



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS VETERINARIAS Y PECUARIAS
ESCUELA DE CIENCIAS VETERINARIAS



DETERMINACIÓN DEL GRADO DE DESEQUILIBRIO DE
LIGAMIENTO EN EL CROMOSOMA ASOCIADO CON
RESISTENCIA A NECROSIS PANCREÁTICA INFECCIOSA EN EL
SALMÓN DEL ATLÁNTICO (SALMO SALAR)

PHILLIP DETTLEFF FAÚNDES

Memoria para optar al Título
Profesional de Médico Veterinario
Departamento de Fomento y
Producción Animal

PROFESOR GUÍA: VÍCTOR MARTÍNEZ MONCADA

SANTIAGO, CHILE
2012



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS VETERINARIAS Y PECUARIAS
ESCUELA DE CIENCIAS VETERINARIAS



DETERMINACIÓN DEL GRADO DE DESEQUILIBRIO DE
LIGAMIENTO EN EL CROMOSOMA ASOCIADO CON
RESISTENCIA A NECROSIS PANCREÁTICA INFECCIOSA EN EL
SALMÓN DEL ATLÁNTICO (SALMO SALAR)

PHILLIP DETTLEFF FAÚNDES

Memoria para optar al Título
Profesional de Médico Veterinario
Departamento de Fomento y
Producción Animal

NOTA FINAL:

	NOTA	FIRMA
PROFESOR GUÍA : VÍCTOR MARTÍNEZ MONCADA
PROFESOR CONSEJERO: MÓNICA DE LOS REYES SOLOVERA
PROFESOR CONSEJERO: PATRICIO RETAMAL MERINO

SANTIAGO, CHILE
2012

**DETERMINACIÓN DEL GRADO DE DESEQUILIBRIO DE
LIGAMIENTO EN EL CROMOSOMA ASOCIADO CON
RESISTENCIA A NECROSIS PANCREÁTICA INFECCIOSA EN
EL SALMÓN DEL ATLÁNTICO (SALMO SALAR)**

**DETERMINATION OF LINKAGE DISEQUILIBRIUM ON THE
CHROMOSOME ASSOCIATED WITH INFECTIOUS
PANCREATIC NECROSIS RESISTANCE IN ATLANTIC
SALMON (SALMO SALAR)**

Phillip James Dettleff Faúndes.

Departamento de Fomento y Producción Animal, Laboratorio de
Investigaciones en Biotecnología y Genómica Animal (FAVET-
INBIOGEN).

RESUMEN

La Necrosis Pancreática Infecciosa (IPN) es una enfermedad viral altamente contagiosa que afecta a los salmónidos, causando grandes mortalidades en individuos jóvenes. Se ha observado que la resistencia a esta enfermedad tiene un fuerte componente genético, identificado por un loci de rasgos cuantitativos (QTL) dentro del grupo de ligamiento 21. El objetivo de este trabajo consiste en identificar huellas de selección asociadas con la resistencia a IPN en poblaciones chilenas de salmón del Atlántico. Para esto se genotiparon cuatro poblaciones, calculando el desequilibrio de ligamiento (LD) dentro del grupo de ligamiento 21, buscando si existían diferencias para la región asociada al QTL. Los resultados mostraron que existen diferencias significativas en esta región en una de las poblaciones. De este modo se podría concluir que ésta población se vio sometida a un proceso selectivo de el o los genes presentes en la región responsable de la resistencia a IPN, modificando la estructura genética de la zona, presentando una huella de selección. Estos resultados muestran que ha existido un proceso selectivo para éste rasgo, que puede ser identificable a nivel del genoma. Sin embargo, es necesario tomar en cuenta la influencia de un proceso de admixtura, por lo que los resultados deben tomarse con precaución.

ABSTRACT

Infectious pancreatic necrosis (IPN) is a highly contagious viral disease of salmonids, causing high mortality in young individuals. It is known that resistance to this disease has a strong genetic component, represented by quantitative trait loci (QTL) in linkage group 21. The aim of this study is to identify signatures of selection associated with resistance to IPN in Chilean populations of Atlantic salmon. Four populations were genotyped, estimating the linkage disequilibrium in linkage group 21, searching differences in the region associated with the QTL. The results showed that one population has significant differences in this region. Thus, one might conclude that this population was subjected to a selective process of the genes in the region responsible for IPN resistance, modifying the genetic structure of the area, presenting a selection mark. These results show that has existed a selective process for this trait, which can be identified at genome level. However, it is necessary to take into account the influence of an admixture process, so the results should be taken with caution.

INTRODUCCION

La necrosis pancreática infecciosa es una enfermedad viral que afecta a los salmónidos, generando importantes mortalidades, en particular en individuos jóvenes (Ozaki *et al*, 2000). A nivel nacional, según datos del Fondo de Investigación Pesquera (FIP), los brotes de IPN han afectado hasta el 48% de las pisciculturas nacionales, provocando enormes pérdidas económicas al sector (FIP, 2003).

El IPN es causado por un virus de doble hebra de ARN perteneciente a la familia Birnaviridae (Houston *et al*, 2008). Este agente produce una necrosis en las células pancreáticas en forma primaria, afectando posteriormente a las células adiposas, comprometiendo en las fases terminales a las células de los túbulos renales, hematopoyéticas y hepáticas, generando letargo y muerte. Se transmite de forma vertical y horizontal, afectando al salmón tanto en la etapa de agua dulce (alevín) como en la etapa post-cría (post-smolt), posterior a la transferencia al agua de mar (Houston *et al*, 2008). Con el fin de combatir ésta infección se han desarrollado vacunas contra el IPN, sin embargo la protección inmunológica es deficiente, existiendo mortalidades significativas a pesar de la inmunización de la población (Moen *et al*, 2009). Adicionalmente, éstas no pueden utilizarse en peces muy jóvenes, quedando desprotegidos contra la enfermedad en la etapa de alevín (Gheyas *et al*, 2010).

En relación con la respuesta de los salmónidos, se ha observado en desafíos experimentales que existe un componente genético asociado con la resistencia a IPN, tanto en trucha arcoíris (Ozaki *et al*, 2000) como en salmón del Atlántico (Houston *et al*, 2008). Conjuntamente, se ha visto que la heredabilidad de esta característica presenta un valor elevado en comparación con la encontrada en la resistencia a otras enfermedades, siendo de 0,38 en la fase marina (Houston *et al*, 2008) y de 0,31 en la fase de agua dulce (Houston *et al*, 2010).

En la misma línea, dos grupos de investigación han identificado un QTL de efecto mayor en la resistencia a IPN ubicado en el grupo de ligamiento 21 del salmón del Atlántico, tanto en estudios con desafíos efectuados en la fase de agua dulce (Houston *et al*, 2010; Gheyas *et al*, 2010), como en la fase marítima (Houston *et al*, 2008; Moen *et al*, 2009). Este QTL explica de manera significativa la varianza genética de la característica, determinándose en desafíos experimentales que la presencia de un haplotipo resistente confiere protección contra el virus

en la fase de alevín (Houston *et al*, 2010), no detectándose mortalidad producida por la enfermedad frente al desafío experimental. Conjuntamente, se ha determinado una alta correlación de los haplotipos resistentes en los desafíos experimentales con respecto a la sobrevivencia en brotes en pisciculturas (Wetten *et al*, 2007), lo cual lo convierte en un excelente rasgo para la selección asistida por marcadores (MAS). Estos resultados han derivado en la inclusión de la característica a los programas de selección de Aqua Gen, Noruega.

