

UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS VETERINARIAS Y PECUARIAS

ESCUELA DE POSTGRADO Y POSTÍTULO



DETECCIÓN DE SEÑALES DE SELECCIÓN A NIVEL GENÓMICO DEL SALMÓN CHINOOK (*Oncorhynchus tshawytscha*) EN SUDAMÉRICA, NUEVA ZELANDA Y NORTEAMÉRICA

Rodrigo Ignacio Marín Nahuelpi

Tesis para optar al
Grado de Magíster en
Ciencias Animales y Veterinarias,
y al título profesional de Médico Veterinario

SANTIAGO – CHILE

2021

UNIVERSIDAD DE CHILE

FACULTAD DE CIENCIAS VETERINARIAS Y PECUARIAS

ESCUELA DE POSTGRADO Y POSTÍTULO



DETECCIÓN DE SEÑALES DE SELECCIÓN A NIVEL GENÓMICO DEL SALMÓN CHINOOK (*Oncorhynchus tshawytscha*) EN SUDAMÉRICA, NUEVA ZELANDA Y NORTEAMÉRICA

Rodrigo Ignacio Marín Nahuelpi

Tesis para optar al
Grado de Magíster en
Ciencias Animales y Veterinarias,
y al título profesional de Médico Veterinario

PROFESOR GUÍA: JOSÉ MANUEL YÁÑEZ LÓPEZ

PROFESOR CO-GUÍA: DANIEL GÓMEZ-UCHIDA

FINANCIAMIENTO: FONDECYT 1191256

SANTIAGO – CHILE

2021

Departamento y Fuentes de Financiamiento:

Esta Tesis de Grado se realizó en el Laboratorio de Genética y Genómica Acuícola, del Departamento de Medicina Preventiva Animal de la Facultad de Ciencias Veterinarias y Pecuarias de la Universidad de Chile, en colaboración con el Laboratorio de Genómica en Ecología, Evolución y Conservación, del Departamento de Zoología de la Universidad de Concepción, y contó con financiamiento PROYECTO FONDECYT 1191256

BIOGRAFÍA

Nacido en La Calera (Región de Valparaíso, Chile) el 22 de febrero de 1994. Criado en Antofagasta junto a su familia desde pequeño. Fue allí donde cursó su educación primaria y la enseñanza media. Estudió en la Escuela Maximiliano Poblete (D-74) desde primero a cuarto básico. Posteriormente ingresó a la Escuela Lidia Moreno, donde cursó desde quinto a octavo básico. Posteriormente ingresó becado al Colegio San Luis, donde cursó desde primero medio a tercero medio. Su último año de formación fue en el Liceo Domingo Herrera Rivera (B-13), también en Antofagasta. Una vez terminada su enseñanza media, ingresó a estudiar Medicina Veterinaria en la Facultad de Ciencia Veterinarias de la Universidad de Chile el año 2013. El año 2018 egresó sin reprobar cursos, graduándose así de Licenciatura en Ciencias Veterinarias y Pecuarias con distinción. Ingresó al programa de continuidad de la Escuela de Postgrado y Postítulo de esta misma facultad, y el año 2019 ingresa al Magister en Ciencias Veterinarias y Animales. Sus estudios fueron dirigidos principalmente al área de genética acuícola y bioinformática, donde a la fecha ha colaborado como coautor de tres publicaciones científicas. La primera publicación, llamada “*Single-Step Genome-Wide Association Study for Resistance to Piscirickettsia salmonis in Rainbow Trout (Oncorhynchus mykiss)*”, fue publicada en la revista *G3 – Genes Genomes and Genetics*, donde estudió la arquitectura genética y la influencia del componente genético sobre la resistencia frente a infecciones por esta bacteria en poblaciones de Trucha Arcoíris. La segunda publicación denominada “*Estimates of Autozygosity Through Runs of Homozygosity in Farmed Coho Salmon*” fue publicada en la revista *Genes*, donde se evaluó la consanguinidad en poblaciones de cultivo de salmón Coho. La tercera publicación denominada “*Detection of selection signatures in the genome of a farmed population of anadromous rainbow trout (Oncorhynchus mykiss)*” fue publicada en la revista *Genomics*, donde se evaluaron señales de selección a nivel molecular por efecto de la domesticación en poblaciones de Trucha Arcoíris. Existe una publicación en preparación denominada “*Genomic signatures of selection support preadaptation of invasive Chinook salmon in multiple continents*”, la cual aún no ha sido publicada.

DEDICATORIA

Esta memoria de título está dedicada a mi familia, amigos, y a todas las personas que tuvieron un rol durante mi formación personal y académica. Todos han dejado una huella importante y han ayudado a llevar a cabo este logro.

AGRADECIMIENTOS

En primer lugar, le agradezco a mi familia entera por todo su esfuerzo en darme las herramientas y apoyo para llevar a cabo mi formación personal y académica. A mi abuela Gilda, a mi madre Alexandra, a Sonia, Macarena, Tabaré, Ethel Bedda, Leila, Evelyn, y a quienes ya no están en lo tangible, pero si en mi memoria, Auristela, Sonia, Enrique y Andrés. Todos aportaron su grano de arena, y siempre estaré agradecido por su esfuerzo, valentía, generosidad y cariño. También quisiera agradecer de manera muy especial a Constanza y su familia, por su apoyo incondicional e incommensurable cariño.

Agradezco al Fondo Nacional de Desarrollo Científico y Tecnológico, por su aporte constante a las ciencias en Chile, y por brindar el financiamiento necesario para llevar a cabo el proyecto de investigación que permitió realizar la presente memoria de título. También a la Facultad de Ciencias Veterinarias y Pecuarias de la Universidad de Chile por entregarme el pensamiento crítico y las herramientas necesarias para desempeñar mi futuro papel como profesional, consciente de mi entorno natural y social. Agradezco, además, a los múltiples programas de becas de los cuales fui beneficiario (Beca Bicentenario y Gratuidad), principalmente por parte del Estado de Chile, pues sin ellos, habría sido imposible para mi costear mi formación profesional.

Agradezco especialmente al GEECLAB y al Doctor Daniel Gómez-Uchida, por brindarme la oportunidad y la confianza de trabajar junto a el en este proyecto. Agradezco a quienes participaron en la recopilación de muestras del salmón Chinook en el sur de Chile, ya que hicieron posible esta tesis. Agradezco al equipo del núcleo milenio INVASAL y por acogerme como investigador joven y brindar su colaboración en mi participación en distintos congresos de los cuales pude participar.

Agradezco a la comisión evaluadora de esta tesis, al Dr. Jurij Wacyk y al Dr. Leonardo Saenz, quienes han aportado constructivamente al desarrollo y conclusión de esta memoria.

Quisiera agradecer especialmente al Dr. José Manuel Yáñez por su confianza, apoyo y por su calidad de persona. Por brindarme la oportunidad de formar parte del laboratorio de Genética y Genómica Acuícola, y la confianza depositada en mi persona para llevar adelante las diferentes tareas y actividades relacionadas con los programas de mejoramiento. Por su bondad en la transmisión de conocimiento, sus consejos y sus consejos para el desarrollo profesional y personal. También a todo el equipo del laboratorio, María Ignacia, María Eugenia, Pablo, Sebastián, David, Jouseph, Carolina, Paulina, Baltasar y Tamara, por ser un excelente equipo de trabajo y amigos.

Quisiera agradecer a mis amigos “los metaleros” por todo lo vivido y compartido con ellos durante mi estadía en la universidad, son grandes personas y los recordaré siempre con mucho cariño; Gianina, Nico, Gazy, Cata, Tomy, Ricardo, Gabriel, Marcelo, Paz, Martin, Daniela, Eddie y Nicole.

INDICE DE CONTENIDOS

1. INTRODUCCIÓN	1
2. REVISIÓN BIBLIOGRÁFICA	3
2.1 Salmón Chinook: Historia y características	3
2.2 Invasiones biológicas y factores genéticos relacionados a la adaptación de una especie foránea	5
2.3 Perspectivas del éxito del salmón Chinook como especie invasora	9
2.4 Pre-adaptación fenotípica y genética	11
2.5 Divergencia y uso de información molecular	13
2.6 Selección natural a nivel molecular	15
2.7 Recombinación	19
2.8 Métodos para la detección de señales de selección	20
3. HIPÓTESIS	22
4. OBJETIVOS	22
4.1 Objetivo General	22
4.2 Objetivos específicos	22
5. MATERIALES Y MÉTODOS	23
5.1 Poblaciones donadoras y naturalizadas	23
5.2 Obtención de datos genómicos	25
5.3 Control de calidad de la información genómica	26
5.4 Estructura poblacional y diferenciación genética	27

5.5 Métodos a utilizar para detectar <i>loci</i> bajo selección	28
5.5.1 BayeScan.....	29
5.5.2 PCAdapt.....	31
5.6 Escaneos genómicos	32
6. RESULTADOS	35
6.1 Genotipificación y control de calidad	35
6.2 Diversidad genética, endogamia y diferenciación poblacional	38
6.3 Análisis de estructura poblacional.....	39
6.4 Escaneos genómicos	41
6.5 Genes candidatos	44
7. DISCUSIÓN	48
7.1 Objetivo 1: ¿Exísten señales de selección en común entre los múltiples casos de poblaciones divergentes?	48
7.2 Objetivo 2: ¿Cuáles son los genes candidatos que se encuentran en o cercanos a los marcadores detectados bajo selección?	53
8. CONCLUSIONES	57
9. BIBLIOGRAFÍA	58

INDICE DE TABLAS

TABLA 1. Región y ríos de origen de las poblaciones donadoras, ríos invadidos, número de individuos por población, número de marcadores a utilizar y tipo de secuenciación.....	24
TABLA 2. Descripción de resultados del control de calidad, ubicación de las poblaciones e indicadores de diversidad genética.....	37
TABLA 3. Diferenciación poblacional (de acuerdo al índice FST) de a pares utilizando a) el set de datos completo (FULL; 4.529); b) set de datos neutral (4.151 SNPs) y C) set de datos atípicos (378 SNPs).....	39
TABLA 4. Error de validación cruzada para cada <i>cluster</i> (K) estimada según ADMIXTURE.....	40
TABLA 5. Número de <i>outliers</i> encontrados por BayeScan para cada prior odds (PO) y PCAdapt con el valor de K estimado para cada caso; y el número de <i>outliers</i> en común entre ambos métodos utilizando un PO = 10.....	42
TABLA 6. Marcadores moleculares con mayor soporte estadístico de estar bajo selección positiva. Se indica la identificación del SNP, el cromosoma, la identidad de los alelos, posición física y la ubicación dentro del genoma de referencia.....	43
TABLA 7. Lista completa de genes dentro de ventanas de 1Mb en el genoma de referencia utilizando como centro los marcadores con mayor poder estadístico de estar bajo selección.....	47
TABLA 8. Frecuencias del alelo menor de los SNPs bajo selección con mayor soporte estadístico dentro de cada población. Se indica la identidad de los alelos (menor / mayor) y la frecuencia del alelo menor en paréntesis.....	52

INDICE DE AYUDAS ILUSTRATIVAS

- FIGURA 1.** Regiones geográficas muestreadas. A) Norteamérica; B) Sudamérica y C) Nueva Zelanda. En amarillo las poblaciones no nativas y en rojo las poblaciones nativas donadoras..... 23
- FIGURA 2.** Gráfico de barras realizado a partir de los valores Q de ADMIXTURE. Cada barra corresponde a un individuo, y cada color corresponde a la proporción Q de cada cluster (K). Las líneas blancas separan las poblaciones. A – B – C corresponde a K = 6 y D – F – G a K = 7..... 38
- FIGURA 3.** Diagrama de Venn con el número de marcadores atípicos en común para todos los casos utilizando ambas metodologías, utilizando un PO = 100..... 41
- FIGURA 4.** Gráficos de Manhattan contruidos con resultados de BayeScan (asumiendo PO = 10). Muestra las regiones genéticas donde se detectaron señales de respuesta a la selección divergente. Cada gráfico corresponde a un caso independiente de poblaciones evolutivamente divergentes en el salmón Chinook. Cada punto es un SNP. El eje X muestra los cromosomas (impares en rojo y pares en azul), el eje Y muestra el umbral de significancia establecido a partir del $-\log_{10}$ del valor q (“q-value” = 0,05)..... 42
- FIGURA 5.** Gráficos de Manhattan contruidos con resultados de PCAdapt. Se muestran las regiones genéticas donde se detectaron señales de respuesta a la selección divergente. Cada gráfico corresponde a un caso independiente de poblaciones evolutivamente divergentes en el salmón Chinook. Cada punto es un SNP. El eje X muestra los cromosomas (impares en rojo y pares en azul), el eje Y muestra el umbral de significancia establecido a partir del $-\log_{10}$ del valor q (“q-value” = 0,05)..... 44

INDICE DE ABREVIATURAS

ADN	: Ácido Desoxirribonucleico
HWE	: Hardy-Weinberg Equilibrium
IBD	: <i>Identity by descent</i>
IPNV	: Infectious Pancreatic Necrosis Virus
LD	: <i>Linkage Disequilibrium</i>
MAF	: Minor allele frequency
PCA	: <i>Principal Component Analysis</i>
PO	: <i>Prior odds</i>
RAD-Seq	: <i>Restriction site associated digestion sequencing</i>
SNP	: <i>Single Nucleotide Polymorphism</i>

RESUMEN

Las invasiones biológicas son procesos complejos en los que intervienen dimensiones ecológicas, evolutivas, ambientales y sociales, y probablemente están mediadas por una combinación de alta presión de propagación, disminución de la resistencia biótica y antecedentes genéticos. La hipótesis de la "pre-adaptación" sugiere que un conjunto de genes no aleatorio provenientes de la población donante puede ser uno de los responsables de la capacidad de fundar nuevas poblaciones fuera de su área de distribución nativa. El éxito de la invasión y el establecimiento de poblaciones autosostenibles de salmón Chinook en distintos continentes constituye una importante oportunidad para poner a prueba esta hipótesis a nivel del genoma. En esta tesis, aprovechamos los recursos genómicos recientemente publicados para el salmón Chinook y pudimos identificar 4529 SNPs bi-alélicos de alta calidad que segregan entre las poblaciones nativas e invasoras contemporáneas para detectar señales de selección a nivel genómico en común entre todos los casos. Realizamos un escaneo del genoma para buscar loci atípicos mediante los programas PCAdapt y BayeScan, utilizando un modelo comparativo pareado que consideraba poblaciones nativas y no nativas donantes, incluyendo poblaciones provenientes de los Grandes Lagos, Nueva Zelanda y las poblaciones invasoras del oeste y este de Sudamérica. Utilizando parámetros conservadores se encontraron 2 valores atípicos comunes en los cromosomas 5 y 14 entre todas las comparaciones de ambos métodos. Los genes candidatos se identificaron buscando 500 kb río arriba y río abajo de cada SNP atípico, según su posición en el genoma de referencia del salmón Chinook (Otsh_v1.0). Los genes candidatos incluyen, entre otros, el *Myo6*, un importante gen previamente encontrado bajo selección, asociado a la resistencia a virus como el *IPNV*, y al crecimiento en el salmón atlántico; el *LSAMP*, un gen que se sugiere que regula la actividad exploratoria, la adaptación a nuevos entornos estresantes, el comportamiento de agresión y que se ha encontrado previamente bajo selección positiva en diferentes especies; el *IGSF11*, un gen que participa en el desarrollo de las células pigmentarias, promueve la migración y la supervivencia de las células del linaje melanóforo y media en el patrón de pigmentación de los adultos en el pez cebra. Pese a que es posible que exista un mayor número de marcadores genéticos realmente bajo selección no capturados en este trabajo, estos resultados apoyan la hipótesis de la "pre-adaptación", incluso cuando se utiliza una densidad de marcadores relativamente baja. Existe un pool de genes putativamente bajo selección positiva divergente entre los múltiples casos evaluados. Este trabajo constituye una primera aproximación para desentrañar las bases moleculares de la adaptación local y la capacidad de invasión del salmón Chinook. Los trabajos futuros deberían tener como objetivo la búsqueda de señales de selección utilizando métodos intrapoblacionales y una mayor densidad de marcadores, lo que podría ayudar a explicar el éxito de la colonización del salmón Chinook en todo el mundo.

SUMMARY

Biological invasions are complex processes involving ecological, evolutionary, environmental, and social dimensions, and are likely mediated by a combination of high propagule pressure, decreased biotic resistance and genetic background. The 'pre-adaptation' hypothesis suggests that a non-random, pre-adapted set of colonists from the donor population may be ultimately responsible for founding a new population outside their native range. The successful invasion and establishment of self-sustaining Chinook salmon populations to different continents constitute an important opportunity to test this hypothesis at genome level. Herein, we took advantage on recently published genome resources for Chinook salmon and were able to identify 4529 high quality bi-allelic SNPs segregating among contemporary native and invasive populations. We performed a genome scan to search for outlier loci using PCAdapt and BayeScan packages, using a paired model considering donor native and non-native populations, including Great Lakes, New Zealand, and western and eastern South America invasive populations. Using conservative parameters 2 common outliers were found at chromosomes 5 and 14 among all comparisons from both methods. Candidate genes were identified by searching 500 kb up and downstream from each outlier SNP, according to their position in Chinook salmon reference genome (Otsh_v1.0). Candidate genes include among others *Myo6*, an important gene previously found under selection, associated with resistance to viruses such as IPNV, and growth in Atlantic salmon; *LSAMP*, a gene suggested to regulate exploratory activity, adaptation to novel stressful environments, aggression behavior and previously found to be under positive selection in different species; *IGSF11*, a gene that participates in pigment cell development, promotes the migration and survival of melanophore lineage cells and mediates adult pigment pattern in zebrafish. There are likely more gene markers that do not follow a neutral pattern of evolution that were not assessable by this work, however these findings support the 'pre-adaptation' hypothesis, even when a relatively low marker density is used. This work constitutes a first approach to disentangle the molecular basis of the local adaptation and invasiveness ability of Chinook salmon. Future works should aim to search for signatures of selection using intrapopulation methods and a wider density of markers, which may help to explain successful colonization of Chinook salmon around the globe.

INTRODUCCIÓN

Las especies del género *Oncorhynchus* (también conocidos como salmones del pacífico), han sido introducidos en diferentes continentes. Por ejemplo, el salmón Chinook (*Oncorhynchus tshawytscha*) ha sido introducido exitosamente a diferentes ríos y/o lagos de Sudamérica, Oceanía y de Norteamérica, donde han logrado establecer poblaciones autosustentables y colonizar ríos diferentes a los cuales fueron originalmente introducidos, por lo cual es considerada una especie invasora. La invasión biológica es un fenómeno complejo que involucra aspectos biológicos, ecológicos, evolutivos e incluso sociales, por lo cual desglosar estos aspectos es fundamental para entender el éxito de la invasión.

Las poblaciones contemporáneas (distribuidas a través del mundo) de salmón Chinook exhiben variación en ciertos rasgos fenotípicos (fecha de retorno, crecimiento en agua dulce, edad de maduración y otros), lo cual sugiere adaptación local. En biología evolutiva, la adaptación local es un fenómeno resultante de la evolución de poblaciones de las especies, donde ciertas poblaciones resultan mejor adaptadas a un determinado ambiente que otros miembros o poblaciones de su misma especie. La variación fenotípica puede ser explicada, en parte, por diferencias en las variantes alélicas y de genotipos presentes en los individuos de cada población. Diferentes fuerzas evolutivas moldean las frecuencias de las variantes alélicas y de los genotipos presentes en una población, como la selección natural, entre otras (migración, mutación y deriva génica). Cuando una variante confiere una ventaja en un ambiente local, la selección natural tiende a aumentar su frecuencia en la población, generando grandes diferencias entre poblaciones adaptadas a localidades diferentes. La identificación de genes o sitios del genoma (*loci*) que han sido o están siendo seleccionados, permite hipotetizar fenotipos, caracteres o mecanismos involucrados con la adaptación de diversas especies.

Los cambios genéticos generados durante el proceso de divergencia, que diferencian a las poblaciones a través del tiempo, pueden ser analizados gracias al desarrollo de tecnologías de secuenciación y genotipificación de marcadores genéticos distribuidos a través del genoma. Entre los marcadores más utilizados se

encuentran los llamados polimorfismos de un solo nucleótido (*Single Nucleotide Polymorphisms*, “SNPs”), que corresponden variaciones dentro de las secuencias de ADN, que afectan a un solo nucleótido. Estos marcadores se encuentran distribuidos a lo largo del genoma de cualquier especie, son heredables, rastreables y generalmente poseen solo dos nucleótidos (bi-alélicos). Utilizando estos marcadores es posible detectar “señales” o “huellas” del proceso de selección basado en las frecuencias alélicas de cada población.

Una de las hipótesis para explicar el éxito adaptativo de diferentes especies animales y vegetales es la hipótesis de “pre-adaptación” de caracteres fenotípicos (Riva Rossi et al., 2012; Ciancio et al., 2015; Elst et al., 2016). Caracteres pre-adaptados son aquellos rasgos que son funcionalmente ventajosos en ambientes no nativos. El establecimiento exitoso del salmón Chinook alrededor de diferentes regiones geográficas del mundo, junto a la disponibilidad de los genotipos de poblaciones donadoras e invasoras, constituye una oportunidad única para estudiar regiones bajo selección en común a través del mundo y para aportar antecedentes a la hipótesis de la “pre-adaptación” a nivel genómico. Para esto, en este proyecto se plantea utilizar escaneos genómicos independientes entre poblaciones divergentes (donador-invasor), con dos metodologías independientes en 8 casos diferentes de divergencia. Los escaneos genómicos son análisis que utilizan la información de cada SNP disponible, para identificar aquellos que muestren un patrón atípico de diferenciación (“*outlier loc*”), es decir, que la diferencia entre la distribución de los alelos de dichos SNPs en ambas poblaciones se explica principalmente por la selección positiva. Además, ya que es posible identificar la posición de un SNP en el genoma, un SNP *outlier* permite identificar genes cercanos al SNP putativo bajo selección que pudieron haber sido seleccionados de forma conjunta. Si a través de los múltiples escaneos genómicos independientes se identifican genes bajo selección positiva en común, entonces podríamos sugerir que estas regiones genómicas albergan genes pre-adaptados, es decir, genes que estarían adaptados previamente a los diferentes hábitats y en consecuencia conferirían ventajas adaptativas y que podrían explicar en parte el éxito adaptativo del Salmón Chinook a diferentes condiciones locales.

Revisión bibliográfica

2.1 Salmón Chinook: Historia y características

El salmón Chinook (*Oncorhynchus tshawytscha*) es una especie anádroma, semélpara y nativa del océano pacífico norte, donde habita de forma nativa diferentes ríos y afluentes de Rusia, Japón, Alaska y toda la costa oeste de Canadá y Estados Unidos (Augerot, 2005; Quinn, 2005). Para reproducirse, la mayoría regresa a sus sitios natales en ríos, tributarios y grandes cuencas hidrográficas (conducta conocida como “*homing*”), sin embargo, es posible la dispersión de individuos y el establecimiento de poblaciones en ríos diferentes al sitio natal (Hendry et al., 2004; Larson et al., 2014).

Esta especie, también conocida como salmón Rey, ha sido exitosamente introducido con fines de pesca recreacional y/o cultivo en diferentes países (Crawford y Muir, 2008). En Chile fue introducido durante la década de 1970, en al menos 5 grandes ríos con desembocadura en el océano pacífico, promovido por políticas gubernamentales e iniciativas privadas con el fin de desarrollar al pesca recreacional y experimentos de “*ocean ranching*” (liberación de ovas fecundadas y/o alevines en ríos para criarse en el mar y la posterior “cosecha” en los ríos) (Becker et al., 2007; Pascual y Ciancio, 2007; Soto et al., 2007). A partir de las poblaciones introducidas a Chile entre los años 1978 y 1989, se lograron establecer de forma exitosa poblaciones por vía artificial (mediadas por el humano) y posteriormente poblaciones naturalizadas (establecidas por migraciones naturales), distribuyéndose actualmente entre la región de La Araucanía hasta Tierra del Fuego (Correa y Gross, 2008; Fernández et al., 2010; Gómez-Uchida et al., 2018). En Argentina, se han descrito poblaciones establecidas en cuencas que desembocan al Océano Atlántico, incluyendo al Río Santa Cruz y Caterina (Ciancio et al., 2005; Riva Rossi et al., 2012). Estas poblaciones se originaron a partir de la dispersión natural de poblaciones artificiales establecidas en Chile, las cuales lograron a través océano Atlántico establecerse en territorio argentino (Correa y Gross, 2008; Gómez-

Uchida *et al.*, 2018), dado que no existen registros de introducción de esta especie a Argentina (Riva-Rossi *et al.*, 2012).

