación Hi-C la preparación de muestras se llevó a cabo utilizando el kit Dovetail Hi-C, seguido de la secuenciación en Illumina NovaSeq.

3.2.2 Ensamble y anotación *de novo* del genoma de referencia

Para obtener el genoma de referencia, se ensamblaron las lecturas de PacBio y Nanopore utilizando el software WENGAN (Di Genova et al., 2021), ensamblador desarrollado para integrar de forma óptima, información de secuencias obtenidas con distintas tecnologías, logrando resultados más completos. WENGAN se caracteriza por generar genomas altamente contiguos, de alta calidad de secuencia y completitud. Los contigos generados por WENGAN, fueron orientados y ordenados en “scaffolds” (andamios) utilizando las lecturas Hi-C. Adicionalmente, se ancló y corrigió el genoma utilizando la información de interacción espacial proporcionada por las secuencias Hi-C con los programas Juicebox (Durand et al., 2016) y 3D-DNA, asegurando así una estructura cromosómica precisa en el ensamblaje final.

Para completar la secuencia del genoma, las lecturas cortas de los 15 individuos fueron alineadas al genoma de referencia previamente construido, utilizando el programa BWA, que mapea eficazmente las lecturas a su ubicación genómica. Para la anotación de variantes y la asignación de funciones y efectos, se utilizó SNPeff, que, en conjunto con la anotación funcional, permite predecir los impactos de las variantes detectadas.

La anotación del genoma para identificar potenciales genes *in silico* se llevó a cabo utilizando AUGUSTUS y GeneMark (Lomsadze et al., 2014). Luego, se identificaron elementos repetitivos con RepeatMasker, y se identificó la función de los genes mediante la búsqueda de homologías funcionales en BLAST. Finalmente, se evaluó la completitud del ensamble con BUSCO utilizando la base de datos mollusca_odb10 junto con la evaluación funcional de los genes obtenidos.

La secuencia del genoma y su anotación estarán disponibles en el repositorio de NCBI una vez realizada la publicación de este en revistas de corriente principal. Se informará a SUBPESCA una vez que se obtenga el número de acceso.

3.3 OE2. Analizar la existencia de diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de *Dosidicus gigas* en el Océano Pacífico Sur Este.

3.3.1 Genotipado mediante tecnología DArT-seq

Para el estudio de las poblaciones naturales, muestras de tejido de 188 individuos representativos de cada localidad y temporada de la ZEE de Chile (**Tabla 1**) fueron enviadas a la empresa Diversity Array Technology Pty LT, Canberra, Australia para ser genotipadas por secuenciación mediante la tecnología DArT-seq². La metodología DArT, integra las técnicas de lectura de secuencias asociadas a sitios de restricción (“restriction-site-associated DNA sequencing”) y la hibridación de ADN (Cruz et al., 2013; Tinker et al., 2009). En la primera etapa se identifican los marcadores SNPs, y en la segunda, a partir de la impresión de las secuencias identificadas en un chip de ADN, se realiza el genotipado de las muestras.

Para cada muestra, el ADN se obtuvo desde el tejido usando el kit comercial NucleoMag (Macherey-Nagel). Se redujo la complejidad del genoma mediante la digestión de ~50 ng de ADN genómico con las enzimas *Pst*I y *Sph*I (Cruz et al., 2013). Posteriormente, se ligaron adaptadores de secuenciación y códigos de identificación mediante PCR. Cada biblioteca, una por cada muestra analizada, fue sometida a 100 ciclos de secuenciación (Single Reads) en el equipo NovaSeq Illumina® 6000. Solicitamos el mismo protocolo empleado en el proyecto FIPA20-20, con el propósito de hacer comparables y poder integrar los resultados de ambos proyectos en el análisis.

Para la identificación y genotipado de marcadores SNPs, el proveedor procesa y filtra las secuencias obtenidas según su calidad (considerando un umbral de Phred de 10 para cada secuencia, estableciendo límites de confianza de un 90% para al menos el 50% de las bases obtenidas). Posteriormente, para la lectura de los códigos de identificación (“barcode”), usó un umbral de calidad Phred de 30, considerando un 99,9% de precisión para el 75% de las bases. Se obtuvieron en promedio 161.464 secuencias por muestra, las que fueron ordenadas y agrupadas según su código de identificación. La búsqueda de variantes alélicas se realizó mediante el análisis y agrupación de todas las secuencias (Pailles et al., 2017). Asimismo, se realizó la evaluación de la distribución mendeliana de los alelos en las muestras analizadas con el objetivo de diferenciar variantes alélicas reales (SNPs verdaderos) de secuencias parálogas. Finalmente, con el fin de evaluar la reproducibilidad de la técnica, 188 réplicas técnicas (100% de todas las muestras analizadas) fueron reprocesadas partiendo desde el ADN genómico hasta la búsqueda de variantes alélicas.

² <https://www.diversityarrays.com/>

Tabla 1. Información de las muestras de Jibia (*Dosidicus gigas*) obtenidas por los proyectos FIPA 2020-20 y FIPA 2023-19 en la ZEE chilena y utilizadas para el análisis de diversidad genética y estructura poblacional.

Fuente	TMP	MZ	Localidad	Región	ID	n	Sexo	n	Talla [cm]	Peso [kg]	Fecha	Lat (S)	Lon (O)
FIPA2020	Verano	Centro	I. Quiriquina	Biobío	Centro_V_08	20	I	20	NA	NA	01-02-08	-36,63	-73,05
			Quintay	Valparaíso					76,5 ± 4,9	NA			
	Otoño	Centro	Sn Antonio	Valparaíso	Centro_O_21	62	H	38	73,2 ± 5,2	NA	05-05-21	-33,19	-71,70
			Lebu	Biobío			M	24		NA			
FIPA2023	Otoño	Norte	Coquimbo	Coquimbo	Norte_O_21	31	H	17	75,1 ± 5,0	NA	04-06-21	-29,97	-71,38
							M	14	75,1 ± 5,1	NA			
	Otoño	Centro	Maguillines	El Maule	Centro_O_24	47	H	25	77,3 ± 8,1	16,87 ± 5,78	13-07-24	-35,29	-72,79
							M	22	65,8 ± 7,4	9466 ± 3,70			
	Otoño	Norte	Guayacán	Coquimbo	Norte_O_24	47	H	1	71	11,85			
							I	45	75,2 ± 14,2	11,94 ± 8,36	04-07-24	-29,74	-71,54
							M	1	80	18,9			
	Verano	Centro	Maguillines	El Maule	Centro_V_24	50	H	26	68,0 ± 4,0	9,81 ± 3,99	28-12-24	-32,14	-71,59
							M	24	69,4 ± 4,0	10,18 ± 2,92			
	Verano	Norte	Pichidangui	Coquimbo	Norte_V_24	44	H	31	76,0 ± 3,3	13,15 ± 6,26			
							I	1	76	15	23-12-24	-32,14	-71,59
							M	12	76,08 ± 2,47	13,06 ± 4,28			

MZ: Macrozona, TMP: Temporada, H: Hembras, M: Machos, I: Indeterminado

3.3.2 Filtrado de datos genotípicos

Para todos los análisis genéticos también se consideraron los genotipos obtenidos durante la ejecución del proyecto FIPA20-20 “Genética poblacional de *Dosidicus gigas* presente en territorio nacional”. En la **Tabla 1** se detalla el origen y características del total de muestras analizadas en el OE2.

La identificación de los paneles de SNPs se realizó utilizando el paquete DartR en el ambiente R (Gruber et al., 2018; Ihaka and Gentleman, 1996; Mijangos et al., 2022). Se aplicaron distintos filtros a los genotipos buscando eliminar sesgos derivados de la técnica de secuenciación, de la calidad de las muestras y de los SNPs identificados, descartando así SNPs falsos positivos y manteniendo SNPs verdaderos. Posteriormente, mediante el uso del paquete SNPrelate, se evaluó el desequilibrio de ligamiento entre marcadores, para eliminar aquellos que presenten esta condición. El resumen de los filtros aplicados se describe en la **Tabla 2**.

Posterior al filtrado, el número mínimo de individuos por grupo fue de 12 (Centro_V_08) (**Tabla 1**). Estos grupos fueron analizados usando los criterios descritos para el uso de marcadores SNP (Aguirre-Liguori et al., 2020; Li et al., 2020; Nazareno et al., 2017; Qu et al., 2020).

Tabla 2. Proceso de filtrado de calidad de datos DArT-Seq.

Filtro	Descripción	Umbral	Paquete
Reproducibilidad	Elimina el loci cuando su reproducibilidad del genotipado independiente de 30 alelos es menor al umbral	0,95	
Monomórficos	Elimina loci monomórficos del set de datos	-	
Profundidad de lectura	Retiene los loci con un número de lecturas entre un mínimo y máximo	3 - 150	
Tasa de genotipado por loci	Retiene los loci genotipados sobre un umbral definido en todas las muestras	0,80	
SNPs Secundarios	Retiene solo un SNP por secuencia	-	
Tasa de genotipado individual	Elimina los individuos que presentan una tasa de genotipado menor al umbral definido	0,80	dartR
Frecuencia del Alelo Menor (MAF)	Retien los loci que poseen una frecuencia del alelo menor ("Minor Allele Frequency") mayor al umbral definido	0,01	
HWE	Elimina los loci que muestran desviaciones de las proporciones establecidas por el equilibrio de Hardy-Weinberg (HWE), en al menos dos grupos.	2	
MD	Elimina los loci que no se genotiparon en un grupo completo	-	
LD	Remueve los loci que se encuentran en Desequilibrio de ligamiento	$r > 0.2$	SNPRelate

3.3.3 Identificación de marcadores putativamente neutrales y bajo selección

La identificación de loci putativamente neutrales y adaptativos se realizó usando los paquetes pcadapt (Luu et al., 2017), fsthet (Flanagan and Jones, 2017) y baypass (Gautier, 2015). Se consideraron putativamente bajo selección todos los loci identificados con al menos una de las tres aproximaciones. Con esta información se separó el panel de marcadores totales, en dos: un panel de marcadores putativamente neutrales y otro de marcadores putativamente bajo selección.

3.3.4 Caracterización de la diversidad genética de *Dosidicus gigas*.

Se estimó la diversidad genética por loci y población utilizando los paquetes *dartR*, *diversity* (Keenan et al., 2013) y *hierfstat* (de Meeûs and Goudet, 2007; GOUDET, 2005) determinando Riqueza Alélica (A_r), Heterocigosidad Observada (H_o), Heterocigosidad Esperada (H_e) e Índice de Consanguinidad (F_{IS}). Se compararon los valores de entre poblaciones con la prueba de Kruskal-Wallis (Hollander and Wolfe, 1973) en combinación con la prueba de Dunn, considerando un nivel alfa del 5% mediante el paquete *ggstatplot* (Patil, 2021).

3.3.5 Descripción de la estructura genética poblacional.

Para la caracterización de la diferenciación genética entre poblaciones se utilizó el Índice de Fijación Global (F_{ST}), que indica el grado en que las frecuencias alélicas de dos poblaciones se diferencian entre sí. Su magnitud se relaciona de forma inversamente proporcional con el flujo de genes. La diferenciación genética, medida como F_{ST} (Weir and Cockerham, 1984), se calculó con el paquete *hierfstat* con un intervalo de confianza del 95% (IC95) estimado a partir de 10.000 iteraciones (Pembleton et al., 2013).

La estructura poblacional, o como se organizan los individuos en tiempo y espacio, se visualizó primeramente mediante el Análisis de Componentes Principales (PCA; “Principal Components Analysis”) que resume la variación dentro y entre grupos (Elhaik, 2022). Se usó el paquete *dartR*, sin incluir la información temporal ni espacial de los grupos analizados. El número de Componentes Principales utilizado se basó en el criterio de $n-1$, donde n es el número de individuos analizados (Miller et al., 2020).

A continuación, se evaluó la estructura poblacional por dos metodologías: En primer término, se realizó un Análisis Discriminante de Componentes Principales (DAPC; “Discriminant Analysis of Principal Components”) con el paquete *Adegenet* v2.1.2 (Jombart, 2008; Jombart et al., 2010; Jombart and Ahmed, 2011). Este análisis maximiza las diferencias entre grupos, minimizando las diferencias dentro de cada grupo, permitiendo una mejor asignación de individuos a grupos predefinidos. El número de clusters se definió *de novo* mediante la función *find.clusters*, en base al criterio de información bayesiana (BIC), calculado con $n-1$ componentes principales (n = número de grupos de muestreo analizados) (Miller et al., 2020; Thia, 2023). Posteriormente, se calculó el número de componentes principales a considerar para el DAPC mediante la estimación del MSE (Mean Squared Error), obtenida con el método de validación cruzada. La validación cruzada se realizó en 1000 réplicas y considerando el número de componentes principales que contribuyen con el 80% de la varianza total.

En segundo lugar, se utilizó *Admixture* v1.3.0 (Alexander et al., 2009), método bayesiano que modela simultáneamente las frecuencias alélicas de la población y sus proporciones de ancestría mediante un estimador de máxima verosimilitud. Se usó el modo no supervisado, sin incluir la información temporal ni espacial de los grupos. El número óptimo de clusters K fue definido en base al error de la validación cruzada (“cross-validation error”) determinado variando K entre 1 y 20, realizando en 20.000 iteraciones (Alexander and Lange, 2011).

3.3.6 Tamaño Efectivo Poblacional y Parentesco Genético

El Tamaño Efectivo de la Población (N_e) se determinó con el programa Ne-estimator v2.01 (Do et al., 2014), basándose en la estimación del desequilibrio de ligamento (LD) que evalúa la asociación no aleatoria entre alelos de distintos loci a través de la covarianza (Hill, 1981; Marandel et al., 2020; Waples and Do, 2008). Consideramos que el sistema de cruce fue de tipo aleatorio, se utilizó un $p_{crítico} = 0,05$ y se consideraron los loci con una frecuencia del alelo menor $MAF < 0,05$. El intervalo de confianza del 95% para el N_e se determinó usando el método de “jackknife” (Waples and Do, 2008).

El parentesco genético entre pares de individuos (R_{xy}) al interior y entre poblaciones (Wang, 2002), se calculó mediante el paquete iRel (Gonçalves da Silva and Russello, 2011). Este coeficiente estima el parentesco de dos individuos basándose en Identidad por Descendencia (IBD; “Identity By Descent”), utilizando el estimador de momentos para dos y cuatro genes (Lynch and Ritland, 1999).

3.4 OE3. Conocer la asociación entre la variabilidad ambiental y las diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de *Dosidicus gigas* en el Océano Pacífico Sur Este.

3.4.1 Estructura espacio-temporal y asociación con variables ambientales de *Dosidicus gigas*

Para estudiar la contribución de las variables ambientales a la variabilidad genética de *Dosidicus gigas* en la ZEE de Chile, es necesario tomar en cuenta la estructura espacial y temporal de los individuos analizados como covariables que corrigen el modelo global (la correlación entre la diferenciación genética poblacional y la distancia geográfica entre zonas de muestreo). Para ello, se modeló en primer lugar la contribución de la estructura espacial a la variación genética global, mediante un análisis de redundancia con el paquete vegan v2.5 en R (Oksanen et al., 2019), utilizando una matriz de coordenadas de distancia transformada en vectores propios de distancia de Moran (dbMEM; Legendre and Fortin, 2010).

Las variables ambientales fueron obtenidas a partir de la base de datos de información meteorológica y climática global de alta resolución espacial Bio-Oracle (Assis et al., 2017; Tyberghein et al., 2012). La base de datos Bio-Oracle³ es una plataforma gratuita y pública de cartografía marina a escala global, que contiene datos de variables geofísicas, bióticas y ambientales históricas del reino marino bentónico y de superficie, obtenidas a partir de ráster SIG. Los datos poseen una resolución espacial de 5 minutos de arco (9,2 km en el ecuador) en series de tiempo decadales. Además, la plataforma contiene variables proyectadas en el tiempo para modelar condiciones futuras hasta el año 2100 (Assis et al., 2024). Esta información permite modelar la distribución y presencia de las especies marinas de acuerdo con características oceanográficas y ambientales (Assis et al., 2017). Se analizaron las variables de salinidad, temperatura, pH, oxígeno disuelto, concentración de

³ <https://www.bio-oracle.org/>

clorofila, silicato, nitrato, fosfato, hierro disuelto y fitoplankton, a nivel superficial, a profundidad media y profundidad máxima, y dirección y velocidad del agua superficial.

Para evaluar la correlación de la variabilidad ambiental con la variabilidad genética, primero, se calculó la colinealidad entre las variables ambientales, obtenidas con el índice de correlación de Pearson utilizando el paquete psych v1.9.12.31 en R (Revelle, 2020). Cuando la correlación entre las variables ambientales fue mayor a $|0.7|$, se mantuvo sólo una variable. A continuación, se estimó el modelo de regresión óptimo entre cada variable ambiental y genética por separado. Se mantuvieron aquellas variables ambientales que contribuyen a la variación genética medida como frecuencias alélicas entre grupos genéticos. Para evaluar la contribución relativa de la estructura espacial y las variables ambientales a la variación genética, se realizó un análisis de redundancia utilizando las variables ambientales de cada macrozona muestreada como variables explicativas del modelo con el paquete vegan v2.5-6 en R (Oksanen *et al.* 2019). Por último, se realizó un análisis de redundancia parcial para obtener la contribución conjunta y por separado de la estructura espacio-temporal y las variables ambientales a la variación genética.

3.4.2 Identificación de loci candidatos de adaptación local en *Dosidicus gigas*

Las variables ambientales que contribuyen significativamente a la variación genética identificadas previamente se usaron para identificar loci candidatos de adaptación local a través del RDA y el modelo mixto de factores latentes (LFMM; Frichot *et al.* 2013). El análisis de RDA se realizó con el paquete de R vegan v2.5-6 (Oksanen *et al.*, 2019). Los loci que explican una porción significativa ($p < 0.01$) de la variación genética, fueron seleccionados como loci candidatos de adaptación local. LFMM es un modelo de regresión que toma en cuenta variables no observadas (factores latentes) que corrigen el modelo (Frichot *et al.*, 2013). Se utilizó la prueba LFMM implementada en el paquete de R LEA (Frichot and François, 2015). El número de factores latentes se estimó a partir de los grupos genéticos obtenidos en los análisis de estructura poblacional ($K=1$). Los loci candidatos fueron identificados según su p-value ($p < 0.05$), ajustado por el método de FDR de Benjamini and Hochberg (1995). Para identificar aquellos SNPs candidatos de adaptación local se consideraron los loci identificados como F_{ST} outliers. Finalmente, se realizó una búsqueda en BLASTn (NCBI; <https://blast.ncbi.nlm.nih.gov/Blast.cgi>) para investigar si los loci candidatos están o no asociados a una respuesta funcional.

3.5 OE4. Asesorar a la Subsecretaría de Pesca y Acuicultura en el contexto de la OROP-PS relacionadas con la genética de *Dosidicus gigas*.

La asesoría a la Subsecretaría de Pesca y Acuicultura en las materias del proyecto, así como en las relacionadas a la biología y genética de *D. gigas* comprometida en la propuesta se materializó en la participación del equipo del proyecto en las actividades que se detallan en resultados.