Los eventos de selección pasados dejan huellas en el genoma de una población, las cuales pueden ser identificadas por medio de distintos marcadores moleculares a través de diferentes procedimientos (Excoffier *et al*, 2009). Un tipo de marcador que puede utilizarse para esto son los microsatélites, los cuales consisten en segmentos cortos de ADN en el que se repite en tándem un fragmento de 2 a 6 nucleótidos un número variable de veces. Aquellos loci que estén bajo selección y que se encuentren estrechamente ligados o en desequilibrio de ligamiento (LD) con un microsatélite, pueden alterar la frecuencia alélica de éste al segregar en forma conjunta, lo cual es conocido como el efecto hitchhiking (Wiehe, 1998). Otro fenómeno de importancia es el desequilibrio de ligamiento, el cual es la propiedad de dos o más loci de no segregar de forma independiente, ya sea que estén en el mismo o en diferentes cromosomas. En el caso de resistencia a IPN los resultados indican que podría haber desequilibrio de ligamiento sustancial entre marcadores en la región del QTL y el polimorfismo subyacente del QTL (Moen *et al*, 2009). Conociendo estos fenómenos, se pueden utilizar los niveles de LD de la población con el objetivo de detectar la presencia de una huella de selección en una región cromosómica (Wiener *et al*, 2003).

El objetivo de este trabajo es identificar huellas de selección asociadas con la resistencia a IPN, con el fin de comprender la composición genética de las poblaciones chilenas de salmón del Atlántico para esta característica. Con este objetivo se determinaron los niveles de LD dentro del grupo de ligamiento que incluye el QTL, y se analizaron los patrones de desequilibrio en la región cromosómica asociada al mismo.

MATERIALES Y MÉTODOS

El proyecto se llevó a cabo en el Laboratorio de Investigaciones en Biotecnología y Genómica Animal de la Facultad de Ciencias Veterinarias y Pecuarias de la Universidad de Chile (FAVET-INBIOGEN).

Poblaciones utilizadas en el estudio

Se utilizaron tres poblaciones comerciales y una naturalizada de la especie *Salmo salar*. Las poblaciones A, B y C han permanecido en cautividad completa durante una serie de generaciones, mientras que la población D es naturalizada, tomada del lago Ranco, X región. El objetivo de cría en las poblaciones comerciales incorpora varias características, tales como tasa de crecimiento, resistencia a enfermedades y características de calidad de canal.

Número de muestras

De las poblaciones comerciales A, B y C se muestrearon 63, 48 y 48 individuos no relacionados, respectivamente. De la población naturalizada (D) se muestrearon 30 individuos. En total se utilizaron 189 muestras para el estudio.

Manejo de las muestras

La extracción del ADN se realizó desde tejido muscular con el kit “DNeasy Blood & Tissue” (QIAGEN), siguiendo las instrucciones del fabricante. La integridad del ADN extraído fue evaluada por medio de electroforesis en gel de agarosa al 1%. La concentración de ADN presente en cada muestra se obtuvo a través de fluorescencia específica por medio del fluorómetro Qubit tm.

Marcadores microsatélites

Con el fin de genotipar las poblaciones en estudio se eligieron marcadores del tipo microsatélites sobre la base de su posición en el mapa del grupo de ligamiento 21, los cuales han confirmado su utilidad en otros estudios (Nikolic *et al*, 2009). Se sabe que este grupo de ligamiento segrega para un QTL de efecto mayor relacionado con la resistencia a IPN, utilizándose 11 marcadores (**Tabla 1**) distribuidos a lo largo del grupo de ligamiento (Reid *et al*, 2005; Moen *et al*, 2009). Los partidores para la reacción en cadena de la polimerasa (PCR) y tamaños de fragmentos de los microsatélites fueron tomados de cGRASP (www.cgrasp.org) o bien de la literatura. Inicialmente se consideraron dos marcadores adicionales (CL18304 y

Rsa476), pero debido a su bajo rendimiento en el genotipado no se tomaron en consideración para los análisis posteriores.

El mapa de ligamiento del grupo 21 utilizado se generó combinando los mapas de machos y hembras disponibles en Asalbase (www.asalbase.org, Phillips *et al*, 2009). Adicionalmente, se tomaron como referencia otros mapas de ligamiento presentes en la literatura (Houston *et al*, 2008; Moen *et al*, 2009). De esta manera se obtuvieron los valores en unidades de recombinación en centiMorgans (cM) entre marcadores, promediando las distancias de los mapas de hembra y macho considerando todos los estudios disponibles (**Tabla 2**).

Genotipado

La amplificación de los fragmentos se realizó por medio de la PCR, considerando la metodología de Schuelke (2000). Para marcar los fragmentos se utilizaron colas M13 marcadas con los fluoróforos 6-FAM, PET, VIC y NED. En una primera etapa se optimizaron los programas de PCR para cada marcador por medio de la visualización de resultados en geles de agarosa al 1%. Se realizó una segunda optimización orientada a una unión más eficiente de la cola M13 a los fragmentos amplificados dentro de la reacción, de modo de obtener un mayor nivel de fragmentos correctamente fluoroforados. Para esto se probaron tres programas de PCR diferentes, modificando la segunda parte del programa, el cual está orientado a la amplificación del fragmento incorporando el partidador fluoroforado, por lo que se modificó el número de ciclos con el fin de obtener un mayor número de fragmentos fluoroforados. Cada uno de estos programas se probó con tres mix distintos, tomando en consideración que la relación óptima entre el partidador sentido y la cola M13 fluoroforada puede variar considerablemente según la especie en estudio (West y Wilson, 2002; Lia *et al*, 2007; Villate *et al*, 2009).

En base a la optimización del párrafo anterior se estableció el protocolo descrito más adelante, el cual mejoró la eficiencia de fragmentos fluoroforados. Cada reacción se realizó en un volumen de 7 μ l conteniendo 1x de buffer $\text{NH}_4 \text{SO}_4$, 0,083 nmoles/ μ l de cada dNTP, 0,25 pmol/ μ l de partidador *sentido*, 1 pmol/ μ l de partidador *anti-sentido*, 1 pmol/ μ l del partidador cola M13 fluoroforada, 0,525 UI de *Taq* ADN polimerasa (Fermentas), 3 mM de MgCl_2 , 0,5 μ g/ μ l de Albumina de suero bovino como coadyuvante y 1,5 ng/ μ l de ADN. La reacción se llevó a cabo en un termociclador Axigen MaxiGene Gradient utilizando el siguiente programa: 7 minutos de denaturación inicial a 94°C; 21 ciclos de 94° C en 30 seg., 60°C (-0,5° por ciclo) por 30 seg., y

72°C por 30 seg.; 20 ciclos de 94°C por 35 seg., 53° C por 55 seg. y 72°C por 30 seg., 1 ciclo de extensión final de 72°C por 30 min. Los fragmentos fluoroforados fueron genotipados por medio de electroforesis capilar en un secuenciador ABI 3730XL. Los tamaños de los fragmentos y alelos fueron determinados por medio del programa GeneMapper 4.0 (Applied Biosystems).