Utilizando registros históricos, junto a tecnologías de secuenciación del ADN, se ha logrado trazar el origen genético de las poblaciones que habitan en Chile y Argentina, revelando que la ancestría de estas poblaciones es polifilética, es decir, de múltiples núcleos genéticos donadores (Correa y Gross, 2008; Niklitschek y Toledo, 2011; Riva Rossi *et al.*, 2012; Ciancio *et al.*, 2015; Correa y Moran, 2017; Gómez-Uchida *et al.*, 2018). Correa y Moran (2017), concluyeron que las poblaciones de los ríos Petrohué y Toltén fueron fundadas a partir de múltiples fuentes. Posteriormente Gómez-Uchida *et al.*, (2018), revisaron los registros históricos disponibles y utilizaron SNPs de los posibles núcleos donadores y no nativos, para concluir que la ancestría donadora predominante del río Toltén y del río Petrohué corresponden a los grupos genéticos de la costa de Oregon-California y al estrecho de Puget (Puget Sound), respectivamente. Estas dos regiones geográficas se ubican en Estados Unidos, Norteamérica. En este continente, el salmón Chinook también fue introducido en la zona de los Grandes Lagos (“Great Lakes”), durante la década de 1960 a partir del núcleo genético de Green River, Washington (Parsons, 1973; Weeder *et al.*, 2005). Posteriormente colonizó diferentes tributarios a través de los grandes lagos, donde hoy existe evidencia de reproducción natural (Connerton *et al.*, 2009; Johnson *et al.*, 2010). Estas poblaciones hacen ciclos migratorios entre ríos y lagos de agua dulce, a diferencia de las poblaciones de Sudamérica y Nueva Zelanda, donde hacen el ciclo migratorio pasa por ríos y mar.

En Nueva Zelanda, el salmón Chinook fue introducido a comienzos del siglo XX, donde en 15 años, a partir de solo un río (Waitaki), logró establecerse en diferentes ríos de South Island de forma exitosa. Estas poblaciones tienen ancestría monofilética, siendo su origen el núcleo genético de Battle Creek River, California (Quinn *et al.*, 2001). Las poblaciones que habitan en los diferentes ríos neozelandeses presentan una alta variación fenotípica de caracteres relacionados

a la probabilidad de supervivencia y éxito reproductivo de los individuos (caracteres como la tasa de crecimiento en agua fresca o de mar, tiempo a la madurez, largo a cierta edad, peso a cierto largo, fecundidad a cierta edad, fechas retorno para desove, y otros), lo cual probablemente se debe a una alta plasticidad fenotípica influenciada por un componente genético que facilitó la expansión del salmón Chinook en Nueva Zelanda (Quinn et al., 2001; Kinnison et al., 2011). En las poblaciones de Sudamérica, la principal diferencia fenotípica está dada por las fechas de retorno a agua dulce (verano u otoño) (Ciancio et al., 2005; Soto et al., 2007).

El salmón Chinook se considera una especie caracterizada por una alta plasticidad fenotípica, ya que presenta grandes diferencias en caracteres relacionados a su ciclo de vida, habiendo individuos que residen un año o más en agua dulce antes de migrar al mar ("*stream-type*") y otros que migran tempranamente antes del año ("*ocean-type*"), además de otras diferencias relacionadas con el largo a cierta edad, peso a cierta edad, edad a la madurez, época de desove, calidad y cantidad de ovas (Quinn et al., 2001; Correa y Gross, 2008; Di Prinzio y Pascual, 2008; Kinnison et al., 2011; Haring et al., 2016; Thorn y Morbey, 2018). Producto la alta variación fenotípica, el salmón Chinook ha sido estudiado en la ecología y biología de la evolución (Quinn, 2005).

2.2 Invasiones biológicas y factores genéticos relacionados a la adaptación de una especie foránea

Cualquier especie introducida por el ser humano en lugares fuera de su rango natural de distribución es una especie invasora. La introducción de especies invasoras conlleva profundas consecuencias en los ecosistemas receptores, como modificar el patrón de distribución natural y abundancia de especies nativas dada la competencia por hábitats y fuentes de alimento, pérdida de servicios ecosistémicos y del retorno económico (Pimentel et al., 2001; Pascual y Ciancio, 2007; Pyšek y Richardson, 2010; Correa y Hendry, 2012; Habit et al., 2015). Por ejemplo, en

especies nativas del sur de Chile, se ha observado una reducción del crecimiento y un menor uso de hábitats, en presencia de especies invasoras (Correa y Hendry, 2012; Vargas et al., 2010).

La capacidad de una especie foránea de adaptarse y colonizar ecosistemas no nativos involucra aspectos biológicos, ecológicos, evolutivos e incluso sociales, por lo cual desglosar estos aspectos es fundamental para entender el establecimiento exitoso y posterior colonización (Vitule et al., 2012; Barrett, 2015; Dlugosch et al., 2015). A pesar de que miles de peces en cientos de especies han experimentado eventos de introducción (Casal, 2006), no está totalmente comprendido el por qué ciertos peces introducidos son capaces de establecerse exitosamente y otros no (Copp et al., 2005). Un marco conceptual del proceso de invasión biológica abordado por Arismendi et al., (2014), apunta a tres dimensiones centrales que actúan de forma concomitante: (1) Dimensión humana; (2) dimensión de interacciones bióticas/abióticas y (3) dimensión del invasor.

Los factores antropogénicos son considerados los principales conductores de una invasión, ya que median la invasión inicial. Los esfuerzos por introducir a la especie (como la captura, almacenaje, transporte y subsecuente liberación de invasores a los ecosistemas receptores), la presión de propágulos y las medidas de respuesta (manejo, contención o erradicación) son consideradas parte de la dimensión humana. Diferentes autores han compilado información respecto a los intentos de introducción de salmónidos a Chile (Basulto, 2003; Correa y Gross, 2008; Correa y Moran, 2017; Gómez-Uchida et al., 2018).

La presión de propágulos hace referencia a la frecuencia y al número de individuos no nativos introducidos, y ha sido considerado uno de los mayores predictores del éxito de una invasión (Hayes y Barry, 2008). En Chile se registran casos de especies exitosas con baja frecuencia e intensidad (Trucha Café), de alta frecuencia y baja intensidad (Salmón Chinook), de alta frecuencia e intensidad (Trucha Arcoíris), sin embargo, otras especies aun con altas frecuencias e intensidad de propágulos han logrado escasa (Salmón Coho) o nula (Salmón del Atlántico) capacidad de formar poblaciones autosustentables (Soto et al., 2001, 2006; Arismendi et al., 2009, 2011,

2014; Schröder y de Leaniz, 2011; Niklitschek et al., 2013; Quiñones et al., 2019; Chalde et al., 2019; Maldonado-Márquez et al., 2020). Si bien no se conoce a cabalidad las causas del por qué, algunos autores han propuesto algunas hipótesis. Arismendi et al., (2014), proponen que el éxito previo de otras especies como *Oncorhynchus mykiss* y *Salmo trutta* (Trucha arcoíris y Trucha marrón, respectivamente), pudo y podría estar limitando el establecimiento de nuevas especies salmónidas, exceptuando al salmón Chinook, dado su menor tiempo de estadía en agua dulce (desde meses hasta poco más de un año versus 2 a 4 años del salmón del Atlántico) y por un solapamiento espacial limitado con las truchas en comparación con el solapamiento del salmón del Atlántico con las truchas (Hansen y Quinn, 1998). Tanto la trucha café como la trucha arcoíris tienden a ser más agresivas y han demostrado ser mejores competidoras que el salmón del Atlántico (Van Zwolle et al., 2012; Houde et al., 2017).

Por otra parte, la domesticación podría influir en la capacidad de establecer exitosamente poblaciones sustentables, sin embargo, su influencia no se conoce a cabalidad, aunque probablemente especies domésticas provenientes de acuicultura sean menos exitosas que aquellas provenientes de poblaciones naturales (Arismendi et al., 2014). Por ejemplo, el salmón del Atlántico doméstico a menudo se reproduce con su conespecífico y su progenie muestra una muy reducida supervivencia en la naturaleza (Glover et al., 2017). Esta teoría ha sido apoyada por estudios que demuestran que truchas de rápido crecimiento provenientes de cultivos domésticos tienen una mayor susceptibilidad a la depredación en comparación con sus conespecíficos de lento crecimiento (Solberg et al., 2020).

La segunda dimensión involucra las interacciones que se generen con la biota nativa, es decir, competidores nativos, predadores, parásitos y patógenos, además entre con otras especies invasoras. Por ejemplo, en el sur de Chile, evidencia indirecta (medida en fuentes de alimento) sugiere que existe competencia entre trucha arcoíris y trucha café (Arismendi y Soto, 2012), y entre salmón del Atlántico y trucha arcoíris (Young et al., 2009). Otro ejemplo es que, en ríos de la Patagonia, las carcasas de salmón Chinook proveen de nutrientes derivados del mar a

ecosistemas de aguadulce, lo cual provee de nutrientes a poblaciones residentes de trucha (Arismendi y Soto, 2012). Además, involucra la resistencia de un ecosistema invadido, que corresponde a las características propias de cada ecosistema invadido que pueden facilitar o no el establecimiento de una población foránea. Involucra aspectos como la competencia entre especies, disponibilidad de alimento, las características bióticas y abióticas (químicas, físicas) del ecosistema en cuestión.

La tercera dimensión incluye aspectos relacionados con el ciclo de vida de la especie, la capacidad migratoria, la plasticidad fenotípica y la tolerancia ambiental. En general, la capacidad de una especie introducida para tolerar condiciones locales depende de su plasticidad fenotípica, del grado de similitud ambiental (entre ambientes nativo e invadido), de tolerar las características fisicoquímicas del ambiente invadido y de los niveles de variación genética neutral y adaptativa (Roman y Darling, 2007; Bell y Gonzalez, 2009). Generalmente, los invasores más exitosos pertenecen a géneros no presentes en los ecosistemas invadidos (Ricciardi y Atkinson, 2004). El éxito de otras especies salmónidas invasoras como la trucha arcoíris y la trucha café ha sido asociado con su alta plasticidad fenotípica (Westley et al., 2013a; Westley et al., 2013b) dada la presencia de poblaciones residentes (Soto et al. 2006), adfluviales (Arismendi et al. 2011) y anádromas (O'Neal y Stanford, 2011). La presión de propágulos, la resistencia ambiental y la influencia de la genética han sido considerados los principales impulsores de invasiones en estudios experimentales (Parker et al., 2013; Szűcs et al., 2014), aunque la presión de propágulos y la resistencia ambiental han sido más estudiadas que los factores genéticos (Colautti et al., 2006; Simberloff, 2009; Parker et al., 2013).

Entre los factores genéticos que podrían aumentar la probabilidad de que las poblaciones subsistan, se encuentra la diversidad genética (Huey et al., 2005). Según Hughes et al., (2008), la diversidad genética se define como “cualquier medida que cuantifique la magnitud de variabilidad genética dentro de una población”. La diversidad genética tiene dos grandes componentes: (1) la abundancia de polimorfismos y (2) el número de alelos. Cuando hay un mayor

número de polimorfismos y de alelos, hay una mayor probabilidad de que se presenten variantes favorables para fenotipos que tengan una mayor probabilidad de sobrevivencia y de reproducirse, en resumen, la eficacia biológica (“*fitness*”).

La heterocigosidad como medida de diversidad genética fue postulada por Nei (1978). La heterocigosidad esperada corresponde a la probabilidad de que al tomar dos alelos de forma aleatoria en una muestra (representativa de la población), estos sean diferentes. La diversidad genética se puede estimar a partir de la heterocigosidad esperada (H_e) bajo el equilibrio Hardy-Weinberg. El equilibrio Hardy-Weinberg establece que, en poblaciones suficientemente grandes para minimizar el efecto de la deriva génica, de genomas diploides, donde hay reproducción aleatoria (panmíxia), sin selección natural, migración ni mutación, la frecuencia de los alelos y de los genotipos puede ser estimada de una generación a otra. Esto raramente ocurre en la realidad, uno o más de los presuntos es generalmente infringido. Para un marcador con dos alelos, como el caso de los SNPs, a partir de las frecuencias de los alelos p y q , la proporción de heterocigotos esperada en la siguiente generación es $2pq$. Las desviaciones significativas de lo esperado (medidas mediante el test exacto de Fisher, usualmente corrigiendo los valores de significancia en función del número de SNPs), pueden ser explicadas por errores de genotipado, o por cualquier otra de las cuatro fuerzas que pueden cambiar las frecuencias alélicas en una población (selección, deriva génica, mutación o migración) (Teo et al., 2007; Turner et al., 2011). La deriva génica es un efecto estocástico dado por el muestreo aleatorio al momento de la reproducción que puede favorecer la pérdida o fijación de alelos en una determinada población, cuya magnitud es dependiente del tamaño poblacional (Gillespie, 2001).

2.3 Perspectivas del éxito del salmón Chinook como especie invasora

Chile es el segundo mayor productor de Salmón del Atlántico a nivel mundial, y los múltiples escapes a lo largo de los años constituyen una alta presión de propágulos, sin embargo, no hay evidencia de que existan poblaciones anádromas autosustentables, es decir, poblaciones que completen su ciclo vital independiente de la intervención humana (Soto et al., 2001, 2006; Niklitschek et al., 2013;

Sepúlveda et al., 2013; Arismendi et al., 2014; Quiñones et al., 2019; FAO, 2020), a diferencia del salmón Chinook que si ha logrado establecer poblaciones autosustentables y colonizar ríos diferentes a los cuales fueron introducidos. Diferentes autores han sugerido que el éxito del salmón Chinook, en términos de lograr establecer poblaciones sustentables y colonizar, se debe a una combinación de factores como la presión de propágulos, la alta plasticidad fenotípica y la baja resistencia de los ecosistemas invadidos (Correa y Gross, 2008, Arismendi et al., 2014). Además, poblaciones con alta diversidad genética tienen mayor probabilidad de subsistir (Huey et al., 2005). Las poblaciones de salmón Chinook establecidas en la Patagonia son genéticamente diversas (en términos de heterocigosidad esperada (H_e)), posiblemente debido a los múltiples eventos de introducción, de múltiples fuentes y la subsecuente hibridación entre los diferentes ancestros genéticos, lo cual junto a la presión de propágulos posiblemente contribuyó a un rápido establecimiento y posterior colonización de ríos vía natural (Becker et al., 2007; Roman y Darling, 2007; Riva Rossi et al., 2012; Arismendi et al., 2014; Gómez-Uchida et al., 2018). Narum et al., (2017), sugieren que la presión de propágulos y el flujo genético ha influido en la mantención de la diversidad genética en poblaciones de salmón Chinook introducidas, dado que no encontró diferencias significativas entre la diversidad de los núcleos donadores y las poblaciones introducidas. El flujo genético es el intercambio de individuos que puede existir entre dos o más poblaciones, lo que permite incorporar vía migración nuevos alelos no presentes en una determinada población, o bien aumentar la frecuencia de alelos que estaban en baja frecuencia.

La naturaleza anádroma, particularmente la conducta de "*homing*" de la especie también puede jugar un papel en el éxito del salmón Chinook (Correa y Gross, 2008). El *homing* (conducta de volver al sitio natal a reproducirse) favorece el establecimiento inicial al congregarse el retorno de adultos en un determinado punto de introducción, de lo contrario los individuos estarían dispersos dificultando la reproducción y el crecimiento. Aunque el salmón Chinook es altamente preciso en volver a su sitio natal, existen individuos que se desvían, lo cual tiene importantes

consecuencias evolucionarias ya que permite fundar nuevas poblaciones, favorece el flujo genético entre poblaciones y permite recolonizar posterior a eventos de extinción local (McDowall, 2001; Quinn et al., 2001; Pess et al., 2012). Ejemplo de esto son las poblaciones del río Santa Cruz en Argentina, originadas probablemente a partir de las poblaciones del Río Pratt en la Patagonia Chilena (Correa y Gross, 2008; Gómez-Uchida *et al.*, 2018), y las poblaciones de los ríos Rakaia y Rangitata, originados a partir de dispersión natural de individuos del río Waitaki en Nueva Zelanda (Quinn et al., 2001). En el contexto de siembra humana de ovas, el “*straying*” es usualmente catalogado negativamente como una falla en el regreso al hogar, sin embargo, en las poblaciones naturales es una característica evolutivamente crucial que complementa el “*homing*” (Keefer y Caudill, 2014). Quinn (1984), postuló que la expresión de estas estrategias se da en función del equilibrio dinámico de las poblaciones naturales. Actualmente se sabe que la calidad y estabilidad de los hábitats pueden influir sobre las tasas de “*homing*” (Hendry et al., 2004; Cram et al., 2013). Además, la evidencia indica que las actividades pesqueras y el manejo de los ríos incrementa la tasa de “*straying*” (Keefer y Caudill, 2014). Por otra parte, peces de criadero tienen mayor propensión a manifestar desviarse (“*stray*”) con relación a sus conespecíficos silvestres (Björnsson et al., 2011). En salmón Chinook particularmente, se ha visto que los peces adultos cuyos gametos de los padres habían sido transportados a la zona de cría eran menos propensos a volver río-arriba a desovar (“*homing*”) (Hard y Heard, 1999), y los animales híbridos se desvían a tasas tres veces más altas que las de la población natal liberada en el mismo lugar (Candy y Beachman, 2000). La hibridación de las múltiples fuentes que dio origen a las poblaciones en Sudamérica podría explicar en alguna medida el potencial invasor de esta especie en este continente.

2.4 Pre-adaptación fenotípica y genética

Una de las hipótesis para explicar el éxito adaptativo de diferentes especies animales y vegetales es la hipótesis de “pre-adaptación” de caracteres fenotípicos (Riva Rossi et al., 2012; Ciancio et al., 2015; Elst et al., 2016). Caracteres pre-

adaptados son aquellos rasgos que son funcionalmente ventajosos en ambientes no nativos. Recientemente, un estudio en aves sugiere que la selección actúa incluso antes de la introducción sobre una amplia gama de rasgos fisiológicos, morfológicos y conductuales que probablemente afectan el éxito de la invasión (Baños-Villalba et al., 2020). Esta hipótesis ha sido poco estudiada a nivel genético en poblaciones invasoras. En estudio reciente en dos especies de carpas invasoras (*Hypophthalmichthys nobilis* y *H. molitrix*), encontraron genes putativamente bajo selección positiva asociado a rasgos relacionados al ciclo de vida, que presumiblemente se desarrollaron en su rango nativo, y que podrían haber facilitado el establecimiento inicial y posterior invasión (Wang et al., 2020). Actualmente, se desconoce si genes específicos han jugado un rol en la capacidad del salmón Chinook de establecer poblaciones autosustentables fuera de sus rangos nativos. Algunos autores sugieren que el rol de las variantes adaptativas y funcionales es clave para explicar la invasión exitosa (Keller y Taylor, 2008).

Identificar los mismos genes (u *loci*) bajo selección en múltiples comparaciones independientes, indicaría que los mismos genes son ventajosos para múltiples ambientes, y que por lo tanto facilitarían el establecimiento y la colonización de esta especie. Correa y Gross (2008) postularon que la conducta del “*homing*” podría jugar un papel en el éxito del salmón Chinook como invasor. Identificar genes bajo selección que estén asociados a esta conducta brindaría soporte a esta hipótesis.

La introducción y posterior colonización del salmón Chinook en diferentes regiones geográficas del planeta, conjunto a la disponibilidad de recursos genómicos y de métodos analíticos, constituyen una oportunidad única de estudiar la hipótesis de pre-adaptación a nivel genómico en el salmón Chinook. Entre los recursos genómicos de uso público se encuentran los provistos por Hecht et al., (2015), que utilizando una tecnología de secuenciación llamada “RAD-seq”, genotipificaron múltiples poblaciones nativas de Norteamérica, dejando a disposición una librería RAD con 21504 secuencias que contienen un SNP ubicable dentro de estas secuencias. Posteriormente Christensen et al., (2018), publicaron el genoma de

referencia del salmón Chinook, haciendo posible contrastar las secuencias RAD y las posiciones de los SNPs dentro de estas secuencias con el genoma de referencia.

Poseer genotipos de los SNPs ubicados en una misma posición en poblaciones o núcleos genéticos donadores e invasores, ofrece la oportunidad única de identificar *loci* bajo selección en múltiples regiones, y descubrir, de existir, cuáles son los genes que podrían estar siendo seleccionados en las diferentes regiones. Esto puede ser desafiante dado el sesgo de determinación (“*ascertainment bias*”), lo cual se refiere a que, dentro de una especie, SNPs descubiertos en una población no necesariamente se encontrarán en otra, dado principalmente por un muestreo no aleatorio y la estructura de las poblaciones utilizadas para el descubrimiento, además por efecto de la recombinación, y el mestizaje entre linajes genéticos puede incrementar la diversidad genética y producir nuevos genotipos (Kelly et al., 2006; Lachance y Tishkoff, 2013).

2.5 Divergencia y uso de información molecular

Las especies introducidas, a medida que van invadiendo y expandiendo su distribución geográfica, desarrollan cambios en la diversidad genética, estructura genética poblacional y también en rasgos fenotípicos (Keller y Taylor, 2008). Por ejemplo, las poblaciones introducidas a Nueva Zelanda (de salmón Chinook) son de ancestría monofilética, donde originalmente los caracteres fenotípicos eran relativamente homogéneos y con el paso de las generaciones, se observó variación en estos caracteres ligados incluso al ciclo de vida (Quinn et al., 2001).

La divergencia genética se refiere a los cambios genéticos que se evidencian en cuando dos o más poblaciones, provenientes de un núcleo genético común, se reproducen de forma aislada en el tiempo (pudiendo llegar incluso a la especiación), que ocurren bajo influencias de la migración (flujo génico), mutación, del azar (deriva génica) y/o de la selección (Keller y Taylor, 2008). El ambiente también podría influir en los cambios que pudieran evidenciarse, ya que diferentes genotipos podrían

tener un diferente *fitness* en diferentes ambientes. Además, un efecto que se conoce como la interacción genotipo-ambiente hace que un mismo genotipo pueda tener diferentes fenotipos en diferentes ambientes.

Los cambios en poblaciones divergentes pueden ser identificados mediante diferentes análisis que utilizan la información del ADN. Con el desarrollo de tecnologías de secuenciación y la utilización de información del ADN (marcadores genéticos) es posible estudiar las poblaciones, y así revelar mecanismos moleculares que subyacen a los caracteres que favorecen la adaptación a ambientes nativos y no nativos (Barson et al., 2015; Hess et al., 2016). Específicamente, la utilización de información genética como polimorfismos de un solo nucleótido (“SNPs” por sus siglas en inglés), han permitido estudiar la diversidad genética de las poblaciones, la arquitectura genética de diferentes caracteres de interés en diferentes especies e identificar *loci* putativos bajo selección, usando información de poblaciones nativas e invasoras de diferentes especies (Allendorf et al., 2010; Chown et al., 2015). Por ejemplo, se han encontrado *loci* bajo selección relacionados al rasgo tiempo de floración en especies herbáceas invasoras como *Mimulus guttatus* (Puzey y Vallejo-Marín, 2014) y *S. austriacum* (Vandepitte et al., 2014). Sin ir más lejos, se han encontrado *loci* bajo selección relacionados a respuesta inmune entre las poblaciones invasoras de salmón Chinook de Sudamérica (S. Narum et al., 2017). Además, *loci* que regulan la edad de maduración y en consecuencia el tiempo de migración han sido influenciados por selección natural (Hess et al., 2016; Prince et al., 2017) y selección artificial (Ayllon et al., 2015). Además, la información genética que brindan marcadores genéticos como microsatélites o los SNPs, ha permitido trazar el origen genético de las múltiples poblaciones de Sudamérica (Correa y Moran, 2017; Gómez-Uchida et al., 2018).

2.6 Selección natural a nivel molecular

Según Darwin y Wallace (1858), la selección positiva es un fenómeno que ocurre cuando ciertos fenotipos incrementan la probabilidad de los individuos de sobrevivir y de reproducirse (*fitness*), por lo cual prevalecen en el tiempo. En términos genómicos, la selección (natural o artificial) se puede clasificar en al menos tres subtipos: selección positiva (direccional), negativa (purificadora) y balanceada. En otras palabras, son diferentes formas de respuesta ante presiones selectivas, donde cada una altera las frecuencias alélicas que conforman una población en el tiempo (Oleksyk et al., 2010). La selección positiva ocurre cuando una determinada variante (ya sea si surge de mutaciones o existentes previamente) proporciona ventajas adaptativas; la selección negativa ocurre cuando las variantes son desventajosas y tienden a ser eliminadas; la selección balanceada depende de la interacción de los alelos en un locus específico y opera de manera que tiende a la mantención de las frecuencias de ambos alelos en la población, generalmente ocurre en heterocigotos que tienen ventajas adaptativas, por efectos de sobredominancia (Oleksyk et al., 2010; Gouveia et al., 2014; López et al., 2015).

Estos modelos de selección constituyen la base de las metodologías desarrolladas para detectar señales de selección, haciendo uso de la premisa de que eventos selectivos alteran los patrones de variación neutral de forma predictiva, por ejemplo, reduciendo la diversidad de haplotipos, aumentando la proporción de alelos raros o incrementando el grado de diferencias en frecuencias alélicas entre poblaciones dependiendo del tipo de selección (Pritchard et al., 2010).