4 Resultados

4.1 Obtención de muestras de *Dosidicus gigas*

El muestreo de la temporada otoño-invierno se realizó durante los meses de julio y agosto de 2024, y el de primavera-verano durante el mes de diciembre del mismo año, ambos en las Macrozonas Norte y Centro dentro de la ZEE de Chile. Los ejemplares fueron obtenidos desde la pesca artesanal, en coordinación con Sindicatos de Pescadores Artesanales pertenecientes a la Macrozona Norte (Región de Coquimbo), y Macrozona Centro (Región del Maule) (**Tabla 3**).

El Durante el muestreo otoño-invierno de la macrozona Norte (Norte_O_24), realizado en la caleta Guayacán, Región de Coquimbo, se obtuvo un total de 50 individuos de un peso y longitud del manto promedio de $15,61 \pm 5,41$ Kg y $76,31 \pm 8,46$ cm, respectivamente (**Tabla 3, Anexo 1**). En esta oportunidad no fue posible determinar el sexo de los individuos ya que no nos autorizaron a abrir los individuos debido a la disminución del valor pagado por las plantas de proceso a los pescadores por los ejemplares intervenidos.

El muestreo otoño-invierno de la macrozona Centro (Centro_O_24) se realizó en el Puerto de Maguillines, Región del Maule, donde se obtuvo un total de 50 individuos, 28 hembras y 22 machos, con un peso y longitud del manto promedio de $13,59 \pm 6,19$ Kg y $72,26 \pm 9,66$ cm, respectivamente (**Tabla 3, Anexo 2**). Las hembras muestreadas fueron más grandes que los machos, alcanzando una longitud promedio del manto de $77,32 \pm 8,12$ cm significativamente mayor que los $65,82 \pm 7,44$ cm de los machos ($t_{48\text{ gl}} = 5.16$, $p\text{-value} = 0,0000047$). Durante el muestreo, en los primeros 50 individuos revisados, se identificó un mayor número de hembras que de machos, obteniendo una proporción sexual de 3:1. Para alcanzar una proporción más cercana de hembras y de machos, se buscó completar las muestras con individuos machos hasta alcanzar, en lo posible, los 25 individuos de cada sexo. Adicionalmente, se almacenaron las gónadas de 21 individuos (10 hembras y 11 machos) en el medio Bouin para realizar histología de dichas muestras y corroborar la asignación de sexo realizada *in situ*, la cual resultó correcta para todas las muestras analizadas (**Anexo 3**).

Durante la temporada primavera-verano, en la Macrozona Norte (Norte_V_24) se realizó el muestreo de 50 ejemplares, 35 hembras y 14 machos, en la caleta Pichidangui, Región de Coquimbo. Los individuos presentaron un tamaño promedio de $75,25 \pm 4,05$ cm y $25,24 \pm 2,46$ Kg para las hembras y $74,85 \pm 3,86$ cm y $13,61 \pm 1,86$ Kg para los machos (**Tabla 3, Anexo 4**). Por último, en el muestreo primavera-verano en la Macrozona Centro se obtuvo 50 ejemplares, 26 hembras y 24 machos desde el Puerto de Maguillines, Región del Maule. Los individuos presentaron un tamaño promedio de $68,03 \pm 3,99$ cm y $10,93 \pm 1,80$ Kg para las hembras y $69,38 \pm 3,94$ cm y $10,62 \pm 1,99$ Kg para los machos (**Tabla 3, Anexo 5**). Los individuos del muestreo de la Macrozona Centro en la temporada de primavera-verano (Centro-V-24) fueron significativamente más pequeños que el resto de los individuos muestreados (Kruskal-Wallis, $p < 0,05$) (**Anexo 6**).

Tabla 3. Número de ejemplares y datos biométricos de *D. gigas* obtenidos por macrozona, región y temporada.

TMP	MZ	Región	Localidad	ID	n	Sexo	n	Talla [cm]	Peso [kg]	Fecha	Lat °S	Lon °O
Otoño Invierno	Norte	Coquimbo	Caleta Guayacán	Norte-O-24	50	H	1	71	11,85 ± 0	04-07-24	-29.7	-71.54
						M	1	80	NA			
						I	48	76,3 ± 8,6	15,7 ± 5,5			
Otoño Invierno	Centro	Maule	Puerto Maguillines	Centro-O-24	50	H	28	77,3 ± 8,1	16,8 ± 5.8	13-07-24	-35.2	-72.79
						M	22	65,8 ± 7,4	9,5 ± 3.7			
Primavera Verano	Norte	Coquimbo	Caleta Pichidangui	Norte-V-24	50	H	35	75,3 ± 4,0	15,2 ± 2.5	23-12-24	-32.1	-71.59
						M	14	74,9 ± 3,9	13,7 ± 1.9			
						I	1	76	15			
Primavera Verano	Centro	Maule	Puerto Maguillines	Centro-V-24	50	H	26	68,0 ± 4,0	10,9 ± 1,8	28-12-24	-34.9	-72.33
						M	24	69,4 ± 3,9	10,6 ± 2,0			

MZ: Macrozona, TPM: Temporada, H: hembras, M: machos, I: Indeterminados. Detalle de individuos en Anexo 1, 2, 3 y 4.

En la Macrozona Sur, *D. gigas* no es una especie objetivo de la pesca artesanal ni industrial, lo que se refleja en los bajos niveles de capturas de esta especie registrados históricamente desde la Región de Los Ríos al Sur (**Figura 3**). Se tomó contacto con el Sr. Aquiles Sepúlveda del Instituto de Investigación Pesquera (INPESCA), y paralelamente, por indicación de la contraparte técnica de SUBPESCA (Sr. Lorenzo Flores), contactamos al Instituto de Fomento Pesquero (IFOP), para intentar obtener muestras de la macrozona Sur. IFOP comprometió un “muestreo de oportunidad”. Las muestras se obtendrían de los ejemplares que aparecen como pesca incidental o fauna acompañante, y serían tomadas por los observadores científicos de IFOP (**Anexo 7**). Con fecha 05/09/2024 se hizo envío del protocolo (**Anexo 8**) y material fungible necesario para la toma de muestras de *D. gigas* a IFOP Puerto Montt (Sr. Héctor Huerta). Hasta la fecha, se han recibido 53 muestras de tejido de *D. gigas* (28 hembras y 25 machos) enviadas por INPESCA (**Anexo 9**). Estas muestras se obtuvieron durante los meses de julio y agosto de 2024 desde la Macrozona Centro (Región del Maule y Región del Biobío) y fueron almacenadas a -20°C. El detalle de las muestras obtenidas por INPESCA se muestran en el **Anexo 9**.

En cuanto a las muestras de otro miembro de la OROP-PS (Ecuador) usadas en el estudio anterior (FIPA 2020-20), SUBPESCA informó en la Reunión de coordinación inicial que éstas no están disponibles para este estudio. Se han hecho gestiones para obtener nuevas muestras, pero debido a las regulaciones internas de Perú y de Ecuador con relación a muestras para análisis genéticos solicitadas por investigadores extranjeros, no fue factible obtenerlas durante el desarrollo del actual proyecto. Esto entendiendo que estas muestras se usaran en presentaciones oficiales ante la OROP-PS, y por lo tanto deben cumplir con toda la normativa de los países donantes.



Figura 3. Estadística pesquera de *Dosidicus gigas* de los últimos 5 años en cada macrozona estudiada en la ZEE de Chile (SUBPESCA 2024).

4.2 OE1. Obtener genoma de referencia sin anotar de *Dosidicus gigas* para estudios genéticos en el Océano Pacífico Sur Este.

4.2.1 Extracción de ADN y construcción de librerías de ADN para secuenciación

El ADN para la secuenciación de genoma completo de referencia de *D. gigas*, se obtuvo desde muestras capturadas en la Macrozona Centro, Puerto de Maguillines en la temporada de otoño-invierno 2024 (**Anexo 2**). Tanto el ADN del ejemplar hembra MO-01 usado para la secuenciación de fragmentos largos en plataformas PacBio HiFi Revio y Oxford Nanopore PromethION, Hi-C y el ADN de los otros individuos (MO-02, 04, 05, 06, 08, 10, 11, 12, 13, 18, 28, 30, 31 y 32) secuenciados para baja cobertura (IcWGS).

Las 15 muestras cumplen con los criterios de control de calidad (QC) necesarios para la secuenciación. Los valores de concentración obtenidos mediante Qubit, y las relaciones 260/280 y 260/230 se muestran en **Anexo 10**. El tamaño promedio de los fragmentos de 15 Mb y el valor DQN promedio de 9,64, reflejan una alta integridad del ADN, asegurando que las muestras son aptas para obtener datos de secuenciación confiables y de alta calidad.

4.2.2 Ensamble y anotación *de novo* del genoma de referencia

En el periodo informado completamos la generación de un ensamblaje cromosómico de *D. gigas* (Jibia_V02) que supera holgadamente los umbrales de calidad comprometidos (contig N50 > 10 Mb; “scaffold” N50 > 30 Mb; QV > 30; BUSCO > 90 %), alcanzando el estándar VGP-16 (Rhie et al. 2021), necesario para ensamblar un genoma *de novo* resuelto por haplotipos y a nivel telómero-a-telómero (T2T).

Se realizó un total de dos corridas de PacBio HiFi a partir de ADN de alto peso molecular de MO-01, obteniendo una cobertura genómica de 64X con lecturas de N50 superior a 18 kb y una calidad promedio mayor a Q28 (1 error cada 1000 pares de bases) (**Anexo 11 – A y C**). Se llevaron a cabo dos corridas de Oxford Nanopore (PromethION), a partir de ADN de alto peso molecular de MO-01, logrando una cobertura de 40X con lecturas de N50 superior a 5 kb (longitudes de hasta 150 kb) y calidades promedio cercanas a Q20 (1 error cada 100 pares de bases) (**Anexo 11 – B y C**). Respecto a la secuenciación Hi-C de dos bibliotecas (MO-01 y MO-02), en el individuo MO-01 obtuvo una calidad superior a Q37 (1 error cada 10000 pares de bases) (**Anexo 12 - A**), aportando 183,6 Gb correspondientes a 1.230 Millones de lecturas, mientras que MO-02 Hi-C generó 206,7 Gb y 1.388 Millones lecturas; en ambas bibliotecas, > 99,5 % de las bases superan Q20 y $\approx$ 98.4 % superan Q30, con una tasa de error de solo 0,02 % (**Anexo 12 - B**).

La combinación de las secuenciaciones de Nanopore y PacBio generó una cobertura estimada de 100X para el genoma de *D. gigas*, que integradas con HIFiasm y ancladas mediante 390 Gb de datos Hi-C de alta fidelidad (**Anexo 11 – C y 12 - B**), produjo un ensamblaje primario de 3.63 Gp con un contig N50 de 39 Mb y “scaffold” N50 de 78 Mb (**Tabla 4**). La Tabla 4 compara las principales métricas de los tres ensamblajes generados para Jibia_V02: el ensamblaje primario y los dos haplotipos hap1 y hap2. Todos muestran un tamaño genómico semejante (3.52–3.63 Gpb) y una contigüidad a nivel de “scaffold”

prácticamente idéntica ($N50 \approx 78$ Mb, $L50 = 20-21$), con “scaffolds” máximos de 104–108 Mb. Sin embargo, el ensamblaje primario es notablemente más continuo a nivel de contig ($N50 = 39$ Mb frente a 19 Mb y 18 Mb). Los tres ensamblajes conservan un contenido GC muy similar ($\sim 33.3\%$) y 29 “scaffolds” alcanzan el estado telómero-a-telómero indicando que son cromosomas completos (29 T2T en el ensamblaje primario y 24-25 por haplotipo, **Tabla 4**) y 17 “scaffolds” que podrían corresponder a brazos cromosómicos, con un error de secuencia $< 1 \times 10^{-4}$ ($QV \approx 40$). La completitud BUSCO es alta y homogénea, alcanzando un 89.4 % de ortólogos completos a nivel genómico (**Figura 4**), esto confirma que tanto el ensamblaje colapsado como los haplotipos resueltos capturan de forma exhaustiva el repertorio génico de un molusco, siendo el primario el más continuo y los haplotipos ligeramente más fragmentados debido al proceso de separación de fases.

Estos resultados sitúan al ensamblaje de *Jibia_V02* en el cuartil superior respecto de 45 genomas de cefalópodos disponibles en NCBI, alcanzando el mayor $N50$ de contigs (~ 40 Mpb) con un tamaño genómico de ~ 3.6 Gpb, superando con holgura a los demás genomas comparados (0–12 Mpb) (**Figura 4**); además lidera la contigüidad dentro del clado de los cefalópodos, con el mayor contig $N50$ y el mayor número de T2T “scaffolds”, seguida por los ensamblajes de la sepia *Eumandya parva* (GCA_963457665.1) y del calamar diamante *Thysanoteuthis rhombus* (GCA_963457665.1) (**Anexo 13**). Estos resultados indican que nuestro ensamblaje *Jibia_V02* alcanza un alto grado de resolución cromosómica y contigüidad entre los genomas cercanos disponibles en NCBI.

Tabla 4. Estadísticas de ensamble para *Dosidicus gigas*.

	Ensamble primario	Haplotipo 1	Haplotipo 2	VGP-2016*
Scaffolds	724	853	448	
Largo total scaffold	3.633.682.487	3.537.358.601	3.526.134.160	
Scaffold N50	78.428.331	78.836.937	78.102.568	>10.000.000
Scaffold L50	21	20	20	
Tamaño máximo scaffold	107.899.994	103.244.024	106.722.806	
Tamaño mínimo scaffold	263	238	892	
cóntigos	888	1.242	838	
Largo total cóntigo	3.633.666.087	3.533.703.687	3.522.338.874	
Contig N50	39.069.295	19.121.515	18.498.182	>1.000.000
Contig L50	36	61	65	
Tamaño máximo cóntigo	88.590.342	59.804.115	52.282.662	
Tamaño mínimo cóntigo	263	238	892	
# gaps en scaffolds	164	389	390	<1.000
Largo total gap en scaffolds	16.400	3.654.914	3.795.286	
Largo promedio gap en scaffolds	100.00	9395.67	9731.50	
Gap N50 en scaffolds	100	1.516.443	408.262	
Gap L50 en scaffolds	82	2	3	
Máximo gap en scaffolds	100	1.549.208	1.144.113	
Mínimo gap en scaffolds	100	100	100	
Contenido GC %	33.29	33.31	33.30	
T2T scaffolds	29	24	25	
BUSCO completitud %	89.37	89.75	87.76	>90%

Las unidades de largo y tamaño de “scaffolds”, cóntigos y “gaps” está expresado en bases nucleotídicas. *VGP-2016: Estándar de calidad de ensamblaje de genoma de referencia *de novo* según el Vertebrate Genoma Project (Rhie et al., 2021).

BUSCO Assessment (Mollusca odb10, n=5,295)

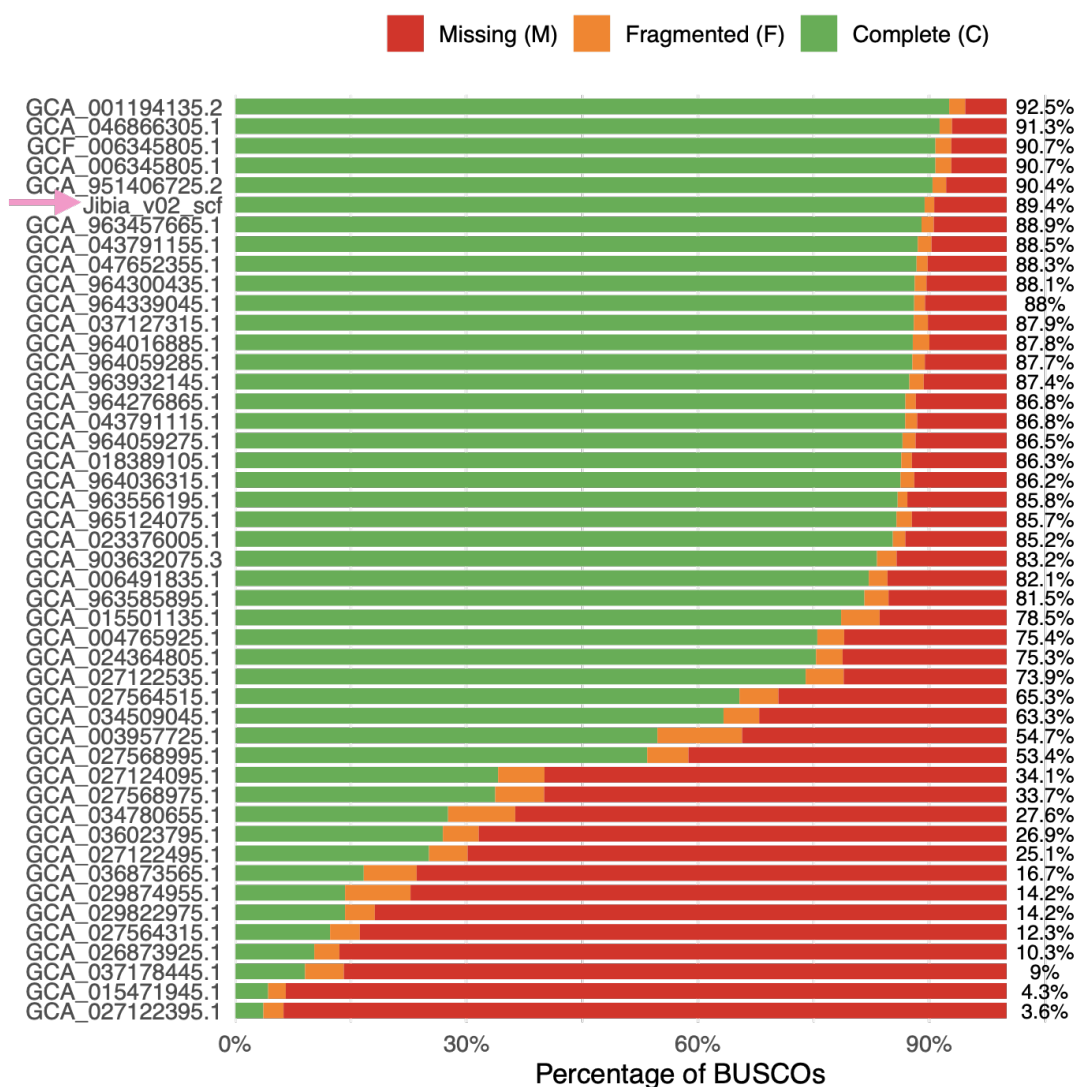


Figura 4. Evaluación BUSCO de 46 ensamblajes de moluscos usando el set mollusca_odb10 (n = 5 295 ortólogos de copia única).