Análisis genéticos generales y reconstrucción de haplotipos

Las frecuencias alélicas por marcador en cada población, Heterocigosidad observada y esperada, número de alelos por locus, desvíos del equilibrio Hardy-Weinberg y los estadísticos F de Wright fueron calculados por medio del programa GENEPOP 4.0 en línea (Raymond y Rousset, 1995).

El haplotipo de cada individuo se reconstruyó utilizando el programa PHASE (Stephens *et al*, 2001; Stephens y Scheet, 2005). Este programa permite reconstruir haplotipos a partir de los genotipos de la población.

Desequilibrio de ligamiento

Para el análisis del LD del conjunto de marcadores con respecto al QTL, la extensión del desequilibrio, considerando la distancia entre los marcadores, se calculó con el estadístico D':

$$D' = \frac{\sum_i \sum_j P(A_i)P(B_j)|D_{ij}|}{D_{ij}^{\max}}$$

Donde, $D_{ij} = P(A_i B_j) - P(A_i)P(B_j)$; $P(A_i)$ es la frecuencia del alelo i en locus A; $P(B_j)$ es la frecuencia del alelo j en locus B; $D_{ij}^{\max} = \min[P(A_i)P(B_j), (1-P(A_i))(1-P(B_j))]$ cuando $D_{ij} < 0$, y $D_{ij}^{\max} = \min[(P(A_i))(1-P(B_j)), (1-P(A_i))(P(B_j))]$ cuando $D_{ij} > 0$.

Los valores de D' se estimaron por medio de la función HAPLOXT desde los haplotipos reconstruidos, la cual se encuentra incluida en el paquete programático GOLD para Windows (Abecasis y Cookson, 2000).

Para el análisis de decaimiento del LD se utilizó la metodología descrita por Heifetz *et al* (2005). Ésta se basa en la ecuación de Sved (1971), estableciendo el decaimiento del LD en relación con la distancia entre marcadores. El modelo utilizado se describe a continuación:

$$LD_{ij} = 1/(1+4b_j d_{ij}) + e_{ij}$$

Donde, LD_{ij} es el desequilibrio de ligamiento observado para el par de marcadores i del conjunto de datos j , d_{ij} es la distancia en Morgans del par de marcadores i en la base de datos j , b_j describe la declinación de LD con la distancia para el conjunto de datos j , y e_{ij} es el residuo aleatorio. Esta función corrige el LD en base a la distancia en unidades de recombinación. El decaimiento por la distancia (b_j) fue estimado para cada población de manera independiente por medio de una regresión no lineal de la función, utilizando el programa Infostat (Di Rienzo *et al*, 2010).

Se dividió el cromosoma en tres regiones, dejando reunidos en una misma región los marcadores asociados al QTL de efecto mayor para resistencia a IPN (Moen *et al.*, 2009; Houston, *et al.*, 2008). Se evaluaron las diferencias en los niveles de LD entre regiones por medio de la metodología descrita por Heifetz *et al* (2005). Se utilizaron los residuos del modelo para evaluar las diferencias entre regiones para cada población por medio de un ANDEVA utilizando la prueba de diferencia menos significativa (LSD) de Fisher.

RESULTADOS

Parámetros genéticos generales

Todos los loci fueron polimórficos, siendo el número promedio de alelos por marcador de 10,6 dentro de un rango de 4 a 24 para el conjunto de poblaciones (**Tabla 3**).

El valor medio de la heterocigosidad observada fue de 52%, variando de 9 a un 86%. La heterocigosidad observada y esperada por locus para las cuatro poblaciones, junto con el análisis de equilibrio Hardy-Weinberg (HW) se presentan en la **Tabla 4**. En ésta se observa que en el 47% de los casos existe un desvío del equilibrio HW, de los cuales más de la mitad tiene un valor de endogamia (F_{is}) elevado **Tabla 5**. Esto sugiere que la causa del desvío del equilibrio HW en una proporción importante de los loci es debido a un déficit de heterocigotos.

Diferenciación genética

Los valores de diferenciación genética entre las poblaciones (F_{st}) indicaron una muy baja diferenciación entre estas (0,033) al compararlas entre sí, considerando todos los marcadores (**Tabla 6**). La diferenciación poblacional presentó considerables variaciones según marcador (**Figura 1**).

Con respecto a los valores de endogamia se calcularon para cada población por marcador (**Tabla 5**), siendo los loci con mayor nivel de endogamia Rsa354 y Ssa0374BSFU, siendo la población B la que presenta los niveles más elevados de F_{IS} .

Niveles de desequilibrio de ligamiento

Para la determinación de los valores de desequilibrio de ligamiento se dividió el grupo de ligamiento 21 en 3 regiones (**Figura 2**). Los valores obtenidos para el decaimiento por la distancia (b_j) fueron $0,17 \pm 0,02$; $0,09 \pm 0,01$; $0,07 \pm 0,01$ y $0,1 \pm 0,01$ para las poblaciones A, B, C y D, respectivamente, siendo el valor de C significativamente distinto al de la población A. Con respecto al b_j de las regiones en cada población, la región 2 presentó los valores más bajos en todas las poblaciones, siendo el menor valor el de la población C (**Tabla 7**). En la **Figura 3** se aprecia el decaimiento del LD en cada una de las poblaciones en base a la distancia.

Tras corregir por la distancia se compararon los niveles de D' calculados a partir de los haplotipos reconstruidos entre las regiones para cada población, por medio de un ANDEVA utilizando la prueba LSD de Fisher. La población C presentó diferencias significativas ($p < 0,01$) entre la región 2 respecto a la región 1 y 3, mientras que en las otras poblaciones no se encontraron diferencias significativas entre regiones (**Figuras 4, 5, 6 y 7**).

DISCUSIÓN

Las poblaciones chilenas de salmón han estado bajo el efecto de procesos de selección natural y comercial a través de las generaciones, modificando su estructura genética en base a distintas presiones de selección. Los programas comerciales de mejoramiento que se han implementado en la salmonicultura han tenido diversos objetivos de selección, siendo importante conocer el efecto generado a nivel genómico producto de la selección en las poblaciones de salmón del Atlántico, de modo de conocer la composición genética de las poblaciones nacionales. Los resultados obtenidos en esta investigación muestran cómo se ha modelado el genoma de las poblaciones chilenas de salmón del Atlántico para la región en la cual se encuentra el QTL responsable de resistencia a IPN.

Existieron diferencias significativas en la población C para la región que posee el QTL de resistencia a IPN, que podrían atribuirse a un proceso de selección actuando dentro de la población (Wiehe, 1998). Con el estudio se logró establecer que la región 2 presenta niveles de LD mayores al resto del cromosoma. Respecto a lo anterior, es conocido que el LD es una

importante huella de selección (Kim y Nielsen, 2004) y que se ha utilizado para identificar huellas generadas tanto por la selección artificial a través de los programas de mejoramiento (Hayes *et al*, 2008), como las generadas producto de la selección natural (Teacher *et al*, 2009). Cuando una mutación favorable surge, y aumenta hacia la fijación en un gen, se genera una ventaja en los genes con los cuales está asociado, aumentando la frecuencia de los genes ligados a la mutación seleccionada. Este proceso se conoce como “*hitchhiking*” genético (Barton, 2000). La selección de la mutación benéfica puede alterar los patrones de variación genética en los alrededores del locus seleccionado vía *hitchhiking* (Wiener *et al*, 2003), promoviéndose un aumento en los valores de LD por medio de ésta vía en los alrededores del gen seleccionado (Kim y Nielsen, 2004). Este fenómeno sería el acontecido en la población C. Una hipótesis para este fenómeno es la aparición de una mutación benéfica, otorgando una ventaja en la resistencia a IPN, ocurriendo la selección del rasgo. Así mismo, se han identificado huellas de selección en otros vertebrados para resistencia frente a virus (Teacher *et al*, 2009), producto de la acción de la selección natural.