Dado que la selección genera cambios en aquellas regiones genómicas que contienen variantes que están bajo efecto de selección (o han sido seleccionadas), es posible medir estos cambios analizando el ADN de individuos a nivel poblacional, constituyendo las denominadas “señales” o “huellas” de que hubo o se mantiene un proceso de selección en aquellas regiones (Qanbari y Simianer, 2014). Por ejemplo, cuando ocurre un proceso de selección positiva sobre un determinado locus, se

espera que los *loci* adyacentes a las variantes benéficas se hereden de manera conjunta, resultando en frecuencias alélicas altas no aleatorias en estas regiones específicas del genoma (Goddard et al., 2009). Al respecto, *loci* o genes adyacentes pueden haber llegado a estar juntos por efecto de diversas causas evolutivas. Por una parte, podrían ser un “bloque sinténico”, definido como regiones cromosómicas que contienen genes derivados de un ancestro común (Tang et al., 2011), los que actualmente se postula que estarían funcionalmente relacionados (Thévenin et al., 2014). También como consecuencia de uno o más de los mecanismos de reorganización cromosomal (duplicaciones, deleciones, inserciones, inversiones, traslocaciones, entre otros) (Lupski y Stankiewicz, 2005).

El proceso donde *loci* cercanos (ligados físicamente a los alelos seleccionados son heredados junto a este se denomina “*hitchhiking*” (Fay y Wu, 2000), donde con el pasar las generaciones, resulta en regiones genómicas con una baja variabilidad río arriba y río abajo de la mutación benéfica, con alta homocigosidad regional y altas frecuencias alélicas, lo que constituye un barrido selectivo (“*Selective sweep*”) (Smith y Haigh, 1974; Haasl y Payseur, 2016).

En el modelo clásico de selección positiva, se asume que, si nuevas mutaciones resultan ser ventajosas en un ambiente determinado, los individuos que portan variantes benéficas tendrán una mayor probabilidad de prosperar y de dejar descendencia respecto a los que no portan esta variable, por lo que a medida que pasan las generaciones, los alelos benéficos tienden a aumentar su frecuencia en la población incluso hasta llegar al punto de ser genotipos fijos en la población, es decir que todos los individuos contengan la misma variante benéfica (Sabeti et al., 2006; Pritchard et al., 2010). Este proceso drástico de selección positiva resulta en señales denominadas “*Hard sweeps*” (Akey, 2009; Bank et al., 2014), que por efecto de *hitchhiking*, presentan un alto desequilibrio de ligamiento (LD por sus siglas en inglés, *linkage disequilibrium*) (entre variantes cercanas y la benéfica), una alta frecuencia haplotípica con una variación concomitante reducida (Pritchard et al., 2010). El desequilibrio de ligamiento hace referencia a una asociación no aleatoria

entre dos alelos, y un haplotipo se refiere a una constitución alélica de múltiples *loci* para un cromosoma dado, el cual puede ser heredado en bloque. Por efecto de *hitchhiking*, se generan ligamiento entre el o los sitios seleccionados con sitios aledaños, por lo cual el sitio encontrado bajo selección puede no ser el realmente seleccionado, sino ser encontrado bajo selección por efecto de ligamiento (Kelley y Swanson, 2008).

Otro tipo, más común, de selección positiva deriva en señales denominadas “*Soft sweeps*”, donde la selección actúa sobre la variación preexistente sometida a deriva génica, donde variantes que pudieron haber sido neutrales o incluso deletéreas, se vuelven favorables debido a la exposición a un nuevo ambiente favorable (Hermisson y Pennings, 2005; Pritchard *et al.*, 2010). En este caso, el alelo seleccionado está presente antes del barrido selectivo, y debido a los eventos de mutación y recombinación que ocurren durante la fase de deriva génica, la región que contiene al alelo seleccionado contiene múltiples haplotipos y no solo uno. En consecuencia, la reducción en la diversidad local es menos pronunciada que en el modelo de *hard sweeps*, donde las frecuencias alélicas de variantes neutrales en ligamiento tampoco son excesivamente altas o bajas (Przeworski *et al.*, 2005) y por el contrario, se esperan frecuencias alélicas intermedias, por lo cual son más difíciles de detectar (Teshima *et al.*, 2006; Pennings y Hermisson, 2006). Además, el desequilibrio de ligamiento será máximo en la región seleccionada, decayendo rápidamente en las zonas flanqueantes (Schridder *et al.*, 2015).

Las *soft sweeps* podrían explicar en gran proporción la respuesta a la selección en las poblaciones naturales, debido a que el proceso adaptativo sería limitado si dependiera solo de la tasa de mutación, además, la mayoría de los fenotipos asociados a la adaptación son de naturaleza poligénica, es decir, son influenciados en cierta medida por la varianza genética de muchos *loci* (Pritchard *et al.*, 2010). La adaptación a menudo ocurre vía acción colectiva de muchos *loci* de efecto pequeño (Olson-Manning *et al.*, 2012), y es más rápida cuando la selección opera sobre la variación genética preexistente, no solo porque las variantes alélicas benéficas ya

están disponibles al momento en que se da un cambio en las presiones selectivas, sino porque pueden estar a altas frecuencias en la población debido a que son más viejas que una nueva mutación, y por lo tanto pudieron haber sido sometida a selección en ambientes pasados (Hermisson y Pennings, 2005; Barrett y Schluter, 2008). Otro escenario posible de selección positiva subyace en los llamados barridos “parciales” o “incompletos”, que consisten en que una mutación benéfica no ha alcanzado la fijación en la población, probablemente debido a que el barrido está en curso al momento del muestreo de la población (Voight et al., 2006; Pritchard et al., 2010). Las metodologías basadas en haplotipos han sido utilizadas en la búsqueda de este tipo de señales (Voight et al., 2006; Sabeti et al., 2007).

Resumiendo, estos fueron los primeros conceptos asociados a señales de selección: 1) *Hard sweeps*: donde nuevas mutaciones con un fuerte efecto sobre el fitness tenderían a incrementar su frecuencia alélica, reduciendo fuertemente la diversidad genética a nivel local, baja diversidad de haplotipos y alto desequilibrio de ligamiento entre las variantes aledañas y, 2) *Soft sweeps*: donde los mismos alelos presentes en la diversidad genética existente (“*background genético*”) se vuelven favorables frente a nuevos ambientes, generando una señal con mayor diversidad de haplotipos y menor desequilibrio de ligamiento.

Hasta hace una década atrás, esta clasificación dicotómica de señales (*hard-* y *soft-sweeps*) tenía relación con el origen de la mutación, siendo la primera mutaciones nuevas, y la segunda provenientes de los cambios ambientales que generaban que las variantes pre-existentes (también referido en la literatura como “*standing genetic variation*” (SGV)) se volvieran benéficas sobre el *fitness*. Hermisson y Pennings (2017) describen un barrido selectivo como los cambios a nivel genómico revelados por ciertos patrones en la diversidad (como cambios drásticos en la frecuencia del alelo benéfico) que son causadas por adaptaciones recientes u presión de selección positiva. Según Hermisson y Pennings (2017), el sello característico de un *hard-sweep* es un ancestro en común muy reciente de todos los alelos benéficos en una muestra determinada. Los autores sugieren que es poco probable que la adaptación

se limite a mutaciones nuevas y en un su marco teórico basado en modelos de coalescencia, señalan que es posible detectar la adaptación y señales de selección desde la SGV, ya sea tipo *hard sweep* o *soft sweep*, que puede ser de origen único o de origen múltiple. La variación ancestral previo al establecimiento de la presión selectiva puede mantenerse si existe recombinación entre el sitio polimórfico y el sitio objeto de selección.

2.7 Recombinación

La recombinación es un reconocido factor que puede influir en los patrones de diversidad genética y patrones de desequilibrio de ligamiento en una población (Voight et al., 2006). La recombinación puede afectar los patrones de señales de selección, y viceversa, la selección y los cambios demográficos pueden influir en la estimación de la tasa de recombinación de la población (Weigand y Leese, 2018). El *crossing-over* es un fenómeno de recombinación que ocurre durante la meiosis, donde dos pares de cromosomas homólogos intercambian segmentos de ADN. Como resultado, pueden generarse gametos que contienen segmentos heredados de las dos ramas ascendientes de ese progenitor, introduciendo variabilidad. Se postula que la recombinación en la meiosis permite reorganizar el material genético de ambos padres para aumentar la variación genética en la progenie. Se ha descrito que la tasa y la distribución del *crossing-over* pueden evolucionar para equilibrar los beneficios de la recombinación y los costes del intercambio ectópico (Kent et al., 2017; Brand et al., 2018). Las tasas de recombinación pueden variar dentro y entre especies, sin embargo, la significancia funcional, genes, mecanismos y causas evolutivas de esta variación no está totalmente comprendida (Ritz et al., 2017).

En el contexto de señales de selección, los procesos de *crossing-over* y recombinación cromosómica promueven la ruptura de los segmentos cromosómicos heredados, reduciendo el tamaño de los segmentos cromosómicos homocigotos a lo largo de las generaciones (Kelley y Swanson, 2008). Por lo tanto, la longitud de los segmentos cromosómicos barridos debido a la presencia de un alelo positivo es

inversamente proporcional al número de generaciones de selección, afectando en consecuencia al tamaño de los fragmentos en estado homocigoto (Kelley y Swanson, 2008). La variabilidad genética se reduce de manera más rápida por un barrido selectivo en regiones con bajas tasas de recombinación en comparación con regiones con alta tasa de recombinación (Weigand y Leese, 2018). De conocerse la tasa de recombinación (por ejemplo, las tasas de recombinación por generación en un mapa de ligamiento medido en centimorgans (cM)), algunos tests estadísticos pueden incorporar esta información y así poder detectar de más adecuada la selección positiva (Weigand y Leese, 2018). Sin embargo, generar un mapa de ligamiento requiere de inversiones en términos de tiempo y presupuesto. Esto no siempre puede llevarse a cabo si por diferentes motivos (como por falta de infraestructura o restricciones normativas) la crianza de la especie objetivo no es posible de llevar a cabo, si se desconocen las relaciones de parentesco, o si existen generaciones solapadas no discretas. Independientemente de la recombinación, en la fase final de un “*hard sweep*”, toda la población queda fijada para la mutación ventajosa, pues ambas cromátidas hermanas poseen la misma variante selectiva (Weigand y Leese, 2018).

2.8 Métodos para la detección de señales de selección

La búsqueda de genes que subyacen a los caracteres fenotípicos de interés puede ser abordado mediante un enfoque denominado *bottom-up* (de abajo hacia arriba; es decir, a partir de la información genética inferir o hipotetizar fenotipos), que corresponde a la utilización de datos genéticos (SNPs en este caso) y a la aplicación de análisis estadísticos con el fin de identificar marcadores putativos bajo selección, sin necesidad de contar con un fenotipo registrado en los individuos a evaluar. Es importante considerar que los estudios que utilizan marcadores genómicos distribuidos en el genoma para identificar genes bajo selección podrían estar representando solo una pequeña proporción de los genes o regiones que realmente están involucradas en la respuesta a la selección (Hohenlohe et al., 2010). Además, el tamaño efectivo poblacional, el número de generaciones bajo selección, la tasa de recombinación, la edad relativa de los alelos neutrales ligados, eventos

demográficos poblaciones como cuellos de botellas, efecto de fundadores, además del tipo e intensidad de selección, pueden afectar la identificación de señales de selección (Gouveia et al., 2014; Utsunomiya et al., 2015).

Para detectar señales de selección se han desarrollado diferentes metodologías estadísticas, donde se comparan los datos observados con los esperados bajo la hipótesis nula de selección neutral, planteada por la teoría neutral de la evolución, que establece que la variación genética dentro de una población sigue un equilibrio Hardy-Weinberg y que es causada solo por la deriva génica (Kimura, 1968). Cambios en el espectro de frecuencias alélicas, que no siguen un patrón esperado, o bien un incremento estadísticamente significativo en las frecuencias de un alelo en específico pueden ser consideradas como una huella de selección en ese *loci* particular (Haasl y Payseur, 2016). La elección apropiada de un método estadístico depende de la escala temporal durante la cual se espera que opere la selección (tiempos de especiación o de evolución intraespecífica), el número de poblaciones a estudiar y de la densidad de marcadores disponibles (Gouveia et al., 2014; Weigand y Leese, 2018). La mayoría de las metodologías desarrolladas apuntan a identificar señales de selección positiva (Vitti et al., 2013). En la detección de señales de selección se asume que las altas presiones de selección (natural o artificial) son la principal explicación para los resultados estadísticamente significativos, aun cuando existen eventos naturales como cuellos de botella (la pérdida de diversidad genética por una reducción drástica en el tamaño poblacional) que pueden llevar a que los patrones observados sean similares a una huella de selección.

HIPÓTESIS

Se detectará un pool común de señales de selección positiva entre los múltiples casos de poblaciones divergentes de salmón Chinook a través del mundo.

OBJETIVOS

Objetivo general

Identificar *loci* y genes candidatos putativamente bajo selección positiva en el Salmón Chinook (*Oncorhynchus tshawytscha*), utilizando polimorfismos de un solo nucleótido e índices de diferenciación poblacional.

Objetivos específicos

1. Detectar *loci* atípicos putativamente bajo selección positiva en común a partir de múltiples escaneos genómicos basados en diferenciación poblacional entre poblaciones representativas de los núcleos ancestrales y poblaciones de regiones no nativas alrededor del mundo.
2. Identificar genes candidatos en el genoma de referencia del salmón Chinook en la región donde se encuentra la posición física de los *loci* atípicos encontrados en común para todos los escaneos.

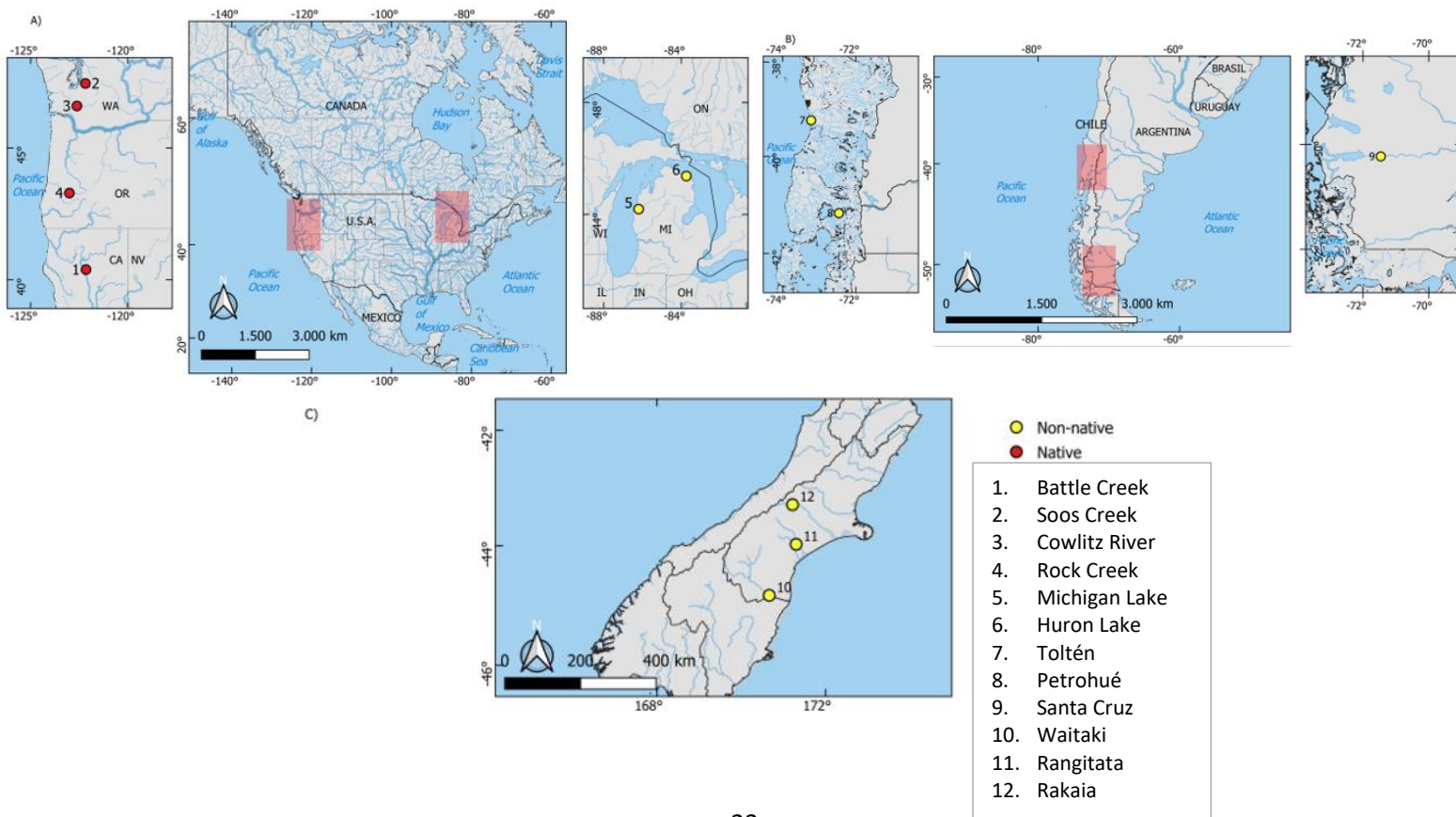
MATERIALES Y MÉTODOS

5.1 Poblaciones donadoras y naturalizadas

Dado el amplio rango de distribución demográfica de las poblaciones norteamericanas y no nativas, además de las dificultades técnicas y económicas relacionadas al muestreo y genotipificación de todas las poblaciones no nativas, se desarrolló un diseño de estudio que utiliza la información disponible publicada (núcleos nativos) y de poblaciones no nativas a genotipificar, gracias a colaboración internacional. Para este estudio se cuenta con poblaciones no nativas de:

- a) Sudamérica: Petrohué, Toltén, Caterina.
- b) Nueva Zelanda: Waitaki, Rakaia, Rangitata
- c) Great Lakes: Huron, Michigan

Figura 1. Regiones geográficas muestreadas. A) Norteamérica; B) Sudamérica y C) Nueva Zelanda. En amarillo las poblaciones no nativas y en rojo las poblaciones nativas donadoras.



El diseño de estudio se basa en comparar la información de genotipos de los núcleos genéticos donadores/ancestrales con los de poblaciones no nativas. Estudios previos han mostrado la ancestría de las poblaciones utilizando información genómica junto a información de registros. Las poblaciones de Sudamérica muestran ancestría polifilética (Riva Rossi et al., 2012; Correa y Moran, 2017; Gómez-Uchida et al., 2018), Nueva Zelanda monofilética (Quinn et al., 2001) y *Great Lakes* monofilética (Peck et al., 1999; Connerton et al., 2009; Suk et al., 2012). En la **Tabla 1**, se muestra el diseño comparativo de esta tesis. Contiene los núcleos ancestrales predominantes de las poblaciones no nativas, el número de SNPs que se espera poder utilizar y la tecnología de secuenciación. Todas las muestras de Sudamérica serán proporcionadas por el Dr. Daniel Gómez-Uchida (GeecLab, Universidad de Concepción, Chile), mientras que todas las muestras de Nueva Zelanda y *Great Lakes* serán proporcionadas gracias a la colaboración la Dra. Mariah Meek (Michigan State University, Estados Unidos).

Tabla 1. Región y ríos de origen de las poblaciones donadoras, ríos invadidos, número de individuos por población, número de marcadores a utilizar y tipo de secuenciación.

Región de origen	Ancestro	Población Invasora	N	Secuenciación
California Central Valley	Battle Creek ^a	Waitaki ^b	42 - 32	RADseq – SeqSNP
California Central Valley	Battle Creek ^a	Rangitata ^b	42 - 29	RADseq – SeqSNP
California Central Valley	Battle Creek ^a	Rakaia ^b	42 - 28	RADseq – SeqSNP
Oregon-California Coast	Rock Creek ^a	Toltén ^c	48 – 50	RADseq – SeqSNP
South Puget Sound	Soos Creek ^a	Petrohué ^c	48 – 60	RADseq – SeqSNP
Lower Columbia River	Cowlitz River ^a	Caterina ^c	45 – 62	RADseq – SeqSNP
South Puget Sound	Soos Creek ^a	Huron Lake ^b	48 – 30	RADseq – SeqSNP
South Puget Sound	Soos Creek ^a	Michigan Lake ^b	48 - 26	RADseq – SeqSNP

^a Genotipos disponibles proporcionados por Hecht *et al.*, (2017). ^b ADN genómico obtenido proporcionado por Dr. Thomas P. Quinn, la Dra. Mariah H. Meek y el Dr. Daniel Gómez Uchida. Estas muestras serán enviadas a la empresa *LGC Biosearch Technologies*, quienes a través del uso de tecnología de secuenciación dirigida (SeqSNP) pueden proporcionar los genotipos para los marcadores descubiertos por Hecht *et al.*, (2015).

5.2 Obtención de datos genómicos

La metodología *RAD-seq* ha ganado popularidad durante los últimos años, debido a su utilidad en especies sin información genética previa, ya que permite identificar miles de SNPs sin necesariamente contar con un genoma de referencia (Davey y Blaxter, 2010; Andrews et al., 2016). Consiste en la utilización de enzimas de restricción que cortan el ADN en sitios específicos, y resultan en la obtención de secuencias (*RAD-tags*), donde a partir de estas, mediante análisis bioinformático, se descubren sitios polimórficos (SNPs) dentro de secuencias conservadas. Los protocolos de extracción de DNA, de preparación de librerías, descripción de los análisis para el descubrimiento de SNPs y las etiquetas RAD están disponibles en Narum et al., (2017) y Hecht et al., (2015). El cromosoma al cual pertenece cada marcador y la posición física de estos, en términos de par de bases (*base pairs: bp*), se obtendrán comparando las secuencias con el genoma de referencia de *O. tshawytscha* proporcionado por Christensen et al., (2018), utilizando BLASTn de NCBI (Camacho et al., 2009).

Una vez obtenidas las posiciones de los marcadores y sus cromosomas correspondientes, se procederá realizar una secuenciación dirigida de las muestras no nativas. La metodología SeqSNP permite la secuenciación de regiones específicas del genoma que contienen los marcadores previamente identificados. Este tipo de secuenciación identifica de manera precisa las variantes nucleotídicas presentes en loci específicos. Se ajusta al presente proyecto debido a que podré disponer de marcadores identificados previamente y utilizarlos para realizar los análisis. Este servicio es brindado por la empresa *LGC Biosearch Technologies*. Es posible que una menor cantidad de SNPs sean descubiertos en las poblaciones no nativas, debido al sesgo de determinación y el mestizaje (*admixture*) entre poblaciones (Kelly et al., 2006; Lachance y Tishkoff, 2013).

5.3 Control de calidad de la información genómica

Una vez obtenidos los datos de genotipos se procederá a realizar un control de calidad en el software PLINK v.1.9 (Purcell et al., 2007), donde mediante línea de comandos se eliminarán individuos cuya tasa de llamado sea menor al 70%, y aquellos marcadores cuya tasa de llamado sea menor al 70%. La tasa de llamado de individuos es un criterio que elimina individuos si contiene menor que el 70% de todos los SNPs correctamente genotipificados. La tasa de llamado de marcadores significa que esos SNPs deben estar en al menos 70% de los individuos correctamente genotipificados, es decir, sin valores faltantes (“*missing values*”). Para estudios de mapeo fino de caracteres de interés, por defecto se utiliza una tasa de llamado individual y de SNPs del 90% como umbral (Anderson et al., 2010). Sin embargo, el diseño de esta tesis considera poblaciones altamente divergentes, por lo cual se espera que hayan SNPs que no sean correctamente genotipados en todas las poblaciones (sesgo de determinación). Diferentes estudios han utilizado una tasa de llamado de hasta un 70% (Von Thaden et al., 2017; Wright et al., 2019).

Posteriormente, en cada población se removerán SNPs cuya frecuencia del alelo menor (MAF) sea menor a 1% y aquellos SNPs que se desvíen del equilibrio Hardy-Weinberg (HWE) corrigiendo según Bonferroni (se divide el valor de significancia (0,05) por el número de SNPs que pasan el control de calidad MAF) (Marees et al., 2018). Este tipo de control de calidad también es utilizado en otros estudios de asociación genómica previamente (Sobota et al., 2015; Kuo, 2017). Para eliminar un marcador, este debe ser eliminado en todas las poblaciones del set de datos por el mismo criterio. Es decir, si en el set de datos RAD-seq, en la población A un marcador tiene una MAF $<0,01$, pero $>0,01$ en la población B y C, este pasa el control de calidad ya que podría ser indicativo de una señal de selección y no un error de genotipificación.

Para evaluar la presencia de individuos duplicados entre los datos, se estimarán los valores de identidad por estado (*Identity by state*; IBS) entre pares de individuos

utilizando PLINK v.1.9 (Purcell *et al.*, 2007). La identidad por estado se define como el número de alelos en común para cada SNP, y se estima entre pares de individuos. Posteriormente el programa identifica el grado de ancestría reciente (*Identity by descent*; IBD), a partir de los valores a nivel genómico de IBS. Posteriormente se construye una estadística de resumen “PI_HAT”, que reconoce gemelos idénticos, o muestras duplicadas cuando su valor es 1.