La anotación de genes comprende 40.072 genes codificantes y 44.725 transcritos (mRNA) (Tabla 5), con una media de 1,1 isoformas por gen. Cada transcrito contiene en promedio 5,6 exones y 4,6 intrones, cifras coherentes con la complejidad génica esperable en moluscos. Aun así, un número considerable de genes (18.156; 45%) carece de intrones (genes de un solo exón). En total se identificaron 252.655 exones y 207.957 intrones, de los cuales el 99% pertenece a regiones codificantes. Los codones de inicio y término fueron detectados prácticamente en todos los mRNA (~44,7 k), lo que respalda la integridad estructural de las predicciones. La evaluación funcional mediante BUSCO a nivel de mRNA arroja un 87,0 % de ortólogos completos, situando la anotación dentro de los estándares

aceptados para genomas de referencia y confirmando que la mayoría del repertorio génico esencial está representado en los modelos generados.

Asimismo, evaluamos la idoneidad del nuevo genoma como referencia transcripcional mapeando 1.286 millones de lecturas RNA-seq provenientes de tres calamares ommastrephídeos (26 muestras de *Sthenoteuthis oualaniensis*, 6 de *D. gigas* y 5 de *Todarodes pacificus*) depositadas en NCBI. Las tasas medias de alineamiento al genoma ensamblado de *D. gigas* mostraron un claro gradiente filogenético: 87 % en lecturas de *D. gigas*, 36 % en *S. oualaniensis* y 15 % en *T. pacificus* mapean al genoma, lo que confirma tanto la solidez del ensamblaje para su propia especie como la disminución esperada de compatibilidad a mayor distancia evolutiva. Antes de la anotación, el genoma fue enmascarado con WindowMasker, identificándose 55 % de contenido repetitivo. Con el genoma enmascarado y la evidencia RNA-seq anterior, ejecutamos BRAKER3, obteniendo los modelos de genes de la **Tabla 5**.

Tabla 5. Estadísticas de la anotación génica del ensamble de *D. gigas*.

Estimación	Valor
Número de genes	40.072
Número de mRNA	44.725
Número de CDS	44.725
Número de exones	252.655
Número de intrones	207.957
Número de codón de inicio	44.703
Número de codón de término	44.707
Número de exones en CDS	252.655
Número de intrones en CDS	207.930
Número de intrones en exones	207.930
Número de intrones en intrones	181.574
Número de genes exon single	18.156
Número de mRNA exon single	18.351
Promedio mRNAs por gen	1,1
Promedio CDSs por mRNA	1,0
Promedio exones por mRNA	5,6
Promedio intrones por mRNA	4,6
Promedio exones por CDS	5,6
Promedio intrones en CDSs por mRNA	4,6
Promedio intrones en exones por mRNA	4,6
Promedio intrones en intrones por mRNA	4,1
BUSCO mRNA complete %	87,01

Para evaluar la utilidad del genoma como referencia para estudios poblacionales, se mapearon las lecturas cortas (short reads) obtenidas en el equipo MGI de 15 individuos adicionales (9 hembras, 6 machos, **Figura 5**). Un total entre ~570 y 950 millones de lecturas fueron alineadas al genoma de referencia, todos los individuos alcanzaron 20–35X de cobertura, 98–99 % de alineación correcta y tasas de duplicados del 2–4 % (**Anexo 14**), demostrando una alta compatibilidad entre las secuencias de los individuos de la especie y el nuevo genoma obtenido, constituyendo una referencia completa para *Dosidicus gigas*. En conjunto, los datos evidencian un rendimiento alto y uniforme, de alta fidelidad en todas las secuenciaciones de lectura corta realizadas.

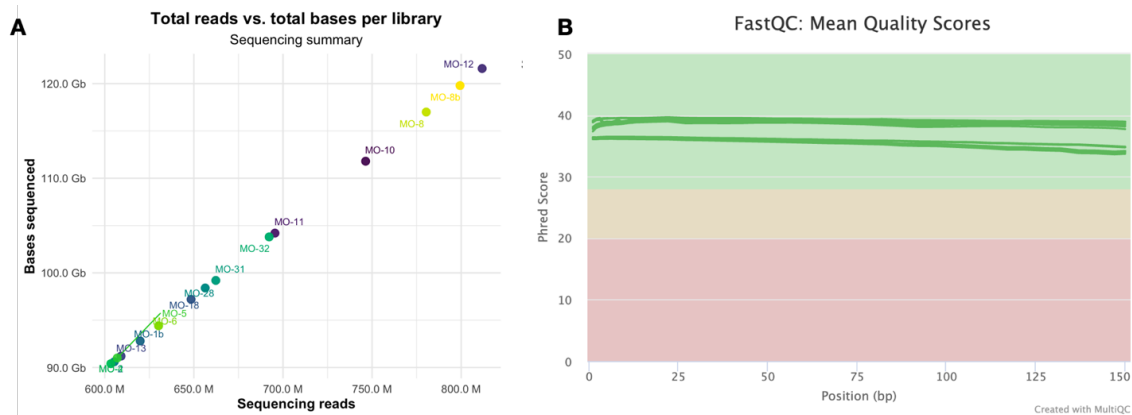


Figura 5. Resumen de la secuenciación y control de calidad de lecturas cortas de 15 individuos de *D. gigas*.

A) Relación entre el número total de lecturas (eje x, millones) y las bases generadas (eje y, gigabases) para cada muestra secuenciada (MO-01 – MO-32). Los puntos muestran una tendencia lineal (600–800 M lecturas $\Rightarrow$ 90–122 Gb), indicando eficiencia de secuenciación homogénea entre individuos. **B)** Perfil de la calidad media por posición a lo largo de las lecturas de 150 pb, reportado por FastQC. Todas las secuenciaciones presentan calidades Phred $\approx$ 37–40 a lo largo de la lectura, permaneciendo dentro de la franja verde ($>$ Q30), sin caídas bruscas hacia los extremos.

Por último, y como resultado adicional de la propuesta, se realizó la identificación del genoma mitocondrial circular completo (19.6 Kb; **Figura 6**), lo que refuerza la completitud del ensamble del genoma de referencia para *Dosidicus gigas*. En conjunto, estos resultados acreditan que hemos cumplido el OE 1: disponemos de un genoma de referencia altamente contiguo, a nivel cromosómico y funcionalmente anotado, apto para cumplir con las obligaciones internacionales contraídas por el estado de Chile, y para aplicaciones regulatorias, de manejo pesquero y de investigación científica en el Océano Pacífico Sur Este.

Este genoma de referencia será la base para la metodología de secuenciación genómica de baja cobertura (lcWGS) acordada en la decimosegunda reunión del comité científico de OROP-PS para realizar los estudios de estructuración poblacionales con marcadores SNPs.

El genoma de referencia generado para *Dosidicus gigas* quedará disponible en el repositorio de NCBI (<https://www.ncbi.nlm.nih.gov/datasets/genome>). Una vez cargado en la base de datos se informará el número de acceso a SUBPESCA para su registro.

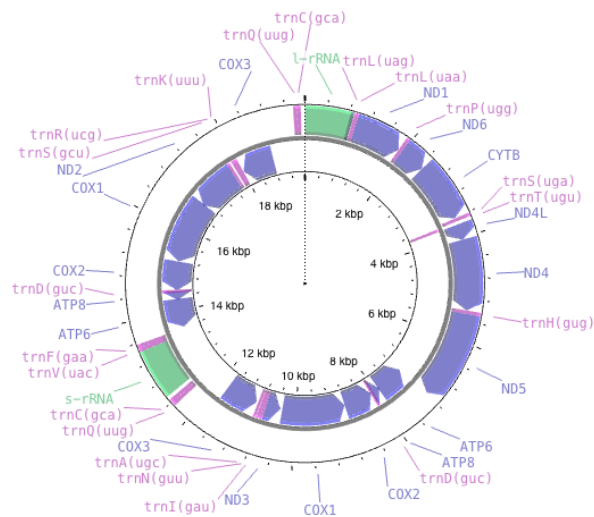


Figura 6. Genoma mitocondrial identificado en el ensamble de *Dosidicus gigas*. El genoma mitocondrial identificado corresponde al contig ptg000596l con una longitud de 19558 pares de bases.

4.3 OE2. Analizar la existencia de diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de *Dosidicus gigas* en el Océano Pacífico Sur Este.

4.4 Diversidad y Estructura Genética en poblaciones naturales

4.4.1 Secuenciación y construcción de base de datos

Para el análisis de diversidad genética espacial y temporal, fueron enviadas a secuenciar mediante DArT-seq un total de 188 muestras de tejido del manto, 47 muestras por cada macrozona (Norte y Centro) y temporada de muestreo (otoño-invierno y primavera-verano).

Resultado del proceso de secuenciación y genotipado se obtuvo un total de 19,840 secuencias, con un 39,47 % de datos perdidos (%MD), para un total de 170 individuos distribuidos en 4 macrozonas (grupos de muestras). Adicionalmente, se consideraron para los análisis de diversidad, 94 individuos genotipados durante la ejecución del proyecto FIPA20-20, los que fueron agrupados por macrozonas y temporada de muestreo que también fueron obtenidas desde la ZEE de Chile. El número de individuos analizados en cada grupo de muestras se detallan en **Tabla 6**.

Tabla 6. Número de muestras por grupo/localidad.

Proyecto	ID Grupo	n
FIPA20-20	Centro_V_08	12
	Centro_O_21	54
	Norte_O_21	28
FIPA23-19	Centro_O_24	42
	Norte_O_24	44
	Centro_V_24	43
	Norte_V_24	41

Una vez aplicado el filtrado de secuencias, la base de datos para el estudio de diversidad y estructura genética en *D. gigas* contó con un número de 1.313 SNPs y un total de 283 individuos pertenecientes a los 7 grupos estudiados. La **Tabla 7** presenta el detalle de los resultados obtenidos durante el filtrado de secuencias para la identificación de SNPs.

Tabla 7. Resultados filtrado preliminar base de datos.

Filtro	Umbral	Nº SNP	Nº Ind	MD
Reproducibilidad	0,95	19.840	288	38,47%
Monomórficos	-	19.487	288	39,69%
Profundidad	3-150	19.366	288	39,80%
Tasa genotipado loci	0,80	6.152	288	7,09%
Secundarios	-	3.842	288	7,60%
Tasa genotipado individual	0,80	3.842	283	7,23%
MAF	0,01	2.054	283	7,59%
HWE	P<0,05	1.862	283	6,91%
Loci en LD	MAF: 0,05	1,313	283	7,52%
Panel global de SNP		1,313	283	7,52%

MD: genotipos faltantes (missing data)

4.4.2 Identificación de loci putativamente bajo selección

Con el objetivo de identificar los loci que están putativamente bajo selección, se analizaron los 1.313 SNPs totales mediante tres métodos (pcadapt, fsthet y BayPass). Resultado del análisis, se encontró un total de 165 loci putativamente bajo selección (19,84% del total de SNPs), de los cuales 5 SNPs fueron seleccionados por pcadapt (0,38 % del total de SNPs), 94 SNPs por fsthet (7,16 % del total de SNPs) y 74 SNPs por BayPass (5,64 % del total de SNPs).

La **Figura 7** presenta un Diagrama de Venn que indica el número y proporción de marcadores identificados por cada método. Al respecto, del total de marcadores putativamente bajo selección, ninguno fue identificado por los tres métodos. Identificados por al menos dos métodos se observaron 8 SNPs, de los cuales 7 loci fueron identificados por fsthet y BayPass (4,20%), y 1 loci por pcadapt y fsthet (0,60%).

Para la construcción del panel de marcadores putativamente bajo selección, se consideraron los 165 SNPs identificados por los tres métodos (7,94% de MD). El panel de marcadores neutrales se conformó con el resto de los marcadores 1.148 SNPs (7,46% de MD).

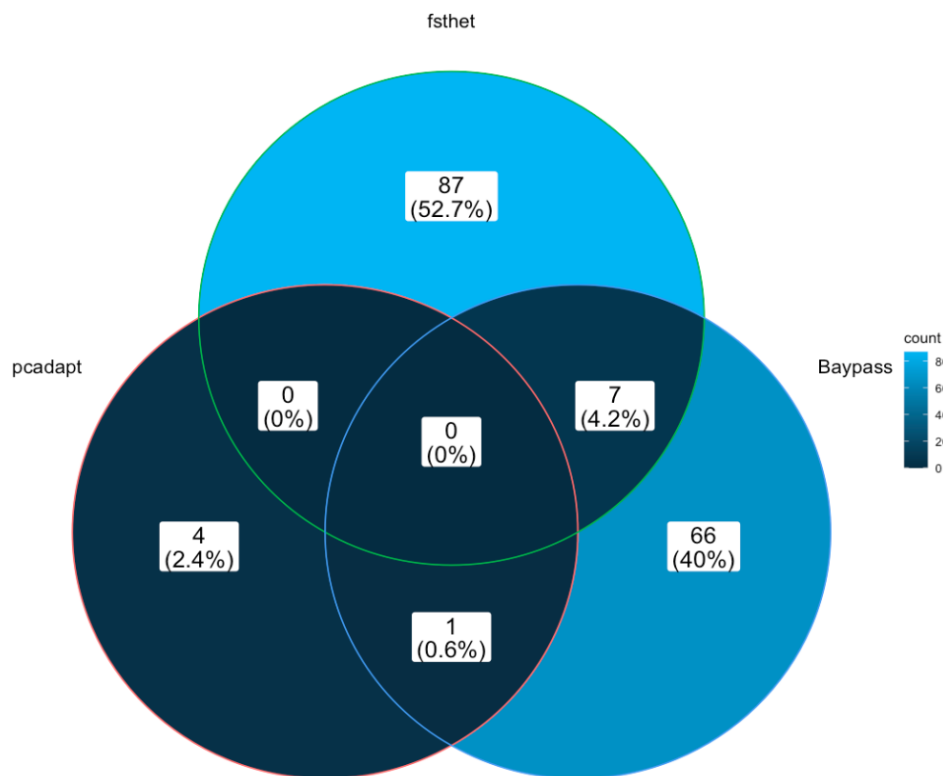


Figura 7. Diagrama de Venn mostrando el resultado de la identificación de loci putativamente bajo selección.

4.4.3 Panel total de marcadores SNPs

En términos de la diversidad genética global, estimada con los 1.313 SNPs, la heterocigosidad media observada (H_o) presentó un valor de 0,0870. La heterocigosidad esperada (H_e) presentó un valor de 0,1338. El índice de consanguinidad (F_{IS}) presentó un valor de 0,3513 (IC95%: 0,3423 – 0,3595) y el índice de diferenciación (F_{ST}) global presentó un valor de 0,0026 (IC95%: -0,0009 – 0,0063). No se observaron grupos con desvío significativos de HWE ($p > 0,05$). Respecto a la Ar, el valor máximo fue encontrado en el grupo Centro_V_08 (1,501) y el mínimo en Centro_O_21. (1,577). La heterocigosidad observada (H_o) fluctuó entre 0,091 (Centro_V_08) y 0,081 (Norte_V_24), y la H_e presentó un valor mínimo en la localidad de Centro_V_08 (0,128) y máximo en Norte_V_24 (0,133). Por su parte, el F_{IS} presentó valores entre 0,233 y 0,360, en los grupos Centro_V_08 y Norte_V_24, respectivamente. El detalle de la diversidad genética a nivel poblacional, estimada con el panel total de 1.313 SNPs se presenta en la **Tabla 8**. Asimismo, el detalle de la diversidad genética estimada con el panel de marcadores neutrales y putativamente bajo selección se muestran en la **Tabla 9** y **Tabla 10**, respectivamente.

Tabla 8. Diversidad genética global por localidad estimada con panel total de 1.313 SNPs.

Grupo	n	<i>Ar</i>	<i>Ho</i>	<i>He</i>	<i>F_{IS}</i>
Centro_V_08	12	1,501	0,091 ^a	0,128 ^b	0,233 ^d
Centro_O_21	54	1,577	0,089 ^b	0,132 ^a	0,303 ^b
Norte_O_21	28	1,553	0,088 ^b	0,130 ^{ab}	0,295 ^{cd}
Centro_O_24	42	1,558	0,083 ^b	0,132 ^{ab}	0,359 ^{bc}
Norte_O_24	44	1,550	0,09 ^b	0,131 ^{ab}	0,269 ^b
Centro_V_24	43	1,554	0,087 ^b	0,129 ^a	0,311 ^a
Norte_V_24	41	1,564	0,081 ^b	0,133 ^a	0,360 ^a

Letras distintas indican diferencias significativas entre poblaciones ($p < 0,05$).

Tabla 9. Diversidad genética neutral por localidad estimada con panel neutral de 1.148 SNPs.

Grupo	n	<i>Ar</i>	<i>Ho</i>	<i>He</i>	<i>F_{IS}</i>
Centro_V_08	12	1.497	0,090 ^b	0,125 ^a	0,223 ^d
Centro_O_21	54	1.581	0,089 ^a	0,132 ^a	0,295 ^b
Norte_O_21	28	1.561	0,088 ^a	0,131 ^{ab}	0,297 ^{cd}
Centro_O_24	42	1.562	0,083 ^a	0,132 ^a	0,364 ^{bc}
Norte_O_24	44	1,55	0,088 ^a	0,129 ^a	0,272 ^b
Centro_V_24	43	1,56	0,087 ^a	0,130 ^a	0,307 ^a
Norte_V_24	41	1.566	0,081 ^a	0,131 ^a	0,357 ^a

Letras distintas indican diferencias significativas entre poblaciones ($p < 0,05$).

Tabla 10. Diversidad genética adaptativa por localidad estimada con panel putativamente bajo selección de 165 SNPs.

Grupo	n	<i>Ar</i>	<i>Ho</i>	<i>He</i>	<i>F_{IS}</i>
Centro_V_08	12	1,524	0,102	0,152	0,308
Centro_O_21	54	1,553	0,088	0,135	0,361
Norte_O_21	28	1,505	0,085	0,127	0,279
Centro_O_24	42	1,530	0,087	0,135	0,321
Norte_O_24	44	1,539	0,102	0,139	0,25
Centro_V_24	43	1,505	0,085	0,127	0,336
Norte_V_24	41	1,556	0,083	0,141	0,378

Letras distintas indican diferencias significativas entre poblaciones ($p < 0,05$).

4.4.4 Diferenciación poblacional

La diferenciación genética medida a través del F_{ST} pareados no muestra un patrón de diferenciación poblacional claro. Se observan valores con diferencias significativas entre los grupos estudiados ($p < 0,05$), sin embargo, los valores de F_{ST} pareados son menores al 0,05% lo que es considerado muy bajo. A nivel espacial no se observa una diferenciación genética clara entre los grupos muestreados en la macrozona centro y norte de la ZEE de Chile. En términos temporales, no se observa diferenciación genética entre los distintos grupos. El detalle de los resultados obtenidos para la diferenciación poblacional con el panel de marcadores total, neutral y putativamente bajo selección se presenta en las **Tabla 11**, **Tabla 12** y **Tabla 13**.

Tabla 11. Valores de F_{ST} pareados estimados con panel total de 1.313 SNPs. F_{ST} pareados (bajo la diagonal) y sus p-values (sobre la diagonal).