Los resultados obtenidos para el decaimiento del LD con la distancia muestran que en la población C éste decae más lentamente. Al realizar el análisis por región se aprecian diferencias significativas entre la región 2 con respecto a las otras regiones, lo que se condice con el hecho de que la población C tenga una diferencia significativa para esta región. Esto se explica por un nivel de LD mayor, asociado a un decaimiento con la distancia menos pronunciado, lo cual podría ser ocasionado por la selección del QTL presente en esta región cromosómica, y que por medio del efecto *hitchhiking* alteraría la conformación de los loci aledaños al QTL (Wiehe, 1998).

El cromosoma se dividió en estas tres regiones con el fin de incluir a todos los marcadores ligados en mayor o menor medida al QTL en una región (Houston *et al*, 2008; Moen *et al*, 2009). Se realizó un análisis alternativo dividiendo la región asociada al QTL en dos regiones de menor tamaño. Los resultados obtenidos fueron muy similares, diferenciándose estas dos subregiones de las otras en la población C (datos no mostrados), lo cual indica que la región 2 tiene un nivel de desequilibrio mayor a través de toda su extensión.

En el análisis de diferenciación poblacional (F_{st}) se aprecia que los valores entre todas las poblaciones son muy pequeños, lo que representa un bajo nivel de diferenciación. Se puede apreciar que la población D es la que más se diferencia de las otras poblaciones para los marcadores que presentan mayores diferencias poblacionales, pero aún en este caso la

diferenciación es muy baja. Este bajo nivel de diferenciación poblacional puede explicarse por el origen común de las poblaciones chilenas, las cuales provienen de poblaciones comerciales europeas originadas por medio de un proceso de admixtura ocurrido entre poblaciones naturales provenientes de diversos ríos. La admixtura consiste en la mezcla de poblaciones de diferente origen, con distintas frecuencias génicas, generando frecuencias diferentes en las poblaciones resultantes (Long, 1991). Es de esperar que a medida que aumente el número de generaciones de las poblaciones, los valores de F_{st} tiendan a aumentar, en particular la diferenciación entre la población naturalizada con las poblaciones comerciales, debido a las diferencias existentes tanto en los sistemas de cruce como en las presiones a las que se ven sometidas .

El análisis de Hardy-Weinberg busca establecer si en una población las frecuencias génicas y genotípicas permanecen constantes de generación en generación, cuando la población cumple con apareamiento aleatorio, herencia mendeliana y ausencia de selección, mutación o migración (Falconer y Mackay, 1996). Los resultados de la prueba HW muestran que una proporción importante de los loci se desvía del equilibrio. En un porcentaje elevado de estos loci, el desvío podría explicarse a causa de un déficit de heterocigotos, presumiblemente dado por un mayor nivel de endogamia, ya que los loci con mayor F_{is} se desvían del equilibrio HW. Los valores de endogamia en las poblaciones (F_{is}) presentaron un amplio rango, tanto entre marcadores como entre poblaciones, sin embargo la población B es la que muestra los valores más elevados (promedio $>0,24$). Estos mayores valores de la población B en los niveles de endogamia se explicarían por el tipo de manejo reproductivo de ésta población comercial, el cuál consiste en cruzamientos cerrados, lo que produciría el aumento en éste parámetro. Si bien una gran proporción de los desvíos del equilibrio HW se explica por el déficit de heterocigotos, es posible que un proceso de selección sobre este grupo de ligamiento, esté alterando las frecuencias génicas y genotípicas dentro de las poblaciones (Wiehe, 1998), en particular en la población C. La selección, al modificar las frecuencias alélicas y genotípicas de una generación a la siguiente, es capaz de producir un desvío del equilibrio HW. Los bajos niveles de heterocigosidad para Ssa0279BSFU, BHMS217, Ssa0374BSFU y Alu333 en la población C podrían explicarse por un proceso selectivo del QTL ubicado en la misma región que estos microsatélites. Pese a lo anterior, no es posible descartar el efecto que pueda tener un proceso previo de admixtura de las poblaciones que dieron origen a las utilizadas en este estudio. Es bien sabido que un proceso de admixtura altera las frecuencias génicas de las poblaciones, generando desviaciones del equilibrio HW (Deng *et al*,

2001), por lo anterior no es posible descartar que los desvíos del equilibrio HW sean causados en cierta medida por la admixtura.

Se ha descrito a nivel poblacional que un exceso en el número de homocigotos puede ser causado por la presencia de alelos nulos. Los alelos nulos ocurren por la mutación en el sitio de unión de los partidores, lo que genera la ausencia de amplificación de un alelo. (Pemberton *et al*, 1995). Para comprobar la existencia de un alelo nulo se pueden utilizar partidores en posiciones diferentes a los originales, secuenciar el marcador junto con las regiones flanqueantes en las cuales se unen los partidores, o utilizar familias comprobadas de padres e hijos, comparando los alelos entre ambos y evidenciando la ausencia de un alelo (causado por un alelo nulo) en el caso en que no haya compatibilidad de padres e hijos (Kline *et al*, 2011). No se han descrito presencia de alelos nulos en ninguna de las publicaciones en las cuales se han utilizado los microsatélites de este estudio (Houston *et al*, 2008; Houston *et al*, 2010; Moen *et al*, 2009; Gheyas *et al*, 2010), por lo que se espera que no exista la presencia de alelos nulos para estos marcadores, sin poder descartar la existencia de nuevas mutaciones en las poblaciones utilizadas en ésta investigación.

El antecedente de que las cuatro poblaciones tengan un origen similar (salmón del Atlántico de origen europeo introducido por la industria acuícola), junto con una baja diferenciación poblacional se relaciona con el proceso de admixtura ocurrido generaciones atrás. Los resultados de la población C muestran la diferenciación de la región en la cual segrega el QTL, dando cuenta de una segregación del QTL asociado a IPN (Houston *et al*, 2008; Moen *et al*, 2009).

Si bien los resultados muestran la existencia de una huella de selección en la población C, éstos deben tomarse con precaución, debido a la admixtura ocurrida previamente. Se ha descrito que la admixtura al generar una mezcla de poblaciones con distintas frecuencias génicas puede producir niveles de LD elevados (Falconer y Mackay, 1996). Debido a esto sería recomendable utilizar otras aproximaciones metodológicas complementarias para detectar huellas de selección, tal como plantea Li *et al* (2010), quién encontró divergencias en los resultados según la aproximación utilizada, siendo recomendable utilizar aproximaciones que consideren diferentes supuestos.

CONCLUSIÓN

La investigación permite inferir la presencia de una huella de selección en la región donde se encuentra el QTL de resistencia a IPN para la población C, dando cuenta de cómo la selección natural y los programas de mejoramiento genético han modelado la conformación estructural de la región asociada a la característica. Estos resultados muestran la utilidad de éste conjunto de microsatélites como candidatos en programas de selección asistida por marcadores con el objetivo de aumentar la resistencia a IPN.