5.4 Estructura poblacional y diferenciación genética

Se evaluará la diversidad genética en términos de heterocigosidad observada (H_o) y esperada (H_e), y el coeficiente de endogamia F utilizando PLINK v.1.9 (Purcell *et al.*, 2007). Para evaluar diferencias en la estructura poblacional se utilizarán dos métodos de agrupación. Se realizará un análisis de componentes principales (PCA) utilizando PLINK v.1.9 (Purcell *et al.*, 2007) y se graficará utilizando el software R (R core team, 2020). Además, se utilizará el programa ADMIXTURE v1.3.0 para inferir las proporciones de ancestría individual (Alexander *et al.*, 2009; Alexander y Lange, 2011). Este programa utiliza un método de agrupamiento (*clustering*) basado en modelos de máxima verosimilitud a partir de las frecuencias alélicas de los SNPs, siendo de mayor velocidad que el método bayesiano implementado en STRUCTURE v.2.3.4 (Pritchard *et al.*, 2000) cuando se tiene una gran cantidad de SNPs e individuos. Se utilizarán parámetros como 200 re-muestreos (*bootstraps*), variando el número de linajes ancestrales (K) desde 1 a 20. Se utiliza un método de validación cruzada (*Ten-fold cross validation*), y se retendrán los K con el menor error de validación cruzada. Esto será con SNPs que hayan pasado el control de calidad compartidos entre todas las poblaciones y con SNPs cuyo valor de desequilibrio de ligamiento sea a lo más 0.4, de tal forma de minimizar los efectos del desequilibrio de ligamiento en los patrones de estructura genética (Kaeuffer *et al.*, 2007). Para estimar el índice de diferenciación poblacional F_{ST} (Weir & Cockerham, 1984) entre todas las poblaciones se utilizará el paquete de R StAMPP (Pembleton, Cogan, y Forster, 2013). La significación se evaluó utilizando 10.000 permutaciones.

5.5 Métodos a utilizar para detectar *loci* bajo selección

Según Qanbari y Simianer, (2014), los métodos se pueden clasificar según la utilización de la información: (1) Estadísticos que utilizan información genómica entre poblaciones y (2) estadísticos que utilizan información genómica dentro de cada población, pudiendo basarse en patrones de desequilibrio de ligamiento, tamaño y frecuencia de haplotipos, diferenciación genética y/o en las frecuencias alélicas. En este trabajo se utilizarán métodos que se basan en la diferenciación genética de poblaciones. Se utilizarán herramientas bioinformáticas para la detección de loci que muestren patrones atípicos (“*outlier*”), es decir, que no siguen un modelo de evolución neutral. En esta tesis se buscarán señales de fuerte selección positiva (*hard sweeps*), debido a distintas razones. Entre ellas se encuentran las siguientes:

- 1) La línea base de variación genética que se utilizará en esta tesis corresponde a genotipos descubiertos a partir de secuenciación RAD en un gran número de poblaciones (46). Si bien son poblaciones relativamente contemporáneas con las poblaciones no nativas, podemos asumir que la variación genética presente disponible en todas las poblaciones constituye un pool representativo de la variación genética existente previo al evento de divergencia. En concordancia con Hermisson y Pennings (2017), será posible detectar este tipo de señales de selección que surgen a partir de la variación genómica existente previo a los eventos de divergencia.
- 2) Poblaciones lejanas geográfica o genéticamente pueden haber sufrido eventos de reordenamiento genético, eventos de deriva génica, cuellos de botella y otros, que podrían afectar la calidad de secuenciación, haciendo que loci se encuentren en ciertas poblaciones y en otras no. En consecuencia, no se espera un gran número de marcadores como para poder realizar análisis para detectar señales basadas en desequilibrio de ligamiento y/o en diversidad de haplotipos, ya que estos requieren una gran densidad de marcadores.

A continuación, se describirán las dos aproximaciones que serán utilizadas en la presente tesis. La justificación de la elección de cada método se describirá junto al método.

5.5.1 Índice de diferenciación genética F_{ST}

El índice F_{ST} de Wright (S. Wright, 1951), ha sido uno de los más utilizados para la detección de regiones genómicas que han sido objeto de selección (Qanbari et al., 2011). Tiene directa relación con el grado de diferenciación genética, en términos de frecuencia de alelos entre poblaciones y una relación indirecta con el grado de semejanza en los individuos dentro de una población. Sus valores fluctúan entre 0 y 1, donde valores bajos de F_{ST} indican que las frecuencias de los alelos son similares entre poblaciones y cuando el valor es alto significa que las frecuencias alélicas entre poblaciones son diferentes, por lo tanto, refleja el grado de diferenciación entre poblaciones en cualquier *locus* dado. La selección negativa y estabilizadora tienden a disminuir los valores de F_{ST} y la selección positiva tiende a incrementar estos valores (Barreiro et al., 2008). Una derivación de esta medida es aquella que compara las varianzas de las frecuencias alélicas dentro de cada población y entre poblaciones, donde se puede entender como una medida de la proporción de la variación genética debido a diferencias en las frecuencias alélicas entre y dentro de las poblaciones (Holsinger y Weir, 2009).

$$F_{ST} = \frac{\sigma_S^2}{\sigma_T^2} = \frac{\sigma_S^2}{\bar{p}(1 - \bar{p})}$$

Donde \bar{p} es la frecuencia media de un alelo en el total de la población, σ_S^2 es la varianza en las frecuencias alélicas en las diferentes subpoblaciones, y σ_T^2 es la varianza de las frecuencias alélicas en la población total.

La distribución de los valores de F_{ST} pueden ser evaluados respecto a su ajuste a la distribución de X^2 y comparar la varianza de esta distribución respecto a la teórica esperada. Si un alelo de un determinado locus es favorecido por la selección, el valor de F_{ST} para ese locus será más alto que en los demás loci, donde la diversidad

genética es causada principalmente por la deriva génica. La deriva afecta a todos los loci del genoma por igual, pero loci bajo selección por lo general revelan patrones atípicos de variación, donde la utilización de marcadores SNPs permite identificar estos loci calificados como atípico u “*outlier*” de la distribución empírica de F_{ST} (Akey et al., 2002). Este método es sensible a los procesos que afectan la distribución de la diversidad nucleotídica, por ejemplo, cuellos de botella o efectos fundadores (Storz, 2005).

Para corregir en cierta medida la tasa de falsos positivos, se han incorporado modelos o escenarios ecológicos realistas en el software BayeScan v.2.1., donde a partir de los valores de F_{ST} para cada *locus*, se estima la probabilidad posterior de que un alelo en particular esté bajo selección positiva. Basado en modelos de regresión logística, la metodología intenta explicar los valores de F_{ST} observados, explorando parámetros de efectos población-específicos (β) y locus-específicos (α). Para loci neutrales, los efectos β son suficientes para explicar el patrón de diferenciación poblacional, mientras que para loci bajo selección es necesario el efecto α . Finalmente, se calcula la razón de probabilidad posterior, es decir, se divide la probabilidad de observar el valor F_{ST} dado un modelo de selección positiva en la probabilidad de haber observado el mismo valor dado un modelo neutral (Foll y Gaggiotti, 2008). Posteriormente se obtiene un indicador (“*Bayes Factor*”) de cuánto más probable es el modelo de selección respecto al modelo neutral.

Entre las ventajas de utilizar esta metodología para nuestro estudio, se encuentran los bajos niveles de error tipo I y II (Narum y Hess, 2011), no requiere un mínimo de marcadores para detectar selección, además asume que las poblaciones en estudio comparten un pool de alelos ancestrales en común, es decir, que provienen de un mismo núcleo genético, lo cual se ajusta en cierta medida al diseño de estudio de esta tesis. Los test basados en F_{ST} tienen una baja tasa de error cuando las muestras provienen de poblaciones que han divergido independiente de su núcleo ancestral (Lotterhos y Whitlock, 2014). Además, el *software* permite manipular las

prioridades a priori (“*prior odds*”) de los modelos neutrales, permitiendo al usuario ser más estricto en la calificación para cada SNP de estar bajo selección.

Utilizar solo una comparación entre solo dos poblaciones puede enmascarar el efecto de la estructura poblacional subyacente, y llevar a encontrar falsos positivos (*loci* detectados bajo selección que fueron simulados como neutrales) (Excoffier et al., 2009). Utilizar múltiples poblaciones podría aumentar la veracidad de la identificación de estas señales (Hohenlohe et al., 2010). Para ser más estrictos en calificar a un SNP bajo selección, se utilizará un método de forma concomitante descrito a continuación.

5.5.2 PCAdapt

PCAdapt es un *software* y paquete de R, desarrollado para identificar señales de adaptación local basado en análisis de componentes principales (PCA), siendo particularmente útil cuando se espera encontrar subestructura o mezcla de núcleos genéticos (*admixture*) en las poblaciones a analizar (Luu, Bazin y Blum, 2017).

Primero, mediante un PCA se identifican los K componentes principales que describen correctamente los datos. Posteriormente utiliza una prueba estadística basada en la distancia de *Mahalanobis* (D), que es una aproximación multidimensional para medir cuan distante se encuentra un punto respecto a la media. La identificación de un *loci outlier* (SNP bajo selección positiva) se basa en el vector de valores Z (*z-scores*) obtenidos cuando se regresan los SNPs sobre K número de componentes principales.

Siendo $z^j = (z_1^j, \dots, z_k^j)$ el vector de K *z-scores*, entre el j° SNP y el primer K componente principal, la distancia de *Mahalanobis* al cuadrado se define como:

$$D_j^2 = (z_j - \bar{z})^T \Sigma^{-1} (z_j - \bar{z})$$

Donde Σ es la matriz (K x K) de covarianza estimada de los z-scores; \bar{z} es el vector de z-scores entre cada SNP y los primeros K componentes principales. Luego, al dividir por un factor de inflación genómica ($\frac{D_j^2}{\lambda}$) se obtienen valores que siguen una distribución de chi cuadrado (χ^2) con K grados de libertad, donde se pueden computar valores que siguen esta distribución y los valores atípicos. Este método se ajusta al diseño del estudio dado que detecta SNPs corregidos por estructura poblacional, permite analizar grupos de poblaciones en específico y considera que puede haber individuos mestizos, que son comunes en poblaciones invasoras en general y en el salmón Chinook (como el caso de las poblaciones de Sudamérica).

5.6 Escaneos genómicos

Para identificar señales de selección entre pares de poblaciones se utilizarán dos métodos complementarios implementados en los *softwares* BAYESCAN v.2.1. (Foll y Gaggiotti, 2008), y PCAdapt en el paquete de R “pcadapt” (Luu, Bazin y Blum, 2017). Para BayeScan realizarán 8 análisis independientes utilizando parámetros conservadores, un *prior odds* de 100, 5000 iteraciones quemadas (*Burn-in*) seguidas por 10.000 iteraciones con un intervalo de adelgazamiento (*Thinning Interval*) de 10 iteraciones. Para considerar un SNP bajo selección positiva, debe tener un valor α positivo, valor $q < 0,05$. Estos son los criterios para considerar, de acuerdo con la interpretación de Jeffreys (revisada por Robert et al., 2009), un Bayes Factor > 32 ($\log_{10} \geq 1,5$) como umbral. Este valor indica que los SNPs están bajo una fuerte selección divergente (positiva).

Para PCAdapt se realizarán las mismas 8 comparaciones independientes. Primero se ejecutará el análisis de componentes principales para determinar el porcentaje de varianza explicada por cada componente. Luego se define el valor óptimo de K basado en el valor más alto de K, posterior a la estabilización de la curva de proporción de varianza explicada por cada K (Luu, Bazin y Blum, 2017). Luego para cada SNP, PCAdapt realiza el *test de Mahalanobis* donde los SNPs calificados

como *outliers* obtienen un valor $p < 0,05$. Adicionalmente, se utilizará una corrección de la distribución de valores p para definir un umbral de detección de *outliers* más conservador, utilizando una tasa de falsos descubiertos de 0,05 utilizando el paquete de R *qvalue* (J.D. Storey et al., 2020).

Cuando se simulan datos para probar la eficacia de los métodos estadísticos, en general se utilizan 3 conceptos para evaluar los resultados: a) Tasa de verdaderos positivos, b) tasa de falsos positivos y c) tasa de falsos descubiertos (*False discovery rate*; FDR). La primera se refiere a los loci correctamente identificados bajo selección cuando han sido simulados bajo selección. Se toman aquellos correctamente identificados dividido por el total de loci bajo selección simulados. La segunda hace referencia a la proporción de falsos positivos, es decir, loci encontrados bajo selección habiendo sido simulados como neutrales. Se dividen los falsos positivos en el total de loci calificados como neutrales. La tasa de falsos descubiertos corresponde a la relación entre los verdaderos positivos y los falsos positivos, dividiendo los falsos positivos entre todos los loci encontrados bajo selección. Por ejemplo, si simulan 100 loci bajo selección y 900 neutrales, y se encuentran 100 verdaderos positivos, la tasa de verdaderos positivos sería 1, por el contrario si se encuentran 45 falsos positivos, la tasa de falsos positivos sería de $45/900 = 0,05$. Si se dan ambos casos, aproximadamente uno de cada tres loci significativos serán verdaderos positivos, dado los 45 loci falsos positivos divididos por los 145 loci en total significativos bajo selección.

Cuando se realiza una prueba estadística y se obtiene un valor p (usualmente $p < 0,05$), significa la probabilidad de aceptar la hipótesis nula siendo que realmente es falsa, es decir, la probabilidad de que sea un falso positivo (5% para un $p < 0,05$). Cuando se realizan múltiples pruebas, como es el caso de la detección de BayeScan o PCAdapt para cada SNP, el valor de 5% puede resultar en un número mayor de falsos positivos. Para abordar este problema, una alternativa corresponde a corregir el valor p por una tasa de falsos descubiertos. La tasa de falsos descubiertos controla el número de falsos positivos en aquellas pruebas que encuentran un valor

significativo. Un valor p ajustado por tasa de falsos descubiertos (también llamado valor q) de 0.05, implica que un 5% de aquellos test significativos resultarán ser falsos positivos (Storey et al., 2005; Storey et al., 2020).

Posteriormente mediante códigos de R, se identificarán los SNPs en común catalogados bajo selección por ambas metodologías y para todas las comparaciones independientes. Mediante un diagrama de *Venn* se graficará el número de *outliers* de acuerdo con la geográfica. Los genes candidatos se buscarán en las regiones colindantes (500 kb (kilo bases) río arriba y río abajo) de las posiciones de aquellos SNPs identificados en común por ambas metodologías en todas las comparaciones. Este criterio y número de kb ha sido utilizado en otros estudios de asociación, dado que puede que el SNP descubierto sea detectado bajo selección por efecto de ligamiento (Castaño-Betancourt et al., 2016; Ciuculete et al., 2017; Justice et al., 2017; Barria et al., 2019).

RESULTADOS

6.1 Genotipificación y control de calidad

En total, 317 individuos correspondientes a ocho poblaciones no nativas fueron genotipificados utilizando el método de secuenciación dirigida SeqSNP (en adelante, “set no nativo”). La secuenciación SeqSNP genera múltiples secuencias de ADN, llamadas lecturas, que deben ser adecuadamente limpiadas y filtradas antes del llamado de variantes, donde se identifica la identidad de los alelos y se determina la posición de los sitios polimórficos. El primer paso fue realizar un *demultiplexing*, que consiste en remover los “códigos de barra” o “*barcodes*” que son secuencias ancladas a las lecturas para identificar a qué muestra corresponde cada lectura. Una vez limpiadas las secuencias se obtuvo ~600 millones de lecturas de 75 pares de bases (pb) de largo, con una media de ~2 millones de lecturas por individuo. Para hacer el llamado de variantes, *LGC genomics* recomienda utilizar el software *Freebayes*, con el cual se logró identificar ~1,2 millones de sitios polimórficos entre todas lecturas de las muestras no nativas. De estos, se logró identificar un total de 8,268 marcadores que compartían una posición en común con el conjunto de datos RAD-seq (en adelante, “set donador”). Los errores de genotipificación por tasa de llamado afectan a lotes de secuenciación, y por esta razón el filtrado de marcadores bajo los criterios de tasa de llamado individual y de marcadores fueron realizados por conjunto de datos, es decir, utilizando los 1958 individuos secuenciados del set donador y los 317 individuos del set no nativo de forma independiente. Posteriormente, se filtraron marcadores bajo criterios MAF y HWE (mencionados en métodos) dentro de cada población.

El resumen del control de calidad (animales y marcadores filtrados según cada criterio y población), localización de la obtención de muestras, indicadores de diversidad genética y de consanguinidad según información genética disponible se muestran en la **Tabla 2**. Brevemente, no se eliminaron individuos ni marcadores bajo el criterio de tasa de llamado en el set donador, mientras que en el set no nativo

se eliminaron 559 SNPs y 17 individuos. Dentro de las poblaciones del set donador, las poblaciones con el mayor y menor número de marcadores con $MAF > 0,01$ y en HWE fueron Rock Creek y Battle Creek, respectivamente. Dentro de las poblaciones del set de datos no nativo, la población con el mayor número de marcadores con $MAF > 0,01$ y en HWE fue la del río Petrohué, mientras que la menor fue la del río Rakaia en Nueva Zelanda.

Dentro del set de datos no nativo, 638 y 54 SNPs fueron eliminados por los filtros de MAF y HWE, respectivamente, en todas las poblaciones; en consecuencia 7.018 SNPs que pasaron el control de calidad. Dentro del set de datos donador, 1.841 y 0 SNPs fueron eliminados MAF y HWE respectivamente, obteniendo un total de 17,178 SNPs que pasaron el control de calidad en este set. Al comparar ambos sets, 5.460 SNPs se ubicaban en la misma posición física dentro de un determinado cromosoma o *scaffold* (segmento de nucleótidos no alineado a ningún cromosoma, pero amplificadas de forma consistente en ensayos RAD)). Al mezclar ambos sets de datos, se identificaron 703 SNPs tri- o tetra- alélicos, los cuales fueron eliminados dejando un set de 4.757 SNPs en común, de los cuales 4.529 se ubican dentro de cromosomas. Finalmente, 483 individuos correspondientes a 12 poblaciones (4 donadoras – 8 no-nativas) de 3 diferentes continentes pasaron el control de calidad y quedaron disponible para los análisis subsecuentes.

Tabla 2. Descripción de resultados del control de calidad, ubicación de las poblaciones e indicadores de diversidad genética.

Set –	N° SNP	Población	Región ^Δ	Lat	Lon	<i>n</i> ¹	<i>n</i> _{QC} ²	Call-rate ³	MAF > 0.01 ⁴	HWE ⁵	<i>H</i> _O ⁶	<i>H</i> _E ⁷	<i>F</i> _{IS} ⁸
									n	n			
RAD-seq	- 19,019	Battle Creek (BAT)	NA	40.4006	-122.1444	42	42	19019	10329	10180	0.14 ± 0.11	0.25 ± 0.16	0.42±0.08
		Soos Creek (SOO)	NA	47.3009	-122.1751	48	48	19019	13239	13219	0.23 ± 0.17	0.24 ± 0.16	0.03±0.04
		Cowlitz (COW)	NA	46.5101	-122.6151	45	45	19019	13633	13614	0.17 ± 0.13	0.23 ± 0.16	0.23±0.13
		Rock (ROC)	NA	43.334	-123.003	48	48	19019	15196	15169	0.18 ± 0.14	0.22 ± 0.16	0.17±0.06
SeqSNP	- 8,268	Waitaki (WTK)	NZ	-44.8372	170.6647	32	31	7709	3923	3621	0.24 ± 0.19	0.24 ± 0.16	0.01±0.03
		Rangitata RTT)	NZ	-43.973	171.3016	29	28	7709	3820	3572	0.24 ± 0.19	0.25 ± 0.16	0.02±0.04
		Rakaia (RKA)	NZ	-43.2946	171.2161	28	28	7709	3693	3458	0.25 ± 0.19	0.25 ± 0.16	0.01±0.04
		Toltén (TLT)	SA	-39.2434	-73.2201	50	49	7709	5437	4959	0.22 ± 0.18	0.22 ± 0.16	0.01±0.05
		Petrohué (PTH)	SA	-41.1791	-72.459	60	54	7709	5790	5371	0.21 ± 0.18	0.21 ± 0.16	0.02±0.08
		Santa Cruz (STC)	SA	-50.2333	-71.456	62	54	7709	4074	3673	0.22 ± 0.19	0.24 ± 0.17	0.05±0.19
		Michigan Lake (MCG)	NA	44.1976	-86.1941	26	26	7709	4931	4700	0.24 ± 0.20	0.24 ± 0.15	-0.01±0.14
		Huron Lake (HRN)	NA	45.4028	-83.7348	30	30	7709	4931	4634	0.31 ± 0.27	0.25 ± 0.16	-0.27±0.21

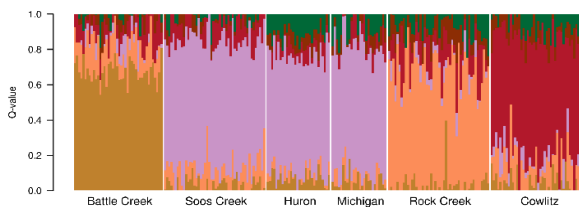
Notas: ¹ Numero de animales genotipificados; ² Número de animales que pasaron el control de calidad; ³ SNPs que pasan el control de tasa de llamado dentro de cada conjunto de datos; ⁴ SNPs con MAF > 0.01; ⁵ SNPs en equilibrio Hardy-Weinberg (corrigiendo según Bonferroni); ⁶ Heterocigosidad observada; ⁷ Heterocigosidad esperada; ⁸ *F*_{IS}: estimación de la consanguinidad según el índice de fijación de Wright (1950). ^Δ Norteamérica (NA) – Sudamérica (SA) – Nueva Zelanda (NZ).

6.2 Análisis de estructura poblacional

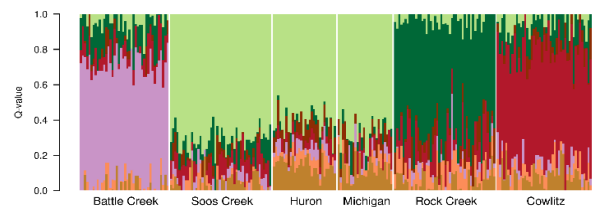
El análisis de agrupación “*clustering*” poblacional fue realizado en el software ADMIXTURE utilizando un set de 668 SNPs neutrales y filtrados según desequilibrio de ligamiento (r^2), seleccionando aquellos que con un r^2 máximo de 0,4 en al menos una población. Los valores de error de validación cruzada (CV-error) sugieren que el número de linajes o grupos genéticos que describe a las muestras es $K = 7$, ya que este K posee el menor error (**Tabla 3**). Sin embargo, existe una mínima diferencia entre $K = 7$ y $K = 6$ (0,00002), lo cual sugiere una dificultad para que el programa encuentre un número único que describa la ancestría de las poblaciones del set de datos. La **Figura 2** muestra la estructura poblacional para $K = 6$ y $K = 7$.

Figura 2. Gráfico de barras realizado a partir de los valores Q de ADMIXTURE. Cada barra corresponde a un individuo, y cada color corresponde a la proporción Q de cada *cluster* (K). Las líneas blancas separan las poblaciones. A – B – C corresponde a $K = 6$ y D – F – G a $K = 7$.

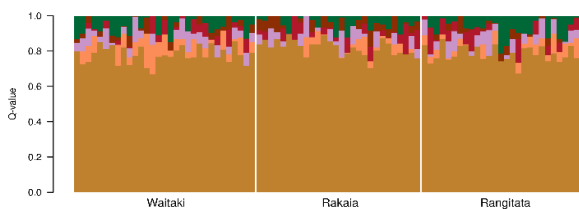
A) Norteamérica para $K = 6$



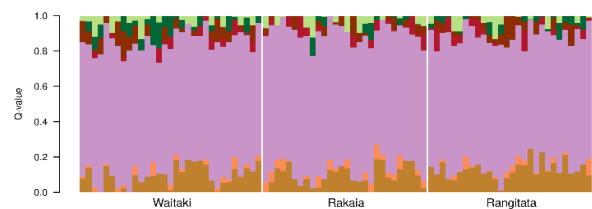
D) Norteamérica para $K = 7$



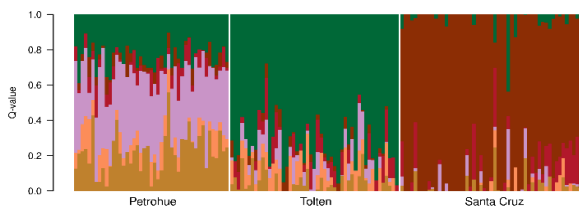
B) Nueva Zelanda para $K = 6$



E) Nueva Zelanda para $K = 7$



C) Sudamérica para $K = 6$



F) Sudamérica para $K = 7$

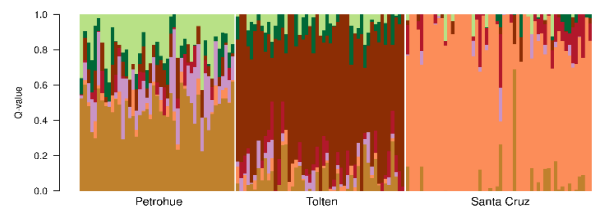


Tabla 3. Error de validación cruzada para cada *cluster* (K) estimada según ADMIXTURE.