-	Centro_V_08	Centro_O_21	Centro_O_24	Centro_V_24	Norte_O_21	Norte_O_24	Norte_V_24
Centro_V_08	-	0,2885	0,026	0,503	0,3217	0,1987	0,4234
Centro_O_21	0,0011	0	0,001	0,0055	0,0001	0,015	0,0024
Centro_O_24	0,0047	0,0039**	0	0,0043	0,0298	0,0256	0,0151
Centro_V_24	0,0001	0,0033**	0,0026**	0	0,0039	0,1152	0,1811
Norte_O_21	0,0009	0,0038**	0,0030*	0,0038*	0	0,0234	0,0032
Norte_O_24	0,0021	0,0028*	0,0018*	0,0011	0,0032*	0	0,1926
Norte_V_24	0,0005	0,0034**	0,0020**	0,0009	0,0041**	0,0008	0

* Indica diferencias significativas, $p < 0,05$. ** Indica diferencias altamente significativas, $p < 0,01$.

Tabla 12. Valores de F_{ST} pareados estimados con panel neutral de 1.148 SNPs. F_{ST} pareados (bajo la diagonal) y sus p-values (sobre la diagonal).

-	Centro_V_08	Centro_O_21	Centro_O_24	Centro_V_24	Norte_O_21	Norte_O_24	Norte_V_24
Centro_V_08	0	0,6405	0,3669	0,9189	0,7521	0,7203	0,8199
Centro_O_21	-0,0007	0	0,0049	0,1726	0,0012	0,1384	0,0296
Centro_O_24	0,0007	0,0025*	0	0,037	0,0791	0,1153	0,0543
Centro_V_24	-0,0028	0,0008	0,0017*	0	0,0369	0,1775	0,6611
Norte_O_21	-0,0014	0,0036**	0,0016	0,0019*	0	0,0923	0,0122
Norte_O_24	-0,0011	0,0009	0,0011	0,0009	0,0016	0	0,3712
Norte_V_24	-0,0018	0,0017*	0,0015	-0,0003	0,0027*	0,0004	0

* Indica diferencias significativas, $p < 0,05$. ** Indica diferencias altamente significativas, $p < 0,01$.

Tabla 13. Valores de F_{ST} pareados estimados con panel putativamente bajo selección de 165 SNPs. F_{ST} pareados (bajo la diagonal) y sus p-values (sobre la diagonal).

-	Centro_V_08	Centro_O_21	Centro_O_24	Centro_V_24	Norte_O_21	Norte_O_24	Norte_V_24
Centro_V_08	0	0,0316	0,0019	0,0357	0,0177	0,012	0,0695
Centro_O_21	0,0128	0	0,029	0,0016	0,0874	0,029	0,0162
Centro_O_24	0,0303**	0,0137*	0	0,0129	0,166	0,0669	0,0648
Centro_V_24	0,0184*	0,0194**	0,0087*	0	0,0133	0,1653	0,0048
Norte_O_21	0,0146*	0,0054	0,0126	0,0162*	0	0,0843	0,1011
Norte_O_24	0,0233*	0,0158*	0,0074	0,0024	0,0143	0	0,0922
Norte_V_24	0,0151	0,0148*	0,0052	0,0087**	0,0131	0,0040	0

* Indica diferencias significativas, $p < 0,05$. ** Indica diferencias altamente significativas, $p < 0,01$.

4.5 Estructura Poblacional

4.5.1 Análisis de Componentes Principales

El Análisis de Componentes Principales (PCA) con el total de marcadores SNP muestra que la variación genética es explicada por los primeros dos componentes, que explican un 1,1% y 1% de la variación, respectivamente (**Anexo 15**). El PCA ejecutado con los marcadores neutrales, muestra que la variación genética es explicada también por los primeros dos componentes, los que explican un 1 de la variación, respectivamente (**Anexo 16**). De la misma forma, el PCA realizado con los marcadores putativamente bajo selección estima que la variación genética es explicada por los primeros dos componentes, con 4,7 % y 2,9% respectivamente (**Figura 8**).

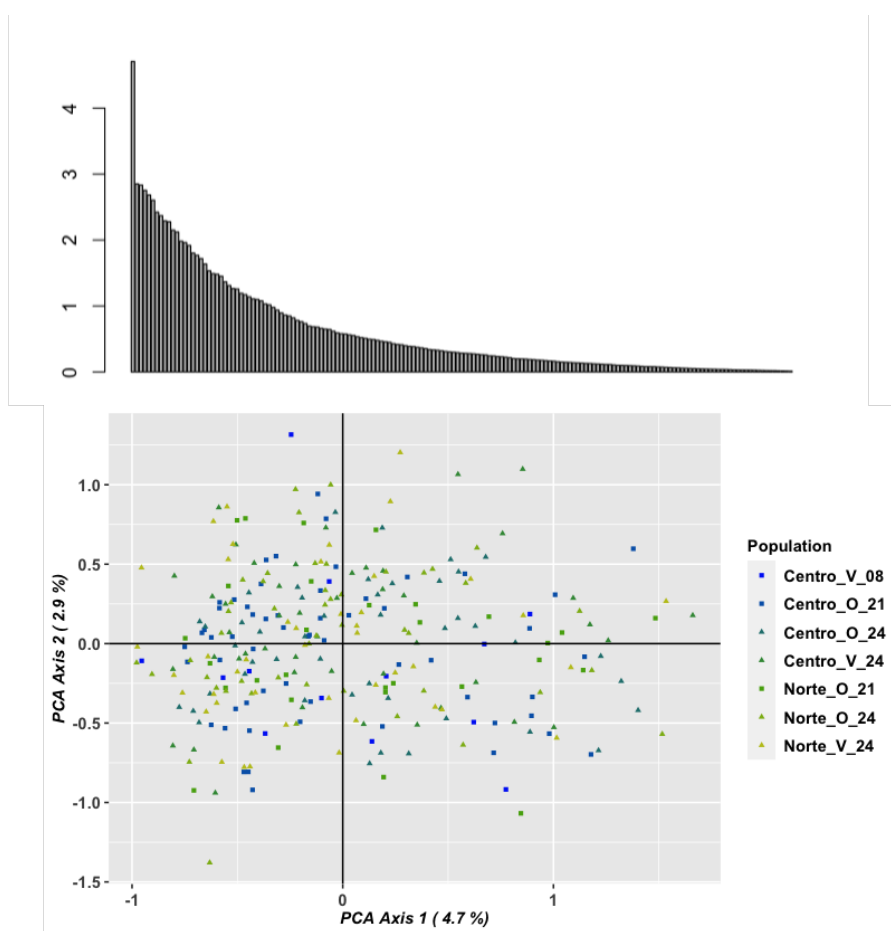


Figura 8. Análisis de Componentes Principales obtenido con 165 SNPs. En el panel superior se encuentra la representación gráfica de los porcentajes de la varianza genética explicada por cada componente principal.

4.5.2 Análisis Discriminante de Componentes Principales (DAPC)

El DAPC realizado con el panel total de marcadores (1.313 SNPs) muestra que no hay distintos grupos en la ZEE de Chile a partir del valor menor de BIC, que fue igual a uno, mientras que el número óptimo de componentes principales retenidos fue de 100. Para el análisis del total de marcadores la primera función discriminante explica el 19,46 % de la varianza genética, mientras que la segunda explica el 12,58 % de la variación. El análisis no separa los grupos estudiados, no obstante, si muestra una tendencia de agrupación entre las muestras del 2008 y 2021 (FIPA20-20), respecto de las muestras obtenidas en invierno y verano del 2024 (FIPA23-19) (**Figura 9**).

El DAPC realizado con el panel neutral de marcadores (1.148 SNPs) muestra que no hay distintos grupos en la ZEE de Chile a partir del valor menor de BIC, que fue igual a uno, mientras que el número óptimo de componentes principales retenidos fue de 60. Para el análisis del total de marcadores la primera función discriminante explica el 14,89 % de la varianza genética, mientras que la segunda explica el 8,11 % de la variación (**Anexo 17**).

El DAPC realizado con el panel de marcadores putativamente bajo selección (165 SNPs) muestra que hay dos grupos distintos en la ZEE de Chile a partir del valor menor de BIC, mientras que el número óptimo de componentes principales retenidos fue de 55. El análisis indica que la primera función discriminante explica el 66,30 % de la varianza genética, mientras que la segunda explica el 4,29 % de la variación. En este caso el DAPC separa los grupos estudiados, mostrando una tendencia de agrupación entre las muestras del 2008 y 2021 (FIPA20-20), separándolas de las muestras obtenidas en las temporadas de otoño-invierno y primavera-verano del 2024 (FIPA23-19) (**Figura 10**).

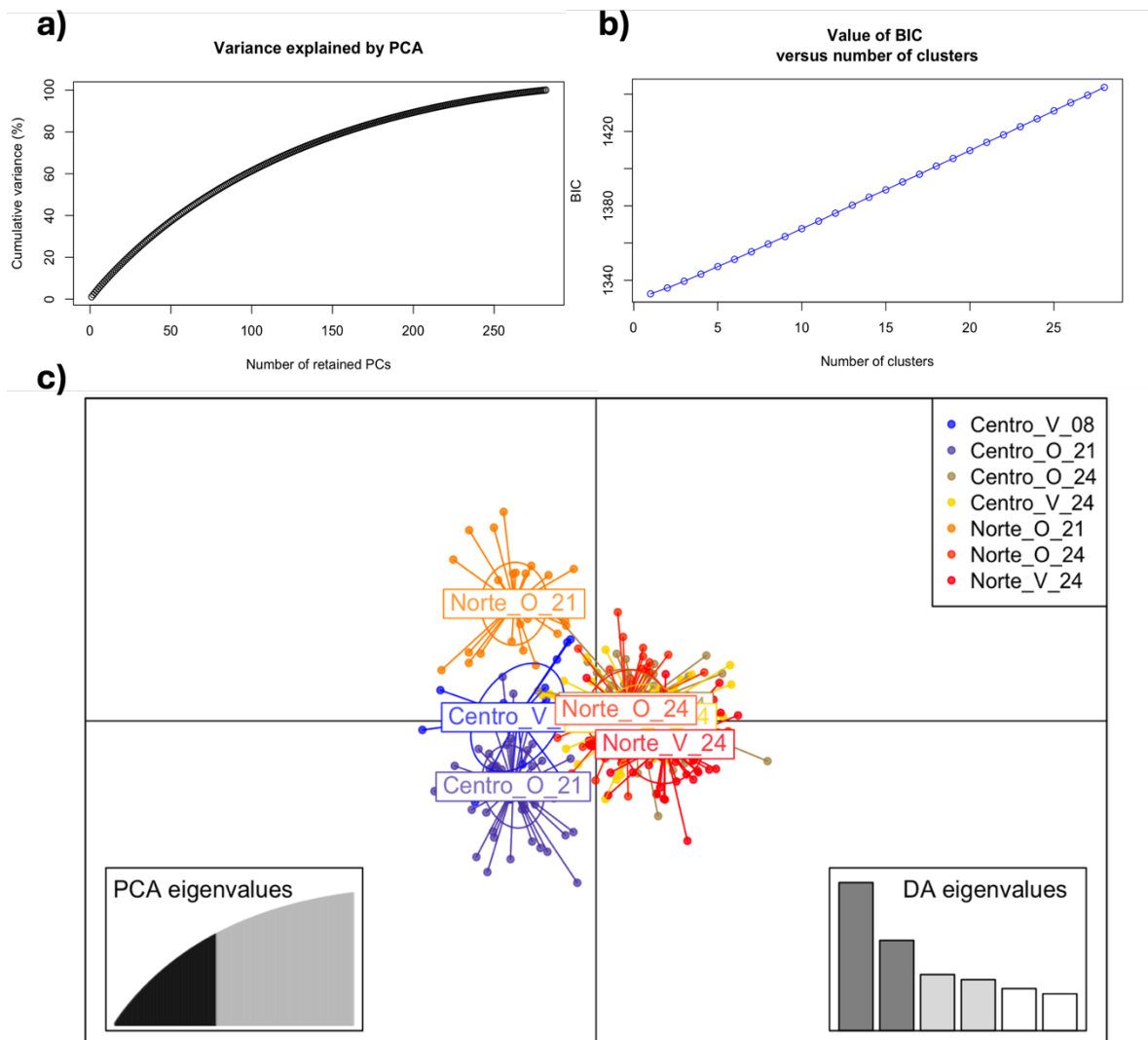


Figura 9. Análisis Discriminante de Componentes Principales realizado a partir del panel total de 1.313 SNPs.

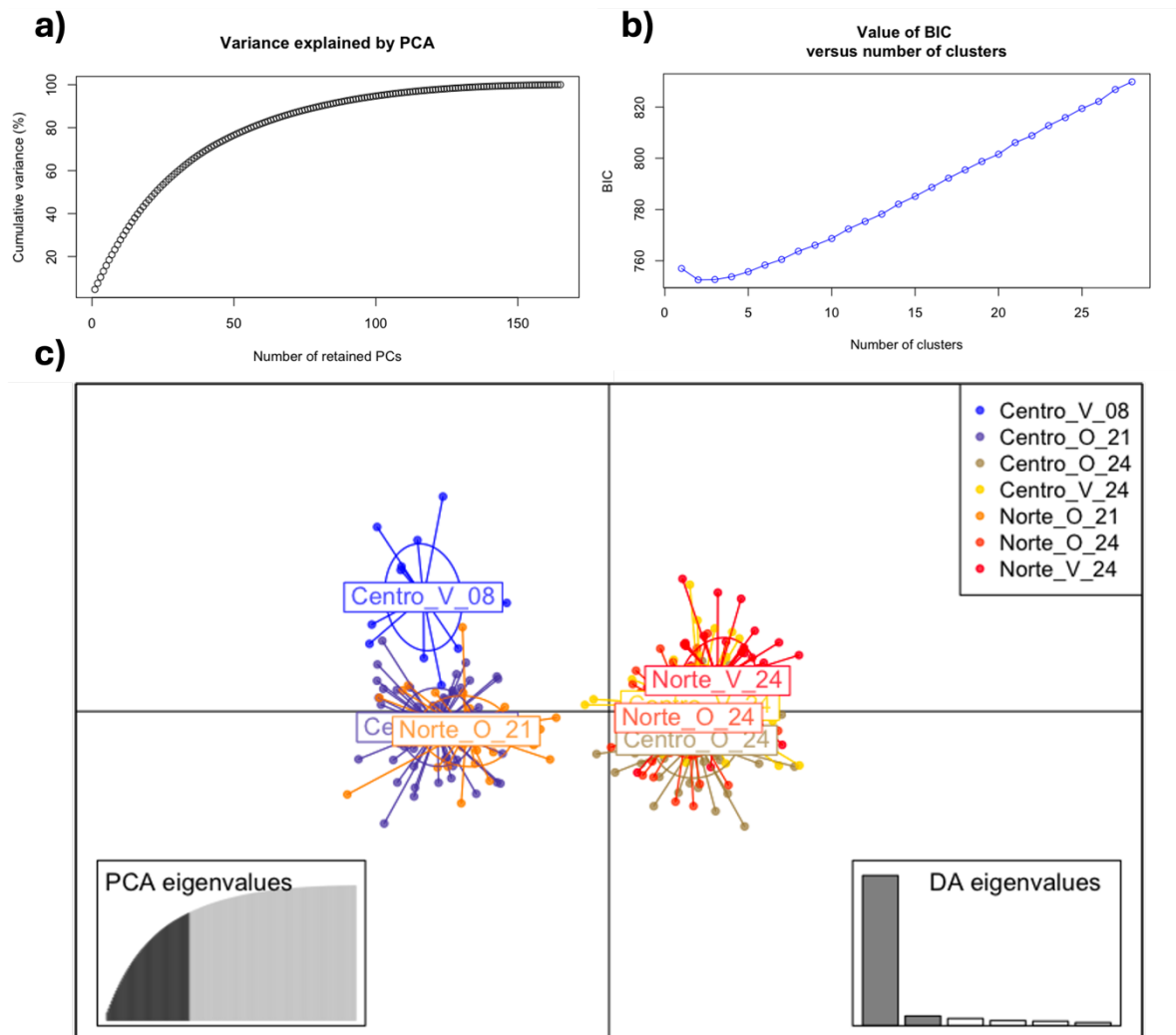


Figura 10. Análisis Discriminante de Componentes Principales realizado a partir del panel putativamente bajo selección de 165 SNPs.

4.5.3 Análisis de Admixture (Ancestría)

Realizada la estimación de ancestría con Admixture, la prueba de error de validación cruzada ("cross-validation error") sugiere que el número óptimo de grupos genéticos fue $K=1$ (CV-error = 0,3538) (**Anexo 18**). El análisis de ancestría realizado con 1.313 SNPs no evidencia la presencia de más de un grupo ancestral (**Figura 11**), al igual que los resultados obtenidos con los 1.148 marcadores neutrales (**Anexo 19**) y 165 SNPs adaptativos (**Anexo 20**).

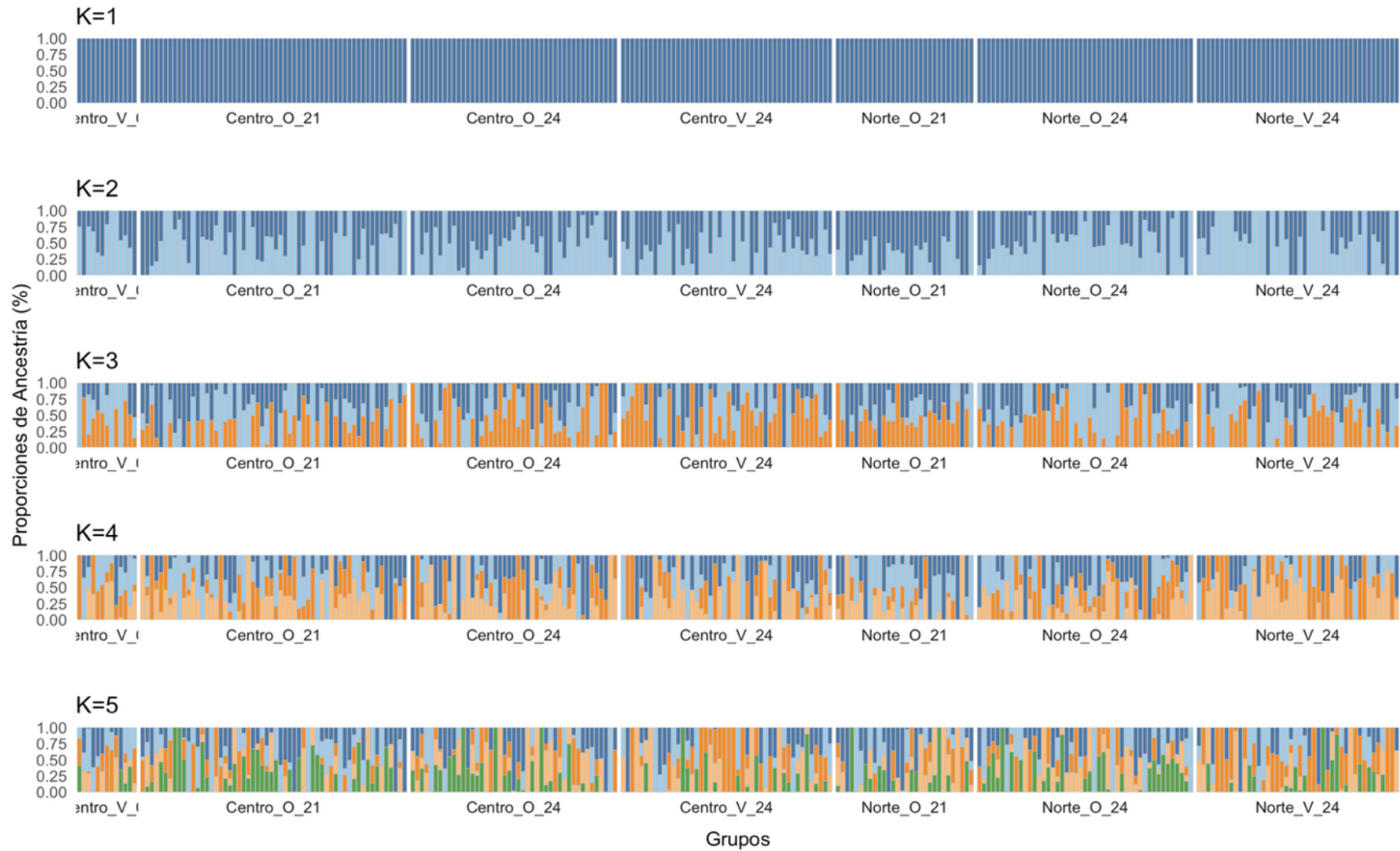


Figura 11. Gráfico de Admixture obtenido con el panel total de 1.313 SNPs.