Se concluye que los elevados valores de LD en la región del QTL para resistencia a IPN, junto con los bajos niveles de heterocigosidad en los microsatélites cercanos al QTL, serían causados por un proceso de selección en la zona del QTL. Los bajos niveles de diferenciación poblacional dan cuenta de lo relacionadas que están las poblaciones de salmón nacionales, y como la admixtura de las poblaciones ancestrales ha condicionado la estructura genética de las poblaciones chilenas. Debido a estos antecedentes los resultados deben ser tomados con precaución, siendo recomendable utilizar otras metodologías alternativas para ratificar la huella de selección encontrada en este estudio.

AGRADECIMIENTOS

Financiado por proyecto FONDECYT 1090632. En primer lugar quiero agradecer al Dr. Víctor Martínez, por la confianza y el apoyo otorgado para participar en este proyecto, y por su ayuda en la comprensión del tema. También quiero agradecer al personal de FAVET-INBIOGEN, en especial a Ana Jedlicki y José Manuel Yañez, por su disposición a resolver dudas y prestar ayuda. También quiero agradecer a Marcia Arriagada por su apoyo y sus consejos y su confianza a través de este proceso.

Finalmente quiero agradecer a mis padres, por su dedicación, confianza y apoyo incondicional, y que me han permitido cumplir los distintos desafíos a nivel profesional y personal, llevándome a finalizar este proceso universitario.

REFERENCIAS

ABECASIS, G; COOKSON, W. 2000. GOLD--graphical overview of linkage disequilibrium
Bioinformatics. 16: 182-183.

ATLANTIC SALMON GENOMIC AND GENETIC SEQUENCE DATABASE.
<www.asalbase.org>.

BARTON, N H. 2000. Genetic hitchhiking. Philosophical Transactions of the Royal Society
Biological Sciences. 355, 1553-1562.

CONSORTIUM FOR GENOMICS RESEARCH ON ALL SALMONIDS PROYECT.
<www.cgrasp.org>.

DENG, H; CHEN, W; RECKER, R. 2001. Population admixture: detection by hardy-weinberg
test and its quantitative effects on linkage-disequilibrium methods for localizing genes
underlying complex traits. Genetics. 157: 885–897.

DI RIENZO, J; CASANOVES, F; BALZARINI, M; GONZALEZ, L; TABLADA, M; ROBLEDO, C.
InfoStat versión 2010. Grupo InfoStat, FCA, Universidad Nacional de Córdoba, Argentina.

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

FALCONER, D S; MACKAY, T F C. 1996. Introduction to Quantitative Genetics. 4 edition.
Editorial Addison Wesley Longman Limited, Harlow, England.

FIP, FONDO DE INVESTIGACIÓN PESQUERA, Chile 2003. Riesgo de introducción de
enfermedades infectocontagiosas en salmónidos. Informe final Proyecto FIP 2001-2008.

GHEYAS, A; HOUSTON, R; MOTA-VELASCO, J; GUY, D; TINCH, A; HALEY, C; WOOLLIAMS, J. 2010. Segregation of infectious pancreatic necrosis resistance QTL in the early life cycle of Atlantic Salmon (*Salmo salar*). *Animal Genetics*. 41: 531-536.

HAYES, B; LIEN, S; NILSEN, H; OLSEN, H; BERG, P; MACEACHERN, S; POTTER, S; MEUWISSEN, T. 2008. The origins of selection signatures on bovine chromosome 6. *Animal Genetics*. 39: 105-111.

HEIFETZ, E; FULTON, J; O'SULLIVAN, E; ZHAO, H; DEKKERS, J; SOLLER, M. 2005. Extent and consistency across generations of linkage disequilibrium in commercial layer chicken breeding populations, *Genetics*. 171: 1173-1181.

HOUSTON, R; HALEY, C; HAMILTON, A; GUY, D; TINCH, A; TAGGART, J; MCANDREW, B; BISHOP, S. 2008. Major QTL affect resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*). *Genetics*.178: 1109-1115.

HOUSTON, R; HALEY, C; HAMILTON, A; GUY, D; MOTA-VELASCO, J; GHEYAS, A; TINCH, A; TAGGART, J; BRON, J; STARKEY, W; MCANDREW, B; VERNER-JEFFREYS, D; PALEY, R; RIMMER, G; TEW, I; BISHOP, S. 2010. The susceptibility of Atlantic salmon fry to freshwater infectious pancreatic necrosis is largely explained by a major QTL. *Heredity*. 105: 318-327.

KIM, Y; NIELSEN, R. 2004. Linkage disequilibrium as a signature of selective sweeps. *Genetics*. 167: 1513–1524.

KLINE, M; HILL, C; DECKER, A; BUTLER, J. 2011. STR sequence analysis for characterizing normal, variant, and null alleles. *Forensic Science International: Genetics*. 5 : 329–332.

LIA, J; ZHANGA, Y; SULLIVANB, M; HONGA, L; HUANGA, L; LUC, F; MCCUTCHANB, T; SU, X. 2007. Typing *Plasmodium yoelii* microsatellites using a simple and affordable fluorescent labeling method. *Molecular Biochemistry Parasitology*. 155(2): 94–102.

LI, MH; ISO-TOURU, T; LAURÉN, H; KANTANEN, J. 2010. A microsatellite-based analysis for the detection of selection on BTA1 and BTA20 in northern Eurasian cattle (*Bos taurus*) populations. *Genetics Selection Evolution*. 42: 32.

LONG, J C. 1991. The genetic structure of admixed populations. *Genetics*. 127: 417-428.

MOEN, T; BARANSKI, M; SONESSON, A; KJOGLUM, S. 2009. Confirmation and fine-mapping of a major QTL for resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*): population-level associations between markers and trait. *BMC Genomics*. 10: 368-382.

NIKOLIC, N; FE`VE, K; CHEVALET, C; HOYHEIM, B; RIQUET, J. 2009. A set of 37 microsatellite DNA markers for genetic diversity and structure analysis of Atlantic salmon. *Journal of Fish Biology*. 74: 458-466.

OZAKI, A; SAKAMOTO, T; KHOO, S; NAKAMURA, K; COIMBRA, M; AKUTSU, T; OKAMOTO, N. 2000. Quantitative trait loci (QTLs) associated with resistance/susceptibility to infectious pancreatic necrosis virus (IPNV) in rainbow trout (*Oncorhynchus mykiss*). *Molecular Genetics and Genomics*. 265: 23-31.

PEMBERTON, JM; SLATE, J; BANCROFT, DR; BARRETT, JA. 1995. Nonamplifying alleles at microsatellite loci: a caution for parentage and population studies. *Molecular Ecology*. 4: 249-52.

PHILLIPS, R; KEATLEY, K; MORASCH, M; VENTURA, B; LUBIENIECKI, K; KOOP, B; DANZMANN, R; DAVIDSON, W. 2009. Assignment of Atlantic salmon (*Salmo salar*) linkage groups to specific chromosomes: Conservation of large syntenic blocks corresponding to whole chromosome arms in rainbow trout (*Oncorhynchus mykiss*). BMC Genetics. 10: 46.

RAYMOND, M; ROUSSET, F. 1995. GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. Heredity. 86: 248-249.