K=1	0.70281	K=11	0.64071
K=2	0.67383	K=12	0.64892
K=3	0.64733	K=13	0.64644
K=4	0.63854	K=14	0.64825
K=5	0.63309	K=15	0.66343
K=6	0.63179	K=16	0.65615
K=7	0.63177	K=17	0.66701
K=8	0.63455	K=18	0.67488
K=9	0.63497	K=19	0.68250
K=10	0.64032	K=20	0.68862

Bajo ambos escenarios, las poblaciones neozelandesas son altamente similares en su estructura genética con la población donadora (BAT) (Fig. 2A y 2D). Lo mismo ocurre con las poblaciones HRN y MCG con su población donadora (SOO) en los mismos escenarios. Para ambos escenarios, PTH comparte los mismos *clusters* que su principal ancestría donadora (SOO), pero con variaciones en la proporción de estos (**Fig. 2C y 2F**). Para K = 6, la población TLT comparte ancestría con la población representativa donadora (ROC) con variación en la proporción del *cluster* verde (**Fig. 2C**), mientras que para K = 7, esta población constituye un *cluster* genético muy poco representado en los demás grupos (**Fig. 2F**). STC comparte en una muy baja proporción ancestría con su población donadora representativa (COW). La semejanza o distancia genética en términos de ancestría o proporción de grupos genéticos realizada en ADMIXTURE es concordante con los valores de F_{ST} , donde STC y COW son altamente diferenciadas, y el resto de los pares de casos donador – no nativo son altamente similares en estructura, es decir, poco diferenciadas.

6.3 Análisis de diferenciación poblacional

Los valores de F_{ST} estimados entre los pares de poblaciones analizados fueron relativamente bajos, excepto para COW - STC ($F_{ST} = 0,12$ para el conjunto neutral, **Tabla 4**). Estos valores son consistentes con el análisis de agrupación que muestra

una marcada diferencia entre las proporciones ancestrales en ambos K. Esto discrepa de los análisis de asignación y los registros existentes que sugieren que la población de origen más probable de STC es el río Columbia inferior - Willamette representado aquí por la población COW (Ciancio et al., 2015; Gomez-Uchida et al., 2018; Riva Rossi et al., 2012). La combinación de una fuerte deriva genética y eventos de cuello de botella puede haber causado el patrón observado, de lo contrario, la fuente putativa de Santa Cruz podría estar mal representada. No obstante, los análisis de asignación y los registros históricos sugieren que el estudio actual dio cuenta de gran parte de la variación genómica que contribuyó a la fundación de cada una de las poblaciones no nativas utilizadas aquí. Además, en algunos escenarios ecológicos, dos poblaciones pueden pertenecer a una población ancestral común, pero los cuellos de botella hacen que el software estime proporciones de mezcla completamente diferentes (Lawson, van Dorp y Falush, 2018).

Tabla 4. Diferenciación poblacional (de acuerdo al índice F_{ST}) de a pares utilizando a) el set de datos completo (FULL; 4.529); b) set de datos neutral (4.151 SNPs) y C) set de datos atípicos (378 SNPs). Todos los valores estimados son significativamente distintos de 0 ($p < 0.0001$).

CASO	A (FULL)	B (NEUTRAL)	C (OUTLIER)
BAT – WTK	0,058	0,023	0,309
BAT – RTT	0,066	0,028	0,327
BAT – RKA	0,067	0,031	0,325
ROC – TLT	0,076	0,052	0,261
SOO – PTH	0,07	0,042	0,286
COW – STC	0,17	0,126	0,419
SOO – HRN	0,044	0,015	0,271
SOO – MCG	0,038	0,009	0,266

6.4 Escaneos genómicos

El número de *outliers* encontrado por *BayeScan* (PO = 10, 30 and 100) y PCAdapt, junto al número de *clusters* (K) utilizado por este último para detectar los marcadores atípicos se encuentran en la **Tabla 5**. Cuando se estableció un PO = 10, se detectó un total de 5 marcadores con un patrón atípico de diferenciación en todos los casos estudiados, mientras que al utilizar un PO = 30 se encontraron 3 y con PO = 100 se encontraron 2 (**Tabla 6; Figura 3**). Los gráficos de Manhattan muestran la distribución cromosomal de los marcadores significativos para cada análisis realizado (**Figura 3 y 4**).

Figura 3. Gráficos de Manhattan contruidos con resultados de *BayeScan* (asumiendo PO = 10). Muestra las regiones genéticas donde se detectaron señales de respuesta a la selección divergente. Cada gráfico corresponde a un caso independiente de poblaciones evolutivamente divergentes en el salmón Chinook. Cada punto es un SNP. El eje X muestra los cromosomas (impares en rojo y pares en azul), el eje Y muestra el umbral de significancia establecido a partir del $-\log_{10}$ del valor q (“*q-value*” = 0,05).

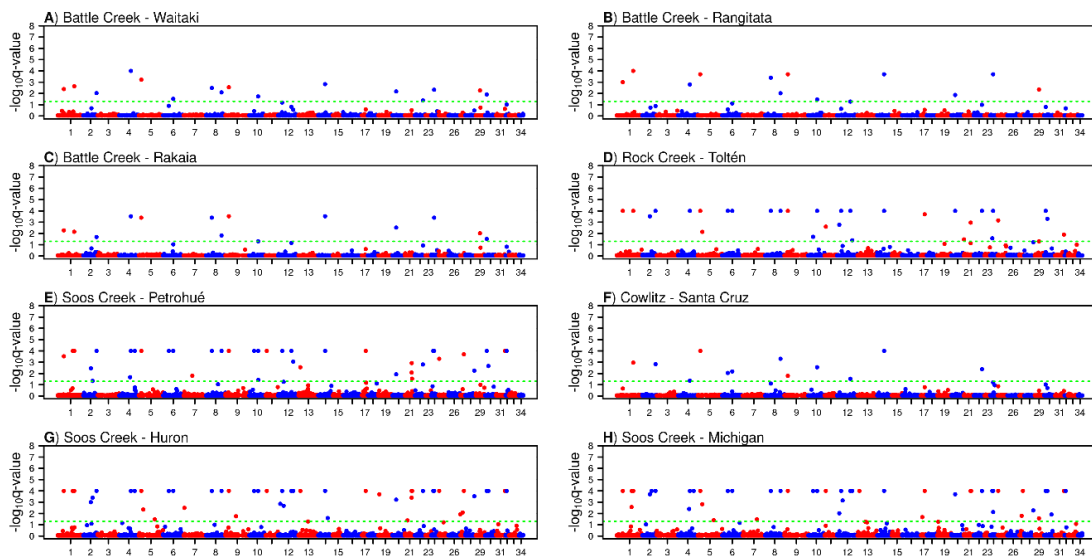


Figura 4. Gráficos de *Manhattan* contruidos con resultados de *PCAdapt*. Muestra las regiones genéticas donde se detectaron señales de respuesta a la selección divergente. Cada gráfico corresponde a un caso independiente de poblaciones evolutivamente divergentes en el salmón Chinook. Cada punto es un SNP. El eje X muestra los cromosomas (impares en rojo y pares en azul), el eje Y muestra el umbral de significancia establecido a partir del $-\log_{10}$ del valor q (“*q-value*” = 0,05).

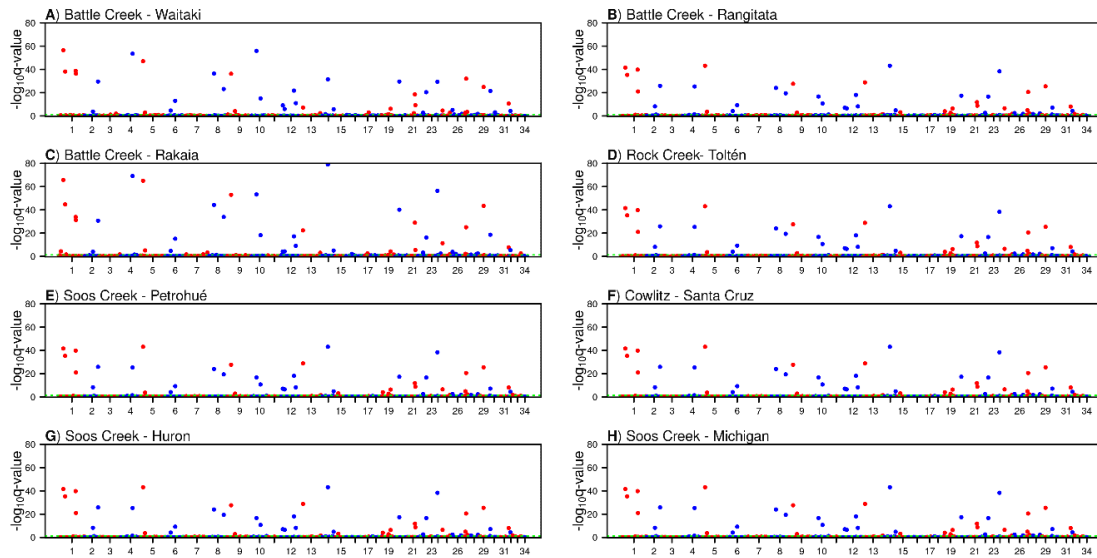


Tabla 5. Número de *outliers* encontrados por *BayeScan* para cada *prior odds* (PO) y *PCAdapt* con el valor de K estimado para cada caso; y el número de *outliers* en común entre ambos métodos utilizando un PO = 10.

Caso	Pob. Nativa	Pob. No Nativa	<i>BayeScan</i>	<i>BayeScan</i>	<i>BayeScan</i>	<i>PCAdapt</i>	<i>PCAdapt</i> K	<i>Outliers</i> en común (PO = 10)
			(PO = 10)	(PO = 30)	(PO = 100)			
1	Battle Creek	Waitaki	16	12	4	50	2	12
2	Battle Creek	Rangitata	11	9	7	52	2	9
3	Battle Creek	Rakaia	13	9	6	62	2	9
4	Rock Creek	Tolten	29	26	23	44	2	25
5	Soos Creek	Petrohué	38	33	28	68	2	33
6	Cowlitz River	Santa Cruz	11	8	5	246	2	8
7	Soos Creek	Huron Lake	46	42	39	72	2	41
8	Soos Creek	Michigan Lake	47	41	38	58	2	41

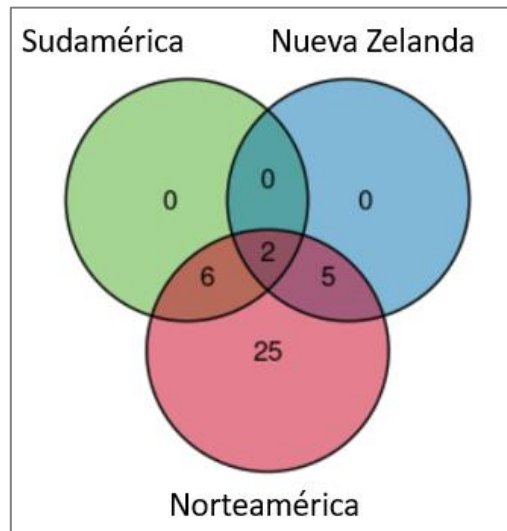
Tabla 6. Marcadores moleculares con mayor soporte estadístico de estar bajo selección positiva. Se indica la identificación del SNP, el cromosoma, la identidad de los alelos, posición física y la ubicación dentro del genoma de referencia.

SNP ID ^a	Cromosoma	Identidad alélica	Posición Física ^b	Ubicación
24210_68 ^{**†}	1	C/T	62089018	Intrón FAM196A
28712_48 [†]	2	C/T	50227374	No codificante
29338_47 ^{***†¶£}	5	C/T	10159551	Intrón MYO6
59021_53 [†]	8	T/C	58310836	No codificante
86653_70 [*]	9	C/T	13211198	Exón DSPP
14927_50 ^{***†¶£}	14	T/C	25896960	Intrón LSAMP

^a Según Hecht et al., 2015; ^b Al alinear sonda al genoma de referencia con BLASTn (Camacho et al. (2009); * Significativo entre todos los casos usando PO = 10; ** Significativo entre todos los casos usando PO = 30; *** Significativo entre todos los casos usando PO = 100; [†] Significativo a nivel regional usando PO = 100 (Norteamérica \cap Sudamérica); [¶] Significativo a nivel regional usando PO = 100 (Norteamérica \cap Nueva Zelanda); [£] Significativo a nivel regional usando PO = 100 (Sudamérica \cap Nueva Zelanda).

Los 2 marcadores con mayor poder estadístico (PO = 100) se ubican en los cromosomas *Ots05* y *Ots14*. Adicional a estos dos, cuando se utiliza PO = 30 se encuentra un marcador en *Ots01*. Cuando se utilizó PO = 10 los 5 marcadores se distribuyen en *Ots01*, *Ots05*, *Ots08*, *Ots09* y *Ots14*. Sin embargo, también se logró detectar señales de selección a nivel regional, es decir, involucrando poblaciones no nativas dentro de cada continente. Por ejemplo, cuando PO = 10, el número de marcadores dentro de Sudamérica, Norteamérica y Nueva Zelanda fue de 11, 40 y 12, respectivamente. Este número disminuye a 8, 38 y 8 cuando se usó PO = 30 y a 5, 35 y 2 cuando PO = 100, respectivamente para los continentes mencionados anteriormente (**Figura 5**).

Figura 5. Diagrama de *Venn* con el número de marcadores atípicos en común para todos los casos utilizando ambas metodologías, utilizando un PO = 100.



Interesantemente, adicional a los 5 marcadores previamente descritos, se encontró un sexto marcador de interés localizado en *Ots02* (**Tabla 6**). Este marcador fue catalogado como bajo selección por ambos métodos en todos los casos que involucraron poblaciones norteamericanas y sudamericanas (es decir, en 5 casos de los 8 independientes) asumiendo una probabilidad a priori del modelo neutral de PO = 100. Bajo este escenario, no fue detectado en las poblaciones de Nueva Zelanda, sin embargo, cuando PO = 30, se detectó en el caso de Battle Creek – Waitaki. Los gráficos de Manhattan muestran la posición y densidad de marcadores en cromosomas específicos que están asociados a rasgos de interés o en este caso, asociados a una fuerte respuesta a la selección positiva (**Figura 3: BayeScan; Figura 4: PCAdapt**).

6.5 Genes candidatos

En las regiones colindantes de los 6 SNPs de mayor importancia (con soporte estadístico y repetitivos en al menos dos regiones), se encontraron genes putativos asociados a rasgos de importancia relacionados con el crecimiento, metabolismo energético, desarrollo neurohormonal, respuesta inmune y a la regulación

conductual de la reproducción. A continuación, se describen los genes que rodean a los 6 SNPs de mayor importancia:

Uno de los marcadores con mayor soporte estadístico de estar bajo selección se encontró en *Ots05* (29338_47), en un intrón del gen *unconventional myosin-VI-like* (*MYO6-like*). Este gen ha sido previamente identificado bajo selección divergente en una especie de *ilish* (*Tenualosa ilisha*) (Asaduzzaman et al., 2020), asociado a rasgos como resistencia al virus de la necrosis pancreática infecciosa (IPNV) (Rodríguez et al., 2019), y asociado junto a otros genes en esta misma región (*FILIP1-like*, *SENP6-like*) a un rasgo ligado al crecimiento (peso eviscerado) en salmón del Atlántico (Yoshida y Yáñez, 2021).

El marcador 24210_68 (*Ots01*: Cromosoma 1 del genoma de referencia de salmón Chinook) se localizó dentro de un intrón del gen *protein FAM196A-like*. Se encuentra localizado en una región que alberga genes de importancia para el metabolismo y en la regulación neurohormonal y conductual de la reproducción de forma conservada a través de diferentes especies fish (*ADAM12*, *FAM196A*, *GNRH3*, *PTPRE*, *EBF3*, *DHX32*) (Kim et al., 2011; Gaillard et al., 2018). *ADAM12* es un gen regulador de la vía de señalización del factor de crecimiento transformante beta (*TgfB*), participa en procesos tisulares y celulares como proliferación, migración, hipertrofia y fusión de mioblastos, procesos clave para la formación y reparación de músculo (Galliano et al., 2000).

El marcador 28712_48 (*Ots02*) se ubica en una región no codificante sin embargo cercana a genes como: *protein FAM19A2-like*: previamente identificado bajo selección en salmón del Atlántico, utilizando una alta densidad de muestreo y de marcadores (Zueva et al., 2014); *Monocarboxylate transporter 1-like* (*MCT1-like*): gen relacionado al transporte de lactato, importante en la respuesta frente a la hiperlactatemia producida en respuesta a la infestación por *Caligus rogercresseyi* (Vargas-Chacoff et al., 2017); *Thyrotropin subunit beta* (*TSHβ*); gen que codifica la subunidad beta de la hormona tirotrópica (*TSH*), hormona involucrada en un

proceso neurohormonal que ocurre durante la migración río abajo llamado “*olfactory imprinting*”, que es fundamental para el reconocimiento de los sitios natales a través de los olores estudiado en juveniles del género *Oncorhynchus* (Ueda et al., 2016) y además está relacionado con la esmoltificación (metamorfosis desde el estado “*parr*” y el estado “*smolt*”, parte del ciclo de vida) (Larsen, Swanson y Dickhoff, 2011). En esta región también se encuentran genes que codifican proteínas de canales de iones (*KCND3-like*, *KCNA3-like*, *KCNA2-like*), ligados con la iono-regulación.

El marcador 59021_53 (*Ots08*), se encuentra en una región no codificante. Cercano a esta región se encuentran genes como *SLC24A4-like*, que ha sido relacionado al peso de cosecha y a la ganancia diaria de peso en salmón del Atlántico (Yoshida y Yáñez, 2021) y *protein FAM 98B-like*, gen que también ha sido asociado a peso corporal en salmón del Atlántico (Yoshida et al., 2017) y bajo selección en salmón Coho de cultivo (López et al., 2020). El marcador 86653_70 (*Ots09*) se localiza en un exón del gen *dentin sialophosphoprotein-like (DSPP)*, gen fundamental para el desarrollo del tejido dental y craneofacial en ratones (Chen et al., 2016) y cercano a genes como *Heat shock protein 30-like (HSP30-like)*, que codifica una co-chaperona altamente expresada en respuesta al estrés por calor agudo en el salmón Chinook, trucha ártica y trucha arcoíris (Currie et al., 2000; Quinn et al., 2011; Tomalty et al., 2015).

En *Ots14* encontramos el marcador 14927_50, quien junto a 29338_47 (*Ots05*) corresponden a los dos marcadores con mayor soporte estadístico. El marcador 14927_50 se ubica en un exón del gen *Limbic system-associated membrane protein-like*, el cual está implicado en la regulación de la orientación espacial y la memoria (Qiu et al., 2010; Catani et al., 2013). Además, es un gen candidato para la tendencia migratoria en trucha café (Lemopoulos et al., 2019). La lista completa de genes dentro de ventanas de 1 Mb considerando cada marcador importante como el centro de la ventana se muestra en la **Tabla 7**.

Tabla 7. Lista completa de genes dentro de ventanas de 1Mb en el genoma de referencia utilizando como centro los marcadores con mayor poder estadístico de estar bajo selección.

<i>CHR - SNP</i>	<i>Genes</i>
1 – 24210_68	<i>PTPRE-like, PGBD4-like (Pseudogene), ADAM12-like, Unc-LOC112251259, EBF3-like, GNRH3, EDFR1-like (Pseudogene), FAM196A-like, CTBP2-like, TEX36, DHX32B, ZRANB1-like.</i>
2 – 28712_48	<i>KCND3-like, CTTNBP2NL, WNT2B-A-like, ST7L, CAPZA1, PPM1H-like, TAF2-like, SLC16A1-like, Unc-LOC112262640, SLC5A8-like, TSHB-like, SLC25A22-like, CD9-like, NGF-like (Pseudogene), KCNA3-like, KCNA2-like.</i>
5 – 29338_47	<i>COL12A1-like, COX7A2L-like, TMEM30A-like, FILIP1L-like, SENP6-like, MYO6-like, IMPG1-like, HTR1D-like, NT5E-like, SNX14-like, SYNCRIP-like, BUB1-like, HTR1E-like, ZNF292-like, GJB7, Unc-LOC112250029, CFAP206, SLC35A1 -like, AKIRIN2-like, Unc-LOC112250032, CNR1-like.</i>
8 – 59021_53	<i>IL17F-like, SPRED1-like, FAM98B-like, PTPN21-like, KCNK10-like, GPR65-like, GPR4-like, BDKRB2, BDKRB1-like, C14ORF132-like, Unc-LOC112256786, Unc-LOC112256787, IL17A-like, STMN4-like, GPN1-like, CCDC88C, GPR68, DGLUCY, SLC24A4-like, CRIM1.</i>
9 – 86653_70	<i>VDAC2-like, SLC25A1-like, PI4KAB, EPR1-like, RNU7-1, CHORDC1-like, RASGRF2-like, EXPO7-like, PBDC1-like, DMTN-like, Unc-LOC112257495, DSPP-like, SH3BP2-like, VLDLR-like, KCNV2-like, HSP30-like, HSPB11-like, TNKS, DUSP4-like, CYP120, 5-HT4-like, PPP3CC-like (Pseudogene), PUM3-like (Pseudogene), CARM1-like (Pseudogene).</i>
14 – 14927_50	<i>RAP1GAP2A-like, Unc-LOC112266825, BMH-like, SLC6A4A-like, Unc-LOC112266828, SSH2-like, ANKRD13B -like, GIT1-like, Unc-LOC112266830, Unc-LOC112266831, DDB_G0292636-like, IGSF11-like, LSAMP-like.</i>

Unc. – Gen no caracterizado

DISCUSIÓN

En esta tesis, se utilizaron marcadores genómicos distribuidos a través del genoma segregando en múltiples poblaciones tanto nativas como naturalizadas de salmón Chinook, con el fin de detectar señales a nivel molecular estadísticamente significativas sugerentes del efecto de una fuerte selección positiva sobre la composición genotípica de las poblaciones en estudio. Este diseño permite aportar antecedentes a la hipótesis de que existen genes pre-adaptados que podrían estar favoreciendo el establecimiento de poblaciones no nativas, ya que, los mismos genes se encontrarían bajo una fuerte selección positiva.

Se incluyeron en los análisis 8 pares de poblaciones independientes incluyendo poblaciones nativas (fuente genética putativa identificada por registros y/o estudios previos) y no nativas (Quinn et al., 2001; Weeder et al., 2005; Correa y Gross, 2008; Correa y Morán, 2017; Gómez-Uchida et al., 2018). Para esto, se utilizaron enfoques genómicos reversos, es decir, sin mediciones directas de fenotipos ni variables ambientales, sino que solo utilizando la información de frecuencias alélicas que brindan los marcadores moleculares tipo SNP segregando en las poblaciones de 3 macro regiones geográficas (Sudamérica, Nueva Zelanda y Norteamérica). De manera complementaria, se evaluó la diversidad y estructura genética de cada población, y la diferenciación genética entre poblaciones.

7.1 Objetivo 1: ¿Existen señales de selección en común entre los múltiples casos de poblaciones divergentes?

Los resultados de este estudio indican que si existen señales de selección en los mismos loci en todos los casos de poblaciones divergentes evaluados. Sin embargo, dado el relativamente bajo número de marcadores utilizados, el número de animales utilizados para realizar las inferencias y la posible influencia de otros factores que pueden modificar las frecuencias alélicas (deriva génica, mutación, migración, etc.), es posible que algunos de estos marcadores puedan ser falsos positivos. Por otra

parte, la ausencia de una mayor cantidad de marcadores bajo selección también podrían ser casos de falsos negativos. Por esta razón, la discusión se centra sobre los marcadores bajo selección con mayor evidencia estadística, es decir, en común a través de los múltiples casos de poblaciones divergentes. Estos, se dividen en primer lugar, aquellos SNPs detectados en común por ambos métodos y entre todos los casos independientes; y, en segundo lugar, aquellos detectados en común dentro de al menos dos regiones. Al realizar los análisis con criterios escépticos y escenarios conservadores, los métodos utilizados sugieren una fuerte selección positiva en al menos 2 SNPs ubicados en diferentes cromosomas (*Otsh05* y *Otsh14*). El patrón atípico de estos dos loci es consistente a través de los 8 casos de poblaciones divergentes, donde el primer alelo se encuentra fijo o en muy alta frecuencia en la población representativa donadora, y el segundo alelo se encuentra fijo o muy alta frecuencia en la población no nativa. A nivel regional, existe un número levemente mayor de marcadores significativos en común (**Tabla 6; Figura 3**). Lo mismo ocurre al aumentar la confianza del modelo de selección versus el modelo neutral (disminuyendo las *prior odds*).

El bajo solapamiento (0,08% de las 4.529 SNPs posicionados en cromosomas) puede explicarse por una combinación de factores que incluye el relativamente bajo número de loci y muestras utilizadas (y, en consecuencia, baja representación de la variación genómica en el análisis), la diferenciación genética entre pares de poblaciones y metodología utilizada. Por una parte, la secuenciación RAD, utilizada para el descubrimiento de sitios polimórficos en múltiples poblaciones nativas (enzima *SbfI*-HF), se caracteriza por una alta cobertura genética, lo cual quiere decir que se tiene confianza en que los loci descubiertos son sitios de variación porque se obtienen muchas lecturas de la misma región. Sin embargo, esta técnica deja “lagunas” o “gaps” en la representación genómica, debido a que los fragmentos que se generan, a partir de los sitios de corte donde actúan las enzimas, no necesariamente son contiguos (Baird et al., 2008). En consecuencia, la densidad de SNP utilizada en este estudio puede no capturar ni evaluar variación genómica adaptativa que albergue señales selección y los loci detectados bajo selección sólo

pueden estar representando una pequeña proporción de las regiones genómicas que están realmente involucradas en la respuesta a la selección (Hohenlohe et al., 2010).