4.5.4 Tamaño Efectivo de la población y Parentesco genético

El tamaño efectivo de la población presentó valores que tendieron en general al infinito. Los menores valores fueron estimados en el grupo Centro_V_24 ($N_e=1.840$). Todos los intervalos de confianza estimados mediante JackKnife presentaron valores inferiores a los 2.000 individuos y superiores tendientes al infinito. Los valores estimados por grupo, junto con los intervalos de confianza de cada estimación, se presentan en la **Tabla 14**.

El parentesco genético (R_{xy}) (**Figura 12**) entre pares de individuos dentro de los siete grupos estudiados presenta valores que sugieren un bajo parentesco al interior de los grupos analizados ($R_{xy} \sim 0$). Las estimaciones entre pares con valores máximos fueron encontradas generalmente entre individuos de un mismo grupo de muestreo. El valor máximo estimado fue de $R_{xy}=0,2241$. Considerando las muestras de todos los grupos en conjunto, fue posible identificar pares de individuos con valores de parentesco entre 0,25 y 0,125, sugiriendo relaciones de abuelos-nietos y primos hermanos (**Anexo 21**).

Tabla 14. Estimadores del Tamaño efectivo de la población.

	CentroV_08	Centro_O_21	Centro_O_24	Centro_V_24	Norte_O_21	Norte_O_24	Norte_V_24
MAF	0,05	0,05	0,05	0,05	0,05	0,05	0,05
Media harmónica de n	10,8	50,2	37,2	37,9	26,2	39,4	37,1
Nº comparaciones	160.793	172.713	165.802	177.764	199.962	182.741	184.136
r ² Global	0,1160	0,0206	0,0281	0,0279	0,0417	0,0266	0,0288
r ² Esperado	0,1171	0,0207	0,0281	0,0277	0,0417	0,0265	0,0285
Ne estimado	∞	∞	∞	1.840,8	∞	3.463,7	1.297,4
IC inf.Paramétrico	∞	9.054,9	2.474,7	910,7	1.281,4	1.235,1	750,1
IC sup. Paramétrico	∞	∞	∞	∞	∞	∞	4.648,5
IC inf. JackKnife	175,9	1.283,1	1.219,7	413,7	498,2	647,3	404,1
IC sup. JackKnife	∞	∞	∞	∞	∞	∞	∞

MAF= Frecuencia del alelo menor; Ne = Tamaño efectivo poblacional; IC = intervalo de confianza del 95%.

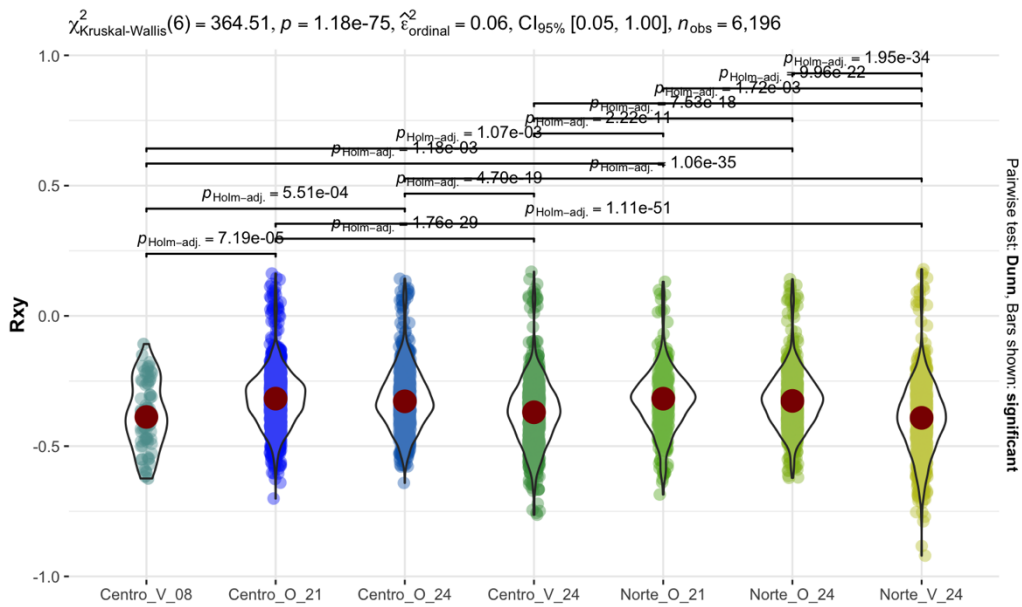


Figura 12. Parentesco genético (Rxy) entre pares de individuos al interior de cada grupo estudiado. Las barras en la parte superior indican las diferencias significativas entre los grupos de muestreo estudiados.

4.6 OE3. Conocer la asociación entre la variabilidad ambiental y las diferencias genéticas temporales utilizando marcadores altamente variables en las poblaciones de *Dosidicus gigas* en el Océano Pacífico Sur Este.

4.6.1 Estructura espacio-temporal y asociación con variables ambientales de *Dosidicus gigas*

Para estudiar la contribución de la distribución espacial a la estructura poblacional, se realizó un análisis de redundancia (RDA) utilizando una matriz de las coordenadas de distancia geográfica transformada como variable explicativa de la variación genética global. De los cuatro vectores geográficos utilizados, ninguno fue seleccionado para el modelo de regresión óptimo, por lo que la variación espacial no se mantuvo como variable predictora de la variación genética en el modelo de regresión final.

Para el estudio de asociación, se descargó un total de 12 variables ambientales, considerando entre ellas la salinidad (ppm), la temperatura (°C), el pH, la velocidad (m/s) y dirección de la corriente superficial del agua, la concentración de clorofila, la concentración de fitoplancton, la concentración de oxígeno disuelto y las concentraciones de silicato, nitrato, fosfato y hierro disuelto. Para cada variable se incluyen medidas de tendencia central (i.e. mínimo, máximo, media y rango) a nivel de superficie, profundidad mínima, profundidad media y profundidad máxima, sumando un total de 288 variables ambientales a incluir en los análisis.

Como primer paso para la selección de las variables ambientales que podrían explicar la variación genética de *Dosidicus gigas* en la ZEE de Chile, se evaluó la colinealidad entre cada una con el fin de eliminar la redundancia resultado de la existencia de correlaciones entre ellas. En este caso, el análisis indica que para cada variable ambiental todas las medidas de tendencia central se encuentran altamente correlacionadas ($r > 0,701$), por lo que mediante la función ordistep se seleccionó la medida de tendencia central más representativa de cada variable ambiental. Posteriormente, se realizó un análisis de redundancia para cada una de las variables con el fin de evaluar su contribución individual a la variación genética. La **Tabla 15** muestra la significancia estadística del análisis de redundancia para las variables ambientales preseleccionadas.

Considerando el resultado de la evaluación de colinealidad entre las variables (**Anexo 22**), se seleccionaron las variables “rango de concentración de fitoplancton a profundidad mínima” y “concentración de clorofila mínima a nivel superficial” para estimar el modelo de regresión final. El análisis de redundancia resultó significativo ($P=0,005$, $R^2_{adj}=0,2209$) utilizando las dos variables (concentración de fitoplancton y concentración de clorofila), sin embargo, sólo la concentración de fitoplancton contribuyó por sí sola de manera significativa ($p=0,002$) a la variación genética (concentración de clorofila, $p=0,064$).

Para el modelo final, se retuvo el rango de concentración de fitoplancton como la variable ambiental explicativa de la variación genética. El análisis de redundancia resultó significativo ($p=0,003$) explicando un 8,39% de la variación genética total, es decir, la variación en la concentración de fitoplancton explica un 8,39% de la variación genética de *D. gigas* en las Macrozonas Norte y Centro del país (**Figura 13**).

La variación en la concentración de fitoplancton fue mayor para los grupos Norte_O_21, Centro_O_21 y Centro_O_08, lo que podría explicar la tendencia de diferenciación temporal observada en el análisis de diferenciación y estructura genética realizado con el panel de 165 SNP putativamente bajo selección. Al corregir este modelo por la variación espacial de *D. gigas* en la costa de Chile, no fue significativo ($p=0,08$), por lo que la variación genética no está correlacionada a la distribución de la especie en Chile.

Tabla 15. Detalles de resultados del RDA y RDA parcial para cada variable ambiental estudiada.

SNPs	Análisis	Variables explicativas		P model	R^2_{adj}
		Espacial	Ambiental		
1.313	RDA	dbMEM-4		0.143	0.0315
	RDA		Salinidad	0.066	0.2028
	RDA		Clorofila	0.046*	0.1926
	RDA		Oxígeno disuelto	0.140	0.0670
	RDA		Temperatura	0.078	0.0762
	RDA		Silicato	0.065	0.1462
	RDA		Hierro disuelto	0.088	0.1544
	RDA		Nitrato	0.076	0.1198
	RDA		Fosfato	0.031*	0.1658
	RDA		pH	0.004**	0.1965
	RDA		Fitoplancton	0.003**	0.0832
	RDA		Dirección del agua	0.088	0.1631
	RDA		Velocidad del agua	0.016*	0.1760
	RDA		Fitoplancton + Clorofila	0.005**	0.2209
	RDA		Rango de concentración Fitoplancton	0.003**	0.0839

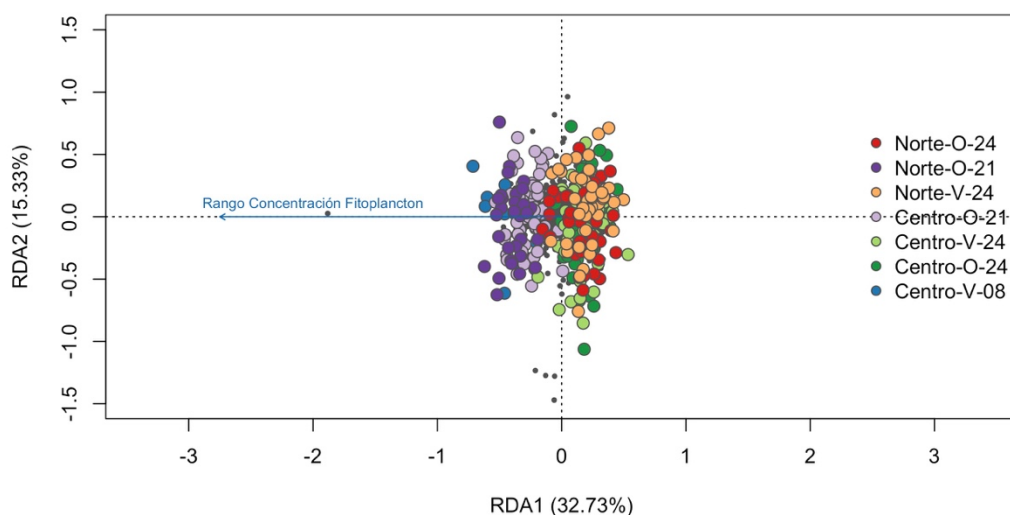


Figura 13. Análisis de Redundancia de la variación genética utilizando un total de 1.313 SNPs y la variable ambiental “rango de concentración de fitoplancton a profundidad mínima” como variable explicativa de la variación genética de *D. gigas*. (RDA model $p=0.003$, $R^2_{adj}= 0.0839$).

4.6.2 Identificación de loci candidatos de adaptación local en *Dosidicus gigas*

Para la identificación de loci candidatos de adaptación local, se utilizó como variable explicativa el rango de concentración de fitoplancton a profundidad mínima. El análisis de redundancia (RDA) identificó 28 SNPs candidatos de adaptación local, correlacionados con la concentración de fitoplancton. Sin embargo, todos los marcadores SNPs identificados presentaron una baja correlación con el predictor ambiental (**Tabla 16**). Por su parte, el método LFMM identificó 1 SNP candidato de adaptación local, el que también fue identificado por el RDA. En consecuencia, entre ambos métodos se identificó un total de 28 SNPs candidatos de adaptación local, asociados a la variación de fitoplancton. La **Figura 14** muestra todos los loci identificados en el presente estudio como putativamente adaptativos o candidatos de adaptación local, en esta es posible observar que 10 SNPs fueron identificados como loci putativamente bajo selección o candidatos de adaptación local a la variación en los niveles de nutrientes (fitoplancton) a lo largo del rango de muestreo.

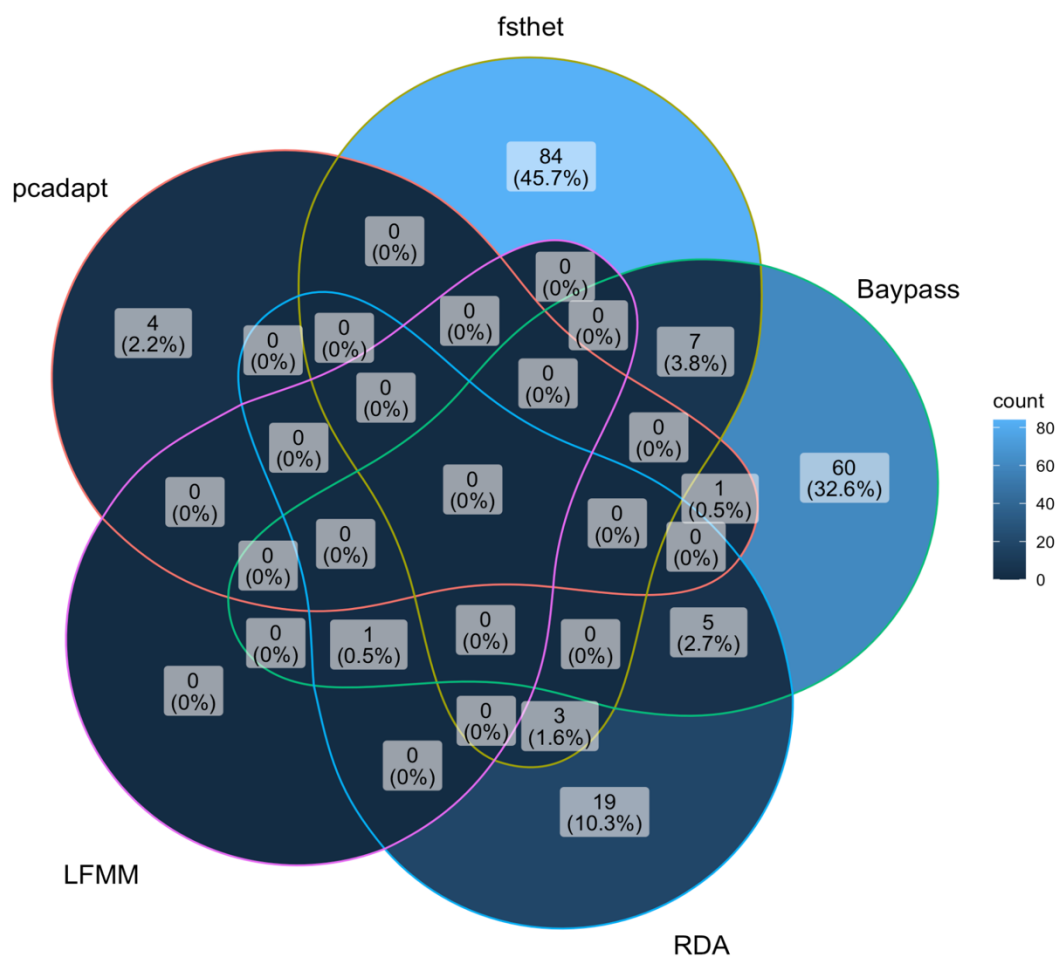


Figura 14. Diagrama de Venn de los loci Outlier identificados por las cinco aproximaciones utilizadas.

Por último, se realizó una búsqueda de BLAST contra la base de datos nucleotídica de NCBI⁴ para identificar potenciales regiones de interés entre los 28 loci candidatos de adaptación local. De este análisis, sólo para 6 de los 28 SNPs se identificó una secuencia homóloga en la base de datos de NCBI, donde fue posible identificar regiones genómicas similares en otras especies de cefalópodos como *Thysanoteuthis rhombus* y *Sepia esculenta* (**Anexo 23**).

⁴ <https://blast.ncbi.nlm.nih.gov/Blast.cgi>

Tabla 16. Correlación de los SNPs candidatos a adaptación local con las variables ambientales predictoras seleccionadas para el modelo óptimo.

N	Eje RDA	SNP	Predictor	Correlación
1	1	100572056-15	Rango concentración de fitoplancton superficial	0.15
2	1	100660738-17	Rango concentración de fitoplancton superficial	-0.12
3	1	100720006-14	Rango concentración de fitoplancton superficial	0.12
4	1	100724716-6	Rango concentración de fitoplancton superficial	0.16
5	1	100739042-15	Rango concentración de fitoplancton superficial	0.11
6	1	100746073-18	Rango concentración de fitoplancton superficial	0.16
7	1	100775969-22	Rango concentración de fitoplancton superficial	-0.16
8	1	100776623-67	Rango concentración de fitoplancton superficial	0.98
9	1	100781183-31	Rango concentración de fitoplancton superficial	0.13
10	1	100859258-30	Rango concentración de fitoplancton superficial	0.12
11	1	100876496-35	Rango concentración de fitoplancton superficial	0.14
12	1	100901302-14	Rango concentración de fitoplancton superficial	0.15
14	2	100455301-64	Rango concentración de fitoplancton superficial	-0.02
15	2	100536765-16	Rango concentración de fitoplancton superficial	0.07
16	2	100542664-48	Rango concentración de fitoplancton superficial	0.08
17	2	100629667-20	Rango concentración de fitoplancton superficial	0.02
18	2	100706494-23	Rango concentración de fitoplancton superficial	0.03
19	2	100790981-29	Rango concentración de fitoplancton superficial	-0.01
20	2	100807518-30	Rango concentración de fitoplancton superficial	0.02
21	2	100807519-32	Rango concentración de fitoplancton superficial	0.05
22	2	100819868-5	Rango concentración de fitoplancton superficial	0.07
23	2	100835806-55	Rango concentración de fitoplancton superficial	-0.01
24	2	100884900-6	Rango concentración de fitoplancton superficial	-0.09
25	2	100900368-37	Rango concentración de fitoplancton superficial	0.02
26	2	100918887-39	Rango concentración de fitoplancton superficial	-0.01
27	2	100922519-25	Rango concentración de fitoplancton superficial	-0.01
28	2	100940047-26	Rango concentración de fitoplancton superficial	0.01

4.7 OE4. Asesorar a la Subsecretaría de Pesca y Acuicultura en el contexto de la OROP-PS relacionadas con la genética de *Dosidicus gigas*.