REID, D; SZANTO, A; GLEBE, B; DANZMANN, RG; FERGUSON, M. 2005. QTL for body weight and condition factor in Atlantic salmon (*Salmo salar*) comparative analysis with rainbow trout (*Oncorhynchus mykiss*) and Arctic charr (*Salvelinus alpinus*). Heredity. 94:166-172.

REXROAD, C; VALLEJO, R. 2009. Estimates of linkage disequilibrium and effective population size in rainbow trout. BMC Genetics. 10: 83.

SCHUELKE, M. 2000. An economic method for the fluorescent labeling of PCR fragments. Nature. 18: 233-234.

STEPHENS, M; SMITH, N; DONNELLY, P. 2001. A new statistical method for haplotype reconstruction from population data. American Journal of Human Genetics. 68: 978-989.

STEPHENS, M; SCHEET, P. 2005. Accounting for decay of linkage disequilibrium in haplotype inference and missing-data imputation. American Journal of Human Genetics. 76: 449-462.

SVED, J. A. 1971. linkage disequilibrium and homozygosity of chromosome segments in finite populations. Theoretical Population Biology. 2: 125-141.

TEACHER, A; GARNER, T; NICHOLS, R . 2009. Evidence for directional selection at a novel major histocompatibility class I marker in wild common frogs (*Rana temporaria*) exposed to a viral pathogen (*Ranavirus*). PLoS ONE. 4: e4616.

VILLATE, L; ESMENJAUD, D; COEDEL, S; PLANTARD, O. 2009. Development of nine polymorphic microsatellite markers for the phytoparasitic nematode *Xiphinema index*, the vector of the grapevine fanleaf virus. Molecular Ecology Resources. 9: 229–232.

WEST, M Y WILSON, G. 2002. Direct fluorescent primers are superior to M13-tailed primers for *Pinus taeda* microsatellites. BioTechniques. 32:46-52.

WETTEN, M; AASMUNDSTAD, T; KJØGLUM, S; STORSET, A. 2007. Genetic analysis of resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*). Aquaculture. 272: 111–117.

WIEHE, T. 1998. The effect of selective sweeps on the variance of the allele distribution of a linked multiallele locus: hitchhiking of microsatellites. Theoretical Population Biology. 53: 272-283.

WIENER, P; BURTON, D; AJMONE-MARSAN, P; DUNNER, S; MOMMENS, G; NIJMAN, I J; RODELLAR, C; VALENTINI, A; WILLIAMS, J L. 2003. Signatures of selection? Patterns of microsatellite diversity on a chromosome containing a selected locus. Heredity. 90. 350–358.

ZHOU, Y; BUI, T; AUCKLAND, L; WILLIAMS, C. 2002. Direct Fluorescent primers are superior to M13-Tailed primers for *Pinus taeda* Microsatellites. BioTechniques. 32: 46-52.

ANEXOS

Tabla 1. Marcadores pertenecientes al grupo de ligamiento 21 del Salmón del Atlántico utilizados en el estudio con sus respectivos partidores. A los partidores sentido se les agregó una cola M13 (CACGACGTTGTAAAACGAC) en el extremo 5'.

Marcador	Partidor <i>sentido</i>	Partidor <i>anti-sentido</i>	Repetición
Rsa354	5'-TGTTACTTAGGGTTGAACG-3'	5'-AGGTTGGAAGTGTGCTG	Dinucleótido
CL68944	5'-CAGCACCACCACCCAACA-3'	5'-TCCATCGGCCTCCCTCTA	Dinucleótido
Omy44TUF	5'-GAGGGTTGGAGTACACAGAAGG-3'	5'-ATGTGGGGACATATTAAGTGGC-3'	Dinucleótido
Ssa0279BSFU	5'-ATGAGAGGATGCTTCTTCAA-3'	5'-AGACGGAACGAGTGTGTTAT-3'	Dinucleótido
BHMS217	5'-GCTGTTCAATTTCTGAGCAG-3'	5'-GACACACCGAATCAGTGC-3'	Dinucleótido
CL13225	5'-GGTTGAGGTCAGGGGGTA-3'	5'-TGTTCAAGTGGCACATTTTGA-3'	Dinucleótido
Ssa0285BSFU	5'-CAGAACACAAACAGAGCTGA-3'	5'-CAACAGGGATCTCTCAACAT-3'	Dinucleótido
Alu333	5'-TTCATAGTCCAAGAACAGTG-3'	5'-GCTGAGTTTACATTACACCTG-3'	Dinucleótido
Ssa0374BSFU	5'-GAGTAGGCAACTGAAACAGG-3'	5'-CAAACCTATTCCCTCACATT-3'	Tetranucleótido
CL10332	5'-GGCTCTTCCACTTGTCCA-3'	5'-TGCTAAGGTGAATGGGTGT-3'	Dinucleótido
Ssa0735BSFU	5'-CTTTTGCAAACCTCAGCTCT-3'	5'-AATATGTCTGTCCGATGAGG-3'	Dinucleótido

Tabla 2. Ubicación dentro del grupo de ligamiento 21 de los microsatélites utilizados en el estudio.

Marcador	Mapa conjunto (cM)	Mapa hembra (cM)	Mapa macho (cM)
Rsa354	5,95	1,1	10,8
CL68944	7,33	3,87	10,8
OmyRGT44TUF	8,1	5,4	10,8
Ssa0279BSFU	10,25	9,7	10,8
BHMS217	11,35	11,9	10,8
CL13225	11,9	13	10,8
Ssa0285BSFU	14,05	17,3	10,8
Alu333	17,3	23,8	10,8
Ssa0374BSFU	17,3	23,8	10,8
CL10332	18,85	26,9	10,8
Ssa0735BSFU	21,15	31,5	10,8

Tabla 3. Número de alelos por locus.

Locus	Población A	Población B	Población C	Población D	Promedio	Total alelos
Rsa354	17	17	14	15	15,75	20
CL68944	5	5	5	5	5	5
OmyRGT44TUF	5	3	3	2	3,25	5
Ssa0279BSFU	4	3	3	3	3,25	4
BHMS217	10	8	9	9	9	12
CL13225	7	4	2	5	4,5	7
Ssa0285BSFU	5	8	8	8	7,25	11
Alu333	6	7	6	5	6	8
Ssa0374BSFU	23	13	17	14	16,75	24
CL10332	8	5	4	7	6	10
Ssa0735BSFU	9	7	6	7	7,25	11
Promedio	9	7,273	7	7,273	7,636	10,636

Tabla 4. Heterocigosidad observada (Ho), esperada (He) y Hardy-Weinberg (p-value) por población.