Es importante considerar, además, que loci atípicos no repetidos a través de todas las poblaciones pueden no estrictamente ser todos falsos positivos, ya que pueden estar implicados en la adaptación a las condiciones específicas de la población, como sugiere el número de valores atípicos encontrados en común dentro de cada región específica (Figura 2). Por otra parte, la diferenciación poblacional entre poblaciones en combinación con la metodología utilizada puede afectar la detección de señales de selección. Tanto los niveles muy bajos como los altos de diferenciación genética entre poblaciones pueden estar afectando al poder de detectar loci bajo selección positiva divergente. La varianza de la distribución de los valores F_{ST} en loci neutrales es mayor cuando el promedio de F_{ST} es mayor. Si el F_{ST} promedio es grande, un locus debe tener un F_{ST} con un valor extremadamente grande para estar en el extremo (o “cola”) de la distribución y, por tanto, ser considerado como bajo selección (Hoban et al., 2016). Los métodos basados en la diferenciación dada por el indicador F_{ST} podrían ser más sensibles a este efecto, que podría explicar la diferencia exacerbada entre el número de loci atípicos encontrados por PCAdapt en comparación con BayeScan en COW-STC, y el menor número de valores atípicos detectados en general por este software en comparación con PCAdapt. El método de diferenciación poblacional implementado en PCAdapt se basa en la estructura poblacional y componentes principales, no en valores F_{ST} . Por lo tanto, por el contrario, a lo que ocurre en BayeScan, cuanto más diferenciadas estén las poblaciones, más valores atípicos encontrará PCAdapt.

Una gran proporción de los rasgos relacionados con el fitness (es decir, fenotipos complejos o medidas que relacionadas con la sobrevivencia - mortalidad y fertilidad) sobre los que actúa la selección natural (con efecto favorable/desfavorable sobre el *fitness*) probablemente no estén controlados por un único locus (Lotterhos y Schaal, 2014), sino más bien son influenciados por muchos loci de pequeño efecto, que a

menudo pasarían desapercibidas utilizando métodos convencionales de valores atípicos (Hancock et al., 2010; Pritchard y Di Rienzo, 2010). Es probable que el enfoque de los valores atípicos sólo descubra loci que tengan un alto efecto sobre fenotipos exitosos, que hayan barrido hasta una alta frecuencia en la población o que estén bajo una fuerte selección divergente entre poblaciones (Pritchard y Di Rienzo, 2010; Lotterhos y Whitlock, 2014).

Aquellos procesos de selección que conducen a pequeños cambios en las frecuencias alélicas no logran ser detectados con la estrategia utilizada en esta tesis, lo cual podría influir sobre el bajo solapamiento entre los diferentes métodos (de Villemereuil et al., 2014). No obstante, los análisis realizados en esta tesis proporcionan los primeros conocimientos sobre las pre-adaptaciones genómicas en el salmón Chinook, y de los alelos putativamente adaptativos que diferencian fuertemente las poblaciones nativas de las no nativas estudiadas.

Los seis SNPs con mayor apoyo estadístico mostraron altos niveles de fijación de alelos opuestos en cada caso emparejado, es decir, la misma identidad alélica fue compartida dentro de las poblaciones nativas y la opuesta dentro de las poblaciones no nativas (**Tabla 8**). La distribución de los alelos dentro de las poblaciones nativas y no nativas puede estar indicando subdominancia (es decir, los heterocigotos tienen menor aptitud que los homocigotos), donde se espera que con el tiempo los alelos se pierdan o se fijen completamente dentro de las poblaciones individuales (Altrock, Traulsen, y Reed, 2011). Un fenómeno similar, pero a nivel de haplotipos, se ha observado en la trucha arcoíris, donde la adaptación paralela ha sido impulsada por un haplotipo conservado en poblaciones geográficamente distantes (Miller et al., 2012). Los autores han predicho que la mayoría de los casos de adaptación paralela en las especies anádromas se logrará a través de la utilización repetida de los mismos alelos adaptativos.

Tabla 8. Frecuencias del alelo menor de los SNPs bajo selección con mayor soporte estadístico dentro de cada población. Se indica la identidad de los alelos (menor / mayor) y la frecuencia del alelo menor en paréntesis.

<i>POP SNP</i>	24210_68	28712_48	29338_47	59021_53	86653_70	14927_50
<u>Battle Creek</u>	0/C (0)	T/C (0.07)	0/C (0)	C/T (0.07)	0/C (0)	0/T (0)
<u>Soos Creek</u>	T/C (0.07)	T/C (0.09)	T/C (0.02)	C/T (0.06)	T/C (0.02)	C/T (0.03)
<u>Cowlitz</u>	0/C (0)	T/C (0.03)	0/C (0)	C/T (0.02)	T/C (0.12)	0/T (0)
<u>Rock Creek</u>	T/C (0.05)	T/C (0.07)	0/C (0)	C/T (0.05)	T/C (0.02)	0/T (0)
<u>Waitaki</u>	C/T (0.03)	0/T (0)	C/T (0.02)	0/C (0)	C/T (0.04)	T/C (0.03)
<u>Ranqitata</u>	0/T (0)	C/T (0.02)	0/T (0)	0/C (0)	0/T (0)	0/C (0)
<u>Rakaia</u>	C/T (0.04)	0/T (0)	0/T (0)	0/C (0)	0/T (0)	0/C (0)
<u>Toltén</u>	0/T (0)	C/T (0.01)	0/T (0)	0/C (0)	0/T (0)	0/C (0)
<u>Petrohué</u>	C/T (0.01)	C/T (0.01)	C/T (0.02)	T/C (0.01)	0/T (0)	0/C (0)
<u>Santa Cruz</u>	C/T (0.01)	0/T (0)	0/T (0)	0/C (0)	0/T (0)	0/C (0)
<u>Michigan</u>	0/T (0)	0/T (0)	0/T (0)	0/C (0)	C/T (0.04)	T/C (0.04)
<u>Huron</u>	0/T (0)	C/T (0.02)	0/T (0)	0/C (0)	0/T (0)	T/C (0.04)

Estos resultados sugieren una fuerte selección positiva entre las múltiples poblaciones divergentes. Sin embargo, sigue siendo posible que resulten ser falsos positivos debido a las otras fuerzas que actúan sobre las frecuencias alélicas poblacionales y que podrían hacer mímica de lo que genera la selección, como los cuellos de botella, la deriva genética, las migraciones inadvertidas o los efectos fundadores. En poblaciones invasoras fundadas por migración natural como la STC, cabría esperar que las fuerzas antes mencionadas pudieran conducir a la fijación de alelos. Pero también esperaríamos que los múltiples eventos de introducción, las presiones de propágulos y la mezcla entre fuentes deberían haber contrarrestado la potencial fijación dentro de las poblaciones no nativas fundadas artificialmente

(Gómez-Uchida et al., 2018; Narum et al., 2017). Además, el modelo de estudio propuesto en este trabajo consideró análisis poblacionales por pares lo cual nos permitió cumplir con el supuesto de evolución independiente (modelo implementado en BayeScan), y maximizar el número de veces en que se repite la evidencia.

7.2 Objetivo 2: ¿Cuáles son los genes candidatos que se encuentran en o cercanos a los marcadores detectados bajo selección?

Curiosamente, a pesar del pequeño número de SNPs identificados en común, algunos genes adyacentes probablemente desempeñan un papel en rasgos fundamentales de la historia de la vida o han sido encontrados bajo selección o asociados a rasgos de importancia mediante enfoques de genómica avanzada.

Entre todos los genes que rodean a los seis SNPs con mayor soporte estadístico de estar bajo selección, se encontraron genes asociados a crecimiento (peso corporal) (*MYO6-like*, *SLC24A4-like*, *FAM98B-like*, *PGBD4-like*), el sistema inmunitario (*MCT1-like*, *IL17F-like*, *IGSF11-like*, *VDAC2-like*), el estrés térmico y termoregulación (*HSP30-like*, *HSPB11-like*, *BDKRB2*), regulación iónica (*KCND3-like*, *KCNA3-like*, *KCNA2-like*, *KCNV2-like*) y genes implicados en el metabolismo, la regulación neurohormonal y conductual de la reproducción en peces (*ADAM12-like*, *FAM196A-like*, *GNRH3*, *PTPRE-like*, *EBF3-like*, *DHX32-like*, *TSH β -like*) (**Tabla 7**). Todas estas características están relacionadas con la adecuación biológica, en consecuencia, estos serían los principales genes candidatos de estar bajo selección. Los hallazgos sugieren que genes relacionados con la aptitud física pueden ser objeto de selección durante la evolución divergente recientes. Estos resultados brindan soporte a la hipótesis de pre-adaptación a nivel molecular, donde los mismos genes presentes previo a los eventos de divergencia, se encuentran bajo selección, jugando un papel muy poco comprendido sobre la capacidad invasora de la especie.

El tamaño y peso corporal pueden tener un fuerte efecto en la adecuación biológica, ya que tiende a estar positivamente correlacionado con la supervivencia en agua dulce, la supervivencia en el mar, la fecundidad, el tamaño de los huevos, el éxito reproductivo y la supervivencia de las crías (García de Leaniz et al., 2007). *MYO6-like* ha sido detectado bajo selección divergente en una especie anádroma de sábalo (*Tenualosa ilisha*) (Asaduzzaman et al., 2020) y asociado a un rasgo relacionado con el peso en el salmón del Atlántico (Yoshida y Yáñez, 2021). Interesantemente, otros autores han logrado, mediante enfoques genómicos que utilizan la información fenotípica de miles de animales y marcadores genéticos, encontrar los mismos genes que se encontraron en este trabajo de tesis, principalmente asociados al crecimiento de juveniles en salmón del Atlántico (*PGBD4-like* (Tsai et al., 2015)); peso de cosecha y ganancia diaria de peso (*SLC24A4-like* (Yoshida y Yáñez, 2021)) y peso corporal (*FAM98B-like* (Yoshida et al., 2017)) en salmón Atlántico adulto.

ADAM12-like es un regulador conocido de la vía de señalización del factor de crecimiento transformante beta ($Tgf\beta$), que media en procesos celulares/tisulares como la proliferación, la migración, la hipertrofia y la invasión, y se ha asociado con la morfogénesis del hocico en cíclidos como *Labeotropheus* spp. (Conith et al., 2018). *ADAM12-like* es crucial para el desarrollo del cartílago-hueso-músculo en el pez cebra (Tokumasu et al., 2016), y con rasgos relacionados con el crecimiento en el pez gato (Li et al., 2018), y en el cerdo (Wang et al., 2016). El papel de este gen en el salmón Chinook no está completamente comprendido; sin embargo, los procesos en los que participa en otras especies son importantes para el desarrollo y la supervivencia de estas. Además, la existencia de genes en un bloque sintético conservado (*ADAM12-like*, *FAM196A-like* y *PTPRE-like*) entre diferentes especies (incluyendo el salmón Chinook, el pez gato, el ser humano y el cerdo) indica que la función de estos genes puede estar conservada evolutivamente (Li et al., 2018).

La variación ambiental en las precipitaciones y la distancia migratoria son los principales impulsores de la adaptación divergente entre las poblaciones del área

de distribución nativa del salmón Chinook (Hecht et al., 2015). Las distancias migratorias que recorren los juveniles para llegar al mar y los adultos para reproducirse significan un coste energético y presiones selectivas (depredadores, adaptación a gradientes químicos/físicos ambientales) (Hecht et al., 2015). La variación de la temperatura del agua también se ha asociado a la divergencia adaptativa (Matala et al., 2014). La tolerancia térmica es crítica para adaptarse a los cambios de temperatura del agua durante la historia de la vida y es fundamental para la supervivencia y la aptitud en el salmón Chinook y otros salmónidos (Eliason et al., 2011; Muñoz et al., 2015). Las *heat-shock proteins* (HSP) son chaperonas que pueden inhibir la agregación de proteínas mal plegadas inducida por el calor. El gen *HSP30* codifica una co-chaperona altamente expresada en respuesta al estrés térmico agudo en el salmón Chinook, el salvelino ártico y la trucha arco iris (Currie et al., 2000; Quinn et al., 2011; Tomalty et al., 2015). En el tejido branquial del salmón Chinook criado a 12 °C y sometido a un choque térmico a 25 °C durante tres horas, se observó una respuesta intensa en la acumulación de ARNm de *HSP30* en comparación con el control (Tomalty et al., 2015). El gen *BDKRB2* juega un rol en la termorregulación mediante la regulación de la vasoconstricción y vasodilatación sanguínea. Este gen en particular ha sido encontrado bajo selección positiva en pingüino emperador, donde los autores sugieren una relación entre la adaptación a la termorregulación y la capacidad de los pingüinos de colonizar áreas sobre su límite máximo térmico (Vianna et al., 2020).

Dos rasgos fundamentales del ciclo vital de *Oncorhynchus* spp. son la migración río abajo de los juveniles y el regreso a casa de los adultos (Ueda, 2012). Durante la residencia en agua dulce y la migración río abajo, los juveniles reconocen características químicas únicas y luego los adultos utilizan estas características químicas para regresar a los arroyos natales, fenómeno conocido como “*downstream imprinting*” (Bourret et al., 2016; Ueda, 2019). Esto aumenta la probabilidad de que los peces encuentren parejas y localicen hábitats que sean favorables tanto para el desove de los adultos como para la supervivencia de los juveniles; por lo tanto, este comportamiento es altamente adaptativo ya que

proporcionan beneficios para la adecuación biológica (Quinn, 2005; Waples et al., 2004). Este comportamiento ha conducido a una amplia divergencia entre las poblaciones dentro de las especies, promueve la adaptación local y permite el extravío no natal, que ha sido fundamental para que diversas especies se expandan a hábitats no explotados (Fraser et al., 2011; Keefer y Caudill, 2014). La impronta juvenil durante la migración río abajo está asociada a la esmoltificación, que es un proceso complejo que está regulado por varias hormonas endocrinas, como la hormona tirotrópica (TSH), entre otras (véase Ueda, 2019). Se piensa que las hormonas del eje cerebro-hipófisis-tiroides ("*Brain-Pituitary-Thyroid*": BPT), incluida la TSH en la hipófisis, desempeñan funciones críticas durante el proceso de impronta olfativa durante la migración río abajo de los juveniles de salmón del Pacífico (Fourcade et al., 2013; Ueda et al., 2016), y con la esmoltificación en el salmón Coho (Larsen et al., 2011). Por otra parte, las hormonas del eje cerebro-hipófisis-gónadal ("*Brain-Pituitary-Gonad*": BPG) como la hormona liberadora de gonadotropina del salmón (GnRH3) está implicada en la maduración gonadal y la recuperación de la memoria olfativa de ambos sexos en el salmón del Pacífico (Ueda, 2019).

El gen *LSAMP-like* ha sido implicado en la regulación de la orientación y la memoria espacial (Catani et al., 2013; Qiu et al., 2010) y como gen candidato a la tendencia migratoria en la trucha marrón (Lemopoulos et al., 2019). La selección en regiones que albergan genes implicados en procesos fundamentales del ciclo vital como la anadromía, la impronta olfativa, el reconocimiento del desove natal y en la maduración sexual puede estar sugiriendo que estos rasgos están pre-adaptados funcionalmente a pesar de no ser rasgos cuantitativos. Sin embargo, no todos los salmónidos anádromos introducidos en la Patagonia han extendido su rango de distribución (Arismendi et al., 2014), lo que enfatiza el punto de que los factores que influyen en el éxito de la colonización pueden ser muy complejos (Dlugosch et al., 2015).

CONCLUSIONES

El presente estudio ha proporcionado una visión sin precedentes de los loci putativamente adaptativos y de la función putativa de sus genes flanqueantes, ya que es uno de los primeros en utilizar el mismo conjunto de datos genómicos para analizar las señales de divergencia entre poblaciones nativas y no nativas. La identificación de SNPs *outliers* tanto por PCAdapt como por BayeScan para 8 ocho análisis independientes por pares proporciona soporte estadístico de una fuerte selección diversificadora en los genes implicados en el rasgo relacionado con la aptitud. Además, estos resultados apoyan la hipótesis de que la selección puede estar actuando sobre genes pre-adaptados existentes previos a los eventos de divergencia, implicados en mecanismos funcionales o rasgos cuantitativos que promueven la aptitud. No es adecuado considerar que el éxito invasor del salmón Chinook se deba exclusivamente a variantes genéticas preadaptadas, sino a una combinación de factores que la incluyen. En consecuencia, la sugerencia es interpretar estos resultados con precaución, ya que la selección detectada aquí puede no ser puramente natural. Dadas las dificultades intrínsecas de la interpretación de los datos de los mapas de selección, serán indispensables herramientas adicionales, como los estudios de asociación de todo el genoma basados en el genotipado de alto rendimiento o bien, la utilización de datos de secuenciación de todo el genoma obtenidos de grandes poblaciones de referencia, para comprender el significado biológico de las señales de selección descubiertas. Futuras investigaciones que utilicen una mayor densidad de marcadores podrían aumentar la potencia estadística y permitir el uso de métodos basados en el LD intrapoblacional, lo que podría mejorar la búsqueda no sólo de una fuerte selección diversificadora, sino también de una selección equilibrada y de pequeños cambios en respuesta a la selección que se produzcan durante la evolución paralela, que es una debilidad inherente a los enfoques de los valores atípicos.

BIBLIOGRAFÍA

AKEY, J. M. 2009. Constructing genomic maps of positive selection in humans: Where do we go from here? *Genome Research*, 19(5), 711–722.

AKEY, J. M.; ZHANG, G.; ZHANG, K.; JIN, L.; SHRIVER, M. D. 2002. Interrogating a high-density SNP map for signatures of natural selection. *Genome Research*, 12(12), 1805–1814.

ALEXANDER, D. H.; LANGE, K. 2011. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinformatics*, 12(1), 246.

ALEXANDER, D. H.; NOVEMBRE, J.; LANGE, K. 2009. Fast model-based estimation of ancestry in unrelated individuals. *Genome Research*, 19(9), 1655–1664.

ALLENDORF, F. W.; HOHENLOHE, P. A.; LUIKART, G. 2010. Genomics and the future of conservation genetics. *Nature Reviews Genetics*, 11(10), 697–709.

ALTROCK, P. M.; TRAUlsen, A.; REED, F. A. 2011. Stability properties of underdominance in finite subdivided populations. *PLoS Computational Biology*, 7(11), 1002260.

ANDERSON, C. A.; PETTERSSON, F. H.; CLARKE, G. M.; CARDON, L. R.; MORRIS, A. P.; ZONDERVAN, K. T. 2010. Data quality control in genetic case-control association studies. *Nature Protocols*, 5(9), 1564–1573.

ANDREWS, K. R.; GOOD, J. M.; MILLER, M. R.; LUIKART, G.; HOHENLOHE, P. A. 2016. Harnessing the power of RADseq for ecological and evolutionary genomics. *Nature Reviews Genetics*, 17(2), 81–92.

ARISMENDI, I.; PENALUNA, B.; DUNHAM, J.; GARCÍA DE LEANIZ, C.; SOTO, D.; FLEMING, I.; GÓMEZ-UCHIDA, D.; GAJARDO, G.; VARGAS, P.; LEÓN-MUÑOZ, J. 2014. Differential invasion success of salmonids in southern Chile: patterns and hypotheses. *Rev Fish Biol Fisheries*, 24, 919–941.

ARISMENDI, I.; SANZANA, J.; SOTO, D. 2011. Seasonal age distributions and maturity stage in a naturalized rainbow trout (*Oncorhynchus mykiss*, Walbaum) population in southern Chile reveal an ad-fluvial life history. *International Journal of Limnology*, 47(2), 133–140.

ARISMENDI, I.; SOTO, D. 2012. Are salmon-derived nutrients being incorporated in food webs of invaded streams? Evidence from southern Chile. *Knowledge and Management of Aquatic Ecosystems*, 405.

ARISMENDI, I.; SOTO, D.; ROOKE, B.; ENALUNA, P.; JARA, C.; LEAL, C.; LEÓN, J. 2009. Aquaculture, non-native salmonid invasions and associated declines of native fishes in Northern Patagonian lakes. *Freshwater Biology*, 54, 1135–1147.

ASADUZZAMAN, M.; IGARASHI, Y.; WAHAB, M.A.; NAHIDUZZAMAN, M.; RAHMAN, M.J.; PHILLIPS, M.J.; HUANG, S.; ASAKAWA, S.; RAHMAN, M.M.; WONG, L.L. 2020. Population Genomics of an Anadromous Hilsa Shad *Tenualosa ilisha* Species across Its Diverse Migratory Habitats: Discrimination by Fine-Scale Local Adaptation. *Genes*, 11, 46.

AUGEROT, X. 2005 *Atlas of Pacific Salmon*. 1st ed. University of California Press, Berkeley.

AYLLON, F.; KJAERNER-SEMB, E.; FURMANEK, T.; WENNEVIK, V.; SOLBERG, M. F.; DAHLE, G.; LASSE TARANGER, G.; GLOVER, K. A.; ALMÉN, M. S.; RUBIN, C. J.; EDVARDBSEN, R. B.; WARGELIUS, A. 2015. The vglI3 Locus Controls Age at Maturity in Wild and Domesticated Atlantic Salmon (*Salmo salar* L.) Males. *PLOS Genetics* 11(11): e1005628.

BAIRD, N. A.; ETTER, P. D.; ATWOOD, T. S.; CURREY, M. C.; SHIVER, A. L.; LEWIS, Z. A.; SELKER, E.U.; CRESKO, W. A.; JOHNSON, E.A. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLOS ONE*, 3 (10), e3376.

BANK, C.; EWING, G. B.; FERRER-ADMETTLA, A.; FOLL, M.; JENSEN, J. D. 2014. Thinking too positive? Revisiting current methods of population genetic selection inference. *Trends in Genetics*, 30(12), pp. 540–546.

BAÑOS-VILLALBA, A.; CARRETE, M.; TELLA, J. L.; BLAS, J.; POTTI, J.; CAMACHO, C.; DIOP, M. S.; MARCHANT, T. A.; CABEZAS, S.; EDELAAR, P. 2020. Selection on individuals of introduced species starts before the actual introduction. *Evolutionary Applications*, 00, pp. 1-13.

BARREIRO, L. B.; LAVAL, G.; QUACH, H.; PATIN, E.; QUINTANA-MURCI, L. 2008. Natural selection has driven population differentiation in modern humans. *Nature Genetics*, 40(3), pp. 340–345.

BARRETT, R. D. H.; SCHLUTER, D. 2008. Adaptation from standing genetic variation. *Trends in Ecology and Evolution*, 23(1), pp. 38–44.

BARRETT, S. C. H. 2015 Foundations of invasion genetics: the Baker and Stebbins legacy. *Molecular Ecology*, 24(9), pp. 1927–1941.

BARRIA, A.; MARÍN-NAHUELPI, R.; CÁCERES, P.; LÓPEZ, M. E.; BASSINI, L. N.; LHORENTE, J. P.; YÁÑEZ, J. M. 2019. Single-step genome-wide association study for resistance to *Piscirickettsia salmonis* in rainbow trout (*Oncorhynchus mykiss*) G3: Genes, Genomes, Genetics, 9(11), pp. 3833-3841.

BARSON, N. J.; AYKANAT, T.; HINDAR, K.; BARANSKI, M.; BOLSTAD, G. H.; FISKE, P.; JACQ, C.; JENSEN, A. J.; JOHNSTON, S. E.; KARLSSON, S.; KENT, M.; MOEN, T.; NIEMELÄ, E.; NOME, T.; NÆSJE, T. F.; ORELL, P.; ROMAKKANIEMI, A.; SÆGROV, H.; URDAL, K.; ERKINARO, J.; LIEN, S.; PRIMMER, C. R. 2015. Sex-dependent dominance at a single locus maintains variation in age at maturity in salmon. *Nature*, 528(7582), pp. 405–408.

BASULTO, S. 2003. El largo viaje de los salmones: una crónica olvidada, propagación y cultivo de especies acuáticas en Chile. Editorial Maval, Ltda.

BECKER, L. A.; PASCUAL, M. A.; BASSO, N. G. 2007. Colonization of the southern Patagonia ocean by exotic Chinook salmon. *Conservation Biology*, 21(5), pp. 1347–1352.

BELL, G.; GONZALEZ, A. 2009. Evolutionary rescue can prevent extinction following environmental change. *Ecology Letters*, 12(9), pp. 942–948.

BJÖRNSSON, B.T.; STEFANSSON, S.O.; MCCORMICK, S. D. 2011. Environmental endocrinology of salmon smoltification. *General and Comparative Endocrinology*. 170(2): pp. 290-298.

BOURRET, S. L.; CAUDILL, C. C.; KEEFER, M. L. 2016. Diversity of juvenile Chinook salmon life history pathways. *Reviews in Fish Biology and Fisheries*, 26(3), pp. 375–403.