4.7.1 Asesorías y Acuerdos OROP-PS

Durante el desarrollo del proyecto, el equipo de investigación entregó periódicamente asesoría técnica a la SUBPESCA en aspectos biológicos y genéticos de la jibia *D. gigas*. Este apoyo fue realizado durante las reuniones periódicas del proyecto (**Anexo 24, 25, 26 y 27**), así como también en las ejecutadas en virtud de la reunión del comité científico (SC-12) de la OROP-PS, realizado en Lima entre el 30 de septiembre y el 6 de octubre de 2024.

En concreto, el equipo técnico mantuvo un canal de comunicación permanente con la contraparte técnica de la SUBPESCA, con el fin de coordinar su participación en las reuniones de preparación previas al SC-12, desarrolladas de forma virtual. Estas reuniones se centraron en asegurar una participación informada de la delegación chilena, respecto a la genética poblacional de la jibia.

Consecuentemente, durante estas reuniones se presentó un re-análisis de los datos genómicos obtenidos en el proyecto FIPA2020-20, estimando nuevamente la diversidad genética y estructura poblacional de la especie, incluyendo métodos para la identificación de huellas de adaptación y el análisis genético diferenciado entre marcadores neutrales y putativamente bajo selección. En ese sentido, se apoyó a la delegación nacional en la interpretación, alcances y limitaciones de las estimaciones de diferenciación y análisis de agrupamiento genético.

Por otro lado, estas instancias preparatorias también permitieron analizar el informe técnico “SNPs for population genomics analysis of *D. gigas*” (SC12-SQ04-PER), presentado por la delegación de Perú para ser discutido durante el SC-12. Este informe presenta resultados de estructuración genético-poblacional en muestras obtenidas desde las pesquerías de la ZEE de Perú mediante el uso marcadores SNPs obtenidos por el método de ddRAD. De su análisis colectivo surgieron preguntas de rigor metodológico que fueron enviadas a la delegación de Perú través de la SUBPESCA (**Anexo 28**).

Durante la reunión de la OROP-PS (SC-12), cuatro miembros del equipo de investigación formaron parte de la delegación oficial de Chile. La Dra. María Angélica Larraín Barth y el Dr. Cristian Araneda Tolosa, asistieron presencialmente a la reunión y participaron de las instancias de discusión de la comisión de genética de *D. gigas*. Mientras que de forma telemática participaron la Dra. Valentina Córdova Alarcón y el Dr. Carlos I. Vargas Manríquez. Durante la instancia se discutió sobre la necesidad de investigar colectivamente la diversidad genética de la especie en toda su distribución geográfica, junto con consensuar los criterios técnicos asociados a la toma de muestras y tecnologías de secuenciación y análisis de datos a emplear. En ese sentido, se alcanzó un acuerdo entre los representantes de las delegaciones de Perú, Ecuador, China y Chile (puntos 217 y 218 del SC12 - Final Report) (**Anexo 29**) que incluye:

- i. La definición de criterios de toma de muestras espacial y temporal: En las áreas E y O (**Figura 15**) se obtendrán hembras en estado de madurez III y IV para cada fenotipo de tamaño por área (n=30) desde 2024 a diciembre de 2025.

- ii. La definición de grupos fenotípicos para la toma de muestras por delegación: Perú en áreas E1 (S+M) y E2 (S+M+L); China en áreas O1 (S+M), O2 (S+M+L), O3 (M+L), y O4 (L); Chile en áreas E3 (L) y E4 (L); y Ecuador en área E1 (S+M).
- iii. La construcción de un formato tipo conjunto (templado) para la obtención de datos biológicos, incluyendo sexo, largo total, estado de madurez gonadal, coordenadas de captura, fecha y lugar.
- iv. El envío de muestras de tejido, o ADN, a un único servicio biotecnológico, previamente acordado (ej. Novogene), para la secuenciación de genoma completo de baja cobertura (low-coverage Whole Genome Sequencing, lcWGS) de las muestras, compartiendo los resultados de con toda la comisión de genética de *D. gigas* de la OROP-PS.
- v. La definición de protocolos para la toma de muestras de tejido, secuenciación de ADN, líneas de trabajo bioinformático para el filtrado de datos y análisis genético-poblacionales, durante el desarrollo de reuniones virtuales durante el año 2025.

Para dar respuesta a los acuerdos alcanzados en la SC-12, se elaboró un protocolo de toma de muestras de tejido de *D. gigas* para la secuenciación de baja cobertura (**Anexo 8**). Se recomienda obtener una muestra de gónada de cada individuo almacenada en buffer Bouin para realizar histología del tejido gonadal con el objetivo de identificar el estado reproductivo de cada individuo. Respecto al muestreo a realizar durante 2025 para la secuenciación de baja cobertura acordada en la SC-12, se obtendrán muestras de la temporada de Invierno de 2025 de la Macrozona Norte y Macrozona Centro de Chile, siguiendo el protocolo descrito en el **Anexo 8**. Estas muestras de tejido serán enviadas a una facilidad de secuenciación que será seleccionada en conjunto con los miembros de la OROP-PS para su secuenciación lc-WGS.

Cabe destacar, que para dar cumplimiento a los objetivos a largo plazo del grupo de Genética de *D. gigas* de la OROP-PS, las muestras de tejido de la ZEE de Chile obtenidas en el presente proyecto quedarán almacenadas en el Laboratorio de Genética y Biotecnología en Acuicultura, del Departamento de Producción Animal de la Facultad de Ciencias Agronómicas de la Universidad de Chile y estarán a disposición de futuras investigaciones. Contacto: Dr. Cristian Araneda Tolosa, [<craraned@uchile.cl>](mailto:craraned@uchile.cl), Profesor Asociado, Universidad de Chile.

Adicionalmente, con los resultados obtenidos en el presente informe, se están preparando dos Working papers para presentar en la reunión del comité científico de la OROP-PS (SC-13), a realizarse entre el 8 y 13 de septiembre de 2025 en la ciudad de Wellington, Nueva Zelanda. Los títulos de los Working papers son:

- i. Whole-genome sequence of the jumbo flying squid *Dosidicus gigas* as a tool for low coverage genome sequencing (lcWGS).
- ii. Genetic diversity and population structure of jumbo flying squid *Dosidicus gigas* estimated with genomic SNPs in the Chilean coast.

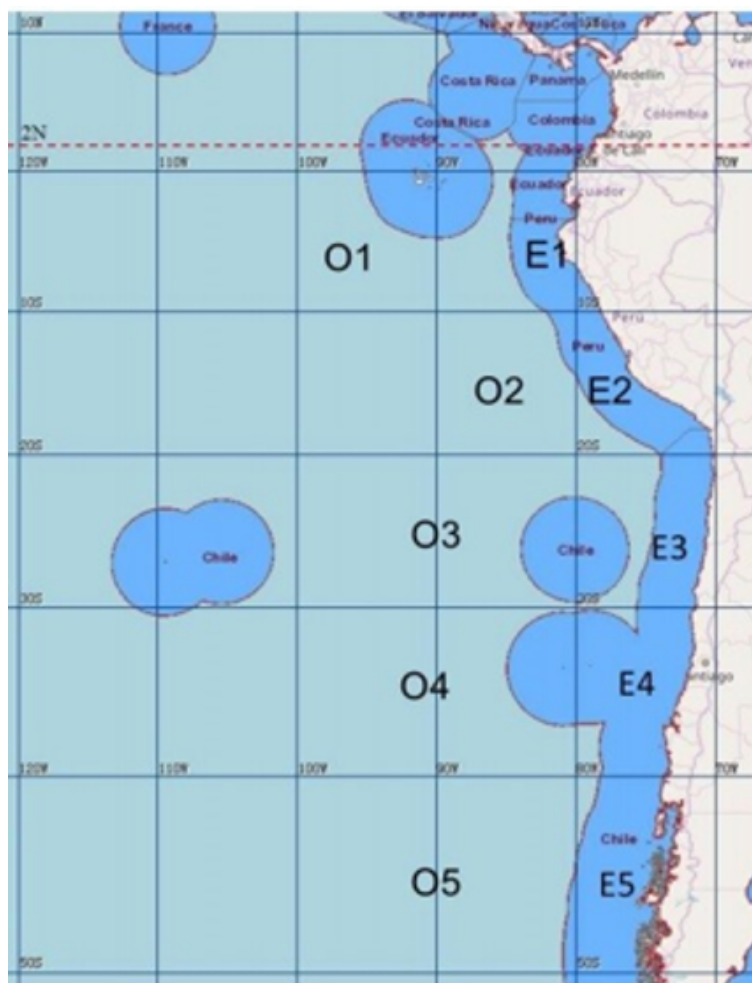


Figura 15. Mapa de zonas de muestreo para el estudio poblacional de *D. gigas* acordado en la OROP-PS SC-12 Lima 2024.

4.7.2 Difusión de resultados

Las actividades de difusión de resultados del proyecto se detallan a continuación:

Taller de difusión de resultados y de capacitación: Esta actividad comprometida en la ejecución del proyecto, estuvo dirigida a funcionarios del FIPA. Y de SUBPESCA. El taller fue programado y ejecutado el viernes 30 de mayo del 2024, en dependencias de SUBPESCA. Tuvo como propósito presentar los resultados del presente proyecto, así como también fortalecer las capacidades institucionales para el análisis e interpretación de resultados genómicos para la gestión del recurso *D. gigas*, contribuyendo así a la toma de decisiones basada en evidencia científica. (Ver video del Taller [AQUI](#) e invitación y foto asistentes en **Anexo 30**).

Las presentaciones del taller correspondieron a: **1.** Secuenciación del genoma completo del calamar de Humboldt (*Dosidicus gigas*), expositor Dr. Cristian Araneda T (**Anexo 31**). **2.** Marcadores Moleculares utilizados en el Estudio del Calamar de Humboldt (*Dosidicus gigas*), expositor Dr. Cristian Araneda T (**Anexo 32**). **3.** Diversidad genética y estructura

poblacional de *D. gigas* estimada con SNPs genómicos, expositor Dr.(c) Carlos Vargas Manríquez (**Anexo 33**). **4.** Asociación entre variabilidad ambiental genética de *D. gigas* en la costa de Chile, expositora Dra. Valentina Córdova Alarcón (**Anexo 34**) y **5.** Asesoría a la SUBPESCA en el contexto de la OROP-PS relacionados con la genética de *D. gigas* (**Anexo 35**).

Simposio XLIV Congreso Ciencias del Mar: “Desarrollo de Herramientas Genómicas aplicadas al manejo de pesquerías transzonales en el Pacífico Sur en el contexto de la OROP-PS” realizado el martes 27 de mayo de 2025, en el marco del XLIV Congreso de Ciencias del Mar (26 al 30 de mayo 2025 en Hotel del Mar Enjoy, Viña del Mar, Valparaíso, Chile). En esta instancia se presentaron los resultados obtenidos en el Objetivo específico 1, 2 y 3 (**Anexo 36**).

5 Discusión

La preservación de la diversidad genética se ha consolidado como un eje central en las políticas internacionales orientadas al uso sostenible y la conservación de los recursos naturales. La diversidad genética es fundamental para la mantención de los equilibrios ecológicos, y se encuentra estrechamente asociada con la capacidad adaptativa de los organismos tanto a nivel individual como poblacional (Ellegren et al., 2016). En animales marinos su estudio ha proporcionado información de gran utilidad para conocer el estado de conservación de las especies, y para apoyar la toma de decisiones en la regulación, administración y manejo de la actividad pesquera, y además de ayudar al desarrollo de herramientas biotecnológicas para monitorear la sostenibilidad de la pesca y acuicultura.

El objetivo del presente estudio fue secuenciar, ensamblar y anotar el genoma nuclear completo de la jibia (*D. gigas*), y describir la diferenciación espacial y temporal de especie en la ZEE de Chile con marcadores genómicos. Los resultados discutidos más adelante corroboran la hipótesis de que la jibia es una especie que posee una alta capacidad migratoria, por lo que se comporta como una especie panmíctica en la ZEE de Chile.

5.1 Genoma completo de *D. gigas*

La construcción del genoma de referencia de *D. gigas*, lograda mediante la combinación de tecnologías de secuenciación Oxford Nanopore y PacBio y complementada con lecturas Hi-C, representa un hito importante en el estudio de la especie. El ensamblaje resultante, con una cobertura total superior a 100X y una calidad de secuencia compatible con “scaffolding” cromosómico Hi-C, permitió obtener un genoma completo de 3,6 Gb con resolución telómero-a-telómero (T2T) en 29 cromosomas, un contenido de GC del 33%, y métricas de contigüidad sobresalientes (N50 de 40 y 78,43 Mb para cóntigos y “scaffolds”, respectivamente). Comparado con otros genomas de calamares disponibles publicamente en NCBI (National Center Biothecnology Information), este ensamblaje presenta la mayor resolución cromosómica y nivel de completitud de un cefalópodo, constituyendo un recurso genómico de referencia para la familia Ommastrephidae. Su calidad es comparable, e incluso superior en términos de contigüidad, a los estándares establecidos por el Vertebrate Genome Project (VGP-2016) (The International Genome 10K Consortium), lo que refuerza la solidez del trabajo. La anotación preliminar, basada en transcriptomas de especies cercanas, reveló una representación génica robusta (40.072 genes). La alta tasa de alineación (98-99%) de las secuencias de 15 individuos adicionales valida aún más la precisión y aplicabilidad del genoma como base para estudios poblacionales y de diversidad genética a lo largo del rango de distribución de *D. gigas*.

Para potenciar la utilidad como referencia de este primer genoma obtenido, se recomienda la incorporación de datos transcriptómicos generados directamente desde tejidos de *D. gigas*, en lugar de depender de especies cercanas como *Todarodes pacificus* y *Sthenoteuthis oualaniensis*, lo que permitiría una anotación más precisa y completa de los genes específicos de la especie. Asimismo, la inclusión de datos epigenómicos de sitios de metilación y multi-ómicos podrían aportar una nueva dimensión para comprender la

respuesta adaptativa mediante la regulación génica de esta especie en diferentes contextos, incluido el cambio climático.

Contar con un genoma completo y de alta resolución ofrece herramientas concretas para fortalecer la gestión y administración pesquera de especies transzonales como *D. gigas*. Esta información permite el desarrollo de marcadores moleculares altamente específicos para el monitoreo genético de las poblaciones explotadas, facilitando la identificación de stocks, el seguimiento de patrones migratorios y la detección temprana de posibles cuellos de botella demográficos. A su vez, se puede alimentar modelos de evaluación de stock para mejorar la precisión, incorporar información sobre conectividad poblacional en tiempo real, y permitir la toma de decisiones más dinámicas y basadas en evidencia genética, en línea con el enfoque ecosistémico de la pesca.

5.2 Diversidad genética y estructura poblacional de *D. gigas* en la costa de Chile

Los resultados obtenidos con el total de marcadores (1.313 SNPs), sugieren que, debido a la capacidad migratoria y las variables oceanográficas y ambientales existentes a lo largo de la distribución de *D. gigas* en Chile, no existe una limitación en el flujo genético de esta especie en la ZEE de Chile. La diversidad genética promedio estimada en siete grupos (600 km de costa) fue de un 8,70% y no presentó estructuración genética, corroborado a partir de los resultados de F_{ST} , PCA, DAPC y agrupamiento Bayesiano (**Tabla 8, Figura 3, Figura 6 y Figura 9**). Estos resultados son concordantes con lo previamente descrito para la especie utilizando marcadores moleculares de DNA mitocondrial (Ibáñez et al., 2016b) y microsatélites (Sanchez et al., 2020b, 2016b), donde se identifica una única unidad poblacional en el territorio nacional. La ausencia de estructura poblacional es esperable para especies con las características de historia de vida de *D. gigas*, como alta fecundidad, alta capacidad de migración, grandes tamaños poblacionales, y ciclos de vida cortos (Ibáñez et al., 2011; Markaida, 2001; Tafur et al., 2010).

Por otra parte, el análisis de los marcadores identificados como adaptativos (165 SNPs), evidencia diferenciación poblacional temporal dentro de la ZEE de Chile, donde los individuos muestreados en 2008 y 2021 (FIPA20-20) se diferencian genéticamente (F_{ST}) entre un 1,4% a un 3% de los individuos muestreados en 2024 (FIPA23-19), independientemente de su ubicación espacial. Esta diferenciación genético-temporal es corroborada por los análisis de PCA y DAPC, donde se evidencian dos grupos genéticos en la ZEE de Chile. En contraste, el análisis de Ancestría (Admixture) evidencia un único grupo genético, lo cual es explicado por la alta tasa de migración que existe en la especie y a la mezcla generacional que existe. Esto último se ve reflejado también en las relaciones de parentesco, las cuales son bajas para todas las comparaciones realizadas, sin embargo, es posible encontrar individuos con un parentesco que sugiere relaciones de abuelos-nietos y primos hermanos (R_{xy} = 0,25; 0,125). A su vez, la presencia de generaciones superpuestas puede tener efectos en la estimación del tamaño efectivo poblacional.

Los estudios genéticos realizados en *D. gigas* a lo largo de la costa del Océano Pacífico, utilizando diferentes marcadores moleculares, han establecido la existencia de dos grupos

genéticos, uno en el Hemisferio Norte (México y EE.UU) y otro en el Hemisferio Sur (Perú-Chile) (Sandoval-Castellanos et al., 2010, 2007; D. Staaf et al., 2010). En particular, en el Hemisferio Sur se ha estudiado la estructura poblacional utilizando DNA mitocondrial (Ibáñez et al., 2011; Pardo - Gandarillas et al., 2025; Sanchez et al., 2020b, 2016b), marcadores microsatélites (Sanchez et al., 2020b, 2016b), y marcadores RAPD (Randomly amplified polymorphic DNA) (Sandoval-Castellanos et al., 2007), y todos los autores coinciden en la identificación de un único grupo genético en las costas de Chile, Perú y Ecuador (Ibáñez et al., 2011; Pardo - Gandarillas et al., 2025; Sandoval-Castellanos et al., 2007). Esperamos que el acuerdo alcanzado en la SC12 de la OROP- PS entre el grupo de genética de la Jibia (**Anexo 29**) para el desarrollo de un estudio genético con marcadores SNPs obtenidos con una metodología común (secuenciación de baja cobertura, lcWGS) desde muestras comprometidas por las delegaciones de Ecuador, Perú, Chile y China, permita resolver de forma conjunta la existencia de una o varias unidades poblacionales de *D. gigas* en el Pacífico Sur Este.