Locus	Población A			Población B			Población C			Población D		
	Ho	He	p-value									
Rsa354	0,586	0,828	0,0000	0,568	0,904	0,0000	0,533	0,823	0,0000	0,643	0,878	0,0041
CL68944	0,725	0,726	0,4221	0,500	0,755	0,0339	0,733	0,716	0,0460	0,667	0,756	0,1648
OmyRGT44TUF	0,258	0,596	0,0000	0,277	0,337	0,2060	0,468	0,530	0,4178	0,400	0,398	1,0000
Ssa0279BSFU	0,190	0,180	1,0000	0,106	0,142	0,0984	0,091	0,089	1,0000	0,167	0,158	1,0000
BHMS217	0,833	0,774	0,3010	0,660	0,794	0,0511	0,625	0,748	0,0002	0,733	0,835	0,0059
CL13225	0,433	0,523	0,0003	0,109	0,145	0,2180	0,104	0,100	1,0000	0,345	0,506	0,0518
Ssa0285BSFU	0,754	0,613	0,0433	0,644	0,756	0,0099	0,681	0,688	0,6670	0,600	0,774	0,0000
Alu333	0,774	0,693	0,0000	0,500	0,692	0,0003	0,523	0,777	0,0002	0,567	0,707	0,2149
Ssa0374BSFU	0,567	0,907	0,0000	0,565	0,900	0,0000	0,488	0,907	0,0000	0,867	0,872	0,7415
CL10332	0,382	0,409	0,1174	0,383	0,426	0,0963	0,333	0,345	0,0318	0,600	0,729	0,0002
Ssa0735BSFU	0,730	0,759	0,0173	0,591	0,775	0,0575	0,681	0,693	0,2946	0,640	0,673	0,7794
PROMEDIO	0,567	0,637		0,446	0,602		0,478	0,583		0,566	0,662	

Tabla 5. Valores de endogamia F_{is} por población para cada locus.

	Población A	Población B	Población C	Población D
Locus	F_{is}	F_{is}	F_{is}	F_{is}
Rsa354	0,2939	0,3745	0,3549	0,2714
CL68944	0,0005	0,3421	-0,0251	0,1199
OmyRGT44TUF	0,5692	0,1808	0,1173	-0,0058
Ssa0279BSFU	-0,0598	0,2532	-0,0269	-0,0584
BHMS217	-0,0774	0,1707	0,1659	0,1236
CL13225	0,1722	0,2512	-0,0444	0,322
Ssa0285BSFU	-0,2321	0,1485	0,0101	0,2278
Alu333	-0,1178	0,2798	0,3297	0,201
Ssa0374BSFU	0,3774	0,3745	0,4646	0,0066
CL10332	0,0659	0,1024	0,034	0,1792
Ssa0735BSFU	0,0388	0,2395	0,0177	0,0507

Tabla 6. Valores de diferenciación poblacional (F_{st}) entre pares de poblaciones.

Población	A	B	C
B	0,0333		
C	0,0311	0,0301	
D	0,0329	0,0403	0,0359

Tabla 7. Valores de declinación de LD con la distancia (b_j) por población para cada región.

Población	Región 1	Región 2	Región 3
A	0,37±0,15	0,17±0,05	0,29±0,21
B	0,27±0,04	0,08±0,02	0,15±0,06
C	0,49±0,14	0,05±0,01	0,21±0,13
D	0,27±0,15	0,12±0,03	0,18±0,08

Figura 1. Valores de diferenciación poblacional (F_{st}) entre pares de poblaciones por marcador.

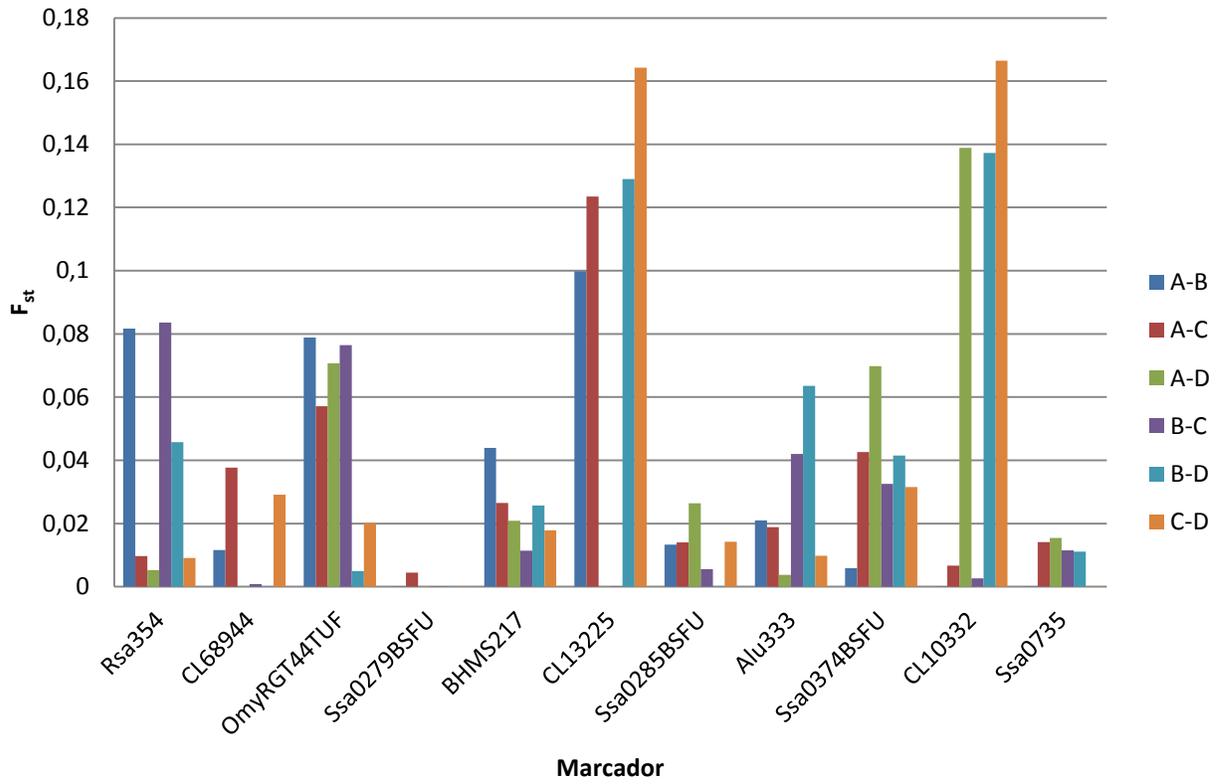


Figura 2. Mapa del grupo de ligamiento 21 dividido en tres regiones.