BRAND, C. L.; CATTANI, M. V.; KINGAN, S. B.; LANDEEN, E. L.; PRESGRAVES, D. C. 2018. Molecular Evolution at a Meiosis Gene Mediates Species Differences in the Rate and Patterning of Recombination, 28(8), pp. 342-344.

CAMACHO, C.; COULOURIS, G.; AVAGYAN, V.; MA, N.; PAPADOPOULOS, J.; BEALER, K.; MADDEN, T. L. 2009. BLAST+: Architecture and applications. *BMC Bioinformatics*, 10.

CANDY, J. R.; BEACHMAN, T.D. 2000. Patterns of homing and straying in southern British Columbia coded-wire tagged chinook salmon (*Oncorhynchus tshawytscha*) populations. *Fisheries Research*, 47(1), pp. 41-56.

CASAL, C. M. V. 2006. Global documentation of fish introductions: The growing crisis and recommendations for action. *Biological Invasions*, 8(1), pp. 3–11.

CASTAÑO-BETANCOURT, M. C.; EVANS, D. S.; RAMOS, Y. F. M.; BOER, C. G.; METRUSTRY, S.; LIU, Y.; DEN HOLLANDER, W.; VAN ROOIJ, J.; KRAUS, V. B.; YAU, M. S.; MITCHELL, B. D.; MUIR, K.; HOFMAN, A.; DOHERTY, M.; DOHERTY, S.; ZHANG, W.; KRAAIJ, R.; RIVADENEIRA, F.; BARRETT-CONNOR, E.; VAN MEURS, J. B. J., et al. 2016. Novel Genetic Variants for Cartilage Thickness and Hip Osteoarthritis. *PLOS Genetics*, 12(10), e1006260.

CATANI, M.; DELL'ACQUA, F.; THIEBAUT DE SCHOTTEN, M. 2013. A revised limbic system model for memory, emotion and behaviour. *Neuroscience and Biobehavioral Reviews*, 37(8), pp. 1724–1737.

CHALDE, T.; NARDI, C. F.; FERNÁNDEZ, D. A. 2019. Early warning: Detection of exotic coho salmon (*Oncorhynchus kisutch*) by environmental DNA and evidence of establishment at the extreme south of patagonia. *Canadian Journal of Fisheries and Aquatic Sciences*, 76(12), pp. 2343–2349.

CHEN, Y.; ZHANG, Y.; RAMACHANDRAN, A.; GEORGE, A. 2016. DSPP Is Essential for Normal Development of the Dental-Craniofacial Complex. *Journal of Dental Research*, 95(3), pp. 302–310.

CHOWN, S. L.; HODGINS, K. A.; GRIFFIN, P. C.; OAKESHOTT, J. G.; BYRNE, M.; HOFFMANN, A. A. 2015. Biological invasions, climate change and genomics. *Evolutionary Applications*, 8(1), pp. 23–46.

CHRISTENSEN, K. A.; LEONG, J. S.; SAKHRANI, D.; BIAGI, C. A.; MINKLEY, D. R.; WITHLER, R. E.; RONDEAU, E. B.; KOOP, B. F.; DEVLIN, R. H. 2018 Chinook salmon (*Oncorhynchus tshawytscha*) genome and transcriptome. *PLOS ONE*, 13(4), e0195461.

CIANCIO, J. E.; PASCUAL, M. A.; LANCELOTTI, J.; ROSSI, C. M. R.; BOTTO, F. 2005. Natural colonization and establishment of a chinook salmon, *Oncorhynchus tshawytscha*, population in the Santa Cruz River, an Atlantic basin of Patagonia. *Environmental Biology of Fishes*, 74(2), pp. 219–227.

CIANCIO, J. E.; RIVA ROSSI, C.; PASCUAL, M.; ANDERSON, E.; GARZA, J. C. 2015. The invasion of an Atlantic Ocean river basin in Patagonia by Chinook salmon: new insights from SNPs. *Biological Invasions*, 17(10), pp. 2989–2998.

CIUCULETE, D. M.; BOSTRÖM, A. E.; VOISIN, S.; PHILIPPS, H.; TITOVA, O. E.; BANDSTEIN, M.; NIKONTOVIC, L.; WILLIAMS, M. J.; MWINYI, J.; SCHIÖTH, H. B. 2017. A methylome-wide mQTL analysis reveals associations of methylation sites with GAD1 and HDAC3 SNPs and a general psychiatric risk score. *Translational Psychiatry*, 7(1), e1002.

COLAUTTI, R. I.; GRIGOROVICH, I. A.; MACISAAC, H. J. 2006. Propagule pressure: A null model for biological invasions. *Biological Invasions*, 8(5), pp. 1023–1037.

CONITH, M. R.; HU, Y.; CONITH, A. J.; MAGINNIS, M. A.; WEBB, J. F.; CRAIG ALBERTSON, R. 2018. Genetic and developmental origins of a unique foraging adaptation in a Lake Malawi cichlid genus. *Proceedings of the National Academy of Sciences of the United States of America*, 115 (27), 7063–7068.

CONNERTON, M. J.; MURRY, B. A.; RINGLER, N. H.; STEWART, D. J. 2009. Majority of age-3 Chinook salmon (*Oncorhynchus tshawytscha*) in Lake Ontario were wild from 1992 to 2005, based on scale pattern analysis. *Journal of Great Lakes Research*, 35(3), pp. 419–429.

COPP, G. H.; BIANCO, P. G.; BOGUTSKAYA, N. G.; EROS, T.; FALKA, I.; FERREIRA, M. T.; FOX, M. G.; FREYHOF, J.; GOZLAN, R. E.; GRABOWSKA, J.; KOVÁČ, V.; MORENO-AMICH, R.; NASEKA, A. M.; PEŇÁZ, M.; POVŽ, M.; PRZYBYLSKI, M.; ROBILLARD, M.; RUSSELL, I. C.; STAKENAS, S. et al. 2005. To be, or not to be, a non-native freshwater fish? *Journal of Applied Ichthyology*, 21(4), pp. 242–262.

CORREA, C.; GROSS, M. R. 2008. Chinook salmon invade southern South America. *Biological Invasions*, 10(5), pp. 615–639.

CORREA, C.; MORAN, P. 2017. Polyphyletic ancestry of expanding Patagonian Chinook salmon populations. *Scientific Reports*, 7(1), pp. 1–13.

CORREA, C.; HENDRY, A. 2012. Invasive salmonids and lake order interact in the decline of puye grande *Galaxias platei* in western Patagonia lakes. *Ecological Applications*, 22(3), pp. 828–842.

CRAM, J. M.; TORGERSEN, C. E.; KLETT, R. S.; PESS, G. R.; MAY, D.; PEARSONS, T. D.; DITTMAN, A.H. 2013. Tradeoffs between homing and habitat

quality for spawning site selection by hatchery-origin Chinook salmon. *Environmental Biology of Fishes*, 96, pp. 109-122.

CRAWFORD, S. S.; MUIR, A. M. 2008. Global introductions of salmon and trout in the genus *Oncorhynchus*: 1870-2007. *Reviews in Fish Biology and Fisheries*, 18(3), pp. 313–344.

CURRIE, S., MOYES, C. D., & TUFTS, B. L. 2000. The effects of heat shock and acclimation temperature on hsp70 and hsp30 mRNA expression in rainbow trout: in vivo and in vitro comparisons. *Journal of Fish Biology*, 56(2), pp. 398–408.

DARWIN, C.; WALLACE, A. 1858. On the Tendency of Species to form Varieties; and on the Perpetuation of Varieties and Species by Natural Means of Selection. *Journal of the Proceedings of the Linnean Society of London. Zoology*, 3(9), pp. 45–62.

DAVEY, J. L.; BLAXTER, M. W. 2010. RADseq: Next-generation population genetics. *Briefings in Functional Genomics*, 9(5–6), pp. 416–423.

KEEFER, M. L.; CAUDILL, C. C. 2014. Homing and straying by anadromous salmonids: A review of mechanisms and rates. *Reviews in Fish Biology and Fisheries*, 24(1), pp. 333–368.

KIM, D.K.; CHO, E.B.; MOON, M.J.; PARK, S.; HWANG, J.I.; KAH, O.; SOWER, S.A.; VAUDRY, H.; SEONG, J.Y. 2011. Revisiting the evolution of gonadotropin-releasing hormones and their receptors in vertebrates: secrets hidden in genomes. *Gen Comp Endocrinol.* 170(1), pp. 68-78.

GARCIA DE LEANIZ, C.; FLEMING, I. A.; EINUM, S.; VERSPOOR, E.; JORDAN, W. C.; CONSUEGRA, S.; QUINN, T. P. 2007. A critical review of adaptive genetic variation in Atlantic salmon: implications for conservation. *Biological Reviews*, 82(2), pp. 173–211.

GOUVEIA, J. J.; DA SILVA, M.; PAIVA, S. R.; DE OLIVEIRA, S. 2014. Identification of selection signatures in livestock species. *Genetics and Molecular Biology*, 37(2), pp. 330–342.

DI PRINZIO, C. Y.; PASCUAL, M. A. 2008. The establishment of exotic Chinook salmon (*Oncorhynchus tshawytscha*) in Pacific rivers of Chubut, Patagonia, Argentina. *Annales de Limnologie - International Journal of Limnology*, 44(1), pp. 25–32.

DLUGOSCH, K. M.; ANDERSON, S. R.; BRAASCH, J.; CANG, F. A.; GILLETTE, H. D. 2015. The devil is in the details: Genetic variation in introduced populations and its contributions to invasion. *Molecular Ecology*, 24(9), 2095–2111.

ELIASON, E. J.; CLARK, T. D.; HAGUE, M. J.; HANSON, L. M.; GALLAGHER, Z. S.; JEFFRIES, K. M.; FARRELL, A. P. 2011. Differences in thermal tolerance among sockeye salmon populations. *Science*, 332(6025), 109–112.

ELST, E. M.; ACHARYA, K. P.; DAR, P. A.; RESHI, Z. A.; TUFTO, J.; NIJS, I.; GRAAE, B. J. 2016. Pre-adaptation or genetic shift after introduction in the invasive species *Impatiens glandulifera*?. *Acta Oecologica*, 70, pp. 60–66.

EXCOFFIER, L.; HOFER, T.; FOLL, M. 2009. Detecting loci under selection in a hierarchically structured population. *Heredity*, 103(4), 285–298.

FAO. 2020. The State of World Fisheries and Aquaculture 2020. Sustainability in action. Food and agriculture organization of the united nations.

FAY, J.; WU, C. 2000. Hitchhiking Under Positive Darwinian Selection. *Genetics*, 155(3), pp. 1405–1413.

FERNÁNDEZ, D. A.; CIANCIO, J.; CEBALLOS, S. G.; RIVA-ROSSI, C.; PASCUAL, M. A. 2010. Chinook salmon (*Oncorhynchus tshawytscha*, Walbaum 1792) in the Beagle Channel, Tierra del Fuego: The onset of an invasion. *Biological Invasions*, 12(9), pp. 2991–2997.

FOLL, M.; GAGGIOTTI, O. 2008. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: A Bayesian perspective. *Genetics*, 180(2), pp. 977–993.

FOURCADE, Y.; CHAPUT-BARDY, A.; SECONDI, J.; FLEURANT, C.; LEMAIRE, C. 2013. Is local selection so widespread in river organisms? Fractal geometry of river networks leads to high bias in outlier detection. *Molecular Ecology*, 22(8), 2065–2073.

FRASER, D. J.; WEIR, L. K.; BERNATCHEZ, L.; HANSEN, M. M.; TAYLOR, E. B. 2011. Extent and scale of local adaptation in salmonid fishes: Review and meta-analysis. *Heredity*, 106(3), pp. 404–420

GAILLARD, A.L.; TAY, B.H.; PÉREZ, D.I.; LAFONT, A.G.; DE FLORI, C.; VISSIO, P.G.; MAZAN, S.; DUFOUR, S.; VENKATESH, B.; TOSTIVINT, H. 2018. Characterization of Gonadotropin-Releasing Hormone (GnRH) Genes From Cartilaginous Fish: Evolutionary Perspectives. *Front. Neurosci.* 12:607.

GALLIANO, M. F.; HUET, C.; FRYGELIUS, J.; POLGREN, A.; WEWER, U. M.; ENGVALL, E. 2000. Binding of ADAM12, a marker of skeletal muscle regeneration, to the muscle-specific actin-binding protein, α -actinin-2, is required for myoblast fusion. *Journal of Biological Chemistry*, 275(18), pp. 13933–13939.

GARCÍA DE LEANIZ, C.; FLEMING, I. A.; EINUM, S.; VERSPOOR, E.; JORDAN, W. C.; CONSUEGRA, S.; QUINN, T. P. 2007. A critical review of adaptive genetic variation in Atlantic salmon: implications for conservation. *Biological Reviews*, 82(2), 173–211. doi: 10.1111/j.1469-185X.2006.00004.x

GILLESPIE, J. H. 2001. Is the population size of a species relevant to its evolution? *Evolution*, 55(11), pp. 2161–2169.

GLOVER, K. A.; SOLBERG, M. F.; MCGINNITY, P.; HINDAR, K.; VERSPOOR, E.; COULSON, M. W.; HANSEN, M. M.; ARAKI, H.; SKAALA, Ø.; SVÅSAND, T. 2017. Half a century of genetic interaction between farmed and wild Atlantic salmon: Status of knowledge and unanswered questions. *Fish and Fisheries*, 18 (5), pp. 890-927.

GODDARD, M. E.; WRAY, N. R.; VERBYLA, K.; VISSCHER, P. M. 2009 Estimating Effects and Making Predictions from Genome-Wide Marker Data. *Statistical Science*, 24(4), pp. 517–529.

GÓMEZ-UCHIDA, D.; CAÑAS-ROJAS, D.; RIVA-ROSSI, C. M.; CIANCIO, J. E.; PASCUAL, M. A.; ERNST, B.; AEDO, E.; MUSLEH, S. S.; VALENZUELA-AGUAYO, F.; QUINN, T. P.; SEEB, J. E.; SEEB, L. W. 2018. Genetic signals of artificial and natural dispersal linked to colonization of South America by non-native Chinook salmon (*Oncorhynchus tshawytscha*) *Ecology and Evolution*, 8(12), pp. 6192–6209.

HAASL, R. J.; PAYSEUR, B. A. 2016. Fifteen years of genome-wide scans for selection: Trends, lessons and unaddressed genetic sources of complication. *Molecular Ecology*, 25(1), pp. 5–23.

HABIT, E.; GONZÁLEZ, J.; ORTIZ-SANDOVAL, J.; ELGUETA, A.; SOBENES, C. 2015. Effects of salmonid invasion in rivers and lakes of Chile. *Ecosistemas*, 24(1), pp. 43–51.

HANCOCK, A. M.; ALKORTA-ARANBURU, G.; WITONSKY, D. B.; DI RIENZO, A. 2010. Adaptations to new environments in humans: The role of subtle allele frequency shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365 (1552), pp. 2459–2468.

HANSEN, L. P.; QUINN, T.P. 1998. The marine phase of Atlantic salmon (*Salmo salar*) life cycle, with comparison to Pacific salmon. *Canadian Journal of Fisheries and Aquatic Sciences*, 55(S1), pp. 104-118.

HARD, J. J.; HEARD, W. R. 1999. Analysis of straying variation in Alaskan hatchery chinook salmon (*Oncorhynchus tshawytscha*) following transplantation. *Canadian Journal of Fisheries and Aquatic Sciences*. 56(4): pp. 578-589.

HARING, M. W.; JOHNSTON, T. A.; WIEGAND, M. D.; FISK, A. T.; PITCHER, T. E. 2016. Differences in egg quantity and quality among hatchery- and wild-origin Chinook salmon (*Oncorhynchus tshawytscha*). *Canadian Journal of Fisheries and Aquatic Sciences*, 73(5), pp. 737–746.

HAYES, K. R.; BARRY, S. C. 2008 Are there any consistent predictors of invasion success? *Biological Invasions*, 10(4), pp. 483–506.

HECHT, B. C.; MATALA, A. P.; HESS, J. E.; NARUM, S. R. 2015. Environmental adaptation in Chinook salmon (*Oncorhynchus tshawytscha*) throughout their North American range. *Molecular Ecology*, 24(22), pp. 5573–5595.

HENDRY, A.; CASTRIC, V.; KINNISON, M.; QUINN, T. 2004. The Evolution of Philopatry and Dispersal: Homing Versus Straying in Salmonids. In *Evolution Illuminated: Salmon and Their Relatives*. pp. 52–91.

HERMISSEON, J.; PENNING, P. S. 2005. Soft sweeps: Molecular population genetics of adaptation from standing genetic variation. *Genetics*, 169(4), pp. 2335–2352.

HERMISSEON, J.; PENNING, P. S. 2017. Soft sweeps and beyond: understanding the patterns and probabilities of selection footprints under rapid adaptation. *Methods in Ecology and Evolution*, 8, pp. 700-716.

HESS, J. E.; ZENDT, J. S.; MATALA, A. R.; NARUM, S. R. 2016. Genetic basis of adult migration timing in anadromous steelhead discovered through multivariate association testing. *Proceedings of the Royal Society B: Biological Sciences*, 283: 20153064.

HOBAN, S., KELLEY, J. L., LOTTERHOS, K. E., ANTOLIN, M. F., BRADBURY, G., LOWRY, D. B., WHITLOCK, M. C. 2016. Finding the genomic basis of local adaptation: Pitfalls, practical solutions, and future directions. *American Naturalist*, 188(4), pp. 379–397.

HOHENLOHE, P. A.; PHILLIPS, P. C.; CRESKO, W. A. 2010. Using Population Genomics to Detect Selection in Natural populations. *International Journal of Plant Science*, 171(9), pp. 1059–1071.

HOLSINGER, K. E.; WEIR, B. S. 2009 Genetics in geographically structured populations: Defining, estimating and interpreting FST. *Nature Reviews Genetics*, 10(9), pp. 639–650.

HOUDE, A. L.; WILSON, C.C.; NEFF, B. D. 2017. Performance of four salmonids species in competition with Atlantic salmon. *Journal of Great Lakes Research*, 43 (1), pp. 211-215.

HUGHES, A. R.; INOUYE, B. D.; JOHNSON, M. T. J.; UNDERWOOD, N.; VELLEND, M. 2008. Ecological consequences of genetic diversity. *Ecology Letters*, 11(6), pp. 609–623).

JOHNSON, J. E.; DEWITT, S. P.; GONDER, D. J. A. 2010. Mass-Marking Reveals Emerging Self Regulation of the Chinook Salmon Population in Lake Huron. *North American Journal of Fisheries Management*, 30(2), pp. 518–529.

JUSTICE, A. E.; WINKLER, T. W.; FEITOSA, M. F.; GRAFF, M.; FISHER, V. A.; YOUNG, K.; BARATA, L.; DENG, X.; CZAJKOWSKI, J.; HADLEY, D.; NGWA, J. S.; AHLUWALIA, T. S.; CHU, A. Y.; HEARD-COSTA, N. L.; LIM, E.; PEREZ, J.; EICHER, J. D.; KUTALIK, Z.; XUE, L. et al. 2017. Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. *Nature Communications*, 8(1), pp. 1–19.

KAEUFFER, R.; RÉALE, D.; COLTMAN, D. W.; PONTIER, D. 2007. Detecting population structure using STRUCTURE software: Effect of background linkage disequilibrium. *Heredity*, 99(4), pp. 374–380.

KENT, T. V.; UZUNOVIĆ, J.; WRIGHT, S. I. 2017. Coevolution between transposable elements and recombination. *Philosophical Transactions, Royal Society B*, 372, 20160458.

KELLER, S. R.; TAYLOR, D. R. 2008. History, chance and adaptation during biological invasion: Separating stochastic phenotypic evolution from response to selection. *Ecology Letters*, 11(8), pp. 852–866.

KELLEY, J. L.; SWANSON, W. J. 2008. Positive Selection in the Human Genome: From Genome Scans to Biological Significance. *Annual Review of Genomics and Human Genetics*, 9(1), pp. 143–160.

- KELLY, D. W.; MUIRHEAD, J. R.; HEATH, D. D.; MACISAAC, H. J.** 2006. Contrasting patterns in genetic diversity following multiple invasions of fresh and brackish waters. *Molecular Ecology*, 15(12), pp. 3641–3653.
- KIMURA, M.** 1968. Evolutionary rate at the molecular level. *Nature*, 217(5129), pp. 624–626.
- KINNISON, M. T.; QUINN, T. P.; UNWIN, M. J.** 2011. Correlated contemporary evolution of life history traits in New Zealand Chinook salmon, *Oncorhynchus tshawytscha*. *Heredity*, 106(3), pp. 448–459.
- KUO, K. H. M.** 2017. Multiple Testing in the Context of Gene Discovery in Sickle Cell Disease Using Genome-Wide Association Studies. *Genomics Insights*, 10(1), 117863101772117.
- LACHANCE, J.; TISHKOFF, S. A.** 2013. SNP ascertainment bias in population genetic analysis. *Bioessays*, 35(9), pp. 780–786.
- LARSEN, D. A.; SWANSON, P.; DICKHOFF, W. W.** 2011. The pituitary-thyroid axis during the parr-smolt transformation of Coho salmon, *Oncorhynchus kisutch*: Quantification of TSH β mRNA, TSH, and thyroid hormones. *General and Comparative Endocrinology*, 171(3), pp. 367–372.
- LARSON, W. A.; SEEB, J. E.; PASCAL, C. E.; TEMPLIN, W. D.; SEEB, L. W.** 2014. Single-nucleotide polymorphisms (SNPs) identified through genotyping-by-sequencing improve genetic stock identification of chinook salmon (*Oncorhynchus tshawytscha*) from western Alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, 71(5), pp. 698–708.
- LAWSON, D. J.; VAN DORP, L.; FALUSH, D.** 2018. A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. *Nature Communications*, 9, 3258.
- LEMOPOULOS, A.; UUSI-HEIKKILÄ, S.; HYVÄRINEN, P.; ALIORAVAINEN, N.; PROKKOLA, J. M.; ELVIDGE, C. K.; VASEMÄGI, A.; VAINIKKA, A.** 2019. Association mapping based on a common-garden migration experiment reveals candidate genes for migration tendency in brown trout. *G3: Genes, Genomes, Genetics*, 9(9), pp. 2887–2896.
- LI, N.; ZHOU, T.; GENG, X.; JIN, Y.; WANG, X.; LIU, S.; LIU, Z.** (2018). Identification of novel genes significantly affecting growth in catfish through GWAS analysis. *Molecular Genetics and Genomics*, 293(3), 587–599.

LOCKE, D. P.; HILLIER, L. W.; WARREN, W. C.; WORLEY, K. C.; NAZARETH, L. V.; MUZNY, D. M.; YANG, S. P.; WANG, Z.; CHINWALLA, A. T.; MINX, P.; MITREVA, M.; COOK, L.; DELEHAUNTY, K. D.; FRONICK, C.; SCHMIDT, H.; FULTON, L. A.; FULTON, R. S.; NELSON, J. O.; MAGRINI, V. et al. 2011 Comparative and demographic analysis of orangutan genomes. *Nature*, 469(7331), pp. 529–533.

LÓPEZ, M.E.; CÁDIZ, M.I.; RONDEAU, E.B.; KOOP, B. F.; YÁÑEZ, J. M. 2021. Detection of selection signatures in farmed coho salmon (*Oncorhynchus kisutch*) using dense genome-wide information. *Scientific Reports* 11, 9685.

LÓPEZ, M. E.; NEIRA, R.; YÁÑEZ, J. M. 2015. Applications in the search for genomic selection signatures in fish. *Frontiers in Genetics*, 5, p. 458.

LOTTERHOS, K. E.; SCHAAL, S. M. 2014. Genome scans for the contemporary response to selection in quantitative traits. *Molecular Ecology*, 23(18), pp. 4435–4437.

LOTTERHOS, K. E.; WHITLOCK, M. C. 2014. Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Molecular Ecology*, 23(9), pp. 2178–2192.

LUPSKI, J. R.; STANKIEWICZ, P. 2005. Genomic Disorders: Molecular Mechanisms for Rearrangements and Conveyed Phenotypes. *PLOS Genetics*, 1(6), e49.

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(1): pp. 67-77.

MALDONADO-MÁRQUEZ, A.; CONTADOR, T.; RENDOLL-CÁRCAMO, J.; MOORE, S.; PÉREZ-TRONCOSO, C.; GÓMEZ-UCHIDA, D.; HARROD, C. 2020. Southernmost distribution limit for endangered Peladillas (*Aplochiton taeniatus*) and non-native coho salmon (*Oncorhynchus kisutch*) coexisting within the Cape Horn biosphere reserve, Chile. *Journal of Fish Biology*, 96(6), pp. 1495–1500.

MATALA, A. P.; ACKERMAN, M. W.; CAMPBELL, M. R.; NARUM, S. R. 2014. Relative contributions of neutral and non-neutral genetic differentiation to inform conservation of steelhead trout across highly variable landscapes. *Evolutionary Applications*, 7(6), 682–701.

MCDOWALL, R. 2001. Anadromy and homing: two life-history traits with adaptive synergies in salmonid fishes? *Fish and Fisheries*, 2(1), pp. 78–85.