Respecto a la diversidad y estructuración genética de recursos de importancia socioeconómica, en las costas de Chile los resultados son contrastantes, existiendo especies que se comportan como una población panmíctica, otras con patrones claros de estructuración y otras donde inicialmente con el uso de DNA mitocondrial la estructuración fue rechazada, para luego ser postulada con el uso de miles de marcadores SNP. La existencia de una única población biológica, utilizando marcadores SNPs se ha reportado en otras especies con alta capacidad de dispersión del Pacífico Sur, como por ejemplo el jurel (*Trachurus murphyi*) (Ferrada Fuentes et al., 2023), la jaiba marmola (*Metacarcinus edwardsii*) (Veliz et al., 2022) y el erizo de espinas largas del Archipiélago de Juan Fernández (*Centrostephanus sylviae*) (Veliz et al., 2021). Adicionalmente esta situación también ha sido reportada con marcadores microsatélites para el bacalao de profundidad (*Dissodicus eleginoides*) (Canales-aguirre and Ferrada-fuentes, 2018) y el congrio dorado (*Genypterus blacodes*) en el Sur de Chile (Canales-Aguirre et al., 2010).

5.3 Asociación ambiental con la variabilidad genética temporal de *D. gigas*

Los resultados obtenidos en relación con la variación ambiental evidencian que la concentración de fitoplancton a profundidad mínima es un factor ambiental significativo (8,3%) en la estructuración genético-temporal de *D. gigas* en las macrozonas Norte y Centro de la ZEE de Chile. Este hallazgo sugiere que las fluctuaciones en la disponibilidad de fitoplancton, probablemente vinculadas a procesos oceanográficos estacionales como los eventos de surgencia que provocan variaciones en la productividad primaria, podrían estar influyendo en la diferenciación temporal observada en los grupos muestreados en 2008 y 2021 en comparación a los grupos muestreados en 2024. Esto, además, podría estar influenciando tanto el reclutamiento como la productividad de la población de *D. gigas* en Chile.

Sin embargo, la ausencia de significancia al corregir por la distribución espacial indica que esta variación genética está asociada a dinámicas ambientales locales y temporales por sobre la distribución geográfica de la especie. Adicionalmente, la identificación de 28 SNPs candidatos de adaptación local asociados al fitoplancton refuerza la hipótesis de que esta variable actúa como un factor selectivo en *D. gigas*. La baja correlación individual de estos marcadores con la variación ambiental sugiere un efecto poligénico. La homología encontrada para alguno de estos loci con especies como *Thysanoteuthis rhombus* y *Sepia esculenta* apunta a mecanismos posiblemente conservados en cefalópodos frente a cambios en la disponibilidad de nutrientes.

En comparación con el ambiente terrestre, el ambiente marino es considerado demográficamente abierto, debido a la falta de barreras físicas que limiten el flujo genético y la alta capacidad de dispersión que poseen los organismos que lo habitan (Olsen et al., 2020). En ese sentido, se ha descrito que las características que poseen un mayor impacto sobre la estructuración genética en especies marinas son el tipo de desarrollo (e.g. directo o indirecto, etapas y duración del ciclo larval), y la capacidad de dispersión activa en la vida adulta (e.g. sésil, sedentaria, altamente migratoria) (Camus, 2001; Lawrence and Fraser, 2020; Palumbi, 2003; Romiguier et al., 2014). Por lo tanto, el conocer los procesos por los cuales es promovida o limitada la conectividad es una pregunta fundamental en ecología y para el manejo pesquero de especies de importancia comercial (Veliz et al., 2022).

5.4 Recomendaciones

Considerando que los resultados obtenidos indican que *D. gigas* constituye una única población panmítica en la ZEE de Chile, las medidas de administración pesquera, como cuotas de extracción o vedas biológicas, pueden ser aplicadas de forma general en las regiones donde se realiza el esfuerzo pesquero.

Se recomienda enriquecer con datos transcriptómicos el genoma de referencia obtenido en este proyecto, para una anotación precisa y completa de los genes específicos de la especie, incluyendo también datos epigenómicos, que aporten información para comprender la respuesta adaptativa mediante la regulación génica y su potencial impacto sobre la productividad de las poblaciones de *D. gigas* a nivel regional y global.

Se sugiere la implementación de un programa de monitoreo periódico (i.e. cada dos años) de la diversidad genética de *D. gigas*, mediante el uso de un panel de marcadores SNPs estándar, con el objetivo de conocer el efecto de la variación ambiental y el cambio climático sobre la diversidad genética en el tiempo. En este sentido, se recomienda desarrollar un panel de SNPs que permita estandarizar las herramientas e indicadores de monitoreo para su implementación en el marco de la OROP-PS, más allá de la ZEE de Chile.

En el mismo sentido, considerando el ciclo de vida y la distribución de la especie en el Océano Pacífico, es fundamental aumentar el rango geográfico de estudio, analizando muestras obtenidas a lo largo de toda su distribución natural con una metodología común, por lo que el seguimiento de los acuerdos realizados durante el SC-12 de la OROP-PS se

torna fundamental para la aplicación conjunta de medidas de administración pesquera para la conservación internacional del recurso *D. gigas*.

6 Conclusiones

OE1: La construcción del genoma completo de referencia para *D. gigas* alcanzó un estándar de calidad comparable, e incluso superior, a los lineamientos establecidos por iniciativas internacionales como el Vertebrate Genome Project (VGP-2026). Se superaron los umbrales de calidad comprometidos y se cumplió con el objetivo de generar un recurso genómico robusto para estudios genéticos en el Océano Pacífico Sur Este. El genoma completo de *D. gigas* tiene un tamaño aproximado de 3,6 Gpb, presenta un alto grado de resolución cromosómica con 29 cromosomas ensamblados de telómero a telómero y 17 “scaffolds” no ensamblados a nivel de cromosoma, un N50 de 40 Mb para cóntigos y 78,43 Mb para “scaffolds”, y 40.072 genes identificados.

OE2: Los resultados obtenidos respecto de la variación genética global y neutral para estudiar la estructuración espacio-temporal de *D. gigas*, confirman lo reportado para la especie en la ZEE de Chile y refuerzan la hipótesis de la existencia de tan solo una población genética en el Océano Pacífico Sur. Esta falta de estructuración genética estaría determinada por la historia de vida de la especie, marcada principalmente por una alta capacidad migratoria a lo largo del sistema de Corrientes de Humboldt.

La variación genética adaptativa evidencia una diferenciación temporal del 1% entre las muestras obtenidas los años 2008 y 2021 (FIPA20-20) con las obtenidas el año 2024 (FIPA23-19). Estos leves cambios en las frecuencias alélicas a lo largo del tiempo probablemente responden a la deriva genética o a la adaptación a cambios en las variables ambientales y oceanográficas observada a través de los años.

OE3: Los resultados de asociación ambiental estiman una asociación significativa, pero con baja correlación, entre 28 marcadores SNPs y la variable concentración de fitoplancton. Esta asociación podría ser producto de la variación espacial y estacional de la productividad primaria, lo que podría estar influenciando la disponibilidad de alimento, el reclutamiento y la productividad de la población de *D. gigas* en la ZEE de Chile.

OE4: El apoyo técnico desarrollado por el equipo de trabajo mejoró notablemente la posición de Chile en la discusión asociada a la genética de esta especie. Consecuentemente, se logró gestionar acuerdos entre los grupos de investigación que representan a los países participantes (i.e. Perú, Ecuador, China y Chile), por lo que se espera que este acuerdo permita realizar estudios genómicos a nivel internacional en el Pacífico Sur a través de la OROP-PS, aportando información imprescindible para la adopción de estrategias de administración para la conservación de *D. gigas* basadas en información científica.

Finalmente, el equipo de investigación ha sostenido una comunicación continua con la contraparte FIPA en distintas instancias, expresado su compromiso para la continuidad del trabajo en el marco de la OROP-PS y el objetivo de fortalecer la administración pesquera en el Pacífico Sur y alcanzar una explotación sostenible del recurso *D. gigas*.

7 Agradecimientos

Los autores del presente informe quisieran agradecer a todas y todos los colaboradores de este proyecto, quienes facilitaron la obtención de muestras y compartieron sus conocimientos sobre la especie y su pesquería. A los Pescadores Artesanales de Caleta Guayacán, Región de Coquimbo, especialmente a Don Guille; A los Pescadores Artesanales de la Caleta Pichidangui, Región de Coquimbo, especialmente a Chelo, Talca y Rumpy; A los Pescadores Artesanales de la Alianza Pesca Maule, Región del Maule, especialmente a Baltazar Encina, Erik Letelier y Alejandro Osses; A los Pescadores Artesanales de Puerto Maguillines, Región del Maule. También agradecemos al Instituto de Investigación Pesquera (INPESCA) por facilitarnos muestras, que no se incluyeron en estos resultados, pero servirán de respaldo para futuros análisis.

8 Referencias bibliográficas

- Aguirre-Liguori, J.A., Luna-Sánchez, J.A., Gasca-Pineda, J., Eguiarte, L.E., 2020. Evaluation of the Minimum Sampling Design for Population Genomic and Microsatellite Studies: An Analysis Based on Wild Maize. *Front Genet* 11. <https://doi.org/10.3389/fgene.2020.00870>
- Albertin, C.B., Simakov, O., Mitros, T., Wang, Z.Y., Pungor, J.R., Edsinger-Gonzales, E., Brenner, S., Ragsdale, C.W., Rokhsar, D.S., 2015. The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature* 524, 220–224. <https://doi.org/10.1038/NATURE14668>;TECHMETA=23,38,45,91;SUBJMETA=181,208,212,631,726;KWRD=EVOLUTION,GENOME,GENOMICS
- Alexander, D.H., Lange, K., 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics* 12, 246. <https://doi.org/10.1186/1471-2105-12-246>
- Alexander, D.H., Novembre, J., Lange, K., 2009. Fast Model-Based Estimation of Ancestry in Unrelated Individuals. *Genome Res* 1655–1664. <https://doi.org/10.1101/gr.094052.109>
- Anderson, C.I.H., Rodhouse, P.G., 2001. Life cycles, oceanography and variability: ommastrephid squid in variable oceanographic environments. *Fish Res* 54, 133–143. [https://doi.org/10.1016/S0165-7836\(01\)00378-2](https://doi.org/10.1016/S0165-7836(01)00378-2)
- Argüelles, J., Tafur, R., 2010a. New insights on the biology of the jumbo squid *Dosidicus gigas* in the Northern Humboldt Current System: Size at maturity, somatic and reproductive investment. *Fish Res* 106, 185–192. <https://doi.org/10.1016/J.FISHRES.2010.06.005>
- Argüelles, J., Tafur, R., 2010b. New insights on the biology of the jumbo squid *Dosidicus gigas* in the Northern Humboldt Current System: Size at maturity, somatic and reproductive investment. *Fish Res* 106, 185–192. <https://doi.org/10.1016/j.fishres.2010.06.005>
- Assis, J., Fernández Bejarano, S.J., Salazar, V.W., Schepers, L., Gouvêa, L., Fragkopoulou, E., Leclercq, F., Vanhoorne, B., Tyberghein, L., Serrão, E.A., Verbruggen, H., De Clerck, O., 2024. Bio-ORACLE v3.0. Pushing marine data layers to the CMIP6 Earth System Models of climate change research. *Global Ecology and Biogeography* 33. <https://doi.org/10.1111/geb.13813>
- Assis, J., Tyberghein, L., Bosch, S., Verbruggen, H., Serrão, E.A., Clerck, O. De, 2017. Bio-ORACLE v2.0: Extending marine data layers for bioclimatic modelling. *Global Ecology and Biogeography* 27, 277–284.
- Bakun, A., Weeks, S.J., 2008. The marine ecosystem off Peru: What are the secrets of its fishery productivity and what might its future hold? *Prog Oceanogr* 79, 290–299. <https://doi.org/10.1016/J.POCEAN.2008.10.027>

- Benjamini, O., Hochberg, Y., 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society* 57, 189–300.
- Camus, P.A., 2001. Biogeografía marina de Chile continental. *Revista chilena de historia natural* 74, 587–617.
- Canales-Aguirre, C.B., Ferrada, S., Hernández, C.E., Galleguillos, R., 2010. Population structure and demographic history of *Gerytherus blacodes* using microsatellite loci. *Fish Res* 106, 102–106. <https://doi.org/10.1016/j.fishres.2010.06.010>
- Canales-Aguirre, C.B., Ferrada-Fuentes, S., 2018. Population genetic structure of Patagonian toothfish (*Dissostichus eleginoides*) in the Southeast Pacific and Southwest Atlantic Ocean 1–20. <https://doi.org/10.7717/peerj.4173>
- Cruz, V.M. V., Kilian, A., Dierig, D.A., 2013. Development of DArT Marker Platforms and Genetic Diversity Assessment of the U.S. Collection of the New Oilseed Crop *Lesquerella* and Related Species. *PLoS One* 8, e64062. <https://doi.org/10.1371/journal.pone.0064062>
- Da Fonseca, R.R., Couto, A., Machado, A.M., Brejova, B., Albertin, C.B., Silva, F., Gardner, P., Baril, T., Hayward, A., Campos, A., Ribeiro, Â.M., Barrio-Hernandez, I., Hoving, H.J., Tafur-Jimenez, R., Chu, C., Frazão, B., Petersen, B., Peñaloza, F., Musacchia, F., Alexander, G.C., Osório, H., Winkelmann, I., Simakov, O., Rasmussen, S., Rahman, M.Z., Pisani, D., Vinther, J., Jarvis, E., Zhang, G., Strugnell, J.M., Castro, L.F.C., Fedrigo, O., Patricio, M., Li, Q., Rocha, S., Antunes, A., Wu, Y., Ma, B., Sanges, R., Vinar, T., Blagoev, B., Sicheritz-Ponten, T., Nielsen, R., Gilbert, M.T.P., 2020. A draft genome sequence of the elusive giant squid, *Architeuthis dux*. *Gigascience* 9, 1–12. <https://doi.org/10.1093/GIGASCIENCE/GIZ152>
- de Meeûs, T., Goudet, J., 2007. A step-by-step tutorial to use HierFstat to analyse populations hierarchically structured at multiple levels. *Infection, Genetics and Evolution* 7, 731–735. <https://doi.org/10.1016/j.meegid.2007.07.005>
- Di Genova, A., Buena-Atienza, E., Ossowski, S., Sagot, M.F., 2021. Efficient hybrid de novo assembly of human genomes with WENGAN. *Nat Biotechnol* 39, 422–430. <https://doi.org/10.1038/S41587-020-00747-W>;SUBJMETA=114,208,212,2302,631,794;KWRD=COMPUTATIONAL+BIOLOGY+AND+BIOINFORMATICS,GENOME+ASSEMBLY+ALGORITHMS,GENOMICS,SOFTWARE
- Do, C., Waples, R.S., Peel, D., Macbeth, G.M., Tillett, B.J., Ovenden, J.R., 2014. NeEstimator v2: re-implementation of software for the estimation of contemporary effective population size (Ne) from genetic data. *Mol Ecol Resour* 14, 209–214.
- Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., Aiden, E.L., 2016. Juicebox Provides a Visualization System for Hi-C Contact Maps with Unlimited Zoom. *Cell Syst* 3, 99–101. <https://doi.org/10.1016/j.cels.2015.07.012>

- Elhaik, E., 2022. Principal Component Analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. *Sci Rep* 12, 14683. <https://doi.org/10.1038/s41598-022-14395-4>
- Ellegren, H., 2014. Genome sequencing and population genomics in non-model organisms. *Trends Ecol Evol* 29, 51–63. <https://doi.org/10.1016/j.tree.2013.09.008>
- Ellegren, H., Galtier, N., Ellegren, N., 2016. Determinants of genetic diversity. *Nature Publishing Group* 17, 422–433. <https://doi.org/10.1038/nrg.2016.58>
- Ferrada Fuentes, S., Galleguillos, R., Herrera-Yáñez, V., Canales-Aguirre, C.B., 2023. Population Genetics of Chilean Jack Mackerel, *Trachurus murphyi* Nichols, 1920, (Pisces, Carangidae), in Waters of the South Pacific Ocean. *Fishes* 8, 162. <https://doi.org/10.3390/fishes8030162>
- FIPA20-20. Ibáñez, C.M., Carrasco, S., Morales, P., Oyarzún, P., Pardo-Gandarillas, C. (2022). Genética poblacional de *Dosidicus gigas* presente en territorio nacional. Informe Final FIPA2020-20, p.130.
- Flanagan, S.P., Jones, A.G., 2017. Constraints on the FST–Heterozygosity Outlier Approach. *Journal of Heredity* 108, 561–573. <https://doi.org/10.1093/jhered/esx048>
- Frichot, E., François, O., 2015. LEA: An R package for landscape and ecological association studies. *Methods Ecol Evol* 6, 925–929. <https://doi.org/10.1111/2041-210X.12382>
- Frichot, E., Schoville, S.D., Bouchard, G., François, O., 2013. Testing for Associations between Loci and Environmental Gradients Using Latent Factor Mixed Models. *MBE* 30, 1687–1699. <https://doi.org/10.1093/molbev/mst063>
- Gautier, M., 2015. Scan for Adaptive Differentiation and Association Analysis with population-specific covariables. *Genetics* 201, 1555–1579.
- Gonçalves da Silva, A., Russello, M.A., 2011. iR el: software for implementing pairwise relatedness estimators and evaluating their performance. *Conserv Genet Resour* 3, 69–71. <https://doi.org/10.1007/s12686-010-9292-4>
- GOUDET, J., 2005. HIERFSTAT, a package for R to compute and test hierarchical F-statistics. *Mol Ecol Notes* 184–186.
- Gruber, B., Unmack, P.J., Berry, O.F., Georges, A., 2018. dartr : An r package to facilitate analysis of SNP data generated from reduced representation genome sequencing. *Mol Ecol Resour* 18, 691–699. <https://doi.org/10.1111/1755-0998.12745>
- Hill, W.G., 1981. Estimation of effective population size from data on linkage disequilibrium. *Genet Res* 38, 209–216. <https://doi.org/10.1017/S0016672300020553>
- Hollander, M., Wolfe, D.A., 1973. *Nonparametric Statistical Methods*. John Wiley & Sons, New York.
- Ibáñez, C., Cubillos, L., Tafur, R., Argüelles, J., Yamashiro, C., Poulin, E., 2011. Genetic diversity and demographic history of *Dosidicus gigas* (Cephalopoda: Ommastrephidae) in the Humboldt Current System. *Mar Ecol Prog Ser* 431, 163–171. <https://doi.org/10.3354/meps09133>