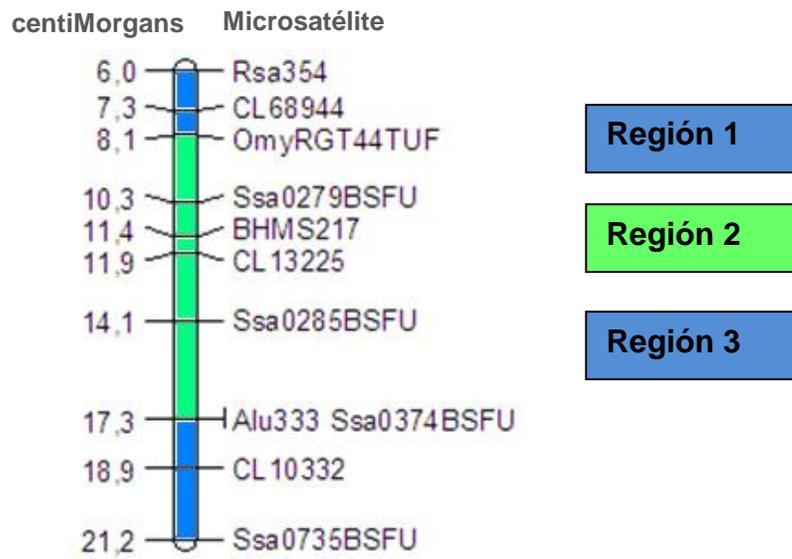
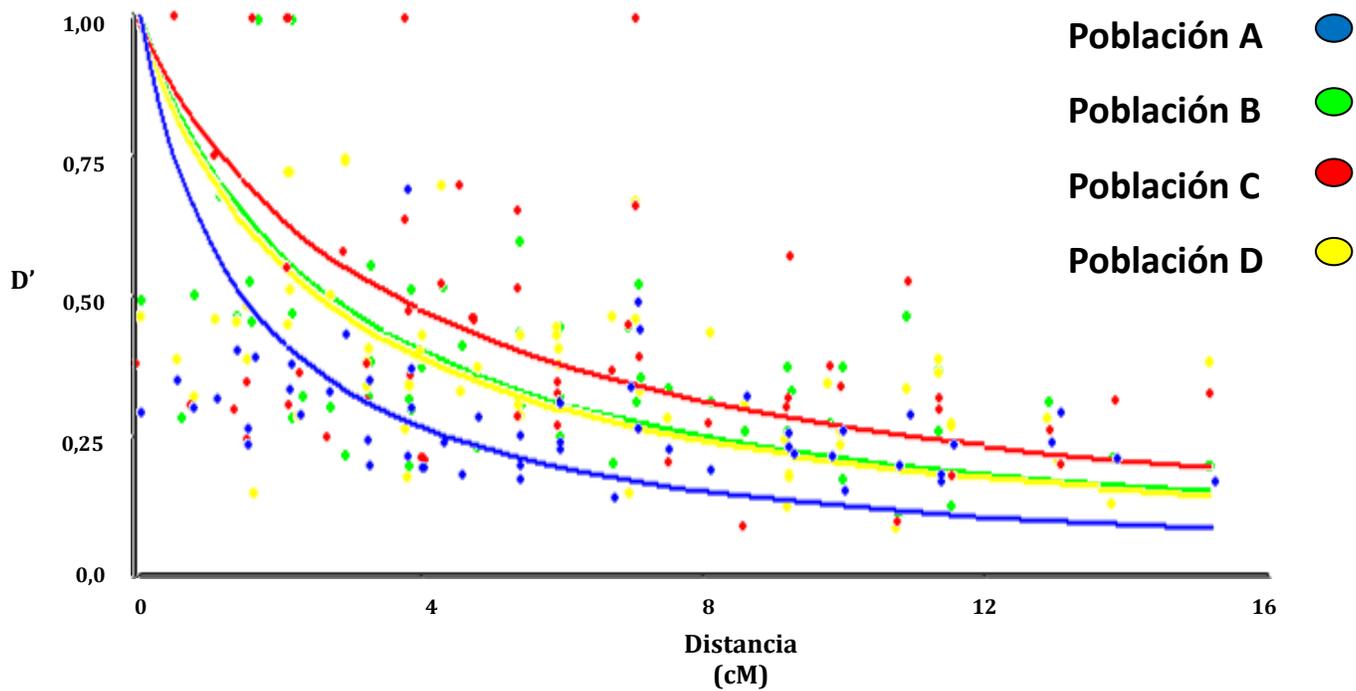
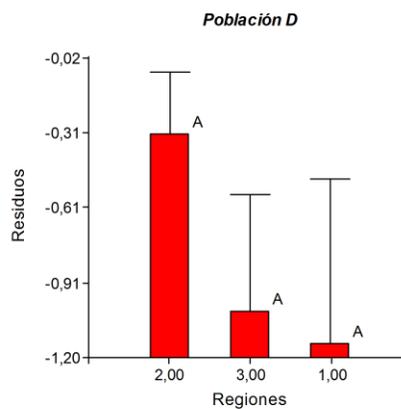
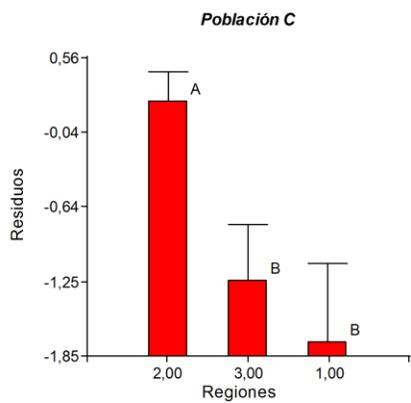
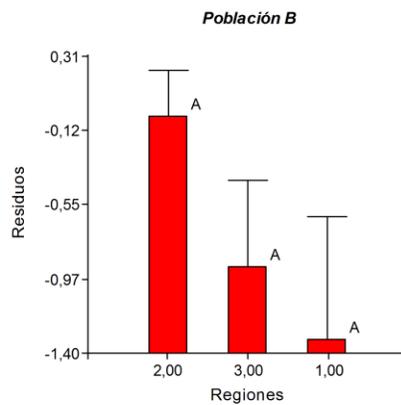
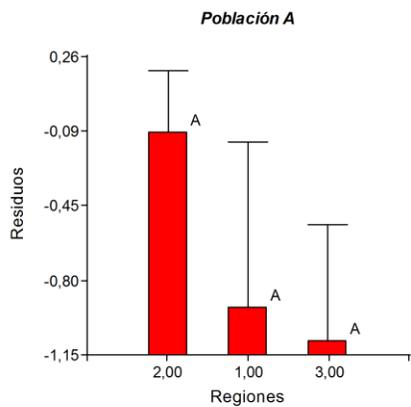


Figura 3. Gráfico de LD (D') y su decaimiento con la distancia (cM) para las cuatro poblaciones.



Figuras 4, 5, 6 y 7. ANDEVA entre regiones en cada población a partir de HaploXT.



GLOSARIO

Admixtura: mezcla de genomas de diferentes poblaciones.

Alélos: diferentes formas del mismo locus que difiere en una o varias bases de la secuencia.

CentiMorgan (cM): como medida de distancia genética, corresponde con un 1% de frecuencia de recombinación.

Desequilibrio de ligamiento: es la propiedad de dos o más loci de no segregar de forma independiente. Se ve afectado por selección, mutación, deriva génica y estructura poblacional. El desequilibrio de ligamiento entre dos marcadores disminuye continuamente debido a la recombinación. La selección produce un aumento del desequilibrio de ligamiento alrededor de la posición seleccionada.

Estadísticos F de Wright: mide la reducción de heterocigosis entre distintos niveles de una jerarquía poblacional. Para definir la estructuración genética entre poblaciones se utiliza el índice de fijación (F_{st}), el cual sirve para indicar el grado en el que dos o más poblaciones se diferencian entre sí. El coeficiente de consanguinidad o endogamia (F_{is}) mide el déficit de heterocigotos en la población, indicando si existe endogamia o exogamia.

Genotipo: La combinación de alélos presentes en un locus o en un individuo.

Heterocigoto: Un individuo con dos diferentes alélos en un locus.

Hitchiking: es el proceso en el cual un alelo puede aumentar su frecuencia a causa de estar ligado a un gen que está siendo seleccionado positivamente.

Homocigoto: Un individuo con dos copias del mismo alélo en un locus.

Loci: plural de locus.

Locus: es una posición fija sobre un cromosoma, como la posición de un gen o de un biomarcador (marcador genético).

Microsatélites: secuencias de ADN en las que un fragmento (de uno hasta seis nucleótidos) se repite de manera consecutiva. La variación en el número de repeticiones crea diferentes alelos los cuales se distinguen entre sí por la longitud total del fragmento.