MILLER, M. R.; BRUNELLI, J. P.; WHEELER, P. A.; LIU, S.; REXROAD, C. E.; PALTÍ, Y.; THORGAARD, G. H. 2012. A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Molecular Ecology*, 21 (2), pp. 237–249.

MUÑOZ, N. J., FARRELL, A. P., HEATH, J. W., & NEFF, B. D. 2015. Adaptive potential of a Pacific salmon challenged by climate change. *Nature Climate Change*, 5(2), pp. 163–166.

NARUM, S.; GALLARDO, P.; CORREA, C.; MATALA, A.; HASSELMAN, D.; SUTHERLAND, B.; BERNATCHEZ, L. 2017. Genomic patterns of diversity and divergence of two introduced salmonid species in Patagonia, South America. *Evolutionary Applications*, 10(4), pp. 402–416.

NARUM, S. R.; HESS, J. E. 2011. Comparison of FST outlier tests for SNP loci under selection. *Molecular Ecology Resources*, 11, pp. 184–194.

NEI, M. 1978. Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics*, 89(3), pp 583-590.

NIKLITSCHKEK, E. J.; SOTO, D.; LAFON, A.; MOLINET, C.; TOLEDO, P. 2013. Southward expansion of the Chilean salmon industry in the Patagonian Fjords: Main environmental challenges. *Reviews in Aquaculture*, 5(3), pp. 172–195.

NIKLITSCHKEK, E.; TOLEDO, P. 2011. “Evaluación cuantitativa del estado trófico de salmónidos de vida libre en el fiordo Aysén, XI región”. Informe Final Proyecto FIP 2008--30.

O’NEAL, S. L.; STANFORD, J. A. 2011. Partial Migration in a Robust Brown Trout Population of a Patagonian River. *Transactions of the American Fisheries Society*, 140(3), pp. 623–635.

OLEKSYK, T. K.; SMITH, M. W.; O’BRIEN, S. J. 2010. Genome-wide scans for footprints of natural selection. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 365(1537), pp. 185–205.

OLSON-MANNING, C. F.; WAGNER, M. R.; MITCHELL-OLDS, T. 2012. Adaptive evolution: Evaluating empirical support for theoretical predictions. *Nature Reviews Genetics*, 13(12), pp. 867–877.

PARKER, J. D.; TORCHIN, M. E.; HUFBAUER, R. A.; LEMOINE, N. P.; ALBA, C.; BLUMENTHAL, D. M.; BOSSDORF, O.; BYERS, J. E.; DUNN, A. M.; HECKMAN,

R. W.; HEJDA, M.; CH JAROŠIČ, V.; JAROŠI, J.; JAROŠIČ, J.; KANAREK, A. R.; MARTIN, L. B.; PERKINS, S. E.; PYČ, P.; SEK, P. et al. 2013. Do invasive species perform better in their new ranges? In *Ecology* 94(5), pp. 985-994.

PARSONS, J. W. 1973. History of salmon in the Great Lakes, 1850-1970. Technical Papers of the Bureau of Sport Fisheries and Wildlife, 68, pp. 1–80.

PASCUAL, M. A.; CIANCIO, J. E. 2007. Introduced Anadromous Salmonids in Patagonia: Risks, Uses, and a Conservation Paradox. In: *Ecological and Genetic Implications of Aquaculture Activities* pp. 333–353.

PASCUAL, M.; BENTZEN, P.; ROSSI, C. R.; MACKEY, G.; KINNISON, M.; WALKER, R. 2001. First Documented Case of Anadromy in a Population of Introduced Rainbow Trout in Patagonia, Argentina. *Transactions of the American Fisheries Society*, 130, pp. 53–67.

PECK, J. W.; JONES, T. S.; MACCALLUM, W. R.; SCHRAM, S. T. 1999. Contribution of Hatchery-Reared Fish to Chinook Salmon Populations and Sport Fisheries in Lake Superior. *North American Journal of Fisheries Management*, 19(1), pp. 155–164.

PEMBLETON, L. W.; COGAN, N. O.; FORSTER, J. W. 2013. StAMPP: an R package for calculation of genetic differentiation and structure of mixed-ploidy level populations. *Molecular Ecology Resources*, 13(5), pp. 946-952.

PESS, G. R.; HILBORN, R.; KLOEHN, K.; QUINN, T. P. 2012. The influence of population dynamics and environmental conditions on pink salmon (*Oncorhynchus gorbuscha*) recolonization after barrier removal in the Fraser River, British Columbia, Canada. *Canadian Journal of Fisheries and Aquatic Sciences*. 69 (5), pp. 970-982.

PIMENTEL, D.; MCNAIR, S.; JANECKA, J.; WIGHTMAN, J.; SIMMONDS, C.; O'CONNELL, C.; WONG, E.; RUSSEL, L.; ZERN, J.; AQUINO, T.; TSOMONDO, T. 2001. Economic and environmental threats of alien plant, animal, and microbe invasions. *Agriculture, Ecosystems and Environment*, 84(1), pp. 1–20.

PRINCE, D. J.; O'ROURKE, S. M.; THOMPSON, T. Q.; ALI, O. A.; LYMAN, H. S.; SAGLAM, I. K.; HOTALING, T. J.; SPIDLE, A. P.; MILLER, M. R. 2017. The evolutionary basis of premature migration in Pacific salmon highlights the utility of genomics for informing conservation. *Science Advances*, 3(8), e1603198.

PRITCHARD, J. K.; DI RIENZO, A. 2010. Adaptation - Not by sweeps alone. *Nature Reviews Genetics*, 11(10), pp. 665–667.

PRITCHARD, J. K.; PICKRELL, J. K.; COOP, G. 2010. The Genetics of Human Adaptation: Hard Sweeps, Soft Sweeps, and Polygenic Adaptation. In *Current Biology*, 20(4), 208-215.

PRZEWORSKI, M.; COOP, G.; WALL, J. D.; PRZEWORSKI, M.; COOP, G.; WALL, J. D. 2005 The Signature of Positive Selection on Standing Genetic Variation. *International Journal of Organic Evolution*, 59(11), 2312–2323.

PURCELL, S.; NEALE, B.; TODD-BROWN, K.; THOMAS, L.; FERREIRA, M. A. R.; BENDER, D.; MALLER, J.; SKLAR, P.; BAKKER, P. I. W. DE, DALY, M. J.; SHAM, P. C. 2007. PLINK: A Tool Set for Whole-Genome Association and Population-Based Linkage Analyses. *American Journal of Human Genetics*, 81, pp. 559–575.

PUZEY, J.; VALLEJO-MARÍN, M. 2014. Genomics of invasion: Diversity and selection in introduced populations of monkeyflowers (*Mimulus guttatus*). *Molecular Ecology*, 23(18), pp. 4472–4485.

PYŠEK, P.; RICHARDSON, D. M. 2010. Invasive Species, Environmental Change and Management, and Health. *Annual Review of Environment and Resources*, 35(1), pp. 25–55.

QANBARI, S.; GIANOLA, D.; HAYES, B.; SCHENKEL, F.; MILLER, S.; MOORE, S.; THALLER, G.; SIMIANER, H. 2011. Application of site and haplotype-frequency based approaches for detecting selection signatures in cattle. *BMC Genomics*, 12(1), 318.

QANBARI, S.; SIMIANER, H. 2014. Mapping signatures of positive selection in the genome of livestock. *Livestock Science*, 166(1), pp. 133–143.

QIU, S.; CHAMPAGNE, D. L.; PETERS, M.; CATANIA, E. H.; WEEBER, E. J.; LEVITT, P.; PIMENTA, A. F. 2010. Loss of limbic system-associated membrane protein leads to reduced hippocampal mineralocorticoid receptor expression, impaired synaptic plasticity, and spatial memory deficit. *Biological Psychiatry*, 68(2), pp. 197–204.

QUINN T.P. 1984. Homing and Straying in Pacific Salmon. In: *Mechanisms of Migration in Fishes*. NATO Conference Series (IV Marine Sciences), 14, pp. 357-362.

QUINN, T. P. 2005. *The Behavior and Ecology of Pacific Salmon and Trout* (Vol. 1).

QUINN, T. P.; KINNISON, M. T.; UNWIN, M. J. 2001. Evolution of chinook salmon (*Oncorhynchus tshawytscha*) populations in New Zealand: Pattern, rate, and process. *Genetica*, 112–113, pp. 493–513.

QUINN, T. P.; UNWIN, M. J.; KINNISON, M. T. 2011. Contemporary divergence in migratory timing of naturalized populations of chinook salmon, *Oncorhynchus tshawytscha*, in New Zealand. *Evolutionary Ecology Research*, 13, pp. 45–54.

QUIÑONES, R. A.; FUENTES, M.; MONTES, R. M.; SOTO, D.; LEÓN-MUÑOZ, J. 2019. Environmental issues in Chilean salmon farming: a review. *Reviews in Aquaculture*, 11(2), pp. 375–402.

R CORE TEAM. 2020. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

RICCIARDI, A.; ATKINSON, S. K. 2004. Distinctiveness magnifies the impact of biological invaders in aquatic ecosystems. *Ecology Letters*, 7(9), pp. 781–784.

RITZ, K. NOOR, M. A.; SINGH, N. D. 2017. Variation in Recombination Rate: Adaptive or Not?. *Trends in genetics*, 33(5), pp. 364-374.

RIVA ROSSI, C. M.; PASCUAL, M. A.; AEDO MARCHANT, E.; BASSO, N.; CIANCIO, J. E.; MEZGA, B.; FERNÁNDEZ, D. A.; ERNST-ELIZALDE, B. 2012. The invasion of Patagonia by Chinook salmon (*Oncorhynchus tshawytscha*): Inferences from mitochondrial DNA patterns. *Genetica*, 140(10–12), pp. 439–453.

RODRÍGUEZ, F.H; FLORES-MARA, R.; YOSHIDA, G.M.; BARRÍA, A.; JEDLICKI, A.M.; LHORENTE, J.P.; REYES-LÓPEZ, F.; YÁÑEZ, J.M. 2019. Genome-Wide Association Analysis for Resistance to Infectious Pancreatic Necrosis Virus Identifies Candidate Genes Involved in Viral Replication and Immune Response in Rainbow Trout (*Oncorhynchus mykiss*). *G3: Genes, Genomes, Genetics*, 9(9), pp. 2897-2904.

ROMAN, J.; DARLING, J. A. 2007. Paradox lost: genetic diversity and the success of aquatic invasions. In *Trends in Ecology and Evolution*, 22(9), pp. 454–464.

SABETI, P. C.; SCHAFFNER, S. F.; FRY, B.; LOHMUELLER, J.; VARILLY, P.; SHAMOVSKY, O.; PALMA, A.; MIKKELSEN, T. S.; ALTSHULER, D.; LANDER, E. S. 2006. Positive natural selection in the human lineage. *Science*, 312(5780), pp. 1614–1620.

SABETI, PARDIS C.; VARILLY, P.; FRY, B.; LOHMUELLER, J.; HOSTETTER, E.; COTSAPAS, C.; XIE, X.; BYRNE, E. H.; MCCARROLL, S. A.; GAUDET, R.;

SCHAFFNER, S. F.; LANDER, E. S.; FRAZER, K. A.; BALLINGER, D. G.; COX, D. R.; HINDS, D. A.; STUVE, L. L.; GIBBS, R. A.; BELMONT, J. W. et al. 2007. Genome-wide detection and characterization of positive selection in human populations. *Nature*, 449(7164), pp. 913–918.

SCHRIDER, D. R.; MENDES, F. K.; HAHN, M. W.; KERN, A. D. 2015. Soft shoulders ahead: Spurious signatures of soft and partial selective sweeps result from linked hard sweeps. *Genetics*, 200(1), pp. 267–284.

SCHRÖDER, V.; DE LEANIZ, C. G. 2011. Discrimination between farmed and free-living invasive salmonids in Chilean Patagonia using stable isotope analysis. *Biological Invasions*, 13(1), pp. 203–213.

SEPÚLVEDA, M.; ARISMENDI, I.; SOTO, D.; JARA, F.; FARIAS, F. 2013. Escaped farmed salmon and trout in Chile: Incidence, impacts, and the need for an ecosystem view. *Aquaculture Environment Interactions* 4(3), pp. 273–283.

SIMBERLOFF, D. 2009. The Role of Propagule Pressure in Biological Invasions. *Annual Review of Ecology, Evolution, and Systematics*, 40(1), pp. 81–102.

SMITH, J. M.; HAIGH, J. 1974. The hitch-hiking effect of a favourable gene. *Genetical Research*, 23(1), pp. 23–35.

SOBOTA, R. S.; SHRINER, D.; KODAMAN, N.; GOODLOE, R.; ZHENG, W.; GAO, Y. T.; EDWARDS, T. L.; AMOS, C. I.; WILLIAMS, S. M. 2015. Addressing population-specific multiple testing burdens in genetic association studies. *Annals of Human Genetics*, 79(2), 136–147.

SOLBERG, M. F.; ROBERTSEN, G.; SUNDT-HANSEN, L. E.; HINDAR, K.; GLOVER, K. 2020. *Scientific Reports*, 10, p. 1929.

SOTO, D.; ARISMENDI, I.; DI PRINZIO, C.; JARA, F. 2007. Establishment of Chinook salmon (*Oncorhynchus tshawytscha*) in Pacific basins of southern South America and its potential ecosystem implications. *Revista Chilena de Historia Natural*, 80(1), pp. 81–98.

SOTO, D.; ARISMENDI, I.; GONZÁLEZ, J.; SANZANA, J.; JARA, F.; JARA, C.; GUZMAN, E.; LARA, A. 2006. Southern Chile, trout and salmon country: Invasion patterns and threats for native species. *Revista Chilena de Historia Natural*, 79(1), pp. 97–117.

SOTO, D.; JARA, F.; MORENO, C. 2001. Escaped salmon in the inner seas, southern Chile: Facing ecological and social conflicts. *Ecological Applications*, 11(6), pp. 1750–1762.

STOREY, J.D.; BASS, A. J.; DABNEY, A.; ROBINSON, D. 2020. qvalue: Q-value estimation for false discovery rate control. R package version 2.20.0.

STOREY, JOHN D.; XIAO, W.; LEEK, J. T.; TOMPKINS, R. G.; DAVIS, R. W. 2005. Significance analysis of time course microarray experiments. *Proceedings of the National Academy of Sciences of the United States of America*, 102(36), pp. 12837–12842.

STORZ, J. F. 2005. Using genome scans of DNA polymorphism to infer adaptive population divergence. *Molecular Ecology* 14(3), pp. 671–688.

SUK, H. Y.; NEFF, B. D.; QUACH, K.; MORBEY, Y. E. 2012. Evolution of introduced Chinook salmon (*Oncorhynchus tshawytscha*) in Lake Huron: Emergence of population genetic structure in less than 10 generations. *Ecology of Freshwater Fish*, 21(2), pp. 235–244.

SZÚCS, M.; MELBOURNE, B. A.; TUFF, T.; HUFBAUER, R. A. 2014. The roles of demography and genetics in the early stages of colonization. *Proceedings of the Royal Society B: Biological Sciences*, 281(1792), 20141073.

TANG, H.; LYONS, E.; PEDERSEN, B.; SCHNABLE, J.C.; PATERSON, A.H.; FREELING, M. 2011. Screening synteny blocks in pairwise genome comparisons through integer programming. *BMC Bioinformatics*. 12(102), pp. 1-11.

TEO, Y. Y.; FRY, A. E.; CLARK, T. G.; TAI, E. S.; SEIELSTAD, M. 2007. On the Usage of HWE for Identifying Genotyping Errors. *Annals of Human Genetics*, 71(5), pp. 701–703.

TESHIMA, K. M.; COOP, G.; PRZEWORSKI, M. 2006 How reliable are empirical genomic scans for selective sweeps? *Genome Research*, 16(6), pp. 702–712.

THÉVENIN, A.; EIN-DOR, L.; OZERY-FLATO, M.; SHAMIR, R. 2014. Functional gene groups are concentrated within chromosomes, among chromosomes and in the nuclear space of the human genome. *Nuclei Acids Research*, 42(15), pp. 9854-9861.

THORN, M. W.; MORBEY, Y. E. 2018. Egg size and the adaptive capacity of early life history traits in Chinook salmon (*Oncorhynchus tshawytscha*). *Evolutionary Applications*, 11(2), pp. 205–219.

TOKUMASU, Y.; IIDA, A.; WANG, Z.; ANSAI, S.; KINOSHITA, M.; SEHARA-FUJISAWA, A. 2016. ADAM12-deficient zebrafish exhibit retardation in body growth at the juvenile stage without developmental defects. *Development, Growth and Differentiation*, 58(4), pp. 409–421.

TOMALTY, K. M. H.; MEEK, M. H.; STEPHENS, M. R.; RINCÓN, G.; FANGUE, N. A.; MAY, B. P.; BAERWALD, M. R. 2015. Transcriptional response to acute thermal exposure in juvenile Chinook salmon determined by RNAseq. *G3: Genes, Genomes, Genetics*, 5(7), pp. 1335–1349.

TSAI, H. Y.; HAMILTON, A.; TINCH, A. E.; GUY, D. R.; GHARBI, K.; STEAR, M. J.; HOUSTON, R. D. 2015. Genome wide association and genomic prediction for growth traits in juvenile farmed Atlantic salmon using a high density SNP array. *BMC Genomics*, 16 (1), 969.

TURNER, S.; ARMSTRONG, L. L.; BRADFORD, Y.; CARLSONY, C. S.; CRAWFORD, D. C.; CRENSHAW, A. T.; DE ANDRADE, M.; DOHENY, K. F.; HAINES, J. L.; HAYES, G.; JARVIK, G.; JIANG, L.; KULLO, I. J.; LI, R.; LING, H.; MANOLIO, T. A.; MATSUMOTO, M. M.; MCCARTY, C. A.; MCDAVID, A. N.; MIREL, D.; PASCHALL, J.; PUGH, E.; RASMUSSEN, L.; WILKE, R.; ZUVICH, R.; RITCHIE, M. D. 2011. Quality control procedures for genome-wide association studies. *Current Protocols in Human Genetics*, 68, pp. 1.19.1-1.19.18.

UEDA, H. 2012. Physiological mechanisms of imprinting and homing migration in Pacific salmon *Oncorhynchus* spp. *Journal of Fish Biology*, 81(2), pp. 543–558.

UEDA, H. 2019. Sensory mechanisms of natal stream imprinting and homing in *Oncorhynchus* spp. *Journal of Fish Biology*, 95(1), pp. 293–303.

UEDA, H., NAKAMURA, S., NAKAMURA, T. 2016. Involvement of hormones in olfactory imprinting and homing in chum salmon. *Scientific Reports*, 6, 21102.

UTSUNOMIYA, Y. T.; PÉREZ O'BRIEN, A. M. P.; SONSTEGARD, T. S.; SÖLKNER, J.; GARCIA, J. F. 2015. Genomic data as the “hitchhiker’s guide” to cattle adaptation: Tracking the milestones of past selection in the bovine genome. *Frontiers in Genetics* 6, 36.

VAN ZWOL, J. A.; NEFF, B. D.; WILSON, C. C. 2012. The effect of competition among three salmonids on dominance and growth during the juvenile life stage. *Ecology of freshwater fish*, 21(4), pp. 533-540.

VANDEPITTE, K.; DE MEYER, T.; HELSEN, K.; VAN ACKER, K.; ROLDÁN-RUIZ, I.; MERGEAY, J.; HONNAY, O. 2014. Rapid genetic adaptation precedes the spread of an exotic plant species. *Molecular Ecology*, 23(9), pp. 2157–2164.

VARGAS, P. V.; ARISMENDI, I.; LARA, G.; MILLAR, J.; PEREDO, S. 2010. Evidencia de solapamiento de micro-hábitat entre juveniles del salmón introducido *Oncorhynchus tshawytscha* y el pez nativo *Trichomycterus areolatus* en el río Allipén, Chile. *Revista de Biología Marina y Oceanografía*, 45(2), pp. 285–292.

VARGAS-CHACOFF, L.; MUÑOZ, J.L.P.; HAWES, C.; OYARZÚN, R.; PONTIGO, J.P.; SARAVIA, J.; GONZÁLEZ, M.P.; MARDONES, O.; LABBÉ, B.S.; MORERA, F.J.; BERTRÁN, C.; PINO, J.; WADSWORTH, S.; YÁÑEZ, A. 2017. Ectoparasite *Caligus rogercresseyi* modifies the lactate response in Atlantic salmon (*Salmo salar*) and Coho salmon (*Oncorhynchus kisutch*). *Veterinary Parasitology*, 243:6-11.

VIANNA, J.A.; FERNANDES, F. A.; FRUGONE, M. J.; FIGUEIRÓ, H. V.;

VITTI, J. J.; GROSSMAN, S. R.; SABETI, P. C. 2013. Detecting Natural Selection in Genomic Data. *Annual Review of Genetics*, 47(1), 97–120.

VITULE, J. R. S.; FREIRE, C. A.; VAZQUEZ, D. P.; NUÑEZ, M. A.; SIMBERLOFF, D. 2012. Revisiting the Potential Conservation Value of Non-Native Species. *Conservation Biology*, 26(6), 1153–1155.

VOIGHT, B. F.; KUDARAVALLI, S.; WEN, X.; PRITCHARD, J. K. 2006. A map of recent positive selection in the human genome. *PLoS Biology*, 4(3), 0446–0458.

VON THADEN, A.; COCCHIARARO, B.; JARAUSCH, A.; JÜNGLING, H.; KARAMANLIDIS, A. A.; TIESMEYER, A.; NOWAK, C.; MUÑOZ-FUENTES, V. 2017. Assessing SNP genotyping of noninvasively collected wildlife samples using microfluidic arrays. *Scientific Reports*, 7, 10768.

WANG, X.; LIU, X.; DENG, D.; YU, M.; LI, X. 2016. Genetic determinants of pig birth weight variability. *BMC Genetics*, 17(1), S15.

WANG, J.; GAUGHAN, S.; LAMER, J. T.; DENG, C.; HU, W.; WACHHOLTZ, M.; QIN, S.; NIE, H.; LIAO, X.; LING, Q.; LI, W.; ZHU, L.; BERNATCHEZ, L.; WANG, C.; LU, G. 2020. Resolving the genetic paradox of invasions: Preadapted genomes and postintroduction hybridization of bigheaded carps in the Mississippi River Basin. *Evolutionary Applications*, 13(2), pp. 263–277.

WAPLES, R. S.; TEEL, D. J.; MYERS, J. M.; MARSHALL, A. R. 2004. Life-history divergence in Chinook salmon: historic contingency and parallel evolution. *Evolution*, 58(2), p. 386.

WEEDER, J. A.; MARSHALL, A. R.; EPIFANIO, J. M. 2005. An Assessment of Population Genetic Variation in Chinook Salmon from Seven Michigan Rivers 30 Years after Introduction. *North American Journal of Fisheries Management*, 25(3), pp. 861–875.

WEIGAND, H.; LEESE, F. 2018. Detecting signatures of positive selection in non-model species using genomic data. *Zoological Journal of the Linnean Society*, 184(2), pp. 528–583.

WESTLEY, P. A. H.; STANLEY, R.; FLEMING, I. A. 2013. Experimental Tests for Heritable Morphological Color Plasticity in Non-Native Brown Trout (*Salmo trutta*) Populations. *PLoS ONE*, 8(11), e80401.

WESTLEY, P. A. H.; WARD, E. J.; FLEMING, I. A. 2013. Fine-scale local adaptation in an invasive freshwater fish has evolved in contemporary time. *Proceedings of the Royal Society B: Biological Sciences*, 280(1751), 20122327.

WRIGHT, B.; FARQUHARSON, K. A.; MCLENNAN, E. A.; BELOV, K.; HOGG, C. J.; GRUEBER, C. E. 2019. From reference genomes to population genomics: Comparing three reference-aligned reduced-representation sequencing pipelines in two wildlife species. *BMC Genomics*, 20(1), 453.

WRIGHT, S. 1951. The genetical structure of populations. *Annals of Eugenics*, 15(1), pp. 323–354.

YOSHIDA, G. M., LHORENTE, J. P., CARVALHEIRO, R., YÁÑEZ, J. M. 2017. Bayesian genome-wide association analysis for body weight in farmed Atlantic salmon (*Salmo salar* L.). *Animal Genetics*, 48(6), pp. 698–703.

YOSHIDA, G.M.; YÁÑEZ, J. M. 2021. Multi-trait GWAS using imputed high-density genotypes from whole-genome sequencing identifies genes associated with body traits in Nile tilapia. *BMC Genomics*, 22(1), 57.

YOUNG, K. A.; STEPHENSON, J.; TERREAU, A.; THAILLY, A. F.; GAJARDO, G.; DE LEANIZ, C. G. 2009. The diversity of juvenile salmonids does not affect their competitive impact on a native galaxiid. *Biological Invasions*, 11(8), pp. 1955–1961.

ZUEVA KJ, LUMME J, VESELOV AE, KENT MP, LIEN S, PRIMMER CR. 2014. Footprints of directional selection in wild Atlantic salmon populations: evidence for parasite-driven evolution? PLoS One, 9(3): e91672.