- Ibáñez, C., Ulloa, P., 2014. Desarrollo de la pesquería y comercialización del calamar *Dosidicus gigas* en Chile. *Amici Molluscarum* 22, 7–14.
- Ibáñez, C.M., Argüelles, J., Yamashiro, C., Sepúlveda, R.D., Pardo-Gandarillas, M.C., Keyl, F., 2016a. Population dynamics of the squids *Dosidicus gigas* (Oegopsida: Ommastrephidae) and *Doryteuthis gahi* (Myopsida: Loliginidae) in northern Peru. *Fish Res* 173, 151–158. <https://doi.org/10.1016/J.FISHRES.2015.06.014>
- Ibáñez, C.M., López-Córdova, D., Zúñiga, A., Arancibia, H., Ibáñez, C.M., López-Córdova, D., Zúñiga, A., Arancibia, H., 2019. Condición reproductiva de la jibia *Dosidicus gigas* en la zona costera del centro-sur de Chile. *Rev Biol Mar Oceanogr* 54, 35–42. <https://doi.org/10.22370/RBMO.2019.54.1.1464>
- Ibáñez, C.M., Poulin, E., 2014. Genetic structure and diversity of squids with contrasting life histories in the Humboldt Current System. *Hidrobiológica* 24, 1–10.
- Ibáñez, C.M., Sepúlveda, R.D., Ulloa, P., Keyl, F., Pardo-Gandarillas, M.C., 2016b. The biology and ecology of the jumbo squid *Dosidicus gigas* (Cephalopoda) in Chilean waters: a review. *Lat Am J Aquat Res* 43, 402–414. <https://doi.org/10.3856/vol43-issue3-fulltext-2>
- Ihaka, R., Gentleman, R., 1996. R : A Language for Data Analysis and Graphics. *Journal of Computational and Graphical Statistics* 5, 299–314.
- Jombart, T., 2008. Adegnet: A R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. <https://doi.org/10.1093/bioinformatics/btn129>
- Jombart, T., Ahmed, I., 2011. adegenet 1.3-1 : new tools for the analysis of genome-wide SNP data. *Bioinformatics Applications Note* 27, 3070–3071. <https://doi.org/10.1093/bioinformatics/btr521>
- Jombart, T., Devillard, S., Balloux, F., Falush, D., Stephens, M., Pritchard, J., Pritchard, J., Stephens, M., Donnelly, P., Corander, J., Waldmann, P., Sillanpää, M., Tang, J., Hanage, W., Fraser, C., Corander, J., Lee, C., Abdool, A., Huang, C., Jombart, T., Jombart, T., Devillard, S., Dufour, A., Pontier, D., Jombart, T., Pontier, D., Dufour, A., McVean, G., Novembre, J., Stephens, M., Patterson, N., Price, A., Reich, D., Price, A., Patterson, N., Plenge, R., Weinblatt, M., Shadick, N., Reich, D., Hotelling, H., Hotelling, H., Pearson, K., Liu, N., Zhao, H., Fisher, R., Lachenbruch, P., Goldstein, M., Aitchison, J., Reymont, R., Beharav, A., Nevo, E., Fraley, C., Raftery, A., Cann, H., Toma, C. de, Cazes, L., Legrand, M., Morel, V., Piouffre, L., Bodmer, J., Bodmer, W., Bonne-Tamir, B., Cambon-Thomsen, A., Ramachandran, S., Deshpande, O., Roseman, C., Rosenberg, N., Feldman, M., Cavalli-Sforza, L., Rosenberg, N., Pritchard, J., Weber, J., Cann, H., Kidd, K., Zhivotovsky, L., Feldman, M., Wang, S., Lewis, C., Jakobsson, M., Ramachandran, S., Ray, N., Bedoya, G., Rojas, W., Parra, M., Molina, J., Gallo, C., Balloux, F., Rosenberg, N., Mahajan, S., Ramachandran, S., Zhao, C., Pritchard, J., Feldman, M., Rambaut, A., Pybus, O., Nelson, M., Viboud, C., Taubenberger, J., Holmes, E., Russell, C., Jones, T., Barr, I., Cox, N., Garten, R., Gregory, V., Gust, I., Hampson, A., Hay, A., Hurt, A., Smith, D., Lapedes, A., Jong, J. de, Bestebroer, T., Rimmelzwaan, G., Osterhaus, A., Fouchier, R., Holmes, E., Ghedin, E., Miller, N., Taylor, J., Bao, Y., George, K.S., Grenfell, B.,

- Salzberg, S., Fraser, C., Lipman, D., Young, J., Palese, P., Benson, D., Karsch-Mizrachi, A., Lipman, D., Ostell, J., Wheeler, D., Larkin, M., Blackshields, G., Brown, N., Chenna, R., McGettigan, P., McWilliam, H., Valentin, F., Wallace, I., Wilm, A., Lopez, R., Waterhouse, A., Procter, J., Martin, D., Clamp, M., Barton, G., Paradis, E., Claude, J., Strimmer, K., Handley, L.L., Manica, A., Goudet, J., Balloux, F., Serre, D., Paabo, S., Corander, J., Marttinen, P., Siren, J., Tang, J., Francois, O., Ancelet, S., Guillot, G., Hunley, K., Healy, M., Long, J., Kittles, R., Weiss, K., Manica, A., Prugnolle, F., Balloux, F., Prugnolle, F., Manica, A., Balloux, F., Romero, I., Manica, A., Handley, L.L., Balloux, F., Amos, W., Hoffman, J., Fraley, C., Raftery, A., Peres-Neto, P., Jackson, D., Somers, K., Saporta, G., Paradis, E., Dray, S., Dufour, A., Schwarz, G., Evanno, G., Regnaut, S., Goudet, J., Jakobsson, M., Rosenberg, N., Chessel, D., Dufour, A., Thioulouse, J., Dray, S., Dufour, A., Chessel, D., Venables, W., Ripley, B., Nei, M., 2010. Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. *BMC Genet* 11, 94. <https://doi.org/10.1186/1471-2156-11-94>
- Keenan, K., McGinnity, P., Cross, T.F., Crozier, W.W., Prod, P.A., 2013. *diversity* : An R package for the estimation and exploration of population genetics parameters and their associated errors. *Methods Ecol Evol* 4, 782–788. <https://doi.org/10.1111/2041-210X.12067>
- Lawrence, E.R., Fraser, D.J., 2020. Latitudinal biodiversity gradients at three levels: Linking species richness, population richness and genetic diversity. *Global Ecology and Biogeography* 29, 770–788. <https://doi.org/10.1111/geb.13075>
- Legendre, P., Fortin, M., 2010. Comparison of the Mantel test and alternative approaches for detecting complex multivariate relationships in the spatial analysis of genetic data. *Mol Ecol Resour* 10, 831–844. <https://doi.org/10.1111/j.1755-0998.2010.02866.x>
- Li, H., Qu, W., Obrycki, J.J., Meng, L., Zhou, X., Chu, D., Li, B., 2020. Optimizing Sample Size for Population Genomic Study in a Global Invasive Lady Beetle, *Harmonia Axyridis*. *Insects* 11, 290. <https://doi.org/10.3390/insects11050290>
- Lomsadze, A., Burns, P.D., Borodovsky, M., 2014. Integration of mapped RNA-Seq reads into automatic training of eukaryotic gene finding algorithm. *Nucleic Acids Res* 42. <https://doi.org/10.1093/NAR/GKU557>
- Lu, G., Luo, M., 2020. Genomes of major fishes in world fisheries and aquaculture: Status, application and perspective. *Aquac Fish* 5, 163–173. <https://doi.org/10.1016/J.AAF.2020.05.004>
- Luu, K., Bazin, E., Blum, M.G.B., 2017. *pcadapt* : an R package to perform genome scans for selection based on principal component analysis. *Mol Ecol Resour* 17, 67–77. <https://doi.org/10.1111/1755-0998.12592>
- Lynch, M., Ritland, K., 1999. Estimation of Pairwise Relatedness With Molecular Markers. *Genetics* 152, 1753–1766. <https://doi.org/10.1093/genetics/152.4.1753>
- Marandel, F., Charrier, G., Lamy, J., Le Cam, S., Lorance, P., Trenkel, V.M., 2020. Estimating effective population size using RADseq: Effects of SNP selection and sample size. *Ecol Evol* 10, 1929–1937. <https://doi.org/10.1002/ece3.6016>

- Markaida, U., 2001. Reproductive biology of jumbo squid *Dosidicus gigas* in the Gulf of California, 1995–1997. *Fish Res* 54, 63–82. [https://doi.org/10.1016/S0165-7836\(01\)00373-3](https://doi.org/10.1016/S0165-7836(01)00373-3)
- Mijangos, J.L., Gruber, B., Berry, O., Pacioni, C., Georges, A., 2022. *dartR* v2: An accessible genetic analysis platform for conservation, ecology and agriculture. *Methods Ecol Evol* 13, 2150–2158. <https://doi.org/10.1111/2041-210X.13918>
- Miller, J.M., Cullingham, C.I., Peery, R.M., 2020. The influence of a priori grouping on inference of genetic clusters: simulation study and literature review of the DAPC method. *Heredity (Edinb)* 125, 269–280. <https://doi.org/10.1038/s41437-020-0348-2>
- Nazareno, A.G., Bemmels, J.B., Dick, C.W., Lohmann, L.G., 2017. Minimum sample sizes for population genomics: an empirical study from an Amazonian plant species. *Mol Ecol Resour* 17, 1136–1147. <https://doi.org/10.1111/1755-0998.12654>
- Nevárez-Martínez, M.O., Hernández-Herrera, A., Morales-Bojórquez, E., Balmori-Ramírez, A., Cisneros-Mata, M.A., Morales-Azpeitia, R., 2000. Biomass and distribution of the jumbo squid (*Dosidicus gigas*; d’Orbigny, 1835) in the Gulf of California, Mexico. *Fish Res* 49, 129–140. [https://doi.org/10.1016/S0165-7836\(00\)00198-3](https://doi.org/10.1016/S0165-7836(00)00198-3)
- Nigmatullin, C.M., Nesis, K.N., Arkhipkin, A.I., 2001. A review of the biology of the jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae). *Fish Res* 54, 9–19. [https://doi.org/10.1016/S0165-7836\(01\)00371-X](https://doi.org/10.1016/S0165-7836(01)00371-X)
- Oksanen, F., Blanchet, G., Friendly, M., Kindt, R., Legendre, P., McGlinn, D., Minchin, P.R., O’Hara, R.B., Simpson, G.L., Solymos, P.M., Stevens, H.H., Szoecs, E., Wagner, H., 2019. *vegan: Community Ecology Package*. R package version 2.5-6.
- Pailles, Y., Ho, S., Pires, I.S., Tester, M., Negrão, S., Schmöckel, S.M., 2017. Genetic Diversity and Population Structure of Two Tomato Species from the Galapagos Islands. *Front Plant Sci* 8. <https://doi.org/10.3389/fpls.2017.00138>
- Palumbi, S.R., 2003. Population genetics, demographic connectivity, and the design of marine reserves. *Ecological Applications* 13, S146–S158.
- Pardo-Gandarillas, M.C., Morales, P., Catalán, J., Carrasco, S.A., Hernández, S., Oyarzún, P.A., Wiff, R., Ibáñez, C.M., 2025. Genetic Population Structure of Jumbo Squid (*Dosidicus gigas*) in the Southeastern Pacific Ocean and Its Implication for Fisheries Management. *Fish Manag Ecol* 32, 59–69. <https://doi.org/10.1111/fme.12782>
- Patil, I., 2021. Visualizations with statistical details: The “ggstatsplot” approach. *J Open Source Softw* 6, 3167. <https://doi.org/10.21105/joss.03167>
- Pembleton, L.W., Cogan, N.O.I., Forster, J.W., 2013. StAMPP: An R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Mol Ecol Resour* 13, 946–952. <https://doi.org/10.1111/1755-0998.12129>
- Qu, W., Liang, N., Wu, Z., Zhao, Y., Chu, D., 2020. Minimum sample sizes for invasion genomics: Empirical investigation in an invasive whitefly. *Ecol Evol* 10, 38–49. <https://doi.org/10.1002/ece3.5677>

- Revelle, W., 2020. psych: Procedures for Psychological, Psychometric, and Personality Research [WWW Document]. Northwestern University, Evanston, Illinois. URL <https://personality-project.org/r/overview.pdf>
- Rocha, F., Guerra, A., González, A.F., 2001. A review of reproductive strategies in cephalopods. *Biological Reviews* 76, 291–304. <https://doi.org/10.1017/S1464793101005681>
- Romiguier, J., Gayral, P., Ballenghien, M., Bernard, A., Cahais, V., Chenuil, A., Chiari, Y., Derrat, R., Duret, L., Faivre, N., Loire, E., Lourenco, J.M., Nabholz, B., Roux, C., Tsagkogeorga, G., Weber, A.A.-T., Weinert, L.A., Belkhir, K., Bierne, N., Glémin, S., Galtier, N., 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. *Nature* 515, 261–263. <https://doi.org/10.1038/nature13685>
- Sanchez, G., Kawai, K., Yamashiro, C., Fujita, R., Wakabayashi, T., Sakai, M., Umino, T., 2020b. Patterns of mitochondrial and microsatellite DNA markers describe historical and contemporary dynamics of the Humboldt squid *Dosidicus gigas* in the Eastern Pacific Ocean. *Rev Fish Biol Fish* 30, 519–533. <https://doi.org/10.1007/s11160-020-09609-9>
- Sanchez, G., Tomano, S., Yamashiro, C., Fujita, R., Wakabayashi, T., Sakai, M., Umino, T., 2016a. Population genetics of the jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) in the northern Humboldt Current system based on mitochondrial and microsatellite DNA markers. *Fish Res* 175, 1–9. <https://doi.org/10.1016/J.FISHRES.2015.11.005>
- Sanchez, G., Tomano, S., Yamashiro, C., Fujita, R., Wakabayashi, T., Sakai, M., Umino, T., 2016b. Population genetics of the jumbo squid *Dosidicus gigas* (Cephalopoda: Ommastrephidae) in the northern Humboldt Current system based on mitochondrial and microsatellite DNA markers. *Fish Res* 175, 1–9. <https://doi.org/10.1016/j.fishres.2015.11.005>
- Sandoval-Castellanos, E., Uribe-Alcocer, M., Díaz-Jaimes, P., 2010. Population genetic structure of the Humboldt squid (*Dosidicus gigas* d'Orbigny, 1835) inferred by mitochondrial DNA analysis. *J Exp Mar Biol Ecol* 385, 73–78. <https://doi.org/10.1016/J.JEMBE.2009.12.015>
- Sandoval-Castellanos, E., Uribe-Alcocer, M., Díaz-Jaimes, P., 2007. Population genetic structure of jumbo squid (*Dosidicus gigas*) evaluated by RAPD analysis. *Fish Res* 83, 113–118. <https://doi.org/10.1016/j.fishres.2006.09.007>
- Seehausen, O., Butlin, R.K., Keller, I., Wagner, C.E., Boughman, J.W., Hohenlohe, P.A., Peichel, C.L., Saetre, G.-P., Bank, C., Brännström, Å., Brelsford, A., Clarkson, C.S., Eroukhmanoff, F., Feder, J.L., Fischer, M.C., Foote, A.D., Franchini, P., Jiggins, C.D., Jones, F.C., Lindholm, A.K., Lucek, K., Maan, M.E., Marques, D.A., Martin, S.H., Matthews, B., Meier, J.I., Möst, M., Nachman, M.W., Nonaka, E., Rennison, D.J., Schwarzer, J., Watson, E.T., Westram, A.M., Widmer, A., 2014. Genomics and the origin of species. *Nat Rev Genet* 15, 176–192. <https://doi.org/10.1038/nrg3644>

- Staaf, D., Ruiz-Cooley, R., Elliger, C., Lebaric, Z., Campos, B., Markaida, U., Gilly, W., 2010. Ommastrephid squids *Sthenoteuthis oualaniensis* and *Dosidicus gigas* in the eastern Pacific show convergent biogeographic breaks but contrasting population structures. *Mar Ecol Prog Ser* 418, 165–178. <https://doi.org/10.3354/meps08829>
- Staaf, D.J., Camarillo-Coop, S., Haddock, S.H.D., Nyack, A.C., Payne, J., Salinas-Zavala, C.A., Seibel, B.A., Trueblood, L., Widmer, C., Gilly, W.F., 2008. Natural egg mass deposition by the Humboldt squid (*Dosidicus gigas*) in the Gulf of California and characteristics of hatchlings and paralarvae. *Journal of the Marine Biological Association of the United Kingdom* 88, 759–770. <https://doi.org/10.1017/S0025315408001422>
- Tafur, R., Keyl, F., Argüelles, J., 2010. Reproductive biology of jumbo squid *Dosidicus gigas* in relation to environmental variability of the northern Humboldt Current System. *Mar Ecol Prog Ser* 400, 127–141. <https://doi.org/10.3354/meps08386>
- Thia, J.A., 2023. Guidelines for standardizing the application of discriminant analysis of principal components to genotype data. *Mol Ecol Resour* 23, 523–538. <https://doi.org/10.1111/1755-0998.13706>
- Tinker, N.A., Kilian, A., Wight, C.P., Heller-Uszynska, K., Wenzl, P., Rines, H.W., Bjørnstad, Å., Howarth, C.J., Jannink, J.-L., Anderson, J.M., Rossnagel, B.G., Stuthman, D.D., Sorrells, M.E., Jackson, E.W., Tuveson, S., Kolb, F.L., Olsson, O., Federizzi, L., Carson, M.L., Ohm, H.W., Molnar, S.J., Scoles, G.J., Eckstein, P.E., Bonman, J.M., Ceplitis, A., Langdon, T., 2009. New DArT markers for oat provide enhanced map coverage and global germplasm characterization. *BMC Genomics* 10, 39. <https://doi.org/10.1186/1471-2164-10-39>
- Tyberghein, L., Verbruggen, H., Pauly, K., Troupin, C., Mineur, F., De Clerck, O., 2012. Bio-ORACLE: a global environmental dataset for marine species distribution modelling. *Global Ecology and Biogeography* 21, 272–281. <https://doi.org/10.1111/j.1466-8238.2011.00656.x>
- Veliz, D., Rojas-Hernández, N., Fibla, P., Dewitte, B., Cornejo-Guzmán, S., Parada, C., 2021. High levels of connectivity over large distances in the diadematid sea urchin *Centrostephanus sylviae*. *PLoS One* 16, e0259595. <https://doi.org/10.1371/journal.pone.0259595>
- Veliz, D., Rojas-Hernández, N., Vega-Retter, C., Zaviezo, C., Garrido, I., Pardo, L.M., 2022. Spatial and temporal stability in the genetic structure of a marine crab despite a biogeographic break. *Sci Rep* 12, 14192. <https://doi.org/10.1038/s41598-022-18368-5>
- Wang, J., 2002. An Estimator for Pairwise Relatedness Using Molecular Markers. *Genetics* 160, 1203–1215. <https://doi.org/10.1093/genetics/160.3.1203>
- Waples, R.S., Do, C., 2008. LDNE: A program for estimating effective population size from data on linkage disequilibrium. *Mol Ecol Resour* 8, 753–756. <https://doi.org/10.1111/j.1755-0998.2007.02061.x>
- Weir, B., Cockerham, C., 1984. Estimating F-Statistics for the Analysis of Population-Structure. *Evolution (N Y)* 38, 1358–1370.

Xu, H., Wang, C., Liu, Y., Liu, B., Li, G., 2021. Development and characterization of 101 SNP markers in jumbo flying squid, *Dosidicus gigas*. Conserv Genet Resour 13, 13–20. <https://doi.org/10.1007/S12686-020-01177-1/TABLES/1>

